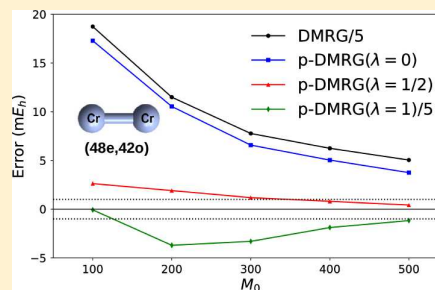


A Perturbative Density Matrix Renormalization Group Algorithm for Large Active Spaces

Sheng Guo,¹ Zhendong Li,¹ and Garnet Kin-Lic Chan^{2*}

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

ABSTRACT: We describe a low cost alternative to the standard variational DMRG (density matrix renormalization group) algorithm that is analogous to the combination of the selected configuration interaction plus perturbation theory (SCI+PT). We denote the resulting method p-DMRG (perturbative DMRG) to distinguish it from the standard variational DMRG. p-DMRG is expected to be useful for systems with very large active spaces, for which variational DMRG becomes too expensive. Similar to SCI+PT, in p-DMRG, a zeroth-order wave function is first obtained by a standard DMRG calculation but with a small bond dimension. Then, the residual correlation is recovered by a second-order perturbative treatment. We discuss the choice of partitioning for perturbation theory, which is crucial for its accuracy and robustness. To circumvent the problem of a large bond dimension in the first-order wave function, we use a sum of matrix product states to expand the first-order wave function, yielding substantial savings in computational cost and memory. We also propose extrapolation schemes to reduce the errors in the zeroth- and first-order wave functions. Numerical results for Cr₂ with a (28e, 76o) active space and 1,3-butadiene with an (22e, 82o) active space reveal that p-DMRG provides ground state energies of a similar quality to variational DMRG with very large bond dimensions but at a significantly lower computational cost. This suggests that p-DMRG will be an efficient tool for benchmark studies in the future.



1. INTRODUCTION

Achieving chemical accuracy (ca. 1 mE_h) in systems with a mix of multireference and dynamic correlations remains a challenging problem in molecular quantum chemistry. While complete active spaces (CAS) with tens of partially filled orbitals can be reliably treated by techniques such as the density matrix renormalization group (DMRG),^{1–11} reaching chemical accuracy in the subsequent description of the dynamic correlation is difficult. The most common technique to treat dynamical correlation in the multireference setting is second-order perturbation theory (PT).^{12–23} However, one often finds that a second-order perturbative treatment is not powerful enough to accurately describe correlations involving some of the moderately correlated nonvalence orbitals in a complex system. For example, in 3d transition metal systems, binding energies and exchange couplings can be substantially in error if the virtual 4d, semicore 3s3p, or valence ligand orbitals are treated only at the second-order perturbative level. The standard remedy is to include these additional moderately correlated orbitals in the multireference active space treatment. However, for complex systems, this can create enormous active spaces that are inaccessible or otherwise impractical even for current DMRG methods.

Recently, selected configuration interaction (SCI) methods^{24–26} have experienced a significant revival.^{27–32} The general idea of selected configuration interaction is quite old, dating back to the CIPSI method,²⁴ and before that to the hand-selected configuration interaction calculations carried out in the earliest days of quantum chemistry.^{33,34} Although modern day SCI methods differ in how they select

determinants, they all share a similar basic strategy. In particular, a small number of determinants are first selected for a variational treatment—in modern calculations, typically 10⁶–10⁸ determinants—and the residual correlation is treated by second-order PT, most commonly using the Epstein-Nesbet (EN) partitioning. Some important recent improvements include the use of stochastic methods to evaluate the second-order energies (E₂) in order to handle large basis sets,^{31,32} as well as the development of more systematic extrapolations with respect to the thresholds in the method. One finds that SCI methods achieve chemical accuracy in the total energy for a variety of small molecule problems using a remarkably small number of variational determinants. However, it is important to observe that the variational CI energy alone is itself usually quite poor. For example, in a heat-bath CI calculation on the chromium dimer (48e, 42o) active space³⁰ popularized in DMRG benchmarks,⁸ the variational CI energy was more than 60 mE_h above the DMRG benchmark result. Instead, it is the second order PT correction that yields the final high accuracy result. In the above case, the total energy error after using the perturbation theory correction is reduced to less than 1 mE_h, a reduction by a factor of almost 100. In other cases, corrections from PT reduce the total energy error by a factor of 10 or more.

The remarkable accuracy of the second-order perturbation correction in selected CI stands in stark contrast to the accuracy of second-order perturbation corrections when used

Received: March 17, 2018

Published: June 21, 2018



with complete active spaces. The physical reason for the difference is that even if the reference wave function is determined exactly (within the complete active space) it is *unbalanced* due to the lack of dynamical correlation. In contrast, although the variational selected CI computes only a quite approximate reference wave function, it is determined in a full, or at least large, space of orbitals, leading to a more balanced reference state. This suggests that the key to an accurate second-order correlation contribution involves balancing the different orbital correlations, rather than describing only the strongest correlations exactly, as in a valence CAS. This observation is independent of choosing selected CI for the reference wave function, and it is the motivation for this work.

In the current paper, we explore how we can use quite approximate, but balanced, variational DMRG reference wave functions computed in large active spaces and correct them efficiently and to high accuracy with second order PT within the same orbital space. We name this technique “perturbatively corrected DMRG” or p-DMRG. In p-DMRG, we represent both the zeroth-order variational reference wave function $|\Psi^{(0)}\rangle$ as well as the first-order perturbative correction $|\Psi^{(1)}\rangle$ in terms of matrix product states (MPS). Note that there are advantages to using a MPS representation, rather than a determinantal expansion, of the variational reference wave function. The MPS representation allows us to construct compact strongly correlated wave functions even where there is little to no determinantal sparsity, for example, in systems with many coupled spins, where there is little sparsity in the coupled low-spin configurations of the system. A second reason is that volume extensivity of the energy is achieved by a matrix product state with a cost $\propto e^{V^{2/3}}$ rather than $\propto e^V$ in configuration interaction. Asymptotically, this makes the variational MPS representation exponentially more compact than a variational determinant expansion and, in practice, allows for a larger number of spatially separated orbitals to be treated.³⁵

Relative to a standard variational DMRG calculation, the cost savings in p-DMRG arise from two sources. First, as described above, the zeroth-order wave function can be computed using a bond dimension M_0 much smaller than is needed to fully converge the variational DMRG calculation. Second, although the bond dimension M_1 for the first-order wave function still needs to be quite large, the first-order wave function is determined by minimizing the Hylleraas functional^{18,36}

$$\begin{aligned}\mathcal{L}[|\Psi_1\rangle] &= \langle \Psi_1 | (\hat{H}_0 - E_0) | \Psi_1 \rangle + 2 \langle \Psi_1 | \hat{V} | \Psi_0 \rangle, \\ \hat{V} &= \hat{H} - \hat{H}_0\end{aligned}\quad (1)$$

which is less expensive than minimizing the variational DMRG energy because the zeroth-order Hamiltonian $\hat{H}_0$ can be chosen to be simpler than the full Hamiltonian $\hat{H}$. For example, if $\hat{H}_0$ is the Fock operator or the Epstein–Nesbet Hamiltonian, then the computational cost to evaluate the Hylleraas functional is a factor of K (where K is the number of orbitals) less than that to evaluate the variational DMRG energy. In addition, since in second-order PT, only the matrix element $\langle \Psi_1 | \hat{V} | \Psi_0 \rangle$ needs to be computed (instead of $\langle \Psi_1 | \hat{H} | \Psi \rangle$ in standard DMRG), we can save a further factor of M_1/M_0 in cost, where we assume M_1 is similar to the bond dimension used in a converged variational DMRG calculation, and $M_1 \gg$

M_0 . The p-DMRG method can still be made exact by gradually increasing M_0 , which thus plays a role analogous to the variational selection threshold in SCI methods. This opens up the possibility to perform extrapolations, similar to those in SCI and in variational DMRG.

