Entanglement estimation in tensor network states via sampling

Noa Feldman,1 Augustine Kshetrimayum,2,3 Jens Eisert,2,3 and Moshe Goldstein1

1Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, 6997801 Tel Aviv, Israel
2Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany
3Helmholtz Center Berlin, 14109 Berlin, Germany

We introduce a method for extracting meaningful entanglement measures of tensor network states in general dimensions. Current methods require the explicit reconstruction of the density matrix, which is highly demanding, or the contraction of replicas, which requires an effort exponential in the number of replicas and which is costly in terms of memory. In contrast, our method requires the stochastic sampling of matrix elements of the classically represented reduced states with respect to random states drawn from simple product probability measures constituting frames. Even though not corresponding to physical operations, such matrix elements are straightforward to calculate for tensor network states, and their moments provide the Rényi entropies and negativities as well as their symmetry-resolved components. We test our method on the one-dimensional critical XX chain and the two-dimensional toric code in a checkerboard geometry. Although the cost is exponential in the subsystem size, it is sufficiently moderate so that - in contrast with other approaches - accurate results can be obtained on a personal computer for relatively large subsystem sizes.

I. INTRODUCTION

Entanglement is the key feature of quantum mechanics that renders it different from classical theories. It takes centre stage in quantum information processing where it plays the role of a resource. The significance of notions of entanglement for capturing properties of condensed matter systems has also long been noted and appreciated. The observation that ground states of gapped phases of matter are expected to feature little entanglement – in fact, they feature what are called area laws for entanglement entropy – is at the basis of tensor networks (TN) methods for accurately describing interacting quantum many-body systems. It has been noted that certain scalings of entanglement measures can indicate the presence of quantum phase transitions. Indeed, the very fact that locally interacting quantum many-body systems tend to be much less entangled than they could possibly be renders TN methods a powerful technique to capture their properties. Maybe most prominently, topologically ordered systems can be regarded as long ranged entangled systems. In addition, detailed information about the scaling of entanglement properties can provide substantial diagnostic information about properties of condensed matter systems.

Accepting that tensor network states often provide an accurate and efficient classical description of interacting quantum systems, the question arises how one can meaningfully read off entanglement properties from tensor network states. This, however, constitutes a challenge. Current entanglement calculation methods in tensor network states in two and higher dimensions are highly impractical even for moderate-size systems, since they require a full reconstruction of the quantum states at heavy computational costs. For Rényi entropies one may instead employ the replica trick, which uses several copies of the RDM (as explained in Section below); this, however, comes with an exponential scaling of the computational effort in the number of copies, often making the calculation unfeasible.

In this work, we develop a method for estimating the entanglement moments of general states represented by tensor network states. We do so by bringing together ideas of tensor network methods with those of random measurement and shadow estimation. In this context, it has been understood that entanglement features can be reliably estimated from expectation values of suitable random measurements.

Here, we bring these ideas to a new level by applying them to quantum states that are classically represented in the first place by tensor networks. The core idea of these methods is basically the following: While the entanglement moments naively require several copies of the system, we can refrain from this requirement by resorting to random sampling. The general protocol is to evolve the system under a random unitary drawn from the Haar measure followed by a measurement of a suitable projector. The process is repeated and moments of the results are averaged over different unitaries, giving as a result entanglement moments or other density-matrix-based properties. The effectiveness of this mindset has been demonstrated experimentally in a number of platforms, including that of cold atoms for Rényi entropies and Rényi negativities.

While these ideas have been further developed into estimation techniques giving rise to classical representations in their own right, we turn these ideas upside down by applying them to quantum systems that are already classically represented by tensor networks. There are some crucial differences that arise in classical representations compared to quantum experiments: First, they are much more suitable for a direct calculation of expectation values, rather than estimating them from sampling from measurements. Second, and importantly, when performing a classical simulation, we are not limited to physically-allowed actions, and specifically, we are not constrained to the application of unitary operators and measurements. This feature is to an extent reminiscent of shadow estimation in that also there, unphysical maps are made use of. It is the estimation procedure itself that is not physical here, however. The method we develop allows for having only a single copy of the simulated state, and at the same time for estimating the entanglement moments based on matrix elements that are naturally calculated. Instead of sampling operators from the Haar measure or some unitary design, we only
need to sample from a simple, finite set of tensor products of independent $d$-dimensional vectors – specifically from what are called frames or spherical complex 1-designs, where $d$ is the Hilbert space of a single site in the system. Furthermore, this simple structure allows our method to be applied to arbitrary system geometries.

The remainder of this work is organized as follows. Section II includes preliminary theoretical background. The Rényi moments we aim to estimate are defined and their relation to standard entanglement measures is discussed in Section II A. Section II B covers the basic ideas of the TN ansatzes we use for estimating entanglement moments of states represented by iPEPS. We discuss the algorithms we used for extracting the reduced density matrix and the naïve method for estimating entanglement moments of states represented by these ansatzes. The solvable models used as benchmarks for testing our method are presented in Section II C. In Section III we explain our proposed algorithm for using random variables for estimating the entanglement moments of TN in general dimension, and study the variance of the estimate in Section IV A from which arises the complexity of an estimation up to a chosen error. We benchmark the algorithm with the ground states of the exactly solvable toric code model, using iPEPS, and the XX chain, using MPS, in Section IV B. Finally, we discuss the results and future steps in the conclusions, Section V. In the appendix, we present variance estimations of the Rényi moments in the general case (Appendix A), as well as specifically in the toric code model, which is used as a benchmark (Appendix B).

II. PRELIMINARIES

A. Entanglement measures

For a quantum system in a pure state $\rho = |\psi\rangle \langle \psi|$, we define for a subsystem the reduced quantum state or reduced density matrix (RDM) as

$$\rho_A := \text{Tr}_\overline{A}(\rho).$$

(1)

The entanglement of the subsystem $A$ with its environment (constituting its complement) $\overline{A}$ is encoded in the RDM. We introduce the moments of the RDM. For a positive integer $n$, the $n$-th RDM moment is defined to be

$$p_n(\rho_A) := \text{Tr}(\rho_A^n),$$

(2)

also referred to as the Rényi moments. On top of being entanglement monotones and hence measures of entanglement in their own right, these moments are used for defining various entanglement measures with useful mathematical properties. The RDM moments can under mild mathematical conditions – be analytically continued to the entanglement measure featuring the strongest interpretation for pure bi-partite quantum states, the von Neumann entanglement entropy, defined as

$$S(\rho_A) := -\text{Tr}(\rho_A \log \rho_A)$$

(3)

for RDMs $\rho_A$, as the 1-Rényi entropy. The von Neumann entropy is obtained in the limit $S(\rho_A) = \lim_{n \rightarrow 1} (1 - n)^{-1} \log(p_n(\rho_A))$. The Rényi moments are especially popular as entanglement indicators since they do not require a full reconstruction of the RDM spectrum. Therefore, they are often easier to either calculate theoretically or measure experimentally than other entanglement measures.

The measures above are appropriate when quantifying the entanglement between a subsystem $A$ and its environment $\overline{A}$ when $A \cup \overline{A}$ is in a pure state. When characterizing the entanglement between two subsystems $A_1$ and $A_2$ whose union $A$ is not necessarily pure, the quantities above will no longer be suitable to quantify entanglement, as they cannot distinguish between the quantum entanglement between $A_1$ and $A_2$ from their entanglement with the environment. One of the best known measures for the entanglement between two subsystems labeled as $A_1$ and $A_2$ is the entanglement negativity based on the partial positive transpose (PPT) criterion, defined as

$$\mathcal{N}(\rho_A) := \frac{\|\rho_A^{T_2}\|_1 - 1}{2},$$

(4)

where $\| \cdot \|_1$ denotes the trace norm, and $\rho_A^{T_2}$ the partial transposition of the degrees of freedom corresponding to $A_2$ in $\rho_A$.

$$\langle i | A_1 \langle j| A_2 \rho_A | k\rangle_{A_1} | l\rangle_{A_2} = \langle i | A_1 \langle j| A_2 \rho_A^{T_2} | k\rangle_{A_1} | l\rangle_{A_2},$$

for all vectors $| i \rangle_{A_1}, | j \rangle_{A_2}, | k \rangle, | l \rangle$ in an orthonormal basis of the Hilbert spaces of $A_1, A_2$, respectively. The usefulness of the negativity as an entanglement measure for two subsystems in a mixed state leads us to define the moments of the partially-traced RDM, further referred to as PT moments. The $n$-th PT moment is defined to be

$$R_n(\rho_A) := \text{Tr}\left((\rho_A^{T_2})^n\right)$$

(5)

for a positive integer $n$. The negativity can be obtained by an analytic continuation of the PT even integer moments by $\|\rho_A^{T_2}\|_1 = \lim_{n \rightarrow 1/2} R_{2n}(\rho_A)$.

