Many-Body Entropies and Entanglement from Polynomially Many Local Measurements

Benoît Vermersch[ORCID],[1,2,3] Marko Ljubotina[ORCID],[4] J. Ignacio Cirac,[5,6] Peter Zoller,[2,3]
Maksym Serbyn[ORCID],[4] and Lorenzo Piroli[ORCID][7]

[1]*Université Grenoble Alpes, CNRS, LPMMC, 38000 Grenoble, France*
[2]*Institute for Theoretical Physics, University of Innsbruck, 6020 Innsbruck, Austria*
[3]*Institute for Quantum Optics and Quantum Information of the Austrian Academy of Sciences,*
6020 Innsbruck, Austria
[4]*Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria*
[5]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany*
[6]*Munich Center for Quantum Science and Technology (MCQST), Schellingstraße 4, D-80799 München, Germany*
[7]*Dipartimento di Fisica e Astronomia, Università di Bologna and INFN, Sezione di Bologna,*
via Irnerio 46, I-40126 Bologna, Italy

Estimating global properties of many-body quantum systems such as entropy or bipartite entanglement is a notoriously difficult task, typically requiring a number of measurements or classical postprocessing resources growing exponentially in the system size. In this work, we address the problem of estimating global entropies and mixed-state entanglement via partial-transposed (PT) moments and show that efficient estimation strategies exist under the assumption that all the spatial correlation lengths are finite. Focusing on one-dimensional systems, we identify a set of approximate factorization conditions (AFCs) on the system density matrix, which allow us to reconstruct entropies and PT moments from information on local subsystems. This identification yields a simple and efficient strategy for entropy and entanglement estimation. Our method could be implemented in different ways, depending on how information on local subsystems is extracted. Focusing on randomized measurements providing a practical and common measurement scheme, we prove that our protocol requires only polynomially many measurements and postprocessing operations, assuming that the state to be measured satisfies the AFCs. We prove that the AFCs hold for finite-depth quantum-circuit states and translation-invariant matrix-product density operators and provide numerical evidence that they are satisfied in more general, physically interesting cases, including thermal states of local Hamiltonians. We argue that our method could be practically useful to detect bipartite mixed-state entanglement for large numbers of qubits available in today's quantum platforms.

Subject Areas: Quantum Information

I. INTRODUCTION

In the context of today's digital quantum technologies [1–9], an outstanding challenge is to devise measurement schemes of many-qubit states which are efficient and yet simple enough to be performed in current noisy intermediate-scale quantum (NISQ) devices [10,11]. This problem has motivated new ideas and protocols to improve our ability to characterize complex quantum states. A notable example is that of the so-called randomized-measurement (RM) toolbox [12–15], which has provided us with novel opportunities to experimentally investigate entanglement,

a cornerstone in both quantum information [16,17] and quantum many-body theory [18,19].

For instance, paralleling earlier experiments based on quantum interference [20–22] (cf. also Refs. [23–26]), pure-state entanglement can be detected by exploiting the two-copy representation of subsystem purities [13,15,27–29], as is now routinely done in various experimental platforms [27,30–33]. RMs have been also applied to study both mixed-state bipartite entanglement based on the estimation of the so-called partial-transposed (PT) moments [34,35], and multipartite entanglement, as characterized by the quantum Fisher information [36–39].

Despite these developments, estimating global properties of quantum systems with a very large number of qubits N remains a difficult task. In particular, while RM approaches to estimate any local observable are more efficient than performing full-state tomography [40–55], estimating global properties requires performing exponentially many

measurements or postprocessing operations. Therefore, it is unfeasible to significantly scale up existing protocols to probe global purities and bipartite entanglement [13,15,27–29,35], raising the question of whether these quantities will be experimentally accessible at all as larger NISQ devices become available.

In this work, we address precisely the problem of estimating global entropies and PT moments in many-qubit systems and show that efficient strategies exist under the assumption that all spatial correlation lengths in the system are finite. This condition encompasses a large class of physically interesting cases, including ground and thermal states of local Hamiltonians and is thus very natural when considering NISQ devices from the point of view of quantum simulation [56–58].

In more detail, we focus on one-dimensional (1D) systems and put forward a strategy for the estimation of entropies and PT moments, requiring only polynomially many measurements and postprocessing operations. Our protocol is provably accurate for states satisfying a set of approximate factorization conditions (AFCs), which express the absence of long-range correlations and which are shown to be a general feature of short-range correlated states. The basic idea, which is conveyed in Fig. 1, is that the AFCs allow one to reconstruct global purities and PT moments from local information. The efficiency of our method is independent of how local information is extracted (for instance, this could be done via full tomography of certain reduced density matrices or via protocols using multiple physical copies, as discussed later). However, we focus on an implementation making use of
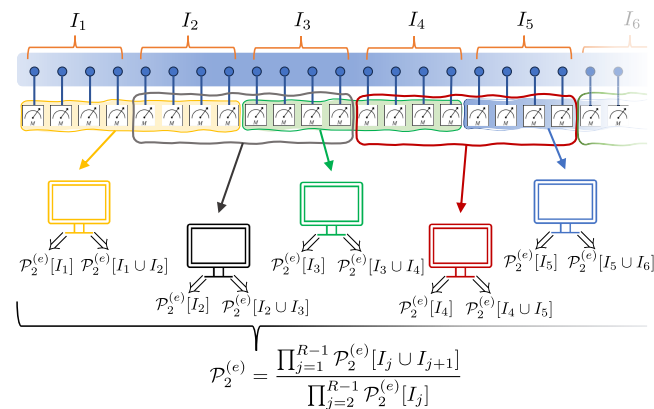

FIG. 1. A 1D quantum system $\mathcal{S}$ is partitioned into adjacent intervals I_j of size k, $|I_j| = k$. In this work, we consider extracting the purities over I_j and $I_j \cup I_{j+1}$ using the classical-shadow approach [15]. Performing M local measurements with respect to randomly chosen bases and classical postprocessing operations, they yield faithful estimators for $\mathrm{Tr}[\rho_I^2]$ denoted by $\mathcal{P}_2^{(e)}[I]$. Such subsystem-purity estimators are then combined to obtain a prediction $\mathcal{P}_2^{(e)}$ for the global purity. The accuracy of the method is controlled by k and the state correlation lengths.

the standard RM toolbox [12]. Our motivation is twofold. On the one hand, there is rich literature studying the statistical errors associated with reconstructing purities and PT moments using RM measurements, allowing us to make our estimates very explicit. On the other hand, RM schemes are a very practical tool, which is now routinely employed in various experimental platforms [12–15].

The proposed approach is different from full tomographic methods relying on prior assumptions on the system state. There, a common strategy is to assume that the system state is described by a sufficiently simple ansatz wave function, and estimate the values of its free parameters. This logic was first put forward in the context of matrix-product states (MPSs) [59–62], while recently extended to matrix-product density operators (MPDOs) [63–67], Gibbs states [68–73], permutation-invariant states [74–76], low-rank states [77], stabilizers [78], and tensor- and neural-network wave functions [79–84]. While these methods may be practically very useful, they also have drawbacks. For instance, as we discuss in more detail in Sec. III A, an accurate estimation of, say, global purities might require reconstructing the state up to exponential precision, thus leading to unpractical overheads in N. On the contrary, our approach does not learn the state wave function, targeting purities and PT moments directly; cf. Fig. 1.

Finally, we mention that our ideas could be extended, in some cases straightforwardly, to probe other types of quantities generally requiring exponentially many measurements, including participation entropies [85–91] or stabilizer Rényi entropies [92], which were recently considered in the many-body setting [92–99]. In addition, while we focus on 1D systems, where analytic and numerical analyses are simpler, we expect that our approach could be generalized to higher spatial dimensions. Therefore, our work also opens up a number of important directions for future research.

The rest of this manuscript is organized as follows. After reviewing a few preliminary notions and tools in Sec. II, we start by introducing the main ideas underlying our approach in Sec. III. To this end, we consider the class of so-called finite-depth quantum-circuit (FDQC) states, which provide an ideal toy model for short-range correlated many-body quantum states. Their minimal structure allows us to remove unnecessary technical complications from the discussion and present the logic of our method in the simplest possible setting. We consider both purities (Sec. III A) and PT moments (Sec. III B), working out efficiency performance guarantees for their estimation.

The most general form of our protocol is presented in Sec. IV. After introducing the AFCs in Sec. IV A, we rigorously derive performance guarantees for the accurate estimation of the purity and PT moments (Sec. IV B) under the assumption that the state to be measured satisfies the AFCs. We then discuss the generality of the AFCs, proving that they are satisfied in translation-invariant MPDOs

(Sec. IV C) and presenting numerical evidence for their validity in thermal states of local Hamiltonian; cf. Sec. IV D. We also present a full classical simulation of the measurement protocol (Sec. IV E), studying in a concrete example the typical number of measurements required to estimate the bipartite purity of area-law pure states.

Next, in Sec. V we discuss a few natural examples of highly mixed states where bipartite entanglement can be detected for large system sizes by estimating just the first few PT moments. Since the first few PT moments can be very simply extracted by our approach, we argue that our results could be practically useful to probe mixed-state entanglement in experimentally available noisy quantum platforms. Finally, we report our conclusions in Sec. VI, while the most technical parts of our work are consigned to several appendixes.

II. ENTANGLEMENT AND RANDOMIZED MEASUREMENTS

A. Mixed-state entanglement and PPT conditions

We consider a system of N qubits denoted by $\mathcal{S}$. Given a region $I \subset \mathcal{S}$, the associated Hilbert space is $\mathcal{H}_I \simeq \mathbb{C}^{\otimes 2|I|}$, where we denoted by $|I|$ the number of qubits in I. We will be interested in the Rényi entropies of the region I,

$$S_n[I] = \frac{1}{1-n} \ln \mathcal{P}_n[I], \qquad (1)$$

where

$$\mathcal{P}_n[I] = \text{Tr}[\rho_I^n], \qquad (2)$$

and ρ_I is the reduced density matrix on the region I. For $n = 2$, $\mathcal{P}_2[I]$ coincides with the purity, which is a simple probe for the subsystem entropy, with $\mathcal{P}_2[I] = 1$ and $\mathcal{P}_2[I] = 2^{-|I|}$ for pure and maximally mixed states, respectively.

Consider now a partition of $\mathcal{S}$ into two disjoint sets $\mathcal{S} = A \cup B$, yielding $\mathcal{H}_{\mathcal{S}} = \mathcal{H}_A \otimes \mathcal{H}_B$. If the system is in a pure state $|\psi\rangle_{AB} \in \mathcal{H}_{\mathcal{S}}$, its bipartite entanglement is quantified by the Rényi entropies $S_n(\rho_A)$ [17]. Conversely, when the state of the system is mixed, its entanglement can be quantified by the logarithmic negativity [100,101]

$$\mathcal{E}(\rho) = \ln \sum_j |\lambda_j|, \qquad (3)$$

where the sum is over all eigenvalues $\{\lambda_j\}_j$ of the operator $\rho_{AB}^{T_A}$, and $(\cdot)^{T_A}$ denotes partial transpose with respect to subsystem A. The spectrum of the PT density matrix, and hence the logarithmic negativity, is completely fixed by the PT moments

$$p_n = \text{Tr}[(\rho_{AB}^{T_A})^n] \qquad (4)$$

for $n = 1, 2, \ldots, \dim(\mathcal{H}_A \otimes \mathcal{H}_B)$. Note that $p_1 = 1$, while the second PT moment coincides with the purity $p_2 = \mathcal{P}_2$ [35].

The importance of the PT moments is twofold. On the one hand, they can be accessed directly via RMs [34,35] (or using quantum interference [102,103]); see Sec. II B. On the other hand, the knowledge of the first few PT moments is enough to certify bipartite entanglement based on the nonpositivity of the partial-transposed density matrix [34,35,104–106], or to detect different types of entanglement structures [107]. In this work, we will consider a particular set of conditions on the PT moments to certify bipartite entanglement, which were derived in Refs. [35,106] and which we call the p_n-positive partial transpose (PPT) conditions. Denoting by SEP the set of separable, i.e., not entangled, states in $\mathcal{S}$, the p_n-PPT conditions take the form

$$\rho \in \text{SEP} \Rightarrow p_n p_{n-2} \geq p_{n-1}^2. \qquad (5)$$

Therefore, when the state of the system ρ violates the p_n-PPT conditions, it is entangled, and the difference $p_{n-1}^2 - p_n p_{n-2}$ is a probe for mixed-state entanglement. Note that, for $n = 3$, Eq. (5) coincides with the relation first derived in Ref. [35]. These conditions are in general not optimal and are a strict subset of those derived in Refs. [105,106]. Still, their simplicity makes them particularly convenient for our purposes.

B. Randomized measurements and classical shadows

The power of RM schemes lies in the fact that they need not be tailored to a specific property of the system. Rather, one performs measurements which are randomly sampled from a fixed ensemble independent of the observable of interest. Subsequently, the outcomes are processed differently depending on the quantity to be estimated [28,29,108,109]. Denoting by ρ the system density matrix, this approach gives us access to all observable expectation values $\text{Tr}[\rho O]$ and, more generally, to multicopy objects of the form $\text{Tr}[\rho^{\otimes n} O]$, where the integer $n \geq 1$ is called the copy (or replica) index.

In this section, we recall the basic aspects of RMs used in our work. While the logic explained in the next sections may be implemented in different ways, we will focus on a set of protocols making use the (local) classical shadows introduced in Ref. [15], a prominent element in the RM toolbox [12]. We briefly recall the main aspects of the formalism, while we refer to Refs. [12,15] for a thorough introduction.

In what follows, we denote by $|0\rangle_j$ and $|1\rangle_j$ the basis elements of the local computational basis corresponding to qubit j, spanning $\mathcal{H}_j \simeq \mathbb{C}^2$. In the classical-shadow framework, one performs a set of M measurements (one per experimental run, each labeled by an integer r) consisting of local unitary operations $\prod_j u_j^{(r)}$ followed by a projective measurement onto the computational basis
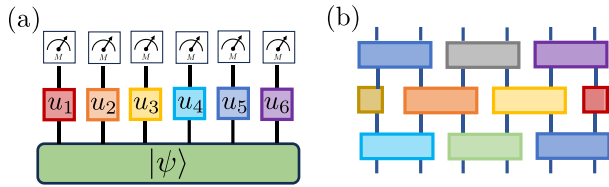
FIG. 2. (a) Within the classical-shadow approach, each measurement process consists of random on-site unitaries followed by local projective measurements. Their outcomes are stored and later postprocessed to construct the classical shadows. (b) Pictorial representation of a local finite-depth quantum circuit. The gates are arranged in a brickwork pattern, forming layers of mutually commuting unitary operations acting on pairs of neighboring qubits. Lower (upper) dangling legs correspond to the input (output) qubits. The depth of the circuit is the number of applied layers. In this picture, the depth is $\ell = 3$.

$$|k_1, \ldots, k_N\rangle = \bigotimes_{j=1}^{N} |k_j\rangle_j, \qquad (6)$$

with $k_j = 0, 1$. The unitaries are sampled from a Haar-random ensemble, identically and independently for each qubit j and experimental run r; see Fig. 2. Denoting by $\{k_j^{(r)}\}$ the set of outcomes of this two-step process, the values $\{k_j^{(r)}\}$ and the unitaries $\{u_j^{(r)}\}$ are used to define the so-called classical shadows

$$\rho_{\mathcal{S}}^{(r)} = \bigotimes_{i \in \mathcal{S}} [3(u_i^{(r)})^\dagger |k_i^{(r)}\rangle\langle k_i^{(r)}| u_i^{(r)} - \mathbb{1}_2], \qquad (7)$$

which can be classically stored in an efficient way.

As mentioned, the measurement protocol does not depend on the observable of interest. Rather, one adapts the postprocessing operations on the classical shadows based on the quantity to be estimated. For instance, given any observable O, an estimator for its expectation value is

$$\hat{o} = \frac{1}{M} \sum_{r=1}^{M} \text{Tr}[O \rho_{\mathcal{S}}^{(r)}]. \qquad (8)$$

It is easy to see that $\hat{o}$ is faithful, i.e., unbiased, while different bounds for the statistical variance of this estimator may be derived depending on the locality properties of O [12,15].

It is important to recall that classical shadows also give access to entropy and PT moments. Let us consider in particular the purity, which is given by Eq. (2) for $n = 2$. For this quantity, the classical shadows allow us to write the estimator [15]

$$\mathcal{P}_2^{(e)}[I] = \frac{1}{M(M-1)} \sum_{r \neq r'} \text{Tr}(\rho_I^{(r)} \rho_I^{(r')}), \qquad (9)$$

where ρ_I is defined as in Eq. (7). $\mathcal{P}^{(e)}[I]$ is a faithful estimator. Note that, alternatively, one can use another

estimator for the purity [13,27,29] that provides robustness against miscalibration errors using the same data. The statistical errors associated with $\mathcal{P}_2^{(e)}[I]$ are quantified by its variance, which can be bounded by [15,35,38]

$$\text{Var}[\mathcal{P}_2^{(e)}[I]] \leq 4\left(\frac{2^{|I|} \mathcal{P}_2[I]}{M}\right) + 2\left(\frac{2^{2|I|}}{M-1}\right)^2, \qquad (10)$$

where $|I|$ denotes the number of qubits in I. This bound is known to be essentially optimal [35,38], telling us that an exponentially large number of measurements is needed to estimate the purity.

The PT moments (4) can be treated similarly. In this case, one constructs the estimator [35]

$$p_n^{(e)}[AB] = \frac{1}{n!} \binom{M}{n}^{-1}$$
$$\times \sum_{r_1 \neq r_2 \neq, \ldots, \neq r_n} \text{Tr}[[\rho_{AB}^{(r_1)}]^{T_A} \cdots [\rho_{AB}^{(r_n)}]^{T_A}]. \qquad (11)$$

Once again, $p_n^{(e)}[AB]$ is faithful, and it is possible to derive explicit bounds on its variance, although it becomes increasingly involved for higher n. For instance, for $n = 3$ one has [35,38]

$$\text{Var}[p_3^{(e)}[AB]] \leq 9\frac{2^{|AB|}}{M} \text{Tr}(\rho_{AB}^4) + 18\frac{2^{3|AB|}}{(M-1)^2} p_2[AB]$$
$$+ 6\frac{2^{6|AB|}}{(M-2)^3}. \qquad (12)$$

Estimating the statistical errors by the variance, Eqs. (10) and (12) imply that, in order to guarantee an accurate reconstruction of the purity and PT moments of the system, exponentially many measurements in its size are needed. As mentioned, this makes it unfeasible to significantly scale up previous experiments making use of this strategy [13,15,27–29,35]. The goal of this work is to show that these limitations may be overcome under assumptions which are very common in the context of many-body physics, and thus also natural from the point of view of quantum simulation. Namely, we will put forward a set of protocols for entropy and entanglement estimation which are provably efficient assuming that all spatial correlation lengths of the state to be measured are finite. We will focus on 1D systems, where analytic and numerical analyses are simpler, although we expect that our approach could be generalized to higher spatial dimensions; see Sec. VI.

III. A TOY MODEL FOR THE ESTIMATION PROTOCOL

In this section, we introduce the main ideas underlying our approach, focusing on a simplified setting where we

can get rid of unnecessary technical complications. We analyze the case of FDQC states, where the state to be measured is prepared by a shallow (local) quantum circuit

$$\rho = U^{(\ell)} \left(\bigotimes_{j=1}^{L} \sigma_j \right) [U^{(\ell)}]^\dagger, \tag{13}$$

where L is length of the system, i.e., the number of qubits (we use the letter L instead of N, as in the previous section, to emphasize that we are focusing on the 1D case). Here, σ_j are arbitrary single-qubit density matrices, while $U^{(\ell)}$ is a local circuit of depth ℓ, namely, $U^{(\ell)} = V_\ell, \dots, V_2 V_1$, where V_j contains quantum gates acting on disjoint pairs of nearest-neighbor qubits; cf. Fig. 2. We will assume that ℓ is fixed, i.e., not increasing with the system size. We do not ask for translation symmetry and, unless specified otherwise, assume open-boundary conditions. The gates making up $U^{(\ell)}$ can be arbitrary; i.e., they need not be taken out of some finite gate set. The FDQC states (13) have a very simple structure from the point of view of many-body physics, but are known to approximate physically interesting states such as MPSs [110,111] and, more generally, ground states of gapped local Hamiltonians [112–115].

Throughout this section, we will assume that we know that the state of the system is exactly of the form (13) for some finite ℓ. We develop a protocol to efficiently estimate the purity and the PT moments of such a state, based on this knowledge. The protocol takes only the depth of the circuit ℓ as an input and does not make use of state tomography. In fact, it will be later generalized replacing the assumption of the FDQC structure with the AFCs, which make no explicit reference to the form of the state wave function.

