



ELSEVIER

14 July 2000

Chemical Physics Letters 324 (2000) 389–402

**CHEMICAL
PHYSICS
LETTERS**

www.elsevier.nl/locate/cplett

The expectation value of the spin operator \hat{S}^2 as a diagnostic tool in coupled cluster theory

The advantages of using UHF-CCSD theory for the description of homolytic dissociation

He Yuan, Dieter Cremer*

Theoretical Chemistry, University of Göteborg, Reutersgatan 2, S-41320 Göteborg, Sweden

Received 13 April 2000

Abstract

Coupled cluster (CC) theory carried out with a spin-unrestricted Hartree–Fock (UHF) reference wave function suffers less from spin contamination because infinite order electron correlation effects covered by the CC method in question reduce spin contamination. For example, UHF-CCSD, contrary to UHF, is not contaminated by a $S + 1$ state. However, the value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$, is strongly influenced by the response of the UHF-CCSD wave function on the spin contamination present at UHF and, therefore, adopts a relatively large value. A reliable diagnostic tool correctly reflecting the influence of spin contamination on the UHF-CCSD energy is obtained by just considering the energy related part of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$. © 2000 Published by Elsevier Science B.V.

1. Introduction

Methods based on an unrestricted Hartree–Fock (UHF) reference wave function benefit from the fact that UHF contrary to restricted HF (RHF) describes homolytic dissociation in many cases qualitatively correct. For the H_2 molecule one can easily show that the UHF wave function includes beside the singlet ground state function also a doubly excited singlet state function, which is needed for a description of the homolytic dissociation process. However, to link both parts together in a single determinant representation, a triplet function has to be added, which leads to the well-known spin contamination problem of UHF descriptions [1]. In other words, when utilizing the advantages of a UHF reference one has to pay as a price that the wave function is no longer an eigenfunction of the spin operator \hat{S}^2 . The UHF ground state energy is contaminated by contributions from higher spin states characterized by total spins $S + 1$, $S + 2$, etc.

Correlation corrected calculations using the UHF reference wave function show that the spin contamination problem reflected by the expectation value of \hat{S}^2 , $\langle \hat{S}^2 \rangle$, is less serious for methods covering a large amount of

* Corresponding author. Fax: +46-31-7735590; e-mail: cremer@theoc.gu.se

dynamic electron correlation [2–7]. The rule of thumb is that the more dynamic electron correlation effects are covered by a particular method, the less problematic spin contamination becomes. This explains for example why the value for $\langle \hat{S}^2 \rangle$ will generally decrease if the method is improved from UHF to unrestricted Møller–Plesset perturbation theory at second order (UMP2) [8–10], unrestricted MP perturbation theory at fourth order (UMP4), which covers single (*S*), double (*D*), triple (*T*), and disconnected quadruple (*Q*) excitations [9], unrestricted coupled cluster theory with *S* and *D* excitations (UHF-CCSD) [10–12], and to unrestricted CC theory with *D* excitations based on Brueckner orbitals (UHF-BCCD) [13] as was observed by various authors [2–5]. Also, density functional theory (DFT) based on an unrestricted Kohn–Sham approach [14], which covers an unspecified, but relatively large amount of dynamic electron correlation effects provides reasonable descriptions of high-spin cases such as doublet radicals or triplet biradicals with $\langle \hat{S}^2 \rangle$ values close to their ideal ones [15–17].

In those cases where the expectation value of \hat{S}^2 is calculated for UHF or a correlation corrected ab initio method based on UHF it was shown that $\langle \hat{S}^2 \rangle$ provides a valuable diagnostic tool by indicating the degree of spin contamination and, by this, the quality of the energy value calculated with an unrestricted method [2–7]. We will show in this work that for UHF-CCSD, $\langle \hat{S}^2 \rangle$ is of limited diagnostic value for the calculated energy. For this purpose, we will analyze $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ by splitting it up into four contributions, each of which will be tested for its usefulness when assessing the accuracy of the UHF-CCSD energy.

We will proceed by showing in Section 2 that under certain circumstances the UHF-CCSD energy is equal to the spin-projected CCSD (PCCSD) energy and that this relationship reflects the accuracy of the UHF-CCSD energy. In Section 3, we will analyze $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ and indicate how the expectation value of \hat{S}^2 relates to properties of the energy. Finally, we will make suggestions on how to use $\langle \hat{S}^2 \rangle$ as a diagnostic tool in CC theory.

2. UHF-CCSD and spin contamination

In view of the fact that highly correlated ab initio methods reduce spin contamination of the UHF wavefunction, it is reasonable to expect rather reliable energies from UHF-CC theory. In Table 1, $\langle \hat{S}^2 \rangle$ expectation values calculated at the UHF, UMP2, and UHF-CCSD level of theory for radicals at their equilibrium geometry confirm this expectation. The small deviation of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ from ideal values is a result of

Table 1
Calculation of $\langle S^2 \rangle$ for some molecules at the UHF, UMP2, and UHF-CCSD level of theory ^a.

$\langle \hat{S}^2 \rangle_{\text{Method}}$	$\text{CH}_2(^3B_1)$	$\text{CH}_2(^1A_1)$	$\text{CH}_3(^2A_2')$			$\text{NH}_2(^2B_1)$		
			$1.0R_e$	$1.5R_e$	$2.0R_e$	$1.0R_e$	$1.5R_e$	$2.0R_e$
$\langle \hat{S}^2 \rangle_{\text{UHF}}$	2.01508	0.71514	0.76129	1.06698	3.15359	0.75789	1.64140	2.54091
$\langle \hat{S}^2 \rangle_{\text{UMP2}}$	2.0050	0.72429	0.75321	0.98702	3.0394	0.75178	1.5221	2.4901
$\langle \hat{S}^2 \rangle_{\text{CCSD}}$	2.00064	0.64146	0.75049	0.77592	0.86121	0.75038	0.87946	1.84380
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}, \hat{A}(1)$	2.00003	0.001593	0.750033	0.75211	−1.54435	0.75003	0.73293	−0.68792
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}, \hat{A}(2)$	2.00002	0.004691	0.750021	0.75323	0.78115	0.75002	0.76539	0.81647
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}, \hat{A}(3)$	–	–	–	0.75324	0.51966	–	0.76542	0.81406
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}, \hat{A}(4)$	–	–	–	–	–	–	–	–
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}, \hat{A}(5)$	2.00002	0.004691	0.750021	0.75324	0.51966	0.75002	0.76542	0.81406
$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{Total}}$	2.00024	0.60102	0.75020	0.75610		0.75015		

^a Identical values are indicated by For the explanation of negative values see Appendix B. Geometry and DZP basis set were taken from Ref. [24] (CH_2), Ref. [25] (CH_3), and Ref. [26] (NH_2).

the fact that CCSD is correct in the space of all S and D excitations (SD space) and all energy contributions resulting from this space should not be affected by spin contamination. The latter should show up in contributions arising from disconnected T and Q (or higher) excitations, which are also covered by the CCSD method [18]. This leads to some interesting properties of UHF-CCSD and spin-projected UHF-CCSD (PCCSD), which help to describe the spin contamination problem in CC theory and, in particular, to answer the question under which conditions UHF-CCSD results are (un)reliable. In the following, we will derive the theory needed to establish some useful connections between CCSD and PCCSD.

