Heriot-Watt University

Heriot-Watt University
Research Gateway

Calculations of potential energy surfaces using Monte Carlo configuration interaction

Coe, Jeremy; Taylor, Daniel; Paterson, Martin J

Link to publication in Heriot-Watt Research Gateway

Calculations of potential energy surfaces using Monte Carlo configuration interaction

Jeremy P. Coe, Daniel J. Taylor, and Martin J. Paterson

Additional information on J. Chem. Phys.

Calculations of potential energy surfaces using Monte Carlo configuration interaction

Jeremy P. Coe, Daniel J. Taylor, and Martin J. Paterson[a)]
Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

We apply the method of Monte Carlo configuration interaction (MCCI) to calculate ground-state potential energy curves for a range of small molecules and compare the results with full configuration interaction. We show that the MCCI potential energy curve can be calculated to relatively good accuracy, as quantified using the non-parallelity error, using only a very small fraction of the full configuration interaction space. In most cases the potential curve is of better accuracy than its constituent single-point energies. We finally test the MCCI program on systems with basis sets beyond full configuration interaction: a lattice of 50 hydrogen atoms and ethylene. The results for ethylene agree fairly well with other computational work while for the lattice of 50 hydrogens we find that the fraction of the full configuration interaction space we were able to consider appears to be too small as, although some qualitative features are reproduced, the potential curve is less accurate. © *2012 American Institute of Physics.* [http://dx.doi.org/10.1063/1.4767052]

I. INTRODUCTION

Full configuration interaction (FCI) enables, in theory, a quantum system to be modelled as accurately as possible within a given basis set, but the rapidly increasing number of configurations means that such calculations are out of reach except for sufficiently small molecules and basis sets. Standard truncation methods, such as considering only single and double excitations with respect to a reference state or restricting excitations to a pre-defined space, can reduce the size of the calculation but often at the expense of the accuracy and consistency of the correlation energy as important configurations may be neglected.

It is acknowledged, however, that a large proportion of the states in many full configuration interaction wavefunctions tend to have practically negligible coefficients and should not be expected to individually contribute substantially to the properties of the wavefunction. Novel approaches to truncated configuration interaction have been developed to attempt to tackle the difficulties posed by the exceedingly rapid growth of the standard full CI space by exploiting this observation and trying to seek out the important states, see Ref. 1 for a recent review. For example these include *a priori* estimates,[2,3] Monte Carlo sampling of a density-matrix renormalization group (DMRG) calculation[4] and estimates from perturbation theory.[5] Another promising method is that of Monte Carlo configuration interaction (MCCI) developed originally by Greer.[6,7] In MCCI a trial wavefunction is randomly augmented with coupled states in an iterative scheme where those states with coefficients smaller than a certain value in the resulting solution of the Schrödinger equation are eventually removed. Even without prior knowledge of the important configurations or molecular orbitals such a procedure can, in principle, result in a compact wavefunction which recreates much of the energy of the FCI wavefunction but using only a small fraction of the FCI states.

Single-reference methods based on coupled-cluster[8] such as coupled-cluster singles doubles (CCSD)[9] are considered to have some of the best efficiencies in calculations for systems in which the FCI expansion would have one dominant configuration and the correlation is considered dynamic. However, they may have problems—especially if the restricted Hartree-Fock determinant is used as a reference—when there are a number of important configurations and the system is deemed multireference. These configurations are associated with static correlation and, e.g., dissociation energies may be modelled incorrectly. For multireference systems the associated static correlation may be accounted for by using methods such as complete active space self-consistent field (CASSCF),[10] but insight is required into the selection of orbitals for the active space and the calculation becomes intractable as the size of the active space increases. If the orbitals have been selected appropriately then perturbation methods (CASPT2) or multi-reference configuration interaction (MRCI) may then be used to account for the remaining dynamic correlation but at relatively high cost. In principle MCCI can give a compact wavefunction which can account for both static and dynamic correlation with minimal user choices: the accuracy and size of the calculation is essentially controlled only by the coefficient cut-off parameter c_{min}.

MCCI has been successfully applied to the single point energy of the water molecule,[6] and the dissociation energy of HF and H_2O.[11] Electronic excitation energies have also been calculated[12] for the first-row atoms beginning with carbon—and also for silicon—with relatively small average errors and numbers of states. While the excitation energies of molecules CH_2, C_2, N_2, and H_2O were considered in Ref. 13 with only

[a)]Electronic mail: m.j.paterson@hw.ac.uk.

tens of meV error and using only a very small proportion of the states compared with FCI.

In this work, we investigate the ability of MCCI to calculate potential energy curves. If the MCCI results can have an almost balanced error across different geometries then we should be able to produce a curve of better accuracy, as quantified by the non-parallelity error, than its constituent single-point energies. We first assess the usefulness of this method on molecules with basis sets for which FCI results exist. We look at potential curves with respect to the dissociation of one hydrogen from HF, BH, and CH_4. The curves for bond stretches are produced for C_2, F_2, N_2, and H_{12}. The potential curves for NH_3 inversion and the model formation of BeH_2 along the reaction coordinate are also calculated. Finally systems beyond the scope of current FCI are also investigated. We calculate a potential curve for H_{50} which we compare with DMRG[14] results, and a potential curve for ethylene as the torsional angle is varied which we compare with other computational results.

II. METHOD

A FCI wavefunction may be written in the notation of second quantization as

$$|\Psi\rangle = c_0 |\Psi_0\rangle + \sum_{i,j} c_i^j a_j^\dagger a_i |\Psi_0\rangle$$
$$+ \sum_{k<i,l<j} c_{ik}^{jl} a_l^\dagger a_k a_j^\dagger a_i |\Psi_0\rangle + \cdots, \qquad (1)$$

where $a_i^\dagger$ (a_i) creates (annihilates) the orbital i in a state. Here letters appearing as subscripts on the coefficients label occupied orbitals in the reference state $|\Psi_0\rangle$ while superscripts are unoccupied orbitals in the reference. A traditional truncation would limit the calculation size by only including substitutions up to a certain substitution level. Limiting the expansion to only the depicted terms in Eq. (1) would correspond to a CISD calculation. MCCI instead randomly augments a wavefunction $|\Psi\rangle = \sum_i c_i |\psi_i\rangle$ with coupled configurations, i.e., states $\langle\psi_j|$ such that $\langle\psi_j|\hat{H}|\Psi\rangle \neq 0$ so that, in principle, the important configurations can eventually be found regardless of their substitution level. Configuration state functions (CSFs) are used in MCCI rather than the more usual Slater determinants (SDs). This means that the MCCI wavefunction fulfils at least one of the requirements of the exact wavefunction: it is an eigenfunction of $\hat{S}^2$. In addition fewer states are needed, e.g., the ratio of SDs to the number of CSFs[15] when $S = M_s = 0$ for 8 electrons and 30 basis functions is around 4.4. One downside is that the calculation of the overlap and Hamiltonian matrix is more complicated than when using SDs.[16]

We briefly describe the full procedure for the MCCI algorithm[6,7] which consists of branching, diagonalization, and pruning. In the branching step, coupled CSFs, created by a symmetry preserving stochastic single or double substitution with equal probability in the current set of CSFs, are added. Branching is always attempted from CSFs with a coefficient greater than a certain value while other CSFs each have probability one half of being used. The coefficients are

then found for the approximate wavefunction expanded in this augmented set of CSFs by solving the generalised eigenvalue problem of $Hc = ESc$. Here H and S are the Hamiltonian and overlap matrix within this subspace. Finally there is a pruning step: added CSFs with coefficients less than a certain value, c_{min}, in the wavefunction expansion are then discarded. These three steps are repeated for a large enough number of iterations so that the energy appears to have converged. In addition, every k iterations (where $k = 10$ in this work) a pruning step is implemented whereby all of the CSFs in the wavefunction are also considered as candidates for removal depending on the magnitude of their coefficient. Furthermore there is no branching or pruning on the last iteration, but a full pruning step is carried out on the penultimate iteration. The program can run in parallel where states that have been added and retained at each step on each processor are shared with all other processors. The results in this work come from MCCI calculations using either eight or twelve processors.

We initialise the procedure with the CSF formed from the restricted Hartree-Fock wavefunction within a given basis. New Hartree-Fock orbitals are calculated at each geometry. For the calculation of the Hartree-Fock molecular orbitals and their one-electron and two-electron integrals we use the MOLPRO package.[17]

III. RESULTS

As the potential is defined up to an additive constant then it is the shape of the potential energy curve that is important. So two curves that differed only by a constant would for all practical purposes be the same. A useful approach to quantify the accuracy of the MCCI potential curve is therefore the non-parallelity error (NPE)[18] defined here for an approximate energy E^{approx} as

$$\text{NPE} = \max_R \left| E_R^{FCI} - E_R^{approx} \right| - \min_R \left| E_R^{FCI} - E_R^{approx} \right|,$$
$$(2)$$

where R ranges over all considered bond lengths.

Hence if we can achieve a balanced accuracy in the energy across the geometries we can recover a potential energy curve with higher accuracy than its single-point energies. We note that this is not necessarily guaranteed as if many points are essentially the FCI energy but there are a few points with large errors then the NPE would be high but the mean single-point energy error could be relatively low. To investigate this we attempt to use a small enough c_{min}, where computationally feasible, to produce a sufficiently accurate curve.

