Contracting projected entangled pair states is average-case hard

Jonas Haferkamp, Dominik Hangleiter, Jens Eisert, and Marek Gluza
Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany

(Received 31 October 2018; revised manuscript received 27 October 2019; published 3 January 2020)

An accurate calculation of the properties of quantum many-body systems is one of the most important yet intricate challenges of modern physics and computer science. In recent years, the tensor network ansatz has established itself as one of the most promising approaches enabling striking efficiency of simulating static properties of one-dimensional systems and abounding numerical applications in condensed matter theory. In higher dimensions, however, a connection to the field of computational complexity theory has shown that the accurate normalization of the two-dimensional tensor networks called projected entangled pair states (PEPS) is #P-complete. Therefore an efficient algorithm for PEPS contraction would allow solving exceedingly difficult combinatorial counting problems, which is considered highly unlikely. Due to the importance of understanding two- and three-dimensional systems the question currently remains: Are the known constructions typical of states relevant for quantum many-body systems? In this work, we show that an accurate evaluation of normalization or expectation values of PEPS is as hard to compute for typical instances as for special configurations of highest computational hardness. We discuss the structural property of average-case hardness in relation to the current research on efficient algorithms attempting tensor network contraction, hinting at a wealth of possible further insights into the average-case hardness of important problems in quantum many-body theory.

DOI: 10.1103/PhysRevResearch.2.013010

I. INTRODUCTION

Determining the properties of quantum many-body systems is of paramount importance in our efforts to understand conductance and thermodynamics of solid-state materials [1,2], designing new sensors and devising novel quantum technologies [3], inferring nuclear processes in stars or the early universe [4,5]. However, oftentimes it is not possible to find degrees of freedom enabling a concise description of a given system in terms of an effective model featuring essentially no interactions. In such a case, there is usually no easy way out but to calculate numerically observables of interest from a Hamiltonian description [6–12]. Here, however, we face a particular challenge namely that the state space of quantum many-body systems demands a number of parameters that grows exponentially with the amount of constituents of the system. If so, even storing the state of the system on a computer becomes impossible and hence one seeks for efficient variational families of states. Tensor networks are a prime example of such an ansatz class [10,13–17]. Despite their spectacular success in one dimension [18–29] as so-called matrix-product states [14,20,30], the most natural tensor network ansatz in two-dimensions, called projected entangled pair states (PEPS) [31], turned out to be burdened by a peculiar difficulty: even to calculate the normalization of PEPS is computationally intractable as has been shown in Ref. [32].

More precisely, the normalization or evaluation of a local expectation value within the PEPS ansatz class is a computational task which is complete for the complexity class #P, i.e., is as hard as any other problem in this class [33–35]. Paradigmatic #P problems consist in counting the solutions to decision problems which are complete for the class NP. Intuitively, counting the solutions to a hard problem can only be harder. Within the current state of knowledge in computer science the optimal runtime for NP-complete problems is unknown. However, it is widely conjectured that there are no algorithms with polynomial runtime solving any NP-complete problem. For the famous SAT-problem, there is even the *exponential-time hypothesis* [36], which conjectures that an exponential runtime is optimal for the problem.

Physically, one can invoke the Church-Turing-Deutsch principle [37] that interprets computations as physical processes. NP has been established to correspond to the cooling of spin glasses [38]. These materials are known to sometimes take an extremely long time to cool down. On the other hand, very many solid-state materials seem to cool down much faster. Indeed, insights in computer science suggest that the hardness of NP-complete problems lies in few tough instances with particularly rugged landscape. Phenomena like this are described in the framework of *average-case* complexity. While many NP-complete problems like 3-SAT are unlikely to be hard on average for uniform distributions [39], average-case hard problems are ubiquitous for the class #P. Recently, first examples directly relevant to demonstrating computational separation between classical and quantum devices have been pointed out [40,41].

There are several approaches for a rigorous theory of average-case complexity. Arguably the most natural is *random self-reducibility*, an immediate consequence of which is that a machine powerful enough to solve, e.g., three quarters of the instances would allow solving all instances. Thus it becomes implausible to find heuristic algorithms that solve significant numbers of instances as the self-reducibility structure would imply efficiency even for those instances that are particularly hard.

In this work, we provide strong complexity theoretical indications that the latter is not the case for generic PEPS due to a random self-reducibility structure that we uncover. This extends the worst-case #P-hardness result [32] to the average case and is an even more challenging obstruction to overcome. Technically, we make an extensive use of the recent insightful work in Ref. [41], where average-case hardness has been established in the context of quantum circuits, and we also employ some of the results established in Ref. [40]. Our main result is the following theorem.

Theorem 1 (Informal). Contracting a subset making up a $\frac{3}{4} + \frac{1}{\mathsf{poly}(N)}$ fraction of the instances drawn from an entrywise Gaussian distribution is #P-hard.

We explain in Sec. IV how this can be generalized to many probability distributions that satisfy an autocorrelation property. In particular a similar statement would hold if the local tensors are drawn from a uniform distribution supported on a bounded region, i.e., the overall "shape" is not crucial as long as the distribution is not infinitely peaked or has unusually broad tails. Firstly, this rules out the possibility that the computational hardness could be hidden in particular instances that are intractable, as it says that one could use the algorithm $\mathcal{O}$ to construct an algorithm $\mathcal{O}'$ that is efficient for all inputs. Secondly, it is important to note that Theorem 1 requires exact computation [32] but a different variant of Theorem 1 shows the following. Approximation up to errors of the form $2^{-\mathsf{poly}(N)}$ for N the system size, is also intractable on average, however, under stronger requirements on the algorithm $\mathcal{O}$. Our choice of the probability distribution is similar to that of Sec. 9.1 of Ref. [40], where the evaluation of the so-called *permanent* is considered which is also a #P-complete computational problem. Note that the result holds for arbitrary graphs as well, though the statement is trivial in one dimension [42].

In certain special instances fast algorithms might still be feasible. For example it is known that matrix-product states admit a polynomial time deterministic contraction algorithm [42]. However, even in two dimensions, this can happen under strong physical assumptions forcing the problem to admit a local structure [43,44]. Additionally, for certain subclasses some heuristic algorithms [42–63] (see Refs. [45,64] for reviews) yield results of practical importance [65–74]. Our average-case hardness result, however, suggests that these approaches could break down even for relevant PEPS instances as otherwise difficult computational problems would admit (quasi-) polynomial algorithms.

Physically, for disordered systems, one would expect any accurate ground state approximation by a PEPS to inherit the randomness of the Hamiltonian [75]. Hence in this setting, we provide evidence of intractability. Oftentimes, however, further physical assumptions are justified: While these completely generic PEPS are relevant for the study of strongly disordered systems, in many practically meaningful settings (in particular in the study of topological order), the relevant PEPS are translation-invariant. Remarkably, a worst-to-average case reduction as described in this paper works just as well for translation-invariant systems but we are unaware of a hardness result in the worst case for such systems.

