

Exponential Thermal Tensor Network Approach for Quantum Lattice Models

Bin-Bin Chen,¹ Lei Chen,¹ Ziyu Chen,¹ Wei Li,^{1,2,*} and Andreas Weichselbaum^{3,4,†}

¹*Department of Physics, Key Laboratory of Micro-Nano Measurement-Manipulation and Physics (Ministry of Education), Beihang University, Beijing 100191, China*

²*International Research Institute of Multidisciplinary Science, Beihang University, Beijing 100191, China*

³*Department of Condensed Matter Physics and Materials Science, Brookhaven National Laboratory, Upton, New York 11973-5000, USA*

⁴*Physics Department, Arnold Sommerfeld Center for Theoretical Physics, and Center for NanoScience, Ludwig-Maximilians-Universität, Theresienstrasse 37, 80333 Munich, Germany*



(Received 30 December 2017; revised manuscript received 12 August 2018; published 26 September 2018)

We speed up thermal simulations of quantum many-body systems in both one- (1D) and two-dimensional (2D) models in an exponential way by iteratively projecting the thermal density matrix $\hat{\rho} = e^{-\beta\hat{H}}$ onto itself. We refer to this scheme of doubling β in each step of the imaginary time evolution as the exponential tensor renormalization group (XTRG). This approach is in stark contrast to conventional Trotter-Suzuki-type methods which evolve $\hat{\rho}$ on a linear quasicontinuous grid in inverse temperature $\beta \equiv 1/T$. As an aside, the large steps in XTRG allow one to swiftly jump across finite-temperature phase transitions, i.e., without the need to resolve each singularly expensive phase-transition point right away, e.g., when interested in low-energy behavior. A fine temperature resolution can be obtained, nevertheless, by using interleaved temperature grids. In general, XTRG can reach low temperatures exponentially fast and, thus, not only saves computational time but also merits better accuracy due to significantly fewer truncation steps. For similar reasons, we also find that the series expansion thermal tensor network approach benefits in both efficiency and precision, from the logarithmic temperature scale setup. We work in an (effective) 1D setting exploiting matrix product operators (MPOs), which allows us to fully and uniquely implement non-Abelian and Abelian symmetries to greatly enhance numerical performance. We use our XTRG machinery to explore the thermal properties of Heisenberg models on 1D chains and 2D square and triangular lattices down to low temperatures approaching ground-state properties. The entanglement properties, as well as the renormalization-group flow of entanglement spectra in MPOs, are discussed, where logarithmic entropies (approximately $\ln \beta$) are shown in both spin chains and square-lattice models with gapless towers of states. We also reveal that XTRG can be employed to accurately simulate the Heisenberg XXZ model on the square lattice which undergoes a thermal phase transition. We determine its critical temperature based on thermal physical observables, as well as entanglement measures. Overall, we demonstrate that XTRG provides an elegant, versatile, and highly competitive approach to explore thermal properties, including finite-temperature thermal phase transitions as well as the different ordering tendencies at various temperature scales for frustrated systems.

DOI: [10.1103/PhysRevX.8.031082](https://doi.org/10.1103/PhysRevX.8.031082)

Subject Areas: Computational Physics,
Condensed Matter Physics

I. INTRODUCTION

Efficient simulations of interacting quantum many-body systems are crucial for a better understanding of correlated materials. In particular, the accurate

computation of thermodynamic quantities including magnetization, heat capacity, magnetic susceptibility, etc., enables a direct comparison to experiments and helps to identify relevant microscopic models. The exotic quantum matter includes Luttinger liquids in one (1D) [1,2] and spin-liquid materials in two dimensions (2D) [3–8]. Besides, the exploration and understanding of the rich and diverse behavior of quantum many-body physics at different energy or, equivalently, temperature scales are interesting from a theoretical perspective. One example are thermal states near a 1D quantum critical point which show universal entropy in the partition function due to emergent conformal symmetry [9–11] in the low-energy regime.

*w.li@buaa.edu.cn

†weichselbaum@bnl.gov

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Another prominent example is the thermal fractionalization in the honeycomb Kitaev spin liquid (KSL) at finite temperature [4], which has been experimentally observed [12] in the proximate KSL material α -RuCl₃ [13,14].

At first glance, the simulation of thermal many-body states seems a task more than challenging. There exist exponentially many excited states in the energy spectrum, many of which possess volume-law entanglement and deny any efficient representation in classical computers. However, it turns out that the ensemble density operator, say, $e^{-\beta H}$ with $\beta \equiv 1/T$ being the inverse temperature, can be efficiently expressed and manipulated in terms of thermal tensor network (TTN) states. The matrix product operator (MPO) is a very natural TTN for describing 1D quantum systems at finite temperature [i.e., (1+1)D], due to the “entanglement” area law in thermal states of both gapped and gapless systems with local interactions. Intuitively, thermal fluctuation effectively “opens” an excitation gap and introduces a finite correlation length in mixed states, rendering an area law in terms of total correlation [15] (as well as operator space entanglement [16]). However, it was estimated that the required MPO bond dimension has an upper bound scaling as $D \sim e^\beta$ [17] which still seems to pose a severe barrier towards obtaining low- T properties.

Nevertheless, on the other hand, various renormalization-group algorithms have been proposed to accurately compute thermodynamics in the (1+1)D problems, practically even down to very low temperatures. These methods include the transfer matrix renormalization group (TMRG) [18–20] and finite-temperature density matrix renormalization group (DMRG) [21] which are based on the traditional DMRG, tensor network algorithms such as the linearized tensor renormalization group (LTRG) [22–24], and the variational projected entangled pair operator method in (2+1)D [25–27]. Besides, a combination of finite-temperature DMRG and Monte Carlo samplings called minimally entangled typical thermal states (METTS) was proposed [28,29] and recently generalized to (2+1)D [30]. The success of these algorithms in (1+1)D, and partly in (2+1)D, strongly suggests that D does not necessarily scale exponentially with β .

To estimate the computational cost in thermal simulations, one can introduce a formal entanglement entropy in the TTN, e.g., in the MPO representation of the mixed state density matrix, as introduced in Refs. [16,31]. It has been revealed that this MPO entanglement saturates for gapped systems and scales logarithmically (as $(c/3)\ln\beta$) for quantum critical spin chains [16,31]. Very recently, two independent works [32,33] deployed conformal field theory (CFT) arguments to show on general grounds that the Renyi entropy of thermal states of effective 1D systems scales as $S_E^{(n)} \sim (c/6)[1 + (1/n)]\ln\beta$. In the limit $n \rightarrow 1$, this implies that also the von Neumann entropy scales like $S_E \sim (c/3)\ln\beta$ for thermal states in 1D on general grounds.

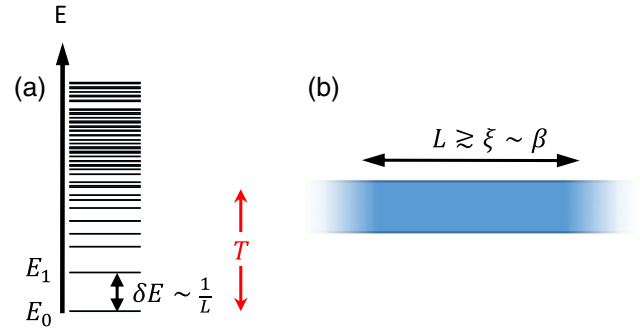


FIG. 1. (a) Finite-size spectra E_s with low-energy level spacing $\delta E \sim 1/L$. By requiring $T > \delta E$ for thermal averaging, this suggests $L \gtrsim \beta$, and therefore a thermal correlation length $\xi \sim \beta$. (b) A large or infinite system has an effective thermal cutoff $\xi \sim \beta$ in system length when measuring local observables. Therefore, provided $\xi \lesssim L$, finite systems can be used, as a very good approximation, to simulate thermal properties in the thermodynamic limit.

Intuitively, this scaling can be understood simply by considering finite-size spectra with many-body low-energy level spacing $\delta E \sim 1/L$, as schematically depicted in Fig. 1. In order to sample a thermal average, it must hold that at the very least $T \gtrsim \delta E$, or equivalently $L \gtrsim \beta$ [in other words, finite β introduces an effective cutoff $\xi \sim \beta$ of system size; see Fig. 1(b)]. Now, given that low-energy states violate a strict area law via logarithmic corrections, one has the block entropy $S_E \simeq (c/6)\ln L + \text{const}$ [34] for individual low-energy states using open boundary conditions (OBCs). By choosing $L = a\beta$ with fixed constant $a \gtrsim 1$, and by going from individual low-energy pure states $|s\rangle$ to a thermal state with weights ρ_s , i.e., $|s\rangle \rightarrow \sum_s \rho_s |s\rangle\langle s|$, the block entropy for the outer product $|s\rangle\langle s|$ acquires another factor of 2. Thermal averaging does not change this scaling due to the subadditivity of the von Neumann entropy (see Appendix A), given the constraint $L = a\beta$ with fixed a . By thermal averaging over a similar set of low-energy states, the MPO block entropy at the center of the system saturates by further increasing $L \gg a\beta$ at a finite value, i.e., is cut off by

$$S_E[\rho(\beta)] \sim \frac{c}{3} \log \beta + \text{const} \quad (1)$$

and, importantly, becomes independent of L . This block entanglement entropy of the thermal state scales similar versus β to the block entropy of a ground-state calculation versus L using periodic boundary conditions (PBCs).

The above intuitive argument fits the holographic picture in terms of the thermal multiscale entanglement renormalization ansatz (MERA) [35], where the minimal surface (of half the system) in thermal MERA, as well as the corresponding (bipartite) entanglement entropy, is argued to be proportional to $\ln\beta$ [36].

Furthermore, the argument of translating the scaling of the entropy in $\ln L$ to that of $\ln \beta$ is completely consistent with the notion within CFT that β and L are equivalent directions connected via a modular \mathcal{S} transformation. This has direct consequences for the conformal TTN framework in $(1+1)\text{D}$, i.e., with one spatial (horizontal) and one imaginary time (vertical) axis. The horizontal transfer matrix $e^{-\tau H}$ across different temperatures has the ground state of H as its dominant eigenvector, which thus contains logarithmic entanglement. For thermodynamics of an infinite-size quantum chain ($L \gg \beta \gg 1$), we therefore expect that the vertical transfer matrix (across different real-space sites) also has a dominant eigenvector with entanglement entropy $S_E \sim \ln(\beta)$. In addition, however, by definition of the partition function, it acquires *intrinsic* PBC in the direction of temperature, which therefore doubles the prefactor in entanglement entropy scaling, in agreement with the earlier arguments.

This logarithmic growth of entropy versus β provides a tight upper bound in efficient thermal simulations [32]. This, together with the constant entanglement for gapped systems, suggests that the bond dimension D only needs to scale algebraically (constantly) as β increases for critical (noncritical) quantum chains, respectively.

Conversely, it directly follows from the above logarithmic scaling of the entanglement entropy that β needs to change significantly on a *relative* and not an absolute scale, in order to see a sizable effect on the entanglement entropy. This suggests for simulations that the numerical grid in β should be *logarithmically* discretized. In particular, as depicted in Fig. 2, by doubling $\beta \rightarrow 2\beta$, one can therefore design an exponentially faster cooling procedure, in contrast to current standard simulation techniques, which linearly evolve the full density matrix [18–24] or the typical sampling states [28–30] in imaginary time. We note that essentially a similar, even though much more involved, strategy was pursued in Refs. [26,27,37]. Their approach was based on a dimensional reduction via a nested contraction of linear Trotter gates, followed by a variational optimization of coarse-graining transformations [38]. In contrast, our approach does not rely on Trotter gates and, hence, is straightforwardly applicable to arbitrary Hamiltonians (1D and 2D). Overall, it represents an extremely simple yet also very efficient approach.

In this work, inspired by the logarithmic MPO entanglement entropy, we propose a one-way exponential tensor renormalization-group (XTRG) scheme along imaginary time. Interestingly, this scheme allows one to draw parallels to the concept of energy scales in the numerical renormalization group (NRG) [39–41]. There also, with every new iteration, the energy scale is reduced by a factor, typically $\gtrsim \sqrt{2}$. Consequently, this also zooms into the low-energy regime in an exponential fashion while dealing only with a very manageable linear number of iterations.

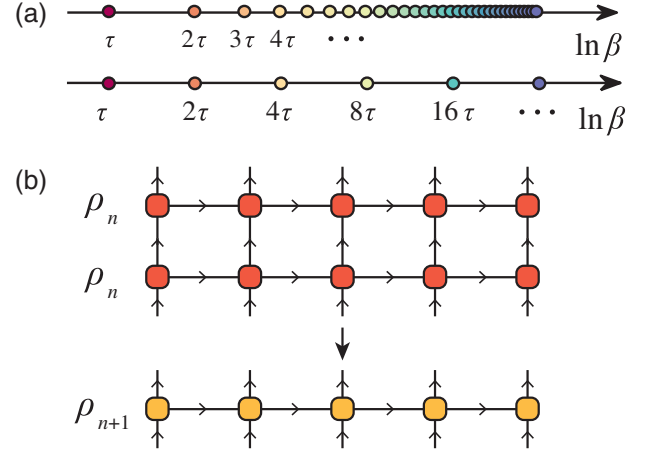


FIG. 2. (a) Linear versus logarithmic temperature scale employed in thermal simulations. (b) A single step in XTRG evolution by projecting MPO ρ_n (at $\beta = 2^n \tau$) to itself. Following common notation, tensor networks are graphically depicted by blocks (i.e., tensors) connected by lines which are to be contracted. Here, vertical lines indicate physical state spaces, whereas horizontal lines indicate virtual or bond state spaces. The exploitation of symmetries, quite generally, mandates directed lines; hence, each line carries an arrow.

We benchmark our results with conventional linear evolution schemes. The results show that, by following the entanglement structure and exploiting the logarithmic temperature scale, one can obtain more accurate results with less cost. By implementing non-Abelian symmetries in the MPO, we can even simulate 2D clusters down to low temperatures with high precision and investigate thermodynamics and related entanglement properties.

The model systems considered here are (anisotropic) spin-half Heisenberg XXZ models

$$H = J \sum_{\langle i,j \rangle} (S_i^x S_j^x + S_i^y S_j^y + \Delta S_i^z S_j^z), \quad (2)$$

both in 1D Heisenberg chains ($\Delta = 1$) of length L , as well as in the 2D square lattice ($\Delta = 1$ and 5) for systems of width W and length L , and thus with a total of $N = WL$ sites, using open (OBC) as well as periodic (PBC) boundary conditions. We include only nearest-neighbor couplings as indicated by the sum $\langle \cdot, \cdot \rangle$. For the purpose of benchmarking, we also consider the spin-half XY chain with $J_z = 0$, i.e.,

$$H = J \sum_{\langle i,j \rangle} (S_i^x S_j^x + S_i^y S_j^y), \quad (3)$$

as it can be mapped to a fermionic tight-binding chain. The XXZ model in Eq. (2) possesses a $U(1)$ symmetry, which restores a larger $SU(2)$ symmetry when $\Delta = 1$. Symmetries, whether non-Abelian or Abelian, are fully exploited,

throughout. We also set $J := 1$ as the unit of energy, unless specified otherwise. Furthermore, we use units $k_B = \hbar = 1$.

The rest of the paper is organized as follows. In Sec. II, we introduce the XTRG scheme with symmetries implemented, as well as an improved series-expansion thermal tensor network (SETTN) method [42] based on a pointwise Taylor expansion algorithm that also exploits the logarithmic temperature scale. The performances of these methods in the simulations of both 1D and 2D quantum many-body systems are presented and compared in Sec. III. In Sec. IV, the entanglement properties of MPOs are investigated, where logarithmic entanglement entropies versus β in the Heisenberg chain and also square-lattice models are discussed. XTRG is also employed to study the finite-temperature phase transition of the 2D Heisenberg XXZ model, where we demonstrate that XTRG can accurately pinpoint the critical temperature.

II. SYMMETRIC THERMAL TENSOR NETWORKS IN LOGARITHMIC TEMPERATURE SCALE

By construction, a thermal density matrix $\rho = e^{-\beta H}$ is a scalar operator and, thus, shares exactly the same symmetries as the Hamiltonian. For example, symmetries are preserved for Trotter-Suzuki-type TTNs [22,23], where every local tensor (storing Boltzmann weights) is symmetric. Similarly, in the series-expansion TTNs, it is also clear that arbitrary H^n 's have exactly the same symmetry as H (any unitary symmetry transformation that leaves H intact also leaves H^n intact), and so does the resulting tensor network representation of $\rho(\beta)$.

Concepts such as spontaneous symmetry breaking apply to individual low-energy (eigen)states but not to a thermal state. Hence, the full exploitation of all symmetries of the Hamiltonian, Abelian and non-Abelian alike, are very natural in XTRG. For finite systems with an open boundary, in particular, XTRG shares the same benefits in efficiency as DMRG in quasi-1D. There is a notable difference, however: As long as the thermal correlation length is clearly smaller than the system size under consideration, local thermal properties in the center of the system can be regarded as in the thermodynamic limit.

A. Symmetric matrix product operator

The explicit implementation of non-Abelian symmetries has been regarded as a standard technique in ground-state DMRG simulations ($T = 0$) [43], which has many important applications including exploring quantum spin liquids in frustrated quantum magnets [44,45], and is also shown to be useful in METTS-type thermal simulations [46,47]. However, the implementations of non-Abelian symmetries in MPOs for finite-temperature simulations are still absent. Here, by virtue of the flexible and versatile QSpace framework [48], we implement non-Abelian SU(2), as

well as Abelian U(1), symmetry in the MPO algorithm and, thus, realize a very efficient thermal renormalization-group (RG) algorithm that can also be applied to 2D problems.

In our MPO-based thermal algorithm, we start by constructing an SU(2)-invariant MPO representation of H . As this construction involves reduced matrix elements in the spirit of the Wigner-Eckart theorem [48], we refer to it as the reduced MPO in contrast to the full MPO when not exploiting non-Abelian symmetries. By switching from a state-based to a multiplet-based description, we can reduce the overall bond dimension from D states to $D^* < D$ multiplets. The reduced MPO representation of H can be constructed by the automata method [49–51] for 1D Hamiltonians and the MPO sum-and-compress scheme [52] for more complicated 2D lattice models. For a Heisenberg chain with nearest-neighbor interactions, a full MPO requires $D_H = 5$ bond states. As these correspond to two singlets and one triplet, i.e., $\underline{1}^2 \oplus \underline{3}^1$, where \underline{d}^n specifies n multiplets of dimension d each, the reduced SU(2)-invariant MPO involves only $D_H^* = 3$ multiplets. For the Heisenberg model on a 2D square lattice, we map a system of width W to a 1D snakelike chain with “long-range” interactions (up to a distance of $2W - 1$). Then, e.g., using OBCs, the full MPO requires $D_H = 3W + 2$ bond states, while the reduced SU(2)-invariant MPO has a significantly more compact representation with only $D_H^* = W + 2$ multiplets ($\underline{1}^2 \oplus \underline{3}^W$). More details on the symmetric MPO representation of total Hamiltonian can be found in Appendix B.

The computational cost in a tensor network algorithm typically scales with some power $\mathcal{O}(D^m)$ aside from other factors concerning number of sites, etc., where for the MPO structure of this paper we encounter $m = [3, \dots, 6]$. By exploiting non-Abelian symmetries, the computational cost can be effectively reduced to $\mathcal{O}((D^*)^m)$, which leads to a gain in numerical efficiency by $\mathcal{O}((D^*/D)^m)$. For a single SU(2) symmetry, it roughly holds, on average, that $D/D^* \sim 3\dots 4$ for spin-1/2 systems. Note also that multiplet dimensions are typically somewhat larger in thermal MPOs as compared to matrix product ground states, which renders us even greater numerical gain of symmetry implementation. The underlying reason for this difference is that an MPO has two physical indices associated with the same site. Therefore, their direct product already also leads to an enlarged effective local spin. With this reasoning in mind, for the sake of readability, we generally quote estimates in numerical efficiency in terms of D , since, after all, $\mathcal{O}(D^m) = \mathcal{O}((D^*)^m)$ with the overall scale factor $(D^*/D)^m \ll 1$ absorbed into the definition of $\mathcal{O}(\ast)$.

B. Exponential tensor renormalization group

For one-dimensional critical systems, the entanglement entropy in the MPO of a thermal state diverges only logarithmically in β . Therefore, to see a sizable effect in

the properties of a thermal state, β must change significantly on a relative and not an absolute scale. For example, a change $\beta \rightarrow a\beta \rightarrow a^2\beta \rightarrow \dots$ with some constant $a > 1$ will change the entanglement by linear increments. This strongly suggests to scale β on a logarithmic and not on a linear scale.

We can take full advantage of this scaling by a novel approach, which we refer to as the exponential tensor renormalization group (XTRG) to simulate quantum many-body systems at finite temperatures with high efficiency and accuracy. We start by preparing an MPO of the (unnormalized) thermal state $\rho(\tau) = e^{-\tau H}$ at exponentially small τ , i.e., at a very high temperature. Then, we can proceed to cool down the system exponentially by multiplying the thermal state with itself,

$$\rho_0 \equiv \rho(\tau) \rightarrow \rho_1 \equiv \rho(\tau) \cdot \rho(\tau) = \rho(2\tau). \quad (4)$$

Feeding the last MPO iteratively into the next step, with $\tau_n \equiv 2^n \tau$ and therefore $\tau_0 \equiv \tau$, we obtain

$$\rho_n \equiv \rho(\tau_n) = \rho_{n-1} \cdot \rho_{n-1}. \quad (5)$$

This approach directly implies an exponential acceleration to reach low temperatures.

Importantly, in the present XTRG scheme, we can easily start from exponentially small τ . For example, for $\tau J = 10^{-3}$ with $J := 1$ the largest local energy scale here given by the Heisenberg coupling strength, we can use an efficient series expansion scheme [cf. Sec. II C]. For τ as small as 10^{-6} , even the simplest lowest-order linear expansion of $e^{-\tau H}$ can suffice, which extremely simplifies initialization even for longer-ranged Hamiltonians which become cumbersome for Trotter-like decompositions or for 2D Hamiltonians in the effective 1D MPO setup. In the latter setup, with minor modifications, the MPO of the Hamiltonian already encodes the essential structure of the thermal state using the *same* bond dimension. A detailed comparison of different initialization strategies, including the Trotter-Suzuki decomposition, SETTN, and simple linear initialization for small τ_0 , is provided in Appendix C.

