Optimal measurement strategies for fast entanglement detection

N. Milazzo,[1,2,*] D. Braun,[1] and O. Giraud[2]

[1]*Institut für theoretische Physik, Universität Tübingen, 72076 Tübingen, Germany*
[2]*LPTMS, CNRS, Université Paris-Sud, Université Paris-Saclay, 91405 Orsay, France*

With the advance of quantum information technology, the question of how to most efficiently test quantum circuits is becoming of increasing relevance. Here we introduce the statistics of lengths of measurement sequences that allows one to certify entanglement across a given bipartition of a multiqubit system over the possible sequence of measurements of random unknown states and identify the best measurement strategies in the sense of the (on average) shortest measurement sequence of (multiqubit) Pauli measurements. The approach is based on the algorithm of truncated moment sequences, which allows one to deal naturally with incomplete information, i.e., information that does not fully specify the quantum state. We find that the set of measurements corresponding to diagonal matrix elements of the moment matrix of the state are particularly efficient. For symmetric states their number increases only like the third power of the number N of qubits. Their efficiency increases rapidly with N, leaving already for $N = 4$ less than a fraction 10^{-6} of randomly chosen entangled states undetected.

I. INTRODUCTION

With the availability of the first small quantum processors, the task of characterizing such processors has become a key challenge. Indeed, long before proving full functionality, one of the major questions that faces a quantum processor is whether it "truly" works quantum mechanically—or could rather be explained by classical processes. Similar questions arise already at the level of a quantum state: Given a physical system in an unknown quantum state, can the statistics arising from it be explained by a classical state? If the state is fully characterized, one can apply nonclassicality measures to find out, but since a mixed quantum state of N qubits is specified by $d = 2^{2N} - 1$ real parameters, it is clear that an answer based on full quantum state tomography quickly becomes impractical. In addition, one can only estimate expectation values based on averages over finitely many measurements that are themselves imperfect, and the resulting uncertainty can lead to nonphysical states in the inversion procedure underlying full quantum state tomography. More robust approaches to state tomography are maximum likelihood estimation of the state [1–4] and Bayesian inference [5,6], which output estimates of the state that are by construction *bona fide* physical states, as well as "self-consistent quantum tomography," which does not necessarily rely on perfect measurements [7], but none of these approaches remedies the efficiency problem.

Recent developments based on compressed sensing make use of prior information on states. They provide a large gain in efficiency, in particular, for the typically low-rank states relevant for quantum information tasks [8–12] or matrix-product states that describe interacting condensed-matter systems in low dimensions [13,14]. Machine learning aimed at

determining by itself what the best measurements are for a certain task, or to recognize entanglement from measurement data, was considered, e.g., in [15–18], but the efficiency of such approaches needs further study. Other proposals include few-copy multiparticle entanglement detection based on probabilistic verification [19,20].

For testing quantum circuits, the approach of randomized benchmarking has emerged [21–24]. Key to this approach is that for estimating fidelities between actual and ideal gate sets, only low moments of the matrix elements are required. In this case, averaging over the full unitary group can be replaced by averaging over a unitary t design [25] or producing required input states by random quantum circuits (see also [26]). References [27] and [28] showed that a small number of parameters of a quantum process can be efficiently obtained, but it is not as clear what the most relevant parameters that should be chosen are.

It is often stated that quantum states and quantum processors are much harder to test and characterize than their classical analogs because of the exponential growth of the Hilbert space [13,29,30]. However, also classically the number of possible memory configurations of N bits increases exponentially as 2^N—and with N of order 10^{13} for a standard laptop computer, it is completely out of the question to test all possible configurations. Costs of integration itself have decreased exponentially according to Moore's law; for the same reason, functional testing of classical memory devices has evolved to the most expensive (because time-consuming) part of the production of integrated memory chips. Functional testing of classical memories has therefore evolved to *testing the most critical known configurations with the goal of demonstrating failure of memory cells as quickly as possible*. "Most critical" depends on the architecture of the chip, and information on its design goes into the design of memory patterns to be tested. For example, a cell on a given bit line

*nadia.milazzo@u-psud.fr

might resist storing a "0," most likely if all other cells on the same bit line contain a "1." In MRAM devices, magnetic stray fields from a set of cells can destabilize others in the vicinity when uniformly polarized, etc.

Quantum information processing may still have a long way to go before such economical pressure on functional testing will be felt. At the moment, rather than showing failure, one would like to prove basic quantum functionalities as quickly as possible. Nevertheless, the principles of classical functional testing can also provide guidance in the current state of affairs on characterizing quantum processors and states: rather than aiming at full quantum tomography, one may want to focus on producing states that are likely to be particularly unstable and show their "functionality" as quickly as possible. In practice, this will require information about the physical realization of the quantum processor, but in the absence of such input, a reasonable target is highly entangled states or, more generally, highly nonclassical states known to be prone to decoherence. Indeed, experimental efforts have early on concentrated on producing such states (see, e.g., [31–35] for states with large numbers of entangled particles).

The question then arises: What is the most efficient measurement strategy to prove that such a state is entangled (or, more generally, nonclassical)? That is, What would you choose to measure first, second, and so on, in order to be able to prove as quickly as possible, with the limited knowledge about the state that you will gain from those measurements, chosen from a given set, that the state is entangled? What are the minimum and average numbers of measurements needed to prove entanglement or, more generally, the statistics on the length of measurement sequences when going down a certain path of measurements?

These are the questions that we start to answer in the present paper. Note that this is not about choosing optimal entanglement witnesses but, rather, about deciding whether or not the intersection of hyperplanes defined by the expectation values of certain observables cuts the set of separable states (see Fig. 4). Perfectly suited for answering these questions is the formalism of *truncated moment sequences* (TMSs) that we introduced in [36] for the analysis of entanglement. The TMS problem aims at finding a probability measure for which only some moments are known. If the probability measure is, furthermore, constrained to be supported on a compact set K, the problem is known as the K-TMS problem. As reviewed below, it can be solved with a hierarchy of flat extensions that maps onto a convex optimization algorithm, using a semidefinite relaxation procedure. Each expectation value can be associated with a moment of a measure, and instead of fixing all moments up to a certain order as in the standard TMS algorithm, one might just specify any set $\mathcal{A}$ of moments. The problem of deciding whether a classical measure that reproduces all these moments exists is then known as the "$\mathcal{A}K$-TMS" problem [37]. It can still be solved with a convex optimization algorithm.

In the present work we exploit this approach in order to obtain the statistics of lengths of measurement sequences in the simplest case of two qubits depending on the chosen measurement strategy. For larger numbers of qubits, the full numerical solution of the $\mathcal{A}K$-TMS problem becomes too demanding, but it turns out that surprisingly efficient sufficient

conditions for entanglement can be obtained for symmetric states from the diagonal matrix elements of the moment matrix used in the approach (see below for a definition). These correspond to certain linear combinations of expectation values of (possibly multipartite) measurements in the Pauli basis and have to be positive for a solution of the $\mathcal{A}K$-TMS problem to exist. Checking the positivity of moment matrices is in fact the first step in the TMS algorithm, and negativity of any of the diagonal matrix elements hence witnesses entanglement. With these we can find numerical estimates of the fraction of randomly drawn states that are already detected as entangled by just measuring the observables corresponding to the diagonal matrix elements of the moment matrix.

Similar ideas for certifying entanglement with incomplete measurements were considered in [38] for continuous variable systems. Here we focus on the statistics of lengths of sequences of measurements for multiqubit systems and the insights that can be drawn from the TMS algorithm, which we review in the next section, before applying it to incomplete measurements.

II. FRAMEWORK AND NOTATION

We now briefly summarize the TMS algorithm approach described in detail in [36], which is the framework for the following sections. The basic idea is to map the quantum entanglement problem onto the mathematically well-studied truncated moment problem. Indeed, finding out whether an arbitrary multipartite state can be decomposed into product states corresponds to finding out about the existence of a probability distribution whose lowest-order moments are fixed. Analytically, the mapping allows one to make use of theorems from the TMS literature providing necessary and sufficient separability conditions; numerically, semidefinite optimization techniques yield an algorithm which gives a certificate of entanglement or separability. The algorithm applies—at least in principle—to arbitrary quantum states with an arbitrary number of constituents and arbitrary symmetries between the subparts. The general case is dealt with in [36]; we only recall here the main key points for the case of symmetric states of qubits, defined as mixtures of symmetric pure states (the latter are invariant under any permutation of the qubits). To do so, we use a convenient representation in terms of symmetric tensors which was introduced in [39], generalizing the Bloch sphere picture of spins-1/2. We can write a generic state ρ of a spin-j state as

$$\rho = \frac{1}{2^N} \sum_{\mu_1,\mu_2,\ldots,\mu_N=0}^{3} X_{\mu_1\mu_2\ldots\mu_N} P_s\big(\sigma_{\mu_1}\otimes\ldots\otimes\sigma_{\mu_N}\big)P_s^{\dagger}, \quad (1)$$

