Coupled cluster theory for high spin, open shell reference wave functions

Peter J. Knowles, Claudia Hampel, and Hans-Joachim Werner

Citation: The Journal of Chemical Physics **99**, 5219 (1993); doi: 10.1063/1.465990
View online: https://doi.org/10.1063/1.465990
View Table of Contents: http://aip.scitation.org/toc/jcp/99/7
Published by the American Institute of Physics

Coupled cluster theory for high spin, open shell reference wave functions

Peter J. Knowles
School of Chemistry and Molecular Sciences, University of Sussex, Falmer, Brighton, BN1 9QJ, United Kingdom

Claudia Hampel and Hans-Joachim Werner
Fakultät für Chemie, Universität Bielefeld, D-33615 Bielefeld, Germany

(Received 8 April 1993; accepted 14 June 1993)

The coupled cluster method restricted to single and double excitations (CCSD) is considered for the case of a spin restricted Hartree–Fock open shell reference determinant. A spin–orbital based formulation, in which the cluster operator spans exactly the minimal first order interacting space, is presented, and computationally optimal working equations are given. In the limit of a large number of closed shell orbitals, the cost is identical to that of an optimum treatment of an equivalent closed shell problem, which is obtained as a special case of the formulation presented. The theory is applied to the calculation of a number of diatomic potential energy functions and compared with spin-unrestricted theory.

I. INTRODUCTION

There has been considerable interest in using the coupled cluster (CC) ansatz[1,2]

$$\Psi=\exp(\hat{T})\Psi_0 \tag{1}$$

as a means for generating approximate molecular electronic wave functions. The reference wave function Ψ_0 is usually a single Slater determinant, while the cluster operator $\hat{T}$ is usually taken to be a linear combination of spin–orbital excitation operators, which may be grouped according to the numbers of spin orbitals excited

$$\hat{T}=\sum_{l=1}^{N}\hat{T}_l, \tag{2}$$

$$\hat{T}_1=\sum_{ia}t_a^i\hat{e}_{ai},$$

$$\hat{T}_2=\frac{1}{4}\sum_{ijab}t_{ab}^{ij}\hat{e}_{ai}\hat{e}_{bj},$$

$$\cdots. \tag{3}$$

In the above, $i,j,k,\ldots$ label spin orbitals which are occupied in Ψ_0, while $a,b,c,\ldots$ label unoccupied spin orbitals. $\hat{e}_{ai}=a^\dagger i$ are the usual spin–orbital excitation operators, which replace spin orbital ϕ_i by ϕ_a. Often Eq. (2) is truncated to $\hat{T}_1+\hat{T}_2$, yielding the CC singles and doubles (CCSD) ansatz; the physical justification lies in the fact that $\hat{T}_1+\hat{T}_2$ spans the space of functions which interact with Ψ_0 through the two body Hamiltonian.

In standard CC theory, equations which determine the cluster amplitudes t_a^i and t_{ab}^{ij} follow by projecting the Schrödinger equation on the left with Ψ_0, $\hat{e}_{ai}\Psi_0$, $\hat{e}_{ai}\hat{e}_{bj}\Psi_0$,

$$\langle\Psi_0|(\hat{H}-E)\exp(\hat{T})|\Psi_0\rangle=0, \tag{4}$$

$$\langle\Psi_0|\hat{e}_{ia}(\hat{H}-E)\exp(\hat{T})|\Psi_0\rangle=0, \tag{5}$$

$$\langle\Psi_0|\hat{e}_{ia}\hat{e}_{jb}(\hat{H}-E)\exp(\hat{T})|\Psi_0\rangle=0. \tag{6}$$

The efficient solution of these equations is now a well understood problem.[3,4]

When Ψ_0 is a closed shell determinant which is a basis for all symmetries present in the exact electronic wave function and consists of pairs of equivalent α and β spin orbitals, it is straightforward to formulate the CCSD theory entirely spin free; the excitation operators $\hat{e}_{ai}$ can without approximation be replaced by their spin summed counterparts

$$\hat{E}_{ai}=\hat{e}_{\tilde{a}\tilde{i}}+\hat{e}_{\bar{a}\bar{i}}, \tag{7}$$

where $\hat{e}_{\tilde{a}\tilde{i}}$ is supposed to make the excitation $\phi_i^\alpha\rightarrow\phi_a^\alpha$ and $\hat{e}_{\bar{a}\bar{i}}$ excites $\phi_i^\beta\rightarrow\phi_a^\beta$. This leads to considerable savings in the computational costs of solving Eqs. (4)–(6).[5–8]

For open shell systems, however, the situation is somewhat more complicated. One avenue of approach is to simply apply the spin–orbital based theory using the spin unrestricted Hartree–Fock (UHF) function as Ψ_0.[9,10] However, it is well known that the problems associated with the fact that the UHF reference function is not a spin eigenfunction can become magnified when electron correlation effects are introduced by Møller–Plesset perturbation theory.[11] This effect is likely to be less important for coupled cluster wave functions and can be further reduced by the use of a spin restricted Hartree–Fock (RHF) Ψ_0 (Refs. 12 and 13); however, even in that case, the operator $\hat{T}$ does not commute with the spin squared $\hat{S}^2$ and so the CC wave function is spin contaminated. Note that this problem also arises in the equivalent linear configuration interaction theory if the space spanned by the functions $\hat{e}_{ai}\hat{e}_{bj}|\Psi_0\rangle$ is used as a basis; this is due to the fact that for high-spin, open shell cases, this space does not include all possible Slater determinants of given M_S which arise from a particular occupancy of spatial orbitals. For instance, in a three electron case with reference function $|\phi_1^\alpha\phi_1^\beta\phi_2^\alpha|$, the determinant $|\phi_a^\alpha\phi_b^\alpha\phi_2^\beta|$ is a triple excitation and is not included in the first order interacting space. This function would be necessary, however, to generate one of the two possible doublet spin eigenfunctions together with the determinants $|\phi_a^\beta\phi_b^\alpha\phi_2^\alpha|$ and $|\phi_a^\alpha\phi_b^\beta\phi_2^\alpha|$. A quartet spin contamination arises if the latter two Slater determinants have coefficients of different magnitude. As will be shown in Sec.

 0021-9606/93/99(7)/5219/9/$6.00

III, this spin contamination can be easily avoided by using the proper minimal first order interacting space, or, equivalently, by applying appropriate projection operators to the residual vector.

Even for cases in which the linear terms in the configuration expansion give a spin eigenfunction, it is the presence of higher powers of $\hat{T}$ which can introduce spin contamination in a nontrivial way. In fact, the spin contamination problem is not as bad as it might seem at first; as discussed by Rittby and Bartlett[12] and Scuseria,[13] the contributions of the spin contaminated parts of the wave function to the energy are automatically projected away by $\langle\Psi_0|(\hat{H}-E)$ in Eq. (4), provided Ψ_0 is a spin eigenfunction. Nevertheless, spin contamination does have an effect on the energy through angular momentum recoupling that arises in powers of the $\hat{T}$ operator. The same spin contamination behavior is seen also in some perturbation theories based on a RHF Ψ_0,[14,15] and also in UHF perturbation theories which involve explicit spin projectors.[16,17]

Another undesirable feature of spin–orbital based CCSD is the cost, which is a factor of about 3 greater than an equivalent closed shell spin coupled formulation.[13] Particularly when dealing with large molecules consisting mostly of doubly occupied orbitals which have little interaction with the open shell, it seems unacceptable to be unable to advantageously exploit spin pairing in any way.