II. DISCUSSION

Before we formalize the above in a rigorous setting, we discuss various aspects of this result.

A. Translation invariance

In many physical applications, e.g., in solid state materials and specifically in systems admitting topological order, the system of interest is translation-invariant. Hence, the data specifying the PEPS efficiently should reflect this symmetry and one would naturally set all local tensors to be equal. In this case, we do not know the corresponding computational problem to be #P-hard, for example the #P-hard instances in Ref. [32] are not translation-invariant. However, our worst-to-average case reduction works just as well in this special case, simply by choosing $(Q^{[v]})_v = (Q)_v$, where Q is drawn from the Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0,\sigma)^{D^4 d}$. The same argument and statement of the main theorem goes through. This leaves us with two mutually exclusive options: If the translation-invariant problem is hard for a complexity class C, then it follows that the problem is C-hard on average in the sense of our main theorem. If the problem is merely in P, then it is enough to find a heuristic for about 3/4 of the inputs to find a full randomized algorithm. On the other hand, if $C = \#\mathsf{P}$, then even the translation-invariant PEPS contraction problem would appear to be average-case intractable. We are unaware of random self-reducibility results for complexity classes other than #P. We thus expect a dichotomy: Either the translation-invariant problem is in P or it is #P-complete.

B. Evaluation precision

As far as we know, it is state of the art in computer science to prove random self-reducibility structures for problems given the promise that $\mathcal{O}$ works with at least exponential precision. In fact, we can improve our main theorem for this case too, at the cost of requiring $\mathcal{O}$ to function with a probability of $1 - \frac{1}{12N}$, where N denotes the system size. The reason for this trade-off is that subtleties arise in the technical steps, where the Berlekamp-Welch algorithm has to be replaced with a noise-resistant method. However, in the bigger picture, it does not seem possible to extend the seminal idea of Lipton to $\mathcal{O}$ working with lower precision. The reason is that th method crucially depends on the extrapolation of polynomials which is highly sensitive to noise. Related questions of precision relaxation are of interest in quantum information theory in the context of searching for quantum speed-ups. Here, certain precision relaxations are conjectured to be average-case hard as well [40,41].

C. Expectation values

The computational problem is concerned with PEPS contractions. The quantity that one computes is the norm of

the respective PEPS. However, in most physical applications the quantities of interest are expectation values of a local observable $\hat{A}$

$$\langle \hat{A} \rangle_\psi = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi / | \psi \rangle}, \tag{1}$$

where $|\psi\rangle$ refers to the PEPS specified by local tensors. Notice that this problem and its unnormalized version have both been proven to be #P-complete in Ref. [32] as well. For any algorithm that uses PEPS normalization as an intermediate step our main theorem is directly of interest and reflects the fundamental structure of the problem at hand. In the general case we can prove a worst-to-average result for this quantity as well. It is easy to see that our discussion of PEPS contraction carries over to the discussion of unnormalized expectation values. We show that a close analog of Theorem 1 holds for this quantity as well. The normalized expectation value is slightly more subtle in the following sense: the analog of the function q is not a polynomial but a rational function q/p where the degrees of both polynomials q and p are bounded by $2N$. We can then use interpolation for rational functions on enough sampling points to obtain the respective coefficients.

D. Implications on practical tensor network algorithms

The results found here have interesting implications to the performance of PEPS contraction algorithms aimed at solving condensed-matter problems [10,14,15]. There are three insights that are important in this respect. Firstly, the results laid out here relate average-case to worst-case complexity. In that, they apply to any tensor network contraction algorithm as the structure of random self-reducibility shows that if a given method $\mathcal{O}$ has trouble at less than a quarter of instances, these can in principle be treated with a small polynomial runtime overhead by our construction of the randomized algorithm $\mathcal{O}'$ (and, for that matter, our results also pertain to algorithms in P). Secondly, it is known that PEPS contraction algorithms often work well in practice for reasonable condensed-matter systems [45,64] which may seem at first sight at odds with the results presented here and in Ref. [32]. For this, one has to acknowledge that many important problems have additional structure that may render the PEPS contraction feasible. Specifically, it has been proven in Ref. [44] that local normalized expectation values of injective PEPS with *uniformly gapped parent Hamiltonian* can be evaluated in quasipolynomial time, i.e., faster than conjectured by the exponential-time hypothesis. Following up on this observation, it seems conceivable that one can devise PEPS algorithms that provide ground states of systems in a trivial phase (possibly even with convergence proofs), by making use of techniques of quasiadiabatic evolution [76,77], applying short circuits to product states as ground states of trivial parents. Having said that, any such approach would require keeping track of ground states of families of Hamiltonians. Thirdly, in most practical algorithms used in practice, in contrast, some initial condition for the PEPS is chosen, which is iteratively refined via sweeps, until a good convergence to the ground state is encountered. In fact, in practice, the PEPS data are initially often chosen randomly, following a refinement in sweeps by iteratively minimizing the energy evaluated from a local Hamiltonian. The results laid out here show that it is crucial to devise meaningful schemes making reasonable choices of these initial conditions. However, our average-case hardness results of PEPS contraction indicate that one should be particularly cautious when choosing such initial states.

III. PROBLEM SETTING

We now come to the technical section of this paper. In this section, we describe the problem in a rigorous setting.

A. Projected entangled pair states

Here we recall the definition of PEPS [52] and review the computational problem from Ref. [32] concerning the contraction of PEPS. We consider a family of graphs $G = (V, E)$ with $|V| = N$. Every vertex v stands for a local spin system described by a Hilbert space $\mathcal{H}_v := \mathbb{C}^d$. The physical Hilbert space is, thus, $\mathcal{H} := \mathcal{H}_v^{\otimes N} = (\mathbb{C}^d)^{\otimes N}$. In the projective construction of PEPS, one thinks of every edge $e \in E$ as a maximally entangled state $\sum_{i=1}^{D} |i\rangle|i\rangle$ in a virtual D-dimensional spin systems. A specific PEPS is then described by linear operators $P^{[v]} : \mathbb{C}^D \otimes \cdots \otimes \mathbb{C}^D \to \mathbb{C}^d$, where the number of copies of $\mathbb{C}^D$ is the number of adjacent edges for v. It is defined as the state vector in $\mathcal{H}$ resulting from the application of all $P^{[v]}$ for all $v \in V$. Note that by this the obtained PEPS is not necessarily normalized. The virtual dimension is assumed to satisfy $D = \mathsf{poly}(N)$ and is called *bond dimension*. In our discussion, it will be crucial to discriminate between the PEPS, which is a state vector in $\mathcal{H}$, and its specification $(P^{[v]})_v$. We will refer to the latter as *PEPS data*. A PEPS is called *translation-invariant* if the local tensors satisfy $P^{[v]} = P^{[w]} = P$ for all $v, w \in V$. These states have already been proven to be immensely useful in condensed matter research but the full regime of applicability is still open. Here, we assume open boundary conditions but our results carry over to the periodic case too.

