

Communication: A flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states

Sandeep Sharma and Garnet Kin-Lic Chan

Citation: *J. Chem. Phys.* **141**, 111101 (2014); doi: 10.1063/1.4895977

View online: <http://dx.doi.org/10.1063/1.4895977>

View Table of Contents: <http://aip.scitation.org/toc/jcp/141/11>

Published by the [American Institute of Physics](#)

Communication: A flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states

Sandeep Sharma and Garnet Kin-Lic Chan

Department of Chemistry, Frick Laboratory, Princeton University, Princeton, New Jersey 08544, USA

(Received 7 August 2014; accepted 8 September 2014; published online 19 September 2014)

We describe a formulation of multi-reference perturbation theory that obtains a rigorous upper bound to the second order energy by minimizing the Hylleraas functional in the space of matrix product states (MPS). The first order wavefunctions so obtained can also be used to compute the third order energy with little overhead. Our formulation has several advantages including (i) flexibility with respect to the choice of zeroth order Hamiltonian, (ii) recovery of the exact uncontracted multi-reference perturbation theory energies in the limit of large MPS bond dimension, (iii) no requirement to compute high body density matrices, (iv) an embarrassingly parallel algorithm (scaling up to the number of virtual orbitals, squared, processors). Preliminary numerical examples show that the MPS bond dimension required for accurate first order wavefunctions scales sub-linearly with the size of the basis. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4895977>]

Commonly, electron correlation is classified into static (or strong) correlation and dynamic (or weak) correlation. The former is essential to capture the qualitative electronic structure, and the latter to generate quantitatively accurate results. In most chemical systems, static correlation involves a small subset of orbitals with degenerate orbital energies on the energy scale of the Coulomb repulsion. In this subspace, also known as the active space, multi-configuration self-consistent field (MCSCF) calculations have traditionally been performed. However, the cost of (numerically exact) MCSCF scales exponentially with the number of active orbitals and is thus restricted to active spaces with about 16 electrons in 16 orbitals. In the last decade, this active space restriction has effectively been removed, without significant numerical errors, with the advent of new near-exact methods such as the density matrix renormalization group (DMRG)^{1–8} and full configuration interaction quantum Monte Carlo (FCIQMC).^{9–11} (More approximate techniques for large active spaces, such as restricted active space and general active space methods,^{12,13} high-order active-space coupled cluster,^{14–17} and variational reduced density matrix methods,^{18–20} have also been advanced.) However, for quantitative chemical accuracy, these new active space methods must still be augmented with techniques to include dynamic correlation, and this remains an important frontier.^{21–25}

The primary technical difficulty in treating dynamic correlation is the large number of orbitals require to converge the short-range electron-electron cusp. Explicit correlation (through R12 and F12 factors)^{26–31} significantly ameliorates, but does not eliminate, the need for large basis sets. One affordable approach to include dynamic correlation from a large number of external orbitals is through perturbation theory. Examples of such multi-reference perturbation methods include complete active space perturbation theory (CASPT2)^{32,33} and n-electron valence state perturbation theory (NEVPT2).^{34,35} However, even when recently combined with DMRG reference functions,^{21,36} these methods have only been used with

moderate active space sizes. The key bottleneck is the need to construct intermediates involving three- and four-body reduced density matrices (RDMs) in the active space. As the number of active orbitals increases, calculating and storing these RDMs become prohibitively expensive.

In this Communication, we describe an alternate formulation of a multi-reference perturbation theory, which we call matrix product state perturbation theory (MPS-PT). The new formulation possesses substantial advantages as follows:

1. It bypasses the need for high-body RDMs in the active space, potentially allowing very large active spaces to be used.
2. Perturbation theory depends strongly on the zeroth-order Hamiltonian H_0 . Changing H_0 (as in CASPT2 versus NEVPT2) usually requires re-deriving and re-implementing non-trivial intermediates. In our formulation we can easily change H_0 without significantly changing the implementation.
3. To reduce the cost of multi-reference perturbation theory, the first order wavefunction is often determined in a restricted space, as in internally contracted CASPT2 or partially and strongly contracted NEVPT2 (PC-NEVPT2, SC-NEVPT2). Aggressive contraction can lead to additional errors which are difficult to estimate *a priori*. We also contract the first order space via the MPS bond dimension, but do so in an automatic and optimal way. This contraction systematically and rapidly converges to the uncontracted result with increasing bond dimension, and the errors of contraction can be robustly estimated.
4. The algorithm is highly parallel and has a modest scaling with the MPS bond dimension and size of basis.