It is important to note that we expect p-DMRG to be useful for a different class of problems than standard DMRG-based multireference perturbation theory such as DMRG-CASPT2^{17,23} or DMRG-NEVPT2.^{18,19,21,22} In particular, we believe the method should be used to target high accuracy calculations (to say $1mE_h$ in the total energy) either in a large active space, including the intermediately correlated orbitals, or to obtain benchmark total energies in small problems, at a cost that is significantly less than that of variational DMRG. This is very different from providing a qualitative treatment of dynamical correlation in very large basis sets, which is the focus of standard DMRG-based multireference PT. Note that p-DMRG differs also from the similarly named DMRG inner space perturbation theory (DMRG-isPT),³⁷ where PT is only used to reduce the cost of the Davidson diagonalization in the DMRG sweeps.

The remainder of the paper is organized as follows. In Section 2.1, we first briefly summarize DMRG in the MPS language and then introduce the p-DMRG algorithm. Two particular pieces needed to establish p-DMRG as an accurate and efficient alternative to variational DMRG are then discussed in the following sections. Specifically, Section 2.2 discusses the choice of $\hat{H}_0$, which is crucial for obtaining high accuracy, while Section 2.3 introduces a way to tackle the large bond dimension M_1 needed to represent the first-order wave function by using a sum of MPS representations. After describing standard benchmark calculations for C_2 and Cr_2 in small active spaces, we carry out two larger benchmark studies using p-DMRG in Section 3: one for Cr_2 in an active space with 28 electrons in 76 orbitals generated by a cc-pVDZ-DK basis, denoted by the notation (28e, 76o), and the other for butadiene in an active space with (22e, 82o) generated by a cc-pVDZ basis. Both sets of calculations demonstrate that in practical problems p-DMRG is substantially more efficient than variational DMRG, obtaining the same benchmark accuracy with greatly reduced cost. Conclusions and outlines for future directions are presented in Section 4.

2. THEORY

2.1. Perturbative Density Matrix Renormalization Group (p-DMRG). Here, we first recapitulate the DMRG algorithm in the MPS language. Interested readers are referred to recent reviews, e.g., refs 11, 38, and 39 for details.

A generic FCI wave function can be written in Fock space as

$$|\Psi\rangle = \sum_{n_1 \cdots n_K} \Psi^{n_1 n_2 \cdots n_K} |n_1 n_2 \cdots n_K\rangle \quad (2)$$

where $|n_1 n_2 \cdots n_K\rangle$ is the occupation basis in the Fock space of K spatial orbitals, and $n_k \in \{0, 1, 2, 3\}$ for the local configuration basis $\{|0\rangle, |k_\beta\rangle, |k_\alpha\rangle, |k_\alpha k_\beta\rangle\}$, respectively. It can be decomposed into a sequential product of matrices associated with different orbitals via successive singular value decompositions (SVDs)

$$|\Psi\rangle = \sum_{n_1 \cdots n_K} A^{n_1}[1] A^{n_2}[2] \cdots A^{n_K}[K] |n_1 n_2 \cdots n_K\rangle \quad (3)$$

where $A^{n_k}[k]$ are matrices, and the symbol $A[k]$ will be used to represent the site tensor as a collection of matrices $A^{n_k}[k]$ for

different n_k . The dimensions of $A^n[k]$ are usually referred to as the bond dimensions, and these take a maximal value of $O(4^{K/2})$ in the middle of the orbital chain.³⁸ The MPS form (eq 3) can be used as a variational ansatz by restricting the maximal bond dimension to a given M , which is then the single parameter that controls the accuracy of the approximation. Clearly, as M approaches $O(4^{K/2})$, the ansatz becomes exact. However, the importance of the MPS ansatz is that for Hamiltonians with local interactions in one dimension the entanglement encoded in MPS with a M with only a very weak dependence on K is sufficient to accurately represent ground and low-energy eigenstates. For real molecules which have a more complicated entanglement structure, the required M is generally much larger than that used in one-dimensional models.⁵

The DMRG algorithm provides an efficient way to variationally optimize MPS that optimizes the tensors site-by-site. For simplicity, we consider here only the single site sweep algorithm. When we optimize the site tensor $A[k]$ at site k , MPS can be recast into a *mixed-canonical* form as

$$|\Psi\rangle = \sum_{n_1 \dots n_K} L^n[1] \dots L^{n_{k-1}}[k-1] C^{n_k}[k] R^{n_{k+1}}[k+1] \dots R^{n_K}[K] |n_1 n_2 \dots n_K\rangle \quad (4)$$

where the set of $L[k]$ are in left canonical form ($\sum_{n_k} L^{n_k} L^{n_k\dagger} = I$), and the set of $R[k]$ are in right canonical form ($\sum_{n_k} R^{n_k} R^{n_k\dagger} = I$). This choice of the left and right canonical gauges makes the renormalized configuration basis $\{|l_{k-1} n_k r_k\rangle\}$ orthonormal, where

$$|l_{k-1}\rangle = \sum_{n_k} (L^{n_k}[1] \dots L^{n_{k-1}}[k-1])_{l_{k-1}} |n_1 \dots n_{k-1}\rangle \quad (5)$$

$$|r_k\rangle = \sum_{n_k} (R^{n_k}[k+1] \dots R^{n_K}[K])_{r_k} |n_{k+1} \dots n_K\rangle \quad (6)$$

that is, $\langle l'_{k-1} | l_{k-1} \rangle = \delta_{l'_{k-1} l_{k-1}}$ and $\langle r'_k | r_k \rangle = \delta_{r'_k r_k}$. The central part $C^{n_k}[k]$ is the wave function to be optimized at site k , and it can be obtained by solving a standard configuration interaction problem in the renormalized configuration basis $\{|l_{k-1} n_k r_k\rangle\}$

$$\sum_{\text{lnr}} H_{l'n'r',\text{lnr}}[k] C_{\text{lnr}}[k] = E C_{\text{lnr}}[k] \quad (7)$$

where $H_{l'n'r',\text{lnr}} = \langle l'_{k-1} n'_k r'_k | \hat{H} | l_{k-1} n_k r_k \rangle$ is the matrix representation of the Hamiltonian $\hat{H}$, and $C_{\text{lnr}}[k]$ is the vectorized version of the tensor $C^{n_k}_{l_{k-1} n_k r_k}[k]$. The multiplication between $H_{l'n'r',\text{lnr}}$ and C_{lnr} dominates the cost of a DMRG calculation and scales as $O(K^3 M^3)$ in total per sweep.^{3,5} This scaling can be understood by noting that $H_{l'n'r',\text{lnr}}$ can always be written as a sum of $O(K^2)$ direct product terms $H_{l'n'r',\text{lnr}} = \sum_{\beta} O^{\beta}_{l'n',\text{ln}} O^{\beta}_{r',r}$ for a generic second quantized Hamiltonian with $O(K^4)$ terms, via the complementary operator technique,^{3,5,40} such that the matrix vector product can be formed by $O(K^2)$ independent matrix multiplications

$$\sigma_{l'n'r'} = \sum_{\text{lnr}} H_{l'n'r',\text{lnr}} C_{\text{lnr}} = \sum_{\beta} \left(\sum_r \left(\sum_{\text{ln}} O^{\beta}_{l'n',\text{ln}} C_{\text{lnr}} \right) O^{\beta}_{r',r} \right) \quad (8)$$