B. PEPS evaluation

PEPS are described by polynomial data only. However, the physical problem we want to tackle remains notoriously difficult in that contraction of PEPS is computationally hard. This is needed for obtaining physical quantities of interest like expectation values of local observables. Specifically, the following computational tasks are the essential ingredients of PEPS contraction algorithms:

Problem 1 (PEPS-contraction). Input: A graph G and corresponding finite PEPS data $(P^{[v]})_v$ describing an unnormalized state $|\psi\rangle$ and with bond dimension $D = \mathsf{poly}(N)$.

Output: $\langle \psi | \psi \rangle$.

Problem 2 (PEPS-contraction:UEV). Input: The same input as in Problem 1 and additionally a local observable $\hat{A}$.

Output: $\langle \psi | \hat{A} | \psi \rangle$.

Problem 3 (PEPS-contraction:NEV). Input: The same input as in Problem 1 and additionally a local observable $\hat{A}$.

Output: $\langle \psi | \hat{A} | \psi \rangle / \langle \psi | \psi \rangle$.

It is one of the key insights in Ref. [32] that these problems are in fact #P-complete for the case that G is a square lattice. In the following, we recall the arguments leading to this observation. The construction uses measurement based quantum

computing [78–80]. Measurement based quantum computing based on cluster states performs a computation by initializing the cluster state on a square lattice and successively applying local sharp (projective) measurements to the local qubits. This is a universal model of a quantum computer and we can use it to encode any quantum circuit in a PEPS with polynomially bounded bond dimension. Notice first that the cluster state is a PEPS with bond dimension $D = 2$. However, the outcome of the quantum computation performed by the measurements depends on the random outcomes. This is dealt with by correcting the outcome with Pauli operators depending on the random outcomes. The PEPS encoding the quantum circuit is now obtained by applying an additional projector $|a\rangle\langle a|$, where a is the outcome that does not give rise to a nontrivial Pauli correction. Hardness follows from encoding the problem of counting solutions for a Boolean formula: Given a Boolean formula f, finding $\#_1(f) := |\{x, f(x) = 1\}|$ is #P-complete.

We prove all results for two canonical choices: The first is to draw entry-wise from a uniform distribution centered around zero and truncated at some chosen threshold σ, which we will denote by $\mathcal{U} = \mathcal{U}_{\mathbb{C}}(0, \sigma)$ and the product distribution by $\mathcal{P}_1 := \mathcal{U}^{D^4dN}$. Almost equivalently we could draw from a Gaussian distribution. We will denote this Gaussian distribution with $\mathcal{P}_2 := \mathcal{G}^{D^4dN} := \mathcal{N}_{\mathbb{C}}(0, \sigma)^{D^4dN}$. This is reminiscent to a discussion about the permanent with entries in the complex numbers in Sec. 9.1 of Ref. [40]. More precisely, we prove the following technical theorems.

Theorem 2 (Worst-to-average reduction). Suppose there exists a machine $\mathcal{O}$ that solves Problem 1 or 2 within precision $2^{-\mathsf{poly}(N)}$ for square lattices in polynomial time with a probability of $1 - \frac{1}{12N}$ over the instance drawn from $\mathcal{P}_i$ for $i = 1, 2$. Then, there exists a machine $\mathcal{O}'$ that solves any instance with precision $2^{-\mathsf{poly}(N)}$ of the respective problem in randomized polynomial time with exponentially high probability.

We will prove this theorem first, as it requires the most technical work. If we do not relax to exponential precision but require perfect arithmetical evaluation of the machine $\mathcal{O}$, we obtain a much stronger worst-to-average reduction:

Theorem 3 (Stronger worst-to-average reduction). Suppose it exists a machine $\mathcal{O}$ that solves Problem 1 or 2 exactly for square lattices in polynomial time with a probability of $\frac{3}{4} + \frac{1}{\mathsf{poly}N}$ drawn from $\mathcal{P}_i$, with $i = 1, 2$. Then, there exists a machine that solves any instance of the respective problem in randomized polynomial time with exponentially high precision.

Finally, requiring perfect evaluation, we obtain a worst-to-average reduction for the normalized expectation value problem as well:

Theorem 4 (Normalized expectation values). Suppose it exists a machine $\mathcal{O}$ that solves Problem 3 exactly for square lattices in polynomial time with a probability of $\frac{3}{4} + \frac{1}{\mathsf{poly}N}$ drawn from P_i with $i = 1, 2$. Then there exists a machine that solves any instance of the respective problem in randomized polynomial time with exponentially high precision.

C. Proof idea

There are several precise mathematical candidates for a definition of *average-case hardness*. We find that PEPS contraction is average-case hard in the same sense as certain combinatorial problems [40,81]: They admit a property called random self-reducibility. A problem is *randomly self-reducible* if the evaluation of any instance x can be reduced to the evaluation of random instances $y_1, \ldots, y_k$ with a bounded probability independent of the input. We will sketch how this is done for the permanent and PEPS giving the essential proof idea, see Ref. [41] for a particularly clear exposition in the context of quantum circuits. The complete argument can be found in Sec. IV.

In a seminal result, Ref. [81] has proven random self-reducibility for the evaluation of the permanent, a function that takes as an input a square matrix and outputs a number. The permanent of an $n \times n$ matrix A over a finite field is defined as the "determinant without signs":

$$\mathrm{perm}(A) := \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)}\,, \tag{2}$$

where S_n is the symmetric group. However, very unlike the determinant, the permanent turns out to yield a difficult combinatorial problem: Its evaluation has been proven to be #P-complete in Ref. [82]. The proof of random-self reducibility is rooted in the algebraic fact that the permanent defines a polynomial of degree n in the entries of its input matrix A. More precisely, the strategy is to take any (hard) instance A that we want to compute, draw a uniformly random matrix B and define

$$E(t) := A + tB\,, \tag{3}$$

for a parameter t in the finite field. Notice that $E(t)$ is uniformly random for any t because B is, even though $E(t)$ and $E(t')$ are correlated. The permanent of these matrices is a polynomial $q(t) := \mathrm{perm}(E(t))$ of degree n. Even if the algorithm $\mathcal{O}$ fails to accurately output $\mathrm{perm}(A)$ it will, by assumption, likely correctly evaluate $q(t_i)$ for a choice of t_i. The idea is to infer $q(0)$ from the values at $\{t_i\}$ via polynomial interpolation. We will explain this step in more detail in the next paragraph for the setting of PEPS.