where σ_0 is the 2×2 identity matrix, σ_1, σ_2, and σ_3 are the Pauli matrices, and P_s is the projector onto the symmetric subspace spanned by the Dicke states $|j,m\rangle$ (eigenstates of pseudoangular momentum component J_z and with total angular momentum quantum number j). They can also be seen as symmetric states of $N = 2j$ spins-1/2 (or qubits). The tensor $X_{\mu_1\mu_2\ldots\mu_N}$ is then given by

$$X_{\mu_1\mu_2\ldots\mu_N} = \text{tr}\big(\rho\sigma_{\mu_1}\otimes\ldots\otimes\sigma_{\mu_N}\big), \quad (2)$$

with $0 \leqslant \mu_i \leqslant 3$. It is real and invariant under permutation of indices and verifies $X_{00...0} = \mathrm{tr}(\rho) = 1$. Moreover, it has the property that

$$\sum_{a=1}^{3} X_{aa\mu_3\mu_4...\mu_N} = X_{00\mu_3\mu_4...\mu_N} \qquad (3)$$

for any choice of the μ_i. A separable pure state can be seen as a spin-coherent state, which in representation (2) has tensor entries $X_{\mu_1\mu_2...\mu_N} = n_{\mu_1} n_{\mu_2} \ldots n_{\mu_n}$, with $n_0 = 1$ and (n_1, n_2, n_3) the unit vector giving the direction of the coherent state on the Bloch sphere. In terms of this tensor representation, a symmetric state is separable if and only if its tensor representation can be written as

$$X_{\mu_1\mu_2...\mu_N} = \sum_i \omega_i n_{\mu_1}^{(i)} n_{\mu_2}^{(i)} \ldots n_{\mu_N}^{(i)}, \quad \omega_i \geqslant 0, \qquad (4)$$

where $n_0^{(i)} = 1$ and $\mathbf{n}^{(i)}$ is the Bloch vector of the single qubit. If we express (4) in the equivalent integral form,

$$X_{\mu_1\mu_2...\mu_N} = \int_K x_{\mu_1} x_{\mu_2} \ldots x_{\mu_N} d\mu(\mathbf{x}), \qquad (5)$$

with $x_0 = 1$, $K = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ the unit sphere, and $d\mu(\mathbf{x}) = \sum_i \omega_i \delta(\mathbf{x} - \mathbf{n}^{(i)})$ a positive measure on K, we can say that a symmetric state is separable if and only if there exists a positive measure $d\mu$ supported by K such that all entries of the tensor $X_{\mu_1\mu_2...\mu_N}$ are given by moments of that measure.

Problems of this type are known as K-TMS problems, or $\mathcal{A}K$-TMS problems in the case of partial knowledge of a state where only a subset of the moments, specified by set $\mathcal{A}$, is known. They can be solved by a semidefinite relaxation procedure. The algorithm proposed in [36] uses indeed semidefinite programming (SDP) and the concept of "extensions," introduced in [40], but based on a matrix of moments and a theorem in the theory of moment sequences. In order to present more clearly the mathematical setting for the $\mathcal{A}K$-TMS problem, we introduce a more compact notation for Eq. (5). For any N-tuple $(\mu_1, \ldots, \mu_N)$ we define a triplet $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ of integers such that $x_{\mu_1} x_{\mu_2} \ldots x_{\mu_N} = x^{\alpha}$, where we use the notation $x^{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$. The degree of the monomial x^{α} is $|\alpha| = \sum_i \alpha_i$. We then set $y_{\alpha} \equiv X_{\mu_1\mu_2...\mu_N}$. The $(y_{\alpha})_{|\alpha| \leqslant d}$ is a TMS, that is, a sequence of moments of μ truncated at degree d. When only a subset $\alpha \in \mathcal{A}$ of these moments is known, we consider the TMS $(y_{\alpha})_{\alpha \in \mathcal{A}}$. With this notation we can rewrite (5) as

$$y_{\alpha} = \int_K x^{\alpha} d\mu(\mathbf{x}). \qquad (6)$$

To a TMS y of degree d, for any integer $k \leqslant d/2$, we can associate a matrix $M_k(y)$ defined by $M_k(y)_{\alpha\beta} = y_{\alpha+\beta}$ with $|\alpha|, |\beta| \leqslant k$, which we call the kth-order moment matrix. A necessary condition for a TMS to admit a representing measure is that the moment matrix of any order be positive semidefinite. A second necessary condition can be obtained from the polynomial constraint $x_1^2 + x_2^2 + x_3^2 = 1$, which defines set K. For even degree d we define a "shifted TMS" of degree $d - 2$, and its moment matrix of order $k - 1$ is called the kth-order localizing matrix of y. It is necessarily positive semidefinite if a TMS admits a representing measure.

Beyond these two necessary conditions, a sufficient condition was obtained in [41] for an even-degree TMS. Namely, if a TMS z of even degree $2k$ is such that

$$\mathrm{rank} M_k(z) = \mathrm{rank} M_{k-1}(z), \qquad (7)$$

then the TMS z admits a representing measure. As the above condition is only sufficient, a TMS admitting a representing measure does not necessarily fulfill it, but one can always search for an extension of it which does. An extension of a TMS y of degree d is defined as any TMS z of degree $2k$ with $2k > d$, such that $z_{\alpha} = y_{\alpha}$ for all $\alpha \in \mathcal{A}$. An extension z is called flat if it satisfies Eq. (7). If z verifies the sufficient conditions above, then it has a representing measure, and so does y as a restriction of z. Then it is possible to formulate a necessary and sufficient condition for the existence of a representing measure as follows.

Theorem. A state ρ is separable if and only if its coordinates $X_{\mu_1\mu_2...\mu_N}$ are mapped to a TMS $(y_{\alpha})_{\alpha \in \mathcal{A}}$ such that there exists a flat extension $(z_{\beta})_{|\beta| \leqslant 2k}$ with $2k > d$ and whose corresponding kth-order moment and localizing matrices are positive semidefinite.

This necessary and sufficient condition can be translated into an algorithm looking for flat extensions of the TMS y associated with a quantum state ρ. One runs the algorithm with the input of the state ρ (which means fixing y_{α} for all $\alpha \in \mathcal{A}$), starting from the lowest possible extension order k. If the corresponding SDP is "infeasible," then the conditions of the theorem are not satisfied and the TMS admits no representing measure $d\mu$, which means that the quantum state whose coordinates are given by y_{α} is entangled. If, on the contrary, the SDP problem is "feasible," then the TMS admits a representing measure, and the corresponding quantum state is separable. The algorithm also extends to the case of nonsymmetric states (see [36] for further detail).

III. UNORDERED MEASUREMENTS

A. Goal

Let us now consider the question raised in Sec. I. Our goal is to identify the smallest set of measurements that should be performed on an unknown spin state to detect that it is entangled. This is possible in a real experiment when many identical copies of the same state are available, so that a different measurement can be performed on each copy. We first discuss the case of symmetric two-qubit states, which, as we see in detail, already presents some complexity. In this case the positive-partial-transpose criterion [42,43] applied to a partially known density matrix would also provide a way of detecting entanglement via SDP. Nevertheless, we use our TMS approach, since it allows for a straightforward generalization to an arbitrary number of qubits, and moreover, it applies SDP to the matrix of moments, whose entries are directly given by measurement results.

B. Symmetries and measurements

For a symmetric two-qubit state ρ, Eq. (2) with $N = 2$ gives

$$X_{\mu_1\mu_2} = \mathrm{tr}\left(\rho \sigma_{\mu_1} \otimes \sigma_{\mu_2}\right) \qquad (8)$$

with $0 \leqslant \mu_i \leqslant 3$ and $(\sigma_0, \sigma_1, \sigma_2, \sigma_3) \equiv (\mathbb{1}, \sigma_x, \sigma_y, \sigma_z)$. In this case the tensor $X_{\mu_1\mu_2}$ reduces to a 4×4 real symmetric matrix. Its 10 entries $X_{\mu_1\mu_2}$ with $\mu_1 \leqslant \mu_2$ can be seen as the result of the measurement of the joint operator $\sigma_{\mu_1} \otimes \sigma_{\mu_2}$. We now ask which are the possible measurements that we can perform and how many there are; the observables considered are the simplest, i.e., Pauli spin operators. Let us denote these inequivalent measurement operators

$$\mathcal{M} = \{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}, M_{zz}\} \quad (9)$$

(we omit the identity operator corresponding to $X_{00} = 1$, and we always order sets of measurements in degree-lexicographic order). For instance, M_x is the measurement of $\mathbb{1}$ on the first qubit and of σ_x on the second one (or the reverse), while M_{xx} is the measurement of the joint operator $\sigma_x \otimes \sigma_x$. Since the tensor $X_{\mu_1\mu_2}$ is such that

$$\sum_{i=1}^{3} X_{ii} = X_{00}, \quad (10)$$

only two of the three diagonal entries are independent, and measuring two of three of the observables M_{xx}, M_{yy}, and M_{zz} yields the third value. Thus, carrying a tomography to its end for a single spin-1 state consists in measuring eight observables in total.