Given an MPO representation for $\rho(\tau_n = \beta/2)$, we can compute the (unnormalized) thermal state at temperature $T = 1/\beta$ via $\rho(\beta) = \rho(\tau_n)^\dagger \rho(\tau_n)$, i.e., by contracting $\rho(\tau_n)$ with its conjugate. This guarantees positivity of the thermal state $\rho(\beta)$ even in the presence of truncation of the MPO for $\rho(\tau_n)$. Furthermore, we can also compute the partition function at $\beta = 2\tau_n$ via $\mathcal{Z}(\beta) = \text{Tr}[\rho(\tau_n)^\dagger \rho(\tau_n)]$ and, thus, gain another factor of 2 to reach lower temperatures. The latter can be simply obtained by computing the Frobenius norm squared of $\rho(\tau_n)$. Not incidentally, many of the features above are directly related with common procedures within the setup of a purified thermal state [18–22,24,26,27,53].

In case the grid of inverse temperature values is too sparse, intermediate values can be easily obtained by shifting the initial value of

$$\tau_0 \rightarrow \tau_0 \cdot 2^z \quad \text{with } z \in [0, 1), \quad (6a)$$

a procedure that is entirely analogous to z -shifts within the NRG. In order to obtain a uniform logarithmic grid over n_z shifts, one may simply choose

$$z_i = \frac{i}{n_z} \quad \text{with } i = 0, \dots, n_z - 1. \quad (6b)$$

Different “ z -shifts” can be computed completely independently from each other and can therefore be efficiently parallelized. Truncation errors are still kept minimal by moving to large β as quickly as possible in an accurate manner. Alternatively, one can obtain intermediate values of β also by computing $\rho_{n,n'} \equiv \rho_{n-1} \cdot \rho_{n'-1}$ for various $n' \leq n$.

C. Series expansion thermal tensor networks

SETTN is a “continuous-time” RG approach for the accurate simulation of quantum lattice models at a finite temperature [42]. By exploiting the series expansion of the density matrix in Eq. (7), SETTN is essentially free of discretization errors, making it distinct from previous Trotter-Suzuki-type RG methods including TMRG [19,20], finite-temperature DMRG [21], LTRG [22–24], and METTS [28], etc. The efficient MPO representations of H^n is the key for the algorithm to work, and both OBC and PBC chain systems can be equally well dealt with in SETTN (here, specifically, we simply use one long-range bond for the simulation of PBCs). Being free of Trotter errors, SETTN has better controllable and uniformly higher accuracy, compared to conventional thermal RG methods.

To initialize ρ_0 for small τ_0 , a series expansion yields [cf. Fig. 3(a)]

$$\rho_0 \equiv \rho(\tau_0) \simeq \sum_{k=0}^{\mathcal{N}_c} \frac{(-\tau_0)^k}{k!} H^k. \quad (7a)$$

The required cutoff order of the expansion is $\mathcal{N}_c \sim N\tau_0$, i.e., proportional to the total number of sites N . In practice, \mathcal{N}_c is determined automatically by allowing only a negligibly small expansion error ($< 10^{-15}$). Therefore, for sufficiently small τ_0 , the initialization of $\rho(\tau_0)$ above is well controlled and accurate, typically resulting in $\mathcal{N}_c \lesssim 10$.

The high-temperature Maclaurin expansion in Eq. (7a) can be employed not only in the initialization stage but also for simulating low-temperature thermal states, as shown in Ref. [42]. Despite its competitive performance, this method still leaves room for further improvement. Since Eq. (7a)

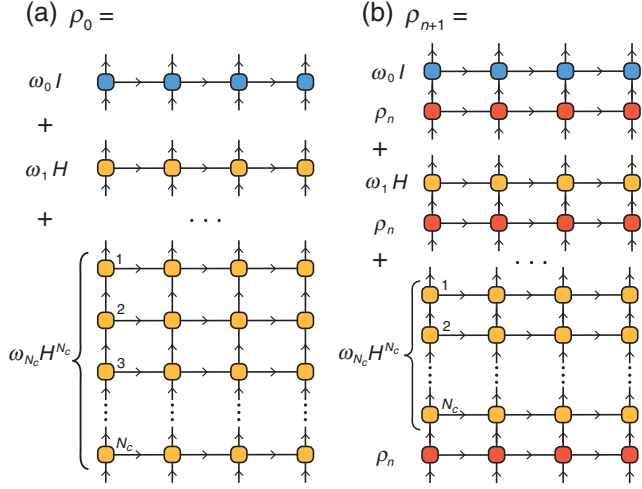


FIG. 3. (a) SETTN initialization of ρ_0 (with $\beta = \tau_0$) using Maclaurin expansion with coefficients $\omega_k \equiv (-\tau_0)^k/k!$ [see Eq. (7a)]. (b) Pointwise SETTN algorithm exploiting the logarithmic β scale, here with coefficients $\omega_k \equiv (\tau_n - \tau_{n+1})^k/k!$ [see Eq. (7b)].

expands ρ around the infinitely high temperature, i.e., $\beta = 0$, the power series in H^n involves large $\mathcal{N}_c \propto N\beta$ for large system size N and low temperature $1/\beta$. The precision of SETTN is limited by the truncation in H^n [see Sec. II D], which generally increases as n increases [42]. In this sense, a pointwise Taylor expansion can help reduce the expansion order \mathcal{N}_c and improve the accuracy, i.e.,

$$\rho(\tau_{n+1}) \simeq \left(\sum_{k=0}^{\mathcal{N}_c} \frac{(\tau_n - \tau_{n+1})^k}{k!} H^k \right) \underbrace{e^{-\tau_n H}}_{\equiv \rho_n}. \quad (7b)$$

Equation (7b) expands the density operator around an arbitrary but fixed τ_n . For generality, the initialization in Eq. (7a) may be viewed as iteration $n = -1$ having $\rho_{-1} = \mathbb{I}$ and $\tau_{-1} = 0$.

Now, given the density operator $\rho(\tau_0)$ obtained by initialization via Eq. (7a), $\rho(a\tau_0)$ with $a > 1$ can be obtained via Taylor expansion around $\rho(\tau_0)$. In particular, this also hold for $a = 2$, which thus may serve as a complementary scheme to the XTRG above. For example, alternative to the plain doubling scheme above, SETTN may be employed to cool down the system and obtain the MPO form of density operators at the inverse temperature grid τ_n . Given ρ_n , the MPO representation of ρ_{n+1} can be expanded as in Eq. (7b). Each term in the summation there can be obtained iteratively by projecting H onto $(H^{n-1}\rho_n)$ and compressing the product. For the overall sum then [cf. Fig. 3(b)], we also employ variational optimization (see Appendix D) to finally arrive at the MPO representation of $\rho_{n+1} \equiv \rho(\tau_{n+1})$.

By repeating this procedure, we also can follow the XTRG protocol to cool down the system along the inverse temperature grid $\tau_n = 2^n \tau_0$. In contrast to the plain doubling

scheme in XTRG, however, in the case of SETTN the step size $\delta\tau \equiv \tau_{n+1} - \tau_n$ can be chosen continuously. In this sense, SETTN is more flexible, as it permits the flexible exploration of thermal properties in the immediate vicinity of temperature τ_n with only a modest cost.

Note that for this *improved* SETTN, as we will refer to it, using an exponentially increasing τ_n series does not acquire exponential acceleration as XTRG does, since in the case of SETTN one still needs to perform projection and compression operations $\mathcal{N}_c \propto \beta N$ times. Nevertheless, from the point of view of SETTN, the exponential τ_n series is computationally preferable to, say, a linear τ_n series as expansion points, since the former can reduce expansion overhead and thus save computational time, in practice, without losing any accuracy (see Appendix E for a detailed comparison).

D. MPO compression and numerical cost

In SETTN, we start with a reduced SU(2)-invariant MPO for H . Then, we iteratively apply the projections H onto $(H^{k-1}\rho_n)$ to obtain $H^k\rho_n$, with Eq. (7) represented by $\rho_{n-1} = \mathbb{I}$. These projections need to be combined with a compression algorithm to reduce the numerical cost in a controlled manner. In the present context, however, truncation by discarded weight is dangerous, since small weights for small τ_0 can affect the accuracy for large τ_n . Hence, we truncate by number of multiplets, throughout. For this, we introduce the control parameters $D_{n,k}^*$, which stand for the maximum number of multiplets D^* to be kept in the k th iterative term when computing ρ_n . For simplicity, we set this parameter constant, i.e., $D_n^* \equiv D_{n,k}^*$, which also stands for the bond dimension of the target state ρ_n . Furthermore, we choose constant $D^* \equiv D_{n>0}^*$ but, for the sake of the analysis, may use a different value for D_0^* for the initialization in Eq. (7a) if specified.

For an extremely small τ_0 (say, as small as 10^{-4} to 10^{-8}), the initialization of $\rho(\tau_0) = e^{-\tau_0 H}$ can be simplified to lowest order, i.e., linear expansion $\rho(\tau_0) \simeq 1 - \tau_0 H$. Having $\mathcal{N}_c = 1$ in Eq. (7a), the result shares the same bond dimension $D_0^* = D_H^*$ as H itself. In contrast, when expanding around finite τ_n as in Eq. (7b), the bond dimension D^* in $\rho(\tau_n)$ typically needs to grow significantly and, therefore, is fixed to some specified $D^* \gg D_0^*$.

The compression of the SETTN projections above can be achieved either by a singular value decomposition (SVD) technique quite similar to that in Ref. [42], apart from the fact that the MPOs here have SU(2) symmetry, or by a variational optimization which can greatly improve numerical efficiency (see Appendix D for more details on related MPO compression techniques). Within SETTN, the cost of either compression scheme scales like $\mathcal{O}(D^3)$. We tested both and found comparable numerical accuracy. Finally, we variationally add up the MPOs for $H^k\rho_n$ with coefficients as in Eqs. (7) to obtain ρ_{n+1} (cf. Appendix D 2).

TABLE I. Time complexity of thermal tensor network methods for a lattice of length L and width W , i.e., a total of $N \equiv WL$ sites, assuming $\beta > 1$ (by default, β and τ are in units of $1/J$).

Methods	XTRG	SETTN	LTRG
Complexity	$\mathcal{O}(\ln(\beta)ND^4)$	$\mathcal{O}(\beta N^2 D^3 D_H)$	$\mathcal{O}(\frac{\beta}{\tau} ND^3 W)$
Relative cost q^a	1	$\mathcal{O}(\frac{\beta N}{D \ln \beta} D_H)$	$\mathcal{O}(\frac{\beta/\tau}{D \ln \beta} W)$

^aXTRG cost is set as the time unit.

In contrast, XTRG projects ρ_n onto itself in Eq. (5). So both MPOs involved have large bond dimension $D^* \gg D_0^*$, D_H^* . Here, a direct SVD compression is numerically costly, $\mathcal{O}(D^6)$, to be compared to $\mathcal{O}(D^3)$ for SETTN. For the XTRG iteration in Eq. (5), we therefore constrain ourselves to a variational compression which scales like $\mathcal{O}(D^4)$ (see Appendix D 1).

More explicitly, we summarize in Table I the time costs of the three algorithms involved in the current discussions. The numerical cost of SETTN for an entire run up to inverse temperature β scales as $\mathcal{O}(\beta N^2 D^3 D_H)$ assuming $\mathcal{N}_c \propto \beta N$ for large β , N with $\beta = \tau_n$ in Eqs. (7), whereas XTRG scales as $\mathcal{O}(\ln(\beta)ND^4)$. We can thus estimate the

relative run time of SETTN over XTRG as $q_S \equiv [(\beta N)/(D \ln \beta)]D_H$. For practical simulations as in Figs. 4 and 5, we find that XTRG calculations are faster than SETTN by more than one order of magnitude. In 1D critical systems, since the required bond dimension scales as $D \sim e^S \sim \beta^c$ ($\lambda \lesssim 1$ for $c = 1$ CFTs, say, spin-1/2 Heisenberg chain; see Ref. [32]). Thus, $q_S \gtrsim [N/(\ln \beta)]D_H$, with $q_S \gg 1$ for N large and $\beta > 1$ (in units of $1/J$). Similarly, also Trotter-Suzuki-type linear thermal RG methods, like the finite-temperature DMRG [21] and LTRG [22,24], with scaling $\mathcal{O}[(\beta/\tau)ND^3]$ (last column in Table I with $W = 1$), are typically much slower by a factor $q_L \simeq [1/(\tau \ln \beta)] \gg 1$ as compared to XTRG.

It is also revealing to compare the efficiency of XTRG with the currently most efficient scheme in 2D systems, i.e., Trotter-Suzuki decomposition plus swap gates [30]. The numerical (time) cost of the latter scheme scales like $\mathcal{O}[(\beta/\tau)ND^3 W]$, where the additional factor W stems from the number of required swap gates which is proportional to the width W . For 2D, however, typically $D \gg \beta$. Therefore, the relative cost of XTRG scales like $q_{2D} \equiv [(\tau D)/(W \beta \ln \beta)]$, resulting in $q_{2D} \sim \mathcal{O}(1)$ [e.g., with $W = 8$, $\tau = 0.05$, $\beta = 50$, and $D^* \sim 2000$ (correspondingly, $D \simeq 8 \times 10^3 - 10^4$) in SU(2) simulations or $D \sim 2000$ in U(1)

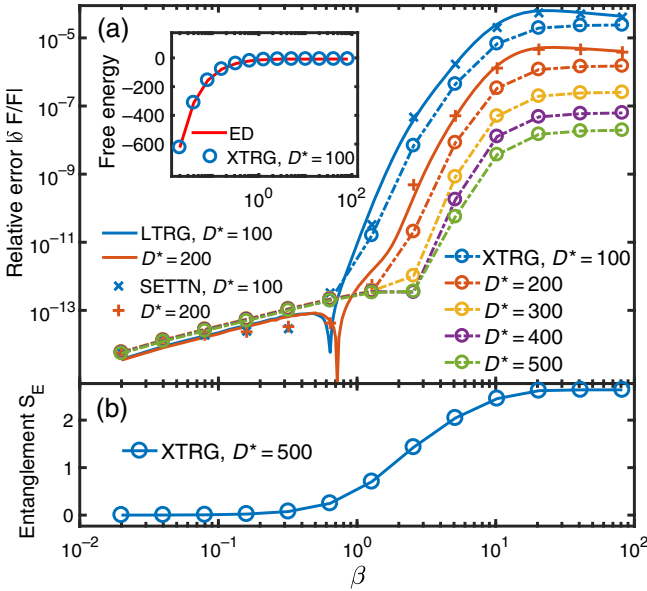


FIG. 4. (a) Relative errors of free energy of an $L = 18$ spin-1/2 Heisenberg chain (PBC), calculated by TTN algorithms including LTRG, XTRG, and SETTN relative to ED, with SU(2) symmetry implemented, throughout. Data for different methods share the same color for the same D^* . Here, $\rho(\tau_0 = 0.01)$ is initialized using SETTN. (b) Bipartite entanglement entropy S_E in the middle of the MPO, which increases monotonically. The time costs of the LTRG calculations are 8 (10) times as long as SETTN and 180 (113) times those of XTRG, in $D^* = 100$ (200) calculations, respectively.

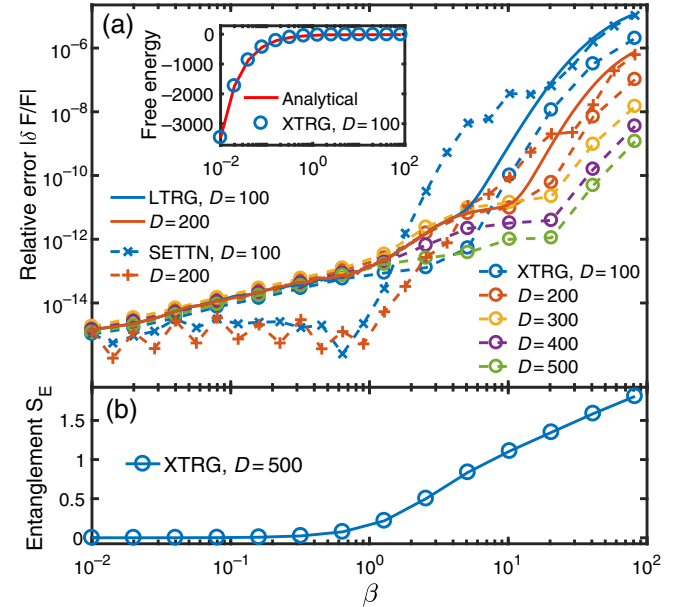


FIG. 5. (a) Relative errors of free energy in an $L = 50$ spin-1/2 XY chain with OBCs, calculated by TTN algorithms including LTRG, XTRG, and SETTN, with U(1) symmetry encoded, relative to the analytical solution. Similar representation as in Fig. 4, otherwise. Here, $\rho(\tau_0 = 0.01)$ is again also initialized using SETTN. (b) Entanglement entropy in the middle of the system. Overall, time costs of LTRG are 7 (7.4) times that of the SETTN run and 363 (183) times that of XTRG, for bond dimension $D = 100$ (200), respectively.

calculations, one obtains $q_{2D} \sim 0.98$]. Nevertheless, XTRG is still clearly advantageous over Trotter gates due to the far fewer truncation steps involved. Besides, XTRG can be simply and efficiently parallelized based on z -shifts [cf. Eq. (6)]. For thermal simulations that are dominated by Trotter error and swap gates in LTRG schemes to reach low temperatures, and not necessarily by the truncation error due to entanglement growth, XTRG may be crucial to reach the lowest temperature scales, e.g., in systems with more than one well-separated physical energy scale. As a very interesting example, in the Kitaev honeycomb model, the gauge flux excitation peak in the specific heat curve locates at a very low temperature, to be referred to as T_l , compared to the high-temperature peak at T_h . To see the T_l peak in the specific heat, one needs to cool down the system till extremely low temperatures ($T_l/T_h \lesssim 10^{-2}$, depending on the coupling constants and boundary conditions [4]) where the exponential acceleration in XTRG can play a very important role. A similar scenario is observed in the 2D triangular Heisenberg lattice, which we discuss in more detail below.

III. BENCHMARK CALCULATIONS: 1D AND 2D HEISENBERG MODELS AT FINITE TEMPERATURE

In this section, we benchmark our XTRG starting from well-understood models such as the 1D Heisenberg chain (Sec. III B) and the 2D square Heisenberg lattice (Sec. III C). We then proceed towards the much less understood and thus much less trivial finite-temperature behaviors of the triangular lattice Heisenberg model (Sec. III D), where, quite generally, the strong local frustration represents a huge challenge to numerical simulations.

A. Thermodynamic quantities

Here, we briefly summarize the thermodynamic quantities that are computed and analyzed in detail below. An equilibrium thermal state is described by the partition function $\mathcal{Z}(\beta) \equiv \text{Tr}[e^{-\beta H}] \equiv \text{Tr}[\rho(\beta)]$. Typical interesting thermodynamic quantities, which constitute important tasks for the TTN algorithms to compute, include

$$f \equiv \frac{F}{N} = -\frac{1}{N\beta} \ln \mathcal{Z}(\beta) \quad \text{free energy}, \quad (8a)$$

$$u \equiv \frac{E}{N} = \frac{\partial(\beta f)}{\partial \beta} = \frac{1}{N} \frac{\text{Tr}[H \cdot \rho(\beta)]}{\mathcal{Z}(\beta)} \quad \text{(internal) energy}, \quad (8b)$$

$$c_V = \frac{\partial u}{\partial T} = -\beta^2 \frac{\partial u}{\partial \beta} \quad \text{specific heat}, \quad (8c)$$

etc. The computation of the free energy f and energy density u is straightforward, where H and $\rho(\beta)$ are expressed as MPOs and the calculations amount to efficient

contractions of tensor networks consisting of these MPOs. The linear derivatives in β for the specific heat as well as the energy density, however, are not very natural for XTRG, which obtains the thermal data on a uniform logarithmic β grid. Therefore, it is more suitable to use $[\partial/(\partial\beta)] = \{\partial/[\beta\partial(\ln\beta)]\}$, i.e.,

$$u = \frac{1}{\beta} \frac{\partial(\beta f)}{\partial \ln \beta}, \quad (9a)$$

$$c_V = -\beta \frac{\partial u}{\partial \ln \beta}, \quad (9b)$$

instead. This is also more stable numerically for small temperatures, since the quotient of numerical differences is divided by T for the specific heat in Eq. (9b) and not by T^2 as in Eq. (8c), and is multiplied by T for the internal energy in Eq. (9). This formula is used to compute the specific heat in Figs. 6 and 7. In order to reduce numerical differential errors, independent calculations with slightly different initial τ values are run in parallel, e.g., using $n_z = 16$ in Eq. (6b), which produces interleaved data points with $\delta z = 1/16$, i.e., $\delta \ln \beta = \delta z \ln 2 \simeq 0.0433$.

Finally, we note that the LTRG approach adopted in this work, e.g., for the data in Figs. 4 and 5 below, is streamlined with the remainder of the TTN procedures used in this work. It differs from the original LTRG algorithm in Refs. [22,24] in that it successively projects the MPO for $\rho(\tau_0)$ to the density operator $\rho(\beta)$ to increase β linearly.