The cost for each multiplication scales as $O(M^3)$; thus, the cost for forming $\sigma_{l'n'r'}$ scales as $O(K^2 M^3)$ at a given site k . In

combination with the cost for building the necessary operators $O^{\beta}_{l'n',\text{ln}}$ and $O^{\beta}_{r',r}$ for representing $H_{l'n'r',\text{lnr}}$, the computational cost for the standard DMRG algorithm using the quantum chemistry Hamiltonian scales as $O(K^3 M^3 + K^4 M^2)$,^{3,5} which, unlike FCI, is a polynomial in K if M can be kept constant as a function of K , as is the case in certain situations, such as in pseudo-one-dimensional molecules.

However, to describe dynamical correlation in a small molecule over length scales too short for locality of correlations to emerge, M needs to scale as $O(K)$ to capture the local double excitations.⁸ This renders the total scaling effectively $O(K^6)$. This limits the number of orbitals that can be treated accurately with reasonable computational resources and time. For instance, as shown in ref 8, a state-of-the-art DMRG calculation on butadiene with an active space (22e, 82o) took 1 day on 42 cores for a single sweep with $M = 3000$. In this scenario, the correlation treatment offered by MPS, where every orbital is treated on an equal footing, is too flexible. Thus, a less general, but more efficient formulation, is clearly desired.

In the p-DMRG method, we assume that MPS with small M_0 has been optimized by the above standard DMRG algorithm, and it is used as the zeroth-order wave function $|\Psi^{(0)}\rangle$. Then, the first-order wave function $|\Psi^{(1)}\rangle$ can be obtained by minimizing the Hylleraas functional (eq 1), which in the exact case is equivalent to solving the first-order equations

$$(\hat{H}_0 - E_0) |\Psi^{(1)}\rangle = -Q \hat{H} |\Psi^{(0)}\rangle, \quad Q = 1 - P, \quad P = |\Psi^{(0)}\rangle \langle \Psi^{(0)}| \quad (9)$$

Note that although the bond dimension of $|\Psi^{(0)}\rangle$ is chosen small, the bond dimension M_1 of $|\Psi^{(1)}\rangle$ arising from eq 9 can be substantially larger, for example, as large as the bond dimension used in a converged variational DMRG calculation. In the following sections, we discuss different definitions of the zeroth-order Hamiltonian $\hat{H}_0$ and how to solve the first-order equation efficiently for the large bond dimensions arising in $|\Psi^{(1)}\rangle$. Note also that although only ground states are considered in the present work it is straightforward to extend the p-DMRG formalism to excited states. In this case, state-averaged DMRG can be used to compute the zeroth-order wave functions for the excited states, and the projector $P = |\Psi^{(0)}\rangle \langle \Psi^{(0)}|$ should be replaced by $P = \sum_i |\Psi_i^{(0)}\rangle \langle \Psi_i^{(0)}|$ to avoid intruder state problems.

2.2. Choices of Zeroth-Order Hamiltonian $\hat{H}_0$. There are several criteria that a good partitioning of $\hat{H}$ must satisfy. First, in order to reduce the computational cost, $\hat{H}_0$ should be as simple as possible. The Fock operator or the diagonal part of $\hat{H}$ in the determinant space used in the EN partition both satisfy this criterion, while the simplest projective definition $\hat{H}_0 = P \hat{H} P + Q \hat{H} Q$ does not. Second, the partition should be free of intruder state problems. The Fock operator generally does not satisfy this criterion (as we have numerically verified) and hence is not discussed further. Instead, we exclusively focus on designing $\hat{H}_0$ based on the idea of the EN partition, as also used in SCI+PT schemes.^{24–32} Third, the partition should give good energies at second order, which requires a balanced treatment of $|\Psi^{(0)}\rangle$ and $|\Psi^{(1)}\rangle$. Fourth, to be used in a spin-adapted DMRG algorithm,⁷ we require a spin-free $\hat{H}_0$. This differs from the partitioning in determinant-based SCI+PT, where $\hat{H}_0$ does not commute with the spin squared operator $\hat{S}^2$, leading to spin contamination in the first-order wave function.

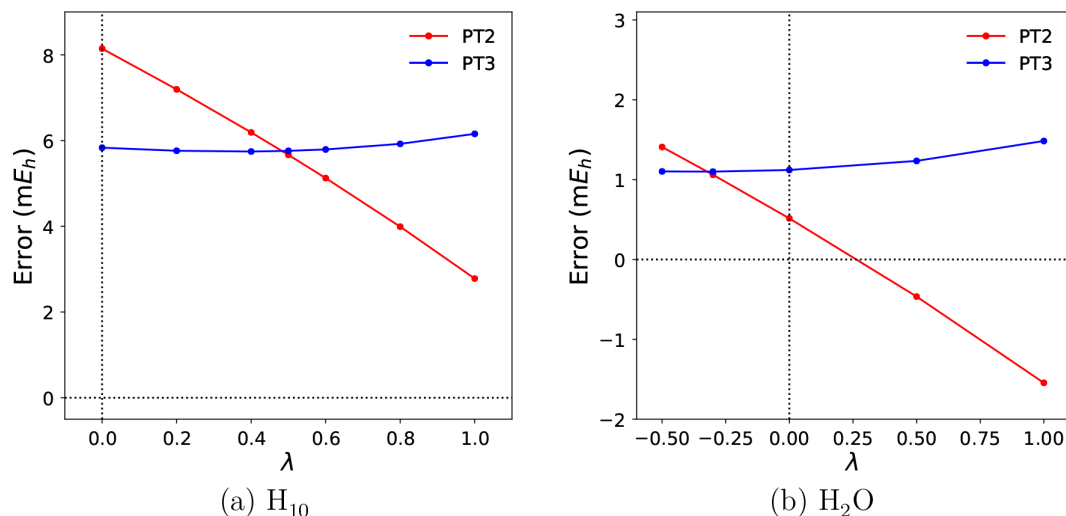


Figure 1. Dependence of the absolute errors of second- and third-order perturbation energy corrections on $E_0(\lambda) = (1 - \lambda) E_{DMRG}^{(0)} + \lambda E_d^{(0)}$: (a) H_{10} with $R(H-H) = 1.0 \text{ \AA}$ and $M_0 = 12$ in a STO-3G basis. (b) H_2O at the equilibrium geometry and $M_0 = 10$ in DZ basis.