An alternative approach is to attempt to use a spin symmetry adapted approach throughout. Instead of using the spin–orbital excitation operators $\hat{e}_{\tilde{a}\tilde{i}}$ and $\hat{e}_{\bar{a}\bar{i}}$, the spin coupled operators $\hat{E}_{ai}$ may be used from the outset. Since the Hamiltonian is spin free, $\hat{T}_1+\hat{T}_2$ built in this way will certainly span the first order interacting space. This approach has been described by Janssen and Schaefer[18] and is also implicit in the symmetry adapted cluster (SAC) treatment of Nakatsuji *et al.*[19–21] However, the rigorous application of this approach is quite complicated; e.g., explicit orbital based equations appear to terminate after only the eighth power of some parts of $\hat{T}$ in the expansion of the exponential operator.[18] Although ultimately such a theory might be expected to be optimal, both in its symmetry properties and its efficiency, it is probably too difficult to program in full; this is particularly true when it is considered that the energy gradient will be even more complicated. There are also other approaches which attempt to retain the spin–orbital formulation, while partially correcting the spin contamination through a projection operator.[12]

Recently, a novel approach to the open shell CCSD problem has been presented, in which it appears possible to begin to approach the efficiency of the closed shell methods, while retaining the simplicity of the spin–orbital formulation.[22,23] In this letter, it is shown how the spin–orbital formulation can be adapted such that its cost is identical to an equivalent closed shell calculation in the limit of a large number of closed shell orbitals. There is no need to invoke the new concepts of rotated spin functions introduced in Refs. 22 and 23, and the theory presented is a true CCSD theory in the sense that $\hat{T}\Psi_0$ spans the first order interacting space and is a spin eigenfunction. Some spin contamination remains, but it is argued that its effects are acceptably small.

II. THE SPIN–ORBITAL CCSD EQUATIONS FOR RHF ORBITALS

The spin–orbital CCSD equations (4)–(6) have been given by several authors in explicit form.[3,4,13,22] In the following, we briefly summarize the working equations for the case of a spin restricted high spin, open shell Hartree–Fock (RHF) reference function. Since in this case only a single set of orbitals is used for α and β spin, only one set of two-electron integrals is needed, and the equations can be formulated explicitly for the different spin combinations. In order to distinguish from calculations based on an unrestricted Hartree–Fock (UHF) reference wave function, we will denote this method RHF-UCCSD. As in configuration interaction[5,24,25] and closed shell coupled cluster theories,[5,8] the equations are compactly written in matrix form, which is also computationally most appropriate. Our equations are in principle equivalent to the ones given by Scuseria,[13] but it should be noted that the latter contains several misprints.

In the following, tilde represents pure alpha quantities, e.g., $\tilde{T}^{ij}_{ab}=t^{\tilde{i}\tilde{j}}_{\tilde{a}\tilde{b}}$, overbar represents pure beta quantities, e.g., $\bar{T}^{ij}_{ab}=t^{\bar{i}\bar{j}}_{\bar{a}\bar{b}}$, and without anything represents mixed quantities with alpha labels first, e.g., $T^{ij}_{ab}=t^{i\bar{j}}_{a\bar{b}}$. Upper case letters denote matrices and lower case letters denote column vectors; different vectors and matrices will be distinguished by superscripts, which always correspond to occupied orbitals. For computational convenience, all matrices are defined in an orthogonal basis spanned by the singly occupied and virtual orbitals. This space is denoted by the labels $a,b,\ldots$. The occupied (internal) orbitals will be denoted by the indices $i,j,\ldots$; note that the two spaces overlap in the space of singly occupied orbitals, which are denoted by the labels $t,u,v\ldots$. The Pauli principle requires that amplitudes corresponding to excitations into open shell α orbitals vanish, i.e., $\tilde{t}^{i}_{t}=0$, $\tilde{T}^{ij}_{st}=0$, $T^{ij}_{ta}=0$, etc. Arbitrary orbitals will be labeled by the indices p, q, r, and s.

The alpha and beta Fock matrices are defined as

$$\tilde{F}_{rs}=H_{rs}+\sum_{l}[2(rs|ll)-(rl|ls)]+\sum_{t}[(rs|tt)-(rt|ts)], \tag{8}$$

$$\bar{F}_{rs}=H_{rs}+\sum_{l}[2(rs|ll)-(rl|ls)]+\sum_{t}(rs|tt), \tag{9}$$

where H_{rs} is the matrix element of the one-electron Hamiltonian and $(pq|rs)$ is the two-electron integral. The two-electron integrals with two occupied orbitals are represented as the matrices of internal Coulomb and exchange operators

$$J^{kl}_{ab}=(ab|kl), \tag{10}$$

$$K^{kl}_{ab}=(ak|lb), \tag{11}$$

and the integrals with three occupied orbital by the vectors

$$k_a^{kli}=(ak|li). \tag{12}$$

The integrals with three and four external orbitals enter via the operators

$$J(\tilde{E}^{ij})_{ab}=\sum_{cr}(ab|rc)\tilde{E}^{ij}_{rc}=\sum_{c}(ab|ic)\tilde{t}^j_c, \tag{13}$$

$$K(\tilde{E}^{ij})_{ab}=\sum_{cr}(ar|bc)\tilde{E}^{ij}_{rc}=\sum_{c}(ai|bc)\tilde{t}^j_c, \tag{14}$$

$$K(\tilde{D}^{ij})_{rs}=\sum_{cd}(rc|ds)\tilde{D}^{ij}_{cd}, \tag{15}$$

where

$$\tilde{E}^{ij}_{rc}=\delta_{rt}\tilde{t}^j_c \tag{16}$$

and $\tilde{D}^{ij}_{rs}$ are modified coefficient matrices, which will be defined below. The operators $J(\bar{E}^{ij})$, $K(\bar{E}^{ij})$, $K(\bar{D}^{ij})$, and $K(D^{ij})$ are defined analogously. As discussed in Ref. 8, these operators can be computed directly from the atomic orbital (AO) integrals, which makes possible integral-direct implementations. The quantities $\tilde{f}^k$ and $K(\tilde{D}^{ij})^k$ represent the kth columns of the matrices $\tilde{F}$ and $K(\tilde{D}^{ij})$, respectively.

For computational convenience, we introduce modified coefficient matrices

$$\tilde{B}^{ik}=\tfrac{1}{2}\tilde{T}^{ik}+\tilde{t}^i\tilde{t}^{k\dagger}-\tilde{t}^k\tilde{t}^{i\dagger}, \tag{17}$$

$$B^{ik}=\tfrac{1}{2}T^{ik}+\tilde{t}^i\bar{t}^{k\dagger}, \tag{18}$$

$$\tilde{C}^{ik}=\tilde{T}^{ik}+\tilde{t}^i\bar{t}^{k\dagger}-\bar{t}^k\tilde{t}^{i\dagger}, \tag{19}$$

$$C^{ik}=T^{ik}+\tilde{t}^i\bar{t}^{k\dagger}, \tag{20}$$

$$\tilde{D}^{ij}=\tilde{C}^{ij}+\tilde{E}^{ij}+\tilde{E}^{ji\dagger}-\tilde{E}^{ij\dagger}-\tilde{E}^{ji}, \tag{21}$$

$$D^{ij}=C^{ij}+\bar{E}^{ij}+\bar{E}^{ji\dagger}. \tag{22}$$

Note that $\tilde{B}^{ik}$ is defined totally antisymmetric, in contrast to the corresponding quantities in previous formulations.[4,13]

Using the above definitions, and assuming summations over repeated indices, the correlation energy is given by

$$E-E_0=\mathrm{Tr}[(\tfrac{1}{2}\tilde{C}^{ij}+\tfrac{1}{2}\bar{C}^{ij}+C^{ij})K^{ji}]+\tilde{f}^{i\dagger}\tilde{t}^i+\bar{f}^{i\dagger}\bar{t}^i. \tag{23}$$

The residual vectors and matrices, which must vanish at convergence, are given by

$$\tilde{v}^i=\tilde{s}^i-\tilde{\beta}_{ik}\tilde{t}^k+\tilde{T}^{ik}\tilde{r}^k+T^{ik}\bar{r}^k, \tag{24}$$

$$\bar{v}^i=\bar{s}^i-\bar{\beta}_{ik}\bar{t}^k+\bar{T}^{ik}\bar{r}^k+T^{ki\dagger}\tilde{r}^k, \tag{25}$$

$$\tilde{V}^{ij}=K^{ij}-K^{ji}+K(\tilde{D}^{ij})+\tilde{\alpha}_{ij,kl}\tilde{C}^{kl}+\tilde{G}^{ij}+\tilde{G}^{ji\dagger}, \tag{26}$$

$$\bar{V}^{ij}=K^{ij}-K^{ji}+K(\bar{D}^{ij})+\bar{\alpha}_{ij,kl}\bar{C}^{kl}+\bar{G}^{ij}+\bar{G}^{ji\dagger}, \tag{27}$$

$$V^{ij}=K^{ij}+K(D^{ij})+\alpha_{ij,kl}C^{kl}+G^{i\bar{j}}+G^{\bar{j}i\dagger}. \tag{28}$$

