



Low entanglement wavefunctions

Garnet Kin-Lic Chan*

We review a class of efficient wavefunction approximations that are based around the limit of low entanglement. These wavefunctions, which go by such names as matrix product states and tensor network states, occupy a different region of Hilbert space from wavefunctions built around the Hartree–Fock limit. The best known class of low entanglement wavefunctions, the matrix product states, forms the variational space of the density matrix renormalization group algorithm. Because of their different structure to many other quantum chemistry wavefunctions, low entanglement approximations hold promise for problems conventionally considered hard in quantum chemistry, and in particular problems which have a multireference or strong correlation nature. In this review, we describe low entanglement wavefunctions at an introductory level, focusing on the main theoretical ideas. Topics covered include the theory of efficient wavefunction approximations, entanglement, matrix product states, and tensor network states including the tree tensor network, projected entangled pair states, and the multiscale entanglement renormalization ansatz. © 2012 John Wiley & Sons, Ltd.

How to cite this article:

WIREs Comput Mol Sci 2012, 2: 907–920 doi: 10.1002/wcms.1095

INTRODUCTION

In wavefunction-based quantum chemistry, we wish to determine the stationary states of the electronic Schrödinger equation. This is a difficult task because the dimension of the Hilbert space grows exponentially with the number of electrons. Thus, any practical calculation must use some simplifying approximations, or ansatz, for the state we wish to study.

Figure 1 illustrates schematically the full Hilbert space of wavefunctions. Complete expansions, for example, via full configuration interaction or full resonance valence bond theory, span the entire Hilbert space at exponential cost. Practical wavefunction approximations are of polynomial cost, and are necessarily biased toward a given portion of the Hilbert space. In opposite corners of the rectangle in Figure 1, we show two families of approximations salient to our discussion. One is biased toward the Hartree–Fock mean-field or independent particle limit. Such wavefunctions are well developed in quantum chemistry and can be systematically made more flexible

by increasing the excitation level in the wavefunction. The other family of wavefunctions, which are the main subject of this review, are biased toward the independent local subsystem, or zero entanglement, limit. These wavefunctions are systematically made more flexible by increasing the amount of encoded entanglement.

Low entanglement wavefunctions first became widely applied with the advent of the density matrix renormalization group (DMRG).^{1–8} Within a few years, it was understood that the DMRG algorithm is an energy minimization algorithm within a class of low entanglement wavefunctions known as matrix product states (MPS).^{5–7,9–15} Decimation in the renormalization group (RG) algorithm is equivalent to encoding limited entanglement in the state, linking the RG picture with quantum information perspectives.^{6,13,15–17} In the last decade, MPSs also have been used in quantum chemistry via the DMRG algorithm.^{8,17–50} Quantum chemical DMRG has significantly widened the scope of *ab initio* calculations with respect to strongly correlated electrons. In particular, it has allowed multireference active space calculations with more than 30, and in some cases up to 100, active orbitals (see Refs 8,22,23,25–30,42,43,47–49 and Ref 7 for a summary of recent calculations).

*Correspondence to: gkc1000@gmail.com

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA

DOI: 10.1002/wcms.1095

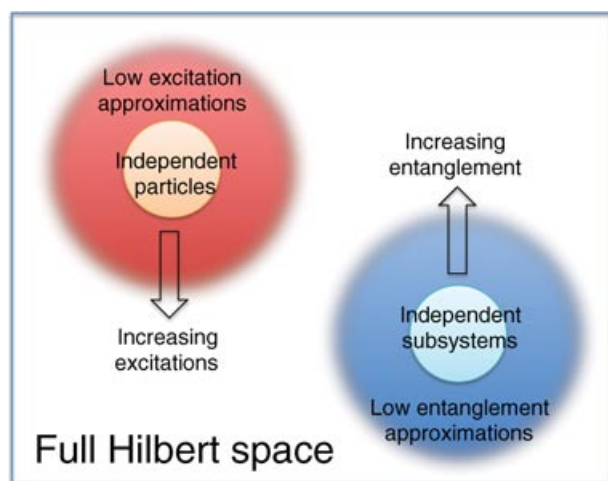


FIGURE 1 | The Hilbert space of many-electron wavefunctions. The independent particle and independent local subsystem limits lie in different regions of the Hilbert space. By increasing the excitation level or entanglement, respectively, in approximations built around these limits, they can be made to span the full Hilbert space.



FIGURE 2 | Area laws express locality in physical systems. In a one-dimensional system, the boundary area is independent of the overall system size, thus entanglement entropy in a physical ground state is expected to be independent of system size, unless at a quantum critical point. In two-dimensional system, the boundary area can scale as the system width, and the entanglement entropy scales similarly, except for critical and topological corrections.

More recently, generalizations of MPSs, which remove some fundamental limitations, for example, with respect to system dimensionality, have been introduced.^{15,51–64} Such tensor network wavefunctions are still in their infancy, and efficient computational techniques to work with them remain under development. Indeed, they have yet to be applied to quantum chemistry (also, see Refs 50, 60, 65 for early work in this direction). Nonetheless, they hold considerable promise due to their increased flexibility over MPSs. We describe them also in this review.

The structure of the discussion is as follows. In section *Wavefunctions and the Orthogonality Catastrophe*, we first consider why, despite the exponential size of the Hilbert space, it is possible to faithfully approximate a quantum state. In sections *Wave-*

functions around the Hartree–Fock Limit and Low Entanglement Wavefunctions, we describe the two families of wavefunctions presented above. We begin with a brief discussion of wavefunctions built around the independent particle limit. Then we turn to our main focus—low entanglement wavefunctions. We introduce the mathematical definition of entanglement, and proceed to describe in detail the theory of MPS. We next describe more general tensor network wavefunctions, including tree tensor networks (TTNs), projected entangled pair states, and the multiscale entanglement renormalization ansatz (MERA). We finish with our conclusions in section *Conclusions*.

WAVEFUNCTIONS AND THE ORTHOGONALITY CATASTROPHE

How do we construct a many-electron wavefunction to represent a quantum state? At first glance, there appears to be a fundamental problem, as the orthogonality catastrophe shows that in a large system, any approximate wavefunction is orthogonal to the state of interest. This might even suggest that a wavefunction is not a valid way to describe large quantum systems (see, e.g., the discussion in Ref 66). To illustrate the catastrophe, consider a set of N noninteracting electrons. For simplicity, we neglect antisymmetry. Then the ground-state wavefunction is a product of N orbitals

$$\Phi = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) \dots \phi_N(\mathbf{r}_N). \quad (1)$$

We imagine representing each orbital ϕ by an approximate ϕ' , with a small error,

$$\langle \phi | \phi' \rangle = 1 - \epsilon. \quad (2)$$

The overlap of the approximate and exact wavefunctions is a product of the overlaps of the orbitals

$$\begin{aligned} \langle \Phi | \Phi' \rangle &= \langle \phi_1 | \phi'_1 \rangle \langle \phi_2 | \phi'_2 \rangle \dots \langle \phi_N | \phi'_N \rangle \\ &= (1 - \epsilon)^N \\ &\sim \exp(-\epsilon N). \end{aligned} \quad (3)$$

We see that the overlap of the true and approximate determinants decreases *exponentially* with the number of particles. This is not essentially changed if include antisymmetry, or add in the interactions between the electrons. This suggests that any approximate wavefunction is a very poor representation of the quantum state in a large systems.

The practical success of quantum chemistry, however, indicates that approximate many-electron wavefunctions do contain useful information. Indeed,

even with a poor overlap, expectation values of operators can be very good. In the above case, for a single-particle operator $\hat{Q} = \sum_i \hat{q}_i$, with expectation value $Q = \langle \Phi | \hat{Q} | \Phi \rangle$,

$$\langle \Phi' | \hat{Q} | \Phi' \rangle = \sum_i \langle \phi'_i | \hat{q}_i | \phi'_i \rangle = Q + \mathcal{O}(\epsilon). \quad (4)$$

We see that the error in the expectation value *per particle* is $\mathcal{O}(\epsilon)$, independent of the number of particles. Thus, wavefunctions with poor overlap can yield meaningful expectation values even in the limit of large particle number, that is, the thermodynamic limit.

