A general second order complete active space self-consistent-field solver for large-scale systems

Qiming Sun[a], Jun Yang[b], Garnet Kin-Lic Chan[a,*]

[a]*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125*
[b]*Department of Chemistry, University of Hong Kong, Pokfulam Road, Hong Kong, China*

Abstract

We present a new second order complete active space self-consistent field implementation to converge wavefunctions for both large active spaces and large atomic orbital (AO) bases. Our algorithm decouples the active space wavefunction solver from the orbital optimization in the microiterations, and thus may be easily combined with various modern active space solvers. We also introduce efficient approximate orbital gradient and Hessian updates, and step size determination. We demonstrate its capabilities by calculating the low-lying states of the Fe(II)-porphine complex with modest resources using a density matrix renormalization group solver in a CAS(22,27) active space and a 3000 AO basis.

Keywords: Second order CASSCF, AO-driven, DMRG-CASSCF, Fe(II)-porphine

1. Introduction

Multiconfigurational electronic structure is widely found across chemistry[1]. The complete active space self-consistent field (CASSCF) is a standard tool to describe multiconfigurational electronic structure problems[1, 2]. The CASSCF wavefunction further forms the starting point for more accurate treatments, including multireference perturbation theory and configuration interaction methods[3]. Because of its importance, much effort has been de-

---

*Corresponding author
    Email address: garnetc@caltech.edu (Garnet Kin-Lic Chan)

voted to efficient CASSCF algorithms in the last decades[4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

A well-known numerical challenge in CASSCF is to converge the self-consistent wavefunction. For this reason, many early investigations focused on second order optimization techniques, which demonstrate superior convergence to pure gradient or super-CI formulations[4, 5, 6, 8, 9, 7, 10, 11, 12]. Unfortunately, these early implementations were optimized for modest AO basis sets, because they transformed the integrals to the current set of CASSCF orbitals in each iteration, incurring significant computational cost and $\mathcal{O}(N^4)$ disk storage. To extend CASSCF algorithms to large AO bases, several strategies have been explored[13, 14, 15, 16]. For example, density-fitted CASSCF[13, 14] and Cholesky decomposition CASSCF[15] both approximate the AO integrals to achieve significant savings in the integral transformation cost and disk storage. GPU-based AO-driven CASSCF implementations[16, 18] further can handle very large numbers of AO functions, although these have not yet been extended to second order optimization. Although AO-driven algorithms typically require more floating point operations than MO-driven approaches, they are favourable for modern computers, due to their low IO and communication costs. In this work, our first motivation is to present a new AO-driven algorithm that can handle large AO basis sets without integral approximations, and also provide second order convergence. Our algorithm may easily be combined with density-fitting or Cholesky decomposition, although this is not a focus of this paper.

A second motivation is associated with the need to extend traditional CASSCF implementations to larger active spaces. In traditional CASSCF, full configuration interaction (FCI) is used as the active space solver. However, due to the exponential scaling of FCI, it is limited to small complete active spaces (CAS), usually no more than CAS(16,16) (16 electrons in 16 orbitals). However, there are now several techniques which can be used to replace the FCI solver[19, 20, 21, 22, 23, 24]. Two of the more commonly used ones are the density matrix renormalization group (DMRG)[19] and full configuration interaction quantum Monte Carlo (FCIQMC)[21, 22]. These can handle correlated active spaces with many tens of orbitals, and in some cases even more[20]. While implementations of DMRG and FCIQMC in the CASSCF algorithm exist [25, 26, 27, 28, 29, 30, 31, 32] they do not yet simultaneously provide second order convergence and the ability to treat very large numbers (i.e. 1000's) of AO's. The implementation we present can be straightforwardly interfaced to any external active space solver and thus fills

this gap. In the current work, we will use FCI and DMRG as the active space solvers. (An earlier FCIQMC-CASSCF calculation, reported in Ref. 30, used the two-step version of our implementation that we describe here).

In section 2, we describe the formulation of our CASSCF algorithm, including the approximate orbital gradient and Hessian updates, and orbital optimization method. In section 3 we carefully study the convergence properties and performance of our algorithm for several benchmark molecules, within our open-source program package PySCF[33]. Finally, as an example of a more challenging large scale problem, we use FCI and DMRG active space solvers and our CASSCF implementation to converge the Fe(II)-porphine singlet, triplet and quintet ground states. Our largest calculation uses a 22 electron, 27 orbital active space and almost 3000 AO basis functions.

2. Algorithm

2.1. Theory

In this section, we first summarize the relevant formulae for the optimization of the CASSCF wavefunction. Given the spin-free electronic Hamiltonian,

$$H = \sum_{ij} h_{ij} E^i_j + \frac{1}{2} \sum_{ijkl} (ij|kl)(E^i_j E^k_l - \delta_{jk} E^i_l) \tag{1}$$

$$E^i_j = a^\dagger_{i\alpha} a_{j\alpha} + a^\dagger_{i\beta} a_{j\beta} \tag{2}$$

the CASSCF energy can be written as a function of the CI coefficients **c** and the unitary orbital transformation matrix **U**,

$$E = H_{ijkl} \Gamma_{ijkl} \tag{3}$$

$$H_{ijkl} = V_{ijkl} U_{pi} U_{qj} U_{rk} U_{sl} \tag{4}$$

$$V_{pqrs} = \frac{1}{2(N_e - 1)} h_{pq} \delta_{rs} + \frac{1}{2(N_e - 1)} h_{rs} \delta_{pq} + \frac{1}{2}(pq|rs) \tag{5}$$

$$\Gamma_{ijkl} = \langle I|(E^i_j E^k_l - \delta_{jk} E^i_l)|J\rangle c_I c_J \tag{6}$$

where the Einstein summation convention is implied. Defining a Lagrangian with normalization constraints for **c** and **U**,

$$F(\mathbf{R}, \mathbf{c}) = E(\mathbf{R}, \Gamma) - \mathcal{E}(\mathbf{c}^\dagger \mathbf{c} - 1) \tag{7}$$

$$\mathbf{U} = \exp(\mathbf{R}) \tag{8}$$

$$\mathbf{R} = -\mathbf{R}^\dagger \tag{9}$$

minimizing the energy is a non-linear optimization problem for $\mathbf{R}^*, \mathbf{c}^*$, where the stationary conditions are

$$\left.\frac{\partial F}{\partial c_I}\right|_{\mathbf{R}^*,\mathbf{c}^*} = 0 \tag{10}$$

$$\left.\frac{\partial F}{\partial R_{pq}}\right|_{\mathbf{R}^*,\mathbf{c}^*} = 0 \tag{11}$$

The starting point for any second order non-linear optimization algorithm is Newton's method. Because the energy is quadratic in the CI coefficients, the Newton step for the CI coefficients, holding the orbitals fixed, is equivalent to solving the standard CI eigenvalue problem

$$\langle I|(H - \mathcal{E})|J\rangle c_J = 0 \tag{12}$$

Similarly, a Newton step for the orbitals, holding the CI coefficients fixed, corresponds to solving the equations

$$\mathcal{H}^{oo}\mathbf{R}^1 + \mathcal{G}^o = 0 \tag{13}$$

$$\mathcal{G}^o_{pq} = \frac{\partial F}{\partial R_{pq}} = \frac{\partial H_{ijkl}}{\partial R_{pq}}\Gamma_{ijkl} \tag{14}$$

$$\mathcal{H}^{oo}_{pq,rs} = \frac{\partial^2 F}{\partial R_{pq}\partial R_{rs}} = \frac{\partial^2 H_{ijkl}}{\partial R_{pq}\partial R_{rs}}\Gamma_{ijkl} \tag{15}$$

The simplest approach to CASSCF optimization is to alternately carry out the Newton steps (12), (13) for the CI coefficients and for the orbitals. This simple alternating scheme is known as the two-step optimization method. Unfortunately, even when the Newton steps are carried out exactly, for example, by using the exact orbital Hessian in Eq. (13), the two-step method suffers from slow convergence due to the neglect of coupling between the CI and orbital optimization problems. It is thus not usually considered a true second order convergent algorithm.

