Nonstochastic algorithms for Jastrow-Slater and correlator product state wave functions

Eric Neuscamman,[1] Hitesh Changlani,[2] Jesse Kinder,[1] and Garnet Kin-Lic Chan[1]
[1]*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA*
[2]*Department of Physics, Cornell University, Ithaca, New York 14853, USA*
(Received 23 April 2011; revised manuscript received 26 August 2011; published 17 November 2011)

Jastrow-Slater and correlator product state wave functions, two classes of quantum many-body wave functions, are commonly studied using Monte Carlo methods with the associated drawbacks of stochastic error. Here we show that efficient nonstochastic algorithms for these wave functions exist, both for observable evaluation and for optimization. The algorithms rely on the structure of these states as a product of local, commuting, invertible operators acting on a simple reference wave function. We describe the nonstochastic energy evaluation and optimization algorithms, and demonstrate them with applications to the Heisenberg and spinless and full Hubbard models. Our results demonstrate that the nonstochastic algorithms yield optimized wave functions and energies very close to those obtained with the variational Monte Carlo algorithm. Such algorithms provide new criteria for identifying new classes of wave functions for efficient computational simulation.

DOI: 10.1103/PhysRevB.84.205132 PACS number(s): 71.27.+a, 71.15.−m, 71.10.Fd, 75.10.Jm

I. INTRODUCTION

Explicit wave-function approximations present a powerful approach to understanding the behavior of quantum many-body systems. For a wave function to be computationally useful, it is not sufficient to simply have a compact form. Instead it must also be possible to efficiently evaluate expectation values. Expectation values formally involve a trace over the Hilbert space, which is of exponential dimension. The evaluation of this trace can often be carried out efficiently with Monte Carlo techniques. However, this introduces statistical error, which must be controlled and which can lead to certain complications, for example in wave-function optimization.

The Jastrow-Slater wave function[1,2] is one of the most common wave functions used together with Monte Carlo sampling in fermionic systems. As the name suggests, the wave function augments a Slater determinant, the appropriate state for a system of noninteracting fermions, with a set of Jastrow factors. Jastrow factors are diagonal operators in the lattice basis and can be thought of as directly introducing correlations between particles in the system. Jastrow-Slater wave functions also form the starting point for more flexible wave functions such as the Jastrow-antisymmetrized geminal power[3–5] and Jastrow-Bardeen-Cooper-Schrieffer[6] wave functions.

Recently, a second class of wave functions, termed correlator product states (CPS),[7] entangled plaquette states (EPS),[8,9] or complete graph tensor networks,[10] have also been studied with Monte Carlo methods. We refer to these wave functions collectively as correlator product states here. These wave functions were constructed as a way to generalize the density-matrix renormalization group (DMRG)[11,12] to higher dimensions, while avoiding the computational cost of recent classes of tensor network wave functions such as the projected entangled pair states (PEPS)[13–23] or multiscale entanglement renormalization ansatz (MERA).[24,25] The correlator product state directly approximates a wave function as a product of correlator amplitudes. Correlators are diagonal operators, and are thus mathematically identical to Jastrow factors. Consequently the correlator product state is also a pure Jastrow wave function. In this work the term correlator will be used to refer to both correlators as used in the CPS, and the Jastrow factors used in the Jastrow-Slater wave function.

While the form of the Jastrow-Slater and CPS wave functions make them a natural fit for Monte Carlo algorithms, it is interesting to ask whether or not statistical algorithms, with their associated drawbacks of statistical error, are the only way to manipulate these states. Here we will show that efficient nonstochastic algorithms exist both to evaluate observables and to optimize the CPS and Jastrow-Slater wave function, so long as we are willing to sacrifice the strict variational principle. The nonstochastic algorithms rely on the common structure of the Jastrow-Slater and CPS wave functions, namely that they are a product of local, commuting, invertible operators acting on a simple reference state. Other wave functions, such as the coupled cluster wave function, also take this form, and the algorithms we describe are similar to the techniques used for observable evaluation and wave-function optimization in the coupled cluster literature.[26]

As we shall see, the nonstochastic methodology we employ can be written in terms of a similarity transform on the Hamiltonian that attempts to convert the system of correlated physical particles into a system of weakly correlated or uncorrelated quasiparticles. In this respect our approach shares many similarities with a number of other methods, in particular the transcorrelated wave equation (TWE),[27] the method of correlated basis functions (CBFs),[28] and coupled cluster theory. In each of these methods, as well as in our own, one must limit the form of the modified Hamiltonian and possibly introduce other approximations in order to achieve computational feasibility. It is the nature of these limitations and approximations that differentiate these methods. In the TWE, the correlating factor (which here we would refer to as a correlator or Jastrow factor) is limited to correlating no more than two particles at a time, allowing one to evaluate the necessary integrals by numerical quadrature. In the CBF approach, the variational energy expression is simplified through a perturbatively motivated truncation based on the ratio of the (supposedly small) volume in which two particles correlate to the total volume of the system, and the resulting expression is minimized using a linear combination of determinants rather than a single determinant. Our approach also uses the concept of a small correlation volume, but not in a perturbative manner. Instead, we explicitly limit the range over which

our "correlating functions" (correlators or Jastrow factors) can operate, which makes the complexity of the operators in the similarity transformed Hamiltonian independent of system size. Unlike the TWE and CBF methods, our method permits the use of many-body (i.e., more than two-body) correlation functions, which is critical when working with CPS tensor networks. Finally, coupled cluster theory uses an approach that is mathematically very similar to ours, although there the form of the correlating function is different and the limitation is placed not on the range of the correlations in terms of real-space distances but rather on the particle excitation rank relative to a reference state. To make a connection with the concept of locality, limiting the excitation rank in coupled cluster can be viewed as limiting the "energetic distance" from the reference.

We begin by reviewing the structure of the CPS and Jastrow wave functions (Sec. II A). We then proceed to describe the energy evaluation (Sec. II B) and optimization algorithms (Sec. II C). Demonstrations of these nonstochastic algorithms are carried out on the Heisenberg model (Sec. III A), the spinless Hubbard model (Sec. III B), and the full Hubbard model (Sec. III C). Our purpose in applying the method to these models is to establish the validity of our approximations, and so in this study we have intentionally limited ourselves to systems for which the correct answers are more or less already known. After these lattice model tests, we finish with our conclusions in Sec. IV.

