AIP The Journal of Chemical Physics

Benchmark studies of variational, unitary and extended coupled cluster methods

Bridgette Cooper and Peter J. Knowles

Citation: The Journal of Chemical Physics **133**, 234102 (2010); doi: 10.1063/1.3520564
View online: http://dx.doi.org/10.1063/1.3520564
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/133/23?ver=pdfcov
Published by the AIP Publishing

THE JOURNAL OF CHEMICAL PHYSICS **133**, 234102 (2010)

Benchmark studies of variational, unitary and extended coupled cluster methods

Bridgette Cooper and Peter J. Knowles[a)]
School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 6 August 2010; accepted 5 November 2010; published online 15 December 2010)

Comparative benchmark calculations are presented for coupled cluster theory in its standard formulation, as well as variational, extended, and unitary coupled cluster methods. The systems studied include HF, N_2, and CN, and with cluster operators that for the first time include up to quadruple excitations. In cases where static correlation effects are weak, the differences between the predictions of molecular properties from each theory are negligible. When, however, static correlation is strong, it is demonstrated that variational coupled cluster theory can be significantly more robust than the traditional ansatz and offers a starting point on which to base single-determinant reference methods that can be used beyond the normal domain of applicability. These conclusions hold at all levels of truncation of the cluster operator, with the variational approach showing significantly smaller errors. © *2010 American Institute of Physics*. [doi:10.1063/1.3520564]

I. INTRODUCTION

The coupled cluster (CC) method is now firmly established as the key approach for computing correlated molecular electronic structure, so long as the underlying reference wavefunction is a reasonable zero-order approximation. Usually, this is the case; for most molecules in most situations, single-determinant Hartree–Fock suffices as the reference. However, where there is strong static correlation, for example in dissociating molecules or excited states, Hartree–Fock is inadequate, and often one would like to proceed using multideterminant reference wavefunctions. Unfortunately, despite significant progress, an unambiguous problem-free multireference coupled cluster theory is not yet available, and one is often forced to use approximate methods that correct the lack of extensivity in the variational multireference configuration interaction approach. Often, one can alternatively proceed using spin-unrestricted Hartree–Fock (UHF) as a reference that includes some static correlation, but it is well known that there are then problems associated with spin contamination artefacts.

It is the purpose of this paper to examine through benchmark calculations the performance of single-reference coupled cluster theory in a few examples where static correlation is strong. Although it is already well known that the standard formulation of single-reference coupled cluster fails to describe situations where, for example, multiple bonds are broken, we investigate also the potential of other formulations, such as variational CC (VCC),[1] unitary CC (UCC),[2–5] and extended CC (ECC).[6] These variants have previously been asserted to be more accurate, although they are in general much more expensive, and the numerical results presented to date give some evidence to support this assertion at the double-excitation level of theory. We extend here the body of numerical evidence by considering connected triple and quadruple excitations, and by considering other measures of incompleteness such as the artificial derivative discontinuities in potential energy surfaces that are introduced in spin-unrestricted theories.

II. METHODS

The coupled cluster wavefunction can be expressed as

$$|\Psi\rangle = \exp(\hat{T})|0\rangle, \tag{1}$$

where $|0\rangle$ is a suitable reference wavefunction, for example the spin-restricted Hartree–Fock (RHF) wavefunction, and $\hat{T}$ is the cluster operator given by

$$\hat{T} = \sum_{I}^{N} \hat{T}_I, \tag{2}$$

$$\hat{T}_1 = \sum_{ia} t_a^i \, a^\dagger \, i, \tag{3}$$

$$\hat{T}_2 = \frac{1}{4} \sum_{ijab} t_{ab}^{ij} \, a^\dagger \, b^\dagger \, j \, i, \tag{4}$$

$$\hat{T}_3 = \frac{1}{3!^2} \sum_{ijkabc} t_{abc}^{ijk} \, a^\dagger \, b^\dagger \, c^\dagger \, k \, j \, i, \tag{5}$$

$$\hat{T}_4 = \frac{1}{4!^2} \sum_{ijklabcd} t_{abcd}^{ijkl} \, a^\dagger \, b^\dagger \, c^\dagger \, d^\dagger \, l \, k \, j \, i. \tag{6}$$

$i, j, k, \ldots$ denote spin-orbitals occupied in the reference function, and $a, b, c, \ldots$ are the virtual orthogonal complement. The operators $a^\dagger \, b^\dagger \, c^\dagger \, \cdots$ and $i, j, k, \ldots$ are the corresponding second-quantized creation and annihilation operators. t_a^i, t_{ab}^{ij}, t_{abc}^{ijk}, and t_{abcd}^{ijkl} are the cluster amplitudes associated with the expansion coefficients.

The exponential form of this trial wavefunction ensures that it is multiplicatively separable for noninteracting subsystems, provided that the reference wavefunction is also separable. If the cluster operator $\hat{T}$ is complete, that is it contains all the possible excitations, then Ψ is capable of coinciding with

---

[a)] Author to whom correspondence should be addressed. Electronic mail: KnowlesPJ@Cardiff.ac.uk.

0021-9606/2010/133(23)/234102/9/$30.00

the exact wavefunction. In practice, $\hat{T}$ is usually truncated to a given excitation level, for example to single and double excitations only, for reasons of computational cost and complexity. We adopt the notation where acronyms such as VCC are appended with one or more of S, D, T and Q to indicate the presence of single, double, triple or quadruple excitation operators.

This trial wavefunction can be used to create expressions for the energy in a variety of ways as discussed below. The way that the different energy expressions are obtained can lead to each of the coupled cluster methods having different desirable properties.[7]

A. Traditional coupled cluster

The most widely used means of obtaining the correlation energy, $E_{\mathrm{corr}} = E_0 - E$, where $E_0 = \langle 0|\hat{H}|0\rangle$, and the coupled cluster wavefunction, is to substitute the trial exponential wavefunction into Schrödinger's equation and project on the left with the reference wavefunction, and with the manifold of excited determinants obtained by acting with the individual cluster operators on the reference.[8] Additionally, the Schrödinger equation is normally premultiplied by $\exp(-\hat{T})$ before projection, introducing a similarity transformed Hamiltonian, $\exp(-\hat{T})\hat{H}\exp(\hat{T})$, and decoupling the energy expression from the set of equations for determining the cluster amplitudes:

$$\langle 0|\exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle = E, \tag{7}$$

$$\langle 0|ijk\cdots c^\dagger b^\dagger a^\dagger \exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle = 0. \tag{8}$$

The similarity transformation can be evaluated in closed form through the Baker–Campbell–Hausdorff commutator expansion, and terminates at fourth order in $\hat{T}$. From this point onwards this method will be refered to as traditional coupled cluster (TCC),[9] and the term coupled cluster will be used to refer to the whole family of methods that use Eq. (1) to generate the wavefunction.