To begin, we start with $\hat{H}_0$ defined as

$$\hat{H}_0 = PE_0P + Q\hat{H}_dQ \quad (10)$$

where $\hat{H}_d$ contains all single and double excitations which do not change the occupation numbers of spatial orbitals

$$\hat{H}_d = \sum_i h_{ii} \hat{E}_{ii} + \frac{1}{2} \sum_{i,j} (iijj) \hat{e}_{ijji} + \frac{1}{2} \sum_{i \neq j} (ijji) \hat{e}_{ijij} \quad (11)$$

with $\hat{E}_{ij} = \sum_{\sigma} a_{i\sigma}^\dagger a_{j\sigma}$ and $\hat{e}_{ijkl} = \sum_{\sigma, \tau} a_{i\sigma}^\dagger a_{j\tau}^\dagger a_{k\tau} a_{l\sigma} = E_{ij} E_{kl} - \delta_{ij} E_{kl}$. $\hat{H}_d$ defined in eq 11 is analogous to the zeroth-order Hamiltonian in the EN partition, but it is spin-free and only block-diagonal rather than diagonal in the determinant basis since it contains additional couplings for determinants with the same spatial occupations due to the $\hat{e}_{ijji}$ operator in the exchange term. Numerical comparisons within a nonspin-adapted DMRG implementation⁴¹ demonstrate that $\hat{H}_d$ and the standard EN partition provide results of very similar quality. For this form of $\hat{H}_0$, when solving eq 9 using the DMRG sweep algorithm, the Hamiltonian and wave function multiplication on the left-hand side (LHS) scales as $O(K^2 M_1^3)$ instead of $O(K^3 M^3)$ in the standard variational DMRG. The construction of the right-hand side (RHS) will scale as $O(K^3 M_1^2 M_0)$ assuming $M_1 \gg M_0$. The cost to build the renormalized operators for the LHS is negligible, as it is only $O(K^2 M_1^2)$, while the corresponding cost for the RHS is $O(K^4 M_1 M_0)$ in total. Thus, compared to the variational DMRG calculation with a similar $M \approx M_1$, we expect a substantial reduction in cost.

In eq 10, we have not yet defined the zeroth-order energy E_0 . There are two natural choices. One is the DMRG energy for $|\Psi^{(0)}\rangle$, viz., $E_{DMRG}^{(0)} = \langle \Psi^{(0)} | \hat{H} | \Psi^{(0)} \rangle$, which is analogous to the choice made in SCI+PT. We observe that the zeroth-order variational energy $E_{DMRG}^{(0)}$ is typically closer to the exact energy than the zeroth-order energies used in SCI+PT. Thus, it is much lower than the lowest energy of the perturbors, which is the lowest eigenvalue of $Q\hat{H}_dQ$, whose eigenstates are relatively uncorrelated. This means that although this choice of E_0 is, in general, numerically stable in that it is free of intruder state problems for typical M_0 (which is large enough to achieve a nonvanishing gap between the zeroth-order state and the

perturbors), the correlation energy recovered is usually too small at the second-order level. The other natural choice $E_d^{(0)} = \langle \Psi^{(0)} | \hat{H}_d | \Psi^{(0)} \rangle$ makes the gap smaller and hence lowers E_2 . However, in this case, the correlation energy can be overestimated, and there is a greater probability of intruder states because there is no guarantee that the lowest eigenvalue of $Q\hat{H}_dQ$ is larger than $E_d^{(0)}$. Therefore, in general, we expect that an interpolation $E_0(\lambda) = (1 - \lambda) E_{DMRG}^{(0)} + \lambda E_d^{(0)}$ between these two limits will provide better performance in terms of stability and accuracy.

Unfortunately, there is no a priori way to determine λ without calculation. One way to define it through a calculation, is through the optimized partitioning method,⁴² where λ is chosen to make $E_3(\lambda) = 0$ or equivalently $E_2(\lambda) + E_3(\lambda)$ stationary, while $E_0(\lambda) + E_1(\lambda) = E_{DMRG}^{(0)}$ is independent of λ . We have explored the dependence of the absolute errors of second- and third-order perturbation theories (PT2 and PT3) on λ as shown in Figure 1 for two small systems, viz., a hydrogen chain H_{10} with $R(H-H) = 1.0 \text{ \AA}$ in a STO-3g basis⁴³ and H_2O at the equilibrium geometry⁵ in the Dunning's DZ basis.⁴⁴ It is clear that as λ increases and $E_0(\lambda)$ approaches $E_d^{(0)}$, $E_2(\lambda)$ is lower, for the reasons discussed above. In contrast, the PT3 energy varies more slowly. However, including PT3 does not always improve the results; e.g., for H_2O , the error of PT2+PT3 is larger than using PT2 alone when $\lambda = 0$. Empirically, we observe that the error obtained at the midpoint $\lambda = 1/2$ is always improved over that obtained with $\lambda = 0$. Hence, in the following, we use this simple choice in addition to the two obvious choices $\lambda = 0$ and $\lambda = 1$.

2.3. Splitting the First-Order Wave Function. In general, for large numbers of orbitals, the bond dimension M_1 required to achieve a given accuracy increases with K . Thus, the dominant scaling when solving for the first-order wave function is dominated by the scaling $O(K^2 M_1^3)$ encountered when computing the LHS of eq 9. A similar computational obstacle arises also in SCI+PT, which gives rise to the memory bottleneck associated with storing all determinants contributing to the first-order wave function. One way to remove this bottleneck is to use a stochastic computation of the perturbation correction, as proposed in refs 31 and 32 for SCI+PT. In the current work, we use a

deterministic approach, where we represent the first-order wave function as a linear combination of MPS,⁴⁵ each with a modest bond dimension.

Specifically, noting that eq 9 is a linear equation, we use the following ansatz

$$|\Psi^{(1)}\rangle = \sum_{i=1}^N |\Psi_i^{(1)}\rangle \quad (12)$$

where each $|\Psi_i^{(1)}\rangle$ is represented by MPS with a fixed bond dimension M_1 and can be determined recursively from the relation

$$(\hat{H}_0 - E_0)|\Psi_i^{(1)}\rangle = |r_i\rangle, \\ |r_i\rangle = -Q\hat{H}|\Psi^{(0)}\rangle - \sum_{j=1}^{i-1} (\hat{H}_0 - E_0)|\Psi_j^{(1)}\rangle \quad (13)$$

The form of the LHS is the same for each i , but the RHS becomes more costly as N increases. When computed from the Hylleraas functional, the largest cost arises from computing the expectation value $\langle \Psi_i^{(1)} | (\hat{H}_0 - E_0) | \Psi_j^{(1)} \rangle$ ($i > j$), and this cost scales as $O(K^2 M_1^3 N^2)$. Thus, using the split ansatz (eq 12) compared with a calculation using a large bond dimension $M'_1 = NM_1$, formally leads to a factor of N reduction in computational cost, as well as a factor of N^2 in memory. However, the representational power of MPS with $M'_1 = NM_1$ is larger than that of a linear combination of N MPS with bond dimension M_1 due to the compressibility of the sum of the MPS representation. For example, in the limiting case of $M_1 = 1$, eq 12 simply becomes a sum of N determinants, while the variational space described by MPS with bond dimension N is of course much larger. Thus, in practice, we try to use M_1 as large as possible given the computational resources and only then use eq 12 to continue the calculations to a larger effective M_1 , which would otherwise be too costly within a single MPS representation. The second order energy $E_2 = \langle \Psi^{(0)} | V | \Psi^{(1)} \rangle = \sum_{i=1}^N \langle E_{2,i} \rangle$ becomes a sum of N terms, where $E_{2,i}$ decays monotonically as i increases. This monotonic decay can be quite systematic, and we explore the possibility to extrapolate the series $\{E_{2,i}\}_{i=1}^N$ for large calculations in Section 3.2.