In the CC approach, the exact wavefunction Ψ_{CC} of the Schrödinger equation is expressed in exponential form [9,10]:

$$\Psi_{\text{CC}} = e^{\hat{T}} |\Phi_0\rangle, \quad (1)$$

where Φ_0 is the HF reference function and \hat{T} the cluster operator. It is customary to expand \hat{T} in terms of one-, two-, and many-particle operators:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_n. \quad (2)$$

The cluster operators, \hat{T}_1 , \hat{T}_2 , etc., of Eq. (2) are expressed as

$$\hat{T}_1 = \sum_{i,a} a_i^a \hat{b}_a^\dagger \hat{b}_i, \quad (3a)$$

$$\hat{T}_2 = \frac{1}{4} \sum_{ij,ab} a_{ij}^{ab} \hat{b}_a^\dagger \hat{b}_i \hat{b}_b^\dagger \hat{b}_j, \quad (3b)$$

or in general

$$\hat{T}_n = \frac{1}{(n!)^2} \sum_{ijk\dots abc\dots} a_{ijk\dots abc\dots}^{abc\dots} \hat{b}_a^\dagger \hat{b}_i \hat{b}_b^\dagger \hat{b}_j \hat{b}_c^\dagger \hat{b}_k \dots, \quad (3c)$$

where subscripts (superscripts) i, j, k, \dots (a, b, c, \dots) denote occupied (virtual) spin orbitals in the reference function while indices p, q, r, \dots are used for general spin orbitals. The S and D amplitudes are given by a_i^a and a_{ij}^{ab} . The operators \hat{b}^\dagger and \hat{b} are creation and annihilation operators, respectively.

The Schrödinger equation for the nonrelativistic Hamiltonian \hat{H} is given by

$$\hat{H} e^{\hat{T}} |\Phi_0\rangle = E_{\text{CC}} e^{\hat{T}} |\Phi_0\rangle. \quad (4a)$$

Alternatively, Eq. (4a) can be written in the form of Eq. (4b) or (4c):

$$(\hat{H} e^{\hat{T}})_C |\Phi_0\rangle = E_{\text{CC}} |\Phi_0\rangle, \quad (4b)$$

$$(\bar{H} e^{\hat{T}})_C |\Phi_0\rangle = \Delta E_{\text{CC}} |\Phi_0\rangle, \quad (4c)$$

where the subscript C indicates connected diagrams \bar{H} refers to the normal product form of \hat{H} defined by

$$\bar{H} = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \sum_{r,s} \{ \hat{b}_r^\dagger \hat{b}_s \} \langle r | \hat{F} | s \rangle + \frac{1}{4} \sum_{r,s,t,u} \{ \hat{b}_r^\dagger \hat{b}_s^\dagger \hat{b}_t \hat{b}_u \} \langle rs | ut \rangle, \quad (5)$$

and the energy ΔE_{CC} is the correlation energy for the HF reference function:

$$\Delta E_{\text{CC}} = E_{\text{CC}} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle. \quad (6)$$

One can consider the operator $(\hat{H} e^{\hat{T}})_C$ of Eq. (4b) as an effective Hamiltonian, which upon acting on the reference function $|\Phi_0\rangle$ yields the total energy E_{CC} :

$$E_{\text{CC}} = \langle \Phi_0 | (\hat{H} e^{\hat{T}})_C | \Phi_0 \rangle = \langle \Phi_0 | \hat{H}_{\text{eff}} | \Phi_0 \rangle. \quad (7)$$

The amplitudes of \hat{T} in \hat{H}_{eff} are determined from the set of energy independent equations obtained by projection of Eq. (4b) onto the set of excitations $\langle \Phi_{ij\dots}^{ab\dots} |$,

$$\langle \Phi_{ij\dots}^{ab\dots} | (\hat{H}e^{\hat{T}})_C | \Phi_0 \rangle = 0. \quad (8)$$

For the case that only S and D excitations are considered, i.e. $\hat{T} = \hat{T}_1 + \hat{T}_2$, Eqs. (7), (8) lead to the CCSD equations:

$$E_{\text{CCSD}} = \langle \Phi_0 | (\hat{H}e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle \quad (9)$$

and

$$\langle \Phi_i^a | (\hat{H}e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle = 0, \quad (10a)$$

$$\langle \Phi_{ij}^{ab} | (\hat{H}e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle = 0. \quad (10b)$$

The expectation value of the spin operator \hat{S}^2 is given by

$$\langle \hat{S}^2 \rangle = \frac{\langle \Psi_{\text{CC}} | \hat{S}^2 | \Psi_{\text{CC}} \rangle}{\langle \Psi_{\text{CC}} | \Psi_{\text{CC}} \rangle} \quad (11)$$

which is generally not used in CC calculations because of the exponential form of the CC wave function Ψ_{CC} (see e.g. the discussion in Ref. [2]). Alternatively, one can add a perturbation $\lambda \hat{S}^2$ to the Hamiltonian \hat{H} [3],

$$\mathcal{H}(\lambda) = \hat{H} + \lambda \hat{S}^2, \quad (12)$$

and evaluate $\langle \hat{S}^2 \rangle$ as a response to an external perturbation λ (at $\lambda = 0$),

$$\langle \hat{S}^2 \rangle = \left. \frac{dE(\lambda)}{d\lambda} \right|_{\lambda=0}. \quad (13)$$

Then, the corresponding CCSD energy is obtained from (14),

$$(\mathcal{H}(\lambda)e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)})_C | \Phi_0 \rangle = E_{\text{CC}}(\lambda) | \Phi_0 \rangle, \quad (14a)$$

$$E_{\text{CCSD}}(\lambda) = \langle \Phi_0 | (\mathcal{H}(\lambda)e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)})_C | \Phi_0 \rangle, \quad (14b)$$

or in another form (when including unlinked diagrams),

$$\langle \Phi_X | \mathcal{H}(\lambda)e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle = E_{\text{CCSD}}(\lambda) \langle \Phi_X | e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle, \quad (15)$$

with $X = S, D$, $\Phi_S = \Phi_i^a$, and $\Phi_D = \Phi_{ij}^{ab}$.

When the reference function Φ_0 for an open-shell system is chosen to be an UHF wave function, spin contamination should be eliminated from the CC wave function Ψ_{CC} with the help of an appropriate spin projection operator \hat{P} [19]:

$$\hat{P} = \prod_{k \neq s} \frac{\hat{S}^2 - k(k+1)}{s(s+1) - k(k+1)}, \quad (16)$$

i.e. the spin-projected function $\Psi_{\text{PCC}} = \hat{P}e^{\hat{T}} | \Phi_0 \rangle$ rather than $\Psi_{\text{CC}} = e^{\hat{T}} | \Phi_0 \rangle$ should be used in the CC calculation. Mostly, it is easier to eliminate just the dominant spin contamination with the annihilation operator \hat{A}_{s+1} where an annihilation operator \hat{A}_{s+i} is generally defined by

$$\hat{A}_{s+i} = \frac{\hat{S}^2 - (s+i+1)(s+i)}{\langle \hat{S}^2 \rangle_0 - (s+i+1)(s+i)} \quad (17a)$$

with

$$\langle \hat{S}^2 \rangle_0 = \langle \Phi_0 | \hat{S}^2 | \Phi_0 \rangle = \langle \hat{S}^2 \rangle_{\text{HF}} \quad (17b)$$

being the HF reference value.

For the singly annihilated CCSD energy, $E_{\text{APCCSD}}^{(S+1)}$, one obtains

$$E_{\text{APCCSD}}^{(S+1)}(\lambda) = \frac{\langle \Phi_0 | \hat{A}_{s+1} \mathcal{Z}(\lambda) e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle} \quad (18a)$$

$$= \frac{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle \langle \Phi_0 | \mathcal{Z}(\lambda) e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle + \sum_X^{S,D} \langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle \langle \Phi_X | \mathcal{Z}(\lambda) e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle} \quad (18b)$$

$$= \frac{E_{\text{CCSD}}(\lambda) \left[\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle + \sum_X^{S,D} \langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle \langle \Phi_X | e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle \right]}{\langle \Phi_0 | \hat{A}_{s+1} e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle} \quad (19)$$

$$= \frac{E_{\text{CCSD}}(\lambda) \sum_X^{|\rangle, S, D} \langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle \langle \Phi_X | e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle} \quad (20)$$

$$= E_{\text{CCSD}}(\lambda), \quad (21)$$

where we have used Eqs. (14), (15) in Eq. (19).