Our current implementation is significantly slower than existing NEVPT2 and CASPT2 implementations for small active spaces but in contrast allows very large active spaces to be used.

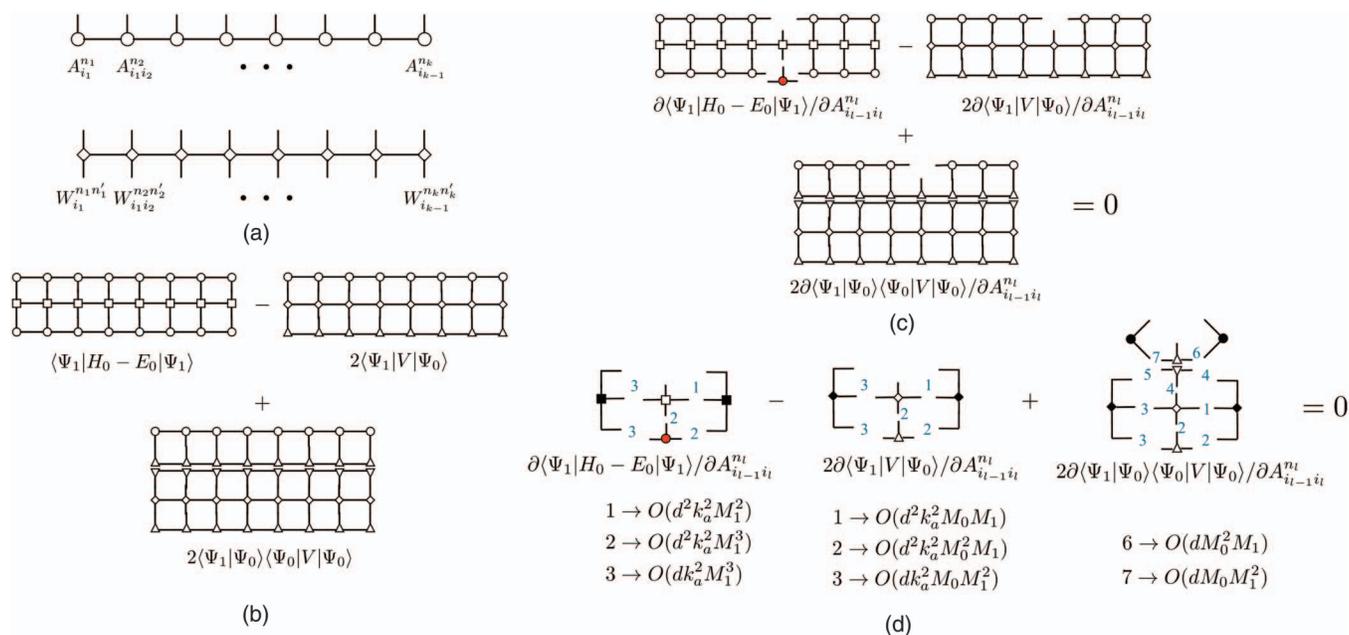


FIG. 1. (a) A matrix product state (MPS) can be represented graphically using a series of three-dimensional tensors, shown with circles, each of which is associated with an orbital. The free index, also known as the physical index (pointing upwards) of the tensors, denotes the occupation of the orbital and the other two indices, known as auxiliary indices, are sequentially contracted. The dimension d of the physical index is 4 for a spatial orbital, and the dimension M of the auxiliary index can be increased to make the MPS arbitrarily flexible. Similarly, a matrix product operator (MPO) can be represented as a series of four-dimensional tensors, with two physical indices and auxiliary indices contracted sequentially. Due to the 2-body nature of the Hamiltonians in quantum chemistry, the auxiliary bond dimension of a MPO is always less than k^2 , where k is the number of orbitals. (b) and (c) The tensors in wavefunctions Ψ_1 and Ψ_0 are represented by circles and triangles, respectively. The zeroth order Hamiltonian H_0 and perturbation operator V are represented by squares and diamonds, respectively. Panel (b) shows the three terms in the Hylleraas functional. Panel (c) shows the partial derivative of the Hylleraas functional $\partial H[\Psi_1] / \partial A_{i_{l-1} i_l}^{n_l}$ that is set to zero in the optimization. The red tensor, $A_{i_{l-1} i_l}^{n_l}$ is the quantity being solved for. (d) Contractions that need to be carried out at each step of the sweep iteration. The corresponding costs (assuming H_0 is Dyall's Hamiltonian⁴¹) are shown. For each term, individual contractions are numbered according to the order in which they are performed. If two contractions have the same number, it means that the two indices on the same tensor are fused to form a larger composite index and then the contraction is performed. The solid tensors represent the results of contracting all the open tensors in panel (c) that are not shown in this figure.

It is well known that perturbation theory can be formulated as a variational problem.^{37,39} For the second order energy, the variational functional is the Hylleraas functional $H[\Psi_1]$, Eq. (1), which is minimized with respect to the first order wavefunction $|\Psi_1\rangle$,