A. Hydrogen dissociation

We first consider varying the length of a bond with hydrogen in BH, HF, and CH_4 and compare the potential energy curves with FCI results from Ref. 19. These potential curves represent one of the simplest dissociations: the breaking of a bond to hydrogen. As such they would be expected to have some multireference character and make an interesting test case for MCCI to be compared with other methods. For these three systems one core orbital is frozen in all the results.
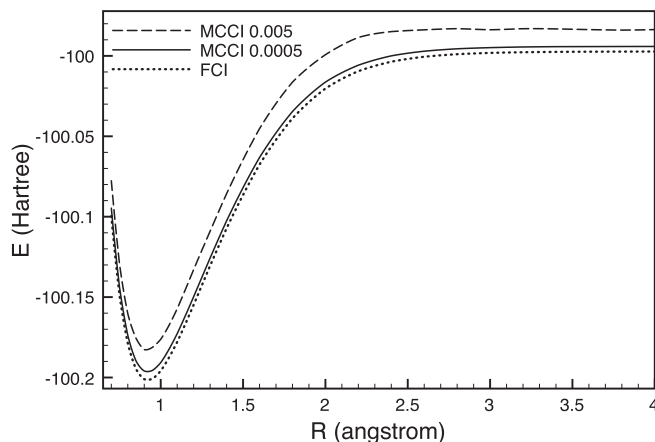
FIG. 1. MCCI and FCI results[19] for energy (Hartree) against bond length R (angstrom) for HF with the 6-31G** basis set and one frozen core orbital.

FIG. 2. Error in two MCCI runs with $c_{min} = 5 \times 10^{-3}$ compared with the FCI results[19] against bond length R (angstrom) for HF with the 6-31G** basis set and one frozen core orbital.

1. HF

We see in Fig. 1 that the shape of the HF FCI curve in a 6-31G* basis is reproduced by MCCI and it appears that for this system, apart from the energy shift, there is not much difference between the two cut-off values. The reduction in NPE by decreasing the cut-off value from $c_{min} = 5 \times 10^{-3}$ to $c_{min} = 5 \times 10^{-4}$ is 5.7 to 1.3 kcal/mol. These NPE compare very favourably with methods based on a single reference where CCSD gives a value of 13.0 kcal/mol, the CCSD(T) NPE is as high as 26.8 kcal/mol, while that of UCCSD(T) is 3.7 kcal/mol.[19] For the multireference approaches considered in Ref. 20, a NPE of 17.69 kcal/mol was found for CASSCF(8, 5) "valence active space" and 4.83 kcal/mol for CASSCF(8, 8) "1:1 active space." CASPT2(8, 5) and CASPT2(8, 8) resulted in NPEs of 2.77 and 0.5 kcal/mol. The NPEs for SOCI(8, 5) and SOCI(8, 8) were 3.20 and 0.04 kcal/mol, respectively. Reference 21 looked into using a minimal active space (one bonding orbital and one anti-bonding orbital for each broken bond) and although their CASSCF(2, 2) had a higher NPE (18.66 kcal/mol) than when using the aforementioned active spaces, the CASPT2 value was lower at 0.47 kcal/mol and the SOCI was better than when using the larger valence active space but, at 0.98 kcal/mol, not as good as when using the 1:1 active space.

The larger cut-off for MCCI required 172.3 CSFs on average, while the smaller needed 1 337. The full configuration space consisted of 3 756 816 SDs[19] so only a very small fraction were used by MCCI to give a NPE of a few kcal/mol (see Table I).

Figures 2 and 3 show that the MCCI results are closest to the FCI when the bond length is large. We also see that rerunning the calculations with a different random number seed does not change the error much and the variation in error seems to be lower as the cut-off is decreased.

As the MCCI potential curves (Fig. 1) appear to have essentially reached dissociation, we may compare MCCI energies at $R = 4$ Å with the total energy from appropriate MCCI calculations on the fragments to indicate how close the MCCI calculation is to size consistency. We find a difference of 4.0 kcal/mol for the larger cut-off and 0.15 kcal/mol for the smaller cut-off.

2. BH

The results for BH were calculated in an aug-cc-pVQZ basis. Preliminary MCCI results for BH are contained in Ref. 1. In Fig. 4, we can see that with a cut-off of 5×10^{-3} the general shape of the MCCI potential curve is qualitatively correct with the minimum in approximately the right place and a well-behaved dissociation although there is a small shoulder in the potential around $R = 3$ Å. The limiting value

TABLE I. MCCI mean CSFs compared with FCI symmetry-adapted SDs.

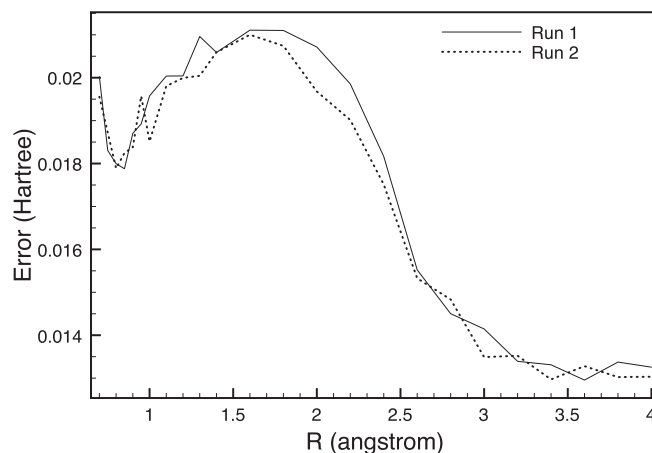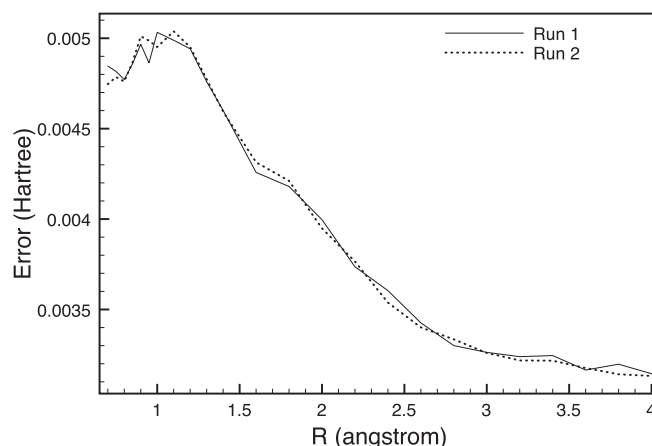| Method	BH states	HF states	CH$_4$ states
MCCI0.005 (CSFs)	333.1	172.3	417.0
MCCI0.0005 (CSFs)	4219.6	1337.0	4272.0
FCI (SDs)	15 132 412	3 756 816	26 755 625
MCCI0.005 fraction	0.0022%	0.0046%	0.0016%
MCCI0.0005 fraction	0.028%	0.036%	0.016%

FIG. 3. Error in two MCCI runs $c_{min} = 5 \times 10^{-4}$ compared with the FCI results[19] against bond length R (angstrom) for HF with the 6-31G** basis set and one frozen core orbital.
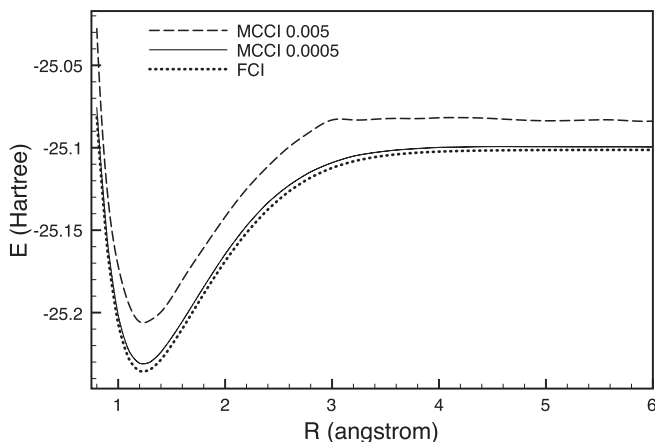
FIG. 4. MCCI and FCI results[19] for energy (Hartree) against bond length R (angstrom) for BH with the aug-cc-pVQZ basis set and one frozen core orbital. Adapted from Ref. 1.

FIG. 5. MCCI ($c_{min} = 5 \times 10^{-4}$) energy error (Hartree) when compared with FCI results against the bond length R (angstrom) for BH with the aug-cc-pVQZ basis set and one frozen core orbital. The NPE is 2.6 kcal/mol.

of the energy is reached too early however. By decreasing c_{min} to 5×10^{-4} the curve is almost that of the FCI shifted by a small energy. The NPE is 22.8 kcal/mol in the former case and reduces to 2.6 kcal/mol with the smaller c_{min} value. For this value of c_{min} the mean number of CSFs was only 4220 compared with around 15×10^{6} Slater determinants in the FCI space (see Table I) while 330 were required on average for the larger case. From Ref. 19, the NPE value for CCSD was 8.1 kcal/mol, that of CCSD(T) was 23.3 kcal/mol while UCCSD(T) had a NPE as low as 3.1 kcal/mol. Multireference methods have also been considered[20] and compared with FCI results for the potential curve of BH but with a smaller range and a cc-pVQZ basis as there were difficulties in convergence when using the aug-cc-pVQZ basis. Although not a direct comparison, we note that the CASSCF(4, 4) "valence active space" there had a NPE of 12.68 kcal/mol, CASSCF(4, 5) "1:1 active space" gave 9.38 kcal/mol, CASPT2(4, 5) gave 3.16 kcal/mol while only the second-order CI (SOCI) results were lower than the MCCI value in this work at 0.29 and 1.54 kcal/mol for active spaces of (4, 4) and (4, 5), respectively.

