

## Multigrid Algorithms for Tensor Network States

Michele Dolfi,<sup>1</sup> Bela Bauer,<sup>2</sup> Matthias Troyer,<sup>1</sup> and Zoran Ristivojevic<sup>3</sup>

<sup>1</sup>*Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland*

<sup>2</sup>*Station Q, Microsoft Research, Santa Barbara, California 93106-6105, USA*

<sup>3</sup>*Laboratoire de Physique Théorique-CNRS, Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France*

(Received 1 April 2012; published 13 July 2012)

The widely used density matrix renormalization group (DMRG) method often fails to converge in systems with multiple length scales, such as lattice discretizations of continuum models and dilute or weakly doped lattice models. The local optimization employed by DMRG to optimize the wave function is ineffective in updating large-scale features. Here we present a multigrid algorithm that solves these convergence problems by optimizing the wave function at different spatial resolutions. We demonstrate its effectiveness by simulating bosons in continuous space and study nonadiabaticity when ramping up the amplitude of an optical lattice. The algorithm can be generalized to tensor network methods and combined with the contractor renormalization group method to study dilute and weakly doped lattice models.

DOI: [10.1103/PhysRevLett.109.020604](https://doi.org/10.1103/PhysRevLett.109.020604)

PACS numbers: 05.10.Cc, 37.10.Jk, 67.85.Hj, 71.27.+a

The optimization of variational wave functions is generally a very difficult problem. In the specific case of matrix-product states (MPS) [1], the density matrix renormalization group algorithm (DMRG) [2–5] often reliably and efficiently optimizes these wave functions to find a good approximation of the ground state. While most efficient in one dimension, it can be applied to medium-sized two-dimensional systems [6] and has been generalized to calculate time-dependent [7–9] and finite temperature properties [10–12].

In systems with multiple length scales, however, the DMRG algorithm often fails to converge, as the local optimizations that are at the core of DMRG are ineffective in optimizing large-scale features of the wave function. Especially in dilute systems where the interparticle distance is large compared to the lattice spacing, the convergence of the density profile can be very slow. Systems with multiple length scales suffering from this problem arise from lattice discretizations of continuum models [13] or in weakly doped lattice models where the hole density exhibits the same convergence problems. The first situation was recently discussed in Ref. [16].

Similar convergence problems are also known in other fields, e.g., when solving partial differential equations [17], lattice field theories [18], or electronic structures [19], and have there been overcome by multigrid approaches. Multigrid methods use a hierarchy of discretizations, as sketched in Fig. 1. Starting from the target problem on the finest grid (or a lattice model), the system is mapped to hierarchy of coarser grids. An approximate solution of the smallest problem on the coarsest grid is then used to initialize optimizations of the problem on the next finer grid, and this process is iterated down to the finest grid. This method can substantially speed up a calculation since the large scale features converge quickly on the coarsest grid, and the following calculations on finer grids only need to

optimize local features at the scale of the respective grid spacing.

In this Letter, we develop a multigrid DMRG (MG DMRG) algorithm to solve the above-mentioned convergence problems in DMRG calculations. As a first application and demonstration of the effectiveness of the algorithm, we present results for bosons in continuous space where MG DMRG enables the study of nonadiabaticities when slowly ramping up the amplitude of an optical lattice.

We start the description of the MG DMRG algorithm by reviewing MPS wave functions on a chain of  $L$  sites

$$|\psi\rangle = \sum_{\sigma} A_1^{\sigma_1} A_2^{\sigma_2} \cdots A_L^{\sigma_L} |\sigma\rangle \quad (1)$$

used in DMRG. They are characterized by a polynomial number  $\propto LM^2$  of variational parameters, the  $M \times M$  matrices  $A_i^{\sigma_i}$ . In one dimension, a good approximation

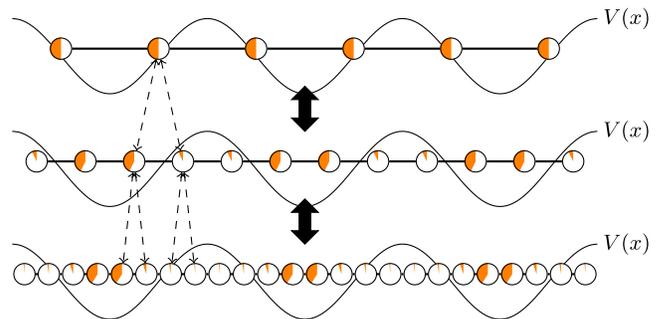


FIG. 1 (color online). Multigrid DMRG illustrated for bosons in an optical lattice  $V(x)$ . The DMRG algorithm converges fast for the rather dense system at the coarsest grid (shown on the top), and this solution is then used to iteratively initialize DMRG calculations on finer grids, which substantially speeds up convergence. The filling of the circles illustrates the probability for a particle to be at that site.