The G^{ij} matrices are given by

$$\tilde{G}^{ij}=\tilde{T}^{ij}\tilde{X}-[K(\tilde{D}^{ij})^k+k^{ijk}-k^{jik}]\tilde{t}^{k\dagger}-\tilde{\beta}_{ik}\tilde{C}^{kj}+\hat{P}_{ij}(\tilde{s}^i\tilde{t}^{j\dagger}+\tilde{T}^{ik}\tilde{Y}^{kj}+T^{ik}Y^{\bar{k}j}), \tag{29}$$

$$G^{i\bar{j}}=T^{ij}\bar{X}-[K(D^{ij})^k+k^{ijk}]\bar{t}^{k\dagger}-\tilde{\beta}_{ik}C^{kj}+\tilde{s}^i\bar{t}^{j\dagger}+\tilde{T}^{ik}Y^{k\bar{j}}+T^{ik}\bar{Y}^{kj}-T^{kj}\tilde{Z}^{ki}, \tag{30}$$

where $\hat{P}_{ij}A^{ij}=A^{ij}-A^{ji}$. The matrices $\bar{G}^{ij}$ and $G^{\bar{i}\bar{j}}$ can be obtained from $\tilde{G}^{ij}$ and $G^{\bar{i}j}$, respectively, by exchanging tilde and overbar quantities and replacing T^{kl}, D^{kl}, and $Y^{\bar{k}\bar{l}}$ by $T^{lk\dagger}$, $D^{lk\dagger}$, and $Y^{k\bar{l}}$ and vice versa. Similar rules hold for the beta–beta or beta–alpha counterparts of the intermediate quantities given below.

The intermediate quantities are the scalars

$$\tilde{\alpha}_{ij,kl}=K^{ij}_{kl}+(k^{lki\dagger}\tilde{t}^j)+(k^{klj\dagger}\tilde{t}^i)+\tfrac{1}{2}\mathrm{Tr}(\tilde{C}^{ij}K^{lk}), \tag{31}$$

$$\alpha_{ij,kl}=K^{ij}_{kl}+(k^{lki\dagger}\bar{t}^j)+(k^{klj\dagger}\bar{t}^i)+\mathrm{Tr}(C^{ij}K^{lk}), \tag{32}$$

$$\tilde{\beta}_{ik}=\tilde{F}_{ik}+(\tilde{t}^{i\dagger}\tilde{f}^k)+\mathrm{Tr}[(\tilde{C}^{il}+C^{il})K^{lk}]+(\bar{t}^l+\bar{t}^l)^\dagger k^{lki}-\tilde{t}^{i\dagger}k^{kli}, \tag{33}$$

vectors

$$\tilde{r}^k=\tilde{f}^k+K^{kl}(\bar{t}^l+\bar{t}^l)-K^{lk}\bar{t}^l, \tag{34}$$

$$\tilde{s}^i=\tilde{f}^i+(H-\tilde{A}^\dagger)\tilde{t}^i+K(\tilde{D}^{ik}+D^{ik})^k-(\tilde{T}^{kl}+T^{kl})k^{lki}, \tag{35}$$

and matrices

$$\tilde{A}=K^{lk}(\tilde{T}^{lk}+T^{lk})^\dagger, \tag{36}$$

$$\tilde{X}=\tilde{F}-\tilde{A}+J(\tilde{E}^{kk}+\bar{E}^{kk})-K(\tilde{E}^{kk})-\tilde{r}^k\tilde{t}^{k\dagger}, \tag{37}$$

$$\tilde{Y}^{kj}=K^{kj}-J^{kj}+K(\tilde{E}^{kj})-J(\tilde{E}^{kj})+K^{kl}(\tilde{B}^{lj}+B^{jl\dagger})-K^{lk}\tilde{B}^{lj}+(k^{lkj}-k^{klj})\tilde{t}^{l\dagger}+\tilde{f}^k\tilde{t}^{j\dagger}, \tag{38}$$

$$Y^{k\bar{j}}=K^{kj}+K(\bar{E}^{kj})+K^{kl}(\bar{B}^{lj}+B^{lj})-K^{lk}B^{lj}-k^{klj}\bar{t}^{l\dagger}+\tilde{f}^k\bar{t}^{j\dagger}, \tag{39}$$

$$\tilde{Z}^{kj}=J^{kj}+J(\tilde{E}^{kj})-(K^{lk}B^{jl}+k^{lkj}\bar{t}^{l\dagger}). \tag{40}$$

In deriving the above equations, we have used the relation

$$\Big[\sum_k K(\tilde{D}^{ik}+D^{ik})^k\Big]_a=\Big[\sum_k K(\tilde{C}^{ik}+C^{ik})^k+K^{ik}(\tilde{t}^k+\bar{t}^k)-J^{ik}\bar{t}^k\Big]_a+[(\tilde{F}-H)\tilde{t}^i]_a, \tag{41}$$

which makes it possible to use a single set of external exchange operators for each pair. This is most important if the external exchange operators are computed directly from the two-electron integrals in the AO basis. These relations hold only for a complete set of pair functions; for calculations in which selected electron pairs are correlated, corrections to the last term of Eq. (41) are necessary.

For a linear (spin-unrestricted) wave function $|\Psi_{\mathrm{CI}}\rangle=(1+\hat{T}_1+\hat{T}_2)|\Psi_0\rangle$, with $\hat{T}_1$ and $\hat{T}_2$ constructed from the operators and amplitudes used in the above approach, the expectation value of $\hat{S}^2$ is given by (summation over all indices is implied)

$$\langle\Psi_{\mathrm{CI}}|\hat{S}^2|\Psi_{\mathrm{CI}}\rangle=M(M+1)+\frac{(\Delta^i_a)^2+(\tilde{\Delta}^{pj}_{ab})^2+(\bar{\Delta}^{ij}_{ar})^2}{\langle\Psi_{\mathrm{CI}}|\Psi_{\mathrm{CI}}\rangle} \tag{42}$$

with

$$\langle\Psi_{CI}|\Psi_{CI}\rangle = 1+(\tilde{t}_a^p)^2+(\bar{t}_r^i)^2+(T_{ar}^{pj})^2 + \tfrac{1}{4}[(\tilde{T}_{ab}^{pq})^2+(\bar{T}_{rs}^{ij})^2], \tag{43}$$

$$\Delta_a^i = \tilde{t}_a^i - \bar{t}_a^i + \sum_t T_{at}^{ti}, \tag{44}$$

$$\tilde{\Delta}_{ab}^{pj} = \tfrac{1}{2}(T_{ab}^{pj} - T_{ba}^{pj} - \tilde{T}_{ab}^{pj}), \tag{45}$$

$$\bar{\Delta}_{ar}^{ij} = \tfrac{1}{2}(T_{ar}^{ij} - T_{ar}^{ji} - \bar{T}_{ar}^{ij}). \tag{46}$$

Here and in the following section, the indices $t,u,v,...$ are used to denote orbitals lying in the singly occupied space; $a,b,c...$ denote only the external orbitals; $i,j,k...$ denote true closed shell orbitals; p,q denote closed and open shell occupied orbitals; and r,s denote open shell and external orbitals.