For the special projection operator, $\hat{P}_S = \{\hat{P}_S^{S,D}, \hat{A}_{s+1}\}$ of Eq. (16), we obtain

$$E_{\text{PCCSD}}^{(S,D)}(\lambda) = \frac{\langle \Phi_0 | \hat{P}_S^{(S,D)} \mathcal{Z}(\lambda) e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle}{\langle \Phi_0 | \hat{P}_S^{(S,D)} e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} | \Phi_0 \rangle} \quad (22)$$

$$= E_{\text{CCSD}}(\lambda) \quad (23)$$

in which $\hat{P}_S^{(S,D)}$ is defined by

$$\hat{P}_S^{(S,D)} = \sum_q^{|\rangle, S, D} \hat{P}_S | \Phi_q \rangle \langle \Phi_q |. \quad (24)$$

Therefore, one can draw the conclusion that in general Eq. (25) holds:

$$E_{\text{CCSD}}(\lambda) = E_{\text{APCCSD}}^{(S+1)}(\lambda) = E_{\text{PCCSD}}^{(S,D)}(\lambda), \quad (25)$$

where the superscript (S,D) indicates that the operator \hat{P}_S acts only in the SD space. Relationship (25), which was first derived by Schlegel [3], reveals that in two situations the UHF-CCSD and PCCSD energy will be identical, namely if a) just the spin annihilation operator \hat{A}_{s+1} is applied or b) the use of the total projection operator is limited to the SD space. This is a consequence of the fact that there is no $S+1$ spin contamination at the UHF-CCSD level, e.g. in the UHF-CCSD description of the homolytic dissociation reaction of a single bond the triplet contamination is suppressed. This can be understood when considering how the triplet

contamination is reduced stepwise in the CCSD iterations. In the first steps of the CCSD iteration, low order MP correlation effects are included [18], which reduce the triplet contamination as is known for the MP2 or MP4 level of theory [2–7]. With an increasing number of iteration steps higher order correlation effects are introduced, which lead to further reduction of the triplet contamination until it completely vanishes in the converged CCSD amplitudes and energy. At this point all (sizeable) infinite order effects in the SD space are included, i.e. the CCSD energy has an FCI quality in the SD space. Remaining spin contaminations are associated with T , Q , etc. excitations, which will only be large if $S + 2$, etc. contaminants play a role.

The quality of the UHF-CCSD energy is a direct result of the total suppression of the $S + 1$ contaminant and one should expect that this quality is directly reflected by the value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$.

3. Properties of $\langle \hat{S}^2 \rangle$ at the UHF-CCSD level of theory

In Fig. 1 (see also Table 2), the expectation value $\langle \hat{S}^2 \rangle$ is shown as a function of the homolytic dissociation parameter R in the case of the FH molecule for different ab initio methods (6-31G basis, see Ref. [2]). The FH molecule was chosen since it had already been studied before by Chen and Schlegel [2] and results of this work can be used. While these authors calculated $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ in an approximate way extending a procedure first suggested by Purvis et al. [5] (PSB), we calculated the correct $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ according to Stanton, who evaluated \hat{S}^2 as a generalized CC expectation value [20]. All calculations were performed with the ab initio programs COLOGNE99 [21] and ACES II [22] using a local version of the latter to apply Stanton's method.

Compared to $\langle \hat{S}^2 \rangle_{\text{UHF}}$, which steeply increases after the RHF/UHF (R/U) instability at 1.27 Å to the limit value of 1 for large R , the $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ value increases only slowly indicating that spin contamination up to $R = 1.8$ Å is negligible. However, in the region, in which the recoupling of the FH bonding electrons from a closed shell singlet to an open-shell singlet electron pair takes place ($1.8 \leq R \leq 3$ Å), $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ increases to a

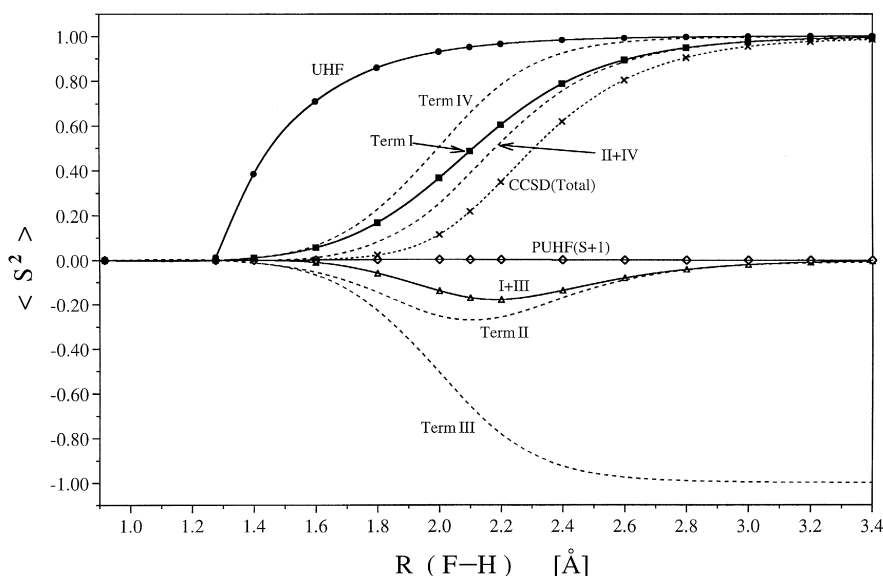


Fig. 1. Changes in $\langle \hat{S}^2 \rangle$ calculated at different levels of theory for the stretching of the F–H bond (basis set: 6-31G). For an explanation of terms I, II, III, IV, see text.

Table 2

Changes in $\langle \hat{S}^2 \rangle$ calculated at different levels of theory for the stretching of the F–H bond ^a.

$r(\text{F-H}) [\text{\AA}]$	UHF	PUHF	UMP2	UMP4	CCSD							
					\hat{A}_{s+1}	Term I	Term II	Term III	Term IV	I + III	II + IV	Total
0.917	0	0	0	0	0	0	0	0	0	0	0	0
1.276	0.0138	0	0	0	0	0	0	0	0	0	0	0
1.400	0.3849	0.0007	0.3282	0.1280	0.0112	-0.0110	-0.0117	0.0117	-0.0005	0.0007	0.0002	0.0002
1.600	0.7098	0.0020	0.6575	0.4807	0.0573	-0.0539	-0.0666	0.0666	-0.0093	0.0127	0.0034	0.0034
1.800	0.8592	0.0033	0.8264	0.7219	0.1687	-0.1452	-0.2273	0.2273	-0.0586	0.0821	0.0235	0.0235
2.000	0.9307	0.0036	0.9122	0.8560	0.3670	-0.2519	-0.5050	0.5050	-0.1380	0.2531	0.1151	0.1151
2.100	0.9513	0.0032	0.9376	0.5970	0.4872	-0.2694	-0.6563	0.6563	-0.1691	0.3869	0.2178	0.2178
2.200	0.9659	0.0027	0.9557	0.9265	0.6045	-0.2556	-0.7814	0.7814	-0.1769	0.5258	0.3489	0.3489
2.400	0.9835	0.0017	0.9778	0.9628	0.7880	-0.1687	-0.9240	0.9240	-0.1360	0.7553	0.6193	0.6193
2.600	0.9923	0.0009	0.9890	0.9813	0.8938	-0.0897	-0.9742	0.9742	-0.0804	0.8845	0.8041	0.8041
2.800	0.9967	0.0004	0.9947	0.9907	0.9481	-0.0442	-0.9901	0.9901	-0.0420	0.9459	0.9039	0.9039
3.000	0.9989	0.0002	0.9975	0.9954	0.9751	-0.0213	-0.9956	0.9956	-0.0205	0.9743	0.9538	0.9538
3.200	1.0000	0.0001	0.9989	0.9977	0.9883	-0.0128	-0.9977	0.9977	-0.0094	0.9849	0.9755	0.9755
3.400	1.0005	0	0.9996	0.0088	0.9947	-0.0097	-0.9986	0.9986	-0.0039	0.9889	0.9850	0.9850