A. Purity estimation: Factorization formula

The starting point of our method is a factorization property for powers of the system density matrix. Let I be any interval of adjacent qubits (I can coincide with the full system $\mathcal{S}$). Consider a partition $I = A \cup B \cup C$, where B separates A and C and denote by $|B|$ the number of qubits in B; cf. Fig. 3(a). Based on the fact that the state is prepared by a finite-depth circuit, Eq. (13), one can prove
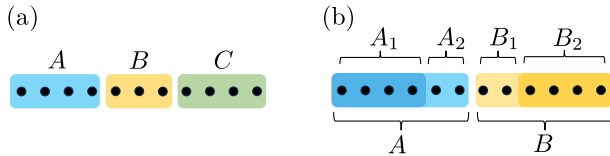
(a)   (b)

FIG. 3. (a) Partition considered in Eq. (14): A 1D interval I is divided into three disjoint regions A, B, and C, where B separates A and C. (b) Partition considered in Eq. (26): A 1D interval I is divided into two halves A and B. Each half is partitioned into two regions $A = A_1 \cup A_2$, $B = B_1 \cup B_2$.

$$\mathrm{Tr}(\rho_I^2) = \frac{\mathrm{Tr}_{AB}(\rho_{AB}^2)\mathrm{Tr}_{BC}(\rho_{BC}^2)}{\mathrm{Tr}_B(\rho_B^2)} \tag{14}$$

for any partition with $|B| \geq 2\ell - 1$, where ρ_X denotes the density matrix reduced to the interval X, and XY is shorthand notation for $X \cup Y$. Now, let $\{I_j\}_{j=1}^R$ be a collection of adjacent intervals covering the system $\mathcal{S}$, with $|I_j| = k \geq 2\ell - 1$, and assume without loss of generality that $R = L/k$ is an integer; cf. Fig. 1. By applying Eq. (14) iteratively, we arrive at

$$\mathrm{Tr}(\rho^2) = \frac{\prod_{j=1}^{R-1} \mathrm{Tr}_{I_j \cup I_{j+1}}(\rho_{I_j \cup I_{j+1}}^2)}{\prod_{j=2}^{R-1} \mathrm{Tr}_{I_j}[\rho_{I_j}^2]}, \tag{15}$$

where ρ is the system density matrix (13). A proof of Eqs. (14) and (15) is given in Appendix A.

Equation (15) is remarkable because its right-hand side (rhs) depends only on the purities of subsystems containing up to $2k$ qubits, and thus gives us a natural basis for an efficient estimation strategy of the global purity. The idea is to first reconstruct the "local" purities $\mathrm{Tr}[\rho_I^2]$ (with either $I = I_j$ or $I = I_j \cup I_{j+1}$), and subsequently plug their estimated values in the rhs of Eq. (15), yielding the global purity estimate. This method is more efficient than directly targeting $\mathrm{Tr}(\rho^2)$, as the local purities can be reconstructed from a number of measurements growing exponentially in k, not in L; see Eq. (10). Below, we give a more precise description of the protocol, proving that an accurate estimation of the purity requires only a polynomial (in L) number of measurements and postprocessing operations.

We consider performing M_I measurements (i.e., experimental runs) to estimate the purity of each interval I, with $I = I_j$ or $I = I_j \cup I_{j+1}$; see Eq. (15). The set of measurements M_I is used to determine only the purity over I, and for simplicity we perform the same number of measurements $M_I = M$ for each interval. Although this increases the total number of experiments M_T by a factor L, since

$$M_T = \sum_I M_I \propto LM, \tag{16}$$

this procedure guarantees that statistical errors associated with distinct regions are independent and facilitate their rigorous analysis. For each interval, we then construct the estimators $\mathcal{P}^{(e)}[I]$ given in Eq. (9), and define

$$r_2^{(e)} = \frac{\prod_{j=1}^{R-1} \mathcal{P}_2^{(e)}[I_j \cup I_{j+1}]}{\prod_{j=2}^{R-1} \mathcal{P}_2^{(e)}[I_j]}, \tag{17}$$

which is our estimator for the global purity. Since the latter purity is typically exponentially small in L, we quantify the accuracy of $r_2^{(e)}$ by the relative error

$$\varepsilon_r = \left| \frac{r_2^{(e)}}{\mathcal{P}_2} - 1 \right|. \tag{18}$$

Next, we seek to bound the number of measurements required to guarantee that ε_r is small with high probability. This problem is solved in Appendix B 1, where we prove the following result: For any arbitrarily small $\delta > 0$, choosing

$$M \geq \max\left\{ 2^{8k}, L^2 \frac{2^{4k+10}}{k^2\delta^2} \right\}, \tag{19}$$

the probability that $|\varepsilon_r| \geq \delta$ satisfies

$$\Pr\left[|\varepsilon_r| \geq \delta\right] \leq \frac{2^{4k+11}L^3}{\delta^2 k^3 M}, \tag{20}$$

where we recall that $k = |I_j|$. The proof of Eq. (20) is based on a careful analysis of how the error on each factor in Eq. (17) affects the global purity, making use of the statistical independence of $\mathcal{P}_2^{(e)}[I]$ for different I and the so-called Chebyshev's inequality. Equations (16), (19), and (20) then imply that polynomially many (in L) measurements and postprocessing operations are enough to accurately estimate the purity, as anticipated. [The polynomial scaling of the postprocessing operations follows straightforwardly from the definition (9), because the number of elements in each sum scales as M^2, where $M = O(L^3)$.] We can rephrase this result in a more transparent way. For a given confidence level $\gamma = \Pr\left[|\varepsilon_r| < \delta\right]$, Eq. (20) implies that it is enough to take a number of measurements

$$M \geq \frac{2^{4k+11}L^3}{\delta^2 k^3 (1 - \gamma)}. \tag{21}$$

Before concluding this section, a few remarks are in order. First, it is important to note that the individual factors in the formula (17) could also be extracted using different RM schemes or even standard tomography for the density matrices ρ_I. Indeed, the possibility of estimating the global purity from polynomially many measurements does not depend on the fact that we are using classical shadows, but rather on the factorization property (17) (since its rhs involves only local subsystems). Still, classical shadows and RMs in general offer several advantages compared to state tomography. For instance, while RM schemes to estimate $\mathrm{Tr}[\rho_I^2]$ require a number of measurements scaling exponentially in $|I|$, the exponents are typically favorable compared to tomography [27,31,116]. In addition, classical shadows make it very easy to rigorously bound statistical errors, facilitating the analyses presented throughout this work.

Second, taking the logarithm of Eq. (15), we obtain

$$S^{(2)}(\rho) = \sum_{j=1}^{R-1} S^{(2)}(\rho_{I_j \cup I_{j+1}}) - \sum_{j=2}^{R-1} S^{(2)}(\rho_{I_j}), \tag{22}$$

where $S^{(2)}(\rho) = -\ln\mathrm{Tr}\rho^2$ is the second Rényi entropy. [We note that similar formulas previously appeared (albeit for the von Neumann, rather than for the Rényi entropy) in the context of approximate Markov-chain states [117–120]; see also Refs. [121–128].] Therefore, our results can equivalently be formulated in terms of Rényi entropies, rather than purities (note that a small relative error on the latter implies a small additive error on the Rényi entropy).

Finally, as the state (13) can be represented exactly as a matrix-product operator (MPO) [129], it is instructive to compare our strategy with those based on MPDO tomography [63–65,65–67]. In many cases, these methods can efficiently provide an estimate of the system density matrix $\rho^{(e)}$ satisfying

$$\|\rho^{(e)} - \rho\|_1 \leq \delta, \tag{23}$$

where $\|\cdot\|_1$ denotes the trace norm, while $\delta = O(L^{-\alpha})$ is a small parameter vanishing polynomially in L. While this approximation allows us to accurately estimate the expectation value of any local observable, it might not be enough to extract the global Rényi-2 entropy. Indeed, given ρ and σ with $\delta = \|\rho - \sigma\|_1$, we have the following bound, which is known to be tight in general [130,131]:

$$|S^{(2)}(\rho) - S^{(2)}(\sigma)| \leq 2^L \left[1 - (1 - 2\delta)^2 - \frac{4\delta^2}{(2^L - 1)} \right]$$
$$\sim 2^L \delta. \tag{24}$$

Therefore, a precise estimation of the Rényi-2 entropy requires an exponentially accurate reconstruction of the system density matrix, typically leading to unpractical overheads in L. (We note that a precise estimation of the von Neumann entanglement entropy, instead, requires only to reconstruct the target state up to an error which decays polynomially in the system size [132]. However, computing the von Neumann entanglement entropy for MPDOs is expected to be computationally hard in general.) Instead, our method gets around this technical issue, as it does not rely on state tomography.

B. The normalized PT-moment estimation

The ideas presented in the previous section may be applied to the PT moments, although a few subtleties must be taken into account. Denoting by $p_n^{(e)}$ our estimate for p_n, one is tempted to ask for a protocol that makes the relative error $|p_n^{(e)}/p_n - 1|$ sufficiently small. This is, however, problematic: Contrary to the moments $\mathcal{P}_n$, it is nontrivial to

bound $|p_n|$ from below by a positive number. In order to get around this issue, we define the normalized PT moment

$$\tilde{p}_n[AB] = \frac{p_n[AB]}{\mathcal{P}_n[A]\mathcal{P}_n[B]}, \tag{25}$$

and, instead of focusing on the relative error, we ask for an accurate estimate of $\tilde{p}_n$ up to a small additive error. The motivation for this choice is twofold. First, as it will be clear later, for FDQC states one can show that $\tilde{p}_n[AB]$ is independent of the system size. Therefore, contrary to the purity, one does not have to deal with numbers which are exponentially small in L. Second, we will see in Sec. V that entanglement certification based on the p_n-PPT conditions requires approximating $\tilde{p}_n[AB]$ up to a small additive error.

In order to estimate the normalized PT moment (25), we once again rely on certain factorization properties of the density matrix. Consider the FDQC state (13) and take a partition of the system as in Fig. 3(b), with $|B_1| = |A_2| = k \geq 2\ell - 1$, where ℓ is the depth of the circuit. As we show in Appendix A 2, one can prove

$$\text{Tr}[(\rho_{AB}^{T_A})^n] = \text{Tr}[(\rho_{A_2B_1}^{T_{A_2}})^n]\frac{\text{Tr}_A[\rho_A^n]\text{Tr}_B[\rho_B^n]}{\text{Tr}_{A_2}[\rho_{A_2}^n]\text{Tr}_{B_1}[\rho_{B_1}^n]}. \tag{26}$$

This equation implies

$$\tilde{p}_n[AB] = s_n[A_2B_1], \tag{27}$$

where we define

$$s_n[XY] = \frac{\text{Tr}[(\rho_{XY}^{T_X})^n]}{\text{Tr}_X[\rho_X^n]\text{Tr}_Y[\rho_Y^n]}. \tag{28}$$

Note that $s_n[A_2B_1]$ depends only on the density matrix of a local subsystem and does not scale with the system size, as anticipated (because of the FDQC structure, the reduced density matrix over the region A_2B_1 is independent of the system size for $L \geq |A_2B_1| + 2\ell - 1$). Therefore, it is possible to obtain an accurate estimate from a number of measurements and postprocessing operations independent of L, for any arbitrary small additive error.

To see this explicitly, we consider the following protocol: We introduce an estimator for $s_n[A_2B_1]$, namely,

$$s_n^{(e)} = \frac{p_n^{(e)}[A_2B_1]}{\mathcal{P}_n^{(e)}[A_2]\mathcal{P}_n^{(e)}[B_1]}. \tag{29}$$

The estimators for the moments $\mathcal{P}_n^{(e)}[A_2]$, $\mathcal{P}_n^{(e)}[B_1]$, and the PT moment $p_n^{(e)}[A_2B_1]$ are defined in Eq. (11). We compute each of them out of M different classical shadows, so that the total number of experimental runs is $M_T = 3M$. This is done so that the statistical errors are independent,

facilitating their analysis. After these steps, we obtain an estimate for $s_n[A_2B_1]$, and thus, due to Eq. (26), for $\tilde{p}_n[AB]$.

Importantly, we can bound the additive statistical error which affects our estimate, namely,

$$\varepsilon_a = |s_n^{(e)} - \tilde{p}_n[AB]|. \tag{30}$$

For instance, focusing for simplicity on the case $n = 3$, we prove in Appendix B 2 the following result: For any small $\delta > 0$ and choosing

$$M \geq 27\frac{2^{11k+9}}{\delta^2}, \tag{31}$$

we have

$$\text{Pr}\left[|s_3^{(e)} - \tilde{p}_3[AB]| \geq \delta\right] \leq 81\frac{2^{11k+9}}{M\delta^2}, \tag{32}$$

where $|A_2| = |B_1| = k \geq 2\ell - 1$, with ℓ being the circuit depth. This result is proved by the same techniques used to derive Eq. (20). Again, it is useful to rephrase our result in terms of the confidence level γ, yielding

$$M \geq 81\frac{2^{11k+9}}{(1-\gamma)\delta^2}. \tag{33}$$

Equation (32) states that a number of measurements not scaling with the system size is enough to guarantee that $\tilde{p}_3$ is approximated with arbitrary precision and high probability. Explicit performance guarantees such as Eq. (32) for higher integer values of n are more cumbersome to derive and will be omitted. Still, based on the analysis presented in Appendix B 2, one can easily see that inequalities of the form (32) hold for higher n too, where the rhs is always independent of L.

IV. FINITE-RANGE CORRELATED STATES

In this section, we present and discuss the most general form of our protocols. First, in Sec. IV A we give the definition of the AFCs, which state that the nth powers of the system density matrix effectively factorize over regions of size smaller than some length scales ξ_n. Our definition might appear arbitrary at first, but we show later that such AFCs hold in large classes of states (see Secs. IV C and IV D). Next, in Sec. IV B we describe a protocol for purity and PT-moment estimation in states satisfying the AFCs. Such protocols take as an input the value of the length scales ξ_n (assumed to be known) and the desired accuracy, yielding as an output the numerical estimations for the purity and PT moments. The protocols require only polynomially many (in L) measurements and postprocessing operations. In addition, we can rigorously prove that the estimation errors are smaller than the desired threshold with high probability. Finally, we discuss the generality of the

AFCs, proving analytically that they hold for MPDOs (Sec. IV C) and presenting numerical evidence for their validity in Gibbs states of local Hamiltonians (Sec. IV D).

A. The approximate factorization conditions

The strategies developed so far rely on exact factorization conditions on the system density matrix, which are related to the sharp light-cone structure of correlation functions in FDQC states [110,133,134]. Away from these ideal cases, Eqs. (14) and (26) do not hold, but it is natural to conjecture that they can be modified taking into account exponential corrections over scales governed by the system correlation lengths. This is done in this section, where we introduce a set of AFCs generalizing Eqs. (14) and (26), respectively.

Given a state ρ, we say that it satisfies the purity AFC if there exist $\alpha_2, \beta_2, k_c > 0$ such that, for any interval $I = A \cup B \cup C$ as in Fig. 3(a), with $|B| = k \geq k_c$, we have

$$\left| \mathrm{Tr}(\rho_I^2)^{-1} \frac{\mathrm{Tr}_{AB}(\rho_{AB}^2)\mathrm{Tr}_{BC}(\rho_{BC}^2)}{\mathrm{Tr}_B(\rho_B^2)} - 1 \right| \leq \alpha_2 e^{-\beta_2 |B|}. \quad (34)$$

Similarly, we say that ρ satisfies the PT-moment AFCs if there exist $\alpha_n, \beta_n > 0$, and an integer k_c such that

$$\left| \tilde{p}_n[AB] - s_n[A_2 B_1] \right| \leq \alpha_n e^{-\beta_n(|A_2|+|B_1|)} \quad (35)$$

for all partitions as in Fig. 3(b) with $|B_1| = |A_2| = k \geq k_c$. Here, $\tilde{p}_n[AB]$ and $s_n[A_2 B_1]$ are defined in Eqs. (25) and (28), respectively. Equations (34) and (35) obviously generalize Eqs. (14) and (26), introducing exponential corrections over length scales $\xi_n = \beta_n^{-1}$.

While the definitions (34) and (35) might appear arbitrary at first, we show later that such AFCs hold in large classes of states (see Secs. IV C and IV D). This statement is very intuitive: On the one hand, we have shown that they hold exactly for FDQC states; on the other hand, in the absence of topological order, ground states are known to be well approximated by FDQC states [112–115]. It is then natural to expect that the AFCs continue to hold as the temperature is increased or as the system is perturbed by incoherent noise, as these ingredients do not introduce long-range quantum correlations. In the next section, we assume that the AFCs hold, and describe a protocol for purity and PT-moment estimation. We stress that we do not need any additional assumption on the state to be measured. For instance, we do not require that it can be represented efficiently by an MPDO.

Our protocols take as an input the values of α_n, β_n, and the desired accuracy. For instance, the purity estimation protocol takes as an input α_2, the correlation length $\xi_2 = \beta_2^{-1}$, and the target threshold relative error δ. Importantly, it is not necessary to know the values of α_2 and β_2 exactly: It is sufficient to have two estimates $\tilde{\alpha}_2, \tilde{\xi}_2$, which upper bound them. This is because if $\alpha_2 \leq \tilde{\alpha}_2$ and

$\xi_2 \leq \tilde{\xi}_2$, then Eq. (34) also holds replacing α_2 and β_2 with $\tilde{\alpha}_2$ and $\tilde{\beta}_2 = \tilde{\xi}_2^{-1}$, respectively. We can choose $\tilde{\alpha}_2$ and $\tilde{\xi}_2$ arbitrarily, but the number of measurements and postprocessing operations increase parametrically with $\tilde{\alpha}_2$ and $\tilde{\xi}_2$; see Eq. (37). Therefore, the more accurately we can estimate the values of α_2 and β_2, i.e., the more information we have on the state to be measured, the more efficiently we can estimate the purity and, similarly, the PT moments. It is important to note that, in this respect, the situation is completely analogous to MPDO tomography. In that case, one assumes that the density matrix of the system ρ can be described accurately by some MPO with a bond dimension D. Such a bond dimension does not need to be known exactly, but only an upper bound for it needs to be known. In practice, one can take α_2 and β_2^{-1} larger than any expected length scale in the system.

Finally, suppose that, for given state ρ, we have an ansatz for α_2 and ξ_2. From the experimental point of view, an important question is whether it is possible to efficiently certify that Eq. (34) holds, with the ansatz values α_2 and ξ_2, for the state to be measured. While at the moment this is an open question, a simple consistency check consists of repeating the purity estimation protocol (explained in Sec. IV B) for increasing values of the input ansatz values α_2 and ξ_2, and checking that the estimated value for the purity does not change, up to the expected precision. If the estimated purity does change, this is a "red flag" signaling the failure of the AFCs. We refer to Sec. VI for further discussion on the possibility to efficiently certify the validity of the AFCs.

B. Estimation protocols for states satisfying the AFCs

First, we consider estimating the global purity of a state satisfying the AFCs (34). We assume the following:

 (i) We know that the system is prepared in a state ρ satisfying the AFCs (34).
 (ii) We know two numerical values α_2 and β_2 for which Eq. (34) is satisfied.

Below, we detail our protocol to efficiently estimate the purity. The proof that the protocol works is nontrivial and is presented in detail in Appendix C. Intuitively, we use the fact that, because of the AFC (34), the estimator $r_2^{(e)}$ defined in Eq. (17) yields a good approximation for the purity, for values of $|I_j|$ which grow mildly (logarithmically) in L.

The protocol takes as an input the values of α_2 and β_2, and consists of the following steps:

 (1) Choose the desired relative error δ on the purity. This can be any arbitrarily small number $\delta > 0$.
 (2) Take a partition of the system as in Fig. 1, where $|I_j| = k$ and choose

$$k \geq \xi_2 \ln(7\alpha_2 L/\delta), \quad (36)$$

where $\xi_2 = \beta_2^{-1}$. The reason for the functional form appearing on the rhs is technical and explained in Appendix C.

(3) As in Sec. III A, perform $M_I = M$ classical-shadow measurements to estimate the purity of each interval I, with $I = I_j$ or $I = I_j \cup I_{j+1}$ [we use the set of measurements M_I to determine only the purity over I, so that the total number of experimental runs is given by Eq. (16)].

(4) From the classical shadows obtained out of the measurements, compute $r_2^{(e)}$ defined Eq. (17). This is the output of our protocol, giving us an estimate for the global purity.

As mentioned, it is intuitive that the output of this protocol $r_2^{(e)}$ is, with high probability, an accurate approximation for the purity. This is proven rigorously in Appendix C, where we show [more precisely, Eq. (37) holds for M sufficiently large, namely, $M \geq \max\{(7\alpha_2 L/\delta)^{8\xi_2 \ln 2}, (L^2 7^2 2^{10}/\delta^2)(7\alpha_2 L/\delta)^{4\xi_2 \ln 2}\}$; see Appendix C]

$$\Pr\left[|(r_2^{(e)}/\mathcal{P}_2) - 1| \geq \delta\right] \leq \frac{7^2 2^{11} L^3}{\delta^2 M} \left(\frac{7\alpha_2 L}{\delta}\right)^{4\xi_2 \ln 2}. \quad (37)$$

Therefore, by a polynomial number of measurements M we can guarantee that the probability that the relative error $|r_2^{(e)}/\mathcal{P}_2 - 1|$ is larger than δ is arbitrarily small. Rephrasing our result in terms of the confidence level γ, we thus obtain

$$M \geq \frac{7^2 2^{11} L^3}{\delta^2(1-\gamma)} \left(\frac{7\alpha_2 L}{\delta}\right)^{4\xi_2 \ln 2}. \quad (38)$$

This formalizes the anticipated result.