We see in Fig. 5 that the energy error for this system at $c_{min} = 5 \times 10^{-4}$ when compared with FCI generally decreases as the bond length increases. We note that the difference between the MCCI energy at $R = 6$ Å and the total MCCI energy for the fragments is 6.6 kcal/mol at $c_{min} = 5 \times 10^{-3}$ and 0.83 kcal/mol at $c_{min} = 5 \times 10^{-4}$.

3. CH$_4$

As in Ref. 19 a tetrahedral geometry was used for methane with bond length 1.086 Å for the three carbon-hydrogen bonds that were not varied. Preliminary MCCI results for CH$_4$ are contained in Ref. 1. Figure 6 shows that, in a 6-31G* basis, MCCI again captures the shape of the FCI potential curve. Now it appears that the larger cut-off may be a little too high in energy at large R compared with the equilibrium geometry. This is quantified using the NPE which is 10.3 kcal/mol for $c_{min} = 5 \times 10^{-3}$ but drops to 0.6 kcal/mol at the smaller cut-off. We note
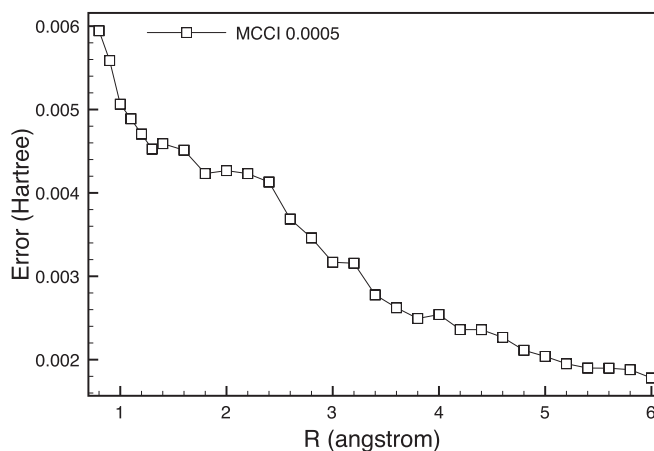
that in Ref. 19, CCSD gave a value of 10.3 kcal/mol, UCCSD gave 5.1 kcal/mol while UCCSD(T) was as low as 3.2 kcal/mol. In Ref. 20, the definition of the 1:1 active space was equivalent to the valence active space in this case and gave CASSCF(8, 8) with a NPE of 6.34 kcal/mol, CASPT2(8, 8) gave 1.56 kcal/mol, CISD[TQ](8, 8) gave 1.33 kcal/mol, while only SOCI(8, 8) had a lower NPE than the MCCI results with 0.3 kcal/mol. The minimal active space results[21] had a NPE of 9.25 kcal/mol for CASSCF(2,2), with CASPT2 at 1.22 kcal/mol and SOCI a little worse with 0.6 kcal/mol. Hence only one of the results from these two works using multireference methods gave a lower NPE for methane.

The FCI space is around 26.7×10^{6} SDs,[19] but in Table I we see that the MCCI results are achieved using only 417 and 4 272 CSFs on average for the larger and smaller c_{min}, respectively.

In Fig. 7, we see that the error in the energy compared with FCI is greatest at very small bond lengths and now the minimum error is at bond lengths a little longer than that at equilibrium, but the errors are all very small. The difference in energies between MCCI applied to the system
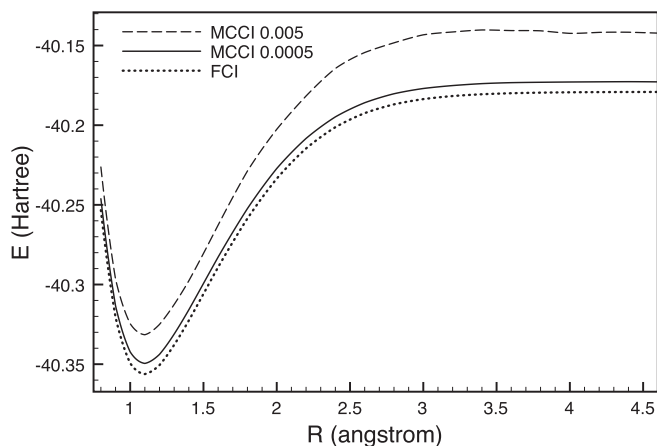

FIG. 6. MCCI and FCI results[19] for energy (Hartree) against one carbon-hydrogen bond length R (angstrom) for CH$_4$ with the 6-31G* basis set and one frozen core orbital. Adapted from Ref. 1.
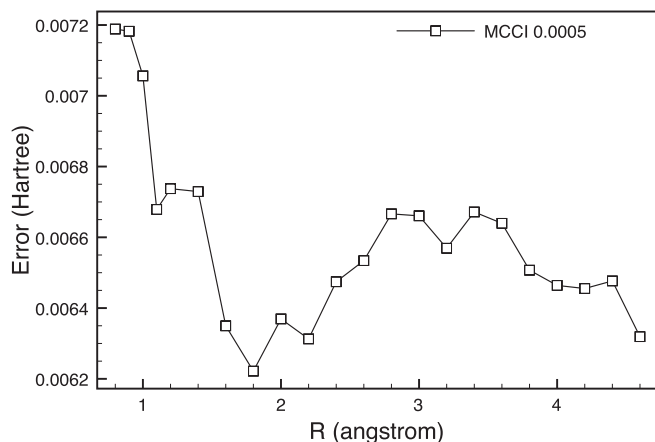
FIG. 7. MCCI ($c_{min} = 5 \times 10^{-4}$) energy error (Hartree) when compared with FCI results against one carbon-hydrogen bond length R (angstrom) for CH_4 with the 6-31G* basis set and one frozen core orbital. The NPE is 0.6 kcal/mol.

FIG. 8. MCCI ($c_{min} = 5 \times 10^{-4}$) and FCI[22] results for energy (Hartree) against the bond length R (angstrom) for C_2 using the 6-31G* basis set with two frozen core orbitals. Here X and B' are $^1\Sigma_g^+$ and B is $^1\Delta_g$. Adapted from Ref. 1.

at $R = 4.6$ Å and the sum of MCCI applied to the fragments is 15 kcal/mol at the larger cut-off and 0.79 kcal/mol at the smaller cut-off. For HF, BH and CH_4 at $c_{min} = 5 \times 10^{-3}$ the difference between the FCI and MCCI energy, at all geometries considered, ranged from around 8.4 kcal/mol to 33.8 kcal/mol while this range was around 1.1 kcal/mol to 4.5 kcal/mol when the cutoff was lowered to $c_{min} = 5 \times 10^{-4}$.

We have seen, for the three systems considered for hydrogen dissociation, that by going to a low enough coefficient cut-off value of $c_{min} = 5 \times 10^{-4}$ we can reproduce the shape of the potential curve to relatively high accuracy: we achieve a non-parallelity error of around a few kcal/mol which is better than the single-reference methods we compare with from the literature. This is accomplished using in the order of hundredths of a percent of the full CI space. The accuracy was also higher than many of the multireference results from the literature. Some CASPT2 and SOCI calculations, using large active spaces or chemical intuition in the construction of a smaller active space, produced a lower non-parallelity error but the number of states used for these calculations was not reported for comparison.

B. Carbon dimer

Next we look at the dissociation of C_2 in a 6-31G* basis with two frozen orbitals. This system is known to be multireference and possess low-lying excited states hence poses a stern test of MCCI. Preliminary MCCI results for C_2 are contained in Ref. 1. FCI results[22] and the MCCI potential curve with $c_{min} = 5 \times 10^{-3}$ are shown in Fig. 8. The MCCI curve accuracy is somewhat difficult to judge by eye due to the many close excited states, but it appears to follow the FCI ground state closely except for bond lengths (R) around 3 Å where it seems that a low lying excited state (B') has been converged upon instead. However, it appears that the three FCI curves are in fact tending to a degenerate state as the system dissociates. In addition, the lowest energy state between about 1.7 and 2.5 Å is B rather than X and MCCI may have converged to this as would be expected: the states have differ-
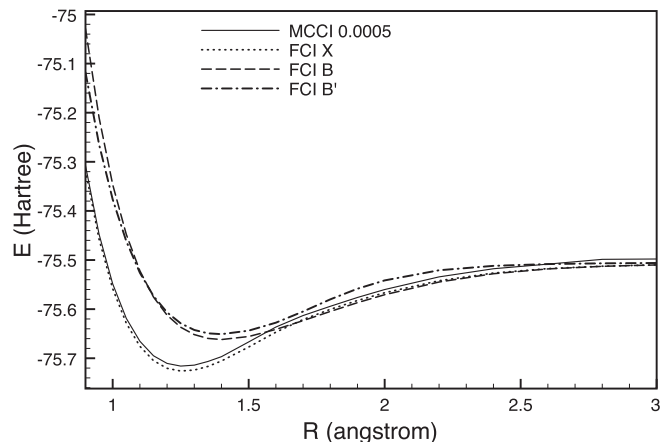
ent symmetry so can cross but have the same spatial symmetry in the abelian subgroup (D_{2h}) used for computation.