The PT moments are not entanglement monotones, but they can be used to detect entanglement between $A_1$ and $A_2$ as well as for estimating the negativity. The popularity of the PT moments as entanglement indicators stems from the fact that they too do not require a full reconstruction of the partially transposed RDM, and are therefore easier to calculate and measure experimentally.

For systems with a conserved charge $Q$, the quantum state of the full system typically commutes with the charge operator,

$$[\rho_A, Q_A] = 0.$$
called charge-resolved or symmetry-resolved moment,

\[ p_n(\rho_A, q) := \text{Tr} (\rho_A(q)^n), \]

again for positive integers \( n \). This definition could be extended to the negativity as well\(^{[51]}\). The study of symmetry-entangled states has drawn much interest lately, both analytically and numerically\(^{[26,53,65]}\), as well as in the development of experimental measurement protocols\(^{[52,54,55,56]}\). It reveals the relation between entanglement and charge and can point to effects such as topological phase transitions\(^{[54,55,51]}\) or to instances of dissipation\(^{[65]}\) in open systems dynamics\(^{[65]}\).

The estimation of symmetry-entangled entanglement can be done based on the analysis in Ref.\(^{[49]}\) We introduce the so-called flux-resolved RDM moments as

\[ p_n(\rho_A, \varphi) := \text{Tr} \left( e^{i\varphi Q_A} \rho_A^n \right), \]

where \( \varphi \in [0, 2\pi) \) can be thought of as an Aharonov-Bohm flux inserted in the replica trick. The symmetry-resolved moments can be extracted from the flux-resolved moments by a Fourier transform according to

\[ p_n(\rho_A, q) := \frac{1}{N_A} \sum_\varphi p_n(\rho_A, \varphi) e^{-iq\varphi}, \]

where \( \varphi = 2\pi k/N_A \) and \( k = 0, \ldots, N_A - 1 \).

**B. Tensor networks**

We will now briefly review the TN tools that are being used to compute the quantities presented above in the remainder of this work by means of sampling techniques.

1. **Matrix product states**

We here consider a one-dimensional spin system, so of a finite local dimension, featuring \( N \) lattice sites. The state vector of the system can be written as

\[ |\psi\rangle = \sum_{\sigma_1, \ldots, \sigma_N} \Psi_{\sigma_1, \ldots, \sigma_N} |\sigma_1, \ldots, \sigma_N\rangle, \]

where \( \Psi \) is the rank \( N \) tensor of the coefficients of \( |\psi\rangle \). \( \Psi \) has \( d^N \) complex amplitudes, where \( d \) is the Hilbert space dimension of a single spin. In an MPS representation we decompose \( \Psi \) into \( N \) different tensors, each corresponding to a single site, as illustrated in Fig.\(^{[1]}\). Each such tensor will have a single index corresponding to the indices of the original tensor, often called the ‘physical leg’ or the ‘physical index’, and two additional indices connecting with the tensors corresponding to the site’s neighbours, often called ‘entanglement legs’ or ‘bond indices’ (with only one bond index for the sites at the edges). Contracting all the bond indices will result in the original tensor \( \Psi \). In many cases one can limit the dimension of the bond index to be some chosen constant \( D \), also known as the bond dimension, and discard the least significant variables. In this way, the number of real parameters will be scaling as \( O(N D^2 d) \), at the cost of getting an approximate representation for the state. States that are expected to be well approximated by such limited tensors obey an entanglement area-law\(^{[53,55]}\) (in fact, this is provably the case for area laws of suitable Rényi entropies\(^{[29]}\)). This MPS decomposition is a widely used method for the simulation of ground states\(^{[70,71]}\) of thermal states\(^{[72,73]}\) and states undergoing a time evolution\(^{[74,75]}\) generated by local Hamiltonians of one-dimensional systems.

For a system partitioned into two contiguous subsystems, the extraction of the spectrum of the RDM, also called the entanglement spectrum, is very natural\(^{[76]}\) and can often be useful in classifying phases of matter in one spatial dimension\(^{[80,81]}\). We note that a decomposition of the system into two tensors, one corresponding to subsystem \( A \) and one to \( \overline{A} \), is built into the decomposition of the systems into site tensors, and that this decomposition can be transformed into the Schmidt decomposition of the state vector

\[ |\psi\rangle = \sum_i \psi_i |i\rangle_A |i\rangle_{\overline{A}}, \]

where \( \{|i\rangle_A\}, \{|i\rangle_{\overline{A}}\} \) are orthonormal bases of \( A, \overline{A} \) respectively. The values \( \psi_i \) are called the Schmidt values, and can be extracted by a singular value decomposition\(^{[52]}\). The RDM is thus

\[ \rho_A = \sum_i |\psi_i|^2 |i\rangle_A \langle i|_A. \]

The RDM eigenvalues are thus the squared absolute values of the Schmidt values, and by obtaining them, we can extract the RDM moment in all ranks \( n \), as well as the von Neumann entropy. Specific techniques have been developed for the extraction of entanglement measures in some additional cases, such as the entanglement of a contiguous subsystem of an infinite system\(^{[83]}\) or the negativity of two contiguous subsystems\(^{[83]}\).

2. **Projected entangled pair states**

For two or higher dimensional lattice systems, the MPS formalism is extended to an ansatz called projected entangled pair states (PEPS)\(^{[90,99]}\). The tensor capturing the state vector of the entire lattice is then decomposed into \( N \) site tensors, each with a single physical index and a bond index for each neighbour of the site in the system. An example for a square lattice is depicted in Fig.\(^{[1]}\), and the generalization to other lattice configurations is straightforward.

The infinite version of PEPS, known as iPEPS\(^{[62]}\) can be used to represent states in the thermodynamic limit in 2D. They are defined by a finite set of tensors repeated all over the lattice with some periodicity. iPEPS have found numerous applications in studying ground states of thermal states\(^{[90,96]}\) and non-equilibrium problems\(^{[97,102]}\) in two spatial dimensions, and have become state of the art numerical technique for studying strongly correlated two-dimensional problems. The technique does not suffer from the infamous sign...
order to get the RDM of a rectangular subsystem, we can construct a boundary tensor product formalism and complex conjugate and placing them back to back as a tensor. Contracting all the bond indices will result in the original tensor. (a) and (b) represent states in a one and two spatial dimensions, respectively. (c) Representing the density matrix of a pure state by the site tensors of the sites in blue is obtained by tracing out the degrees of freedom of the sites in grey, which in turn is obtained by contracting the physical legs of all tensors. (d) The RDM of the sites in black is $D^2$, while that of the legs in green, connecting the boundary tensors to each other, is $\chi$, the bond dimension of the environment.

The pure quantum state $\rho = |\psi\rangle \langle \psi|$ of the quantum system can be obtained by taking a PEPS state vector and its Hermitian conjugate and placing them back to back as a tensor product, as depicted for PEPS in Fig. 1. We now examine a rectangular subsystem $A$ with $N_A = w_1 \times w_2$ sites, where $w_1 \leq w_2$ (as will be the notation throughout this work). In order to get the RDM of $A$ as defined in Eq. 1, the degrees of freedom of $\mathcal{A}$ need to be traced out. This can be obtained by contracting the physical legs of all tensors corresponding to sites in $\mathcal{A}$ with the physical legs of the same tensor in the complex conjugate. We get a RDM composed of site tensors for the tensors in $A$, and boundary tensors resulting from the tensors in $\mathcal{A}$, as depicted in Fig. 1e. Such boundary tensors can be approximately computed for an infinite system. We remark here that exactly contracting PEPS tensors is a computationally hard problem (in worst case complexity and for meaningful probability measures even in average case), and therefore, we have to rely on approximation algorithms such as the corner transfer matrix renormalization group algorithm, tensor techniques, higher order tensor renormalization group method, or others. It is also known that those PEPS that are ground states of uniformly gapped parent Hamiltonians – which are interesting in the condensed matter context – can actually be contracted in quasi-polynomial time. In this work, we make use of the boundary MPS technique: We create a one-dimensional TN representing the boundary of the (supposedly infinite) system, and multiply it by the ‘traced out’ tensors indicated in Fig. 1. The boundary bond dimension is limited to a constant dimension $\chi$. This process is then repeated until the one-dimensional boundary tensors are converged, resulting in a one-dimensional boundary as depicted in Fig. 1.