We sketch how the worst to average-case reduction works for PEPS contractions. For a detailed and formal proof we refer to Sec. IV. A major difference to Lipton's result for the permanent is that we work over the complex numbers, for which there is no uniform distribution. Instead, we work with an entry-wise Gaussian distribution.

Intuitively, we scramble independently the individual tensors. Given a hard instance $(P^{[v]})_v$, we draw random PEPS data $(Q^{[v]})_v$ and define

$$(R(t)^{[v]})_v := t(P^{[v]})_v + (1-t)(Q^{[v]})_v. \tag{4}$$

Thus $(R(0)^{[v]})_v = (Q^{[v]})_v$ and $(R(1)^{[v]})_v = (P^{[v]})_v$. Notice that PEPS data and PEPS since the above definition has nothing to do with the addition of the corresponding states This choice of a scrambled operator is suitable for us because it allows us to deal with a subtlety arising from the fact that the PEPS data $(R(t)^{[v]})_v$ is not Gauss-random even though $(Q^{[v]})_v$ is. This is different to the setting of Ref. [81] but has been worked out for boson sampling [40], where it was shown that the difference is immaterial for small t. This carries over to our case as we discuss in Sec. IV.

IV. PROOFS

We can now provide rigorous proofs for Theorems 2–4.

A. Proof of Theorem 2

Before we turn to presenting the proof, we state a modification of Lemma 48 in Ref. [40]. Let us denote with $\mathcal{N}_{\mathbb{C}}(\mu,\sigma)$ the normal distribution over the complex numbers with mean μ and standard deviation σ. The lemma establishes that products of normal distributions with small mean are close to a product of the standard normal distribution with zero mean.

Lemma 5 (Autocorrelation of Gaussian distributions). For the distributions

$$\mathcal{D}_1 := \mathcal{N}_{\mathbb{R}}(0,(1-\varepsilon)^2\sigma)^M, \tag{5}$$

$$\mathcal{D}_2 := \prod_{i=1}^{M}\mathcal{N}_{\mathbb{R}}(v_i,\sigma) \tag{6}$$

with $v\in\mathbb{C}^M$, it holds that

$$||\mathcal{D}_1-\mathcal{N}_{\mathbb{R}}(0,\sigma)^M||\leqslant 2M\varepsilon, \tag{7}$$

$$||\mathcal{D}_2-\mathcal{N}_{\mathbb{R}}(0,\sigma)^M||\leqslant \frac{1}{\sigma}||v||_1, \tag{8}$$

where $||.||$ denotes the total variation distance and $v\in\mathbb{C}^M$. The same result holds if we substitute $\mathcal{N}$ with $\mathcal{U}$.

Proof of Lemma 5. We prove the lemma for the Gaussian case. The uniform can be obtained similarly. We obtain with the triangle inequality for the total variation distance:

$$||\mathcal{D}_1-\mathcal{G}^M||\leqslant M||\mathcal{N}_{\mathbb{R}}(0,(1-\varepsilon)^2\sigma)-\mathcal{N}_{\mathbb{R}}(0,\sigma)||. \tag{9}$$

With the relation between total variation distance and L^1 norm, we obtain

$$\begin{aligned}
&||\mathcal{D}_1-\mathcal{G}^M||\\
&\leqslant \frac{M}{2}\int_{-\infty}^{\infty}\left|\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{x^2}{2\sigma^2}}-\frac{1}{\sqrt{2\pi\sigma}(1-\varepsilon)}e^{-\frac{x^2}{2\sigma^2(1-\varepsilon)^2}}\right|dx\\
&=\frac{M}{2\sqrt{2\pi}\sigma(1-\varepsilon)}\int_{-\infty}^{\infty}\left|(1-\varepsilon)e^{-\frac{x^2}{2\sigma^2}}-e^{-\frac{x^2}{2\sigma^2(1-\varepsilon)^2}}\right|dx\\
&\leqslant\frac{M\varepsilon}{2\sqrt{2\pi}\sigma(1-\varepsilon)}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2\sigma^2}}\\
&\quad+\frac{M}{2\sqrt{2\pi}\sigma(1-\varepsilon)}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2\sigma^2}}-e^{-\frac{x^2}{2\sigma^2(1-\varepsilon)^2}}dx\\
&=\frac{M\varepsilon}{2(1-\varepsilon)}+\frac{M}{2(1-\varepsilon)}-\frac{M}{2}=\frac{M\varepsilon}{1-\varepsilon}\leqslant 2M\varepsilon.
\end{aligned} \tag{10}$$

The second inequality follows using again the triangle inequality:

$$\begin{aligned}
||\mathcal{D}_2-\mathcal{G}^M||&\leqslant\sum_{i=1}^{M}||\mathcal{N}_{\mathbb{R}}(v_i,\sigma)-\mathcal{N}_{\mathbb{R}}(0,\sigma)||\\
&=\sum_{i=1}^{M}\frac{1}{2}\int_{-\infty}^{\infty}\left|\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-v_i)^2}{2\sigma^2}}-\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}\right|dx\\
&=\sum_{i=1}^{M}\frac{1}{2\sqrt{2\pi}}\int_{-\infty}^{\infty}\left|e^{-\frac{(x-v_i/\sigma)^2}{2}}-e^{-\frac{x^2}{2}}\right|dx\\
&\leqslant\sum_{i=1}^{M}\frac{|v_i|}{\sigma}=\frac{||v||_1}{\sigma},
\end{aligned} \tag{11}$$

where the last inequality follows from a straightforward calculation. ■

Proof of Theorem 2. For simplicity, we set $\sigma=1$. Furthermore, we restrict to the case of Problem 1 as the proof for the case of Problem 2 is completely analogous. Consider Problem 1 and a hard instance defined by the data $(P^{[v]})_v$, e.g., the encoding of a Boolean function as was done in Ref. [32]. It suffices to consider a $(P^{[v]})_v$ with all matrix entries being bounded by 1 as all instances constructed in Ref. [32] admit this form. Furthermore, we draw PEPS data from the standard Gaussian distribution entrywise, denoted as $(Q^{[v]})_v\sim\mathcal{G}^{D^4dN}$. Analogously to Lipton [81], we define

$$(R(t)^{[v]})_v := t(P^{[v]})_v+(1-t)(Q^{[v]})_v. \tag{12}$$

Now, let $|\psi(t)\rangle$ denote the PEPS corresponding to these data. In analogy to the discussion of the permanent, we define the function $q(t):=\langle\psi(t)|\psi(t)\rangle$. Notice that this function is a polynomial in t with degree $r=2N$, which scales polynomially in the input length. Before we can apply Theorem 8, we have to deal with the fact that the $(R(t)^{[v]})_v$ are not distributed according to the Gaussian distribution. We will need only very small t bounded by some $\varepsilon>0$, such that the difference between the respective distributions is immaterial. Specifically, the $(R(t)^{[v]})_v$ tensors are distributed according to

$$\mathcal{D}=\prod_{i=1}^{D^4dN}\mathcal{N}_{\mathbb{C}}(tp_i,(1-t)^2). \tag{13}$$

