Optimization of device-independent witnesses of entanglement depth from two-body correlators

J. Tura [1,*] A. Aloy,[2] F. Baccari,[2] A. Acín,[2,3] M. Lewenstein,[2,3] and R. Augusiak[4]

[1]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany*
[2]*ICFO - Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*
[3]*ICREA, Pg. Lluis Companys 23, 08010 Barcelona, Spain*
[4]*Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotnikòw 32/46, 02-668 Warsaw, Poland*

In a recent work [A. Aloy *et al.*, Phys. Rev. Lett. **123**, 100507 (2019)] we have considered the characterization of entanglement depth, from a device-independent perspective, in a quantum many-body system. We have shown that the inequalities introduced by J. Tura *et al.* [Science **344**, 1256 (2014)] can be used to obtain device-independent witnesses of entanglement depth (DIWEDs) and that they enjoy two key properties that allow one to compute their k-producibility bounds more efficiently for larger system sizes, as well as yielding experimentally friendlier device-independent witnesses of entanglement depth: they involve at most two-body correlators and they are permutationally invariant. While the main aim of our previous work was to illustrate the main ideas and applicability of the method, here we outline the details and complement its findings with detailed analysis and further case studies. Specifically, we consider the problem of finding the k-producible bounds of such DIWEDs under different assumptions. Not surprisingly, with the weakest assumptions, we can compute k-producible bounds only for a relatively small number of parties; however, we can still learn interesting features from these solutions that motivate the search on larger systems under the assumption that these features persist. This allows us to tackle the case where the system size eventually reaches the thermodynamic limit.

I. INTRODUCTION

Assessing entanglement in many-body systems is a fundamental question that requires dedicated tools [1] (see also [2,3]). On the one hand, full tomography on the state is impractical. On the other hand, a practical criterion should be as easy to measure as possible; for instance, it is desirable that such a criterion consists of few-body correlators (where *few* should ideally be *two*). Furthermore, the lower the number of assumptions on the criterion being used, the better. With the advent of the device-independent (DI) formalism for quantum information processing [4], entanglement witnesses have been naturally upgraded to their DI version [5–8]. Recently, we have considered the characterization of entanglement depth in a many-body system from a device-independent perspective [9]. Our criteria are based on the families of Bell inequalities introduced in [10], yielding device-independent witnesses of entanglement depth (DIWEDs). Such DIWEDs naturally inherit desirable properties that allow (i) to compute their k-producibility bounds more efficiently for larger system sizes and (ii) they are experimentally friendlier to measure, as they involve at most two-body correlators and are permutationally invariant.

DIWEDs are stronger than usual entanglement witnesses, in the sense that they are derived without relying on assumptions about the physical systems (such as the dimension of the underlying Hilbert space) nor the measurements that are performed on them [5]. In addition, in the many-body regime,

_____
*jordi.tura@mpq.mpg.de

where a loophole-free Bell test may be very difficult to achieve in practice, the DIWEDs we consider in this work can also be indirectly measured via trusted collective measurements and second moments thereof, opening the way to probe them in situations beyond the current state of the art [11,12].

Determining how many particles share genuinely multipartite entanglement constitutes a good and intuitive measure of the entanglement strength present in the system. While in [9] our aim was to illustrate the key ideas and the applicability of our method, here we analyze the construction of such DIWEDs in detail, providing further case studies. We discuss the different methods to derive k-producible bounds under different sets of assumptions. Interestingly, some of the features that we observe in the smallest systems (where one can make the weakest assumptions) motivate the search on larger system sizes under the assumption that these features persist, eventually reaching the thermodynamic limit. In this work, we also compare the efficiency of our DIWEDs against other witnesses of entanglement depth, witnesses of nonlocality depth and actual experimental data.

The paper is organized as follows. In Sec. II we set the notation and recall the basic concepts this work is built upon. In Sec. III we discuss the optimization problem for finding k-producibility bounds in many-body systems under different families of Bell inequalities, from more general to the case where they are permutationally invariant and consisting of one- and two-body correlators only. In Sec. IV we describe in detail the two complementary methods that we use to characterize the k-producibility bounds on the DIWEDs we present: a variational method and a certificate of optimality based on a partial solution of the quantum marginal problem. In Sec. V

we present two case studies and the numerical results that stem from our work; the first one without extra assumptions and the second one assuming the features that numerics suggest. In Sec. VI we perform an asymptotic analysis that allows us to delve fully into the many-body regime, obtaining the scaling of different k-producibility bounds, where k is a function of the system size. In Sec. VII we discuss how, contrary to existing DIWEDs, the ones we present here can be accessed through experimentally friendly observables, such as collective measurements and moments thereof. In Sec. VIII we compare the entanglement depth criteria we derive in this work with previous existing criteria. Finally, we conclude and discuss further research directions in Sec. IX.

II. PRELIMINARIES

We begin this section by introducing the concepts that will be used throughout all the paper. We review the relevant separability definitions and some natural extension to the multipartite case in Sec. II A. In Sec. II B we review the use of Bell nonlocality for entanglement detection. In particular, in Sec. II B 1 we recall the definition of a Bell experiment in the multipartite case, and then in Sec. II B 2 we provide the basics of Bell inequalities involving two-body correlators.

A. Separability notions

Let us begin by considering a bipartite quantum system shared by Alice (A) and Bob (B), which is described by a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let us denote by $\mathscr{B}(\mathcal{H})$ the set of bounded linear operators acting on $\mathcal{H}$. Then, a mixed state, represented by a density matrix $\rho \in \mathscr{B}(\mathcal{H})$, is a positive-semidefinite operator ($\rho \succeq 0$) of unit trace, $\mathrm{Tr}(\rho) = 1$. A mixed state ρ is said to be separable if it cannot be written in the following form [13]:

$$\rho = \sum_i \lambda_i \rho_i^{(A)} \otimes \rho_i^{(B)}, \tag{1}$$

where λ_i form a convex combination (they are non-negative numbers that sum up to 1). Equation (1) has the following operational interpretation: Alice and Bob, in distant laboratories, can prepare a separable state ρ from scratch just by preagreeing on a common strategy (given by the λ_i's) and preparing locally the corresponding states $\rho_i^{(A)}$ ($\rho_i^{(B)}$) [13,14].

In the multipartite setting, the notion of separability admits many generalizations [15,16]. A natural one is the so-called biseparability. Let us consider a set of n parties, now indexed by $[n] := \{0, \ldots, n-1\}$. Following the above interpretation, a biseparable state is defined as a state that can be produced by allowing the parties to join in two (nonempty) subsets and produce the state in the same fashion. Moreover, the parties are allowed to join in different groups over different realizations, as long as they never join all together. Then, the density matrix that describes the quantum state produced by this procedure is given by

$$\rho = \sum_{A \subset [n], A \neq [n], \emptyset} \lambda_A \sum_i \lambda_i^A \rho_i^{(A)} \otimes \rho_i^{(A^c)}, \tag{2}$$

where $A^c := [n] \setminus A$ is the complement of A in $[n]$ and the λ_A's form convex combinations. This notion can be immediately

generalized (see, e.g., Refs. [15,17]) to that of k separability by simply considering the possible partitions of $[n]$ into k pairwise disjoint, nonempty subsets, i.e., $A^1, \ldots, A^k \subsetneq [n]$ such that $A^i \cap A^j = \emptyset$ if $i \neq j$, and $\cup_{i=1}^k A^i = [n]$ and $A^i \neq \emptyset$. Let us denote by $\mathscr{P}_{\mathrm{sep}}^k$ the set of partitions of $[n]$ into k subsets satisfying such conditions. Then, a k-separable state is described by a density matrix that admits the following decomposition:

$$\rho = \sum_{\mathcal{P} \in \mathscr{P}_{\mathrm{sep}}^k} \lambda_{\mathcal{P}} \sum_i \lambda_i^{\mathcal{P}} \bigotimes_{A^j \in \mathcal{P}} \rho_i^{(A^j)}. \tag{3}$$

As we shall argue later on in the paper, in Sec. V, the definition of biseparability introduced in Eq. (2) [or, more generally, the definition of k separability in Eq. (3)], although being a natural generalization of Eq. (1), is not ideal from an experimental point of view. In fact, the notion of k producibility is much better suited to our work, which we define in what follows.

If, instead of restricting the size of the partitions of $[n]$ to consist of k proper nonempty subsets, we restrict the size of the elements in the partition to contain, at most, k elements, we obtain the definition of k producibility (see, e.g., Refs. [18,19]). Let $|S|$ denote the number of elements in a set S. Let $\mathscr{P}_{\mathrm{prod}}^k$ be the set of partitions $\mathcal{P}$ of $[n]$ satisfying $A_i \subsetneq [n]$, $A_i \cap A_j = \emptyset$ if $i \neq j$, $|A_i| \leqslant k$ and $\cup_{A_i \in \mathcal{P}} A_i = [n]$ for all $A_i \in \mathcal{P}$. Then, a multipartite quantum state is said to be k producible if, and only if, it admits the following decomposition:

$$\rho = \sum_{\mathcal{P} \in \mathscr{P}_{\mathrm{prod}}^k} \lambda_{\mathcal{P}} \sum_i \lambda_i^{\mathcal{P}} \bigotimes_{A_j \in \mathcal{P}} \rho_i^{(A_j)}. \tag{4}$$

A quantum state that cannot be written in the form of Eq. (4) is said to have an entanglement depth at least $k + 1$. Operationally, the interpretation is that strictly more than k parties would need to join together in order to prepare such a state from scratch. We shall denote by $\mathscr{D}_k$ the set of k-producible states [cf. Eq. (4)]. This definition is clearly preferable as k indicates the maximal degree of genuinely multipartite entanglement that the parties may be sharing. In Fig. 1 we schematically outline the differences between such separability notions.

We note that the strongest form of multipartite entanglement, genuinely multipartite entanglement (GME), is given by those states that are not $(n-1)$ producible or, equivalently, not biseparable. Exemplary quantum states that are GME include the GHZ state [20,21] and all entangled symmetric states [22–24].

B. Bell nonlocality for entanglement detection

It is widely known that every state producing nonlocal correlations must be entangled [13]. Therefore, Bell inequalities constitute the natural tool to certify entanglement in a device-independent way. Much work has been devoted to their reinterpretation as device-independent entanglement witnesses [5–7] (see also [8] for a recent complementary approach). On one hand, Bell inequalities can sometimes identify the state producing their maximal violation, thus self-testing the state solely from the statistics it produces in a Bell experiment [25–36]. Therefore, if one manages to self-test a
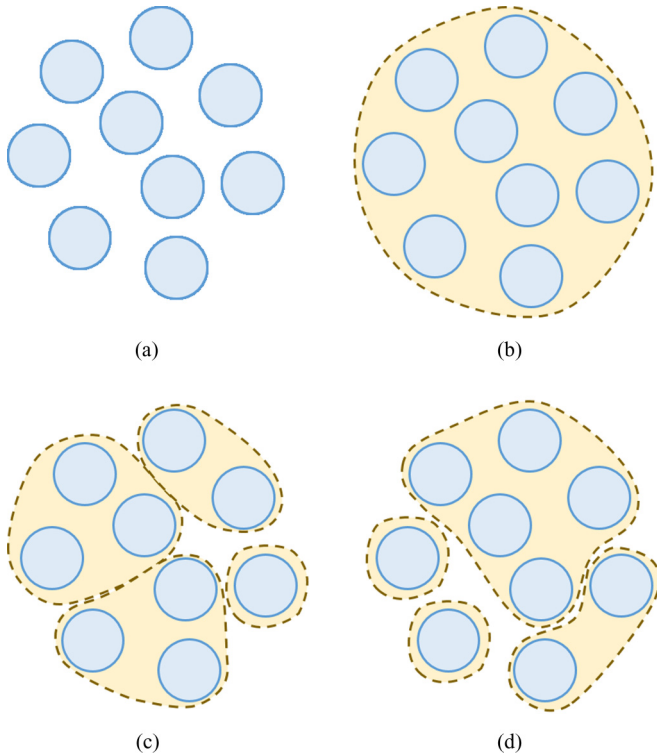
FIG. 1. Cartoon picture illustrating the notions of k producibility and l separability. Operationally, the parties (represented by blue circles) in different groups (represented by shapes enclosing a subset of the parties) are allowed to perform local operations and classical communication between different groups. Within each group, different parties can prepare any state they want. Note that we consider a framework where, at every round, groupings may change [cf. Eq. (3) and Eq. (4)], albeit maintaining the restrictions imposed by k and l. Note that k limits the maximal number of parties in any group, while l limits the total number of groups. (a) Fully separable, one producible; (b) GME, nine producible; (c) four separable, three producible; (d) four separable, five producible.

quantum state of a certain entanglement depth, the goal we pursue in this work is met. However, that only works for very specific states that are known beforehand, as such tests are designed exploiting special properties of the state one wants to self-test.