C. Entanglement measures computed from reduced states

The entanglement measures presented in Section II A can be extracted for two contiguous systems in an MPS as presented in Section II B 1, as well as in additional specific cases in one and two spatial dimensions. However, for a general dimension and partition, there is no efficient way known to quantify the entanglement. A straightforward method can be contracting the tensors such that the RDM is obtained explicitly to then obtain its spectral decomposition. However, the explicit RDM is of dimension $d^{N_A} \times d^{N_A}$, which comes along with substantial computational effort and which imposes a strong restriction on the accessible system sizes.

That being said, the $n$-th RDM or PT moments defined in Eqs. (2, 5) can be calculated in polynomial time in $N_A$ using $n$ copies of the system tensors as depicted in Fig. 2 for a two-dimensional PEPS. The space complexity required for performing this multiplication for an MPS scales as $O(d^{2n}D^{2n} + D^{4n})$. The space complexity for a two-dimensional PEPS is given by $O(\chi^{2n}D^{2(w_1+1)n} + D^{8n}d)$, where $\chi$ is the bond dimension of the environment as depicted in Fig. 1, and $w_1$ is the short edge of the rectangular system, as defined in Section II B 2 above. For $n > 1$, the exponential dependence of the cost on $n$ quickly makes it prohibitively large, despite the fact that for a narrow system (constant $w_1$),
the time complexity is linear in $N_A$ and the space complexity does not depend on $N_A$ (except for the possible dependence of $D$ on $N_A$, as can sometimes happen in finite systems).

D. Benchmark models

Before turning to presenting the actual sampling method for computing entanglement measures in quantum systems captured by tensor networks, we here first present the models we used for benchmarking our method: The two-dimensional gapped toric code model on a square lattice and the one-dimensional gapless XX model.

1. The toric code model

The first benchmark model we elaborate on is the analytically solvable toric code model on a square lattice. The toric code, introduced by Kitaev, transferring insights from topological quantum field theory to the realm of quantum spin systems, is a model of spins on a square lattice with local dimension $d = 2$. The spins live on the edges of the lattice rather than its nodes. The Hamiltonian of the model is given by

$$H = -J_s \sum_s \otimes_{i \in s} \sigma_i^x - J_p \sum_p \otimes_{i \in p} \sigma_i^z. \quad (13)$$

$s$ in the equation above represents the set of edges around a single node in the lattice (a star) and $p$ represents the set of edges forming a plaquette in the lattice, as shown in Fig. [3].

The ground state of toric code model displays several important properties, among which are topological order, which leads to robustness to local errors, making it an important candidate for fault-tolerant error correction code. For $J_s, J_p > 0$, the ground state vector of the model with open boundary conditions is known and can be written as

$$|\psi_0\rangle = \prod_s \frac{(1 + \otimes_{i \in s} \sigma_i^x)}{2} |0\rangle \otimes N, \quad (14)$$

where $N$ is the number of all sites in the system. In the limit $N \to \infty$, the iPEPS representation of the infinite toric code ground state is given in Refs. [5] and [113]. A set of two site tensors, $T_A$ and $T_B$, are repeated infinitely such that all of the nearest-neighbours of a site represented by $T_A$ are of the form $T_B$ and vice versa. The bond dimension of all bond indices of $T_A$ and $T_B$ is $D = 2$.

For a subsystem of the infinite system in the state defined in Eq. (14), the density-matrix-based measures can be analytically calculated. This relies on the symmetry of the ground state under the application of $\otimes_{i \in s} \sigma_i^z$ for all stars $s$ and of $\otimes_{i \in p} \sigma_i^z$ for all plaquettes $p$. Due to this symmetry, the RDM is block diagonal, where the size of each block equals the order of the group generated by each operator, which is 2 for the operators above. Considering the fact that all non-zero blocks are identical, as can be seen from Eq. (14), the eigenvalues of the RDM can be extracted analytically. Note that the symmetry mentioned above is not utilized in the numerical method, so as to make our performance results applicable to general analytically-unsolvable models, which do not possess such local symmetries. Here, we estimate the 2nd, 3rd and 4th RDM moments, as well as the 3rd PT moment, for a checkerboard-like partition of a square subsystem ($w_1 = w_2 = w$), as shown for $w = 6$ in Fig. [3]. We study the cases of $w = 4, 6, 8$. Note that while the toric code displays an area law type entanglement structure, here the entire system is in the area and therefore a volume law is reached. Such extensive partitions were shown to be interesting for the study of topological phases in Refs. [115-116] and following works (a more traditional geometry is studied in Appendix [15]). For such systems, the $n$-th RDM moment is shown to be

$$p_n(\rho_A) = R_n(\rho_A) = 2^{-w(w^2/2-1)(n-1)}. \quad (15)$$

The log of the moment deviates from an area law by an additive constant term, reflecting the topological order of the model [117-118]. Note that for the toric code ground state $p_3 = R_3$ due to the structure of the RDM discussed above. However, we compute an estimate for $R_3$ based on the generally-applicable estimator defined in Eq. (25) for completeness.

2. The XX model

While suited for high dimensions, we note that our method is blind to the dimensionality of the system, and will apply to one-dimensional systems in precisely the same way as it would for higher dimensions. Therefore, we can use one-dimensional models as benchmark models for testing the system. We test our model on the ground state of the one-dimensional XX model captured by the local Hamiltonian

$$H = J \sum_i \sigma_i^+ \sigma_{i+1}^- + \text{h.c.}, \quad (16)$$

where $i$ stands for a site in the system. The Hamiltonian can be seen as a Hamiltonian of non-interacting fermions by virtue of the Jordan-Wigner transformation [119] and is thus analytically solvable [120]. We compute the ground state of a system of length $2l$ and extract the 2nd, 3rd and 4th RDM moments of a contiguous half of the system. In contrast to the toric code, the
XX model is gapless in the absence of a large magnetic field and can be well approximated as a conformal system. For such systems, the RDM moments of a subsystem when the total system is in the ground state is to a good approximation predicted to be\cite{121,122}

\[
p_n(\rho_A) \sim \frac{1 + n^{-1}}{6} N_A^{-c(n-1)/6},
\]

(17)

where \(c = 1\) is the conformal charge. As opposed to the toric code ground state RDM, which is composed of 2 \(\times\) 2 blocks, the XX ground state is not as structured, and the performance of the method is harder to predict. As such, the XX model ground state is a good complement to the toric code ground state in the study of the method’s performance.

### III. METHOD

We now turn to describing the method that is at the heart of this work. The core idea is that with suitable stochastic sampling techniques, one can more resource-efficiently estimate entanglement properties of systems captured by tensor networks. Inspired by the growing body of methods based on random unitaries\cite{10,27} and described in Section I, we now turn to describe our random-variables-based method for estimating the entanglement contained in a TN state. As mentioned in the introduction, our method differs from the protocols that are routinely implemented in experiments by two key aspects: First, we do not base the protocol on sampling local measurement results, which are cumbersome to extract from TNs, but on sampling of expectation values, which can be naturally calculated in TNs. Note that while actual sampling from MPS can be performed exactly\cite{123}, sampling from PEPS is shown to be computationally hard, in the worst as well as average case\cite{124,125}. The second difference between our method and the experimental protocols is that we do not have to limit ourselves to physically allowed processes, and specifically, our random operations are neither unitary nor quantum channels, which allows for a significant simplification of the protocol.