Note that $r_2^{(e)}$ defined above is a faithful estimator for the quantity

$$r_2 = \frac{\prod_{j=1}^{R-1} \mathrm{Tr}_{I_j \cup I_{j+1}}(\rho_{I_j \cup I_{j+1}}^2)}{\prod_{j=2}^{R-1} \mathrm{Tr}_{I_j}[\rho_{I_j}^2]}, \quad (39)$$

but because Eq. (15) does not hold exactly, it is not a faithful estimator of the global purity $\mathcal{P}_2$. Still, it is a good approximation, as it is clear from the relation

$$\left|\frac{r_2}{\mathcal{P}_2} - 1\right| \leq \frac{4\alpha_2 L}{k} e^{-\beta_2 k}, \quad (40)$$

which holds for

$$k \geq \xi_2 \ln(2\alpha_2 L), \quad (41)$$

and follows directly from Eq. (34); cf. Appendix C. In practice, the rhs of Eq. (40) makes it necessary to consider intervals I_j whose length k grows logarithmically in L to keep the error below the desired threshold.

A similar protocol works for the PT moments. In this case, a number of measurements, which does not scale with the system size, is enough to guarantee arbitrary accuracy with high probability; cf. Appendix C. We note that the results of this section imply that the number of postprocessing operations to estimate the purity and PT moments is also at most polynomial in L. This is because the estimators (9) and (11) are constructed summing a number of terms which are polynomial in M, and hence, at most polynomial in L.

In the rest of this section, we discuss the generality of the AFCs. We prove analytically that they hold for the important class of translation-invariant MPDOs and provide numerical evidence that they are also satisfied by thermal states of local Hamiltonians. These results showcase the generality and versatility of our approach.

C. Proof of AFCs in matrix-product density operators

We now focus on MPDOs, a very general class of states providing accurate approximations for several short-range correlated density matrices, including thermal states of local Hamiltonians [60,135]. We study the translation-invariant case, where we can provide analytic results. We recall that a translation-invariant MPDO σ is defined by a four-index tensor $A_{a,b}^{j,k}$, with j, $k = 0$, 1 and $a, b = 0, 1, \ldots, \chi - 1$, where χ is its bond dimension [129]. For a system size L, its matrix elements read

$$\langle i_1, \ldots, i_L | \sigma_L | j_1, \ldots, j_L \rangle = A_{a_1,a_2}^{i_1,j_1} A_{a_2,a_3}^{i_2,j_2} \cdots A_{a_L,a_1}^{i_L,j_L}, \quad (42)$$

where repeated indices are summed over. This defines a family of normalized states (we assume that the local tensors A generate a positive operator for all system sizes L, namely, $\sigma_L \geq 0$ for all L)

$$\rho_L = \frac{\sigma_L}{\mathrm{Tr}[\sigma_L]} \quad (43)$$

Using the standard notation from tensor-network theory [129], we can represent Eq. (42) as

$$\sigma_L = \quad \cdots -\!\!\bigcirc\!\!-\!\!\bigcirc\!\!-\!\!\bigcirc\!\!-\cdots \quad \quad \cdots -\!\!\bigcirc\!\!-\!\!\bigcirc\!\!-\!\!\bigcirc\!\!-\cdots \quad (44)$$

Here, each circle corresponds to a tensor $A_{a,b}^{j,k}$. The lower and upper legs are associated with input and output degrees of freedom, respectively, while the remaining ones correspond to the "virtual" indices $a_j = 0, \ldots, \chi - 1$. Note that the contracted legs indicate pairs of indices which are summed over.

We first focus on the purity. We are able to show that MPDOs satisfy the purity AFCs under a few technical assumptions, which encode the fact that all correlation lengths are finite but are otherwise very general.

Technically, we impose a few conditions on the spectrum and eigenstates of the transfer matrices (note that τ_n is not Hermitian and that, for $n = 1$, we recover the standard transfer matrix [135])

$$\tau_n = n \left\{ \begin{array}{c} \vcenter{\hbox{}} \end{array} \right. ,\qquad (45)$$

where the top and bottom legs are contracted. Informally, we ask that τ_1 and τ_2 have nondegenerate eigenvalues with largest absolute value, and that the corresponding eigenstates are not orthogonal. These conditions are general in the sense that one needs to choose fine-tuned examples to violate them. We refer the reader to Appendix D for a precise statement and a detailed discussion.

Under these assumptions, we show that ρ_L satisfies the purity AFCs. Note that, different from the rest of this work, here we are considering periodic boundary conditions. Since Eq. (34) was introduced for open-boundary conditions, we consider directly its global version (40), which can immediately be generalized to the periodic case. Clearly, the expression for r_2 must be modified with respect to Eq. (39): Taking into account periodic boundary conditions, we replace it with

$$r_2 = \frac{\prod_{j=1}^{R} \mathrm{Tr}_{I_j \cup I_{j+1}} (\sigma_{I_j \cup I_{j+1}}^2)}{\prod_{j=1}^{R} \mathrm{Tr}_{I_j} [\sigma_{I_j}^2]}, \qquad (46)$$

where we assume without loss of generality that $R = L/k$ is an integer, while the normalization factors $\mathrm{Tr}[\sigma^2]$ cancel out.

The task of proving Eq. (40) [using the definition (46)] is carried out in Appendix D. There, we derive the following statement: Given the family of MPDOs ρ_L as above, there exist $\zeta > 0$, and $C > 0$ (independent of L) such that, for any integer $k \geq k_{\min}$, with

$$k_{\min} = \max \{1, \zeta \ln(20 C \chi^2 L)\}, \qquad (47)$$

we have

$$\left| \frac{r_2}{\mathcal{P}_2} - 1 \right| \leq \chi^2 (80C + 32) \frac{L}{k} e^{-k/\zeta} \qquad (48)$$

for all $L \geq \max\{\zeta \ln(2^5 \chi^2), 4k + \zeta \ln(2\chi)\}$. The correlation length ζ and the constant C depend on both the eigenvalues and eigenvectors of the transfer matrices τ_1 and τ_2; cf. Appendix D. Therefore, the purity approximate factorization formula (40) holds with $\alpha_2 = \chi^2(20C + 8)$, and $\beta_2 = 1/\zeta$. In turn, using the results of the previous

sections, this implies that the purity of MPDOs can be estimated efficiently using our approach [note that, with these identifications, Eq. (36) implies $k \geq k_{\min}$, where $k_{\min}$ is defined in Eq. (47)].

Similarly, the AFCs for PT moments could be verified analytically for bulk-translation-invariant MPDOs with suitable open-boundary conditions. This is slightly more involved, as a few additional technical assumptions on the boundaries are needed. For this reason, we do not discuss this explicitly. Instead, the AFCs for PT moments will be numerically analyzed in detail in the next section, together with the purity AFCs, for the physically interesting case of Gibbs states of local Hamiltonians in a few concrete examples.

D. Numerical study of AFCs in Gibbs states

As mentioned, it is natural to conjecture that the AFCs are very general, as they encode the fact that all the spatial correlation lengths in the system are finite. In this section, we provide numerical evidence supporting this claim, studying thermal states of two prototypical 1D local Hamiltonians. We focus in particular on the Ising model with transverse and longitudinal fields

$$H_{\mathrm{I}} = -\frac{1}{4} \sum_{j=1}^{L-1} [\sigma_j^z \sigma_{j+1}^z + h_x \sigma_j^x + h_z \sigma_j^z], \qquad (49)$$

and the so-called XXZ Heisenberg spin chain

$$H_{\mathrm{XXZ}} = -\frac{1}{4} \sum_{j=1}^{L-1} [\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \Delta \sigma_j^z \sigma_{j+1}^z]. \qquad (50)$$

We study the corresponding Gibbs states

$$\rho_\beta = \frac{\exp(-\beta H)}{Z_\beta}, \qquad (51)$$

where Z_β is a normalization factor.

Note that the Hamiltonians H_I and H_{XXZ} are integrable for $h_z = 0$ [136] and all values of Δ [137], respectively, although we do not expect that integrability plays any role in the following discussion. Note also that both Hamiltonians feature second-order quantum phase transitions at zero temperature: For H_I the critical line is $h_z = 1$, while H_{XXZ} displays a critical phase for $|\Delta| < 1$. However, we are interested in finite-temperature states, and since in 1D the temperature always introduces a finite length scale, we similarly expect that quantum criticality does not play any role [138].

In order to assess the validity of the AFC for the purity, we test its global version (40). To this end, we numerically compute $r_2(k)$, as defined in Eq. (39), for increasing values of the interval size $k = |I_j|$, and the global purity $\mathcal{P}_2$. Then, we evaluate the corresponding relative error

$$\varepsilon_r(k,L,\beta) = \left| \frac{r_2(k)}{\mathcal{P}_2} - 1 \right|, \tag{52}$$

where we make the dependence on k, β, and L explicit. The calculations are carried out using the iTensor library [140], by first approximating the thermal states by MPOs of finite bond dimension denoted by χ, and subsequently taking powers and traces of such MPOs. For each quantity, this procedure is repeated for increasing χ, until convergence is reached. We refer the reader to Appendix E for further details.

For fixed values of L, we first test the exponential dependence in k of $\varepsilon_r(k,L,\beta)$ predicted by Eq. (40). Examples of our numerical data are reported in Fig. 4, showing a very clear exponential decay. For the Ising model, we choose a nonzero value of the longitudinal field in order to break integrability. Note that, for very large k, we see an apparent plateau. We attribute this behavior to the finite numerical precision of our computations (note the very small values of ε_r corresponding to these plateaus). It is interesting to note that the decay rate for the XXZ chain is faster than the Ising model, although higher bond dimensions are needed in order to accurately approximate the corresponding Gibbs state by an MPO. In any case, we see that quite small values of k are enough to obtain a very good approximation of the purity. We repeat the calculations for different values of the Hamiltonian parameters and β, finding consistent results.

Next, we study the functional form of $\varepsilon_r(k,L,\beta)$. From the numerical point of view, it is not straightforward to test that it is asymptotically bounded by the rhs of Eq. (40). Therefore, we proceed differently. Given an arbitrary fixed $\delta > 0$, we define the minimal region size required to achieve a specified precision δ,

$$k^*(\delta,L) = \min\{k : \varepsilon_r(k,L,\beta) < \delta\}. \tag{53}$$

Assuming that ε_r satisfies an asymptotic bound of the form (40), it is easy to show that $k^*(\delta,L)$ grows at most logarithmically in L; namely, there exist α_2 and β_2 such that

$$k^*(\delta,L) \leq \frac{1}{\beta_2} \ln(\alpha_2 L/\delta). \tag{54}$$

When this condition holds, the estimation protocol explained in Sec. IV B is efficient. Importantly, compared to Eq. (35), Eq. (54) is straightforward to test by our numerical computations.

We have verified that Eq. (54) is satisfied for several values of the temperature and the Hamiltonian parameters. An example of our numerical results is given in Fig. 5. For the Ising model, we see a very clear logarithmic scaling. For the Heisenberg chain, higher bond dimensions are required to achieve the same accuracy. Accordingly, we were not able to simulate sufficiently large system sizes to obtain a clear logarithmic behavior. Still, our results were always consistent with a growth of $k^*(\delta,L)$ which is at most logarithmic in L; cf. Fig. 5. Thus, our numerical data fully support the validity of Eq. (54). In fact, note that the values of $k^*(\delta,L)$ are found to be very small in practice. This is true even when the required bond dimension for MPO calculations is large, as in the case of the XXZ Heisenberg chain. This might make our approach useful already for relatively small system sizes, even if compared against MPO tomographic methods.

Finally, we study the AFC for the PT moments (34). To this end, we define the additive error

$$\varepsilon_a(k,L,\beta) = |\tilde{p}_n[AB] - s_n[A_2 B_1]|, \tag{55}$$

where we make the dependence on k, β, and L explicit. Because the numerical cost of evaluating the trace of the nth power of an MPO increases with n, we restrict our analyses to $n = 3$.
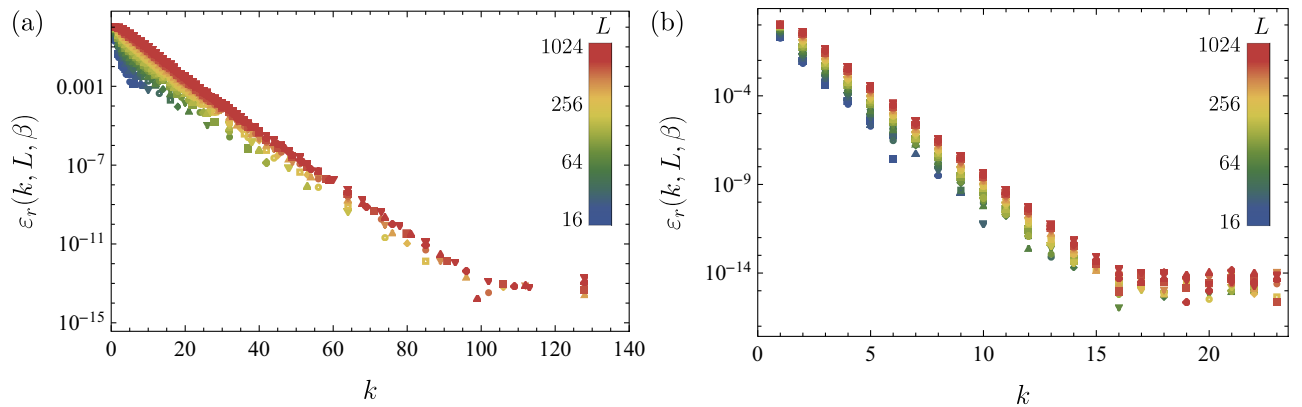

FIG. 4. Relative error $\varepsilon_r(k,L,\beta)$ defined in Eq. (52) as a function of k for $\beta = 2$ and different values of L. In (a) and (b), we report numerical data corresponding to the Ising and XXZ spin chains, respectively. In the former case, we choose $h_x = 1.1$, $h_z = -0.04$, while the maximal bond dimension used in the computation is $\chi = 32$. In the latter case, we set $\Delta = 2$, while $\chi = 64$.
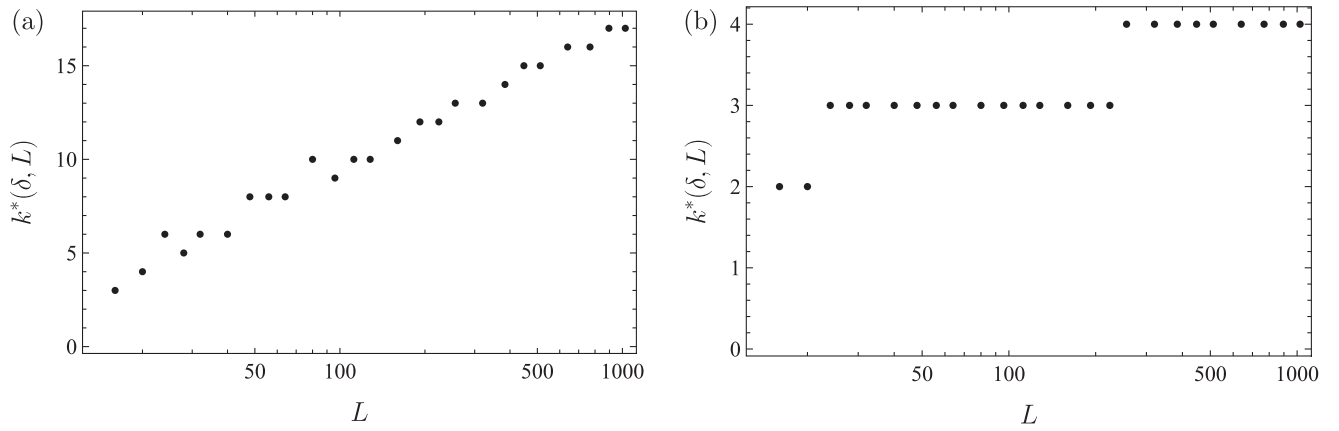
FIG. 5.   Numerical results for the minimal region size $k^*(\delta, L)$ required to achieve a specified precision δ of purity estimation; see Eq. (53). In the plots, we choose $\delta = 0.01$ and $\beta = 2$, while (a) and (b) correspond to the Ising and XXZ spin chains, respectively. The Hamiltonian parameters and bond dimensions are the same as in Fig. 4.

In Fig. 6, we report an example of our data for $\varepsilon_a(k, L, \beta)$ as a function of k, for increasing system sizes. From the plots, two things are apparent. First, for fixed L, the additive error displays a clear exponential decay. Second, its values are independent of L (for values of k much smaller than the system size). This behavior is perfectly consistent with Eq. (35). Note that, as in the previous plots, we observe irregular behavior for sufficiently large k, where ε_a approaches a plateau and displays an L dependence. We interpret these effects as arising due to numerical inaccuracies (note the very small values of ε_a). Similarly, when k approaches $L/2$ we observe L-dependent behavior, which can be attributed to finite-size effects. We have repeated our calculations for different β and Hamiltonian parameters (in each case, we checked that the bond dimension was large enough), always finding consistent results.

E. Practical feasibility of purity estimation

In Sec. IV B, we have shown that, for states satisfying the AFCs, the purity and the PT moments may be estimated by a number of measurements growing only polynomially in the system size. Still, the degrees of the polynomials in our bounds are high; cf., for example, Eq. (33). Accordingly, one may wonder whether our approach could be applied in practice in experiments with current repetition rates. It is important to note, however, that our analysis was mainly concerned with establishing the polynomial dependence rigorously, so that our bounds may likely be improved by more refined mathematical derivations. More importantly, as we have stressed, our protocol may be improved by choosing different ways of estimating the local purities and PT moments, a simple variant being the method of "common randomized measurements" [48]. These improvements are expected to lead to polynomial bounds with lower exponents.

Going beyond general bounds, it is important to illustrate how our estimation protocol performs in practice. This is done in this section, where we exemplify its feasibility by presenting a classical simulation of the full protocol in an explicit case. For simplicity, we focus on random MPSs [60,135] of N qubits and consider the purity of the reduced density matrix over a half system. The choice
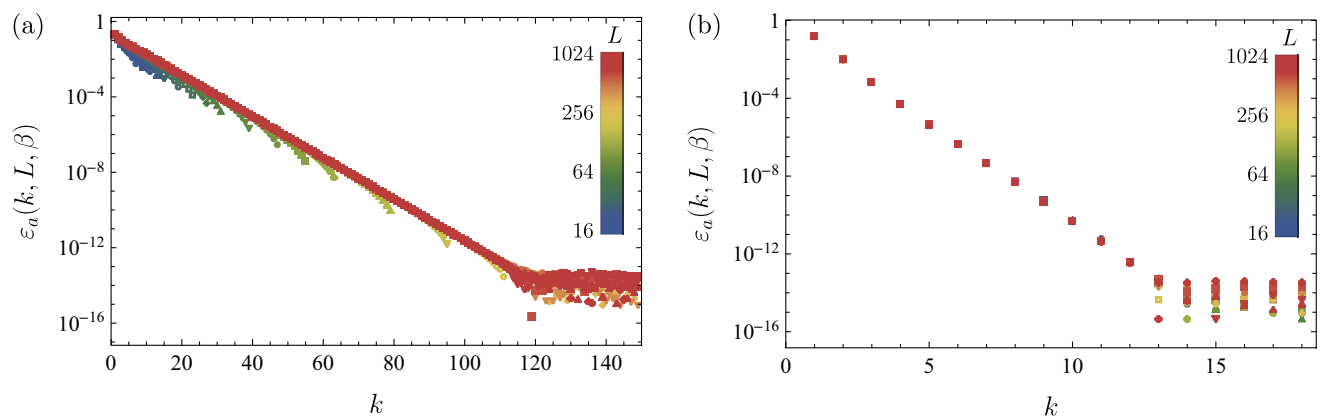

FIG. 6.   Additive error $\varepsilon_a(k, L, \beta)$ defined in Eq. (55) as a function of k, for $\beta = 2$ and different values of L. In (a) and (b) we report numerical data corresponding to the Ising and XXZ spin chains, respectively. The Hamiltonian parameters and bond dimensions are chosen as in Fig. 4.

of pure MPSs is motivated by the fact that it makes it easier to classically simulate the probability distribution of the measurement outcomes. At the same time, random states model typical physical states without allowing for fine-tuning. Our simulations show that the number of measurements needed to reconstruct the purities is orders of magnitude smaller compared to our theoretical bounds, demonstrating that our protocol is experimentally feasible in current quantum platforms.