II. THEORY

A. CPS and Jastrow-Slater wave functions

Consider an arbitrary quantum wave function on k lattice sites. This can be written as

$$|\Psi\rangle = \sum_{n_1 n_2 \cdots n_k} \Psi_{n_1 n_2 \cdots n_k} |n_1 n_2 \cdots n_k\rangle = \sum_{\mathbf{n}} \Psi_{\mathbf{n}} |\mathbf{n}\rangle, \tag{1}$$

where $\mathbf{n}$ denotes the vector of occupancies $n_1 n_2 \cdots n_k$. In a spin-1/2 system,

$$|n_i\rangle \in \{|\uparrow\rangle, |\downarrow\rangle\}, \tag{2}$$

while in a fermion system,

$$|n_i\rangle \in \{|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}. \tag{3}$$

Here, we describe the CPS and Jastrow-Slater approximations to $|\Psi\rangle$. We first define a Jastrow factor. A Jastrow factor is an operator that is diagonal in the lattice basis $|\mathbf{n}\rangle$. It is usually written as an exponential of an expansion in number operators,

$$\hat{J} = \exp\left(j + \sum_i j_i \hat{n}_i + \sum_{ij} j_{ij} \hat{n}_i \hat{n}_j + \cdots\right). \tag{4}$$

In most applications, the Jastrow amplitudes $j, j_i, j_{ij} \ldots$, are restricted to be real numbers, which makes the Jastrow factor $\hat{J}$ positive definite. For full generality, however, we should regard the amplitudes as possibly being complex.

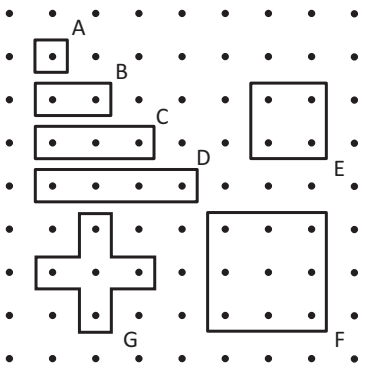

FIG. 1. Examples of correlators on the square lattice. A: one-site correlator, similar to a Gutzwiller factor. B: nearest-neighbor two-site correlator. C: three-site line correlator. D: four-site line correlator. E: four-site square correlator. F: nine-site square correlator. G: five-site cross correlator.

A correlator is also an operator that is diagonal in the lattice basis. It can be viewed as a Jastrow factor written in nonexponential form. A single correlator acts on a domain of sites. For example, a correlator on sites i, j takes the form

$$\hat{c}_{ij} = \sum_{n_i n_j} |n_i n_j\rangle\, c_{n_i n_j}\, \langle n_i n_j| = \sum_{n_i n_j} c_{n_i n_j} \hat{P}_{n_i n_j}, \tag{5}$$

where $c_{n_i n_j}$ are the correlator amplitudes, and we have introduced the projection operator $\hat{P}_{n_i n_j} = |n_i n_j\rangle\langle n_i n_j|$. The equivalence of the correlator and Jastrow factor is seen by recognizing that a two-site correlator $\hat{c}_{ij}$ is exactly representable by a two-site Jastrow factor $\exp(j + j_i \hat{n}_i + j_{ij} \hat{n}_i \hat{n}_j)$, so long as the Jastrow amplitudes are allowed to be complex. The choice of using the exponentiated Jastrow representation or the correlator representation is a matter of numerical expediency. In the current work, we henceforth use the term "correlator" to refer to both representations.

A correlator can be chosen to act on an arbitrary number of sites (see Figs. 1 and 2). Such a general correlator is written as

$$\hat{c}_\lambda = \textstyle\sum_{\mathbf{n}_\lambda} c_{\mathbf{n}_\lambda} \hat{P}_{\mathbf{n}_\lambda}, \tag{6}$$

where $\mathbf{n}_\lambda$ is the occupancy vector of the sites in the domain of the correlator, and $\hat{P}_{\mathbf{n}_\lambda}$ is the corresponding projector.

Both Jastrow factors and correlators can be applied to reference wave functions $|\Phi\rangle$ to generate approximations to $|\Psi\rangle$. We start with the correlator product state (CPS). This is obtained from the uniform reference function $|\Phi_U\rangle$, an equally weighted sum over the lattice quantum basis,

$$|\Phi_U\rangle = \sum_{\mathbf{n}} |\mathbf{n}\rangle. \tag{7}$$

The summation in Eq. (7) may be chosen with symmetry constraints. For example, in this work, we always use a uniform reference such that for spin systems, the summation in Eq. (7) refers to states only with given S_z, while for fermionic system, to states with only given N and given S_z.

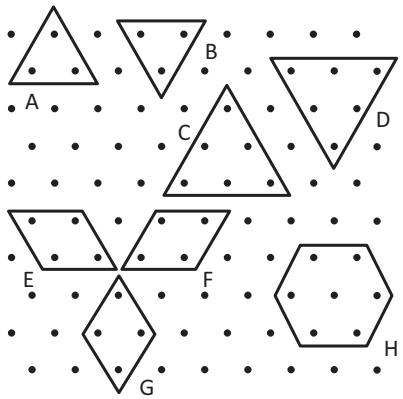

FIG. 2. Examples of correlators on the triangular lattice. A and B: the two orientations of three-site triangle correlators. C and D: the two orientations of six-site triangle correlators. E, F, and G: the three orientations of four-site rhombus correlators. H: a seven-site hexagon correlator.

The CPS is then obtained by applying a product of correlators to $|\Phi_U\rangle$,

$$|\Psi\rangle = \hat{C}|\Phi_U\rangle = \prod_\lambda \hat{c}_\lambda |\Phi_U\rangle. \tag{8}$$

The correlator amplitudes $c_{\mathbf{n}_\lambda}$ provide a product approximation to the wave-function amplitudes $\Psi_{\mathbf{n}}$ in Eq. (1),

$$\Psi_{\mathbf{n}} = \prod_\lambda c_{\mathbf{n}_\lambda}, \tag{9}$$

and this expression can be evaluated efficiently in a time proportional to the number of correlators. Note that the domains λ of the different correlators will usually contain overlapping sites. For example, for a CPS wave function in one dimension with nearest-neighbor correlators, we have

$$|\Psi\rangle = \hat{c}_{12}\hat{c}_{23}\hat{c}_{34}\cdots\hat{c}_{k-1k}|\Phi_U\rangle. \tag{10}$$

By using correlators that cover increasingly larger numbers of sites, CPSs become an exact family of states.