Further examination of the original argument also shows that the catastrophe is not actually as severe as it first appears. We assumed that we compute each orbital to within an error of ϵ , regardless of how large the system is. However, we could try to represent each orbital *more* accurately as the number of electrons increases. Indeed, to defeat the exponential loss of overlap we need the error in each orbital to decrease such as

$$\langle \phi | \phi' \rangle \sim 1 - \frac{\epsilon}{N}. \quad (5)$$

In this case, the approximate wavefunction can retain unit overlap with the true state. For the scheme to be practical, we must be able to determine an orbital to a given error of ϵ in a time polynomial in ϵ^{-1} . Current methods to solve the single-particle molecular Schrödinger equation (e.g., by Gaussian basis set expansion) satisfy this criterion.^{67–70} Thus, even using the total wavefunction overlap as a measure of accuracy, it is possible in some cases to faithfully represent the state of a large system with a wavefunction.

Now that we have established that approximate wavefunctions can in principle provide meaningful descriptions of N electron systems, we still have the problem of determining appropriate approximations for physical problems. We first describe wavefunctions that are built around the Hartree–Fock limit, then move to our main focus of low entanglement wavefunctions.

WAVEFUNCTIONS AROUND THE HARTREE–FOCK LIMIT

Wavefunctions built around the Hartree–Fock mean-field limit are widely used in quantum chemistry. The Hartree–Fock wavefunction is the determinant which minimizes the energy $E = \langle \Phi | H | \Phi \rangle / \langle \Phi | \Phi \rangle$ ^{71,72}. Starting from a determinant allows a separation of orbitals into an occupied space (labeled by i, j, \dots) and a virtual space (labeled by a, b, \dots), thus defining excita-

tions relative to the determinant. In the standard excitation hierarchy, one expands a general wavefunction as a linear combination of excitations around the Hartree–Fock reference,

$$|\Psi\rangle = |\Phi_0\rangle + \sum_{ia} c_i^a |\Phi_i^a\rangle + \sum_{i>j, a>b} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \dots \quad (6)$$

By mixing in successively higher excitations, one covers the full quantum Hilbert space. Compact wavefunction approximations are constructed by restricting the generality of the excitation expansion. In truncated configuration interaction and perturbation theories,^{71,72} we restrict the maximum excitation level in the expansion (6), whereas truncated coupled cluster theory rewrites the expansion first in cluster form^{71–74}

$$|\Psi\rangle = \exp(\hat{T})|\Phi_0\rangle, \quad (7)$$

and truncates the excitation level in the cluster operator \hat{T} ,

$$\hat{T} = \sum_{ia} t_i^a a_a^\dagger a_i + \sum_{i>j, a>b} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_i a_j + \dots \quad (8)$$

There are other ways to construct wavefunctions around the Hartree–Fock limit aside from an excitation expansion. Another class of compact wavefunctions is obtained by applying diagonal operators, known as Jastrow factors (sometimes known as correlation factors), to the determinant.^{75–77} Jastrow factors take the general form

$$\hat{f} = \sum_i j_i \hat{n}_i + \sum_{i>j} j_{ij} \hat{n}_i \hat{n}_j + \dots, \quad (9)$$

where \hat{n}_i is a number operator $a_i^\dagger a_i$ in some basis, not necessarily the molecular orbital basis of the determinant. The Jastrow determinant wavefunction is then

$$|\Psi\rangle = \hat{f}|\Phi\rangle. \quad (10)$$

As successively more terms are included in the Jastrow factor, the Jastrow determinant wavefunction can also cover the full quantum Hilbert space.

In both the excitation and Jastrow expansions, the choice of starting determinant is not unique. For example, the Hartree–Fock energy minimization has many local minima, each of which may form a suitable reference. Some solutions of the energy minimization do not preserve the symmetry of the original Hamiltonian. Symmetry-broken solutions can often be associated with different kinds of electronic phases and can thus be chosen to reflect the

nature of the final state we wish to capture. Breaking spin symmetry yields unrestricted and general Hartree–Fock determinants, while breaking number symmetry leads to the Bardeen–Cooper–Schrieffer (BCS) wavefunction.⁷⁸ (We refer to the BCS wavefunction also as a determinant for simplicity, although it is a determinant of quasiparticles, not electrons.) In molecular systems, the BCS wavefunction is typically projected to restore the particle number symmetry, and the projected states are known as the antisymmetrized geminal power^{79,80} or (if they break spin symmetry also) Pfaffian wavefunctions.⁸¹

Despite the flexibility in the choice of starting reference, there are some situations when no good starting determinant can be found. Determinant wavefunctions are eigenfunctions of one-particle effective Hamiltonians

$$\hat{h} = \sum_{ij} t_{ij} a_i^\dagger a_j + \Delta_{ij} (a_i a_j + a_j^\dagger a_i^\dagger), \quad (11)$$

where the first term is a number conserving one particle term, and the second is the pairing field associated with breaking particle number symmetry in the BCS wavefunction. The quality of the starting reference is governed by the magnitude of the perturbation $\hat{V} = \hat{H} - \hat{h}$, where \hat{H} is the full electronic Hamiltonian of the problem. Strongly correlated electronic structure, commonly found in transition metal chemistry and excited electronic states, is characterized by a large magnitude of \hat{V} relative to the energy spacing of \hat{h} .

For small numbers of strongly correlated electrons, one can augment the starting reference determinant by a linear combination of starting determinants. This gives rise to multireference wavefunction approximations. Perturbative, configuration interaction, and coupled cluster theories from a multireference starting point can be formulated,^{72,82–86} as can multireference Jastrow wavefunctions.⁸⁷ However, multireference approximations are compact only so long as the number of starting references remains polynomial in the system size. This is not the case for large numbers of strongly correlated electrons. We therefore turn our attention to a different class of wavefunction approximations that seems well suited to strong correlation problems: the low entanglement wavefunctions.

LOW ENTANGLEMENT WAVEFUNCTIONS

Valence Bond Theory

The motivations for constructing wavefunctions by encoding entanglement between local objects are il-

lustrated by the stretched two-electron bond in the hydrogen molecule. Using an excitation expansion such as Eq. (6), the wavefunction at large separation may be written as

$$|\Psi\rangle = 2^{-1/2} (\mathcal{A}[\sigma_g(1)\alpha\sigma_g(2)\beta] - \mathcal{A}[\sigma_u(1)\alpha\sigma_u(2)\beta]), \quad (12)$$

where σ_g and σ_u are the bonding and antibonding molecular orbitals. The stretched bond wavefunction requires a multireference approximation as there is no dominant determinant. At large separation, however, it seems more natural to build the wavefunction from the atomic states. Denoting the atomic orbitals as $1s_a$ and $1s_b$, we write

$$|\Psi\rangle = \frac{1}{2} [1s_a(1)1s_b(2) + 1s_a(2)1s_b(1)] \times [\alpha_a(1)\beta_b(2) - \beta_b(1)\alpha_a(2)]. \quad (13)$$

The coupling between the atomic spins reflects strong correlations. For example, if we measure spin α on atom *A*, the corresponding spin measurement on atom *B* will yield spin β . This type of strong correlation is known as entanglement, and is associated with the chemical bond. (Note that as discussed above, strong correlation is commonly defined with respect to a perturbative partitioning of the Hamiltonian; strong correlation and entanglement are thus not synonyms. One of the attractions of entanglement as a descriptor, rather than strong correlation, is the ability to define precise mathematical measures of entanglement, as illustrated in the next section). However, in a molecule, not all atoms are bonded to all other atoms, that is, there are only spin couplings of the above form along specific chemical bonds. This suggests that compact wavefunction approximations may be constructed by encoding only the entanglement between local states to reflect the bonding network of the molecule.