^a For the explanation of terms I, II, III, and IV, see text.

value somewhat smaller than 1, which could be interpreted in such a way that the UHF-CCSD wave function still suffers from a triplet contamination. This seems to be in contradiction to Eq. (25) and, therefore, the value of \hat{S}^2 has to be analyzed in detail. We will do this by evaluating \hat{S}^2 as a response property [2] according to Eqs. (12), (13), however using the spin-projected CCSD wave function (26) and energy (27) rather than the corresponding UHF-CCSD quantities thus exploiting relationship (25).

$$\Psi_{\text{APCCSD}} = \hat{A}_{s+1} e^{\hat{T}_1 + \hat{T}_2} |\Phi_0\rangle, \quad (26)$$

$$E_{\text{APCCSD}} = \frac{\langle \Phi_0 | \hat{A}_{s+1} (\hat{H} e^{\hat{T}_1 + \hat{T}_2})_C |\Phi_0\rangle}{\langle \Phi_0 | \hat{A}_{s+1} |\Phi_0\rangle}. \quad (27)$$

(For the simplification of the denominator on the right side of Eq. (18a), see Ref. [23].) Hence, $\langle \hat{S}^2 \rangle_{\text{APCCSD}}$ can be written as:

$$\langle \hat{S}^2 \rangle_{\text{APCCSD}} = \left. \frac{d E_{\text{APCCSD}}^\lambda}{d \lambda} \right|_{\lambda=0} = \left. \frac{d}{d \lambda} \left(\frac{\langle \Phi_0 | \hat{A}_{s+1} e^{-(\hat{T}_1(\lambda) + \hat{T}_2(\lambda))} \hat{H}(\lambda) e^{\hat{T}_1(\lambda) + \hat{T}_2(\lambda)} |\Phi_0\rangle}{\langle \Phi_0 | \hat{A}_{s+1} |\Phi_0\rangle} \right) \right|_{\lambda=0} \quad (28a)$$

$$= \frac{\langle \Phi_0 | \hat{A}_{s+1} \left(-(\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right) e^{-(\hat{T}_1 + \hat{T}_2)} \hat{H} e^{\hat{T}_1 + \hat{T}_2} |\Phi_0\rangle}{\langle \Phi_0 | \hat{A}_{s+1} |\Phi_0\rangle} + \frac{\langle \Phi_0 | \hat{A}_{s+1} e^{-(\hat{T}_1 + \hat{T}_2)} \hat{H}^\lambda e^{\hat{T}_1 + \hat{T}_2} |\Phi_0\rangle}{\langle \Phi_0 | \hat{A}_{s+1} |\Phi_0\rangle} \\ + \frac{\langle \Phi_0 | \hat{A}_{s+1} e^{-(\hat{T}_1 + \hat{T}_2)} \hat{H} e^{\hat{T}_1 + \hat{T}_2} (\hat{T}_1^\lambda + \hat{T}_2^\lambda) |\Phi_0\rangle}{\langle \Phi_0 | \hat{A}_{s+1} |\Phi_0\rangle}. \quad (28b)$$

Using the fact that

$$e^{-\hat{T}} \hat{H} e^{\hat{T}} = (\hat{H} e^{\hat{T}})_C \quad (29)$$

one obtains Eq. (30)

$$\langle \hat{S}^2 \rangle_{\text{APCCSD}} = \frac{\langle \Phi_0 | \hat{A}_{s+1} (\hat{H}^\lambda e^{\hat{T}_1 + \hat{T}_2})_c | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle} + \frac{\langle \Phi_0 | \hat{A}_{s+1} \left(-(\hat{T}_1^\lambda + \hat{T}_2^\lambda) (\hat{H} e^{\hat{T}_1 + \hat{T}_2})_c + (\hat{H} e^{\hat{T}_1 + \hat{T}_2})_c (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right) | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle} \quad (30)$$

$$= \frac{\langle \Phi_0 | \hat{A}_{s+1} (\hat{H}^\lambda e^{\hat{T}_1 + \hat{T}_2})_c | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle} + \frac{\langle \Phi_0 | \hat{A}_{s+1} \left[(\hat{H} e^{\hat{T}_1 + \hat{T}_2})_c, (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right] | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle}. \quad (31)$$

Considering that $\hat{H}^\lambda = \hat{S}^2$, and inserting the identity $\hat{I} = \sum_X^{0,S,D,\dots} |\Phi_X\rangle \langle \Phi_X|$ after operator \hat{A}_{s+1} , the value of $\langle \hat{S}^2 \rangle_{\text{APCCSD}} = \langle \hat{S}^2 \rangle_{\text{CCSD}}$ (see Eq. (25)) can be expressed as a response property composed of the four terms I, II, III, and IV:

$$\langle \hat{S}^2 \rangle_{\text{APCCSD}} = \langle \Phi_0 | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_c | \Phi_0 \rangle \quad (32a, \text{I})$$

$$+ \langle \Phi_0 | \left[(\hat{H} e^{\hat{T}_1 + \hat{T}_2})_c, (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right] | \Phi_0 \rangle \quad (32b, \text{II})$$

$$+ \frac{\sum_X^{S,D} \langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle \langle \Phi_X | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_c | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle} \quad (32c, \text{III})$$

$$+ \frac{\sum_X^{S,D} \langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle \langle \Phi_X | \left[(\hat{H} e^{\hat{T}_1 + \hat{T}_2})_c, (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right] | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}_{s+1} | \Phi_0 \rangle}. \quad (32d, \text{IV})$$

Terms II and IV in Eqs. (32b), (32d) arise from the fact that the CCSD wave function does not fulfill the Hellmann–Feynman theorem, i.e. the perturbation $\lambda \hat{S}^2$ changes the cluster amplitudes:

$$\hat{T}_i(\lambda) = \hat{T}_i + \lambda \hat{T}_i^\lambda \quad \text{for } i = 1, 2, \quad (33)$$

where \hat{T}^λ denotes the cluster amplitude derivative with regard to the perturbation parameter λ .

The value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ given in Eqs. (32) is numerically equivalent to the generalized expectation value derived by Stanton [20] for UHF-CCSD wave functions. The advantage of Eqs. (32) is that it decomposes $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ into four terms, two of which result from the form of the energy operator (12) (CCSD energy related terms I and III) and two of which are related to the CCSD wave function (terms II and IV) because they reflect the response of the wave function on the perturbation caused by spin contamination. One could also speak of Hellmann–Feynman terms I and III and response terms II and IV.