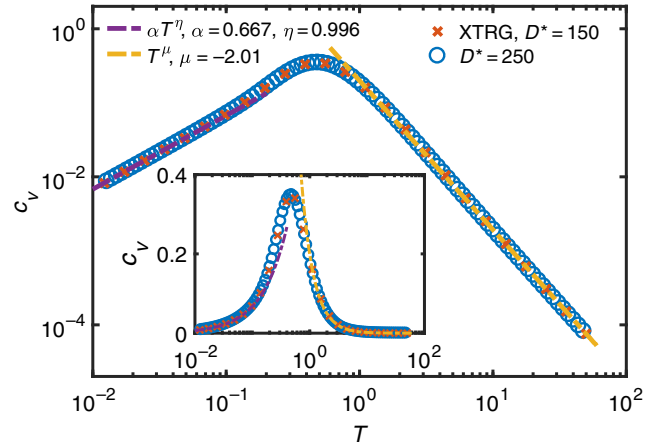


FIG. 6. Specific heat of a Heisenberg chain of length $L = 300$, with up to $D^* = 250$ multiplets retained. For low temperatures $T \ll 1$, a universal linear behavior versus T is observed, i.e., $c_V = \alpha T^\eta$ with fitted exponent $\eta = 0.996$ and slope $\alpha \simeq 2/3$, in the regime $T \leq 0.025$. Exploiting the fact that $\alpha = [(\pi c)/(3v)]$ ($v = \pi/2$ for a spin-1/2 Heisenberg chain), we extract the central charge $c \simeq 1$. For large T , the specific heat shows a universal $1/T^2$ temperature dependence (the fit shown is performed for $T > 15$). The inset shows the same data on a linear vertical scale.

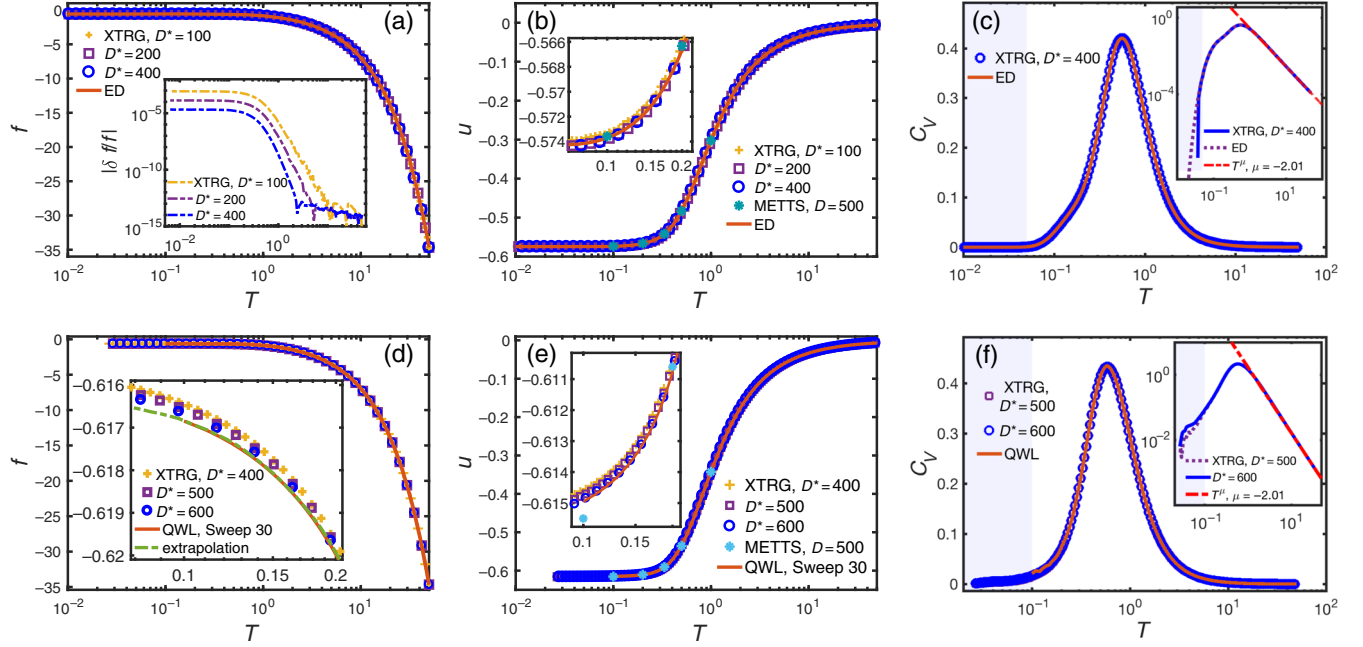


FIG. 7. Free energy f , internal energy u , and specific heat c_V of the isotropic Heisenberg square lattice on an $L = 4$, $W = 4$ [upper panels (a)–(c)] and $L = 16$, $W = 5$ [lower panels (d)–(f)] square lattice. The XTRG results are in very good agreement with those of ED for a 4×4 lattice and quantum Wang-Landau (QWL) (expansion order 1200, sweep number 30, down to $T/J = 0.1$) for the 16×5 lattice. With D^* the number of retained bond multiplets in the $SU(2)$ MPO and D the corresponding number of individual $U(1)$ states, data is shown for D^* (D) equal to $D^* = 400$ (1462), $D^* = 500$ (1832), and $D^* = 600$ (2400), resulting in a maximal ratio $D/D^* \simeq 4$. The maximal truncation errors, i.e., at largest β , are $\delta\rho \sim 1.4 \times 10^{-5}$ for a 4×4 lattice at $D^* = 400$ and $\delta\rho \sim 1.9 \times 10^{-4}$ for a 16×5 lattice at $D^* = 600$. These truncation errors directly scale with the relative error in the partition function and, thus, the free energy. Extrapolating $1/D^* \rightarrow 0$ by a quadratic polynomial based on $D^* \sim 300$ – 600 data results in perfect agreement (green dashed line) with QWL data in (d). The low-temperature regions in (c),(f) are shaded where we have a limited accuracy of c_V . The insets in (c),(f) fit the high- T specific heat with T^μ , resulting in $\mu = -2.01$ in both cases (based on the last eight points, i.e., $T > 30$).

Since the Trotter-Suzuki decomposition is not involved in the procedure, it is thus free of Trotter error.

B. Heisenberg chain

First, we benchmark XTRG results with conventional linear evolution in Fig. 4, where an 18-site spin-1/2 Heisenberg (PBC) chain [cf. Eq. (2)] is calculated up to $\beta \simeq 82$. From Fig. 4(a), it is clear that the accuracy in the low- T regime gets continuously improved as D^* increases in XTRG. Starting from a fixed $\tau = 0.01$, XTRG reaches a precision as good as 10^{-8} also at the lowest temperatures for $D^* = 500$ as compared to exact diagonalization (ED) data. By keeping the same bond dimensions, LTRG and SETTN have almost the same accuracy in all temperature regimes. When compared to XTRG, they are of similar accuracies only at high temperatures ($\beta \lesssim 1$) but are clearly less accurate in the low- T region, where truncation errors dominate. These remarkable results suggest that, as XTRG targets the low- T properties much faster than LTRG as well as SETTN, due to the (much) fewer truncation steps and its algorithmically much simpler setup, it also gains better results.

The MPO entanglement, as defined in Sec. IV and measured in the center of the chain, is plotted in Fig. 4(b),

which offers a quantitative estimate of computational complexity. The entanglement data suggests that truncation errors start to develop for $\beta \gtrsim 1$ and the simulation errors stop increasing due to the convergence of entanglement for $\beta \gtrsim 10$. This is also clearly reflected in the overall error in physical quantities such as the free energy in Fig. 4(a).

Besides the Heisenberg chain, we also benchmark XTRG, LTRG, and SETTN for an XY chain [cf. Eq. (3)] with size $L = 50$, where analytical solutions are available (Appendix F). As shown in Fig. 5, again XTRG gets better results than LTRG and SETTN, and the accuracy in the low-temperature regime also improves continuously as we increase bond dimensions D . This simulation on a longer XY chain again confirms that increasing β exponentially fast not only improves the efficiency but also gains in accuracy.

Besides the free energy, we also calculate the specific heat of a spin-1/2 Heisenberg chain of length $L = 300$, utilizing the XTRG algorithm with a bond dimension up to $D^* = 250$. As shown in Fig. 6, in the low- T region, the specific heat of the system shows a universal linear relation versus temperature, as indicated by the polynomial fitting (purple dashed line) with $\eta \simeq 1$. In addition, the fitted slope is also in perfect agreement with the well-known value $2/3$ from the CFT prediction [54], from which we extract the

central charge $c \simeq 1$. On the other hand, in the high- T regime, the specific heat is also universal and decays as $1/T^2$ (a polynomial fit in the log-log scale, depicted by the yellow dashed line, yields an exponent $\mu = -2.01$). This exponent can be confirmed by a high-temperature expansion up to the second order, which approximates the energy as

$$E = \frac{\text{Tr}(e^{-\tau H})}{\mathcal{Z}(\tau)} \simeq \frac{1}{\mathcal{Z}^0} \left[\text{Tr}H - \text{Tr}(H^2)\tau + \frac{(\text{Tr}H)^2}{\mathcal{Z}^0}\tau + \mathcal{O}(\tau^2) \right],$$

where $\mathcal{Z}^0 = \text{Tr}(\mathbb{I})$, with the high- T limit $c_V \sim 1/T^2$.

C. Square-lattice Heisenberg model

Symmetric TTN methods, including the XTRG and the SETTN, can be conveniently employed to calculate 2D systems, with minor adaptations. We map the 2D clusters into a 1D snake shape and prepare the MPO representation of this Hamiltonian (with “long-range” interactions) as elaborated in Appendix B. Other than that, one follows exactly the same line as in 1D simulations and can represent the density matrix of the 2D systems accurately in terms of MPOs.

Here, we perform calculations on 2D clusters and benchmark the calculations with ED for small (OBC) systems (4×4) in Figs. 7(a)–7(c) and quantum Monte Carlo (QMC) for larger systems (16×5) in Figs. 7(d)–7(f). With non-Abelian symmetries implemented in the highly efficient XTRG algorithms, we obtain high-quality data till quite low temperatures, which were not accessible before by other thermal RG algorithms.

In Figs. 7(a)–7(c), we show the free energy, energy, and specific heat results of a 4×4 square-lattice Heisenberg (SLH) model. Very nice agreement between XTRG and ED data is observed in all three plots. As seen in the inset in Fig. 7(a), the relative accuracy is quite high, i.e., 10^{-4} (10^{-5}) for $D^* = 200$ (400) at low temperatures ($T \leq 0.05$). The energy density shown in Fig. 7(b) is obtained by taking derivatives of interleaved XTRG free energy data [cf. Eq. (9)]. The error in the energy density is small even down to $T \leq 0.05$, as seen in the inset in Fig. 7(b) by enlarging the low- T region. XTRG data differs from the ED results only in the fourth digit for $D^* = 200$ and are bounded by a numerical differentiation error only for $D^* = 400$. In Fig. 7(c), we show our results of the specific heat, which is calculated by taking derivatives of energy data as in Eq. (9b). The inset plots c_V on a log-log scale, from which we again observe an algebraic behavior ($1/T^2$) at high temperatures. In the low- T region, it shows a very rapid (exponential) decay versus the temperature. This decay can even be observed in the dashed region in Fig. 7(c), although there the XTRG data departs from ED results due to a lack of accuracy.

For a 16×5 SLH, which is far beyond the scope of ED calculations, we compare our XTRG results to

those of quantum Wang-Landau (QWL) simulations [55] in Figs. 7(d)–7(f). We run the calculation down to $T = 0.025$. For the smallest temperature $T = 0.1$ for which we have well-converged QWL reference data at a comparable numerical cost, the error in the $D^* = 600$ data for the free energy is approximately 2×10^{-4} . Since the truncation error is generally larger on a 16×5 lattice, we extrapolate the free energy in Fig. 7(d) to $1/D^* \rightarrow 0$ and observe a perfect agreement with the QWL data, with the error further reduced by about an order of magnitude. Besides the free energy, in Figs. 7(e) and 7(f), we also show that the energy density u as well as specific heat c_V all have very good accuracy. In the high- T region, c_V again shows a $1/T^2$ relation, as shown in the inset in Fig. 7(f) and as already discussed with Fig. 6. In the very low temperatures [dashed region in Fig. 7(f)], though, our accuracy becomes somewhat limited, as the derivative to obtain c_V develops minor wiggles at the lowest temperatures, seen on the log-scale in the inset. However, the upturn does not appear to be due to numerical inaccuracies but appears to be physical, in the sense that, similar to Fig. 7(c), it is a precursor before the finite-size spectral gap sets in.

In Figs. 7(b) and 7(e), for comparison, we also include METTS data exploiting $U(1)$ symmetry only [30]. The METTS results also show good agreement with our other methods in both cases, apart from the fact that the METTS energy data is not strictly variational, i.e., could be even lower than the (quasi)exact value. As for the 16×5 plot in Fig. 7(e), note that $T = 0.1$ is currently the typical lowest temperature that 2D METTS simulations can reach [30] at computational resources that are comparable to XTRG. However, the current 2D METTS involves many swap gates and needs at least a few hundreds of samples. In contrast, our XTRG method, with $SU(2)$ symmetry implemented, is much more efficient and can reach much lower temperatures with great accuracy in 2D.

D. Triangular lattice Heisenberg model

Since Anderson’s famous conjecture on the resonating valence bond (RVB) state in the frustrated triangular lattice Heisenberg (TLH) model [56,57], TLH continues to intrigue people today [58–68]. Although the ground state turns out to be a 120° ordered magnetic state [61,63,64], finite-temperature properties of TLH appear to be anomalous, in that they cannot be described by renormalized classical behavior, as one would expect for systems with a magnetically ordered ground state and that thermal data “extrapolates” to a disordered state [58,60,65,66,68]. TLH materials have been realized experimentally [69–72]. Therefore, the theoretical understanding of its thermodynamic properties becomes more pressing. So far, however, low-temperature simulations of TLH were hindered by the lack of sufficiently powerful numerical approaches.

Besides nonfrustrated systems such as the SLH, XTRG can also be applied to frustrated magnets like TLH. In

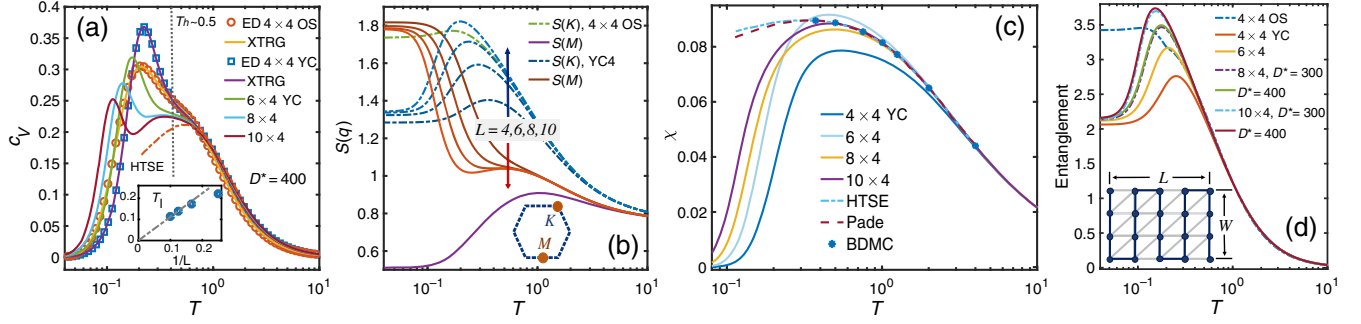


FIG. 8. Thermodynamics of the triangular lattice Heisenberg model defined on two $W = 4$ geometries, OS with OBCs on both directions and Y cylinder (YC) with various lengths L , i.e., PBCs along the vertical direction [see the inset in (d) for our specific choice of the YC geometry, as well as the TLH lattice layout in Fig. 9(f) below]. The presented quantities include (a) specific heat c_V , (b) static structure factor $S(q)$ at $q = K$ and $q = M$ at the boundary of the first Brillouin zone (see the inset), (c) uniform magnetic susceptibility χ , and (d) thermal entanglement versus temperatures. The inset in (a) depicts the temperature T_l of the peak in c_V at a lower temperature versus inverse length $1/L$, with the dashed line a guide for the eye. Note that the high-temperature scale in c_V stays at $T_h \sim 0.5$. The HTSE and Padé data is taken from Refs. [58,59] and BDMC from Ref. [60].

Fig. 8, we present exemplary XTRG results of TLH on width $W = 4$ systems [see the inset in Fig. 8(d)], including open strips (OS) and Y cylinders (YC). To be specific, we consider YC $4 \equiv$ YC($W = 4$) geometry with various aspect ratios $L/W = 1-2.5$, as well as a 4×4 OS for comparison. The YC geometry we adopt in practice is shown in the inset in Fig. 8(d) and also in Figs. 9(f)–(i) after a proper transformation to restore the corresponding triangular lattice geometry. Our realization is essentially equivalent to conventional YC in previous DMRG studies [73] at $T = 0$. We cool down the system from high temperatures to as low as $T/J = 0.03$, and $D^* = 400$ multiplets are kept in the following calculations, to ensure convergence versus bond dimensions.

Although YC4 shows clear finite-size effects in the ground state at $T = 0$, nevertheless, we can observe how correlations gradually develop as we lower the temperatures, with relevant finite- T physics in the thermodynamic limit at larger T down to temperatures where finite-size effects set in. In Fig. 8(a), we plot the specific heat of a 4×4 OS and various YC4 lattices. By comparing Fig. 8(a) to corresponding c_V in Fig. 7(c) of the SLH, we observe quite distinct features in the TLH case. Already on the 4×4 lattice, either OS or YC, the specific heat c_V exhibits a peak, whose location is denoted by T_l , and a shoulder structure at a higher temperature $T_h \sim 0.5 \sim J$. For YC4 lattices with increasing length $L = 6, 8, 10$, these features in c_V develop into a pronounced two-peak structure where the plateaulike feature for $L = 4$ develops into a broad peak located around the same stable value of T_h . Note that T_h is also about the same temperature scale where the data from high-temperature expansion (HTSE) shows a round peak [58]. In stark contrast, the peak at T_l moves towards lower temperatures. Its value scales like $T_l \sim 1/L$ as seen in the inset in Fig. 8(a) and, therefore, is clearly linked to the finite system size. Overall, the stark qualitative difference between the data for TLH in Fig. 8(a) as compared to

SLH in Fig. 7(c) may be ascribed to finite-temperature effects of magnetic frustration.

To gain a better understanding, we analyze the static structure factor

$$S(q) = \sum_j e^{-iq \cdot r_{0j}} \langle \mathbf{S}_0 \cdot \mathbf{S}_j \rangle, \quad (10)$$

where $r_{0j} \equiv r_j - r_0$ with r_j the lattice location of site j . In practical calculations, site 0 is fixed in the system center, while j runs over the whole lattice. Therefore, by inversion symmetry of the TLH, $S(q)$ is a real number. Note that also the structure factor above can be conveniently and efficiently obtained via the expectation value of single q -dependent MPO of bond dimension $D_S^* = 2$.

The static structure factor for the TLH is analyzed at specific values of q in Fig. 8(b). However, before discussing these in detail, let us look at the static structure factor over the entire Brillouin zone, as presented in Fig. 9 for a YC4 system at $L = 8$ along with the bond energy texture, at various temperatures (including $T = 0$, obtained by DMRG). In agreement with the previous discussion, Fig. 9(a) indicates the existence of three regions, separated by the two temperature scales T_l and T_h as introduced with Fig. 8(a). For $T \gg T_h$, the system is in a “spin gas” paramagnetic phase, with a featureless structure factor and bond texture in Figs. 9(b) and 9(f), respectively. In Fig. 9(c), we can observe that the intensities of $S(q)$ at $q = K \equiv \pm[(2\pi)/3](1, \sqrt{3})$, $\pm[(2\pi)/3](-1, \sqrt{3})$, and $\pm[(2\pi)/3](2, 0)$ become prominent in the intermediate region $T_l < T < T_h$, representing the development of strong 120° -ordering correlations.

In addition to that, there also emerges an extended region with considerable intensity around $q = M \equiv \pm 2\pi[0, (1/\sqrt{3})]$, $\pm\pi[1, -(1/\sqrt{3})]$, and $\pm\pi[1, (1/\sqrt{3})]$, near which rotonlike excitations were reported both theoretically [59,65,66,68] and experimentally [74]. The

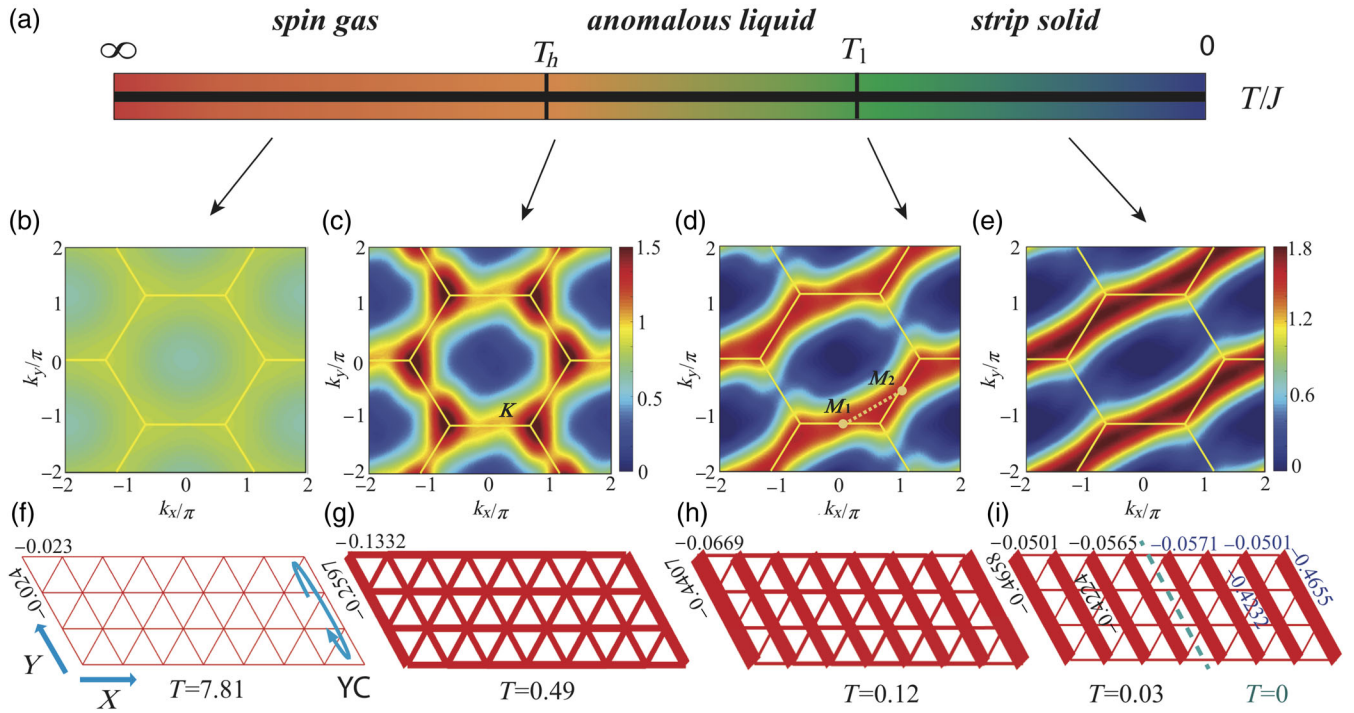


FIG. 9. (a) Finite-temperature phase diagram of TLH on a YC4 lattice, which consists of high- T “gas,” intermediate- T liquid, and low- T solid states. (b)–(e) show the static structure factor at four representative temperature points (left to right, from high temperature to low), and (f)–(i) exhibit corresponding bond textures at finite T , and the values are the bond energies to whose absolute value the thickness of nearest-neighbor bonds are proportional. In (i), we also show the bond energy distribution at $T = 0$ calculated by DMRG, where the agreement between XTRG results (left half) at low- T and DMRG data (right half) is apparent, and one can recognize the strip solid structure quite distinctly. Dashed lines are guide for the eye, representing in (e) the line connecting two inequivalent M points (labeled as M_1 and M_2) with enhanced intensity. In (f), we label the X and Y directions, as well as the way we wrap the lattice into a Y cylinder.