The mean number of CSFs employed was ~6900 for MCCI here, while FCI results from Ref. 22 required 52 407 353 symmetry-adapted Slater determinants. This gives a ratio of 0.013%. The NPE is 4.9 kcal/mol for MCCI here which is substantially better than methods based on a single reference. From Ref. 22, CCSD had a NPE of 24.3 kcal/mol and that of CCSD(T) was 61.3 kcal/mol, highlighting the instability of coupled cluster at large R. Unrestricted methods behaved better at dissociation but were less good at intermediate bond lengths resulting in a NPE for UCCSD(T) of 21.6 kcal/mol. CISDTQ gave a curve with 16.6 kcal/mol for the NPE. It is worth noting that a CISDTQ, without symmetry, and using Slater determinants would be expected to have around 3×10^6 terms. Even allowing for the reduction in the size of the space through spatial symmetry and the use of CSFs, the MCCI result with ~6900 CSFs on average and a better NPE shows the usefulness of the MCCI approach here. Valence active space CASSCF in Ref. 23 used 660 determinants and had a NPE of 5.4 kcal/mol for (SA)-CASSCF. (EOM)CCSD gave 24.3 kcal/mol and CR-(EOM)CCSD(T), III gave 13.5 kcal/mol for the NPE. CISD[TQ] had a NPE of 16 kcal/mol in that work and used 87 415 determinants. In fact only two of the multireference techniques investigated in Ref. 23 had a lower NPE: (SA)-CASSCF MRCI with 0.4 kcal/mol using 270 338 determinants and SA-CASPT2 with a NPE of 3.8 kcal/mol.

A wavefunction for the carbon dimer found using FCI quantum Monte Carlo (FCIQMC) in Slater determinant space in Ref. 24 required 9.15×10^6 walkers at 0.9 Å in this 6-31G* basis to give almost FCI results although the point at $R = 3$ Å was neglected. We note that with a cc-pVDZ basis and requiring that only states with zero angular momentum were allowed they calculated very similar results to MRCI, and with lower energy for $R > 1.15$ Å than when using 6-31G*, with 2×10^6 walkers compared with a standard FCI space of 4.74×10^9 Slater determinants. This suggests that the performance of MCCI here could perhaps be improved
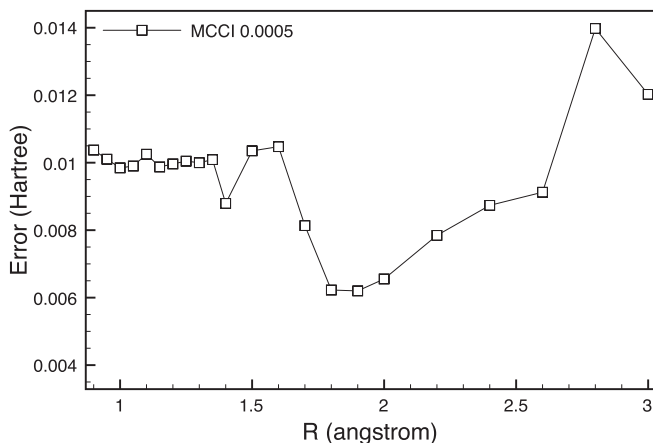
FIG. 9. MCCI ($c_{min} = 5 \times 10^{-4}$) energy error (Hartree) when compared with FCI[22] results against the bond length R (angstrom) for C_2 using the 6-31G* basis set with two frozen core orbitals. The NPE is 4.9 kcal/mol.

if we also discriminate between states using their angular momentum in addition to their symmetry and spin.

We note that MCCI potential curve for the carbon dimer is of slightly higher accuracy, in a sense, than its constituent single-point energy calculations as the mean single-point error was around 6.0 kcal/mol. In addition, if we exclude the two points around $R = 3$ Å, that seem to have converged to a low-lying excited state, then the single point energy error only reduces a little to 5.7 kcal/mol but the NPE is now just 2.7 kcal/mol. In Fig. 9, we see that the error does not vary too much, as summarised by the NPE, but is highest approaching the largest bond length. Interestingly it is lowest not at equilibrium but at bond lengths from around 1.8 Å to 2 Å, which in this case may be due to the lowest energy state being reached, not the original ground-state. This slightly lowers the mean single-point error in this case, and will slightly raise the NPE. At 3 Å, we compare the MCCI energy for the full system with that of the sum of the MCCI energies of the fragments and find a difference of 3.2 kcal/mol. We note that as the energy of the system using MCCI is higher than the energy of the fragments here then as the system has perhaps not sufficiently dissociated the difference is perhaps smaller than if a larger R had been considered.

1. Analysis of the C_2 MCCI wavefunction

At the equilibrium geometry for C_2 the MCCI wavefunction has states with as many as 7 substitutions from the Hartree-Fock reference and the multiconfigurational nature is evident from the largest ten coefficients of the CSFs in Table II. By $R = 2.0$ Å, the dominant configuration is now a double substitution of the Hartree-Fock reference which could suggest that the lowest lying state here (B), not the equilibrium ground-state curve (X), may have been converged upon as expected or could just be an aspect of the multiconfigurational nature which is is more apparent here: the largest two coefficients are -0.636 and 0.557. By $R = 3.0$ Å, the largest coefficient belongs to a CSF which differs to the major contributor in the previous two cases. This may be linked to the

TABLE II. Ten largest coefficients of the C_2 MCCI wavefunction at $R = 1.25$ Å with the number of substitutions with respect to the Hartree-Fock reference listed by spin.

Coefficient	α substitutions	β substitutions
-0.830	0	0
0.331	1	1
-0.184	1	1
-0.180	1	1
-0.179	1	1
-0.160	0	2
0.158	1	1
0.158	1	1
-0.142	0	1
0.106	2	2

suggestion from the potential curves that here a different excited state is converged upon at large R.

C. F_2

In Fig. 10, we now display the MCCI potential curve for F_2 in a cc-pVDZ basis set without any frozen orbitals. The MCCI results can be seen to follow the shape of the full valence CI results of Ref. 25 despite the MCCI calculation not freezing any orbitals. We can also see how the Hartree-Fock energy increases much too rapidly as the bond length becomes large.

We note that the MCCI with $c_{min} = 5 \times 10^{-4}$ required around 3 577 CSFs on average for the depicted points. The FCI space with two frozen orbitals would be expected to comprise 4.3×10^{11}. Even though only a very small fraction of this space is used (8.3×10^{-7}), when compared with the full valence CI points the NPE for MCCI is 6.2 kcal/mol.

In Fig. 11, we display the difference between the MCCI and Full valence CI for the latter's points in the range displayed in Fig. 10. Here the MCCI error tends to decrease with increasing bond length for the points displayed. Comparing the MCCI energy of the system at $R = 3$ Å with the sum of
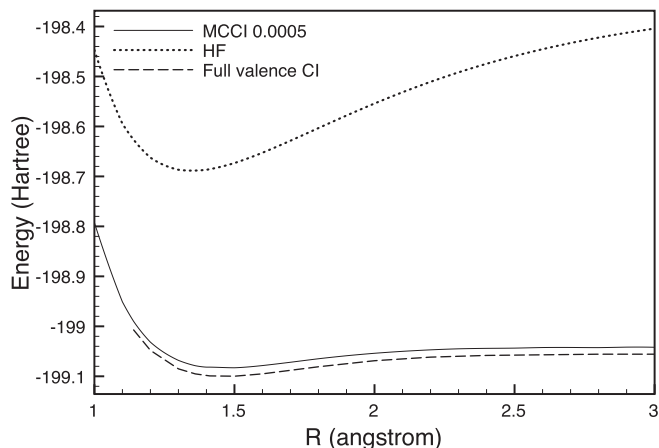

FIG. 10. MCCI (two frozen orbitals) with $c_{min} = 5 \times 10^{-4}$, Full valence CI[25] and Hartree-Fock results for energy (Hartree) against bond length R (angstrom) for F_2 using the cc-pVDZ basis set.
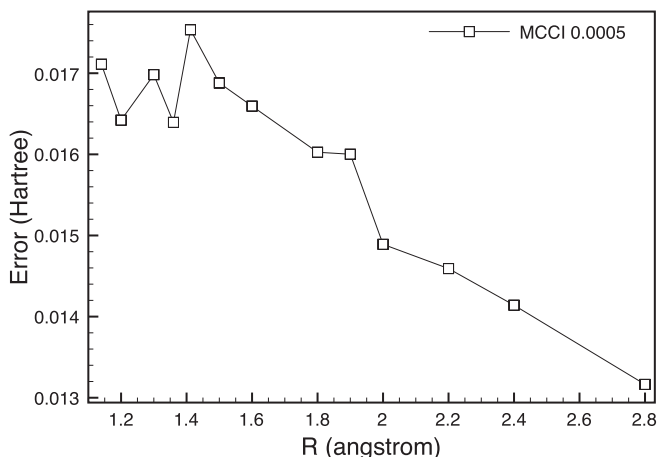
FIG. 11. Energy error of MCCI (two frozen orbitals) with $c_{min} = 5 \times 10^{-4}$ when compared with Full valence CI[25] against bond length R (angstrom) for F_2 using the cc-pVDZ basis set. The NPE is 6.2 kcal/mol.

the MCCI energies of the constituent atoms, we find a difference of 7.3 kcal/mol.