$$H[\Psi_1] = \langle \Psi_1 | H_0 - E_0 | \Psi_1 \rangle + 2\langle \Psi_1 | QV | \Psi_0 \rangle. \quad (1)$$

Here H_0 , E_0 , and $|\Psi_0\rangle$ are, respectively, the zeroth order Hamiltonian, energy, and wavefunction, V is the perturbation ($H - H_0$), and Q the projector onto the space orthogonal to $|\Psi_0\rangle$. It can be easily verified that at the minimum, the wavefunction $|\Psi_1\rangle$ satisfies the familiar equation

$$(H_0 - E_0)|\Psi_1\rangle = -QV|\Psi_0\rangle. \quad (2)$$

In MPS-PT2 we minimize the Hylleraas functional, while expressing $|\Psi_1\rangle$ as a MPS. MPS forms a complete variational space, and the quality of the MPS solution is controlled by a single parameter “ M ,” the dimension of the auxiliary bond in the MPS (see Figure 1). Overlaps ($\langle \Psi_1 | \Psi_2 \rangle$) and transition elements of operators between two MPSs ($\langle \Psi_1 | O | \Psi_2 \rangle$) can be evaluated in $O(M^3)$ central processing unit (CPU) time. The Hylleraas functional can be minimized with respect to a MPS of arbitrary bond dimension using a sweep algorithm analogous to that in the DMRG.⁴⁰ In the limit of large M , the solution converges to the exact (uncontracted) first order

wavefunction (and second order energy). From the first order wavefunction, the third order correction can also be calculated due to Wigner's $2n + 1$ rule, as

$$E_3 = \langle \Psi_1 | V | \Psi_1 \rangle - E_1 \langle \Psi_1 | \Psi_1 \rangle. \quad (3)$$

A general MPS representing $|\Psi_1\rangle$ is shown in Eq. (4), where n_i is the occupation of orbital i ,

$$|\Psi_1\rangle = \sum_{\{n\}, i_1 \dots i_{k-1}} A_{i_1}^{n_1} A_{i_1 i_2}^{n_2} \dots A_{i_{k-1}}^{n_k} |n_1 n_2 \dots n_k\rangle. \quad (4)$$

Eq. (4) is conveniently expressed graphically as a tensor network, as shown in panel (a) of Figure 1 (see Refs. 42–44 for a detailed explanation of the graphical notation). Similarly, any operator Ω can be written in matrix product operator (MPO) form,

$$\Omega = \sum_{\{n'\}, i_1 \dots i_{k-1}} W_{i_1}^{n_1 n'_1} W_{i_1 i_2}^{n_2 n'_2} \dots W_{i_{k-1}}^{n_k n'_k} |n_1 n_2 \dots n_k\rangle \langle n'_1 n'_2 \dots n'_k|, \quad (5)$$

and this is also expressed graphically in panel (a) of Figure 1. With this notation, the Hylleraas functional is shown in panel (b) of Figure 1, where the operators and wavefunctions have been represented as MPOs and MPSs, respectively, and the corresponding tensor networks are contracted to obtain the final expression.