(anomalous) enhancement of $S(q)$ around these M points for temperatures $T \sim T_h$, which can be more clearly observed still in Fig. 8(b), may thus be related with roton excitations. By association with T_h , these are also linked to the round peak or shoulder in the specific heat c_V . The corresponding bond texture in Fig. 9(g) reveals much stronger spin-spin correlation than those in paramagnetic phase, while it still maintains a uniform pattern that respects the TLH lattice symmetry, like a liquid state. Thus, we dub the finite- T region between T_l and T_h the anomalous quantum liquid phase.

As T is lowered further down to less than $T_l \sim 0.14$, the system in Fig. 9 undergoes a rapid crossover. Because of the finite YC circumference, tightly bound RVB rings of length W form, as can be clearly observed in the bond textures in Figs. 9(h) and 9(i). There emerges a vertical strip pattern around the cylindrical circumference at $T \lesssim T_l$ [Fig. 9(h)], which becomes strongly dominating for lower temperatures [Fig. 9(i)]. Note that the low-temperature ($T/J = 0.03$) bond texture [left half of Fig. 9(h)] is in perfect agreement with $T = 0$ DMRG results [right half of Fig. 9(h), glued together at the dashed center line]. The agreement between bond energies is better than 0.1% for most bonds, indicating that the XTRG already effectively

reached the $T = 0$ regime. This is further reflected in the average energy per site $e_0 = -0.53027(1)$, for a given YC4 system at $L = 8$ and $T/J = 0.03$, which is in perfect agreement with, i.e., just above the DMRG ground-state energy $e_0 = -0.53034$ [with $D = 4000$ U(1) states retained]. Similarly, also the structure factor is in very good agreement, e.g., with $S(M) = 1.82$ at $T = 0$ DMRG to be compared to $S(M) = 1.79$ at $T/J = 0.03$ XTRG.

As a consequence of the tightly bound rings around the circumference of the YC4 cylinders, the correlations become significantly enhanced along a line connecting two of the three initially equivalent M points in the Brillouin zone [dashed line in Fig. 9(d)]. This enhancement strongly competes with the triangular magnetization associated with the K points [Fig. 9(d)], such that the structure factor $S(M)$ at these points eventually dominates over $S(K)$ [Fig. 9(e) and also Fig. 8(b)]. This competition gives rise to the low- T peak of the specific heat curve in Fig. 8(a). From Fig. 8(b), it can be observed that, almost exactly at T_l , the two correlations cross and $S(M)$ becomes larger than $S(K)$ for $T < T_l$. Here, the symmetry breaking across the initially equivalent M points is due to the finite-size cylindrical structure.

The results of magnetic susceptibility χ of various YC4 lattices are shown in Fig. 8(c). It can be observed that $W = 4$ data can produce results in agreement with the HTSE [58] and bold diagrammatic Monte Carlo (BDMC) [60], as well as Padé approximation results [59], down to $T \sim 0.4$ (for YC 10×4), despite a very limited circumference. This observation is quite surprising, which suggests a very small finite-size effect and correlation length $\xi (\lesssim 1)$ for $T/J \geq 0.4$, in accordance with previous studies [58].

In Fig. 8(d), we plot the bipartite entanglement entropy S_E in the purified state. Since S_E differs from bond to bond, we show there the maximal value among all bonds in the effective 1D chain structure adopted in the calculation [inset in Fig. 8(d)]. It can be clearly seen in Fig. 8(d) that this low-temperature entanglement reaches a relatively small value around $S_E = 2.16$. DMRG simulation shows that the corresponding entanglement $S_E \simeq 1.1$, which, as expected, is roughly one-half of the entanglement in the low- T mixed state.

The tightly bound stripes in the low- T regime in a given YC4 system lead to a strongly reduced entanglement entropy when cutting the system exactly in between the stripes, i.e., vertically in the inset in Fig. 8(d), where the entanglement is reduced, e.g., down to $S_E \sim 0.9$ at $T = 0.03$. Therefore, at low temperatures $T < T_l$, the system is close to a direct product of tightly bound uniform 1D rings around the circumference of the YC4 cylinder. These four-site rings favor a well-known RVB ground state $|\text{RVB}\rangle = [1/(2\sqrt{3})](\phi_{12}\phi_{34} - \phi_{14}\phi_{23})$, where $\phi_{ij} \equiv |\uparrow_i\downarrow_j - \downarrow_i\uparrow_j\rangle$ constitutes a valence bond, and the lowest excited state is separated by a significant energy gap $\Delta \sim J$. These tightly bound rings are disrupted when opening the boundary. Therefore, in the OS geometry, e.g., in the 4×4 data in Fig. 8(b), $S(K)$ remains strong and never crosses with $S(M)$ at $M = \pm 2\pi[0, (1/\sqrt{3})]$ as T is lowered. However, for the 4×4 OS, the dominant weight of the structure factor turns out to be located at one out of three types of Néel antiferromagnetic order, i.e., $S(M)$ with $M = \pm(\pi, -\pi/\sqrt{3})$ [data not shown in Fig. 8]. The entanglement also increases monotonically on the OS lattice, until it saturates due to the finite system size [Fig. 8(d)].

Recently, BDMC has been employed to explore thermal properties of TLH, and it was found that the extrapolated ground state is disordered, via a particular quantum-to-classical mapping based on their best thermal data (with temperatures down to $T/J \simeq 0.375$) [60]. The anomalous thermodynamic behavior is in contradiction with the ground-state 120° ordering, which reveals that the true low-temperature regime has not been reached. To fully understand the finite-temperature anomaly and to resolve the above apparent contradiction, a more extensive XTRG survey of the TLH, e.g., on wider cylinders down to low temperatures, is required. This is beyond the scope of this

paper and, thus, will be reported elsewhere [75]. For the purpose of this paper, nevertheless, we demonstrate that XTRG provides a highly competitive approach that allows one to tackle complex and challenging problems.

IV. ENTANGLEMENT IN THERMAL TENSOR NETWORKS

A. Thermal entanglement renormalization-group flow

The entanglement measure in a thermal state is more complicated as compared to a ground state due to the interplay of classical correlation and quantum entanglement. Among various definitions, we take a very natural and most relevant measure of the entanglement in practice, i.e., entanglement in the normalized “superstate” [76]

$$|\Psi(\beta)\rangle \equiv \frac{1}{\sqrt{\mathcal{Z}(\beta)}} \underbrace{\{e^{-(\beta/2)H}\}}_{\equiv |\tilde{\Psi}(\beta)\rangle} \quad (11)$$

which vectorizes the MPO for $e^{-\beta H/2}$. In other words, the MPO is simply transformed into a matrix product state (MPS) with doubled local state spaces. Then the partition function is equivalent to the overlap of the unnormalized superstate $\mathcal{Z}(\beta) = \langle \tilde{\Psi}(\beta) | \tilde{\Psi}(\beta) \rangle$, whereas $\langle \Psi(\beta) | \Psi(\beta) \rangle = 1$. Note that this definition is a specific (and most natural) choice of purification [21,77], which in some other context is also called the thermofield double (TFD) state $|\Psi(\beta)\rangle = [1/\sqrt{\mathcal{Z}(\beta)}]e^{-(\beta/2)E_n}|n, \bar{n}\rangle$, where E_n is the eigenenergy of eigenstate $|n\rangle$ and its duplicate $|\bar{n}\rangle$ in the auxiliary state space [53,78–80]. Here, for simplicity of notation, by the entanglement entropy or the entanglement spectrum (ES) of the MPO or the thermal state, we refer to precisely these quantities obtained from the underlying TFD or, equivalently, the purified and normalized state in Eq. (11). Specifically, the entanglement spectrum is given by the eigenspectrum of $\mathcal{H}_{\text{ES}} \equiv -\ln \mathcal{R}$, where \mathcal{R} is the “superdensity” matrix of the purified thermal state $|\Psi(\beta)\rangle$ as in Eq. (11).

Note that the MPO entanglement analyzed here is not directly related to the entanglement of purification, which is defined as the minimal value among various purification schemes [81,82]. Nevertheless, through the optimal truncation via orthogonal state spaces in the XTRG (and also LTRG), one is simultaneously optimizing the superstate overlap (i.e., partition function), as well as this MPO (TFD) entanglement. Therefore, this MPO entanglement, as well as some other measures, like the mutual information, quantifies the resources required to perform efficient thermal simulations and, thus, have attracted recent interest [32,83,84].

We start by analyzing the entanglement spectra of the thermal state for a spin-1/2 Heisenberg chain, from which we can also compute its entanglement block entropy S_E . By lowering the temperatures, one generates a RG flow that

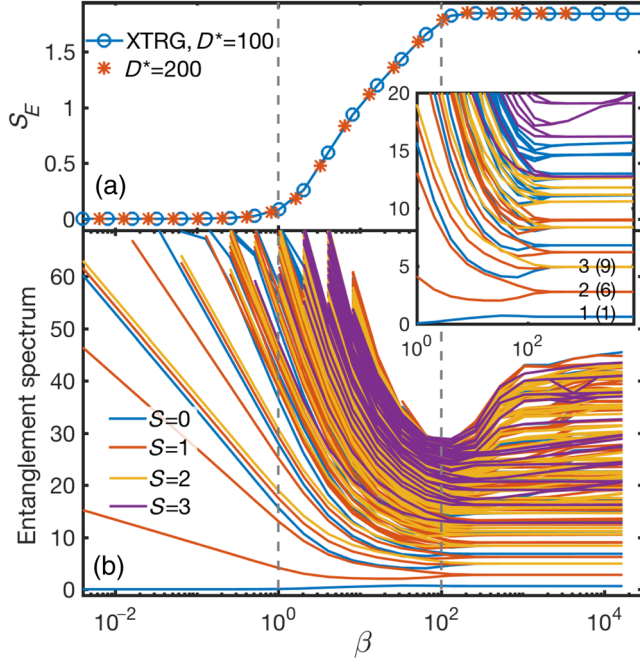


FIG. 10. (a) Bipartite MPO entanglement entropy S_E across the center of the system of a spin-1/2 Heisenberg chain (length $L = 100$, OBC). (b) Entanglement spectra obtained in the center of the system versus a wide range of temperatures presented as an RG flow in energy scales, where lines from the same symmetry sectors, i.e., $S = 0, 1, 2, \dots$, are plotted in the same color. Vertical markers depict different temperature regimes (see text). The inset enlarges the region around $\beta = 100$, and the labels d^* (d) indicate the degeneracy in the RG fixed point ES in terms of individual multiplets (or states), respectively.

directly reflects different physical regimes of the system at various temperatures, i.e., energy scales. In Fig. 10, we show the entanglement RG flow over a very wide range of temperatures. We can vary β over 7 orders of magnitude, which thus reaches far beyond Trotter-Suzuki-type calculations. The RG flow reveals three distinct regimes, demarcated by vertical dashed lines in Fig. 10: (i) a low-entanglement region $\beta \lesssim 1$, (ii) an intermediate region $1 \lesssim \beta \lesssim 100$ where entanglement rises quickly, and (iii) the saturation region for $\beta \gtrsim 100$ where the ES flows to a fixed point, either converging to the ground state of a physically gapped state in the thermodynamic limit or resolving the gap of finite size level spacing.

When approaching the low-energy fixed point ES, lines systematically merge into groups with larger degeneracy as seen in the inset in Fig. 10. Given that the entropy already converges to a finite value at the lowest temperatures, this suggests that the low-energy fixed-point spectrum must be related to the tensor product space of two copies of the ground state (bra and ket) which naturally results in systematically enlarged degeneracies.

For the remainder of this section, we focus on the entanglement entropy in both 1D chain and 2D lattice

models over a wide range of temperature scales. Interesting logarithmic behaviors are observed, which intimately relate to (gapless) low-energy excitations yet also suggest efficient computational complexity of thermal simulations for the specific model systems considered.

B. Universal entanglement behavior in (1+1)D conformal thermal states

In Fig. 11, we plot the entanglement entropy in a spin-1/2 Heisenberg chain. By simulating twice the length ($L = 200$) as in Fig. 10, we can observe a logarithmic divergence in the low-temperature region ($L \gtrsim \beta \gg 1$). The logarithmic entropy was already observed in the past and related to the computational complexity of finite-temperature simulations [16,31]. More recently, this was further analyzed by numerical simulations on Renyi entropy and also conformal field theory (CFT) analysis [32]. Notably, the finite-temperature entanglement calculation in Fig. 11 provides a convenient and accurate way to extract the central charge c of CFT: Without going into ground-state calculation at $T = 0$ [85–87], one can fit the MPO entanglement at finite temperatures. In Fig. 11, we observe that, by fitting $L = 200$ data, the estimate of central charge c is already very accurate ($c = 0.999 \simeq 1$). For this, we fit the data to the CFT prediction $S_E = (c/3) \ln \beta + \text{const}$ [32]. Importantly, by having the system length sufficiently

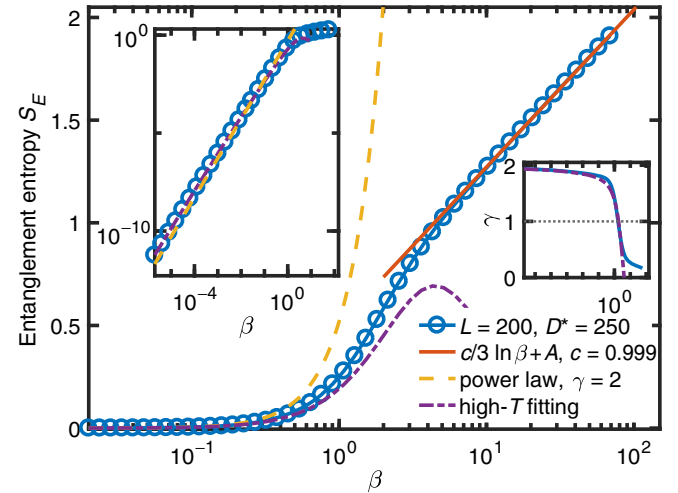


FIG. 11. Bipartite MPO entanglement S_E across the center of the system in an OBC Heisenberg chain of length $L = 200$. At low temperatures ($\beta \geq 10$), S_E diverges logarithmically. The line in the main panel for large β (small T) is a fitting to $S_E = c/3 \ln \beta + \text{const}$, with central charge $c \simeq 0.999$ determined by fitting the range $\beta > 12$. The left inset shows the same data on a log-log plot, which emphasizes a power law of entanglement versus β for $\beta \lesssim 1$. The purple dashed lines represent the fit in Eq. (13) with $\alpha = 0.05$, while the yellow dashed lines represent β^2 as a guide to the eye. The right inset shows the slope of the log-log data in the left inset, i.e., the power-law exponent γ of the entanglement versus β .

large, the physics of the thermal state in the center of the system is effectively short-ranged by a thermal correlation length. In this sense, the simulation of the central charge in Fig. 11 does not yet see the finite open boundary condition, which is in stark contrast to the evaluation of central charge using ground-state properties.

Universal features of the entanglement property appear also at large temperatures. As seen in the left inset in Fig. 11, the MPO entanglement shows a power-law behavior for $\beta \lesssim 1$. The slope γ on the log-log plot at very high temperatures is analyzed in the right inset in Fig. 11, which suggests a power-law exponent $\gamma \approx 2$ for $\beta \lesssim 1$. The growth in the entanglement S_E , however, slows down strongly for $S_E \gtrsim 1$, i.e., $\beta \gtrsim 1$, where γ drops significantly below 1 as seen in the right inset in Fig. 11. The entanglement behaviors at high temperatures can be understood from a lowest, i.e., first-order, expansion of the density operator, $\rho(\tau) = \mathbb{I} - \tau H$. The singular value spectrum of this MPO is given by the vector $s = [1, \alpha\tau]$, with α another numerical vector. The resulting normalized “density matrix” of the supervector in Eq. (11) has eigenvalues $r_i = s_i^2 / \sum_j s_j^2$ with lowest-order thermal contributions $\propto \tau^2$. With the von Neumann entropy $S_E = -\sum_i r_i \ln(r_i)$ and

$$\gamma(\tau) \equiv \frac{d \ln S}{d \ln \tau}, \quad (12a)$$

one obtains that

$$\gamma_0 \equiv \lim_{\tau \rightarrow 0^+} \gamma(\tau) = 2. \quad (12b)$$

For simplicity, one may consider a single value for the vector α , resulting in the two normalized weights

$$(r_1, r_2) = \frac{1}{1 + \alpha^2 \tau^2} (1, \alpha^2 \tau^2) \quad (13a)$$

with von Neumann entropy

$$S(\tau; \alpha) = -\sum_{i=1}^2 r_i \ln r_i. \quad (13b)$$

A subsequent one-parameter fitting of $S(\tau; \alpha)$ with respect to α to the actual entanglement entropy S_E for $\tau = \beta \ll 1$ nicely reproduces the high- T entanglement data, as shown in the right inset in Fig. 11 for $\alpha = 0.05$. The slope γ decreases monotonically, starting from $\gamma = 2$ and undergoing a sharp decrease around $\beta \sim 1$. Note, however, that the convergence towards the power-law exponent of $\gamma = 2$ for small τ is extremely slow, as also clearly supported by the simple asymptotic analysis above. For example, for $\tau = 10^{-3}$, one has only $\gamma \approx 1.94$.

Nevertheless, it follows from the generality of the above asymptotic argument that the exponent $\gamma = 2$ for

infinitesimal τ is universal. It should hold for any Hamiltonian and therefore, in particular, also in arbitrary dimensions. Furthermore, given that, by construction, the exponent $\gamma = 2$ holds only for $\beta \ll 1$ where $S_E \ll 1$, the growth of the entropy of the MPO may be considered sublinear in this regime, in the sense that the entropy grows slower than linear for infinitesimal τ , having $\lim_{\beta \rightarrow 0^+} [(dS_E)/(d\beta)] = 0$.

C. Logarithmic entanglement in thermal states of 2D Heisenberg model

The low-temperature entanglement of the thermal state saturates for gapped quantum chains and grows only polynomially for critical ones. This directly implies excellent numerical efficiency in 1D quantum systems, since the required bond dimension grows at most polynomially with inverse temperature β [16,31] rather than exponentially as originally estimated [17]. We take this as a motivation to explore the MPO entanglement of the Heisenberg magnet on the square lattice and take it also as an indicator of computational complexity for the latter.

In Fig. 12, we plot the entanglement property versus inverse temperature β , for the SLH of various system sizes, ranging from width $W = 2$ to $W = 6$. As shown in the inset in Fig. 12, in the high-temperature regime, one still

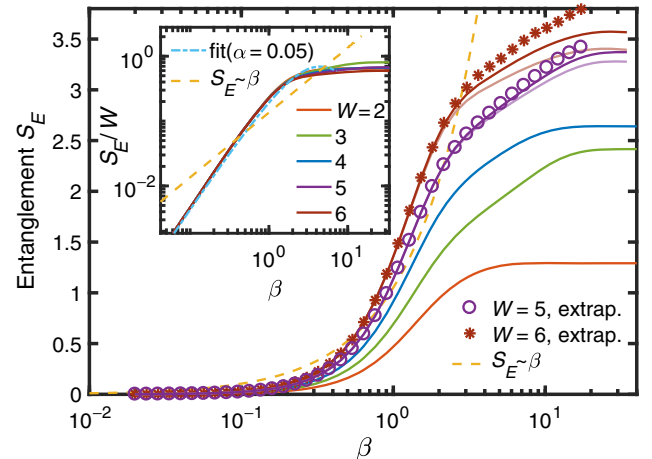


FIG. 12. Bipartite MPO entanglement across the center of the system in the SLH, up to width $W = 6$. The system length is fixed to $L = 10$ for width $W = 2, 3, 4, 5$ and to $L = 12$ for $W = 6$. For $W \leq 4$, the entanglement entropy in the center of the system is well converged by retaining $D^* = 500$ bond multiplets. For $W = 5$ and 6 , we show data with $D^*(D) = 300$ (1000) and 550 (2000) (color matched lines) as well as extrapolated data (symbols) based on a quadratic polynomial extrapolation in $1/D^* \rightarrow 0$ that includes additional data between $D^* = 300$ and $D^* = 550$. The inset presents the entropy per leg (S_E/W) on a log-log scale, where the power-law increase for $\beta \ll 1$ again clearly follows the large-temperature fit in Eq. (13) with approximate exponent $\gamma \approx 2$, using $\alpha = 0.05$. The dashed line shows $S_E \sim \beta$ as a guide to the eye for reference.

recovers the universal power law $\gamma \simeq 2$; cf. Eq. IV B. Specifically, the data is very well fitted by the function in Eq. (13) with exactly the same parameter $\alpha = 0.05$ as in Fig. 11 for the 1D quantum chain.