The Jastrow-Slater fermion wave function is obtained by applying a set of correlators (i.e., Jastrow factors) to a Slater determinant reference of orbitals $|\Phi_D\rangle = \det|\phi_1\phi_2\cdots\phi_N|$,

$$|\Psi\rangle = \hat{C}|\Phi_D\rangle. \tag{11}$$

The wave-function amplitudes in this case are given by

$$\Psi_{n_1 n_2 \ldots n_k} = \prod_\lambda c_{\mathbf{n}_\lambda} \times \det|\phi_1(r_1)\phi_2(r_2)\cdots\phi_N(r_N)|, \tag{12}$$

where $r_1, r_2 \ldots r_N$ label the positions of the N sites occupied in the occupancy vector $|n_1 n_2 \cdots n_k\rangle$. The determinant contribution to the wave-function amplitude can be evaluated efficiently in $O(N^3)$ time. As in the case of the CPSs, Jastrow-Slater wave functions become an exact family of states if the correlators (Jastrows) cover larger and larger numbers of sites.

Both the CPS and Jastrow-Slater wave functions have so far been used in conjunction with Monte Carlo algorithms. In variational Monte Carlo,[29,30] the energy is written as

$$E = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \sum_{\mathbf{n}} \frac{|\Psi_{\mathbf{n}}|^2}{\langle\Psi|\Psi\rangle} E_{\mathrm{L}}(\mathbf{n}), \tag{13}$$

where the local energy $E_{\mathrm{L}}(\mathbf{n})$ is defined by

$$E_{\mathrm{L}}(\mathbf{n}) = \textstyle\sum_{\mathbf{n}'} \frac{\Psi_{\mathbf{n}'}}{\Psi_{\mathbf{n}}} \langle\mathbf{n}|H|\mathbf{n}'\rangle. \tag{14}$$

As long as $\Psi_{\mathbf{n}}$ can be evaluated efficiently, which is the case for both the CPS and Jastrow-Slater wave functions, a Markov chain can be used to sample the probability distribution $|\Psi_{\mathbf{n}}|^2/\langle\Psi|\Psi\rangle$ and to efficiently compute the overall energy as an average of the sampled local energies. The wave functions may also be variationally optimized using stochastic estimates for the gradient of the energy with respect to the wave-function amplitudes. Variational Monte Carlo is a stochastic approach to using the CPS and Jastrow-Slater wave functions. A disadvantage of stochastic algorithms is the presence of statistical error, which must be controlled by increasing the sample size. We now show how statistical noise can be avoided by using nonstochastic algorithms both to evaluate the energy of, and to optimize the amplitudes of, CPS and Jastrow-Slater wave functions.

B. Nonstochastic energy evaluation

The idea behind our approach to nonstochastic energy evaluation is the following. First, we assume that the CPS or Jastrow wave function $|\Psi\rangle = \hat{C}|\Phi\rangle$ (we use $|\Phi\rangle$ to refer to either the uniform or determinant reference) is an eigenstate of the Hamiltonian H,

$$H\hat{C}|\Phi\rangle = E\hat{C}|\Phi\rangle. \tag{15}$$

The energy is then obtained as the asymmetric expectation value,

$$E = \langle\Phi|\hat{C}^{-1}H\hat{C}|\Phi\rangle = \langle\Phi|\bar{H}|\Phi\rangle, \tag{16}$$

where we define the similarity transformed effective Hamiltonian, $\bar{H} = \hat{C}^{-1}H\hat{C}$.

Efficient nonstochastic energy evaluation now reduces to whether or not $\langle\bar{H}\rangle$ can be efficiently obtained with the reference state $|\Phi\rangle$, where $|\Phi\rangle$ is a uniform reference in the case of CPSs, or a determinant in the case of the Jastrow-Slater wave function. By efficient, we mean that the cost is polynomial in the lattice size. The standard expectation value $\langle\Phi|H|\Phi\rangle$ can be efficiently evaluated in either case because the individual terms in the Hamiltonian act on a small number of sites, independent of lattice size. For example, the Heisenberg and Hubbard Hamiltonians contain terms that only act at most on a pair of sites. The corresponding terms in the effective Hamiltonian $\bar{H}$ act on a larger number of sites due to the similarity transformation by the correlators. However, if the size and range of the correlators are independent of lattice size, then even after similarity transformation, the terms in $\bar{H}$ still act on a number of sites that will be independent of lattice size. Consequently, the expectation value $\langle\bar{H}\rangle$ can still be evaluated at a cost scaling polynomially with respect to the lattice size. As we shall see, the magnitude of the evaluation

cost (although not its scaling with respect to system size) is strongly dependent on the range of the correlators. One of the questions we seek to answer in this work is what the practical limits for correlator range are and what accuracy can be achieved within these limits.

For a more concrete view of the nonstochastic energy evaluation, consider the energy contribution E_{xy} from a hopping operator $a_x^\dagger a_y$ associated with sites x and y,

$$E_{xy} = \langle\Phi|\hat{C}^{-1}a_x^\dagger a_y \hat{C}|\Phi\rangle. \tag{17}$$

We can divide the correlators in $\hat{C}$ (and their inverses in $\hat{C}^{-1}$) into two classes: those that involve (touch) sites x or y and those that do not. Denote the product of all the correlators that involve x or y as $\hat{C}_{xy}$, and the product of the remaining correlators as $\hat{C}_{\overline{xy}}$. Then we have

$$\hat{C} = \hat{C}_{xy}\,\hat{C}_{\overline{xy}}. \tag{18}$$

As an example, consider the hopping operator $a_3^\dagger a_4$, and the one-dimensional CPS in Eq. (10). Then, $\hat{C}_{34}$ is given by

$$\hat{C}_{34} = \hat{c}_{23}\hat{c}_{34}\hat{c}_{45}. \tag{19}$$

Because $\hat{C}_{\overline{xy}}$ only contains correlators that do *not* involve sites x or y, it can be commuted past $a_x^\dagger a_y$ in Eq. (17) to cancel with its corresponding inverse $\hat{C}_{\overline{xy}}^{-1}$,

$$\begin{aligned} E_{xy} &= \langle\Phi|\hat{C}_{xy}^{-1}\,\hat{C}_{\overline{xy}}^{-1}\,a_x^\dagger a_y\,\hat{C}_{xy}\,\hat{C}_{\overline{xy}}|\Phi\rangle \\ &= \langle\Phi|\hat{C}_{xy}^{-1}a_x^\dagger a_y \hat{C}_{xy}|\Phi\rangle. \end{aligned} \tag{20}$$

Thus the similarity transform of $a_x^\dagger a_y$ involves only $\hat{C}_{xy}$, not the whole $\hat{C}$ operator. The correlators in $\hat{C}_{xy}$ define a cluster of k_{xy} sites, where the size depends on the sizes and ranges of the correlators and the geometry of the cluster, but is independent of lattice size as long as the correlators are local. The transformed hopping operator $\hat{C}_{xy}^{-1}a_x^\dagger a_y\hat{C}_{xy}$ now acts on k_{xy} sites. The energy contribution E_{xy} thus requires evaluating the expectation value of a k_{xy} site operator with the reference function $|\Phi\rangle$.