Term I corresponds to the transition expectation value suggested by PSB [5] and can be expressed as

$$\langle \Phi_0 | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_c | \Phi_0 \rangle = \langle \Phi_0 | \hat{S}^2 | \Phi_0 \rangle + \langle \Phi_0 | \{ \hat{S}^2 \} e^{\hat{T}_1 + \hat{T}_2}_c | \Phi_0 \rangle \quad (34a)$$

$$= \langle \hat{S}^2 \rangle_0 + \langle \Phi_0 | \hat{S}^2 | \Phi_S \rangle + \langle \Phi_0 | \hat{S}^2 | \Phi_D \rangle. \quad (34b)$$

The relatively large $\langle \hat{S}^2 \rangle_{\text{UHF}}$ value is reduced by the second and the third term of Eq. (34b), which for the FH dissociation are always negative and reflect corrections introduced by the *S* and *D* cluster amplitudes [2]. As shown in Fig. 1, term I is somewhat larger than the correct value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ for all *R*, but otherwise describes the behavior of the latter for increasing *R* surprisingly well. The difference UHF – I (Fig. 1) becomes

relatively large beyond the R/U instability at 1.27 Å where the D -term in (34b) represents the dominant correction.

Actually, the large value of term I contradicts the fact that the CCSD energy does not suffer from a $S + 1$ contamination. However, term I has to be seen in connection with term III, which was directly calculated in this work. For this purpose, the matrix elements $\langle \Phi_X | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle$ ($X = S, D$) of Eqs. (32) are written in a computable form using the normal order form of the spin operator \hat{S}^2 (see Appendix A). The computational cost for evaluating terms (32a) and (32c) is $O(M^6)$ where M denotes the number of basis functions.

Term III (Fig. 1, Table 2) is comparable in absolute magnitude with term I but has opposite sign. Using the definition (17) for the annihilation operator \hat{A}_{s+1} one can easily show that the matrix element $\langle \Phi_0 | \hat{A}_{s+1} | \Phi_X \rangle$ must always be positive while the denominator of term III is either 1 or close to 1. Hence, the sign of term III is determined by the value of the matrix elements $\langle \Phi_X | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle$, which complement the corrections of $\langle \hat{S}^2 \rangle_{\text{UHF}}$ introduced by CC theory (see term I) in the SD space.

The sum of I + III is always negative, but otherwise relatively small (Fig. 1) and shows a typical dependence on the parameter $R(\text{FH})$. In the region $1.6 < R < 3$ Å, term I + III significantly deviates from 0 possessing a minimum at 2.2 Å. This is exactly the region, in which the UHF-CCSD energy deviates from the corresponding FCI energy where the deviation is largest at 2.2 Å (see, e.g. Ref. [2]). Hence, the energy related part of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ (sum of terms I and III) has the diagnostic value, which normally is associated with $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ itself. It shows that the influence of spin contamination on the UHF-CCSD energy is generally small but becomes largest in the region in which the spin recoupling process (closed-shell singlet \rightarrow open-shell singlet) takes place.

It remains to clarify why the value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ is much larger and no longer so much of a diagnostic value for the UHF-CCSD energy. For this purpose, we first derive term IV. The derivative of Eq. (14a) with respect to λ is given by Eq. (35):

$$\left\{ \left[(\hat{H} e^{\hat{T}_1 + \hat{T}_2})_C, (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right] + (\hat{H}^\lambda e^{\hat{T}_1 + \hat{T}_2})_C \right\} | \Phi_0 \rangle = E_{\text{CC}}^\lambda | \Phi_0 \rangle. \quad (35)$$

Projection from the left by $\langle \Phi_X |$ (with $X \neq 0$) leads to

$$\langle \Phi_X | (\hat{H}^\lambda e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle + \langle \Phi_X | \left[(\hat{H} e^{\hat{T}_1 + \hat{T}_2})_C, (\hat{T}_1^\lambda + \hat{T}_2^\lambda) \right] | \Phi_0 \rangle = 0. \quad (36)$$

If the perturbation $\hat{H}^\lambda = \hat{S}^2$, then from Eq. (36) one can show that terms III and IV cancel each other which holds both for the UHF-CCSD and the projected UHF-CCSD wave function provided $\hat{P}_S = \{ \hat{P}_S^{(S,D)}, \hat{A}_{s+1} \}$ is used. Once term III is calculated, IV is also known.

Term II was determined by subtracting from $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ evaluated according to Stanton [20] the sum I + III + IV = I. While term II is rather small and of comparable magnitude to the sum I + III (Fig. 1), the wave function related terms II + IV are clearly responsible for the value of $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ and its dependence on the distance $R(\text{FH})$. The relatively large values of \hat{S}^2 are a result of the response of the UHF-CCSD wave function on the spin contamination introduced by the UHF wave function. Hence, $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ still possesses a diagnostic value, however for the homolytic dissociation reaction this is predominantly, with regard to the strong response of the wave function, as a result of large spin contamination. For example, the UHF wave function has 50% triplet character for large R , which requires a major change in the UHF-CCSD wave function brought about by the S and D amplitudes and reflected by a $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ value close to 1.

4. \hat{S}^2 as a diagnostic in UHF-CC theory

In Table 1, the $\langle \hat{S}^2 \rangle$ values of the triplet biradical $\text{CH}_2(^3B_1)$ and the doublet radicals $\text{CH}_3(^2A_2')$ and $\text{NH}_2(^2B_1)$ calculated at their equilibrium geometries with a VDZP basis set are listed [24–26]. As an example for a closed shell singlet molecule with large contamination at the UHF level, the $\langle \hat{S}^2 \rangle$ value of $\text{CH}_2(^1A_1)$ is also given. Finally, the doublet radicals $\text{CH}_3(^2A_2')$ and $\text{NH}_2(^2B_1)$ with stretched geometries ($1.5 r_e$ and $2 r_e$: all

AH bonds are stretched to 1.5 and 2 times the equilibrium value) are included because breaking of two or three single bonds at the same time can lead to significant spin contamination.

For the radicals at equilibrium geometries, both the $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I}}$, the $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ and the true $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ value are close to the ideal values of 2 or 0.75 where the energy related value is clearly the closest indicating that the $S + 1$ state was the major contaminant at UHF, which is deleted at UHF-CCSD. There is some small influence from the response of the wave function reflected by the fact that $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{Total}}$ is slightly larger than $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ (Table 1).

The UHF closed shell singlet state $\text{CH}_2(^1A_1)$ possesses a large triplet contaminant as reflected by a $\langle \hat{S}^2 \rangle_{\text{UHF}}$ value of 0.71 (Table 1). At the UHF-CCSD level, $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I}}$ is 0.64 and, by this, close to the true $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ value of 0.60. However, this does not imply that the corresponding energy is unreliable since the large $\langle \hat{S}^2 \rangle_{\text{CCSD}}$ is dominated by the strong response of the UHF-CCSD wave function on the perturbation $\lambda \hat{S}^2$. The energy related part $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ is 0.004 thus indicating that the UHF-CCSD energy does not suffer from a serious spin contamination problem. This is confirmed by a S-T splitting of 12.3 kcal/mol calculated at UHF-CCSD/VDZP for CH_2 , which is close to the corresponding FCI value of 12.0 kcal/mol [24].

For the stretched geometries, $\langle \hat{S}^2 \rangle_{\text{UHF}}$ values are between 1 and 3.1 (Table 1) indicating contamination by the $S + 1$ and the $S + 2$ ($S + 3$) contaminant typical of a simultaneous breaking of two (three) AH single bonds. At UMP2, the large values for \hat{S}^2 are only slightly reduced while stronger reduction is indicated by the corresponding UHF-CCSD values. The energy related term $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ takes values at 0.75, -1.54 (CH_3), 0.73, and -0.69 (NH_2 , Table 1) indicating that the $S + 2$ contaminant is not excluded at the UHF-CCSD level. As shown in Appendix B, negative \hat{S}^2 values are a result of the action of the cluster operators \hat{T}_1 and \hat{T}_2 , which is equivalent to projecting out the $S + 1$ contaminant.