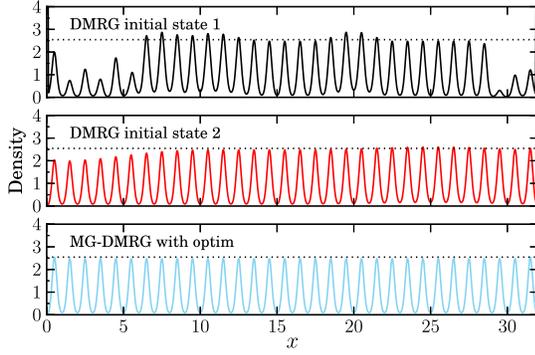


FIG. 2 (color online). Density profiles of a continuum system of bosons in an optical lattice consisting of  $L = 32$  unit cells. The top and middle panels show the nonconverged results obtained for  $N = 32$  grid points per unit cell after 12 sweeps of the DMRG algorithm with two different initial states: initial state 1 is a random state, initial state 2 is obtained from an infinite-size growing procedure as implemented in the ALPS DMRG code [24]. The bottom panel shows the multigrid result.

for low-energy states can be obtained by MPS wave functions with a fixed or, at most, polynomially growing  $M$  [20,21].

While optimizing MPS wave functions to obtain a variational estimate for the ground state is a hard nonlinear problem, the DMRG algorithm is very effective in many cases. It iteratively optimizes one or two of the matrices  $A_i^{\sigma_i}$  while keeping all other matrices fixed and sweeps back and forth along a (quasi-) one-dimensional system until convergence is achieved. For a recent review and implementation see Ref. [5]. It can, however, get trapped in local minima of this nonlinear optimization problem or become very slow, especially for the dilute systems considered here. As an example, see the badly converged density profiles obtained by standard DMRG approaches in Fig. 2.

In our implementation of the MG DMRG algorithm, we start by constructing the target lattice model and a hierarchy of models on coarser grids. Starting from the coarsest level we optimize the wave function and interpolate it to the next finer level, repeating this procedure until we reach the target system. Many generalizations are possible, for example, iterating the procedure by going back to coarser levels or starting from the finest instead of coarsest level.

The restriction operation maps a system to a coarser grid, merging  $n$  (typically  $n = 2$ ) sites into one. The model, given by the Hamiltonian  $H$  and defined in the local basis  $\{\sigma\}$ , is mapped to a restricted model  $\tilde{H}$  in a truncated local basis  $\{\tilde{\sigma}\}$  for the  $n$  merged sites. The truncation, denoted by an isometry  $T_{\tilde{\sigma}_1, \dots, \tilde{\sigma}_n}^{\sigma_1, \dots, \sigma_n}$ , is straightforward for continuum models, and an approach for lattice models will be discussed below. Any error due to the truncation will be corrected when returning to finer scales, as long as we stay in the same phase. As illustrated in Fig. 3 the restriction transforms a matrix product state  $A_1, \dots, A_n$  into

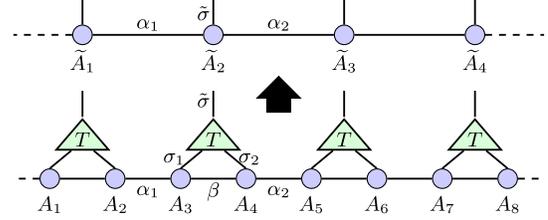


FIG. 3 (color online). The restriction of an MPS state reducing it from eight to four sites. The contraction along the indices  $\beta$ ,  $\sigma_1$ , and  $\sigma_2$  creates a new tensor  $\tilde{A}_{\alpha_1 \alpha_2}^{\tilde{\sigma}}$  with a new index  $\tilde{\sigma}$  but keeping the old connections to the neighboring sites ( $\alpha_1$  and  $\alpha_2$ ).

$$\tilde{A}_{\alpha_1 \alpha_2}^{\tilde{\sigma}} = A_{\alpha_1 \beta_1}^{\sigma_1} A_{\beta_1 \beta_2}^{\sigma_2} \cdots A_{\beta_{n-1} \alpha_2}^{\sigma_n} T_{\tilde{\sigma}_1, \dots, \tilde{\sigma}_n}^{\sigma_1, \dots, \sigma_n}. \quad (2)$$