To see the explicit dependence of the evaluation of $\langle a_x^\dagger a_y\rangle$ on the reference function, we separate $\hat{C}_{xy}$ into its amplitude and projection operator components,

$$\hat{C}_{xy} = \sum_{\mathbf{n}_{xy}} C_{\mathbf{n}_{xy}}\hat{P}_{\mathbf{n}_{xy}}, \tag{21}$$

where $\mathbf{n}_{xy}$ is the occupancy vector for the cluster of sites defined by $\hat{C}_{xy}$. The expectation values of the projection operators define a many-body reduced density-matrix (RDM) γ on k_{xy} sites,

$$\gamma_{\mathbf{n}_{xy},\mathbf{n}'_{xy}} = \langle\Phi|\hat{P}_{\mathbf{n}_{xy}}a_x^\dagger a_y\hat{P}_{\mathbf{n}'_{xy}}|\Phi\rangle. \tag{22}$$

The evaluation of γ depends on the form of the reference function $|\Phi\rangle$. In the case of the CPS, each element of the RDM is obtained in $O(1)$ time (even with particle number and S_z restrictions). For a Jastrow-Slater wave function, the corresponding RDM element can be evaluated as a determinant of one-body RDM elements in $O(k_{xy}^3)$ time. Once the RDM is obtained, the combination with the amplitudes is independent of the reference. The expectation value E_{xy} becomes

$$E_{xy} = \sum_{\mathbf{n}_{xy}\,\mathbf{n}'_{xy}} C_{\mathbf{n}_{xy}}^{-1}\,\gamma_{\mathbf{n}_{xy},\mathbf{n}'_{xy}}\,C_{\mathbf{n}'_{xy}}. \tag{23}$$

Due to the summation over the occupancy vector, the cost of Eq. (23) is exponential in the cluster size, k_{xy}, but not in the lattice size. Note that the cost is effectively the cost of a single summation over $\mathbf{n}_{xy}$, rather than the formal double summation shown above, because the sparsity of γ means that it has $O(d^{k_{xy}})$ nonzero elements, where d is dimension of a single site. For sufficiently short-range correlators, which lead to small clusters, the summation can be carried out affordably. As each operator in the Hamiltonian involves a similar contribution, the entire energy may be efficiently evaluated.

The nonstochastic energy evaluation algorithm relies on little of the detailed structure of the Jastrow-Slater and CPS wave functions. The key steps require only that (i) $\hat{C}$ is made of a product of local, commuting, invertible operators, (ii) only a small number of these operators do not commute with a given term in the Hamiltonian, and (iii) the reference function $|\Phi\rangle$ is sufficiently simple that the expectation values in Eq. (22) can be efficiently obtained. This structure is obeyed by many other wave functions, such as the coupled cluster wave function[26] (although there locality refers to the particle rank of excitations out of some reference function rather than a distance in real space), and indeed for any wave function with this structure, an efficient nonstochastic energy evaluation algorithm may be formulated.

C. Nonstochastic wave-function optimization

Above we showed that we can evaluate the approximate energy of a CPS or Jastrow-Slater wave function in a nonstochastic way, as long as the individual correlators or Jastrow factors do not cover too many sites. For a complete calculation, we need to also determine the correlator or Jastrow amplitudes. Since the nonstochastic energy is not variational, we cannot obtain the optimal parameters by a minimization of the energy. Instead, we require that the CPS or Jastrow wave function satisfies a set of nonlinear projected Schrödinger equations. Solving these equations yields exactly the same solution as the minimum of the variational energy if the wave function provides an exact parametrization, although this is not the case for approximate CPS and Jastrow wave functions.

As in the previous section, we assume that the CPS or Jastrow wave function is a true eigenstate of the Hamiltonian, Eq. (15). From the Schrödinger equation, we need to obtain a set of nonlinear equations equal in number to the number of amplitudes in $\hat{C}$. We obtain sufficient equations by projecting with the bras $\langle\Phi|\hat{P}_{\mathbf{n}_\lambda}$, where $\hat{P}_{\mathbf{n}_\lambda}$ are the projectors used to define the correlator operators in Eq. (6). Applying these bra states gives us

$$\langle\Phi|\hat{P}_{\mathbf{n}_\lambda}(\bar{H} - E)|\Phi\rangle = R_{\mathbf{n}_\lambda} = 0. \tag{24}$$

By requiring the residuals $R_{\mathbf{n}_\lambda}$ to vanish we determine all the correlator amplitudes $c_{\mathbf{n}_\lambda}$. As with energy evaluation, the expectation value in Eq. (24) can be obtained in a nonstochastic

TABLE I. Comparison of nonstochastic and variational MC results for the antiferromagnetic spin-$\frac{1}{2}$ Heisenberg model on the 8×8 square lattice with periodic boundary conditions and total $S_z = 0$. The energies per site are reported in units of J. The variational Monte Carlo error is reported relative to the essentially exact stochastic series expansion (SSE) (Ref. 31). Translationally invariant correlators are employed, with each of the two sublattices having independent correlators. When optimizing the correlators, Marshall's sign rule is used as an initial guess. The acronyms NS and VNS refer to the nonstochastic and variational energies, respectively, of the wave function resulting from a nonstochastic optimization, while %D indicates the percent difference between these energies and the variational energy of the VMC wave function.

Correlators	VMC	NS	%D	VNS	%D	VMC % Error
Nearest neighbor	−0.6534(6)	−0.6883	−5.34	−0.6501(2)	0.51	2.98
Four-site squares	−0.6617(3)	−0.6659	−0.64	−0.6561(2)	0.84	1.76
Five-site crosses	−0.6637(3)	−0.6449	2.82	−0.6573(2)	0.97	1.46
Nine-site squares	−0.6699(2)	−0.6651	0.72	−0.6689(1)	0.15	0.53

manner by tracing over clusters of sites associated with each similarity transformed term in the effective Hamiltonian $\bar{H}$.