On the other hand, numerous works have been performed that certify the degree of entanglement present in the state [37–39], often genuine multipartite entanglement [5,40,41]. However, the question of assessing the entanglement depth of a many-body system in a device-independent manner remains broadly unexplored, although remarkable steps have been made towards this direction [6,7], yet still focused on few-body systems.

A natural approach to tackle this question for larger values of the entanglement depth and system sizes consists in designing Bell inequalities for nonlocality depth [42]: broadly speaking, one considers hybrid local-nonlocal models of a form similar to Eq. (3), where parties in the same group are allowed to produce any no-signaling (possibly supraquantum) probability distribution. Analogous to the fact that a quantum state violating an ordinary Bell inequality cannot be fully separable, one exploits the property that a state violating such a nonlocality depth Bell inequality (of depth, say, k) cannot be k producible.

In this work, however, we consider the case where the parties that form the same group can only produce quantum correlations. This difference is conceptually significant and it also poses new technical challenges: the set of quantum correlations is much less characterized than its no-signaling counterpart. For instance, the set of no-signaling correlations forms a polytope and it is therefore characterizable by a finite number of linear inequalities, whereas the set of quantum correlations is not even closed [43] and it has a geometrically much richer—albeit complicated—structure [44]. Despite the increased technical difficulty, constraining the strength of the correlations produced by the parties to be quantum correlations yields DIWEDs which are much tighter than those considered in [42], as we will show in Fig. 8.

1. Bell experiment

Due to the fact that Bell-like inequalities constitute a key element in the derivation of our DIWEDs, in this section we recall the notion of Bell experiment. It involves n spatially separated parties sharing a multipartite resource (e.g., a multipartite entangled state). Each party has m measurements at their disposal and each measurement can yield d different outcomes. Let us denote by $\mathbf{a} = (a_0, \ldots, a_{n-1})$ and $\mathbf{x} = (x_0, \ldots, x_{n-1})$ their outputs (outcomes) and inputs (choices of measurements), respectively. By performing these measurements many times, the probability distributions $p(\mathbf{a}|\mathbf{x})$ are estimated from the collected statistics. The parties are allowed to communicate prior to the experiment, but during every round they cannot know any measurement choice from the rest. Also, the parties do not make any assumptions about the internal working of their devices and only the statistics describing the outcomes, given the inputs, are relevant. We denote such a Bell scenario by (n, m, d).

The no-signaling principle guarantees that the marginal probabilities do not depend on the input of other parties (therefore preventing signaling of information by changing the measurement choice). In particular, in this work, we shall consider only the one- and two-body marginals at sites i and (i, j), respectively, given by

$$p_i(a_i|x_i), \quad p_{ij}(a_i a_j|x_i x_j). \tag{5}$$

Of special interest is the multipartite Bell scenario for which $d = 2$. In this case, we can assume that the outcomes are labeled ± 1 without loss of generality. Let us denote by $M_k^{(i)}$ the correlator corresponding to the expectation value of the kth observable of the ith party ($k \in [m]$, $i \in [n]$), which is defined as

$$M_k^{(i)} = p_i(1|k) - p_i(-1|k). \tag{6}$$

Similarly, we consider the two-body correlator corresponding to the kth and lth observables of the ith and jth parties, respectively, which is given by

$$M_{kl}^{(i,j)} = p_{ij}(a_i a_j = 1|kl) - p_{ij}(a_i a_j = -1|kl). \tag{7}$$

The notion of correlator can be generalized to any number $p \leqslant n$ of parties, yielding the p-body correlators $M_{k_1,\ldots,k_p}^{i_1,\ldots i_p}$,

where $k_j \in [m]$ and $i_j \in [n]$, thus yielding the generic form of a Bell inequality $I - \beta_C \geqslant 0$, where

$$I := \sum_{p=1}^{n} \sum_{k_j \in [m]} \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \alpha_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)} M_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)}, \qquad (8)$$

with $\alpha_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)} \in \mathbb{R}$ and $\beta_C = \min_{\mathrm{LHV}} I$ is the so-called classical bound, i.e., the minimum value that I can take over any local hidden variable (LHV) model. Note that this is equivalent to the correlations that arise from one-producible states; i.e., product states.

2. Bell nonlocality from symmetric two-body correlators

We can now consider the following symmetric one- and two-body correlators, respectively defined as [10]

$$\mathcal{S}_k := \sum_{i \in [n]} M_k^{(i)}, \quad \mathcal{S}_{kl} := \sum_{i \neq j \in [n]} M_{kl}^{(i,j)}. \qquad (9)$$

In this work, we focus on Bell inequalities involving only the correlators introduced in Eq. (9). These inequalities have recently appeared in numerous contexts [9–12,42,45–50] and they can be written as

$$I := \sum_{k \in [m]} \alpha_k \mathcal{S}_k + \sum_{k \leqslant l \in [m]} \alpha_{kl} \mathcal{S}_{kl}, \qquad (10)$$

with $\alpha_k, \alpha_{kl} \in \mathbb{R}$ and that satisfy $I - \beta_C \geqslant 0$ on any local hidden variable theory, where β_C is the so-called classical bound.

By measuring I, one can extract information about the existence of Bell correlations in the system [10,45,48] (as has been performed in recent experiments [11,12]), even on their nonlocality depth [42]. Therefore, we aim at drawing conclusions on the entanglement depth of the system, given the value of I that is measured. To this aim, we want to find a sequence of values β_k such that

$$I - \beta_k \geqslant 0 \qquad (11)$$

is satisfied for every possible measurement on any k-producible state. We then say that Eq. (11) constitutes a DIWED.

III. OPTIMIZATION OVER k-PRODUCIBLE STATES

In this section, we present a general method to find β_k. In Sec. III A we describe the method in its full generality, which applies to any Bell inequality in the $(n, 2, 2)$ scenario. However, the method can be further improved under certain additional hypotheses: for instance, if the inequality consists of few-body correlators or it enjoys additional symmetries. Therefore, in Sec. III B we particularize it for the case of two-body correlators. In Sec. III C we add the permutation invariance property and, motivated by the numerical results we shall present in Sec. V, in Sec. III C we include the additional properties suggested by numerical results.

A. Generic case

Consider an $(n, 2, 2)$ scenario and a Bell inequality I of the form of Eq. (8). We begin by noting that, to every Bell inequality I, one can assign a Bell operator $\mathcal{B} \in \mathscr{B}(\mathcal{H})$, such that

$$\mathcal{B} := \sum_{p=1}^{n} \sum_{k_j \in [m]} \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \alpha_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)} \mathcal{M}_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)}, \qquad (12)$$

where

$$\mathcal{M}_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)} = \bigotimes_{t=1}^{p} \mathcal{M}_{k_t}^{(i_t)}, \qquad (13)$$

and $\mathcal{M}_k^{(i)} \in \mathscr{B}(\mathcal{H}_i)$ denotes the kth observable performed by the ith party. Note that, with a Bell operator defined as Eq. (12), the expectation value of I for a quantum state ρ and local measurements $\mathcal{M}_k^{(i)}$ is automatically given by $I = \mathrm{Tr}[\rho \mathcal{B}]$.

However, observe that the dimension of the local Hilbert spaces $\mathcal{H}_i$ is not fixed. Therefore, in general, one would need to carry the optimization for β_k for Hilbert spaces of increasing dimension (see, e.g., Ref. [51]). This poses, in general, a formidable challenge, not only from the numerical point of view [52,53], as an efficient description of the set of quantum correlations is unlikely to be found (see, e.g., Refs. [54,55]). It is even undecidable to determine whether a linear-system game admits a perfect finite-dimensional quantum strategy or a limit of finite-dimensional quantum strategies [43].

Fortunately, in the $(n, 2, 2)$ case every Bell operator is isometrically equivalent to a Bell operator for qubits ($\mathscr{H}_i = \mathbb{C}^2$) and, furthermore, one can assume without loss of generality that the local measurements $\mathcal{M}_k^{(i)}$ are performed on the XZ plane on the Bloch sphere [56] due to Jordan's lemma [57]. Hence every local measurement can be parametrized as $\mathcal{M}_k^{(i)} = \cos(\theta_{i,k}) \sigma_x^{(i)} + \sin(\theta_{i,k}) \sigma_z^{(i)}$. We shall encode as $\boldsymbol{\theta}$ the vector containing all $\theta_{i,k}$ and we shall therefore denote the Bell operator $\mathcal{B}$ as $\mathcal{B}(\boldsymbol{\theta})$ when we want to make this dependence explicit.

Let us therefore consider $\rho \in \mathscr{D}_k$, with $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$. The optimization of β_k is such that

$$\beta_k = \min_{\rho \in \mathscr{D}_k, \boldsymbol{\theta}} \mathrm{Tr}[\rho \mathcal{B}(\boldsymbol{\theta})]. \qquad (14)$$

Observe that, for every value of $\boldsymbol{\theta}$, one can assume without loss of generality that $\beta_k(\boldsymbol{\theta})$ is attained at a pure state, which corresponds to a fixed partition $\mathcal{P} \in \mathscr{P}_{\mathrm{prod}}^k$ [cf. Eq. (4)]. Therefore, the optimization of β_k can be carried over pure states by fixing their k partition:

$$\beta_k = \min_{\mathcal{P} \in \mathscr{P}_{\mathrm{prod}}^k} \beta_k^{\mathcal{P}}, \qquad (15)$$

where

$$\beta_k^{\mathcal{P}} = \min_{|\Psi_{\mathcal{P}}\rangle, \boldsymbol{\theta}} \langle \Psi_{\mathcal{P}} | \mathcal{B}(\boldsymbol{\theta}) | \Psi_{\mathcal{P}} \rangle \qquad (16)$$

and $|\Psi_{\mathcal{P}}\rangle = \bigotimes_{A \in \mathcal{P}} |\psi_A\rangle$.

In what follows, we show that it is, in principle, possible to find the solutions to the optimization problem Eq. (16) exactly. This implies that we can find β_k without running into local minima issues. To do that, we use arguments that apply to systems of polynomial equations. In particular, for every k-partition $\mathcal{P}$, there is a system of polynomial equations under polynomial equality constraints that defines $\beta_k^{\mathcal{P}}$.

To derive such a system, let us define the variables $x_{i,k} \equiv \cos(\theta_{i,k})$ and $y_{i,k} \equiv \sin(\theta_{i,k})$. Clearly, they are subject to the constraint $x_{i,k}^2 + y_{i,k}^2 = 1$ for every $i \in [n]$ and $k \in [2]$. On the other hand, the quantum state $|\psi_{\mathcal{P}}\rangle$ can be assumed to be real, because $\mathcal{B}$ is a real, symmetric operator . Therefore, we can expand $|\psi_A\rangle = (\psi_0^A, \ldots, \psi_{2^{|A|}-1}^A)^T$ for every $A \in \mathcal{P}$, where $\sum_{i=0}^{2^{|A|}-1} [\psi_i^A]^2 = 1$.

Therefore, $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ is a polynomial of $4n + \sum_{A \in \mathcal{P}} 2^{|A|}$ variables, of degree 2 in the ψ_i^A variables and degree at most n in the $x_{i,k}$ and $y_{i,k}$ variables, subject to $2n + |\mathcal{P}|$ equality constraints. One can determine the critical points of $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ subject to the above constraints using the Lagrange multipliers method. Defining $g_{i,k} = x_{i,k}^2 + y_{i,k}^2 - 1$ and $h_A = \sum_{i=0}^{2^{|A|}-1} [\psi_i^A]^2 - 1$ and associating the dual variables $\lambda_{i,k}, \mu_A$ to each of them, one builds the Lagrangian function

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\psi}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle + \sum_{i,k} \lambda_{i,k} g_{i,k} + \sum_A \mu_A h_A.$$

(17)

In practice, one may want to fix one of the measurements, e.g., $\mathcal{M}_0^{(i)} = \sigma_x$ (equivalently, $x_{i}, 0 = 1$ for all $i \in [n]$), since one can apply a local rotation to the measurements and the inverse rotation to the quantum state without changing the value of $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ [45], thus reducing the number of variables in the optimization.