Equation (13) is also the starting point for valence bond (VB) theory. In VB theory, the wavefunction is expanded in a valence bond basis, corresponding to the classical resonance structures of chemistry. For example, for three hydrogen atoms, there are three neutral doublet resonance structures. However, transforming to the valence bond basis does not by itself yield a compact wavefunction, as the basis is exponentially large. For a compact parametrization, we have to limit the number of resonance structures in the expansion. If only a single-resonance structure is used, we obtain a method known as generalized valence bond theory.^{88,89} Moving beyond a single-resonance structure theory in a scalable, fully general, and compact way, however, proves quite difficult.^{90–92} Low entanglement wavefunctions change the focus from

the bonds to the local states and provide a more practical framework for compact approximations. We now turn to a mathematical discussion of entanglement.

Entanglement

Bipartite entanglement,⁹³ or the entanglement between two systems, is defined as follows: let systems 1, 2 be spanned by sets of (orthonormal) states $\{|n_1\rangle\}$, $\{|n_2\rangle\}$, respectively. Any state in the combined system 1, 2 is expanded in the product space $\{|n_1 n_2\rangle\}$ as

$$|\Psi\rangle = \sum_{n_1 n_2} \psi_{n_1 n_2} |n_1 n_2\rangle. \quad (14)$$

(Here we work in a second quantized formulation, where fermion antisymmetry is expressed in the commutation relations between operators, rather than by antisymmetrizing products of states). The bipartite entanglement entropy is defined from a singular value decomposition of the coefficients $\psi_{n_1 n_2}$. Writing this as a matrix ψ , then

$$\psi = U \sigma V^T, \quad (15)$$

where σ is the diagonal matrix of singular values. The entanglement entropy is

$$S = - \sum_i \sigma_i^2 \log_2 \sigma_i^2. \quad (16)$$

The entanglement entropy quantifies the correlations between the systems. If the wavefunction is a product wavefunction, there is only a single nonzero singular value, and the entanglement entropy is zero. In the maximally entangled state, all singular values are the same. The maximally entangled state is thus written as

$$|\Psi\rangle = \frac{1}{\sqrt{M}} \sum_i |r_i l_i\rangle, \quad (17)$$

where $\{|r_i\rangle\}$ and $\{|l_i\rangle\}$ denote the basis in which the entangling is performed, and the entanglement entropy is $\log_2 M$. The hydrogen molecule at infinite separation is in a maximally entangled state. Considering the spin part of the wavefunction

$$|\Psi_{\text{spin}}\rangle = \frac{1}{2} [\alpha(1)\beta(2) - \beta(1)\alpha(2)], \quad (18)$$

we find that the entanglement entropy is $\log_2 2 = 1$, that is, there is one pair of entangled spins.

Any wavefunction can be written in a diagonal form similar to the maximally entangled state. From Eq. (15), the matrices U and V define a transformation of the original bases. We define $\{|l_i\rangle\}$ and $\{|r_i\rangle\}$

through

$$|r_i\rangle = \sum_{n_1} U_{n_1 i} |n_1\rangle, \quad (19)$$

$$|l_i\rangle = \sum_{n_2} V_{n_2 i} |n_2\rangle. \quad (20)$$

In terms of the transformed basis, the wavefunction then assumes a diagonal form

$$|\Psi\rangle = \sum_i \sigma_i |r_i l_i\rangle. \quad (21)$$

Matrix Product States

When constructing a wavefunction from the states of two local systems, the wavefunction can always be written in the diagonal form $|\Psi\rangle = \sum_i \sigma_i |r_i l_i\rangle$. Relating this to the original basis $|n_1 n_2\rangle$, this gives

$$|\Psi\rangle = \sum_{n_1 n_2, i} A_{i_1}^{n_1} A_{i_2}^{n_2} |n_1 n_2\rangle, \quad (22)$$

where we have for notational purposes absorbed the singular values in Eq. (15) into the definitions of the A matrices. This form is very suggestive, as we can think of the index i (which we call an auxiliary index) as entangling the original states $\{|n_1\rangle\}$, $\{|n_2\rangle\}$, in the two systems.

We can now try to generalize the above to more than two local systems. Consider three systems 1, 2, and 3. We imagine entangling system 1 with system 2, and then system 2 with system 3. The natural generalization of Eq. (22) is to use two entangling auxiliary indices i_1, i_2 , and to write the wavefunction as

$$\begin{aligned} |\Psi\rangle &= \sum_{n_1 n_2 n_3} \sum_{i_1 i_2} A_{i_1}^{n_1} A_{i_1 i_2}^{n_2} A_{i_2}^{n_3} |n_1 n_2 n_3\rangle \\ &= \sum_{n_1 n_2 n_3} A^{n_1} A^{n_2} A^{n_3} |n_1 n_2 n_3\rangle, \end{aligned} \quad (23)$$

where in the last line we have switched to a matrix notation, and A^{n_1} and A^{n_3} are thin matrices (a row vector and a column vector, respectively). The multiplication of the matrices A^{n_1} , A^{n_2} , A^{n_3} , yields the wavefunction coefficient $\Psi^{n_1 n_2 n_3}$ associated with the configuration $|n_1 n_2 n_3\rangle$. Since the wavefunction coefficient is obtained as a matrix product, this wavefunction approximation is known as a *matrix product state* (MPS).^{5–7,9–15} The accuracy and complexity of the MPS is controlled by the dimension of the auxiliary indices i_1, i_2 of the matrices. If we fix this at some dimension M , we say that the MPS is of dimension M .

We can generalize the MPS to k systems, where we entangle them in a sequential manner,

$$|\Psi\rangle = \sum_{\mathbf{n}} \mathbf{A}^{n_1} \mathbf{A}^{n_2} \dots \mathbf{A}^{n_k} |\mathbf{n}\rangle, \quad (24)$$

where \mathbf{n} denotes the configuration $|n_1 n_2 \dots n_k\rangle$. Note that the MPS construction is general and does not depend on the definition of the local system or the nature of the constituent states $|\mathbf{n}\rangle$. For example, we may take each system to be an atom (in which case the local states span the Fock space of each atom, where by Fock space we mean the full Hilbert space across all particle number sectors). Or, we could take each system to be as small as a single orbital, spanned by the four states, $|\downarrow\rangle, |\alpha\rangle, |\beta\rangle, |\alpha\beta\rangle$. To emphasize this generality, we now introduce the term *site* as a synonym for local system. This is the terminology primarily used in the MPS and DMRG literature. The wavefunction approximation in Eq. (24) is thus an MPS wavefunction for k sites.

One way to analyze the accuracy of the MPS approximation is to consider how much (bipartite) entanglement entropy can be described by such a state, if we imagine dividing our k sites into two sets. If the dimension of the auxiliary indices is M , then the maximum entanglement expressed by the MPS for any cut between the sites is $\log_2 M$, irrespective of the number of sites in the problem. We might, however, more naturally expect the true entropy to capture in a physical problem to scale with the size of the problem.

The link between entanglement entropy, system size, and dimensionality is given rigorously by so-called area laws.^{94–96} Whereas thermodynamic entropy is an extensive quantity (i.e., it is simply proportional to volume), entanglement entropy is believed to scale only with the size (i.e., area) of the *boundary* between the two systems used to define the entropy. At an intuitive level, this expresses locality. Given local interactions in the Hamiltonian, then in the ground state, only quantum degrees of freedom along the boundary can be entangled with each other, and thus the entropy scales with boundary size. The boundary size depends on the overall system size in a dimensionally dependent way. In one dimension, any division of the system yields a boundary size independent of system size, thus entanglement entropy in a one-dimensional ground state is independent of system size. However, in a two- or higher-dimensional system, the boundary of a division of the system can scale like the width L of the system, and thus the entanglement entropy scales as the system width (see Figure 2).

As we argued above, for an MPS of dimension M , the maximum entanglement entropy $\log_2 M$

is independent of system size, and we see that this is clearly appropriate to a one-dimensional system ground state. If we apply an MPS to describe a two or higher-dimensional system, however, to capture the growing entanglement entropy, the matrix dimensions must grow exponentially with the boundary area, that is, $M \sim e^L$.

