UC Berkeley
UC Berkeley Previously Published Works

Title

A deterministic alternative to the full configuration interaction quantum Monte Carlo method.

Permalink

Journal

ISSN

Authors

Tubman, Norm M
Lee, Joonho
Takeshita, Tyler Y
et al.

Publication Date

DOI

Peer reviewed

A deterministic alternative to the full configuration interaction quantum Monte Carlo method

Norm M. Tubman, Joonho Lee, Tyler Y. Takeshita, Martin Head-Gordon, K. Birgitta Whaley[1]

University of California, Berkeley, Berkeley, CA 94720, USA

(Dated: June 20, 2016)

Development of exponentially scaling methods has seen great progress in tackling larger systems than previously thought possible. One such technique, full configuration interaction quantum Monte Carlo, is a useful algorithm that allows exact diagonalization through stochastically sampling determinants. The method derives its utility from the information in the matrix elements of the Hamiltonian, along with a stochastic projected wave function, to find the important parts of Hilbert space. However, the stochastic representation of the wave function is not required to search Hilbert space efficiently, and here we describe a highly efficient deterministic method that can achieve chemical accuracy for a wide range of systems, including the difficult Cr_2 molecule. We demonstrate for systems like Cr_2 that such calculations can be performed in just a few cpu hours which makes it one of the most efficient and accurate methods that can attain chemical accuracy for strongly correlated systems. In addition our method also allows efficient calculation of excited state energies, which we illustrate with benchmark results for the excited states of C_2.

Introduction: The scope of traditional approaches to full configuration interaction (FCI) has been limited to simple diatomic molecules [1,2], and there has been little progress in diagonalizing spaces much larger than a billion determinants.[3–5] However, recent progress in alternative approaches to FCI problems has increased the scope of FCI beyond simple diatomic molecules. Two techniques in particular have been important in this progress, full configuration quantum Monte Carlo (FCIQMC)[6], and density matrix renormalization group (DMRG).[7–9] Both algorithms provide unique advantages, with DMRG being the definitive method for systems in which one can identify degrees of freedom with low levels of entanglement[10,11], and FCIQMC showing promise for molecules and extended systems in two or more dimensions.[12,13] The success of DMRG and FCIQMC in quantum chemistry is highlighted by their recent applications to unprecedented large-size determinant spaces while also achieving chemical accuracy.[12,14–22]

The FCIQMC method is a useful technique with a few limitations which include biased sampling and comparatively computationally expensive simulations. The biased sampling is a result of the initiator approximation[23] generally employed in FCIQMC calculations, which additionally limits the space in which determinants can be sampled. We will denote the initiator algorithm as iFCIQMC. The initiator approximation can cause systematic errors in iFCIQMC calculations that can be unexpectedly large, as evidenced by comparison to variational DMRG results.[24] The need for this approximation is related to the Monte Carlo sampling and not necessarily related to power of the technique which we suggest is finding important determinants.

Here we suggest an alternative to stochastic sampling in favor of a completely deterministic version of the FCIQMC technique that efficiently samples the determinant space. We denote this deterministic algorithm by Adaptive Sampling CI (ASCI). ASCI does not have the same problem with quantum entanglement growth that DMRG has for 2 and 3 dimensional systems, and is also appears faster than FCIQMC in computational efficiency. Our approach also surpasses other techniques for finding energetically important determinant subspaces. The majority of traditional CI methods encode relevant physical degrees of freedom based on excitation levels from a reference determinant.[25] These excitation-based methods can also suffer from inaccuracy as they miss important parts of determinant space. Within the CI framework several promising ways to circumvent this problem have been suggested that focus on selected CI approaches where one selects relevant determinants based on different criteria.[26–41] After presenting the ASCI method, we establish a connection between FCIQMC and various selected CI approaches. We then apply the ASCI method to Cr_2, a classic hard problem for many computational electronic structure methods.[24] Finally, we demonstrate that excited states are straightforward to calculate with ASCI. We note that the calculation of excited states within stochastic FCIQMC is possible, but requires specialized techniques[42], or stochastic orthogonalization between walker sets, which is quite different from the method described here.