The principal objection to TCC is that it is not variational, that is it does not provide an upper bound to the exact energy. We will see below that in some examples where electron correlation is strong, for example when multiple chemical bonds are broken, TCCSD can give energies that are significantly below those from full configuration interaction (FCI). It then becomes desirable to examine alternatives to projection that have the potential to be more robust.

B. Variational coupled cluster

An alternative ansatz is to calculate the energy as an expectation value:

$$E = \min_T \frac{\langle 0|\exp(\hat{T}^\dagger)\hat{H}\exp(\hat{T})|0\rangle}{\langle 0|\exp(\hat{T}^\dagger)\exp(\hat{T})|0\rangle} \tag{9}$$

$$\equiv \min_T \langle 0|\exp(\hat{T}^\dagger)\hat{H}\exp(\hat{T})|0\rangle_{\mathrm{linked}}. \tag{10}$$

This method, termed here as VCC, provides an upper bound to the exact energy. A related approach is the expectation value coupled cluster (XCC) (Ref. 1) ansatz, in which the amplitudes are derived by projection rather than variationally. VCC has not been widely implemented because factorial scaling of the computational cost makes it impracticable except for small systems. Unlike TCC there is no convenient truncation of terms in the energy expression, because of the presence of the $\exp(\hat{T}^\dagger)$ to the left of the Hamiltonian operator. This means that the energy expression does not terminate at fourth order as in TCC; for the first, quotient, form of Eq. (9) one would have to include all terms up to $\hat{T}^N$ where N is the number of electrons, while the many-body expansion in just linked diagrams (10) does not terminate at any finite power of $\hat{T}$. Van Voorhis and Head-Gordon[10] reported a "brute-force" implementation of VCCD in the quotient formulation and applied it to several situations of strong correlation in small molecules. The present paper is similar in spirit, but investigates cluster operators including for the first time up to connected quadruple excitations.

There are several choices for truncating the number of terms in the above expectation value coupled cluster functional to make approximate VCC methods. Bartlett and Noga[1] have truncated the exponential to make a series of methods that are correct to a given perturbation order. This truncation is done so that each approximation satisfies the General Hellmann Feynmann theorem to aid the calculation of molecular properties. These truncated energy functionals are still extensive, but have lost the property of being an upper bound to the correct energy.[1]

C. Unitary coupled cluster

The UCC method uses an anti-Hermitian operator in the exponential[2–5] to generate a unitary transformation:

$$|\Psi\rangle = \exp(\hat{\sigma})|0\rangle, \tag{11}$$

$$\hat{\sigma} = \hat{T} - \hat{T}^\dagger, \tag{12}$$

$$E = \min_T \langle 0|\exp(-(\hat{T} - \hat{T}^\dagger))\hat{H}\exp(\hat{T} - \hat{T}^\dagger)|0\rangle, \tag{13}$$

where the denominator in the variational quotient is unity, since the wavefunction is normalized by construction. This method has the disadvantage of a nonterminating Taylor series expansion as the unitary operator contains both an excitation operator $\hat{T}$ and a deexcitation operator $\hat{T}^\dagger$, but the series typically converges rapidly.[3] If the contributions to the UCC energy are grouped in terms of the powers of $\hat{T}$ then it can be shown that the zero, first, and second order terms are similar to the variational coupled cluster. Beyond this level, the UCC becomes more complicated in comparison to the VCC energy expressions.[9]

A means of creating a truncation hierarchy is to require that the energy is correct through to some order of perturbation theory.[1,11,12] Hoffmann and Simons[13] have developed a UCC method by truncating the Hausdorff expansion of the unitary coupled cluster energy functional to second order in the t amplitudes:

$$\left(H + [H,\sigma] + \tfrac{1}{2}[[H,\sigma],\sigma]\right)|0\rangle \approx E|0\rangle. \tag{14}$$

Pal[14] has pointed out the advantage gained in property calculations through using a unitary ansatz.

D. Extended coupled cluster

The energy expression for the ECC method[6] again has its starting point in the VCC energy Eq. (9):

$$E = \frac{\langle 0|\exp(\hat{T}^\dagger)\hat{H}\exp(\hat{T})|0\rangle}{\langle 0|\exp(\hat{T}^\dagger)\exp(\hat{T})|0\rangle} \tag{15}$$

$$= \langle\omega|\exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle, \tag{16}$$

where the bra state is given by

$$\langle\omega| = \frac{\langle 0|\exp(\hat{T}^\dagger)\exp(\hat{T})}{\langle 0|\exp(\hat{T}^\dagger)\exp(\hat{T})|0\rangle} \tag{17}$$

$$= \langle 0|\exp(\hat{\Sigma}^\dagger), \tag{18}$$

where $\hat{\Sigma}^\dagger$ is an effective deexcitation operator. This leads to the following equation for the energy:

$$E = \langle 0|\exp(\hat{\Sigma}^\dagger)\exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle, \tag{19}$$

where $\hat{\Sigma}$ and $\hat{T}$ are varied independently to make E stationary, although in fact the optimal Σ is uniquely defined by $\hat{T}$, and can be obtained as the solution of linear equations if $\hat{T}$ is known.[15] Piecuch and Bartlett have presented an alternative ansatz for obtaining the cluster operators[16] by projecting the double-similarity-transformed Hamiltonian against the manifold of singly and doubly excited determinants. Pal and co-workers.[17–21] have analyzed the computation of response properties within the ECC formalism.

Extensive analysis, and benchmark calculations of the performance, of ECCSD have been presented by Piecuch and co-workers.[22–25] Some of these calculations are precursors of the computations that we present in Sec. III, for example the stretching of the N_2 molecule using TCCSD, ECCSD, and these methods with perturbative corrections for the effects of connected triple and quadruple excitations.