Our aim is to find the probability that a state is detected as entangled if only the result of measurements of a certain subset of these eight observables is known. Let us first observe that these probabilities should not depend on the choice of the reference frame for the axes along which the measurement is performed. As a consequence, the results for equivalent measurements in different directions should be the same. We therefore consider only sets which are nonequivalent under permutation of the axes, that is, sets that are unchanged under transpositions $\{P_{xy}, P_{xz}, P_{yz}\}$, which exchange two axes, and cyclic permutations P_{yzx} and P_{zxy}.

We consider all possible nonequivalent sets of k measurements, with $1 \leqslant k \leqslant 8$, disregarding the order of measurements within a set. For sets of length $k = 1$ we can easily see that the nonequivalent measurements are only three: M_x, M_{xx}, and M_{xy}. Indeed, the local measurements M_x, M_y, and M_z are equivalent, as well as the two-qubit "diagonal" measurements M_{xx}, M_{yy}, and M_{zz} [giving the diagonal entries of matrix $(X_{\mu\nu})_{1 \leqslant \mu, \nu \leqslant 3}$] and, also, the two-qubit "off-diagonal" measurements M_{xy}, M_{xz}, and M_{yz} (giving its off-diagonal entries). For $k = 2$, there are 28 possible pairs, among which only 9 are inequivalent, namely, $\{M_x, M_y\}$, $\{M_x, M_{xx}\}$, $\{M_x, M_{xy}\}$, $\{M_x, M_{yy}\}$, $\{M_x, M_{yz}\}$, $\{M_{xx}, M_{xy}\}$, $\{M_{xx}, M_{yy}\}$, $\{M_{xx}, M_{yz}\}$, $\{M_{xy}, M_{xz}\}$. We denote by m_k the number of nonequivalent sets of k measurements, and we list them in Table I. The corresponding complete lists of measurements for all k are given in Appendix A.

For each k, our question reduces to finding out which set of measurements, among the m_k possible ones, yields the highest entanglement detection probability. Note that performing k measurements is not exactly equivalent to having k fixed moments. Indeed, since moments are related by Eq. (10), measuring M_{xx} and M_{yy} fixes the three moments X_{11}, X_{22}, and

TABLE I. First row of data: Number m_k of nonequivalent unordered sets of measurements for $1 \leqslant k \leqslant 8$. Second row of data: Number m_k of nonequivalent ordered sequences of measurements for $1 \leqslant k \leqslant 8$.

				k				
	1	2	3	4	5	6	7	8
m_k (unordered)	3	9	19	26	23	14	5	1
m_k' (ordered; M_{xx} fixed)	1	5	26	128	524	1604	3228	3228

X_{33}. Any measurement set of length k containing both M_{xx} and M_{yy} will in fact correspond to a TMS with $k + 1$ moments fixed. We therefore always discard M_{zz} from the measurement sets.

C. Set probabilities

In terms of the TMS algorithm, performing a measurement means obtaining a value of a tensor entry $X_{\mu_1\mu_2...\mu_N}$ or, equivalently, of a moment y_α. Performing k measurements means that the k moments y_α corresponding to these measurements are fixed, as well as all moments obtained via relation (3).

For a given number k of measurements, we indicate a specific set of measurements among the m_k possible ones as $\{M\}_I$. For instance, if $k = 3$, we could have $I = \{x, y, zz\}$, which corresponds to the set of measurements $\{M_x, M_y, M_{zz}\}$.

If we consider a fixed k and a fixed subset $\{M\}_I$ of the set of observables $\mathcal{M}$, we denote the sample space of outcomes of the $\mathcal{A}K$-TMS algorithm applied to the moments $(y_\alpha)_{\alpha \in \mathcal{A}}$ of an entangled state as Ω_I. It contains two possible outcomes, to which a probability can be assigned: detecting the state as entangled (if the associated SDP is infeasible, i.e., if the state is entangled), with probability $P(E, \{M\}_I)$; or *not* detecting it as entangled (if the SDP is feasible, i.e., if the state with such moments fixed is still compatible with a separable state), with probability $P(\bar{E}, \{M\}_I)$. To shorten the notation we may denote $P(E, \{M\}_I)$ as $p_I^{(k)}$, which entails $P(\bar{E}, \{M\}_I) = 1 - p_I^{(k)}$.

These probabilities can be estimated by running the TMS algorithm for each k and each I, testing all the m_k possible sets of measurements. Note that $p_I^{(k)}$ always increases, in the sense that $p_J^{(k')} \leqslant p_I^{(k)}$ for $J \subset I$. Indeed, the probability of not detecting entanglement with more and more measurements decreases with the number of measurements. In other words, fixing more moments y_α reduces the probability of finding a measure μ with such moments. Once all eight measurements are done the state is fixed uniquely, so that for entangled states $p_I^{(8)} = 1$. To estimate the values for the probabilities $p_I^{(k)}$, we sample states from the set of symmetric two-qubit states. We generated 5×10^4 random states drawn from the Hilbert-Schmidt ensemble of matrices $\rho = \frac{GG^\dagger}{\text{tr}(GG^\dagger)}$, with G a complex matrix with independent Gaussian entries (following [44]). Among them were 1843 separable states that we discarded, implying the normalization condition $p_I^{(8)} = 1$ for full tomography. For each measurement set $\{M\}_I$ and each entangled state in our sample the TMS algorithm was run with the corresponding moments fixed; the results for the probabilities $p_I^{(k)}$ are reported in Figs. 1 and 2.
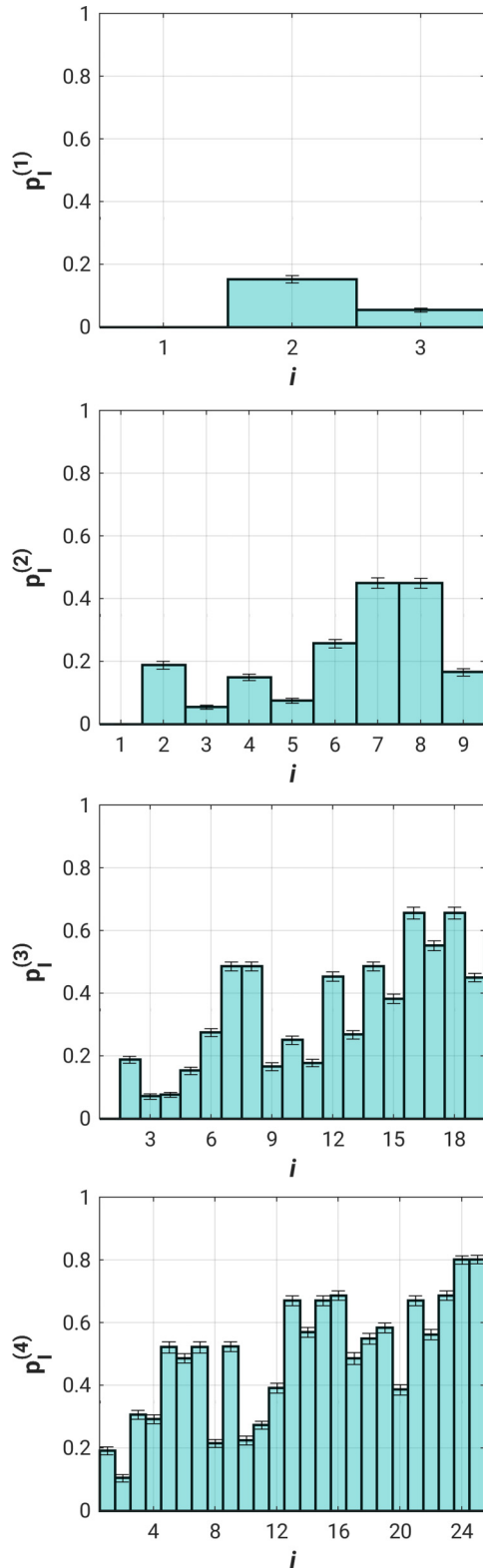
FIG. 1. Probabilities $p_I^{(k)}$ of detecting entanglement in a symmetric state of two qubits with measurement set I of cardinality k, $1 \leqslant k \leqslant 4$, as a function of the label i of set I ($1 \leqslant i \leqslant m_k$). The associated error bars represent the difference between the maximum and the minimum of the fluctuations observed for 1000 different samples of size 4×10^4 randomly extracted from the initial sample considered. The set of measurements $\{M\}_I$ corresponding to each label is listed in Appendix A.
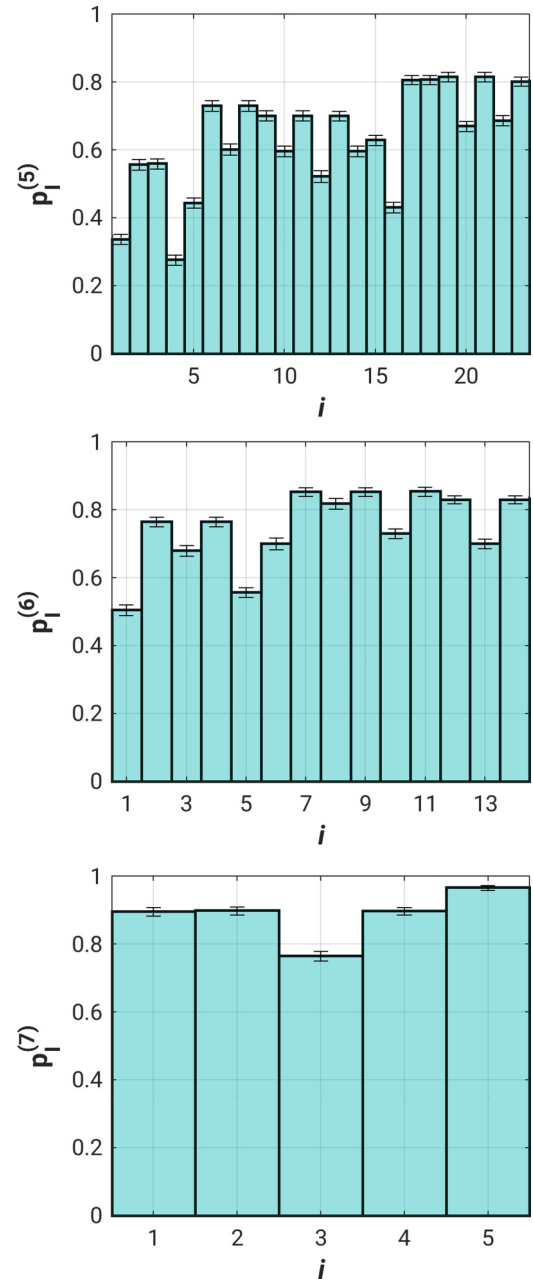