A path to a deterministic algorithm: In the initial development of FCIQMC, one of the original improvements on the method was to take part of the projection step and make it deterministic.[20] In this work, we go further and develop a completely deterministic algorithm. Our approach here is to find important determinants in the same manner as FCIQMC, i.e. to sample determinants based on the absolute value of the ground state wave function amplitudes. The FCIQMC technique was originally presented as a projector method in imaginary time and we use this approach to motivate our method. We start by expanding a wave function in the space of determinants,

$$\psi(\tau) = \sum_i C_i(\tau)|D_i\rangle, \qquad (1)$$

and the propagator in imaginary time,

$$-\frac{dC_i}{d\tau} = (H_{ii} - E)C_i + \sum_{i \neq j} H_{ij}C_j, \qquad (2)$$

which has an asymptotic solution of a stationary state with $\frac{dC_i}{d\tau} = 0$. In these equations τ is the imaginary time, i is an

index for determinants, D_i are determinants, C_i is the amplitude associated with determinant D_i, and H_{ij}, is the Hamiltonian matrix element $\langle D_i|H|D_j\rangle$. In FCIQMC, the parameter E is a free parameter that controls the population of walkers. Here we will consider E to be the ground state energy or our best approximation thereof. The power of FCIQMC is that it ignores the unimportant parts of determinant space, finds important determinants, and samples them according to their amplitudes. For a stationary state we can solve for the individual coefficients as $C_i = -\frac{\sum_{i\neq j} H_{ij}C_j}{(H_{ii}-E)}$.

The RHS of this equation captures all aspects of the FCIQMC algorithm. The H_{ij} in the numerator corresponds to the spawning step (transition moves between determinants), the sign of H_{ij} in the numerator and specifically summing over positive and negative terms corresponds to the annihilation step (cancelation of positive and negative walkers), and the denominator corresponds to the death/cloning step (adding or removing walkers from the simulation). The key to turning FCIQMC into a non-stochastic algorithm is to remove the stochastic sampling and replace it with a deterministic ranking of important determinants. In both approaches the Hamiltonian matrix elements and a wave function are needed. In FCIQMC the wave function is represented by the distribution of walkers at any given step, whereas for our deterministic algorithm, we use an approximate wave function at each iteration as follows,

$$C_i^1 = \frac{\sum_{j\neq i} H_{ij}C_j^0}{(H_{ii}-E)}. \tag{3}$$

In this equation, we have labeled the coefficients of the initial wave function as C^0, and the output coefficients as C^1. The C^0 and C^1 coefficients correspond to different determinant spaces, where D^1 is the determinant space that includes everything in D^0 plus all single and double excitations from the determinants in D^0. Thus for any good approximation to the ground state wave function we can use Eq. 3 to determine the importance of determinants in a much larger space that what is initially included in D^0. This C^1 coefficients are essentially a first-order perturbation estimate for CI coefficients in the Epstein-Nesbet perturbation theory.[31]

The Deterministic Algorithm: The ASCI algorithm is defined by two determinant subspaces: a core space of size *cdets* and a target space of size *tdets*. The *core space* determines the number of terms j we include in the sum in Eq. 3. This is to say we select cdet determinants with the largest C_j^0, and consider only those to be non-zero in the sum of equation Eq. 3. The space to be searched is the set of single and double excitations of the core set of determinants. Since our objective is to find the determinants with the largest amplitudes, we in general only need to search determinants connected to those with large amplitudes. We illustrate the use of this approximation with numerical tests in the next section. The *target space* contains the top *tdet* determinants, as determined from the ranking, and is the rank of the matrix diagonalized in each iteration.

Initalize: set size of the core space, D^{core}, to size (*cdets*) and the target space, D^{target}, to size (*tdets*). Set the starting wave function to the Hartree Fock wave function ($C_0 = 1$, $C_{i>0} = 0$; E=E_{HF}).

(1) Evaluate the perturbed wave function amplitudes over all the single and double substitutions, D^{SD}, from the core space, D^{core}.

$$A_i = \frac{\sum_{j\neq i}^{core} H_{ij}C_j}{H_{ii}-E}. \tag{4}$$

(2) From D^{core} and D^{SD}, use the corresponding coefficients $\{C_i\}$ and $\{A_i\}$ to select the *tdet* largest absolute values to define the new target space, D^{target}.