The more sophisticated, one-step, optimization methods aim to approximate the joint CI and orbital Newton step, corresponding to solving

$$\begin{pmatrix} \mathcal{H}^{cc} & \mathcal{H}^{co} \\ \mathcal{H}^{oc} & \mathcal{H}^{oo} \end{pmatrix} \begin{pmatrix} \mathbf{c}^1 \\ \mathbf{R}^1 \end{pmatrix} + \begin{pmatrix} \mathcal{G}^c \\ \mathcal{G}^o \end{pmatrix} = 0 \tag{16}$$

where the Hessian matrices are

$$\mathcal{H}_{IJ}^{cc} = \frac{\partial^2 F}{\partial c_I \partial c_J} = \langle I|(H - \mathcal{E})|J \rangle \tag{17}$$

$$\mathcal{H}_{I,pq}^{co} = \mathcal{H}_{pq,I}^{oc} = \frac{\partial^2 F}{\partial c_I \partial R_{pq}} = \frac{\partial H_{ijkl}}{\partial R_{pq}} \frac{\partial \Gamma_{ijkl}}{\partial c_I} \tag{18}$$

Here, the first row of the coupled equations (16)

$$\mathcal{H}^{cc}\mathbf{c}^1 + \mathcal{H}^{co}\mathbf{R}^1 + \mathcal{G}^c = 0 \tag{19}$$

can be rewritten as a CI response problem

$$\mathbf{H}^0\mathbf{c}^1 + \mathbf{H}^R\mathbf{c}^0 = E^0\mathbf{c}^1 \tag{20}$$

since

$$(\mathcal{H}^{co}\mathbf{R}^1)_I = H_{ijkl}^R \langle I|(E_j^i E_l^k - \delta_{jk}E_l^i)|J \rangle c_J^0 \tag{21}$$

$$\mathcal{G}_I^c = \langle I|(\mathbf{H}^0 - E^0)|J \rangle c_J^0 = 0 \tag{22}$$

where the first order Hamiltonian $\mathbf{H}^R$ is obtained from the chain rule

$$H_{ijkl}^R = \frac{\partial H_{ijkl}}{\partial R_{pq}} R_{pq}^1 = V_{pjkl}R_{pi}^1 + V_{ipkl}R_{pj}^1 + V_{ijpl}R_{pk}^1 + V_{ijkp}R_{pl}^1 \tag{23}$$

The second row of Eq. (16)

$$\mathcal{H}^{oo}\mathbf{R}^1 + \mathcal{H}^{oc}\mathbf{c}^1 + \mathcal{G}^o = 0 \tag{24}$$

can be interpreted as the orbital Newton problem with dressed gradients

$$\mathcal{H}^{oo}\mathbf{R}^1 = -\tilde{\mathcal{G}}^o \tag{25}$$

$$\tilde{\mathcal{G}}_{pq}^o = \mathcal{G}_{pq}^o + \mathcal{H}^{oc}\mathbf{c}^1 = \mathcal{G}_{pq}^o + \frac{\partial H_{ijkl}}{\partial R_{pq}}\Gamma_{ijkl}^1 \tag{26}$$

$$\Gamma_{ijkl}^1 = \frac{\partial \Gamma_{ijkl}}{\partial c_I}c_I^1 \tag{27}$$

The CI coefficient and orbital optimization problems are thus coupled through the first order $\mathbf{H}^R$ in Eq. (20) and the first order 2-particle density matrix Γ^1 in Eq. (25).

In principle, in the one-step method, the true CI Newton step requires solving the response equation (20) exactly. This is how some early versions of one-step optimization in CASSCF were implemented. However, if an iterative procedure is used to determine the CI eigenstate in Eq. (12), then a single (or few) steps of the same iterative procedure, with the modified Hamiltonian $\mathbf{H}^0 + \mathbf{H}^R$ and initial eigenstate guess of $\mathbf{c}^0$, can be used to determine an approximate $\mathbf{c}^1$. For example, a single Davidson iteration[34] with these quantities yields

$$\mathbf{c}^1 \approx -[\text{diag}(\mathbf{H}^0 - E^0)]^{-1}\mathbf{H}^R\mathbf{c}^0 \tag{28}$$

as an approximate solution of Eq. (20). The well-known MCSCF implementation by Werner and Knowles[9], uses this type of approximation. In our implementation, we also use a few iterations of the active space solver to determine an approximate update $\mathbf{c}^1$. The first order 2-particle density matrix is then computed by finite difference

$$\Gamma^1_{ijkl} \approx \Gamma_{ijkl}[\mathbf{c}^0 + \mathbf{c}^1] - \Gamma_{ijkl}[\mathbf{c}^0] \tag{29}$$

Importantly, this mechanism decouples the orbital optimization from the active space solver implementation in each Newton step, with the two communicating solely by passing the 2-particle density matrix and active space Hamiltonian. This allows us to easily plug-in different iterative active space solvers, so long as they can provide the 2-particle density matrix.

A single CI and orbital Newton step provides $\mathbf{c}^1$ and $\mathbf{R}^1$. We then need to update all quantities that depend on the new CI coefficients and new orbitals. This involves transforming H to the new set of orbitals in Eq. (4), solving for the new CI eigenstate in Eq. (12) and computing the new orbital gradients and Hessians from Eqs. (14) and (15). However, we only perform an exact update of all these quantities every few Newton steps, as it is computationally very expensive. Instead for most steps, we use only an approximate update of the quantities. We call the steps with approximate updates, microiterations. Every 3 - 4 microiterations, a macroiteration is carried out where all quantities are updated exactly. We next describe the different kinds of approximate updates used in the microiterations.

2.2. Microiterations

In each microiteration, we perform approximate updates both for the CI and orbital parts of the optimization problem. The quality of the update approximation is important, as it can affect the rate of CASSCF convergence.

We first discuss the orbital update.We have considered two frameworks for approximation. The first is the dynamic-expansion-point (DEP) scheme (as used for example, in Ref. 9) where we compute the new Hamiltonian matrix elements $\bar{H}_{ijkl}$

$$\bar{H}_{ijkl} = V_{pqrs}\bar{U}_{pi}\bar{U}_{qj}\bar{U}_{rk}\bar{U}_{sl} \tag{30}$$

$$\bar{U} = \exp(\mathbf{R}^1) \tag{31}$$

and define the updated orbital gradient and orbital Hessian from

$$\mathcal{G}^o_{pq} = \left.\frac{\partial F}{\partial R_{pq}}\right|_{\mathbf{R}^1} = \frac{\partial \bar{H}_{ijkl}}{\partial R_{pq}}\Gamma_{ijkl} \tag{32}$$

$$\mathcal{H}^{oo}_{pq} = \left.\frac{\partial^2 F}{\partial R_{pq}\partial R_{rs}}\right|_{\mathbf{R}^1} = \frac{\partial^2 \bar{H}_{ijkl}}{\partial R_{pq}\partial R_{rs}}\Gamma_{ijkl} \tag{33}$$

$$\tag{34}$$

The approximate updates in the DEP framework consist of approximating $\bar{H}$ to reduce the costs of the 4-index integral transformation (30). By dividing $\bar{U}$ into two parts

$$\bar{U} = 1 + \mathbf{T} \tag{35}$$

an approximate $\bar{H}$ can be defined up a given order in $\mathbf{T}$. For example, the first order approximate update (DEP1) corresponds to

$$\bar{H}_{ijkl} = V_{ijkl} + V_{pjkl}T_{pi} + V_{ipkl}T_{pj} + V_{ijpl}T_{pk} + V_{ijkp}T_{pl} \tag{36}$$

The exact update is recovered at fourth order, where the complete 4-index integral transformation is carried out. The transformation in DEP4 has the same cost as the integral transformation in the two-step CASSCF optimization method.

