Class of Quantum Many-Body States That Can Be Efficiently Simulated

G. Vidal

School of Physical Sciences, The University of Queensland, QLD 4072, Australia

(Received 3 December 2006; published 12 September 2008)

We introduce the multiscale entanglement renormalization ansatz, a class of quantum many-body states on a D-dimensional lattice that can be efficiently simulated with a classical computer, in that the expectation value of local observables can be computed exactly and efficiently. The multiscale entanglement renormalization ansatz is equivalent to a quantum circuit of logarithmic depth that has a very characteristic causal structure. It is also the ansatz underlying entanglement renormalization, a novel coarse-graining scheme for many-body quantum systems on a lattice.

DOI: 10.1103/PhysRevLett.101.110501 PACS numbers: 03.67.Lx, 03.65.Ud, 03.67.Hk

A better understanding of quantum entanglement has enabled significant progress in the numerical simulation of quantum many-body systems over the past few years [1–4]. Building on the density matrix renormalization group [5]—a well-established technique for systems on a 1D lattice—new insight from quantum information science has led, e.g., to efficient algorithms to simulate time evolution [1] and address 2D systems [2].

A key ingredient of such algorithms is the use of a network of tensors to efficiently represent quantum many-body states. Examples of tensor networks are matrix product states (MPSs) [6] for 1D systems, tree tensor networks (TTNs) [3] for systems with a tree shape, and projected entangled-pair states (PEPSs) [2] for 2D systems and beyond. The three structures differ in the graph that defines how the tensors are interconnected into a network: The graphs for MPSs, TTNs, and 2D PEPSs are, respectively, a chain, a tree, and a 2D lattice. Importantly, from these tensor networks the expectation value of local observables can be computed efficiently. But whereas from a MPS and a TTN such calculations are exact, from a PEPS local expectation values can be obtained efficiently only after a number of approximations.

In this Letter, we present the multiscale entanglement renormalization ansatz (MERA), a structure that efficiently encodes quantum many-body states of D-dimensional lattice systems and from which local expectation values can be computed exactly. A MERA consists of a network of *isometric* tensors in $D+1$ dimensions, where the extra dimension can be interpreted in two alternative ways: either as the time of a peculiar class of quantum computations or as parametrizing different length scales in the system, according to successive applications of a lattice coarse-graining procedure known as entanglement renormalization [7]. The MERA is a promising candidate to describe emergent quantum phenomena, including quantum phase transitions, quasiparticle excitations, and topological order. Here we establish its connection to entanglement renormalization and explore some of its basic properties: the efficient contractibility of the network, leading to an efficient evaluation of local expectation values; the inherent support of power-law correlations; the versatility to adapt to both the local and the global structure of the system's underlying lattice; and the ability to assimilate symmetries such as invariance under translations or rescaling—which result in substantial additional gains in computational efficiency.

Let us consider a lattice $\mathcal{L}$ in D spatial dimensions consisting of N sites, where each site $s \in \mathcal{L}$ is described by a complex vector space $\mathbb{V}$ of finite dimension m. We study states $|\Psi\rangle \in \mathbb{V}^{\otimes N}$ of the lattice $\mathcal{L}$ that are generated by a special type of quantum circuit $\mathcal{C}$ (Fig. 1). Each site in $\mathcal{L}$ is identified with one outgoing wire of $\mathcal{C}$, labeled with the same index s. In order to further characterize the class of circuits $\mathcal{C}$, we need to introduce the notion of a causal cone.

Definition.—The *causal cone* $\mathcal{C}^{[s]}$ of an outgoing wire s of circuit $\mathcal{C}$ (correspondingly, a site s of $\mathcal{L}$) is the subcircuit causally connected with s, that is, the subset of wires and gates that influence the state of s. The *width* of the causal cone $\mathcal{C}^{[s]}$ is the maximum number of wires involved in a time slice $\mathcal{C}^{[s]}_{\theta}$, where θ is a discrete time label and wires

