

# The *ab-initio* density matrix renormalization group in practice

Roberto Olivares-Amaya,<sup>1,a)</sup> Weifeng Hu,<sup>1,a)</sup> Naoki Nakatani,<sup>1,2</sup> Sandeep Sharma,<sup>1</sup> Jun Yang,<sup>1</sup> and Garnet Kin-Lic Chan<sup>1</sup>

<sup>1</sup>Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA

<sup>2</sup>Catalysis Research Center, Hokkaido University, Kita 21 Nishi 10, Sapporo, Hokkaido 001-0021, Japan

(Received 14 September 2014; accepted 19 December 2014; published online 15 January 2015)

The *ab-initio* density matrix renormalization group (DMRG) is a tool that can be applied to a wide variety of interesting problems in quantum chemistry. Here, we examine the density matrix renormalization group from the vantage point of the quantum chemistry user. What kinds of problems is the DMRG well-suited to? What are the largest systems that can be treated at practical cost? What sort of accuracies can be obtained, and how do we reason about the computational difficulty in different molecules? By examining a diverse benchmark set of molecules:  $\pi$ -electron systems, benchmark main-group and transition metal dimers, and the Mn-oxo-salen and Fe-porphine organometallic compounds, we provide some answers to these questions, and show how the density matrix renormalization group is used in practice. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4905329>]

## I. INTRODUCTION

The density matrix renormalization group algorithm (DMRG), introduced by White,<sup>1,2</sup> has by now been widely applied in quantum chemistry.<sup>3–15</sup> Developing from its roots in condensed matter, its earliest application to chemical problems used the semi-empirical  $\pi$ -electron Pariser-Parr-Pople Hamiltonian.<sup>16–18</sup> White and Martin introduced the first efficient formulation of the DMRG algorithm for *ab-initio* Hamiltonians (using some algorithmic contributions from Xiang<sup>19</sup>) and the *ab-initio* density matrix renormalization group method was born. Since then, many groups have independently implemented and improved on the *ab-initio* DMRG algorithm. Some of these improvements include parallelization,<sup>8,20</sup> non-Abelian symmetry and spin-adaptation,<sup>7,21–23</sup> orbital ordering,<sup>5,24–26</sup> and optimization,<sup>9,27–29</sup> more sophisticated initial guesses,<sup>5,24,25,30,31</sup> better noise algorithms,<sup>5,32</sup> extrapolation procedures,<sup>5,33,34</sup> response theories,<sup>35,36</sup> as well as the combination of the DMRG with various other quantum chemistry methods such as perturbation theory,<sup>37</sup> canonical transformations,<sup>38</sup> configuration interaction,<sup>39</sup> and relativistic Hamiltonians.<sup>40</sup>

In the ecosystem of quantum chemistry, the DMRG occupies a unique spot. On one hand, because the accuracy depends on a single, essentially continuously, tunable parameter—the number of renormalized states  $M$ —calculations can be converged and extrapolated in a simple way to the exact full configuration interaction limit. As a result, its earliest applications were as a proxy for full configuration interaction (FCI) in modest sized molecules where FCI would be too expensive. Examples include the early water<sup>34</sup> and nitrogen<sup>41</sup> calculations that used relatively small (but still inaccessible to FCI) basis sets, as well as more recent benchmarks, such as the exact treatment of the Be<sub>2</sub> molecule<sup>42</sup> at the complete basis set limit. On the other hand, the matrix product state (MPS) wavefunction<sup>43–46</sup> underlying the DMRG is not based on an

excitation expansion around a single-reference, and is thus well suited to non-mean-field, or strongly correlated, electronic structure, such as arising in transition-metal chemistry. Here, the DMRG has so far been applied in a complete active space setting,<sup>8,38,47–51</sup> starting with the early work of Reiher and coworkers,<sup>47</sup> to the latest calculations on systems as large as the bioinorganic Mn<sub>4</sub>Ca core of photosystem II by Kurashige *et al.*,<sup>52</sup> and the ubiquitous [4Fe-4S] biological iron-sulfur complexes by Sharma *et al.*,<sup>53</sup> these have active spaces in excess of 50 orbitals. Finally, the internal structure of the MPS means that it is uniquely suited to pseudo-one-dimensional correlation, and in the *ab-initio* chemical context, the DMRG has been used since its inception to study ground- and excited-states of  $\pi$ -conjugated molecules,<sup>27,54,55</sup> such as the polyacenes<sup>55</sup> and graphene nanoribbons,<sup>49</sup> as well as other one-dimensional systems, such as atomic chains and rings.<sup>54,56</sup> The entanglement structure of the MPS has even generated a niche in interpretative quantum chemistry.<sup>52,57–60</sup>

With the advent of publicly available quantum chemistry DMRG codes,<sup>10,20–22</sup> the DMRG is transitioning into a method available not only to expert developers but also to the informed user. As with all complex methods, there is some degree of experience required to use it effectively. The purpose of this article is therefore to answer the following questions:

- What sort of molecules can the DMRG be practically applied to?
- What sort of accuracies can be obtained and at what cost? What are the typical sizes of systems (e.g., number of active orbitals) that can be treated with practical computational resources?
- How do we reason about the accuracy of DMRG calculations for different molecules?
- How is a DMRG calculation best specified (e.g., in terms of starting orbitals and their order)?

We answer these questions both from theoretical reasoning and by applying the method to a “representative” set of

<sup>a)</sup>R. Olivares-Amaya and W. Hu contributed equally to this work.

molecules for DMRG applications:  $\pi$ -conjugated systems (polyenes and arenes), benchmark dimers ( $C_2$  and  $Cr_2$ ) in several basis sets (with more than 100 orbitals), and organo-metallic complexes (Fe-porphine and a Mn-oxo-salen model of Jacobsen's catalyst) in large active spaces (with more than 40 active orbitals). Importantly, the calculations we describe are all run in a completely black-box fashion using the default settings of our BLOCK code, and thus should be easily replicated. The BLOCK code is currently available in the MOLPRO,<sup>61</sup> ORCA,<sup>62</sup> and Q-CHEM<sup>63</sup> distributions, and can also be freely downloaded from our weblink.<sup>64</sup>

In Secs. II A–II E, we give a self-contained overview of the DMRG algorithm and MPS. We describe both the theory and how to reason about it with respect to orbital choices, ordering, accuracy, and cost, from the viewpoint of the user. We further define some of the choices and settings used in our particular implementation of the DMRG algorithm in BLOCK, with respect to the sweep and noise schedule. The various issues are then explored in the context of the set of “representative” molecules in Sec. III. The summary of our findings and recommendations is given in Sec. IV.

## II. THEORETICAL BACKGROUND

### A. The DMRG wavefunction and optimization algorithm

As with other wavefunction methods in chemistry, the DMRG is based on an approximate wavefunction ansatz. This ansatz is known as the MPS. A MPS is a non-linear wavefunction ansatz, built as a product of variational objects for each orbital in the basis. Technically, the DMRG refers to the combination of a specific, efficient, and self-consistent optimization algorithm (the sweep algorithm) with the variational MPS ansatz. In fact, historically, the DMRG is discussed primarily in terms of the optimization algorithm,<sup>1,2</sup> which derives from the original numerical renormalization group,<sup>65</sup> rather than in terms of its wavefunction structure. However, the two pictures are closely intertwined because in practice the existence of an efficient optimization scheme for the highly non-linear ansatz is a key to the success of the DMRG in practice. The close relationship between the wavefunction and optimization method in DMRG recalls the close relationship between the Hartree-Fock Slater determinant approximation and the self-consistent field algorithm used to optimize it. Indeed, in quantum chemistry, we often interpret “SCF” to mean Slater determinant in an eigenvalue (e.g., ground-state) context, and similarly “DMRG” is often synonymous with MPS for eigenvalue problems. Note that the parallel terminology is not the only similarity between DMRG and Hartree-Fock theory; the product structure of the MPS is the source of further useful analogies, discussed in detail in Refs. 66–69.

In practice, a DMRG calculation will often use two types of matrix product states: a *one-site* and a *two-site* MPS. (Site is here synonymous with orbital, in an occupation number representation.) These refer to different but related ansatz. The one-site MPS is easier to reason about and converge to high precision, but optimization with this ansatz alone often ends up trapped in local minima<sup>32</sup> due to the inability to change

“quantum numbers” (a type of internal wavefunction symmetry) during the DMRG sweep, as described in Sec. II C. A standard strategy is therefore to start the DMRG calculation using the more flexible two-site MPS and then to switch near convergence to the one-site MPS.

To define a MPS, we are first required to order the orbitals  $\phi_1 \dots \phi_k$ . Then, the one-site MPS is defined as

$$|\Psi\rangle = \sum_n \mathbf{A}^{n_1} \mathbf{A}^{n_2} \dots \mathbf{A}^{n_k} |n_1 \dots n_k\rangle. \quad (1)$$

Here,  $n_i$  denotes the occupancy of orbital  $\phi_i$ . It takes four values corresponding to the states  $|-\rangle$ ,  $|\uparrow\rangle$ ,  $|\downarrow\rangle$ , and  $|\uparrow\downarrow\rangle$ . For a given index  $n_i$ ,  $\mathbf{A}^{n_i}$  is then an  $M \times M$  matrix, and each  $\mathbf{A}$  is correspondingly a 3-index tensor.  $\mathbf{A}^{n_i}$  contains information on the wavefunction coefficients for that particular choice of occupancy of orbital  $\phi_i$ . The first and last  $\mathbf{A}$ 's are slightly special, as the dimensions of  $\mathbf{A}^{n_1}$  and  $\mathbf{A}^{n_k}$  are  $1 \times M$  and  $M \times 1$ , respectively; this ensures that for a given occupancy string  $n_1 \dots n_k$ , performing the vector, matrix, matrix ... vector product defined by  $\mathbf{A}^{n_1} \mathbf{A}^{n_2} \dots \mathbf{A}^{n_k}$  yields a scalar wavefunction amplitude, corresponding to the coefficient of the determinant  $|n_1 \dots n_k\rangle$ .

We see that the DMRG is a kind of product wavefunction, but unlike a Slater determinant, there is a variational object associated with each *orbital in the basis*, rather than with each electron. The variational freedom and thus accuracy of a DMRG wavefunction is controlled by the single dimensional parameter  $M$ , referred to as the “number of renormalized states,” or MPS bond dimension. In total, there are  $O(M^{2k})$  variational parameters in the wavefunction. When  $M^2 \sim N_d$  (where  $N_d$  is the total number of possible determinants), the DMRG wavefunction becomes exact, but typically the convergence with  $M$  is rapid and a much smaller  $M$  (in the range of 1000–100 000) is used for chemical accuracy in practice.