D. N_2

The nitrogen molecule is well known for exhibiting multireference character at long bond lengths and here methods based on a single reference can fail, see, e.g., Ref. 26. Hence this is another challenging system to test if MCCI can produce a potential energy curve with balanced error. The MCCI potential curve for a cc-pVDZ basis with two frozen core orbitals is displayed in Fig. 12 and compared with FCI results from a number of sources.[27–29] We see that the general shape of the curve is recovered by MCCI when using a cut-off of $c_{min} = 10^{-3}$ and the gap appears fairly constant except at very small bond lengths. For the values for which FCI results could be found, depicted in Fig. 12, the NPE is 6.6 kcal/mol, while the MCCI CSFs ranged from about a thousand to almost five thousand as the bond length increased with a mean value of 2 854. This is in contrast to a SD space of 4.3×10^9 for a
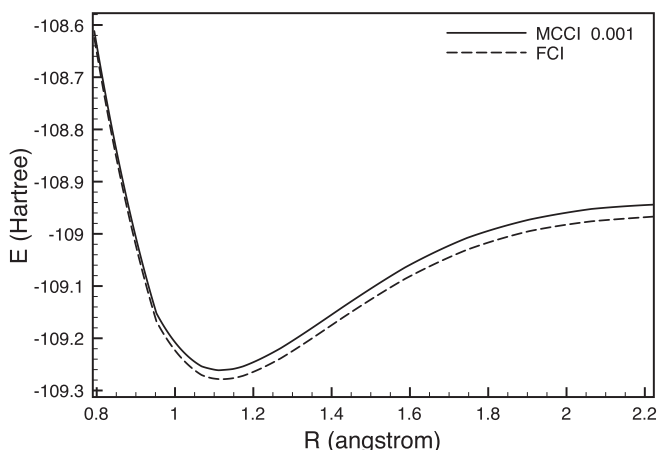

FIG. 12. MCCI ($c_{min} = 10^{-3}$) and FCI[27–29] energy (Hartree) against bond length R (angstrom) for N_2 using the cc-pVDZ basis set with two frozen cores.

FCI when ignoring spatial symmetries. The 200 iterations required from around 3 min at small R to 1.3 h as R became large using 8 processors.

For the six bond lengths considered in Ref. 29 for this system using the DMRG, the NPE in mE_h ranged from 0.4 to 0.01 for DMRG1000 and DMRG4000, respectively, while CCSD(T) was 172.724 and CCSD 57.754. Using a CAS(6, 6) wavefunction as a reference, MRCI and MRCC had 0.464 and 0.732 mE_h respectively, but the number of states used was not specified. While the results of our MCCI calculation with $c_{min} = 10^{-3}$ gave $5.37 mE_h$ but using only a few thousand CSFs on average. The variational parameters of DMRG allow a rough comparison with the size of the space. The number of parameters is known to be $O(M^2 k)$ (see for example Ref. 30) and so for the 26 non-frozen orbitals used there are of the order of 26×10^6 parameters for DMRG1000 ($M = 1000$). Reference 31 uses the same set of points to compare the accuracy of multireference coupled cluster approaches, again with a CAS(6, 6) reference. For state-specific multireference CCSD (SS-MRCCSD) with delocalised orbitals they find a NPE of 12.0 mE_h, single-root MR Brillouin-Wigner CCSD (srMRBWCCSD) had a NPE of 19.4 mE_h, and MR averaged quadratic CC MRAQCC gave 1.1 mE_h.

In Ref. 21, the 13 points of Ref. 27 are used. Over these points we calculate the MCCI NPE as 5.9 kcal/mol. The CASSCF(10, 10) "1:1 active space" calculation[21] had the highest NPE at 22.93 kcal/mol of the methods they considered for this system, while the CASSCF(10, 8) "valence active space" result had a NPE of 15.03 kcal/mol. Using their minimal active space, defined with knowledge of which bonds were broken, resulted in a CASSCF(6, 6) with a NPE of 14.59 kcal/mol. Interestingly the CASPT2(6, 6) had the largest NPE of the three active spaces considered for CASPT2 at 5.2 kcal/mol, unlike their results for HF and CH_4, while CASPT2(10, 10) had the smallest at 1.88 kcal/mol. The SOCI were all below one kcal/mol but it was not clear how many states were used to achieve this result. MCCI with the cut-off employed is hence relatively close in accuracy to the CASPT2 results where knowledge of which bonds are broken has been used and it appears that large active spaces for CASPT2 are needed to reduce the NPE to below 2 kcal/mol.

In Fig. 13, we see that the difference between the MCCI and FCI energy tends to increase with the bond length for the nitrogen molecule.

We have also calculated a potential curve using a cc-pVTZ basis. This is displayed in Fig. 14 and shows how the dissociation is again well-behaved for this larger range of points. Without symmetry considerations the FCI space is around 10^{17} SDs while the mean number of CSFs required for the MCCI curve was around 4 600 at $c_{min} = 10^{-3}$ and around 10 600 at $c_{min} = 5 \times 10^{-4}$. The very large space precludes a comparative FCI calculation. However we may make an approximate comparison with FCI results[32] which used an ANO[4s3p1d] basis with two frozen cores. The difference in energy between the MCCI results at the equilibrium bond length of 1.098 Å and at 4 Å is 0.381 Hartree at the smallest cut-off used. This compares reasonably well with the approximate dissociation energy FCI results of Ref. 32. There an approximate equilibrium bond length of 2.1 Bohr was used to
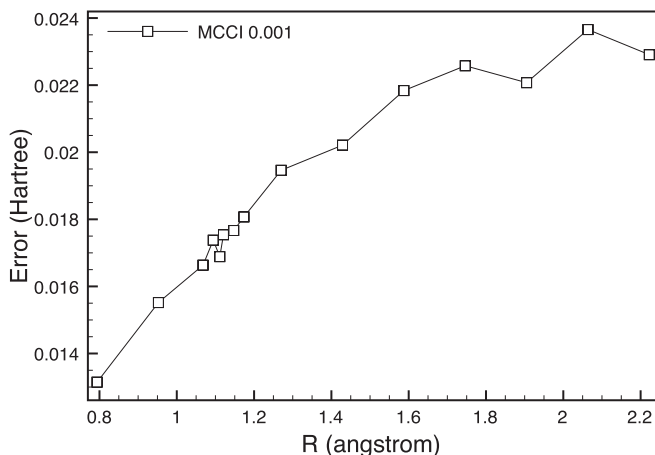
FIG. 13. Error (Hartree) in the MCCI ($c_{min} = 10^{-3}$) calculation when compared with the FCI against bond length R (angstrom) for N_2 using the cc-pVDZ basis set with two frozen cores. The NPE is 6.6 kcal/mol.

give 0.321 Hartree and 40 Bohr was used to calculate the energy at dissociation. When comparing the MCCI energy of the system at $R = 4$ Å with the sum of the MCCI energies of the appropriate fragments we find a difference of 56.8 kcal/mol when using $c_{min} = 10^{-3}$ and a difference of 41.8 kcal/mol when using $c_{min} = 5 \times 10^{-4}$. This suggests that the atoms are treated more accurately when using MCCI with HF MOs on this multireference system with a basis that is beyond current FCI. However, we note that the curve appears to dissociate correctly even with the largest difference seen in this work between the molecule essentially at dissociation and the fragments suggesting that MCCI is the least size-consistent in this case.

To give a somewhat fairer comparison with the FCI results of Ref. 32 we remove the f orbitals from a cc-pVTZ basis, and attempt to use the same geometries as Ref. 32, but keep the MCCI cut-off value the same and do not freeze any core orbitals. For $R = 2.1$ Bohr, we find an energy of -109.296 Hartree using 12 664 CSFs, but the MCCI results were much too high in energy at $R = 40$ Bohr (-93.66 Hartree
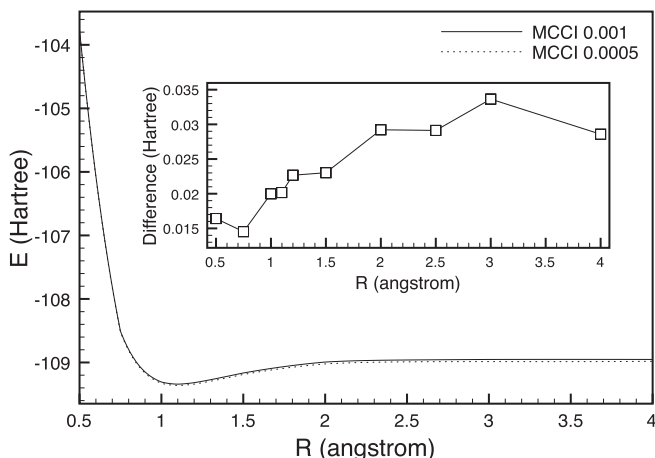

FIG. 14. MCCI energy (Hartree) against bond length R (angstrom) for N_2 using the cc-pVTZ basis set with a cut-off of $c_{min} = 10^{-3}$ or $c_{min} = 5 \times 10^{-4}$. Inset: Energy difference between the MCCI results using the two cut-offs.

using 5, 029 CSFs) which we attribute to the restricted Hartree-Fock (RHF) reference being qualitatively wrong. The amount of states needed to compensate for this and produce a sensible MCCI energy using the RHF molecular orbitals is then perhaps much too large for the cut-off employed. However, we saw that when using a full cc-pVTZ basis the potential curve dissociated as one would expect in Fig. 14 when the largest bond length was 4 Å. Hence we note that the problems resulting in a very poor reference at extremely large bond length could be circumvented, to a degree, by calculating the curve until approximate convergence toward the energy as dissociation where the smaller bond distance means that RHF is not quite as incorrect. To this end we use $R = 7.56$ Bohr (4 Å) and find the energy to be -108.959 with 19 156 CSFs. This gives an approximate dissociation energy of 0.337 Hartree. The MCCI result therefore has an approximate error of around 10 kcal/mol. The symmetry adapted FCI space was around 9.68×10^9 while the mean number of CSFs used by MCCI for the two points was 15 910. Hence here only around $1.6 \times 10^{-4}\%$ of the FCI symmetry adapted space was used.