III. PARTIALLY SPIN ADAPTED COUPLED CLUSTER THEORY

As discussed in Sec. I, the above equations are obtained by using a cluster operator and projection set which span the first order interacting space generated by the operators $\hat{e}_{\tilde{a}\tilde{i}}$, $\hat{e}_{\bar{a}\bar{i}}$, $\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\tilde{b}\bar{j}}$, $\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{b}\bar{j}}$, and $\hat{e}_{\bar{a}\bar{t}}\hat{e}_{\bar{b}\bar{j}}$. In fully spin coupled theory,[18] it is recognized that the Hamiltonian operator is spin free, and therefore these operators may be replaced by the smaller set $\hat{E}_{ai}$, $\hat{E}_{at}$, $\hat{E}_{ti}$, and their products. A simpler theory, including some but not all of the spin coupling, may be obtained by using the operators $\hat{E}_{ai}$, $\hat{e}_{\tilde{a}\tilde{i}}$, $\hat{e}_{\bar{t}\bar{i}}$, and their products; because the orbitals ϕ_i^α are occupied and ϕ_i^β are unoccupied in Ψ_0, the wave function should be spin adapted for CISD configuration expansions which are linear in these operators, but not for the nonlinear CCSD case, where products of $\hat{T}$ can give a spin contaminated contribution to the wave function, as discussed in Sec. I. This ansatz has been suggested previously and is denoted "partially spin adapted" (PSA) CCSD by Janssen and Schaefer.[18] In Ref. 18, it was suggested that this scheme might be used in conjunction with a basis of spin adapted excited states, or a suitable projection operator in order to avoid the spin contamination problem. It is suggested here that this approach be adopted without projection operators; for large scale calculations, an explicit orbital based method is absolutely necessary, since methods which rely on explicit computation of spin function algebra will always be too slow. Some spin contamination remains, but this will only negligibly influence the computed energies.

We note further that the suggestion to use $\hat{E}_{ai}$, $\hat{e}_{\tilde{a}\tilde{i}}$, $\hat{e}_{\bar{t}\bar{i}}$, and all their binary products, given explicitly in Eq. (3.23) of Ref. 18, as the basis for the $\hat{T}$ operator is not quite correct. On applying the Hamiltonian operator to the RHF reference wave function, we obtain

$$\begin{aligned}\hat{H}|\Psi_0\rangle = \Bigg\{ & E_{RHF} + \sum_i^{doc}\sum_t^{soc} \bar{F}_{ti}\hat{e}_{\bar{t}\bar{i}} + \sum_i^{doc}\sum_a^{vir} \frac{1}{2}(\tilde{F}_{ai}+\bar{F}_{ai})\hat{E}_{ai} + \sum_t^{soc}\sum_a^{vir} \tilde{F}_{at}\hat{e}_{\tilde{a}\tilde{t}} + \frac{1}{2}\sum_{ij}^{doc}\sum_{tu}^{soc}(ti|uj)\hat{e}_{\bar{t}\bar{i}}\hat{e}_{\bar{u}\bar{j}} \\ & + \sum_{ij}^{doc}\sum_a^{vir}\sum_u^{soc}(ai|uj)\hat{E}_{at}\hat{e}_{\bar{u}\bar{j}} + \frac{1}{2}\sum_{ij}^{doc}\sum_{ab}^{vir}(ai|bj)\hat{E}_{ai}\hat{E}_{bj} + \sum_i^{doc}\sum_t^{soc}\sum_{ab}^{vir}(at|bj)\hat{e}_{\tilde{a}\tilde{t}}\hat{E}_{bj} + \frac{1}{2}\sum_{tu}^{soc}\sum_{ab}^{vir}(at|bu)\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\tilde{b}\tilde{u}} \\ & + \sum_{tu}^{soc}\sum_i^{doc}\sum_a^{vir}(at|ui)\left[\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{u}\bar{i}} - \frac{1}{2}\delta_{tu}(\hat{e}_{\tilde{a}\tilde{i}}-\hat{e}_{\bar{a}\bar{i}})\right]\Bigg\}|\Psi_0\rangle. \end{aligned} \tag{47}$$

The coefficients of $\hat{e}_{\tilde{a}\tilde{i}}$ and $\hat{e}_{\bar{a}\bar{i}}$ — $\tilde{F}_{ai}$ and $\bar{F}_{ai}$, respectively — differ by the open shell exchange $\Sigma_t^{soc}(at|ti)$, and so to span the first order space, it is necessary to use both $\hat{e}_{\tilde{a}\tilde{i}}$ and $\hat{e}_{\bar{a}\bar{i}}$, not just $\hat{E}_{ai}$. However, the same integrals also appear as the coefficients of the diagonal "semi-internal" double excitations $\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{t}\bar{i}}$. The net result is that there are two independent interacting spin eigenfunctions for the spatial configuration $\cdots i^1\cdots t^1\cdots a^1$, which may be generated by the operators $\hat{e}_{\tilde{a}\tilde{i}}+\hat{e}_{\bar{a}\bar{i}}$ and $\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{t}\bar{i}}-\frac{1}{2}\hat{e}_{\tilde{a}\tilde{i}}+\frac{1}{2}\hat{e}_{\bar{a}\bar{i}}$; in principle, any linear combination of these two would also suffice, but this choice gives two first order interacting configurations which are mutually orthogonal. Thus a correct set of operators, which exactly spans the first order interacting space, is

$$\hat{e}_{\bar{t}\bar{i}},\ \hat{E}_{ai},\ \hat{e}_{\tilde{a}\tilde{t}},\ \hat{e}_{\bar{t}\bar{i}}\hat{e}_{\bar{u}\bar{j}},\ \hat{E}_{at}\hat{e}_{\bar{t}\bar{j}},\ \hat{E}_{at}\hat{E}_{bj},\ \hat{E}_{at}\hat{e}_{\tilde{b}\tilde{u}},\ \hat{e}_{\tilde{a}\tilde{t}}\hat{E}_{bj},$$

$$(\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{u}\bar{i}} - \tfrac{1}{2}\delta_{tu}\hat{e}_{\tilde{a}\tilde{i}} + \tfrac{1}{2}\delta_{tu}\hat{e}_{\bar{a}\bar{i}}). \tag{48}$$

Using these operators, the cluster operator $\hat{T}$ can be written in the form

$$\hat{T} = \tilde{t}_a^p\hat{e}_{\tilde{a}\tilde{p}} + \bar{t}_r^i\hat{e}_{\bar{r}\bar{i}} + T_{tu}^{ij}\hat{e}_{\bar{t}\bar{i}}\hat{e}_{\bar{u}\bar{j}} + T_{at}^{ij}\hat{E}_{at}\hat{e}_{\bar{t}\bar{j}} + T_{ab}^{ij}\hat{E}_{ai}\hat{E}_{bj} + T_{ab}^{tj}\hat{e}_{\tilde{a}\tilde{t}}\hat{E}_{bj} + T_{au}^{tj}\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\bar{u}\bar{j}} + T_{ab}^{tu}\hat{e}_{\tilde{a}\tilde{t}}\hat{e}_{\tilde{b}\tilde{u}} \tag{49}$$

with the restrictions

$$\tilde{t}_a^i = t_a^i - \frac{1}{2}\sum_t T_{at}^{ti}, \tag{50}$$

$$\bar{t}_a^i = t_a^i + \frac{1}{2}\sum_t T_{at}^{ti}. \tag{51}$$

Equating $\hat{T}$ to the corresponding spin unrestricted cluster operator used in Sec. II yields the additional relations

$$\tilde{T}_{ab}^{tu} = T_{ab}^{tu}, \tag{52}$$

$$\bar{T}_{tu}^{ij} = T_{tu}^{ij}, \tag{53}$$

$$\tilde{T}^{pj}_{ab}=T^{pj}_{ab}-T^{pj}_{ba}, \tag{54}$$

$$\bar{T}^{ij}_{ar}=T^{ij}_{ar}-T^{ji}_{ar}. \tag{55}$$

Setting further $t^i_t=\frac{1}{2}\bar{t}^i_t$ and $t^i_a=\frac{1}{2}\tilde{t}^i_a$, we obtain a unique set of amplitudes $T^{pq}_{rs}(p\geq q)$ and t^p_r to be solved for (note that for $p=q$, the matrices T^{pp} are symmetric and only the elements T^{pp}_{rs} with $r\geq s$ are independent parameters). The number of independent amplitudes is exactly the same as in a fully spin adapted formulation and about three times smaller than in the spin-unrestricted theory of Sec. II. From Eq. (42), it is obvious that a linear wave function $\Psi_{\mathrm{RCISD}}=(1+\hat{T})\,|\Psi_0\rangle$, with coefficients which satisfy Eqs. (50)–(55), is fully spin adapted. Thus, spin contamination in the coupled cluster wave function $\Psi_{\mathrm{RCCSD}}=e^{\hat{T}}\,|\Psi_0\rangle$ can arise only via products of the cluster amplitudes.