The two-site DMRG wavefunction is defined similarly, but the  $\mathbf{A}$  tensors of a special pair of adjacent (“two”) sites are fused together to become a four-index tensor  $\mathbf{T}$

$$|\Psi\rangle = \sum_n \mathbf{A}^{n_1} \mathbf{A}^{n_2} \dots \mathbf{T}^{n_{i+1}} \dots \mathbf{A}^{n_k} |n_1 \dots n_k\rangle. \quad (2)$$

For each pair of occupancies  $n_i$  and  $n_{i+1}$ ,  $\mathbf{T}^{n_{i+1}}$  is an  $M \times M$  matrix. The  $\mathbf{T}$  tensor in the two-site MPS gives it slightly more variational freedom than the one-site MPS. This allows for an optimization of the quantum numbers as described below, which helps to avoid local minima. During the DMRG sweep the boundary for the special pair of sites is iterated from the first (last) pair of orbitals to the last (first) during the DMRG forwards (backwards) sweep optimization; each boundary corresponds to a single step of the sweep. Thus in the DMRG sweep there is not a unique two-site MPS, but rather  $k-1$  of them, depending on where we choose to place the two sites  $i+1$ . When we refer to the DMRG energy from a two-site wavefunction, we therefore mean the energies obtained at the lowest point in the sweep. This is often near to when the two sites are at the middle of the lattice.

## B. Sweeps and operators

In the DMRG sweep algorithm, the energy is variationally optimized with respect to a *single*  $\mathbf{A}$  (or  $\mathbf{T}$  in the case of the two-site MPS) tensor at a time. This is analogous to Hartree's original SCF procedure in Hartree-Fock theory where a single orbital is updated at a time, and is also conceptually related to "alternating" methods in multi-linear algebra. The sweep traverses tensors from left to right (forwards) and from right to left (backwards) and multiple sweeps are carried out to converge the energy. To obtain the updated  $\mathbf{A}^{n_i}$  matrices at each site, we solve an effective Schrödinger equation with an effective Hamiltonian (analogous to a Fock equation for each site) for the  $O(M^2)$  coefficients in  $\mathbf{A}^{n_i}$ . Because  $M^2$  can be quite large, and we only need the lowest eigenvalue and eigenvector (if we are interested in a single state), we use the Davidson algorithm.<sup>70</sup> The wavefunction solution is the dominant computational cost in a DMRG calculation. It has a formal cost scaling of  $O(M^3k^3)$  per sweep, where  $M^3$  derives from matrix multiplication, although in practice  $O(M^2)$  scaling is often observed due to the block-sparse structure of the  $\mathbf{A}^{n_i}$  matrices deriving from quantum numbers, as is used in this work.

The effective Hamiltonian at each site is defined by tracing the Hamiltonian in the energy expectation value  $\langle \Psi | H | \Psi \rangle$  with tensors from all other sites in the bra and ket. The efficient construction of the effective Hamiltonian requires decomposing it into contributions from all orbitals to the left of the current optimized tensor (the left block), and orbitals to the right of the current optimized tensor (the right block). These intermediates (renormalized operators) are stored in memory and on disk, and form the dominant memory cost of the DMRG calculation:  $O(M^2k^2)$  memory and  $O(M^2k^3)$  disk.

Because the sweeps are variational in nature, the DMRG energy at each sweep decreases monotonically. (Strictly speaking, monotonicity is guaranteed only with the one-site algorithm but is usually observed with the two-site algorithm also except when very close to final convergence.) It is more efficient to carry out the initial DMRG sweeps with small values of  $M$  to approximately converge the DMRG wavefunction, and then to carry out later sweeps with larger  $M$ . The choice of successive increasing  $M$ 's leads to a *sweep schedule*. The default sweep schedule in **Block**, as used in all our calculations here, is listed in Table I. In the calculation, the desired final  $M$  is input, and the sweep schedule is followed up to the desired  $M$ . The first sweep (the warmup sweep) is a little bit special as we need to provide an initial guess for the  $\mathbf{A}$  tensors in the MPS. In **Block**, we build the  $\mathbf{A}$  tensors for the warmup sweep corresponding to a small number of low-energy determinants.<sup>5</sup>

## C. Quantum numbers, symmetries, and noise

Global symmetries, including Abelian symmetries (such as particle number, point-group symmetry, and  $S_z$  symmetry) as well as non-Abelian symmetries (such as the  $SU(2)$  or  $S^2$  symmetry, or non-Abelian point-group symmetry) can be efficiently used with the DMRG algorithm and our **Block** implementation. We can ensure that the underlying MPS transforms according to an irrep of a global symmetry, by assigning

TABLE I. Default schedule for a DMRG calculation.

M	No. of iterations	Noise
500	8	$1 \times 10^{-4}$
1000	8	$5 \times 10^{-5}$
2000	4	$5 \times 10^{-5}$
3000	4	$5 \times 10^{-5}$
...	4	$5 \times 10^{-5}$
Max. $M$	4	$5 \times 10^{-5}$
Max. $M$	2	0

irreducible representations to the left and right matrix indices of  $\mathbf{A}^{n_i}$ . In this context, the irreps associated with the left and right matrix indices are referred to as "quantum numbers." The global irrep then implies a block-sparse structure in the  $\mathbf{A}^{n_i}$  matrices. The global molecular symmetry thus imply symmetries of local transformations of the matrices.

Note that symmetry does not itself dictate the *sizes* of the non-zero blocks of  $\mathbf{A}^{n_i}$ . In principle, these sizes should be optimized during the optimization of the wavefunction, as an additional discrete optimization problem. A poor distribution of block sizes in the tensors is the main cause of the local minimum problem in DMRG, where a calculation appears to converge to a state of too high energy. This phenomenon is sometimes also referred to as having incorrect, or "losing" quantum numbers (i.e., not having the appropriate irrep labels for the states).<sup>20</sup> One way to detect an incorrect convergence to a local minimum is to increase  $M$ . A calculation with incorrect quantum numbers at small  $M$  will exhibit a very sudden lowering of the energy at larger  $M$  when the correct quantum number sector is finally recovered.

So far in the literature, the discrete optimization of block sizes has not been addressed in a very sophisticated way in DMRG algorithms. Since the original formulation it was recognized that the two-site MPS allowed for a limited local variation of block sizes for the particular choice of the "two sites."<sup>2,5</sup> When coupled with perturbative noise which introduces random quantum numbers into the wavefunction, a DMRG sweep with the two-site wavefunction provides some ability to dynamically search different block distributions.<sup>32</sup> Using the perturbative noise is particularly important in the initial sweeps in the DMRG, as different block distributions can be easily ignored due to the small total number of states  $M$ . However, the noise should be gradually reduced in later sweeps in the schedule to avoid affecting the final converged answer. In conjunction with the default sweep schedule in **Block**, we use an accompanying noise schedule as shown in Table I. For the final sweeps at the desired maximum  $M$ , the noise is set to zero.

## D. Truncation error and extrapolation

Because the two-site MPS has additional variational freedom over the one-site MPS, there is, in general, an error from projecting from a two-site MPS to a one-site MPS, which only vanishes in the limit of large  $M$ . This "truncation error" (equivalent to the discarded weight in the density matrix in the traditional DMRG language) provides an estimate of the accuracy

of the DMRG wavefunction. It was empirically observed<sup>5,33</sup> that the truncation error is almost linearly proportional to the error in the DMRG energy. In an exact (i.e., for sufficiently large  $M$ ) calculation, the truncation error becomes zero, and thus extrapolating to zero error gives an estimate of the “exact” energy for infinite  $M$ .

As a practical note, the “noise” in the initial sweeps introduces random error into the DMRG wavefunction and can cause problems with extrapolation when using the data at small  $M$ . One can avoid this by carrying out a few sweeps at each  $M$  without noise before proceeding to higher  $M$ .<sup>10</sup> Alternatively, after the calculation is converged, we can carry out a “reverse schedule” where  $M$  starts at its maximum value and is then *lowered* to its smallest value, to obtain good data at small  $M$  (which is often biased by the warmup sweep). We call this the *reverse schedule*. Using energies from the reverse schedule can sometimes lead to more accurate linear extrapolations of the DMRG energy. Comparison of the extrapolations using the standard schedule and reverse schedule data is shown in Fig. 1.

### E. Orbital choice and ordering in DMRG

Choosing an orbital space for a DMRG calculation, much like the selection of an active space in other wavefunction methods, requires some trial and error. However, unlike many other common wavefunction methods, a MPS is not invariant to orbital rotations within the active space, except at large  $M$ .<sup>27,28</sup> Orbital re-ordering can be viewed as a large rotation. Although the DMRG calculation becomes less sensitive to these choices as  $M$  increases, at fixed  $M$ , a good choice of orbitals and ordering greatly improves the accuracy of a DMRG calculation in practice.

In principle, the best ordering and choice of orbitals minimizes the maximum entanglement at any cut of the DMRG orbital lattice (the lattice being the one-dimensional ordering of orbitals). Indeed, for a MPS wavefunction with  $M$  renormalized states, there is an upper bound of  $\log_2 M$  on the amount of entanglement that can be captured by such a state. A related

criterion is to minimize entanglement (defined in some way) between distant orbitals in the lattice.

Entanglement is distinct from quantum chemical correlation. One way to reason about ground-state entanglement is to view it as generated by the one- and two-electron parts of the Hamiltonian, as we project from a simple product state of orbitals to the ground-state with the operator  $\exp(-\beta H)$ ,  $\beta \rightarrow \infty$ . If we start with very localized orbitals such as orthogonalized atomic orbitals, the one-electron part of the Hamiltonian will cause the orbitals to mix due to electron delocalization (to reduce the “spread” in the energy domain) and this can generate a large amount of entanglement. The “one-electron” entanglement effects can be removed by transforming to molecular orbitals. On the other hand, the two-electron part of the Hamiltonian generates entanglement also, which can typically be minimized by working with a more local (in real-space) representation, as familiar from local correlation methods. A reasonable compromise is to use split-localized orbitals, which correspond to orbitals where the occupied and virtual orbitals are separately localized.

Once the orbitals have been determined, we still need to choose an order. (One could, in principle, optimize both the shapes and the order of the orbitals directly via energy minimization: in complete active space terminology, this would correspond to “active-active” rotations. We do not consider these in this work.) The ordering should place orbitals which are most entangled (such as orbitals which are close in space, or pairs of bonding and anti-bonding orbitals) close together. In quasi-one-dimensional molecules (such as chain, strip, or ring-like molecules), there is an obvious ordering that follows the connectivity of the molecule. However, finding the optimal ordering in general is NP-hard.