(3) Form and diagonalize H in the target space.

(4) The lowest eigenvalue is the new energy E. The largest *cdet* amplitudes, by magnitude, define the new core space. If the energy is not converged, return to step (1). For simulations presented here, we halted the simulations when the energy converged to 10^{-5} Ha. We have included a convergence plot of Cr_2 for (tdets=1,000,000,cdets=5,000) in Fig. 1.

Figure 1. (Color online): Convergence of the energy with the ASCI algorithm for Cr_2 in the SV basis, frozen core with (24e,30o), for (tdets=1,000,000,cdets=5,000). The energy is calculated relative to our final energy (-2086.4038876399 Ha) in the last step (step 7).

The determinants found at the end of the simulation will in general be the most important determinants for the ground state wave function. Unlike FCIQMC this technique has no population control bias, initiator bias, and no sign problem. The technique described here provides an inherently variational energy. However, it is possible to extend the accuracy of the technique with perturbation theory, which comes at the expense of the energy no longer being variational.[43] We use the following second order correction to the energy,

$$E_{PT2} = \sum_i \frac{\langle\psi|H|D_i\rangle}{H_{ii}-E}. \tag{5}$$

We note that other perturbative corrections can be considered.[43] For the largest systems considered here, only a modest amount of memory is needed, and all parts except the diagonalization step can be trivially parallelized.

Discussion: When performing an ASCI calculation the first few steps involve the exploration of higher order excitations. If the starting wave function is a single determinant, then each step increases the maximum number of excitations that have been explored by 2. The coefficients A, as calculated in step (1), can span all single and double excitations from the current wave function, and is generally very large. Truncating the coefficients that are calculated for A in the main part of the self-consistent loop allows us to consider applying our algorithm to large systems. To maintain size consistency, the value of *tdets* and *cdets* will have to grow with system size. For the systems considered in this work, it was possible to converge the value of *cdets*, as the energy with respect to this parameter can be extrapolated by running with a few different values.

Apart from the diagonalization step, the most computationally expensive task is determining A and finding its largest

values. This can be split among many processors, with the only non-trivial communication occurring during the final aggregation of the final values.

Further improvements are possible by generating the natural orbitals after an initial run and using them to recalculate the electron integrals. The natural orbitals can be generated at any point in the simulation from the current best wave function. The natural orbitals are generally thought to produce compact representations of a wave function.[44] For a difficult system like Cr_2, we find the use of natural orbitals to be crucial to obtain accurate energies.

Before we present our results we consider numerical tests for our approximation of determining the magnitude of A. In general, the coefficients of the ground state wave function will span many orders of magnitude, and searching over the determinants with the smallest amplitude coefficients is inefficient. Using *cdets* as a parameter to limit the search to only the important determinants appears to be a well controlled approximation that can be converged for all the system we tested. This approach has similarities to the FCIQMC initiator approximation.[23]

We demonstrate the accuracy of this approximation by considering the CN molecule in an STO-3G basis at 1.194 Å (6240 determinants). We show that we can find the most important determinants in Hilbert space without having to perform a diagonalization over a larger determinant space. Table I provides a comparison of the most important determinants found with different values of *cdets* and *tdets* to those obtained from a full diagonalization over the entire space. For a simulation of 100 core determinants and a target space of 200 determinants (100/200) we found 182 of the top 200 determinants (91%). The remaining 9% of missed determinants were found to be close in amplitude to the determinants that replaced them in the target space. Similar result can be seen for for all the simulations presented. These results suggest that for some simulations that the search algorithm isn't highly dependent on the size of the core space and the small percentage of determinants that are missed by the algorithm are replaced by determinants that are similar in importance. Thus we argue that extremely high accuracy is not needed in determining the ranking order. The errors of this approximation is even further reduced when perturbation corrections are used, which can correct for any important determinants that were missed.