The details of our simulation are as follows. We initialize the system in a random MPS with complex coefficients and bond dimension χ [140]. Once the random MPS is constructed, we save a copy of it and simulate the RM process (in particular, we imagine that we have experimental access to many identical copies of the state). In our simulation, we consider a slightly more general measurement scheme compared to the one detailed in Sec. II B. In particular, we make use of the RMs described in Ref. [29] where each random unitary u_j is reused for n_M local measurements. Therefore, denoting by n_U the number of distinct unitaries, the total number of experiments is $n_U n_M$ (so that the classical-shadow approach described in Sec. II B is a special case with $n_M = 1$, $n_U = M$). Each simulated experimental run produces a bit string $\mathbf{s}^{(m,k)} = (s_1^{m,k}, \ldots, s_L^{m,k})$, with $m = 1, \ldots, n_U$, $k = 1, \ldots, n_M$. Note that, numerically, we can obtain the correct probability distribution for the measurement outcomes using the perfect-sampling algorithm for MPSs [141]. We have chosen to simulate the RM scheme with reuse of local unitaries, as this scheme is more efficient to simulate, and it has been used in a number of recent experiments [27,30–33].

Finally, we collect the simulated measurement outcomes to construct the purity estimator (17). Instead of Eq. (9), we use the following estimator for the local purities:

$$\mathcal{P}_2^{(e)}[I] = \frac{2^N}{n_U n_M (n_M - 1)} \sum_{m=1}^{n_U} \sum_{\substack{k,k'=1 \\ k \neq k'}}^{n_M} (-2)^{-D[\mathbf{s}_I^{(m,k)}, \mathbf{s}_I^{(m,k')}]}, \quad (56)$$

where $D[\mathbf{s}_I^{(m,k)}, \mathbf{s}_I^{(m,k')}]$ is the Hamming distance between the bit stings $\mathbf{s}_I^{(m,k)}$ and $\mathbf{s}_I^{(m,k')}$ on subsystem I. Compared to Eq. (9), Eq. (56) achieves a more accurate estimate for many repetitions n_M using a few sets of distinct unitaries u_M and is expected to be more robust against miscalibration errors [12]. We also checked numerically that our results were qualitatively unchanged when using estimators based on classical shadows.

Using the above simulation procedure, we studied the estimator $r_2^{(e)}$ defined in Eq. (29) as a function of the number of measurements $n_U n_M$. We computed in particular the statistical error $\varepsilon_s = r_2^{(e)}/r_2$, where r_2 is the rhs of Eq. (15) and the full relative error $\varepsilon = \varepsilon_r$ defined in Eq. (18). In Fig. 7, we report our data for $\chi = 2$, 3, $k = 4$, 6, and system sizes up to $L = 96$ (each plot corresponds to an average error over 20 instances of the random MPS).

From the plots, we see that roughly 10^4 measurements are enough to obtain a relative error below 10% for system sizes up to $L \sim 100$ (note that in this range of the parameters, statistical and full errors are seen to be numerically close, signaling that the deviations from exact AFCs
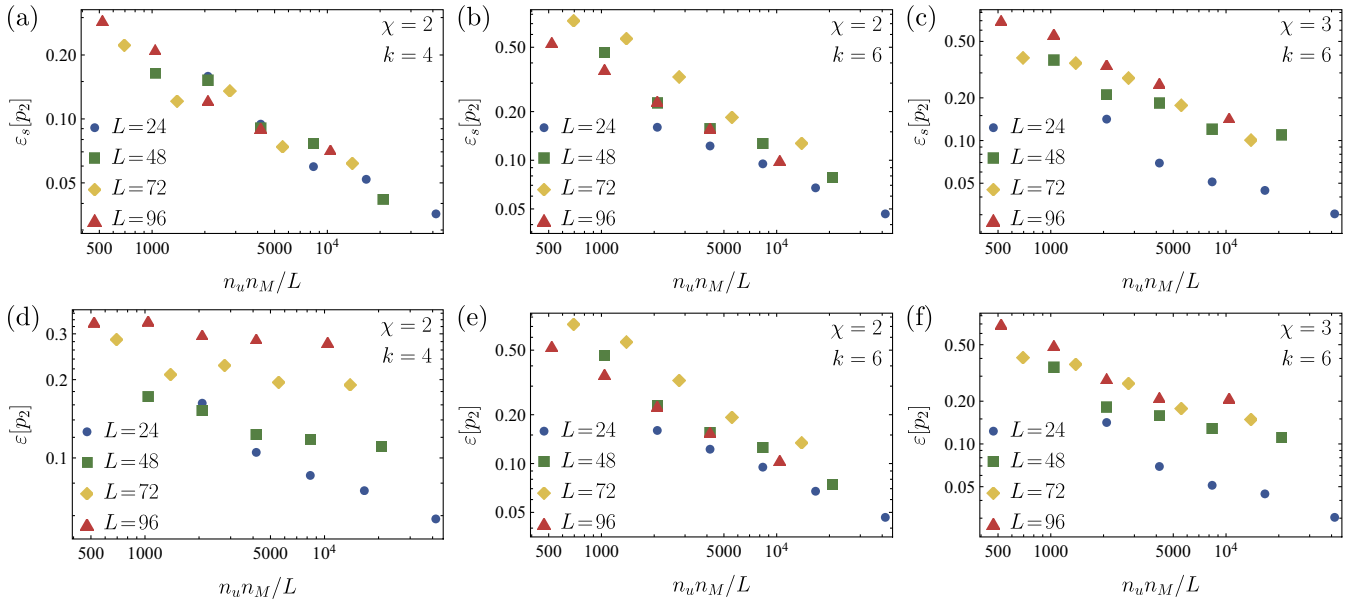

FIG. 7. Statistical and full relative error on the bipartite purity in random MPSs, as a function of the number of experimental runs $n_U n_M$, with fixed $n_M = 1000$. For each plot, we report the bond dimension χ of the random MPS and the length of the intervals $|I_j| = k$ in which the system is partitioned in our estimation protocol. Plots (a), (b), and (c) show the statistical error ε_s, while plots (d), (e), and (f) show the full relative error ε, cf. the main text for the definitions.

are small for $k = 6$, both for $\chi = 2$ and $\chi = 3$). In addition, roughly 10^5 measurements are needed to reach an error around 5% for system sizes up to $L \sim 50$ (note that, for these sizes, a direct application of RM schemes to estimate global purities is unfeasible). As announced, these numbers are orders of magnitude smaller than what was predicted by our general bounds, suggesting that these bounds are far from being tight.

We stress that the number of measurements performed in our simulations are of the order of those performed in current trapped-ion and superconducting qubit experiments; see, e.g., Refs. [27,30–33,39]. The results reported in this section thus substantiate the claim that our protocol is practically useful for global purity and entanglement estimation in today's quantum platforms.

V. MIXED-STATE ENTANGLEMENT DETECTION

In this section, we study the p_n-PPT conditions for states satisfying the AFCs. In Sec. V A, we first introduce a set of quantities, which we call $f_n(\rho)$, displaying the following properties: (i) $f_n(\rho) < 0$ if and only if ρ violates the p_n-PPT conditions, and (ii) $f_n(\rho)$ can be estimated efficiently if ρ satisfies the AFCs, representing natural probes for mixed-state entanglement. Next, in Sec. V B, we compute f_n in a few concrete cases including the Gibbs states of local Hamiltonians and identify the values of the system parameters for which $f_2, f_3 < 0$. The results of this section show that our approach can be successful in detecting entanglement even for large numbers of qubits and highly mixed states, thus being practically useful in many situations routinely encountered in available quantum platforms.

A. The p_n-PPT conditions and mixed-state entanglement probes

Given the density matrix ρ_{AB} on a bipartite system, its logarithmic negativity (3) can be extracted from the knowledge of all PT moments p_n, with $n = 1, \ldots, \dim(\mathcal{H}_A \otimes \mathcal{H}_B)$. Therefore, while our method is efficient to estimate the individual PT moments, reconstructing the logarithmic negativity is still hard, as it requires an exponential number of them.

Luckily, the p_n-PPT conditions introduced in Sec. II A allow us to detect mixed-state entanglement from only a few PT moments. For example, the first two nontrivial elements in the family (5) involve only the PT moments up to $n = 5$. Explicitly, they read

$$p_3 - \text{PPT}, \qquad p_3 p_1 \geq p_2^2; \qquad (57a)$$

$$p_5 - \text{PPT}, \qquad p_3 p_5 \geq p_4^2. \qquad (57b)$$

The conditions (57) can be equivalently rewritten in terms of the quantities

$$f_3(\rho_{AB}) = \tilde{p}_3 \tilde{p}_1 - \tilde{p}_2^2 \frac{\mathcal{P}_2^2[A]\mathcal{P}_2^2[B]}{\mathcal{P}_1[A]\mathcal{P}_1[B]\mathcal{P}_3[A]\mathcal{P}_3[B]}, \qquad (58a)$$

$$f_5(\rho_{AB}) = \tilde{p}_5 \tilde{p}_3 - \tilde{p}_4^2 \frac{\mathcal{P}_4^2[A]\mathcal{P}_4^2[B]}{\mathcal{P}_3[A]\mathcal{P}_3[B]\mathcal{P}_5[A]\mathcal{P}_5[B]}, \qquad (58b)$$

where $\tilde{p}_n$ are the normalized PT moments defined in Eq. (25). Indeed, f_3 and f_5 are negative if and only if the inequalities (57) are violated.

Crucially, $f_3(\rho_{AB})$ and $f_5(\rho_{AB})$ can be estimated efficiently if ρ_{AB} satisfies the AFCs. Indeed, they depend only on $\mathcal{P}_n$ and the normalized PT moments $\tilde{p}_n$, both of which can be estimated using the protocol introduced in Sec. IV; see also Ref. [142]. We note that such estimation protocols allow us to reconstruct $f_n(\rho_{AB})$ up to any arbitrary small additive error. Namely, for any arbitrary positive constant δ, one can construct an estimator $f_n^{(e)}$, with $|f_n^{(e)} - f_n| \leq \delta$ by polynomially many (in L) measurements and postprocessing operations. This is easily seen recalling that $\tilde{p}_n$ ($\mathcal{P}_n$) can be estimated up to any additive (relative) error, and using that

$$|\tilde{p}_n[I]| \leq 1, \qquad \frac{(\mathcal{P}_n[I])^2}{\mathcal{P}_{n-1}[I]\mathcal{P}_{n+1}[I]} \leq 1, \qquad (59)$$

where the first inequality from $|\tilde{p}_n[I]| \leq \tilde{p}_2[I] = \mathcal{P}_2[I]$, while the second follows from Hölder's inequality.

At this point, it is important to ask whether the quantities f_n are practically useful in the many-body context considered in this work. For example, one could be worried that the conditions (57) are never violated for large L (even if the logarithmic negativity is nonzero), or that f_n is exponentially small in L for highly mixed states, thus requiring an exponential accuracy for its estimation. To address this question, consider a family of states $\rho_{AB}^{(L)}$ defined for increasing system sizes L and satisfying the AFCs. Suppose that we extract $f_n(\rho_{AB}^{(L)})$ using our estimation protocol for a fixed approximation error δ_n. Then, the p_n-PPT conditions (57) allow us to successfully detect entanglement if

$$f_3(\rho_{AB}^{(L)}) \leq -C_3, \qquad (60a)$$

$$f_5(\rho_{AB}^{(L)}) \leq -C_5, \qquad (60b)$$

where $C_3, C_5 > 0$ are constants (independent of L) with $C_n \geq 2\delta_n$. Indeed, in these cases the statistical inaccuracy is smaller (with high probability) than the amount by which $f_n(\rho_{AB}^{(L)})$ is negative, allowing us to certify entanglement.

In the next subsection, we study a few natural examples of states satisfying the AFCs showing that Eq. (60) holds either for all values of L or up to very large system sizes.

Therefore, we exhibit concrete examples where f_n are provably useful for entanglement detection.

B. Mixed-state entanglement detection in FDQC and Gibbs states

We start by studying the relations (60) in the simplest case of FDQC states (13). We consider a family of states where the single-qubit density matrices in Eq. (13) are parametrized as

$$\sigma_i(\gamma) = \gamma|a_i\rangle\langle a_i| + (1-\gamma)|b_i\rangle\langle b_i|. \qquad (61)$$

Here, $|a_i\rangle$, $|b_i\rangle$ are arbitrary orthonormal states, while $0 \leq \gamma \leq 1/2$ plays the role of a depolarizing parameter controlling both purity and bipartite entanglement (γ is not to be confused with the symbol previously used for confidence levels). We ask for which values of γ the conditions (57) are violated.

This problem is analyzed in detail in Appendix F. We prove that, for generic choices of the unitary gates, there exist $\gamma_3, \gamma_5 > 0$ such that Eqs. (60a) and (60b) are satisfied, respectively, for all $\gamma \leq \gamma_3/L$ and $\gamma \leq \gamma_5/L^{1/3}$. Now, choosing in particular $\gamma_L = \gamma_5/L^{1/3}$, we have

$$\begin{aligned} \mathrm{Tr}[\rho^{(L)}(\gamma_L)^2] &= [\gamma_L^2 + (1-\gamma_L)^2]^L \\ &\leq \exp[-2(\gamma_L - \gamma_L^2)L] \\ &\leq \exp[-\gamma_L L] = \exp[-\gamma_5 L^{2/3}], \quad (62) \end{aligned}$$

namely,

$$S_2[\rho^{(L)}(\gamma_L)] \geq \gamma_5 L^{2/3}. \qquad (63)$$

Therefore, $\rho^{(L)}(\gamma_L)$ provides an example where the p_5-PPT condition is violated by a finite amount C_5 and, at the same time, the global Rényi-2 entropy grows polynomially, albeit sublinearly, in the system size [143].

As a final interesting example, let us consider once again finite-temperature states of local Hamiltonians. In this case, the purity decays exponentially in L, and we expect that the conditions (58) fail to detect the presence of entanglement in the thermodynamic limit. In the following, we give evidence that they can nevertheless be useful when considering systems of large but finite size.

To support this claim, we study thermal states in the Ising model (49), and numerically compute the coefficients (60) for different temperatures and system sizes. The computations are performed using tensor-network methods, following the same strategy explained in Sec. IV D. An example of our results is reported in Fig. 8. In these plots, we have fixed small but otherwise arbitrary positive numbers C_3 and C_5, and identified the values of β and L for which $f_3(\beta, L) \leq -C_3$ and $f_5(\beta, L) \leq -C_5$.

The data corresponding to the p_3-PPT condition are reported in Fig. 8(a), which clearly shows that the values of the temperature $T = \beta^{-1}$ for which the entanglement can be detected decrease with L, consistent with our previous analysis of FDQC states. On the other hand, Fig. 8(b) shows that $f_5(\beta, L) < -C_5$ up to values of the temperature ($T \simeq 0.1$) that do not significantly depend on L, at least up to very large system sizes ($L \simeq 1024$). We expect that this is a finite-size effect, namely, that increasing L further, the values of T for which entanglement can be detected will
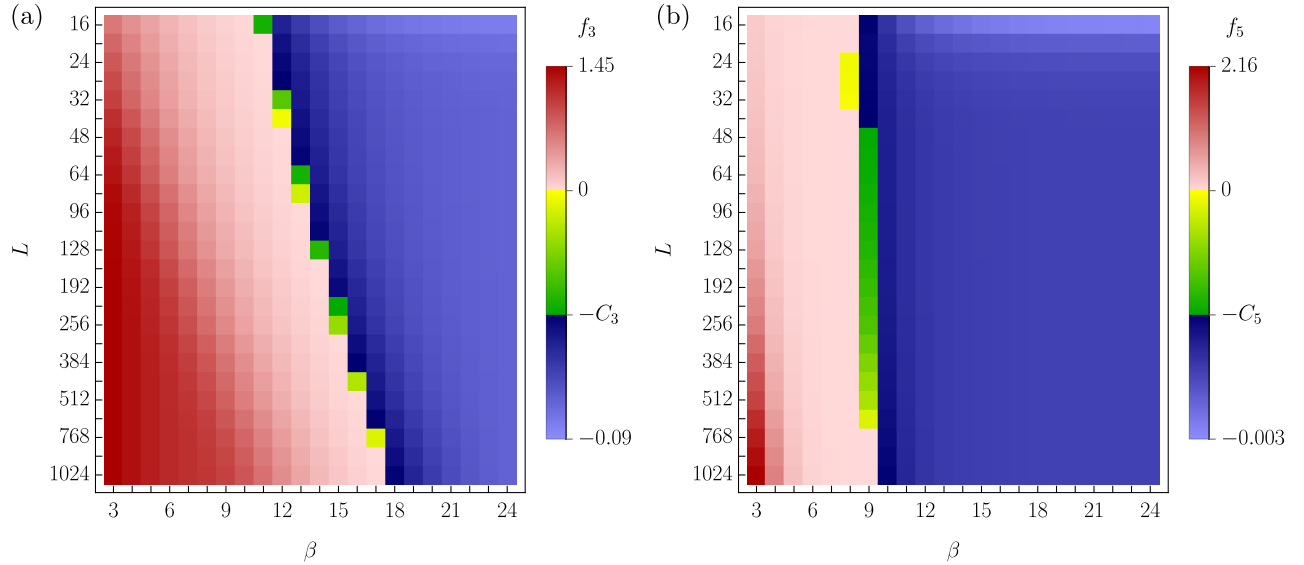

FIG. 8.   Density plots for the coefficients $f_3(\beta, L)$ [plot (a)] and $f_5(\beta, L)$ [plot (b)] as a function of β and L. The data correspond to thermal states in the Ising model (49), with $h_x = 1.1$ and $h_z = -0.04$. Red (blue) regions indicate the values of β and L for which $f_3(\beta, L)$, $f_5(\beta, L)$ are larger (smaller) than some small values $-C_3 = 0.01$ and $-C_5 = 0.001$, respectively.

decrease. Still, this example shows how the p_n-PPT conditions may be able to detect entanglement even for highly mixed states and up to very large system sizes, making the estimation protocol for PT moments practically very useful.

VI. OUTLOOK

We have studied the problem of estimating global entropies and entanglement in many-qubit states, assuming the validity of certain AFCs which encode the fact that all the spatial correlations lengths in the system are finite. We have shown that the AFCs allow one to reconstruct entropies and PT moments from local information which can be extracted efficiently using standard RM schemes. Exploiting this fact, we have devised a very simple strategy for entropy and entanglement estimation which is provably accurate, requiring polynomially many local measurements and postprocessing operations. We have discussed the generality of the AFCs, providing both analytic and numerical evidence for their validity in different classes of states, including MPDOs and thermal states of local Hamiltonians.

As we have argued, the protocols proposed in this work could be practically useful to detect bipartite mixed-state entanglement for large numbers of qubits in today's quantum platforms. This is especially true when considering current NISQ devices from the point of view of quantum simulation, since the AFCs hold under assumptions that are very common in the context of many-body physics. In addition, our work has also raised important questions and opens up a number of interesting directions.

First, our protocols assume that the state to be measured satisfies the AFCs. Therefore, an important question from the experimental perspective is whether it is possible to devise a complementary (and still efficient) strategy to certify their validity, similar to what can be done in the case of MPS [61] or entanglement-Hamiltonian tomography [69].

As we have discussed in Sec. IV A, a simple but effective consistency check can be carried out by repeating our protocols multiple times, each time with an increased ansatz value for the correlation lengths. If the estimated purities or PT moments change beyond the expected precision over different repetitions, we obtain a red flag signaling the failure of the AFCs in the state to be measured. This process, however, does not allow one to rigorously rule out the presence of long-range correlations, as this would require running the estimation protocols using, as an input, correlation lengths which are proportional to the system size. We envision that a way to tackle the certification problem could be to identify a hierarchy of properties which are strictly stronger than the AFCs and for which rigorous certification protocols can be found more easily. At the moment, however, this remains an open problem.

Next, while we have focused on 1D systems, a very natural direction is to extend our approach to higher spatial dimensions. In this case, additional technical complications arise, but we expect that efficient estimation strategies for the purity and PT moments exist under similar conditions encoding the absence of long-range correlations. We stress that this problem is particularly important, since tomographic methods for many-qubit states (based, for instance, on tensor or neural networks) are much less developed in higher dimensions.

Finally, we mention that our approach could also be applied to probe other quantities which generally require exponentially many measurements. Straightforward examples include the participation entropies [85–89] and the so-called stabilizer Rényi entropies [92], but we expect that similar ideas could be implemented in other contexts. One example is that of cross-platform verification protocols [144], which involve probing the (possibly exponentially small) overlaps between different density matrices. Similarly, efficient estimation strategies for the so-called Loschmidt echo [145] also appear to be within the reach of the techniques developed in this work. While these problems require dealing with additional technical subtleties compared to the analysis of the entropy and PT moments, we believe that the latter are not substantial and could be overcome. We leave the study of these applications to future research.