Nonetheless, reasonable heuristic algorithms are available. These start with by defining a matrix, or *entanglement metric*. A rigorous measure of entanglement between two orbitals is the mutual information (pair entanglement entropy) first investigated by Rissler and White.<sup>26</sup> However, this requires knowledge of the final wavefunction. A much simpler proxy, which has some of the correct properties, is the exchange integral between the two orbitals  $K_{ij} = \int \int d\mathbf{r}_1 d\mathbf{r}_2 r_{12}^{-1} \phi_i^*(\mathbf{r}_1) \phi_j(\mathbf{r}_2) \phi_j^*(\mathbf{r}_1) \phi_i(\mathbf{r}_2)$ , which measures their proximity and spatial overlap. (The one-electron Hamiltonian was initially used in Ref. 5 and provides an alternative simple metric.) The goal is then to optimize the ordering such that  $\sum_{ij} |K_{ij}|$  (or some related function) is minimized.

In our Block code, we have two default orderings: a Fiedler vector order and a genetic algorithm optimized order. The Fiedler vector is a graph theoretic technique which provides good approximations to the graph min-cut problem, which is what we are interested in.<sup>71–73</sup> Graph techniques had earlier been examined in the DMRG ordering problem,<sup>5</sup> but a detailed study of the Fiedler vector was first presented by Barcza *et al.* in Ref. 25. The derived ordering can be obtained as follows. First, we assume that the positions of the  $k$  orbitals on the DMRG orbital lattice are continuous 1D variables  $x_1 \dots x_k$ . Then, we measure the distance between orbitals on the lattice by

$$F(\{x\}) = \sum_{ij} (x_i - x_j)^2 |K_{ij}|. \quad (3)$$

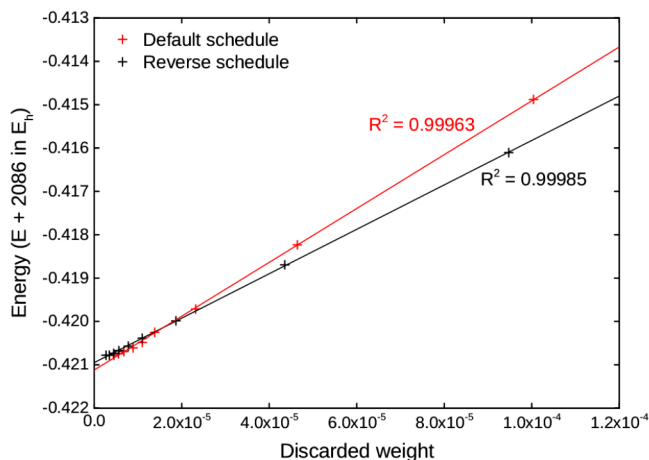


FIG. 1. Extrapolation of  $\text{Cr}_2$  (24e, 30o) energy (basis and active space described in Sec. III) using the default schedule and reverse schedule. Note that there is more noise in the data in the default schedule, as shown by the error of the fit.

We now minimize  $F(x)$  with respect to the vector of positions  $\{x\}$ . To prevent rigid translation of all “coordinates,” we fix  $\sum_i x_i = 0$ , and to prevent a trivial solution where  $x = 0$ , we impose normalization  $\sum_i x_i^2 = 1$ . We therefore solve  $\min F(\{x\})$  subject to  $\sum_i x_i = 0$ ;  $\sum_i x_i^2 = 1$ , a simple linear algebra problem, equivalent to finding the second lowest eigenvector of the graph Laplacian  $\mathbf{L} = \mathbf{D} - \mathbf{K}$  where  $\mathbf{D}$  is a diagonal matrix with entries  $D_{ii} = \sum_j |K_{ij}|$  (the first eigenvector is simply the constraint condition).<sup>74,75</sup> Sorting the values of the vector coordinates  $\{x\}$  then gives the Fiedler ordering. The advantage of this method is that it is trivial to obtain from a small matrix diagonalization, and since our optimization metric  $|K_{ij}|$  is in any case approximate, it is not necessary to obtain the true global optimum. For this reason, the Fiedler ordering is the default ordering in the BLOCK code.

We can look for additional improvement in the metric by using a genetic algorithm (GA). Genetic algorithms are global optimization algorithms based on an analogy to natural selection<sup>76</sup> and were first employed in the DMRG orbital ordering problem by Moritz *et al.*<sup>77</sup> In our work, we use as our cost function

$$F(x) = \frac{\sum_{ij} |K_{ij}| D_{ij}^2(x)}{\sum_{ij} |K_{ij}|}, \quad (4)$$

where  $D_{ij}(x)$  is the distance matrix that depends on the ordering  $x$ . To implement the GA, we need crossover and/or mutation operators upon selection. The crossover operator generates a new ordering from two different input ordering strings. The implementation in BLOCK uses a *partially matched crossover* (PMX), which works well for permutation strings.<sup>78,79</sup> The mutation operator modifies a string randomly, and in our case, we pick two parts of string randomly and swap them to create a new string. We have further chosen the probabilities of selection (by trial and error) to avoid local minima for the orbital ordering problem. We start the GA using the Fiedler vector as an initial guess, which greatly reduces the optimization time in the GA. For all cases we tested, the solution with and without the Fiedler vector guess results in the same GA minimum.

## F. Accuracy and computational cost

As discussed above, once the orbitals and ordering are defined, the accuracy of the DMRG calculation is controlled by a single parameter  $M$ , the number of renormalized states. The memory and CPU costs are summarized as follows:

- **Memory:** The scaling with the number of states  $M$  and orbitals  $k$  is roughly  $O(M^2 k^2)$ , while disk usage is  $O(M^2 k^3)$ . DMRG calculations are very memory intensive. The largest calculation in this work is the butadiene calculation (22e, 82 orbs,  $C_1$  symmetry), which used 850 Gb of memory and 17 Tb of disk. However, note that the parallel DMRG algorithm used in the BLOCK code<sup>20</sup> distributes both memory and disk storage across different nodes, and thus even modest sized computational clusters provide quite substantial resources.
- **CPU:** DMRG CPU times scale as  $O(M^3 k^3)$  per sweep. For small  $M$  (say, less than 1000), and for the initial

DMRG warmup sweep, our implementation has a large  $M^2$  overhead associated with large sets of quantum numbers which can dominate the computational cost. Using the parallelization strategy in Ref. 20, we observe reasonable parallelization up to a number of cores equal to the number of orbitals. For the largest calculation in this work (butadiene, 22e, 82 orbs,  $C_1$  symmetry), a single sweep at  $M = 3000$  (which gave an energy below CCSDT) took 25 h on 42 cores.

In almost all cases in chemistry, the DMRG energy converges almost exponentially with  $M$  (empirically, a convergence of  $\exp(\kappa(\ln M)^a)$ , where  $a \approx 2$  is observed<sup>5,80</sup>); however, the exponent  $\kappa$  of the exponential is molecule dependent. To understand  $\kappa$  and the rate of convergence in different molecules, we can make a few general statements. As discussed above,  $M$  determines the maximum amount of entanglement by the MPS wavefunction. In the simple case of (non-critical) 1D system where every orbital is equivalent (e.g., a chain of hydrogen atoms in a minimal orthogonalized atomic basis) the ground-state entanglement is controlled only by the gap of the system and is otherwise independent of length. In this case,  $M$  can be held fixed and will give a fixed accuracy per electron regardless of system size, and the only scaling of the method derives from the significant Hamiltonian matrix elements. This is why the DMRG and MPS are informally said to give a polynomial complexity solution for correlation in one-dimensional systems. Moving to two-dimensional systems in a similar atomic basis (i.e., where every orbital is equivalent), the amount of entanglement grows *linearly* with the width of the system (for systems that satisfy the area law). This requires  $M$  to depend *exponentially* on width, which is why the DMRG is said to have an exponential scaling with the width of a 2D system.

However, in many molecular calculations, we may be more interested in the computational scaling with respect to increasing the basis size (for fixed energy accuracy and molecular size). In this case, as one is increasing the basis, the orbitals entering are not all equivalent; for example, one might have in the basis valence orbitals and correlating Rydberg like orbitals. These typically have very different occupancies and contributions to the wavefunction and energy. In a large basis, the molecular wavefunction, even in strongly correlated systems, tends to be somewhat “concentrated” around the occupied and valence orbitals. Linear combinations of a few (say  $N_{\text{large}}$ ) large weighted determinants, as is the case for many multireference problems in molecules, can only generate  $\log N_{\text{large}}$  entanglement and are thus easy to capture with a MPS. It can be more effort to capture the manifold of double excitations that, in small molecules, provide the bulk of the dynamical correlation contributions. Assuming that the reordering of orbitals results in the occupied and virtual orbitals being distributed randomly in the DMRG lattice, this generates  $O(\log n_{\text{occ}} n_{\text{virt}})$  entanglement, and thus  $M \propto n_{\text{occ}} n_{\text{virt}}$ . Thus, we expect a worst case scaling of  $O(n_{\text{virt}}^6)$ , for fixed accuracy in the dynamical correlation as the basis size increases, for a fixed molecular size. (As the molecular size increases, the locality results in a lower scaling, as reasoned about in the paragraphs above.)

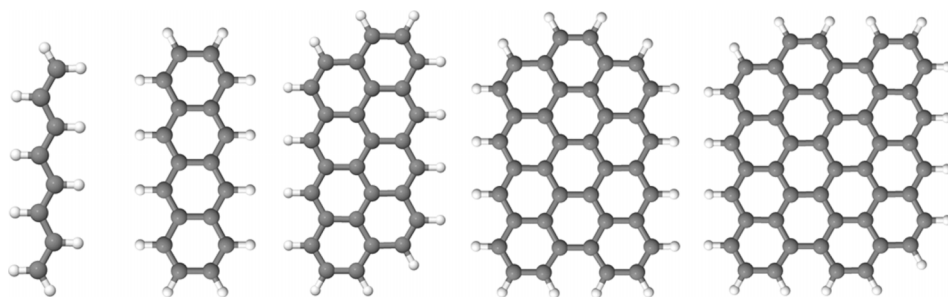


FIG. 2. Arenes studied here. We label them from left to right by their “width” (1-5).

