Isometric Tensor Network States in Two Dimensions

Michael P. Zaletel[1] and Frank Pollmann[2,3]
[1]Department of Physics, University of California, Berkeley, California 94720, USA
[2]Technische Universität München, Physics Department T42, 85747 Garching, Germany
[3]Munich Center for Quantum Science and Technology (MCQST), 80799 München, Germany

(Received 26 February 2019; revised manuscript received 27 September 2019; published 24 January 2020)

Tensor-network states (TNS) are a promising but numerically challenging tool for simulating two-dimensional (2D) quantum many-body problems. We introduce an isometric restriction of the TNS ansatz that allows for highly efficient contraction of the network. We consider two concrete applications using this ansatz. First, we show that a matrix-product state representation of a 2D quantum state can be iteratively transformed into an isometric 2D TNS. Second, we introduce a 2D version of the time-evolving block decimation algorithm for approximating of the ground state of a Hamiltonian as an isometric TNS—which we demonstrate for the 2D transverse field Ising model.

DOI: 10.1103/PhysRevLett.124.037201

Overcoming the exponential growth of complexity when simulating quantum many-body systems is one of the most challenging goals in computational physics. For ground state properties of one-dimensional (1D) systems this challenge was answered by the density matrix renormalization group (DMRG) algorithm, which provides an essentially exact numerical solution of gapped 1D lattice models [1] and field theories [2]. Subsequently understood as a variational method over the class of matrix-product states (MPS) [3,4] its success follows from the ability of MPS to adequately capture the area-law entanglement characteristic of gapped ground states [5,6]. A central goal has been to generalize the success of DMRG to higher dimensions. For certain classes of states, this is achieved by so-called tensor-network states (TNS) whose connectivity reflects the geometry of many-body entanglement [7,8]. However, while evaluating properties of 1D MPS is highly efficient, *exactly* evaluating properties of TNS in higher dimensions is generically exponentially hard. Consequently, there has been a long-standing effort to determine the best way to numerically approximate TNS contractions in order to minimize the variational energy of TNS for a given Hamiltonian. Progress has been made for two-dimensional (2D) systems by introducing a number of algorithms to manipulate and optimize TNS for various lattice models [7–22]. However, at this point it is fair to say that the "right" way to generalize 1D DMRG is not yet agreed upon.

In this work, we study a restriction of the TNS ansatz, which we dub "isoTNS," which allows for highly efficient contraction of the network. When collapsing either the rows or columns of the 2D network it reduces to the canonical form of a 1D MPS [23,24]. As a result, any 1D MPS algorithm, such as DMRG [1] or the time dependent block decimation (TEBD) [23], can be turned into a 2D algorithm by applying it in a nested loop with respect to the rows and columns of the 2D isoTNS. While the ansatz we discuss is known to some practitioners [25], and is related to a previous work on correlated contour states [26], it does not seem to have been studied in practice. Here, we introduce a key procedure for manipulating isoTNS, the "Moses move" (MM), and demonstrate its utility with two concrete applications: First, we show that a 1D MPS representation of a 2D quantum state can be iteratively transformed into an isoTNS, and examine the resulting entanglement properties. Second, we implement a "TEBD^2" algorithm and use it to approximate the ground state of the 2D transverse field Ising model as an isoTNS.

The isometric tensor-network ansatz.—We first review the canonical form of a 1D MPS (see Refs. [27,28] for more details). Suppressing the indices of all tensors, the MPS for an N-site chain takes the form $\Psi = T^1 T^2 \cdots T^N$. Here, each T^a is a rank-3 tensor which we view as a $\chi_{a-1} \times \chi_a$ matrix in an "ancilla space" whose entries are vectors in the d-dimensional single-site Hilbert space of site a. Multiplication of the matrices implicitly comes with a tensor product over the single-site Hilbert spaces, producing an N-site wave function. At the boundaries, $\chi_0 = \chi_N = 1$. For any contiguous region of spins $V = a{:}b$, the partial contraction $T^{V\to\partial V} \equiv T^a \cdots T^b$ is a linear map from the Hilbert space $\mathcal{H}_V$ of the subregion to the $\chi_{a-1} \times \chi_b$ dimensional Hilbert space $\mathcal{H}_{\partial V}$ of the ancillas dangling from the boundary of the region. The "canonical form with ℓ-site center" is defined by requiring that the boundary map $T^{V\to\partial V}$ is an isometry if $V = 1{:}a$ for $a < \ell$ or $V = a{:}N$ for $a > \ell$. Recall a map is an isometry if $T^\dagger_{V\to\partial V} T_{V\to\partial V} = \mathbb{1}_{\partial V}$, while $T_{V\to\partial V} T^\dagger_{V\to\partial V} = P_V$ is a projection operator. The isometry condition ensures the ancillas on ∂V form an orthonormal sub-basis for V. In what