Benchmark results: Our main goal in presenting this analysis is to demonstrate that a deterministic method is capable of exploring determinant space in a similar manner to FCIQMC. We did not focus on algorithmic speed and we only provide timings here to suggest an upper bound of what can be expected. For the results presented here we built our own implementation of this algorithm and incorporated tools from various codes[45–47] and electron integrals from various packages.[48,49]

C_2 *with the cc-pVDZ basis set*: For C_2 simulations, the convergence of energies to chemical accuracy was easily achieved, and neither the use of natural orbitals nor perturbation corrections are needed. In comparison with the exact results[4], we were able to achieve an accuracy of 1 mHa using a diagonalization no larger than 200,000 determinants. The

cdets	tdets	energy	top 200	top 400	top 800
100	200	-91.17389	182 (91%)	200	200
100	400	-91.17637	200	373 (93%)	400
100	800	-91.17749	200	400	724 (91%)
200	400	-91.17657	200	387 (97%)	400
200	800	-91.17753	200	400	758 (95%)
400	800	-91.17755	200	400	776 (97%)
FCI		-91.17767			

Table I. Test of the CN molecule in finding important determinants. The energies are in units of Ha. The columns with 'top 200; is the number of determinants we found that agree with the top 200 determinants, by amplitude, of the exact answer. Likewise for 'top 400' and 'top 800'. The total FCI space is 6240 determinants for STO-3G at 1.194 Å.

cdets	tdets	energy
4,000	10,000	-75.71842
4,000	50,000	-75.72626
4,000	100,000	-75.72786
8,000	100,000	-75.72795
4,000	200,000	-75.72878
10,000	200,000	-75.72891
20,000	200,000	-75.72895
20,000	300,000	-75.72928
15,000	500,000	-75.72953
FCI	27,900,000	-75.72985

Table II. Energy of C_2 molecule in units of Hartrees, at bond length 1.27273 Å. The size of the determinant space is given by the D_{2h} point group and with a frozen core. The benchmark results for this molecule with cc-pVDZ basis set is given from the following reference.[4,6]

total computer time for a simulation of this size was less than 2 cpu hours. Results are presented in Table II, using different values of *tdets* and *cdets*. We also calculated the binding curve for C_2 with our method and compare to FCI results[50] in Fig. 2. The FCI results were reported to 0.1 mHa accuracy in a previous benchmark. We compare to a subset of these points and our ASCI+PT2 results agree exactly to this level of accuracy. We also ran a cc-pVTZ simulation at the equilibrium distance, and found the ASCI+PT2 energy to be -75.80919 Ha in comparison to the DMRG(M=6000) result of -75.809285 Ha, a difference of 0.1 mHa.

Figure 2. (Color online): C_2 binding curve. The blue curve is the exact FCI result[50], the green curve is the ASCI simulation and the red curve is the ASCI+PT2. For all ASCI calculations tdets and cdets were fixed to be 300,000 and 30,000 respectively.

Cr_2 *with the SV basis set*: Fig. 3 shows the convergence of our results for Cr_2. In order to make a comparison with previous studies[24], Cr_2 calculations were carried out with the SV basis set[51] at 1.5 Å with 24 active electrons in 30 orbitals and a frozen core. A compact representation of the wave function

in this system is dependent on having a good set of orbitals. As part of our algorithm, we run a preliminary calculation with a target space of 100,000 determinants, after which we calculated the natural orbitals. The resulting natural orbitals were used to recalculate the integrals for the production run. The total energy for our most accurate simulation converged to within 16 mHa of the predicted full CI basis set energy.[24]

Figure 3. (Color online): Cr_2 energy as a function of tdet size, with the largest target space going up to 1 million determinants. The plotted energies have been shifted by 2086 Ha. The PT2 correction line represents our result with the added perturbation theory correction, which brings our final energy to within 1 mHa of the predicted exact result. The first data point starts with 10,000 determinants. Our best energy without the perturbation correction is -2086.40388 Ha. With the perturbation correction it is -2086.4203 Ha. The DMRG benchmark is -2086.420948 Ha.[24]

Figure 4. (Color online): Cr_2 binding curve. The green curve shows the ASCI+PT2 energies and the red curve shows the DRMG energies for (24e,30o). For all ASCI calculations tdets and cdets were fixed to be 300,000 and 30,000 respectively. The blue curve are ASCI energies without the frozen core (48e,42o). The extrapolated DMRG energy for (48e,42o) at 1.5 Å is -2086.44478 Ha[24] compared to -2086.44325 Ha with ASCI+PT2.[52]

A perturbation theory analysis[43] was performed using equation 5, which results in a final energy within 1 mHa of the predicted exact result. The timing for the largest simulation ($tdets=10^6$), including the initial run for calculating the natural orbitals but not including the perturbation correction, requires approximately 7 cpu hours when run on a single core of a 2.40 GHz Intel Xeon processor. We expect further improvements will speed up the ASCI simulations significantly.