E. BeH$_2$

The model reaction for the formation of BeH$_2$ was put forward in Ref. 33 to investigate CCSD. There CCSD was found to describe the system well even when the system had multireference character with two states. Recently, BeH$_2$ has been considered as a test system for internally contracted multireference coupled cluster (ic-MRCC).[34] In the aforementioned work the beryllium atom is at the origin and the hydrogen atoms have, in Bohr, the coordinates x, y and x, $-y$ where $y = 2.54 - 0.46x$ and $x \in [0, 4]$. We run the calculations with no frozen orbitals and note that the full CI space is around 4×10^6 SDs when neglecting spatial symmetries, while the MCCI wavefunction with $c_{min} = 10^{-3}$ consisted of CSFs ranging from 386 at $x = 0$ to a maximum of 1002 around the transition state at $x = 2.8$ with a mean value of 628. Full CI results were calculated using MOLPRO.[17, 35, 36] We see that the match between the MCCI and the FCI is both qualitatively correct and quantitatively appears very good in sharp contrast to the incorrect behaviour of Hartree-Fock here (Fig. 15). The good match of the MCCI with the FCI is shown in the NPE value of 0.628 kcal/mol despite the wavefunction only being a few-hundredths of a percent of the size of the FCI SD space. We note that in Ref. 34 the NPE, when using a slightly different basis set, was found to be 0.653 kcal/mol for ic-MRCCSD.

We see in Fig. 16 that the difference between the FCI and MCCI energy is smallest at $x = 4$ Bohr and increases as the reactants approach (x decreases) until about $x = 2$ Bohr from then on it decreases as x decreases. We note that the maximum error does not correspond to the transition state.

F. Ammonia inversion

We investigate the ability of MCCI to reproduce the potential curve of NH$_3$ as its trigonal pyramid inverts by moving through a planar structure. We use the cc-pVDZ basis and
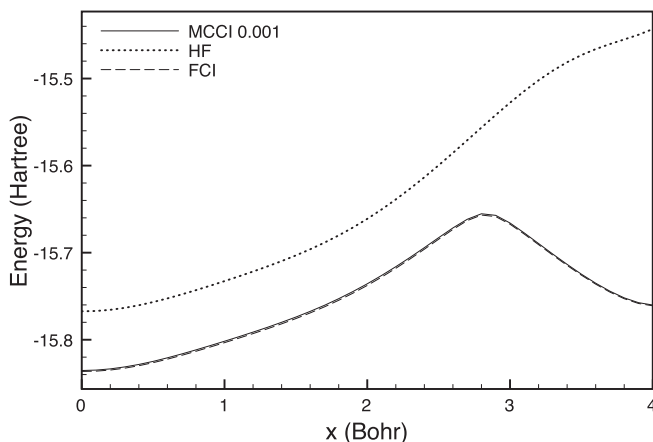
FIG. 15. MCCI ($c_{min} = 10^{-3}$), FCI and Hartree-Fock energy (Hartree) against the reaction coordinate x (Bohr) for BeH$_2$ using the cc-pVDZ basis.

FIG. 17. MCCI ($c_{min} = 10^{-3}$), FCI, and Hartree-Fock energy (Hartree) against angle (θ) for NH$_3$ using the cc-pVDZ basis and one frozen core orbital.

freeze one orbital. The NH bond length is 1.025 Å and the hydrogens are at 120° to each other. The NH bond makes an angle θ with a line passing through the nitrogen and perpendicular to the hydrogen plane.

The MCCI curve is displayed in Fig. 17 and compared with Hartree-Fock and FCI results which were calculated using MOLPRO.[17,35,36] The plot is mirrored around the line $x =$ 90. We see that both MCCI with $c_{min} = 10^{-3}$ and Hartree-Fock recover the shape of the FCI curve with the minima and transition point in approximately the correct place. The MCCI curve appears to be a little too flat as it approaches the planar structure in that the energy does not change as much as the FCI or Hartree-Fock between 85° and 90°. Despite this, for the 11 points from 40° to 90°, the NPE is 2.4 kcal/mol while Hartree-Fock has a NPE of 9.9 kcal/mol. The number of CSFs ranged from 1226 to 1824 with a mean value of 1629. This is in contrast to the size of the FCI SD space when neglecting symmetry of $\sim 4 \times 10^8$.

The difference in energy between the MCCI and FCI calculation is displayed in Fig. 18. For this system we see that the error generally decreases as the planar transition state is approached.
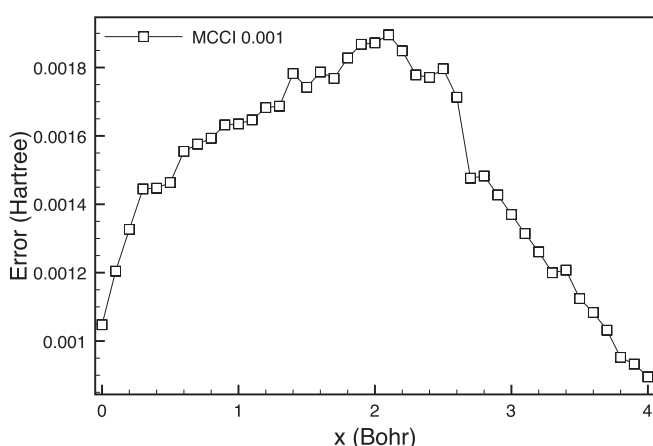
G. Hydrogen lattice

We now consider a linear chain of hydrogen atoms. Such a system would perhaps be expected to be more suitable for modelling using techniques appropriate for strongly-correlated one-dimensional lattice systems such as DMRG[14] methods which can in principle scale linearly with the size of the system; a linear 50 hydrogen system has been considered using the DMRG for quantum chemistry in Ref. 37 with a STO-6G basis. The 1D hydrogen chain therefore poses an interesting challenge for MCCI particularly as the correlation progresses from dynamic to a large amount of, what could be considered, static correlation as the distances between the atoms increase: the FCI results using MOLPRO[17,35,36] for a chain of 12 hydrogens in the STO-6G basis have only two coefficients greater than 0.05 (−0.110 and 0.981) at 1.0 Bohr separation compared with 19 at 4.2 Bohr where the largest is only 0.22.

In Fig. 19, we see that for a chain of 12 hydrogens the Hartree-Fock energy increases much too fast as the distance between hydrogen atoms increases. MCCI, with $c_{min} = 10^{-3}$,

FIG. 16. Error (Hartree) in the MCCI ($c_{min} = 10^{-3}$) energy compared with FCI against the reaction coordinate x (Bohr) for BeH$_2$ using the cc-pVDZ basis. The NPE is 0.63 kcal/mol.
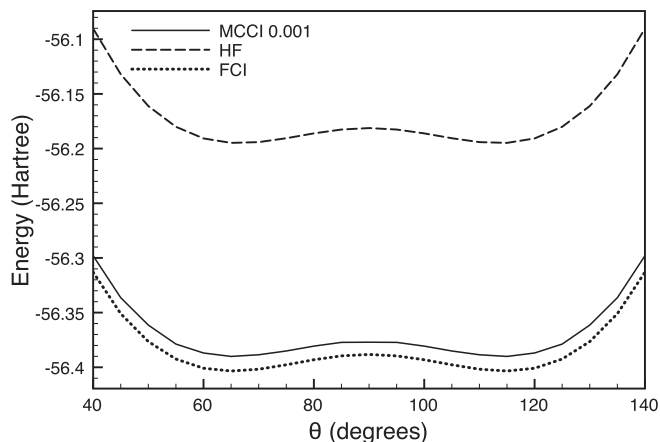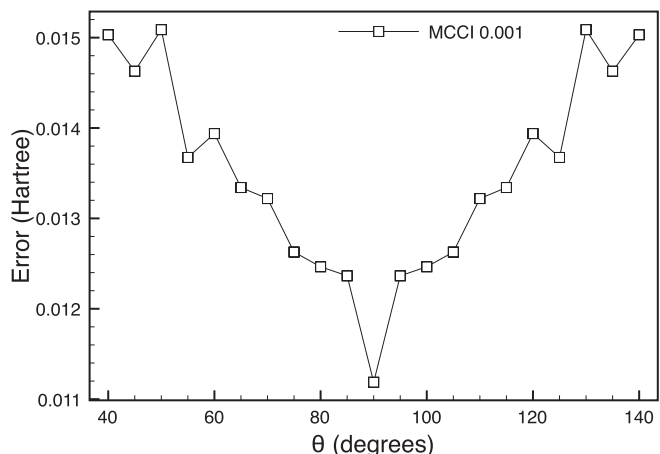