Thus, from a triangle inequality and Lemma 5, we obtain

$$||\mathcal{D}-\mathcal{G}^{D^4dN}||\leqslant(4D^4dN+2D^4dN)\varepsilon=(6D^4dN)\varepsilon \tag{14}$$

for $|t|\leqslant\varepsilon$, by identifying $\mathbb{C}$ with $\mathbb{R}^2$. It will suffice to set

$$\varepsilon:=\frac{\delta}{6D^4dN} \tag{15}$$

and $\delta:=\frac{1}{12N}$. This implies that for a small enough inverse polynomial ε, we can make the total variation distance polynomially small. Let $\{t_i\}_{i\in[r+1]}$ be the set of $r+1$ equidistant points in $[0,\varepsilon]$. We will now use the assumption from the theorem's statement that the machine $\mathcal{O}$ works for a $1-\delta$ fraction of the instances drawn from $\mathcal{G}^{D^4dN}$. Using (14), we obtain for the success probability of the machine evaluating at the points t_i accurately up to within precision $2^{-\mathrm{poly}N}$

$$\begin{aligned}
&\Pr[|\mathcal{O}((R^{[v]})_v(t_i))-q(t_i)|\leqslant 2^{-\mathrm{poly}N}]\\
&\quad\geqslant 1-\delta-||\mathcal{D}-\mathcal{G}^{D^4dN}||\\
&\quad\geqslant 1-2\delta,
\end{aligned} \tag{16}$$

where we used that the total variation distance is an upper bound on the difference in probability the two distributions could possibly assign to an event.

Finally, we obtain the probability of $r+1$ consecutive successful evaluations as

$$\begin{aligned}&\Pr[|\{i \in [r+1], |\mathcal{O}(t_i) - q(t_i)| \leqslant 2^{-\mathsf{poly}N}\}| = r+1] \\ &= (1-2\delta)^{r+1} = \left(1 - \frac{1}{6N}\right)^{r+1} \\ &\geqslant 1 - \frac{2N+1}{6N} = \frac{2}{3} - \frac{1}{6N}, \end{aligned} \tag{17}$$

by Bernoulli's inequality. Here, we abbreviated $\mathcal{O}((R^{[v]})_v(t_i))$ with $\mathcal{O}(t_i)$. Given the evaluation values at the t_i, we can solve for the coefficients and obtain a polynomial $\tilde{q}$ which satisfies $|\tilde{q}(t_i) - q(t_i)| \leqslant 2^{-\mathsf{poly}N}$ for all t_i with high probability. The machine $\mathcal{O}'$ then evaluates $\tilde{q}(1)$, which is an estimate for $q(1) = \langle\psi/|\psi\rangle$.

To bound the error on this estimate we will use two powerful results: The first on noisy extrapolations and the second on noisy interpolations of polynomials. A version of the following lemma was proven in Ref. [83], see also Sec. 9.1 in Ref. [40].

Lemma 6 (Paturi). Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial of degree r and suppose $|p(x)| \leqslant \Delta$ for all x such that $|x| \leqslant \varepsilon$. Then, $|p(1)| \leqslant \Delta e^{2r(1+1/\varepsilon)}$.

The following result was proven in Rakhmanov [84].

Theorem 7 (Rakhmanov). Let E_k denote the set of k equidistant points in $(-1, 1)$. Then, for a polynomial $p : \mathbb{R} \to \mathbb{R}$ with degree r such that $|p(y)| \leqslant 1$ for all $y \in E_k$, it holds that

$$|p(x)| \leqslant C \log\left(\frac{\pi}{\arctan\left(\frac{k}{r}\sqrt{\mathcal{R}^2 - x^2}\right)}\right) \tag{18}$$

with

$$|x| \leqslant \mathcal{R} := \sqrt{1 - \frac{r^2}{k^2}}. \tag{19}$$

We will use the second result to bound the error between the points and then use the first result to bound the error on $\tilde{q}(1)$. For the proof, we shift the polynomial p such that the intervall of interest is centered around the origin. Furthermore, we can straightforwardly implement that we work with a smaller interval. We obtain that

$$\mathcal{R} = \sqrt{1 - \frac{r^2}{(r+1)^2}}\frac{\varepsilon}{2} = \sqrt{\frac{4N+1}{(2N+1)^2}}\frac{\varepsilon}{2}. \tag{20}$$

Restricting to the strict subinterval $[-\frac{\mathcal{R}}{2}, \frac{\mathcal{R}}{2}]$, we can apply Theorem 7 and obtain the following bound for all $t \in [-\frac{\mathcal{R}}{2}, \frac{\mathcal{R}}{2}]$,

$$\begin{aligned}|p(t)| &\leqslant 2^{-\mathsf{poly}N} C \ln\left(\frac{\pi}{\arctan\left(\frac{k}{r}\sqrt{\mathcal{R}^2 - x^2}\right)}\right) \\ &\leqslant 2^{-\mathsf{poly}N} C \ln\left(\frac{\pi}{\arctan(2\mathcal{R})}\right) \leqslant 2^{-\frac{1}{2}\mathsf{poly}N}. \end{aligned} \tag{21}$$

Finally, we can apply Lemma 6. This yields the desired bound on the difference between the estimate $\tilde{q}(1)$ and the actual value $q(1)$:

$$\begin{aligned}|\tilde{q}(1) - q(1)| = |p(1)| &\leqslant 2^{-\frac{1}{2}\mathsf{poly}N + 4\log_2(e)N(1+2/\mathcal{R})} \\ &= 2^{-\mathsf{poly}'N} \end{aligned} \tag{22}$$

for a sufficiently large poly. Finally, we remark that the success probability can be exponentially amplified by repeating the above procedure polynomially many times because of the Chernoff bound. ■

B. Proof of Theorem 3

The superior bound in Theorem 3 follows from the fact that we can invoke the Berlekamp-Welch algorithm in the interpolation step. The latter is a provably correct algorithm for the interpolation of polynomials due to Ref. [85]. Compare also Bouland *et al.* [41].