### III. RESULTS

#### A. Arenes

We begin with the arenes, shown in Fig. 2. These conjugated molecules can be viewed as finite strips of graphene of various “widths.” We used octatetraene (8 carbon atoms), tetracene (18 carbon atoms), and three larger analogous polyaromatic hydrocarbons with 28, 38, and 48 carbon atoms, respectively, corresponding to widths 1–5. The range of widths allows us to study the dependence of the DMRG energy convergence on molecular dimensionality. To provide an idealized test system for orbital shapes and orbital ordering, we limit this particular study to the  $\pi$ -electron system in a minimal (STO-3G) basis.<sup>81</sup> We used optimized restricted Hartree-Fock (RHF/STO-3G) geometries. Within the  $\pi$ -active spaces, corresponding to (8e, 8o), (18e, 18o), (28e, 28o), (38e, 38o), and (48e, 48o) for 1–5, respectively, we considered several orbital choices: canonical Hartree-Fock orbitals, split-localized orbitals (that resemble two-centre  $\pi$ -bonding and anti-bonding orbitals), and fully-localized orbitals (orthogonalized  $p$  atomic orbitals). We localized orbitals using the Pipek-Mezey method.<sup>82</sup>

For each orbital choice, the subsequent ordering was determined using both the Fiedler vector and GA optimization schemes described in Sec. II E. These orderings are shown for localized orbitals in Fig. 3. It is known that a good DMRG ordering in 2D sheets is a simple “snake-like” pseudo-1D ordering traversing each row.<sup>83</sup> This is reproduced in the smaller arenes both by the Fiedler and GA orderings. For arene 5, which has more of a square aspect ratio, the Fiedler and GA

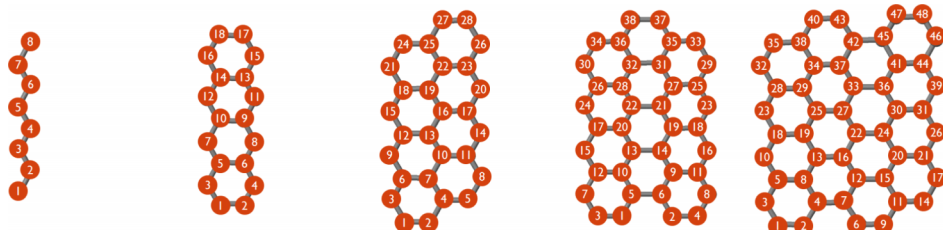
orderings produce orderings that roughly go diagonally across the arene, and the resulting orderings are slightly different.

Fig. 4 shows the energy convergence for arene 3 for different orbital choices and orderings. This convergence is representative of the other arenes. The differences between the Fiedler and GA ordering results are very small; the effect of orbital choice is far more significant. Due to the valence only nature of the active space, we observe fastest convergence with fully-localized orbitals (atomic-like  $p$  orbitals), even though the energy at small  $M$  is relatively poor. In contrast, the canonical Hartree-Fock orbitals give good energies at very small  $M$  (since the Hartree-Fock limit is captured by  $M = 1$ ) but only slowly converge the energy as  $M$  is increased. Split-localized orbitals are a good compromise, capturing the Hartree-Fock limit at  $M = 1$  and giving better energies than the localized orbitals at very small  $M$ , while giving rapid convergence to the exact result that is only slightly slower than the fully localized basis.

Note in Fig. 4 the near-exponential convergence of the energy error with  $M$ , typical of the DMRG. As discussed in Sec. II F, a more careful analysis predicts that the energy error follows  $\ln|\delta E| \sim \kappa(\ln M)^2$ .<sup>5</sup> This behaviour is seen from the plot in Fig. 5. Finally, the linear relationship between the discarded weight and the energy is shown in Fig. 6. This demonstrates the simple extrapolation of the energy to zero discarded weight.

In Fig. 7, we study the DMRG energy convergence as a function of the arene “width.” As expected, the  $M$  required for a given accuracy grows roughly exponentially with width, reflecting the exponentially increasing set of fluctuations across any “boundary” that cuts across the system. However even for

Fiedler



Genetic algorithm

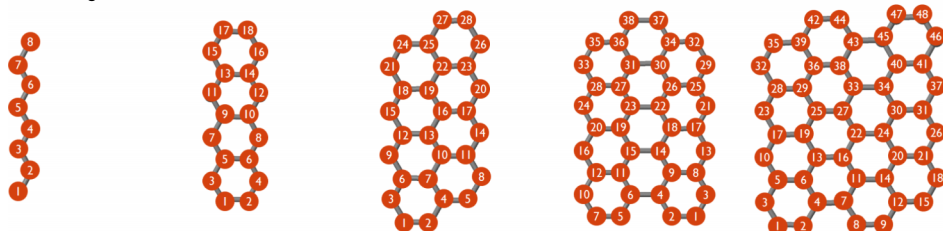


FIG. 3. Fiedler and GA orderings for the arenes 1–5 using localized orbitals.

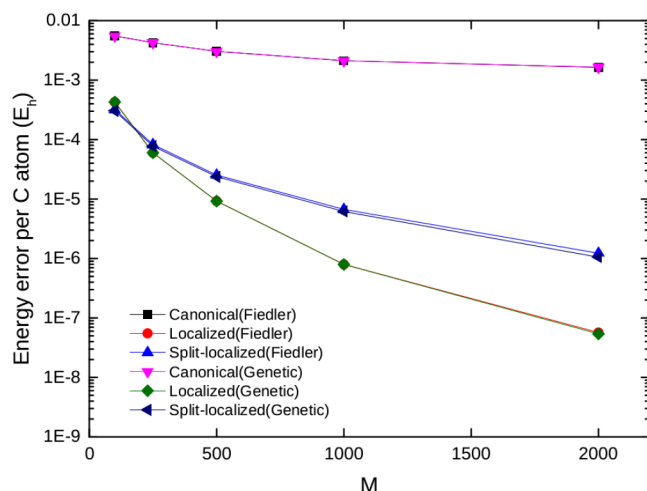


FIG. 4. Error in the energy per C atom for arene **3** ( $\pi$  active space, 28 electrons in 28 orbitals, STO-3G basis) for different orbital choices and orderings.

the largest arene, we see that it is possible to converge the active space energy per carbon atom in 48 orbitals nearly exactly (to 0.01  $mE_h$ ) with an accessible  $M < 100\,000$ . (As a technical note, it should be noted here that because we use  $S^2$  symmetry in our DMRG code, our  $M$  values refer to the number of multiplets retained. The above would thus correspond to  $M = 200\,000 - 300\,000$  states in a code with only  $S_z$  symmetry.)

## B. Dimers

In our next test case, we consider two diatomics,  $C_2$  and  $Cr_2$ . Diatomics provide a contrast to the  $\pi$ -electron system of the arenes as there is little meaningful spatial locality in such a small molecule. Further here, we carry out calculations with all electrons and larger basis sets. This provides a wide range of energies and weights in the wavefunction, unlike in the arenes, and allow us to explore the use of the DMRG for dynamical correlation.

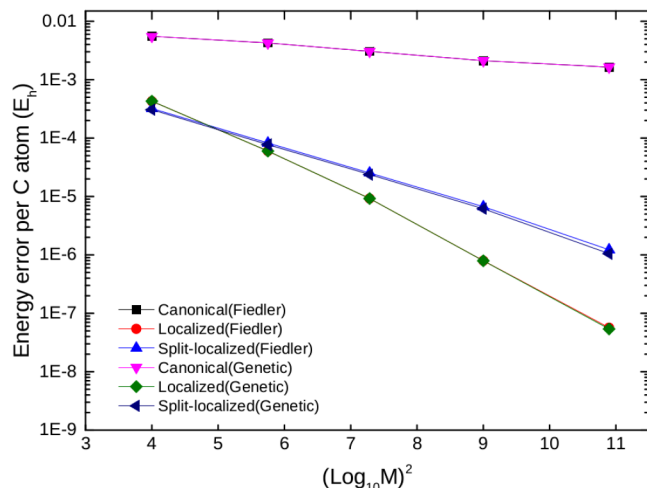


FIG. 5. Energy error per C atom for arene **3** ( $\pi$  active space, 28 electrons in 28 orbitals, STO-3G basis) as a function of  $(\log_{10} M)^2$ .

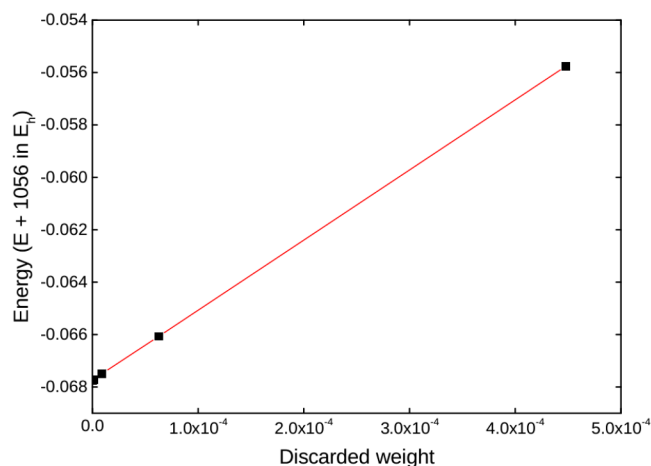


FIG. 6. Linear relationship between the DMRG energy and discarded weight in calculation of arene **3**, using localized orbitals and Fiedler ordering.

We start with  $C_2$ . We use the equilibrium bond distance of 1.24253 Å (the same as in the initiator full configuration interaction quantum Monte Carlo (*i*-FCIQMC) calculations in Ref. 84) and perform calculations in Dunning's cc-pVDZ, cc-pVTZ basis, and cc-pVQZ basis sets. We carry out both all electron (AE) calculations (corresponding to  $(12e, 28o)$ ,  $(12e, 60o)$ , and  $(12e, 110o)$  spaces for cc-pVDZ, cc-pVTZ, and cc-pVQZ, respectively). We first examine the effect of different orderings. There is no natural ordering in a diatomic molecule, unlike for localized orbitals in the arenes. However, we can still reorder the orbitals using the Fiedler and genetic algorithms as for the arenes. The corresponding energies are shown in Table II. Also shown in Table II are energies obtained using the reverse sweep schedule where  $M$  is decreased from its maximum value of 5000 to 500. These are much better at smaller  $M$  than the energies obtained in the standard sweep, evidence that the warmup sweep is relatively poor in the standard schedule and also that additional important quantum numbers are detected during the sweeps at larger  $M$  that were not picked at smaller  $M$ .