3. RESULTS

3.1. Benchmark: C_2 and Cr_2 . To test the performance of p-DMRG for various choices of H_0 , we examined two diatomic molecules, C_2 and Cr_2 , for which variational DMRG results are available in the literature.⁸ The same molecules were also studied in recent Heat-Bath CI plus PT calculations.³⁰ For these two molecules, we used canonical Hartree–Fock orbitals with D_{2h} symmetry and ordered them using genetic ordering as used in ref 8. The zeroth-order DMRG wave functions were computed in a default forward sweep where M_0 was increased gradually, using the BLOCK code.^{5,7}

Figure 2 shows the p-DMRG results for C_2 at the equilibrium bond length of 1.24253 Å in the cc-pVTZ basis set.⁴⁶ All electrons were correlated corresponding to an orbital space of (12e, 60o). The absolute errors are given relative to the essentially exact variational DMRG value.⁸ The second-order perturbation energies were calculated at an effective $M_1 = \infty$ by extrapolating with discarded weight from $M_1 = 5000, 4000, 3000$ in reverse sweep mode.⁸ We first note the significance of the perturbation correction; to compare the variational DMRG and p-DMRG calculations as a function of

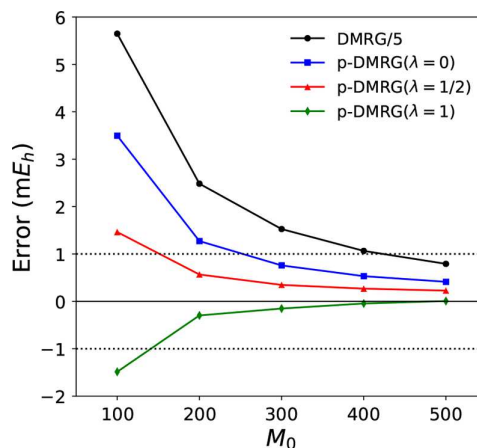


Figure 2. Absolute errors in zeroth-order DMRG energies $E_{DMRG}^{(0)}$ and perturbation corrections $E_{DMRG}^{(0)} + E_2(\lambda)$ with different H_0 for C_2 at the equilibrium bond length of 1.24253 Å. The errors are calculated relative to the converged variational DMRG energy in ref 8. The errors of zeroth-order DMRG energies are divided by 5 to put all curves into the same figure.

M_0 on the same plot, we had to divide the variational error by 5. We also see that in this dynamic correlation dominated system, the performance of the zeroth-order Hamiltonian with $\lambda = 1$ is quite good. Using $E_{DMRG}^{(0)}$ as E_0 instead underestimates the correlation energy. Here, $\lambda = 1/2$ also yields reasonable errors which reach chemical accuracy already for the very small variational DMRG calculation with $M_0 = 200$. We see that in the absence of intruder state problems, p-DMRG with different choices of λ all converge to the same ground state energy as M_0 increases, but the accuracy when M_0 is small can be quite different. For this reason, it is important to choose λ , such that one obtains good accuracy already with small M_0 , to obtain significant computational savings.

Next, we consider a more challenging example, Cr_2 , at three bond distances, $R = 1.50$ Å, which have been previously benchmarked by variational DMRG,⁸ the equilibrium bond length⁴⁷ $R = 1.68$ Å, and $R = 2.20$ Å. We used the Ahlrichs' SV basis set⁴⁸ and correlated all electrons. The resulting orbital space is (48e, 42o). The second-order perturbation energies were calculated at $M_1 = \infty$ by extrapolation from $M_1 = 8000, 7000, 6000$ (in reverse sweep mode). The p-DMRG results are shown in Figure 3. It is clear that Cr_2 is much more challenging than C_2 since all the DMRG and p-DMRG errors for a given M_0 are larger than those for C_2 with the same M_0 . At the equilibrium geometry, using $E_d^{(0)}$ ($\lambda = 1$) in p-DMRG leads to relatively larger errors due to a near-intruder state, while at $R = 1.50$ Å, the second-order energy is unphysically large. Using the midpoint energy $\lambda = 1/2$ as E_0 is a dramatic improvement compared with both $\lambda = 1$ and $\lambda = 0$ (the latter leads to an underestimation of the correlation energy). With M_0 equal to 300 or 400, the p-DMRG($\lambda = 1/2$) reaches chemical accuracy for $R = 1.50$ Å and $R = 1.68$ Å, with the perturbation correction again providing a large improvement of the variational energy. For $R = 2.20$ Å, where the correlation is stronger, the p-DMRG with $\lambda = 1/2$ converges slower compared with the other two bond distances but is still the best among different choices of λ investigated here. Thus, in the rest of this work, we always use $\lambda = 1/2$.

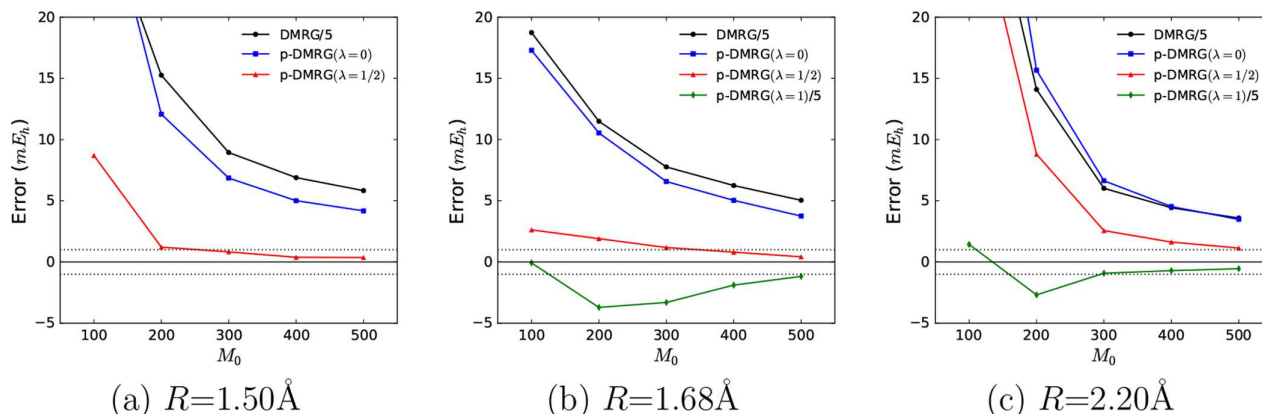


Figure 3. Absolute errors in zeroth-order DMRG energies $E_{\text{DMRG}}^{(0)}$ and perturbation corrections $E_{\text{DMRG}}^{(0)} + E_2(\lambda)$ with different $\hat{H}_0$ for Cr_2 . Both the errors of DMRG energies and p-DMRG($\lambda = 1$) were divided by 5 to fit all curves on the same figure. For $R = 1.50 \text{ \AA}$, p-DMRG($\lambda = 1$) suffers from intruder state problems.