For the inset in Fig. 12, since the entanglement entropy satisfies the area law, i.e., $S_E \propto W$, we divide the entropy S_E by the width W . For high temperatures $\beta \lesssim 1$, this collapses the data for different system widths on top of each other, indeed demonstrating the universal area law in this regime. For intermediate and low temperatures, $\beta > 1$, deviations from the strict scaling collapse of the area law can be observed.

In the main panel in Fig. 12, the entropy S_E changes gradually into a logarithmic divergence versus $\ln \beta$ as W increases for both even and odd widths. This suggests that simulations are also efficient with increasing β in the 2D setting, while bearing in mind an additive constant term to the entropy that is proportional to the width (note that, in order to satisfy the area law, the entropy data for $2 \leq \beta \leq 20$ is separated roughly by equal vertical offsets when incrementing the width for $W \geq 3$). Since the calculations of the entropy in the system center are not fully converged for the wider systems, we also extrapolate the width $W = 5$ and 6 systems in $1/D^* \rightarrow 0$. This actually further reinforces the regime of logarithmic increase of S_E versus $\ln \beta$ for $2 \leq \beta \leq 20$.

A similar additive logarithmic scaling of the entanglement entropy ($\propto \ln W$) in the 2D Heisenberg model has been also found numerically via QMC calculations in the ground state [80,88–92]. The coefficient of logarithmic correction is argued to be universal [93] (proportional to the number of Goldstone modes in the system). In the 2D Heisenberg model here, according to ED and DMRG studies of the tower of states (TOS) in the energy and the entanglement spectra [94,95], respectively, the relevant low-energy TOS has characteristic level spacing that scales as $1/N$ with $N = WL$ the total number of sites. This is in contrast to spin-wave excitations (Goldstone modes) which have characteristic level spacing that scales with the inverse linear system size, i.e., $1/L$.

These $1/N$ TOS excitations are responsible for the logarithmic entanglement at $T = 0$ and possibly also relate to the $\ln \beta$ scaling of the entropy observed in the present study. In Fig. 12, we restrict the length to be as small as $L = 10$, which suggests that the magnon excitations are gapped out at temperatures as low as $T = 1/30$ – $1/10$. Therefore, the relevant energy scale is likely only TOS modes with an energy level spacing $\propto 1/N \sim 1/50$, which is smaller than the temperatures in the regime where logarithmic scaling $S_E \sim \ln \beta$ is observed.

The logarithmic growth in the entanglement in Fig. 12, as well as the low- T specific heat behavior, is quite remarkable. As we show shortly, they differ qualitatively from the anisotropic case $\Delta \neq 1$. In the isotropic case, the entanglement curves do not show any sign of singularity at

any finite T , which suggests the absence of a phase transition at $T \neq 0$. The low-temperature specific heat c_V gets significantly enhanced when the width W is increased from 4 to 5, but it still grows monotonically with increasing β , e.g., with no sign of a singularity at finite β . This result is, of course, in complete agreement with the celebrated finite-temperature Mermin-Wagner theorem [96].

D. XTRG and thermal phase transitions in 2D

Lastly, we analyze the 2D anisotropic XXZ model on the square lattice, i.e., Eq. (2) with $J = -1$ and $\Delta \neq 1$. There exists a finite-temperature phase transition at the critical temperature T_c towards a gapped low-energy ferromagnetically ordered phase. Note that, while Trotter-like methods have to necessarily work straight through a thermal phase transition point, which becomes singularly hard already with the first one encountered, our XTRG can *jump across* phase transition points [26] by reducing temperatures by a factor of 2.

In the following, we show that XTRG can be employed to simulate such a model with nonzero T_c with high accuracy. In particular, we determine the phase transition point using various quantities including the block entanglement entropy and specific heat, as well as the Binder ratio in spin fluctuations. We choose the same model parameters as in Ref. [30], i.e., $\Delta = 5$, where the critical temperature is estimated as $T_c^*/\Delta = 0.56 \pm 0.01$. This is in agreement with the exact result $T_c/\Delta = 0.56$ in Ref. [97] from QMC for much larger system sizes, which we also take as a reference for our data below.

Our results are presented in Fig. 13, where we also make a comparison to QMC data explicitly generated by the ALPS loop code [55]. Figure 13(a) shows the landscape of the block entanglement entropy across each bond of the MPO density matrices with a decreasing temperature. Because the low-temperatures phase is gapped, prominent peaks are present around the critical temperature $T_c = 0.56$, with only a weak dependence on the snakelike serial ordering of our 2D system, otherwise. Figure 13(b) shows a top view of the same data where the temperature T_c^S of the maxima of the entropy data S_E (dashed dotted line) agrees very well with the exact QMC value T_c (vertical dashed line).

To see this more clearly, we also take a cut in the center of the system [while shifted by half a column to avoid local minimal; see the caption of Fig. 13(c)]. In Fig. 13(c), we show S_E versus temperature T for various system widths W . From this we observe that the peak position T_c^S in S_E versus T quickly approaches T_c as W increases. Specifically, for $W = 8$ we already have $T_c^S \simeq 0.552$ with an error $|T_c^S - T_c| < 0.01$. Similar calculations are also performed using cylindrical boundary conditions, as shown in Appendix G (Fig. 19), where we also see a quite accurate agreement with T_c for rather small W .

We also resort to other more standard thermal quantities including specific heat c_V [Fig. 13(c)] and Binder ratio

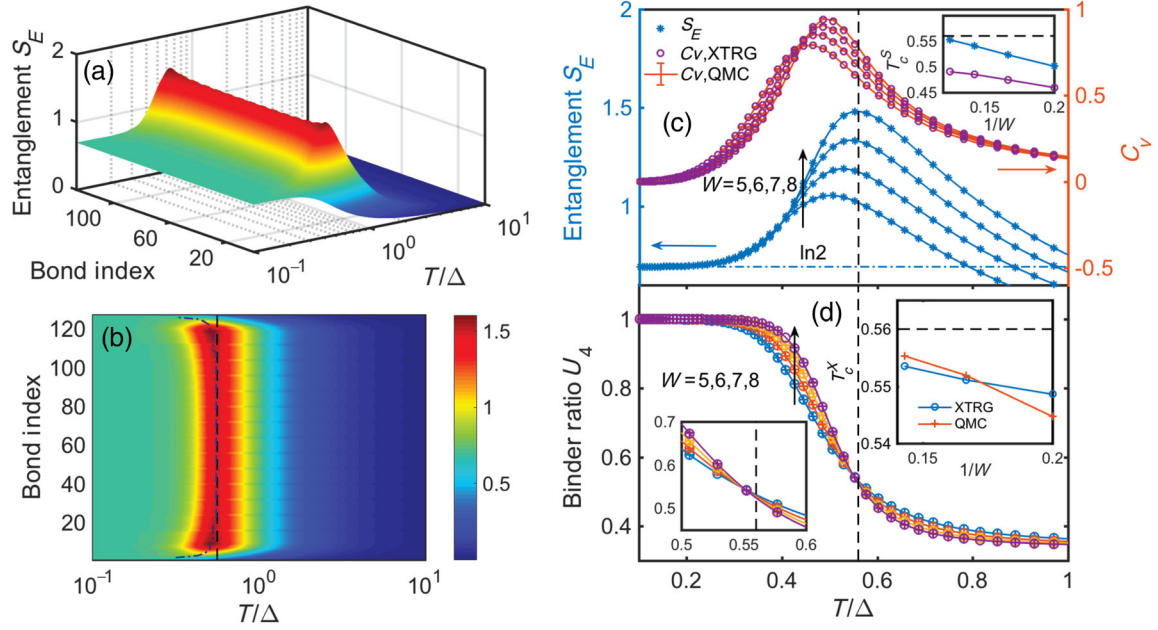


FIG. 13. Finite-temperature phase transition in the XXZ Heisenberg model at $\Delta = 5$, keeping $D = 600$ states in U(1) XTRG. (a) Entanglement landscape versus temperatures T/Δ and bond indices on a $L = 16$, $W = 8$ open square lattice which exhibits pronounced peaks around the exact transition temperature $T_c = 0.56$ [97] [dashed black lines in (b)–(d)]. The dash-dotted line in (b) depicts the maximal entropy point for each bond, which converges to $T_c^S \simeq 0.552$ in the central region of the system. (c) Block entanglements S_E at bond $[(N + W)/2]$ cutting across the center of the system (where $N = LW$ and $[\cdot]$ means the integer part) and also the specific heat c_V with a comparison to QMC data. The data is shown for various widths $W = 5, 6, 7, 8$ with a fixed aspect ratio, using $L = 2W$. Both S_E and c_V curves show peaks near T_c (vertical dashed line). The horizontal dashed line represents constant $\ln 2$, manifesting the global \mathbb{Z}_2 symmetry which can be spontaneously broken at low temperatures if applying a small pinning field, otherwise. The peak positions T_c^S of the S_E and c_V are collected and shown versus $1/W$ in the inset. The value of T_c^S from S_E suffers smaller finite-size corrections than that obtained from the specific heat. (d) Binder ratio U_4 for various system sizes $L = 2W$. XTRG data is plotted with open circles and looper QMC data with crosses with perfect agreement between the two data sets. The cross point determines T_c^X . By enlarging the data in the lower left inset, we obtain the crossing points $T_c^X(W)$ for pairs of consecutive system sizes W , as plotted versus $1/W$ in the upper right inset, again with good agreement between the XTRG and the QMC data. From this we obtain $T_c^X \simeq 0.554$ for our largest systems, which still trends towards T_c (horizontal dashed line) with increasing system size and, hence, agrees with T_c to within an error bar of 1%.

[Fig. 13(d)] to pinpoint the critical temperature. Our results for c_V by XTRG are in perfect agreement with QMC simulations, as shown in Fig. 13(c). The specific heat also exhibits a peak near T_c . As is well known from numerous QMC simulations, however, similar to other physical observables such as the spin susceptibility, the specific heat suffers significant finite-size corrections. So this often provides only a first rough estimate for T_c in practical numerics, where even the finite-size extrapolation to the thermodynamic limit often has larger error bars still, as is also the case here [cf. the inset in Fig. 13(c)].

The lowest-order, finite-size corrections in standard thermal quantities, however, can be eliminated by taking ratios of expectation values. In practice, a common way to pinpoint T_c more precisely is the Binder ratio [98,99]

$$U_4 \equiv \frac{\langle (S_{\text{tot}}^z)^2 \rangle_\beta^2}{\langle (S_{\text{tot}}^z)^4 \rangle_\beta}, \quad (14)$$

where $S_{\text{tot}}^z = \sum_i S_i^z$ is the total magnetization in the z -direction, with $\langle S_{\text{tot}}^z \rangle = 0$. We calculate U_4 with both

XTRG and QMC the looper code and find perfect agreement. Note that the Binder ratio U_4 can be conveniently calculated using an exact MPO representation for the total magnetization (operator) $S_{\text{tot}}^z = \sum_i S_i^z$ of bond dimension $D = 2$. By taking the product of S_{tot}^z and performing compression (without any essential truncations), one can obtain compact representations of $(S_{\text{tot}}^z)^2$ and $(S_{\text{tot}}^z)^4$ with $D = 3$ and $D = 5$ MPOs, respectively. With these MPOs, one can evaluate the finite-temperature expectation values $\langle \cdot \rangle_\beta$ of the two operators required for the Binder ratio in Eq. (14).

From the data as in Fig. 13(d), we collect the crossing points T_c^X for two consecutive curves for widths W and $W + 1$. The data is summarized in the inset in Fig. 13(d). By considering the data for the largest system to be the most accurate given the aspect ratio $L/W = 2$, we obtain $T_c^X \simeq 0.554$ from both the XTRG and QMC data. By considering the minor trend still with increasing system size, we find agreement between this T_c^X and T_c to within 1%, which thus serves as a very good estimate for the thermodynamic limit.

From the above, we conclude that XTRG can be used to determine T_c of the thermal phase transition very accurately. Furthermore, the maximum in the entanglement entropy S_E itself can already provide a good estimate for the critical temperature. Here, in particular, it clearly outperforms conventional thermal quantities like the specific heat c_V , as seen in the upper inset in Fig. 13(d). Nevertheless, since S_E grows linearly with W for $T > T_c$ but stays constant for $T < T_c$ in a given system (the system nearly becomes a product state for $T \ll T_c$), the peak around T_c^S is rather round. This might lead to systematic offsets in T_c^S in the thermodynamic limit but needs further studies.

V. CONCLUSIONS AND OUTLOOK

Inspired by the logarithmic growth of entanglement of purified (thermofield double) states, we propose an exponential speedup of thermal simulation. This thermal tensor network algorithm employs an MPO form of the density operator and proceeds via doubling of the density matrix $\rho(\beta)$ along the imaginary time evolution. We show that this XTRG method gains both accuracy and efficiency, in thermal simulations of the Heisenberg models. We also implement this idea of logarithmic temperature setup in a pointwise SETTN algorithm. Also, there we get more efficient and accurate results than previous Maclaurin SETTN [42].

We apply XTRG and SETTN to efficiently simulate thermal states of 1D and 2D Heisenberg spin models, obtain accurately the thermodynamic quantities including the free energy, energy, and specific heat, and study their low- and high-temperature behaviors. We have also investigated the temperature dependence of entanglement properties S_E in the MPO and observed logarithmic entropies $S_E \simeq a + b \ln \beta$ with constants a and b at low temperatures not only in gapless quantum chains but also in the SLH at a fixed system size due to gapless TOS modes.

We applied XTRG to a 2D Heisenberg models with a thermal phase transition, as well as the TLH with spin frustration. The results demonstrate the efficiency of the present algorithm, which is capable to show rich finite- T physics, especially for those system with spin frustration. It will be interesting to see how useful the present algorithm can be in exploring more 2D challenging systems, such as the kagome and the $J_1 - J_2$ Heisenberg models on the square lattice, as well as interacting fermionic systems.

With XTRG, however, we are not limited to high temperatures but can simulate down to much lower temperatures than previously anticipated [30]. The present MPO algorithms may be improved in several directions, including combining them with METTS samplings at low T or linked-cluster expansion to reduce finite-size effects, etc. Moreover, XTRG may also be straightforwardly combined with efforts to reduce block entanglement entropy further by operating with disentanglers on the auxiliary state space

in the purified scheme [81,82], all of which certainly deserve further exploration.

ACKNOWLEDGMENTS

The authors thank Benedikt Bruognolo for nicely providing the METTS data and for very constructive discussions. B.-B.C. and W.L. acknowledge helpful discussions on related topics with Han Li and Yun-Jing Liu, and A.W. also with Thomas Barthel. W.L. is indebted to Shou-Shu Gong for providing the DMRG data of YC4 TLH. This work was supported by the National Natural Science Foundation of China (Grants No. 11504014, No. 11834014, No. 11474015, No. 61227902, and No. 11774018) and the Beijing Key Discipline Foundation of Condensed Matter Physics. B.-B.C., L.C., and W.L. thank the hospitality of Arnold Sommerfeld Center for Theoretical Physics, University of Munich, where this work was finished. A.W. acknowledges support from the German Research Foundation (DFG) WE4819/2-1 and WE4819/3-1 until December 2017 and since then by U.S. DOE under Contract No. DE-SC0012704.

B.-B.C. and L.C. contributed equally to this work.

APPENDIX A: ENTANGLEMENT ENTROPY IN THERMAL STATES

The block entropy in the center $\ell \sim L/2$ of an individual low-energy pure state $|s\rangle$ scales like $S_E(|s\rangle) \simeq (c/6) \log \ell + \text{const}$ [86,87]. The MPO block entropy for the outer product $\hat{\rho}_s \equiv |s\rangle\langle s|$ then acquires a factor of 2, i.e., $S_E(\hat{\rho}_s) \simeq (c/3) \log \ell + \text{const}$. Now, by going to a thermal state $\hat{\rho} = \sum_s \rho_s |s\rangle\langle s|$ with weights $\sum_s \rho_s = 1$, this scaling does not change. More precisely, due to the subadditivity of entanglement entropy, one obtains an upper bound $S_E(\hat{\rho}) \lesssim (c/3) \log \ell + \text{const}$.

More explicitly, the block entropy of the thermal density matrix changes most for the *worst case* that its spectrum is altered from $q_i^{(s)}$ for some fixed s to the set $\rho_i^s \equiv \rho_s q_i^{(s)}$. The corresponding block entropy for a cut across the center of the system has the upper bound

$$\begin{aligned} S_E(\hat{\rho}) &\lesssim -\sum_{s,i} \rho_i^s \log \rho_i^s \\ &= \underbrace{-\sum_s \rho_s \log \rho_s}_{\equiv S(\{\rho_s\}) \simeq \text{const}} + \sum_s \rho_s \underbrace{\left(-\sum_i q_i^{(s)} \log q_i^{(s)} \right)}_{\equiv S_E(\hat{\rho}_s) \simeq \frac{c}{3} \ln \ell + \text{const}} \\ &\simeq \frac{c}{3} \ln \ell + \text{const}, \end{aligned} \quad (\text{A1})$$

having used $\sum_s \rho_s = 1$, as well as $\sum_i q_i^{(s)} = 1$ for all s . Here, $S(\{\rho_s\})$ is the entropy of the weight distribution ρ_s .

Now, Eq. (A1) provides an upper bound. As argued with Fig. 1, in order to sample thermal averages, one requires

that the system size must be larger than the thermal correlation length. For physical systems whose finite-size spectra have a low-energy level spacing that scales like the inverse system size, this is achieved by choosing $L = a\beta$ with $a \gtrsim 1$ sufficiently large but constant. Therefore, one can substitute $\ell = L/2 \rightarrow \beta$ in Eq. (A1), resulting in the overall block entropy of the thermal state

$$S_E(\hat{\rho}) \simeq \frac{c}{3} \log \beta + \text{const.} \quad (\text{A2})$$

The block entropy of the thermal state thus saturates, since the system length is effectively cut off by the thermal correlation length ξ , resulting in $\ell \sim \min(L, \xi)$ in the thermodynamic limit $L \rightarrow \infty$. See also Refs. [32,33] for a more rigorous derivation based on CFT arguments. Furthermore, note that even Eq. (A2) can be still considered an upper estimate, since, in the purification scheme, the thermal state allows one to minimize block entanglement entropy by disentangling operations on the auxiliary state space [81,82].

APPENDIX B: SYMMETRY-INVARIANT MATRIX PRODUCT OPERATOR FOR HAMILTONIANS

In this Appendix, we discuss our approach to the implementation of both Abelian as well as non-Abelian symmetries into the MPO representation of a given Hamiltonian. Conceptually, non-Abelian symmetries proceed the same way as Abelian symmetries, as we explain below. The actual implementation is based on the framework of the tensor library QSpace [48] that can deal with Abelian and non-Abelian symmetries such as $SU(N)$ or the symplectic symmetry $Sp(2N)$ on a generic footing.

In order to emphasize the generality of the argument, we frequently use the notation q for a label of a generic irreducible representation (irrep) of a given symmetry. Here, specifically, it may stand for either the spin projection S^z or spin S label in the case of a $U(1)$ [an $SU(2)$] symmetry, respectively. For this reason, we also refer to $q = 0$ only as the scalar representation even if for $U(1)$ symmetry all symmetry multiplets are actually one-dimensional and, in that sense, scalars. Examples for scalar operators are the full Hamiltonian, as well as all of its terms in its sum including local one-site terms. Vacuum states transform like a scalar multiplet.

The dual representation q^* of some given irrep q is defined by the unique representation that allows one to form a scalar, i.e., with Clebsch-Gordan coefficients (CGCs) $(q, q^*; 0)$. These CGCs when properly normalized define a unitary matrix $U^{[q]}$, which is referred to as $1j$ symbol by analogy, e.g., to $3j$ symbols for the $SU(2)$ spin symmetry, with the difference that here only a single irrep label is concerned. While $SU(2)$ symmetry is self-dual, i.e., $q^* = q$, Abelian $U(1)$ symmetries are not, since one has $q^* = -q$ such that $q + q^* = q'$ properly adds up to the

scalar representation $q' = 0$. For self-dual symmetries, such as $SU(2)$, however, one must be careful in that $U^{[q]}$, when written simply reduced to a unitary matrix of rank-2, becomes indistinguishable from $(U^{[q]})^\dagger = U^{[q^']}$, which, however, may differ by a sign, e.g., for half-integer spins in the case of $SU(2)$. Importantly, $1j$ symbols allow one to revert arrows in lines in a tensor network by inserting $\mathbb{I} = U^{[q]^\dagger} U^{[q]}$.

1. Automata approach

First, we briefly recapitulate the automata approach for constructing MPOs of the Hamiltonian [49–51]. Consider, for example, the quantum Ising chain $H = \sum_i S_i^x S_{i+1}^x - h S_i^z$, which provides a simple example of a Hamiltonian with a single nearest-neighbor interaction term together with a local term (here with magnetic field strength h). We need to compute and store the matrix elements of the spin operators $\{S_x, S_z\}$. Together with the identity operator \mathbb{I} , these form a basis of local operators that enter a rank-4 tensor $T_{\alpha, \alpha', \sigma, \sigma'}$, where by rank we refer to the number of indices (or legs in a graphical depiction) of a given tensor. The tensor T is the elementary local tensor of the MPO, with σ the local state space of a given site and α the virtual bond states that tie together the MPO. The tensor T has the same form for every site due to the translational invariance of the Hamiltonian [assuming an open boundary condition, the open virtual indices of the T tensors for the first and last sites are contracted (“capped”) with a start and a stop state, respectively; see below].