Prolongation is the inverse of restriction and maps a solution from a coarse grid to a finer one. The isometry is inverted, and  $T^{-1}$  replaces one index by  $k$  new indices. From this tensor, we can recover a (nonunique) MPS representation by repeatedly applying singular value decompositions and splitting it into matrices  $A_1, \dots, A_n$  (see Fig. 4). It has turned out to be useful to perform a standard DMRG update on the newly obtained matrices immediately after prolongation while keeping the rest of the system on the coarse-grained lattice [22].

As a first application, we apply MG DMRG to bosons in a one-dimensional continuum system with an external optical lattice potential,  $V(x) = V_0 \cos^2(kx)$ , with  $k = \pi/a$  and  $a$  the size of a unit cell. The continuous-space Hamiltonian describing spinless bosons interacting through a  $\delta$  potential in a system of  $L$  unit cells and length  $La$  is

$$\hat{H} = \int_0^{La} dx \hat{\psi}^\dagger(x) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \hat{\psi}(x) + \frac{g}{2} \int_0^{La} dx \hat{\psi}^\dagger(x) \hat{\psi}^\dagger(x) \hat{\psi}(x) \hat{\psi}(x), \quad (3)$$

where a boson is created at position  $x$  with the field operator  $\hat{\psi}^\dagger(x)$ , satisfying the usual commutation relations. We express energies in units of the recoil energy  $E_r = \frac{\hbar^2 k^2}{2m}$ . The interaction  $g$  is conveniently parametrized by the dimensionless coupling  $\gamma = mg/\hbar^2 n$ , where  $n$  is the density.

In deep optical lattices, the low-energy sector of the model can be mapped to an effective single band

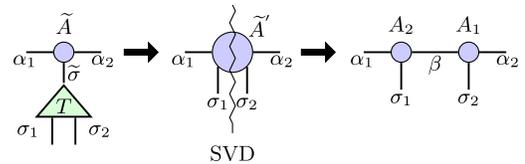


FIG. 4 (color online). Prolongation operation doubling the number of sites: First, we transform the basis  $\tilde{\sigma}$  of the initial matrix  $\tilde{A}$  into two new local bases  $\sigma_1$  and  $\sigma_2$ . With a singular value decomposition, we then split the rank-4 tensor  $\tilde{A}'$  into two matrices  $A_1$  and  $A_2$ .

Hubbard model with one site per unit cell. We are, however, interested also in weak optical lattices and thus, discretize the continuum model on a grid with  $N$  points per unit cell and spacing  $\Delta x = a/N$ . To discretize the Hamiltonian Eq. (3) on this lattice, we replace the Laplacian by a second order finite difference approximation and replace field operators by lattice annihilation and creation operators  $\hat{\psi}^\dagger(x = (i + 1/2)\Delta x) = \frac{1}{\sqrt{\Delta x}} \hat{c}_i^\dagger$ . We end up with a Hubbard-like model in a spatially varying potential,

$$\hat{H}(\Delta x) = -t(\Delta x) \sum_i [\hat{c}_i^\dagger \hat{c}_{i+1} + \text{H.c.}] + \sum_i \mu_i(\Delta x) \hat{n}_i + \frac{U(\Delta x)}{2} \sum_i \hat{n}_i(\hat{n}_i - 1), \quad (4)$$

with  $t(\Delta x) = (\hbar^2/2m)/\Delta x^2$ ,  $U(\Delta x) = g/\Delta x$ , and  $\mu_i(\Delta x) = V((i + 1/2)\Delta x) + 2(\hbar^2/2m)/\Delta x^2$ . A similar lattice model can be formulated for fermions.

With this definition of the Hamiltonian for arbitrary  $\Delta x$ , the implementation of MG DMRG is straightforward. The matrix elements of the isometry  $T$  are

$$T_{\tilde{\sigma}, \sigma_1, \dots, \sigma_n}^{\tilde{\sigma}'} = \frac{\delta(\tilde{\sigma}, \sigma_1 + \dots + \sigma_n)}{\sqrt{\sum_{\sigma_1, \dots} \delta(\tilde{\sigma}, \sigma_1' + \dots + \sigma_n')}}}, \quad (5)$$

where  $\sigma_i$ ,  $\sigma_i'$  and  $\tilde{\sigma}$  are particle-number eigenstates and we truncate the maximum occupation of a site at  $N_{\max}$ , i.e.,  $\sigma_i, \tilde{\sigma} \in \{0, \dots, N_{\max}\}$ . Because of particle-number conservation,  $T$  is a  $N_{\max} \times N_{\max}^n$  block-diagonal matrix. Note that we start from a coarse-grained lattice and perform only prolongations.