A corresponding minimal set of coupled equations can be obtained by projecting the Schrödinger equation onto the space of functions generated by applying the operators (48) to the reference function Ψ_0. Since the configurations obtained in this way are not orthonormal, it is more convenient to project the Schrödinger equation onto the equivalent contravariant space,[5,8] defined such that its members Φ^r_p and Φ^{rs}_{pq} satisfy

$$\langle\Phi^r_p|\,\hat{T}\,|\Psi_0\rangle=t^p_r, \tag{56}$$

$$\langle\Phi^{rs}_{pq}|\,\hat{T}\,|\Psi_0\rangle=\mathcal{T}^{pq}_{rs} \tag{57}$$

with

$$\mathcal{T}^{pq}=\tfrac{1}{6}(\tilde{T}^{pq}+\bar{T}^{pq}+2T^{pq}+2T^{qp\dagger}+T^{pq\dagger}+T^{qp}). \tag{58}$$

Here we assume that all elements of the matrices $\tilde{T}^{pq}$, $\bar{T}^{pq}$, and T^{pq}, which correspond to exclusion principle violating configurations, are zero. Insertion of Eqs. (52)–(55) into Eq. (58) shows that each unique element $\mathcal{T}^{pq}_{rs}$ is proportional to a coefficient T^{pq}_{rs} with a particular normalization factor, namely,

$$\begin{aligned}
&\mathcal{T}^{ij}_{tu}=\tfrac{1}{6}T^{ij}_{tu}, \quad \mathcal{T}^{ij}_{at}=\tfrac{1}{2}T^{ij}_{at}, \quad \mathcal{T}^{ij}_{ta}=\tfrac{1}{2}T^{ji}_{at},\\
&\mathcal{T}^{ij}_{ab}=T^{ij}_{ab}, \quad \mathcal{T}^{it}_{ab}=T^{ti}_{ba}, \quad \mathcal{T}^{ti}_{ab}=\tfrac{1}{2}T^{ti}_{ab},\\
&\mathcal{T}^{ti}_{au}=\tfrac{1}{3}T^{ti}_{au}, \quad \mathcal{T}^{tu}_{ab}=\tfrac{1}{6}T^{tu}_{ab}.
\end{aligned} \tag{59}$$

Note that $\mathcal{T}^{pq}=\mathcal{T}^{qp\dagger}$, while a corresponding relation is not valid for the original T^{pq}. In the following, it will be convenient to use the renormalized coefficients $\mathcal{T}^{pq}_{rs}$ as a set of independent parameters. The contravariant functions which satisfy Eqs. (56)–(58) are

$$\Phi^r_p=\tfrac{1}{2}\hat{E}_{rp}\,|\Psi_0\rangle, \tag{60}$$

$$\Phi^{rs}_{pq}=\tfrac{1}{6}(2\hat{E}_{rp}\hat{E}_{sq}+\hat{E}_{sp}\hat{E}_{rq})\,|\Psi_0\rangle. \tag{61}$$

The resulting set of partially spin adapted coupled cluster equations is derived straightforwardly by making appropriate linear combinations of the residual vectors and matrices given in Eqs. (24)–(28) and replacing $\tilde{t}^p_a$, $\bar{t}^p_r$, $\tilde{T}^{pq}_{ab}$, and $\bar{T}^{ij}_{rs}$ by the relations given in Eqs. (50)–(55) and (59). In terms of the quantities in Eqs. (24)–(28), the partially spin adapted residuals take the general form

$$v^p=\tfrac{1}{2}(\tilde{v}^p+\bar{v}^p), \tag{62}$$

$$\mathcal{V}^{pq}=\tfrac{1}{6}(\tilde{V}^{pq}+\bar{V}^{pq}+2V^{pq}+2V^{qp\dagger}+V^{qp}+V^{pq\dagger}) \quad (p\geq q). \tag{63}$$

As above for the coefficient vectors and matrices, we assume that all elements of the vectors $\tilde{v}^p$ and $\bar{v}^p$ and matrices $\tilde{V}^{pq}$, $\bar{V}^{pq}$, and V^{pq}, which correspond to exclusion principle violating configurations, are zero.

In the following, we will present the explicit equations for the computationally most demanding cases. The theory will be presented in full detail elsewhere.[26] For the sake of simplicity in later expressions, we define

$$t^{p,-1}=\tfrac{1}{2}(\tilde{t}^p-\bar{t}^p), \tag{64}$$

$$\mathcal{T}^{pq,-1}=\tfrac{1}{6}(\tilde{T}^{pq}+2T^{qp\dagger}+T^{pq\dagger}-\bar{T}^{pq}-2T^{pq}-T^{qp}). \tag{65}$$

Note that the elements of these vectors and matrices can be expressed in terms of the independent parameters $t^{p,1}$ and $\mathcal{T}^{pq,1}$. In particular, $t^{i,-1}_a=-\Sigma_t T^{ti}_{at}$, $t^{t,-1}_a=t^{t,1}_a=\frac{1}{2}\tilde{t}^i_a$, $t^{i,-1}_t=-t^{i,1}_t=-\frac{1}{2}\bar{t}^i_t$. For convenience, we also set $t^{p,1}=t^p$ and $\mathcal{T}^{ij,1}=\mathcal{T}^{ij}$. Corresponding matrices $\mathcal{C}^{ij,\mu}$ and $\mathcal{D}^{ij,\mu}$ are defined similarly. The residual for the single excitations now takes the simple form

$$v^p=s^{p,1}-\beta_{pq,\mu}t^{q,\mu}+(2\mathcal{T}^{pq,\mu}-\mathcal{T}^{qp,\mu})r^{q,\mu}, \tag{66}$$

where summation over $\mu=\pm 1$ is implied. The intermediate quantities are defined as

$$A^\mu=K^{qp}(2\mathcal{T}^{pq,\mu}-\mathcal{T}^{qp,\mu}), \tag{67}$$

$$F^\mu=\tfrac{1}{2}(\tilde{F}+\mu\bar{F}), \tag{68}$$

$$f^{p,\mu}=\tfrac{1}{2}(\tilde{f}^p+\mu\bar{f}^p), \tag{69}$$

$$r^{p,\mu}=f^{p,\mu}+[(1+\mu)K^{pq}-K^{qp}]t^{q,\mu}, \tag{70}$$

$$\begin{aligned}
s^{p,\mu}=f^{p,\mu}+Ht^{p,\mu}&-\frac{1+\mu}{2}A^{\nu\dagger}t^{p,\nu}-\frac{1-\mu}{2}A^{\nu\dagger}t^{p,-\nu}\\
&+[2K(\mathcal{D}^{qp,\mu\dagger})-K(\mathcal{D}^{pq,\mu\dagger})]^q\\
&-(2\mathcal{T}^{sr,\mu}-\mathcal{T}^{rs,\mu})^\dagger k^{srp},
\end{aligned} \tag{71}$$

$$\begin{aligned}
\beta_{pq,\mu}=F_{pq,\mu}&+\frac{1+\mu}{2}t^{p,\nu\dagger}f^{q,\nu}+\frac{1-\mu}{2}t^{p,\nu\dagger}f^{q,-\nu}+\mathrm{Tr}[(2\mathcal{C}^{rp,\mu}\\
&-\mathcal{C}^{pr,\mu})K^{qr}]+t^{r,\mu\dagger}[(1+\mu)k^{rqp}-k^{qrp}].
\end{aligned} \tag{72}$$