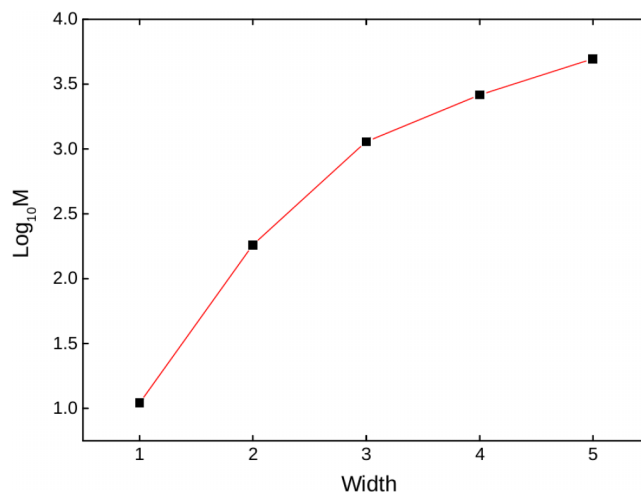


FIG. 7.  $M$  to obtain a set accuracy of 0.01  $mE_h$  in the energy per carbon atom, as a function of arene width. Calculations use Fiedler ordering and localized orbitals.

TABLE II. DMRG energies ( $E + 75$  in  $E_h$ ) using Energy Fiedler and GA orderings for the ground-state of  $C_2$  (all-electron, cc-pVQZ). Reverse indicates energies from a reverse schedule. Note that the reverse schedule GA energy at  $M = 5000$  is below the default schedule GA energy at  $M = 5000$ . This is due to the noise setting in the default schedule, which is slightly too large for this molecule at large  $M$ .

M	Energy E ( $E_h$ )	Genetic E ( $E_h$ )	Fiedler E ( $E_h$ )	Reverse genetic E ( $E_h$ )
500	-0.846 422	-0.839 705	-0.838 761	-0.853 141
1000	-0.851 769	-0.847 877	-0.849 565	-0.855 803
2000	-0.854 534	-0.853 348	-0.853 724	-0.856 866
3000	-0.855 732	-0.855 104	-0.855 446	-0.857 135
4000	-0.856 425	-0.855 890	-0.856 238	-0.857 231
5000	-0.856 762	-0.856 386	-0.856 645	-0.857 256
Extrapolated				-0.857 28

Note also that the energy at  $M = 5000$  for the reverse schedule (GA ordering) is about  $0.9 mE_h$  lower than the  $M = 5000$  energy (GA ordering) in the default schedule. This difference is entirely due to the noise in the default sweeps, which is turned off for the reverse schedule. (Recall that noise is used to ensure more robust convergence in the early sweeps.) We see that in this molecule, the default setting for the noise in Table I should be decreased at larger  $M$  to achieve faster convergence. This is unsurprising as the sweep parameters have been chosen for general utility rather than for specific molecules and illustrates the gains in convergence that could in principle be made by optimizing the sweep settings on a per molecule basis. However, the detrimental effect of the non-optimal noise remains quite small and is unimportant on the chemical scale of  $1 mE_h$ .

For the GA ordering, we further give converged energies (accurate to better than  $0.05 mE_h$ ) for the  $1s$  frozen core (FC) calculations in the same bases. This allows direct comparison (shown in Table III) to  $i$ -FCIQMC calculations in the literature.<sup>84</sup> The  $i$ -FCIQMC energies reported in Ref. 84 for the TZ and QZ bases are in rough agreement with the DMRG energies. However, the variational DMRG energies are clearly lower and have significantly smaller error bounds. The difference between the DMRG and  $i$ -FCIQMC energy corresponds to the remaining initiator error in the  $i$ -FCIQMC calculation, as we have also found in other comparisons between the two techniques.<sup>42</sup>

In  $C_2$ , we are using basis sets with many virtual orbitals, thus much of the correlation we describe is dynamic. This,

TABLE III. DMRG energies of  $C_2$  (all-electron and frozen core, in cc-pVDZ (DZ), cc-pVTZ (TZ), and cc-pVQZ (QZ) bases). Energy ( $E + 75$ ) in  $E_h$ .  $i$ -FCIQMC energies TZ, QZ, from Ref. 84 shown for comparison, estimated statistical error in brackets.

M	DZ	TZ	QZ
500	-0.731 449	-0.807 296	-0.853 141
1000	-0.731 856	-0.808 662	-0.855 803
2000	-0.731 945	-0.809 123	-0.856 866
4000	-0.731 958	-0.809 260	-0.857 231
6000	...	-0.809 285	...
(FC) DMRG	-0.728 556	-0.785 054	-0.802 671
(FC) $i$ -FCIQMC	-0.7287(8)	-0.7849(3)	-0.8025(1)

together with the lack of spatial locality in the system, means (as described in Sec. II F) that the  $M$  required for a given accuracy in a molecule should scale roughly as  $n_{occ}n_{virt}$  to capture the dominant doubles excitations, i.e., the required  $M$  should scale linearly with the number of electrons and basis size. The  $M$  required to achieve an accuracy of  $0.01 mE_h$  is plotted against  $n_{occ}n_{virt}$  in Fig. 8. We observe a good linear-fit to  $n_{occ}n_{virt}$ , confirming that the entanglement is dominated by the double excitations, and that correlation in this molecule is mainly dynamic in character.

We next consider the  $Cr_2$  dimer.  $Cr_2$ , with its formal hexuple bond, is a long-standing favourite of quantum chemistry, due to the difficulty in adequately describing the shelf-like nature of its potential energy curve.<sup>85</sup> This unusual curve arises from a combination of non-dynamic (strong) correlation and dynamic correlation in large basis sets.  $Cr_2$  has been studied previously with DMRG methods, which (in combination with perturbation theory) give a very accurate description of the  $Cr_2$  potential energy curve.<sup>37</sup>

Here, for benchmark purposes, we use smaller basis sets where it is possible to converge the DMRG energy “exactly,” i.e., to beyond chemical accuracy. We consider  $Cr_2$  at the Cr-Cr distance of  $1.5 \text{ \AA}$  with the Ahlrichs’ SV basis,<sup>86</sup> both

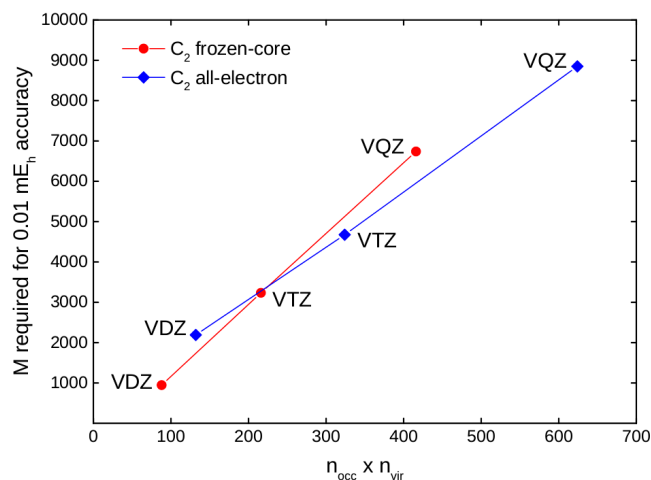


FIG. 8.  $M$  required to obtain  $0.01 mE_h$  accuracy. This shows a linear dependence on  $n_{occ}n_{virt}$  as is expected for a system dominated by dynamic correlation.

TABLE IV. DMRG energies ( $E + 2086$  in  $E_h$ ) of  $\text{Cr}_2$  in a SV basis. All electron ( $48e$ ,  $42o$ ) and active space ( $24e$ ,  $30o$ ) results are shown, as obtained from the default and reverse schedule. Extrapolated DMRG energies from Fig. 9 and CCSD(T), CCSDTQ energies shown for comparison.

M	(24e, 30o)				(48e, 42o)			
	Default		Reverse		Default		Reverse	
	DW	E ( $E_h$ )	DW	E ( $E_h$ )	DW	E ( $E_h$ )	DW	E ( $E_h$ )
500	$1.02 \times 10^{-4}$	-0.414 935	$9.47 \times 10^{-5}$	-0.416 106	$2.73 \times 10^{-4}$	-0.177 377	$4.78 \times 10^{-4}$	-0.421 928
1000	$4.59 \times 10^{-5}$	-0.418 219	$4.36 \times 10^{-5}$	-0.418 693	$2.18 \times 10^{-4}$	-0.426 115	$2.89 \times 10^{-4}$	-0.432 057
2000	$2.34 \times 10^{-5}$	-0.419 707	$1.87 \times 10^{-5}$	-0.419 986	$1.34 \times 10^{-4}$	-0.436 820	$1.51 \times 10^{-4}$	-0.438 127
3000	$1.38 \times 10^{-5}$	-0.420 246	$1.10 \times 10^{-5}$	-0.420 386	$9.24 \times 10^{-5}$	-0.439 855	$9.59 \times 10^{-5}$	-0.440 324
4000	$1.10 \times 10^{-5}$	-0.420 480	$7.78 \times 10^{-6}$	-0.420 570	$7.14 \times 10^{-5}$	-0.441 211	$6.92 \times 10^{-5}$	-0.441 458
5000	$9.26 \times 10^{-6}$	-0.420 609	$5.65 \times 10^{-6}$	-0.420 672	$5.55 \times 10^{-5}$	-0.441 927	$5.41 \times 10^{-5}$	-0.442 146
6000	$6.69 \times 10^{-6}$	-0.420 686	$4.44 \times 10^{-6}$	-0.420 735	$4.50 \times 10^{-5}$	-0.442 439	$4.37 \times 10^{-5}$	-0.442 607
7000	$5.58 \times 10^{-6}$	-0.420 745	$3.44 \times 10^{-6}$	-0.420 776	$3.61 \times 10^{-5}$	-0.442 792	$3.66 \times 10^{-5}$	-0.442 933
8000	$4.55 \times 10^{-6}$	-0.420 774	$2.69 \times 10^{-6}$	-0.420 780	$3.15 \times 10^{-5}$	-0.443 334	$3.08 \times 10^{-5}$	-0.443 173
Extrapolated			0.0	-0.420 948			0.0	-0.444 784
CCSD(T)		-0.398 638				-0.422 229		
CCSDTQ		-0.406 696				-0.430 244		

correlating all electrons (a ( $48e$ ,  $42o$ ) space) as well as a ( $24e$ ,  $30o$ ) active space subset, used in previous DMRG benchmarks.<sup>21,37</sup> We used GA ordering for all calculations.