An important related approach in which the similarity-transformed Hamiltonian is projected against an alternative space to the set of determinants that interact with $|0\rangle$ through $\hat{T}$ is the method of moments and related methods.[26–30] Although we do not consider these methods explicitly in the calculations described below, their numerical performance at the single and double excitation level is already well documented.[26–30]

E. Quadratic coupled cluster

The ECC method gives rise to a hierarchy of methods can be derived by truncating $\exp(\hat{\Sigma}^\dagger)$ to different powers of $\hat{\Sigma}^\dagger$. Thus, the TCC energy expression can be written in the form

$$E = \langle 0|(1+\hat{\Sigma}^\dagger)\exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle, \tag{20}$$

and variation with respect to $\hat{\Sigma}$ yields the projected Schrödinger equation governing $\hat{T}$. The next higher order of $\exp(\hat{\Sigma}^\dagger)$ to a quadratic expression gives the quadratic coupled cluster (QCC) energy expression:[31]

$$E = \langle 0|\big(1+\hat{\Sigma}^\dagger+\tfrac{1}{2}\hat{\Sigma}^{\dagger 2}\big)\exp(-\hat{T})\hat{H}\exp(\hat{T})|0\rangle. \tag{21}$$

This method leads to coupled amplitudes equations, that is the cluster amplitudes for both $\hat{\Sigma}^\dagger$ and $\hat{T}$ must be solved together. QCCD was implemented in Ref. 31, and later extended to include single excitation operators under the acronym QECCSD.[22–25]

F. Implementation for benchmark calculations

For the purpose of benchmarking each of these coupled cluster variants, the methods were implemented within the FCI program in the MOLPRO package.[32] Although previous implementations of ECCSD have been through the use of diagrammatic methods to generate an efficient code,[22,23] such an approach is not an option for VCC, where the diagram series is in principle infinite, and in practice will terminate to numerical precision only at high order. A diagrammatic approach would also be difficult for the higher excitation levels that we consider in this work. For these reasons, we adopt a general brute-force approach. The FCI method uses a complete set of Slater determinants that increases in dimension factorially with the number of electrons. Thus, benchmarks are possible only for relatively small systems and small orbital basis sets. However, because there is an explicit representation of the complete space from which wavefunctions are constructed, arbitrary quantities can be constructed and manipulated.

In standard FCI calculations, the ground-state wavefunction is found through an iterative approach to the lowest eigenvector of the Hamiltonian matrix using the Davidson algorithm,[33] based on successive applications of the Hamiltonian matrix on a Krylov sequence of trial vectors,

$$\mathbf{v}_i = \mathbf{H}\,\mathbf{c}_i. \tag{22}$$

$\mathbf{H}\,\mathbf{c}$ is evaluated using resolution-of-the-identity methods[34,35]

$$(\mathbf{H}\,\mathbf{c})_I = \frac{1}{2}\sum_{KJ}\sum_{ijkl}(ij|kl)\langle I|i^\dagger j|K\rangle\,\langle K|k^\dagger l|J\rangle c_J. \tag{23}$$

For the purpose of benchmarking coupled cluster theories, we evaluate the action of the cluster operators on wavefunctions in the same way as the Hamiltonian:

$$\hat{T}_2\left(\sum_J c_J|J\rangle\right) = \sum_{IJK}\sum_{ijab}|I\rangle\,\langle I|a^\dagger i|K\rangle\,\langle K|b^\dagger j|J\rangle\,c_J\,t^{ij}_{ab}. \tag{24}$$

The other required quantity is the projection of the residual vector on the manifold of excited determinants, which has the form of a transition density matrix:

$$V^{ij}_{ab} = \langle 0|i^\dagger j^\dagger ba\,\hat{H}\sum_I|I\rangle c_I. \tag{25}$$

With these basic building blocks, each of the coupled cluster variants (TCC, VCC, UCC, and ECC) can be constructed by multiple applications of the cluster operators. Theories that are in principle infinite series can in practice be summed without difficulty to machine precision with a finite number of terms because of the inverse factorial factor in the expansion of the exponential. The implementation supports closed- and open-shell reference wavefunctions, including different orbital sets for α and β spins, and cluster operators up to and including $\hat{T}_4$. Cluster operators and residual vectors are represented as instances of a single object (implemented as a FORTRAN-90 module), with associated methods that perform the operations described above. The implementation details, and even the excitation level of the operator, are hidden inside the object, and not exposed to the higher-level code implementing a particular method, allowing that code to remain uncluttered. The approach is much less elegant than other general code generators,[36] but has the advantage that it is highly modular, allowing each new theory to be simulated in just a few lines of high-level code. For example, in the case of ECC (which includes QCC as a special case), 30 FORTRAN statements are used to express the algorithm for the energy and residual.

We note that this implementation is highly effective in the safe prototyping of new methods, but is not at all designed to be computationally efficient. Except for ansätze such as TCC where the series terminate at finite order, the resource scaling (arithmetical operations and storage) is necessarily factorial in the number of electrons since repeated application of the cluster operator eventually visits the entire FCI space.

III. RESULTS AND DISCUSSION

It is intuitive, and explicitly demonstrated in earlier calculations,[10] that when static electron correlation effects are not strong, VCC, and the other variants, are expected to give essentially the same results as TCC, at a given level of truncation of the cluster operator. We therefore focus on a number of examples of small molecules where static correlation is strong, in order to address the question of whether there are differences between the performance of the methods when, for example, covalent bonds are broken.

A. Potential energy curve for hydrogen fluoride

As a first example of strong static correlation when a chemical bond is broken, we examine hydrogen fluoride. Calculations performed on HF with TCC including up to octuple excitations have shown that as the bond length is increased higher orders of excitation are need to accurately describe the potential energy curve.[36,37] Previous calculations on hydrogen fluoride comparing traditional, variational, and quadratic coupled cluster methods with only double excitations have shown that there is little difference between these methods.[31]

Calculations were performed using a cc-pVDZ basis,[38] and with all ten electrons correlated. Table I gives the results of these calculations with the FCI energy given as differences from the asymptotic energy, and the CC results given as differences from the FCI values. Figure 1 shows the potential energy curve of HF for each of the coupled cluster methods in the range 1−3 Å, as well as differences from FCI values over this range.

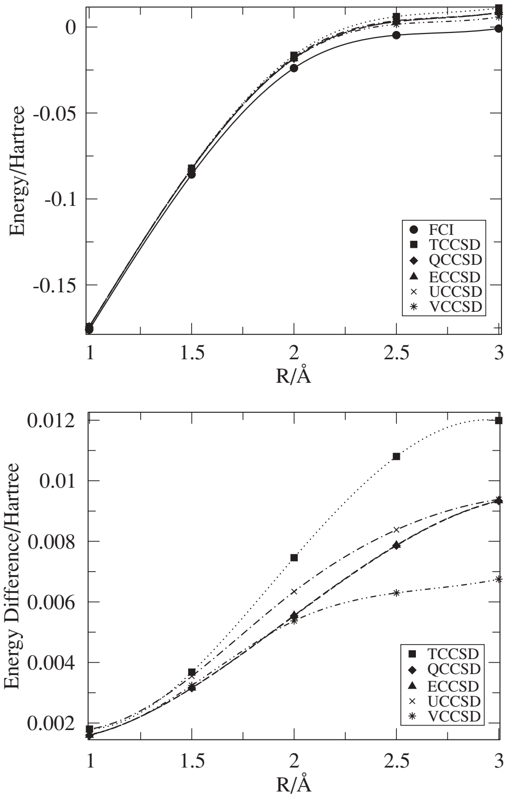

FIG. 1. Potential energy curve for HF, with differences from FCI below.