In Fig. 4, we plot the binding curve of Cr_2 for DMRG(M=1600) and ASCI+PT2 for (24e,30o). The DMRG(M=1600) calculations are taken from a previous benchmark study.[53] We also plot the ASCI+PT2 for the non-frozen core calculation of (48e,42o).[52] The ASCI results were calculated with (tdets=300,000,cdets=30,000). At equilibrium both ASCI results match the extrapolated DMRG results to less than 1mHa accuracy, as given in a previous references.[24] For the (24e,30o) calculation, ASCI is more accurate than the DMRG(M=1600). As the bond is stretched, it is clear ASCI is losing some accuracy with regards to DMRG(M=1600),

To demonstrate the distribution of the determinants located by ASCI we plot a histogram of excitations from the dominant determinants for Cr_2 in Fig. 5. The ratio between different excitations sectors does not change much in increasing $tdet$ values from 10^5 to 10^6. For this range, the quadruple excitations generally make up half the wave function, while higher excitations above the quadruples make up roughly 30% of the wave function.

Excited states: We are also able to calculate excited states with our technique, as they are obtained automatically within the diagonalization procedure. For a 6-31G* basis of C_2, we compare against previous FCI simulations for the first two excited states.[54] For a simulation at a distance of 1.25 Å with cdets=10^4 and tdets=800000, we have the fol-

Figure 5. (Color online): The distribution of determinants of Cr_2 by excitations from the dominant determinant. The x-axis is excitation level from the dominant determinant, and the y-axis is the fraction of determinants. Plots (a), (b), (c), (d) have tdets = 100k, 400k, 800k, and 1 million respectively. Simulations with different tdet spaces (for the ones shown here) have roughly the same fractional importance in the different excitation levels. Thus even for our smallest tdet simulations, large determinant excitations are important.

lowing energies for the ground state and the first two excited states $(-75.7256, -75.6345, -75.6271)$ in units of Ha. The exact results are $(-75.725995, -75.636861, -75.628883)$.[54] Thus, although the accuracy of the excited states is not as good as the ground state, it is still straightforward to achieve chemical accuracy (which is generally defined as 0.0016 Ha or 1 kcal/mol).

We can improve the accuracy of excited states by noting that any eigenstate can be used in Eq. 3. Thus we can find and rank determinants by their importance to individual excited states. The excited state optimization can be done simultaneously with the ground state method, or in a state-by-state bootstrap method. For the simultaneous optimization algorithm, we determine a set of important determinants to retain for each excited state. There is a separate search step for each excited state, but one diagonalization step that combines all retained determinants. In contrast, the bootstrap method would converge each excited state one by one, where at each step determinants would be added in specifically for the targeted excited state. The use of natural orbitals averaged over various excited states may also be used to improve the algorithm.[37,55] A detailed study of the targeted excited state technique will be presented elsewhere.

Connection to selected CI: As mentioned earlier, the ASCI method may also be considered a variant of selected CI techniques[31,32,35,39] in which there is considerable current interest.[26,41,43,56,57] Our method employs first-order perturbation coefficients for selecting determinants, and is thus related to the CI method, CIPSI (Configuration Interaction by Perturbation with multiconfigurational zeroth-order wave functions Selected by Iterative process).[31] Another related technique is the Λ+SD-CI[26] method which uses a one-step energy criteria, and a one-step approach together with our Eq. 3 in order to find important determinants. Despite these similarities, none of these algorithms have been pushed to achieve chemical accuracy for strongly correlated systems, and do not appear to have been benchmarked against FCIQMC or DMRG. The largest Λ+SD-CI calculations included roughly 50,000 determinants and attained 1–3 mHa accuracy for C_2 6-31G* (comparable to our results in Table II). As shown in this work, it is easy to go more than an order of magnitude in accuracy using our iterative scheme, without significantly increasing the computational effort. The largest selected CI techniques we are aware of have been extended up to 4 million determinants[58], but for systems in which no benchmarks exist. Monte Carlo Configuration Interaction method (MCCI), is another related technique that can be used to solve CI problems with random sampling in determinant space, however, it has been benchmarked far less than other techniques.[59,60]