Theorem 8 (Berlekamp-Welch [85]). Let q be a degree-r polynomial over any field $\mathbb{F}$. Suppose we are given k pairs of elements $\{(x_i, y_i)\}_{i=1,\ldots,k}$ with all x_i distinct with the promise that $y_i = q(x_i)$ for at least $\max(r+1, (k+r)/2)$ points. Then, one can recover q exactly in $\mathsf{poly}(k, r)$ deterministic time.

As explained in Sec. III C, we arrive at a polynomial $q(t) = \langle\psi(t)/|\psi(t)\rangle$ of degree $r = 2N$. Instead of $r+1$ queries to the machine $\mathcal{O}$, we query it $k = \mathsf{poly}(N)$ times. Berlekamp-Welch requires that at least $\frac{k+r}{2}$ of obtained k data points are correct in order to reconstruct the polynomial. We furthermore assume that $k > r$. From Markov's inequality and the union bound, we obtain

$$\begin{aligned}&\Pr\left[|\{i, \mathcal{O}(t_i) = q(t_i)\}| \geqslant \frac{k+r}{2}\right] \geqslant 1 - \frac{2\mathbb{E}}{k-r} \\ &\geqslant 1 - \frac{2\left(\frac{1}{4} - \frac{1}{\mathsf{poly}N}\right)k}{k-r} = 1 - \frac{k}{2(k-r)} + \frac{2k}{\mathsf{poly}(N)(k-r)} \\ &= \frac{1}{2} - \frac{r}{2(k-r)} + \frac{2k}{\mathsf{poly}(N)(k-r)}, \end{aligned} \tag{23}$$

where we abbreviate the expectation value in question with $\mathbb{E}$. Thus, by choosing k polynomially large, we obtain an expression that is polynomially close to $1/2$. Again, by repeating the procedure a polynomial number of times and taking a majority vote we can amplify this probability exponentially. With this probability, the Berlekamp-Welch algorithm outputs q exactly and we can simply evaluate $q(1)$ without having to worry about the error of extrapolation. It seems appropriate to point out that we are in fact not drawing data from the Gaussian distribution in this case but from a discrete analog of it. However, this does not change the details of our analysis.

C. Proof of Theorem 4

We know that the function we are interested in can be described by the quotient of two polynomials of degree at most $r = 2N$. This leaves us with $4N+1$ unknown coefficients. There is an equivalent of the Berlekamp-Welch algorithm for rational functions [86]. Invoking this algorithm, the proof proceeds analogously to the proof of Theorem 3.

V. EXPONENTIAL DEPENDENCE ON PEPS DATA

The argument in the main text emphasizes the demanding precision that is required when specifying the PEPS data. In this section, we stress that this is not merely done for complexity-theoretic reasons: A pair of states can be defined by very similar PEPS data, while their norms can be vastly

different. In fact, to specify the norm of a PEPS, one needs exponential precision in the PEPS data, as a moment of thought reveals. This is already true in one spatial dimension for matrix product states. Take $D = 2$, $d = 2$, an a translation-invariant open boundary condition MPS, so that the vertex set V is that of N sites, E reflecting nearest neighbor interactions. The linear operators $P^{[v]} = P$ are for all v defined by

$$P^{[v]} = \sum_{i=1,2} \sum_{\alpha,\beta=1}^{D} A[i]_{\alpha,\beta} |i\rangle\langle\alpha, \beta|, \tag{24}$$

where for the state vector $|\psi\rangle$ we take

$$A[0] := \mathrm{diag}(1, 0), \quad A[1] := \mathrm{diag}(0, 1). \tag{25}$$

The boundary conditions are taken open, as in the main text, and fixed by vector $|0\rangle$ and the respective dual. Obviously, this is a representation of the product $|0, \ldots, 0\rangle$ with norm $\langle\psi/|\psi\rangle = 1$. For $|\phi\rangle$, we choose

$$B[0] := \mathrm{diag}(1, 0), \quad B[1] := \mathrm{diag}(\eta, 1), \tag{26}$$

with the same boundary conditions, for some $\eta > 0$. It is still straightforward to compute the norm, invoking the transfer operator

$$\mathbb{E} := B[0] \otimes B^*[0] + B[1] \otimes B^*[1] = \mathrm{diag}(1 + \eta^2, \eta, \eta, 1). \tag{27}$$

This gives

$$\langle\phi/|\phi\rangle = \langle 0|\mathbb{E}^N|0\rangle = (1 + \eta^2)^N. \tag{28}$$

Clearly, for the two states to feature norms that are the same up to a constant, an in N exponentially small $\eta > 0$ is required. In fact, even for a bond dimension $D = 1$ one could have come to a similar conclusion. However, $|\psi\rangle$ and $|\phi\rangle$ are even vastly different in their entanglement properties, the latter featuring an entanglement entropy of a symmetrically bisected chain that is extensive in N.

VI. OUTLOOK

In this work, we presented the first average-case complexity result in the context of quantum many-body systems, specifically tensor network states. Our main result is structural, namely we prove that the hard instances of PEPS contraction make up a significant fraction of all instances. Physically, this means that contraction of PEPS with random tensors is likely to be computationally hard to accurately evaluate. Conceptually, we establish structural similarities to the evaluation of the permanent. Our results hold under the assumption of accurate or exponential precision. In Sec. V, we stress that also on physical grounds, to demand exponential precision is very much reasonable. However, in a physical context it is often sufficient to evaluate observables up to polynomial precision. The major *open problem* is thus to extend the presented analysis to this case. For PEPS contractions establishing such a result would have direct practical implications. Furthermore, we are not aware of any #P-completeness result for translation-invariant PEPS. Thus the general *open question* should be: what are the instances of PEPS for which known contraction methods have convergence guarantees? It is our hope that further research at the interface between computer science and quantum many-body physics will provide exciting insights to this question.

ACKNOWLEDGMENTS

We thank Adam Bouland, Bill Fefferman, Juani Bermejo-Vega, Christopher Chubb, Ingo Roth, Alexander Nietner, Augustine Kshetrimayum, Frederik Hahn, and Radu Curticapean for valuable discussions and comments. In particular we are grateful to Adam Bouland and Bill Fefferman for pointing out to us the result of Rakhmanov [84]. This work was supported by the DFG (EI 519/14-1, EI 519/7-1, EI 519/15-1, and CRC 183 B1), the ERC (TAQ), and the Templeton Foundation.

[1] A. Altland and B. D. Simons, *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, UK, 2010).

[2] Y. V. Nazarov and Y. M. Blanter, *Quantum Transport: Introduction to Nanoscience* (Cambridge University Press, Cambridge, UK, 2009).

[3] A. Acin, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, The European quantum technologies roadmap, New J. Phys. **20**, 080201 (2018).

[4] F. Weber, Strange quark matter and compact stars, Prog. Part. Nucl. Phys. **54**, 193 (2005).