The DMRG energies and weights are shown in Table IV with CCSD(T) and CCSDTQ calculations for comparison. We see a large difference between the CCSDTQ and DMRG energy indicating the importance of high-order correlations in this complex system. The linear extrapolation of the energies in Table IV is plotted in Figure 9 from which obtain the estimated exact energies. The error bars from the fitting are  $\pm 0.007$  and  $\pm 0.19 \text{ m}E_h$  for the ( $24e$ ,  $30o$ ) and ( $48e$ ,  $42o$ ) active spaces, respectively, probably an underestimate. Alternatively, a conservative rule of thumb used in some DMRG extrapolations<sup>87</sup> is to assign the error bar as 1/5 the extrapolation energy, which comes to  $\pm 0.034$  and  $\pm 0.32 \text{ m}E_h$  for the two active spaces, respectively, probably overestimates. Either way, the energies are very accurate. Further, with the maximum  $M = 8000$ , even the *unextrapolated* energies are within  $0.2 \text{ m}E_h$  of the estimated exact result in the ( $24e$ ,  $30o$ ) active space, and within  $1.6 \text{ m}E_h$  of the estimated exact result in the ( $48e$ ,  $42o$ ) active space.

### C. Butadiene

We next consider a larger polyatomic molecule, butadiene. Butadiene has been of continuing interest to quantum chemists, due to the correlation effects in the  $\pi$ -conjugated system, and has recently been the target of “exact” methods including *i*-FCIQMC<sup>88</sup> and composite high-order coupled cluster approaches.<sup>89</sup>

Unlike in our earlier model arene calculations, here we correlate all electrons except for a frozen  $1s$  core. We carry out calculations in the same ANO-L-pVDZ basis<sup>106</sup> as used in Ref. 88. This generates a space of ( $22e$ ,  $82o$ ). This is representative of a large “all-electron” exact calculation with the DMRG. We considered both restricted HF canonical orbital as

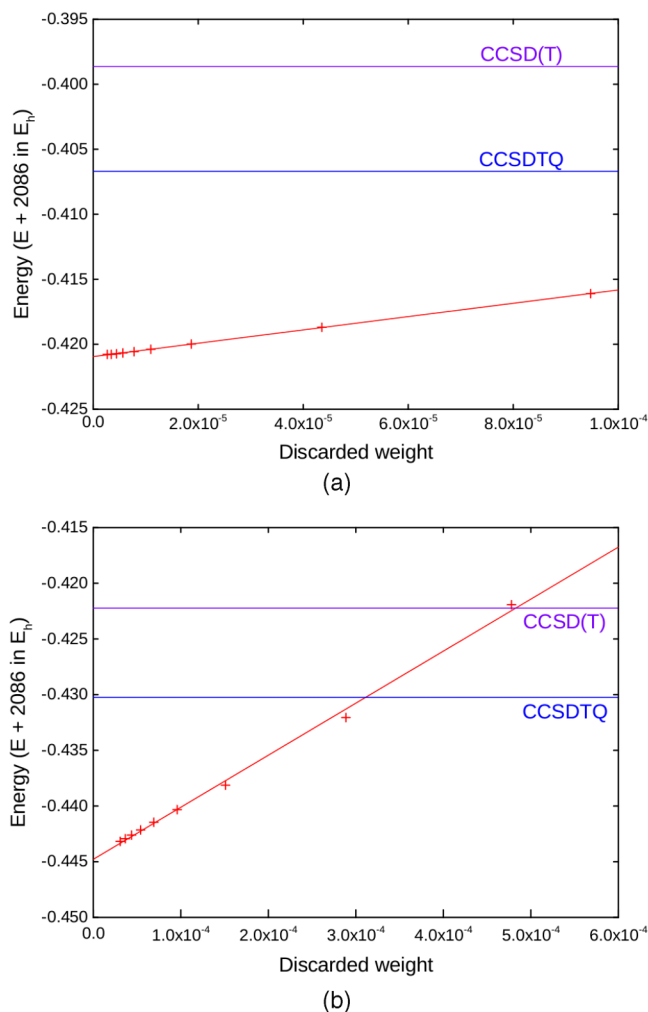


FIG. 9. Extrapolations of ( $24e$ ,  $30o$ ) (a) and ( $48e$ ,  $42o$ ) (b)  $\text{Cr}_2$  DMRG energies using the reverse schedule and GA ordering, and comparisons to CCSD(T) and CCSDTQ energies.

TABLE V. Comparison of the DMRG sweep energies ( $E + 155$  in  $E_h$ ) from the reverse schedule for butadiene (22 electrons in 82 orbitals, ANO-L-pVDZ basis) using canonical, MP2 natural orbitals, and split-localized MP2 natural orbitals.

M	Type of orbital			
	Canonical	Canonical Split-localized	MP2 natural	MP2 nature Split-localized
500	-0.500 593	-0.550 526	-0.505 492	-0.552 366
1000	-0.516 424	-0.554 396	-0.523 147	-0.555 348
2000	-0.531 067	-0.556 172	-0.537 489	-0.556 669

well as MP2 natural orbitals, at the geometry given in Ref. 88. Orbitals were ordered by the GA.

Energies using the different orbital choices up to a modest  $M = 2000$  are given in Table V. We observe that the natural orbitals perform better than the Hartree-Fock orbitals, and the split-localized orbitals perform better than the canonical orbitals. Clearly for a molecule of this size locality is already very important. We thus use only the split-localized natural orbitals for the larger  $M$  calculations on this molecule.

The DMRG energies using the split-localized natural orbitals are given in Table VI together with comparison coupled cluster and *i*-FCIQMC energies. We find that we already begin to surpass CCSDT accuracy by about  $M = 2000$ . By  $M = 6000$ , we are more than 1  $mE_h$  below the CCSDT energy, giving an estimate of quadruple and higher order effects. These calculations, however, required considerable resources, especially because there was no point-group symmetry in the split-localized basis: a single  $M = 3000$  sweep in  $C_1$  symmetry required 25 h on 42 cores.

The total energy of butadiene in the ANO-L-pVDZ basis was very recently the subject of a large scale *i*-FCIQMC study. We note that our final DMRG energy is more than 8  $mE_h$  below the *i*-FCIQMC energy. Since the DMRG energy is variational, this means that the *i*-FCIQMC energy is too high, due to initiator bias. The results reported by Daday *et al.* in Ref. 88 used  $1 \times 10^9$  walkers and were limited by the available memory. Clearly, larger numbers of walkers are necessary to remove the initiator bias.

TABLE VI. DMRG sweep energies ( $E + 155$  in  $E_h$ ) of butadiene (default schedule) in an ANO-L-pVDZ basis compared to various other methods (*i*-FCIQMC results are from Ref. 88). Energy in  $E_h$ .

M	Energy
500	-0.550 134
1000	-0.554 182
2000	-0.555 899
3000	-0.556 543
4000	-0.556 874
5000	-0.557 050
6000	-0.557 178
CCSD(T)	-0.555 002
CCSDT	-0.555 959
<i>i</i> -FCIQMC	-0.5491(4)

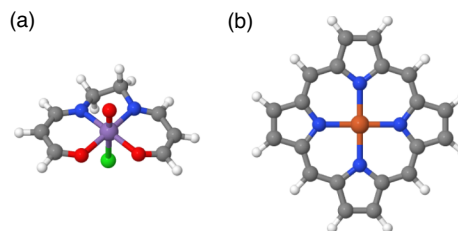


FIG. 10. oxo-Mn(II)salen and Fe(II)-porphine.

## D. Organometallics

We now turn to some more chemically complex systems, as exemplified by organometallic complexes. As a representative example, we have chosen the oxo-Mn(salen) system (Fig. 10), an analogue of Jacobsen's catalyst, considered by Ivanic *et al.* using complete active space self-consistent field (CASSCF)<sup>90,91</sup> and a number of subsequent studies including a recent DMRG-CASSCF study.<sup>51</sup> We also consider Fe(II)-porphine, the prototype for biological metalloporphyrins (Fig. 10). Both these molecules are too large to compute all-electron, full basis, DMRG calculations converged to the FCI limit. We thus restrict ourselves here to active-space DMRG calculations. In more realistic studies, such active-space calculations would be augmented by a further more approximate dynamical correlation treatment using perturbation theory,<sup>8,92,93</sup> configuration interaction,<sup>55,94</sup> or canonical transformation theory.<sup>38</sup>

We first consider oxo-Mn(salen). This molecule is used as a simple model for Jacobsen's epoxidation catalyst.<sup>95,96</sup> The ordering of the lowest spin states is considered important as different reaction paths have been posited depending on the spin state.<sup>97</sup> In our DMRG calculations, we used the 6-31G(d) basis set<sup>81,98,99</sup> and computed the restricted open-shell Hartree-Fock (ROHF) orbitals for the triplet state in the  $C_1$  point group symmetry. Further, the active space was defined by ROHF molecular orbitals with the following dominant atomic orbital characters: (1) five Mn 3*d* orbitals, (2) 2*p<sub>z</sub>* of the equatorial C, N and O atoms, giving 10  $\pi$  orbitals, (3) 2*p<sub>x</sub>* and 2*p<sub>y</sub>* of the equatorial O and N atoms for the Mn-N and Mn-O  $\sigma$  bonds, (4) 2*p<sub>x</sub>*, 2*p<sub>y</sub>* and 2*p<sub>z</sub>* of the axial Cl atoms, (5) an axial O 2*p<sub>z</sub>*, and combination of the axial O 2*p<sub>x</sub>*, 2*p<sub>y</sub>* orbitals, responsible for the Mn-O  $\sigma$  and  $\pi$  bonds. The doubly, singly, and unoccupied orbitals were then separately split-localized, and we carried out DMRG(-CI) calculations in this active-space using up to  $M = 2000$  and GA ordered orbitals. (To keep the benchmarks simple and reproducible we do not consider DMRG-CASSCF calculations here.) The results are shown in Table VII. We find that with only 24 orbitals, the active space energy can be converged to tens of microHartrees with a modest  $M = 2000$ , a rather modest calculation with this number of orbitals. In this active space, we find that the singlet is lowest in energy, with a very small singlet-triplet energy gap of 0.6  $mE_h$ . This is reminiscent of the nearly degenerate ground state picture found in previous theoretical works,<sup>51,97</sup> although a quantitative ordering requires augmentation by dynamic correlation.