We tested whether higher contaminants can be projected out with the help of an approximate projection operator constructed from the annihilation operators,

$$\hat{A}(m) = \prod_{i=1}^m \hat{A}_{s+i}, \quad (37)$$

so that Eqs. (32) takes the form:

$$\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}} = \langle \hat{S}^2 \rangle_{\text{APCCSD}}^{(S,D)} = \langle \Phi_0 | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle + \frac{\sum_{S,D} \langle \Phi_0 | \hat{A}(m) | \Phi_y \rangle \langle \Phi_y | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle}{\langle \Phi_0 | \hat{A}(m) | \Phi_0 \rangle}, \quad (38)$$

where index m is stepwise increased from 1 to 5.

The data in Table 1 reveal that the largest effect is obtained by applying $\hat{A}(2)$ after which $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ no longer changes suggesting some $S + 2$ contamination as one should expect for simultaneous breaking of two (three) AH bonds. However, for CH_3 at $2 R_e$ application of $\hat{A}(3)$ is needed to get a constant $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ value (Table 1), which confirms that for the breaking of three AH bonds even the $S + 3$ contaminant plays a role.

Some interesting conclusions result from these calculations.

(1) UHF-CCSD is well-suited to correctly describing single bond breaking because it is free of the $S + 1$ contaminant typical of and problematic for UHF, UMP n ($n \leq 4$), and all UHF-based correlation corrected ab initio methods without infinite order effects.

(2) The quality of the UHF-CCSD energy is reflected by the energy related term $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$, which correctly indicates that region of bond breaking suffering most (recoupling region) and least (small R and large R) from spin contamination. Hence, $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{I+III}}$ rather than $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{Total}}$ should be used as a diagnostic tool for the accuracy of the UHF-CCSD energy.

(3) The response of the UHF-CCSD wave function will be large if the UHF wave function has a large spin contamination. This is reflected by the wave function related term $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{II+IV}}$, which dominates the $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{Total}}$ value. Hence, a large $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{\text{Total}}$ value as for example for large distances R in the case of the homolytic

dissociation of a single bond is simply the reflection of the fact that the UHF-CCSD wave function does not fulfill the Hellmann–Feynman theorem.

(4) The energy related term $\langle \hat{S}^2 \rangle_{\text{CCSD}}^{I+III}$ reveals that the UHF-CCSD energy will be of less accuracy if two or more bonds are broken and $S + 2$, $S + 3$, etc. contaminants are included into the wave function. The latter can be projected out by the approximate projection operator $\hat{A}(m)$, which helps to identify the most serious spin contaminations.

Since UHF-CCSD theory, contrary to UHF theory, does not suffer from the $S + 1$ contaminant, this level of theory should be appropriate for describing singlet biradicals such as p-benzyne [27]. However, if spin contaminants $S + 2$, $S + 3$, etc. become important at the UHF-CCSD level of theory (CH_3 (${}^2A'_2$) or NH_2 (2B_1) at $2R_e$, see Table 1), then the use of spin projection methods will be advisable. Alternatively, CCSDT or BCCD can be employed since these methods will also reduce the $S + 2$ contaminant. Work is in progress to investigate the performance of PCCSD methods on the one hand and CCSDT on the other hand [23].

Acknowledgements

Valuable discussions with Professor J. Gauss (Mainz, Germany), Professor John F. Stanton (Austin, USA), Dr. Zhi He (Lund, Sweden), and Dr. Jürgen Gräfenstein (Göteborg, Sweden) are acknowledged. This work was supported by the Swedish Natural Science Research Council (NFR), Stockholm, Sweden. Calculations were carried out with the supercomputers of the Nationellt Superdator Centrum (NSC) in Linköping, Sweden. D.C. thanks the NSC for a generous allotment of computer time.

Appendix A

In the following, we give explicit expressions for the matrix elements $\langle \Phi_y | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle$ ($y = S, D$) of Eq. (32c).

$$\begin{aligned} \langle \Phi_i^a | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle &= -\tilde{S}_a^i + \sum_{b,\bar{c}} S_{a\bar{c}} S_{b\bar{c}} a_i^b - \sum_{i,\bar{c}} S_{j\bar{c}} S_{i\bar{c}} a_j^a - \sum_{\bar{k},\bar{c}} S_{a\bar{c}} S_{i\bar{k}} a_{\bar{k}}^{\bar{c}} - \sum_{j\bar{b}} \tilde{S}_b^j (a_{ij}^{ab} - a_i^b a_j^a) \\ &\quad - \sum_{\bar{k},\bar{c}} \tilde{S}_c^{\bar{k}} a_{i\bar{k}}^{a\bar{c}} - \sum_{\bar{k},b\bar{c}} S_{a\bar{c}} S_{b\bar{k}} (a_{i\bar{k}}^{b\bar{c}} + a_i^b a_{\bar{k}}^{\bar{c}}) + \sum_{j\bar{k},\bar{c}} S_{j\bar{c}} S_{i\bar{k}} (a_{j\bar{k}}^{a\bar{c}} + a_j^a a_{\bar{k}}^{\bar{c}}), \quad (\text{A.1}) \\ \langle \Phi_{ij}^{ab} | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle &= \sum_P (-1)^P P(a/b) \sum_{\bar{c}d} S_{b\bar{c}} S_{d\bar{c}} a_{ij}^{ad} - \sum_P (-1)^P P(i/j) \sum_{k,\bar{c}} S_{k\bar{c}} S_{j\bar{c}} a_{ik}^{ab} \\ &\quad - \sum_P (-1)^P P(i/j|a/b) \sum_{\bar{i},\bar{d}} S_{b\bar{d}} S_{j\bar{i}} a_{i\bar{i}}^{a\bar{d}} + \sum_{l,d} \tilde{S}_d^l \left(\sum_P (-1)^P P(a/b) a_{ij}^{ad} a_l^b \right. \\ &\quad \left. + \sum_P (-1)^P P(i/j) a_{il}^{ab} a_j^d \right) - \sum_{\bar{i},\bar{c}\bar{d}} \sum_P (-1)^P P(a/b) S_{b\bar{d}} S_{c\bar{i}} \left(a_{ij}^{ac} a_{\bar{i}}^{\bar{d}} \right. \\ &\quad \left. + \sum_{P'} (-1)^{P'} P'(i/j) a_{i\bar{i}}^{a\bar{d}} a_j^c \right) + \sum_{k\bar{l},\bar{d}} \sum_P (-1)^P P(i/j) S_{k\bar{d}} S_{j\bar{l}} \left(a_{ik}^{ab} a_{\bar{l}}^{\bar{d}} \right. \\ &\quad \left. + \sum_{P'} (-1)^{P'} P'(a/b) a_{i\bar{l}}^{a\bar{d}} a_k^b \right) \\ &\quad + \sum_{k\bar{l},\bar{c}\bar{d}} S_{k\bar{d}} S_{l\bar{c}} \left[\sum_P (-1)^P P(i/j) a_{ik}^{ab} (a_{j\bar{l}}^{a\bar{d}} + a_j^c a_{\bar{l}}^{\bar{d}}) \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_P (-1)^P P(a/b) a_{ij}^{ac} (a_{kl}^{b\bar{d}} + a_k^b a_l^{\bar{d}}) \\
& - \sum_P (-1)^P P(i/j|a/b) (a_{ik}^{ac} - a_i^c a_k^a) a_{jl}^{b\bar{d}} \Big], \tag{A.2}
\end{aligned}$$