In order to solve the simultaneous set of equations (24), we have taken two approaches. In the first case, we use a standard Newton-Raphson procedure to find the simultaneous zeros of the residuals $R_{\mathbf{n}_\lambda}$. This requires evaluating the Jacobian matrix $\partial R_{\mathbf{n}_\lambda}/\partial c_{\mathbf{n}_\mu}$. Alternatively, we may take the approach of constructing and diagonalizing a local Hamiltonian for each correlator's amplitudes. The local Hamiltonian and overlap matrices that determine $c_{\mathbf{n}_\lambda}$ are defined as

$$\bar{H}_{\mathbf{n}_\lambda,\mathbf{n}'_\lambda} = \left\langle \Phi | \hat{P}_{\mathbf{n}_\lambda}\, c_{\mathbf{n}_\lambda}\, \bar{H}\, c^{-1}_{\mathbf{n}'_\lambda}\, \hat{P}_{\mathbf{n}'_\lambda} | \Phi \right\rangle, \tag{25}$$

$$S_{\mathbf{n}_\lambda,\mathbf{n}'_\lambda} = \langle \Phi | \hat{P}_{\mathbf{n}_\lambda}\, \hat{P}_{\mathbf{n}'_\lambda} | \Phi \rangle. \tag{26}$$

The correlator amplitudes are obtained from solving an eigenvalue problem for each correlator $\hat{c}_\lambda$,

$$\sum_{\mathbf{n}'_\lambda} \bar{H}_{\mathbf{n}_\lambda,\mathbf{n}'_\lambda} c_{\mathbf{n}'_\lambda} = E \sum_{\mathbf{n}'_\lambda} S_{\mathbf{n}_\lambda,\mathbf{n}'_\lambda} c_{\mathbf{n}'_\lambda}. \tag{27}$$

The local Hamiltonian matrix $\bar{H}_{\mathbf{n}_\lambda,\mathbf{n}'_\lambda}$ depends on the amplitudes of all the correlators, $\hat{c}_{\mu\neq\lambda}$. Thus after each correlator amplitude is obtained from the respective eigenvalue problem (27), the local Hamiltonians are updated, and the procedure is iterated until convergence is achieved.

Much as in the case of energy evaluation, the formulation of the amplitude equations relies only on generic elements of the product structure of the Jastrow-Slater and CPS wave functions. By analogy with methods for the coupled cluster wave function, we can also write down a nonstochastic algorithm to obtain expectation values of arbitrary operators. Starting from the amplitude equations, we first define a Lagrangian as

$$L = \langle \Phi | \bar{H} + \sum_{\mathbf{n}_\mu} \Lambda_{\mathbf{n}_\mu} \hat{P}_{\mathbf{n}_\mu} \left(\bar{H} - \langle \Phi | \bar{H} | \Phi \rangle \right) | \Phi \rangle, \tag{28}$$

from which the amplitude equations arise from the stationary conditions

$$\frac{\partial L}{\partial \Lambda_{\mathbf{n}_\mu}} = 0. \tag{29}$$

The values of the Lagrange multipliers $\Lambda_{\mathbf{n}_\mu}$ are found by requiring the Lagrangian to be stationary with respect to the correlator variables $c_{\mathbf{n}_\mu}$,

$$\frac{\partial L}{\partial c_{\mathbf{n}_\mu}} = 0. \tag{30}$$

Then, derivatives of the Lagrangian with respect to the Hamiltonian's parameters define reduced density matrices as

$$\Gamma_{ij} = \langle \Phi | \overline{a^\dagger_i a_j} + \sum_{\mathbf{n}_\mu} \Lambda_{\mathbf{n}_\mu} \hat{P}_{\mathbf{n}_\mu} (\overline{a^\dagger_i a_j} - \langle \Phi | \overline{a^\dagger_i a_j} | \Phi \rangle) | \Phi \rangle, \tag{31}$$

where $\overline{a^\dagger_i a_j} = \hat{C}^{-1} a^\dagger_i a_j \hat{C}$. The two body reduced density matrix Γ_{ijkl} is defined likewise using $\overline{a^\dagger_i a^\dagger_j a_l a_k}$. These density matrices allow us to obtain expectation values of arbitrary one- and two-body operators.

III. RESULTS

In this section we report results for a number of benchmark calculations in which we test the assumptions inherent to

TABLE II. Comparison of nonstochastic and variational MC results for the antiferromagnetic spin-$\frac{1}{2}$ Heisenberg model on the 6×6 triangular lattice with periodic boundary conditions and total $S_z = 0$. The energies per site are reported in units of J. Translationally invariant correlators are employed, with each of the three sublattices having independent correlators. When optimizing the correlators, the solution to the classical Heisenberg model is used as an initial guess. Here NS, VNS, and %D have the same meaning as in Table I.

Correlators	VMC	NS	%D	VNS	%D
Nearest neighbor	−0.5184(2)	−0.5253	−1.33	−0.5183(1)	0.02
Three-site triangles	−0.5184(2)	−0.5253	−1.33	−0.5183(1)	0.03
Four-site rhombuses	−0.5383(3)	−0.5353	0.55	−0.5108(2)	5.11
Six-site triangles	−0.5419(2)	−0.5289	2.40	−0.5353(1)	1.21
Seven-site hexagons	−0.5435(2)	−0.5390	0.83	−0.5397(1)	0.69

our nonstochastic approach. It is not within the scope of our study to investigate new phenomena with our technique. Instead we limit ourselves to systems for which the correct results are more or less known in order to assess the performance of our approach. To simplify the tabulation of data, we have used the abbreviations NS (nonstochastic energy of the nonstochastic wave function), VNS (variational energy of the nonstochastic wave function), and %D (percent difference from the variational energy of the variational wave function). Where appropriate, we include the statistical uncertainty of the final digit in parentheses.

A. Antiferromagnetic Heisenberg model

We have applied the correlator product state to the antiferromagnetic spin-$\frac{1}{2}$ Heisenberg model on a periodic 8×8 square lattice and a periodic 6×6 triangular lattice using both the nonstochastic and variational Monte Carlo (MC) frameworks. The Hamiltonian is written as

$$H = J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j, \tag{32}$$

where $\langle ij \rangle$ indicates nearest-neighbor pairs and $J > 0$. The results are summarized in Tables I and II. In the case of the square lattice, we have essentially exact stochastic series expansion (SSE) results with which to compare.[31] While the accuracy of the CPS ansatz is not the main question we are studying here (such studies can be found in Refs. 7–9), we see that both the nonstochastic and variational Monte Carlo energies are within 2% of the SSE result for four-site square correlators and about 1% for nine-site square correlators. The more central question here is the relative difference between the nonstochastic and variational Monte Carlo energies. We see that in all cases, the relative difference is comparable to the intrinsic energy error associated with the wave function, and in one case it is significantly smaller. As a measure of the variational quality of the nonstochastic results, we have also computed the variational energy of the wave functions resulting from our nonstochastic optimization (to save space in tabulation we term this the VNS energy). We see that in most cases this energy is very similar to the minimal variational energy, especially in the case of the largest correlators (nine-site squares and seven-site hexagons).