Equation (17) gives rise to a system of polynomial equations, where one of their solutions (the minimal among the real ones) will correspond to $\beta_k^{\mathcal{P}}$. Bézout's Theorem allows us to upper bound their number of solutions in the generic case [58]. Recall that two algebraic curves of degrees m and n intersect at mn points (which may be real, complex, or at infinity) and cannot meet at more than mn points unless they have a common component, which generically will not be the case.

Finding the solutions to a system of polynomials consisting of p equations and p unknowns has been studied for a long time [59], yet it remains an area of intense research [60]. Since we have $p > 2$ equations, we have to look at generalizations of Bézout's theorem. Generalizations based on the homotopy method (see, e.g., Refs. [61–63]) tell us that, generically, the number of solutions will be the product of the degrees of the p polynomials.

Hence we need to count how many curves of which degree we have. On the one hand, we will have an equation for each variable involved in $\mathcal{L}$, which is $n + n + \sum_{A \in \mathcal{P}} 2^{|A|} + n + |\mathcal{P}| \leqslant 3n + (n/k)2^k + n/k \leqslant n(2^k/k + 4)$. On the other hand, the degree of every polynomial will be upper bounded by the degree of $\mathcal{L}$, which is at most $n + 2$. Hence we obtain an upper bound to the number of solutions

$$(n+2)^{n(2^k/k+4)}.$$

(18)

To find these solutions, one needs to calculate a Gröbner basis [60,64] in order to solve the system of polynomials arising from (17). The computation of a Gröbner basis has, in the worst case, exponential complexity in the number of variables. In very pathological cases, this complexity can even become doubly exponential [65]. However, in practice, most computer algebra routines find them in times nowhere near these complexity bounds [65,66].

To conclude this section, let us briefly comment on the nongeneric case. The most extreme and well-known example we can think of is that of a system of linear equations, where all the polynomial degrees are 1; therefore, one would expect generically one solution according to the above argument. Of course, systems of linear equations with nontrivial kernel arise in a variety of contexts. However, a random ε perturbation of the system will bring it to the generic case. This shows how the underlying structure of polynomial systems is extremely rigid, giving rise in the worst case to the above-mentioned bounds. However, for the purposes of our work, one can always take an ε perturbation of the coefficients $\alpha_{k_1,\ldots,k_p}^{(i_1,\ldots,i_p)}$ of the DIWED, if necessary. Furthermore, the aim of this section is to show that it is possible, in principle, to find exactly the k-producible bound of a DIWED of the form of Eq. (12).

B. Two-body case

In this section, we particularize the method shown in Sec. III A to two-body Bell inequalities, and show that these give an exponential advantage in the number of solutions (18).

Let us begin by writing the Bell operator $\mathcal{B}$ in this case:

$$\mathcal{B} = \sum_{k \in [2]} \sum_{i \in [n]} \alpha_k^{(i)} \mathcal{M}_k^{(i)} + \sum_{k,l \in [2]} \sum_{i \neq j \in [n]} \alpha_{k,l}^{(i,j)} \mathcal{M}_k^{(i)} \otimes \mathcal{M}_l^{(j)}.$$

(19)

When computing the quantity $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ it is now easy to see that each of the terms in $\mathcal{B}$ has support onto one or two of the states $|\psi_A\rangle$ that form $|\Psi_{\mathcal{P}}\rangle$. Therefore, expanding Eq. (19) yields in this case a polynomial of much lower degree, albeit with the same number of variables. In particular, the degree of $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ is now at most 2 in the $\boldsymbol{x}, \boldsymbol{y}$ variables. Hence all the equations stemming from Eq. (17) have degree at most 4, which yields an exponential improvement in the number of solutions, namely

$$4^{n(2^k/k+4)}.$$

(20)

Therefore, if the Bell operator consists of two-body correlators, $\langle \Psi_{\mathcal{P}} | \mathcal{B} | \Psi_{\mathcal{P}} \rangle$ can now be computed efficiently for a constant k and large n, as it requires $O(2^{2k}n^2) = O(n^2)$ steps. In addition, note that in this case we can mildly relax the assumption that k is a constant and still preserve the polynomial complexity of the algorithm: as long as $k \in O(\log n)$, the overall evaluation time will still be polynomial in n.

C. Two-body symmetric case

In this section, we add the additional requirement that the Bell inequality I should be permutationally invariant and built from one- and two-body correlators [10]. We shall see that the expectation value of the Bell operator can be computed in this case in a much more efficient way. Motivated by the numerical results in Sec. V, in Sec. III C we also study the case in which all parties in the same region $A \in \mathcal{P}$ measure the same set of measurements, and when the state $|\psi_A\rangle$ is symmetric.

In the case of a permutationally invariant two-body Bell inequality, the coefficients $\alpha_k^{(i)}$ and $\alpha_{kl}^{(i,j)}$ in Eq. (19) do not depend on i or j. Hence the Bell operator can be expressed as

$$\mathcal{B} = \sum_{k \in [2]} \alpha_k \sum_{i \in [n]} \mathcal{M}_k^{(i)} + \sum_{k \leqslant l \in [2]} \alpha_{kl} \sum_{i \neq j \in [n]} \mathcal{M}_k^{(i)} \otimes \mathcal{M}_l^{(j)}. \quad (21)$$
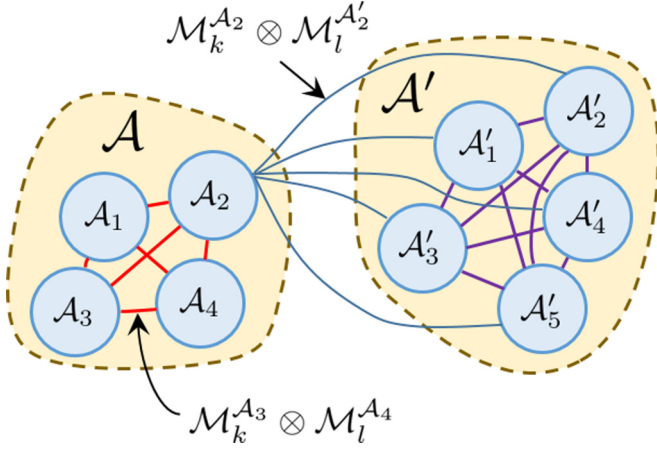
FIG. 2. Interpretation of Eq. (22). We have represented two groups of parties, $\mathcal{A}$ and $\mathcal{A}'$, forming a partition $\mathcal{P} \in \mathscr{P}^k_{\text{prod}}$ in the yellow shapes with discontinuous contour. Within each group there are several parties represented by the blue circles. The one-body terms of Eq. (22) are not represented, as they correspond to each blue circle (party). The two-body terms are divided in two parts: those that involve two parties belonging to the same group (red lines for $\mathcal{A}$, purple lines for $\mathcal{A}'$) and those that involve two parties that belong to different groups (here, for clarity of the figure, we have only depicted the terms that connect $\mathcal{A}_2$ from $\mathcal{A}$ to any party in $\mathcal{A}'$, but of course there would be many more lines crossing different regions). Since the state $|\Psi_\mathcal{P}\rangle$ is product with respect to different regions, expectation values for the lines that connect different regions will factorize.

Its expectation value over $|\Psi_\mathcal{P}\rangle$ can be computed as

$\langle\Psi_\mathcal{P}|\,\mathcal{B}\,|\Psi_\mathcal{P}\rangle$

$$= \sum_{A\in\mathcal{P}}\left(\sum_k \alpha_k \underbrace{\langle\psi_A|\mathcal{B}^A_k|\psi_A\rangle}_{\text{one-body terms}} + \sum_{k\leqslant l}\alpha_{kl}\underbrace{\langle\psi_A|\mathcal{B}^A_{kl}|\psi_A\rangle}_{\text{same region terms}}\right)$$
$$+ \sum_{A\neq A'\in\mathcal{P}}\left(\sum_{k\leqslant l}\alpha_{kl}\underbrace{\langle\psi_A|\mathcal{B}^A_k|\psi_A\rangle\langle\psi_{A'}|\mathcal{B}^{A'}_l|\psi_{A'}\rangle}_{\text{crossed region terms}}\right), \quad (22)$$

where the terms of the Bell operator have been grouped in the following way:

$$\mathcal{B}^A_k := \sum_{i\in A}\mathcal{M}^{(i)}_k, \quad \mathcal{B}^A_{kl} := \sum_{i\in A, i'\in A\setminus\{i\}}\mathcal{M}^{(i)}_k\otimes\mathcal{M}^{(i')}_l. \quad (23)$$

In Fig. 2 we graphically represent Eq. (22).

Optimization over symmetric regions

In this section, we are interested in finding an expression for the quantity $\langle\Psi_\mathcal{P}|\,\mathcal{B}\,|\Psi_\mathcal{P}\rangle$ under the assumption that the Bell operator consists of two-body correlators and it is permutationally invariant on each region $A\in\mathcal{P}$. The last assumption is motivated by the numerical results that we shall present in Sec. IV A.

It is well known that a permutationally invariant operator $\mathcal{B}$ acting on a Hilbert space $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$ can be block diagonalized into invariant subspaces that correspond to different spin components with their corresponding multiplicities, a result also known as Schur-Weyl duality. For a permutationally

invariant operator of n qudits, the size of these blocks is $O\left(\binom{n+d-1}{d-1}\right)$ [67–71].

For instance, consider an operator $\mathcal{B} \in \mathscr{B}((\mathbb{C}^2)^{\otimes 2})$ acting on a two qubit space, which is permutationally invariant. Then, $\mathcal{B}$ is block diagonalized in the basis spanned by the Dicke states and the antisymmetric Bell state, $V^\dagger\mathcal{B}V = \mathcal{B}_{3\times 3}\oplus\mathcal{B}_{1\times 1}$:

$$V^\dagger\mathcal{B}V = \left(\begin{array}{c|c}\mathcal{B}_{3\times 3} & 0 \\ \hline 0 & \mathcal{B}_{1\times 1}\end{array}\right),$$

$$V = \begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 0 & 1 & 0\end{pmatrix}. \quad (24)$$

Let us recover the form of the Bell operator, from Eq. (22),

$$\mathcal{B} = \sum_{A\in\mathcal{P}}\left(\sum_k\alpha_k\mathcal{B}^A_k + \sum_{k\leqslant l}\alpha_{kl}\mathcal{B}^A_{kl}\right)$$
$$+ \sum_{A\neq A'\in\mathcal{P}}\sum_{k\leqslant l}\alpha_{kl}\mathcal{B}^A_k\otimes\mathcal{B}^{A'}_l. \quad (25)$$

Let V_A be a matrix consisting of the Schur-Weyl basis vectors that block diagonalize the Ath block arranged in columns. Then, by denoting $V = \bigotimes_{A\in\mathcal{P}}V_A$, we have that

$$V^\dagger\mathcal{B}V = \sum_{A\in\mathcal{P}}\left(\sum_k\alpha_k V^\dagger_A\mathcal{B}^A_k V_A + \sum_{k\leqslant l}\alpha_{kl}V^\dagger_A\mathcal{B}^A_{kl}V_A\right)$$
$$+ \sum_{A\neq A'\in\mathcal{P}}\sum_{k\leqslant l}\alpha_{kl}\left(V^\dagger_A\mathcal{B}^A_k V_A\right)\otimes\left(V^\dagger_{A'}\mathcal{B}^{A'}_l V_{A'}\right). \quad (26)$$

Note that $V_A V^\dagger_A = V^\dagger_A V_A = \mathbb{1}_{\mathcal{H}_A}$. Therefore, any k-producible pure state $|\Psi\rangle$ is expressed as $V^\dagger|\Psi\rangle$ in the tensor product basis of the blocks.