In the closed shell case, all $t^{i,-1}$ and $\mathcal{T}^{ij,-1}$ are zero and only the $\mu=1$ terms survive. The equations are then identical to those given in Ref. 8. At first sight, it appears as if the computational effort would be twice as large in the open shell case, since all quantities are needed for $\mu=\pm 1$. However, this is not quite true, since

$$\mathcal{T}^{ij,-1}_{ab}=0. \tag{73}$$

Hence, the costly matrix operations in Eqs. (67) and (71) involve only the small rectangular blocks $\mathcal{T}^{ij,-1}_{rt}$ and $\mathcal{T}^{ij,-1}_{tr}$. The most expensive contributions in the above equations are the operators $K(\mathcal{D}^{ij,1})_{rp}$, which require of the order of N^3m^3 operations. The additional exchange operators $K(\mathcal{D}^{ij,-1})$ require much less effort, since they can be written as

$$K(\mathscr{D}^{ij,-1})_{rp}=\sum_{ab}\mathscr{C}^{ij,-1}_{ab}(ra|bp)+\sum_{at}[\mathscr{D}^{ij,-1}_{at}(ra|tp)+\mathscr{D}^{ij,-1}_{ta}(rt|ap)]+\sum_{tu}\mathscr{D}^{ij,-1}_{tu}(rt|up)$$
$$=\sum_{ab}\mathscr{C}^{ij,-1}_{ab}(ra|bp)+\sum_{ta}J^{tp}_{ra}\mathscr{D}^{ij,-1}_{at}+\sum_{ts}K^{tp}_{rs}\mathscr{D}^{ij,-1}_{ts}. \quad (74)$$

Using

$$\mathscr{C}^{ij,-1}_{ab}=t^{i,1}_a t^{j,-1}_b-\tfrac{1}{3}t^{i,-1}_a t^{j,1}_b-\tfrac{2}{3}t^{j,-1}_a t^{i,1}_b, \quad (75)$$

$$K(t^{i,1}_d t^{j,-1}_b)_{rs}=\sum_a t^{i,1}_a\sum_b t^{j,-1}_b(ra|bs), \quad (76)$$

the contribution of the integrals with three external orbitals can be written as

$$\sum_{ab}\mathscr{C}^{ij,-1}_{ab}(ra|bp)=K((t^{i,1}t^{j,-1\dagger})_{\mathrm{ext}})_{rp}-\frac{1}{3}K((t^{j,1}t^{i,-1\dagger})_{\mathrm{ext}})_{pr}-\frac{2}{3}K((t^{i,1}t^{j,-1\dagger})_{\mathrm{ext}})_{pr}, \quad (77)$$

where the subscript "ext" indicates that only the external parts of the matrix are involved. The evaluation of the operators $K((t^{i,1}t^{j,-1\dagger})_{\mathrm{ext}})_{rs}$ corresponds to a generalized partial integral transformation. Since either r or s is internal, it requires only of the order of N^3m^2 operations, as compared to the N^3m^3 operations needed to compute the operators $K(\mathscr{D}^{ij,1})_{rp}$.

The residuals for the double excitations can be derived along similar lines, but quite a large number of cases appear and it is difficult to formulate them compactly. Here, we will concentrate on the case for excitations from closed shell orbitals into virtual orbitals. Clearly, for large basis sets and only a few open shell orbitals, these will be computationally most demanding. All further terms will involve at least one open shell shell index, and therefore, one fewer virtual or closed shell index, and will require only little additional computational cost.

The residuals $\mathscr{V}^{ij}_{ab}$ can be written as

$$\mathscr{V}^{ij}_{ab}=K^{ij}_{ab}+K(\mathscr{D}^{ij,1})_{ab}+\alpha^{+}_{ij,pq}\mathscr{C}^{pq,1}_{ab}+\alpha^{-}_{ij,pq}t^{p,-1}_a t^{q,-1}_b+G^{ij}_{ab}+G^{ji}_{ba} \quad (78)$$

with

$$G^{ij}_{ab}=\gamma_{ij,pq}t^{p,-1}_a t^{q,1}_b-[K(\mathscr{D}^{ij,1})_{ap}+k^{ijp}_a]t^{p,1}_b-K(\mathscr{D}^{ij,-1})_{ap}t^{p,-1}_b-\beta_{ip,1}\mathscr{C}^{pj,1}_{ab}-\beta_{ik,-1}(t^{k,-1}_a t^{j,1}_b-\tfrac{1}{3}t^{k,1}_a t^{j,-1}_b-\tfrac{2}{3}t^{j,-1}_a t^{k,1}_b)$$
$$-\beta_{it,-1}\mathscr{C}^{tj}_{ab}+s^{i,1}_a t^{j,1}_b-\tfrac{1}{3}s^{i,-1}_a t^{j,-1}_b-\tfrac{2}{3}s^{j,-1}_a t^{i,-1}_b+\left[\mathscr{T}^{ij}X+(2\mathscr{T}^{ip}-\mathscr{T}^{pi})Y^{pj}-\frac{1}{2}\mathscr{T}^{pi}Z^{pj}-(\mathscr{T}^{pi}Z^{pj})^{\dagger}\right]_{ab}. \quad (79)$$

The coefficients $\alpha^{\pm}_{ij,pq}$ and $\gamma_{ij,pq}$ are linear combinations of $\tilde{\alpha}_{ij,pq}$, $\bar{\alpha}_{ij,pq}$, and $\alpha_{ij,pq}$ given in Eqs. (31) and (32). For instance, for i, j, k, and l closed, one obtains

$$\alpha^{+}_{ij,kl}=\tfrac{1}{6}(\tilde{\alpha}_{ij,kl}-\tilde{\alpha}_{ij,lk}+\bar{\alpha}_{ij,kl}-\bar{\alpha}_{ij,lk}+2\alpha_{ij,kl}+2\alpha_{ji,lk}+\alpha_{ji,kl}+\alpha_{ij,lk}),$$
$$\alpha^{-}_{ij,kl}=\tfrac{1}{9}(\tilde{\alpha}_{ij,kl}-\tilde{\alpha}_{ij,lk}+\bar{\alpha}_{ij,kl}-\bar{\alpha}_{ij,lk}-\alpha_{ij,kl}-\alpha_{ji,lk}+\alpha_{ji,kl}+\alpha_{ij,lk}),$$
$$\gamma_{ij,kl}=\tfrac{1}{6}(\tilde{\alpha}_{ij,kl}-\tilde{\alpha}_{ij,lk}-\bar{\alpha}_{ij,kl}+\bar{\alpha}_{ij,lk}+2\alpha_{ij,kl}-2\alpha_{ji,lk}+\alpha_{ji,kl}-\alpha_{ij,lk}). \quad (80)$$

These terms require only small computational effort and are not further considered in this paper. The intermediate matrices X, Y^{pj}, and Z^{pj} are defined as

$$X_{cb}=\tfrac{1}{2}(\tilde{X}+\bar{X})_{cb},\quad X_{tb}=\bar{X}_{tb}, \quad (81)$$

$$Y^{kj}_{cb}=\tfrac{1}{4}(Y^{k\bar{j}}+Y^{\bar{k}j}+\tilde{Y}^{kj}+\bar{Y}^{kj})_{cb},\quad Y^{kj}_{ub}=\tfrac{1}{2}(Y^{\bar{k}j}+\bar{Y}^{kj})_{ub},$$
$$Y^{ti}_{cb}=\tfrac{1}{2}(Y^{t\bar{j}}+\tilde{Y}^{tj})_{cb},\quad Y^{tj}_{ub}=0, \quad (82)$$

$$Z^{kj}_{cb}=\tfrac{1}{6}(Y^{k\bar{j}}+Y^{\bar{k}j}-\tilde{Y}^{kj}-\bar{Y}^{kj}+2\tilde{Z}^{kj}+2\bar{Z}^{kj})_{cb},$$
$$Z^{kj}_{ub}=\tfrac{1}{3}(Y^{\bar{k}j}-\bar{Y}^{kj}+2\tilde{Z}^{kj})_{ub}, \quad (83)$$
$$Z^{tj}_{cb}=\tfrac{1}{3}(Y^{t\bar{j}}-\tilde{Y}^{tj}+2\tilde{Z}^{tj})_{cb},\quad Z^{tj}_{ub}=\tilde{Z}^{tj}_{ub}.$$