As expected, the nonstochastic energy is not variational, and for small correlators it tends to be slightly below the variational energy. More surprisingly, the convergence in accuracy for the nonstochastic energy is not monotonic with correlator size. We see both larger deviations from the variational Monte Carlo results, as well as lower accuracy in the total energy, for five-site crosses (in the square lattice) and six-site triangles (in the triangular lattice), than for some smaller correlators. We find that square and rhombus correlators do particularly well on the square and triangular lattices, respectively, with relative differences from the variational energy all within 1%. We speculate that these correlators' success is due in part to the fact that the amplitude equations used in their optimization [Eq. (24)] are constructed using projectors that share the translational symmetry of the underlying lattice.

One appealing aspect of the nonstochastic method for CPS on periodic spin lattices is that the cost of the method can be made independent of the lattice size by using translationally invariant correlators and by taking advantage of the uniform reference's particularly simple RDM elements [Eq. (22)]. As seen in Table VI, the nonstochastic method was much faster than variational MC for small correlators (five-site crosses/four-site rhombuses and smaller), while being slower for larger correlators due to the exponential increase of the cost with correlator size.

B. Spinless Hubbard model

We have studied a 20-site (4×5) spinless Hubbard lattice with open boundary conditions using the Jastrow-Slater wave function and the nonstochastic algorithms. We have also performed variational Monte Carlo calculations for comparison.

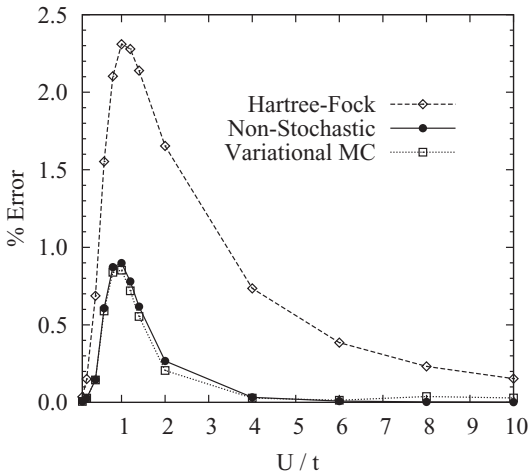

FIG. 3. Energy errors for the 20-site (4×5) spinless Hubbard model with ten particles and open boundary conditions. The nonstochastic and variational MC methods used four-site square Jastrow factors (correlators). Exact results were computed using the sparse diagonalization routine in the ALPS program (Ref. 32).

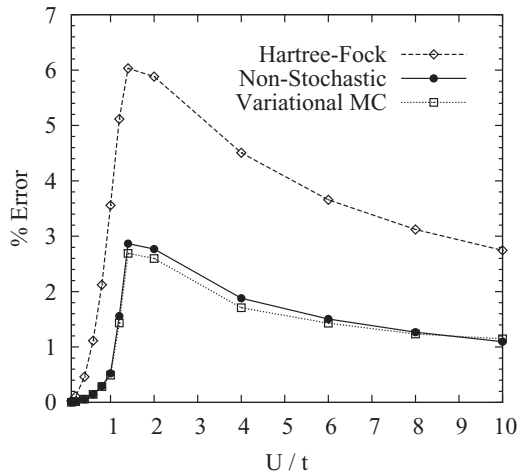

FIG. 4. Ground-state energy errors for the 20-site (4×5) spinless Hubbard model with nine particles and open boundary conditions. The nonstochastic and variational MC methods used four-site square Jastrow factors (correlators). Exact results were computed using the sparse diagonalization routine in the ALPS program (Ref. 32).

TABLE III. Total ground-state energies, in units of t, for the 4×5 spinless Hubbard lattice at half filling with open boundary conditions. Both the nonstochastic and variational MC methods use four-site square Jastrow factors (correlators). Exact results were computed using the sparse diagonalization routine in the ALPS program (Ref. 32). Here NS and %D have the same meaning as in Table I.

U/t	VMC	NS	%D	Hartree-Fock	Exact
0.1	−13.8158(1)	−13.8157	0.00	−13.8117	−13.8166
0.2	−13.2438(2)	−13.2437	0.00	−13.2276	−13.2475
0.4	−12.1368(4)	−12.1367	0.00	−12.0709	−12.1544
0.6	−11.0611(2)	−11.0591	0.02	−10.9539	−11.1267
0.8	−10.0884(2)	−10.0850	0.03	−9.9598	−10.1737
1.0	−9.2273(2)	−9.2230	0.05	−9.0917	−9.3066
1.2	−8.4712(2)	−8.4660	0.06	−8.3382	−8.5326
1.4	−7.8072(2)	−7.8022	0.06	−7.6827	−7.8507
2.0	−6.2529(2)	−6.2491	0.06	−6.1622	−6.2658
4.0	−3.6141(1)	−3.6139	0.00	−3.5885	−3.6151
6.0	−2.4988(1)	−2.4989	−0.01	−2.4895	−2.4991
8.0	−1.9000(1)	−1.9007	−0.03	−1.8963	−1.9007
10.0	−1.5304(1)	−1.5308	−0.03	−1.5285	−1.5308

While we could treat much larger lattices, we chose this lattice size in order to compare to exact results. The Hamiltonian for the spinless Hubbard model is

$$H = \sum_{\langle ij \rangle} -t(a_i^\dagger a_j + a_j^\dagger a_i) + U a_i^\dagger a_i a_j^\dagger a_j, \tag{33}$$

in which $a_i^\dagger$ and a_i are the fermionic particle creation and destruction operators on site i, and $\langle ij \rangle$ represents nearest neighbors.

Results for half filling and single hole doping are presented in Figs. 3 and 4 and Tables III and IV, respectively. We find that the difference between the nonstochastic and variational Monte Carlo energies is small for all ratios of U/t at both half filling and single hole doping. At half filling, the largest difference is 0.06%, while for single hole doping, it is 0.18%. The energy errors of the Jastrow-Slater form (compared to the exact energy) are less than 1% for all values of U/t at half filling, and below 3% for single hole doping. We see that the difference between the nonstochastic and variational Monte Carlo energies is here much smaller than the intrinsic error associated with the quality of the wave function.