FIG. 2. Probabilities $p_I^{(k)}$ for $5 \leqslant k \leqslant 7$; same as Fig. 1.

Some probabilities appear to be equal. This is, for instance, the case for probabilities labeled 16 and 18 for $k = 3$. This is a consequence of an additional symmetry due to the linear equations that measurement results must satisfy. In the case where $k = 3$, labels 16 and 18 correspond to $\{M_{xx}, M_{xy}, M_{yy}\}$ and $\{M_{xx}, M_{xz}, M_{yy}\}$, respectively. Since, as we have mentioned, knowing the result of any two diagonal measurements gives the third one because of Eq. (10), the information acquired by measuring the observables corresponding to labels 16 and 18 is equivalent, and therefore the probabilities must be equal.

The optimal choice of measurements $\{M\}_{I_{opt}}$ at fixed k corresponds to the sets giving the highest probability of detecting entanglement. For $k = 1$ the highest value of $p_I^{(1)}$ corresponds to measurement 2, $\{M_{xx}\}$. For $k = 2$ it corresponds to 7,

$\{M_{xx}, M_{yy}\}$. For $k = 3$ the highest values correspond to two measurements: 16, $\{M_{xx}, M_{xy}, M_{yy}\}$, and 18, $\{M_{xx}, M_{xz}, M_{yy}\}$. For $k = 4$ it corresponds to 23, $\{M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$, 24, $\{M_{xx}, M_{xy}, M_{xz}, M_{yz}\}$, and 25, $\{M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$. Again, the degeneracy of the optimal set reflects the equivalence of the corresponding sets once (10) is taken into account. For $k \geqslant 2$, the sets $\{M\}_{I_{\text{opt}}}$ in fact correspond to cases where measuring two observables fixes three moments.

D. Quantumness

For a fixed set of measurements M_I one can ask whether the rate of detected entangled states depends on *how quantum* a state is. For an arbitrary state ρ, quantumness may be defined in several different ways; we follow here the definition given in [45], based on spin-coherent states. These are a generalization of the usual coherent states of the harmonic oscillator used in quantum optics to spins; they correspond to spin states which minimize a particular uncertainty relation, and they move as classical phase space points under a Hamiltonian linear in the angular momentum operators [46,47]. As any spin-1/2 pure state $|\phi\rangle$ has this property, an arbitrary N-qubit spin-coherent state can be defined as $|\phi\rangle^{\otimes N}$ with $|\phi\rangle$ a one-qubit state.

Quantumness is then defined as the Hilbert-Schmidt distance to the convex set $\mathcal{C}$ of classical spin states [48], that is, the ensemble of all density matrices which can be expressed as a mixture of spin-coherent states with positive weights (or in other words the set $\mathcal{C}$ is the convex hull of spin-coherent states). Namely, the quantumness $Q(\rho)$ is given by

$$Q(\rho) = \min_{\rho_c \in \mathcal{C}} \|\rho - \rho_c\|, \qquad (11)$$

where $\|O\| = \sqrt{\text{Tr}(O^\dagger O)}$ is the Hilbert-Schmidt norm. For all ρ the property $Q(\rho) \geqslant 0$ holds, with equality for classical states $\rho \in \mathcal{C}$. Results are reported in Fig. 3, up to $k = 4$ for the optimal sets of measurements $\{M\}_{I_{\text{opt}}}$ given above. We can observe that the rate of detected entangled states increases with the quantumness of the states; in other words, the more quantum a state is, the faster it is detected as entangled.

IV. ORDERED MEASUREMENTS

A. The setting

In the previous section we assumed that k observables are measured among the eight possible ones and that the TMS algorithm is subsequently run. Of course, we can imagine a different experimental protocol where we would perform a measurement, run the TMS algorithm with a single moment fixed, and then, only in the case where the state is not detected as entangled, perform a second measurement and run the TMS algorithm again with two moments fixed, and so on, until entanglement is detected or full tomography is achieved. In this setting, we need to distinguish the $k!$ different ordered arrangements of each k-element subset of $\mathcal{M}$.

In the following, we call an ordered sequence of measurements a *path*, and we denote it γ. To distinguish it from a set, we denote it as a tuple with parentheses, such as (M_x, M_y, M_{xz}). A path of length k can be alternatively seen as a list of k sets of increasing size given by the restriction
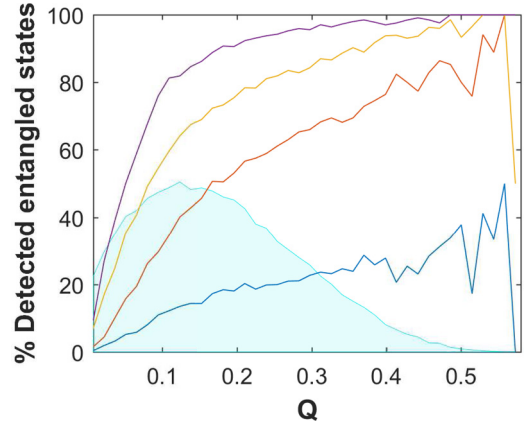

FIG. 3. Percentage of detected entangled states for the optimal sets of measurements $\{M\}_I$ for $k = 1$ to 4 (solid lines from bottom to top), as a function of the quantumness for symmetric states of two qubits. The shaded area in the background represents the distribution of the quantumness Q (bin width, 0.015) of the total number of states (multiplied by a factor of 2×10^{-2}); the first bin contains entangled states with Q between 10^{-4} and 10^{-2}. The distribution shows that there are very few states for the highest values of quantumness, which explains the large statistical errors at maximum quantumness.

of the path to the first k' observables with $1 \leqslant k' \leqslant k$. For instance, for $k = 3$ the path (M_x, M_{xz}, M_y) can be seen as the list $\{M_x\}, \{M_x, M_{xz}\}, \{M_x, M_y, M_{xz}\}$ (as usual we write sets in lexicographical order since the order within a set does not matter).

Considering all 8! paths of length 8 would require an exceedingly long computational time. For this reason, we slightly simplify the problem by fixing the first measurement to perform. The most reasonable choice, looking at the results in Fig. 1, is to fix it as a diagonal observable M_{xx}, M_{yy}, or
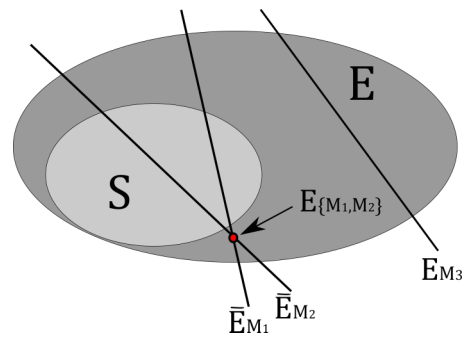

FIG. 4. Two-dimensional sketch of the sets involved. S, separable states; E, entangled states. We consider an arbitrary state in region E. Fixing one moment means restricting the set of compatible states to a hyperplane (one of the three lines in the sketch). Hyperplanes which cross the set of separable states contain both entangled and separable states, thus measuring the observable M_1 or M_2 alone is not enough to detect entanglement. Fixing both, on the other hand, restricts the set of compatible states to a region (a point in the sketch) outside S, i.e., observables $\{M_1, M_2\}$ together detect a fraction of states as entangled (E). The third line instead does not cross the set of separable states, meaning that measuring M_3 suffices to detect entanglement (which we denote E_{M_3}).