This may suggest that using an MPS to simulate two- and three-dimensional systems is a bad idea. It is indeed possible to devise states with more general entanglement structures than the MPS, which naturally capture the area law in higher dimensions and such states are discussed in section *General Tensor Network States*. However, crucially, the MPS captures the correct structure of entanglement for one of the spatial dimensions, and thus the exponential cost is associated only with the remaining dimensions, that is, an area. This is a clear improvement over full configuration interaction, where the cost scales exponentially with the system *volume*. In practice, therefore, MPS offers significant computational advantages when simulating systems even with two- or three-dimensional connectivities.⁹⁷

We now discuss a few miscellaneous topics related to MPS: algorithms, their interpretation as projected wavefunctions, and finally their graphical representation.

MPS Algorithms

MPSs are compatible with many efficient computational algorithms. For example, overlaps and expectation values can be obtained efficiently.^{5,6,15} We illustrate this with the computation of the norm. The norm of a matrix product state is

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \sum_{\mathbf{n}} (\mathbf{A}^{n_1 \dagger} \mathbf{A}^{n_2 \dagger} \dots \mathbf{A}^{n_k \dagger}) \otimes (\mathbf{A}^{n_1} \mathbf{A}^{n_2} \dots \mathbf{A}^{n_k}) \\ &= \mathbf{E}^1 \mathbf{E}^2 \dots \mathbf{E}^k, \end{aligned} \quad (25)$$

where the \mathbf{E} matrices are of dimension $M^2 \times M^2$,

$$\mathbf{E}^i = \sum_{n_i} \mathbf{A}^{n_i \dagger} \otimes \mathbf{A}^{n_i}. \quad (26)$$

The evaluation of the norm is itself a matrix product. From Eq. (25), it would appear to take $\mathcal{O}(M^4 k)$ time; however, using the fact that the \mathbf{E} matrices have an outer product structure, it requires only $\mathcal{O}(M^3 k)$ time. By a similar construction, one finds that expectation values of local operators can also be evaluated in $\mathcal{O}(M^3 k)$ time.⁶

The expectation value structure in Eq. (25) is reminiscent of a transfer matrix expression in statistical mechanics, where the \mathbf{E} matrices are the transfer operators. This structure means that correlation

functions such as $\langle \hat{Q}_i \hat{Q}_j \rangle$ (where \hat{Q}_i, \hat{Q}_j act on sites i and j , respectively) decay in MPS exponentially like $\exp(-\lambda|i-j|)$. λ is related to the smallest eigenvalue of the transfer operator.^{6,12} As M (and thus the dimension of E) is increased, it is possible to reproduce smaller and smaller λ , thus capturing longer range correlations.

The (matrix) product structure of MPS means that they are a generalization of a mean-field theory.¹⁴ Unlike Hartree–Fock theory, which is a mean-field theory for independent electrons, MPSs represent a mean-field theory for each of the local systems or sites. By minimizing the variational MPS energy with respect to the individual matrix components A^n , we obtain a Schrödinger equation for each component,

$$\mathbf{H} \cdot \mathbf{a} = E\mathbf{a}, \quad (27)$$

where \mathbf{H} is an effective Hamiltonian and \mathbf{a} denotes A^n flattened into a vector form. This is analogous to the orbital equation in Hartree–Fock theory. As described in *Introduction*, MPS arose originally in conjunction with the DMRG. From the viewpoint of the current discussion, we see the DMRG as an algorithm which works within the space of MPS to variationally minimize the energy with respect to a single component A^n at a time. The equations solved at each step of the DMRG are the effective Schrödinger equations (27), and the sweep algorithm of the DMRG is then analogous to the traditional self-consistent field algorithm of Hartree, where each orbital is determined sequentially, and the effective Fock operator is updated one orbital at a time until self-consistency is reached.^{6,10,14}

Projected Wavefunctions

There is another way to view MPS that is often useful and which connects to some traditional quantum chemistry ideas. In this language, the MPS arises as a projection from a reference function. The particles appearing in the reference are not the actual electrons, but are fictitious or auxiliary particles.^{13,15} In other contexts, the idea of introducing fictitious particles to describe a state is sometimes known as using a shadow wavefunction.⁹⁸

Consider, the MPS for two sites

$$|\Psi\rangle = \sum_{n_1 n_2, i} A_i^{n_1} A_i^{n_2} |n_1 n_2\rangle. \quad (28)$$

Formally, we can introduce auxiliary states $|r_i l_i\rangle$. We then find that the MPS can be written as a set of projections

$$|\Psi\rangle = \hat{A}_1 \hat{A}_2 \sum_i |r_i l_i\rangle, \quad (29)$$

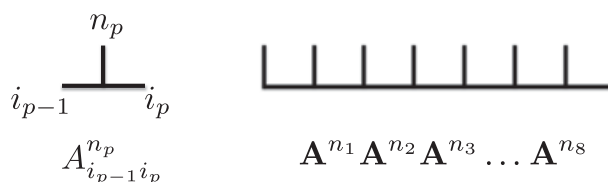


FIGURE 3 | Graphical representation of matrix product states. (Left) For every state of a site $|n_p\rangle$, there is a matrix A^{n_p} . The indices of the matrix, i_{p-1} and i_p , are referred to as auxiliary indices and are of dimension M . Note that the leftmost and rightmost matrices are, in practice, row and column vectors. In this way, the product over all the matrices yields a scalar, the wavefunction coefficient $\Psi_{n_1 n_2 \dots n_k}$. (Right) To obtain the matrix product state approximation to the wavefunction coefficient $\Psi_{n_1 n_2 \dots n_k}$, we contract over all the auxiliary indices in the product $A^{n_1} A^{n_2} \dots A^{n_k}$. Graphically, the contraction is represented by joining together the auxiliary index lines.

where we have defined projection operators \hat{A}_1, \hat{A}_2 as

$$\hat{A}_1 = \sum_{i n_1} A_i^{n_1} |n_1\rangle \langle r_i|, \quad (30)$$

$$\hat{A}_2 = \sum_{i n_2} A_i^{n_2} |n_2\rangle \langle l_i|. \quad (31)$$


The state that is being projected from is a maximally entangled state of auxiliary particles $\sum_i |r_i l_i\rangle$. In general, the reference for an MPS for k sites can be viewed as a simple product of entangled states (one set for each auxiliary index) operated on by a set of commuting projection operators

$$|\Psi\rangle = \hat{A}_1 \hat{A}_2 \dots \hat{A}_k \left(\sum_{i_1} |r_{i_1} l_{i_1}\rangle \right) \times \left(\sum_{i_2} |r_{i_2} l_{i_2}\rangle \right) \dots \left(\sum_{i_{k-1}} |r_{i_{k-1}} l_{i_{k-1}}\rangle \right). \quad (32)$$

The view of the MPS as a set of commuting operators acting on a simple product reference function formally connects the MPS construction to other kinds of wavefunctions studied in quantum chemistry such as the coupled cluster wavefunction (which is a product of a set of commuting excitation operators on a single-determinant reference) although this analogy has yet to be much explored.

Graphical Representation

The proliferation of indices in the MPS means that it is often more convenient to use a graphical rather than algebraic representation.⁶ The graphical representation of an MPS is shown in Figure 3. A matrix A^n is represented by an inverted T-shaped object with lines going left and right (the two auxiliary indices),

$$E_{i_{p-1} i_p i'_{p-1} i'_p}^p = \sum_{n_p} A_{i_{p-1} i_p}^{n_p} A_{i'_{p-1} i'_p}^{n_p}$$


$$\langle \Psi | \Psi \rangle = E^1 E^2 E^3 \dots E^8$$

FIGURE 4 | (Bottom) A norm can be represented graphically by fusing the vertical indices of the matrix product state (MPS) bra (pointing downward) with the vertical indices of the MPS ket (pointing upward), denoting summation over all the state indices, n_1, n_2, \dots, n_k . (Top) Performing a state summation at each site leads to an intermediate matrix E of dimension M^2 , and the norm evaluation is a product of the E matrices. (Note that the leftmost and rightmost matrices are, in practice, row and column vectors.)

and a vertical line (the site index n). Lines that are connected represent indices that are summed over.