Note added. Recently, a work by Gondolf *et al.* [146] showed that any translation-invariant, finite-range Hamiltonian in 1D at any inverse temperature $\beta > 0$ satisfies the purity AFCs, in complete agreement with our numerical results of Sec. IV D and further substantiating the generality of our approach.
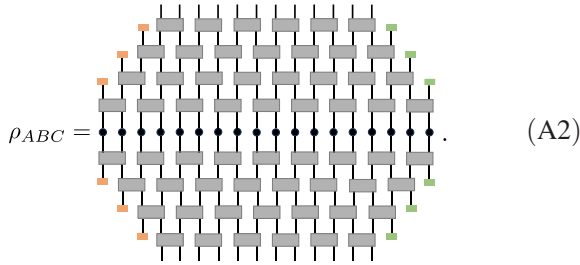
APPENDIX A: FACTORIZATION CONDITION FOR FDQC STATES

1. The purity

In this section, we show that the FDQC states (13) satisfy the factorization condition (14). In fact, we prove the stronger condition

$$
\begin{aligned}
&\mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B(\rho_{ABC}^{\otimes n})] \\
&= \frac{\mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B(\rho_{AB}^{\otimes n})] \otimes \mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B(\rho_{BC}^{\otimes n})]}{\mathrm{Tr}_B(\rho_B^n)},
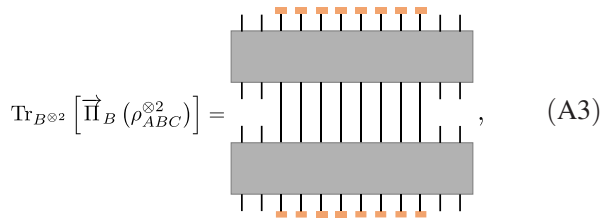\end{aligned} \tag{A1}
$$

where we introduce the n-copy cyclic permutation operator $\vec{\Pi}_B|i_1, \ldots, i_n\rangle = |i_n, i_1, \ldots, i_{n-1}\rangle$, and we assume $|B| \geq (2\ell - 1)$, where ℓ is the circuit depth. We make use of a graphical derivation. We show the case $n = 2$, $\ell = 4$, and $|B| = 8$ for concreteness, but it is clear that the proof generalizes for all values of n and ℓ. The reduced density matrix over the interval $\mathcal{I} = A \cup B \cup C$ reads
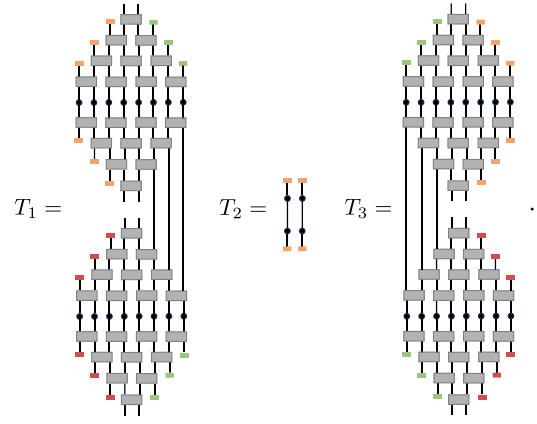

$$\tag{A2}$$

Following the standard tensor-network notation, the lower (upper) dangling legs correspond to the input (output) qubits. In addition, here and in the following, the upper and lower legs in the same column, which are marked with a small rectangle of the same color, are traced over. Finally, black dots correspond to the density matrices σ_j.

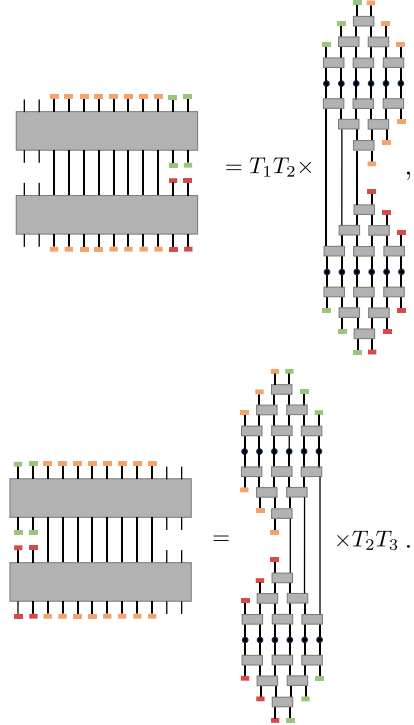The left-hand side of Eq. (A1) can be represented as

$$\tag{A3}$$

where dangling legs are not contracted, while the gray rectangle is a shorthand notation for the reduced density matrix (A2). Using unitarity of the gates, we get $\mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B(\rho_{ABC}^{\otimes n})] = T_1 T_2 T_3$, where
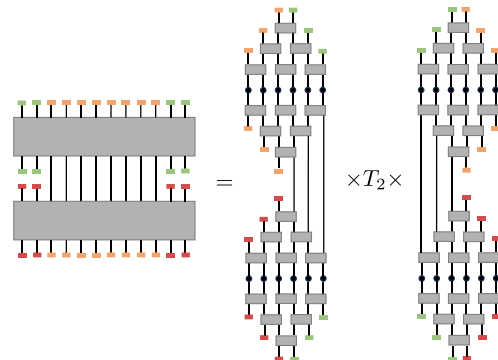

Note that T_1 and T_3 are operators, while T_2 is a number. Let us compare this graphical expression with the one on the rhs of Eq. (A1). The first and second terms in the numerator read, respectively,

Similarly, the denominator in Eq. (A1) reads

Putting everything together and simplifying the common factors in the numerator and denominator, we arrive at Eq. (A1). Finally, using the fact that $\mathrm{Tr}(\rho_X^n) = \mathrm{Tr}(\vec{\Pi}_X \rho_X^{\otimes n})$, and the relation $\vec{\Pi}_{ABC} = \vec{\Pi}_A \otimes \vec{\Pi}_B \otimes \vec{\Pi}_C$, we arrive at

$$\mathrm{Tr}(\rho_{ABC}^n) = \frac{\mathrm{Tr}_{AB}(\rho_{AB}^n)\mathrm{Tr}_{BC}(\rho_{BC}^n)}{\mathrm{Tr}_B(\rho_B^n)}. \tag{A4}$$

For $n = 2$, this proves Eq. (14).

Next, we prove Eq. (15). Let us consider a partition of the whole system $\mathcal{S} = A \cup B \cup C$, where $|B| = |C| = k \geq 2\ell - 1$. We first make use of Eq. (14). Setting now $A^{(0)} := A$, $B^{(0)} := B$, $C^{(0)} := C$ we can iterate this operation, each time splitting the interval $A^{(j)} \cup B^{(j)}$ into three adjacent regions $A^{(j+1)}$, $B^{(j+1)}$, and $C^{(j+1)}$ with $|B^{(j+1)}| = |C^{(j+1)}| = k$. We continue this procedure until $|A^R| = k$, which gives us Eq. (15).

2. The PT moments

The goal of this section is to prove Eq. (26).

We start by recalling the expression for the PT moments [102]

$$\mathrm{Tr}[(\rho_{AB}^{T_A})^n] = \mathrm{Tr}[\vec{\Pi}_A \vec{\Pi}_B (\rho_{AB}^{\otimes n})], \tag{A5}$$

where we introduce the n-copy forward (backward) cyclic permutation operator in the n-replica space defined by

$$\begin{aligned} \vec{\Pi}_A |i_1, \ldots, i_n\rangle &= |i_n, i_1, \ldots, i_{n-1}\rangle, \\ \vec{\Pi}_B |i_1, \ldots, i_n\rangle &= |i_2, i_3, \ldots, i_1\rangle. \end{aligned} \tag{A6}$$

Using this notation, it is easy to see that Eq. (A1) implies

$$\begin{aligned} &\mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B (\rho_{ABC}^{\otimes n})] \\ &= \frac{\mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B (\rho_{AB}^{\otimes n})] \otimes \mathrm{Tr}_{B^{\otimes n}}[\vec{\Pi}_B (\rho_{BC}^{\otimes n})]}{\mathrm{Tr}_B(\rho_B^n)}. \end{aligned} \tag{A7}$$

Equation (A7) can be proved by taking the partial transpose with respect to $A^{\otimes n}$ and $C^{\otimes n}$ on both sides of Eq. (A1), followed by complex conjugation. Next, let us take a bipartition of the system into A and B, each divided into two intervals $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$ with $|A_2| = |B_1| = k$; cf. Fig. 3(b). Making use of Eq. (A1), we have

$$\begin{aligned} \mathrm{Tr}[(\rho_{AB}^{T_A})^n] &= \mathrm{Tr}_{A^{\otimes n}} \mathrm{Tr}_{B_2^{\otimes n}} \vec{\Pi}_A \vec{\Pi}_{B_2} \mathrm{Tr}_{B_1^{\otimes n}} \vec{\Pi}_{B_1} \rho^{\otimes n} \\ &= \frac{\mathrm{Tr}_{A^{\otimes n} \otimes B_1^{\otimes n}}(\vec{\Pi}_A \vec{\Pi}_{B_1} \rho_{AB_1}^{\otimes n})}{\mathrm{Tr}_{B_1}(\rho_{B_1}^n)} \\ &\quad \times \frac{\mathrm{Tr}_{B_1^{\otimes n} \otimes B_2^{\otimes n}}(\vec{\Pi}_{B_1} \vec{\Pi}_{B_2} \rho_{B_1 B_2}^{\otimes n})}{\mathrm{Tr}_{B_1}(\rho_{B_1}^n)}. \end{aligned} \tag{A8}$$

Using

$$\begin{aligned} &\mathrm{Tr}_{A^{\otimes n} \otimes B_1^{\otimes n}}(\vec{\Pi}_A \vec{\Pi}_{B_1} \rho_{AB_1}^{\otimes n}) \\ &= \mathrm{Tr}_{A_1^{\otimes n}} \mathrm{Tr}_{B_1^{\otimes n}} \vec{\Pi}_{A_1} \vec{\Pi}_{B_1} \mathrm{Tr}_{A_2^{\otimes n}} \vec{\Pi}_{A_2} \rho_{A_1 A_2 B_1}^{\otimes n} \\ &= \mathrm{Tr}_{A_1^{\otimes n} \otimes A_2^{\otimes n}}(\vec{\Pi}_{A_1} \vec{\Pi}_{A_2} \rho_{A_1 A_2}^{\otimes n}) \\ &\quad \times \mathrm{Tr}_{A_2^{\otimes n} \otimes B_1^{\otimes n}}(\vec{\Pi}_{A_2} \vec{\Pi}_{B_1} \rho_{A_2 B_1}^{\otimes n})[\mathrm{Tr}_{A_2}(\rho_{A_2}^n)]^{-1}, \end{aligned} \tag{A9}$$

we finally arrive at Eq. (26).

APPENDIX B: STATISTICAL-ERROR ANALYSIS IN FDQC STATES

1. Statistical-error analysis for the purity

In this section we prove Eq. (20). We consider the protocol explained in Sec. III A and denote by $\mathcal{P}_2^{(e)}[I]$ the estimates for $\mathcal{P}_2[I] = \mathrm{Tr}(\rho_I^2)$ obtained from the classical-shadow approach. These local purities are estimated with a nonzero relative error, i.e., $\mathcal{P}_2^{(e)}[I]/\mathcal{P}_2[I] = (1 + \varepsilon^I)$. We define

$$\varepsilon = \max\{|\varepsilon^I| : I \in \{I_j\}_j \cup \{I_j \cup I_{j+1}\}_j\}. \tag{B1}$$

We take $r_2^{(e)}$ defined in Eq. (17) as the experimental estimate for the global purity $\mathcal{P}_2 = \mathrm{Tr}(\rho^2)$, choosing the intervals I_j with $|I_j| = k \geq 2\ell - 1$, where ℓ is the circuit depth, so that Eq. (15) holds.

We start our proof from a few preliminary lemmas.

Lemma 1. Suppose $\varepsilon \leq 1/2$. Then,

$$\left| \frac{r_2^{(e)}}{\mathcal{P}_2} - 1 \right| \leq e^{(4L/k)\varepsilon} - 1. \tag{B2}$$

Proof. Setting $\mathcal{P}_2^{(e)}[I] = \mathcal{P}_2[I](1 + \varepsilon^I)$, recalling $R = L/k$, where $k = |I_j|$ and using Eq. (15), we have

$$\frac{r_2^{(e)}}{\mathcal{P}_2} = \frac{\prod_{j=1}^{R-1}(1 + \varepsilon^{I_j \cup I_{j+1}})}{\prod_{j=2}^{R-1}(1 + \varepsilon^{I_j})}, \tag{B3}$$

as so, using $|\varepsilon^I| \leq \varepsilon$,

$$\begin{aligned} \frac{r_2^{(e)}}{\mathcal{P}_2} &\leq \frac{(1+\varepsilon)^{R-1}}{(1-\varepsilon)^{R-2}} \leq \left[\frac{1+\varepsilon}{1-\varepsilon}\right]^R, \\ \frac{r_2^{(e)}}{\mathcal{P}_2} &\geq \frac{(1-\varepsilon)^{R-1}}{(1+\varepsilon)^{R-2}} \geq \left[\frac{1-\varepsilon}{1+\varepsilon}\right]^R. \end{aligned} \tag{B4}$$

Since $\varepsilon \leq 1/2$, and using $(1 + x) \leq e^x$, we have

$$\frac{1+\varepsilon}{1-\varepsilon} = 1 + \frac{2\varepsilon}{1-\varepsilon} \leq 1 + 4\varepsilon \leq e^{4\varepsilon}. \tag{B5}$$

Therefore,

$$e^{-(4L/k)\varepsilon} \leq \frac{r_2^{(e)}}{\mathcal{P}_2} \leq e^{(4L/k)\varepsilon}. \qquad (B6)$$

Finally, using

$$(1 - e^{-a}) \leq e^a - 1, \qquad (B7)$$

we obtain Eq. (B2). ∎

Lemma 2. Let $M \geq 2^{4|I|}$. Then,

$$\Pr\left(|\varepsilon^I| \geq \alpha\right) \leq \frac{2^{2|I|+3}}{\alpha^2 M}. \qquad (B8)$$

Proof. We start from the bound (10). Since $M \geq 2^{4|I|}$, we have in particular $M \geq 4$, and Eq. (10) implies

$$\mathrm{Var}[\mathcal{P}_2^{(e)}[I]] \leq 4\left(\frac{2^{|I|}\mathcal{P}_2[I]}{M}\right) + 4\left(\frac{2^{2|I|}}{M}\right)^2. \qquad (B9)$$

Since by hypothesis $M \geq 2^{4|I|}$, while $\mathcal{P}_2[I] \geq 2^{-|I|}$, we have

$$4\left(\frac{2^{2|I|}}{M}\right)^2 \leq 4\left(\frac{2^{|I|}\mathcal{P}_2[I]}{M}\right), \qquad (B10)$$

and so

$$\mathrm{Var}[\mathcal{P}_2^{(e)}[I]] \leq 8\frac{2^{|I|}\mathcal{P}_2[I]}{M}. \qquad (B11)$$

Equation (B8) then follows using Chebyshev's inequality and $\mathcal{P}_2[I] \geq 2^{-|I|}$. ∎

We are ready to state and prove the main result of this section.

Theorem 1. Let $0 \leq \delta \leq 1$ and take

$$M \geq \max\left\{2^{8k}, L^2\frac{2^{4k+10}}{k^2\delta^2}\right\}, \qquad (B12)$$

where $k = |I_j| \geq 2$ (the case $k = 1$ is trivial). Then,

$$\Pr\left[|(r_2^{(e)}/\mathcal{P}_2) - 1| \geq \delta\right] \leq \frac{2^{4k+11}L^3}{\delta^2 k^3 M}. \qquad (B13)$$

Proof. Recalling the definition (B1), note that if $\varepsilon < (k/8L)\delta$, then $\varepsilon < 1/2$ and also $e^{(4L/k)\varepsilon} - 1 < \delta$. Using Lemma 1, this implies

$$\Pr[|(r_2^{(e)}/\mathcal{P}_2) - 1| \geq \delta] \leq \Pr[\varepsilon \geq (k/8L)\delta]$$
$$= 1 - \Pr[\varepsilon < (k/8L)\delta]. \qquad (B14)$$

In more detail, the validity of the first inequality can be seen as follows: Suppose $\varepsilon < (k/8L)\delta$. Then necessarily

$\varepsilon < 1/2$ and so also $|r_2^{(e)}/\mathcal{P}_2 - 1| < \delta$ (by Lemma 1). Therefore, the set of cases in which $|r_2^{(e)}/\mathcal{P}_2 - 1| \geq \delta$ must be contained in the set of cases in which $\varepsilon \geq (k/8L)\delta$. In a formula, this is the first line of Eq. (B14).

From the definition of ε and Lemma 2, we have

$$\Pr[\varepsilon < x] = \prod_j (1 - \Pr[|\varepsilon^{I_j}| \geq x])$$
$$\times \prod_j (1 - \Pr[|\varepsilon^{I_j \cup I_{j+1}}| \geq x])$$
$$\geq \left[\left(1 - \frac{2^{2k+3}}{x^2 M}\right)\left(1 - \frac{2^{4k+3}}{x^2 M}\right)\right]^R, \qquad (B15)$$

where we used that the random variables ε^{I_j} are statistically independent. Therefore,

$$\Pr[\varepsilon < x] \geq \left(1 - \frac{2^{4k+3}}{x^2 M}\right)^{2R}. \qquad (B16)$$

Setting $x = \delta k/(8L)$, we have by hypothesis $2^{4k+3}/x^2 M \leq 1/2$. Therefore, using $1 - z \geq e^{-2z}$ for $0 \leq z \leq 1/2$, we obtain

$$\Pr[\varepsilon < (k/8L)\delta] \geq \exp\left[-\frac{2^{4k+11}L^3}{M\delta^2 k^3}\right]. \qquad (B17)$$

Plugging this into Eq. (B14), and using $1 - e^{-z} \leq z$, we finally obtain Eq. (B13). ∎

2. Statistical-error analysis for the PT moments

In this section we prove Eq. (32). Considering the same protocol and using the same notation as Sec. III B, we set

$$p_3^{(e)}[A_2 B_1] = p_3[A_2 B_1] + \varepsilon^{A_2 B1}, \qquad (B18a)$$

$$\mathcal{P}_3^{(e)}[A_2] = \mathcal{P}_3[A_2](1 + \varepsilon^{A_2}), \qquad (B18b)$$

$$\mathcal{P}_3^{(e)}[B_1] = \mathcal{P}_3[B_1](1 + \varepsilon^{B_1}). \qquad (B18c)$$

Note that $\varepsilon^{A_2 B_1}$ is an additive error, while ε^{A_2}, ε^{B_1} are relative errors. We also define

$$\varepsilon = 2^{2|A_2|+2|B_1|} \max\{|\varepsilon^{A_2 B_1}|, |\varepsilon^{A_2}|, |\varepsilon^{B_1}|\}. \qquad (B19)$$

In the following, we will set $|A_2| = |B_1| = k$, where $k \geq 2\ell - 1$ with ℓ being the circuit depth. With this choice, Eq. (27) holds.

For clarity, we organize the proof into lemmas.

Lemma 3. Suppose $\varepsilon \leq 1$. Then,

$$|s_3^{(e)}[A_2 B_1] - s_3[A_2 B_1]| \leq 16\varepsilon, \qquad (B20)$$

where $s_n[A_2 B_1]$ is defined in Eq. (27).