As a benchmark for the multigrid algorithm, we consider an optical lattice with  $V_0/E_r = 6$  and  $1/\gamma = 0.1$  corresponding to the insulating phase [23] at unit filling. Our MG DMRG simulations were performed with up to  $N = 128$  lattice sites per unit cell ( $\Delta x = 0.0078125$ ) with  $N_{\max} = 2$ , keeping  $M = 200$  states and using 12 sweeps of single-site updates at each level. We also perform standard DMRG simulations starting from either a random initial state, a state obtained from an infinite-size growing procedure, or a few steps of imaginary time evolution (not shown). The infinite-size growing procedure is commonly used to obtain good initial states for one-dimensional systems and has been proven to be very efficient in most cases. We use the implementation of the ALPS DMRG code [24], which performs the growing procedure on a state with very small bond dimension and increases the bond dimension linearly with the number of sweeps thereafter.

Figure 2 shows the density profile obtained with the three approaches. Clearly, the standard DMRG approaches are trapped in configurations with a globally nonhomogeneous density distribution, and further sweeps are not effective in redistributing the particles. The multigrid

method, on the other hand, achieves a symmetric distribution, since it performs the optimization of the large-scale features on the initial coarse mesh with just  $N = 2$  sites per unit cell, where convergence is very fast. Subsequent calculations on finer lattices are initialized with the prolonged solution of the coarser lattice. This is already close to the ground state, and only the local fine-scale features of the wave function need to be optimized.

The better convergence of the MG DMRG is also reflected in the energies shown in Fig. 5. While, for modest discretizations, standard approaches yield energies comparable to MG DMRG, they encounter severe convergence problems for smaller values of  $\Delta x$  where the difference between multiple scales of the dilute system become more and more important. The most reliable method is MG DMRG combined with optimization in the prolongation.

MG DMRG opens new interesting applications for DMRG that have not been accessible before. As an example, we combine MG DMRG with time evolution [7–9] to study heating caused by nonadiabaticity when ramping up the amplitude of an optical lattice. We start from the ground state of a homogeneous system of length  $L = 16$  and  $N = 16$  grid points per unit cell calculated by MG DMRG. We evolve it in time using a fourth-order Trotter decomposition with  $\Delta t = 0.01\hbar/E_r$ . Nonadiabaticities due to ramping at a finite speed cause heating, and we plot the energy difference to the ground state in Fig. 6 for three different ramp profiles and several total ramping times. For the calculation of the ground-state energies in weak optical lattices, MG DMRG was used. We observe that as the ramp speed decreases, differences in ramp shape are less important than the total ramping time, indicating that the

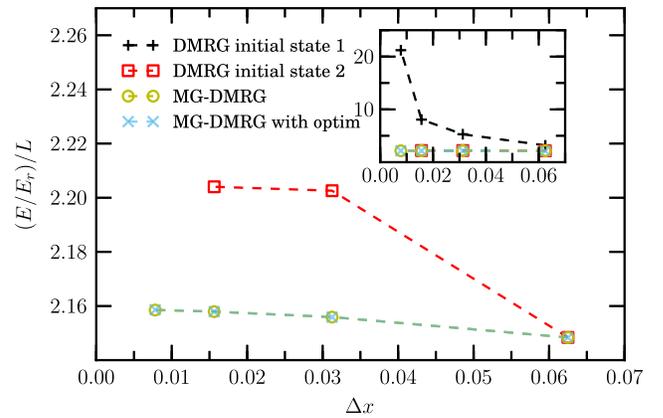


FIG. 5 (color online). Comparison of the energies in a bosonic optical lattice ( $V_0/E_r = 6$ ,  $1/\gamma = 0.1$ ) with  $L = 32$  unit cells discretized with increasing discretization  $N = 16, 32, 64, 128$  obtained with different strategies: DMRG with initial state 1 optimizing an initial random state, DMRG with initial state 2 initializing the system with an infinite size procedure and linearly increasing the number of states (results obtained with the ALPS DMRG code [24]), MG DMRG, and MG DMRG combined with local optimization in the prolongations.