To be concrete, each tensor T contains $D_H^2 d^2$ matrix elements, where d is the dimension of the local state space σ and D_H is the bond dimension of the virtual state space α . Every matrix element of T in the indices (α, α') is linked to a local operator with matrix indices (σ, σ') . It is therefore natural to group the relevant local operators into an (orthogonal) set that we will also index below. For the Ising model above, for example, the relevant set of local operators is given by $\{\mathbb{I}, S_x, S_z\}$.

The virtual bond state space is given by a start state $[\alpha = 1 \text{ or, equivalently, } (1, 0, 0, \dots)^T]$ and a stop state $[\alpha = 2 \text{ or, equivalently, } (0, 1, 0, \dots)^T]$, followed by $\alpha = 3, \dots, m_{\text{int}} + 2$, which assigns an index position to every one of the m_{int} interaction terms in the Hamiltonian that stretches across a given bond in the MPO (strictly speaking, m_{int} corresponds to the number of operators that need to be stored across a given bond, which may be less than the number of elementary interaction terms in the Hamiltonian if interaction terms can be grouped by factoring out specific operators). Hence, the dimension of the virtual bond state space is given by $D_H = 2 + m_{\text{int}}$. For example, for the Ising model above, the i th bond in the system in between sites $i' \leq i$ and $j' \geq i + 1$ carries the single interaction term $S_i^x S_{i+1}^x$; hence, $m_{\text{int}} = 1$ and $D_H = 3$.

The general strategy then for setting up the MPO with respect to the specific example of the Ising model is as follows: Starting from the left end, the bond state space, i.e., the automaton, is initialized in the start state ($\alpha = 1$). This state is carried through the MPO (therefore, $T_{1,1} = \mathbb{I}$) until an interaction term in the Hamiltonian occurs, say, at site i , which brings the automaton into the state $\alpha = 3$ (therefore, $T_{1,3} = S^x$). Having only nearest-neighbor terms, the subsequent T tensor, e.g., at site $i + 1$, immediately brings down the automaton to the end state $\alpha = 2$ (therefore, $T_{3,2} = S^x$). By having completed the interaction term, the automaton stays in that state (hence, $T_{2,2} = \mathbb{I}$). Overall, what has been encoded this way is simply the interaction term $\mathbb{I}_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_{i-1} \otimes S_i^z \otimes S_{i+1}^z \otimes \mathbb{I}_{i+2} \otimes \cdots \otimes \mathbb{I}_L$. With the same line of argument, the local transverse field term, say, at site i , is described by $T_{1,2} = -hS^z$, which directly brings the automaton from the start into the end state. By translational invariance, there is nothing special about site i , though. Therefore, all of the matrix elements of the tensor T specified above must hold for every site. Given these local tensors in MPO, the summation (trace) over geometric indices α is equivalent to adding up all the interaction terms in the total Hamiltonian.

2. From super-MPS to MPO

In the presence of global continuous symmetries, all state spaces must be organized into symmetry multiplets. Naturally, this also implies a directedness of lines in a tensor network. From the point of view of a given tensor, the direction on its lines indicates the bra or ket nature of these state spaces which, in pictorial language, is equivalent to legs (lines) entering or leaving a given tensor, respectively. For U(1) symmetries, it implies that the sum of all charges that enter a tensor must be exactly equal to the sum of all charges leaving it. For SU(2) symmetries, the fusion of all ingoing lines must result in a symmetry sector that exactly matches a symmetry sector resulting from the fusion of all outgoing lines. If all lines are ingoing, the tensor must be scalar in that the (skipped) outgoing line transforms as a singleton index that transforms like the vacuum state (and vice versa, if all lines are outgoing).

In contrast to the Ising model above, which has no simple continuous symmetry for $h \neq 0$, let us continue with the model system of interest in this work, the (anisotropic) Heisenberg model [compare Eq. (2), using $J := 1$]

$$\hat{H} = \sum_{\langle i,j \rangle} \underbrace{\hat{S}_i^x \hat{S}_j^x + \hat{S}_i^y \hat{S}_j^y + \Delta \hat{S}_i^z \hat{S}_j^z}_{=\frac{1}{2}(\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+)} \quad (\text{B1})$$

where we temporarily introduce hats on top of operators, in order to differentiate them from symmetry labels (e.g., \hat{S}^z versus S^z). The model in Eq. (B1) is U(1) symmetric, as it preserves S_{tot}^z . In the isotropic case $\Delta = 1$, it becomes SU(2) spin symmetric. Then, the spin operators need to be grouped into a spinor

$$\hat{\mathbf{S}} \equiv \begin{pmatrix} -\frac{1}{\sqrt{2}} \hat{S}^+ \\ \hat{S}^z \\ +\frac{1}{\sqrt{2}} \hat{S}^- \end{pmatrix} \quad (\text{B2})$$

such that Eq. (B1) can be rewritten in SU(2)-invariant form,

$$\hat{H} = \sum_{\langle i,j \rangle} \hat{\mathbf{S}}_i^\dagger \cdot \hat{\mathbf{S}}_j. \quad (\text{B3})$$

Note that the relative weights and signs in Eq. (B2) are important for consistency with standard conventions on SU(2) spin multiplets. In particular, the operators in the spinor in Eq. (B2) exactly represent, top to bottom, the states $S^z = (+1, 0, -1)$ of an $S = 1$ spin multiplet (e.g., see Ref. [48]). In contrast, the Hermitian set of operators $(\hat{S}^x, \hat{S}^y, \hat{S}^z)$ does not. However, using Eq. (B2), the dagger on one of the spinors in Eq. (B3) is important.

In general, a local operator acting on some physical site is a spinor, i.e., a collection of operators that transforms like some multiplet q [cf. Fig. 14(a)]. It can be written as the irreducible operator (IROP) $\hat{X}^{[nq;q_z]}$, where the composite index $(nq; q_z)$ naturally specifies entire state spaces [48] or here an operator space: The index n differentiates between local IROPs that transform according to the same irreducible representation q . By the definition of an index, $n = 1, 2, \dots$, we therefore also introduce an arbitrary but fixed order to the local operators. The label q_z , finally, fully differentiates the operators within a given spinor [48]. For example, within SU(2), q_z simply stands for S_z .

The matrix elements of IROPs are determined via the Wigner-Eckart theorem. For a generic spinor, this IROP acquires a third dimension, which indexes the operators in the irreducible set. Scalar operators then are special. With one ingoing and one outgoing index, the third index having $q = 0$ is a trivial singleton dimension that may safely be skipped. In this sense, scalar operators can be reduced to rank 2 and are block diagonal. For U(1) spin symmetry, for example, scalar operators are the identity operator $\hat{\mathbb{I}}$ or the spin projection operator \hat{S}^z . In contrast, the operators \hat{S}^\pm carry $q = \pm \frac{1}{2}$ hence switch between symmetry sectors, and therefore are not considered scalar operators.

All local operators eventually can be combined into a single rank-3 tensor \hat{X} [cf. Figs. 14(a) and 14(b)]. The third index then represents the state space $|nq; q_z\rangle$ [48] of the ‘‘supervectors’’ $X^{[nq;q_z]}$. For efficiency, the set of local operators should be orthogonal in the sense

$$\text{tr}[(X^{[nq;q_z]})^\dagger X^{[n'q';q'_z]}] \propto \delta_{nn'} \delta_{q,q'} \delta_{q_z,q'_z}, \quad (\text{B4})$$

with arbitrary normalization, otherwise. This is also in the spirit of an orthogonal local (super)state space of a (super-)MPS. Conversely, assuming that the tensor T of the MPO is given, note that the intermediate supervector index that

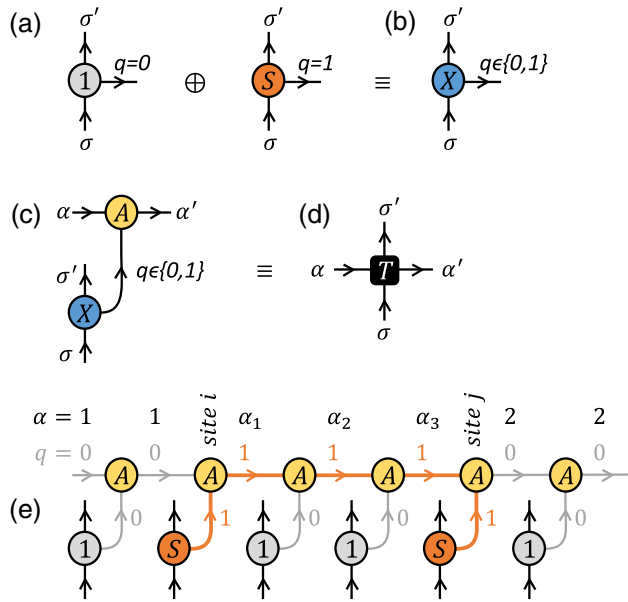


FIG. 14. (a) The local identity and spin operator for the SU(2) spin-symmetric Heisenberg model, which transform according to $S \equiv q = 0$ and $q = 1$, respectively. All relevant local operators, including the identity operator, can be combined into the rank-3 tensor X [(b)]. (c) The super-MPS described by the rank-3 tensor A can be contracted with the operator index in X to form the rank-4 local tensor T in the SU(2)-invariant MPO [(d)]. (e) Typical sequence that occurs in the construction of the MPO via the super-MPS that describes a specific individual interaction term $\hat{S}_i \cdot \hat{S}_j$. This demonstrates how the quantum number of the IROP S simply stretches like a string (red line) along the virtual bonds in between the two sites i and j where the spin operators act. The values $\alpha_1, \alpha_2, \dots$ on given A -tensors are reserved for this very specific interaction term, where, in general, the indices $\alpha_i > 2$ are not all the same. Note that the arrows on the red line are reversed with respect to site j , which thus needs to incorporate the Clebsch-Gordan coefficients $(1, 1; 0)$ that combine the $S = 1$ multiplet of the spinor S with its dual (also $S = 1$) into a scalar. It is this CGC on the A tensor of site j within the super-MPS that takes properly care of the dagger in the scalar product $\hat{S}_i \cdot \hat{S}_j^\dagger$. Similarly, also the coupling strength J is encoded with the A tensor in the super-MPS.

connects the super-MPS with the local operators \hat{X} [cf. Fig. 14(c)] may also be generated by the reverse operation of splitting off the local state space (σ, σ') from the tensor T via SVD. Then, by construction, the operators in \hat{X} would be orthonormal.

For both conceptual and implementational transparency, we can construct an MPO as a super-MPS of operators. By this, we mean that the local state spaces of the super-MPS are “superstates” that actually refer to a set of orthogonal local operators [e.g., see Figs. 14(a) and 14(b)]. By finally contracting the super-MPS (rank-3 tensors) with the local operators along the intermediate index, this leads to the final rank-4 tensors T that constitutes the MPO [cf. Figs. 14(c) and 14(d)]. Note that the intermediate index

also specifies an arbitrary but fixed order of the local operators (“supervectors”).

Now the structure of an interaction term as in Eq. (B3) is generic: A nonscalar irreducible operator $\hat{X}_i^{[q]}$ acting on site i must be paired up, i.e., contracted on the spinor index into a scalar term of the Hamiltonian, with another operator $(\hat{Y}_j^{[q]})^\dagger$ acting on site j that transforms according to exactly the same irreducible representation (typically, $\hat{Y} = \hat{X}$; here, we also ignore three- or more-site interactions). This observation holds for both Abelian and non-Abelian symmetries.

The construction of the super-MPS that encodes the MPO is greatly simplified by the simple bilinear structure of two-site interactions as in Eq. (B1) or Eq. (B3). In particular, the super-MPS can be built completely analogous to the automata approach above while paying simple attention to symmetry sectors. The start ($\alpha = 1$) and the stop ($\alpha = 2$) states on the virtual bonds transform like scalars (i.e., have $q = 0$), whereas the bond states $\alpha > 2$ directly inherit the symmetry labels from the underlying IROPs in the two-site interactions [cf. Fig. 14(e)].

For the Heisenberg model in Eq. (B1), the set of local operators is given by $\hat{X} = \{\hat{\mathbb{1}}, \hat{S}^z, \hat{S}^+, \hat{S}^-\}$ for the U(1) symmetric setup and by $\hat{X} = \{\hat{\mathbb{1}}, \hat{\mathbf{S}}\}$ for the SU(2) symmetric setup. In either case, the set of local operators is orthogonal as in Eq. (B4). Note also that, while in the U(1) symmetric case the daggered operator $(\hat{S}^+)^{\dagger} = \hat{S}^-$ appears in the set, such is not the case for the SU(2) symmetric case, since SU(2) is self-dual, and therefore $\hat{\mathbf{S}}_i^{\dagger} \cdot \hat{\mathbf{S}}_j = \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j^{\dagger}$. Specifically, with $U^{[1]} \propto (1, 1; 0)$ a unitary transformation that corresponds to the Clebsch-Gordan coefficients which combine a spin $S = 1$ with its dual (again $S = 1$) into a singlet [cf. $1j$ symbol earlier], the spin-spin interaction can be written as $\hat{\mathbf{S}}_i^{\dagger} \cdot \hat{\mathbf{S}}_j = \sum_{r,r'=1}^3 \hat{S}_i^r U_{rr'}^{[1]} \hat{S}_j^{r'}$. Therefore, the action of the dagger on one of the spin operators can be transferred via the unitary $1j$ symbol into the definition of the super-MPS itself, proper sign convention on $U^{[1]}$ implied.

For more complicated cases, like the snake MPO representation of 2D Heisenberg Hamiltonian, longer-range interactions need to be included. This is straightforward in the automata construction above yet requires that the bond dimension D_H increases [see also Fig. 14(e)]. The period of the translational invariance of the MPO also increases from 1 to the width W of the system and, hence, requires at least W different A tensors in the super-MPS [cf. Figs. 14(c) and 14(d)].

Once the super-MPS is obtained, one can use standard MPS techniques to check whether it can be compressed. An important ingredient here is that the local supervector space is orthogonal, indeed [cf. Eq. (B4)]. If the bond dimension D_H can be reduced at no cost, i.e., by discarding singular values that are strictly zero, the super-MPS and subsequently the MPO contain inefficiencies that may be

simply removed with an improved setup of the super-MPS itself. On the other hand, for long-ranged systems, the bond dimension D_H may simply become too large, in practice, for an exact representation of the Hamiltonian. In this case, standard MPS truncation techniques may be employed on the level of the super-MPS itself. Here, a uniform normalization of the supervector space (i.e., the local operators) is advised such that standard MPS techniques are directly applicable without any further ado. Alternatively, one may truncate on the level of the MPO, either by SVD or variational techniques. The latter is unavoidable for MPO products or sums in any case, as discussed next.

APPENDIX C: INITIALIZATION IN THE XTRG ALGORITHM

In this Appendix, we compare three different initializations of $\rho(\tau_0)$ in the XTRG algorithm. The quality of our initial $\rho(\tau_0)$ for small τ_0 is measured by estimating the relative error of the free energy $|\delta F(\beta)/F(\beta)|$, starting from exponentially small $\beta = 2\tau_0$ (i.e., the first data point after initialization at $\beta = \tau_0$) down to intermediate temperatures $\beta \gtrsim 10$.

1. Series expansion versus Trotter-Suzuki initialization

First, we compare the Trotter-Suzuki initialization with series expansion [Eq. (7) in the main text] followed by XTRG. Trotter-Suzuki decomposition breaks $e^{-\tau_0 H}$ into a product of local evolution gates, which for nearest-neighbor spin-1/2 chains within first order can be represented as an MPO with bond dimension $D_0^* = 2$ ($D_0 = 4$), comprised of $\frac{1}{2} \otimes \frac{1}{2} = \underline{1}^1 \oplus \underline{3}^1$. In Fig. 15, we plot the relative errors of the free energy after Trotter-Suzuki initialization at two values of $\tau_0 = 0.1$ and $\tau_0 = 0.01$. At $\beta = 2\tau_0$, the respective errors $|\delta F/F|$ are 10^{-5} and 10^{-9} , respectively.

Interestingly, first-order Trotter-Suzuki initialization manages to arrive at an MPO with a bond dimension that is *lower* than what is required for the actual representation of the Hamiltonian itself. But, as a consequence, the overall errors are also larger. For comparison, nevertheless, we also show data initialized via SETTN at the same $D_0^* = 2$ ($D_0 = 4$; green data). The resulting errors are much lower: 10^{-10} and 10^{-15} for $\tau_0 = 0.1$ and 0.01 , respectively.

By increasing the initial bond dimension to D_H^* and slightly above, i.e., $D_0^* = 3, 4$ ($D_0 = 5, 8$), as seen in Fig. 15, SETTN initialization as in Eq. (7) can offer generally better accuracy. The initial inaccuracy represents a systematic error that also propagates along the XTRG procedure towards lower temperatures. For example, for $\beta \sim 10$, the SETTN initialization with $D_0^* = 2$ is still an order of magnitude more accurate as compared to the Trotter-initialized data and orders of magnitude more accurate if D_0 is only marginally increased to $D_0^* = 3, 4$ relative to the $D^* = 150$ of the subsequent XTRG.

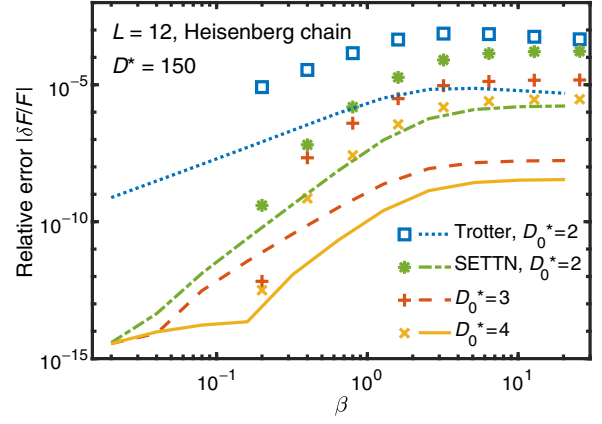


FIG. 15. Comparison between the first-order Trotter-Suzuki and the series-expansion initialization schemes on an $L = 12$ Heisenberg chain. Two sets of data are shown from calculations starting with $\tau_0 = 0.1$ (symbols) and $\tau_0 = 0.01$ (lines). After the initialization, the MPO $e^{-\tau_0 H}$ is fed into an XTRG evolution, where the number of retained bond states is set to $D^* = 150$. The same color represents the same type of initialization; i.e., blue represents Trotter initialization, while green, red, and yellow represent initialization by SETTN with bond dimensions $D_0^* = 2, 3, 4$ ($D_0 = 4, 5, 8$), respectively.

2. Linear initialization $\rho(\tau_0) \simeq \mathbb{I} - \tau_0 H$

By sweeping over several orders of energy scales, the XTRG algorithm permits a simple linear initialization $\rho(\tau_0) \simeq \mathbb{I} - \tau_0 H$ at basically infinitesimal, i.e., exponentially small, τ_0 . In the following, we analyze the effect of the initial τ_0 in more detail.

In Fig. 16, we compare the initialization of $\rho(\tau_0)$ with SETTN [Fig. 16(a)] and linear expansion [Fig. 16(b)], respectively. We benchmark the accuracy by analyzing the relative errors of the free energy $\delta F/F$ with various initial $\tau_0 = 10^{-8}, \dots, 10^{-2}$. In Fig. 16(a), the initial state $\rho(\beta = \tau_0)$ is constructed via SETTN [cf. Eq. (7)] and constrained to the bond dimension D_0^* for $\beta = \tau_0$ only. For all other points at lower temperatures $\beta_n \equiv 2^n \tau_0$ with $n = 1, 2, \dots$, we use $D^* = 200$. With this setup, the data in Fig. 16(a) at the lowest temperatures (largest β) exhibits a similar level of accuracy, irrespective of the choice of the initial $\tau_0 \lesssim 10^{-4}$ for $D_0^* \geq D_H^* = 6$. Note that $D_0^* = D_H^*$ is the minimal bond dimension to represent the lowest-order linear expansion $\rho(\tau_0) \simeq \mathbb{I} - \tau_0 H$. The long extended straight slopes for $\tau_0 = 10^{-6}$ or $\tau_0 = 10^{-8}$ are simply $\propto \beta$, as indicated by the guide to the eye (purple dashed line), which just indicates that the accuracy is limited by the accumulated double precision error of the calculation. The strong upturn for $\beta \gtrsim 0.1$ then is where the truncation error sets in. Using a larger initial τ_0 , e.g., $\tau_0 = 10^{-4}$ (yellow curve), the error accumulates more strongly, which implies that $D_0^* = 6$ already starts to affect the accuracy at larger β . Clearly, for $\tau \gtrsim 10^{-4}$ a larger D_0^* , i.e., higher-order terms in the series expansion, are required to maintain accuracy. For

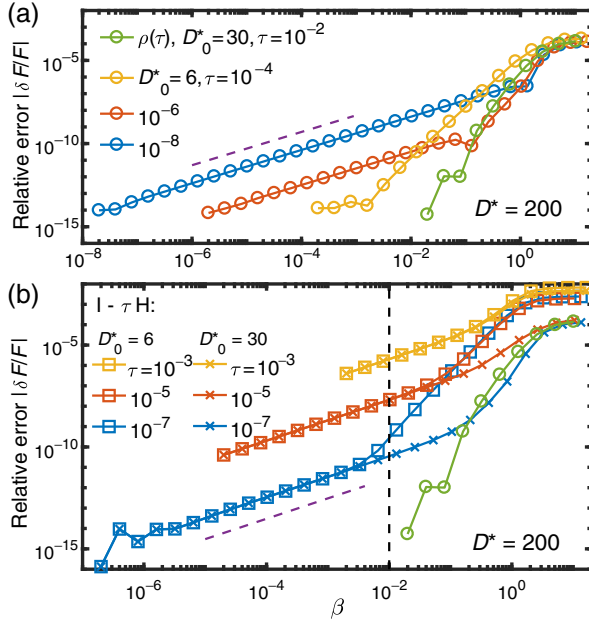


FIG. 16. (a) Relative error of free energy for the Heisenberg model on a 4×4 open square lattice ($D_H^* = 6$), with various initial τ and D_0^* values in the XTRG scheme. Here, the $\rho(\beta = \tau_0)$ is initialized via SETTN [cf. Eq. (7)]. (b) Similar to (a) but using the lowest-order, i.e., linear, expansion of $\rho(\tau_0) = \mathbb{I} - \tau_0 H$ for the initialization of $\rho(\tau_0)$ instead of SETTN [only the green curve at $\tau_0 = 10^{-2}$ uses SETTN and is copied from (a) for direct reference]. Here, we use $D^* = D_0^*$ for all temperatures down to $\beta = 10^{-2}$ (black dashed line), where we switch to $D^* = 200$. Purple dashed lines indicate linear behavior, i.e., $|\delta F/F| \propto \beta$.

example, the data for $\tau_0 = 10^{-2}$ and $D_0^* = 30$ again shows good accuracy in comparison.