The other framework in which to define approximate updates is the fixed expansion point (FEP) scheme. In FEP, the orbital gradients and Hessians are defined directly by an expansion in $\mathbf{R}^1$

$$\mathcal{G}^o \rightarrow \mathcal{G}^o_{pq} + \frac{\partial^2 F}{\partial R_{pq}\partial R_{rs}} \cdot R^1_{rs} + \dots \tag{37}$$

$$\mathcal{H}^{oo} \rightarrow \mathcal{H}^{oo}_{pq,rs} + \frac{\partial^3 F}{\partial R_{pq}\partial R_{rs}\partial R_{tu}} \cdot R^1_{tu} + \dots \tag{38}$$

The approximate updates in the FEP framework correspond to truncating the order of the above expansion. For example, in the simplest FEP1 approximation, we only update the gradients using the non-updated Hessian matrix elements

$$\mathcal{G}^o + \mathcal{H}^{oo}\mathbf{R}^1 \tag{39}$$

which corresponds to keeping the orbital Hessian frozen within each macro iteration. We can view the gradients in FEP1 to be an approximation to the gradients in DEP1, obtained by replacing $\mathbf{T}$ with $\mathbf{R}$.

To update the CI part, we need to update the Hamiltonian that defines the first order CI problem (23). Analogous approximations within the DEP/FEP framework can be formulated for the Hamiltonian update. For example, Eq. (36) is also the definition of the updated CI Hamiltonian within the DEP1 approximation.

2.3. Orbital optimization

So far we have been discussing the determination of the Newton step. However, directly following a Newton step is problematic in highly non-quadratic optimizations, as the steps can be unbounded, and in fact are not even guaranteed to go towards the minimum. This is a well-known problem which is often seen in orbital optimization in CASSCF.

A commonly employed technique to modify the Newton step for the orbitals, is to use the augmented Hessian (AH) method with step-size control. Here, a modified Newton step is obtained by solving the eigenvector equation

$$\begin{pmatrix} 0 & \mathcal{G}^\dagger \\ \mathcal{G} & \mathcal{H} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{R}^1 \end{pmatrix} = \varepsilon \begin{pmatrix} 1 \\ \mathbf{R}^1 \end{pmatrix} \tag{40}$$

In the standard use of this method, the rotation direction $\mathbf{R}^1$ is first obtained as the lowest eigenvector of Eq. (40).This direction provides an interpolation between steepest descent and the full Newton step. Once $\mathbf{R}^1$ is determined, then an approximate line-search is performed along $\mathbf{R}^1$ to take an appropriate step size.

In our current algorithm, we also use the augmented Hessian method, and we solve Eq. (40) using the Davidson method[34]. However, rather than first determining the orbital search direction $\mathbf{R}^1$ from solving Eq. (40), and then carrying out a separate line-search, we combine these two procedures, in a co-iteration. Each co-iteration corresponds to an AH Davidson iteration,

followed by an update of the orbital gradient in the FEP1 approximation. For example after i Davidson iterations, the new gradient is updated as

$$\mathcal{G}^o_{[i+1]} = \mathcal{G}^o_{[i]} + \mathcal{H}^{oo}\mathbf{R}^1_{[i]} \tag{41}$$

where $\mathbf{R}^1_{[i]}$ is the approximate Davidson AH solution at the ith iteration. We ensure that each $\mathbf{R}^1_{[i]}$ is sufficiently small, by introducing a scale parameter $\lambda \geq 1$ in the AH equations, where λ is chosen such that the largest element in the scaled $\mathbf{R}^1_{[i]}$ is smaller than a predefined threshold (0.03 in our current implementation)

$$\begin{pmatrix} 0 & \mathcal{G}^\dagger_{[i]} \\ \mathcal{G}_{[i]} & \mathcal{H} \end{pmatrix} \begin{pmatrix} 1 \\ \lambda\mathbf{R}^1_{[i]} \end{pmatrix} = \varepsilon \begin{pmatrix} 1 \\ \lambda\mathbf{R}^1_{[i]} \end{pmatrix} \tag{42}$$

Accumulating the small steps $\mathbf{R}^1_{[i]}$ from the co-iterations, we obtain the full orbital rotation,

$$\mathbf{R}^1 = \mathbf{R}^1_{[0]} + \mathbf{R}^1_{[1]} + \mathbf{R}^1_{[2]} + \ldots \tag{43}$$

Because the gradients are updated, the quantities entering into the AH equations change at each co-iteration. Table 1 shows how $\mathcal{H}$, $\mathcal{G}$ and $\mathbf{R}^1_{[i]}$ as a function of the iteration number.

We choose λ such that every step $\mathbf{R}^1_{[i]}$ is small, but in our numerical tests, this small stepsize does not lead to low efficiency. Instead, the accumulated small steps effectively provide the ability to take large total rotations in the orbital space. We thus achieve a good compromise between larger steps for greater efficiency, and small steps for greater robustness.

2.4. Computational costs

The two main computational costs to consider are the memory usage and the operation count. Memory usage is quite different in MO-driven and AO-driven CASSCF optimization algorithms. In the traditional MO-driven algorithm, all integrals are transformed to the MO representation in each macroiteration, and then they are reused in the microiterations. Although the MO-driven approach requires less CPU resources, it requires more I/O than the AO-driven algorithm when there are a large number of core orbitals. For example, to evaluate the contraction $\mathcal{H}\mathbf{R}^1$ in the DEP1 approximation, the MO-driven algorithm requires the two-electron integrals $(AA|**)$, $(A*|A*)$, $(CC|VV)$, $(CV|CV)$, $(CV|AC)$, $(CV|AV)$, $(AV|CC)$, $(AC|VV)$ (see the matrix elements in the supplemental material), where the letters C, A, V stand for core, active and external orbitals, and the symbol "$*$"

stands for any of these kinds of orbitals. However, the integrals $(CC|VV)$, $(CV|CV)$, $(CV|AC)$, $(CV|AV)$, $(AV|CC)$, $(AC|VV)$ are associated with the contractions to the core or the active space 1-particle density matrices. These can instead be evaluated in a direct-SCF style AO-driven J and K (Coulomb and exchange matrix) build. As such, the AO-driven algorithm only requires two kinds of integrals $(AA|**)$, $(A*|A*)$ to be computed in each macroiteration. For the various orders of the DEP/FEP approximations, Table 2 summarizes the memory requirements in the MO-driven and AO-driven implementations. Due to the lower memory requirements, the AO-driven DEP1 approximation is favoured in our general CASSCF implementation.

The CPU costs of the CASSCF algorithm are more difficult to optimize than the memory costs. In the one-step algorithm, the demanding CPU steps are the expensive macro iteration (which invoke the 4-index integral transformation and the accurate solution for the CI eigenstate), the CI response problem in the micro iteration, and the contraction operation $\mathcal{H}^{oo}\mathbf{R}^1$ in the orbital optimization problem. The CI response problem in the micro iteration can be a dominant cost when an expensive active space solver, such as DMRG or FCIQMC is used. However, as we have mentioned above, we can solve the CI response equation approximately, by solving the CI eigenstate problem approximately. Using a FCI active space solver, 2 - 5 Davidson iterations are usually enough to generate a sufficiently accurate first order density matrix to generate the dressed orbital gradients (26), while using a DMRG active space solver, 2 - 3 sweeps are usually sufficient.

The $\mathcal{H}^{oo}\mathbf{R}^1$ contraction in the AO-driven algorithm is expensive because solving the AH equations in the orbital optimization requires many $\mathcal{H}^{oo}\mathbf{R}^1$ contractions, and every $\mathcal{H}^{oo}\mathbf{R}^1$ contraction requires a build of the entire J and K matrices. One workaround is to gradually increase the accuracy of the solution of the AH problem in the optimization. The AH equation is first only approximately solved in a small Davidson subspace, until the CASSCF wavefunction is close to convergence.