#### A. Sampling random vectors

In what follows, random vectors \(|v\rangle \in \mathbb{C}^d\) drawn from appropriate probability measures will feature with the property such that

\[
E(|v\rangle \langle v|) = I,
\]

(18)

where \(E\) refers to the average over the chosen probability measure. This is up to the normalization that is only different by a factor of \(d^{1/2}\) than what is commonly called a frame or a spherical complex 1-design\cite{126,127}. This convention is helpful in what follows. The set of vectors can be a discrete or a continuous set. We use the random variable to get simple estimators for the entanglement quantifiers based on Rényi moments of RDM of subsystems \(A\), each site \(\alpha \in A\) of which corresponds to a system of local dimension \(d\). In this setting, we consider random vectors

\[
|V\rangle = \bigotimes_{\alpha \in A} |v^{[\alpha]}\rangle \in \mathbb{C}^{d^{|A|}},
\]

(19)

where \(|v^{[\alpha]}\rangle \in \mathbb{C}^d\) are vectors drawn in an i.i.d. fashion as in Eq. (18), one for each site \(\alpha\). Naturally

\[
E(|V\rangle \langle V|) = I
\]

(20)

still holds true in this multi-partite setting.
B. Estimators of entanglement measures

By applying these random vectors to the reduced density matrix, we obtain an estimator of the second entanglement moment, also referred to as the purity, from expressions of the form

\[ \hat{p}_2(\rho_A) = \langle V | \rho_A | V' \rangle^2. \]  

(21)

Indeed, averaging over the (independent) random vectors \( |V \rangle, |V' \rangle \) drawn from a product probability measure as defined in Eq. (19), we consistently obtain the second moment, also referred to as the purity, from expressions of the matrix, we obtain an estimator of the second entanglement moment calculation is repeated \( M \) times the number of repetitions \( s \). The estimator of the negativity moment is obtained as the expectation

\[ E(\hat{p}_2(\rho_A)) = E(\langle V | \rho_A | V' \rangle \langle V' | \rho_A | V \rangle) \]

(22)

where \( |V(1)\rangle, \ldots, |V(n)\rangle \) are drawn in an i.i.d. fashion from the same probability measure. For the PT moments, perform the partial transposition with respect to subsystem \( A_2 \). Specifically, we define, for \( i, j = 1, \ldots, n \), pairs of product vectors as

\[ |V(i,j)\rangle := \otimes_{\alpha \in A_1} |\hat{v}^{(\alpha)}(i)\rangle \otimes_{\beta \in A_2} |\hat{v}^{(\beta)}(j)\rangle, \]

(24)

so that the correct ordering of random product vectors can be reflected. The estimator of the negativity moment is obtained by

\[ \hat{R}_n(\rho_A) = \langle V(1,n) | \rho_A | V(2,n-1) \rangle \langle V(2,n-1) | \rho_A | V(3,n-2) \rangle \times \ldots \langle V(n-1,2) | \rho_A | V(n,1) \rangle \langle V(n,1) | \rho_A | V(1,n) \rangle, \]

(25)

so that

\[ E(\hat{R}_n(\rho_A)) = \text{Tr} \left( (\rho A)^n \right) = R_n(\rho_A). \]

(26)

C. Candidate probability measures

The required property of the random vectors, captured in Eq. (18), can be naturally obtained in a wealth of ways: After all, all that is required is to have up to normalization a spherical 1-design property. Still, since we do not require the vectors to necessarily constitute a spherical complex 2-design, the second moments will depend on the specific choice of the probability measure. For example, this can be done by choosing the vectors randomly out of some orthogonal basis, or several orthogonal bases. For prime dimension \( d \), the clock and shift operators, Weyl operators, or simply \( d \)-dimensional Pauli matrices, are defined to be the operators

\[ Z_d := \sum_{i=0}^{d-1} \omega^i |i\rangle \rangle, X_d := \sum_{i=0}^{d-1} |i\rangle \rangle \langle \mod(i + 1, d)|, \]

(27)

where \( \omega := e^{2\pi i/d} \). The \( j \)-th normalized eigenvector of the \( i \)-th \( d \)-dimensional Pauli matrix is denoted by \( |p(i,j)\rangle \), and we note that for prime \( d \), the number of non-commuting Pauli matrices is \( d+1 \). We then compare two possible distributions that are particularly practical in the context given: First is the ‘full-basis’ distribution,

\[ |v\rangle \in \left\{ \sqrt{d} |p(i,j)\rangle \right\}_{i=1,\ldots,d+1,j=1,\ldots,d}, \]

(28)

The vectors are normalized such that Eq. (18) is obeyed. The second distribution is referred to as ‘partial-basis’, in which we sample from the eigenbasis of only \( d \) Pauli matrices

\[ |v\rangle \in \left\{ \sqrt{d} |p(i,j)\rangle \right\}_{i=1,\ldots,d,j=1,\ldots,d}. \]

(29)

For non-prime \( d \), the Pauli matrices can be defined to be tensor products of the matrices in Eq. (27) in the dimensions of the factors of \( d \). In this case, the vector distributions defined as in Eqs. (28), (29), but with the eigenbases of the independent products of the clock and shift operators. We compare the two probability measures (and discuss why it is sufficient to only consider these distributions) in Secs. [III D] and [IV B] and in Appendix [A] below. For states represented efficiently by TN, the partial-basis distribution (with an optimized basis choice, as detailed in Section [II C]) turns out to be more efficient, and therefore most of the presented results were obtained using this method.

D. Required number of repetitions

The method suggested makes use of random quantum states for the estimation of entanglement measures. When drawing random vectors from the probability measures indicated above, one finds for the probability of deviating from the expectation to be bounded by

\[ \Pr \left( |\hat{p}_n - E(\hat{p}_n)| \geq k \sigma(\hat{p}_n) \right) \leq \frac{1}{k^2}, \]

(30)
for real \( k \) and \( \sigma(\hat{p}_n)^2 := \text{Var}(\hat{p}_n) \). This is true by virtue of Chebychev’s inequality, a large deviation bound. We here and in the following suppress the dependence on \( \rho_A \). For \( M \) repetitions, the variance \( \text{Var}_M \) of the estimator of the mean \( E(\hat{p}_n) \) is given by

\[
\text{Var}_M := \frac{\text{Var}(\hat{p}_n)}{M}.
\]

Since then

\[
\varepsilon^2 := \frac{\text{Var}_M}{p_n^2} = \frac{\text{Var}(\hat{p}_n)}{M p_n^2},
\]

for a given \( \varepsilon > 0 \), the number of required repetitions scales as \( M = \frac{\text{Var}(\hat{p}_n)}{p_n^2 \varepsilon^2} \).

The characteristics of the method lead us to expect an exponential dependence of the required number of repetitions \( M \) on system size (whilst as can be seen below, a weak one). We thus define the scaling factor \( \xi_n \) by

\[
\frac{\text{Var}(\hat{p}_n(\rho_A))}{p_n(\rho_A)^2} = : \xi_n^{N_A}.
\]

We use the notation \( \xi_n \) for the scaling factor of \( \hat{R}_n \) as well, since the scaling factors for both properties are expected to behave similarly.

In Appendix A, we show that the variance of the estimators defined in Eqs. (23) and (25) is given by

\[
\text{Var}(\hat{p}_n) = \text{Tr}\left((\rho_A^{n/2} \mathcal{E})^n\right) - p_n(\rho_A)^2
\]

and

\[
\text{Var}(\hat{R}_n) = \text{Tr}\left(\left((\rho_A^{T_2})^{n/2} \mathcal{E}\right)^n\right) - R_n(\rho_A)^2,
\]

where

\[
\mathcal{E} := \mathbb{E}\left(\left|V\right>\left< V\right|\right)\left|V\right>\left< V\right|.
\]

\(\xi_3 \) is shown in Appendix A, the squared coefficient of variation in a

product state \( \rho = \otimes_{\alpha \in A} |\psi_{\alpha}\rangle\langle \psi_{\alpha}| \) in a maximally mixed
case (i.e., subsystem $A$ being maximally entangled with the rest of the system), $\rho_A = \mathbb{I}/d^{N_A}$, are

$$\text{Var}(\hat{p}_n) = \left(\frac{2d}{d+1}\right)^{nN_A} - 1$$

(42)

and

$$\text{Var}(\hat{p}_n) = \left(\frac{2d}{d+1}\right)^{(n-1)N_A} - 1,$$

(43)

respectively. The same results apply for the PT moments. For example, for $d = 2$,

$$\xi_n = \left(\frac{4}{3}\right)^n$$

(44)

for a product state and

$$\xi_n = \left(\frac{4}{3}\right)^{n-1}$$

(45)

for a maximally mixed state. Note that the scaling factors obtained below for the benchmark models, as displayed in Table I, are in-between these two extreme cases.