It is of particular interest the case where $|\psi_A\rangle$ is fully supported on an invariant subspace of V_A (e.g., the symmetric space of A). In this case, $|\psi_A\rangle$ is described by, at most, $|A|$ real parameters. Therefore, the number of local minima is upper bounded by

$$4^{3|\mathcal{P}|+n/k\cdot k} \leqslant 4^{n(1+3/k)}. \quad (27)$$

If, in addition, all regions $|A|$ have the same cardinal k, and $|\psi_A\rangle = |\psi_{A'}\rangle$ for every $A, A'\in\mathcal{P}$, then the above bound is dramatically reduced to

$$4^{4+k}. \quad (28)$$

This last case is of particular importance in the asymptotic analysis developed in Sec. VI.

IV. METHODS

In this section we describe the two main methods that allow us to construct the DIWEDs defined in Eq. (11) from a particular two-body permutationally invariant Bell inequality of the form given in Eq. (10). In Sec. IV A we give a variational approach that allows one to find a good upper bound β^U_k to

β_k. In Sec. IV B we construct a certificate that provides a lower bound β_k^L to β_k, thus complementing the bound given in Sec. IV A. Hence $\beta_k^L \leqslant \beta_k \leqslant \beta_k^U$.

A. Variational upper bound

To find an upper bound to the optimization given by Eq. (16), it is sufficient to find a k-producible vector $|\Psi\rangle$ and a set of measurements $\mathcal{M}_k^{(i)}$. Indeed, any expectation value provided by a k-producible state on some Bell operator upon performing any set of measurements leads, by construction, to an upper bound to β_k.

Our aim here is to describe in detail a method that gives a good guess to the optimal solution β_k, and we do so by constructing a k-producible pure state and finding a suitable set of measurements. The steps of the algorithm are as follows.

(1) For every partition $\mathcal{P} \in \mathscr{P}_{\text{prod}}^k$,

(2) pick a random k-producible pure state $|\Psi\rangle$ and random initial measurement settings, parametrized by $\boldsymbol{\theta}$.

(3) Use a seesaw iteration method (see, e.g., Refs. [51,72]) to improve the choice of $\boldsymbol{\theta}$, while keeping $|\Psi\rangle$ fixed.

The seesaw iteration technique consists in doing several sweeps across $[n]$ and optimizing the measurement parameters one site at a time. Thus the seesaw iteration fixes the measurements $\mathcal{M}_k^{(i)}$ of all parties except for the jth one. Then, one can write a cost function

$$E_j(\boldsymbol{\theta}_j) = \langle\Psi| \mathcal{B}(\boldsymbol{\theta}_j; \boldsymbol{\theta}_0, \ldots, \hat{\boldsymbol{\theta}}_j, \ldots, \boldsymbol{\theta}_{n-1}) |\Psi\rangle \qquad (29)$$

that only depends on $\boldsymbol{\theta}_j$, while the rest of them ($\boldsymbol{\theta}_i$, $i \neq j$) simply become parameters describing the fixed $\mathcal{M}_k^{(i)}$.

Finding $\boldsymbol{\theta}_j$ can be done using stochastic gradient descent methods. However, we note that, since Eq. (29) is linear in $\mathcal{M}_k^{(j)}$, it can be expressed as a semidefinite program whose variables are the positive-operator valued measure (POVM) elements of $\mathcal{M}_k^{(j)}$, thus obtaining the minimum of Eq. (29) in one iteration:

$$
\begin{aligned}
\min_{\Pi_0^{(k)}} \quad & \text{Tr}[|\psi\rangle\langle\psi| \mathcal{B}(\boldsymbol{\theta})] \\
\text{such that} \quad & 0 \preccurlyeq \Pi_0^{(k)} \preccurlyeq \mathbb{1}, \\
& \mathcal{M}_k^{(j)} = \Pi_0^{(k)} - \Pi_1^{(k)}, \\
& \Pi_0^{(k)} + \Pi_1^{(k)} = \mathbb{1}.
\end{aligned}
\qquad (30)
$$

We keep the value of the measurements stemming from Eq. (30) as the updated value of $\boldsymbol{\theta}_j$, and we move to the next value of j. Note that the values of $\langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$ form at each step a monotonically decreasing function, thus converging to a local minimum. We go to step (4) once the sequence has converged within numerical accuracy.

Alternatively, we note that the optimal choice of $\boldsymbol{\theta}$ minimizing $\langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$ with $|\psi\rangle$ fixed can be also found using stochastic gradient descent techniques. However, we discourage its use, as the number of free parameters grows with n resulting in a worse performance than the seesaw approach, and a much higher sensitivity to the initial point $\boldsymbol{\theta}$.

(4) In this step we find a k-producible state that improves the value of $\langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$, while keeping the measurement settings, parametrized by $\boldsymbol{\theta}$, fixed. Like in the previous step, a seesaw iteration method is possible, by fixing the value of all $|\psi_i\rangle$, except for $i = j$, where j enumerates the elements of $\mathcal{P}$. Note that, in this case, Eq. (25) enables us to find $|\psi_j\rangle$ as

the eigenvector of minimal eigenvalue of

$$
\begin{aligned}
\tilde{\mathcal{B}}_j = {} & \sum_k \alpha_k \mathcal{B}_k^{A_j} + \sum_{k \leqslant l} \alpha_{kl} \mathcal{B}_{kl}^{A_j} \\
& + \sum_{k \leqslant l} \alpha_{kl} \sum_{A_i \neq A_j \in \mathcal{P}} \langle \mathcal{B}_k^{A_i} \rangle \mathcal{B}_l^{A_j} + \mathcal{B}_k^{A_j} \langle \mathcal{B}_l^{A_i} \rangle. \qquad (31)
\end{aligned}
$$

Like in step (3), the value of j is iterated back and forth until $\langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$ converges within numerical accuracy.

(5) Repeat steps (3) and (4) until the value of $\langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$ lies within numerical accuracy; i.e., the result from both steps (3) and (4) lie already within numerical accuracy.

(6) Define $\beta_k^U(\mathcal{P}) = \langle\Psi| \mathcal{B}(\boldsymbol{\theta}) |\Psi\rangle$. If there are partitions left, go to step (1).

We note that it is not necessary to explore all $\mathcal{P} \in \mathscr{P}_{\text{prod}}^k$, as one can define a partial order in $\mathscr{P}_{\text{prod}}^k$ induced by inclusion: formally, for every $\mathcal{P}, \mathcal{Q} \in \mathscr{P}_{\text{prod}}^k$, we say that $\mathcal{P} \preccurlyeq \mathcal{Q}$ if, and only if, for every $A \in \mathcal{P}$, there exists a $B \in \mathcal{Q}$ such that $A \subseteq B$. Then it is clear that $\beta_k^U(\mathcal{P}) \geqslant \beta_k^U(\mathcal{Q})$ if $\mathcal{P} \preccurlyeq \mathcal{Q}$.

Hence, since $\mathscr{P}_{\text{prod}}^k$ is a poset, it is sufficient to pick $\mathcal{P}$ in step (1) from its minimal elements.

(7) Output $\beta_k^U = \min_{\mathcal{P}} \beta_k^U(\mathcal{P})$.

As the end result of the algorithm, we obtain a value $\beta_k^U \geqslant \beta_k$. Thanks to Jordan's lemma we do not need to increase the Hilbert space dimension of the states, nor consider measurements outside of the XZ plane. Recall, however, that the seesaw optimization method, although it allows one to tackle multilateral optimization instances by splitting them into several unilateral approaches, therefore giving fast convergence rates, is a greedy algorithm which can be stuck at local minima. This depends highly on the landscape of the objective function and the number of local minima it has. In Sec. III we have discussed some of the worst-case bounds for the number of local minima, which gives an indication on the number of initial points necessary to ensure convergence with high probability.

B. Lower bound certificate

Here we provide a complementary method to the one introduced in Sec. IV A to lower bound β_k by means of a semidefinite program based on a relaxation of the quantum marginal problem.

Recall that the optimal value for β_k is achieved at a certain partition $\mathcal{P}$, which we shall assume from here on to be fixed (we simply run the method for every minimal element of $\mathscr{P}_{\text{prod}}^k$, as in Sec. IV A, and keep the best value). Let us recall that, given an arbitrary multipartite quantum state $\rho \in \mathscr{B}(\mathcal{H})$, its expectation value on the Bell operator can be computed as [cf. Eq. (25)]

$$
\begin{aligned}
& \text{Tr}[\mathcal{B}(\boldsymbol{\theta})\rho] \\
& = \sum_{A_j \in \mathcal{P}} \left(\sum_k \alpha_k \text{Tr}[\mathcal{B}_k^{A_j} \rho_{A_j}] + \sum_{k \leqslant l} \alpha_{kl} \text{Tr}[\mathcal{B}_{kl}^{A_j} \rho_{A_j}] \right) \\
& \quad + \sum_{A_j \neq A_{j'} \in \mathcal{P}} \sum_{k \leqslant l} \alpha_{kl} [\mathcal{B}_k^{A_j} \otimes \mathcal{B}_l^{A_{j'}} \rho_{A_j \cup A_{j'}}], \qquad (32)
\end{aligned}
$$

where ρ_A is the reduced state of ρ on the parties that constitute $A \subseteq [n]$.

Furthermore, ρ is a k-separable state by assumption. Therefore, since $\mathcal{P}$ is fixed, ρ must be separable across every cut $A_j | A_{j'}$. Therefore, we can find a lower bound to β_k by considering the following problem:

$$\min \mathrm{Tr}[\mathcal{B}(\boldsymbol{\theta})\rho]$$
$$\text{such that } \rho_{A_j} \succeq 0, \ \rho_{A_j \cup A_{j'}} \succeq 0,$$
$$\mathrm{Tr}[\rho_{A_j}] = \mathrm{Tr}[\rho_{A_j \cup A_{j'}}] = 1,$$
$$\mathrm{Tr}_{A_j}[\rho_{A_j \cup A_{j'}}] = \rho_{A_j},$$
$$\rho_{A_j \cup A_{j'}} \text{ is separable across } A_j | A_{j'}. \quad (33)$$

The optimization presented in Eq. (33) is almost a semidefinite problem. Unfortunately, it is *NP* hard to determine if a mixed quantum state is entangled or separable [73]. However, the problem can be brought to a semidefinite programming form by allowing the feasible set to be an outer approximation to that of Eq. (33).

An outer approximation to the set of separable states that behaves particularly well in semidefinite programming instances is the so-called positivity under partial transposition (PPT) [74]. Recall that a separable state ρ can be expressed according to Eq. (1). Then, one can define its PPT version, which is the result of the map $(\mathbb{1} \otimes T)[\rho] =: \rho^\Gamma$, where T is the transposition operator with respect to the computational basis. If ρ is of the form of Eq. (1), then $\rho^\Gamma \succeq 0$.

Hence we can now consider the relaxation of Eq. (33) in its semidefinite programing form:

$$\min \mathrm{Tr}[\mathcal{B}(\boldsymbol{\theta})\rho]$$
$$\text{such that } \rho_{A_j} \succeq 0, \ \rho_{A_j \cup A_{j'}} \succeq 0,$$
$$\mathrm{Tr}[\rho_{A_j}] = \mathrm{Tr}[\rho_{A_j \cup A_{j'}}] = 1,$$
$$\mathrm{Tr}_{A_j}[\rho_{A_j \cup A_{j'}}] = \rho_{A_j},$$
$$(\rho_{A_j \cup A_{j'}})^\Gamma \succeq 0. \quad (34)$$

The relaxation arising from Eq. (34) can be tightened by considering stronger, yet numerically more expensive, approximations to the separable set [75,76]. However, we shall see in Sec. V indications why it is not necessary to do so for inequalities of the form Eq. (10) with $m = 2$.

The convergence of Eq. (34) guarantees that the set of marginals with which we compute $\mathrm{Tr}[\mathcal{B}(\boldsymbol{\theta})\rho]$ is actually a minimum. Although strong duality need not hold in general for semidefinite programs, the cases where it does not hold are quite pathological [77]; weak duality (i.e., a feasible solution to the dual) would still give a lower bound to the minimum. Check hence, for each choice of measurement settings, described by $\boldsymbol{\theta}$, we obtain a value $\beta_k^L(\boldsymbol{\theta}) \leqslant \beta_k^U(\boldsymbol{\theta})$. In particular, for the optimal choice of $\boldsymbol{\theta}$, we get a lower bound $\beta_k^L \leqslant \beta_k$.