In the sweep algorithm, each tensor $[A^n]$ of the MPS wavefunction $|\Psi_1\rangle$ is optimized in sequence, keeping the other tensors fixed. This converts the nonlinear optimization of the MPS (which is multi-linear in its parameters) to a series of linear equations that we solve using the conjugate gradient algorithm. At step l of the sweep, we take the partial derivative of the Hylleraas functional with respect to the set of tensor elements $A_{i_{l-1}i_l}^{n_l}$ of $|\Psi_1\rangle$, and set it to zero. The resulting linear equation is shown graphically in panel (c) of Figure 1 and has 3 terms, corresponding to the partial derivatives of the three terms in the Hylleraas functional. The first involves a contraction of $A_{i_{l-1}i_l}^{n_l}$ (the quantity being solved for, shown with a red circle) with a 6-index tensor, formed by contracting all tensors in the operator $H_0 - E_0$ and $|\Psi_1\rangle$ with $A_{i_{l-1}i_l}^{n_l}$ removed. Similarly, the second and third terms involve the tensors of V , $|\Psi_0\rangle$ and all tensors of $|\Psi_1\rangle$ with $A_{i_{l-1}i_l}^{n_l}$ removed. The resulting linear equation for $A_{i_{l-1}i_l}^{n_l}$ is in $O(dM^2)$ unknowns. Solving it by conjugate gradient requires only $O(M^3)$ time (due to the special structure of the operator-vector product) and yields the current best estimate of the tensor $A_{i_{l-1}i_l}^{n_l}$. It is clear that the tensor so obtained depends on the values of all the other tensors $[A^n]$ in $|\Psi_1\rangle$. Thus we sweep through all the orbitals of $|\Psi_1\rangle$ solving a linear algebra problem at each step $l = 1 \dots k$ to optimize each tensor, while keeping the others fixed. This algorithm is essentially the same as the standard DMRG sweep algorithm, the only difference being that instead of an eigenvalue problem, a linear equation is solved at each step.

The computational cost of the above sweep can be expressed in detail in terms of the auxiliary bond dimensions of $|\Psi_1\rangle$ and $|\Psi_0\rangle$ (M_1 and M_0 , respectively) and the number of active and external orbitals (k_a and k_v , respectively). For concreteness, we consider H_0 to be Dyall's Hamiltonian,⁴¹ as used in NEVPT2. All the contractions to be performed at each sweep iteration, together with their respective (leading order in M , k_a , k_v) computational costs, are shown in panel (d) of Figure 1. Contractions 1, 2, and 3 of the first two terms have the same leading scaling with respect to the number of orbitals even though the operator in the first term (denoted by squares) is a Dyall Hamiltonian H_0 , which does not couple the active and external spaces, and the operator V in the second term (denoted by diamonds) couples the two spaces. The low scaling of the second term is because components of V with more than 2 external indices do not contribute to the second order energy and can be neglected. The resulting MPO for V , under this restriction, can then always be represented with $O(k_a^2)$ auxiliary bond dimension, the same as the Dyall H_0 . The costs for contractions 1–5 in the third term are not shown because they are used to construct the bottom diagram, which is a constant that does not change during the sweep, and thus has a subleading dependence on k_v . The final computational cost of a single sweep is $O(k_v k_a^2 M_1^3 + k_v k_a^2 M_0 M_1^2 + k_v k_a^2 M_0^2 M_1)$, where we have multiplied the most expensive costs in each individual step shown in panel (d) of Figure 1 by k_v (the number of virtual orbitals) to reflect the total number of steps in the sweep, assuming $k_v + k_a \sim O(k_v)$. We have also dropped the dependence on d , which is a fixed constant (usually 4). (For brevity, we have not considered the cost of terms with

subleading (lower than cubic) dependence on M . These only become important when k_v is very large. For example, forming the MPO for V (through DMRG complementary operator techniques⁴⁵) costs $O(k_v^2 k_a^2 M_0 M_1)$ /sweep, which becomes important when $k_v > M$.)