Table 1. Energy ($E+2099$ in E_h) of Cr_2 Obtained with Standard Variational DMRG in cc-pVDZ-DK Basis Set^a

M	8000	10 000	12 000	14 000	16 000	∞ (extrapolated)
E (default schedule)	−0.8957	−0.8991	−0.9024	−0.9047	−0.9061	−0.9195 ± 0.0027
E (reverse schedule)	−0.8980	−0.9015	−0.9040	−0.9058	−0.9071	−0.9192 ± 0.0024

^aExtrapolated DMRG energy of Cr atom is $-1049.93254(4)E_h$.

3.2. Cr_2 with (28e, 76o) Orbital Space. As a first example of a larger calculation, we study the ground state energy of Cr_2 at $R = 1.68 \text{ \AA}$ with the cc-pVDZ-DK basis set.⁴⁹ Scalar relativistic effects were included through the spin-free X2C Hamiltonian.^{50–53} We used natural orbitals obtained from a CASSCF with a (12e, 12o) active space in the DMRG and p-DMRG calculations. The 1s, 2s, and 2p natural orbitals were not included in the (p)-DMRG calculations, leading to an orbital space with (28e, 76o). The standard variational DMRG energies for Cr_2 and the Cr atom are shown in Table 1. P-DMRG energies for Cr_2 are shown in Table 2. As an empirical

Table 2. Energy ($E+2099$ in E_h) of Cr_2 Obtained with p-DMRG in cc-pVDZ-DK Basis Set^a

M_0	1000	2000	3000	4000
$E_{\text{DMRG}}^{(0)}$	−0.8346	−0.8617	−0.8743	−0.8818
$E_2^{[1]}$	−0.0607	−0.0323	−0.0196	−0.0130
$E_2^{[2]}$	−0.0652	−0.0371	−0.0243	−0.0173
$E_2^{[3]}$	−0.0671	−0.0396	−0.0268	−0.0195
$E_2^{[4]}$	−0.0682	−0.0409	−0.0282	−0.0209
$E_2^{[5]}$	−0.0690	−0.0418	−0.0293	−0.0219
$E_2^{[\infty]}$	−0.0734	−0.0492	−0.0386	−0.0323
$E_{\text{DMRG}}^{(0)} + E_2^{[\infty]}$	−0.9080	−0.9109	−0.9129	−0.9141
Δ_2/Δ_0^b	0.141	0.157	0.157	0.157

^a $E_2^{[i]} = \sum_{j=1}^i E_{2,j}$ represents the accumulated second-order perturbation energy for the sum of the first i first-order MPS. $E_2^{[\infty]}$ represents the extrapolated energy for $M_1 = \infty$. The final extrapolated p-DMRG energy with respect to M_0 is $E_\infty = -2099.9201E_h$. ^b $\Delta_0 = E_{\text{DMRG}}^{(0)} - E_\infty$, $\Delta_2 = E_{\text{DMRG}}^{(0)} + E_2^{[\infty]} - E_\infty$.

estimate,⁸ the extrapolation error bar in the variational DMRG is assigned as 1/5 of the difference between the extrapolation energy and the energy with the largest $M = 16000$.

As shown in Table 1, the standard variational DMRG energy converges very slowly with respect to M . Even at $M = 16000$, the variational DMRG energy is above the extrapolated energy

by about $10mE_h$, while the DMRG energy at $M = 8000$ is about $20mE_h$ above. Similarly, unlike in the p-DMRG calculation with (48e, 42o), it is hard to converge $|\Psi_1\rangle$ with respect to bond dimension using a single MPS. Thus, in this system, we used the split ansatz (eq 12) to represent $|\Psi_1\rangle$. We chose the bond dimension of each split MPS to be $M_1 = 7500$. In Table 2, the accumulated second-order perturbation energies, $E_2^{[i]} = \sum_{j=1}^i E_{2,j}$ for the sum of the first i first-order MPS, is shown for the first five terms in the split. We also see slow convergence, for example, at $M_0 = 3000$, adding an additional MPS in the sum only lowers the energy by about $1mE_h$ (after the second term in the sum). In fact, we found that even after summing over 10 MPS (when $M_0 = 3000$), the change in E_2 for each subsequent MPS was as large as $0.3mE_h$. Thus, extrapolation is also needed to estimate a converged E_2 .

To carry out the extrapolation, we used the linear relation between $\ln |\delta E|$ and $(\ln M)^2$ described in refs 5 and 54. Figure 4 shows the accumulated energies $E_2(M = NM_1) \triangleq E_2^{[N]}$ as a function of $(\ln M)^2$ as well as the fitted curves $E_2(M) = E_2^{[\infty]} + Ae^{-\kappa(\ln M)^2}$ using the first 5 (red solid) and 10 (blue dashed) points. We see that using the first 5 points is sufficient to obtain a good extrapolation. The extrapolated $E_2^{[\infty]}$ from 5 points is $-0.03861E_h$, which differs from that using 10 points ($-0.03845E_h$) by only $0.16mE_h$. Using such an extrapolation leads to substantial computational savings. The full set of extrapolated results $E_2^{[\infty]}$ are listed in Table 2. It is notable that the p-DMRG energy at $M_0 = 1000$ with the first five basis functions, $E_{\text{DMRG}}^{(0)} + E_2^{[5]}$, is $-0.9036E_h$, which is already close to the variational DMRG result with $M_0 = 12000$. Using the extrapolated E_2 , the p-DMRG energies with every M_0 are lower than the variational DMRG results with $M = 12000$, the largest M used.

To obtain a fully converged energy, we further need to extrapolate the variational bond dimension $M_0 \rightarrow \infty$. The need for two extrapolations are similar to the dual extrapolation in the original heat-bath CI+PT,³⁰ where one extrapolation is for

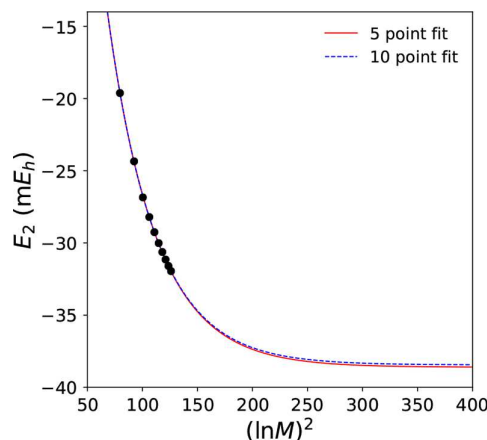


Figure 4. Accumulated energies $E_2(M = NM_1) \triangleq E_2^{[N]}$ as a function of $(\ln M)^2$ and the fitted curves $E_2(M) = E_2^{[\infty]} + Ae^{-\kappa(\ln M)^2}$ using the first 5 (red solid) and 10 (blue dashed) splitting functions for $M_0 = 3000$ and $M_1 = 7500$.

the exact PT2 energy, while the other is to extrapolate the CI energy to zero selection threshold. To carry out this second extrapolation, we observe that the ratio Δ_2/Δ_0 , where $\Delta_2 = E_{\text{DMRG}}^{(0)} + E_2^{[\infty]} - E_\infty$ and $\Delta_0 = E_{\text{DMRG}}^{(0)} - E_\infty$ is almost perfectly constant for different M_0 , as seen Table 2. This relation allows us to estimate E_∞ . The estimated E_∞ , using the largest three M_0 , is $-2099.9201E_h$, which is in agreement with the extrapolated variational DMRG results to within $1mE_h$, and within the extrapolation error bars. Compared to the atomic energies, we obtain a binding energy at this geometry of 1.50 eV, which is in fortuitously good agreement with the experimental value of 1.47 eV.³⁵ This demonstrates how, in practice, p-DMRG can be used as a cheaper alternative to variational DMRG to estimate an exact ground state energy even in a fairly complicated system.