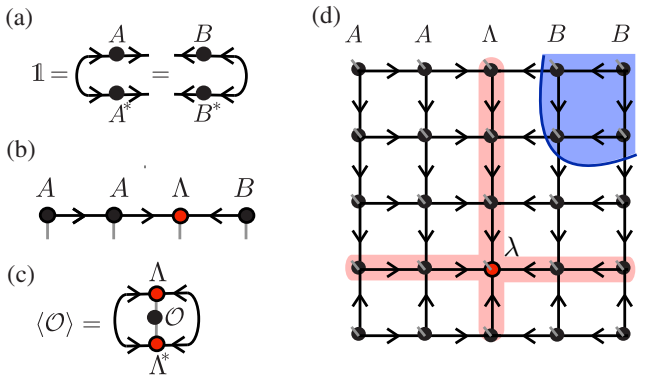

FIG. 1. Schematic representation of the canonical form in 1D and 2D. (a) Left and right isometries are represented by arrows whose orientation indicates whether $A^\dagger A = BB^\dagger = \mathbb{1}$. We view the isometry as an RG-like procedure from the large Hilbert space (incoming arrows) to the smaller one (outgoing arrows). In the case of higher-rank tensors, the contraction $A^\dagger A = \mathbb{1}$ is always over all the incoming arrows. (b) A 1D MPS can be brought into a mixed canonical form with orthogonality center Λ. Note that each dangling physical index implicitly has an incoming arrow. (c) Expectation values of local operators can be directly obtained from Λ. (d) 2D canonical form with "orthogonality hypersurfaces" Λ (column and row highlighted in red). The orthogonality center λ is marked by a red dot. In blue we indicate an example of a subregion with only outgoing arrows, whose boundary map is consequently an isometry.

follows, we denote the isometry conditions graphically by assigning arrows to the tensors as shown in Fig. 1(a) [29–31]. A convenient notation for the representation of MPS with ℓ-site center is to distinguish the tensors A, Λ, B and write

$$\Psi = A^1 A^2 \cdots A^{\ell-1} \Lambda^\ell B^{\ell+1} \cdots B^N \tag{1}$$

as shown in Fig. 1(b). It is easy to verify that the canonical form is satisfied if and only if each A^a, B^a is individually an isometry from the left and right, respectively. $T_{V\to\partial V}$ is an isometric boundary map if and only if the boundary ∂V has only outgoing arrows. On the other hand, a region with only incoming arrows, like the Λ, is precisely the wave function of the system expressed in an orthonormal basis, so it is called an *orthogonality center*. In particular, $\|\Psi\| = \|\Lambda\|$ and any site-ℓ expectation value can be locally computed as $\langle\Psi|O^\ell|\Psi\rangle = \langle\Lambda|O^\ell|\Lambda\rangle$, as seen in Fig. 1(c), because the A, B tensors in its exterior contract to 1 by the isometry condition.

Once the canonical form is understood as a restriction on the boundary maps, it can naturally be generalized to higher dimensions. By analogy to Eq. (1), we demand that each row and column of the TNS is an isometry, as indicated in Fig. 1(d). This constraint can be satisfied by further demanding that each tensor is an isometry from a physical and two ancilla legs to the remaining two ancillas according to the direction of the arrows indicated. This gives a causal structure to the tensor network, though in our convention time flows opposite to the direction of the arrows. As in 1D, there is a set of spacelike hypersurfaces with only outgoing arrows whose "past" defines the wave function in a orthonormal basis and whose "future" is an isometric boundary map. An expectation value $\langle\Psi|O|\Psi\rangle$ depends only on the tensors in the past of the insertion O. This is more than an analogy: the network between two spacelike surfaces defines a Kraus decomposition of a quantum channel relating the boundary ancilla.