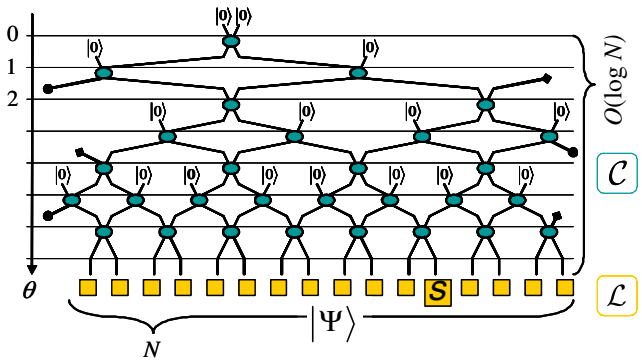

FIG. 1 (color online). Example of quantum circuit $\mathcal{C}$ that transforms the product state $|0\rangle^{\otimes N}$ into the state $|\Psi\rangle$ of lattice $\mathcal{L}$, case $D=1$. This circuit contains $2N-1$ gates organized in $O(\log N)$ layers that are labeled by a discrete time θ. Notice the periodic boundary conditions.

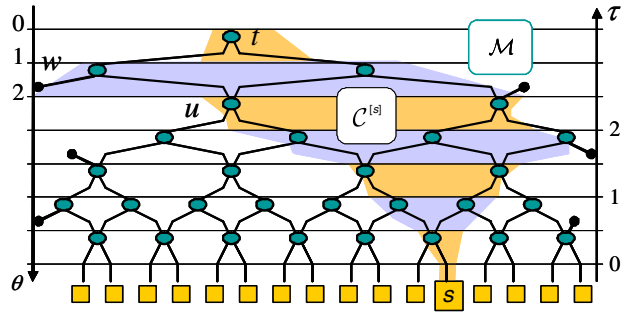

FIG. 2 (color online). The MERA $\mathcal{M}$ is made of three types of tensors: t, w, and u; see Eqs. (1) and (2). Its causal structure is inherited from that of the circuit $\mathcal{C}$: The causal cone $\mathcal{C}^{[s]}$ for site s has bounded width. When reversing the arrow of time θ, $\mathcal{M}$ corresponds to $O(\log N)$ entanglement renormalization transformations labeled by τ.

that are in a product state $|0\rangle$ are not counted in the width of $\mathcal{C}^{[s]}$.

The defining feature of circuit $\mathcal{C}$ is that the width of the causal cone of any of its N outgoing wires is bounded by a constant that is independent of N. For instance, in the example of Fig. 1, corresponding to a lattice in $D = 1$ dimensions, the causal cone of any outgoing wire involves at most four wires at any time θ, as highlighted in Fig. 2. Circuits with this property can be built for any lattice $\mathcal{L}$ in any spatial dimension D. Also, given a lattice $\mathcal{L}$, there are still many circuits with bounded causal cone width. For a cubic lattice in D dimensions, Fig. 3 shows a possible construction where causal cones have at most $3^{D-1}4$ wires at any given time θ. The above constraint in the causal cone is easily seen to imply that the depth of $\mathcal{C}$ is at most $O(N)$. If, in addition, we assume that the circuit is organized in a pattern that resembles a tree as in Fig. 1, then its depth is $O(\log N)$.

The MERA is a network of tensors $\mathcal{M}$ made from any such circuit $\mathcal{C}$ with a few cosmetic changes. For concreteness, we consider the circuit of Fig. 1. Each unitary gate u of $\mathcal{C}$ gives rise to a tensor of the MERA $\mathcal{M}$. But incoming wires in the fixed state $|0\rangle$ are absorbed, producing three

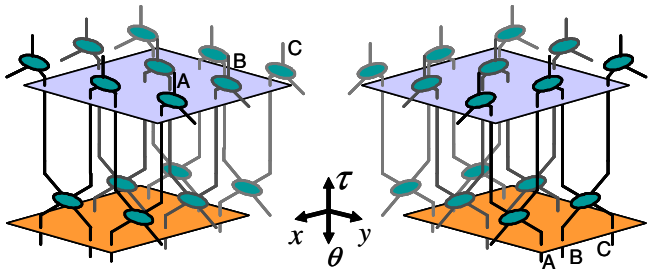

FIG. 3 (color online). Detail of a causal cone of a simple MERA construction for a square lattice, $D = 2$. 3×3 wires (bottom left) are mapped into 3×3 wires (top right) by layers of tensors that act along the y and x directions [some stages involve 3×4 wires]. The generalization to $D > 2$ is straightforward.