Proof. We start from

$$s_3^{(e)}[A_2B_1] = \frac{p_3[A_2B_1] + \varepsilon^{A_2B_1}}{\mathcal{P}_3[A_2](1+\varepsilon^{A_2})\mathcal{P}_3[B_1](1+\varepsilon^{B_1})}, \quad \text{(B21)}$$

which gives us

$$\begin{aligned}
|s_3^{(e)}&[A_2B_1] - s_3[A_2B_1]| \\
&\leq \frac{|p_3[A_2B_1]|}{\mathcal{P}_3[A_2]\mathcal{P}_3[B_1]} \frac{|\varepsilon^{A_2}| + |\varepsilon^{B_1}| + |\varepsilon^{A_2}\varepsilon^{B_1}|}{(1+\varepsilon^{A_2})(1+\varepsilon^{B_1})} \\
&\quad + \frac{|\varepsilon^{A_2B_1}|}{\mathcal{P}_3[A_2]\mathcal{P}_3[B_1](1+\varepsilon^{A_2})(1+\varepsilon^{B_1})}.
\end{aligned} \quad \text{(B22)}$$

Denoting by λ_j the eigenvalues of $\rho_{A_2B1}^{T_{A_2}}$, we have

$$|p_3[A_2B_1]| \leq \sum_j |\lambda_j|^3 \leq \sum_j \lambda_j^2 = \mathcal{P}[A_2B_1] \leq 1, \quad \text{(B23)}$$

where we used $\lambda_j \in [-1/2, 1]$ [147] and $\text{Tr}[(\rho_{AB}^{T_A})^2] = \text{Tr}[\rho_{AB}^2]$ [35]. In addition, since $\varepsilon \leq 1$, clearly $|\varepsilon^{A_2}| \leq 1/2$, $|\varepsilon^{B_1}| \leq 1/2$ and so also $|\varepsilon^{A_2}\varepsilon^{B_1}| < |\varepsilon^{A_2}|$. Putting everything together, and using $\mathcal{P}_3[I] \leq 2^{-2|I|}$, we finally arrive at Eq. (B20). ∎

Lemma 4. Let $M \geq 3 \times 2^{8k}$. Then,

$$\Pr\left(|\varepsilon^{A_2}| \geq \alpha\right) \leq 27 \frac{2^{3k}}{\alpha^2 M}, \quad \text{(B24a)}$$

$$\Pr\left(|\varepsilon^{B_1}| \geq \alpha\right) \leq 27 \frac{2^{3k}}{\alpha^2 M}, \quad \text{(B24b)}$$

$$\Pr\left(|\varepsilon^{A_2B_1}| \geq \alpha\right) \leq 27 \frac{2^{2k}}{\alpha^2 M}. \quad \text{(B24c)}$$

Proof. We start from the bound [38]

$$\begin{aligned}
\text{Var}[p_3^{(e)}[AB]] &\leq 9 \frac{2^{|AB|}}{M} \text{Tr}(\rho_{AB}^4) + 18 \frac{2^{3|AB|}}{(M-1)^2} p_2[AB] \\
&\quad + 6 \frac{2^{6|AB|}}{(M-2)^3}.
\end{aligned} \quad \text{(B25)}$$

Since $M \geq 3 \times 2^{8k}$ (and $k \geq 1$), then

$$\frac{18}{(M-1)^2} \leq \frac{27}{M^2}, \qquad \frac{6}{(M-2)^3} \leq \frac{9}{M^3}. \quad \text{(B26)}$$

Therefore,

$$\begin{aligned}
\text{Var}[p_3^{(e)}[AB]] &\leq 9 \frac{2^{|AB|}}{M} \text{Tr}(\rho_{AB}^4) + 27 \frac{2^{3|AB|}}{M^2} p_2[AB] \\
&\quad + 9 \frac{2^{6|AB|}}{M^3}.
\end{aligned} \quad \text{(B27)}$$

Since $M \geq 3 \times 2^{8k}$, we have

$$\begin{aligned}
9 \frac{2^{|AB|}}{M} \text{Tr}(\rho_{AB}^4) &\geq 27 \frac{2^{3|AB|}}{M^2} p_2[AB], \\
9 \frac{2^{|AB|}}{M} \text{Tr}(\rho_{AB}^4) &\geq 9 \frac{2^{6|AB|}}{M^3}.
\end{aligned} \quad \text{(B28)}$$

In the first line, we have used that $p_2[I] = \text{Tr}[\rho_I^2]$ and Hölder's inequality, which guarantees

$$\text{Tr}[\rho_I^2]/\text{Tr}[\rho_I^4] \leq 2^{|I|}/\text{Tr}[\rho_I^2] \leq 2^{2|I|}, \quad \text{(B29)}$$

while in the second line, we have used $\text{Tr}(\rho_{AB}^4) \geq 2^{-3|AB|} = 2^{-6k}$. Putting it all together, we get

$$\text{Var}[p_3^{(e)}[AB]] \leq 27 \frac{2^{2k}}{M} \text{Tr}(\rho_{AB}^4). \quad \text{(B30)}$$

Similarly, we have [38]

$$\begin{aligned}
\text{Var}[\mathcal{P}_3[A_2]] &\leq 9 \frac{2^{|A_2|}}{M} \text{Tr}(\rho_{A_2}^4) + 18 \frac{2^{3|A_2|}}{(M-1)^2} \text{Tr}(\rho_{A_2}^2) \\
&\quad + 6 \frac{2^{6|A_2|}}{(M-2)^3} \\
&\leq 27 \frac{2^k}{M} \text{Tr}(\rho_{A_2}^4).
\end{aligned} \quad \text{(B31)}$$

Using $\text{Tr}(\rho_{A_2}^4)/[\text{Tr}(\rho_{A_2}^3)]^2 \leq 1/\text{Tr}(\rho_{A_2}^3) \leq 2^{2k}$, we obtain

$$\text{Var}\left[\frac{\mathcal{P}_3^{(e)}[A_2]}{\mathcal{P}_3[A_2]}\right] \leq 27 \frac{2^k}{M} \frac{\text{Tr}(\rho_{A_2}^4)}{[\text{Tr}(\rho_{A_2}^3)]^2} \leq 27 \frac{2^{3k}}{M}, \quad \text{(B32)}$$

and analogously for $\mathcal{P}_3^{(2)}(B_1)$. Equations (B24) follow using Chebyshev's inequality. ∎

We are ready to state and prove the main result of this section, yielding Eq. (32).

Theorem 2. Let $0 \leq \delta \leq 1$ and take

$$M \geq 27 \frac{2^{11k+9}}{\delta^2}. \quad \text{(B33)}$$

Then,

$$\Pr\left[|s_3^{(e)}[A_2B_1] - s_3[A_2B_1]| \geq \delta\right] \leq 81 \frac{2^{11k+9}}{M\delta^2}. \quad \text{(B34)}$$

Proof. Recall the definition (B19) and note that if $\varepsilon \leq \delta/16$, then trivially $\varepsilon < 1$ and so, by Lemma 3, $|s_3^{(e)} - s_3| \leq \delta$. This implies

$$\begin{aligned}
\Pr[|s_3^{(e)} - s_3| \geq \delta] &\leq \Pr[\varepsilon \geq \delta/16] \\
&= 1 - \Pr[\varepsilon < \delta/16].
\end{aligned} \quad \text{(B35)}$$

In more detail, the validity of the first inequality can be seen as follows: Suppose $\varepsilon < \delta/16$. Then trivially $\varepsilon < 1$ and so necessarily also $|s_3^{(e)} - s_3| < \delta$ (by Lemma 3). Therefore, the set of cases in which $|s_3^{(e)} - s_3| \geq \delta$ is contained in the set of cases for which $\varepsilon \geq \delta/16$. In a formula, this is the first line of Eq. (B35).

From the definition of ε and Lemma 4, we have

$$
\begin{aligned}
\Pr[\varepsilon < x] &= (1 - \Pr[|\varepsilon^{A_2}| \geq x2^{-4k}]) \\
&\quad \times (1 - \Pr[|\varepsilon^{B_1}| \geq x2^{-4k}])(1 - \Pr[|\varepsilon^{A_2 B_1}| \geq x2^{-4k}]) \\
&\geq \left(1 - 27\frac{2^{11k}}{x^2 M}\right)^3,
\end{aligned} \tag{B36}
$$

where we used that the random variables ε^{I_j} are statistically independent. Setting $x = \delta/16$, we have by hypothesis $27 \times 2^{11k}/x^2 M < 1/2$. Therefore, using $1 - z \geq e^{-2z}$ for $0 \leq z \leq 1/2$, we obtain

$$
\Pr[\varepsilon < \delta/16] \geq \exp\left[-81\frac{2^{11k+9}}{M\delta^2}\right]. \tag{B37}
$$

Plugging into Eq. (B35), and using $1 - e^{-z} \leq z$, we finally obtain Eq. (B34). ∎

APPENDIX C: STATISTICAL-ERROR ANALYSIS FOR STATES SATISFYING THE AFCs

In this appendix, we prove the main claims reported in Sec. IV B. Let ρ be a state satisfying Eq. (34) for all partitions as in Fig. 3(a), with $|B| = k \geq k_c$ (and open-boundary conditions). We first prove Eq. (37).

We start with the following:

Lemma 5. Let

$$
k \geq \frac{\ln(\alpha_2 L/\delta)}{\beta_2}, \tag{C1}
$$

for $0 \leq \delta \leq 1/2$. Then, defining r_2 as in Eq. (39), we have

$$
\left|\frac{r_2}{\text{Tr}(\rho^2)} - 1\right| \leq 4\delta. \tag{C2}
$$

Proof. First, applying iteratively Eq. (34), we get (recalling that $R = L/k$)

$$
\frac{r_2}{\text{Tr}(\rho^2)} = \prod_{j=2}^{R-1}(1 + \varepsilon_j), \tag{C3}
$$

with $|\varepsilon_j| \leq \alpha_2 e^{-\beta_2 k}$, and so

$$
\left|\frac{r_2}{\text{Tr}(\rho^2)} - 1\right| = \left|\prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1\right|. \tag{C4}
$$

We have

$$
\prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1 \leq (1 + \alpha_2 e^{-\beta_2 k})^{R-2} - 1
$$
$$
\leq \exp\left[\alpha_2(L/k)e^{-\beta_2 k}\right] - 1. \tag{C5}
$$

Analogously,

$$
\prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1 \geq (1 - \alpha_2 e^{-\beta_2 k})^{R-2} - 1. \tag{C6}
$$

Because of Eq. (C1), we have $\alpha_2 e^{-\beta_2 k} \leq 1/2$. Therefore, using $1 - z \geq e^{-2z}$ for $0 \leq z \leq 1/2$, we have

$$
\prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1 \geq \exp\left[-2\alpha_2(L/k)e^{-\beta_2 k}\right] - 1. \tag{C7}
$$

Therefore,

$$
e^{-y} - 1 \leq \prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1 \leq e^{y/2} - 1 \leq e^y - 1, \tag{C8}
$$

where $y = 2\alpha_2(L/k)e^{-\beta_2 k}$. Since $1 - e^{-y} \leq e^y - 1$, this implies

$$
\left|\prod_{j=2}^{R-1}(1 + \varepsilon_j) - 1\right| \leq \exp\left[2\alpha_2(L/k)e^{-\beta_2 k}\right] - 1. \tag{C9}
$$

Finally, due to Eq. (C1), we have $2\alpha_2(L/k)e^{-\beta_2 k} \leq 1$. Therefore, using $e^z - 1 \leq 2z$ for $0 \leq z \leq 1$, we arrive at

$$
\left|\frac{r_2}{\text{Tr}(\rho^2)} - 1\right| \leq \frac{4\alpha_2 L}{k}e^{-\beta_2 k} \leq 4\delta. \tag{C10}
$$

∎

Next, we prove Eq. (37), via the following:

Theorem 3. Let $0 \leq \delta \leq 1/2$ and set

$$
k = \frac{\ln(\alpha_2 L/\delta)}{\beta_2}. \tag{C11}
$$

Choosing

$$
M \geq \max\left\{\left(\frac{\alpha_2 L}{\delta}\right)^{\frac{8\ln 2}{\beta_2}}, \frac{L^2 2^{10}}{\delta^2}\left(\frac{\alpha_2 L}{\delta}\right)^{\frac{4\ln 2}{\beta_2}}\right\}, \tag{C12}
$$

and recalling the definition (17), we have

$$
\Pr\left[|(r_2^{(e)}/\mathcal{P}_2) - 1| \geq 7\delta\right] \leq \frac{2^{11} L^3}{\delta^2 M}\left(\frac{\alpha_2 L}{\delta}\right)^{\frac{4\ln 2}{\beta_2}}. \tag{C13}
$$

Proof. Set $r_2^{(e)}/r_2 = (1 + \varepsilon_1)$ and $r_2/\mathcal{P}_2 = (1 + \varepsilon_2)$, where $r_2^{(e)}$ is defined in Eq. (17). Using Lemma 5 (and that $\delta \le 1/2$), we have $|\varepsilon_2| \le 4\delta$ and so

$$\left| \frac{r_2^{(e)}}{\mathcal{P}_2} - 1 \right| = \left| \frac{r_2^{(e)}}{r_2} \frac{r_2}{\mathcal{P}_2} - 1 \right| \le 4\delta + 3|\varepsilon_1|. \quad \text{(C14)}$$

If $|\varepsilon_1| < \delta$, then $|r_2^{(e)}/\mathcal{P}_2 - 1| < 7\delta$. Therefore, the set of cases in which $|r_2^{(e)}/\mathcal{P}_2 - 1| \ge 7\delta$ is contained in the set of cases in which $|\varepsilon_1| \ge \delta$. In a formula,

$$\Pr[|(r_2^{(e)}/\mathcal{P}_2) - 1| \ge 7\delta] \le \Pr[|(r_2^{(e)}/r_2) - 1| \ge \delta|]. \quad \text{(C15)}$$

Thanks to Eq. (C12), we can use Theorem 1, yielding

$$\Pr[|(r_2^{(e)}/r_2) - 1| \ge \delta|] \le \frac{2^{4k+11}L^3}{\delta^2 k^3 M} \le \frac{2^{4k+11}L^3}{\delta^2 M}$$
$$\le \frac{2^{11}L^3}{\delta^2 M} \left(\frac{\alpha_2 L}{\delta} \right)^{\frac{4\ln 2}{\beta_2}}, \quad \text{(C16)}$$

which completes the proof. ∎

Note that, in order to recover Eqs. (36) and (37), we simply rename $\delta' = 7\delta$.

Finally, we present a technical result showing that the PT moments can be estimated efficiently, assuming that the state to be measured satisfies the AFCs. We focus for simplicity on the case $n = 3$ and follow a protocol similar to that of Sec. III B. Namely, we take $s_3^{(e)}[A_2 B_1]$ defined in Eq. (29) as our estimator for $\tilde{p}_3[AB]$, and, for any $0 \le \delta \le 1$, we choose

$$k = \frac{\ln(2\alpha_3/\delta)}{2\beta_3}. \quad \text{(C17)}$$

Finally, we perform M measurements to estimate $p_3[A_2 B_1]$, $\mathcal{P}_3[A_2]$, and $\mathcal{P}_3[B_1]$ each (so the total number is $M_T = 3M$). We can then prove the following:

Theorem 4. For any $0 \le \delta \le 1/2$, set

$$k = \frac{\ln(2\alpha_3/\delta)}{2\beta_3}. \quad \text{(C18)}$$

If

$$M \ge 27 \frac{2^9}{\delta^2} \left(\frac{2\alpha_3}{\delta} \right)^{\frac{11\ln 2}{2\beta_3}}, \quad \text{(C19)}$$

then

$$\Pr[|s_3^{(e)} - \tilde{p}_3| \ge \delta] \le \frac{81 \times 2^9}{\delta^2 M} \left(\frac{2\alpha_3}{\delta} \right)^{\frac{11\ln 2}{2\beta_3}}. \quad \text{(C20)}$$

Proof. Our estimator for the normalized PT moment $\tilde{p}_3[AB]$ is

$$s_3^{(e)} = \frac{p_3^{(e)}[A_2 B_1]}{\mathcal{P}_3^{(e)}[A_2] \mathcal{P}_3^{(e)}[B_1]}. \quad \text{(C21)}$$

First, suppose $|s_3 - s_3^{(e)}| < \delta/2$. Then, using Eqs. (35) and (C18), we have

$$|s_3^{(e)} - \tilde{p}_3| \le |s_3^{(e)} - s_3| + |s_3 - \tilde{p}_3|$$
$$\le (\delta/2) + (\delta/2) = \delta. \quad \text{(C22)}$$

Therefore, the set of cases in which $|s_3^{(e)} - \tilde{p}_3| \ge \delta$ is contained in the set of cases in which $|s_3^{(e)} - s_3| \ge \delta/2$, namely,

$$\Pr[|s_3^{(e)} - \tilde{p}_3| \ge \delta] \le \Pr[|s_3^{(e)} - s_3| \ge \delta]. \quad \text{(C23)}$$

Finally, thanks to Eq. (C19), we can apply Theorem 2, yielding

$$\Pr[|s_3^{(e)} - \tilde{p}_3| \ge \delta] \le \frac{81 \times 2^9}{\delta^2 M} \left(\frac{2\alpha_3}{\delta} \right)^{\frac{11\ln 2}{2\beta_3}}. \quad \text{(C24)}$$

∎

APPENDIX D: AFCs AND MPDOs

In this appendix, we prove that the purity AFCs hold for MPDOs. To this end, we assume the following conditions on the transfer matrices (45):

(A) The matrices τ_1 and τ_2 admit the spectral decomposition

$$\tau_1 = \sum_{j=0}^{\chi-1} \lambda_j |R_j^{(1)}\rangle \langle L_j^{(1)}|, \quad \text{(D1a)}$$

$$\tau_2 = \sum_{j=0}^{\chi^2-1} \mu_j |R_j^{(2)}\rangle \langle L_j^{(2)}|, \quad \text{(D1b)}$$

where we assume $|\lambda_0| > |\lambda_j|, |\mu_0| > |\mu_j|$ for all $j \ge 1$; i.e., τ_1, τ_2 have a trivial Jordan form and a finite gap. (The assumption that τ_1 and τ_2 can be diagonalized is purely technical and not necessary. However, we keep it here as it makes the analysis simpler, and it is in any case quite general.) Note that we can also assume without loss of generality that $\lambda_0 = 1$ (which implies $\mu_0 > 0$, since $\rho_L > 0$ for all L). Finally, $|R_j^{(n)}\rangle$ and $\langle L_j^{(n)}|$ are the left and right eigenstates; i.e., they are vectors on the left and right virtual indices of τ_1 and τ_2 which are normalized such that

$$\langle L_j^{(n)}|R_k^{(n)}\rangle = \delta_{j,k}, \quad n = 1, 2. \quad \text{(D2)}$$

(B) We further need to assume the technical conditions

$$(\langle L_0^{(1)}| \otimes \langle L_0^{(1)}|)|R_0^{(2)}\rangle \neq 0, \qquad \text{(D3a)}$$

$$\langle L_0^{(2)}|(|R_0^{(1)}\rangle \otimes |R_0^{(1)}\rangle) \neq 0, \qquad \text{(D3b)}$$

which again are quite general, as orthogonality requires fine-tuning.

We use the same notations and assumptions as in Sec. IV C, so that

$$\mathcal{P}_2 = \frac{\mathrm{Tr}[\sigma^2]}{(\mathrm{Tr}[\sigma])^2} \qquad \text{(D4)}$$

and

$$r_2 = \frac{\prod_{j=1}^R \mathrm{Tr}_{I_j \cup I_{j+1}}(\sigma^2_{I_j \cup I_{j+1}})}{\prod_{j=1}^R \mathrm{Tr}_{I_j}[\sigma^2_{I_j}]}, \qquad \text{(D5)}$$

where $|I_j| = k$, and $R = L/k$ is an integer.

First, we introduce the correlation lengths

$$e^{-\frac{1}{\zeta_1}} = \frac{\max_{j>0}\{|\lambda_j|\}}{\lambda_0}, \qquad e^{-\frac{1}{\zeta_2}} = \frac{\max_{j>0}\{|\mu_j|\}}{\mu_0}, \qquad \text{(D6)}$$

and also

$$\zeta = \max(\zeta_1, \zeta_2). \qquad \text{(D7)}$$

Next, given the spectral decomposition in Eqs. (D1), we define

$$C = \max_{\substack{(j,k,l) \\ \neq(0,0,0)}} \left\{ \frac{|\langle L_j^{(1)}|\langle L_k^{(1)}|R_l^{(2)}\rangle\langle L_l^{(2)}|R_j^{(1)}\rangle|R_k^{(1)}\rangle|}{|\langle L_0^{(1)}|\langle L_0^{(1)}|R_0^{(2)}\rangle\langle L_0^{(2)}|R_0^{(1)}\rangle|R_0^{(1)}\rangle|} \right\}. \qquad \text{(D8)}$$

Note that the denominator is nonvanishing because of Eqs. (D3). We can now state the main result of this section.