3.3. Butadiene with (22e, 82o) Active Space. The final system we consider is 1,3-butadiene. This system has been studied by many accurate methods including high-order coupled cluster theory,⁵⁶ initiator full configuration interaction quantum Monte Carlo (*i*-FCIQMC),^{57,58} and stochastic heat-bath CI (SHCI).⁵⁹ Benchmark energies have been reported using variational DMRG.⁸ We used the same basis ANO-L-VDZP[3s2p1d]/[2s1p]⁶⁰ as used in previous studies.^{8,56,57,59} All electrons except for a frozen 1s core were correlated, leading to an orbital space with (22e, 82o). We used split-localized canonical orbitals for the p-DMRG calculations, ordered by genetic ordering.⁸ In the p-DMRG calculations, the first-order MPS was split into five parts, and each part had a bond dimension $M_1 = 3000$. We used the same extrapolation procedures as used for Cr_2 in the previous section. The computed energies are shown in Table 3. Because of the prohibitive computational cost, the extrapolated variational DMRG was not reported in ref 8. However, it can be seen that $E_{\text{DMRG}}^{(0)} + E_2^{[\infty]}$ for $M_0 = 2000$ is already lower than the variational DMRG energy for $M = 6000$. Thus, we expect the exact ground state energy should be even lower. Further using extrapolation for M_0 , we obtain an estimated exact energy of $-155.557567E_h$, which is lower than the $M_0 = 2000$ p-DMRG energy by only $0.25mE_h$. A similar extrapolation using the previous DMRG energies gives $-155.5578E_h$, which is consistent with the extrapolated p-DMRG energy. The very

Table 3. Energy ($E+155$ in E_h) of Butadiene with (22e, 82o) Active Space

DMRG-PT			
M_0	$E_{\text{DMRG}}^{(0)}$	$E_{\text{DMRG}}^{(0)} + E_2^{[\infty]}$	Δ_2/Δ_0^a
500	-0.552593^b	-0.556038	0.308
1000	-0.555438^b	-0.556887	0.319
2000	-0.556713^b	-0.557318	0.292
∞		-0.557567	
$M = 4000^c$		-0.556874	
$M = 5000^c$		-0.557050	
$M = 6000^c$		-0.557178	
$M = \infty^c$		-0.5578	
CCSD(T) ^d		-0.555002	
CCSDT ^d		-0.555959	
<i>i</i> -FCIQMC ^e		$-0.5491(4)$	
<i>i</i> -FCIQMC ^f		$-0.5578(10)$	
SHCI ^g		$-0.5582(1)$	

^a $\Delta_0 = E_{\text{DMRG}}^{(0)} - E_\infty$, $\Delta_2 = E_{\text{DMRG}}^{(0)} + E_2^{[\infty]} - E_\infty$. ^bZeroth-order DMRG energies are slightly different from the previous DMRG energies in ref 8 due to the use of different orbitals and optimization schedules. Split-localized canonical orbitals were used in this work, and $E_{\text{DMRG}}^{(0)}$ are for converged one-site MPS calculations without any noise. ^cDMRG results from ref 8. The extrapolation is performed based on $E(M) = E^{[\infty]} + Ae^{-\kappa(\ln M)^2}$. ^dRef 56. ^eRef 57. ^fRef 58. Perturbation correction was used based on FCIQMC. ^gRef 59.

recent HCI+PT⁵⁹ and FCIQMC⁵⁸ results are also within $1mE_h$ compared with the extrapolated variational DMRG and p-DMRG energies. Thus, we expect the extrapolated p-DMRG energy to be very close to the exact ground state energy and at least within chemical accuracy.

4. CONCLUSION

In this work, we defined a p-DMRG method that uses perturbation theory within the DMRG framework to efficiently target exact energies in large orbital spaces where not all orbitals are strongly correlated. Using a carefully defined zeroth-order Hamiltonian, and with extrapolation procedures, we found that p-DMRG can indeed provide benchmark quality energies as accurate as those obtained in far more expensive standard variational DMRG calculation. Future work will be carried out to perform benchmark studies using p-DMRG for the kinds of strongly correlated problems where there are a large number of intermediately correlated, as well as strongly correlated orbitals, and which currently lie beyond the capabilities of the practical variational DMRG calculations.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gkc1000@gmail.com.

ORCID

Sheng Guo: 0000-0002-1083-1882

Zhendong Li: 0000-0002-0683-6293

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation through CHE 1665333. Additional support for development of the BLOCK program used in this work was provided by OAC 1657286. Z.L. is supported by the Simons

Collaboration on the Many-Electron Problem. G.K.-L.C. is a Simons Investigator in Physics.