$$\begin{aligned}
\langle \Phi_{ij}^{a\bar{b}} | (\hat{S}^2 e^{\hat{T}_1 + \hat{T}_2})_C | \Phi_0 \rangle = & -S_{i\bar{b}} S_{aj} - \sum_c S_{aj} S_{c\bar{b}} a_i^c + \sum_k S_{kj} S_{i\bar{b}} a_k^a - \sum_{\bar{c}} S_{a\bar{c}} S_{i\bar{b}} a_j^{\bar{c}} + \sum_{\bar{k}} S_{aj} S_{i\bar{k}} a_{\bar{k}}^{\bar{b}} \\
& - \sum_{k,c} S_{kj} S_{c\bar{b}} (a_{ik}^{ac} - a_i^c a_k^a) - \sum_{\bar{l}, \bar{d}} S_{a\bar{d}} S_{i\bar{l}} (a_{j\bar{l}}^{b\bar{d}} - a_j^{\bar{d}} a_{\bar{l}}^{\bar{b}}) + \sum_{c, \bar{d}} S_{a\bar{d}} S_{c\bar{d}} a_{ij}^{c\bar{b}} \\
& - \sum_{k, \bar{d}} S_{k\bar{d}} S_{i\bar{d}} a_{kj}^{a\bar{b}} - \sum_{k, \bar{c}} S_{k\bar{c}} S_{k\bar{b}} a_{ij}^{a\bar{c}} + \sum_{k, \bar{l}} S_{kj} S_{k\bar{l}} a_{i\bar{l}}^{a\bar{b}} - \sum_{c, \bar{d}} S_{a\bar{d}} S_{c\bar{b}} (a_{ij}^{a\bar{c}} + a_i^c a_j^{\bar{d}}) \\
& - \sum_{k, \bar{l}} S_{kj} S_{i\bar{l}} (a_{k\bar{l}}^{a\bar{b}} + a_k^a a_{\bar{l}}^{\bar{b}}) + \sum_{\bar{l}, c} S_{aj} S_{i\bar{c}} (a_{i\bar{l}}^{c\bar{b}} + a_i^c a_{\bar{l}}^{\bar{b}}) + \sum_{k, \bar{d}} S_{k\bar{d}} S_{i\bar{b}} (a_{kj}^{a\bar{d}} + a_k^a a_j^{\bar{d}}) \\
& + \sum_{k,c} \tilde{S}_c^k (a_{ij}^{c\bar{b}} a_k^a + a_{kj}^{a\bar{b}} a_i^c) + \sum_{\bar{l}, \bar{d}} \tilde{S}_{\bar{d}}^{\bar{l}} (a_{i\bar{l}}^{a\bar{b}} a_{\bar{j}}^{\bar{d}} + a_{i\bar{j}}^{a\bar{d}} a_{\bar{l}}^{\bar{b}}) \\
& - \sum_{k, c\bar{d}} S_{k\bar{d}} S_{c\bar{b}} [a_{ij}^{a\bar{d}} a_k^c + a_j^{\bar{d}} (a_{ik}^{ac} - a_i^c a_k^a) - a_{kj}^{a\bar{d}} a_i^c - a_{ij}^{c\bar{d}} a_k^a] \\
& + \sum_{k\bar{l}, c} S_{kj} S_{c\bar{l}} [a_{i\bar{l}}^{a\bar{b}} a_k^c + a_{\bar{l}}^{\bar{b}} (a_{ik}^{ac} - a_i^c a_k^a) - a_{i\bar{l}}^{c\bar{b}} a_k^a - a_{k\bar{l}}^{a\bar{b}} a_i^c] \\
& - \sum_{\bar{l}, c\bar{d}} S_{a\bar{d}} S_{c\bar{l}} [(a_{j\bar{l}}^{b\bar{d}} - a_j^{\bar{d}} a_{\bar{l}}^{\bar{b}}) a_i^c + a_{ij}^{c\bar{b}} a_{\bar{l}}^{\bar{d}} - a_{i\bar{l}}^{c\bar{b}} a_j^{\bar{d}} - a_{ij}^{c\bar{d}} a_{\bar{l}}^{\bar{b}}] \\
& + \sum_{k\bar{l}, c} S_{k\bar{d}} S_{i\bar{l}} [(a_{j\bar{l}}^{b\bar{d}} - a_j^{\bar{d}} a_{\bar{l}}^{\bar{b}}) a_k^a + a_{kj}^{a\bar{b}} a_{\bar{l}}^{\bar{d}} - a_{k\bar{j}}^{a\bar{d}} a_{\bar{l}}^{\bar{b}} - a_{k\bar{l}}^{a\bar{b}} a_{\bar{j}}^{\bar{d}}] \\
& - \sum_{k\bar{l}, c\bar{d}} S_{k\bar{d}} S_{c\bar{l}} [(a_{ik}^{ac} - a_i^c a_k^a) (a_{j\bar{l}}^{b\bar{d}} - a_j^{\bar{d}} a_{\bar{l}}^{\bar{b}}) + a_{i\bar{l}}^{a\bar{d}} a_{k\bar{j}}^{c\bar{b}} - a_{ij}^{a\bar{d}} (a_{k\bar{l}}^{c\bar{b}} + a_k^c a_{\bar{l}}^{\bar{b}}) \\
& - a_{i\bar{l}}^{a\bar{b}} (a_{k\bar{j}}^{c\bar{d}} + a_k^c a_{\bar{j}}^{\bar{d}}) - a_{ij}^{c\bar{b}} (a_{k\bar{l}}^{a\bar{d}} + a_k^a a_{\bar{l}}^{\bar{d}}) - a_{k\bar{j}}^{a\bar{b}} (a_{i\bar{l}}^{c\bar{d}} + a_i^c a_{\bar{l}}^{\bar{d}}) \\
& + (a_{ij}^{c\bar{d}} + a_i^c a_j^{\bar{d}}) (a_{k\bar{l}}^{a\bar{b}} + a_k^a a_{\bar{l}}^{\bar{b}}) + (a_{k\bar{j}}^{a\bar{d}} + a_k^a a_{\bar{j}}^{\bar{d}}) \\
& \times (a_{i\bar{l}}^{c\bar{b}} + a_i^c a_{\bar{l}}^{\bar{b}}) - 2a_i^c a_{\bar{j}}^{\bar{d}} a_k^a a_{\bar{l}}^{\bar{b}}]. \tag{A.3}
\end{aligned}$$

A bar indicates a β spin orbital. The symbol $\sum_P (-1)^P P(a/b)$ denotes summation over the identity permutation and the permutation interchanging labels a and b. The quantities \tilde{S}_a^i and $\tilde{S}_{\bar{c}}^{\bar{k}}$ are defined by

$$\tilde{S}_a^i = \sum_{\bar{l}} S_{a\bar{l}} S_{i\bar{l}}, \tag{A.4}$$

$$\tilde{S}_{\bar{c}}^{\bar{k}} = \sum_j S_{j\bar{c}} S_{j\bar{k}}, \tag{A.5}$$

with $S_{p\bar{q}}$ being

$$S_{p\bar{q}} = \int \phi_p \hat{S}_+ \bar{\phi}_q d\tau. \tag{A.6}$$

Appendix B

In the following, it is shown why $\langle \hat{S}^2 \rangle$ may become negative after applying the annihilation operator $\hat{A}(1)$.

The UHF-CCSD wavefunction $|\Psi\rangle$ can be expanded as a linear combination of $|\phi_i\rangle$, where functions $|\phi_i\rangle$ are eigenfunctions of \hat{S}^2 :

$$|\Psi\rangle = C_0|\phi_0\rangle + C_1|\phi_1\rangle + C_2|\phi_2\rangle + \dots + C_n|\phi_n\rangle, \quad (\text{B.1})$$

$$\hat{S}^2|\phi_0\rangle = S_0(S_0 + 1)|\phi_0\rangle, \quad (\text{B.2})$$

$$\hat{S}^2|\phi_1\rangle = (S_0 + 1)(S_0 + 2)|\phi_1\rangle, \quad (\text{B.3})$$

$$\hat{S}^2|\phi_2\rangle = (S_0 + 2)(S_0 + 3)|\phi_2\rangle, \quad (\text{B.4})$$

...