Explicit forms of these quantities can be obtained straightforwardly by inserting the quantities given in Eqs. (38)–(40). As examples, we present the external blocks of the matrices X, Y^{kj}, and Z^{kj},

$$X_{cb}=[F^1-A^1+2J(E^{pp})-K(E^{pp})-t^{p,\mu}t^{p,\mu\dagger}]_{cb}, \quad (84)$$

$$Y^{kj}_{cb}=[K^{kj}-\tfrac{1}{2}J^{kj}+K(E^{kj})-\tfrac{1}{2}J(E^{kj})+(2K^{kp}-K^{pk})\mathscr{B}^{pj}-(k^{kpj}-\tfrac{1}{2}k^{pkj})t^{p,1\dagger}+f^{k,1}t^{j,1\dagger}]_{cb}, \quad (85)$$

$$Z^{kj}_{cb}=[J^{kj}+J(E^{kj})-K^{pk}\mathscr{B}^{pj}-k^{pkj}t^{p,1\dagger}-\tfrac{2}{3}f^{k,-1}t^{j,-1\dagger}]_{cb}, \quad (86)$$

with

$$E^{kj}=\tfrac{1}{2}(\tilde{E}^{kj}+\bar{E}^{kj}), \quad (87)$$

$$\mathscr{B}^{pj}=\tfrac{1}{4}(\tilde{B}^{pj}+\bar{B}^{pj}+B^{jp\dagger}+B^{pj}), \quad (88)$$

$$\bar{\mathscr{B}}^{pj} = \tfrac{1}{6}(B^{jp\dagger} + B^{pj} - \tilde{B}^{pj} - \bar{B}^{pj} + 2B^{jp} + 2B^{pj\dagger}). \quad (89)$$

The forms of Eqs. (85) and (86) are identical to the closed shell ones presented in Ref. 8, except for the last term in Z^{kj}, which disappears in the closed shell case.

In the above treatment, all quantities can be expressed by the minimal set of cluster amplitudes t^p and $\mathscr{T}^{pq}$. Most numerical factors appearing in the open shell equations have been incorporated in the definitions of these matrices [cf. Eqs. (58) and (59)], which leads to equations which are very similar to the closed shell case. The only difference between the closed and open shell cases is the appearance of the quantities for $\mu = -1$. In the closed shell case, the latter quantities vanish and all equations reduce to those given in Ref. 8. As compared to a closed shell calculation, all additional terms in the equations scale in cost as the first or higher power of the number of open shell orbitals, and therefore require only small additional computation time. In particular, the number of matrix multiplications appearing in Eqs. (79), (85), and (86) is the same in the open and closed shell cases, namely, $4m^3$. Also, in both cases, only one external exchange operator $K(\mathscr{D}^{ij})$ is needed per pair. Since these are the only terms which scale with N^4m^2, their evaluation takes most of the time in calculations with large basis sets. This means that in the limit of a large number of closed shell orbitals, and for large atomic basis sets, the cost of evaluating the residuals is virtually the same as for a corresponding fully closed shell calculation. As compared to the spin-unrestricted method described in Sec. II, the reduction in computer time should be almost a factor of 3. Thus the formulation presented here may be considered to be optimally efficient. It should also be noted that the most significant savings in computational cost can be achieved by implementing only the simple equations (79) and (84)–(86), which are similar to the closed shell case. The remaining terms which involve open shell orbitals can be calculated without much overhead from the quantities explicitly given in Sec. II.

Partial reduction of the $\hat{T}$ operator towards the minimum required to span the first order space has been described by Jayatilaka and Lee.[22,23] Their approach, involving rotation of the spin space for the open shell orbitals, amounts to linearly combining the spin–orbital excitation operators, as also occurs in this work, followed by recognition that certain of the rotated operators cause excitations on Ψ_0 which do not lie in the first order space. However, in that work, the number of independent amplitudes to be solved for and the computational effort are approximately half that of the general spin-unrestricted formulation, compared with the optimum factor of $\frac{1}{3}$ which appears in the formulation which we are recommending. Furthermore, the resulting equations as given explicitly in the Appendix of Ref. 23 are extremely complex. We therefore believe that our method is not only more efficient, but also simpler to implement.

IV. TEST CALCULATIONS

In this section, we will investigate the effect of spin contamination in the unrestricted (RHF-UCCSD) and partially restricted (RCCSD) methods. At present, we are not able to compare our partially spin-restricted CCSD with a fully spin-coupled approach. However, a good indication of the importance of spin contamination can be obtained by comparing the unrestricted and partially spin-restricted methods and by computing the expectation value of $\hat{S}^2$ as defined in Eq. (42). As mentioned earlier, for linear configuration expansions (CISD), our new approach is fully spin adapted and yields identical results to other CI methods which use the first-order interacting space as a basis.[25] In our partially spin adapted RCCSD method, spin contamination can only occur via nonlinear terms in the wave function. Accordingly, the spin contamination computed from Eq. (42) is always zero for the RCISD and RCCSD methods. However, as discussed in Sec. II, spin contamination can occur in the linear terms of

TABLE I. A comparison of RCCSD and RHF-UCCSD spectroscopic constants[a] for some diatomic molecules.

Molecule	Method[b]	E_{r_e} (a.u.)	r_e (Å)	B_e	ω_e	$\omega_e x_e$
CH						
	RHF	−38.279 332 0	1.1039	14.881	3044.0	55.3
	UCCSD	−38.413 323 6	1.1183	14.499	2871.1	63.2
	RCCSD	−38.413 138 6	1.1182	14.501	2871.9	63.1
	Exp.[c]		1.1199	14.457	2858.5	63.0
NH						
	RHF	−54.977 552 4	1.0175	17.322	3548.7	66.1
	UCCSD	−55.147 078 4	1.0346	16.753	3308.9	77.4
	RCCSD	−55.146 686 7	1.0343	16.762	3313.5	76.7
	Exp.[c]		1.0362	16.699	3282.3	78.3
OH						
	RHF	−75.421 275 1	0.9494	19.727	4056.2	74.7
	UCCSD	−75.650 265 2	0.9668	19.023	3783.5	83.9
	RCCSD	−75.650 033 4	0.9667	19.027	3785.5	83.6
	Exp.[c]		0.9697	18.911	3737.8	84.9
CN						
	RHF	−92.228 237 2	1.1273	2.053	2453.8	11.8
	UCCSD	−92.564 364 4	1.1646	1.924	2146.6	12.3
	RCCSD	−92.563 161 3	1.1640	1.925	2148.6	12.7
	Exp.[c]		1.1718	1.8997	2068.6	13.1
NO						
	RHF	−129.301 490 2	1.1125	1.824	2247.6	12.0
	UCCSD	−129.724 332 3	1.1439	1.726	1986.6	13.0
	RCCSD	−129.723 730 0	1.1435	1.727	1992.1	12.8
	Exp.[c]		1.1508	1.696	1904.1	14.1
CF						
	RHF	−137.230 372 0	1.2518	1.463	1409.0	11.1
	UCCSD	−137.620 052 6	1.2705	1.420	1331.6	10.5
	RCCSD	−137.619 691 7	1.2705	1.420	1331.4	10.5
	Exp.[c]		1.2718	1.417	1308.1	11.1
O_2						
	RHF	−149.669 637 7	1.1511	1.591	2006.0	9.5
	UCCSD	−150.141 733 8	1.1970	1.471	1677.8	10.3
	RCCSD	−150.140 618 4	1.1965	1.473	1680.3	10.3
	Exp.[c]		1.2075	1.446	1580.2	12.0

[a]All values in cm^{-1} unless otherwise noted; UCCSD denotes unrestricted coupled cluster with RHF reference.
[b]All results for Dunning (Ref. 27) VQZ basis (*5s4p3d2f/4s3p2d*).
[c]Reference 28.