There are a special row and column Λ [highlighted in red in Fig. 1(d)] which has only incoming arrows, which is hence the 1D "orthogonality hypersurface" of the TNS. Because its exterior is an isometry from the physical to the incoming ancillas, Λ is the wave function of the system in an orthonormal basis. Hence, Λ can be treated just like an MPS and can itself put into 1D canonical form (consequently *its* orthogonality center tensor λ can be moved freely using the standard 1D algorithm). Tracing over the left or right ancillas of Λ results in a density matrix which is isospectral to the reduced density matrix of the right or left, e.g., $\rho_L \sim \Lambda\Lambda^\dagger$, so Λ encodes the entanglement spectrum. For any operator O inside Λ, $\langle\Psi|O|\Psi\rangle = \langle\Lambda|O|\Lambda\rangle$, e.g., there is a dimensional reduction to a 1D expectation value which can be computed efficiently without further approximations via standard MPS algorithms. This is in stark contrast to generic TNS where expectation values require an approximate contraction of the entire network using, e.g., boundary MPS [8] or corner transfer matrices [9,10]. Furthermore, any variationally optimal compression of the orthogonality hypersurface Λ is variationally optimal for the global state. Note that by our choice of isometries, the resulting orthogonality hypersurface Λ has minimal entanglement for each vertical cut and is expected to follow a 1D area law. This entanglement is different from the vertical entanglement of the *full* many-body wave function: it differs by the action of the isometries, which contain the 2D area-law entanglement.

It is an interesting and open question how the variational power of an isoTNS differs from that of a generic TNS. One restriction is that many of its correlations must decay exponentially, because any two-point function along the orthogonality hypersurface can be reduced to that of the MPS Λ, which must have exponentially decaying correlations. In contrast, a generic 2D TNS can represent power-law correlations. On the other hand, we have shown that any string-net state, thought to represent all 2D topological orders with gappable edges, can explicitly be put into isoTNS form [32].

Shifting the orthogonality hypersurface.—The canonical form is only useful for computational purposes if the orthogonality hypersurface Λ can be moved throughout the network efficiently. In 1D, for example, the basic move $\Lambda^\ell B^{\ell+1} = A^\ell \Lambda^{\ell+1}$ can be accomplished by any orthogonal matrix factorization, i.e., QR or a singular value

decomposition (SVD). In 2D we need to solve the same equation but with A, Λ, B entire columns of the TNS. Using QR or SVD is hopeless, as it will destroy the locality required to express Λ as an MPS. The key insight is that the canonical form can be preserved under a unitary insertion $(A^\ell U^\dagger)(U\Lambda^{\ell+1})$. We propose to use this ambiguity to choose A^ℓ such that it "disentangles" $\Lambda^{\ell+1}$, so that $\Lambda^{\ell+1}$ has an efficient (low rank) MPS form.

It is actually sufficient to solve a simpler auxiliary problem: decompose $\Lambda^\ell = A^\ell\Lambda$, where Λ is a wave function with only ancilla degrees of freedom (d.o.f.) (a "zero-column" wave function). The start and end points of the problem are shown in sequences (i) and (v) of Fig. 2(a). This move will be sufficient to move Λ^ℓ throughout the network, because we can tack the zero-column wave function onto the right in order to obtain the one-column wave function, $\Lambda^{\ell+1} = \Lambda B^{\ell+1}$.

We can solve $\Lambda^\ell \approx A^\ell\Lambda$ as a variational problem, sweeping back and forth through the tensors to minimize $\|\Lambda^\ell - A^\ell\Lambda\|$ while respecting the isometry condition on A and reducing the bond dimension of Λ [33]. Interestingly, however, we find a *single* unzipping sweep based on disentangling provides a solution very close to the variational one [33], but is far quicker. This "Moses move" (MM) is illustrated in sequences (i) to (v) of Fig. 2(a).