Before spin annihilation, $\langle \hat{S}^2 \rangle$ is always non-negative:

$$\langle \hat{S}^2 \rangle = \langle \Psi | \hat{S}^2 | \Psi \rangle \quad (\text{B.5})$$

$$= C_0^2 \langle \phi_0 | \hat{S}^2 | \phi_0 \rangle + C_1^2 \langle \phi_1 | \hat{S}^2 | \phi_1 \rangle + C_2^2 \langle \phi_2 | \hat{S}^2 | \phi_2 \rangle + \dots + C_n^2 \langle \phi_n | \hat{S}^2 | \phi_n \rangle \quad (\text{B.6})$$

$$= C_0^2 S_0(S_0 + 1) + C_1^2 (S_0 + 1)(S_0 + 2) + C_2^2 (S_0 + 2)(S_0 + 3) + \dots \quad (\text{B.7})$$

If one inserts the annihilation operator,

$$\hat{A}_{s+1} = \frac{\hat{S}^2 - (S_0 + 1)(S_0 + 2)}{S_0(S_0 + 1) - (S_0 + 1)(S_0 + 2)}, \quad (\text{B.8})$$

into Eq. (B.5) one obtains

$$\langle \hat{S}^2 \rangle^{A_{s+1}} = \langle \Psi | \hat{S}^2 \hat{A}_{s+1} | \Psi \rangle = \langle \Psi | \hat{S}^2 | \Psi_{A_{s+1}} \rangle. \quad (\text{B.9})$$

Operator \hat{A}_{s+1} may change not only the value but also the sign of coefficients C_2, C_3, \dots, C_n :

$$\begin{aligned} \hat{A}_{s+1}|\Psi\rangle &= C_0 \hat{A}_{s+1}|\phi_0\rangle + C_1 \hat{A}_{s+1}|\phi_1\rangle + C_2 \hat{A}_{s+1}|\phi_2\rangle + \dots \\ &= C_0|\phi_0\rangle + C_2 \frac{(S_0 + 2)(S_0 + 3) - (S_0 + 1)(S_0 + 2)}{S_0(S_0 + 1) - (S_0 + 1)(S_0 + 2)}|\phi_2\rangle \\ &\quad + C_3 \frac{(S_0 + 3)(S_0 + 4) - (S_0 + 1)(S_0 + 2)}{S_0(S_0 + 1) - (S_0 + 1)(S_0 + 2)}|\phi_3\rangle + \dots \end{aligned} \quad (\text{B.10})$$

$$= C_0|\phi_0\rangle - C_2 \frac{S_0 + 2}{S_0 + 1}|\phi_2\rangle - C_3 \frac{2S_0 + 5}{S_0 + 1}C_2|\phi_3\rangle - \dots \quad (\text{B.11})$$

By substituting Eq. (B.11) into (B.9), Eq. (B.12) results:

$$\begin{aligned} \langle \hat{S}^2 \rangle^{A_{s+1}} &= \langle \Psi | \hat{S}^2 \hat{A}_{s+1} | \Psi \rangle = C_0^2 S_0(S_0 + 1) - C_2^2 \frac{S_0 + 2}{S_0 + 1} (S_0 + 1)(S_0 + 2) \\ &\quad - C_3^2 \frac{2S_0 + 5}{S_0 + 1} (S_0 + 2)(S_0 + 3) - \dots \end{aligned} \quad (\text{B.12})$$

Comparison of Eqs. (B.12), (B.7) reveals that the value of $\langle \hat{S}^2 \rangle$ after application of \hat{A}_{s+1} may become negative in the case of $S_0 = 0$ or relatively large coefficients C_2 and C_3 . Application of \hat{A}_{s+1} projects out the $S + 1$ contamination from $|\phi_1\rangle$ but can also amplify higher spin contaminants.

References

- [1] T. Amos, G.G. Hall, *Proc. R. Soc. (London)* A 263 (1961) 483.
- [2] W. Chen, H.B. Schlegel, *J. Chem. Phys.* 101 (1994) 5957.
- [3] H.B. Schlegel, *J. Phys. Chem.* 92 (1988) 3075.
- [4] H.B. Schlegel, *J. Chem. Phys.* 84 (1986) 4530.
- [5] G.D. Purvis, H. Skeino, R.J. Bartlett, *Collect. Czech. Chem. Commun.* 53 (1988) 2203.
- [6] M. Rittby, R.J. Bartlett, *J. Phys. Chem.* 92 (1988) 3033.
- [7] J.S. Andrews, D. Jauatilaka, R.G. Bone, N.C. Handy, R.D. Amos, *Chem. Phys. Lett.* 183 (1991) 423.
- [8] C. Møller, M.S. Plesset, *Phys. Rev.* 46 (1934) 618.
- [9] For a recent review, see D. Cremer in *Encyclopedia of Computational Chemistry*, H.F. Schaefer (Ed.), Wiley, New York, 1997.
- [10] P. Carsky, M. Urban, *Ab initio Calculations, Methods and Applications in Chemistry*, Lecture Notes in Chemistry 16, Springer, Heidelberg, 1980.
- [11] G.D. Purvis III, R.J. Bartlett, *J. Chem. Phys.* 76 (1982) 1910.
- [12] For a recent review, see J. Gauss in *Encyclopedia of Computational Chemistry*, H.F. Schaefer (Ed.), Wiley, New York, 1997.
- [13] N.C. Handy, J.A. Pople, M. Head-Gordon, K. Raghavachari, G.W. Trucks, *Chem. Phys. Lett.* 164 (1989) 185.
- [14] P. Hohenberg, W. Kohn, *Phys. Rev.* 136 (1964) B 864; W. Kohn, L.J. Sham, *J. Phys. Rev.* 140 (1965) A1133.
- [15] J. Wang, A.D. Becke, V. Smith, *J. Chem. Phys.* 102 (1995) 3477.
- [16] J. Pople, P.M.W. Gill, N.C. Handy, *Int. J. Quant. Chem.* 56 (1995) 303.
- [17] J. Gräfenstein, D. Cremer, to be published.
- [18] Z. He, D. Cremer, *Int. J. Quant. Chem. Quant. Chem. Symp.* 25 (1991) 43.
- [19] P.-O. Löwdin, *Phys. Rev.* 97 (1955) 1509.
- [20] J.F. Stanton, *J. Chem. Phys.* 101 (1994) 371.
- [21] E. Kraka, J. Gräfenstein, Y. He, J. Gauss, F. Reichel, L. Olsson, Z. Konkoli, Z. He, D. Cremer, Program package COLOGNE 99, Göteborg University, Göteborg, 1999.
- [22] J.F. Stanton, J. Gauss, J.D. Watts, W.J. Lauderdale, R.J. Bartlett, *Int. J. Quant. Chem. Symp.* 26 (1992) 879.
- [23] Y. He, D. Cremer, *Theor. Chem. Acc.*, submitted for publication.
- [24] C.W. Bauschlicher Jr., P.R. Taylor, *J. Chem. Phys.* 85 (1986) 6510.
- [25] C.W. Bauschlicher Jr., P.R. Taylor, *J. Chem. Phys.* 86 (1986) 5600.
- [26] J.D. Watts, J. Gauss, R.J. Bartlett, *J. Chem. Phys.* 98 (1993) 8718.
- [27] J. Gräfenstein, A. Hjerpe, E. Kraka, D. Cremer, *J. Phys. Chem. A* 104 (2000) 1748.