kinds of tensors: (i) The top tensor $t = u|0\rangle|0\rangle$ of $\mathcal{M}$ has two indices and is normalized to 1:

$$(t)_{\mu\nu} \equiv (u)^{\alpha\beta}_{\mu\nu}|_{\alpha,\beta=0}, \qquad \sum_{\mu\nu}(t^*)_{\mu\nu}(t)_{\mu\nu} = 1; \quad (1)$$

(ii) tensors in every second row are *isometries* $w = u|0\rangle$:

$$(w)^{\alpha}_{\mu\nu} \equiv (u)^{\alpha\beta}_{\mu\nu}|_{\beta=0}, \qquad \sum_{\mu\nu}(w^*)^{\alpha}_{\mu\nu}(w)^{\alpha'}_{\mu\nu} = \delta_{\alpha\alpha'}; \quad (2)$$

(iii) the rest of the tensors in $\mathcal{M}$ are unitary gates u:

$$\sum_{\mu\nu}(u^*)^{\alpha\beta}_{\mu\nu}(u)^{\alpha'\beta'}_{\mu\nu} = \sum_{\mu\nu}(u^*)^{\mu\nu}_{\alpha\beta}(u)^{\mu\nu}_{\alpha'\beta'} = \delta_{\alpha\alpha'}\delta_{\beta\beta'}, \quad (3)$$

that we call *disentanglers*. Notice that in this example storing $\mathcal{M}$ requires computational space that grows as $O(m^4 N)$, that is, linearly in N, given that there are $2N - 3$ tensors and each tensor depends on at most m^4 parameters.

Similarly, in the case of a generic lattice $\mathcal{L}$ in D spatial dimensions, the MERA is an efficient representation for certain pure states $|\Psi\rangle$ of $\mathcal{L}$ that consists of a tensor network $\mathcal{M}$ in $D + 1$ dimensions with two properties: (i) Tensors are constrained to be unitary or isometric as in Eqs. (1)–(3); (ii) each open wire s, associated to one site of the underlying lattice $\mathcal{L}$, has a causal cone $\mathcal{C}^{[s]}$ with bounded width. As a consequence of its peculiar causal structure, the reduced density matrix of a small number of lattice sites can be computed exactly with remarkably small cost. In what follows, p_1 and p_2 are finite integers that depend on the choice of $\mathcal{C}$, and we assume that the underlying circuit has depth $O(\log N)$.

Lemma 1.—The density matrix $\rho^{[s]} \equiv \mathrm{tr}_{\bar{s}}(|\Psi\rangle\langle\Psi|)$ for one site $s \in \mathcal{L}$ can be obtained from the MERA with computational effort that scales as $O(m^{p_1} \log N)$.

Proof.—Let σ_θ denote the reduced density matrix for the time slice $\mathcal{C}^{[s]}_\theta$ of the causal cone $\mathcal{C}^{[s]}$. As illustrated in Figs. 4 and 5 for the particular MERA of Fig. 2, $\sigma_{\theta+1}$ can be computed from σ_θ, with a cost polynomial in m and independent of N. Iterating, we obtain a sequence $\{\sigma_1, \sigma_2, \ldots, \sigma_\Theta\}$ of $\Theta = O(\log N)$ density matrices, where $\rho^{[s]} = \sigma_\Theta$. This completes the proof.

Lemma 2.—The two-site density matrix $\rho^{[s_1 s_2]}$ can be computed with time $O(m^{p_2} \log N)$.

Proof.—Again, the causal cone $\mathcal{C}^{[s_1 s_2]}$ has logarithmic depth and a width independent of N (see Fig. 7), and the reduced density matrix $\sigma_{\theta+1}$ for time slice $\mathcal{C}^{[s_1 s_2]}_{\theta+1}$ can be obtained from σ_θ for $\mathcal{C}^{[s_1 s_2]}_\theta$ with a cost $O(m^{p_2})$.