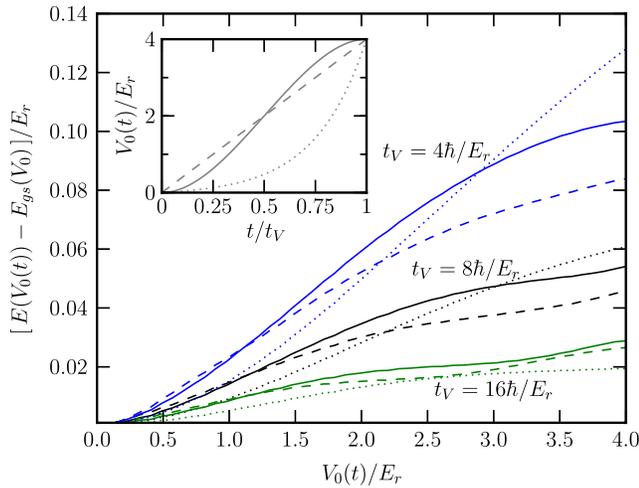


FIG. 6 (color online). Heating due to finite ramping speeds in a system of length  $L = 16$  for  $1/\gamma = 0.08$  with  $M = 400$ . Different lines show the energy difference to the ground state while ramping up to  $V_0/E_r = 0 \rightarrow 4$ . (inset) Ramping functions used in the time evolution: linear,  $V_0(t)/E_r = 4t/t_V$  (dashed); exponential,  $V_0(t)/E_r = 4[\exp(4t/t_V) - 1]/[\exp(4) - 1]$  (dotted); and  $s$  like,  $V_0(t)/E_r = 4[3(t/t_V)^2 - 2(t/t_V)^3]$  (solid).

exact shape of the ramp profiles play a minor role in experiments, and experimentalists should focus on determining optimal ramping times.

In DMRG simulations of weakly doped  $t$ - $J$  or Hubbard ladder models [25–27], the hole density shows similar convergence problems as seen above for dilute particle systems. In particular, it has been observed that for six holes in more than  $2 \times 64$  sites the standard DMRG algorithm fails to distribute the three bound hole pairs evenly over the ladder [28], and MG DMRG can be of use here. The restriction step of mapping the model to a coarser lattice is, however, not as straightforward as in continuum models. We propose to use the contractor renormalization (CORE) method [29] to find a good approximation of the model in the reduced Hilbert space of the coarser models, and to iterate this procedure in further restriction steps. For the specific case of doped ladder models, the first step maps two-site rungs or four-site plaquettes to a hardcore boson model for the hole pairs or an extended plaquette model containing hole pairs, magnons, and holes [30]. Further restriction steps map to simpler bosonic models for the hole pairs, as illustrated in Fig. 7. After prolongation back to the full lattice model, the ground-state wave function can be further improved by repeating the multigrid scheme. Now one can use knowledge of the approximate ground state to perform the restrictions of the basis instead of using CORE. Details of this method and results of this approach will be published elsewhere.

We point out that MG DMRG is a fundamentally different approach from the one taken by tree-tensor networks [31] or the multiscale entanglement renormalization ansatz (MERA) [32,33]. In those approaches, a new class of wave

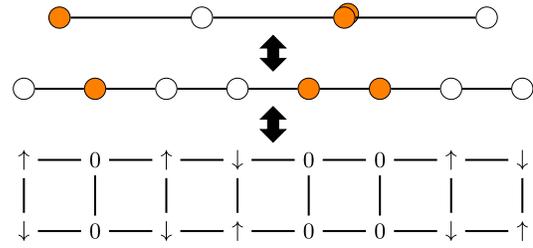


FIG. 7 (color online). Coarse graining a ladder of spin-1/2 fermions into a system of hard-core bosons. Spin singlets are mapped to empty sites and hole pairs to hard-core bosons in the first step. Further restriction steps of the bosonic model are similar to dilute bosonic models discussed above.

functions is proposed that describes the system at several levels of coarse graining, and all levels are optimized simultaneously, which can still suffer from convergence problems at fine scales. Instead, our approach relies on standard matrix-product states which can be optimized and evaluated much more easily and much faster but uses a hierarchical coarse graining to achieve a faster and more reliable optimization than standard DMRG.