[5] I. Arsene, I. Bearden, D. Beavis, C. Besliu, B. Budick, H. Bøggild, C. Chasman, C. Christensen, P. Christiansen, J. Cibor *et al.*, Quark-gluon plasma and color glass condensate at RHIC? The perspective from the BRAHMS experiment, Nucl. Phys. A **757**, 1 (2005).

[6] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. **68**, 13 (1996).

[7] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Continuous-time Monte Carlo methods for quantum impurity models, Rev. Mod. Phys. **83**, 349 (2011).

[8] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Functional renormalization group approach to correlated fermion systems, Rev. Mod. Phys. **84**, 299 (2012).

[9] U. Schollwöck, The density-matrix renormalization group, Rev. Mod. Phys. **77**, 259 (2005).

[10] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[11] F. Alet and N. Laflorencie, Many-body localization: An introduction and selected topics, C. R. Phys. **19**, 498 (2018).

[12] T. DeGrand and C. DeTar, *Lattice Methods for Quantum Chromodynamics* (World Scientific, Singapore, 2006).

[13] M. Fannes, B. Nachtergaele, and R. F. Werner, Finitely correlated states on quantum spin chains, Commun. Math. Phys. **144**, 443 (1992).

[14] F. Verstraete, V. Murg, and J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization

group methods for quantum spin systems, Adv. Phys. **57**, 143 (2008).
[15] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, Ann. Phys. **349**, 117 (2014).
[16] J. C. Bridgeman and C. T. Chubb, Hand-waving and interpretive dance: an introductory course on tensor networks, J. Phys. A **50**, 223001 (2017).
[17] J. Eisert, M. Cramer, and M. B. Plenio, Area laws for the entanglement entropy, Rev. Mod. Phys. **82**, 277 (2010).
[18] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems, Phys. Rev. Lett. **93**, 207204 (2004).
[19] D. Gobert, C. Kollath, U. Schollwöck, and G. Schütz, Real-time dynamics in spin-1/2 chains with adaptive time-dependent density matrix renormalization group, Phys. Rev. E **71**, 036102 (2005).
[20] F. Verstraete and J. I. Cirac, Matrix product states represent ground states faithfully, Phys. Rev. B **73**, 094423 (2006).
[21] S. R. Manmana, S. Wessel, R. M. Noack, and A. Muramatsu, Strongly Correlated Fermions After a Quantum Quench, Phys. Rev. Lett. **98**, 210405 (2007).
[22] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas, Nat. Phys. **8**, 325 (2012).
[23] T. Fukuhara, A. Kantian, M. Endres, M. Cheneau, P. Schauß, S. Hild, D. Bellem, U. Schollwöck, T. Giamarchi, C. Gross *et al.*, Quantum dynamics of a mobile spin impurity, Nat. Phys. **9**, 235 (2013).
[24] F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, Entanglement spectrum of a topological phase in one dimension, Phys. Rev. B **81**, 064439 (2010).
[25] J. H. Bardarson, F. Pollmann, and J. E. Moore, Unbounded Growth of Entanglement in Models of Many-Body Localization, Phys. Rev. Lett. **109**, 017202 (2012).
[26] C. Karrasch, J. H. Bardarson, and J. E. Moore, Finite-Temperature Dynamical Density Matrix Renormalization Group and the Drude Weight of Spin-1/2 Chains, Phys. Rev. Lett. **108**, 227206 (2012).
[27] T. Prosen, Open XXZ Spin Chain: Nonequilibrium Steady State and a Strict Bound on Ballistic Transport, Phys. Rev. Lett. **106**, 217206 (2011).
[28] J. P. Ronzheimer, M. Schreiber, S. Braun, S. S. Hodgman, S. Langer, I. P. McCulloch, F. Heidrich-Meisner, I. Bloch, and U. Schneider, Expansion Dynamics of Interacting Bosons in Homogeneous Lattices in One and Two Dimensions, Phys. Rev. Lett. **110**, 205301 (2013).
[29] I. Arad, Z. Landau, U. Vazirani, and T. Vidick, Rigorous RG algorithms and area laws for low energy eigenstates in 1D, Commun. Math. Phys. **356**, 65 (2017).
[30] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, Quantum Inf. Comput. **7**, 401 (2007).
[31] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States, Phys. Rev. Lett. **96**, 220601 (2006).
[32] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, The Computational Complexity of PEPS, Phys. Rev. Lett. **98**, 140506 (2007).
[33] S. Arora and B. Barak, *Computational Complexity: A Modern Approach* (Cambridge University Press, Cambridge, UK, 2009).
[34] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2002).
[35] S. Aaronson, *Quantum Computing Since Democritus* (Cambridge University Press, Cambridge, UK, 2013).
[36] D. Lokshtanov, D. Marx, and S. Saurabh, Lower bounds based on the exponential time hypothesis, Bull. Eur. Assoc. Theor. Comput. Sci. EATCS **105**, 41 (2011).
[37] D. Deutsch, Quantum theory, the Church–Turing principle and the universal quantum computer, Proc. R. Soc. London A **400**, 97 (1985).
[38] F. Barahona, On the computational complexity of Ising spin glass models, J. Phys. A **15**, 3241 (1982).
[39] J. Feigenbaum and L. Fortnow, On the random-self-reducibility of complete sets, in *Proceedings of the Sixth Annual Structure in Complexity Theory Conference* (IEEE, Piscataway, NJ, 1991).
[40] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theor. Comput. **9**, 143 (2013).
[41] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, Quantum supremacy and the complexity of random circuit sampling, Nat. Phys. **15**, 159 (2019).
[42] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, Phys. Rev. A **74**, 022320 (2006).
[43] A. Anshu, I. Arad, and A. Jain, How local is the information in tensor networks of matrix product states or projected entangled pairs states, Phys. Rev. B **94**, 195143 (2016).
[44] M. Schwarz, O. Buerschaper, and J. Eisert, Approximating local observables on projected entangled pair states, Phys. Rev. A **95**, 060102(R) (2017).
[45] M. Lubasch, J. I. Cirac, and M.-C. Banuls, Unifying projected entangled pair states contractions, New J. Phys. **16**, 033014 (2014).
[46] M. Levin and C. P. Nave, Tensor Renormalization Group Approach to Two-Dimensional Classical Lattice Models, Phys. Rev. Lett. **99**, 120601 (2007).
[47] Z.-C. Gu and X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order, Phys. Rev. B **80**, 155131 (2009).
[48] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Second Renormalization of Tensor-Network States, Phys. Rev. Lett. **103**, 160601 (2009).
[49] H. H. Zhao, Z. Y. Xie, Q. N. Chen, Z. C. Wei, J. W. Cai, and T. Xiang, Renormalization of tensor-network states, Phys. Rev. B **81**, 174411 (2010).
[50] S. Yang, Z.-C. Gu, and X.-G. Wen, Loop Optimization for Tensor Network Renormalization, Phys. Rev. Lett. **118**, 110504 (2017).
[51] G. Evenbly, Algorithms for tensor network renormalization, Phys. Rev. B **95**, 045117 (2017).
[52] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, arXiv:cond-mat/0407066.