TABLE I. The potential energy curve of HF, FCI energies in hartrees, and coupled cluster methods as differences from FCI values in millihartree (mhartree).

R/Å	FCI	TCCSD	VCCSD	UCCSD	ECCSD	QCCSD
3.0	– 0.000852	11.990	6.749	9.385	9.343	9.333
2.5	– 0.004682	10.802	6.296	8.381	7.872	7.859
2.0	– 0.023859	7.459	5.373	6.341	5.552	5.536
1.5	– 0.085758	3.680	3.240	3.547	3.157	3.154
1.0	– 0.176039	1.800	1.621	1.797	1.596	1.595

The results show that each of the CCSD methods perform comparably with VCCSD for single bond breaking. At short distances, there is little difference between the coupled cluster methods, all of which have small errors from FCI of 1–3 mhartree. At the longer bond lengths, differences start to be seen between the different coupled cluster methods, where TCCSD performs least well, with errors ten times larger than at the equilibrium bond length. In contrast, VCCSD performs significantly better with an error of 7 mhartree. Unitary, extended, and quadratic coupled cluster perform slightly worse than VCCSD at long bond lengths, but somewhat better than TCCSD.

B. Symmetric stretching of water

The simultaneous stretching of both OH bonds in water has been extensively studied,[10,31,37,39–41] and provides a more challenging case than single bond breaking. Previously[10,31] benchmarks have been performed on H_2O with variational and quadratic coupled cluster methods with only double excitations in the cluster operator. These results showed much better agreement with FCI results and both were a considerable improvement over TCCD results. Here these are extended upon to include both single and double excitations

TABLE II. The potential energy curve of H_2O, FCI energies in hartrees, and coupled cluster methods as differences from FCI values in mhartrees.

R/Å	FCI	TCCSD	VCCSD	UCCSD	ECCSD	QCCSD
0.967	– 0.351813	1.646	1.405	1.439	1.401	1.401
1.450	– 0.232412	5.852	4.918	5.094	4.915	4.914
1.933	– 0.124534	9.222	9.914	10.373	10.025	10.033

and also compared with unitary and extended coupled cluster methods.

Energy calculations were performed at three points along the symmetric stretching mode of a water molecule. The points were at the equilibrium bond length r_e, $1.5r_e$, and $2r_e$. The calculations were performed in the 6–21G basis set, with all ten electrons correlated, and single and double excitations in the cluster operator.

The results of these calculations are shown in Table II, where the FCI values are given relative to the asymptotic energy of an oxygen atom and two hydrogen atoms, and the coupled cluster values are differences from the FCI values. The results show that all the CCSD methods perform well at the equilibrium bond length with errors between 1.4 and 1.7 mHartrees in comparison with the FCI values. At this bond length, TCCSD has larger errors than variational, unitary, extended, and quadratic coupled cluster. As the bonds are stretched, all CCSD methods show a strongly increasing error, reflecting the fact that the static correlation is increasing, and therefore it becomes harder for approximate correlation methods to account for all the static correlation present. Interestingly, TCCSD has slightly smaller errors in comparison with VCCSD at $2r_e$. This may be a reflection of the fact that TCCSD is not variationally constrained.

C. Potential energy curve of N_2

The potential energy curve for the dissociation of N_2 is frequently studied,[23,27,42–55] as it involves the breaking of a triple bond. This is an extremely challenging problem for approximate methods for calculating correlation energies as it is an example where there is strong static correlation at stretched bond lengths.

Previous benchmarks with TCCSD using a RHF reference wavefunction have shown that this method does not accurate describe the breaking of the triple bond.[51] TCCSD calculations at stretched bond lengths lie below the FCI values and the magnitude of the error can be more than 100 mhartree.[56] Multireference configuration interaction and coupled cluster methods perform much better across the potential energy curve in comparison with their single reference counterparts and give more accurate equilibrium properties such as vibrational frequencies for N_2.[45,47] There is strong evidence, however, that single-reference methods can be used. The VCCD calculation in the STO-3G basis set[10] demonstrated that the constraint of the variational principle prevents the catastrophe suffered by TCCD at long bond lengths, but that at short bond lengths VCCD and TCCD are similar; the consequence is that the VCCD potential is close to FCI. Further calculations have shown that QCCD produces similar results to VCCD;[31] that QCCSD is a reasonable approximation to ECCSD, which itself reasonably approximates FCI.[22,23] These conclusions hold in a larger DZ basis set, where QCCSD is demonstrated to be a good approximation to FCI.[22]

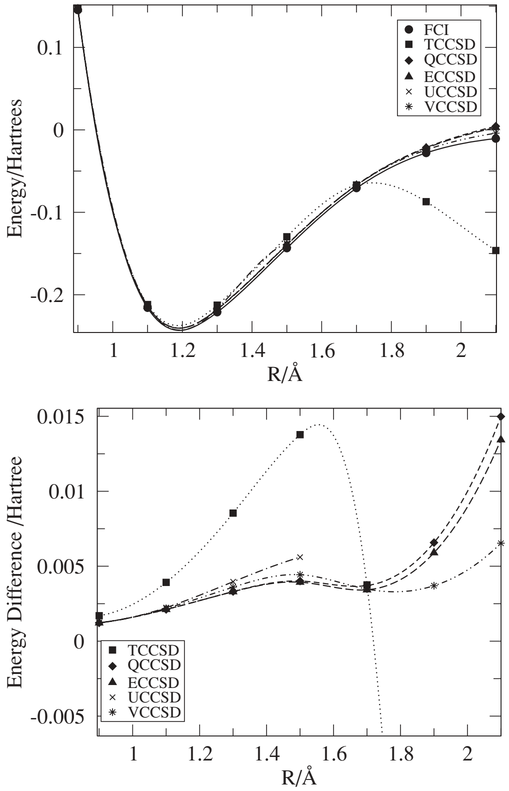

FIG. 2. Potential energy curve of N_2 in hartrees, with single and double excitations in the cluster operator, with differences from FCI below.