Our algorithm can be easily combined with other optimization schemes for DMRG, such as using time evolving block decimation [8,34] to directly simulate the thermodynamic limit. One can also easily generalize the restriction Eq. (2) and prolongation to tensors of higher rank in order to apply the multigrid scheme to other tensor network states, e.g., multiscale entanglement renormalization [32,33], projected entangled pair states (PEPS) [35], and infinite PEPS [36].

This project was supported by a grant of ETH Zurich. The calculations were performed with the MAQUIS-DMRG code, developed with support of the Swiss High Performance and High Productivity Computing (HP2C) initiative and based on the ALPS libraries [24]. Z. R. acknowledges the ANR Grant No. 09-BLAN-0097-01/2.

- [1] S. Östlund and S. Rommer, *Phys. Rev. Lett.* **75**, 3537 (1995).
- [2] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [3] S. R. White and R. M. Noack, *Phys. Rev. Lett.* **68**, 3487 (1992).
- [4] U. Schollwöck, *Rev. Mod. Phys.* **77**, 259 (2005).
- [5] U. Schollwöck, *Ann. Phys. (N.Y.)* **326**, 96 (2011).
- [6] E. Stoudenmire and S. R. White, *Annu. Rev. Condens. Matter Phys.* **3**, 111 (2012).
- [7] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, *J. Stat. Mech.* (2004) P04005.
- [8] G. Vidal, *Phys. Rev. Lett.* **91**, 147902 (2003).
- [9] S. R. White and A. E. Feiguin, *Phys. Rev. Lett.* **93**, 076401 (2004).
- [10] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 207204 (2004).

- [11] A. E. Feiguin and S. R. White, *Phys. Rev. B* **72**, 220401 (2005).
- [12] S. R. White, *Phys. Rev. Lett.* **102**, 190601 (2009).
- [13] While a recently developed approach allows the simulation of continuum systems without introducing a lattice [14], the standard DMRG method on a fine mesh is currently more robust and accurate [15].
- [14] F. Verstraete and J. I. Cirac, *Phys. Rev. Lett.* **104**, 190405 (2010).
- [15] M. Dolfi, Master's thesis, ETH Zurich, 2011.
- [16] E. Stoudenmire, L. Wagner, S. White, and K. Burke, [arXiv:1107.2394](https://arxiv.org/abs/1107.2394).
- [17] K. Stüben, *J. Comput. Appl. Math.* **128**, 281 (2001).
- [18] J. Goodman and A. D. Sokal, *Phys. Rev. Lett.* **56**, 1015 (1986).
- [19] M. Heiskanen, T. Torsti, M. J. Puska, and R. M. Nieminen, *Phys. Rev. B* **63**, 245106 (2001).
- [20] F. Verstraete and J. I. Cirac, *Phys. Rev. B* **73**, 094423 (2006).
- [21] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, *Phys. Rev. Lett.* **100**, 030504 (2008).
- [22] This requires a representation of the Hamiltonian on a nonuniform grid.
- [23] H. P. Büchler, G. Blatter, and W. Zwerger, *Phys. Rev. Lett.* **90**, 130401 (2003).
- [24] B. Bauer *et al.*, *J. Stat. Mech.* (2011) P05001.
- [25] E. Dagotto, J. Riera, and D. Scalapino, *Phys. Rev. B* **45**, 5744 (1992).
- [26] M. Troyer, H. Tsunetsugu, and T. M. Rice, *Phys. Rev. B* **53**, 251 (1996).
- [27] R. M. Noack, N. Bulut, D. J. Scalapino, and M. G. Zacher, *Phys. Rev. B* **56**, 7162 (1997).
- [28] T. Siller, M. Troyer, T. M. Rice, and S. R. White, *Phys. Rev. B* **63**, 195106 (2001).
- [29] C. J. Morningstar and M. Weinstein, *Phys. Rev. Lett.* **73**, 1873 (1994).
- [30] E. Altman and A. Auerbach, *Phys. Rev. B* **65**, 104508 (2002).
- [31] Y.-Y. Shi, L.-M. Duan, and G. Vidal, *Phys. Rev. A* **74**, 022320 (2006).
- [32] G. Vidal, *Phys. Rev. Lett.* **99**, 220405 (2007).
- [33] G. Vidal, *Phys. Rev. Lett.* **101**, 110501 (2008).
- [34] G. Vidal, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [35] F. Verstraete and J. I. Cirac, [arXiv:cond-mat/0407066](https://arxiv.org/abs/cond-mat/0407066).
- [36] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, *Phys. Rev. Lett.* **101**, 250602 (2008).