REFERENCES

- (1) White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **1992**, *69*, 2863–2866.
- (2) White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *48*, 10345–10356.
- (3) White, S. R.; Martin, R. L. Ab initio quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* **1999**, *110*, 4127–4130.
- (4) Mitrushenkov, A. O.; Fano, G.; Ortolani, F.; Linguerr, R.; Palmieri, P. Quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* **2001**, *115*, 6815–6821.
- (5) Chan, G. K.-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.
- (6) Legeza, Ö.; Röder, J.; Hess, B. A. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2003**, *67*, 125114.
- (7) Sharma, S.; Chan, G. K.-L. Spin-adapted density matrix renormalization group algorithms for quantum chemistry. *J. Chem. Phys.* **2012**, *136*, 124121.
- (8) Olivares-Amaya, R.; Hu, W.; Nakatani, N.; Sharma, S.; Yang, J.; Chan, G. K.-L. The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* **2015**, *142*, 034102.
- (9) Keller, S.; Dolfi, M.; Troyer, M.; Reiher, M. An efficient matrix product operator representation of the quantum chemical Hamiltonian. *J. Chem. Phys.* **2015**, *143*, 244118.
- (10) Yanai, T.; Kurashige, Y.; Mizukami, W.; Chalupský, J.; Lan, T. N.; Saitow, M. Density matrix renormalization group for ab initio Calculations and associated dynamic correlation methods: A review of theory and applications. *Int. J. Quantum Chem.* **2015**, *115*, 283–299.
- (11) Chan, G. K.-L.; Keselman, A.; Nakatani, N.; Li, Z.; White, S. R. Matrix product operators, matrix product states, and ab initio density matrix renormalization group algorithms. *J. Chem. Phys.* **2016**, *145*, 014102.
- (12) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- (13) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons, Inc., 1996; pp 219–331.
- (14) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multi-reference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
- (15) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.* **2001**, *350*, 297–305.
- (16) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138–9153.
- (17) Kurashige, Y.; Yanai, T. Second-order perturbation theory with a density matrix renormalization group self-consistent field reference function: Theory and application to the study of chromium dimer. *J. Chem. Phys.* **2011**, *135*, 094104.
- (18) Sharma, S.; Chan, G. K.-L. Communication: A flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states. *J. Chem. Phys.* **2014**, *141*, 111101.
- (19) Guo, S.; Watson, M. A.; Hu, W.; Sun, Q.; Chan, G. K.-L. N-Electron Valence State Perturbation Theory Based on a Density Matrix Renormalization Group Reference Function, with Applications to the Chromium Dimer and a Trimer Model of Poly(p-Phenylenevinylene). *J. Chem. Theory Comput.* **2016**, *12*, 1583–1591.
- (20) Sokolov, A. Y.; Chan, G. K.-L. A time-dependent formulation of multi-reference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 064102.
- (21) Freitag, L.; Knecht, S.; Angeli, C.; Reiher, M. Multireference Perturbation Theory with Cholesky Decomposition for the Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2017**, *13*, 451–459. PMID: 28094988.
- (22) Sokolov, A. Y.; Guo, S.; Ronca, E.; Chan, G. K.-L. Time-dependent N-electron valence perturbation theory with matrix product state reference wavefunctions for large active spaces and basis sets: Applications to the chromium dimer and all-trans polyenes. *J. Chem. Phys.* **2017**, *146*, 244102.
- (23) Nakatani, N.; Guo, S. Density matrix renormalization group (DMRG) method as a common tool for large active-space CASSCF/CASPT2 calculations. *J. Chem. Phys.* **2017**, *146*, 094102.
- (24) Huron, B.; Malrieu, J.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745–5759.
- (25) Buenker, R. J.; Peyerimhoff, S. D. Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theor. Chem. Acc.* **1974**, *35*, 33–58.
- (26) Harrison, R. J. Approximating full configuration interaction with selected configuration interaction and perturbation theory. *J. Chem. Phys.* **1991**, *94*, 5021–5031.
- (27) Schriber, J. B.; Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106.
- (28) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112.
- (29) Liu, W.; Hoffmann, M. R. iCI: Iterative CI toward full CI. *J. Chem. Theory Comput.* **2016**, *12*, 1169–1178.
- (30) Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.
- (31) Sharma, S.; Holmes, A. A.; Jeanmairet, G.; Alavi, A.; Umrigar, C. J. Semistochastic Heat-bath Configuration Interaction method: selected configuration interaction with semistochastic perturbation theory. *J. Chem. Theory Comput.* **2017**, *13*, 1595–1604.
- (32) Garniron, Y.; Scemama, A.; Loos, P.-F.; Caffarel, M. Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory. *J. Chem. Phys.* **2017**, *147*, 034101.
- (33) Foster, J.; Boys, S. Canonical configurational interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300.
- (34) Bender, C. F.; Davidson, E. R. Studies in configuration interaction: The first-row diatomic hydrides. *Phys. Rev.* **1969**, *183*, 23.
- (35) Hachmann, J.; Cardoen, W.; Chan, G. K.-L. Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group. *J. Chem. Phys.* **2006**, *125*, 144101.
- (36) Hylleraas, E. A. Über den Grundterm der Zweielektronenprobleme von H⁻, He, Li⁺, Be⁺⁺ usw. *Eur. Phys. J. A* **1930**, *65*, 209–225.
- (37) Ren, J.; Yi, Y.; Shuai, Z. Inner Space Perturbation Theory in Matrix Product States: Replacing Expensive Iterative Diagonalization. *J. Chem. Theory Comput.* **2016**, *12*, 4871–4878.
- (38) Schollwöck, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **2011**, *326*, 96–192.
- (39) Keller, S.; Reiher, M. Spin-adapted matrix product states and operators. *J. Chem. Phys.* **2016**, *144*, 134101.
- (40) Xiang, T. Density-matrix renormalization-group method in momentum space. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *53*, R10445.
- (41) Li, Z.; Chan, G. K.-L. Spin-Projected Matrix Product States: Versatile Tool for Strongly Correlated Systems. *J. Chem. Theory Comput.* **2017**, *13*, 2681–2695.

- (42) Surján, P.; Szabados, A. Optimized partitioning in perturbation theory: Comparison to related approaches. *J. Chem. Phys.* **2000**, *112*, 4438–4446.
- (43) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.* **1969**, *51*, 2657–2664.
- (44) Dunning, T. H., Jr. Gaussian Basis Functions for Use in Molecular Calculations. I. Contraction of (9s5p) Atomic Basis Sets for the First-Row Atoms. *J. Chem. Phys.* **1970**, *53*, 2823–2833.
- (45) Huang, R.-Z.; Liao, H.-J.; Liu, Z.-Y.; Xie, H.-D.; Xie, Z.-Y.; Zhao, H.-H.; Chen, J.; Xiang, T. A generalized Lanczos method for systematic optimization of tensor network states. *arXiv:1611.09574*.
- (46) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (47) Bondybey, V. E.; English, J. H. Electronic structure and vibrational frequency of Cr₂. *Chem. Phys. Lett.* **1983**, *94*, 443–447.
- (48) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (49) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.
- (50) Liu, W. Ideas of relativistic quantum chemistry. *Mol. Phys.* **2010**, *108*, 1679–1706.
- (51) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *ChemPhysChem* **2011**, *12*, 3077–3094.
- (52) Peng, D.; Reiher, M. Exact decoupling of the relativistic Fock operator. *Theor. Chem. Acc.* **2012**, *131*, 1081.
- (53) Li, Z.; Xiao, Y.; Liu, W. On the spin separation of algebraic two-component relativistic Hamiltonians. *J. Chem. Phys.* **2012**, *137*, 154114.
- (54) Chan, G. K.-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.
- (55) Casey, S. M.; Leopold, D. G. Negative ion photoelectron spectroscopy of chromium dimer. *J. Phys. Chem.* **1993**, *97*, 816–830.
- (56) Watson, M. A.; Chan, G. K.-L. Excited States of Butadiene to Chemical Accuracy: Reconciling Theory and Experiment. *J. Chem. Theory Comput.* **2012**, *8*, 4013–4018.
- (57) Daday, C.; Smart, S.; Booth, G. H.; Alavi, A.; Filippi, C. Full Configuration Interaction Excitations of Ethene and Butadiene: Resolution of an Ancient Question. *J. Chem. Theory Comput.* **2012**, *8*, 4441–4451.
- (58) Blunt, N. S. An efficient and accurate perturbative correction to initiator full configuration interaction quantum Monte Carlo. *arXiv:1804.09528*.
- (59) Chien, A. D.; Holmes, A. A.; Otten, M.; Umrigar, C. J.; Sharma, S.; Zimmerman, P. M. Excited States of Methylene, Polyenes, and Ozone from Heat-Bath Configuration Interaction. *J. Phys. Chem. A* **2018**, *122*, 2714–2722. PMID: 29473750.
- (60) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor. Chim. Acta* **1990**, *77*, 291–306.