FIG. 18. Error (Hartree) in the MCCI ($c_{min} = 10^{-3}$) energy when compared with the FCI against angle (θ) for NH$_3$ using the cc-pVDZ basis and one frozen core orbital. The NPE is 2.4 kcal/mol.
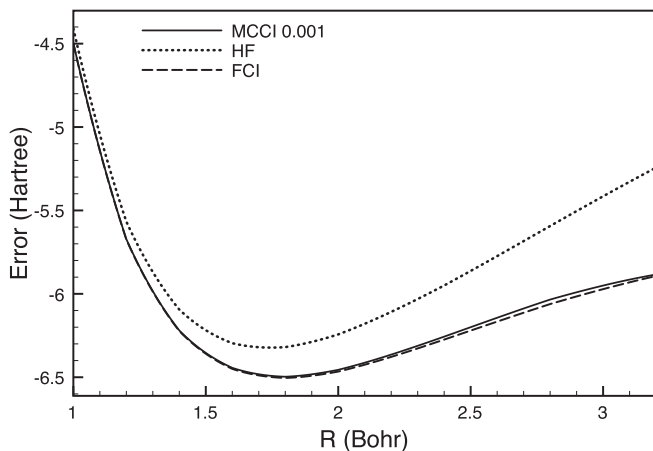
FIG. 19. MCCI ($c_{min} = 10^{-3}$), FCI, and Hartree-Fock energy (Hartree) against distance between atoms R (Bohr) for a chain of 12 hydrogen atoms with a STO-6G basis.

FIG. 21. MCCI ($c_{min} = 5 \times 10^{-4}$), Hartree-Fock, LDMRG(500) (see Ref. 37), and MCCI ($c_{min} = 5 \times 10^{-4}$) using orthogonal atomic orbitals (AOs). Energies (Hartree) against distance between atoms R (Bohr) for a chain of 50 hydrogen atoms when using a STO-6G basis.

is a fairly good match with the FCI results. We note that there was a numerical issue with the MCCI results for $R > 3.2$ Bohr when using MOs. Here MCCI produced an energy slightly lower than the FCI energy. However, when using orthogonal atomic orbitals we found a MCCI energy of -5.694 Hartree at 4.2 Bohr compared with the FCI result of -5.699 Hartree, this seemed to suggest that the use of very poor MOs in MCCI for this challenging system can lead to numerical problems when using CSFs in MCCI. The full configuration space consists of 853 776 determinants while the MCCI wavefunction ranged from 342 at $R = 1$ Bohr to a maximum of 6477 by $R = 2.8$ Bohr. The mean value was 2701 CSFs. The NPE for the displayed results was 15.8 kcal/mol. In Fig. 20, we display the difference between the MCCI and FCI energies. We see that the MCCI energy error increases from its value at $R = 1$ Bohr until $R = 2.8$ Bohr.

We did not satisfactorily calculate a potential curve (Fig. 21) for the 50 hydrogen system compared with DMRG based results of Ref. 37 which we suggest as due to the extremely large configuration space (10^{28}) of which we only included up to ~80 000 states (~10^{-21}% of the full space)
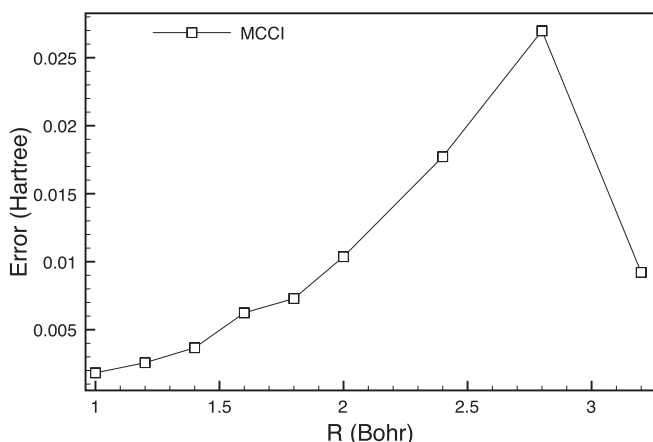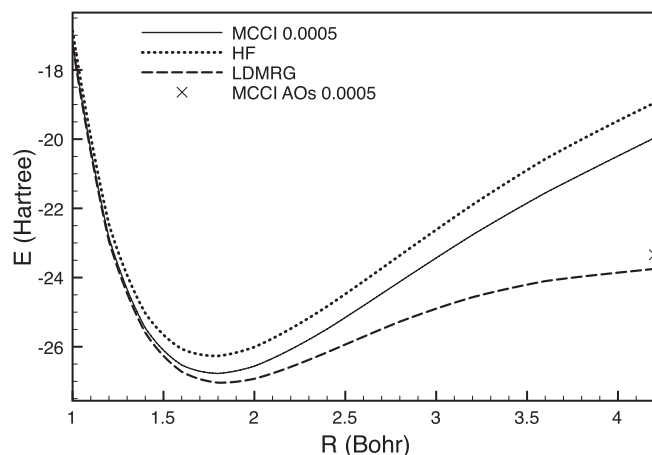

FIG. 20. Error (Hartree) in the MCCI ($c_{min} = 10^{-3}$) energy compared with the FCI against distance between atoms R (Bohr) for a chain of 12 hydrogen atoms with a STO-6G basis. The NPE is 15.8 kcal/mol.

which even assuming that the fraction of states needed for an accurate potential curve decreases with the size of the space still appears to be much too small a sample. We note that the DMRG results used localized orbitals while the MCCI potential curve used HF MOs. We can see in Fig. 21 that the MCCI curve not only improves upon that of Hartree-Fock in this basis and retains the correct equilibrium bond length but also still displays incorrect behaviour at large R. This suggests that it does not sufficiently account for the static correlation for this system at the value of c_{min} used (5×10^{-3}). We note that the MCCI and Hartree-Fock curve for $R \lesssim 1.8$ Bohr appear to have moderately good agreement with the DMRG results which may be expected as the correlation is essentially dynamic here.

As R becomes larger we would expect the system to be closer to a collection of non-interacting hydrogen atoms so the atomic orbitals might be expected to be a more efficient basis than the restricted Hartree-Fock molecular orbitals. We create orthogonal atomic orbitals using the Gram-Schmidt procedure starting with the leftmost atomic orbital. The use of CSFs allow us to put one electron in each atomic orbital and still have a correct $S = 0$ spin function. At $R = 4.2$ Bohr, we find that we have an energy which is 0.4 Hartree above the DMRG result, but the MCCI program does not seem able to easily improve upon this single CSF as no other configurations are then found. Such an approach would not work well for slightly smaller R where both dynamic and static correlation is important—for $R = 3.2$ Bohr this approach yields a difference of 1.2 Hartree—and suggests that a larger fraction of the configuration space would need to be explored and that approximate natural orbitals could be useful.

H. Ethylene torsional angle

We now look at ethylene for which we use $R_{C-C} = 1.325$ Å, $R_{C-H} = 1.090$ Å, and $\angle HCC = 120.252°$. Using a cc-pVDZ basis with $c_{min} = 10^{-3}$ and vary the torsional angle. We find the torsional barrier in ethylene to be 75.52 kcal/mol which compares relatively favourably with the
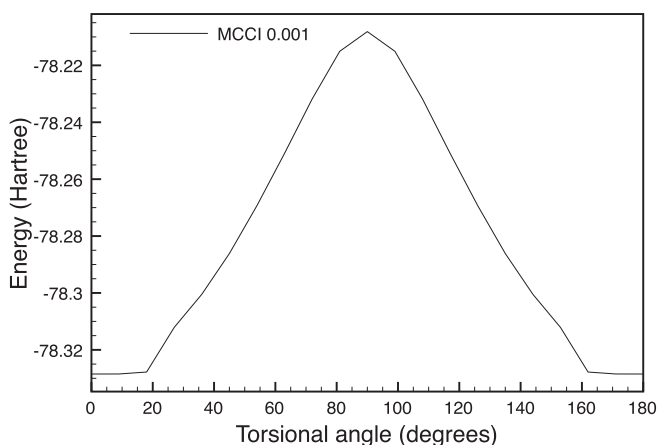
FIG. 22. Energy against torsional angle (degrees) for ethylene with a cc-pVDZ basis and $c_{\min} = 10^{-3}$.

TABLE III. Table showing the mean CSFs to FCI SDs (without symmetry considerations) ratio, mean CSFs to FCI CSFs (without symmetry considerations), non parallelity error (NPE) (kcal/mol) and mean single-point error (MSPE) (kcal/mol), at the smallest cut-off used, for the systems investigated for which FCI results were available.

System	% FCI SD space	% FCI CSF space	NPE	MSPE
BH	0.007%	0.021%	2.6	2.1
HF	0.014%	0.056%	1.3	2.6
CH_4	0.012%	0.049%	0.6	4.1
C_2	0.003%	0.013%	4.9	6.0
F_2	$8.3 \times 10^{-7}\%$	$4.9 \times 10^{-6}\%$	6.2	9.0
N_2	$6.6 \times 10^{-5}\%$	$3.2 \times 10^{-4}\%$	6.6	11.9
H_{12}	0.32%	1.2%	15.8	6.0
BeH_2	0.016 %	0.052%	0.63	0.94
NH_3	$4 \times 10^{-4}\%$	$1.7 \times 10^{-3}\%$	2.4	8.4

cc-pVDZ results from Ref. 38. There the barrier was found to be 68.2 kcal/mol when using CASSCF(12,12) while CASPT2 lowered this to 65.5 kcal/mol. The Hartree-Fock barrier was as high as 111.8 kcal/mol. We note that it is not clear whether exactly the same geometry was used and also whether it was kept fixed, as in our results, or made more realistic by being allowed to relax at 90° in the quoted literature values (i.e., fully geometry optimizing the transition state structure).