Once τ_0 is small enough, plain simple lowest-order linear expansion suffices. This is analyzed in Fig. 16(b), where we replace series expansion for the initialization of $\rho(\tau_0)$ by the MPO for $\mathbb{I} - \tau_0 H$. This is an extremely convenient initialization that can be simply derived from the MPO for the Hamiltonian. In particular, the MPO for this initial $\rho(\tau_0)$ can be exactly represented with bond dimension $D_0^* = D_H^*$.

Starting with $\beta = \tau_0 \lll 10^{-2}$, XTRG can be used to exponentially decrease the temperature down to values of $\beta \approx 10^{-2}$, which still may be considered part of the initialization of ρ at large temperatures before actual physical energy scales set in. Therefore, while we always have $D_0^* = D_H^*$ for the very first step, by definition, we can also constrain D_0^* to the values specified for the entire range $\beta \leq 10^{-2}$ [vertical dashed line in Fig. 16(b)]. For $\beta > 10^{-2}$, we allow the MPO to grow up to dimension $D^* = 200$ to capture the quickly growing MPO entanglement.

Using $D_0^* = D_H^*$ up to $\beta = 10^{-2}$ [squares in Fig. 16(b)], the data already significantly deteriorates at the largest $\beta \gg 1$ by about 2 orders of magnitude. This shows that $D_0^* = D_H^*$ up to $\beta = 10^{-2}$ is simply too small, as it introduces

systematic errors. By increasing D_0^* modestly, i.e., $D_0^* = 30$, which is still about an order of magnitude smaller than what is required for large β , the systematic errors reduce dramatically. Starting with $\tau_0 \lesssim 10^{-4}$, the accuracy at large β competes with a careful higher-order expansion for finite τ_0 [for reference, we replot the data from the $D_0^* = 30$ series-expansion initialization starting with $\tau_0 = 10^{-2}$ from Fig. 16(a)]. Similar to Fig. 16(a), for our data sets with smallest initial τ_0 , we see wide ranges where the numerical error is simply $\propto \beta$ and, hence, given by the accumulated double precision error and, in particular, not the truncation error.

The initialization procedure in Fig. 16(b) above starts from infinitesimally small $\beta = \tau$ and works its way up exponentially using XTRG to $\beta = 10^{-2}$. Using $D_H^* < D_0^* \ll D_{(\text{final})}^*$, the overall numerical cost of the entire calculation is strongly dominated by the simulation of $\beta \gg 1$. For example, for $D_0^* = 6$ (30) and $\tau_0 = 10^{-7}$, the cost for $\beta \leq 10^{-2}$ is about 7% and 11% of the total calculation, respectively. In this sense, the procedure above is an extremely simple, efficient, and accurate initialization for finite-temperature calculations. The algorithm works for arbitrary topologies of Hamiltonians, including long-range interactions or higher-dimensional systems.

Nevertheless, since the series-expansion scheme serves as a systematic way to provide accurate initialization at finite $\tau_0 < 0.1$, in this work, we stick to initializing the density operator $\rho(\tau_0)$ with our already existing codes on series expansion.

APPENDIX D: COMPRESSION OF MATRIX PRODUCT OPERATORS

MPO compression is of key importance and frequently used in the XTRG and SETTN algorithm, to compress the *product* or *sum* of two MPOs. The overall procedure follows standard MPS strategies, where the implications of Abelian or non-Abelian symmetries can be largely put aside as an extremely convenient benefit of using the QSpace tensor library.

Overall, the variational method is preferable due to its higher efficiency, while the direct SVD compression is also useful as long as the bond dimension D is manageable. In the following, we focus on the variational compression but also provide details of the SVD compression along the way.

1. Compression of MPO product

Consider the product of two MPSs $A * B = C$, with bond dimensions D_a , D_b , and D_c , respectively. The representation of the product is exact if $D_c = D_a D_b$, which, however, is typically numerically costly. Therefore, C needs to be compressed in an efficient and numerically controlled manner. A typical example in this paper is the representation of $C := H^n X = H * (H^{n-1} X)$, with $X = \mathbb{I}$ or $\rho(\beta)$ in SETTN, where we need to project $A := H$ to a previously,

e.g., iteratively, obtained MPO for $B := H^{n-1}X$ and then compress the “fat” MPO to bring down its bond dimensions. Similarly, in XTRG, one needs to apply $\rho(\beta)$ onto itself, i.e., $A = B := \rho(\beta)$, in order to reach the density matrix $C := \rho(2\beta)$ at half the temperature.

The direct SVD compression is straightforward [42] but computationally costly. Given the product of two MPOs, as depicted pictorially in Fig. 2(b), horizontal lines are fused pairwise with respect to the same horizontal bond position into a single “fat” index of dimension $D_a D_b$ and then truncated via SVD. Using symmetries, Abelian or non-Abelian symmetries alike, the fusion step includes a simple tensor product of two state spaces. For this, it is important that the arrows along the virtual bond state spaces (horizontal lines) are parallel and point all in the same direction [Fig. 2(b)].

The computational cost of SVD compression scales as $\mathcal{O}(D_a^3 D_b^3)$. For the series expansions, say, when constructing $H^n \rho$ with $D_a = D_H$ and $D_b = D$, this is still relatively cheap, $\mathcal{O}(D^3)$, given that typically $D_H \ll D$. In contrast, the numerical cost becomes quickly prohibitive for XTRG, since with $D_a = D_b = D$ the cost scales as $\mathcal{O}(D^6)$.

The variational method can significantly reduce the numerical cost and, therefore, is mainly adopted in the present study. For this, we use a two-site update similar to standard DMRG procedures to allow the adaptive adjustment of bond dimensions. This is particularly important when exploiting symmetries, Abelian and non-Abelian alike, to optimally adapt bond dimensions with respect to each individual symmetry sector. This way, irrelevant symmetry sectors drop out automatically, whereas possibly new relevant symmetry sectors can emerge or get strengthened.

For the variational approach, we minimize the cost function (Frobenius norm squared),

$$\begin{aligned} & \|(A * B) - C\|_F^2 \\ &= C^\dagger C - (A * B)^\dagger C - C^\dagger (A * B) + \text{const.} \end{aligned} \quad (\text{D1})$$

We then take the partial derivative with respect to the product of two adjacent local tensors $C_i C_{i+1}$ in the full MPO of C . This results in the linear system of equations

$$\frac{\partial [C^\dagger C]}{\partial (C_i C_{i+1})^*} = \frac{\partial [C^\dagger (A * B)]}{\partial (C_i C_{i+1})^*}, \quad (\text{D2})$$

to be solved iteratively for $i = 1, \dots, L - 1$, with L the length of the MPO (while all simulations in this work are based on real numbers, for the simplicity of the derivation, nevertheless, we assume complex numbers). Both sides of Eq. (D2) can be expressed as fully contracted tensor networks, except for the missing tensors C_i^* and C_{i+1}^* (i.e., with “punched holes”), as shown in Fig. 17(a). Therefore, both sides of Eq. (D2) represent tensors of

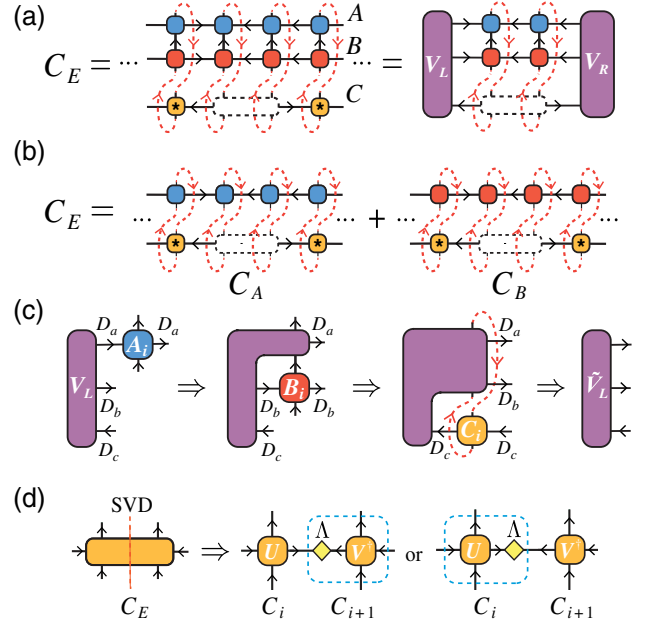


FIG. 17. (a) The environment tensor C_E in two-site variational optimization of compressing the product of two MPOs into a single MPO, i.e., $C := A * B$ [cf. Eq. (D2)]. Arrows on the horizontal lines demonstrate orthonormalization or, equivalently, canonicalization of the MPO. The environment C_E can be evaluated by contracting the cluster which consists of two environment tensors V_L and V_R , as well as relevant MPO tensors. (b) The environment tensor C_E in two-site variational scheme of rewriting the sum of two MPOs into a single MPO, i.e., $C := A + B$ [cf. Eq. (D4)]. An asterisk inside a box indicates a “dagged” tensor. (c) Update the left environment tensor V_L with three local tensor A_i , B_i , and C_i of MPOs A , B , and C , respectively. The computational costs of the three substeps (from left to right) are $\mathcal{O}(D_a^2 D_b D_c)$, $\mathcal{O}(D_a D_b^2 D_c)$, and $\mathcal{O}(D_a D_b D_c^2)$, respectively. V_R can be updated in a similar way. (d) Update of the local tensor T_i^* and T_{i+1}^* of MPO C , by performing SVD on either environment tensor C_E or $(C_A + C_B)$.

rank 6. For a canonicalized MPO C , where all lines in C are directed towards the *orthogonality center*, here, at sites $(i, i + 1)$, the left-hand side is simply $C_i C_{i+1}$. The right-hand side defines the generalized overlap tensor C_E [cf. Fig. 17(a)]. Equation (D2) therefore directly states the solution for $(C_i C_{i+1})$ to the local optimization problem with respect to sites $(i, i + 1)$.

To obtain C_E , one needs to iteratively update the left or right environment tensors V_L/V_R of the three-MPO product until the structure in the right-hand side in Fig. 17(a) reached. As shown in Fig. 17(c), to update the V_L/V_R tensors, we iteratively contract the local tensors A_i , B_i , and C_i of MPO A , B , and C with it. The procedures of updating V_L/V_R constitute the most time-consuming ones, which scale as $\mathcal{O}(D^4)$ (assuming $D_a = D_b = D_c =: D$), when we square ρ and compress it in XTRG. Nevertheless, this is still computationally much cheaper as compared to the

SVD compression of $\mathcal{O}(D^6)$ cost as briefly discussed above.

Given the optimized product $C_i C_{i+1} = C_E$, we need to split C_E into the actual product shape $C_i C_{i+1}$ by performing SVD, $C_E = U \Lambda V^\dagger$ as shown schematically in Fig. 17(d). To gauge the MPO C in a canonical form, C_i is updated with U , and C_{i+1} with ΛV^\dagger in a left-to-right sweep, while in a right-to-left sweep, the matrix Λ is contracted with U and thus associated with C_i , instead. One typically needs only a few full sweeps (left to right and vice versa) to converge the cost function and obtain the optimal MPO for C . In this work, typically at most $\lesssim 4$ sweeps were sufficient.

2. Compression of MPO sum

The summation of MPOs is an essential technique, e.g., in the series expansion $\rho(\beta) = \sum_{n=0}^N [(-\beta)/n!] H^n$. Here, we generalize standard procedures for the addition of MPS to MPO. For this, we also resort to a two-site variational approach as illustrated in Fig. 17(b). To find an optimal MPO for $C = A + B$, where the generalization to more than two vectors is straightforward, we minimize the cost function,

$$\begin{aligned} & \| (A + B) - C \|_F^2 \\ &= C^\dagger C - (A + B)^\dagger C - C^\dagger (A + B) + \text{const.} \end{aligned} \quad (\text{D3})$$

Again, we take the partial derivative with respect to the product of two adjacent local tensors $C_i C_{i+1}$ of the MPO C , resulting in the linear system of equations,

$$\frac{\partial(C^\dagger C)}{\partial(C_i C_{i+1})^*} = \frac{\partial[C^\dagger (A + B)]}{\partial(C_i C_{i+1})^*}, \quad (\text{D4})$$

to be solved iteratively for $i = 1, \dots, L - 1$. Again, using a canonicalized MPO for C , Eq. (D4) simply reduces to $C_i C_{i+1} = C_A + C_B$, with C_A , etc., generalized overlap matrices to be computed iteratively [cf. Fig. 17(b)]. The remainder of the algorithm proceeds exactly the same as the compression of MPO products above.

APPENDIX E: SERIES EXPANSION TENSOR NETWORK SIMULATIONS WITH LINEAR VERSUS EXPONENTIAL β GRID

In this Appendix, we compare SETTN calculations with three schemes of selecting expansion point set: (M) the Maclaurin scheme which expands $\rho(\beta)$ around $\beta = 0$ —i.e., there is only one expansion point in the set; (L) pointwise Taylor expansion around a linear β set, $\beta = n\tau_0$ with n an integer; (X) exponential set $\beta = 2^n \tau_0$. The results are summarized in Fig. 18.

In Fig. 18(a), we compare the accuracy of the above grid setups (M, L, X) in the calculation of the free energy for an $L = 16$ Heisenberg chain. Compared to (M), both (L) and

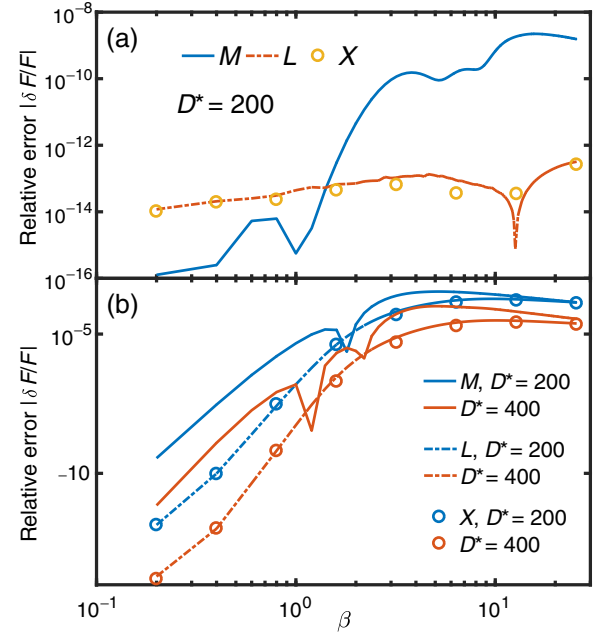


FIG. 18. Relative errors of the free energy fully within SETTN in (a) an $L = 16$ Heisenberg chain and (b) a 4×4 Heisenberg model on the square lattice (OBC). The data is computed along various β grids, including Maclaurin (M) expansion around $\beta = 0$ and Taylor expansion in linear (L) and exponential (X) β scales. In (a), we use $D^* = 200$, resulting in relative CPU run times $M:L:X = 2.34:5.25:1$. In (b), we use $D^* = 200, 400$. For $D^* = 200$, this results in relative CPU run times $M:L:X = 2.1:4.95:1$ and, for $D^* = 400$, $M:L:X = 1.4:4.2:1$.

(X) are clearly superior, as they gain 4 orders of magnitude in accuracy for the largest β .

We also compare these three grid setups within SETTN to the more challenging system of a 4×4 SLH. Here, due to the significantly larger truncation errors across all approaches, the gain of (L, X) over (M) is significantly less pronounced, as seen in Fig. 18(b). To reduce the truncation error, the number of virtual bond states would have to be increased significantly from the $D^* = 200$ or even 400 in the present calculation. As for the numerical efficiency, although having very good accuracy, the computational overhead of the linear scheme (L) is significant in any case. In contrast, the logarithmic scheme (X) strongly reduces the computational cost (see the caption of Fig. 18 for explicit numbers) without losing any accuracy. By minimizing the number of intermediate and thus also truncation steps by moving to large β in the fastest possible way, the numerical errors of (X) are on the lower end.

APPENDIX F: EXACT SOLUTION OF THE XY CHAIN AT FINITE TEMPERATURE

For completeness, we provide an analytical expression of the partition function in a 1D XY model. The Hamiltonian is given by

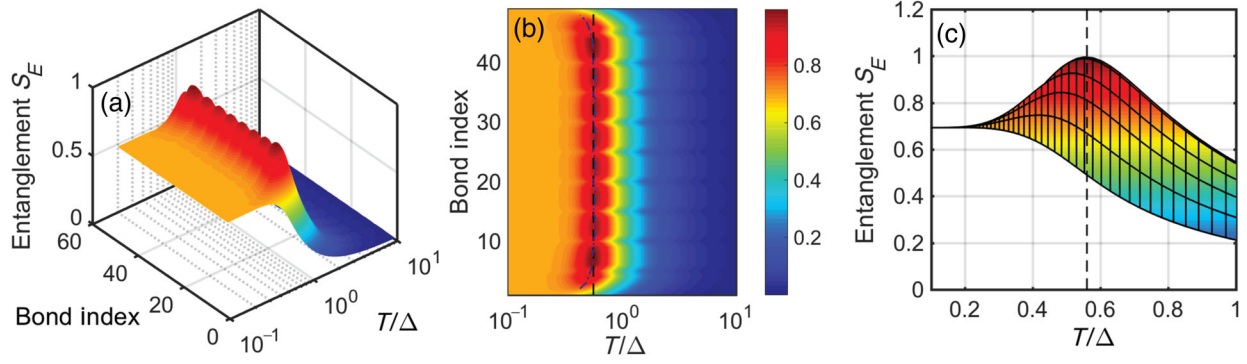


FIG. 19. (a) Block entanglement landscapes versus temperature T/Δ and bond index for the XXZ model on a square lattice (using $\Delta = 5$, $L = 10$, $W = 5$ keeping up to $D = 500$ states, i.e., similar analysis as in Fig. 13 but for a smaller system size). (b) Top view and (c) side view versus T/Δ of the same data as in (a). The black dashed line represents the exact value $T_c = 0.56$, and the dash-dotted line is tracking the “ridge” of entanglement landscape (i.e., maxima versus T on each bond) in (a); the latter coincides with T_c in the central bonds of the system. Incidentally, the *envelope* of entanglement curves also shows a peak around T_c (with the resolution $\delta T/\Delta \simeq 0.02$).

$$\begin{aligned}
 H &= J \sum_i (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y) \\
 &\equiv \frac{J}{2} \sum_i (S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+). \quad (\text{F1})
 \end{aligned}$$

Exploiting the Jordan-Wigner transformation, the Hamiltonian can be mapped onto a plain fermionic tight-binding chain,

$$H = \frac{J}{2} \sum_i (c_i^\dagger c_{i+1} + \text{H.c.}) = \sum_{k=1}^L \epsilon_k c_k^\dagger c_k, \quad (\text{F2})$$

which for an OBC is diagonalized by the one-particle eigenstates,

$$\begin{aligned}
 c_k &= \sqrt{\frac{2}{L+1}} \sum_{i=1}^L c_i \sin\left(\frac{k\pi}{L+1} i\right), \\
 \epsilon_k &= J \cos\left(\frac{k\pi}{L+1}\right), \quad (\text{F3})
 \end{aligned}$$

where $c_i^{(\dagger)}$ ($c_k^{(\dagger)}$) are fermionic annihilation (creation) operators at site i (“momentum” k), respectively [9,11].

The partition function is fully determined by the dispersion ϵ_k ,

$$\mathcal{Z}_{\text{OBC}}^{XY} = \prod_{k=1}^L (1 + e^{-\beta \epsilon_k}) \quad (\text{F4})$$

with free energy $F = -T \ln \mathcal{Z}$.

APPENDIX G: ENTANGLEMENT MEASUREMENTS OF XXZ MODEL ON THE CYLINDER GEOMETRY

The entanglement landscape, along with its top and side views, for the 2D XXZ model on a square-lattice cylinder ($L = 10$, $W = 5$, $\Delta = 5$) is analyzed in Fig. 19. Similar to

the analysis of the larger system in Fig. 13, we can again observe prominent peaks which can help to pinpoint the critical temperature T_c [vertical dashed lines in Figs. 19(b) and 19(c)]. In the top view in Fig. 19(b), we again see that the peak position T_c^S [represented as the dash-dotted lines in Fig. 19(b)] in the very center of the system deviates slightly from $T_c = 0.56$ in that it approaches from above, i.e., $T_c^S > T_c$. This is different from Fig. 13 in the main text, where T_c^S approaches T_c from the low-temperature side. The underlying reason is the different boundary condition: cylindrical BC here versus fully open BC in Fig. 13.

Interestingly, we find that T_c^S determined from cylinder geometry again serves as a good estimate for T_c . As shown in Fig. 19(c), we can see that the profile of entanglement curves overall (i.e., selecting the maximal entanglement over bonds for any given temperature), dubbed the entanglement envelope, shows a peak in close proximity to $T_c = 0.56$.

-
- [1] J.-S. Xiang, C. Chen, W. Li, X.-L. Sheng, N. Su, Z.-H. Cheng, Q. Chen, and Z.-Y. Chen, *Criticality-Enhanced Magnetocaloric Effect in Quantum Spin Chain Material Copper Nitrate*, *Sci. Rep.* **7**, 44643 (2017).
 - [2] B. Lake, D. A. Tennant, J.-S. Caux, T. Barthel, U. Schollwöck, S. E. Nagler, and C. D. Frost, *Multispinon Continua at Zero and Finite Temperature in a Near-Ideal Heisenberg Chain*, *Phys. Rev. Lett.* **111**, 137205 (2013).
 - [3] L. Balents, *Spin Liquids in Frustrated Magnets*, *Nature (London)* **464**, 199 (2010).
 - [4] J. Nasu, M. Udagawa, and Y. Motome, *Thermal Fractionalization of Quantum Spins in a Kitaev Model: Temperature-Linear Specific Heat and Coherent Transport of Majorana Fermions*, *Phys. Rev. B* **92**, 115122 (2015).
 - [5] Y. Shen *et al.*, *Evidence for a Spinon Fermi Surface in a Triangular-Lattice Quantum-Spin-Liquid Candidate*, *Nature (London)* **540**, 559 (2016).