We have extended these earlier investigations by further STO-3G, ten-active electron computations. The potential energy curves obtained when single and double excitations were included in the cluster operator are shown in Fig. 2, and errors from FCI values shown in Table III; for TCCSD, ECCSD, and QCCSD, these results are already known,[22,23] but are presented here for completeness. The VCCSD results are similar to those of the earlier VCCD (Ref. 10) computations. Figure 3 shows the potential energy curve with quadruple excitations in the cluster operator, as well as VCCSD and VCCSDT for

TABLE III. The potential energy curve of N_2, FCI energies in hartree, and coupled cluster methods as differences from FCI values in millihartree.

R/Å	FCI	TCCSD	VCCSD	UCCSD	ECCSD	QCCSD
0.9	0.145308	1.70	1.23	1.24	1.22	1.22
1.1	– 0.215807	3.92	2.19	2.24	2.11	2.12
1.3	– 0.221133	8.54	3.60	3.96	3.30	3.31
1.5	– 0.143463	13.78	4.43	5.60	3.93	4.01
1.7	– 0.070611	3.75	3.47		3.43	3.71
1.9	– 0.028038	– 59.09	3.70		5.90	6.58
2.1	– 0.010584	– 135.75	6.53		13.43	14.99

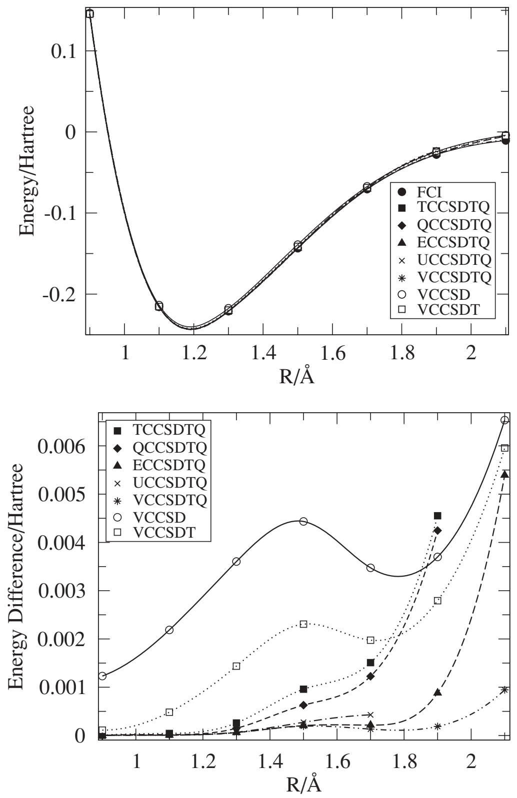

FIG. 3. Potential energy curve of N_2 in hartrees, with up to quadruple excitations in the cluster operator, with differences from FCI below.

comparison. Table IV shows the errors in comparison with FCI for traditional and variational coupled cluster methods with increasing level of excitation included in the cluster operator up to quadruple excitations. Full treatment of triples and quadruples is new.

The results show that with single and double excitations included in the cluster operator each of the coupled cluster methods perform well at equilibrium bonds lengths, with errors in the range of $1-4$ mhartree. However, Fig. 2 shows the dramatic failure of TCCSD in breaking the triple bond in nitrogen. The energies beyond 1.7 Å lie below the FCI values, and the TCCSD energy curve shows an energy barrier to forming the nitrogen triple bond. The TCCSD method gives errors at stretched bond lengths of tens and even hundreds of mhartree.

TABLE IV. Differences from FCI for traditional and variational coupled cluster methods with triple and quadruple excitations in the cluster operator for N_2. Energy errors are given in mhartrees.

R/ Å	TCCSD	TCCSDT	TCCSDTQ	VCCSD	VCCSDT	VCCSDTQ
0.9	1.70	0.49	0.01	1.23	0.11	7.5×10^{-4}
1.1	3.92	2.05	0.04	2.19	0.48	8.1×10^{-3}
1.3	8.54	6.15	0.26	3.60	1.44	0.06
1.5	13.78	11.62	0.96	4.43	2.31	0.19
1.7	3.75	3.06	1.51	3.47	1.97	0.13
1.9	− 59.09	− 59.84	4.55	3.70	2.79	0.19
2.1	− 135.72	− 136.42	− 14.01	6.52	5.96	0.95

In contrast the VCCSD curve follows the FCI energy curve more closely, and as expected does not dip below the FCI results. The errors in comparison with FCI values for VCCSD stay in the order of mhartrees, even when the bond is quite stretched at 2.1 Å, the error is only 6.55 mhartree. ECC and QCC also show better agreement with the FCI and give similar curves to the VCCSD method, with errors again a few mhartrees. QCCSD performs considerably better than TCCSD, with the maximum error in the range tested of 14.99 mhartree, and the values for the energies do not go below the FCI values. Unfortunately, our numerical procedures for solving the nonlinear equations for each method have not proved sufficiently robust to achieve convergence beyond 2.1 Å for any of the methods, and UCCSD fails to converge after 1.5 Å.

Figure 3, where up to quadruple excitations are included in the cluster operator for each of the coupled cluster methods, shows that all the methods now give a qualitatively correct potential energy curve, and there are no more fictious barriers to the formation of a triple bond. VCCSDTQ still outperforms TCCSDTQ, as the errors from FCI are still smaller for VCCSDTQ, and the last point for TCCSDTQ still lies below the FCI value although the error is now in the tens of mhartree rather than over 100.

Table IV compares the effect of including higher excitations in the TCC and VCC methods. From these data, the inclusion of triple excitations is clearly not enough to improve the quality of the TCC or VCC energy curves produced. While the errors at equilibrium are reduced by a factor of around 4 in the VCC case, at long bond lengths the errors are comparable with those obtained with the inclusion of up to double excitations only. For VCCSDT the error at 2.1 Å is just 0.56 mhartree smaller than with single and double excitations only. Further inclusion of quadruple excitations, however, leads to a dramatic reduction in the errors for both methods. At equilibrium, errors are on the μHartree scale with quadruple excitations included, however, at long bond lengths the errors are much increased. With the inclusion of quadruple excitations TCCSDTQ performs comparably with VCCSD.

These results shows that for cases with strong static correlation, VCC and related methods are significantly better than TCC, for a given level of excitation in the cluster operator. We have shown that QCCSD is a signficant improvement on TCCSD for cases with strong static correlation and as it has been previously shown that the QCCD method scales as n^7 with the system size,[31] whereas there is a cost scaling of n^{10} for TCCSDTQ (Ref. 57) using a method that is on the way to being more variational may be preferable to including higher levels of excitation when correlation is strong.