The central subproblem of the MM [Fig. 2(b)] takes in the orthogonality center $|\lambda\rangle$, which by grouping legs is a tripartite state $|ABC\rangle$ on the top, lower left, and lower right d.o.f., and "splits" it into a four-partite state $|AB_LB_RC\rangle$. More precisely, we look for the splitting isometry $a^\dagger : B \to B_L \otimes B_R$, where $a^\dagger a = \mathbb{1}$ and $B_{L/R}$ have dimension χ, such that $|AB_LB_RC\rangle = a^\dagger|ABC\rangle$ has minimal entanglement $S_{AB_L:B_RC}$. This is closely related to finding the entanglement of purification [34] of ρ_{AC}. To do so we make an initial (suboptimal) guess for the isometry a_0 chosen so that B_LB_R includes the χ^2 highest weight states in B, and then parametrize the optimal choice as $a = a_0U^\dagger$ for a unitary U acting on B_LB_R. We choose U to minimize $S_{AB_L:B_RC}$ or its Renyi generalization [35], a well-defined optimization problem [36,37]. The resulting a comprise the isometries in A^ℓ, allowing us to successively unzip Λ^ℓ into $A^\ell\Lambda$. We do not have a rigorous proof regarding the success of the MM but instead we will consider two practical numerical tests in the following.

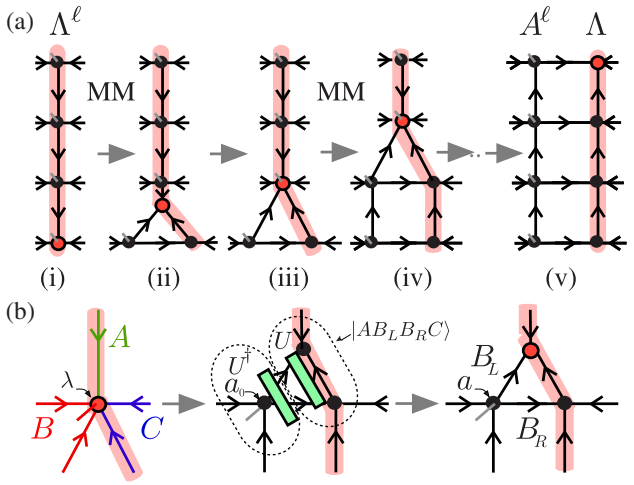

FIG. 2. The Moses move. (a) The orthogonality hypersurface Λ^ℓ is split into the product of a left isometry A^ℓ and a zero-column state Λ with no physical indices. The unzipping is performed by successively applying the splitting procedure shown in panel (b). The legs of the center site λ are grouped into a tripartite state $|ABC\rangle$ which is "split" into three tensors in two steps: first find $|ABC\rangle \approx a_0U^\dagger|AB_LB_RC\rangle$ for an initial guess of the isometry a_0 and unitary U which minimizes the entanglement across the vertical bond highlighted in red; second set $a = a_0U^\dagger$ and split $|AB_LB_RC\rangle$ in two via SVD. The resulting a comprise the tensors in A^ℓ, and the choice of U will produce a Λ with minimal vertical entanglement.

MPS to isoTNS.—Given a ground state wave function $|\Psi\rangle$ on an $L_x \times L_y$ strip, we propose an iterative algorithm to put $|\Psi\rangle$ into an isoTNS which we test for the transverse Ising model $H = -\sum_{\langle i,j\rangle}\sigma_i^z\sigma_j^z - g\sum_i\sigma^x$ with Pauli matrices σ^μ. To implement it numerically, we consider a strip with $L_y \gg L_x$ and use DMRG to obtain the ground state as a 1D MPS $\Lambda^{1:L_x}$ where each "site" contains the L_x spins of the corresponding row [Fig. 3(a)]. As described in Fig. 3, the MM can then be used to iteratively peal off columns of the wave function $\Lambda^{\ell:L_x} = A^\ell\Lambda^{\ell+1:L_x}$, producing an isoTNS. The algorithm is exponentially difficult in L_x (since Ψ is obtained as an MPS), but serves as a check on the ansatz independent of a ground state search scheme. Using an ancilla dimension $\chi = 6$ for the isometries, the error $\||\Psi_{\mathrm{MPS}}\rangle - |\Psi_{\mathrm{isoTNS}}\rangle\|^2$ is 2×10^{-6} per site at $g = 3.5$ (paramagnetic phase), $L_y = 20$, $L_x = 6$, $\chi_{\mathrm{MPS}} = 128$, obtained in about ten minutes on a laptop.