C. Full Hubbard model

We have also studied the Hubbard model at half filling with open boundary conditions, in one and two dimensions, using the nonstochastic and variational Monte Carlo algorithms for the Jastrow-Slater wave function. The Hubbard Hamiltonian is

$$H = -t \sum_{\langle ij \rangle} \sum_{\sigma=\uparrow,\downarrow} (a_{i\sigma}^\dagger a_{j\sigma} + a_{j\sigma}^\dagger a_{i\sigma}) + U \sum_i a_{i\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow}^\dagger a_{i\downarrow}, \tag{34}$$

in which $a_{i\uparrow(\downarrow)}^\dagger$ and $a_{i\uparrow(\downarrow)}$ are the fermionic creation and destruction operators for particles with spin $\uparrow$ ($\downarrow$), and $\langle ij \rangle$ refers to nearest neighbors.

Since the fermions have spin, there are several choices of Slater determinant possible. We use as our Slater determinant

TABLE IV. Total ground-state energies, in units of t, for the 4×5 spinless Hubbard lattice with single hole doping and open boundary conditions. Both the nonstochastic and variational MC methods use four-site square Jastrow factors (correlators). Exact results were computed using the sparse diagonalization routine in the ALPS program (Ref. 32). Here NS and %D have the same meaning as in Table I.

U/t	VMC	NS	%D	Hartree-Fock	Exact
0.1	−13.8554(1)	−13.8554	0.00	−13.8522	−13.8558
0.2	−13.4339(1)	−13.4341	0.00	−13.4214	−13.4359
0.4	−12.6194(4)	−12.6195	0.00	−12.5685	−12.6268
0.6	−11.8411(2)	−11.8410	0.00	−11.7262	−11.8583
0.8	−11.0985(3)	−11.0976	0.01	−10.8940	−11.1303
1.0	−10.3921(3)	−10.3886	0.03	−10.0718	−10.4434
1.2	−9.6579(4)	−9.6456	0.13	−9.2971	−9.7983
1.4	−8.9483(4)	−8.9322	0.18	−8.6411	−9.1958
2.0	−7.4486(3)	−7.4356	0.17	−7.1974	−7.6473
4.0	−4.9105(3)	−4.9019	0.18	−4.7707	−4.9959
6.0	−3.8238(2)	−3.8209	0.08	−3.7374	−3.8793
8.0	−3.244(1)	−3.2430	0.03	−3.1820	−3.2846
10.0	−2.885(1)	−2.8865	−0.05	−2.8384	−2.9185

TABLE V. Total ground-state energies in units of t for the Hubbard model at half filling with open boundary conditions and total $S_z = 0$. The DMRG results used $m = 1600$ renormalized states. The Jastrow factors (correlators) employed were three-site lines for the one-dimensional lattices and four-site squares for the 4×5 lattice. See Sec. III C for details. Here NS, VNS, and %D have the same meaning as in Table I.

Lattice	VMC	NS	%D	VNS	%D	RHF	DMRG
				$U/t = 2$			
1×14	−11.240(1)	−11.241776	−0.02	−11.236(1)	0.04	−10.133544	−11.279897
1×18	−14.591(1)	−14.592961	−0.01	−14.587(1)	0.03	−13.219131	−14.653987
1×22	−17.947(2)	−17.946260	0.00	−17.938(1)	0.05	−16.307287	−18.029379
4×5	−19.917(1)	−19.920320	−0.02	−19.915(1)	0.01	−18.800678	−20.127521
				$U/t = 4$			
1×14	−7.556(1)	−7.631100	−0.99	−7.501(1)	0.73	−3.133544	−7.672349
1×18	−9.770(3)	−9.842409	−0.74	−9.695(2)	0.77	−4.219131	−9.965398
1×22	−11.968(4)	−12.042344	−0.62	−11.891(3)	0.64	−5.307287	−12.259082
4×5	−13.350(1)	−13.384297	−0.26	−13.316(2)	0.25	−8.800678	−14.404488

the restricted Hartree-Fock (RHF) Slater determinant. While better energies could be obtained with an unrestricted or generalized Slater determinant, the restricted Hartree-Fock determinant is sufficient for the comparison between the nonstochastic and variational Monte Carlo algorithms that is our primary concern.

Results for the ratios $U/t = 2$ and $U/t = 4$ are presented in Table V. For $U/t = 2$, the bare RHF Slater determinant produces energies in error by 6–10%, which are reduced to 1% or less after the inclusion of Jastrow factors, optimized either through the nonstochastic or variational Monte Carlo algorithms. Importantly, the nonstochastic energies lie within the error bars of the variational CPS energies for both the one- and two-dimensional lattices. For the case of $U/t = 4$, the RHF reference is qualitatively incorrect with relative errors as high as 60%. The inclusion of Jastrow factors reduces the error to 2% and 7% in one and two dimensions, respectively. Despite the poor quality of the wave function in this problem, the nonstochastic energy reproduces the variational CPS energy quite well, with the relative differences in Table V never exceeding 1%. This is much less than the intrinsic error due to the quality of the wave function.

As seen in Table VI, the cost of the nonstochastic method relative to variational Monte Carlo is less favorable than for the Heisenberg model. The reasons for this slowdown are twofold. First, the site dimension d is twice as large for the Hubbard model, and second, the evaluation of the RDM elements (while still polynomial scaling) is much higher for the Slater determinant than for the uniform reference. We note, however, that the cost scaling of the nonstochastic method is $O(N^2)$, while that of variational MC is $O(N^3)$ assuming that the number of samples needed grows linearly with system size. Thus for very large lattices, the nonstochastic approach may offer significant efficiencies.

Finally, for the 22-site chain with $U/t = 4$, we investigated the effect of varying the size of the Jastrow factor. As shown in Fig. 5, both the nonstochastic and variational energies improve when extending the Jastrow factor from one to four sites.

TABLE VI. Computational time taken for various optimizations. Times are reported in total CPU hours, defined as the wall clock time multiplied by the number of processing cores used.