M_{zz}, since for $k = 1$ it detects the largest fraction of entangled states. Up to relabeling of the axes, we can take M_{xx} as the first element since, as before, we only keep nonequivalent paths. To find these paths, we define a canonical representation of a path γ of length k by considering its equivalent list of k sets of length k'. For each of these sets we choose the first one in lexicographical order among the ones that are obtained by relabeling of the axes. The list of k sets obtained in this way is the canonical representation of γ. Two paths are equivalent if they have the same canonical representation. We report the number m'_k of nonequivalent paths of length k in Table I, where, e.g., for $k = 2$ the nonequivalent sequences will be (M_{xx}, M_x), (M_{xx}, M_y), (M_{xx}, M_{xy}), (M_{xx}, M_{yy}), and (M_{xx}, M_{yz}).

B. Path probabilities

We now show how to retrieve the results for this more general case from the $p_I^{(k)}$ obtained in the previous section. The probability of detecting a state as entangled after the first measurement, say M_1, is $P(E, \{M_1\})$, given in the previous section. The probability of detecting a state as entangled after the second measurement, say M_2, is then $P(E, \{M_1, M_2\}|\bar{E}, \{M_1\})$, which is the probability of detecting entanglement with the second measurement given that it was not detected with the first one. This quantity now depends on which measurement is performed first. This is illustrated in Fig. 4. Using the theorem of total probability, we have

$$P(E, \{M_1, M_2\}) = P(E, \{M_1\})P(E, \{M_1, M_2\}|E, \{M_1\})$$
$$+ P(\bar{E}, \{M_1\})P(E, \{M_1, M_2\}|\bar{E}, \{M_1\}).$$
(12)

Since $P(\bar{E}, \{M_1\}) = 1 - P(E, \{M_1\})$ and $P(E, \{M_1, M_2\}|E, M_1) = 1$ we get

$$P(E, \{M_1, M_2\}|\bar{E}, \{M_1\}) = \frac{P(E, \{M_1, M_2\}) - P(E, \{M_1\})}{1 - P(E, \{M_1\})}.$$
(13)

Thus, the conditional probability we are looking for can be expressed solely in terms of the $p_I^{(k)}$ from the previous section.

Then let $\gamma = (M_1, \ldots, M_8)$ be a path of length $k = 8$. We define

$$q^{(k)}(\gamma) = P(E, \{M_1, \ldots, M_k\}|\bar{E}, \{M_1, \ldots, M_{k-1}\}) \quad (14)$$

as the probability of detecting entanglement at step k in γ given that no entanglement was detected up to step $k - 1$. By a reasoning similar to the one leading to Eq. (13), we can express $q^{(k)}(\gamma)$ in terms of the $p^{(k)}(\gamma) \equiv P(E, \{M_1, \ldots, M_k\})$, as

$$q^{(k)}(\gamma) = \frac{p^{(k)}(\gamma) - p^{(k-1)}(\gamma)}{1 - p^{(k-1)}(\gamma)}. \quad (15)$$

In particular, since $p^{(0)} = 0$ (as nothing is measured, and hence detected as entangled, at level 0), we have $q^{(1)} = p^{(1)}$. Inverting (15) one obtains $p^{(k)}(\gamma)$ in terms of $q^{(k)}(\gamma)$ as $p^{(k)}(\gamma) = \sum_{j=1}^{k} q^{(j)}(\gamma) \prod_{n=j+1}^{k} (1 - q^{(n)}(\gamma))$.

A third natural probability to consider is related to our measurement algorithm, where we perform TMS calculations at step k only if the state was compatible with a
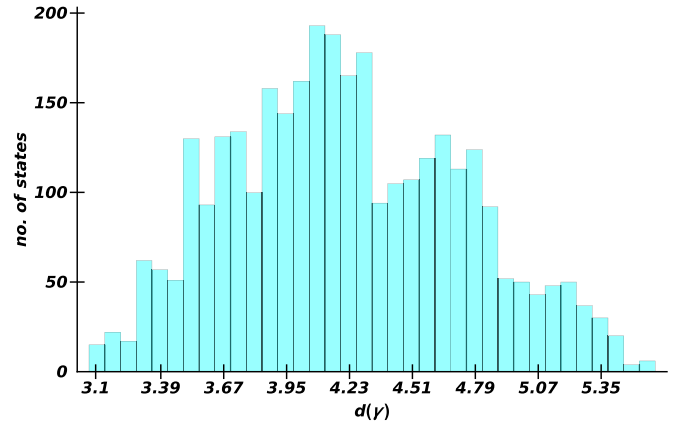

FIG. 5. Distribution (with bin width 0.07) of lengths $d(\gamma)$ of measurement sequences γ of symmetric states of two qubits resulting in detection of entanglement between the minimum value of 3.07 and the maximum one of 5.61.

separable state. We define $r^{(k)}(\gamma)$ as the probability of stopping exactly at the kth level when measurements are taken along path γ. It can be written as the joint probability $P(E, \{M_1, \ldots, M_k\} \cap \bar{E}, \{M_1, \ldots, M_{k-1}\})$. Using the identity $P(A \cap B) = P(A|B)P(B)$, $r^{(k)}(\gamma)$ can be expressed as $q^{(k)}(\gamma)(1 - p^{(k-1)}(\gamma))$. It can be rewritten in terms of $q^{(k)}(\gamma)$ or $p^{(k)}(\gamma)$ as

$$r^{(k)}(\gamma) = q^{(k)}(\gamma) \prod_{j=1}^{k-1} (1 - q^{(j)}(\gamma)) = p^{(k)}(\gamma) - p^{(k-1)}(\gamma).$$

(16)

C. The best path

Using (15) and (16) and the numerical results in the previous section, we can obtain a numerical estimate of the $q^{(k)}(\gamma)$ and the $r^{(k)}(\gamma)$ for all possible paths. The optimal path γ_{best} is the one that detects as quickly as possible (on average) whether the state is entangled. To identify γ_{best} among all possible ones we define the average depth at which our algorithm stops as

$$d(\gamma) = \sum_{k=1}^{8} k r^{(k)}(\gamma). \quad (17)$$

Expressing Eq. (17) in words, $d(\gamma)$ gives the number of measurements that one needs to perform, on average, to detect a state as entangled, following the path γ. Each path will be characterized by this number, and in particular, the shortest path will be given by

$$\gamma_{\text{best}} = \arg \min_{\gamma \in S} d(\gamma). \quad (18)$$

The distribution of $d(\gamma)$ over all 3228 paths of length 8 for symmetric states of two qubits is reported in Fig. 5. The minimum value found for $d(\gamma)$ is $d = 3.07$, while the maximum value is 5.61. The minimum value is degenerate and corresponds to three optimal paths. Although these three paths do not have the same canonical representation they lead to the same value because of condition (10).
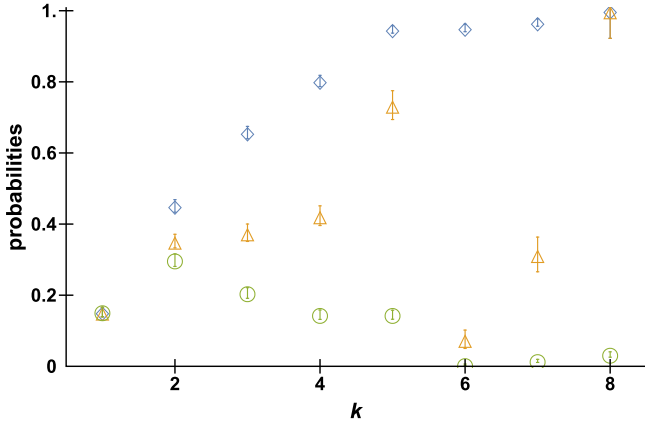
FIG. 6. Probabilities $p^{(k)}(\gamma_{\text{best}})$ (blue diamonds), $q^{(k)}(\gamma_{\text{best}})$ (orange triangles), and $r^{(k)}(\gamma_{\text{best}})$ (green circles). Error bars represent the statistical errors derived from those of the $p_I^{(k)}$; see Fig. 1.

If one considered that knowing two diagonal moments is equivalent to knowing them all and included in the symmetrization the third diagonal moment once the first two are measured, there would be a unique optimal path. We report here one of the three equivalent optimal paths, $\gamma_{\text{best}} = (M_{xx}, M_{yy}, M_{xz}, M_{yz}, M_{xy}, M_x, M_y, M_z)$; choosing this path, one only needs to perform (on average) three measurements to detect a state as entangled. These three measurements give access to the two diagonal moments [and thus all of them via (10)] and one of the off-diagonal ones. The probabilities relative to this best path are shown in Fig. 6.