More interesting is the behavior of the "vertical" (top and bottom) and "horizontal" (left and right) entanglement of the resulting isoTNS. At each step ℓ the orthogonality hypersurface $\Lambda^{\ell+1:L_x}$ makes a "$\urcorner$" shape, running up the right and over the top. In Fig. 3(b) we show the entanglement entropy $S_\ell(y)$ for cuts along Λ, and find S_ℓ decrease with ℓ. If the underlying phase has area law $S_R = s|\partial R| + \cdots$, for $y \sim L_y/2$ we hope S_ℓ goes as $S_\ell \approx s(L_x - \ell) + \cdots$. If not, the isometric columns A^ℓ are not removing their share of the entanglement and the algorithm will fail in the thermodynamic limit. In Fig. 3(c), we see that after the initial delay the algorithm begins to remove remarkably close to s entanglement per iteration. The initial delay is expected, because any two vertically entangled d.o.f. will individually have some horizontal extent. Until their entire support is to the left of Λ, the isometries A cannot remove them. The residual horizontal entanglement is left behind in the top region of Λ. As hoped for, the horizontal entanglement is of the order of s, and for $\ell = L_x$ we find S_ℓ smoothly matches up between the right and

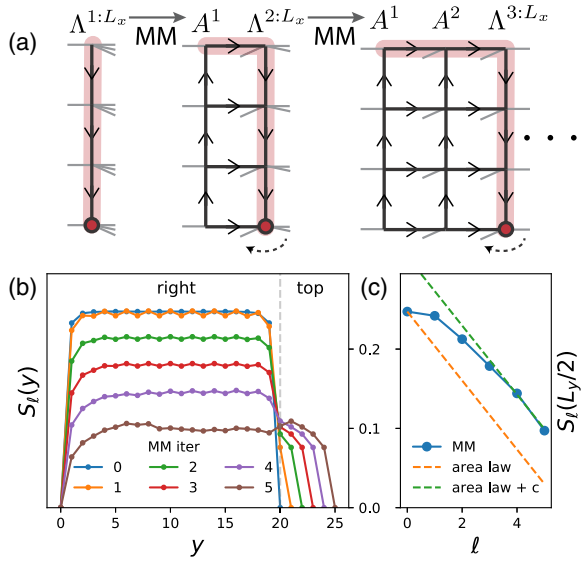

FIG. 3. The MPS to isoTNS algorithm: (a) The MPS $\Lambda^{1:L_x}$ for an $L_x \times L_y$ strip is fed into the MM by treating the legs of the first column as the left ancilla and the remaining columns as the right ancilla to obtain $\Lambda^{1:L_x} = A^1\Lambda^{2:L_x}$. The renormalized wave function $\Lambda^{2:L_x}$ is then reshaped by viewing the legs of the second column as physical, and its vertical arrows are reversed downwards using the standard 1D MPS canonicalization algorithm. Applying the MM again, we can repeat to obtain a canonical TNS. (b) Entanglement entropy S_ℓ for the sequence of orthogonality hypersurfaces (highlighted in red) after ℓ iterations. y runs from bottom right, to top right, to top left. (c) S_ℓ for a cut at $y \sim L_y/2$, compared against the bulk area law determined from DMRG.