As for the partial-basis distribution, as explained in detail in Appendix A, the highest variance, hence the largest additive sampling error, for both the RDM moments and PT moments arises for

$$\rho = \otimes_{\alpha \in A} |\psi_E\rangle \langle \psi_E|_\alpha,$$

(46)

where $|\psi_E\rangle$ stands for a state vector of the form

$$|\psi_E\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^d [e^{i\phi_j} |j\rangle],$$

(47)

with an equal magnitude of the amplitude for each state in the computational basis. In this case, the contribution of each site to the first term of the variance is

$$\xi_n = \left(\left(d+4\left(d^2/2\right)^n\right)^{N_A},$$

(48)

and the overall variance is given by

$$\left(\left(d+4\left(d^2/2\right)^{N_A}\right)^n - 1.$$  

(49)

For example, for $d = 2$, the scaling factor in such a case is $\xi_n = (3/2)^n$. The best case is $\rho_A = |0\rangle \langle 0|^{N_A}$ (or any other basis state in the computational basis), in which the variance is 0. Both cases are completely disentangled, and the moments equal 1. Therefore, the former of these is not maximal in terms of the squared coefficient of variation, which is determined by the ratio between the standard deviation and the expected value. However, the analysis of the variance itself already serves to demonstrate that the choice of basis for the vector $|\psi\rangle$ can have a significant impact on the variance and due to that on the performance of the algorithm, as discussed in Section IV C below. The coefficient of variation in

Figure 7. Relative error in the estimation of (a) the 2nd-4th RDM moments and the 3rd PT moment of the toric code ground state for a checkerboard partition ($N_A = w^2/2$), and (b) The 2nd-4th RDM moment and the 3rd symmetry-resolved RDM moment for $q = 0$ of the XX model ground state, based on Eq. (23) for the RDM moments, on Eq. (25) for the PT moment and (51) for the symmetry-resolved moment. Here, we use the partial-basis distribution, Eq. (29). The normalization of $M$ by the scaling factor $\xi_n$ in the horizontal axis is extracted from Fig. 8 and presented in Table I. The presented plots are averaged over 20 different permutations of the estimation results.
the maximally mixed case in this method is calculated in Appendix A and equals

$$\frac{\text{Var}(\hat{p}_n)}{p_n^2} = \left( d + \frac{d(d+1)}{2} 2^n \right) N_A. \quad (50)$$

The two cases represented in Eqs. (42), (43), and (48) above, in which the variance can be calculated exactly, are not promised to be the best or worst case for the two distribution methods. For a two-qubit system we have performed a gradient descent search for the extreme cases in both distributions, where a basis optimization (as described in Section IV C) has been included in the partial-basis method. Fig. 5 presents the results, which strongly support the hypothesis that the maximally- and minimally-mixed cases are indeed the extreme cases for the method’s performance.

The best distribution choice is therefore case-dependent: For moderate or large $n$ and highly mixed cases, the full-basis distribution is advantageous. For small $n$ or weakly entangled cases, the partial-basis method with a smart basis choice (as discuss in Section IV C) is more beneficial. Based on Eq. (50), the partial-basis seems to have a poor performance on highly entangled (mixed) states. However, states represented efficiently by TN feature an entanglement that exhibits an area law, which restrains the entanglement of relevant states to begin with. Note that even in this worst case our method is still favourable compared to the time required for an exact diagonalization of the RDM, which scales as $O(d^{3N_A})$, for intermediate $n$. The variances calculated above compare favorably with the variances in the experimental sampling-based protocols [10,24], as calculated in Ref. [24]. For $d = 2$, the relative variances for $n = 2, 3$ are shown to be $\text{Var}(\hat{p}_2)/p_2^2 \geq (8/p_2^2) \max \{ 2^{N_A} p_2, 2^{1.5N_A} \}$ and $\text{Var}(\hat{p}_3)/p_3^2 \geq (39/p_3^2) \max \{ 2^{N_A} p_2, 2^{1.5N_A} p_2, 2^{2N_A} \}$.

It would be interesting to use the analysis above as a basis for a study regarding the number of local samples required for the estimation of Rényi moments in general. Naively, the moments are defined as a function of the full RDM, which has $d^{2N_A}$ elements, hence should require a comparable number of samples. However, an extraction of a single degree of freedom is expected to require a smaller number of samples, as is the case in small $n$s in our method. Such analysis may point to the amount of information contained in Rényi moments of different ranks.

E. Symmetry-resolved RDM moments

Using the locality of the phase operator from Eq. (8), $e^{i\varphi Q_A} = \otimes_{i \in A} e^{i\varphi Q_i}$, we can extract the flux-resolved moments from the TN, and substitute them into Eq. (9) to get the symmetry-resolved moments. The estimator of the flux-resolved moment is similar to the estimator of the full moments in Eq. (27) and is obtained from

$$\hat{p}_n(\rho_A, \varphi) = \langle V^{(1)} | e^{i\varphi Q_A} \rho_A | V^{(2)} \rangle \langle V^{(2)} | \rho_A | V^{(3)} \rangle \times \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle. \quad (51)$$

The estimator of the charge-resolved moment is obtained from

$$\hat{p}_n(\rho_A, q) = \sum_{\varphi} e^{-i\varphi q} \langle V^{(1)} | e^{i\varphi Q_A} \rho_A | V^{(2)} \rangle \ldots \times \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle. \quad (52)$$

A similar analysis of symmetry-resolved PT moments has been done in Ref. [52] and the extension to their estimation is natural. Below we estimate the symmetry-resolved RDM moments for the XX model and its conserved total $S_z$. For this model, the symmetry-resolved moments can be obtained exactly following Ref. [49]. The expected and extracted results for $q \mapsto p_n(\rho_A, q)$ for $n = 2, 3$ are displayed in Fig. 6.

IV. TESTING THE METHOD AGAINST THE BENCHMARK MODELS

A. Specific tests

We have tested the model against the exactly solvable two-dimensional toric code model and one-dimensional XX model as detailed in Section III D, employing the TensorNetwork library [25]. The precision of the estimation for both models as a function of $M$, the number of samples of the expressions in Eqs. (21), (23), (25), and (51) is shown in Fig. 7. The results are obtained using the partial-basis distribution, Eq. (29), and are optimized based on the analysis in Section IV C below. In order to reduce the numerical noise in the dependence of the precision in $M$, we averaged this dependence over several permutations of the $M$ repetitions. While the required number of repetitions $M$ (for a given allowed error $\varepsilon > 0$) is exponential in system size (as discussed above in Section III D), it has a relatively small base $\xi_n$. When considering the significant decrease in required memory space, our method can become advantageous for systems around $N_A = 20$, for which the method described in Section II B can become too heavy in memory demands for a standard computer workstation.

B. Variance estimation

Here we follow the analysis of the scaling factor $\xi_n$ in Section III D and estimate the scaling factors in the benchmark models, in order to get some idea regarding the variance in the general case. The scaling factors $\xi_n$ obtained for the toric code and XX ground states, for $n = 2, 3, 4$ are estimated numerically. In Appendix B we demonstrate an exact calculation of the expressions in Eqs. (34) and (35) for a narrow strip-like system in the toric code model, and show that the resulting expressions agree with the numerically estimated scaling factors.

We emphasize that the models tested are not specifically suitable for the method, and are not expected to have a low scaling factors based on Eqs. (34) and (35). The estimated scaling factor $\xi_n$ can therefore be considered typical.

Fig. 8 presents the estimated variances of the benchmark models in the full-basis distribution and partial-basis distri-
Figure 8. Numerical estimation of the variance of (a) the 2nd-4th RDM moments and the 3rd PT moment of the toric code ground state for a checkerboard partition; (b) the 2nd-4th RDM moment and the 3rd symmetry-resolved RDM moment for \( q = 0 \) of the XX model ground state. Here, we mainly use the partial-basis distribution, Eq. (29), and add results for \( p_n \) in the full-basis distribution, Eq. (28), which displays a very similar performance in the more strongly entangled toric code model, and a worst behaviour in the log-like entangled XX model. In the vertical axis title, \( p \) stands for the estimated moment in all cases. The dependence of the variance on system size \( \xi \), defined in Eq. (33), is extracted from the linear fit in the figure and shown in Table I. Note that for the toric code ground state shown in (a), the behaviours of the 3rd RDM moment and the 3rd PT moment are expected to be identical, and are slightly different only due to the random nature of our protocol.