Rewriting $d(\gamma)$ in terms of $p^{(k)}(\gamma)$ we get

$$d(\gamma) = 8p^{(8)}(\gamma) - p^{(7)}(\gamma) - p^{(6)}(\gamma) - \ldots - p^{(1)}(\gamma). \quad (19)$$

It turns out that choosing measurements according to γ_{best} coincides (within the error bars) with choosing for each k, $1 \leqslant k \leqslant 8$, the best set of measurements, i.e., the one with the highest probability of detecting entanglement at a given level (highest $p_I^{(k)}$ among the m_k possibilities for each k). This is not obvious, and it is not always the case: a counterexample is given by a binary tree of depth 4 in which the random probabilities satisfy the same constraint as in our case, i.e., $p^{(k-1)}(\gamma) \leqslant p^{(k)}(\gamma)$, and the four paths have probabilities $(0.57, 0.62, 0.76)$, $(0.57, 0.62, 0.95)$, $(0.57, 0.68, 0.77)$, $(0.57, 0.68, 0.78)$. It is easily verified that the best path, with $d(\gamma) = 1.8$, is the second one, which at depth 2 does not have the highest $p^{(2)}(\gamma)$, so the minimal $d(\gamma)$ does not always correspond to the path with the highest $p^{(k)}(\gamma)$ at each step.

In practice, joint measurements such as M_{xx} might be more challenging to implement than two single measurements M_x, as qubits need first a unitary operation to entangle them and then two local measurements. In such a case, one might modify (17) with another factor for each path that takes such additional costs into account. Also, we base our analysis on average values of measurement outcomes which we took as known with arbitrary precision. This is, of course, an idealization. In practice, only a finite number of measurements can be performed, leading to statistical error bars for each moment. These can, in principle, be taken into account in

TABLE II. Number m_k of nonequivalent unordered sets of measurements of two qubits for $1 \leqslant k \leqslant 15$.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
m_k	3	10	30	69	132	205	254	254	205	132	69	30	10	3	1

the TMS algorithm, but they increase the computational time. Both of these points are beyond the scope of the present paper.

V. NONSYMMETRIC CASE

So far we have restricted ourselves to symmetric states of two qubits. Let us now consider the generic case of arbitrary two-qubit nonsymmetric states. In this case we can still exploit the TMS algorithm, with the following differences [36]. The bipartite state acts on the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of Hilbert spaces and each of them now has its own set of variables x, y, and z; we label these variables (x_i, y_i, z_i), with $i = 1, 2$. The compact set K is now the product of two Bloch spheres. The set $\mathcal{M}$ of possible measurements is

$$\mathcal{M} = \big\{ M_{x_1}, M_{y_1}, M_{z_1}, M_{x_2}, M_{x_1 x_2}, M_{y_1 x_2}, M_{z_1 x_2}, M_{y_2},$$
$$M_{x_1 y_2}, M_{y_1 y_2}, M_{z_1 y_2}, M_{z_2}, M_{x_1 z_2}, M_{y_1 z_2}, M_{z_1 z_2} \big\}. \quad (20)$$

For example, M_{x_1} is the measurement of $\sigma_x \otimes \mathbb{1}$ and $M_{x_1 x_2}$ is the measurement of the joint operator $\sigma_x \otimes \sigma_x$. Up to relabeling the variables for each qubit, some sets of measurement operators are equivalent. The number m_k of nonequivalent sets of measurements is obtained by applying the 36 possible permutations on the (x_i, y_i, z_i). This number is reported in Table II for $1 \leqslant k \leqslant 15$.

The number m_k increases rapidly with k, and so does the size of the moment matrices considered in the TMS algorithm: indeed, because of condition (7), the algorithm always searches at least for the first extension; in both cases (symmetric and nonsymmetric) the smallest extension corresponds to the moment matrix of order 2. In the symmetric case it is a 10×10 matrix, while in the nonsymmetric case it already becomes a 28×28 matrix which contains all the monomials up to degree 4 for the set of six variables x_i, y_i, z_i, with $i = 1, 2$, i.e., 210 moments, versus 35 in the symmetric case. For the previous reasons computational times become an issue in the nonsymmetric case. Nevertheless, we could estimate probabilities up to $k = 5$, running the TMS algorithm over a database of 50 000 nonsymmetric two-qubit random states. What we observe is that no state is detected as entangled with only one measurement, a tiny fraction($\sim$1%) is detected as entangled by the combination of two measurements $\{M_{x_1 x_2}, M_{y_1 y_2}\}$, and the largest fraction of states detected as entangled for $3 \leqslant k \leqslant 5$ is given, respectively, by the set of measurements $\{M_{x_1 x_2}, M_{y_1 y_2}, M_{z_1 z_2}\}$ ($\sim$10%), $\{M_{x_1 x_2}, M_{x_1 y_2}, M_{y_1 x_2}, M_{z_1 z_2}\}$ ($\sim$12%), $\{M_{x_1 x_2}, M_{x_1 y_2}, M_{y_1 x_2}, M_{y_1 y_2}, M_{z_1 z_2}\}$ ($\sim$23%). This is a big difference compared to the symmetric case, in which we could detect $\sim$15% of the states as entangled with a single measurement, $\sim$40% already with two measurements, and almost all states with five measurements.

VI. HIGHER SPIN-j

Going back to the case of symmetric states, we can also get an idea of how complexity changes for higher spin sizes; indeed, the size of the set $\mathcal{M}$ in the symmetric case corresponds to the sum of the number of monomials in three variables up to degree $d = 2j + 1$, where j is the spin size. These numbers form the sequence of triangular numbers $T_n = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$; we can then write that m_k for any spin-j is

$$m_k = \binom{\sum_{n=1}^{2j+1} T_n - 1}{k}, \tag{21}$$

where we subtract 1 since the first element of $\mathcal{M}$ is always the identity. However, in this case, we can still have some information looking at the expression for the tensor representation of a separable state in (4). Indeed, for an even number of qubits (integer spins) we can look at the diagonal tensor entries, which are defined as the entries of the form $X_{\mu_1...\mu_j\mu_1...\mu_j}$ with $0 \leqslant \mu_i \leqslant 3$. These correspond to terms of the form $\sum_j \omega_j (n_{\mu_1}...n_{\mu_j})^{2j}$; it follows that for a separable state these entries are positive, since the n_{μ_i} are real and $\omega_j \geqslant 0$. Therefore measuring a negative value for any of the corresponding measurement operators means detecting entanglement; we indicate the operators corresponding to the diagonal entries of the tensor $X_{\mu_1\mu_2...\mu_N}$ with $\{D\}_I$. We can then restrict our investigation for an integer spin-j to these 4^j observables, which are further reduced by symmetry to $\binom{j+3}{3}$. We report in Fig. 7 the number of entangled states that are not detected by any of the observables $\{D\}_I$ for spin size $1 \leqslant j \leqslant 5$ (for each j we used a sample of 10^6 random states from which we again removed the separable ones). The number of undetected entangled states decreases with the spin size j and already at $j = 4$ all the states in the sample are detected; we can also observe that restricting the analysis to these observables already gives significant information for spin-1 and spin-2 and almost-complete information for spin-3. Moreover, we can also compare these observables to see which is the most efficient measurement to perform as we did for the spin-1 case; to estimate the corresponding $p_I^{(1)}$, we again only consider sets which are nonequivalent
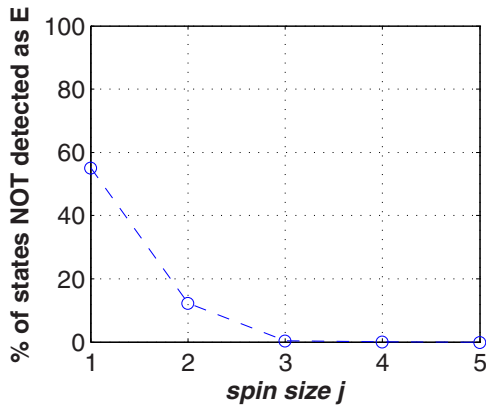

FIG. 8. Comparison of the nonequivalent diagonal observables $\{D\}_I$ for spin-j, $2 \leqslant j \leqslant 5$: probabilities $p_I^{(1)}$ as a function of the label i of set I. The highest values are reached, respectively, for D_{xxyy}, D_{xxyyzz}, $D_{xxxxxxyy}$, and $D_{xxxxyyyy}$ (where the last term corresponds to the measurement of $\mathbb{1}^{\times 2} \otimes \sigma_x^{\otimes 4} \otimes \sigma_y^{\otimes 4}$; see Appendix B for the full list).

under permutation of the axes, performing the transformations $\{P_{xy}, P_{xz}, P_{yz}, P_{yzx}, P_{zxy}\}$ described in Sec. III B. The results are shown in Fig. 8. The question arises whether similarly efficient measurements can be found for half-integer spin-j. It was recently shown in [49] how the positive-partial-transpose separability criterion for symmetric states of multiqubit systems can be formulated in terms of matrix inequalities based on the tensor representation in Eq. (2). It is possible to construct a matrix T from the tensor representation of the state and show that it is similar to the partial transpose of the density matrix written in the computational basis. In the case of spin-3/2 this matrix is an 8×8 Hermitian matrix given