By running also a seesaw optimization on $\boldsymbol{\theta}$, we find that the optimal value of $\beta_k^L(\boldsymbol{\theta})$ is obtained by the same value of $\boldsymbol{\theta}$ that yields the optimal value $\beta_k^U(\boldsymbol{\theta})$ whenever the initial value of $\boldsymbol{\theta}$ is close to the one yielding $\beta_k^U(\boldsymbol{\theta})$. Note that for values of $\boldsymbol{\theta}$ that are too far away from the optimal measurement parameters, it could happen that $\beta_k^L(\boldsymbol{\theta}') > \beta_k^U(\boldsymbol{\theta})$ (for instance, if $\boldsymbol{\theta}'$ corresponds to commuting observables, one can only produce the classical bound of the Bell inequality). However,

one can discard all the measurement parameters yielding such a contradiction to also rapidly reduce the parameter space, in a similar spirit in which branch and bound methods operate: for instance, any $\boldsymbol{\theta}'$ giving a lower bound certificate strictly above any variational solution $\beta_k^U(\boldsymbol{\theta})$ automatically discards $\boldsymbol{\theta}'$ [78]. Additionally, for the lower bound certificate, which uses the PPT condition as a key ingredient to relax the separability condition, a possible way to strengthen it (in cases where β_k^L might not be sufficiently close to β_k^U) would be to consider tests that also rule out PPT-bound entanglement. For instance, it is known that mutually unbiased bases with internal permutations lead to entanglement witnesses that detect bound entanglement[36,79–81].

V. NUMERICAL RESULTS

In this section, we present the numerical results we have calculated, obtaining DIWEDs for different PIBIs and number of parties. In Sec. V A we discuss the results obtained without any additional hypotheses, and we extract some properties that we conjecture are likely to be preserved in the many-body regime. In Sec. V B we perform the numerical analysis with these additional assumptions.

A. Unconstrained optimization

We have performed numerical studies up to $n = 10$ parties using the seesaw method outlined in Sec. IV A with the following inequalities [10]:

$$-2\mathcal{S}_0 + \tfrac{1}{2}\mathcal{S}_{00} - \mathcal{S}_{01} + \tfrac{1}{2}\mathcal{S}_{11} \geqslant \beta_k \quad (35)$$

and

$$(n \bmod 2)(n-1)(n\mathcal{S}_0 + \mathcal{S}_1) + \binom{n}{2}\mathcal{S}_{00} + n\mathcal{S}_{01} - \mathcal{S}_{11} \geqslant \beta_k. \quad (36)$$

The reason for these choices is the following.

Experiments have already been performed with Eq. (35) [11,12]. Therefore, we can apply our method to draw conclusions about the entanglement depth that was present in such experiments. Concerning Eq. (36), it is an inequality that is tailored to the half-filled Dicke state [45], therefore making it potentially relevant for experimental settings in which such states are prepared (e.g., via spin-changing collisions [82]) and it is a generalization of CHSH [83] when $n = 2$ is used.

To find β_k for Eq. (35) and Eq. (36), we have represented the quantum state in the computational basis (storing the whole state vector) and we have added no constraint among measurement settings, except setting $\mathcal{M}_0^{(i)} = \sigma_z$ for every $i \in [n]$. Therefore, the relevant measurement parameters are θ_i for each $i \in [n]$, thus parametrizing $\mathcal{M}_1^{(i)} = \cos\theta_i \sigma_z + \sin\theta_i \sigma_x$. This assumption incurs in no loss of generality, since the measurements of the ith party can be simultaneously rotated to any measurements $\tilde{\mathcal{M}}_0^{(i)} = \cos(c_i)\sigma_z + \sin(c_i)\sigma_x$ and $\tilde{\mathcal{M}}_1^{(i)} = \cos(c_i + \theta_i)\sigma_z + \sin(c_i + \theta_i)\sigma_x$ by means of a local unitary rotation of angle c_i. This comes at the expense of rotating the state with a single-qubit rotation of angle $-c_i$ applied to the ith site, but this does not change the entanglement-depth properties of $|\Psi_{\mathcal{P}}\rangle$ [45].

We have run the seesaw method with more than 10^4 initial measurement settings, uniformly distributed in the $[0, 2\pi]^n$ hypercube, and k-producible states, where the initial state at each region is chosen uniformly at random according to the Haar measure.

We observe the following behavior. For every partition $\mathcal{P}$ which is a maximal element of $\mathscr{P}_{\text{prod}}^k$ with respect to the partial order $\preccurlyeq$, the optimal point is reached when the state at each region is permutationally invariant (in particular, it is a superposition of Dicke states), and the parties at each region perform the same measurement settings. The measurement settings among different regions, however, are not always the same, but they differ less if the size of these regions is similar.

Both the variational method and the complementary semidefinite programming certificate lead to the same results, within numerical accuracy; i.e., $\beta_k^U - \beta_k^L < 10^{-7}$. The numerical accuracy limitation is dominated by the accuracy of the SDP solver, which is of the order of 10^{-7}. We have tried our methods with a variety of SDP solvers, including SeDuMi [84], Mosek [85], and SDPT3 [86], with CVX [87,88] and Yalmip [89] as parsers, all leading to similar results.

It is worth noting that our results show a similar behavior to that observed in [10,45], where the optimal point is reached when all the parties in the same region use the same θ_i. The case treated in [10,45] corresponds to the trivial partition $\mathcal{P} = \{[n]\}$; therefore, our method recovers the original result.

B. Optimization under additional hypotheses

Motivated by the behavior observed in Sec. V A, we have performed further numerical studies for $n, k \gg 10$, under the following assumptions.

(i) The optimal measurement settings leading to β_k are, modulo local unitary rotations, the same on each region $A \in \mathcal{P}$.

(ii) The optimal state is of the form $|\Psi_{\mathcal{P}}\rangle = |\psi_{A_1}\rangle \otimes \cdots \otimes |\psi_{A_{|\mathcal{P}|}}\rangle$, where each $|\psi_A\rangle$ is a superposition of symmetric states of the same spin length. In the case of Eq. (35) and Eq. (36), this corresponds to states of the maximal spin length (Dicke states). These superpositions involve only real coefficients; therefore, they can be described exactly with $|A|$ real parameters (including normalization).

In the case of Eq. (35), we observe that, for k-producible bounds of large k, the state $|\psi_A\rangle$ can be well approximated (after suitable local unitary rotations [45]) by a superposition of Dicke states of the form

$$|\psi_A\rangle \approx \sum_{j=0}^{|A|} c_j \left|D_{|A|}^j\right\rangle, \tag{37}$$

where $c_j = e^{-(j-\mu_A)^2/4\sigma_A}/\sqrt[4]{2\pi\sigma_A}$ and $|D_m^j\rangle$ is the Dicke state of m qubits with j excitations.

We believe this behavior arises due to the following reasons: Eq. (35) is maximally violated by this kind of state, and the maximal quantum violation, normalized to the classical bound, is monotonically increasing in magnitude, tending to a constant value. Therefore, the optimization to find β_k favors
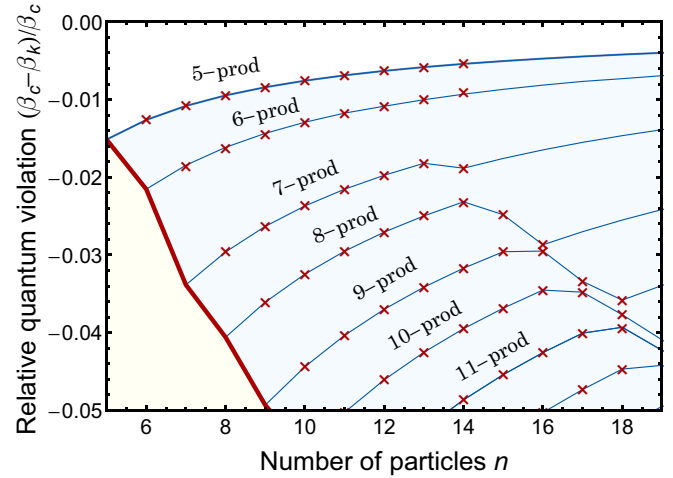

FIG. 3. DIWED bounds for the two-body PIBI presented in Eq. (35) up to 20 particles. Each line represents a k-producible bound. The wavylike behavior of the bounds comes from the fact that Eq. (35) has no quantum violation for less than five parties [10]. Therefore, the optimal partition $\mathcal{P}$ for every (n, k) tries to avoid forming groups of four parties or less. The lines correspond to the variational solution (connected for better visibility) and the crosses to the lower bound certificate obtained via a semidefinite program. We see that, for practical effects, both values always coincide.

this family of states on each region, which become better as the size of the region increases. In addition, Eq. (35) has no quantum violation for less than five parties. This gives rise to the wavelike behavior for small values of n in Fig. (3) (see also [9]). Furthermore, we see that, due to this reason, larger regions give the optimal value of β_k, for which we conjecture that the most balanced partition (every $A \in \mathcal{P}$ has k elements, except perhaps for one if n is not a multiple of k) is the one that produces β_k. We analyze the asymptotic behavior of the DIWED arising from Eq. (35) under this assumption in Sec. VI.
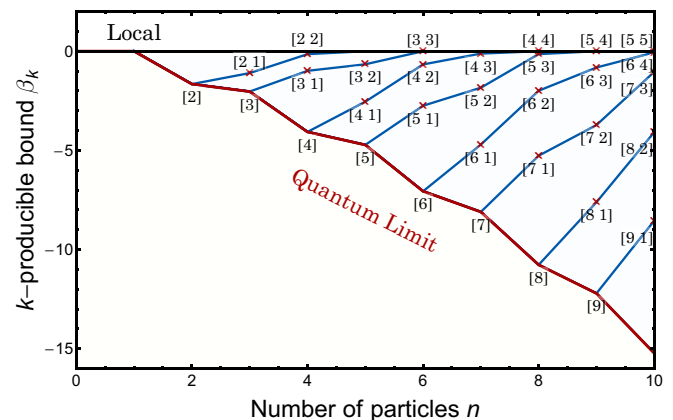

FIG. 4. Entanglement depth bounds for inequality (36) up to 10 particles. Each line represents a k-producible bound and each cross the corresponding lower-bound certificate. We also label the partition yielding β_k next to the corresponding point in the plot. In some cases, different partitions yield the same bound, for instance, {7, 2} and {7, 1, 1}. We only label one of them for clarity. We observe that both methods converge up to numerical precision.
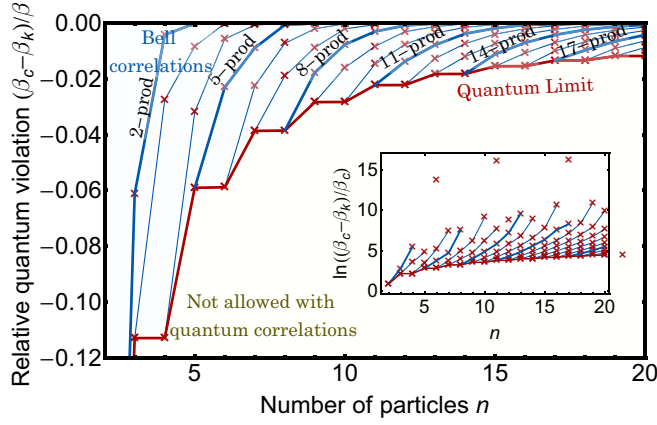
FIG. 5. Entanglement depth bounds for inequality (36) up to 20 particles relative to the classical bound. The inset shows how the precision required to certify entanglement depth increases with the number of particles, as the relative quantum limit approaches zero with the system size. In some cases, the variational method would give zero, whereas the lower bound certificate would yield a value close to zero with negative sign, due to numerical precision. This explains the three isolated points in the inset.

In the case of Eq. (36), its behavior is shown in Fig. (4), and normalized to the classical bound in Fig. (5). In this case, we see that the optimal partition is more difficult to predict, which we attribute to the fact that the relative quantum violation of Eq. (36) decreases with the system size [10,45], although the optimal state tends asymptotically to the half-filled Dicke state $|D_n^{\lceil n/2 \rceil}\rangle$.