In our implementation, we carry out 4 microiterations for each macroiteration and 4 - 6 Davidson (co)-iterations per microiteration. Thus, the ratio of the number of macroiterations, microiterations and J and K builds, is roughly 1:4:20. It should be noted that our parameters have not been carefully optimized to minimize the total number of macroiterations. Instead, we have used conservative settings to provide robust convergence. Optimal settings will be investigated in future studies.

3. Numerical studies

We next assess the performance of our CASSCF algorithm in a variety of small molecules and challenging systems. We first tested the convergence properties as a function of the approximate orbital gradient and Hessian DEP/FEP updates. The computational efficiency was measured by the number of macroiterations, microiterations, and J and K builds. Next, to test the capabilities of the algorithm in a large problem, we optimized the singlet, triplet, and quintet states of the Fe(II)-porphine molecule in a (8e, 11o) active space using the cc-pVDZ (439 AO functions), cc-pVTZ (956 AO functions), cc-pVQZ (1784 AO functions), and cc-pV5Z bases (2997 AO functions). Finally, we demonstrate the flexibility of our implementation by carrying out a larger active space study (22e, 27o) using a DMRG active space solver.

The general CASSCF optimization algorithm and FCI solver were implemented in the open-source PySCF program package[33]. The DMRG solver was provided by the BLOCK code[35]. All tests were executed on a workstation equipped with 2 Xeon E5-2670 CPUs (16 CPU cores @ 2.5 GHz) and 64 GB memory.

3.1. DEP and FEP approximations

We first summarize the different DEP/FEP approximate updates that we tested in Table 3. At the lowest level of FEP1, we only consider the contribution of the first order $\mathbf{R}^1$ to the orbital gradients and CI Hamiltonian. In the DEP1 and DEP2 approximations, the orbital gradients and CI Hamiltonian are evaluated up to first and second order in $\mathbf{T}$, respectively. The contributions of $\mathbf{T}$ to the orbital Hessian are not included in the FEP1, DEP1, and DEP2 approximations. The treatment of the orbital gradients and CI Hamiltonians in DEP1+, DEP2+ is the same as in the DEP1, DEP2 approximations, while higher order terms in the orbital Hessian are included. In DEP4+, all quantities are exactly updated, because 4th order is the highest expansion order for the Hamiltonian matrix elements.

Table 4 presents the convergence performance of the 1-step optimization method, for different DEP and FEP approximations, in a variety of simple systems. Regardless of the approximation, the 1-step method is always superior to the 2-step method. The number of macroiterations required in the 1-step method is only 1/3 to 1/4 of that required by the 2-step method. Except for the two systems CH_2 and O_3, the different approximations present similar comparative convergence behavior across the different molecules. For

11

CH$_2$ and O$_3$, the lower order approximations FEP1 and DEP1 behave slightly worse than DEP2 and DEP4.

Figure 1 visualizes the convergence of the total energy against the macroiteration number for CH$_2$ CAS(6,14)/cc-pVDZ. The six different approximate updates used in the microiterations all show similar convergence properties, and they all require 9 macroiterations to converge the total energy. One microiteration of the 1-step method shows similar convergence gains to one macroiteration of the 2-step method, despite being much cheaper computationally. Since we used 4 microiterations for every macroiteration in the 1-step method, each macroiteration of the 1-step method effectively performs like 4 macroiterations in the 2-step method. As shown by the horizontal dotted lines in the figure, the DEP1 approximation needed 2 macroiterations in the 1-step method to converge the energy error from 10^{-3} to 10^{-6} Hartrees, while the 2-step method using the same DEP1 approximation, took 9 macroiterations to achieve the same.

Figure 2 shows the convergence behaviour for the singlet ground state of the Al$_4$O$_2$ cluster (see supplemental material for the geometry) using a CAS(12,12) active space and the aug-cc-pVDZ basis. The active space is chosen using a DMET-like procedure[36, 37, 38, 39]. First, the AO's were localized with a meta-Löwdin orthogonalization[37], then the 6 oxygen $2p$ orbitals were selected as the "impurities". The resulting 6 impurity and 6 bath orbitals obtained as the DMET space were chosen to be the initial active space. Except when using the FEP1 approximate update, the 1-step optimization converges in 7 macroiterations, about 1/3 of the 22 macroiterations required by the 2-step method. Using the DEP1 approximation, converging the energy error from 10^{-3} to 10^{-6} Hartrees required 3 macroiterations of the 1-step method, and 11 macroiterations of the 2-step method.

Figure 3 show the convergence behaviour of the optimization algorithm for the ground-state ($^1A_1$) of the HN$_3$ molecule (see supplemental material for the geometry) using a CAS(10,10) active space and the cc-pVTZ basis. We use 10 Hartree-Fock orbitals (5 HOMO's, 5 LUMO's) to form the initial active space. We observed a large change in the character of the active space and slow initial convergence for both the 1-step and 2-step methods. Depending on the approximate update used in the microiteration, the 1-step method needed 7 - 9 macroiterations to converge.

Overall, our tests demonstrate that the low level approximate update (DEP1) performs almost as well as the exact orbital update (DEP4+) with our current CASSCF algorithm settings. However, it should be noted that

12

the total number of macroiterations has not been optimized for the case of the high-level DEP approximations. For example, it is always possible to converge DEP4+ in one macro iteration, with a very large number of microiterations, because all quantities are updated exactly in the microiterations. However, we cannot use too many microiterations with DEP1 because the errors in the approximate updates will quickly accumulate. To obtain the best runtime performance, a trade-off has to be made between the DEP level (for memory/disk usage) and the number of macroiterations (for CPU time). In the next section, we use the DEP1 approximation because of its low memory requirements, which allows us to treat a large number of orbitals.

3.2. CASSCF and DMRG-CASSCF study of Fe(II)-porphine

The Fe(II)-porphine ground state[40, 41, 42, 43, 44, 45, 46, 47, 48, 49] has long been a target of multireference quantum chemistry studies. There has been much debate about the ordering of the lowest spin states. Density functional approximations tend to predict a triplet ground state, while many CASSCF and multireference perturbation calculation with small active spaces argue for a quintet ground-state[42, 44, 45, 50].

We carried out CASSCF calculations at the triplet geometry[45] with D_{2h} symmetry in the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV5Z bases. The active space was initialized with the DMET procedure and consisted of Fe $3d$ and $4d$ meta-Löwdin orthogonalized orbitals and one optimal bath orbital[37], giving 11 orbitals in total. The DMET bath orbital was chosen as the most important bath orbital from the spin-free ROHF density matrix, which is strongly entangled with the Fe $3d$ and $4d$ orbitals. Based on the Hartree-Fock density matrix, we assigned 8 electrons to the active space. The population density was 6.14 electrons on the impurity (Fe 3d, 4d) orbitals, and 1.88 on the bath orbital. Table 5 gives the energies of some of the low lying singlet, triplet, and quintet states. With our initial guess, good convergence was found for all states except for the $^3B_{1g}$ state. For the $^3B_{1g}$ state, within 4 - 5 iterations, the optimization appeared to approach a very flat region, with small gradients and energy changes of about 10^{-6} Hartrees between macroiterations. However, after 5 - 10 more macroiterations, the optimization left this region, and then rapidly converged to a solution that was about 1 mE_{h} lower than the flat region. The converged solution has significant $4s$ components, which are not part of the initial guess.

Finally, as an example of a CASSCF calculation with a larger active space, we extended our active space to 27 orbitals, containing the Fe $3d, 4d, 4s$ shells

13

(11 orbitals), 4 N $2p_z$ orbitals, N $2p_x$ and $2p_y$ orbitals (for the Fe-N σ bonds) and the most important 8 DMET bath orbitals generated using the above impurity orbitals. The 8 DMET bath orbitals constituted 4 ligand π orbitals and 4 Fe-N anti-bonding ligand orbitals. To tackle this large active space, we used a DMRG active space solver, with a bond dimension of $M = 1000$. For the AO basis, we used the cc-pVDZ, cc-pVQZ, and cc-pV5Z bases. The energies of the triplet and quintet ground-states are given in Table 5. In this larger active space, we find that ^5A_g is the lowest state and is about 6 mE_h lower than $^3B_{3g}$ state, irrespective of basis.