We next discuss the Fe(II)-porphine system. This is another complex where the correct spin-ordering of the states remains

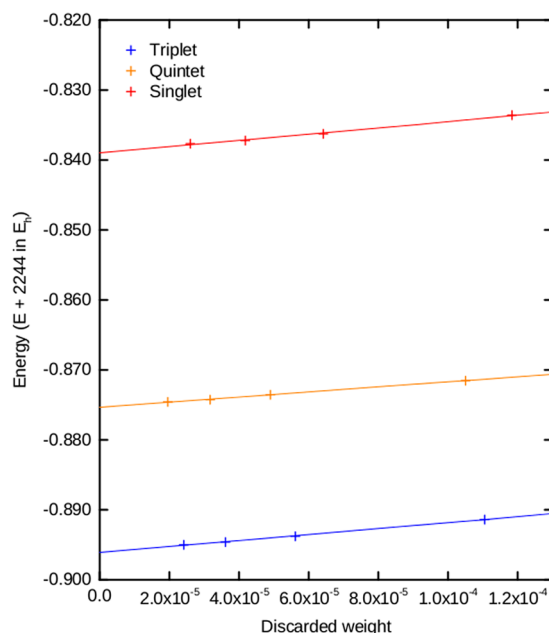
TABLE VII. DMRG energies ( $E + 2251$  in  $E_h$ ) and gaps for oxo-Mn(salen) using a  $(32e, 24o)$  active space and 6-31g(d) basis.

M	Singlet		Triplet	
	DW	Energy	DW	Energy
500	$1.33 \times 10^{-5}$	-0.304 421	$1.46 \times 10^{-5}$	-0.303 777
1000	$4.51 \times 10^{-6}$	-0.304 648	$5.20 \times 10^{-6}$	-0.304 045
2000	$1.19 \times 10^{-6}$	-0.304 712	$1.40 \times 10^{-6}$	-0.304 128

somewhat uncertain.<sup>92,100,101</sup> Here, we have considered a large active space, consisting of all Fe  $3d$  and  $4d$  orbitals (the additional  $4d$  shell gives the important double-shell correlations) and the full set of ligand  $\sigma$  and  $\pi$  orbitals. In the  $D_{2h}$  symmetry, we started from the quintet ROHF orbitals at the triplet geometry described by Groenhof *et al.*<sup>102</sup> with the cc-pVDZ basis set.<sup>103–105</sup> We classified orbitals by their dominant atomic or bond character and defined an active space of  $(44e, 44o)$ , containing molecular orbitals with the following character: (1) Fe  $3d$  and  $4d$  orbitals (10 orbitals), (2)  $2p_z$  orbitals of C and N atoms giving 24  $\pi$  orbitals, (3)  $\sigma$  bonds between Fe  $4p_x, 4p_y$ , and N  $2p_x$  and  $2p_y$  orbitals (10 orbitals). Finally, all orbitals were split-localized and reordered using the genetic algorithm.

Table VIII presents the energies obtained using the reverse schedule. The corresponding extrapolations are shown in Fig. 11. We see that although convergence is not rapid in this very large active space, it is still possible to compute all the electronic states to within about  $1.5 mE_h$  of the estimated exact result at  $M = 4000$ . The difficulty of the calculation is similar to that of the all-electron  $\text{Cr}_2$  calculation, which correlated 48 electrons in 40 orbitals. Within this active space, we find that the energy order is triplet < quintet < singlet. The largest calculations (in  $C_1$  symmetry due to split localization) with  $M = 4000$ , e.g., for the singlet state, took 15 h per sweep with 40 cores, and are representative of expensive active space DMRG calculations.

Many interesting organometallic complexes involve multiplet open shell transition metal species. What are the prospects of extending DMRG calculations to such systems? The largest DMRG calculations to date involve 4 open shell transition metal centers with bridging ligands (such as in the  $\text{Mn}_4\text{Ca}$  cluster of the oxygen evolving complex in Ref. 52, or the [4Fe-4S] biological redox cofactor in Ref. 53). While such systems are at the frontier of DMRG calculations today, as Ref. 53 shows, calculations with  $M$  up to 7500, together with

FIG. 11. Extrapolation of the triplet and quintet ground-states of Fe(II)-porphine in a  $(44e, 44o)$  active space.

extrapolation, yield energies accurate to about  $1 mE_h$ , within chemical accuracy.

#### IV. CONCLUSIONS

In this work, we presented an overview of the theory and practice of the *ab-initio* density matrix renormalization group. In modern implementations, such as in the BLOCK code, the *ab-initio* DMRG is implemented in a black-box manner so that it can be used in a fashion similar to other quantum chemistry methods. The user needs only to specify the desired target number of renormalized states, and the choice of active space and orbitals. With these specifications, all other aspects of the DMRG calculation, such as orbital ordering, perturbative noise, sweep schedule, convergence thresholds, and extrapolations, can be handled automatically by the program.

We have examined the behaviour of the DMRG from the user perspective in a variety of different molecular settings: arenes, diatomics, polyatomics, and organometallics. We summarize our recommendations thus, which are as follows:

- For most systems, it is important to balance spatial and energy locality in the orbitals by using split-localized orbitals.

TABLE VIII. DMRG energies ( $E + 2244$  in  $E_h$ ) for Fe(II)-porphine,  $(44e, 44o)$  active space, quintet and triplet, using GA ordering and cc-pVDZ basis.

M	Singlet		Triplet		Quintet	
	DW	E ( $E_h$ )	DW	E ( $E_h$ )	DW	E ( $E_h$ )
1000	$1.18 \times 10^{-4}$	-0.833 592	$1.11 \times 10^{-4}$	-0.891 395	$1.05 \times 10^{-4}$	-0.871 543
2000	$6.42 \times 10^{-5}$	-0.836 259	$5.62 \times 10^{-5}$	-0.893 772	$4.90 \times 10^{-5}$	-0.873 551
3000	$4.18 \times 10^{-5}$	-0.837 218	$3.61 \times 10^{-5}$	-0.894 605	$3.17 \times 10^{-5}$	-0.874 250
4000	$2.60 \times 10^{-5}$	-0.837 683	$2.42 \times 10^{-5}$	-0.895 003	$1.95 \times 10^{-5}$	-0.874 581
Extrapolated		-0.839 108		-0.896 092		-0.875 330

- Fiedler vector orbital orderings (default in the BLOCK code) perform well.
- For large molecules, we can reason about the number of renormalized states  $M$  required to achieve a given accuracy in terms of the effective “width” of the system.
- For dynamic correlation,  $M$  scales roughly linearly with the number of virtuals (basis size) for fixed accuracy.
- All-electron basis calculations with  $n_{elec}n_{orb}$  of around 2000 (e.g., 20 electrons in 100 orbitals, or 40 electrons in 50 orbitals), converged to chemical accuracy or better and without symmetry, can be considered accessible with cluster computational resources with 50 or so computational cores, with sufficient memory and disk. (Of course, larger calculations are possible with more resources!)

We have focused here on relatively simple benchmarks, but of course, many more applications of the DMRG to different systems can be envisaged. We have not explicitly discussed here, for example, excited states, or large polynuclear bioinorganic clusters, although these have been significant targets of study with the DMRG in the literature.<sup>42,50,52,53</sup> With the simple recommendations above, we believe such systems can now be studied not only by the specialist, but by the general user.

## ACKNOWLEDGMENTS

This work was primarily supported by NSF Grant No. OCI-1265278. Additional funding was provided by NSF Grant No. CHE-1265277. The authors acknowledge Helen van Aggelen for useful discussions regarding plot designs, Peter Knowles and Andy May for their help with implementing BLOCK within MOLPRO, Kantharuban Sivalingam for his help with implementing BLOCK within ORCA, and Yihan Shao, for his help with implementing BLOCK within Q-CHEM.

<sup>1</sup>S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).

<sup>2</sup>S. R. White, *Phys. Rev. B* **48**, 10345 (1993).

<sup>3</sup>S. R. White and R. L. Martin, *J. Chem. Phys.* **110**, 4127 (1999).

<sup>4</sup>A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, *J. Chem. Phys.* **115**, 6815 (2001).

<sup>5</sup>G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).

<sup>6</sup>Ö. Legeza, J. Röder, and B. Hess, *Phys. Rev. B* **67**, 125114 (2003).

<sup>7</sup>D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 014107 (2008).

<sup>8</sup>Y. Kurashige and T. Yanai, *J. Chem. Phys.* **130**, 234114 (2009).

<sup>9</sup>H.-G. Luo, M.-P. Qin, and T. Xiang, *Phys. Rev. B* **81**, 235129 (2010).

<sup>10</sup>S. Wouters, W. Poelmans, P. W. Ayers, and D. Van Neck, *Comput. Phys. Commun.* **185**, 1501 (2014).

<sup>11</sup>K. H. Marti and M. Reiher, *Z. Phys. Chem.* **224**, 583 (2010).

<sup>12</sup>Ö. Legeza, T. Rohwedder, R. Schneider, and S. Szalay, “Tensor Product Approximation (DMRG) and Coupled Cluster Method in Quantum Chemistry,” in *Many-Electron Approaches in Physics, Chemistry and Mathematics*, edited by V. Bach and L. Delle Site (Springer International Publishing, 2014), pp. 53–76.

<sup>13</sup>Y. Kurashige, *Mol. Phys.* **112**, 1485 (2014).

<sup>14</sup>G. K.-L. Chan and D. Zgid, *Annu. Rep. Comput. Chem.* **5**, 149 (2009).

<sup>15</sup>G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 (2011).

<sup>16</sup>Z. Shuai, J. Bredas, S. Pati, and S. Ramasesha, *Proc. SPIE* **3145**, 293 (1997).

<sup>17</sup>D. Yaron, E. Moore, Z. Shuai, and J. Bredas, *J. Chem. Phys.* **108**, 7451 (1998).

<sup>18</sup>G. Fano, F. Ortolani, and L. Ziosi, *J. Chem. Phys.* **108**, 9246 (1998).

<sup>19</sup>T. Xiang, *Phys. Rev. B* **53**, R10445 (1996).

<sup>20</sup>G. K.-L. Chan, *J. Chem. Phys.* **120**, 3172 (2004).

<sup>21</sup>S. Sharma and G. K.-L. Chan, *J. Chem. Phys.* **136**, 124121 (2012).

<sup>22</sup>S. Wouters, P. A. Limacher, D. Van Neck, and P. W. Ayers, *J. Chem. Phys.* **136**, 134110 (2012).

<sup>23</sup>S. Sharma, preprint [arXiv:1408.5868](https://arxiv.org/abs/1408.5868) (2014).

<sup>24</sup>Ö. Legeza and J. Sólyom, *Phys. Rev. B* **68**, 195116 (2003).

<sup>25</sup>G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher, *Phys. Rev. A* **83**, 012508 (2011).

<sup>26</sup>J. Rissler, R. M. Noack, and S. R. White, *Chem. Phys.* **323**, 519 (2006).

<sup>27</sup>D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **128**, 144117 (2008).