More generally, the reduced density matrix $\rho^{[s_1 \cdots s_q]}$ for q sites can be computed with cost $O(m^p \log N)$, where p grows exponentially in q. The above lemmas readily imply that the expectation value of local observables, such as two-site correlators

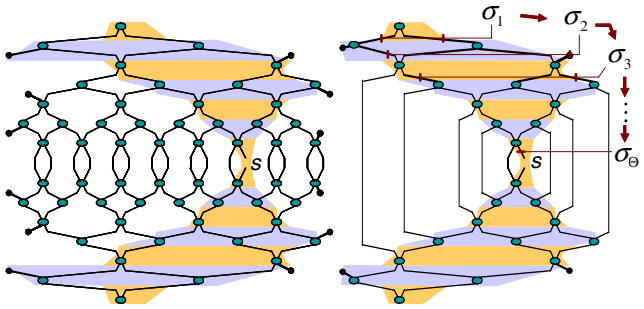

FIG. 4 (color online). Left: Tensor network for $\rho^{[s]} = \mathrm{tr}_{\mathcal{L}\setminus s}|\Psi\rangle\langle\Psi|$ in Lemma 1. Right: The gates outside the causal cone of s can be removed by using the unitarity constraints of Eqs. (2) and (3). The resulting tensor network can now be contracted by sequentially computing $O(\log N)$ density matrices $\{\sigma_1, \sigma_2, \ldots\}$.

$$C_2(s_1, s_2) \equiv \langle\Psi|a^{[s_1]}b^{[s_2]}|\Psi\rangle = \mathrm{tr}(\rho^{[s_1 s_2]}a^{[s_1]}b^{[s_2]}), \quad (4)$$

can be computed efficiently and exactly from the MERA.

We now show how the MERA relates to a real-space renormalization group transformation. We reverse the arrow of time in the quantum circuit $\mathcal{C}$, and, starting from $|\Psi_0\rangle \equiv |\Psi\rangle$, we consider the sequence of states $\{|\Psi_0\rangle, |\Psi_1\rangle, |\Psi_2\rangle, \ldots\}$, where $|\Psi_\tau\rangle$ is the [nonfactorizable part of the] state of $\mathcal{C}$ at time $\theta = \Theta - 2\tau$ (Fig. 6). Notice that $|\Psi_{\tau+1}\rangle$ is obtained from $|\Psi_\tau\rangle$ by applying one layer of disentanglers u and one layer of isometries w. In other words, $\mathcal{M}$ implements entanglement renormalization transformations [7]. More generally, we can use the MERA to transform the sites of lattice $\mathcal{L}_0 \equiv \mathcal{L}$, as well as linear operators defined on $\mathcal{L}$ such as a Hamiltonian H_0,

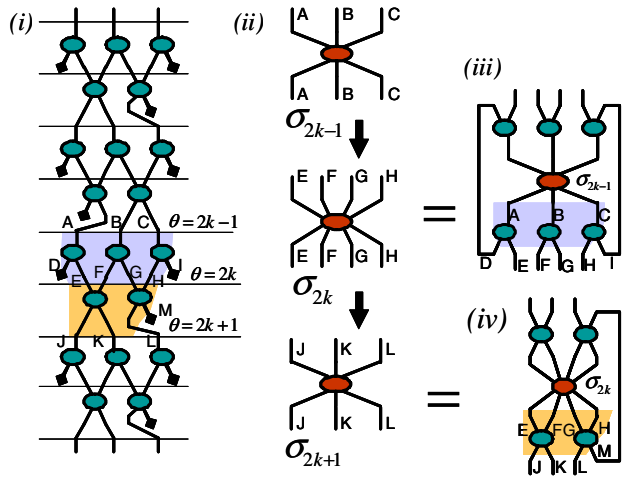

FIG. 5 (color online). Efficient computation of $\rho^{[s]}$, Lemma 1. (i) Part of the causal cone $\mathcal{C}^{[s]}$ of $\rho^{[s]}$. (ii) Sequence of density matrices $\sigma_{2k-1} \to \sigma_{2k} \to \sigma_{2k+1}$ represented as tensors with 6 or 8 indices. (iii) σ_{2k} expressed in terms of σ_{2k-1} and isometries w. (iv) σ_{2k+1} expressed in terms of σ_{2k} and disentanglers u. By contracting all of the indices shared by two tensors in (iii) and (iv), σ_{2k} and σ_{2k+1} are obtained with cost $O(m^{p_1})$.