Theorem 5. Take $|I_j| = k$ with

$$k \geq k_{\min} = \max\{1, \zeta \ln(20C\chi^2 L)\}. \qquad \text{(D9)}$$

Then, for all

$$L \geq \max\{\zeta \ln(2^5\chi^2), 4k + \zeta \ln(2\chi)\}, \qquad \text{(D10)}$$

we have

$$\left|\frac{r_2}{\mathcal{P}_2} - 1\right| \leq \chi^2(80C + 32)\frac{L}{k}e^{-k/\zeta}. \qquad \text{(D11)}$$

Proof. Using Eqs. (D1), we have

$$\mathcal{P}_2 = \frac{\sum_{j=0}^{\chi^2-1}\mu_j^L}{(\sum_{j=0}^{\chi-1}\lambda_j^L)^2} = \frac{\mu_0^L}{\lambda_0^{2L}}(1 + \tilde{\varepsilon}), \qquad \text{(D12)}$$

where

$$\tilde{\varepsilon} = \frac{1 + \sum_{j=1}^{\chi^2-1}(\mu_j/\mu_0)^L}{(1 + \sum_{j=1}^{\chi-1}(\lambda_j/\lambda_0)^L)^2} - 1. \qquad \text{(D13)}$$

Since by hypothesis $L \geq \zeta_1 \ln(2\chi)$, we have $\chi e^{-L/\zeta_1} \leq 1/2$, and so

$$|\tilde{\varepsilon}| \leq 4(\chi^2 e^{-L/\zeta_2} + 2\chi e^{-L/\zeta_1} + \chi^2 e^{-2L/\zeta_1})$$
$$\leq 2^4\chi^2 e^{-L/\zeta}. \qquad \text{(D14)}$$

On the other hand,

$$\mathrm{Tr}[\sigma_I^2] = \sum_{j,k=0}^{\chi-1}\lambda_j^{L-|I|}\lambda_k^{L-|I|}(\langle L_j^{(1)}| \otimes \langle L_k^{(1)}|)$$
$$\times \left(\sum_{l=0}^{\chi^2-1}\mu_l^{|I|}|R_l^{(2)}\rangle\langle L_l^{(2)}|\right)(|R_j^{(1)}\rangle \otimes |R_k^{(1)}\rangle). \qquad \text{(D15)}$$

Therefore, recalling the definition (D8), we have

$$\mathrm{Tr}[\sigma_I^2] = \lambda_0^{2(L-|I|)}\mu_0^{|I|}\langle L_0^{(1)}|\langle L_0^{(1)}|R_0^{(2)}\rangle$$
$$\times \langle L_0^{(2)}|R_0^{(1)}\rangle|R_0^{(1)}\rangle(1 + \varepsilon^I), \qquad \text{(D16)}$$

where

$$|\varepsilon^I| \leq 2C\chi e^{-(L-|I|)/\zeta_1} + C\chi^2 e^{-|I|/\zeta_2}$$
$$+ C\chi^2 e^{-2(L-|I|)/\zeta_1} + 2C\chi^3 e^{-(L-|I|)/\zeta_1}e^{-|I|/\zeta_2}$$
$$+ C\chi^4 e^{-2(L-|I|)/\zeta_1}e^{-|I|/\zeta_2}. \qquad \text{(D17)}$$

Since by hypothesis $L \geq 4k + \zeta \ln(2\chi)$, it is easy to verify that all five terms above are upper bounded by $C\chi^2 e^{-|I|/\zeta}$ (recall that $|I| \leq 2k$ for all I), and so

$$|\varepsilon^I| \leq 5C\chi^2 e^{-|I|/\zeta} \leq 5C\chi^2 e^{-k/\zeta}, \qquad \text{(D18)}$$

where we used that either $|I| = k$ or $|I| = 2k$, and so $|I| \geq k$. Therefore,

$$\frac{\prod_{j=1}^R \mathrm{Tr}_{I_j \cup I_{j+1}}(\sigma^2_{I_j \cup I_{j+1}})}{\prod_{j=1}^R \mathrm{Tr}_{I_j}[\sigma^2_{I_j}]}$$
$$= \frac{\mu_0^L}{\lambda_0^{2L}}\left(\frac{\prod_{j=1}^R(1 + \varepsilon^{I_j \cup I_{j+1}})}{\prod_{j=1}^R(1 + \varepsilon^{I_j})}\right), \qquad \text{(D19)}$$

with $|\varepsilon^I| \leq 5C\chi^2 e^{-k/\zeta}$. Combining this with Eq. (D12), we arrive at

$$\frac{r_2}{\mathcal{P}_2} - 1 = \left(\frac{\prod_{j=1}^R (1 + \varepsilon^{I_j \cup I_{j+1}})}{\prod_{j=1}^R (1 + \varepsilon^{I_j})}\right)(1 + \tilde{\varepsilon})^{-1} - 1$$

$$= (1 + \tilde{\varepsilon})^{-1}\left[\frac{\prod_{j=1}^R (1 + \varepsilon^{I_j \cup I_{j+1}})}{\prod_{j=1}^R (1 + \varepsilon^{I_j})} - 1 - \tilde{\varepsilon}\right], \quad \text{(D20)}$$

where $\tilde{\varepsilon}$ is given in Eq. (D13).

Next, we define

$$\varepsilon = \max \left\{|\varepsilon^I| : I \in \{I_j\}_j \cup \{I_j \cup I_{j+1}\}_j\right\}. \quad \text{(D21)}$$

By hypothesis $k \geq k_{\min}$ [$k_{\min}$ is given in Eq. (D9)]. Therefore, using Eq. (D18) we have $|\varepsilon^I| \leq 1/2$ for all $|I|$ and so also $\varepsilon \leq 1/2$. Therefore, we can apply the derivation in Lemma 1 to show

$$\left|\frac{\prod_{j=1}^R (1 + \varepsilon^{I_j \cup I_{j+1}})}{\prod_{j=1}^R (1 + \varepsilon^{I_j})} - 1\right| \leq e^{(4L/k)\varepsilon} - 1. \quad \text{(D22)}$$

Since by hypothesis $L \geq \xi \ln(2^5 \chi^2)$, we also have $\tilde{\varepsilon} \leq 1/2$ [cf. Eq. (D14)], and so Eq. (D20) yields

$$\left|\frac{r_2}{\mathcal{P}_2} - 1\right| \leq 2(e^{(4L/k)\varepsilon} - 1) + 2|\tilde{\varepsilon}|. \quad \text{(D23)}$$

Finally, we note that Eq. (D9) implies that $(4L/k)\varepsilon \leq 1$, and using $e^z - 1 \leq 2z$ for $0 \leq z \leq 1$, we arrive at

$$\left|\frac{r_2}{\mathcal{P}_2} - 1\right| \leq 80\chi^2 C(L/k)e^{-k/\zeta} + 2^5 \chi^2 e^{-L/\zeta}$$

$$\leq \chi^2 (80C + 32)\frac{L}{k}e^{-k/\zeta}. \quad \text{(D24)}$$

∎

This theorem proves Eq. (48), under the condition (47). Therefore, the approximate factorization property (40) holds for MPDOs, with the identification

$$\alpha_2 = (20C + 8)\chi^2, \qquad \beta_2 = 1/\zeta. \quad \text{(D25)}$$

APPENDIX E: DETAILS ON THE NUMERICAL COMPUTATIONS

In this appendix, we provide further details on the numerical computations performed to obtain the data presented in Secs. IV D and V. As mentioned, the calculations are carried out using the iTensor library [140] by first approximating the thermal states by MPOs of bond dimension χ and subsequently taking powers and traces of the density matrices represented in this way. For each quantity, we have always verified that the results were stable upon increasing the bond dimension χ. An example of our data is reported in Fig. 9, where we study $\varepsilon_r(k, L, \beta)$ defined in Eq. (52) as a function of the bond dimension χ used to approximate the thermal state. In general, we have found that relatively small bond dimensions are enough in the quantum Ising chain, while larger bond dimensions are required in order to observe convergence in the Heisenberg model.
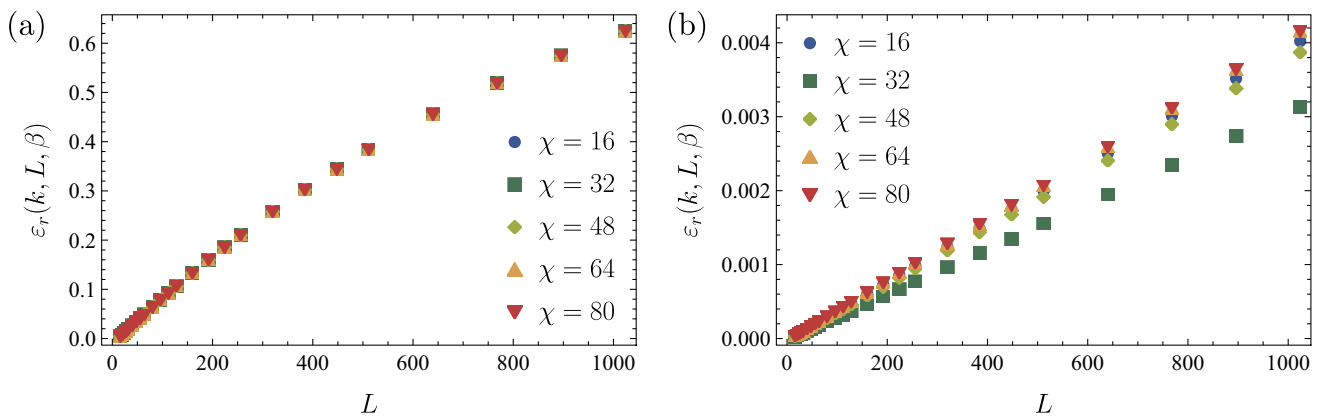

FIG. 9. Scaling of the relative error $\varepsilon_r(k, L, \beta)$ defined in Eq. (52) as a function of the bond dimension χ used to approximate the thermal state. In (a), we show data for the Ising chain while (b) shows data for the XXZ chain. In both cases, we use the same parameters as in Fig. 4. Data shown for $k = 4$, other values of k show similar behavior with $\chi = 16$ being sufficient for convergence in the Ising model, while the XXZ model appears to be well converged from $\chi = 64$.

APPENDIX F: TECHNICAL DETAILS ON THE PPT CONDITIONS

The goal of this appendix is to identify the conditions under which the class of states $\rho^{(L)}(\gamma)$ introduced in Sec. V B satisfy Eqs. (60) for all system sizes L and some suitable constants C_3, C_5 independent of L.

For concreteness, let the circuit depth ℓ be even, $\ell = 2k$ with $k \geq 1$ (a similar discussion holds for $\ell = 2k + 1$) and take a bipartition of the system as in Fig. 10. We group neighboring sites into blocks containing $q = \ell - 1$ qubits, forming a new superqudit associated with a Hilbert space of dimension $d = 2^{\ell - 1}$. It is easy to show that the circuit can be rewritten as a depth-2 quantum circuit acting on the super qudits; cf. Fig. 10. Therefore (assuming without loss of generality $R = L/q$ is an integer),

$$\rho_L(\gamma) = V^{(2)} \left(\bigotimes_{j=1}^{L/q} \omega_j \right) [V^{(2)}]^\dagger, \tag{F1}$$

where $\omega_j = \otimes_{i=1}^q \sigma_{jq+i}$, while

$$V^{(2)} = \left(\prod_j V_{2j,2j+1} \right) \left(\prod_j V_{2j+1,2j+2} \right) \tag{F2}$$

is the depth-2 FDQC acting on the superlattice.

As it is manifest from Fig. 10, the region A contains all superqudits with labels from 1 to $L/2q$, while B contains all those with labels from $L/2q + 1$ to $R = L/q$. Define the sets of superqudits $S_1 = \{R/2 - 1, R/2\}$, $S_2 = \{R/2 + 1, R/2 + 2\}$, and

$$\tilde{\rho}_{S_1 S_2} = W \omega_{R/2-1} \otimes \omega_{R/2} \otimes \omega_{R/2+1} \otimes \omega_{R/2+2} W^\dagger, \tag{F3}$$

where $W = V_{R/2,R/2+1} V_{R/2-1,R/2} V_{R/2+1,R/2+2}$. Note that $\tilde{\rho}_{S_1 S_2}$ is different from the reduced density matrix over $S_1 \cup S_2$. Using the graphical representation for the blocked circuits and Eq. (27), it is simple to show

$$\tilde{p}_n[AB] = \frac{\mathrm{Tr}[(\tilde{\rho}_{S_1 S_2}^{T_{S_1}})^n]}{\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^n]\mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^n]} =: \tilde{s}_n, \tag{F4}$$

where $\tilde{\rho}_{S_1} = \mathrm{Tr}_{S_2}[\tilde{\rho}_{S_1 S_2}]$ and $\tilde{\rho}_{S_2} = \mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1 S_2}]$. Therefore, $\tilde{p}_n[AB]$ coincides with the normalized PT moments of the state $\tilde{\rho}_{S_1 S_2}$ supported on four superqudits. Next, it is also easy to compute

$$\mathcal{P}_n[A] = \mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^n][\gamma^n + (1 - \gamma)^n]^{L/2 - 2q},$$
$$\mathcal{P}_n[B] = \mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^n][\gamma^n + (1 - \gamma)^n]^{L/2 - 2q}. \tag{F5}$$

Finally, setting

$$\tilde{t}_3 = \frac{\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^2]^2 \mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^2]^2}{\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^3]\mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^3]}, \tag{F6}$$

$$\tilde{t}_5 = \frac{\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^4]^2 \mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^4]^2}{\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^3]\mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^3]\mathrm{Tr}_{S_1}[\tilde{\rho}_{S_1}^5]\mathrm{Tr}_{S_2}[\tilde{\rho}_{S_2}^5]}, \tag{F7}$$

we arrive at

$$f_3 = \tilde{s}_3 - \tilde{s}_2^2 \tilde{t}_3 \frac{[\gamma^2 + (1 - \gamma)^2]^{2L - 8q}}{[\gamma^3 + (1 - \gamma)^3]^{L - 4q}}, \tag{F8a}$$

$$f_5 = \tilde{s}_5 \tilde{s}_3 - \tilde{s}_4^2 \tilde{t}_5 \frac{[\gamma^4 + (1 - \gamma)^4]^{2L - 8q}}{[\gamma^3 + (1 - \gamma)^3]^{L - 4q}[\gamma^5 + (1 - \gamma)^5]^{L - 4q}}. \tag{F8b}$$

Now, define

$$K_3 = \max\{\tilde{s}_3(\gamma) - \tilde{s}_2^2(\gamma)\tilde{t}_3(\gamma) : 0 \leq \gamma \leq 1/4\}, \tag{F9}$$

$$K_5 = \max\{\tilde{s}_5(\gamma)\tilde{s}_3(\gamma) - \tilde{s}_4^2(\gamma)\tilde{t}_5(\gamma) : 0 \leq \gamma \leq 1/4\}, \tag{F10}$$

and

$$H_3 = \max\{|\tilde{s}_2^2(\gamma)\tilde{t}_3(\gamma)| : 0 \leq \gamma \leq 1/4\}, \tag{F11}$$

$$H_5 = \max\{|\tilde{s}_4^2(\gamma)\tilde{t}_5(\gamma)| : 0 \leq \gamma \leq 1/4\}. \tag{F12}$$

The constants K_3, K_5, H_3, and H_5 depend on the specific choices of the unitary gates forming $U^{(\ell)}$. For nonentangling gates, K_3 and K_5 are positive, but for generic choices
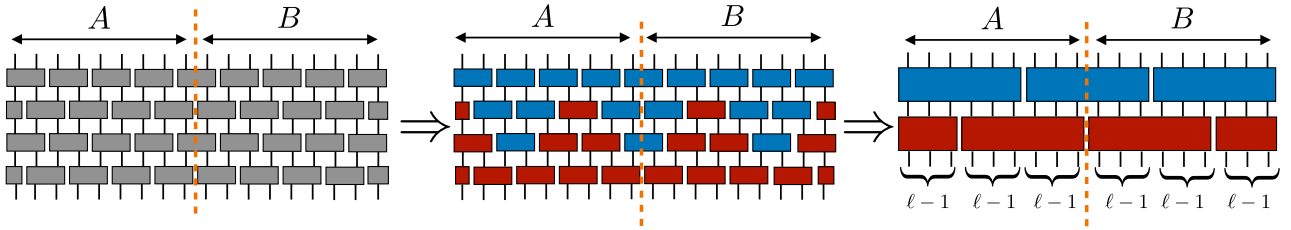

FIG. 10. Any FDQC of depth ℓ can be rewritten as a depth-2 quantum circuit after grouping $\ell - 1$ qubits into a single qudit. In the middle panel, we highlight with different colored sets of gates defining the two-qudit gates in the grouped lattice (colored rectangles in the right panel). In all panels, a dashed orange line separates the regions A and B, defining the bipartition of the system.

of gates, one has $K_3 < 0$ and $K_5 < 0$. We are ready to state our main result.

Theorem 6. Suppose K_3, $K_5 < 0$. Then, assuming without loss of generality $\gamma \leq 1/4$ and defining

$$C_l = \frac{|K_l|}{2}, \quad \gamma_3 = -\frac{K_3}{4H_3}, \quad \gamma_5 = \left(-\frac{K_5}{8H_5}\right)^{1/3}, \quad \text{(F13)}$$

we have

$$0 \leq \gamma \leq \gamma_3/L \Rightarrow f_3 \leq -C_3, \quad \text{(F14)}$$

while

$$0 \leq \gamma \leq \gamma_5/L^{1/3} \Rightarrow f_5 \leq -C_5, \quad \text{(F15)}$$

for all system sizes L.

Proof. We start from Eq. (F8) and note

$$\frac{[\gamma^2 + (1-\gamma)^2]^2}{\gamma^3 + (1-\gamma)^3} = 1 - \varepsilon_3, \quad \text{(F16)}$$

with $0 \leq \varepsilon_3 \leq \gamma$ for $0 \leq \gamma \leq 1/4$. Therefore,

$$f_3 \leq K_3 + \tilde{s}_2^2 \tilde{t}_3 [1 - (1-\varepsilon_3)^{L-4q}]$$
$$\leq K_3 + H_3 [1 - (1-\gamma)^{L-4q}]. \quad \text{(F17)}$$

Using $1 - \gamma \geq e^{-2\gamma}$ for $0 \leq \gamma \leq 1/4$, we get

$$f_3 \leq K_3 + H_3 [1 - e^{-2\gamma(L-4q)}] \leq K_3 + 2H_3 \gamma L. \quad \text{(F18)}$$

Hence, if $\gamma \leq \gamma_3/L$, we finally arrive at

$$f_3 \leq K_3/2 = -C_3. \quad \text{(F19)}$$

Analogously, we have

$$\frac{[\gamma^4 + (1-\gamma)^4]^2}{[\gamma^3 + (1-\gamma)^3][\gamma^5 + (1-\gamma)^5]} = 1 - \varepsilon_5, \quad \text{(F20)}$$

with $0 \leq \varepsilon_3 \leq 2\gamma^3$ for $0 \leq \gamma \leq 1/4$. Therefore,

$$f_5 \leq K_5 + \tilde{s}_4^2 \tilde{t}_5 [1 - (1-\varepsilon_5)^{L-4q}]$$
$$\leq K_3 + H_5 [1 - (1-2\gamma^3)^{L-4q}]. \quad \text{(F21)}$$

Using $1 - x \geq e^{-2x}$ for $0 \leq x \leq 1/4$, we get

$$f_5 \leq K_5 + H_5 [1 - e^{-4\gamma^3(L-4q)}] \leq K_5 + 4H_5 \gamma^3 L. \quad \text{(F22)}$$

Hence, if $\gamma \leq \gamma_5/L^{1/3}$, we arrive at

$$f_5 \leq K_5/2 = -C_5. \quad \text{(F23)}$$

∎

[1] R. Blatt and C. F. Roos, *Quantum simulations with trapped ions*, Nat. Phys. **8**, 277 (2012).

[2] C. Gross and I. Bloch, *Quantum simulations with ultracold atoms in optical lattices*, Science **357**, 995 (2017).

[3] F. Schäfer, T. Fukuhara, S. Sugawa, Y. Takasu, and Y. Takahashi, *Tools for quantum simulation with ultracold atoms in optical lattices*, Nat. Rev. Phys. **2**, 411 (2020).

[4] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, *Superconducting qubits: Current state of play*, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[5] M. Morgado and S. Whitlock, *Quantum simulation and computing with Rydberg-interacting qubits*, AVS Quantum Sci. **3**, 023501 (2021).

[6] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao, *Programmable quantum simulations of spin systems with trapped ions*, Rev. Mod. Phys. **93**, 025001 (2021).

[7] Y. Alexeev, D. Bacon, K. R. Brown, R. Calderbank, L. D. Carr, F. T. Chong, B. DeMarco, D. Englund, E. Farhi, B. Fefferman *et al.*, *Quantum computer systems for scientific discovery*, PRX Quantum **2**, 017001 (2021).

[8] E. Pelucchi, G. Fagas, I. Aharonovich, D. Englund, E. Figueroa, Q. Gong, H. Hannes, J. Liu, C.-Y. Lu, N. Matsuda *et al.*, *The potential and global outlook of integrated photonics for quantum technologies*, Nat. Rev. Phys. **4**, 194 (2022).

[9] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, *Semiconductor spin qubits*, Rev. Mod. Phys. **95**, 025003 (2023).

[10] J. Preskill, *Quantum computing in the NISQ era and beyond*, Quantum **2**, 79 (2018).

[11] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. A. Eriksson, Kai-Mei C. Fu *et al.*, *Quantum simulators: Architectures and opportunities*, PRX Quantum **2**, 017003 (2021).

[12] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, *The randomized measurement toolbox*, Nat. Rev. Phys. **5**, 9 (2023).

[13] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, *Statistical correlations between locally randomized measurements: A toolbox for probing entanglement in many-body quantum states*, Phys. Rev. A **99**, 052323 (2019).

[14] P. Cieśliński, S. Imai, J. Dziewior, O. Gühne, L. Knips, W. Laskowski, J. Meinecke, T. Paterek, and T. Vértesi, *Analysing quantum systems with randomised measurements*, arXiv:2307.01251.

[15] H.-Y. Huang, R. Kueng, and J. Preskill, *Predicting many properties of a quantum system from very few measurements*, Nat. Phys. **16**, 1050 (2020).

[16] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865 (2009).

[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).

[18] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, J. Phys. A **42**, 504005 (2009).

[19] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, Rev. Mod. Phys. **82,** 277 (2010).

[20] R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, and M. Greiner, *Measuring entanglement entropy in a quantum many-body system*, Nature (London) **528,** 77 (2015).

[21] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, *Quantum thermalization through entanglement in an isolated many-body system*, Science **353,** 794 (2016).

[22] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, *Measuring the Rényi entropy of a two-site Fermi-Hubbard model on a trapped ion quantum computer*, Phys. Rev. A **98,** 052334 (2018).

[23] C. M. Alves and D. Jaksch, *Multipartite entanglement detection in bosons*, Phys. Rev. Lett. **93,** 110501 (2004).

[24] J. Cardy, *Measuring entanglement using quantum quenches*, Phys. Rev. Lett. **106,** 150404 (2011).