<sup>28</sup>D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 144116 (2008).

<sup>29</sup>D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 144115 (2008).

<sup>30</sup>G. Moritz and M. Reiher, *J. Chem. Phys.* **124**, 034103 (2006).

<sup>31</sup>G. Moritz and M. Reiher, *J. Chem. Phys.* **126**, 244109 (2007).

<sup>32</sup>S. R. White, *Phys. Rev. B* **72**, 180403 (2005).

<sup>33</sup>Ö. Legeza and G. Fáth, *Phys. Rev. B* **53**, 14349 (1996).

<sup>34</sup>G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **118**, 8551 (2003).

<sup>35</sup>J. J. Dorando, J. Hachmann, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 184111 (2009).

<sup>36</sup>N. Nakatani, S. Wouters, D. Van Neck, and G. K.-L. Chan, *J. Chem. Phys.* **140**, 024108 (2014).

<sup>37</sup>Y. Kurashige and T. Yanai, *J. Chem. Phys.* **135**, 094104 (2011).

<sup>38</sup>T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, *J. Chem. Phys.* **132**, 024105 (2010).

<sup>39</sup>M. Saitow, Y. Kurashige, and T. Yanai, *J. Chem. Phys.* **139**, 044118 (2013).

<sup>40</sup>S. Knecht, Ö. Legeza, and M. Reiher, *J. Chem. Phys.* **140**, 041101 (2014).

<sup>41</sup>G. K.-L. Chan, M. Kállay, and J. Gauss, *J. Chem. Phys.* **121**, 6110 (2004).

<sup>42</sup>S. Sharma, T. Yanai, G. H. Booth, C. J. Umrigar, and G. K.-L. Chan, *J. Chem. Phys.* **140**, 104112 (2014).

<sup>43</sup>T. Nishino and K. Okunishi, *J. Phys. Soc. Jpn.* **64**, 4084 (1995).

<sup>44</sup>S. Rommer and S. Östlund, *Phys. Rev. B* **55**, 2164 (1997).

<sup>45</sup>J. Dukelsky, M. Martin-Delgado, T. Nishino, and G. Sierra, *Europhys. Lett.* **43**, 457 (1998).

<sup>46</sup>U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).

<sup>47</sup>K. H. Marti, I. M. Ondík, G. Moritz, and M. Reiher, *J. Chem. Phys.* **128**, 014104 (2008).

<sup>48</sup>E. Neuscamman, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **132**, 024106 (2010).

<sup>49</sup>W. Mizukami, Y. Kurashige, and T. Yanai, *J. Chem. Theory Comput.* **9**, 401 (2013).

<sup>50</sup>F. Liu, Y. Kurashige, T. Yanai, and K. Morokuma, *J. Chem. Theory Comput.* **9**, 4462 (2013).

<sup>51</sup>S. Wouters, T. Bogaerts, P. Van Der Voort, V. Van Speybroeck, and D. Van Neck, *J. Chem. Phys.* **140**, 241103 (2014).

<sup>52</sup>Y. Kurashige, G. K.-L. Chan, and T. Yanai, *Nat. Chem.* **5**, 660 (2013).

<sup>53</sup>S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan, *Nat. Chem.* **6**, 927 (2014).

<sup>54</sup>J. Hachmann, W. Cardoen, and G. K.-L. Chan, *J. Chem. Phys.* **125**, 144101 (2006).

<sup>55</sup>J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, *J. Chem. Phys.* **127**, 134309 (2007).

<sup>56</sup>E. Fertitta, B. Paulus, G. Barcza, and Ö. Legeza, *Phys. Rev. B* **90**, 245129 (2014).

<sup>57</sup>K. Boguslawski, P. Tecmer, Ö. Legeza, and M. Reiher, *J. Phys. Chem. Lett.* **3**, 3129 (2012).

<sup>58</sup>K. Boguslawski, P. Tecmer, G. Barcza, Ö. Legeza, and M. Reiher, *J. Chem. Theory Comput.* **9**, 2959 (2013).

<sup>59</sup>M. Mottet, P. Tecmer, K. Boguslawski, Ö. Legeza, and M. Reiher, *Phys. Chem. Chem. Phys.* **16**, 8872 (2014).

<sup>60</sup>P. Tecmer, K. Boguslawski, Ö. Legeza, and M. Reiher, *Phys. Chem. Chem. Phys.* **16**, 719 (2014).

<sup>61</sup>H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs 2012, see <http://www.molpro.net>.

<sup>62</sup>F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 73 (2012).

<sup>63</sup>Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng *et al.*, *Mol. Phys.* **113**, 184 (2015).

<sup>64</sup>BLOCK homepage, <http://www.princeton.edu/chemistry/chan/software/dmrg/>.

<sup>65</sup>K. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975).

<sup>66</sup>G. K.-L. Chan, *Phys. Chem. Chem. Phys.* **10**, 3454 (2008).

- <sup>67</sup>J. M. Kinder, C. C. Ralph, and G. K.-L. Chan, "Analytic Time Evolution, Random Phase Approximation, and Green Functions for Matrix Product States," in *Quantum Information and Computation for Chemistry*, Advances in Chemical Physics Vol. 154, edited by S. Kais (John Wiley & Sons, Inc., Hoboken, NJ, 2014), pp. 179–192.
- <sup>68</sup>N. Nakatani and G. K.-L. Chan, *J. Chem. Phys.* **138**, 134113 (2013).
- <sup>69</sup>S. Wouters, B. Verstichel, D. Van Neck, and G. K.-L. Chan, *Phys. Rev. B* **90**, 045104 (2014).
- <sup>70</sup>E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975).
- <sup>71</sup>M. Fiedler, *Czech. Math. J.* **23**, 298 (1973), available at <http://dml.cz/dmlcz/101356>.
- <sup>72</sup>M. Fiedler, *Czech. Math. J.* **25**, 607 (1975), available at <http://dml.cz/dmlcz/101356>.
- <sup>73</sup>M. Juvan and B. Mohar, *Discrete Appl. Math.* **36**, 153 (1992).
- <sup>74</sup>M. W. Newman, Ph.D. thesis, University of Manitoba, 2000.
- <sup>75</sup>T. Biyikoglu, J. Leydold, and P. F. Stadler, *Laplacian Eigenvectors of Graphs: Perron–Frobenius and Faber–Krahn Type Theorems*, Lecture Notes in Mathematics (Springer, 2007).
- <sup>76</sup>D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley Professional, 1989).
- <sup>77</sup>G. Moritz, B. A. Hess, and M. Reiher, *J. Chem. Phys.* **122**, 024107 (2005).
- <sup>78</sup>L. Davis, "Applying adaptive algorithms to epistatic domains," in *IJCAI'85 Proceedings of the 9th International Joint Conference on Artificial Intelligence* (Morgan Kaufmann Publishers Inc., San Francisco, 1985), Vol. 1, pp. 162–164, available at <http://dl.acm.org/citation.cfm?id=1625135.1625164>.
- <sup>79</sup>L. Davis, "Job Shop Scheduling with Genetic Algorithm," in *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, edited by J. Grefenstette (Psychology Press, Taylor & Francis Group, 2014), pp. 136–140.
- <sup>80</sup>G. K.-L. Chan and P. Ayers, *J. Stat. Phys.* **109**, 289 (2002).
- <sup>81</sup>W. J. Hehre, *J. Chem. Phys.* **51**, 2657 (1969).
- <sup>82</sup>J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- <sup>83</sup>E. Stoudenmire and S. R. White, *Annu. Rev. Condens. Matter Phys.* **3**, 111 (2012).
- <sup>84</sup>G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **135**, 084104 (2011).
- <sup>85</sup>B. O. Roos, *Collect. Czech. Chem. Commun.* **68**, 265 (2003).
- <sup>86</sup>A. Schäfer, H. Horn, and R. Ahlrichs, *J. Chem. Phys.* **97**, 2571 (1992).
- <sup>87</sup>S. R. White, private communication.
- <sup>88</sup>C. Daday, S. Smart, G. H. Booth, A. Alavi, and C. Filippi, *J. Chem. Theory Comput.* **8**, 4441 (2012).
- <sup>89</sup>M. A. Watson and G. K.-L. Chan, *J. Chem. Theory Comput.* **8**, 4013 (2012).
- <sup>90</sup>J. Ivanic, *J. Chem. Phys.* **119**, 9377 (2003).
- <sup>91</sup>J. Ivanic, J. R. Collins, and S. K. Burt, *J. Phys. Chem. A* **108**, 2314 (2004).
- <sup>92</sup>B. O. Roos, V. Veryazov, J. Conradie, P. R. Taylor, and A. Ghosh, *J. Phys. Chem. B* **112**, 14099 (2008).
- <sup>93</sup>S. Vancouillie, H. Zhao, V. T. Tran, M. F. A. Hendrickx, and K. Pierloot, *J. Chem. Theory Comput.* **7**, 3961 (2011).
- <sup>94</sup>K. Boguslawski, K. H. Marti, and M. Reiher, *J. Chem. Phys.* **134**, 224101 (2011).
- <sup>95</sup>W. Zhang, J. L. Loebach, S. R. Wilson, and E. N. Jacobsen, *J. Am. Chem. Soc.* **112**, 2801 (1990).
- <sup>96</sup>E. N. Jacobsen, W. Zhang, A. R. Muci, J. R. Ecker, and L. Deng, *J. Am. Chem. Soc.* **113**, 7063 (1991).
- <sup>97</sup>Y. G. Abashkin, J. R. Collins, and S. K. Burt, *Inorg. Chem.* **40**, 4040 (2001).
- <sup>98</sup>M. M. Francl, *J. Chem. Phys.* **77**, 3654 (1982).
- <sup>99</sup>V. A. Rassolov, J. A. Pople, M. A. Ratner, and T. L. Windus, *J. Chem. Phys.* **109**, 1223 (1998).
- <sup>100</sup>H. Chen, W. Lai, and S. Shaik, *J. Phys. Chem. B* **115**, 1727 (2011).
- <sup>101</sup>M. Swart, *Int. J. Quantum Chem.* **113**, 2 (2013).
- <sup>102</sup>A. R. Groenhof, M. Swart, A. W. Ehlers, and K. Lammertsma, *J. Phys. Chem. A* **109**, 3411 (2005).
- <sup>103</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>104</sup>N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **123**, 64107 (2005).
- <sup>105</sup>N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **125**, 074110 (2006).
- <sup>106</sup>P.-O. Widmark, P.-Å. Malmqvist, and B. O. Roos, *Theor. Chim. Acta.* **77**, 291 (1990).