- [6] S.-H. Do *et al.*, *Majorana Fermions in the Kitaev Quantum Spin System α -RuCl₃*, *Nat. Phys.* **13**, 1079 (2017).
- [7] S. Yamashita, Y. Nakazawa, A. Ueda, and H. Mori, *Thermodynamics of the Quantum Spin Liquid State of the Single-Component Dimer Mott System $\kappa - h_3(\text{Cat} - \text{EDT} - \text{TTF})_2$* , *Phys. Rev. B* **95**, 184425 (2017).
- [8] Z. A. Kelly, M. J. Gallagher, and T. M. McQueen, *Electron Doping a Kagome Spin Liquid*, *Phys. Rev. X* **6**, 041007 (2016).
- [9] H.-H. Tu, *Universal Entropy of Conformal Critical Points on a Klein Bottle*, *Phys. Rev. Lett.* **119**, 261603 (2017).
- [10] W. Tang, L. Chen, W. Li, X. C. Xie, H.-H. Tu, and L. Wang, *Universal Boundary Entropies in Conformal Field Theory: A Quantum Monte Carlo Study*, *Phys. Rev. B* **96**, 115136 (2017).
- [11] L. Chen, H.-X. Wang, L. Wang, and W. Li, *Conformal Thermal Tensor Network and Universal Entropy on Topological Manifolds*, *Phys. Rev. B* **96**, 174429 (2017).
- [12] S.-H. Do, S.-Y. Park, J. Yoshitake, J. Nasu, Y. Motome, Y. S. Kwon, D. T. Adroja, D. J. Voneshen, K. Kim, T.-H. Jang, J.-H. Park, K.-Y. Choi, and S. Ji, *Majorana Fermions in the Kitaev Quantum Spin System α -RuCl₃*, *Nat. Phys.* **13**, 1079 (2017).
- [13] K. W. Plumb, J. P. Clancy, L. J. Sandilands, V. V. Shankar, Y. F. Hu, K. S. Burch, H.-Y. Kee, and Y.-J. Kim, *α -RuCl₃: A Spin-Orbit Assisted Mott Insulator on a Honeycomb Lattice*, *Phys. Rev. B* **90**, 041112 (2014).
- [14] A. Banerjee, C. A. Bridges, J.-Q. Yan, A. A. Aczel, L. Li, M. B. Stone, G. E. Granroth, M. D. Lumsden, Y. Yiu, J. Knolle, S. Bhattacharjee, D. L. Kovrizhin, R. Moessner, D. A. Tennant, D. G. Mandrus, and S. E. Nagler, *Proximate Kitaev Quantum Spin Liquid Behaviour in a Honeycomb Magnet*, *Nat. Mater.* **15**, 733 (2016).
- [15] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area Laws for the Entanglement Entropy*, *Rev. Mod. Phys.* **82**, 277 (2010).
- [16] M. Žnidarič, T. Prosen, and I. Pižorn, *Complexity of Thermal States in Quantum Spin Chains*, *Phys. Rev. A* **78**, 022103 (2008).
- [17] M. B. Hastings, *Solving Gapped Hamiltonians Locally*, *Phys. Rev. B* **73**, 085115 (2006).
- [18] R. J. Bursill, T. Xiang, and G. A. Gehring, *The Density Matrix Renormalization Group for a Quantum Spin Chain at Non-Zero Temperature*, *J. Phys. Condens. Matter* **8**, L583 (1996).
- [19] X. Wang and T. Xiang, *Transfer-Matrix Density-Matrix Renormalization-Group Theory for Thermodynamics of One-Dimensional Quantum Systems*, *Phys. Rev. B* **56**, 5061 (1997).
- [20] T. Xiang, *Thermodynamics of Quantum Heisenberg Spin Chains*, *Phys. Rev. B* **58**, 9142 (1998).
- [21] A. E. Feiguin and S. R. White, *Finite-Temperature Density Matrix Renormalization Using an Enlarged Hilbert Space*, *Phys. Rev. B* **72**, 220401 (2005).
- [22] W. Li, S.-J. Ran, S.-S. Gong, Y. Zhao, B. Xi, F. Ye, and G. Su, *Linearized Tensor Renormalization Group Algorithm for the Calculation of Thermodynamic Properties of Quantum Lattice Models*, *Phys. Rev. Lett.* **106**, 127202 (2011).
- [23] S.-J. Ran, W. Li, B. Xi, Z. Zhang, and G. Su, *Optimized Decimation of Tensor Networks with Super-Orthogonalization for Two-Dimensional Quantum Lattice Models*, *Phys. Rev. B* **86**, 134429 (2012).
- [24] Y.-L. Dong, L. Chen, Y.-J. Liu, and W. Li, *Bilayer Linearized Tensor Renormalization Group Approach for Thermal Tensor Networks*, *Phys. Rev. B* **95**, 144428 (2017).
- [25] P. Czarnik, L. Cincio, and J. Dziarmaga, *Projected Entangled Pair States at Finite Temperature: Imaginary Time Evolution with Ancillas*, *Phys. Rev. B* **86**, 245101 (2012).
- [26] P. Czarnik and J. Dziarmaga, *Variational Approach to Projected Entangled Pair States at Finite Temperature*, *Phys. Rev. B* **92**, 035152 (2015).
- [27] P. Czarnik, M. M. Rams, and J. Dziarmaga, *Variational Tensor Network Renormalization in Imaginary Time: Benchmark Results in the Hubbard Model at Finite Temperature*, *Phys. Rev. B* **94**, 235142 (2016).
- [28] S. R. White, *Minimally Entangled Typical Quantum States at Finite Temperature*, *Phys. Rev. Lett.* **102**, 190601 (2009).
- [29] E. M. Stoudenmire and S. R. White, *Minimally Entangled Typical Thermal State Algorithms*, *New J. Phys.* **12**, 055026 (2010).
- [30] B. Bruognolo, Z. Zhu, S. R. White, and E. Miles Stoudenmire, *Matrix Product State Techniques for Two-Dimensional Systems at Finite Temperature*, *arXiv:1705.05578*.
- [31] T. Prosen and I. Pižorn, *Operator Space Entanglement Entropy in a Transverse Ising Chain*, *Phys. Rev. A* **76**, 032316 (2007).
- [32] T. Barthel, *One-Dimensional Quantum Systems at Finite Temperatures Can Be Simulated Efficiently on Classical Computers*, *arXiv:1708.09349*.
- [33] J. Dubail, *Entanglement Scaling of Operators: A Conformal Field Theory Approach, with a Glimpse of Simulability of Long-Time Dynamics in $1+1d$* , *J. Phys. A* **50**, 234001 (2017).
- [34] A. Bazavov, Y. Meurice, S.-W. Tsai, J. Unmuth-Yockey, Li-Ping Yang, and Jin Zhang, *Estimating the Central Charge from the Rényi Entanglement Entropy*, *Phys. Rev. D* **96**, 034514 (2017).
- [35] G. Evenbly and G. Vidal, *Tensor Network Renormalization Yields the Multiscale Entanglement Renormalization Ansatz*, *Phys. Rev. Lett.* **115**, 200401 (2015).
- [36] B. Swingle, *Entanglement Renormalization and Holography*, *Phys. Rev. D* **86**, 065007 (2012).
- [37] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, *Coarse-Graining Renormalization by Higher-Order Singular Value Decomposition*, *Phys. Rev. B* **86**, 045139 (2012).
- [38] M. Levin and C. P. Nave, *Tensor Renormalization Group Approach to Two-Dimensional Classical Lattice Models*, *Phys. Rev. Lett.* **99**, 120601 (2007).
- [39] K. G. Wilson, *The Renormalization Group: Critical Phenomena and the Kondo Problem*, *Rev. Mod. Phys.* **47**, 773 (1975).
- [40] R. Bulla, T. A. Costi, and T. Pruschke, *Numerical Renormalization Group Method for Quantum Impurity Systems*, *Rev. Mod. Phys.* **80**, 395 (2008).
- [41] A. Weichselbaum, *Tensor Networks and the Numerical Renormalization Group*, *Phys. Rev. B* **86**, 245124 (2012).
- [42] B.-B. Chen, Y.-J. Liu, Z. Chen, and W. Li, *Series-Expansion Thermal Tensor Network Approach for Quantum Lattice Models*, *Phys. Rev. B* **95**, 161104 (2017).

- [43] I. P. McCulloch and M. Gulácsi, *The Non-Abelian Density Matrix Renormalization Group Algorithm*, *Europhys. Lett.* **57**, 852 (2002).
- [44] S. Depenbrock, I. P. McCulloch, and U. Schollwöck, *Nature of the Spin-Liquid Ground State of the $s = 1/2$ Heisenberg Model on the Kagome Lattice*, *Phys. Rev. Lett.* **109**, 067201 (2012).
- [45] S.-S. Gong, W. Zhu, D. N. Sheng, O. I. Motrunich, and M. P. A. Fisher, *Plaquette Ordered Phase and Quantum Phase Diagram in the Spin- $\frac{1}{2}$ $J_1 - J_2$ Square Heisenberg Model*, *Phys. Rev. Lett.* **113**, 027201 (2014).
- [46] B. Bruognolo, J. von Delft, and A. Weichselbaum, *Symmetric Minimally Entangled Typical Thermal States*, *Phys. Rev. B* **92**, 115105 (2015).
- [47] M. Binder and T. Barthel, *Symmetric Minimally Entangled Typical Thermal States for Canonical and Grand-Canonical Ensembles*, *Phys. Rev. B* **95**, 195148 (2017).
- [48] A. Weichselbaum, *Non-Abelian Symmetries in Tensor Networks: A Quantum Symmetry Space Approach*, *Ann. Phys. (Berlin)* **327**, 2972 (2012).
- [49] G. M. Crosswhite and D. Bacon, *Finite Automata for Caching in Matrix Product Algorithms*, *Phys. Rev. A* **78**, 012356 (2008).
- [50] F. Fröwis, V. Nebendahl, and W. Dür, *Tensor Operators: Constructions and Applications for Long-Range Interaction Systems*, *Phys. Rev. A* **81**, 062337 (2010).
- [51] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, *Matrix Product Operator Representations*, *New J. Phys.* **12**, 025012 (2010).
- [52] C. Hubig, I. P. McCulloch, and U. Schollwöck, *Generic Construction of Efficient Matrix Product Operators*, *Phys. Rev. B* **95**, 035129 (2017).
- [53] F. Schwarz, I. Weymann, J. von Delft, and A. Weichselbaum, *Nonequilibrium Steady-State Transport in Quantum Impurity Models: A Thermofield and Quantum Quench Approach Using Matrix Product States*, arXiv:1708.06315 [Phys. Rev. Lett. (to be published)].
- [54] I. Affleck, *Universal Term in the Free Energy at a Critical Point and the Conformal Anomaly*, *Phys. Rev. Lett.* **56**, 746 (1986).
- [55] B. Bauer et al., *The ALPS Project Release 2.0: Open Source Software for Strongly Correlated Systems*, *J. Stat. Mech.* (2011) P05001.
- [56] P. W. Anderson, *Resonating Valence Bonds: A New Kind of Insulator?*, *Mater. Res. Bull.* **8**, 153 (1973).
- [57] P. W. Anderson, *The Resonating Valence Bond State in La_2CuO_4 and Superconductivity*, *Science* **235**, 1196 (1987).
- [58] N. Elstner, R. R. P. Singh, and A. P. Young, *Finite Temperature Properties of the Spin-1/2 Heisenberg Antiferromagnet on the Triangular Lattice*, *Phys. Rev. Lett.* **71**, 1629 (1993).
- [59] W. Zheng, R. R. P. Singh, R. H. McKenzie, and R. Coldea, *Temperature Dependence of the Magnetic Susceptibility for Triangular-Lattice Antiferromagnets with Spatially Anisotropic Exchange Constants*, *Phys. Rev. B* **71**, 134422 (2005).
- [60] S. A. Kulagin, N. Prokof'ev, O. A. Starykh, B. Svistunov, and C. N. Varney, *Bold Diagrammatic Monte Carlo Method Applied to Fermionized Frustrated Spins*, *Phys. Rev. Lett.* **110**, 070601 (2013).
- [61] B. Bernu, P. Lecheminant, C. Lhuillier, and L. Pierre, *Exact Spectra, Spin Susceptibilities, and Order Parameter of the Quantum Heisenberg Antiferromagnet on the Triangular Lattice*, *Phys. Rev. B* **50**, 10048 (1994).
- [62] B. Bernu and G. Misguich, *Specific Heat and High-Temperature Series of Lattice Models: Interpolation Scheme and Examples on Quantum Spin Systems in One and Two Dimensions*, *Phys. Rev. B* **63**, 134409 (2001).
- [63] L. Capriotti, A. E. Trumper, and S. Sorella, *Long-Range Néel Order in the Triangular Heisenberg Model*, *Phys. Rev. Lett.* **82**, 3899 (1999).
- [64] S. R. White and A. L. Chernyshev, *Néel Order in Square and Triangular Lattice Heisenberg Models*, *Phys. Rev. Lett.* **99**, 127004 (2007).
- [65] W. Zheng, J. O. Fjærestad, R. R. P. Singh, R. H. McKenzie, and R. Coldea, *Excitation Spectra of the Spin- $\frac{1}{2}$ Triangular-Lattice Heisenberg Antiferromagnet*, *Phys. Rev. B* **74**, 224420 (2006).
- [66] W. Zheng, J. O. Fjærestad, R. R. P. Singh, R. H. McKenzie, and R. Coldea, *Anomalous Excitation Spectra of Frustrated Quantum Antiferromagnets*, *Phys. Rev. Lett.* **96**, 057201 (2006).
- [67] Z. Zhu, P. A. Maksimov, S. R. White, and A. L. Chernyshev, *Topography of Spin Liquids on a Triangular Lattice*, *Phys. Rev. Lett.* **120**, 207203 (2018).
- [68] J. Alicea, O. I. Motrunich, and M. P. A. Fisher, *Theory of the Algebraic Vortex Liquid in an Anisotropic Spin- $\frac{1}{2}$ Triangular Antiferromagnet*, *Phys. Rev. B* **73**, 174430 (2006).
- [69] Y. Shirata, H. Tanaka, A. Matsuo, and K. Kindo, *Experimental Realization of a Spin-1/2 Triangular-Lattice Heisenberg Antiferromagnet*, *Phys. Rev. Lett.* **108**, 057205 (2012).
- [70] H. D. Zhou, C. Xu, A. M. Hallas, H. J. Silverstein, C. R. Wiebe, I. Umegaki, J. Q. Yan, T. P. Murphy, J.-H. Park, Y. Qiu, J. R. D. Copley, J. S. Gardner, and Y. Takano, *Successive Phase Transitions and Extended Spin-Excitation Continuum in the $s = 1/2$ Triangular-Lattice Antiferromagnet $\text{Ba}_3\text{CoSb}_2\text{O}_9$* , *Phys. Rev. Lett.* **109**, 267206 (2012).
- [71] T. Susuki, N. Kurita, T. Tanaka, H. Nojiri, A. Matsuo, K. Kindo, and H. Tanaka, *Magnetization Process and Collective Excitations in the $s = 1/2$ Triangular-Lattice Heisenberg Antiferromagnet $\text{Ba}_3\text{CoSb}_2\text{O}_9$* , *Phys. Rev. Lett.* **110**, 267201 (2013).
- [72] Y. Cui, J. Dai, P. Zhou, P. S. Wang, T. R. Li, W. H. Song, J. C. Wang, L. Ma, Z. Zhang, S. Y. Li, G. M. Luke, B. Normand, T. Xiang, and W. Yu, *Mermin-Wagner Physics, (h, t) Phase Diagram, and Candidate Quantum Spin-Liquid Phase in the Spin-1/2 Triangular-Lattice Antiferromagnet $\text{Ba}_8\text{CoNb}_6\text{O}_{24}$* , *Phys. Rev. Mater.* **2**, 044403 (2018).
- [73] W.-J. Hu, S.-S. Gong, W. Zhu, and D. N. Sheng, *Competing Spin-Liquid States in the Spin-1/2 Heisenberg Model on the Triangular Lattice*, *Phys. Rev. B* **92**, 140403 (2015).
- [74] S. Ito, N. Kurita, H. Tanaka, S. Ohira-Kawamura, K. Nakajima, S. Itoh, K. Kuwahara, and K. Kakurai, *Structure of the Magnetic Excitations in the Spin-1/2 Triangular-Lattice Heisenberg Antiferromagnet $\text{Ba}_3\text{CoSb}_2\text{O}_9$* , *Nat. Commun.* **8**, 235 (2017).
- [75] L. Chen, D.-W. Qu, S.-S. Gong, J. von Delft, A. Weichselbaum, and W. Li (to be published).

- [76] M. Zwolak and G. Vidal, *Mixed-State Dynamics in One-Dimensional Quantum Lattice Systems: A Time-Dependent Superoperator Renormalization Algorithm*, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [77] F. Verstraete and J. I. Cirac, *Renormalization Algorithms for Quantum-Many Body Systems in Two and Higher Dimensions*, [arXiv:cond-mat/0407066](https://arxiv.org/abs/cond-mat/0407066).
- [78] Y. Takahashi and H. Umezawa, *Thermo Field Dynamics*, *Int. J. Mod. Phys. B* **10**, 1755 (1996).
- [79] J. Maldacena, *Eternal Black Holes in Anti-de Sitter*, *J. High Energy Phys.* 04 (2003) 021.
- [80] A. B. Kallin, M. B. Hastings, R. G. Melko, and R. R. P. Singh, *Anomalies in the Entanglement Properties of the Square-Lattice Heisenberg Model*, *Phys. Rev. B* **84**, 165134 (2011).
- [81] B. M. Terhal, M. Horodecki, D. W. Leung, and D. P. DiVincenzo, *The Entanglement of Purification*, *J. Math. Phys. (N.Y.)* **43**, 4286 (2002).
- [82] J. Hauschild, E. Leviatan, J. H. Bardarson, E. Altman, M. P. Zaletel, and F. Pollmann, *Finding Purifications with Minimal Entanglement*, [arXiv:1711.01288](https://arxiv.org/abs/1711.01288).
- [83] P. Nguyen, T. Devakul, M. G. Halbasch, M. P. Zaletel, and B. Swingle, *Entanglement of Purification: From Spin Chains to Holography*, *J. High Energy Phys.* 01 (2018) 098.
- [84] T. Takayanagi and K. Umemoto, *Holographic Entanglement of Purification*, *Nat. Phys.* **14**, 573 (2018).
- [85] C. Holzhey, F. Larsen, and F. Wilczek, *Geometric and Renormalized Entropy in Conformal Field Theory*, *Nucl. Phys.* **B424**, 443 (1994).
- [86] P. Calabrese and J. Cardy, *Entanglement Entropy and Conformal Field Theory*, *J. Phys. A* **42**, 504005 (2009).
- [87] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, *Entanglement in Many-Body Systems*, *Rev. Mod. Phys.* **80**, 517 (2008).
- [88] H. F. Song, N. Laflorencie, S. Rachel, and K. Le Hur, *Entanglement Entropy of the Two-Dimensional Heisenberg Antiferromagnet*, *Phys. Rev. B* **83**, 224410 (2011).
- [89] S. Humeniuk and T. Roscilde, *Quantum Monte Carlo Calculation of Entanglement Rényi Entropies for Generic Quantum Systems*, *Phys. Rev. B* **86**, 235116 (2012).
- [90] B. Kulchytskyy, C. M. Herdman, S. Inglis, and R. G. Melko, *Detecting Goldstone Modes with Entanglement Entropy*, *Phys. Rev. B* **92**, 115146 (2015).
- [91] D. J. Luitz, X. Plat, F. Alet, and N. Laflorencie, *Universal Logarithmic Corrections to Entanglement Entropies in Two Dimensions with Spontaneously Broken Continuous Symmetries*, *Phys. Rev. B* **91**, 155145 (2015).
- [92] N. Laflorencie, D. J. Luitz, and F. Alet, *Spin-Wave Approach for Entanglement Entropies of the $J_1 - J_2$ Heisenberg Antiferromagnet on the Square Lattice*, *Phys. Rev. B* **92**, 115126 (2015).
- [93] M. A. Metlitski and T. Grover, *Entanglement Entropy of Systems with Spontaneously Broken Continuous Symmetry*, [arXiv:1112.5166](https://arxiv.org/abs/1112.5166).
- [94] V. Alba, M. Haque, and A. M. Läuchli, *Entanglement Spectrum of the Two-Dimensional Bose-Hubbard Model*, *Phys. Rev. Lett.* **110**, 260403 (2013).
- [95] F. Kolley, S. Depenbrock, I. P. McCulloch, U. Schollwöck, and V. Alba, *Entanglement Spectroscopy of $SU(2)$ -Broken Phases in Two Dimensions*, *Phys. Rev. B* **88**, 144426 (2013).
- [96] N. D. Mermin and H. Wagner, *Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models*, *Phys. Rev. Lett.* **17**, 1133 (1966).
- [97] S. Göttel, S. Andergassen, C. Honerkamp, D. Schuricht, and S. Wessel, *Critical Scales in Anisotropic Spin Systems from Functional Renormalization*, *Phys. Rev. B* **85**, 214406 (2012).
- [98] K. Binder, *Critical Properties from Monte Carlo Coarse Graining and Renormalization*, *Phys. Rev. Lett.* **47**, 693 (1981).
- [99] K. Binder, *Finite Size Scaling Analysis of Ising Model Block Distribution Functions*, *Z. Phys. B* **43**, 119 (1981).