### C. Dependence on basis choice

In the partial-basis distribution, as shown above, the largest and smallest additive variance values both correspond to a completely disentangled case, and the difference between the two stems from the single-particle basis choice alone. This can be understood by the decomposition

\[
\mathcal{E}^{[\alpha]} = \mathbb{1} \otimes \mathbb{1} + \frac{1}{2} (\sigma^x \otimes \sigma^x + \sigma^y \otimes \sigma^y),
\]

as can be seen from Eqs. (40) and (41), that demonstrates the orientation dependence of \( \mathcal{E} \) in this case (in contrast with

\[
\mathcal{E}^{[\alpha]} = \mathbb{1} \otimes \mathbb{1} + \frac{1}{3} (\sigma^x \otimes \sigma^x + \sigma^y \otimes \sigma^y + \sigma^z \otimes \sigma^z)
\]

in the full-basis case). For a translationally invariant system, one can expect that the optimal basis choice for each site will be the same. We can now attempt to decrease the variance by finding the basis for which \( \text{Var}(\hat{p}_1) \) is minimal, and use this basis for the estimation of higher moments. We have tested this idea against the first moment of the two benchmark models, by rotating the random vectors

\[
|\psi^{[\alpha]}\rangle \rightarrow e^{i\phi \sigma^x} e^{i\theta \sigma^y} |\psi^{[\alpha]}\rangle,
\]

and finding the best basis, as demonstrated in Fig. 9. We plot the scaling factor for \( n = 1, \xi_1 \), for the two benchmark models as a function of the basis choice. We then compare the best and worst choices of \( \phi, \theta \in [0, 2\pi) \) and extract the variance of the higher moments in the corresponding bases, as summarized in Table I. We can see that the \( p_1 \) case acts as a good indicator for the basis choice of the moments in higher \( n \)s, and allows for a smart basis choice which decreases the variance.
We have demonstrated our method with the iPEPS representation of the toric code ground state and the MPS representation of the XX ground state, and compared the results with analytical calculations. The method can be readily used for any tensor network ansatz representing a spin or bosonic system, and provide information on the entanglement of systems that were formerly unreachable by today’s computers due to a strong exponential dependence of the memory space in the moment degree $n$. Additionally, our method is advantageous for nontrivial partitions in one or higher spatial dimensions, such as the checkerboard partition\cite{wei2023} or random partition\cite{ho2023} for which the moments are hard to calculate even for one-dimensional MPS.

We compare two options for the random distribution, where each of the methods turns out to be suitable for different cases. For small $n$, the scaling of required samples number $M$ with system size $N_A$ turns out to be lower than the scaling of RDM size, and can have implications regarding the information contained in these moments. It would be interesting to try and develop a sampling-based method which incorporates nonphysical operations and compare its performance with ours. The analysis of such a protocol may shed more light on the power of Rényi moments.

The method should be suitable to fermionic PEPS\cite{maldacena2016} and can be generalized to additional Rényi measures, such as participation entropies, used for the detection of many-body localization\cite{boschi2023}. Exploring the possibility of derandomizing the algorithm, similarly to the recent results of Huang \textit{et al.}\cite{huang2023a} would also be interesting. In contrast to setting of shadow estimation, the very quantum state is already classically efficiently represented, and computing overlaps with suitable random vectors gives rise to an effective estimation of entanglement properties. Now that the option to use non-physical sampling has been opened, it can be expanded to various platforms, including experimental setting with a vectorized density matrix. It is the hope that this work contributes to the program of exploiting the power of random measurements in quantum physics, even in situations where the sampling scheme itself is not reflected by physical operations.

**ACKNOWLEDGMENTS**

We thank I. Arad, G. Cohen, and E. Zohar for very useful discussions. Our work has been supported by the Israel Science Foundation (ISF) and the Directorate for Defense Research and Development (DDR&D) grant No. 3427/21 and by the US-Israel Binational Science Foundation (BSF) Grants No. 2016224 and 2020072. N. F. is supported by the Azrieli Foundation Fellows program. J. E. has been supported by the DFG (CRC 183, project B01). This work has also received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 817482 (PASQuanS).

---

**Table I**. The extracted scaling factor defined in Eq. (33), $\xi_n$, for $n = 2, 3, 4$ for the best and worst case in the toric code ground state extracted from Fig. 9 and the XX model ground state. The scaling factors obtained in the full-basis method are displayed for comparison. The worst and best case in $\xi_1$ are shown to predict well the dependence of $\xi_n$ on the basis choice.

| $\varphi, \theta$ | Toric code best case | Toric code worst case | XX best case | XX worst case |
|------------------|----------------------|----------------------|--------------|---------------|
| $\xi_1$ | 1.05 | 1.07 | 1.24 | 1.30 |
| $\xi_2$ | 1.23 | 1.36 | 1.54 | 1.73 |
| [Full-basis] | 1.26 | 1.51 |
| $\xi_3$ | 1.80 | 1.89 | 1.65 | 1.98 |
| [Full-basis] | 1.81 | 1.77 |
| $\xi_4$ | 2.75 | 2.91 | 2.09 | 2.50 |
| [Full-basis] | 2.81 | 2.34 |

---

**V. CONCLUSIONS AND OUTLOOK**

In high-dimensional TN states, the naive computation of entanglement is highly sensitive to the size of the system (when explicitly extracting the RDM) or the bond dimension of the site tensors and boundaries (when performing the replica trick). We developed a method for estimating RDM moments in Eq. (2) and PT moments in Eq. (5) of such systems, as well as their symmetry-resolved components in Eqs. (7) and (8), without fully reconstructing the density matrix or contracting several copies of the state. The method uses randomization in order to correlate separate copies of the TN state, allowing for the estimation of properties that are defined using more than one copy of the RDM. Though we are inspired by recent experimental protocols\cite{wei2023, ho2023}, we developed a completely new algorithm which is suitable to classical simulations, takes advantage of their strengths such as ability to estimate the expectation value of non-Hermitian operators, and avoids their weakness in sampling the outcomes of random measurements.
Appendix A: Full derivation of the variance

Below we derive Eqs. (21), (23) and (25) of the main text, and use them to find density matrices which extremize the variance. First, we write the expression for the RDM moments estimator explicitly. In the above coordinate independent fashion, this derivation is straightforward. The $n$-th RDM moment is obtained as the expectation of

$$
\hat{R}_n(\rho_A) = \langle V^{(1)} | \rho_A | V^{(2)} \rangle \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle ,
$$

(A1)

where again, the product state vectors $|V^{(1)}\rangle, \ldots, |V^{(n)}\rangle$ are drawn in an i.i.d. fashion from the same probability measure. In expectation, we find

$$
\mathbb{E}(\hat{R}_n(\rho_A)) = \mathbb{E}(\langle V^{(1)} | \rho_A | V^{(2)} \rangle \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle )
$$

(A2)

$$
= \mathbb{E}(\text{Tr}(\langle V^{(1)} | \langle V^{(1)} | \rho_A | V^{(2)} \rangle \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle ))
$$

$$
= \text{Tr}(\rho_A^T)^n ,
$$

(A3)

Similarly, for the PT moments, one can make use of random vectors of the form

$$
|V^{(i,j)}\rangle = \otimes_{\alpha \in A_1} |v^{(i)}_{\alpha}\rangle \otimes_{\beta \in A_2} |v^{(j)}_{\beta}\rangle ,
$$

(A4)

for $i, j = 1, \ldots, n$, so that the estimator of the negativity moment

$$
\hat{p}_n(\rho_A) = \langle V^{(1)} | \rho_A | V^{(2)} \rangle \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle ,
$$

(A5)

$$
\mathbb{E}(\hat{p}_n(\rho_A)) = \mathbb{E}(\langle V^{(1)} | \rho_A | V^{(2)} \rangle \ldots \langle V^{(n)} | \rho_A | V^{(1)} \rangle )
$$

(A6)

$$
= \mathbb{E}(\text{Tr}(\langle V^{(1)} | \otimes (\langle V^{(1)} | \rho_A \otimes \rho_A ) (|V^{(2)}\rangle \otimes |V^{(2)}\rangle ) \ldots (\langle V^{(n)} | \otimes (\langle V^{(n)} | \rho_A \otimes \rho_A )) (|V^{(1)}\rangle \otimes |V^{(1)}\rangle )))
$$

(A7)

In expectation, this is $\mathbb{E}(\hat{p}_n(\rho_A)^2) = \text{Tr}((\rho_A^\otimes 2 \mathcal{E})^n)$ with

$$
\mathcal{E} = \mathbb{E}(\langle V \otimes |V\rangle \langle V \otimes |V\rangle )
$$

(A8)

and hence $\mathcal{E} = \otimes_{\alpha \in A} \mathcal{E}^{[\alpha]}$, where

$$
\mathcal{E}^{[\alpha]} = \mathbb{E}(\langle |v^{[\alpha]}\rangle \otimes |v^{[\alpha]}\rangle \langle v^{[\alpha]} | \otimes \langle v^{[\alpha]} |)
$$

(A9)

In this way, one finds the expectation for the variance

$$
\text{Var}(\hat{p}_n(\rho_A)) = \text{Tr}((\rho_A^\otimes 2 \mathcal{E})^n) - p_n^2.
$$

(A10)