Lattice	Correlators	Nonstochastic	Variational MC
8×8 Sq. Heis.	nearest neighbor	2.8×10^{-5}	1.4×10^{0}
8×8 Sq. Heis.	four-site squares	1.9×10^{-4}	1.4×10^{0}
8×8 Sq. Heis.	five-site crosses	2.7×10^{-2}	1.7×10^{0}
8×8 Sq. Heis.	nine-site squares	4.1×10^{2}	3.1×10^{0}
6×6 Tr. Heis.	nearest neighbor	9.6×10^{-4}	3.8×10^{0}
6×6 Tr. Heis.	three-site triangles	8.7×10^{-4}	2.6×10^{0}
6×6 Tr. Heis.	four-site rhombuses	1.3×10^{0}	6.1×10^{0}
6×6 Tr. Heis.	six-site triangles	4.0×10^{1}	4.1×10^{0}
6×6 Tr. Heis.	seven-site hexagons	2.1×10^{1}	2.4×10^{0}
1×14 Hub. $U/t = 2$	three-site lines	6.7×10^{-1}	2.3×10^{0}
1×18 Hub. $U/t = 2$	three-site lines	1.5×10^{0}	2.5×10^{0}
1×22 Hub. $U/t = 2$	three-site lines	3.7×10^{0}	2.8×10^{0}
4×5 Hub. $U/t = 2$	four-site squares	1.2×10^{3}	3.8×10^{0}
1×14 Hub. $U/t = 4$	three-site lines	1.1×10^{0}	2.0×10^{0}
1×18 Hub. $U/t = 4$	three-site lines	3.1×10^{0}	2.1×10^{0}
1×22 Hub. $U/t = 4$	three-site lines	4.9×10^{0}	2.4×10^{0}
4×5 Hub. $U/t = 4$	four-site squares	1.2×10^{3}	5.6×10^{0}

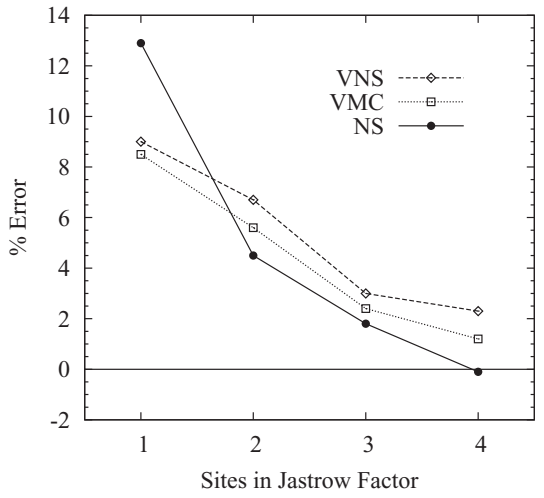

FIG. 5. Energy errors (relative to $m = 1600$ DMRG) in the one-dimensional Hubbard model with $U/t = 4$ for line-shaped Jastrow factors of different sizes. The lattice is a 22-site chain with open boundary conditions. The reference function is the restricted Hartree-Fock determinant. NS and VNS have the same meaning as in Table I.

We observe in all cases that the nonstochastic and variational energies differ by an amount significantly less than the intrinsic energy error of the wave function, except for the four-site Jastrows, where the nonstochastic energy lies slightly (-0.1%) below the true energy, while the variational energy is above (1.2%) and thus the difference between the nonstochastic and variational energies is almost exactly the same as the intrinsic variational energy error.

IV. CONCLUSIONS

We have shown that efficient nonstochastic algorithms exist both to evaluate the energy and expectation values of Jastrow-Slater and correlator product state wave functions, as well as to optimize the wave-function parameters. We have tested our methods in three models: the spin-$\frac{1}{2}$ antiferromagnetic Heisenberg model, the spinless Hubbard model, and the full Hubbard model. While unlike the variational Monte Carlo energy, the nonstochastic energy is not a strict upper bound, the difference between the two energies is comparable to and often significantly less than the intrinsic error associated with the quality of the wave function. In practice we find that the nonstochastic algorithms are faster than the variational Monte Carlo algorithms for small correlator (or Jastrow) sizes, but become more expensive for larger correlators.

The nonstochastic algorithms we have described rely on the mathematical form of the Jastrow-Slater and correlator product state wave functions as a product of local, commuting, invertible operators acting on a simple reference wave function. Any wave function with this mathematical form may be studied with analogous efficient nonstochastic techniques. This possibility can guide the construction of efficient new classes of wave functions in the future.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation through the NSF Center for Molecular Interfacing as well as Grants No. CHE-0645380 and No. CHE-1004603.

[1] R. Jastrow, Phys. Rev. **98**, 1479 (1955).
[2] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[3] A. C. Hurley, J. E. Lennard-Jones, and J. A. Pople, Proc. R. Soc. London, Ser. A **220**, 446 (1953).
[4] P. W. Anderson, Science **235**, 1196 (1987).
[5] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).
[6] J. P. Bouchand and C. Lhuillier, Z. Phys. B **75**, 283 (1989).
[7] H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **80**, 245116 (2009).
[8] F. Mezzacapo, N. Schuch, M. Boninsegni, and J. I. Cirac, New J. Phys. **11**, 083026 (2009).
[9] F. Mezzacapo and J. I. Cirac, New J. Phys. **12**, 103039 (2010).
[10] K. H. Marti, B. Bauer, M. Reiher, M. Troyer, and F. Verstraete, New J. Phys. **12**, 103008 (2010).
[11] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[12] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[13] F. Verstraete and J. I. Cirac, e-print arXiv:cond-mat/0407066 (to be published).
[14] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. A **75**, 033605 (2007).
[15] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **101**, 250602 (2008).
[16] Z.-C. Gu, M. Levin, and X.-G. Wen, Phys. Rev. B **78**, 205116 (2008).
[17] H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **101**, 090603 (2008).
[18] Z. Y. Xie, H. C. Jiang, Q. N. Chen, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **103**, 160601 (2009).
[19] P. Chen, C.-Y. Lai, and M.-F. Yang, J. Stat. Mech. (2009) P10001.
[20] R. Orús and G. Vidal, Phys. Rev. B **80**, 094403 (2009).
[21] P. Corboz, R. Orús, B. Bauer, and G. Vidal, Phys. Rev. B **81**, 165104 (2010).
[22] C. V. Kraus, N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. A **81**, 052338 (2010).
[23] I. Pižorn and F. Verstraete, Phys. Rev. B **81**, 245110 (2010).
[24] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).
[25] M. Rizzi, S. Montangero, and G. Vidal, Phys. Rev. A **77**, 052328 (2008).
[26] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[27] S. F. Boys and N. C. Handy, Proc. R. Soc. A **311**, 309 (1969).
[28] J. W. Clark and P. Westhaus, Phys. Rev. **141**, 833 (1966).
[29] W. L. McMillan, Phys. Rev. **138**, A442 (1965).
[30] M. P. Nightingale and C. J. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry*, NATO Advanced Study Institute Series C (Kluwer Academic, Boston, 1999), p. 525.
[31] A. W. Sandvik, Phys. Rev. B **56**, 11678 (1997).
[32] A. Albuquerque *et al.*, J. Magn. Magn. Mater. **310**, 1187 (2007).