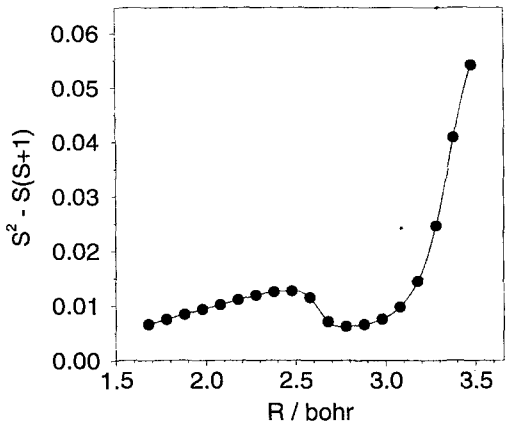

FIG. 1. Spin contamination in RHF-UCCSD calculations on CN. The deviation of the expectation value of $\hat{S}^2$ [given by Eq. (42)] from $S(S+1)$ is plotted against bond length R/bohr.

the RHF-UCISD and RHF-UCCSD wave functions, and the expectation value of $\hat{S}^2-M_S(M_S+1)$ can be used as a measure of the importance of this effect.

In order to compare the RHF-UCCSD and RCCSD methods, we have computed potential energy functions for a number of diatomic molecules. For each molecule, the energies were evaluated at ten geometries, and spectroscopic constants were obtained from polynomial fits of the eighth degree. Variations of the fit had negligible effect on the results. As bases, we used Dunning's[27] polarized correlation-consistent valence quadruple sets (5*s*4*p*3*d*2 *f* for C, N, O, and F and 4*s*3*p*2*d* for H). For the RHF-UCCSD method, the $\hat{S}^2$ expectation values as defined in Eq. (42) have also been computed.

The results are summarized in Table I. It is found that in all cases, the differences between both methods are fairly small. In most cases, the RCCSD yields energies which are less than 0.5 mhartree higher than the RHF-UCCSD method. Only for CN and O_2, larger differences of 1.2 and 1.1 mhartrees, respectively, are found. However, even in these cases, the potential energy functions are closely parallel, as indicated by the similarity of the spectroscopic constants. Even in the case of CN, where spin contamination is known to be strong in second-order unrestricted Møller–Plesset (UMP2) treatments, the ω_e values differ by less than 5 cm^{-1}. It should be noted that the intent of these calculations was not to reproduce the experimental data as accurately as possible. As is well known, connected triple excitations and also *g* functions, which were not included in the present calculations, make a significant contribution. Therefore, at present, no conclusion is possible about the question which method gives better agreement with experiment in the limit of converged basis sets. It can be concluded, however, that our new RCCSD method yields very similar results to the RHF-UCCSD method, which is computationally three times more demanding. It is very likely that the effect of the remaining spin contamination in the RCCSD wave function is negligibly small.

Figure 1 shows for the RHF-UCCSD method the approximate spin contamination $S^2-M_S(M_S+1)$ for the CN radial as a function of the bond distance R. The S^2 expectation value has been computed according to Eq. (42), and hence only the linear terms in the wave function are included. It is found that the spin contamination increases monotonically up to $R=2.5$ bohr; for longer distances, it first decreases and then steeply increases. This behavior is attributed to strong changes of the electronic structure in this region. Apparently, the RHF-UCCSD approximation works well as long as the spin adapted RHF determinant dominates the wave function, but breaks down if further open shell configurations become important.

TABLE II. Singlet–triplet splittings[a] for CH_2.

Basis[b]	RHF	UCCSD	RCCSD
VDZ (3*s*2*p*1*d*/2*s*1*p*)	25.40	12.78	12.64
VTZ (4*s*3*p*2*d*1*f*/3*s*2*p*1*d*)	25.07	11.15	10.95
VQZ (5*s*4*p*3*d*2 *f*1*g*/4*s*3*p*2*d*1*f*)	24.87	10.52	10.30
V5Z (6*s*5*p*4*d*3*f*2*g*/5*s*4*p*3*d*2 *f*)	24.79	10.34	10.12
V5Z (6*s*5*p*4*d*3*f*2*g*1*h*/5*s*4*p*3*d*2 *f*1*g*)	24.79	10.30	10.08
AVDZ (4*s*3*p*2*d*/3*d*2*p*)	24.25	11.27	11.11
AVTZ (5*s*4*p*3*d*2 *f*/4*s*3*p*2*d*)	24.69	10.58	10.37
AVQZ (6*s*5*p*4*d*3*f*2*g*/5*s*4*p*3*d*2 *f*)	24.74	10.32	10.10
AV5Z (7*s*6*p*5*d*4*f*3*g*/6*s*5*p*4*d*3*f*)	24.75	10.27	10.05

[a]All values in kcal/mol; UCCSD denotes unrestricted coupled cluster with RHF reference.
[b]Correlation consistent polarized valence basis sets of Dunning (Ref. 27). The AV*x*Z basis sets are derived from the V*x*Z basis by adding one diffuse function of each type (Ref. 29).

As another test, we computed the singlet–triplet splitting of CH_2 using both methods. The correlation consistent basis sets of Dunning[27] were employed, which allow a systematic study of the basis set convergence. From the convergence pattern seen in Table II, we expect the results obtained with the largest basis sets to be within 0.1 kcal/mol of the basis set limits. Again, the differences between the RHF-UCCSD and RCCSD methods are quite small—in fact, in most cases, smaller than the basis set error. Since the triplet state is lower in energy than the singlet state and the RCCSD correlation energy is smaller than the RHF-UCCSD correlation energy, the RCCSD splitting is about 0.2 kcal/mol smaller than the RHF-UCCSD one. The final RCCSD value of 10.0 kcal/mol is still significantly above the experimental value of 9.0 kcal/mol. The remaining error is mainly attributed to the contributions of connected triple excitations, which were not included in the present set of calculations.

Since we have not yet fully implemented the RCCSD equations, we have simulated the new method by using the RHF-UCCSD method with the restrictions in Eqs. (50)–(55). The development of a new RCCSD program is in progress at present, and more applications and timing data will be presented in a future publication.[26]

ACKNOWLEDGMENTS

The authors thank Dr. Lee, Dr. Jayatilaka, Dr. Scuseria, Dr. Pople, and Dr. Handy for stimulating discussions.

This work was supported by the Deutsche Forschungsgemeinschaft, the German Fonds der Chemischen Industrie, and the Science and Engineering Research Council.

[1] J. Čížek, J. Chem. Phys. **45**, 4526 (1966).
[2] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).
[3] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[4] G. Scuseria and H. F. Schaefer III, J. Chem. Phys. **90**, 3700 (1989).
[5] P. Pulay, S. Saebø, and W. Meyer, J. Chem. Phys. **81**, 1901 (1984).
[6] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, J. Chem. Phys. **86**, 2881 (1987).
[7] G. E. Scuseria, C. L. Janssen, and H. F. Schaefer III, J. Chem. Phys. **89**, 7382 (1988).
[8] C. Hampel, K. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).
[9] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).
[10] R. J. Bartlett, J. Chem. Phys. **93**, 1697 (1989).
[11] R. H. Nobes, J. A. Pople, L. Radom, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **138**, 481 (1987).
[12] M. Rittby and R. J. Bartlett, J. Phys. Chem. **92**, 3033 (1988).
[13] G. E. Scuseria, Chem. Phys. Lett. **176**, 27 (1991).
[14] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. A. Pople, Chem. Phys. Lett. **186**, 130 (1991).
[15] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **187**, 21 (1991).
[16] H. B. Schlegel, J. Chem. Phys. **84**, 4530 (1986).
[17] P. J. Knowles and N. C. Handy, J. Phys. Chem. **92**, 3097 (1988).
[18] C. Janssen and H. F. Schaefer III, Theor. Chim. Acta **79**, 1 (1991).
[19] H. Nakatsuji and K. Hirao, J. Chem. Phys. **68**, 2053 (1978).
[20] H. Nakatsuji, Chem. Phys. Lett. **67**, 329 (1979).
[21] H. Nakatsuji, Chem. Phys. Lett. **67**, 334 (1979).
[22] D. Jayatilaka and T. J. Lee, Chem. Phys. Lett. **199**, 211 (1992).
[23] D. Jayatilaka and T. J. Lee, J. Chem. Phys. **98**, 9734 (1993).
[24] H.-J. Werner and E.-A. Reinsch, J. Chem. Phys. **76**, 3144 (1982).
[25] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).
[26] H.-J. Werner, C. Hampel, and P. J. Knowles, J. Chem. Phys. (to be published).
[27] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[28] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
[29] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).