In Table 5, we show the efficiency of the CASSCF optimization in these Fe(II)-porphine calculations, as measured by the number of J and K matrix evaluations, microiterations, and macroiterations. Depending on the type of calculation, the time-dominant step differs. For small active spaces, the CPU time for the CI problem is negligible. The J and K matrix evaluation is also more costly than the integral transformation. For example, it takes 49 hours to converge the ^5A_g state for CAS(8,11)/cc-pVQZ on a 16-core node, in which about 63.7 % of the time is used to evaluate J and K matrices, and 35 % of the time is spent on the macroiterations (in integral transformation). When the active space is larger and handled by a more expensive active space solver, the cost of the macroiterations and microiterations both increase. In the DMRG-CAS(22,27) calculation, 41 % of the time was spent on macroiterations (with about 1/5 of the time in the DMRG solver and 4/5 of the time in integral transformation). In the microiterations, over 8 % of the time was spent in the DMRG solver.

4. Conclusion

In this work, we presented a general second order CASSCF implementation for large scale calculations. We used an AO-driven approach to handle large AO bases, and formulated our Newton steps to decouple the CI solver from the orbital optimization in each microiteration, thus allowing ready interfacing to modern active space solvers, such as the DMRG and FCIQMC. Further, to achieve greater efficiency, we proposed several approximate updates of the orbital and CI gradient and Hessians, as well as a co-iterative augmented Hessian algorithm to determine the orbital step. We assessed the numerical performance of the general CASSCF solver with different approximate updates, and with 1-step and 2-step optimization, in a variety of small molecules, and in a larger case-study of the Fe(II)-porphine low-lying states.

14

Using our algorithm, we showed that we could converge a DMRG-CASSCF calculation using a (22,27) active space and 3000 AO basis functions with only modest resources.

5. Acknowledgments

[1] M. W. Schmidt, M. S. Gordon, The construction and interpretation of mcscf wavefunctions, Annu. Rev. Phys. Chem. 49 (1) (1998) 233–266. arXiv:http://dx.doi.org/10.1146/annurev.physchem.49.1.233, doi:10.1146/annurev.physchem.49.1.233.
URL http://dx.doi.org/10.1146/annurev.physchem.49.1.233

[2] B. O. Roos, The complete active space self-consistent field method and its applications in electronic structure calculations, Adv. Chem. Phys. 69 (1987) 399–445. doi:10.1002/9780470142943.ch7.
URL http://dx.doi.org/10.1002/9780470142943.ch7

[3] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, R. Shepard, Multiconfiguration self-consistent field and multireference configuration interaction methods and applications, Chem. Rev. 112 (1) (2012) 108–181. arXiv:http://pubs.acs.org/doi/pdf/10.1021/cr200137a, doi:10.1021/cr200137a.
URL http://pubs.acs.org/doi/abs/10.1021/cr200137a

[4] B. H. Lengsfield, General second order mcscf theory: A density matrix directed algorithm, J. Chem. Phys. 73 (1) (1980) 382–390. doi:10.1063/1.439885.
URL http://link.aip.org/link/?JCP/73/382/1

[5] H.-J. Werner, W. Meyer, A quadratically convergent multiconfiguration-self-consistent field method with simultaneous optimization of orbitals and ci coefficients, J. Chem. Phys. 73 (5) (1980) 2342–2356. doi:http://dx.doi.org/10.1063/1.440384.
URL http://scitation.aip.org/content/aip/journal/jcp/73/5/10.1063/1.440384

[6] B. O. Roos, P. R. Taylor, P. E. Siegbahn, A complete active space {SCF} method (casscf) using a density matrix formulated super-ci approach, Chem. Phys. 48 (2) (1980) 157 – 173. doi:http://dx.doi.org/10.1016/0301-0104(80)80045-0.
URL http://www.sciencedirect.com/science/article/pii/0301010480800450

[7] J. Olsen, D. L. Yeager, P. Jørgensen, Optimization and characterization of a multiconfigurational self-consistent field (mcscf) state, Adv. Chem. Phys. 54 (1983) 1–176. doi:10.1002/9780470142783.ch1.
URL `http://dx.doi.org/10.1002/9780470142783.ch1`

[8] H.-J. A. Jensen, P. Jørgensen, A direct approach to second-order mcscf calculations using a norm extended optimization scheme, J. Chem. Phys. 80 (3) (1984) 1204–1214. doi:http://dx.doi.org/10.1063/1.446797.
URL `http://scitation.aip.org/content/aip/journal/jcp/80/3/10.1063/1.446797`

[9] H.-J. Werner, P. J. Knowles, A second order multiconfiguration scf procedure with optimum convergence, J. Chem. Phys. 82 (11) (1985) 5053–5063. doi:http://dx.doi.org/10.1063/1.448627.
URL `http://scitation.aip.org/content/aip/journal/jcp/82/11/10.1063/1.448627`

[10] H.-J. Werner, Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration reference configuration-interaction methods, Adv. Chem. Phys. 69 (1987) 1–62. doi:10.1002/9780470142943.ch1.
URL `http://dx.doi.org/10.1002/9780470142943.ch1`

[11] R. Shepard, The multiconfiguration self-consistent field method, Adv. Chem. Phys. 69 (2007) 63–200. doi:10.1002/9780470142943.ch2.
URL `http://dx.doi.org/10.1002/9780470142943.ch2`

[12] H. Nakano, K. Hirao, A quasi-complete active space self-consistent field method, Chem. Phys. Lett. 317 (12) (2000) 90 – 96. doi:http://dx.doi.org/10.1016/S0009-2614(99)01364-0.
URL `http://www.sciencedirect.com/science/article/pii/S0009261499013640`

[13] S. Ten-no, S. Iwata, Multiconfiguration self-consistent field procedure employing linear combination of atomic-electron distributions, J. Chem. Phys. 105 (9) (1996) 3604–3611. doi:http://dx.doi.org/10.1063/1.472231.
URL `http://scitation.aip.org/content/aip/journal/jcp/105/9/10.1063/1.472231`

[14] W. Győrffy, T. Shiozaki, G. Knizia, H.-J. Werner, Analytical energy gradients for second-order multireference perturbation theory using density fitting, J. Chem. Phys. 138 (10) (2013) 104104.

doi:http://dx.doi.org/10.1063/1.4793737.
URL `http://scitation.aip.org/content/aip/journal/jcp/138/10/10.1063/1.4793737`

[15] F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. Sánchez de Merás, H. Koch, Accurate ab initio density fitting for multiconfigurational self-consistent field methods, J. Chem. Phys. 129 (2) (2008) 024113. doi:http://dx.doi.org/10.1063/1.2953696.
URL `http://scitation.aip.org/content/aip/journal/jcp/129/2/10.1063/1.2953696`

[16] E. G. Hohenstein, N. Luehr, I. S. Ufimtsev, T. J. Martínez, An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units, J. Chem. Phys. 142 (22) (2015) 224103. doi:http://dx.doi.org/10.1063/1.4921956.
URL `http://scitation.aip.org/content/aip/journal/jcp/142/22/10.1063/1.4921956`

[17] I. Kim, S. M. Parker, T. Shiozaki, Orbital optimization in the active space decomposition model, J. Chem. Theory Comput. 11 (8) (2015) 3636–3642. arXiv:http://dx.doi.org/10.1021/acs.jctc.5b00429, doi:10.1021/acs.jctc.5b00429.
URL `http://dx.doi.org/10.1021/acs.jctc.5b00429`

[18] B. S. Fales, B. G. Levine, Nanoscale multireference quantum chemistry: Full configuration interaction on graphical processing units, J. Chem. Theory Comput. 11 (10) (2015) 4708–4716. arXiv:http://dx.doi.org/10.1021/acs.jctc.5b00634, doi:10.1021/acs.jctc.5b00634.
URL `http://dx.doi.org/10.1021/acs.jctc.5b00634`