The graphical representation is useful not only for representing the states, but also operators and expectation values. The computation of an overlap is shown in Figure 4. The graphical representation clearly shows the nature of the index contractions. This becomes even more essential when considering states with a more general entanglement structure than the MPSs, to which we now turn.

General Tensor Network States

The limitation of MPS is that they can only efficiently encode a sequential structure of entanglement. For arbitrary systems, more general encodings can be explored. This is a rapidly expanding area of investigation, and many questions remain to be answered. Here we discuss three promising kinds of tensor networks: TTN,^{57,58,60} which extend MPSs and maintain a similar computational efficiency, but which still restrict the structure of the entanglement; projected entangled pair states (PEPS),^{15,52,54,55} which are a very general way to encode a network of entanglement; and the MERA,^{61–64} which combines some of the practical strengths of the TTN while achieving some of the formal advantages of PEPS.

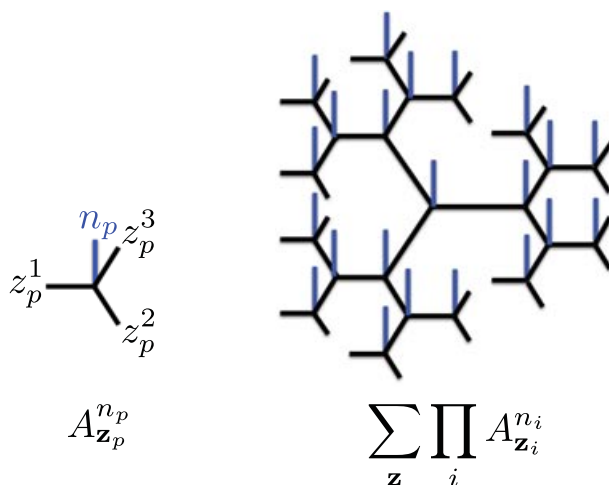


FIGURE 5 | (Left) For every local state $|n_p\rangle$, a tree tensor network (TTN) associates a tensor with z auxiliary indices. Here shown is a TTN where each tensor has three auxiliary indices. (Right) The wavefunction coefficient $\Psi^{n_1 n_2 \dots n_k}$ is obtained by contracting the auxiliary indices according to the tree connectivity

Tree Tensor Networks

In an MPS, a local system is represented by a matrix A^n with two auxiliary indices, which allows entangling to the left and right neighbors. In the generalization to a tree structure, a local system is represented by a tensor with z auxiliary indices, which allows couplings to z neighbors. The TTN state can be written as

$$|\Psi\rangle = \sum_{\mathbf{n}} \sum_{\mathbf{z}} \prod_i A_{\mathbf{z}_i}^{n_i} |\mathbf{n}\rangle, \quad (33)$$

where \mathbf{z}_i denotes the set of z indices on each tensor, and $\sum_{\mathbf{z}}$ denotes the contraction with the indices of neighboring tensors according to the connectivity of the tree, shown in Figure 5. The accuracy of the TTN is controlled by the dimensions of the auxiliary indices \mathbf{z} , which we can denote by M .

One advantage of the TTN over the MPS is that the maximum distance between two sites in a tree is logarithmic in the total number of sites, as opposed to linear in the case of the MPS. This allows the TTN to describe states with correlation functions, which decay algebraically with respect to site separation, as opposed to exponentially in the case of MPS, as described in section *MPS Algorithms*.

A second advantage of the TTN is that, just as in the MPS, expectation values can be efficiently obtained. By contracting the tensors from out to in, we see that the cost for calculating a norm is $O(M)^{z+1}k$. While the scaling of a $z > 2$ TTN computation is higher than that for a corresponding MPS ($z = 2$), the ability of TTN to better represent long-range

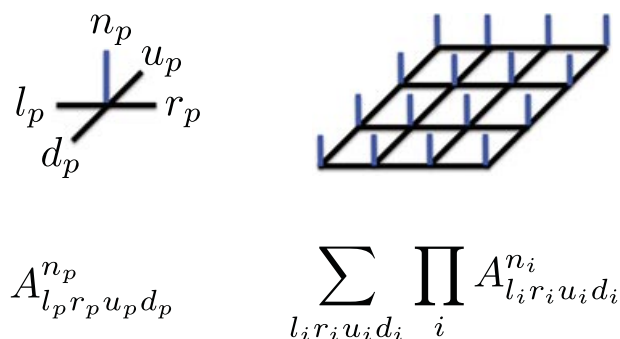


FIGURE 6 | Projected entangled pair states (PEPS) are associated with arbitrary networks of entanglement. Shown here is a representation of the PEPS for a square lattice of sites. Each local state $|n_p\rangle$ is then associated with a tensor with four auxiliary indices l, r, u, d that point along the bonds of the square lattice. The PEPS approximation for the wavefunction coefficient $\Psi^{n_1 n_2 \dots n_k}$ is obtained by contracting all the auxiliary indices.

correlations means that we expect a TTN will reach the accuracy of a given MPS calculation with smaller auxiliary dimension of the tensors. Preliminary investigations appear to confirm this.⁶⁰

Projected Entangled Pair States

The TTN and the MPS possess a special connectivity, with no cycles between the tensors in the wavefunctions. This is why these states are compatible with very simple and efficient algorithms for calculating norms and expectation values. As one contracts the expectation value network, shown in Figure 4 for the MPS, one proceeds through intermediates where the number of dangling bonds is independent of system size.

States with arbitrary networks of entanglement on the other hand, including cycles, are known as PEPS. To illustrate the PEPS construction, we consider an appropriate state to describe a square lattice of sites. In the PEPS, each site is associated with a tensor A_z^n , where z denotes the four auxiliary indices l, r, u, d that point toward the neighbors in the directions left, right, up, and down. In this notation, the PEPS wavefunction takes exactly the same form as the TTN (33), except the auxiliary indices are contracted according to the ‘bonds’ of the square lattice (see Figure 6).

The name PEPS arises from the projected wavefunction interpretation of the state, described in section *Projected Wavefunctions* for MPSs. Every site is surrounded by four ‘bonds’ on which we place maximally entangled states of auxiliary particles, $\sum_m |\eta\bar{\eta}\rangle$, and each site is associated with a projector operator which projects from the auxiliary particles onto the

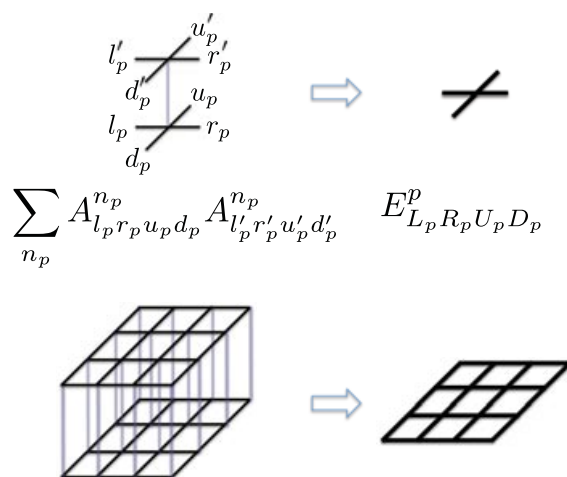


FIGURE 7 | (Bottom) Projected entangled pair states (PEPS) norm evaluation can be represented by fusing a PEPS bra lattice (pointing downward) with a PEPS ket lattice (pointing upward). The result is a square lattice which must be contracted over the auxiliary indices. (Top) Each contraction of the local states $|n_p\rangle$ results in an E tensor with four auxiliary indices, each of dimension M^2 .

physical particles,

$$\hat{A} = \sum_{nlrud} A_{lrud}^n |n\rangle \langle \eta_l \eta_r \eta_u \eta_d|. \quad (34)$$

The whole PEPS wavefunction is the product of such commuting projection operators (associated with the sites) on a reference, which is a product of entangled pairs on the bonds.