The above costs assume the Dyall H_0 . A similar analysis applies to the CASPT2 H_0 , which takes the form of a one-body Fock operator multiplied by many-body projectors. The many-body projectors lead to additional overlap computations of the form $\langle \Psi_0 | \Psi_1 \rangle$, but these only cost $O(k_a M_0^2 M_1 + k_a M_0 M_1^2)$ for the entire sweep, which is independent of k_v . The one-body Fock operator leads to a lower overall computational cost for the first term in the Hylleraas functional involving H_0 . However, the second term involving V results in the same leading cost as for NEVPT2 above. Additionally, note that because we do not proceed via special reduced density matrix intermediates, the computations for both the Dyall H_0 in NEVPT and for the CASPT2 H_0 involve the same basic contractions of MPOs and MPSs and thus it is simple to change the zeroth order Hamiltonian in the implementation. This is similar to the simplicity of implementing fully uncontracted determinant algorithms, except here contraction is automatically provided by the finite bond dimension of the MPS. Here, we would also like to point out that this approach can be extended to excited state calculations by a small modification to the current procedure.³⁸

We have implemented the above MPS-PT sweep algorithm to evaluate the NEVPT2 and NEVPT3 energy in the BLOCK code.⁴⁶ For computational robustness, we have implemented the two-site version of the sweep algorithm (see Refs. 40 and 42 for a discussion of the difference between one-site and two-site sweep algorithms). In the limit of large M_1 , our MPS-PT algorithm will yield the uncontracted NEVPT2 and NEVPT3 energies. We have parallelized the algorithm exactly in the manner of the DMRG module of the BLOCK code, where each auxiliary bond contributing to the MPO (arising from a term in the operator/complementary operator product) is evaluated on a different processor. As explained previously, the auxiliary bond dimension of both H_0 and V is $O(k_a^2)$, thus the code is expected to efficiently parallelize over $O(k_a^2)$ processors. We refer the reader to the detailed description of DMRG parallelization in Ref. 47. There is an additional “embarrassing” parallelization over $O(k_v^2)$ processors which we mention (but have not implemented). This arises by breaking the summation in V into k_v^2 individual terms V_i . Each V_i produces a mutually orthogonal state when acting on $|\Psi_0\rangle$ and thus the Hylleraas functional can be computed and minimized for each V_i separately, giving a final second-order energy that is just the sum of these terms.

Symmetries of the Hamiltonian are easily utilized in an analogous fashion to the DMRG module of the BLOCK code, which currently implements many symmetries including particle number, Abelian and non-Abelian point groups, and SU(2) (spin symmetry).

We next illustrate various aspects of the theory by performing calculations on the N_2 bond dissociation benchmark using a full-valence CAS of 10 electrons in 8 orbitals and the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets.⁴⁸ Table I shows the energies of the N_2 molecule with various NEVPT2 and

TABLE I. Energy ($E+108.0$ in E_h) for a MPS-PT calculation performed with a cc-pVDZ basis set for the $^1\Sigma_g$ state of N_2 molecule at various bond-lengths. A (10e, 8o) active space was used. PC-NEVPT2, and SC-NEVPT2 results are given for comparison. The last four columns show the MPS-PT2 energy calculated using different auxiliary bond dimensions M_1 .