FIG. 7. Percentage of entangled states not detected by any of the negative outcomes of the measurements $\{D\}_I$ corresponding to diagonal entries of the tensor $X_{\mu_1\mu_2...\mu_N}$ as a function of the spin size j.
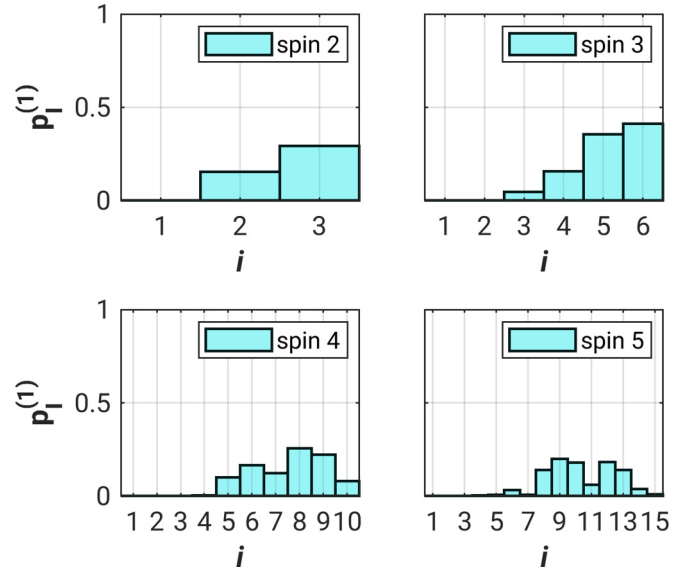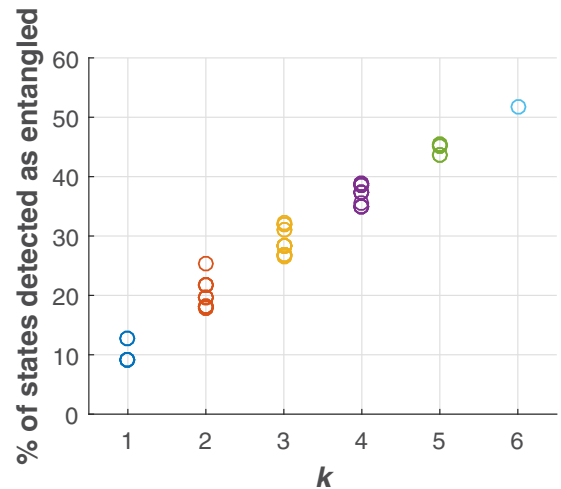

FIG. 9. Entanglement detection probabilities based on the negativity of the $\binom{6}{k}$ subsets of the set $\{X_{011} - X_{113}, X_{011} + X_{113}, X_{022} - X_{223}, X_{022} + X_{223}, X_{033} - X_{333}, X_{033} + X_{333}\}$ for $k = 1, \ldots, 6$, where the tensor $X_{\mu_1\mu_2\mu_3}$ represents a spin-3/2.

by $T_{\mu i,\nu i'} = \sum_{\tau=0}^{3} X_{\tau\mu\nu}\sigma_{i,i'}^{\tau}$, where $\sigma_{i,i'}^{\tau}$ are the Pauli-matrix components, and its positivity is a necessary and sufficient classicality criterion; as a consequence, the positivity of the diagonal entries is a necessary condition for a separable state. We can again restrict our investigation to the corresponding observables $\{D\}_I$, but this time it implies the measurement of sets of two observables. Indeed, in terms of the tensor entries $X_{\mu_1\mu_2\mu_3}$, the diagonal entries in T are $X_{000} \pm X_{003}$, $X_{011} \pm X_{113}$, $X_{022} \pm X_{223}$, $X_{033} \pm X_{333}$, so we need to compare pairs of outcomes. Recalling that $X_{000} = 1$, we can neglect the first entry, since the condition $-1 \leqslant X_{003} \leqslant 1$ is always satisfied. The results of such investigation for the other six pairs and for their combinations [all the $\binom{6}{k}$ sets, with $2 \leqslant k \leqslant 6$] are reported in Fig. 9. As before, we can gain already relevant information from this restricted analysis.

VII. CONCLUSIONS

In summary, we have studied the statistics of lengths of measurement sequences for multiqubit systems that allow one to detect entanglement without any prior information about the state, for both unordered sets of measurements and ordered ones (i.e., measurement paths). For symmetric states of two qubits, we have identified the best measurement path that results, on average over all randomly chosen entangled states, in a proof of entanglement with 3.07 measurements (compared to the 8 measurements needed for full tomography in this case). For larger numbers N of qubits in symmetric states, we found that measurements based on the diagonal matrix elements of the moment matrix of the state become very efficient in detecting entanglement. Their number increases like N^3, and already at $N = 8$ qubits the number of states *not* detected as entangled has decreased to about 10^{-6} or smaller. For nonsymmetric states, substantially larger numbers of measurements are needed to detect entanglement with certainty: at least two measurements are needed for two-qubit states, resulting in only about a 1% detection probability, however. With five measurements the probability increases to about 23%. The work is based on the truncated moment sequence algorithm that naturally allows one to deal with missing data. It is very flexible and can be easily adapted to experimentally relevant ensembles of states and other side conditions, such as sets of measurements that can be implemented or more elaborate cost functions.

ACKNOWLEDGMENTS

APPENDIX A: UNORDERED MEASUREMENT SETS

We list here all the m_k unique sets of k measurements for $1 \leqslant k \leqslant 8$.

$k = 1$	
1	M_x
2	M_{xx}
3	M_{xy}

(Continued).

$k = 1$	
$k = 2$	
1	$\{M_x, M_y\}$
2	$\{M_x, M_{xx}\}$
3	$\{M_x, M_{xy}\}$
4	$\{M_x, M_{yy}\}$
5	$\{M_x, M_{yz}\}$
6	$\{M_{xx}, M_{xy}\}$
7	$\{M_{xx}, M_{yy}\}$
8	$\{M_{xx}, M_{yz}\}$
9	$\{M_{xy}, M_{xz}\}$
$k = 3$	
1	$\{M_x, M_y, M_z\}$
2	$\{M_x, M_y, M_{xx}\}$
3	$\{M_x, M_y, M_{xy}\}$
4	$\{M_x, M_y, M_{xz}\}$
5	$\{M_x, M_z, M_{yy}\}$
6	$\{M_x, M_{xx}, M_{xy}\}$
7	$\{M_x, M_{xx}, M_{yy}\}$
8	$\{M_x, M_{xx}, M_{yz}\}$
9	$\{M_x, M_{xy}, M_{xz}\}$
10	$\{M_x, M_{xy}, M_{yy}\}$
11	$\{M_x, M_{xy}, M_{yz}\}$
12	$\{M_x, M_{xz}, M_{yy}\}$
13	$\{M_x, M_{yy}, M_{yz}\}$
14	$\{M_z, M_{xx}, M_{yy}\}$
15	$\{M_{xx}, M_{xy}, M_{xz}\}$
16	$\{M_{xx}, M_{xy}, M_{yy}\}$
17	$\{M_{xx}, M_{xy}, M_{yz}\}$
18	$\{M_{xx}, M_{xz}, M_{yy}\}$
19	$\{M_{xy}, M_{xz}, M_{yz}\}$
$k = 4$	
1	$\{M_x, M_y, M_z, M_{xx}\}$
2	$\{M_x, M_y, M_z, M_{xy}\}$
3	$\{M_x, M_y, M_{xx}, M_{xy}\}$
4	$\{M_x, M_y, M_{xx}, M_{xz}\}$
5	$\{M_x, M_y, M_{xx}, M_{yy}\}$
6	$\{M_x, M_y, M_{xx}, M_{yz}\}$
7	$\{M_x, M_z, M_{xx}, M_{yy}\}$
8	$\{M_x, M_y, M_{xy}, M_{xz}\}$
9	$\{M_x, M_z, M_{xz}, M_{yy}\}$
10	$\{M_x, M_y, M_{xz}, M_{yz}\}$
11	$\{M_x, M_z, M_{xy}, M_{yy}\}$
12	$\{M_x, M_{xx}, M_{xy}, M_{xz}\}$
13	$\{M_x, M_{xx}, M_{xy}, M_{yy}\}$
14	$\{M_x, M_{xx}, M_{xy}, M_{yz}\}$
15	$\{M_x, M_{xx}, M_{xz}, M_{yy}\}$
16	$\{M_x, M_{xx}, M_{yy}, M_{yz}\}$
17	$\{M_x, M_{xy}, M_{xz}, M_{yy}\}$
18	$\{M_x, M_{xy}, M_{xz}, M_{yz}\}$
19	$\{M_x, M_{xy}, M_{yy}, M_{yz}\}$
20	$\{M_z, M_{xx}, M_{yy}, M_{yz}\}$
21	$\{M_x, M_{xz}, M_{yy}, M_{yz}\}$
22	$\{M_z, M_{xx}, M_{xz}, M_{yy}\}$
23	$\{M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$
24	$\{M_{xx}, M_{xy}, M_{xz}, M_{yz}\}$
25	$\{M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
$k = 5$	
1	$\{M_x, M_y, M_z, M_{xx}, M_{xy}\}$
2	$\{M_x, M_y, M_z, M_{xx}, M_{yy}\}$

<div style="display:flex">

(Continued).