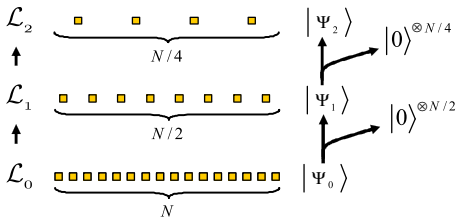

FIG. 6 (color online). Renormalization group transformation associated to the MERA. The tensors in $\mathcal{M}$ define a sequence $\{\mathcal{L}_0, \mathcal{L}_1, \ldots\}$ of increasingly coarser lattices in states $\{|\Psi_0\rangle, |\Psi_1\rangle, \ldots\}$.

so as to obtain a sequence of increasingly coarse-grained lattices $\{\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \ldots\}$ and corresponding effective Hamiltonians $\{H_0, H_1, H_2, \ldots\}$. An operator defined on site $s \in \mathcal{L}$ is mapped into an operator contained in the causal cone $\mathcal{C}^{[s]}$, so that *local* operators in $\mathcal{L}_0$ remain local in $\mathcal{L}_\tau$.

Numerical evidence for 1D lattices [7] shows that a MERA encodes, in a markedly more efficient way than a MPS [8], accurate approximations to the ground state near to or at quantum criticality. At a critical point, correlators $C_2(s_1, s_2)$ decay as a power law with the distance r between sites s_1 and s_2. Next we argue that the MERA supports algebraic decay of correlations. This is the case in any dimension D.

It follows from the geometry of the λ-shaped causal cone $\mathcal{C}^{[s_1 s_2]}$ that we can compute $\rho^{[s_1 s_2]}$ from the density matrix $\rho_{\bar\tau}$ corresponding to a block of contiguous sites of $\mathcal{L}_{\bar\tau}$, where $\bar\tau \approx \log_2 r$, by means of a sequence of density matrices $\{\rho_{\bar\tau}, \ldots, \rho_2, \rho_1, \rho^{[s_1 s_2]}\}$; see Fig. 7. That is, $\rho^{[s_1 s_2]}$ is obtained from $\rho_{\bar\tau}$ after $O(\log r)$ transformations. If, as it can be argued in a (scale-invariant) critical ground state, each of these transformations reduces correlations by a constant factor $z < 1$, we readily obtain a power-law decay for $C_2(s_1, s_2)$:

$$C_2(s_1, s_2) \approx z^{\log r} = r^{-q}, \qquad q \equiv \log\frac{1}{z} > 0. \quad (5)$$

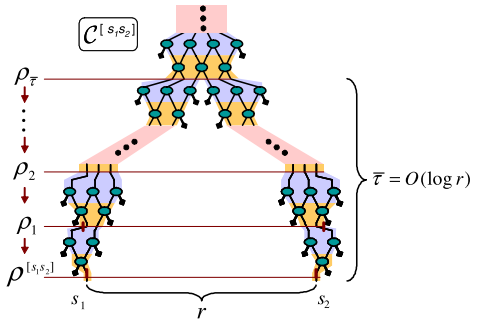

FIG. 7 (color online). Causal cone $\mathcal{C}^{[s_1 s_2]}$ for $D = 1$. A small square at the end of an open wire indicates a trace over the corresponding degrees of freedom. The states $\rho^{[s_1 s_2]}$ are obtained from $\rho_{\bar\tau}$ through a sequence of $O(\log r)$ transformations. Similar constructions hold for $D > 1$.

This property, valid in any spatial dimension D, is enabled by the fact that two sites separated a distance r in $\mathcal{L}$ are connected through a path of length $O(\log r)$ in $\mathcal{M}$, and it indicates that a MERA is particularly well suited to describe states with quasi-long-range order, such as critical ground states. As demonstrated in Refs. [7,9], the MERA can also reproduce the $\log L$ scaling of entanglement for a block of L contiguous sites that characterizes critical ground states in $D = 1$ dimensions [10] and the boundary law $L^{(D-1)}$ for a hypercube of L^D sites in $D > 1$ dimensions.

For the sake of concreteness, we have focused on a specific MERA for a $D = 1$ lattice and indicated a possible extension to a square (cubic) lattice for $D > 1$. A generic lattice can be addressed by properly modifying circuit $\mathcal{C}$ while keeping its characteristic causal structure. For instance, in $D = 2$ dimensions, a specific MERA can be engineered to represent states of a triangular lattice, or of a lattice with random vacancies or linear or bulk defects, or to account for a variety of boundary conditions (e.g., plane, cylinder, sphere, or torus). In addition, the number of levels m can vary throughout $\mathcal{M}$. When used as the basis for numerical simulations, adjusting the MERA to the specifics of a problem leads to computational gains.