Here it is relevant that the frames made use of do not necessarily have to constitute complex spherical 2-designs for estimating the above entanglement measures in an unbiased fashion, so that the average does not necessarily resemble that of the Haar average, and may depend on the ensemble. We now see why it is meaningful to consider the two probability measures specified above: Drawing vectors from the eigenbasis of a single Pauli matrix, i.e., randomly sampling RDM elements in this basis, would give, for example if we drew from the eigenbasis of $Z$,

$$
\mathcal{E}^{[\alpha]}_{\delta_{i,k},\delta_{j,l}} = \delta_{i,j} \delta_{k,l} + \delta_{i,k} \delta_{j,l} + \delta_{i,l} \delta_{j,k} - 2\delta_{i,k} \delta_{j,l} \delta_{i,l}.
$$

(A11)

This matrix equals the matrix obtained above, with added positive terms, hence, can only increase the variance. The same argument can be made for any number of Pauli matrices between 2 and $d - 1$. The Pauli matrices constitute a unitary $1$-design, which is the requirement for them to be a universal measure for the estimators in Eqs. (21), (23), (25), and (27). Therefore, it is unnecessary to consider additional distributions. In the maximally mixed case, $\rho_A = (I_d/d)^\otimes N_A$, the normalized first term is

$$
\text{Tr}
\begin{pmatrix}
(\rho_A^\otimes 2 \mathcal{E})^n
\end{pmatrix}
\over p_n^2
= \left(\frac{2d}{d+1}\right)^{(n-1)N_A}
$$

(A12)
for the full-basis and

\[
\frac{\text{Tr} \left( \left[ \rho_A^\otimes 2 \mathcal{E} \right]^n \right)}{p_n^A} = \left( \frac{d + \left( \frac{d}{2} \right)^2}{d^2} \right)^{N_A}.
\]  \hspace{1cm} (A13)

for the partial-basis, respectively. In a product state, \( \rho_A = \otimes_{\alpha \in A} |\psi\rangle_\alpha \langle \psi| \alpha \) in the full-basis distribution, one finds

\[
\frac{\text{Tr} \left( \left[ \rho_A^\otimes 2 \mathcal{E} \right]^n \right)}{p_n^A} = \left( \frac{2d}{d + 1} \right)^{nN_A}.
\]  \hspace{1cm} (A14)

The performance of the partial-basis distribution for a product state can be analyzed as follows: \( \mathcal{E}^{[\alpha]} \) is a block diagonal matrix, with \( d \) blocks of the form \( B_1 = (1) \) and \( \left( \frac{d}{2} \right) \) blocks of the form

\[
B_2 = \begin{pmatrix}
1 & 1 \\
1 & 1
\end{pmatrix}.
\]  \hspace{1cm} (A15)

The largest eigenvalue of \( \prod_{\alpha \in A} \mathcal{E}^{[\alpha]} \) is thus \( 2^{N_A} \), and corresponds to an eigenvector of the form \( \otimes_{\alpha \in A} |i, j\rangle \), where \( i \) and \( j \) correspond to the spin of the site \( \alpha \) in the two copies and \( i, j = 1, \ldots, d \), for \( i \neq j \). However, since \( \rho_A^\otimes 2 \) is constructed from two identical copies of \( \rho_A \), vectors of the form above cannot be the only contributors to the RDM. The RDM with the largest possible variance will have the form

\[
\rho_A = \otimes_{\alpha \in A} \frac{1}{d} \begin{pmatrix}
1 & \ldots & 1 \\
\vdots & \ddots & \vdots \\
1 & \ldots & 1
\end{pmatrix}.
\]  \hspace{1cm} (A16)

Then, the contribution of each site to the first term of the variance is

\[
\xi_n = \left( \left( d + 4 \left( \frac{d}{2} \right) \right) / d^2 \right)^n.
\]  \hspace{1cm} (A17)

The RDM with smallest possible variance will be, for example,

\[
\rho_A = |0\rangle \langle 0|^{\otimes N_A},
\]  \hspace{1cm} (A18)

or any other product state in the computational basis. In this case, the first term of the variance sums up to 1 and \( \text{Var}(\hat{p}_n) = 0 \). In both cases \( A \) is disentangled from its environment, which demonstrates that the variance of the estimated value depends on the basis choice for the vectors \( |\psi\rangle \).

The variance of the flux-resolved moment estimators of the complex valued random variable defined in Eq. (51) can similarly be computed from

\[
|\hat{p}_n(\rho_A, \varphi)|^2 = (V_1|e^{i\varphi Q_A} \rho_A|V_2)(V_2|\rho_A|V_3) \ldots (V_n|\rho_A|V_1) \\
\times (V_1|\rho_A|V_n) \ldots (V_3|\rho_A|V_2)(V_2|e^{-i\varphi Q_A} \rho_A|V_1) \\
= \text{Tr} \left( \left( \rho_A e^{i\varphi Q_A} \otimes e^{-i\varphi Q_A} \rho_A \right) (\rho_A^\otimes 2 \mathcal{E})^{n-1} \right),
\]  \hspace{1cm} (A19)

which is bounded from above by the variance for the non-resolved case, as the sum of the first term is composed of terms with the same absolute values, but with added phases. The estimator for the variance of the symmetry-resolved moments is therefore also bounded by

\[
\text{Var}(\hat{p}_n(q)) = \frac{1}{N_A} \sum_{\varphi} \text{Var}(\hat{p}_n(\varphi)) \\
\leq \frac{1}{N_A} \sum_{\varphi} \text{Var}(\hat{p}_n) = \text{Var}(\hat{p}_n).
\]  \hspace{1cm} (A20)
Appendix B: Explicit variance calculation for the toric code

Here, we demonstrate how the variance can be calculated exactly in the toric code model for a subsystem shaped as a narrow strip (Fig. 10a) for the partial-basis distribution, and compare it to the extracted variance. We start from Eqs. (34), (40) and (41), and the decomposition

$$\mathcal{E}^\alpha = I + \frac{1}{2} \sigma^x \otimes \sigma^x + \frac{1}{2} \sigma^y \otimes \sigma^y,$$

(B1)

which applies for the partial-basis method in $d = 2$. The full $\mathcal{E}$ matrix can be written as

$$\mathcal{E} = \sum_s \otimes_{\alpha} S_{\alpha},$$

where $S$ is an $N_A$-long configuration of the operators $I, \frac{1}{2} \sigma^x \otimes \sigma^x, \frac{1}{2} \sigma^y \otimes \sigma^y$. We use the local symmetry of the toric code ground state

$$\otimes_{i \in s} \sigma^x_\alpha \rho = \otimes_{i \in p} \sigma^x_\alpha \rho = \rho \otimes_{i \in s} \sigma^x_\alpha = \rho \otimes_{i \in p} \sigma^x_\alpha = \rho,$$

(B2)

in order to distinguish configurations $S$ which will contribute to Eq. (34). An allowed configuration $S$ will contain any number of star $(\otimes_{i \in s} \sigma^x_\alpha)$ or plaquette $(\otimes_{i \in p} \sigma^x_\alpha)$ operators, also called the stabilizers. Pairs of some operator which commutes with the stabilizers and act on two sides of the same copy of $\rho \otimes \rho$ are also allowed, as well as combinations of operators which can be transformed into such pairs by a multiplication of the operators by stabilizers. This is illustrated in Fig. 10b.

We calculate the variance for a narrow system of dimension $2 \times 2l$, as depicted in Fig. 10a. Such a system is composed of a chain of $l$ contiguous stars. For a moment of rank $n$, we think of $n$ copies of the subsystem, and write a $3^n$-dimensional vector of combinations of operators $[I, \sigma^x \otimes \sigma^x, \sigma^y \otimes \sigma^y]$ on a single star in all $n$ copies. We now write a transfer matrix $T^{(n)}$ which takes the operator combinations on the $l$-th star to the contributing combinations on the $(l + 1)$-th star: $T^{(n)}_{lj}$.

| $n$ | Exact variance | Estimated variance |
|-----|----------------|--------------------|
| 1   | $O(1.016^{N_A})$ | $O(1.016^{N_A})$  |
| 2   | $O(1.225^{N_A})$ | $O(1.220^{N_A})$  |
| 3   | $O(1.549^{N_A})$ | $O(1.64^{N_A})$   |

Table II. Exact dependence of the variance on system size for the toric code model as calculated in by the transfer matrix method compared to the variance estimated from sampling of the expressions in Eqs. (34) and (35).

equals the contribution of an operator combination with operators $j$ on the $(l + 1)$-th star to the variance in Eq. (34), given that the combination on the $l$-th star is $i$, where the symmetries in Eq. (52) are considered, as well as the $j$ factors in Eq. (51). For clarity, we give a specific example in Fig. 10c-d.