VI. ASYMPTOTIC ANALYSIS

In this section we perform an asymptotic analysis of the DIWED arising from Eq. (35), in the limit of large n with the following assumptions, motivated by our findings in Sec. V.

(i) The optimal partition $\mathcal{P} \in \mathscr{P}_{\text{prod}}^k$ is given by $\lfloor k/n \rfloor$ groups of k parties, plus a smaller partition of $n \mod k$ elements, if k does not divide n.

(ii) The optimal measurement settings on each group of parties are the same for each region (up to local unitaries).

(iii) The measurement settings on different regions converge to the same (up to local unitaries) as the system becomes bigger.

Let us remark that the DIWED arising from Eq. (35) does not detect entanglement depths for $k < 5$ due to the nature of Eq. (35), which has no quantum violation for less than five parties. Moreover, the contribution of the smaller group fades away as n grows, so we will consider k to be a multiple of n for large values of n. Therefore, we define $m = |\mathcal{P}| = n/k$. Note that there is no loss of generality in ignoring the asymptotic analysis of the cases where m does not divide n; this has also been observed in other works (see, e.g., Ref. [90]).

Since we assume all parties to perform the same set of measurements, we can write $\boldsymbol{\theta} \equiv (\varphi, \theta)$, which parametrize the measurement operators as $\mathcal{M}_0^{(i)} = \cos \varphi \sigma_z^{(i)} + \sin \varphi \sigma_x^{(i)}$ and $\mathcal{M}_1^{(i)} = \cos \theta \sigma_z^{(i)} + \sin \theta \sigma_x^{(i)}$. There will be a value of $\theta - \varphi$ in which the state at each region can be parametrized

as Eq. (37), with $\sigma_A =: \sigma$ and $\mu_A =: \mu$ being independent of $A \in \mathcal{P}$.

Let us begin by noting the following decomposition of the expectation value of a two-body Bell operator of the form of Eq. (19) on a k-producible state $|\Psi_\mathcal{P}\rangle$:

$$\langle \mathcal{B} \rangle = \sum_{A \in \mathcal{P}} \langle \mathcal{B}^A \rangle + \sum_{A \neq A' \in \mathcal{P}} \frac{\gamma}{2} \langle \mathcal{S}_0^A \rangle \langle \mathcal{S}_0^{A'} \rangle + \delta \langle \mathcal{S}_0^A \rangle \langle \mathcal{S}_1^{A'} \rangle + \frac{\varepsilon}{2} \langle \mathcal{S}_1^A \rangle \langle \mathcal{S}_1^{A'} \rangle, \tag{38}$$

where $\mathcal{S}_k^A$ and $\mathcal{B}^A$ are the respective restrictions of $\mathcal{S}_k$ and $\mathcal{B}$ on A.

In particular, for the particular Bell inequality Eq. (35) we have

$$\langle \mathcal{B} \rangle = \sum_{A \in \mathcal{P}} \langle \mathcal{B}^A \rangle + \frac{1}{2} \sum_{A \neq A' \in \mathcal{P}} \left(\langle \mathcal{S}_0^A \rangle - \langle \mathcal{S}_1^A \rangle \right) \left(\langle \mathcal{S}_0^{A'} \rangle - \langle \mathcal{S}_1^{A'} \rangle \right). \tag{39}$$

Now we need to recall two results from [45].

(i) The quantum limit β_n of Eq. (35) for a system of n parties behaves asymptotically as

$$\beta_n = -\frac{5}{2} n + \frac{\sqrt{3}}{2} n^{1/2} - \frac{3}{2} + o(1). \tag{40}$$

The derivation of Eq. (40) uses a refinement from the approximation made in [45]. Here we use that the σ parameter in Eq. (37) can be taken as $\sigma = \sqrt{n/48}$ for large n and not just $\Theta(\sqrt{n})$. The details can be found in Appendix.

(ii) The one-body reduced density matrix corresponding to the optimal state can be approximated as

$$\rho_1 = \frac{n}{(n-1)} |+\rangle \langle +| + \begin{pmatrix} \frac{-1-2c}{2(n-1)} & 0 \\ 0 & \frac{2c-1}{2(n-1)} \end{pmatrix} + o(1), \tag{41}$$

where $c = \mu - n/2 = 1/(4 \cos \theta)$ and $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$.

It is now clear how to approximate the k-producible bound for large n and k, by combining Eq. (39) with Eq. (40) and Eq. (41). For the sake of clarity, we use the following measurement parameters: $\varphi = \pi - \theta$ and $\theta = 5\pi/6$.

On the one hand, the expectation value $\langle \mathcal{B}^A \rangle$ on a region of size k will asymptotically converge to

$$\langle \mathcal{B}^A \rangle = -\frac{5}{2} k + \frac{\sqrt{3}}{2} k^{1/2} - \frac{3}{2} + o(1). \tag{42}$$

On the other hand, the expectation values of $\langle \mathcal{S}_0^A \rangle$ and $\langle \mathcal{S}_1^A \rangle$ will asymptotically tend to

$$\langle \mathcal{S}_0^A \rangle = k[2c \cos(\theta)/(k-1) + \sin \theta k/(k-1)] \tag{43}$$

and

$$\langle \mathcal{S}_1^A \rangle = k[-2c \cos(\theta)/(k-1) + \sin \theta k/(k-1)]. \tag{44}$$

Therefore, with the optimal asymptotic value of $\theta = 5\pi/6$ we obtain

$$\langle \mathcal{S}_0^A \rangle - \langle \mathcal{S}_1^A \rangle = \frac{k}{k-1}. \tag{45}$$
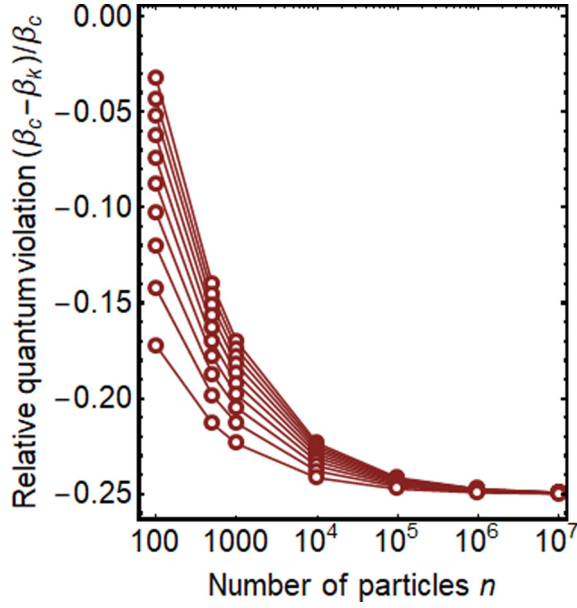
FIG. 6. Asymptotic k-producibility bounds for the DIWED in Eq. (35) for $k = n/m$, where m is an integer ranging from 1 (lowest line, coinciding with the maximal quantum violation of the Bell inequality) to 10. In case of not exact divisibility, the value of n that is a multiple of m and is closest to a power of 10 has been chosen. For this asymptotic computation, we have taken a state of the form Eq. (37) and we have numerically optimized μ_A and σ_A. The optimal value converges to all μ_A and σ_A equal, since all elements in $\mathcal{P}$ have the same cardinality. We also observe that taking $m = O(1)$ does not allow in the large n limit for an experimentally robust result, and a different scaling is therefore needed, as shown in Fig. 7.

Hence the asymptotic limit for a partition of $m = n/k$ elements reads

$$\beta_k \approx m\left(-\frac{5}{2}k + \frac{\sqrt{3}}{2}k^{1/2} - \frac{3}{2}\right) + \binom{m}{2}\left(\frac{k}{k-1}\right)^2. \quad (46)$$

In terms of its relative quantum violation $\tilde{\beta}_k := (\beta_c - \beta_k)/\beta_c$, we have that

$$\tilde{\beta}_k = -\frac{1}{4} + \frac{\sqrt{3m}}{4}n^{-1/2} + \frac{m^2 - 4m}{4}n^{-1} + O(n^{-3/2}). \quad (47)$$

In Fig. 6 and Fig. 7 we have shown the bounds obtained for large n and different k-producibility bounds, where $k = n/m$ and $m \in [10]$ (Fig. 6) or $m = n^{1/x}$ for several integer values of x (Fig. 7). We see that as n grows the bounds tend to the relative maximal quantum violation $-1/4$ as $1/\sqrt{n}$, and the approximation holds as long as $m = O(\sqrt[3]{n})$. Note that for much larger values of m [e.g., $m = O(n)$], this corresponds to a small value for k, for which the approximation in Eq. (42) could break down.

VII. EXPERIMENTAL REALIZATIONS

Although the DIWEDs introduced in this work allow one to certify certain entanglement depths by performing a Bell test on spatially separated parties, a Bell test in the many-body regime is much more demanding. The full potential of our DIWEDs can be unlocked by means of Bell correlation
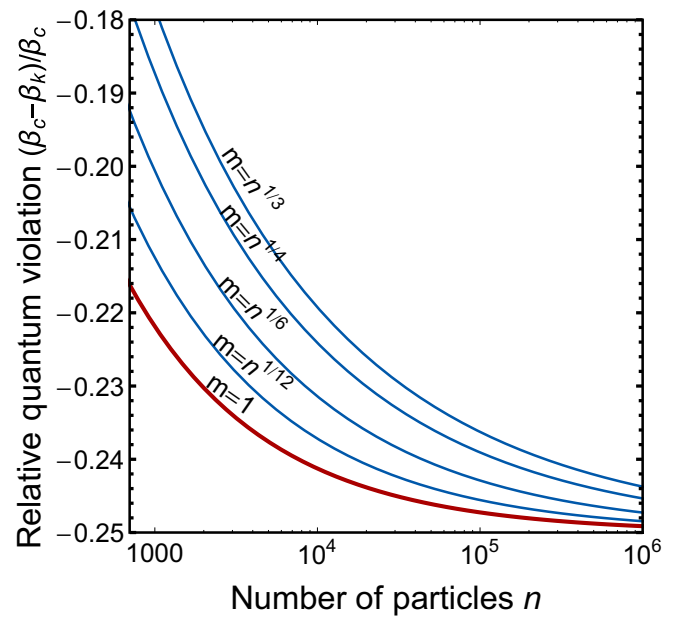

FIG. 7. Asymptotic bounds of the DIWED stemming from Eq. (35) for different values of $k = n/m$, according to Eq. (47). Although all the curves converge to the maximal quantum violation in the thermodynamic limit, these curves would allow for entanglement depth detection in experimentally realistic parameter regimes [11,12].

witnesses, which only require trusted collective measurements [11], thus certifying the depth of entanglement in the state without an explicit characterization of the latter. The goal of this section is to rewrite our DIWEDs in terms of trusted collective measurements, which can be already performed in current experiments [11,12].

When all parties perform the same measurements, it is clear that $\mathcal{S}_k$ can be directly mapped to a collective spin measurement along the direction of $\mathcal{M}_k$, and it is also easy to see that the expectation value of $\mathcal{S}_{kk}$ can be obtained by measuring the second moments of such collective spin measurement along the direction of $\mathcal{M}_k$ [45]. The reason for that is simply that one can rewrite $\mathcal{S}_{kk} = (\mathcal{S}_k)^2 - n$. The $\mathcal{S}_{01}$ term is, however, more problematic, since one would apparently need to individually address the different parties in the system in order to measure it. One possibility is to estimate the second moments of the collective spin component along the directions given by $(\mathcal{M}_0 \pm \mathcal{M}_1)/\sqrt{2}$ [45], but this approach could be somewhat unsatisfactory, as evaluating the expectation value of the Bell inequality then requires measuring along two extra directions.