The number of CSFs increased from around 5900 to about 11 200 as the torsional angle increased to 90°. The mean number of CSFs was 8250 which is $8.3 \times 10^{-12}\%$ of the size of the 10^{17} SD FCI space when neglecting possible spatial symmetries. In Fig. 22, the results have been mirrored about the line $x = 90$. The lack of a cusp at 90° when using multireference perturbation but its occurrence when using single reference methods such as CCSD was noted in Ref. 39 and we see here that a cusp appears to be avoided as the gradient of the MCCI curve decreases as it approaches 90°.

IV. SUMMARY

We have seen that potential energy curves for small systems may be calculated by MCCI to approaching chemical accuracy, when compared with full configuration interaction results, even when the system is known to be multiconfigurational. Table III shows how NPE values of often around a few kcal/mol can be achieved using MCCI to construct a wavefunction that uses only a very small percentage of the size of the FCI space. For all the systems but BH, which has fairly similar NPE and mean single-point energy error, and the chain of 12 hydrogens, which had the largest fraction of states but worst NPE, we achieved a better NPE than the mean single point error demonstrating that MCCI can indeed be used for calculating a potential curve to higher accuracy than its constituent single-point energies. We note that in general for a fixed $c_{\min} > 0$ MCCI would not be expected to be size-consistent nor size-extensive, but the potential curves calculated here suggest, that for a low enough $c_{\min}$, MCCI is sufficiently size-consistent to correctly describe the approach to dissociation of many of the systems considered in this work. Quantifying the size-consistency by comparing the MCCI en-

ergy at the largest bond length and lowest cut-off used with the sum of the MCCI energy of the fragments reveals a difference of less than 10 kcal/mol for most of the systems considered and less than 1 kcal/mol for those undergoing a single hydrogen dissociation; however, N_2 in a cc-pVTZ basis had a difference of 41.8 kcal/mol.

We note that as the space became larger the fraction used tended to decrease, for example, the results for the nitrogen and fluorine molecules in Table III, which we would hope may be a general feature. We acknowledge that the NPE was larger for the nitrogen and fluorine molecules when the space became larger but comparing, e.g., nitrogen with BH we see that the NPE is only around 2.5 times larger but the fraction of the space is around 100 times smaller while the reduction in the space is even more pronounced for fluorine and the increase in NPE a little smaller. We note that the NH_3 FCI has more than six times the states of BH but MCCI uses a fraction around 100 times smaller to give a slightly lower NPE. We also saw that nitrogen in a cc-pVDZ basis used 2854 on average of a possible 10^9 states while in a cc-pVTZ basis without frozen orbitals we could produce a well-behaved curve using just 10 600 states, at the smallest cut-off considered, of the much larger 10^{17} FCI space, although a NPE value was not available in the latter case. We acknowledge that a fixed value of $c_{\min}$ restricts the size of the MCCI space so that the maximum fraction possible must decrease as the FCI space enlarges. However, our MCCI results are well within this limit and we note that the number of MCCI states appears to increase much more slowly than the FCI space.

However, even with the possibility of a decrease in the fraction of the FCI space required as the system became very much larger, we found that a chain of 50 hydrogens was to prove too challenging for MCCI. Here we saw that although the equilibrium geometry was correct this was not an improvement on the Hartree-Fock results and for longer bond lengths, where static correlation would be considered important, the curve, although lower in energy than that of Hartree-Fock, increased too quickly. We suggest that this is due to the very large size of the configuration space ($\sim 10^{29}$) and the very challenging system for standard quantum chemistry methods partly due to its strongly correlated nature: restricted Hartree-Fock is a very poor approximation at large bond lengths here

and H_{12} required the largest fraction of the FCI space and gave the largest NPE. We note that as a single CSF using orthogonalised atomic orbitals was much closer to the DMRG energy here then the results could perhaps be improved by the use of approximate natural orbitals and allowing the algorithm to explore more of this space but this strongly correlated system may still require many states and seems to be more amenable to modelling using other methods. Finally, the potential curve for the isomerisation of ethylene was found to be relatively smooth and the calculated barrier compared fairly favourably with other computational results. Although we calculated potential surfaces, for which FCI results exist, to relatively high accuracy using just a very small fraction of the FCI space, we have noted two possible limitations of MCCI when applied to potential energy surfaces. The first is that when the cut-off is not sufficiently small a potential curve may not be smooth due to the stochastic nature of the algorithm. However, we were able to produce smooth curves with reasonable cut-offs that resulted in only a very small fraction of the FCI space being used. The second limitation is when many states of a large FCI space are required due to a combination of the reference being very qualitatively wrong and the system being considered strongly correlated, e.g., in the case of restricted Hartree-Fock used with N_2 at extremely large bond lengths or H_{50} at moderately large bond lengths, then not enough states at a feasible cut-off can be included to compensate for the poor choice of molecular orbitals. For N_2 it seemed that this could be remedied to some degree by considering a smaller bond length where the system had almost converged to the energy at dissociation. However, this was not possible for H_{50} so we suggested that different orbitals, perhaps approximate natural orbitals, might be used to try and improve the description at bond lengths away from the equilibrium geometry. To have a better chance of working well with extremely large FCI spaces we also suggest that MCCI could perhaps be used with a relatively large cut-off to find a good starting point for multireference perturbation theory.

ACKNOWLEDGMENTS

[1] J. P. Coe and M. J. Paterson, "Novel truncated and stochastic approaches to configuration interaction," in *Recent Research Developments in Chemical Physics* (Transworld Research Network, Kerala, India, 2012), Vol. 6, pp. 41–65.

[2] L. Bytautas and K. Ruedenberg, Chem. Phys. **356**, 64 (2009).

[3] M. C. Troparevsky and A. Franceschetti, J. Phys.: Condens. Matter **20**, 055211 (2008).

[4] K. Boguslawski, K. H. Marti, and M. Reiher, J. Chem. Phys. **134**, 224101 (2011).

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

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

[7] L. Tong, M. Nolan, T. Cheng, and J. C. Greer, Comput. Phys. Commun. **131**, 142 (2000).

[8] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, 2009); T. D. Crawford and H. F. Schaefer III, "An introduction to coupled cluster theory for computational chemists," in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (Wiley-VCH, New York, 2000), Vol. 14, pp. 33–136.

[9] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[10] P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, J. Chem. Phys. **74**, 2384 (1981).

[11] J. C. Greer, J. Chem. Phys. **103**, 7996 (1995).

[12] J. A. Larsson, L. Tong, T. Cheng, M. Nolan, and J. C. Greer, J. Chem. Phys. **114**, 15 (2001).

[13] W. Győrffy, R. J. Bartlett, and J. C. Greer, J. Chem. Phys. **129**, 064103 (2008).

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

[15] J. Paldus, J. Chem. Phys. **46**, 2769 (1967).

[16] F. E. Harris, J. Chem. Phys. **61**, 5321 (1974).

[17] H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, 2010, see http://www.molpro.net.

[18] X. Li and J. Paldus, J. Chem. Phys. **103**, 1024 (1995).

[19] A. Dutta and C. D. Sherrill, J. Chem. Phys. **118**, 1610 (2003).

[20] M. L. Abrams and C. D. Sherrill, J. Phys. Chem. A **107**, 5611 (2003).

[21] J. S. Sears and C. D. Sherrill, Mol. Phys. **103**, 803 (2005).

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

[23] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).

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

[25] L. Bytautas, T. Nagata, M. S. Gordon, and K. Ruedenberg, J. Chem. Phys. **127**, 164317 (2007).

[26] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, Mol. Phys. **105**, 1335 (2007).

[27] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen, J. Chem. Phys. **113**, 6677 (2000).

[28] S. R. Gwaltney, E. F. C. Byrd, T. V. Voorhis, and M. Head-Gordon, Chem. Phys. Lett. **353**, 359 (2002).

[29] G. K.-L. Chan, M. Kallay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).

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

[31] S. Das, D. Mukherjee, and M. Kállay, J. Chem. Phys. **132**, 074103 (2010).

[32] E. Rossi, G. L. Bendazzoli, S. Evangelisti, and D. Maynau, Chem. Phys. Lett. **310**, 530 (1999).

[33] G. D. Purvis III, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).

[34] F. A. Evangelista and J. Gauss, J. Chem. Phys. **134**, 114102 (2011).

[35] P. J. Knowles and N. C. Handy, Chem. Phys. Lett. **111**, 315 (1984).

[36] P. J. Knowles and N. C. Handy, Comput. Phys. Commun. **54**, 75 (1989).

[37] J. Hachmann, W. Cardoen, and G. K.-L. Chan, J. Chem. Phys. **125**, 144101 (2006).

[38] M. Piris, X. Lopez, F. Ruipérez, J. M. Matxain, and J. M. Ugalde, J. Chem. Phys. **134**, 164102 (2011).

[39] R. K. Chaudhuri, K. F. Freed, S. Chattopadhyay, and U. S. Mahapatra, J. Chem. Phys. **128**, 144304 (2008).