$k = 5$	
3	$\{M_x, M_y, M_z, M_{xx}, M_{yz}\}$
4	$\{M_x, M_y, M_z, M_{xy}, M_{xz}\}$
5	$\{M_x, M_y, M_{xx}, M_{xy}, M_{xz}\}$
6	$\{M_x, M_y, M_{xx}, M_{xy}, M_{yy}\}$
7	$\{M_x, M_y, M_{xx}, M_{xy}, M_{yz}\}$
8	$\{M_x, M_z, M_{xx}, M_{xz}, M_{yy}\}$
9	$\{M_x, M_y, M_{xx}, M_{xz}, M_{yy}\}$
10	$\{M_x, M_y, M_{xx}, M_{xz}, M_{yz}\}$
11	$\{M_x, M_z, M_{xx}, M_{xy}, M_{yy}\}$
12	$\{M_x, M_z, M_{xx}, M_{yy}, M_{yz}\}$
13	$\{M_x, M_y, M_{xy}, M_{xz}, M_{yz}\}$
14	$\{M_x, M_z, M_{xy}, M_{xz}, M_{yy}\}$
15	$\{M_x, M_z, M_{xy}, M_{yy}, M_{yz}\}$
16	$\{M_x, M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$
17	$\{M_x, M_{xx}, M_{xy}, M_{xz}, M_{yz}\}$
18	$\{M_x, M_{xx}, M_{xy}, M_{yy}, M_{yz}\}$
19	$\{M_x, M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
20	$\{M_x, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
21	$\{M_z, M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
22	$\{M_z, M_{xx}, M_{xy}, M_{yy}, M_{yz}\}$
23	$\{M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
$k = 6$	
1	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{xz}\}$
2	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{yy}\}$
3	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{yz}\}$
4	$\{M_x, M_y, M_z, M_{xx}, M_{xz}, M_{yy}\}$
5	$\{M_x, M_y, M_z, M_{xy}, M_{xz}, M_{yz}\}$
6	$\{M_x, M_y, M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$
7	$\{M_x, M_y, M_{xx}, M_{xy}, M_{xz}, M_{yz}\}$
8	$\{M_x, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$
9	$\{M_x, M_z, M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
10	$\{M_x, M_y, M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
11	$\{M_x, M_z, M_{xx}, M_{xy}, M_{yy}, M_{yz}\}$
12	$\{M_x, M_z, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
13	$\{M_x, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
14	$\{M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
$k = 7$	
1	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}\}$
2	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yz}\}$
3	$\{M_x, M_y, M_z, M_{xx}, M_{xz}, M_{yy}, M_{yz}\}$
4	$\{M_x, M_y, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
5	$\{M_x, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$
$k = 8$	
1	$\{M_x, M_y, M_z, M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}\}$

APPENDIX B: NONEQUIVALENT DIAGONAL OBSERVABLES

We list here all the nonequivalent observables D_I for spin-j, $2 \leqslant j \leqslant 5$.

$j = 2$	
1	D_{xx}
2	D_{xxxx}
3	D_{xxyy}
$j = 3$	
1	D_{xx}
2	D_{xxxx}
3	D_{xxyy}
4	D_{xxxxxx}
5	D_{xxxxyy}
6	D_{xxyyzz}
$j = 4$	
1	D_{xx}
2	D_{xxxx}
3	D_{xxyy}
4	D_{xxxxxx}
5	D_{xxxxyy}
6	D_{xxyyzz}
7	$D_{xxxxxxxx}$
8	$D_{xxxxxxyy}$
9	$D_{xxxxyyyy}$
10	$D_{xxxxyyzz}$
$j = 5$	
1	D_{xx}
2	D_{xxxx}
3	D_{xxyy}
4	D_{xxxxxx}
5	D_{xxxxyy}
6	D_{xxyyzz}
7	$D_{xxxxxxxx}$
8	$D_{xxxxxxyy}$
9	$D_{xxxxyyyy}$
10	$D_{xxxxyyzz}$
11	$D_{xxxxxxxxxx}$
12	$D_{xxxxxxxxyy}$
13	$D_{xxxxxxyyyy}$
14	$D_{xxxxxxyyzz}$
15	$D_{xxxxyyyyzz}$

</div>

[1] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).

[2] R. Blume-Kohout, Phys. Rev. Lett. **105**, 200504 (2010).

[3] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).

[4] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio, New J. Phys. **15**, 125004 (2013).

[5] R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).

[6] J. Rau, Phys. Rev. A **82**, 012104 (2010).

[7] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Corcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, Phys. Rev. A **87**, 062119 (2013).

[8] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).

[9] M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert, arXiv:1111.0853.

[10] M. Ohliger, V. Nesme, and J. Eisert, New J. Phys. **15**, 015024 (2013).

[11] M. Kliesch, R. Kueng, J. Eisert, and D. Gross, arXiv:1701.03135.

[12] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, New J. Phys. **14**, 095022 (2012).

[13] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).

[14] B. P. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. S. Buyskikh, A. J. Daley, M. Cramer *et al.*, Nat. Phys. **13**, 1158 (2017).

[15] C. Ferrie, Phys. Rev. Lett. **113**, 190404 (2014).

[16] B. Wang, arXiv:1709.03617.

[17] S. Lu, S. Huang, K. Li, J. Li, J. Chen, D. Lu, Z. Ji, Y. Shen, D. Zhou, and B. Zeng, Phys. Rev. A **98**, 012315 (2018).

[18] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).

[19] V. Saggio, A. Dimić, C. Greganti, L. A. Rozema, P. Walther, and B. Dakić, arXiv:1809.05455.

[20] A. Dimić and B. Dakić, npj Quantum Info. **4**, 11 (2018).

[21] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).

[22] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki *et al.*, Phys. Rev. Lett. **109**, 080505 (2012).

[23] J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried *et al.*, Phys. Rev. Lett. **108**, 260503 (2012).

[24] G. Chiribella, G. M. D'Ariano, and M. Roetteler, New J. Phys. **15**, 103019 (2013).

[25] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).

[26] L. Arnaud and D. Braun, Phys. Rev. A **75**, 062314 (2007).

[27] C. T. Schmiegelow, A. Bendersky, M. A. Larotonda, and J. P. Paz, Phys. Rev. Lett. **107**, 100502 (2011).

[28] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. A **80**, 032116 (2009).

[29] A. I. Lvovsky and M. G. Raymer, Rev. Mod. Phys. **81**, 299 (2009).

[30] J. G. Titchener, M. Gräfe, R. Heilmann, A. S. Solntsev, A. Szameit, and A. A. Sukhorukov, npj Quantum Info. **4**, 19 (2018).

[31] H. Häffner, W. Hansel, C. F. Roos, J. Benhelm, D. Chek-al kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. Schmidt *et al.*, Nature **438**, 643 (2005).

[32] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hansel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).

[33] T. S. Iskhakov, I. N. Agafonov, M. V. Chekhova, and G. Leuchs, Phys. Rev. Lett. **109**, 150502 (2012).

[34] R. McConnell, H. Zhang, J. Hu, S. Ćuk, and V. Vuletić, Nature **519**, 439 (2015).

[35] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, Science **352**, 1297 (2016).

[36] F. Bohnet-Waldraff, D. Braun, and O. Giraud, Phys. Rev. A **96**, 032312 (2017).

[37] J. Nie, Found. Comput. Math. **14**, 1243 (2014).

[38] A. Mari, K. Kieling, B. M. Nielsen, E. S. Polzik, and J. Eisert, Phys. Rev. Lett. **106**, 010403 (2011).

[39] O. Giraud, D. Braun, D. Baguette, T. Bastin, and J. Martin, Phys. Rev. Lett. **114**, 080401 (2015).

[40] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **71**, 032333 (2005).

[41] L. Fialkow and J. Nie, J. Funct. Anal. **263**, 1682 (2012).

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

[43] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[44] K. Zyczkowski, K. A. Penson, I. Nechita, and B. Collins, J. Math. Phys. **52**, 062201 (2011).

[45] O. Giraud, P. Braun, and D. Braun, New J. Phys. **12**, 063005 (2010).

[46] R. H. Dicke, Phys. Rev. **93**, 99 (1954).

[47] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6**, 2211 (1972).

[48] O. Giraud, P. Braun, and D. Braun, Phys. Rev. A **78**, 042112 (2008).

[49] F. Bohnet-Waldraff, D. Braun, and O. Giraud, Phys. Rev. A **94**, 042343 (2016).