PEPS wavefunctions can be constructed to satisfy the correct entanglement area laws in two and three dimensions. Consequently, they are expected to provide compact descriptions of the ground states of almost arbitrary physical systems. However, the ability to represent the wavefunction compactly does not guarantee that computations of observables are efficient. For the square lattice PEPS above, the evaluation of an expectation value requires contracting a square lattice network of E-type tensors, as shown in Figure 7, where $E_{zz} = \sum_n A_z^n A_z^n$. As we contract column by column, we form intermediates where the number of dangling bonds is proportional to the system width. The cost of forming these intermediates, and thus the cost of the contraction, grows exponentially with system width. However, for physical systems, it appears to be possible to approximate these intermediates by simpler forms. For example, a two column contracted intermediate can be accurately approximated by a single-column intermediate with slightly larger auxiliary dimension (Figure 8). This means that an iterative approximation algorithm can be constructed where the evaluation of an

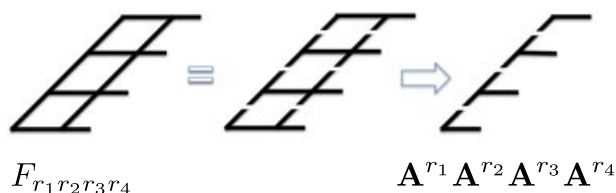


FIGURE 8 | Naive contraction of the square lattice of E tensors that arises from the projected entangled pair states (PEPS) norm evaluation is of exponential cost. Instead, we consider first a contraction of two E columns, starting from the left edge. We then view the dangling bonds (pointing toward the right, here r_1, r_2, r_3, r_4) as the local state indices of a fictitious wavefunction $F^{r_1 r_2 r_3 r_4}$. We can imagine approximating such a wavefunction by a matrix product state (MPS) of a small auxiliary dimension, $A^{r_1} A^{r_2} A^{r_3} A^{r_4}$. This effectively approximates the two column structure by a single-column structure. Repeating this for the remaining columns of the E-square lattice, we can contract the full PEPS from left to right efficiently. The final top-to-bottom contraction involves only a single column and is of the same complexity as an MPS norm evaluation.

observable reduces finally to a single-column contraction, as in the case of an MPS observable evaluation.^{55,56}

Multiscale Entanglement Renormalization Ansatz

The MERA is a flexible entanglement structure that is closely related to the original real space renormalization group used in statistical problems. For simplicity, we describe the MERA in one dimension, although analogous constructions can be considered in two and three dimensions.

MERAs consist of a set of operators known as isometries, which perform coarse graining, and disentanglers, which are unitary rotations. We first discuss the isometries. We divide the k sites under consideration into blocks. For concreteness, we take each block to contain three sites. The first block spans the space $\{|n_1 n_2 n_3\rangle\}$. We now coarse grain from the Hilbert space of the three site block $\{|n_1 n_2 n_3\rangle\}$ to a single effective site with a reduced Hilbert space $\{|n'_1\rangle\}$ of specified dimension M . The operator which performs this projection is an isometry,

$$\hat{W} = \sum_{n'_1, n_1 n_2 n_3} |n'_1\rangle W_{n'_1 n_1 n_2 n_3} \langle n_1 n_2 n_3|, \quad (35)$$

as illustrated in Figure 9. In a similar way, we coarse grain all the three site blocks onto effective sites, thus giving a lattice with $k/3$ effective sites, each with an M -dimensional Hilbert space. Once all the blocks have been coarse grained, the coarse graining can be iterated. For example, we can use another isometry to map a block of three coarse-grained sites to a single,

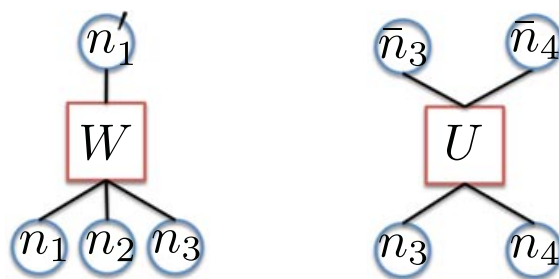


FIGURE 9 | The multiscale entanglement renormalization ansatz is constructed from two types of tensors: isometries W , which map the states of a block of sites to a single effective sites; and disentanglers U , which perform a unitary rotation of the basis of two or more sites.

even more coarse-grained site $\{|n'_1 n'_2 n'_3\rangle\} \rightarrow \{|n''_1\rangle\}$. After $\log_3 k$ levels of coarse graining, we obtain a final effective site with an M -dimensional Hilbert space, in which the ground state of the entire problem is expressed. The iterated isometries are connected in a tree structure (of different nature to the TTN described above) and the coefficient of the isometries \hat{W} are the variational parameters of the state.

The problem with using only isometries in a tree structure is that the coarse graining does not treat the sites uniformly, and thus neglects some entanglement that may be present in the ground state. For example, in the above, the effective coarse-grained states in the first two blocks $\{|n'_1\rangle\}, \{|n'_2\rangle\}$ build in entanglement within the first and second blocks of three sites in the ground state, but do not between sites 3 and 4, which are on the boundaries of the blocks. The idea behind the MERA is that one can incorporate this missing entanglement between the block boundaries, by applying a unitary operator to these sites.⁶¹ For example, in the above example, we would use a unitary operator which acts on the product space of sites 3 and 4. The MERA consists of a layer of disentanglers, which first handle the entanglement between the boundaries of the blocks, alternating with layers of isometries, which handle the entanglement within the blocks,

$$|\Psi\rangle = (\hat{W}_l \hat{U}_1 \hat{W}_2 \hat{U}_2 \dots \hat{W}_l) \sum_{\mathbf{n}} |\mathbf{n}\rangle, \quad (36)$$

where l is $\log_3 k$, the number of layers in the MERA. The MERA structure is illustrated in Figure 10.

The MERA possesses several unique strengths. First, the presence of disentanglers allows the MERA to correctly describe the entanglement area laws in any dimension, even including the corrections to area laws that are present in gapless one-dimensional systems. Second, the close relationship between the MERA and a tree structure means that unlike the

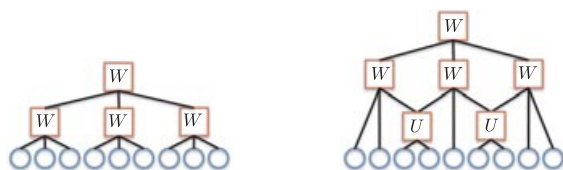


FIGURE 10 | (Left) The original real space RG is associated with a tree structure for the wavefunction, where a layer of sites is subjected to successive coarse grainings (isometries, W). However, the coarse grainings within the blocks neglect to take into account entanglement at the boundaries of the blocks. (Right) This is rectified within the MERA by first performing a disentangling operation U between the block boundaries before coarse graining.

PEPS, expectation values and observables can be exactly computed in polynomial time.⁶¹ Nonetheless, the computational effort of using the MERA remains high, and applications remain at a very early stage of development.

CONCLUSIONS

In this review, we have described the basic theory behind low entanglement wavefunction approximations. Because low entanglement wavefunctions occupy a region of Hilbert space far from the Hartree–

Fock mean-field limit, they offer the intriguing possibility to efficiently model chemical problems that challenge standard quantum chemistry methods. In recent years, practical evidence of this has been provided by the DMRG algorithm (based on low entanglement wavefunctions known as MPSs). This has been used to describe large multireference, strong correlation problems involving transition metals and excited states. In our discussion, in addition to MPSs, we have also described more general low entanglement approximations, such as TTNs, PEPS, and MERA. These states are considerably more complex than MPSs and their properties are not yet completely understood. Nonetheless, it is clear at the mathematical level that they overcome many of the formal limitations of MPSs, particularly in relation to the encoding of entanglement as a function of dimensionality. A key remaining challenge is the development of practical numerical algorithms to work with these flexible states. This is an area where we can expect the wealth of knowledge in quantum chemistry in computational algorithms for complex wavefunctions to play a major part. Numerical computation with general low entanglement approximations is thus positioned to be an exciting possibility for quantum chemistry to explore in this coming decade.