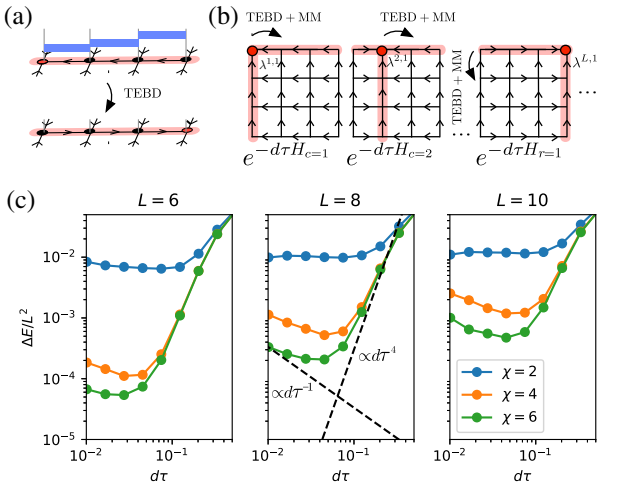

FIG. 4. The TEBD^2 algorithm: (a) Trotterization of $e^{-\tau H_{r/c}}$ into a product of two-site terms acting on a single row or column of the isoTNS. (b) To complete one time step, the 1D update is applied to all rows or columns by first applying the TEBD sweep and then sequentially shifting the orthogonality center $\lambda^{c,r}$ using the MM. Note that the update of one row or column reverses the arrows twice, e.g., the TEBD sweep moves $\lambda^{c,r}$ from the top to the bottom and then MM moves it up again. (c) Error densities of the energy of the transverse field Ising model with $g = 3.5$ for different system sizes and maximal bond dimensions χ as function of the Trotter step size $d\tau$.

top regions, despite the seemingly anisotropic nature of the algorithm.

TEBD^2 *algorithm.*—We now propose a Trotterized time stepper for isoTNS which can be used to obtain the ground state by imaginary-time evolution. Assuming a nearest-neighbor interaction, we split the Hamiltonian into terms acting on columns and rows, $H = \sum_{c=1}^{L_x} H_c + \sum_{r=1}^{L_y} H_r$. We Trotterize according to $e^{-\tau H} \approx \prod_r e^{-\tau H_r} \prod_c e^{-\tau H_c}$ as illustrated in Fig. 4(a). As for the TEBD update in 1D, the TEBD^2 can be easily improved to second order. We start in canonical form with the orthogonality center $\lambda^{1,1}$ at site c, $r = 1$, 1. The evolution $e^{-\tau H_{c=1}}$ is then applied to column Λ^1 by calling the standard 1D TEBD algorithm [23] at a cost $\propto \chi^6$. We then use the MM to bring the orthogonality center over by one column, to $\lambda^{2,1}$ at a cost $\propto \chi^7$, and apply $H_{c=2}$, and so on, bringing the orthogonality center to $\lambda^{L_x,1}$ (in contrast, the full update of an unconstrained PEPS costs χ^{12} [8]). Applying $e^{-\tau H_r}$ analogously brings the center to λ^{L_x,L_y}, and we repeat to bring λ counterclockwise around the four corners to complete the time step. Within a sweep the algorithm is literally two nested versions of 1D TEBD, hence the name TEBD^2.

To benchmark TEBD^2, we return to the transverse field Ising model. Figure 4(c) shows the energy density obtained from TEBD^2 relative to numerically exact results from large scale 1D-DMRG simulations at $g = 3.5$. If the evolution were exact the energy would decrease monotonically as the Trotter step $d\tau$ is decreased. However, the MM has a small truncation error ϵ_{MM}, and we see that the resulting energy has a minimum in $d\tau$. For a pth order Trotter step, the energy error should be $\Delta E = a\epsilon_{\mathrm{MM}}/d\tau + bd\tau^{2p}$ (in our implementation $p = 2$) [38], in agreement with the observed minima. A similar effect is also observed in the full update of TNS, and can be partially remedied by using a variational update instead of imaginary-time evolution [16,17]. The minimum energy converges towards the exact result as the bond dimension χ is increased.

Conclusions.—We introduced an isometric TNS ansatz which results in a canonical form that allows for 1D MPS algorithms to be efficiently adapted to 2D. To numerically benchmark the ansatz, we first demonstrated that an MPS representation of the ground state of the 2D transverse field Ising model can be efficiently transformed into an isoTNS. Second, we implemented a TEBD^2 algorithm and showed that it efficiently finds an approximation of the ground state of the 2D TFI model within the isoTNS form.

We thank B. Bauer, M. Fishman, J. Haah, S. Lin, R. Mong, M. Stoudenmire, A. Turner, XL. Qi, and

F. Verstraete for enlightening discussions. We are indebted to R. Mong for pointing out the biblical origin of the Moses move algorithm. F. P. acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Research Unit FOR 1807 through Grants No. PO 1370/2-1, TRR80, Germany's Excellence Strategy—EXC-2111–390814868, and the European Research Council (ERC) under the European Union as Horizon 2020 research and innovation program (Grant No. 771537). M. Z. was funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05-CH11231 through the Scientific Discovery through Advanced Computing (SciDAC) program (KC23DAC Topological and Correlated Matter via Tensor Networks and Quantum Monte Carlo). This work was finished at the Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1607611.