We can also compare the VCCSDT and VCCSDTQ results with ECCSD with perturbative TQ corrections.[22] At a bond length of 2.1167 Å, the QCCSD(TQ) and ECCSD(TQ) errors are 3.60, 2.95 mhartree, respectively,[22] significantly more than our VCCSDTQ error of 0.95 mhartree at 2.1 Å. The QCCSD(T) and ECCSD(T) errors are slightly less, and are smaller than with VCCSDT, indicating some overcorrection that partially compensates for omitted higher-order terms.

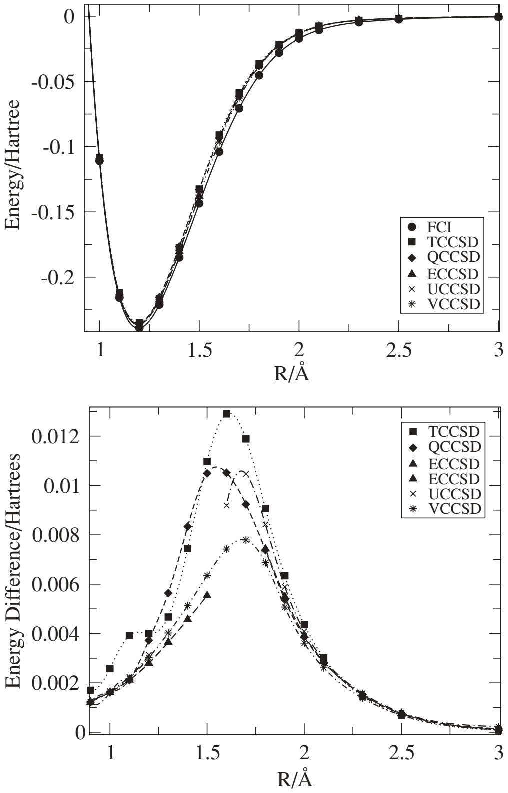

FIG. 4. Potential energy curves for each of the coupled cluster method using an UHF reference function for N_2 and differences from FCI below.

An alternative to using the RHF wavefunction as the reference is to adopt a spin-unrestricted UHF wavefunction. Then, at dissociation, the reference wavefunction can be a product of two ^4S N atom wavefunctions (although not with pure spin symmetry), and it is hoped that the long range correlation will be described better across the whole geometry range when CC theory is used.

Figure 4 illustrates the results of each of the coupled cluster methods for N_2 with an UHF reference wavefunction. Calculations were performed with the minimal STO-3G basis set. At long bond lengths all the coupled cluster methods give good results, with the right dissociation energy, in contrast to where a RHF wavefunction was used.

From these results it is clear that the largest errors are seen at short and intermediate bond lengths between 1 and 2 Å. The smallest errors are seen with VCCSD and ECCSD, with TCCSD performing considerably poorer. There are missing data points in the UCCSD and ECCSD graphs as in this region our numerical procedures failed to converge.

D. Energetic effects of spin contamination in N_2

In the above results, the region after 1 Å is the start of the onset of the divergence of UHF from RHF. Therefore, the large errors may be arising from the coupled cluster wavefunctions still containing spin contamination. Consideration of the variational principle produces the conclusion that whilst the RHF potential curve is smooth in all its derivatives, the UHF curve contains a discontinuity in its second derivative at the point of divergence. Methods based on an UHF reference that are not fully variational in all parameters have additionally a discontinuous first gradient.[58] The exception, of course, is that in the exact (FCI) wavefunction there will be no discontinuities. Therefore the size of the discontinuity in approximate wavefunction methods can give a means of assessing how close the approximate wavefunction is to the exact wavefunction, as well as an idea of the magnitude of the spin contamination remaining in the method.

TABLE V. The size of the discontinuity in the first and second derivatives of the energy for each of the coupled cluster methods, with single and double excitations in the cluster operator.

Method	First derivative	Second derivative
FCI	– 0.0003	0.002
HF	– 0.0004	5.16
TCC	0.021	
VCC	– 0.002	
UCC	0.000	
ECC	– 0.001	
QCC	– 0.015	
UMP2	1.346	
UMP4	– 0.533	

The magnitude of the derivative discontinuities was calculated by first evaluating numerically the derivative of the energy at a number of points between 1.11 and 1.13 Å, a range that includes the location of the knot. The results of the numerical differential were then extrapolated to find the value of the derivative at the location of the root, 1.1217 Å, using an interpolating polynomial. From these results a difference in the extrapolated value was seen from points either side of the root, and this discrepancy was taken as the size of the discontinuity.

Table V shows the size of the discontinuity in the energy derivatives for UHF and each of the coupled cluster methods with single and double excitations. The results were calculated for N_2 in the STO-3G basis set. For comparison the discontinuity was also calculated for Møller–Plesset perturbation theory at second (UMP2) and fourth (UMP4) orders. The FCI value gives an indication of the numerical error in the calculation.

From these results it can be seen that there is no discontinuity in the first derivative of the energy with the UHF method, as is indicated by the value being similar to that obtained by the FCI results. The size of the discontinuity in the second derivative, however, is very large at 5.161 hartree / Å^2.

The results show that there is a considerable amount of spin contamination in the traditional coupled cluster method, which has a value of 0.0207 hartree / Å for the size of the discontinuity in the first derivative of the energy. The variational coupled cluster and extended coupled cluster have a similar level of energy artefact from spin contamination that is almost ten times smaller than in TCC, while UCCSD happens to give a very much smaller value.

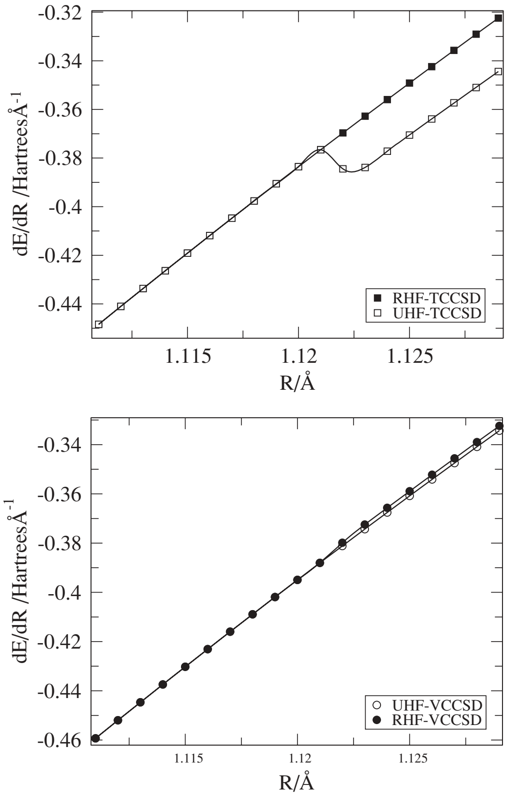

FIG. 5. The values of the derivative of the potential energy function for N_2 as the bond length is changed. Upper panel: TCC; lower panel: VCC.