ACKNOWLEDGMENTS

The author acknowledges support from the NSF programs CHE-0645380 and CHE-1004603.

REFERENCES

- White SR. Density-matrix algorithms for quantum renormalization groups. *Phys Rev B* 1993, 48:10345–10356.
- White SR. Density matrix formulation for quantum renormalization groups. *Phys Rev Lett* 1992, 69:2863–2866.
- Schollwöck U. The density-matrix renormalization group. *Rev Mod Phys* 2005, 77:259–315.
- Hallberg KA. New trends in density matrix renormalization. *Adv Phys* 2006, 55:477–526.
- Chan GK-L, Dorando JJ, Ghosh D, Hachmann J, Neuscamman E, Wang H, Yanai T. An introduction to the density matrix renormalization group ansatz in quantum chemistry. In: Wilson S, Groot PJ, Mariani J, Delgado-Barrio G, Piecuch P, eds. *Frontiers in Quantum Systems in Chemistry and Physics*. Progress in Theoretical Chemistry and Physics. Vol. 18. Springer Netherlands, Dordrecht, The Netherlands; 2008, 49–65.
- Schollwöck U. The density-matrix renormalization group in the age of matrix product states. *Ann Phys* 2010, 326:96–213.
- Chan GK-L, Sharma S. The density matrix renormalization group in quantum chemistry. *Annu Rev Phys Chem* 2011, 62:465–481.
- Marti KH, Reiher M. The density matrix renormalization group algorithm in quantum chemistry. *Z Phys Chem* 2010, 224:583–599.
- Nishino T, Okunishi K. Product wave function renormalization group. *J Phys Soc Jpn* 1995, 64:4084–4087.
- Dukelsky J, Martin-Delgado MA, Nishino T, Sierra G. Equivalence of the variational matrix product method and the density matrix renormalization group applied to spin chains. *Europhys Lett* 1998, 43:457–462.
- Nishino T, Hikihara T, Okunishi K, Hieida Y. Density matrix renormalization group: introduction from

- a variational point of view. *Int J Mod Phys B* 1999, 13:1–26.
12. Rommer S, Östlund S. Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys Rev B* 1997, 55:2164–2181.
 13. Verstraete F, Porras D, Cirac JI. Density matrix renormalization group and periodic boundary conditions: a quantum information perspective. *Phys Rev Lett* 2004, 93:227205.
 14. Chan GK-L. Density matrix renormalisation group Lagrangians. *Phys Chem Chem Phys* 2008, 10:3454–3459.
 15. Verstraete F, Murg V, Cirac JI. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv Phys* 2008, 57:143–224.
 16. Vidal G. Efficient simulation of one-dimensional quantum many-body systems. *Phys Rev Lett* 2004, 93:40502.
 17. Legeza Ö, Sólyom J. Optimizing the density-matrix renormalization group method using quantum information entropy. *Phys Rev B* 2003, 68:195116.
 18. White SR, Martin RL. Ab initio quantum chemistry using the density matrix renormalization group. *J Chem Phys* 1999, 110:4127–4130.
 19. Daul S, Ciofini I, Daul C, White SR. Full-CI quantum chemistry using the density matrix renormalization group. *Int J Quantum Chem* 2000, 79:331–342.
 20. Chan GK-L, Head-Gordon M. Highly correlated calculations with a polynomial cost algorithm: a study of the density matrix renormalization group. *J Chem Phys* 2002, 116:4462–4476.
 21. Chan GK-L. An algorithm for large scale density matrix renormalization group calculations. *J Chem Phys* 2004, 120:3172–3178.
 22. Chan GK-L, Head-Gordon M. Exact solution (within a triple-zeta, double polarization basis set) of the electronic Schrödinger equation for water. *J Chem Phys* 2003, 118:8551–8554.
 23. Chan GK-L, Kállay M, Gauss J. State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve. *J Chem Phys* 2004, 121:6110–6116.
 24. Chan GK-L, Van Voorhis T. Density-matrix renormalization-group algorithms with nonorthogonal orbitals and non-Hermitian operators, and applications to polyenes. *J Chem Phys* 2005, 122:204101.
 25. Hachmann J, Cardoen W, Chan GK-L. Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group. *J Chem Phys* 2006, 125:144101.
 26. Hachmann J, Dorando JJ, Avilés M, Chan GK-L. The radical character of the acenes: a density matrix renormalization group study. *J Chem Phys* 2007, 127:134309.
 27. Dorando JJ, Hachmann J, Chan GK-L. Targeted excited state algorithms. *J Chem Phys* 2007, 127:084109.
 28. Ghosh D, Hachmann J, Yanai T, Chan GK-L. Orbital optimization in the density matrix renormalization group, with applications to polyenes and β -carotene. *J Chem Phys* 2008, 128:144117.
 29. Dorando JJ, Hachmann J, Chan GK-L. Analytic theory of response in the density matrix renormalization group. *J Chem Phys* 2009, 130:184111.
 30. Kurashige Y, Yanai T. High-performance ab initio density matrix renormalization group method: applicability to large-scale multireference problems for metal compounds. *J Chem Phys* 2009, 130:234114.
 31. Legeza Ö, Röder J, Hess BA. Controlling the accuracy of the density-matrix renormalization-group method: the dynamical block state selection approach. *Phys Rev B* 2003, 67:125114.
 32. Legeza Ö, Röder J, Hess BA. QC-DMRG study of the ionic-neutral curve crossing of LiF. *Mol Phys* 2003, 101:2019–2028.
 33. Legeza Ö, Sólyom J. Quantum data compression, quantum information generation, and the density-matrix renormalization-group method. *Phys Rev B* 2004, 70:205118.
 34. Luo H-G, Qin M-P, Xiang T. Optimizing Hartree-Fock orbitals by the density-matrix renormalization group. *Phys Rev B* 2010, 81:235129.
 35. Marti KH, Ondík IM, Moritz G, Reiher M. Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *J Chem Phys* 2008, 128:014104.
 36. Mitrushenkov AO, Fano G, Ortolani F, Linguerri R, Palmieri P. Quantum chemistry using the density matrix renormalization group. *J Chem Phys* 2001, 115:6815–6821.
 37. Mitrushenkov AO, Linguerri R, Palmieri P, Fano G. Quantum chemistry using the density matrix renormalization group. II. *J Chem Phys* 2003, 119:4148–4158.
 38. Moritz G, Hess BA, Reiher M. Convergence behavior of the density-matrix renormalization group algorithm for optimized orbital orderings. *J Chem Phys* 2005, 122:024107.
 39. Moritz G, Reiher M. Decomposition of density matrix renormalization group states into a Slater determinant basis. *J Chem Phys* 2007, 126:244109.
 40. Moritz G, Reiher M. Construction of environment states in quantum-chemical density-matrix renormalization group calculations. *J Chem Phys* 2006, 124:034103.
 41. Moritz G, Wolf A, Reiher M. Relativistic DMRG calculations on the curve crossing of cesium hydride. *J Chem Phys* 2005, 123:184105.
 42. Yanai T, Kurashige Y, Ghosh D, Chan GK-L. Accelerating convergence in iterative solutions of large