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[2] F. Verstraete and J. I. Cirac, Phys. Rev. Lett. **104**, 190405 (2010).
[3] S. Ostlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
[4] T. N. J. Dukelsky, M. A. Martn-Delgado, and G. Sierra, Europhys. Lett. **43**, 457 (1998).
[5] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).
[6] M. B. Hastings, J. Stat. Mech. (2007) P08024.
[7] N. Maeshima, Y. Hieida, Y. Akutsu, T. Nishino, and K. Okunishi, Phys. Rev. E **64**, 016705 (2001).
[8] F. Verstraete and J. Cirac, arXiv:cond-mat/0407066.
[9] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **65**, 891 (1996).
[10] T. Nishino and K. Okunishi, J. Phys. Soc. Jpn. **66**, 3040 (1997).
[11] M. Levin and C. P. Nave, Phys. Rev. Lett. **99**, 120601 (2007).
[12] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).
[13] H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **101**, 090603 (2008).
[14] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **103**, 160601 (2009).
[15] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang, and T. Xiang, Phys. Rev. B **86**, 045139 (2012).
[16] P. Corboz, Phys. Rev. B **94**, 035133 (2016).
[17] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Phys. Rev. B **94**, 155123 (2016).
[18] M. C. Bañuls, D. Pérez-García, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. A **77**, 052306 (2008).
[19] M. T. Fishman, L. Vanderstraeten, V. Zauner-Stauber, J. Haegeman, and F. Verstraete, Phys. Rev. B **98**, 235148 (2018).
[20] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Science **358**, 1155 (2017).
[21] P. Corboz and F. Mila, Phys. Rev. B **87**, 115144 (2013).
[22] H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang, Phys. Rev. Lett. **118**, 137202 (2017).
[23] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).
[24] D. Perez-Garcia, M. M. Wolf, M. Sanz, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 167202 (2008).
[25] G. Evenbly, M. Fishman, N.Schuch, F. Verstraete, and G. Vidal (private communication).
[26] S. Richter, Construction of states on two-dimensional lattices and quantum cellular automata, Ph.D. Thesis, Osnabrück, 1994.
[27] F. Verstraete, V. Murg, and J. I. Cirac, Adv. Phys. **57**, 143 (2008).
[28] U. Schollwöck, Ann. Phys. (Amsterdam) **326**, 96 (2011).
[29] E. M. Stoudenmire and S. R. White, Phys. Rev. B **87**, 155137 (2013).
[30] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Phys. Rev. B **94**, 165116 (2016).
[31] M. Bal, M. M. Rams, V. Zauner, J. Haegeman, and F. Verstraete, Phys. Rev. B **94**, 205122 (2016).
[32] T. Soejima, K. Siva, N. Bultinck, S. Chatterjee, F. Pollmann, and M. P. Zaletel, arXiv:1908.07545 [Phys. Rev. B (to be published)].
[33] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.037201 for details about the variational solution of $\Lambda^{\ell} = A^{\ell}\Lambda$ and benchmark of the Moses move.
[34] B. M. Terhal, M. Horodecki, D. W. Leung, and D. P. DiVincenzo, J. Math. Phys. (N.Y.) **43**, 4286 (2002).
[35] Results closest to the variational optimum are obtain if we use Renyi index $\alpha < 1$; see Supplemental Material [33].
[36] G. Evenbly and G. Vidal, J. Stat. Phys. **157**, 931 (2014).
[37] J. Hauschild, E. Leviatan, J. H. Bardarson, E. Altman, M. P. Zaletel, and F. Pollmann, Phys. Rev. B **98**, 235163 (2018).
[38] Letting Δ be a typical energy scale and ΔE the energy density, one step of imaginary-time evolution decreases the energy density by $\Delta E \Delta d\tau$, while the Moses move truncation increases it by $\epsilon_{\mathrm{MM}}\Delta$. Thus at long times the MM introduces $\Delta E \propto \epsilon_{\mathrm{MM}}/d\tau$.