[25] A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, *Measuring entanglement growth in quench dynamics of bosons in an optical lattice*, Phys. Rev. Lett. **109,** 020505 (2012).

[26] D. A. Abanin and E. Demler, *Measuring entanglement entropy of a generic many-body system with a quantum switch*, Phys. Rev. Lett. **109,** 020504 (2012).

[27] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, *Probing Rényi entanglement entropy via randomized measurements*, Science **364,** 260 (2019).

[28] S. J. van Enk and C. W. J. Beenakker, *Measuring Trρⁿ on single copies of ρ using random measurements*, Phys. Rev. Lett. **108,** 110503 (2012).

[29] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, *Rényi entropies from random quenches in atomic Hubbard and spin models*, Phys. Rev. Lett. **120,** 050406 (2018).

[30] K. J. Satzinger, Y.-J. Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth *et al.*, *Realizing topologically ordered states on a quantum processor*, Science **374,** 1237 (2021).

[31] R. Stricker, M. Meth, L. Postler, C. Edmunds, C. Ferrie, R. Blatt, P. Schindler, T. Monz, R. Kueng, and M. Ringbauer, *Experimental single-setting quantum state tomography*, PRX Quantum **3,** 040310 (2022).

[32] D. Zhu, Z. P. Cian, C. Noel, A. Risinger, D. Biswas, L. Egan, Y. Zhu, A. M. Green, C. H. Alderete, N. H. Nguyen *et al.*, *Cross-platform comparison of arbitrary quantum states*, Nat. Commun. **13,** 6620 (2022).

[33] J. C. Hoke, M. Ippoliti, E. Rosenberg, D. Abanin, R. Acharya, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw *et al.*, *Measurement-induced entanglement and teleportation on a noisy quantum processor*, Nature (London) **622,** 481 (2023).

[34] Y. Zhou, P. Zeng, and Z. Liu, *Single-copies estimation of entanglement negativity*, Phys. Rev. Lett. **125,** 200502 (2020).

[35] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, *Mixed-state entanglement from local randomized measurements*, Phys. Rev. Lett. **125,** 200501 (2020).

[36] M. Cerezo, A. Sone, J. L. Beckey, and P. J. Coles, *Sub-quantum Fisher information*, Quantum Sci. Technol. **6,** 035008 (2021).

[37] M. Yu, D. Li, J. Wang, Y. Chu, P. Yang, M. Gong, N. Goldman, and J. Cai, *Experimental estimation of the quantum Fisher information from randomized measurements*, Phys. Rev. Res. **3,** 043122 (2021).

[38] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch, *Quantum Fisher information from randomized measurements*, Phys. Rev. Lett. **127,** 260501 (2021).

[39] V. Vitale, A. Rath, P. Jurcevic, A. Elben, C. Branciard, and B. Vermersch, *Estimation of the quantum Fisher information on a quantum processor*, arXiv:2307.16882.

[40] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, *Quantum tomography via compressed sensing: Error bounds, sample complexity and efficient estimators*, New J. Phys. **14,** 095022 (2012).

[41] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, *Sample-optimal tomography of quantum states*, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (2016), pp. 913–925, https://ieeexplore.ieee.org/document/7956181.

[42] R. O'Donnell and J. Wright, *Efficient quantum tomography*, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (2016), pp. 899–912, 10.1145/2897518.2897544.

[43] M. Paini and A. Kalev, *An approximate description of quantum states*, arXiv:1910.10543.

[44] C. Hadfield, *Adaptive Pauli shadows for energy estimation*, arXiv:2105.12207.

[45] H.-Y. Huang, R. Kueng, and J. Preskill, *Efficient estimation of Pauli observables by derandomization*, Phys. Rev. Lett. **127,** 030503 (2021).

[46] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, *Measurements of quantum Hamiltonians with locally-biased classical shadows*, Commun. Math. Phys. **391,** 951 (2022).

[47] A. Rath, R. van Bijnen, A. Elben, P. Zoller, and B. Vermersch, *Importance sampling of randomized measurements for probing entanglement*, Phys. Rev. Lett. **127,** 200503 (2021).

[48] B. Vermersch, A. Rath, B. Sundar, C. Branciard, J. Preskill, and A. Elben, *Enhanced estimation of quantum properties with common randomized measurements*, PRX Quantum **5,** 010352 (2024).

[49] T.-C. Yen, A. Ganeshram, and A. F. Izmaylov, *Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates*, npj Quantum Inf. **9,** 14 (2023).

[50] M. Arienzo, M. Heinrich, I. Roth, and M. Kliesch, *Closed-form analytic expressions for shadow estimation with brickwork circuits*, Quantum Inf. Comput. **23,** 961 (2023).

[51] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, *Scalable and flexible classical shadow tomography with tensor networks*, Quantum **7,** 1026 (2023).

[52] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan, *Shallow shadows: Expectation*

estimation using low-depth random Clifford circuits, **133**, 020602 (2024).

[53] M. Ippoliti, Y. Li, T. Rakovszky, and V. Khemani, *Operator relaxation and the optimal depth of classical shadows*, Phys. Rev. Lett. **130**, 230403 (2023).

[54] M. Ippoliti, *Classical shadows based on locally-entangled measurements*, Quantum **8**, 1293 (2024).

[55] G. García-Pérez, M. A. C. Rossi, B. Sokolov, F. Tacchino, P. K. Barkoutsos, G. Mazzola, I. Tavernelli, and S. Maniscalco, *Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms*, PRX Quantum **2**, 040342 (2021).

[56] I. M. Georgescu, S. Ashhab, and F. Nori, *Quantum simulation*, Rev. Mod. Phys. **86**, 153 (2014).

[57] F. Tacchino, A. Chiesa, S. Carretta, and D. Gerace, *Quantum computers as universal quantum simulators: State-of-the-art and perspectives*, Adv. Quantum Technol. **3**, 1900052 (2020).

[58] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, *Practical quantum advantage in quantum simulation*, Nature (London) **607**, 667 (2022).

[59] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Matrix product state representations*, Quantum Inf. Comput. **7**, 401 (2007).

[60] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, *Matrix product states and projected entangled pair states: Concepts, symmetries, theorems*, Rev. Mod. Phys. **93**, 045003 (2021).

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

[62] B. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. Buyskikh, A. Daley, M. Cramer *et al.*, *Efficient tomography of a quantum manybody system*, Nat. Phys. **13**, 1158 (2017).

[63] Z. Qin, C. Jameson, Z. Gong, M. B. Wakin, and Z. Zhu, *Stable tomography for structured quantum states*, arXiv:2306.09432.

[64] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, *Scalable reconstruction of density matrices*, Phys. Rev. Lett. **111**, 020401 (2013).

[65] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio, *A scalable maximum likelihood method for quantum state tomography*, New J. Phys. **15**, 125004 (2013).

[66] M. Holzäpfel, M. Cramer, N. Datta, and M. B. Plenio, *Petz recovery versus matrix reconstruction*, J. Math. Phys. (N.Y.) **59**, 042201 (2018).

[67] A. Lidiak, C. Jameson, Z. Qin, G. Tang, M. B. Wakin, Z. Zhu, and Z. Gong, *Quantum state tomography with tensor train cross approximation*, arXiv:2207.06397.

[68] C. Kokail, R. van Bijnen, A. Elben, B. Vermersch, and P. Zoller, *Entanglement hamiltonian tomography in quantum simulation*, Nat. Phys. **17**, 936 (2021).

[69] M. K. Joshi, C. Kokail, R. van Bijnen, F. Kranzl, T. V. Zache, R. Blatt, C. F. Roos, and P. Zoller, *Exploring large-scale entanglement in quantum simulation*, Nature (London) **624**, 539 (2023).

[70] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar, *Sample-efficient learning of interacting quantum systems*, Nat. Phys. **17**, 931 (2021).

[71] A. Anshu and S. Arunachalam, *A survey on the complexity of learning quantum states*, arXiv:2305.20069.

[72] C. Rouzé and D. S. França, *Learning quantum manybody systems from a few copies*, Quantum **8**, 1319 (2024).

[73] E. Onorati, C. Rouzé, D. S. França, and J. D. Watson, *Efficient learning of ground & thermal states within phases of matter*, arXiv:2301.12946.

[74] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, *Permutationally invariant quantum tomography*, Phys. Rev. Lett. **105**, 250403 (2010).

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

[76] C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, *Experimental comparison of efficient tomography schemes for a six-qubit state*, Phys. Rev. Lett. **113**, 040503 (2014).

[77] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, *Quantum state tomography via compressed sensing*, Phys. Rev. Lett. **105**, 150401 (2010).

[78] A. Montanaro, *Learning stabilizer states by Bell sampling*, arXiv:1707.04012.

[79] S. S. Kuzmin, V. I. Mikhailova, I. V. Dyakonov, and S. S. Straupe, *Learning the tensor network model of a quantum state using a few single-qubit measurements*, Phys. Rev. A **109**, 052616 (2024).

[80] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Neural-network quantum state tomography*, Nat. Phys. **14**, 447 (2018).

[81] L. Zhao, N. Guo, M.-X. Luo, and P. Rebentrost, *Provable learning of quantum states with graphical models circuits*, arXiv:2309.09235.

[82] M. Rieger, M. Reh, and M. Gärttner, *Sample-efficient estimation of entanglement entropy through supervised learning*, Phys. Rev. A **109**, 012403 (2024).

[83] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, *Reconstructing quantum states with generative models*, Nat. Mach. Intell. **1**, 155 (2019).

[84] T. Schmale, M. Reh, and M. Gärttner, *Efficient quantum state tomography with convolutional neural networks*, npj Quantum Inf. **8**, 115 (2022).

[85] D. J. Luitz, N. Laflorencie, and F. Alet, *Participation spectroscopy and entanglement Hamiltonian of quantum spin models*, J. Stat. Mech. (2014) P08007.

[86] J.-M. Stéphan, S. Furukawa, G. Misguich, and V. Pasquier, *Shannon and entanglement entropies of one- and twodimensional critical wave functions*, Phys. Rev. B **80**, 184421 (2009).

[87] J.-M. Stéphan, G. Misguich, and V. Pasquier, *Rényi entropy of a line in two-dimensional Ising models*, Phys. Rev. B **82**, 125455 (2010).

[88] F. C. Alcaraz and M. A. Rajabpour, *Universal behavior of the Shannon mutual information of critical quantum chains*, Phys. Rev. Lett. **111**, 017201 (2013).

[89] J.-M. Stéphan, *Shannon and Rényi mutual information in quantum critical spin chains*, Phys. Rev. B **90**, 045424 (2014).

[90] P. Sierant and X. Turkeshi, *Universal behavior beyond multifractality of wave functions at measurement-induced phase transitions*, Phys. Rev. Lett. **128**, 130605 (2022).

[91] X. Turkeshi and P. Sierant, *Error-resilience phase transitions in encoding-decoding quantum circuits*, Phys. Rev. Lett. **132**, 140401 (2024).

[92] L. Leone, S. F. E. Oliviero, and A. Hamma, *Stabilizer Rényi entropy*, Phys. Rev. Lett. **128**, 050402 (2022).

[93] S. F. E. Oliviero, L. Leone, and A. Hamma, *Magic-state resource theory for the ground state of the transverse-field Ising model*, Phys. Rev. A **106**, 042426 (2022).

[94] T. Haug and L. Piroli, *Quantifying nonstabilizerness of matrix product states*, Phys. Rev. B **107**, 035148 (2023).

[95] T. Haug and L. Piroli, *Stabilizer entropies and nonstabilizerness monotones*, Quantum **7**, 1092 (2023).

[96] G. Lami and M. Collura, *Nonstabilizerness via perfect Pauli sampling of matrix product states*, Phys. Rev. Lett. **131**, 180401 (2023).

[97] P. S. Tarabunga, E. Tirrito, T. Chanda, and M. Dalmonte, *Many-body magic via Pauli-Markov chains–from criticality to gauge theories*, PRX Quantum **4**, 040317 (2023).

[98] P. Niroula, C. D. White, Q. Wang, S. Johri, D. Zhu, C. Monroe, C. Noel, and M. J. Gullans, *Phase transition in magic with random quantum circuits*, arXiv:2304.10481.

[99] X. Turkeshi, M. Schirò, and P. Sierant, *Measuring nonstabilizerness via multifractal flatness*, Phys. Rev. A **108**, 042408 (2023).

[100] G. Vidal and R. F. Werner, *Computable measure of entanglement*, Phys. Rev. A **65**, 032314 (2002).

[101] M. B. Plenio, *Logarithmic negativity: A full entanglement monotone that is not convex*, Phys. Rev. Lett. **95**, 090503 (2005).

[102] H. A. Carteret, *Noiseless quantum circuits for the Peres separability criterion*, Phys. Rev. Lett. **94**, 040502 (2005).

[103] J. Gray, L. Banchi, A. Bayat, and S. Bose, *Machine-learning-assisted many-body entanglement measurement*, Phys. Rev. Lett. **121**, 150503 (2018).

[104] H. A. Carteret, *Estimating the entanglement negativity from low-order moments of the partially transposed density matrix*, arXiv:1605.08751.

[105] A. Neven, J. Carrasco, V. Vitale, C. Kokail, A. Elben, M. Dalmonte, P. Calabrese, P. Zoller, B. Vermersch, R. Kueng *et al.*, *Symmetry-resolved entanglement detection using partial transpose moments*, npj Quantum Inf. **7**, 152 (2021).

[106] X.-D. Yu, S. Imai, and O. Gühne, *Optimal entanglement certification from moments of the partial transpose*, Phys. Rev. Lett. **127**, 060504 (2021).

[107] J. Carrasco, M. Votto, V. Vitale, C. Kokail, A. Neven, P. Zoller, B. Vermersch, and B. Kraus, *Entanglement phase diagrams from partial transpose moments*, Phys. Rev. A **109**, 012422 (2024).

[108] L. Knips, J. Dziewior, W. Kłobus, W. Laskowski, T. Paterek, P. J. Shadbolt, H. Weinfurter, and J. D. Meinecke, *Multipartite entanglement analysis from random correlations*, npj Quantum Inf. **6**, 51 (2020).

[109] A. Ketterer, N. Wyderka, and O. Gühne, *Characterizing multipartite entanglement with moments of random correlations*, Phys. Rev. Lett. **122**, 120505 (2019).

[110] L. Piroli, G. Styliaris, and J. I. Cirac, *Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter*, Phys. Rev. Lett. **127**, 220503 (2021).

[111] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, *Preparation of matrix product states with log-depth quantum circuits*, Phys. Rev. Lett. **132**, 040404 (2024).

[112] X. Chen, Z.-C. Gu, and X.-G. Wen, *Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order*, Phys. Rev. B **82**, 155138 (2010).

[113] M. B. Hastings, *Classifying quantum phases with the Kirby torus trick*, Phys. Rev. B **88**, 165114 (2013).

[114] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen, *Quantum information meets quantum matter–from quantum entanglement to topological phase in many-body systems*, arXiv:1508.02595.

[115] B. Zeng and X.-G. Wen, *Gapped quantum liquids and topological order, stochastic local transformations and emergence of unitarity*, Phys. Rev. B **91**, 125121 (2015).

[116] A. Rath, V. Vitale, S. Murciano, M. Votto, J. Dubail, R. Kueng, C. Branciard, P. Calabrese, and B. Vermersch, *Entanglement barrier and its symmetry resolution: Theory and experimental observation*, PRX Quantum **4**, 010318 (2023).

[117] D. Poulin and M. B. Hastings, *Markov entropy decomposition: A variational dual for quantum belief propagation*, Phys. Rev. Lett. **106**, 080403 (2011).

[118] K. Kato and F. G. Brandao, *Quantum approximate Markov chains are thermal*, Commun. Math. Phys. **370**, 117 (2019).

[119] F. G. Brandão and M. J. Kastoryano, *Finite correlation length implies efficient preparation of quantum thermal states*, Commun. Math. Phys. **365**, 1 (2019).

[120] I. H. Kim, *Markovian matrix product density operators: Efficient computation of global entropy*, arXiv:1709.07828.

[121] P. Svetlichnyy, S. Mittal, and T. Kennedy, *Matrix product states and the decay of quantum conditional mutual information*, J. Math. Phys. (N.Y.) **65**, 022201 (2024).

[122] P. Svetlichnyy and T. Kennedy, *Decay of quantum conditional mutual information for purely generated finitely correlated states*, J. Math. Phys. (N.Y.) **63**, 072201 (2022).

[123] D. Haag, F. Baccari, and G. Styliaris, *Typical correlation length of sequentially generated tensor network states*, PRX Quantum **4**, 030330 (2023).

[124] C.-F. Chen, K. Kato, and F. G. Brandão, *Matrix product density operators: When do they have a local parent Hamiltonian?*, arXiv:2010.14682.

[125] I. H. Kim, *Entropy scaling law and the quantum marginal problem*, Phys. Rev. X **11**, 021039 (2021).

[126] I. H. Kim, *Entropy scaling law and the quantum marginal problem: Simplification and generalization*, arXiv:2109.11688.

[127] T. Kuwahara, K. Kato, and F. G. S. L. Brandão, *Clustering of conditional mutual information for quantum Gibbs states above a threshold temperature*, Phys. Rev. Lett. **124**, 220601 (2020).

[128] T. Kuwahara, A. M. Alhambra, and A. Anshu, *Improved thermal area law and quasilinear time algorithm for quantum Gibbs states*, Phys. Rev. X **11**, 011047 (2021).

[129] P. Silvi, F. Tschirsich, M. Gerster, J. Jünemann, D. Jaschke, M. Rizzi, and S. Montangero, *The tensor networks anthology: Simulation techniques for many-body quantum lattice systems*, SciPost Phys. Lect. Notes **8** (2019). 10.21468/SciPostPhysLectNotes.8

[130] K. M. Audenaert, *A sharp continuity estimate for the von Neumann entropy*, J. Phys. A **40**, 8127 (2007).

[131] Z. Chen, Z. Ma, I. Nikoufar, and S. Fei, *Sharp continuity bounds for entropy and conditional entropy*, Sci. China Phys. Mech. Astron. **60**, 020321 (2017).

[132] M. A. Nielsen, *Continuity bounds for entanglement*, Phys. Rev. A **61**, 064301 (2000).

[133] T. Farrelly, *A review of quantum cellular automata*, Quantum **4**, 368 (2020).

[134] L. Piroli and J. I. Cirac, *Quantum cellular automata, tensor networks, and area laws*, Phys. Rev. Lett. **125**, 190402 (2020).

[135] J. I. Cirac, D. Perez-Garcia, N. Schuch, and F. Verstraete, *Matrix product density operators: Renormalization fixed points and boundary theories*, Ann. Phys. (Amsterdam) **378**, 100 (2017).

[136] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, England, 1999), Vol. 12.

[137] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge University Press, Cambridge, England, 1997), Vol. 3.

[138] An interesting question is whether our approach could be extended to pure critical states. Although in this case there is no spatial length scale, it is known that they can be accurately approximated by MPSs with bond-dimension scaling polynomially in the system size [139], suggesting that our approach could be generalized to include that case as well. We leave this problem for future work.

[139] F. Verstraete and J. I. Cirac, *Matrix product states represent ground states faithfully*, Phys. Rev. B **73**, 094423 (2006).

[140] M. Fishman, S. R. White, and E. M. Stoudenmire, *The ITENSOR software library for tensor network calculations*, SciPost Phys. Codebases **4** (2022).

[141] A. J. Ferris and G. Vidal, *Perfect sampling with unitary tensor networks*, Phys. Rev. B **85**, 165146 (2012).

[142] Note that, while in Sec. IV B we considered only the purity $\mathcal{P}_2$, similar AFCs can be defined for $\mathcal{P}_n$ with $n > 2$, and hence, similar estimation protocols work to estimate higher moments. Note also that, for the case of FDQCs, the proof of exact factorization conditions for $\mathcal{P}_2$ can be trivially extended to any $\mathcal{P}_n$ with $n > 2$; cf. Appendix A.

[143] It is interesting to note that the p_3-PPT condition allows us to detect only bipartite entanglement in states with $O(1)$ entropy. Therefore, the p_5-PPT condition improves this result by allowing the entropy to scale as $O(L^{2/3})$. It is then natural to conjecture that higher-n PPT conditions lead to a further improvement of the maximal scaling to $O(L^\alpha)$, with α approaching 1 in the large-n limit.

[144] A. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller, *Cross-platform verification of intermediate scale quantum devices*, Phys. Rev. Lett. **124**, 010504 (2020).

[145] H. T. Quan, Z. Song, X. F. Liu, P. Zanardi, and C. P. Sun, *Decay of Loschmidt echo enhanced by quantum criticality*, Phys. Rev. Lett. **96**, 140604 (2006).

[146] P. Gondolf, S. O. Scalet, A. Ruiz-de Alarcon, A. M. Alhambra, and A. Capel, *Conditional independence of 1D Gibbs states with applications to efficient learning*, arXiv:2402.18500.

[147] S. Rana, *Negative eigenvalues of partial transposition of arbitrary bipartite states*, Phys. Rev. A **87**, 054301 (2013).