r_e (Å)	PC-NEVPT2	SC-NEVPT2	MPS-PT3 $M_1 = 800$	M_1 used in MPS-PT2			
				800	300	200	100
1.8	-1.1430	-1.1396	-1.1626	-1.1432	-1.1431	-1.1428	-1.1403
2.0	-1.2420	-1.2389	-1.2620	-1.2422	-1.2421	-1.2418	-1.2382
2.2	-1.2487	-1.2458	-1.2689	-1.2490	-1.2487	-1.2483	-1.2432
2.4	-1.2129	-1.2103	-1.2333	-1.2132	-1.2129	-1.2123	-1.2066
2.7	-1.1342	-1.1318	-1.1549	-1.1345	-1.1341	-1.1332	-1.1258
3.0	-1.0591	-1.0568	-1.0807	-1.0594	-1.0589	-1.0578	-1.0484
3.6	-0.9648	-0.9631	-0.9896	-0.9651	-0.9641	-0.9624	-0.9430
4.2	-0.9350	-0.9342	-0.9621	-0.9352	-0.9341	-0.9324	-0.9055

MPS-PT variants at several bond-lengths in the cc-pVDZ basis set. The MPS-PT2 energy is converged to a few tens of μE_h with $M_1 = 800$ and always gives an energy lower than the partially contracted NEVPT2 theory (since it is converging to the fully uncontracted result) although the energy difference is not large. The error of various standard flavors of contracted NEVPT2 theory, PC-NEVPT2 and SC-NEVPT2 (evaluated using MOLPRO⁴⁹ and MPS-PT2 calculations with smaller $M_1 = 100$ –300, are plotted against increasing N_2 bond-length in Figure 2. We see that the value of M_1 required to accurately describe the first order wavefunction $|\Psi_1\rangle$ increases somewhat with increasing bond length. This can be understood by realising that the first-order wavefunction contains contributions from $O(N_a k_a^2 k_v^2)$ determinants (where N_a is the number of significant determinants in the active space). As N_a increases at increased bond-lengths, the first-order wavefunction becomes more complicated and requires slightly larger M . Nonetheless, $M \ll N_a k_a^2 k_v^2$ and except for $M = 100$, the MPS-PT error remains small at all bond-lengths.

When one views the DMRG algorithm in the traditional renormalization group language, each tensor in the MPS (in canonical form) can be viewed as mapping from $4M$ states (for $d = 4$) to the M most significant renormalized states.

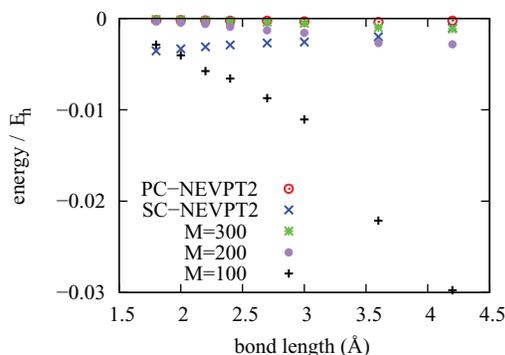


FIG. 2. Error in the second-order energy (in E_h) for various flavors of NEVPT2 and MPS-PT2 calculations performed with cc-pVDZ basis sets for the $^1\Sigma_g$ state of N_2 molecule at several bond lengths. A (10e, 8o) active space was used for these calculations. The graph shows that the value of M_1 required to capture the correct second-order energy using the MPS-PT2 theory increases with the N_2 bond length, most likely due to increase in the multi-reference character of the wavefunction. Nonetheless, for $M_1 > 100$, the error remains quite small at all bond-lengths.

In the two-site variant of the algorithm, the weight of these discarded $3M$ states measures the severity of the truncation performed in the renormalization step. The discarded weight goes to zero in the limit of large M and for smaller M usually exhibits a linear relationship with the energy error.^{40,50} Figure 3 shows that after an initial rapid relaxation in the energy, the error in the MPS-PT2 indeed scales linearly with the corresponding discarded weight in $|\Psi_1\rangle$. Thus the relationship between the energy error and discarded weight can be used to gauge the convergence of the MPS-PT2 energy with M , just as in the DMRG algorithm.