[19] G. K.-L. Chan, S. Sharma, The density matrix renormalization group in quantum chemistry, Annu. Rev. Phys. Chem. 62 (1) (2011) 465–481. arXiv:http://dx.doi.org/10.1146/annurev-physchem-032210-103338, doi:10.1146/annurev-physchem-032210-103338.
URL `http://dx.doi.org/10.1146/annurev-physchem-032210-103338`

[20] S. Sharma, T. Yanai, G. H. Booth, C. J. Umrigar, G. K.-L. Chan, Spectroscopic accuracy directly from quantum chemistry: Application to ground and excited states of beryllium dimer, J. Chem. Phys. 140 (10) (2014) 104112. doi:http://dx.doi.org/10.1063/1.4867383.
URL `http://scitation.aip.org/content/aip/journal/jcp/140/10/10.1063/1.4867383`

[21] G. H. Booth, A. J. W. Thom, A. Alavi, Fermion monte carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space, J. Chem. Phys. 131 (5) (2009) 054106. doi:http://dx.doi.org/10.1063/1.3193710.
URL http://scitation.aip.org/content/aip/journal/jcp/131/5/10.1063/1.3193710

[22] G. H. Booth, D. Cleland, A. Alavi, D. P. Tew, An explicitly correlated approach to basis set incompleteness in full configuration interaction quantum monte carlo, J. Chem. Phys. 137 (16) (2012) 164112. doi:http://dx.doi.org/10.1063/1.4762445.
URL http://scitation.aip.org/content/aip/journal/jcp/137/16/10.1063/1.4762445

[23] K. D. Vogiatzis, G. Li Manni, S. J. Stoneburner, D. Ma, L. Gagliardi, Systematic expansion of active spaces beyond the casscf limit: A gasscf/splitgas benchmark study, J. Chem. Theory Comput. 11 (7) (2015) 3010–3021. arXiv:http://dx.doi.org/10.1021/acs.jctc.5b00191, doi:10.1021/acs.jctc.5b00191.
URL http://dx.doi.org/10.1021/acs.jctc.5b00191

[24] J. Fosso-Tande, T.-S. Nguyen, G. Gidofalvi, A. E. DePrince, Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods, J. Chem. Theory Comput. 12 (5) (2016) 2260–2271. arXiv:http://dx.doi.org/10.1021/acs.jctc.6b00190, doi:10.1021/acs.jctc.6b00190.
URL http://dx.doi.org/10.1021/acs.jctc.6b00190

[25] D. Ghosh, J. Hachmann, T. Yanai, G. K.-L. Chan, Orbital optimization in the density matrix renormalization group, with applications to polyenes and -carotene, J. Chem. Phys. 128 (14). doi:http://dx.doi.org/10.1063/1.2883976.
URL http://scitation.aip.org/content/aip/journal/jcp/128/14/10.1063/1.2883976

[26] D. Zgid, M. Nooijen, The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space, J. Chem. Phys. 128 (14). doi:http://dx.doi.org/10.1063/1.2883981.
URL http://scitation.aip.org/content/aip/journal/jcp/128/14/10.1063/1.2883981

[27] T. Yanai, Y. Kurashige, D. Ghosh, G. K.-L. Chan, Accelerating convergence in iterative solution for large-scale complete active space self-

consistent-field calculations, Int. J. Quant. Chem. 109 (10) (2009) 2178–2190. doi:10.1002/qua.22099.
URL `http://dx.doi.org/10.1002/qua.22099`

[28] Y. Ma, H. Ma, Assessment of various natural orbitals as the basis of large active space density-matrix renormalization group calculations, J. Chem. Phys. 138 (22). doi:http://dx.doi.org/10.1063/1.4809682.
URL `http://scitation.aip.org/content/aip/journal/jcp/138/22/10.1063/1.4809682`

[29] S. Wouters, T. Bogaerts, P. Van Der Voort, V. Van Speybroeck, D. Van Neck, Communication: Dmrg-scf study of the singlet, triplet, and quintet states of oxo-mn(salen), J. Chem. Phys. 140 (24). doi:http://dx.doi.org/10.1063/1.4885815.
URL `http://scitation.aip.org/content/aip/journal/jcp/140/24/10.1063/1.4885815`

[30] R. E. Thomas, Q. Sun, A. Alavi, G. H. Booth, Stochastic multiconfigurational self-consistent field theory, J. Chem. Theory Comput. 11 (11) (2015) 5316–5325. arXiv:http://dx.doi.org/10.1021/acs.jctc.5b00917, doi:10.1021/acs.jctc.5b00917.
URL `http://dx.doi.org/10.1021/acs.jctc.5b00917`

[31] G. Li Manni, S. D. Smart, A. Alavi, Combining the complete active space self-consistent field method and the full configuration interaction quantum monte carlo within a super-ci framework, with application to challenging metal-porphyrins, J. Chem. Theory Comput. 12 (3) (2016) 1245–1258. arXiv:http://dx.doi.org/10.1021/acs.jctc.5b01190, doi:10.1021/acs.jctc.5b01190.
URL `http://dx.doi.org/10.1021/acs.jctc.5b01190`

[32] Y. Ma, S. Knecht, S. Keller, M. Reiher, Second-order self-consistent-field density-matrix renormalization group, arXiv:1611.05972 [physics.chem-ph]arXiv:1611.05972.

[33] Q. Sun, Python module for quantum chemistry program, `https://github.com/sunqm/pyscf.git` (2014).

[34] E. R. Davidson, The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, J. Comput. Phys. 17 (1) (1975) 87 – 94. doi:10.1016/0021-9991(75)90065-0.
URL `http://www.sciencedirect.com/science/article/pii/0021999175900650`

[35] S. Sharma, G. K.-L. Chan, Spin-adapted density matrix renormalization group algorithms for quantum chemistry, J. Chem. Phys. 136 (12). doi:http://dx.doi.org/10.1063/1.3695642.
URL `http://scitation.aip.org/content/aip/journal/jcp/136/12/10.1063/1.3695642`

[36] G. Knizia, G. K.-L. Chan, Density matrix embedding: A simple alternative to dynamical mean-field theory, Phys. Rev. Lett. 109 (2012) 186404. doi:10.1103/PhysRevLett.109.186404.
URL `http://link.aps.org/doi/10.1103/PhysRevLett.109.186404`

[37] Q. Sun, G. K.-L. Chan, Exact and optimal quantum mechanics/molecular mechanics boundaries, J. Chem. Theory Comput. 10 (9) (2014) 3784–3790. arXiv:http://dx.doi.org/10.1021/ct500512f, doi:10.1021/ct500512f.
URL `http://dx.doi.org/10.1021/ct500512f`

[38] S. Wouters, C. A. Jiménez-Hoyos, Q. Sun, G. K.-L. Chan, A practical guide to density matrix embedding theory in quantum chemistry, J. Chem. Theory Comput. 12 (6) (2016) 2706–2719. arXiv:http://dx.doi.org/10.1021/acs.jctc.6b00316, doi:10.1021/acs.jctc.6b00316.
URL `http://dx.doi.org/10.1021/acs.jctc.6b00316`

[39] E. R. Sayfutyarova, G. Knizia, Q. Sun, G. K.-L. Chan, Automatic construction of molecular active spaces from atomic valence orbitalsIn preparation.