[53] R. Orús and G. Vidal, Infinite time-evolving block decimation algorithm beyond unitary evolution, Phys. Rev. B **78**, 155117 (2008).

[54] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Gradient methods for variational optimization of projected entangled-pair states, Phys. Rev. B **94**, 155123 (2016).

[55] M. T. Fishman, L. Vanderstraeten, V. Zauner-Stauber, J. Haegeman, and F. Verstraete, Faster methods for contracting infinite 2D tensor networks, Phys. Rev. B **98**, 235148 (2018).

[56] T. Nishino, Density matrix renormalization group method for 2D classical models, J. Phys. Soc. Jpn. **64**, 3598 (1995).

[57] T. Nishino, K. Okunishi, and M. Kikuchi, Numerical renormalization group at criticality, Phys. Lett. A **213**, 69 (1996).

[58] T. Nishino and K. Okunishi, Corner transfer matrix algorithm for classical renormalization group, J. Phys. Soc. Jpn. **66**, 3040 (1997).

[59] R. Orús and G. Vidal, Simulation of two-dimensional quantum systems on an infinite lattice revisited: Corner transfer matrix for tensor contraction, Phys. Rev. B **80**, 094403 (2009).

[60] R. Orús, Exploring corner transfer matrices and corner tensors for the classical simulation of quantum lattice systems, Phys. Rev. B **85**, 205117 (2012).

[61] A. Kshetrimayum, M. Rizzi, J. Eisert, and R. Orús, A Tensor Network Annealing Algorithm for Two-Dimensional Thermal States, Phys. Rev. Lett. **122**, 070502 (2019).

[62] I. Markov and Y. Shi, Simulating quantum computation by contracting tensor networks, SIAM J. Comp. **38**, 963 (2008).

[63] I. Arad and Z. Landau, Quantum computation and the evaluation of tensor networks, SIAM J. Comp. **39**, 3089 (2010).

[64] S.-J. Ran, E. Tirrito, C. Peng, X. Chen, L. Tagliacozzo, G. Su, and M. Lewenstein, *Tensor Network Contractions: Methods and Applications to Quantum Many-Body Systems*, Lecture Notes in Physics, Vol. 964 (Springer, 2000).

[65] P. Corboz, R. Orús, B. Bauer, and G. Vidal, Simulation of strongly correlated fermions in two spatial dimensions with fermionic projected entangled-pair states, Phys. Rev. B **81**, 165104 (2010).

[66] P. Corboz, S. R. White, G. Vidal, and M. Troyer, Stripes in the two-dimensional t-J model with infinite projected entangled-pair states, Phys. Rev. B **84**, 041108(R) (2011).

[67] I. Niesen and P. Corboz, A tensor network study of the complete ground state phase diagram of the spin-1 bilinear-biquadratic Heisenberg model on the square lattice, SciPost Phys. **3**, 030 (2017).

[68] Y. H. Matsuda, N. Abe, S. Takeyama, H. Kageyama, P. Corboz, A. Honecker, S. R. Manmana, G. R. Foltin, K. P. Schmidt, and F. Mila, Magnetization of $SrCu_2(BO_3)_2$ in Ultrahigh Magnetic Fields up to 118 T, Phys. Rev. Lett. **111**, 137204 (2013).

[69] A. Kshetrimayum, T. Picot, R. Orús, and D. Poilblanc, Spin-1/2 kagome XXZ model in a field: Competition between lattice nematic and solid orders, Phys. Rev. B **94**, 235146 (2016).

[70] S. Yan, D. A. Huse, and S. R. White, Spin-liquid ground state of the S = 1/2 Kagome Heisenberg antiferromagnet, Science **332**, 1173 (2011).

[71] D. Poilblanc, J. I. Cirac, and N. Schuch, Chiral topological spin liquids with projected entangled pair states, Phys. Rev. B **91**, 224431 (2015).

[72] T. Picot, M. Ziegler, R. Orús, and D. Poilblanc, Spin-S kagome quantum antiferromagnets in a field with tensor networks, Phys. Rev. B **93**, 060407(R) (2016).

[73] H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang, Gapless Spin-Liquid Ground State in the $S = 1/2$ Kagome Antiferromagnet, Phys. Rev. Lett. **118**, 137202 (2017).

[74] J.-Y. Chen, L. Vanderstraeten, S. Capponi, and D. Poilblanc, Non-Abelian chiral spin liquid in a quantum antiferromagnet revealed by an iPEPS study, Phys. Rev. B **98**, 184409 (2018).

[75] T. B. Wahl, A. Pal, and S. H. Simon, Efficient Representation of Fully Many-Body Localized Systems Using Tensor Networks, Phys. Rev. X **7**, 021018 (2017).

[76] S. Bachmann, S. Michalakis, B. Nachtergaele, and R. Sims, Automorphic equivalence within gapped phases of quantum lattice systems, Commun. Math. Phys. **309**, 835 (2012).

[77] M. B. Hastings, Quasi-adiabatic continuation for disordered systems: Applications to correlations, Lieb-Schultz-Mattis, and Hall conductance, arXiv:1001.5280.

[78] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer, Phys. Rev. Lett. **86**, 5188 (2001).

[79] R. Raussendorf, D. E. Browne, and H. J. Briegel, Measurement-based quantum computation with cluster states, Phys. Rev. A **68**, 022312 (2003).

[80] D. Gross and J. Eisert, Novel Schemes for Measurement-Based Quantum Computation, Phys. Rev. Lett. **98**, 220503 (2007).

[81] R. Lipton, *New Directions in Testing: Distributed Computing and Cryprography* (American Mathematical Society, Providence, RI, 1991).

[82] L. G. Valiant, Quantum cicuits that can be simulated classically in polynomial time, SIAM Comput. **31**, 1229 (1979).

[83] R. Paturi, On the degree of polynomials that approximate symmetric Boolean functions, in *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing* (ACM Press, New York, 1992), pp. 468–474.

[84] E. A. Rakhmanov, Bounds for polynomials with a unit discrete norm, Ann. Math. **165**, 55 (2007).

[85] L. Welch and E. Berlekamp, Error correction for algbraic block codes, US Patent, 4,633,470 (1986).

[86] R. Movassagh, Efficient unitary paths and quantum computational supremacy: A proof of average-case hardness of Random Circuit Sampling, arXiv:1810.04681.