- active-space self-consistent field calculations. *Int J Quantum Chem* 2009, 109:2178–2190.
43. Yanai T, Kurashige Y, Neuscamman E, Chan GK-L. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J Chem Phys* 2010, 132:024105.
 44. Zgid D, Nooijen M. Obtaining the two-body density matrix in the density matrix renormalization group method. *J Chem Phys* 2008, 128:144115.
 45. Zgid D, Nooijen M. The density matrix renormalization group self-consistent field method: orbital optimization with the density matrix renormalization group method in the active space. *J Chem Phys* 2008, 128:144116.
 46. Zgid D, Nooijen M. On the spin and symmetry adaptation of the density matrix renormalization group method. *J Chem Phys* 2008, 128:014107.
 47. Neuscamman E, Yanai T, Chan GK-L. Strongly contracted canonical transformation theory. *J Chem Phys* 2010, 132:024106.
 48. Mizukami W, Kurashige Y, Yanai T. Novel quantum states of electron spins in polycarbenes from ab initio density matrix renormalization group calculations. *J Chem Phys* 2010, 133:091101.
 49. Barcza G, Legeza Ö, Marti KH, Reiher M. Quantum-information analysis of electronic states of different molecular structures. *Phys Rev A* 2011, 83:012508.
 50. Marti KH, Reiher M. New electron correlation theories for transition metal chemistry. *Phys Chem Chem Phys* 2011, 13:6750–6759.
 51. Nishino T, Okunishi K, Hieida Y, Maeshima N, Akutsu Y. Self-consistent tensor product variational approximation for 3D classical models. *Nucl Phys B* 2000, 575:504–512.
 52. Nishino T, Hieida Y, Okunishi K, Maeshima N, Akutsu Y, Gendiar A. Two-dimensional tensor product variational formulation. *Prog Theor Phys* 2001, 105:409–417.
 53. Maeshima N, Hieida Y, Akutsu Y, Nishino T, Okunishi K. Vertical density matrix algorithm: a higher-dimensional numerical renormalization scheme based on the tensor product state ansatz. *Phys Rev E* 2001, 64:016705.
 54. Kraus CV, Schuch N, Verstraete F, Cirac JL. Fermionic projected entangled pair states. *Phys Rev A* 2010, 81:052338.
 55. Verstraete F, Cirac JL. Renormalization algorithms for quantum-many body systems in two and higher dimensions. *Arxiv preprint cond-mat/0407066*, 2004.
 56. Murg V, Verstraete F, Cirac JL. Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states. *Phys Rev A* 2007, 75:033605.
 57. Shi Y, Duan L, Vidal G. Classical simulation of quantum many-body systems with a tree tensor network. *Phys Rev A* 2006, 74:22320.
 58. Tagliacozzo L, Evenbly G, Vidal G. Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law. *Phys Rev B* 2009, 80:235127.
 59. Vidal G. Class of quantum many-body states that can be efficiently simulated. *Phys Rev Lett* 2008, 101:110501.
 60. Murg V, Verstraete F, Ö. Legeza, Noack RM. Simulating strongly correlated quantum systems with tree tensor networks. *Phys Rev B* 2010, 82:205105.
 61. Vidal G. Entanglement renormalization. *Phys Rev Lett* 2007, 99:220405.
 62. Evenbly G, Vidal G. Entanglement renormalization in two spatial dimensions. *Phys Rev Lett* 2009, 102:180406.
 63. Corboz P, Vidal G. Fermionic multiscale entanglement renormalization ansatz. *Phys Rev B* 2009, 80:165129.
 64. Vidal G. Entanglement renormalization: an introduction. *Arxiv preprint arXiv:0912.1651*, 2009.
 65. Marti KH, Bauer B, Reiher M, Troyer M, Verstraete F. Complete-graph tensor network states: a new fermionic wave function ansatz for molecules. *New J Phys* 2010, 12:103008.
 66. Kohn W. Nobel lecture: Electronic structure of matter wave functions and density functionals. *Rev Mod Phys* 1999, 71:1253–1266.
 67. Schmidt MW, Ruedenberg KE. Effective convergence to complete orbital bases and to the atomic Hartree–Fock limit through systematic sequences of gaussian primitives. *J Chem Phys* 1979, 71:3951–3962.
 68. Kutzelnigg W. Theory of the expansion of wavefunctions in a gaussian basis. *Int J Quantum Chem.* 1994, 51:447–463.
 69. Pahl FA, Handy NC. Plane waves and radial polynomials: a new mixed basis. *Mol Phys* 2002, 100:3199–3224.
 70. Harrison RJ, Fann GI, Yanai T, Gan Z, Beylkin G. Multiresolution quantum chemistry: basic theory and initial applications. *J Chem Phys* 2004, 121:11587–11598.
 71. Szabo A, Ostlund NS. *Modern Quantum Chemistry*. Mineola, NY: Dover Publications; 1996.
 72. Helgaker T, Olsen J, Jorgensen P. *Molecular Electronic Structure Theory*. Chichester, UK: John Wiley & Sons; 2000.
 73. Bartlett RJ, Musial M. Coupled-cluster theory in quantum chemistry. *Rev Mod Phys* 2007, 79:291–352.
 74. Bartlett RJ. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annu Rev Phys Chem* 1981, 32:359–401.
 75. Jastrow R. Many-body problem with strong forces. *Phys. Rev.* 1955, 98:1479–1484.
 76. Umrigar CJ, Wilson KG, Wilkins JW. Optimized trial wave functions for quantum Monte Carlo calculations. *Phys Rev Lett* 1988, 60:1719–1722.

77. Boys SF, Handy NC. The determination of energies and wavefunctions with full electronic correlation. *Proc R Soc London A: Math Phys Sci* 1969, 310:43–61.
78. Blaizot JP, Ripka G. *Quantum Theory of Finite Systems*. Cambridge, MA: MIT Press; 1986.
79. Coleman AJ. Structure of fermion density matrices. II. Antisymmetrized geminal powers. *J Math Phys* 1965, 6:1425–1431.
80. Weiner B, Goscinski O. Calculation of optimal generalized antisymmetrized geminal-power functions and their associated excitation spectrum. *Phys Rev A* 1980, 22:2374–2391.
81. Bajdich M, Mitas L, Drobn G, Wagner LK, Schmidt KE. Pfaffian pairing wave functions in electronic-structure quantum monte carlo simulations. *Phys Rev Lett* 2006, 96:130201.
82. Jeziorski B, Paldus J. Valence universal exponential ansatz and the cluster structure of multireference configuration interaction wave function. *J Chem Phys* 1989, 90:2714–2731.
83. Mahapatra US, Datta B, Bandyopadhyay B, Mukherjee D. State-specific multi-reference coupled cluster formulations: two paradigms. *Adv Quantum Chem* 1998, 30:163–193.
84. Evangelista FA, Gauss J. An orbital-invariant internally contracted multireference coupled cluster approach. *J Chem Phys* 2011, 134:114102.
85. Hanrath M. An exponential multireference wavefunction ansatz. *J Chem Phys* 2005, 123:084102.
86. Evangelista FA, Simmonett AC, Allen WD, Schaefer III HF, Gauss J. Triple excitations in state-specific multireference coupled cluster theory: Application of mk-mrcc sdt and mk-mrccsdt-n methods to model systems. *J Chem Phys* 2008, 128:124104.
87. Luchow A. Quantum monte carlo methods. *Comput Mol Sci* 2011, 1:388–402.
88. Goddard III WA, Dunning Jr TH, Hunt WJ, Hay PJ. Generalized valence bond description of bonding in low-lying states of molecules. *Acc Chem Res* 1973, 6:368–376.
89. Van Voorhis T, Head-Gordon M. Connections between coupled cluster and generalized valence bond theories. *J Chem Phys* 2001, 115:7814–7821.
90. Voter AF, Goddard III WA. A method for describing resonance between generalized valence bond wavefunctions. *Chem Phys* 1981, 57:253–259.
91. Cooper DL, Gerratt J, Raimondi M. Applications of spin-coupled valence bond theory. *Chem Rev* 1991, 91:929–964.
92. Hiberty PC, Shaik S. A survey of recent developments in ab initio valence bond theory. *J Comput Chem* 2007, 28:137–151.
93. Audretsch J. *Entangled Systems: New Directions in Quantum Physics*. John Wiley & Sons; 2007.
94. Eisert J, Cramer M, Plenio MB. Colloquium: area laws for the entanglement entropy. *Rev Mod Phys* 2010, 82:277–306.
95. Hastings MB. An area law for one-dimensional quantum systems. *J Stat Mech* 2007, 2007:08024.
96. Hastings MB. Entropy and entanglement in quantum ground states. *Phys Rev B* 2007, 76:035114.
97. Stoudenmire EM, White SR. Studying two dimensional systems with the density matrix renormalization group. *Ann Rev Cond Mat Phys* 2011, 3:111–128.
98. Reatto L, Masserini GL. Shadow wave function for many-boson systems. *Phys Rev B* 1988, 38:4516–4522.