With these definitions, the first term in the left side of Eq. (34) for a subsystem of $2 \times 2l$ sites is $\langle c_0 | T^{(n)} | c_0 \rangle$, where $| c_0 \rangle$ is the vector of allowed contributions for the edges of the subsystem as can be deduced from the ground states of the toric code model in Eq. (14). The dependence of this term in the system size is thus $O(\lambda_{\text{max}}^d)$, where

$$\lambda_{\text{max}} := \|T^{(n)}\|,$$

(B3)
is the largest eigenvalue of the Hermitian $T^{(n)}$. One may, in fact, work with equivalent but much smaller transfer matrices, by considering only the two left sites of a star rather than the whole star. This allows decreasing the transfer matrix to dimension $3^{2n} \times 3^{2n}$. For $n \leq 3$ the matrix $T^{(n)}$ can be extracted and diagonalized exactly. We have performed an estimation of the variance for such strip-like systems, similarly to the one done in Fig. 8 in the main text. The estimated variances are displayed in Fig. 11. We have compared the results obtained exactly using the transfer matrix to the numerical variances and got a good agreement, demonstrating the accuracy of our PEPS calculations, as can be seen in the Table II.
eralization of the accessible entanglement entropy,” Phys. Rev. Lett. 121, 150501 (2018)
52. E. Cornfeld, M. Goldstein, and E. Sela, “Imbalance entanglement: Symmetry decomposition of negativity,” Phys. Rev. A 98, 032302 (2018)
53. E. Cornfeld, L. A. Landau, K. Shhtengel, and E. Sela, “Entanglement spectroscopy of non-Abelian anyons: Reading off quantum dimensions of individual anyons,” Phys. Rev. B 99, 115429 (2019)
54. N. Feldman and M. Goldstein, “Dynamics of charge-resolved entanglement after a local quench,” Phys. Rev. B 100, 235146 (2019)
55. S. Fraenkel and M. Goldstein, “Symmetry resolved entanglement: Exact results in 1D and beyond,” J. Stat. Mech. 2020, 033106 (2020)
56. S. Murciano, G. Di Giulio, and P. Calabrese, “Symmetry resolved entanglement in gapped integrable systems: A corner transfer matrix approach,” SciPost Phys. 8, 46 (2020)
57. L. Capizzi, P. Ruggiero, and P. Calabrese, “Symmetry resolved entanglement entropy of excited states in a CFT,” J. Stat. Mech. 2020, 073101 (2020)
58. G. Parez, R. Bonsignori, and P. Calabrese, “Quasiparticle dynamics of symmetry-resolved entanglement after a quench: Examples of conformal field theories and free fermions,” Phys. Rev. B 103, L041104 (2021)
59. M. T. Tan and S. Ryu, “Particle number fluctuations, Rényi entropy, and symmetry-resolved entanglement entropy in a two-dimensional Fermi gas from multidimensional bosonization,” Phys. Rev. B 101, 235169 (2020)
60. Y. Fuji and Y. Ashida, “Measurement-induced quantum criticality under continuous monitoring,” Phys. Rev. B 102, 054302 (2020)
61. D. Azses, R. Haenel, Y. Naveh, R. Raussendorf, E. Sela, and E. G. Dalla Torre, “Identification of symmetry-protected topological states on noisy quantum computers,” Phys. Rev. Lett. 125, 120502 (2020)
62. D. Azses and E. Sela, “Symmetry-resolved entanglement in symmetry-protected topological phases,” Phys. Rev. B 102, 235157 (2020)
63. B. Estienne, Y. Ikhef, and A. Morin-Duchesne, “Finite-size corrections in critical symmetry-resolved entanglement,” SciPost Phys. 10, 54 (2021)
64. S. Zhao, C. Northe, and R. Meyer, “Symmetry-resolved entanglement in Ads3/CFT2 coupled to $U(1)$ Chern-Simons theory,” JHEP 2021, 30 (2021)
65. V. Vitale, A. Elben, R. Kueng, A. Neven, J. Carrasco, B. Kraus, P. Zoller, P. Calabrese, B. Vermersch, and M. Dalmonte, “Symmetry-resolved dynamical purification in synthetic quantum matter,” arXiv:2101.07814 (2021).
66. S. Fraenkel and M. Goldstein, “Entanglement measures in a nonequilibrium steady state: Exact results in one dimension,” SciPost Phys. 11, 85 (2021)
67. A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, R. Khemani, J. Léonard, and M. Greiner, “Probing entanglement in a many-body-localized system,” Science 364, 256-260 (2019)
68. J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, “Classical simulation of infinite-size quantum lattice systems in two spatial dimensions,” Phys. Rev. Lett. 101, 250602 (2008)
69. N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, “Entropy scaling and simulability by matrix product states,” Phys. Rev. Lett. 100, 030504 (2008)
70. S. R. White, “Density matrix formulation for quantum renormalization groups,” Phys. Rev. Lett. 69, 2863–2866 (1992)
71. G. Vidal, “Efficient simulation of one-dimensional quantum many-body systems,” Phys. Rev. Lett. 93, 040502 (2004)
72. F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, “Matrix product density operators: Simulation of finite-temperature and dissipative systems,” Phys. Rev. Lett. 93, 207204 (2004)
73. M. Zwolak and G. Vidal, “Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm,” Phys. Rev. Lett. 93, 207205 (2004)
74. H. Weimer, A. Kshetrimayum, and R. Orús, “Simulation methods for open quantum many-body systems,” Rev. Mod. Phys. 93, 015008 (2021)
75. G. Vidal, “Efficient classical simulation of slightly entangled quantum computations,” Phys. Rev. Lett. 91, 147902 (2003)
76. S. R. White and A. E. Feiguin, “Real-time evolution using the density matrix renormalization group,” Phys. Rev. Lett. 93, 076401 (2004)
77. R. Orús and G. Vidal, “Infinite time-evolving block decimation algorithm beyond unitary evolution,” Phys. Rev. B 78, 155117 (2008)
78. S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, “Time-evolution methods for matrix-product states,” Ann. Phys. 411, 167998 (2019)
79. U. Schollwöck, “The density-matrix renormalization group in the age of matrix product states,” Ann. Phys. 326, 96 – 192 (2011)
80. F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, “Entanglement spectrum of a topological phase in one dimension,” Phys. Rev. B 81, 064349 (2010)
81. N. Schuch, D. Perez-Garcia, and I. Cirac, “Classifying quantum phases using matrix product states and projected entangled pair states,” Phys. Rev. B 84, 165139 (2011).
82. F. Pollmann, E. Berg, A. M. Turner, and M. Oshikawa, “Symmetry protection of topological phases in one-dimensional quantum spin systems,” Phys. Rev. B 85, 075125 (2012)
83. A. Kshetrimayum, Hong-Hao Tu, and R. Orús, “All spin-1 topological phases in a single spin-2 chain,” Phys. Rev. B 91, 205118 (2015)
84. J. I. Cirac, D. Perez-Garcia, N. Schuch, and F. Verstraete, “Matrix product states and projected entangled pair states: Concepts, symmetries, and theorems,” (2020), arXiv:2011.12127
85. P. Ruggiero, V. Alba, and P. Calabrese, “Entanglement negativity in random spin chains,” Phys. Rev. B 94, 035152 (2016)
86. F. Verstraete and J. I. Cirac, “Renormalization algorithms for quantum-many body systems in two and higher dimensions,” e-print cond-mat/0407066 (2004)
87. J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, “Classical simulation of infinite-size quantum lattice systems in two spatial dimensions,” Phys. Rev. Lett. 101, 250602 (2008)
88. P. Corboz, “Improved energy extrapolation with infinite projected entangled-pair states applied to the two-dimensional Hubbard model,” Phys. Rev. B 93, 045116 (2016)
89. H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang, “Gapless spin-liquid ground state in the $S = 1/2$ Kagome antiferromagnet,” Phys. Rev. Lett. 118, 137202 (2017)
90. T. Picot, M. Ziegler, R. Orús, and D. Poilblanc, “Spin-$S$ Kagome quantum antiferromagnets in a field with tensor networks,” Phys. Rev. B 93, 060407 (2016)
91. A. Kshetrimayum, T. Picot, R. Orús, and D. Poilblanc, “Spin-1/2 Kagome XXZ model in a field: Competition between lattice nematic and solid orders,” Phys. Rev. B 94, 235146 (2016)
92. A. Kshetrimayum, C. Balz, B. Lake, and J. Eisert, “Tensor network investigation of the double layer Kagome compound Ca$_3$Ti$_2$Cr$_2$O$_{12}$,” Ann. Phys. 421, 168292 (2020)