In particular, the symmetries of $|\Psi\rangle$ can be assimilated into the MERA. An internal symmetry, such as $SU(2)$ invariance, results in a series of constraints for the tensors in $\mathcal{M}$ [11], which then depend on fewer parameters. More substantial gains are obtained when $|\Psi\rangle$ is invariant under translations (that is, cyclic shifts by one lattice site in a system with periodic boundary conditions), since all of the tensors in a layer of $\mathcal{M}$ can be chosen to be the same and the MERA depends on $O(m^4 \log_2 N)$ parameters instead of $O(m^4 N)$. The most dramatic simplification occurs for states that are invariant under entanglement renormalization transformations, even in an infinite lattice [7]. Now all disentanglers and isometries in $\mathcal{M}$ are identical and the MERA depends on just $O(m^4)$ parameters. Some 1D critical ground states are numerically seen to belong to this class [7].

We conclude with a few pointers to future work. On the one hand, most techniques to simulate quantum systems with a MPS can be generalized to a MERA. This includes algorithms to compute the ground and thermal states and to simulate time evolution. Importantly, in order to update $\mathcal{M}$ after a unitary gate $v^{[s_1 s_2]}$ acts on sites s_1 and s_2, only the tensors in the causal cone $\mathcal{C}^{[s_1 s_2]}$ need to be modified, with a cost logarithmic in N.

On the other hand, the potential of the MERA goes beyond representing individual states. Notice that, by feeding the incoming wires, labeled by r, of quantum circuit $\mathcal{C}$ with an arbitrary product state $\otimes_{r=1}^{N} |\phi_r^{[r]}\rangle$ instead of $\otimes_{r=1}^{N} |0^{[r]}\rangle$, we can generate a different entangled state $|\Psi_{\{\phi_r\}}\rangle$ for the lattice, this one represented by a MERA that differs from the original one only in the isometric tensors w and such that, given the unitarity of $\mathcal{C}$, fulfills

$$\langle \Psi_{\{\phi_r\}} | \Psi_{\{\phi'_r\}} \rangle = \prod_r \langle \phi_r^{[r]} | \phi_r'^{[r]} \rangle. \tag{6}$$

This can be used to encode, in just one (generalized) MERA, not only the ground state of a Hamiltonian H but also its quasiparticle excitations. Moreover, in systems with topological order, this property allows one to represent sets of topologically inequivalent ground states in a single MERA, by storing the topological information on the top tensor t [12].

The author acknowledges support from the Australian Research Council, No. FF0668731 and No. DP0878830.

[1] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003); **93**, 040502 (2004); S. R. White and A. E. Feiguin, *ibid.* **93**, 076401 (2004); A. J. Daley *et al.*, J. Stat. Mech. (2004) P04005.

[2] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066; F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).

[3] Y. Y. Shi, L. M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).

[4] See, for instance, F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004); M. Zwolak and G. Vidal, *ibid.* **93**, 207205 (2004); B. Paredes, F. Verstraete, and J. I. Cirac, *ibid.* **95**, 140501 (2005); F. Verstraete, D. Porras, and J. I. Cirac, *ibid.* **93**, 227205 (2004); D. Porras, F. Verstraete, and J. I. Cirac, Phys. Rev. B **73**, 014410 (2006); T. Osborne, Phys. Rev. A **75**, 032321 (2007); **75**, 042306 (2007); arXiv:cond-mat/0605194.

[5] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); Phys. Rev. B **48**, 10 345 (1993); U. Schollwoeck, Rev. Mod. Phys. **77**, 259 (2005).

[6] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992); S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).

[7] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).

[8] In a MPS, correlations between the left and right halves of a chain are accounted for through m_{MPS} terms in the Schmidt decomposition. A MERA can account for about $2m\log_2 N$ terms in the Schmidt decomposition. Thus, it takes a MPS with $m_{\mathrm{MPS}} \approx 2m\log_2 N$ to represent a state stored with a MERA of m-level wires.

[9] G. Evenbly and G. Vidal, arXiv:0710.0692v2; arXiv:0801.2449v1.

[10] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003); P. Calabrese and J. Cardy, J. Stat. Mech. 0406 (2004) P002; A. R. Its, B.-Q. Jin, and V. E. Korepin, J. Phys. A **38**, 2975 (2005).

[11] S. Singh *et al.* (to be published).

[12] M. Aguado and G. Vidal, Phys. Rev. Lett. **100**, 070404 (2008).