Figure 4 shows the relationship between the size of the basis set and the value of M_1 required to obtain a second-order energy within 1 m E_h of the exact fully uncontracted MPS-PT2 energy. We see that, encouragingly, the M_1 required scales sub-linearly with the size of the basis. Indeed, the entanglement in the basis set limit is bounded by a constant (the dimension of the active orbital Hilbert space). Thus, M_1 must asymptotically be independent of k_v .

To summarize, we have shown that minimizing the Hylleraas functional in the space of wavefunctions described by a MPS can be used to obtain accurate upper bounds to the exact multi-reference second order perturbation theory energy. With little additional computational overhead, third

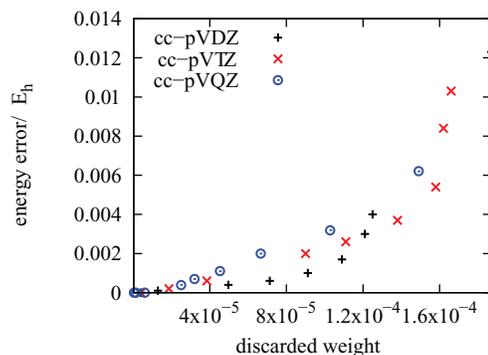


FIG. 3. Error in the second-order energy (in E_h) versus the discarded weight for a MPS-PT2 calculation performed with cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets for the $^1\Sigma_g$ state of N_2 molecule at a bond length of 2 Å. A (10e, 8o) active space was used for these calculations. After an initial rapid relaxation of the energy one recovers the linear relationship between the energy and the discarded weight.

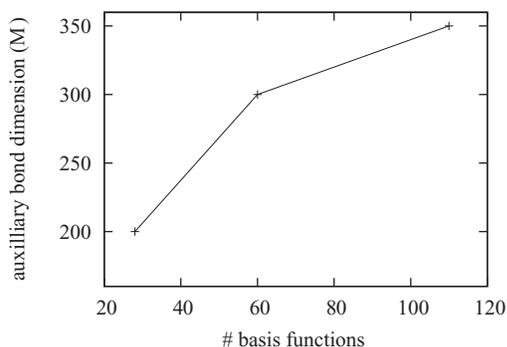


FIG. 4. The size of the auxiliary bond dimension M of the first order wavefunction required to obtain a second-order energy within 1 mE_h of the exact fully uncontracted MPS-PT2 energy.

order corrections to the energies can also be obtained. No active space density matrices are required, allowing large active spaces to be treated, and our implementation can be easily modified to accommodate any reasonable zeroth order hamiltonian. Parallelism is also easily exploited. In benchmark calculations we have shown that this approach efficiently recovers the uncontracted perturbation theory energy, beyond existing internally contracted multi-reference perturbation results. Finally, we observe that the bond dimension of the first order wavefunction required to obtain chemical accuracy scales sub-linearly with the size of the basis set, which is promising for future applications with very large basis sets.

This work was supported by the US National Science Foundation (NSF) through Grant No. NSF-CHE-1265277. Additional support for software development was provided through Grant No. NSF-OCI-1265278. S.S. would also like to thank Brecht Verstichel for many helpful discussions.

¹S. R. White and R. L. Martin, *J. Chem. Phys.* **110**, 4127 (1999).

²G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 (2011).

³Ö. Legeza, R. Noack, J. Sólyom, and L. Tincani, *Computational Many-Particle Physics*, Lecture Notes in Physics Vol. 739, edited by H. Fehske, R. Schneider, and A. Weie (Springer, Berlin, Heidelberg, 2008), pp. 653–664.

⁴K. H. Marti and M. Reiher, *Phys. Chem. Chem. Phys.* **13**, 6750 (2011).

⁵Y. Kurashige and T. Yanai, *J. Phys. Chem.* **130**, 234114 (2009).

⁶S. Wouters and D. Van Neck, “The density matrix renormalization group for *ab initio* quantum chemistry,” e-print [arXiv:1407.2040v1](https://arxiv.org/abs/1407.2040v1) (2014).

⁷Y. Kurashige, G. K.-L. Chan, and T. Yanai, *Nat. Chem.* **5**, 660 (2013).