This large difference between the size of the discontinuity in the energy for TCCSD and VCCSD can be seen in Fig. 5. There is a clear break in the TCCSD energy derivative when the UHF wavefunction is used, and there is a significant lowering of the value of the energy derivative after the onset of the UHF solution. In comparison, the difference is much smaller with the RHF and UHF VCCSD results being very similar over the range of the onset of the UHF solution. Table V also shows that there is much less spin contamination in all the coupled cluster methods in comparison to the Møller–Plesset perturbation methods, UMP2 and UMP4, which are long known to be disastrous methods in this and other strongly spin-contaminated systems.[59,60]

E. Cyano radical

The UHF wavefunction for the cyano radical (CN) is heavily spin contaminated even at the equilibrium geometry. Previously,[59] it has been found that the UMP series for the cyano radical converges slowly, thus UMP2 electron affinities for this species give poor agreement compared to experiment. The slow convergence of the UMP series has been rationalized in terms of spin contamination in the wavefunction.[61–63] It has also been shown that there is small but negligible spin contamination in UHF-CCSD and ROHF-CCSD wavefunctions, even in cases where the spin contamination in the UHF reference is high when considering molecules at the equilibrium structures.[64] It has been further shown that spin contamination can also give very large errors in the vibrational frequencies of diatomic molecules,[65,66] with the errors related to the geometric derivative of $\langle \hat{S}^2 \rangle$. As the wavefunction gets contaminated with higher energy states the potential energy surface will rise too steeply, giving rise to large values of $d\langle S^2 \rangle / dR$, and thus the vibrational energies will be too high. The equilibrium bond length can be too short, and the energy too high in molecules with a large amount of spin contamination, and these properties related directly to the difference $\langle \hat{S}^2 \rangle - S(S+1)$.

TABLE VI. Harmonic vibrational wavenumber, bond lengths and energies for CN^-.

Method	ω_e / cm^{-1}	r$_e$ / Å	E$_{CN^-}$ / Hartree
HF	2559.7	1.1623	− 90.9376 70
FCI	2134.1	1.2117	− 91.0831 79
TCCSD	2189.1	1.2066	− 91.0780 11
VCCSD	2170.7	1.2082	− 91.0793 58
UCCSD	2179.3	1.2075	− 91.0789 79
ECCSD	2167.6	1.2084	− 91.0794 93
QCCSD	2167.7	1.2084	− 91.0794 94

Tables VI, VII, and VIII show the results of benchmark computations of spectroscopic constants and electron affinity of CN, using both RHF and UHF reference wavefunctions. All calculation were performed with the STO-3G minimal basis set, and with all electrons correlated. The energies were calculated for each of the methods at five points around the minimum, with a quartic fit to obtain the equilibrium bond length, r_e, the equilibrium energies, and the harmonic vibrational wavenumber. Electron affinities were obtained from corresponding calculations on CN^- and with inclusion of computed zero-point vibrational energies. In the STO-3G basis set, there are large absolute errors (including negative electron affinities), but as in previous work,[59] comparison with the FCI result is still meaningful.

RHF strongly overestimates the bond strength of the CN radical, and all of the CCSD methods also show bond lengths that are too short, and frequencies that are too large. In TCCSD, the error is considerable (93 cm^{-1}), but in VCCSD it is much smaller (38 cm^{-1}). UCCSD, ECCSD, and QCCSD all give similar results to VCCSD.

UHF strongly underestimates the bond strength of the radical, but on applying any of the CCSD methods with an UHF reference, results of the same order of magnitude as with the RHF reference (i.e., slight overbinding) are obtained. In general, the errors are somewhat smaller than with the RHF

TABLE VII. Harmonic vibrational wavenumber, bond lengths, energies, and electron affinities for CN with RHF reference wavefunction.

Method	ω_e / cm^{-1}	r$_e$ / Å	E$_{CN}$ / Hartree	EA / eV
RHF	2534.2	1.1649	− 90.9975 43	− 1.631
FCI	2043.7	1.2294	− 91.1801 28	− 2.644
TCCSD	2136.9	1.2257	− 91.1727 32	− 2.581
VCCSD	2081.3	1.2257	− 91.1746 12	− 2.598
UCCSD	2091.5	1.2249	− 91.1735 87	− 2.580
ECCSD	2077.8	1.2249	− 91.1749 28	− 2.602
QCCSD	2079.0	1.2259	− 91.1749 01	− 2.602

TABLE VIII. Harmonic vibrational wavenumber, bond lengths, energies, and electron affinities for CN with UHF reference wavefunction.

Method	ω_e / cm^{-1}	r$_e$ / Å	E$_{\mathrm{CN}}$ / Hartree	EA / eV
UHF	1562.0	1.2346	− 91.0263 90	− 2.476
FCI	2043.7	1.2294	− 91.1801 28	− 2.644
TCCSD	2098.2	1.2264	− 91.1760 25	− 2.673
VCCSD	2063.8	1.2273	− 91.1755 38	− 2.624
UCCSD	2095.9	1.2248	− 91.1739 45	− 2.589
ECCSD	2063.8	1.2270	− 91.1754 73	− 2.618
QCCSD	2087.5	1.2256	− 91.1750 33	− 2.605

reference, but the trend is the same: VCCSD and ECCSD have significantly smaller errors of about 20 cm^{-1}.

The picture is less clear with respect to absolute energies and the electron affinity. Each of the coupled cluster methods gives electron affinities quite similar to FCI because of cancellation of the overbinding errors in both the CN radical and CN anion. The electron affinities when an UHF reference wavefunction is used for the radical, for each of the coupled cluster methods, are somewhat closer to the FCI values than with a RHF reference.

IV. CONCLUSIONS

In this paper, we have examined through numerical experiments the performance of alternative coupled cluster ansätze. The results show clearly that where there are no strong static correlation effects, traditional projection coupled cluster gives essentially the same results as variational, extended, and unitary CC methods. However, when static correlation is strong, for example as a bond is broken, there are substantial differences, and VCC appears as a significantly better theory, even when the reference wavefunction is very poor. In particular, the variational principle can provide a helpful constraint in cases where TCC gives energies that are spuriously below FCI.