[40] W. D. Edwards, B. Weiner, M. C. Zerner, On the low-lying states and electronic spectroscopy of iron(ii) porphine, J. Am. Chem. Soc. 108 (9) (1986) 2196–2204. arXiv:http://dx.doi.org/10.1021/ja00269a012, doi:10.1021/ja00269a012.
URL `http://dx.doi.org/10.1021/ja00269a012`

[41] C. Rovira, K. Kunc, J. Hutter, P. Ballone, M. Parrinello, Equilibrium geometries and electronic structure of iron-porphyrin complexes: a density functional study, J. Phys. Chem. A 101 (47) (1997) 8914–8925. arXiv:http://dx.doi.org/10.1021/jp9722115, doi:10.1021/jp9722115.
URL `http://dx.doi.org/10.1021/jp9722115`

[42] Y.-K. Choe, T. Nakajima, K. Hirao, R. Lindh, Theoretical study of the electronic ground state of iron(ii) porphine. ii, J. Chem. Phys. 111 (9) (1999) 3837–3845. doi:http://dx.doi.org/10.1063/1.479687.
URL http://scitation.aip.org/content/aip/journal/jcp/111/9/10.1063/1.479687

[43] K. Oliveira, M. Trsic, Comparative theoretical study of the electronic structures and electronic spectra of fe2+-, fe+3-porphyrin and free base porphyrin, J. Mol. Struct. (THEOCHEM) 539 (13) (2001) 107 – 117. doi:http://dx.doi.org/10.1016/S0166-1280(00)00778-8.
URL http://www.sciencedirect.com/science/article/pii/S0166128000007788

[44] K. Pierloot, The caspt2 method in inorganic electronic spectroscopy: from ionic transition metal to covalent actinide complexes, Mol. Phys. 101 (13) (2003) 2083–2094. arXiv:http://www.tandfonline.com/doi/pdf/10.1080/0026897031000109356, doi:10.1080/0026897031000109356.
URL http://www.tandfonline.com/doi/abs/10.1080/0026897031000109356

[45] A. R. Groenhof, M. Swart, A. W. Ehlers, K. Lammertsma, Electronic ground states of iron porphyrin and of the first species in the catalytic reaction cycle of cytochrome p450s, J. Phys. Chem. A. 109 (15) (2005) 3411–3417. arXiv:http://dx.doi.org/10.1021/jp0441442, doi:10.1021/jp0441442.
URL http://dx.doi.org/10.1021/jp0441442

[46] M.-S. Liao, J. D. Watts, M.-J. Huang, Assessment of the performance of density-functional methods for calculations on iron porphyrins and related compounds, J. Comput. Chem. 27 (13) (2006) 1577–1592. doi:10.1002/jcc.20458.
URL http://dx.doi.org/10.1002/jcc.20458

[47] H. Chen, W. Lai, S. Shaik, Multireference and multiconfiguration ab initio methods in heme-related systems: What have we learned so far?, J. Chem. Phys. 115 (8) (2011) 1727–1742. arXiv:http://dx.doi.org/10.1021/jp110016u, doi:10.1021/jp110016u.
URL http://dx.doi.org/10.1021/jp110016u

[48] M. Swart, Spin states of (bio)inorganic systems: Successes and pitfalls, Int. J. Quant. Chem. 113 (1) (2013) 2–7. doi:10.1002/qua.24255.
URL http://dx.doi.org/10.1002/qua.24255

[49] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, G. K.-L. Chan, The ab-initio density matrix renormalization group in practice, J. Chem. Phys. 142 (3) (2015) 034102. doi:http://dx.doi.org/10.1063/1.4905329.
URL http://scitation.aip.org/content/aip/journal/jcp/142/3/10.1063/1.4905329

[50] S. Vancoillie, H. Zhao, V. T. Tran, M. F. A. Hendrickx, K. Pierloot, Multiconfigurational second-order perturbation theory restricted active space (raspt2) studies on mononuclear first-row transition-metal systems, J. Chem. Theory Comput. 7 (12) (2011) 3961–3977. arXiv:http://dx.doi.org/10.1021/ct200597h, doi:10.1021/ct200597h.
URL http://dx.doi.org/10.1021/ct200597h

Table 1: Co-iterative algorithm for orbital optimization

Davidson iteration	AH iteration	AH matrix elements Hessian	Gradient	Davidson space size
0		$\mathcal{H}$	$\mathcal{G}_{[0]}$	1
1		$\mathcal{H}$	$\mathcal{G}_{[0]}$	2
2	$\mathbf{R}^1_{[0]}$	$\mathcal{H}$	$\mathcal{G}_{[1]} = \mathcal{G}_{[0]} + \mathcal{H}\mathbf{R}^1_{[0]}$	3
3	$\mathbf{R}^1_{[1]}$	$\mathcal{H}$	$\mathcal{G}_{[2]} = \mathcal{G}_{[1]} + \mathcal{H}\mathbf{R}^1_{[1]}$	4
4	$\mathbf{R}^1_{[2]}$	$\mathcal{H}$	$\mathcal{G}_{[3]} = \mathcal{G}_{[2]} + \mathcal{H}\mathbf{R}^1_{[2]}$	5
5	$\vdots$	$\mathcal{H}$	$\vdots$	6

Table 2: The types of two-electron integrals required by different approximations for MO-driven and AO-driven algorithms.

	MO-driven	AO-driven
FEP1	$(AA|**), (A*|A*),$ $(CC|VV), (CV|CV), (CV|AC),$ $(CV|AV), (AV|CC), (AC|VV)$	$(AA|**), (A*|A*)$
DEP1	$(AA|**), (A*|A*),$ $(CC|C*), (CC|V*), (CV|C*), (CV|A*), (AV|C*),$ $(CA|C*), (CA|V*)$	$(AA|**), (A*|A*)$
DEP2	$(A*|**), (CC|**), (CV|**)$	$(A*|**)$
DEP4	$(**|**)$	$(**|**)$

Notation	Description
FEP1	Frozen $\mathcal{H}^{oo}$. First order $\mathbf{R}$ expansion for $\mathcal{G}^o$ and CI Hamiltonian
DEP1	Frozen $\mathcal{H}^{oo}$. First order $\mathbf{T}$ expansion for $\mathcal{G}^o$ and CI Hamiltonian
DEP2	Frozen $\mathcal{H}^{oo}$. Second order $\mathbf{T}$ expansion for $\mathcal{G}^o$ and CI Hamiltonian
DEP1+	First order $\mathbf{T}$ expansion for $\mathcal{H}^{oo}$, $\mathcal{G}^o$ and CI Hamiltonian
DEP2+	Second order $\mathbf{T}$ expansion for $\mathcal{H}^{oo}$, $\mathcal{G}^o$ and CI Hamiltonian
DEP4+	Exact $\mathcal{H}^{oo}$, $\mathcal{G}^o$ and CI Hamiltonian

Table 3: DEP/FEP approximations for microiteration updates.

Table 4: Number of macroiterations required for convergence using different DEP/FEP approximations. The convergence threshold is 10^{-8} Hartrees.

	N_2	CO	HF	C_2	C_2	O_3	NO_2	CH_2	HCHO	C_6H_6
CASSCF	(10,8)	(10,8)	(8,9)	(8,8)	(8,8)	(12,9)	(5,6)	(6,14)	(12,10)	(6,6)
State	$^1\Sigma_{g+}$	$^1\Sigma_{g+}$	$^1\Sigma_{g+}$	$^1\Sigma_{g+}$	$^3\Pi_{u+}$	$^1A_1$	$^2A_1$	$^3B_2$	$^1A_1$	^1A_g
DEP4+	3	4	5	3	3	3	5	5	6	3
DEP2+	3	4	5	3	3	3	5	5	6	3
DEP1+	3	4	5	3	3	4	5	5	6	3
DEP2	3	4	5	3	3	3	5	5	6	3
DEP1	3	4	5	3	3	4	5	5	6	3
FEP1	3	4	5	3	3	4	5	7	6	3
2-step	10	13	17	9	11	9	17	22	21	7