Let us briefly recall the approach in [11] in order to show how one would measure a DIWED in the many-body regime [specifically, in a Bose-Einstein condensate (BEC)] using trusted collective measurements. This will allow us to compare the performance of our DIWEDs to other existing entanglement and nonlocality criteria, as we show in Fig. 8 and in Sec. VIII. Note that the Bell inequality that was actually used in [11] belongs to the same equivalence class as Eq. (35), as one can take all its coefficients to be non-negative by renaming measurement outcomes appropriately [45]. Here we
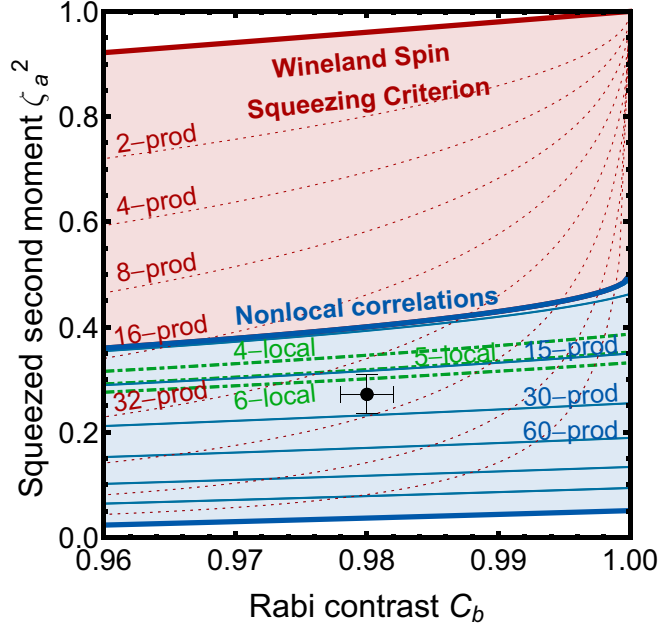
FIG. 8. Witnesses of entanglement depth derived from trusted collective measurements. The blue lines follow Eq. (50) for different values of k. In green we have plotted the nonlocality depth bounds for Eq. (35) for the values of $k = 4, 5, 6$. Finally, in red, the non-device-independent witnesses of entanglement depth derived from the Wineland spin squeezing criterion.

construct the entanglement depth witness for Eq. (35) to be consistent with the rest of the paper.

Using the same hypotheses as in [11], namely, that the system is characterized by a quantum mechanical description, that the experimental calibration of the measuremetns is trusted and that the particles do not communicate nor interact through channels unaccounted for, one can derive an entanglement depth witness that bypasses the problems due to the $\mathcal{S}_{01}$ term. The steps that we now follow are essentially identical to the ones in [11], substituting the classical bound β_c by the k-producible bound β_k, and changing one of the signs of the measurement directions, due to the inequalities considered here and in [11] being two different representatives of the same equivalence class. Nevertheless, in the interest of the current section being self-contained, we reproduce them here for completeness.

First, one associates to the ith observer a spin-1/2 particle and considers that the measurements are spin projections along a direction d in the Bloch sphere; i.e., $\mathcal{M}_d^{(i)} = 2\hat{s}^{(i)} \cdot d \equiv \boldsymbol{\sigma}^{(i)} \cdot d$, where $\boldsymbol{\sigma}^{(i)} := \{\sigma_x^{(i)}, \sigma_y^{(i)}, \sigma_z^{(i)}\}$ formally represents a vector formed by the Pauli matrices and $\hat{s}^{(i)}$ is the individual spin operator acting on the ith site. It is worth noting that the spin-1/2 description corresponds to the lowest energy levels of the atoms and that higher energy levels as well as further degrees of freedom such as atomic motion are neglected in this model of description.

Second, one defines the total spin observable as the sum of the individual spins along a given direction: $\hat{S}_d = d \cdot \sum_i \hat{s}^{(i)}$. Consider now two unit vectors a and n and define a new vector $m = 2(a \cdot n)a - n$. We note that m is also a unit vector and it is the result of applying a symmetry transformation to n

with respect to the symmetry axis defined by a. By setting the measurement directions to be $\mathcal{M}_0^{(i)} = \mathcal{M}_n^{(i)}$ and $\mathcal{M}_1^{(i)} = \mathcal{M}_m^{(i)}$, the interpretation of a is clear, as it is the bisector of the angle formed between the two measurement directions.

In terms of the total spin component, we note the identities $\langle \hat{S}_n \rangle = \langle \mathcal{S}_0 \rangle / 2$ and $16(a \cdot n)^2 \langle \hat{S}_a^2 \rangle = \langle \mathcal{S}_{00} \rangle + 2\langle \mathcal{S}_{01} \rangle + \langle \mathcal{S}_{11} \rangle + 4n(a \cdot n)^2$.

Now, we would like to obtain a term of the form $\langle \mathcal{S}_{00} \rangle - 2\langle \mathcal{S}_{01} \rangle + \langle \mathcal{S}_{11} \rangle$, in order to map the two-body terms of Eq. (35) to the Bell correlation witness we are constructing. This corresponds to the symmetry $\mathcal{M}_0^{(i)} \leftrightarrow -\mathcal{M}_0^{(i)}$ (renaming the outcomes of the zeroth observable of every party). Thus let us define $d = -n$. Then $m = -2(a \cdot d)a + d$ and in a similar fashion we arrive at the new identity, $16(a \cdot d)^2 \langle \hat{S}_a^2 \rangle = \langle \mathcal{S}_{00} \rangle - 2\langle \mathcal{S}_{01} \rangle + \langle \mathcal{S}_{11} \rangle + 4n(a \cdot d)^2$.

Hence, by normalizing to the classical bound, we can now turn our DIWED Eq. (35) into the following witness, which only consists of first and second moments of (trusted) collective spin measurements:

$$\hat{W} = -2\hat{S}_d/n + 4(a \cdot d)^2 \hat{S}_a^2/n - (a \cdot d)^2 - \frac{\beta_k}{2n}\mathbb{1}, \quad (48)$$

which satisfies the inequality $\langle \hat{W} \rangle \geqslant 0$ on k-producible states. Thus an expectation value $\langle \hat{W} \rangle < 0$ certifies an entanglement depth of at least $k + 1$ particles being genuinely multipartite entangled (see Fig. 8).

VIII. COMPARISON TO OTHER ENTANGLEMENT CRITERIA

In order to compare with other entanglement criteria, let us recall the definition of the scaled collective spin, also known as spin contrast, in the direction d: $C_d := \langle 2\hat{S}_d/\hat{n} \rangle$, where $\hat{n}$ is the particle number operator. Note that one has to take into consideration that the number of particles may vary among different experimental realizations [11]. However, $[\hat{S}_d, \hat{n}] = 0$, thus allowing one to rewrite $C_d = \langle 2\hat{S}_d \rangle / n$. Let us also recall the definition of the scaled second moment collective spin in the direction a, $\zeta_a^2 := \langle 4\hat{S}_a^2/\hat{n} \rangle = \langle 4\hat{S}_a^2 \rangle / n$, for the same reason. Therefore, we have that the expectation value of $\hat{W}$ defined in can be recast as

$$\langle \hat{W} \rangle = -C_d + (a \cdot d)^2 \zeta_a^2 - (a \cdot d)^2 - \frac{\beta_k}{2n}. \quad (49)$$

In this form, an algebraic manipulation [11] allows us to compare the performance of our DIWED with other entanglement criteria such as Wineland's [91,92] (see Fig. 8). In order to do so, let us decompose the direction d into three orthonormal vectors a, b, c such that a is the squeezed axis (hence we can assume $C_a \approx 0$ and we only need to sweep across the b direction). Therefore, the inequality $\langle \hat{W} \rangle \geqslant 0$ becomes [cf.]

$$\zeta_a^2 \geqslant \frac{1}{2}\left[1 + \frac{\beta_k}{2n} - \sqrt{\left(1 - \frac{\beta_k}{2n}\right)^2 - C_b^2} \right]. \quad (50)$$

Furthermore, in Fig. 8, we also compared the performance of the Bell correlation depth witnesses from [42] with that of our DIWEDs. Since in [42] one takes the Bell inequalities

of the form of Eq. (35) and derives a nonlocality depth based on the assumption that parties in the same group can produce any nonsignaling (hence potentially supraquantum) probability distribution, here we obtain tighter entanglement depth bounds because we restrict the power of the parties in the same group to produce only quantum correlations.

IX. CONCLUSIONS AND OUTLOOK

In this work we have extended the analysis and results of [9], presenting in detail the optimization procedure to derive DIWEDs from two-body permutationally invariant Bell inequalities. We have compared our findings to other existing criteria for determining entanglement depth and our methods yield better results (tighter bounds) when compared to methods based on stronger assumptions [42], yet worse than methods that are not device independent [91,92].

We see that a key ingredient in our method is the use of Jordan's lemma, which applies to Bell inequalities with two binary observables per site. Therefore, it is not obvious how to extend our method to some more general families of Bell inequalities than the ones we here consider, for instance, with more measurement settings per site [47]. A straightforward generalization of the method to these more complex Bell scenarios could be made, at the price of shifting to the semi-device-independent paradigm, because one would need to know the dimension of the physical systems in advance [93]. A natural next step, however, is the case of Bell inequalities with the same features, and with higher-order (but constant) correlators, which can be obtained without having to resort to the polytope approach [48]. In this case we would expect to obtain even tighter DIWEDs with essentially an analogous analysis as the one we performed here. Also, we leave as an open problem the possibility of generalizing our approach to nonlinear device-independent witnesses of entanglement depth, that is, witnesses based on nonlinear Bell expressions such as those studied in Refs. [48,94] which are more effective tools to characterize entanglement depth for broader sets of correlations.

Another interesting feature we observe is that the optimal state tends to be symmetric at each partition. Recently, tremendous experimental progress has been achieved in measuring entanglement across different wells of a spin-squeezed BEC [95–97] (see also [98]), which poses the question about how well the DIWEDs we here present can certify the entanglement in these systems, or how robust are the DIWED bounds for more realistic experimental conditions, such as finite temperature [50].

Implementing these DIWEDs in the multipartite scenario poses the same technical challenges as performing a multipartite Bell test. To date, loophole-free Bell tests exist only for the bipartite scenario [99–101] (see also [102]). However, the DIWEDs presented here inherit the property that their expectation value can be estimated via trusted collective observables and higher moments thereof [11,12], which, for instance, makes them applicable in quantum systems in which individual particles cannot be addressed.

Throughout this work, our method to produce a lower bound certificate via semidefinite programming is currently based on the PPT condition as a relaxation of the separability condition. However, we expect that the gap between the variational solution and the certificate should not close in the general case, as fulfilling the PPT criterion is in general not sufficient to guarantee separability, even in the symmetric case [22,23], or even under stronger assumptions [103–105]. Therefore, a natural next step in order to tighten these lower bound certificates is the use of stronger entanglement criteria (either using the symmetry condition in PPT states or using stronger, albeit computationally harder relaxations in the form of semidefinite programs [76]).

APPENDIX: ASYMPTOTIC ANALYSIS DETAILS

In this section we motivate where the $\sigma = \sqrt{n/48}$ approximation, which we use in Sec. VI, comes from. In [45] an asymptotic analysis of the value of the maximal quantum violation of Eq. (35) was performed. However, only the scaling was relevant for that study. Here we determine the coefficient of the second-order term exactly.

Let us recall that the expectation value of the optimal state [cf. Eq. (37)] violating Eq. (35) can be expressed as [45]

$$\langle \mathcal{B} \rangle = \left(\frac{\beta_c}{n} - \frac{B}{2} + e^{-1/8\sigma} A' \right) n + (2B\sigma - A^2/2B + e^{-1/8\sigma} A') + O(\sigma/n), \tag{A1}$$

where $A = 2\cos\theta$, $B = 4\cos^2\theta$, and $A' = -2\sin\theta$, with $\theta = 5\pi/6$ being the optimal asymptotic values. To determine the width σ of the Gaussian superposition of Dicke states we use these values and we find that σ must fulfill

$$e^{-1/8\sigma}(n+1) = 48\sigma^2, \tag{A2}$$

which is a transcendental equation. However, since the scaling for σ will be $\sqrt{n}$ in the first order, we can for large values of n ignore the $e^{-1/8\sigma}$ term and therefore approximate $\sigma = \sqrt{(n+1)/48} \approx \sqrt{n/48}$, which is what we use in our asymptotic analysis.

[1] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).

[2] G. Vitagliano, P. Hyllus, I. L. Egusquiza, and G. Tóth, Phys. Rev. Lett. **107**, 240502 (2011).