For the future development of ASCI and other selected CI techniques, it is important to consider how such methods are different from standard CI methods. The difference is largely due to the construction of the Hamiltonian. Selected CI techniques need unique data structures in order to construct the Hamiltonian efficiently.[58,61] A previous study demonstrated that much larger scale simulations, than what we presented here, is possible for selected CI techniques.[58] We are currently considering various data structures used previously[58,62] and new structures, to determine the best way to scale up our simulations.

We also note that another group contacted us about improving the techniques presented here, and their work has since been published.[63] Their results show interesting behavior with regards to how many determinants are needed to converge a calculation with respect to increasing system size. Understanding this behavior will ultimately be important in determining just how far we can push ASCI to simulate larger systems.

Conclusions: We have shown that the underlying dynamics of FCIQMC can be used to generate a deterministic algorithm that can be efficiently used to calculate both ground and excited states of chemical systems. We have applied this technique to a known difficult problem in electronic structure theory, the Cr_2 molecule, and shown that chemical accuracy can be achieved with the cpu power available on any modern computer. Our results suggest that the ASCI method (and selected CI methods in general) should be considered as a state of the art CI method in both accuracy and efficiency. The importance of ASCI and FCIQMC is evident as DMRG and post-DMRG methods are not yet well suited for simulations in two and three dimensions. The ASCI method also distinguishes itself from FCIQMC in that excited states and other properties, such as the 2-RDM, are inherently easy to calculate.[64,65]

I. ACKNOWLEDGEMENTS

REFERENCES

[1] Z. Gan and R. Harrison, in *Supercomputing, 2005. Proceedings of the ACM/IEEE SC 2005 Conference* (2005) pp. 22–22.

[2] Z. Gan, D. J. Grant, R. J. Harrison, and D. A. Dixon, J. Chem. Phys. **125**, 124311 (2006).

[3] J. Olsen, O. Christiansen, H. Koch, and P. Jorgensen, J. Chem. Phys. **105**, 5082 (1996).

[4] M. L. Leininger, W. D. Allen, H. F. Schaefer, and C. D. Sherrill, J. Chem. Phys. **112**, 9213 (2000).

[5] J. Olsen, P. Jorgensen, H. Koch, A. Balkova, and R. J. Bartlett, J. Chem. Phys. **104**, 8007 (1996).

[6] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[7] S. R. White, Phys. Rev. B **48**, 10345 (1993).

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

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

[10] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[11] E. Stoudenmire and S. R. White, Annual Review of Cond. Mat. Phys. **3**, 111 (2012).

[12] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, Nature **493**, 365 (2013).

[13] J. J. Shepherd, G. H. Booth, and A. Alavi, J. Chem. Phys. **136**, 244101 (2012).

[14] C. Daday, S. Smart, G. H. Booth, A. Alavi, and C. Filippi, J. Chem. Theory Comput. **8**, 4441 (2012).

[15] R. E. Thomas, G. H. Booth, and A. Alavi, Phys. Rev. Lett. **114**, 033001 (2015).

[16] D. Cleland, G. H. Booth, C. Overy, and A. Alavi, J. Chem. Theory Comput. **8**, 4138 (2012).

[17] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).

[18] W. Mizukami, Y. Kurashige, and T. Yanai, J. Chem. Theory Comput. **9**, 401 (2013).

[19] S. Sharma, K. Sivalingam, F. Neese, and G. K. L. Chan, Nature Chem. **6**, 927 (2014).

[20] F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012).

[21] A. Holmes, H. J. Changlani, and C. J. Umrigar, ArXiv e-prints (2015), arXiv:1512.03757 [cond-mat.str-el].

[22] M. H. Kolodrubetz, J. S. Spencer, B. K. Clark, and W. M. Foulkes, J. Chem. Phys. **138**, 024110 (2013).

[23] D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[24] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, J. Chem. Phys. **142**, 034102 (2015).

[25] C. D. Sherrill and H. F. S. III. (Academic Press, 1999) pp. 143 – 269.