⁸S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan, *Nat. Chem.* (published online 2014).

⁹G. H. Booth, D. Cleland, A. Alavi, and D. P. Tew, *J. Chem. Phys.* **137**, 164112 (2012).

¹⁰G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 54106 (2009).

¹¹G. H. Booth, S. D. Smart, and A. Alavi, *Mol. Phys.* **112**, 1855 (2014).

¹²J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).

¹³P. A. Malmqvist, A. Rendell, and B. O. Roos, *J. Phys. Chem.* **94**, 5477 (1990).

¹⁴M. Kállay and P. R. Surján, *J. Chem. Phys.* **115**, 2945 (2001).

¹⁵D. W. Small and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **13**, 19285 (2011).

¹⁶P. Piecuch, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **110**, 6103 (1999).

¹⁷J. Olsen, *J. Chem. Phys.* **113**, 7140 (2000).

¹⁸H. Nakatsuji and K. Yasuda, *Phys. Rev. Lett.* **76**, 1039 (1996).

¹⁹D. A. Mazziotti, *Chem. Rev.* **112**, 244 (2012).

²⁰B. Verstichel, H. van Aggelen, D. Van Neck, P. W. Ayers, and P. Bultinck, *Phys. Rev. A* **80**, 032508 (2009).

²¹Y. Kurashige and T. Yanai, *J. Chem. Phys.* **135**, 094104 (2011).

²²S. Sharma, T. Yanai, G. H. Booth, C. J. Umrigar, and G. K.-L. Chan, *J. Chem. Phys.* **140**, 104112 (2014).

²³T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, *J. Chem. Phys.* **132**, 024105 (2010).

²⁴M. Saitow, Y. Kurashige, and T. Yanai, *J. Chem. Phys.* **139**, 044118 (2013).

²⁵D. Ma, G. Li Manni, and L. Gagliardi, *J. Chem. Phys.* **135**, 044128 (2011).

²⁶S. Ten-no, *Theor. Chem. Acc.* **131**, 1070 (2012).

²⁷M. Torheyden and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008).

²⁸L. Kong, F. A. Bischoff, and E. F. Valeev, *Chem. Rev.* **112**, 75 (2011).

²⁹W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006).

³⁰T. Yanai and T. Shiozaki, *J. Chem. Phys.* **136**, 84107 (2012).

³¹T. B. Adler, H.-J. Werner, and F. R. Manby, *J. Chem. Phys.* **130**, 054106 (2009).

³²K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990).

³³B. O. Roos, *Acc. Chem. Res.* **32**, 137 (1999).

³⁴C. Angeli, R. Cimraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, *J. Chem. Phys.* **114**, 10252 (2001).

³⁵I. Schapiro, K. Sivalingam, and F. Neese, *J. Chem. Theory Comput.* **9**, 3567 (2013).

³⁶F. Liu, Y. Kurashige, T. Yanai, and K. Morokuma, *J. Chem. Theory Comput.* **9**, 4462 (2013).

³⁷O. Sinanoğlu, *J. Chem. Phys.* **34**, 1237 (1961).

³⁸O. Sinanoğlu, *Phys. Rev.* **122**, 491 (1961).

³⁹E. Hylleraas, *Z. Phys.* **65**, 209 (1930).

⁴⁰G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).

⁴¹K. G. Dyall, *J. Chem. Phys.* **102**, 4909 (1995).

⁴²U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).

⁴³G. K.-L. Chan, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 907 (2012).

⁴⁴R. Orús, *Ann. Phys.* **349**, 117 (2014).

⁴⁵T. Xiang, *Phys. Rev. B* **53**, 10445 (1996).

⁴⁶S. Sharma and G. K.-L. Chan, *J. Chem. Phys.* **136**, 124121 (2012).

⁴⁷G. K.-L. Chan, *J. Chem. Phys.* **120**, 3172 (2004).

⁴⁸T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).

⁴⁹H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 242 (2012).

⁵⁰Ö. Legeza and G. Fáth, *Phys. Rev. B* **53**, 14349 (1996).