[3] F. Troiani and I. Siloi, Phys. Rev. A **86**, 032330 (2012).

[4] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[5] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Phys. Rev. Lett. **106**, 250404 (2011).

[6] Y.-C. Liang, D. Rosset, J.-D. Bancal, G. Pütz, T. J. Barnea, and N. Gisin, Phys. Rev. Lett. **114**, 190401 (2015).

[7] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Phys. Rev. Lett. **111**, 030501 (2013).

[8] P.-S. Lin, J.-C. Hung, C.-H. Chen, and Y.-C. Liang, Phys. Rev. A **99**, 062338 (2019).

[9] A. Aloy, J. Tura, F. Baccari, A. Acín, M. Lewenstein, and R. Augusiak, Phys. Rev. Lett. **123**, 100507 (2019).

[10] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, Science **344**, 1256 (2014).

[11] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, Science **352**, 441 (2016).

[12] N. J. Engelsen, R. Krishnakumar, O. Hosten, and M. A. Kasevich, Phys. Rev. Lett. **118**, 140401 (2017).

[13] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[14] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Commun. Math. Phys. **328**, 303 (2014).

[15] S. Szalay and Z. Kökényesi, Phys. Rev. A **86**, 032341 (2012).

[16] S. Szalay, Phys. Rev. A **92**, 042329 (2015).

[17] F. Levi and F. Mintert, Phys. Rev. Lett. **110**, 150402 (2013).

[18] O. Gühne and M. Seevinck, New J. Phys. **12**, 053002 (2010).

[19] S. Wölk and O. Gühne, New J. Phys. **18**, 123024 (2016).

[20] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, 1989), pp. 69–72.

[21] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, Nature (London) **403**, 515 (2000).

[22] J. Tura, R. Augusiak, P. Hyllus, M. Kuś, J. Samsonowicz, and M. Lewenstein, Phys. Rev. A **85**, 060302(R) (2012).

[23] R. Augusiak, J. Tura, J. Samsonowicz, and M. Lewenstein, Phys. Rev. A **86**, 042316 (2012).

[24] M. Huber, P. Erker, H. Schimpf, A. Gabriel, and B. Hiesmayr, Phys. Rev. A **83**, 040301(R) (2011).

[25] D. Mayers and A. Yao, Quantum Inf. Comput. **4**, 273 (2004).

[26] A. Coladangelo, K. T. Goh, and V. Scarani, Nat. Commun. **8**, 15485 (2017).

[27] J. Kaniewski, Phys. Rev. A **95**, 062323 (2017).

[28] C. Bamps and S. Pironio, Phys. Rev. A **91**, 052111 (2015).

[29] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, Phys. Rev. Lett. **121**, 180503 (2018).

[30] J. Kaniewski, Phys. Rev. Lett. **117**, 070402 (2016).

[31] W.-H. Zhang, G. Chen, P. Yin, X.-X. Peng, X.-M. Hu, Z.-B. Hou, Z.-Y. Zhou, S. Yu, X.-J. Ye, Z.-Q. Zhou, X.-Y. Xu, J.-S. Tang, J.-S. Xu, Y.-J. Han, B.-H. Liu, C.-F. Li, and G.-C. Guo, npj Quantum Inf. **5**, 4 (2019).

[32] F. Baccari, R. Augusiak, I. Šupić, J. Tura, and A. Acín, arXiv:1812.10428v1.

[33] T. H. Yang and M. Navascués, Phys. Rev. A **87**, 050102 (2013).

[34] A. Salavrakos, R. Augusiak, J. Tura, P. Wittek, A. Acín, and S. Pironio, Phys. Rev. Lett. **119**, 040402 (2017).

[35] J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, D. Bonneau, J. W. Silverstone, Q. Gong, A. Acín, K. Rottwitt, L. K. Oxenløwe, J. L. O'Brien, A. Laing, and M. G. Thompson, Science **360**, 285 (2018).

[36] J. Kaniewski, I. Supic, J. Tura, F. Baccari, A. Salavrakos, and R. Augusiak, arXiv:1807.03332v1.

[37] C. Schwemmer, L. Knips, M. C. Tran, A. de Rosier, W. Laskowski, T. Paterek, and H. Weinfurter, Phys. Rev. Lett. **114**, 180501 (2015).

[38] L. Knips, C. Schwemmer, N. Klein, M. Wieśniak, and H. Weinfurter, Phys. Rev. Lett. **117**, 210504 (2016).

[39] M. C. Tran, M. Zuppardo, A. de Rosier, L. Knips, W. Laskowski, T. Paterek, and H. Weinfurter, Phys. Rev. A **95**, 062331 (2017).

[40] M. Zwerger, W. Dür, J.-D. Bancal, and P. Sekatski, Phys. Rev. Lett. **122**, 060502 (2019).

[41] J. Barreiro, J.-D. Bancal, P. Schindler, D. Nigg, M. Hennrich, T. Monz, N. Gisin, and R. Blatt, Nat. Phys. **9**, 559 (2013).

[42] F. Baccari, J. Tura, M. Fadel, A. Aloy, J.-D. Bancal, N. Sangouard, M. Lewenstein, A. Acín, and R. Augusiak, Phys. Rev. A **100**, 022121 (2019).

[43] W. Slofstra, arXiv:1703.08618v2.

[44] K. T. Goh, J. Kaniewski, E. Wolfe, T. Vértesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, Phys. Rev. A **97**, 022104 (2018).

[45] J. Tura, R. Augusiak, A. Sainz, B. Lücke, C. Klempt, M. Lewenstein, and A. Acín, Ann. Phys. (NY) **362**, 370 (2015).

[46] S. Pelisson, L. Pezzè, and A. Smerzi, Phys. Rev. A **93**, 022115 (2016).

[47] S. Wagner, R. Schmied, M. Fadel, P. Treutlein, N. Sangouard, and J.-D. Bancal, Phys. Rev. Lett. **119**, 170403 (2017).

[48] M. Fadel and J. Tura, Phys. Rev. Lett. **119**, 230402 (2017).

[49] D.-L. Deng, Phys. Rev. Lett. **120**, 240402 (2018).

[50] M. Fadel and J. Tura, Quantum **2**, 107 (2018).

[51] K. F. Pál and T. Vértesi, Phys. Rev. A **82**, 022116 (2010).

[52] M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, and R. F. Werner, J. Math. Phys. **52**, 012102 (2011).

[53] T. Fritz, Rev. Math. Phys. **24**, 1250012 (2012).

[54] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, Nat. Commun. **6**, 6288 (2015).

[55] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).

[56] B. Toner and F. Verstraete, arXiv:quant-ph/0611001v1.

[57] C. Jordan, Bull. Soc. Math. France **3**, 103 (1875).

[58] J. Harris, *Algebraic Geometry* (Springer, New York, 1992).

[59] E. Noether and B. L. van der Waerden, *Die Alternative Bei Nichlinearen Gleichungen*, Nachrichten der Gesselschaft der Wissenschaften zu Gottingen No. 77-87 (Math. Phys. Klasse, Philadelphia, USA, 1928).

[60] B. Sturmfels, *Solving Systems of Polynomial Equations (CBMS Regional Conference Series in Mathematics)* (American Mathematical Society, Providence, RI, 2002).

[61] A. H. Wright, Math. Comput. **44**, 125 (1985).

[62] C. B. Garcia and T. Y. Li, SIAM J. Numer. Anal. **17**, 540 (1980).

[63] M. Kojima and S. Mizuno, Math. Program. **25**, 131 (1983).

[64] B. Sturmfels, Not. Am. Math. Soc. **52**, 1199 (2005).

[65] J.-C. Faugère, J. Pure Appl. Alg. **139**, 61 (1999).

[66] J. C. Faugère, in *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation - ISSAC02* (ACM Press, NY, USA, 2002).

[67] T. Moroder, P. Hyllus, G. Tóth, C. Schwemmer, A. Niggebaum, S. Gaile, O. Gühne, and H. Weinfurter, New J. Phys. **14**, 105001 (2012).

[68] J. T. i Brugués, *Characterizing Entanglement and Quantum Correlations Constrained by Symmetry* (Springer International Publishing, Berlin, 2017).

[69] A. W. Harrow, arXiv:quant-ph/0512255v1.

[70] M. Christandl, arXiv:quant-ph/0604183v1.

[71] K. Audenaert (unpublished).

[72] R. F. Werner and M. M. Wolf, Quantum Information & Computation **3**, 1 (2001).

[73] L. Gurvits, in *Proceedings of the Thirty-fifth ACM Symposium on Theory of Computing - STOC'03* (ACM Press, NY, USA, 2003).

[74] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[75] A. W. Harrow, A. Natarajan, and X. Wu, Commun. Math. Phys. **352**, 881 (2017).

[76] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004).

[77] *Semidefinite Optimization and Convex Algebraic Geometry* edited by G. Blekherman , P. A. Parrilo, and R. R. Thomas (SIAM, Philadelphia, PA, 2013).

[78] F. Baccari, C. Gogolin, P. Wittek, and A. Acín, arXiv:1808.01275v1.

[79] M. A. Jafarizadeh, M. Mahdian, A. Heshmati, and K. Aghayar, Eur. Phys. J. D **50**, 107 (2008).

[80] C. Spengler, M. Huber, S. Brierley, T. Adaktylos, and B. C. Hiesmayr, Phys. Rev. A **86**, 022311 (2012).

[81] B. C. Hiesmayr and W. Löffler, Phys. Scr. **T160**, 014017 (2014).

[82] B. Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, and C. Klempt, Phys. Rev. Lett. **112**, 155304 (2014).

[83] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[84] J. F. Sturm, Opt. Meth. Softw. **11**, 625 (1999).

[85] The MOSEK optimization software.

[86] K. C. Toh, M. Todd, R. Tütüncü, and R. H. Tutuncu, Opt. Meth. Softw. **11**, 545 (1998).

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

[88] M. C. Grant and S. P. Boyd, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, Vol. 371, edited by V. D. Blondel, S. P. Boyd, and

H. Kimura (Springer-Verlag Limited, Berlin, 2008), pp. 95–110, https://web.stanford.edu/~boyd/papers/graph_dcp.html.

[89] J. Lofberg, in *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)* (IEEE, New York, 2004).

[90] F. Fröwis, P. C. Strassmann, A. Tiranov, C. Gut, J. Lavoie, N. Brunner, F. Bussières, M. Afzelius, and N. Gisin, Nat. Commun. **8**, 907 (2017).

[91] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys. Rev. A **50**, 67 (1994).

[92] A. S. Sørensen and K. Mølmer, Phys. Rev. Lett. **86**, 4431 (2001).

[93] Y.-C. Liang, T. Vértesi, and N. Brunner, Phys. Rev. A **83**, 022108 (2011).

[94] A. Salles, D. Cavalcanti, A. Acin, D. Perez-Garcia, and M. M. Wolf, Quantum Inf. Comput. **10**, 0703 (2010).

[95] M. Fadel, T. Zibold, B. Décamps, and P. Treutlein, Science **360**, 409 (2018).

[96] P. Kunkel, M. Prüfer, H. Strobel, D. Linnemann, A. Frölian, T. Gasenzer, M. Gärttner, and M. K. Oberthaler, Science **360**, 413 (2018).

[97] K. Lange, J. Peise, B. Lücke, I. Kruse, G. Vitagliano, I. Apellaniz, M. Kleinmann, G. Tóth, and C. Klempt, Science **360**, 416 (2018).

[98] D. K. Shin, S. S. Hodgman, B. M. Henson, T. Wasak, J. Chwedenczuk, and A. G. Truscott, arXiv:1811.05681v1.

[99] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Nature (London) **526**, 682 (2015).

[100] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman *et al.*, Phys. Rev. Lett. **115**, 250402 (2015).

[101] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, Phys. Rev. Lett. **115**, 250401 (2015).

[102] T. B. B. T. Collaboration, Nature (London) **557**, 212 (2018).

[103] R. Quesada and A. Sanpera, Phys. Rev. A **89**, 052319 (2014).

[104] R. Quesada, S. Rana, and A. Sanpera, Phys. Rev. A **95**, 042128 (2017).

[105] J. Tura, A. Aloy, R. Quesada, M. Lewenstein, and A. Sanpera, Quantum **2**, 45 (2018).