[26] F. A. Evangelista, J. Chem. Phys. **140**, 124114 (2014).

[27] C. F. Bender and E. R. Davidson, Phys. Rev. **183**, 23 (1969).

[28] S. R. Langhoff, S. T. Elbert, and E. R. Davidson, Int. J. Quantum Chem. **7**, 999 (1973).

[29] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33 (1974).

[30] R. J. Buenker, S. D. Peyerimhoff, and W. Butscher, Mol. Phys. **35**, 771 (1978).

[31] B. Huron, J. P. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[32] S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).

[33] R. Cimiraglia and M. Persico, J. Comput. Chem. **8**, 39 (1987).

[34] F. Illas, J. Rubio, J. M. Ricart, and P. S. Bagus, J. Chem. Phys. **95**, 1877 (1991).

[35] R. J. Harrison, J. Chem. Phys. **94**, 5021 (1991).

[36] J. Daudey, J. Heully, and J. Malrieu, J. Chem. Phys. **99**, 1240 (1993).

[37] F. Neese, J. Chem. Phys. **119**, 9428 (2003).

[38] R. Roth, Phys. Rev. C **79**, 064324 (2009).

[39] N. Ben Amor, F. Bessac, S. Hoyau, and D. Maynau, J. Chem. Phys. **135**, 014101 (2011).

[40] E. Giner, A. Scemama, and M. Caffarel, J. Chem. Phys. **142**, 044115 (2015), http://dx.doi.org/10.1063/1.4905528.

[41] W. Liu and M. R. Hoffmann, J. Chem. Theory Comput. **12**, 1169 (2016).

[42] G. H. Booth and G. K.-L. Chan, J. Chem. Phys. **137**, 191102 (2012).

[43] G. Emmanuel, S. Anthony, and C. Michel, Can. J. Chemistry **91**, 879 (2013).

[44] A. Szabo and N. Ostlund, *Modern Quantum Chemistry* (Dover, 1982).

[45] J. S. Spencer, N. S. Blunt, W. A. Vigor, F. D. Malone, W. M. C. Foulkes, J. J. Shepherd, and A. J. W. Thom, ArXiv e-prints (2014), arXiv:1407.5407 [cs.SE].

[46] A. Scemama and E. Giner, ArXiv e-prints (2013), arXiv:1311.6244 [physics.comp-ph].

[47] K. Wu and H. Simon, SIAM J. Sci. Comput. **22**, 602 (2000).

[48] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 556 (2012).

[49] F. Neese, Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 73 (2012).

[50] G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **135**, 084104 (2011).

[51] A. Schafer, H. Horn, and R. Ahlrichs, J. Chem. Phys. **97**, 2571 (1992).

[52] See supplemental material at [URL will be inserted by AIP] for inforation pertaining to the orbitals used for the (48e,42o) calculation at 1.5 Å.

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

[54] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **121**, 9211 (2004).

[55] Y. Shu, E. G. Hohenstein, and B. G. Levine, J. Chem. Phys. **142**, 024102 (2015).

[56] M. Caffarel, E. Giner, A. Scemama, and A. Ramirez-Solis, J. Chem. Theory Comput. **10**, 5286 (2014).

[57] M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, J. Chem. Phys. **144**, 151103 (2016).

[58] P. Stampfuß and W. Wenzel, J. Chem. Phys. **122**, 024110 (2005).

[59] J. C. Greer, J. Chem. Phys. **103**, 1821 (1995).

[60] J. P. Coe and M. J. Paterson, J. Chem. Phys. **137**, 204108 (2012).

[61] P. G. Szalay, T. Muller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2012).

[62] B. Engels, M. Hanrath, and C. Lennartz, C. Comput. Chem. **25**, 15 (2001).

[63] J. B. Schriber and F. A. Evangelista, J. Chem. Phys. **144**, 161106 (2016).

[64] R. E. Thomas, D. Opalka, C. Overy, P. J. Knowles, A. Alavi, and G. H. Booth, J. Chem. Phys. **143**, 054108 (2015).

[65] N. S. Blunt, S. D. Smart, G. H. Booth, and A. Alavi, J. Chem. Phys. **143**, 134117 (2015).