Table 5: Number of macro iterations, micro iterations and J,K calls to converge the singlet, triplet, and quintet states of Fe(II)-porphine with CASSCF(8,11) and DMRG-CASSCF(22,27)

| State	Active space	Basis	Energy	macro	micro	J,K calls
^1A_g	CAS(8,11)	DZ	-2244.7656583	4	13	63
		TZ	-2244.9928018	5	15	74
		QZ	-2245.0550513	5	15	72
$^3B_{1g}$	CAS(8,11)	DZ	-2244.8155591	18	54	274
		TZ	-2245.0429961	10	34	201
		QZ	-2245.1050841	9	33	197
		5Z	-2245.1187001	9	32	203
$^3B_{3g}$	CAS(8,11)	DZ	-2244.8113231	5	15	74
		TZ	-2245.0378914	6	18	93
		QZ	-2245.0999279	6	18	100
^5A_g	CAS(8,11)	DZ	-2244.8291051	5	17	88
		TZ	-2245.0559164	6	18	91
		QZ	-2245.1180998	6	21	126
		5Z	-2245.1316309	6	19	92
$^5B_{2g}$	CAS(8,11)	DZ	-2244.8204477	7	22	113
		TZ	-2245.0474655	5	15	76
		QZ	-2245.1095986	6	19	102
$^3B_{1g}$	CAS(22,27)	DZ	-2245.0006085	8	26	136
		QZ	-2245.2917119	7	22	121
		5Z	-2245.3060865	8	20	94
$^3B_{3g}$	CAS(22,27)	DZ	-2244.9974501	8	29	145
^5A_g	CAS(22,27)	DZ	-2245.0062936	8	28	148
		QZ	-2245.2974638	10	33	198
		5Z	-2245.311861	8	29	149
$^5B_{2g}$	CAS(22,27)	DZ	-2244.9985422	8	29	150
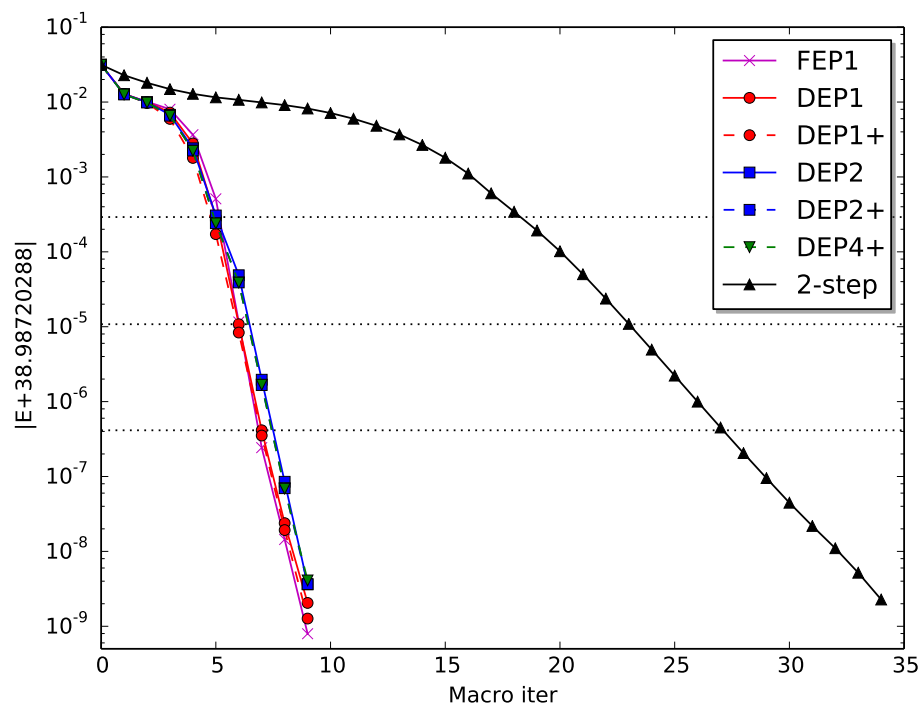
Figure 1: Convergence of the energy of the $^1A_1$ state of CH_2 molecule with CAS(6,14)/cc-pVDZ.
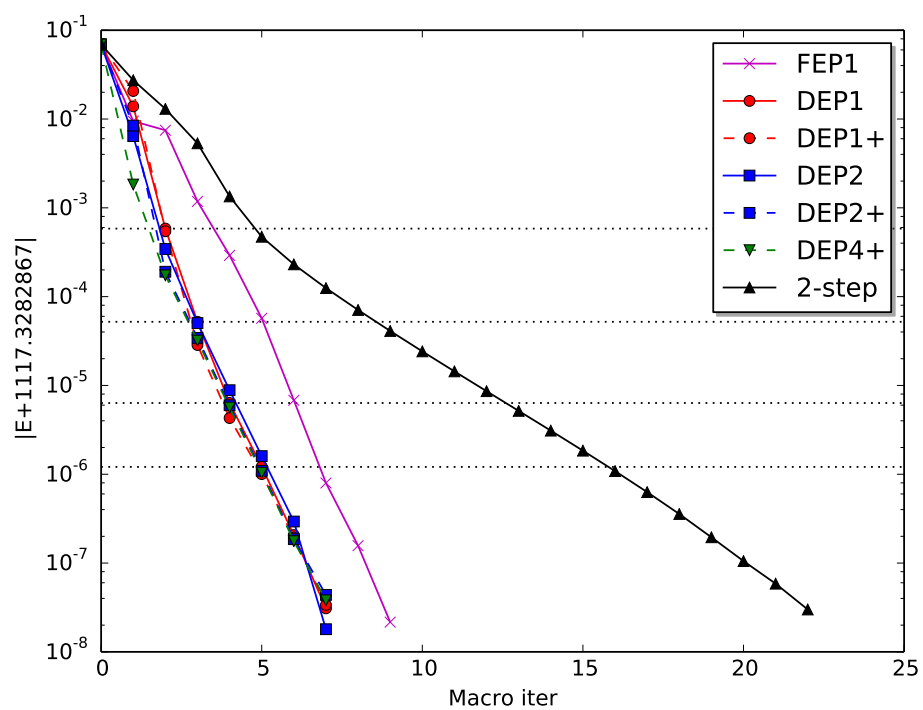
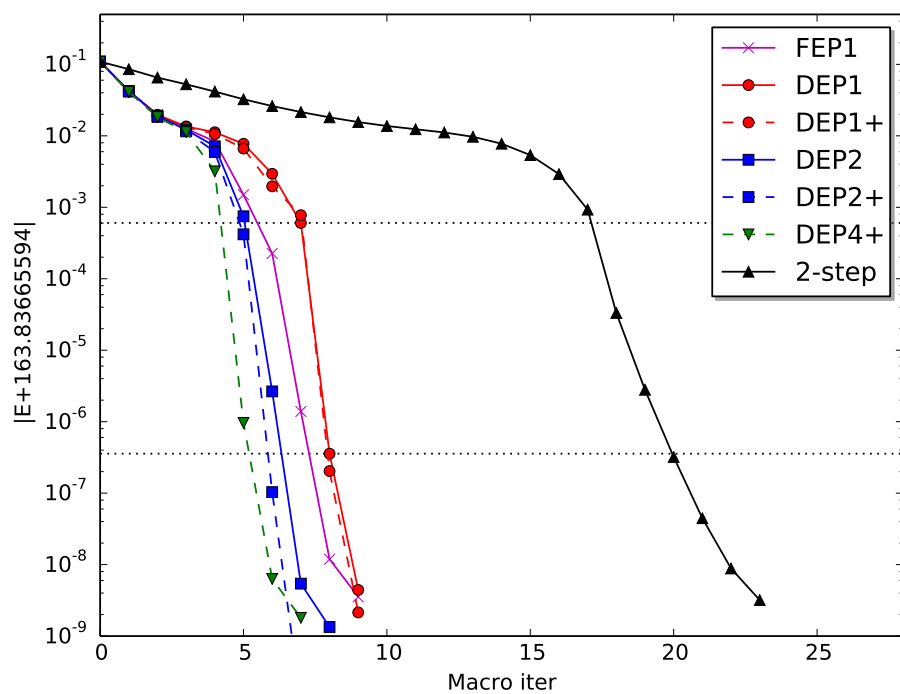Figure 2: Convergence of the ground-state energy of Al_4O_2 with CAS(12,12)/aug-cc-pVDZ.

Figure 3: Convergence of the ground-state energy of HN_3 with CAS(10,10)/cc-pVTZ.