VCCSD is a completely impractical theory, and even the more pragmatic ECCSD has polynomial computational cost scaling at a much higher order of magnitude than TCCSD. However, these results give impetus to a search for approximations to VCCSD that retain some of its advantageous properties.

ACKNOWLEDGMENTS

This work was performed using the computational facilities of the Advanced Research Computing at Cardiff (ARCCA) Division, Cardiff University, and was supported in part by EPSRC (GB).

[1] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).
[2] W. Kutzelnigg, J. Chem. Phys. **77**, 3081 (1982).
[3] W. Kutzelnigg and S. Koch, J. Chem. Phys. **79**, 4315 (1983).
[4] W. Kutzelnigg, J. Chem. Phys. **82**, 4166 (1984).
[5] H. F. Schaefer, *Methods of Electronic Structure Theory* (Plenum, New York, 1977).
[6] J. Arponen, Ann. Phys. **151**, 311 (1983).
[7] P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).
[8] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[9] W. Kutzelnigg, Theor. Chem. Acc. **80**, 349 (1991).
[10] T. Van Voorhis and M. Head-Gordon, J. Chem. Phys. **113**, 8873 (2000).
[11] R. J. Bartlett, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **155**, 133 (1989).
[12] J. D. Watts, G. W. Trucks, and R. J. Bartlett, Chem. Phys. Lett. **157**, 359 (1989).
[13] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).
[14] S. Pal, Theor. Chem. Acc. **66**, 207 (1984).
[15] B. Jeziorski and R. Moszynski, Int. J. Quantum Chem. **48**, 161 (1993).
[16] P. Piecuch and R. J. Bartlett, Adv. Quantum Chem. **34**, 295 (1999).
[17] S. Pal, Phys. Rev. A **34**, 2682 (1986).
[18] S. Pal, Phys. Rev. A **42**, 4385 (1990).
[19] K. Ghose, P. Nair, and S. Pal, Chem. Phys. Lett. **211**, 15 (1993).
[20] B. Kumar, N. Vaval, and S. Pal, Chem. Phys. Lett. **295**, 189 (1998).
[21] P. Manohar, N. Vaval, and S. Pal, Chem. Phys. Lett. **387**, 442 (2004).
[22] P.-D. Fan, K. Kowalski, and P. Piecuch, Molec. Phys. **103**, 2191 (2005).
[23] P.-D. Fan and P. Piecuch, Adv. Quantum Chem. **51**, 1 (2006).
[24] P. Piecuch, I. S. O. Pimienta, P.-D. Fan, and K. Kowalski, ACS Symp. Ser. **958**, 37 (2007).
[25] P. Piecuch, K. Kowalski, P.-D. Fan, and I. S. O. Pimienta, Prog. Theor. Chem. Phys. **12**, 119 (2010)
[26] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
[27] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 5644 (2000).
[28] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. McGuire, Int. Rev. Phys. Chem. **21**, 527 (2002).
[29] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P.-D. Fan, M. Lodriguito, M. McGuire, S. Kucharski, T. Kuś, and M. Musiał, Theor. Chem. Acc. **112**, 349 (2004).
[30] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).
[31] T. Van Voorhis and M. Head-Gordon, Chem. Phys. Lett. **330**, 585 (2000).
[32] MOLPRO, a package of *ab initio* programs version 2009.1, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, and M. Schütz *et al.* (2009), see http://www.molpro.net.
[33] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[34] P. E. M. Siegbahn, Chem. Phys. Lett. **109**, 417 (1984).
[35] P. J. Knowles and N. C. Handy, Chem. Phys. Lett. **111**, 315 (1984).
[36] M. Kállay and P. R. Surján, J. Chem. Phys. **113**, 1359 (2000).
[37] S. Hirata and R. J. Bartlett, Chem. Phys. Lett. **321**, 216 (2000).
[38] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[39] C. W. Bauschlicher, Jr. and P. R. Taylor, J. Chem. Phys. **85**, 2779 (1986).
[40] R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **98**, 97 (1983).
[41] J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987).
[42] W. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987).
[43] C. W. Bauschlicher, Jr. and S. R. Langhoff, J. Chem. Phys. **86**, 5595 (1987).
[44] R. J. Bartlett, J. Phys. Chem. **93**, 1697 (1989).
[45] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **94**, 1264 (1991).
[46] X. Li and J. Paldus, Chem. Phys. Lett. **286**, 145 (1998).
[47] J. Krogh and J. Olsen, Chem. Phys. Lett. **344**, 578 (2001).
[48] P. Piecuch, S. A. Kucharski, and K. Kowalski, Chem. Phys. Lett. **344**, 176 (2001).
[49] I. Hubac, J. Pittner, and P. Carsky, J. Chem. Phys. **112**, 8779 (2000).
[50] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, Mol. Phys. **105**, 1335 (2007).
[51] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen, J. Chem. Phys. **113**, 6677 (2000).
[52] G. K. L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).
[53] M. Hanrath, Theor. Chem. Acc. **121**, 187 (2008).
[54] A. Engels-Putzka and M. Hanrath, Mol. Phys. **107**, 143 (2009).
[55] M. Musiał and R. J. Bartlett, J. Chem. Phys. **122**, 224102 (2005).
[56] L. Meissner, S. Hirata, and R. J. Bartlett, Theor. Chem. Acc. **116**, 440 (2006).
[57] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **97**, 4282 (1992).
[58] P. O. Löwdin, Phys. Rev. **97**, 1509 (1955).
[59] R. H. Nobes, J. A. Pople, L. Radom, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **138**, 481 (1987).
[60] R. H. Nobes, D. Moncrieff, M. M. Wah, L. Radom, P. M. W. Gill, and J. A. Pople, Chem. Phys. Lett. **182**, 216 (1991).
[61] W. Chen and H. B. Schlegel, J. Chem. Phys. **101**, 5957 (1994).
[62] H. Yuan and D. Cremer, Chem. Phys. Lett. **324**, 389 (2000).
[63] P. M. W. Gill, J. A. Pople, L. Radom, and R. H. Nobes, J. Chem. Phys. **89**, 7307 (1988).
[64] J. Stanton, J. Chem. Phys. **101**, 371 (1994).
[65] F. Jensen, Chem. Phys. Lett. **169**, 519 (1990).
[66] J. M. L. Martin, J. P. François, and R. Gijbels, Chem. Phys. Lett. **166**, 295 (1990).