Andreas Köhn and David P. Tew

Additional information on J. Chem. Phys.

Towards the Hartree–Fock and coupled-cluster singles and doubles basis set limit: A study of various models that employ single excitations into a complementary auxiliary basis set

Andreas Köhn[1,a)] and David P. Tew[2]

[1]*Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany*
[2]*School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom*

In explicitly correlated coupled-cluster singles and doubles [CCSD(F12)] calculations, the basis set incompleteness error in the double excitations is reduced to such an extent that the error in the Hartree–Fock energy and the error in the single excitations become important. Using arguments from perturbation theory to systematically truncate the coupled-cluster singles and CCSD(F12) Lagrangians, a series of coupled-cluster models are proposed and studied that reduce these basis set incompleteness errors through additional single excitations into a complementary auxiliary basis. Convergence with model and size of complementary basis is rapid and there appears to be no need to go beyond second-order models. Our iterative second-order approach is a slight improvement over the existing noniterative approach, but its main advantage is that it is suitable for response theory. © *2010 American Institute of Physics.* [doi:10.1063/1.3291040]

I. INTRODUCTION

F12 explicitly correlated methods have emerged as a practical solution to the basis set problem for correlated wave function methods such as the second-order Møller–Plesset perturbation theory (MP2),[1–8] second-order multireference perturbation theory,[9] and coupled-cluster singles and doubles with noniterative triples correction [CCSD(T)].[10–17] In these methods, the slow X^{-3} basis set convergence of the standard expansion of the wave function in excited state Hartree–Fock determinants is eliminated by incorporating geminal basis functions with explicit r_{12} dependence.[18,19] The F12 geminals enter the parametrization of double excitations and are chosen to resemble the spherically averaged pair-correlation holes at short to medium r_{12} distances and, in particular, to efficiently reproduce the correlation cusp. The small set of geminals alone can recover as much as 70% of the doubles' contribution to the correlation energy[20] and the standard orbital expansion is well suited to the description of the remaining 30%. For example, at least 98% of the MP2 correlation energy is obtained using the MP2-F12 method with only a double-zeta orbital basis set (OBS) and, amazingly, the basis set error in a MP2-F12 calculation is often dominated by that of the Hartree–Fock energy.[14,21] In CCSD(T)-F12 calculations, the total basis set error also contains contributions from the parametrization of the singles and triples excitations. One of us has recently extended F12 methods to incorporate geminal basis functions into the triples excitation operator, which is necessary to accelerate the basis set convergence of the (T) energy.[17] This article is concerned with reducing the basis set error in the Hartree–Fock and singles correlation contributions to the CCSD-F12 energy.

Several approaches have been forwarded in the recent literature. In the context of dual basis MP2,[22] Wolinski and Pulay treated the Fock matrix elements between the Hartree–Fock basis and an orthogonal complementary orbital basis as a perturbation, and derived a simple second-order correction to reduce the basis set error in the Hartree–Fock energy.[23] A more elaborate analysis of such corrections was recently given by Gill and co-workers.[24]

Noga *et al.*[25] introduced an analogous approach into F12 theory, using the one-electron component of the F12 geminals for the auxiliary single excitations. They also investigated several choices of zeroth-order Fock operator. Much improved results were obtained by Adler *et al.*,[13] who expanded the auxiliary singles into the complementary auxiliary basis set (CABS) of the F12 theory, which is used primarily for the evaluation of three- and four-electron integrals using an approximate resolution of the identity.[4] Regarding the partitioning of the zeroth-order Fock operator and the precise form of the perturbative correction, a consensus seems to have been reached between the two research groups.[14,26]

In this work we extend these approaches to coupled-cluster theory, where the auxiliary singles act to reduce the basis set error in the correlation energy as well as the Hartree–Fock energy. We adopt the CABS singles approach: Although the one-electron component of the F12 geminals is suitable for improving the parametrization of singles excitations in a correlation treatment, it is inappropriate for reducing the basis set error in a Hartree–Fock wave function, which is the most important contribution (*vide infra*). A fully coupled CABS singles CCSD-F12 method would offer almost no computational saving over using the unified orbital plus complementary auxiliary (CA) basis as the Hartree–Fock basis in a CCSD-F12 calculation since N^6 contractions with three externals in the unified basis occur. Using pertur-

TABLE I. Index conventions used in this work.

$p, q, r, \ldots$	Orbitals in finite basis
$i, j, k, \ldots$	Occupied orbitals
$a, b, c, \ldots$	Virtual orbitals in finite basis
$\kappa, \lambda, \mu, \ldots$	Orbitals in formally complete basis
$\alpha', \beta', \gamma', \ldots$	Complementary virtual orbitals of formally complete basis
$a', b', c', \ldots$	CABS representation of complementary virtual orbitals

bation arguments similar to those used to derive the CCSD(F12) model,[27] we define a series of approximations to the full CABS singles CCS and CCSD(F12) methods, which we dub Sn and CCSD(F12)-Sn. We perform a systematic study of the effect of the model and the quality of the CA basis and demonstrate that our second-order models essentially eliminate the basis set error from the parametrization of the single excitations, while only involving terms that scale linearly with the size of CABS in the N^6 contractions.

II. CORRECTION MODELS

We partition the normal-ordered electronic Hamiltonian according to

$$H = E_0 + F^{(0)} + F^{(1)} + G^{(1)}, \tag{1}$$

where $E_0 = \langle 0|H|0\rangle$ is the energy of the reference determinant $|0\rangle$. Using the index conventions in Table I and the short hands $a^p_q = a^p a_q$ and $a^{pr}_{qs} = a^p a^r a_s a_q$, where a^p and a_q are creation and annihilation operators, respectively, we can express the Fock operator F as

$$F^{(0)} = f^q_p a^p_q + f^{\beta'}_a a^a_{\beta'} + f^b_{\alpha'} a^{\alpha'}_b + f^{\beta'}_{\alpha'} a^{\alpha'}_{\beta'}, \tag{2}$$

$$F^{(1)} = f^i_{\alpha'} a^{\alpha'}_i + f^{\alpha'}_i a^i_{\alpha'}, \tag{3}$$

and the fluctuation potential G as

$$G^{(1)} = \tfrac{1}{4} g^{\lambda\nu}_{\kappa\mu} a^{\kappa\mu}_{\lambda\nu}, \tag{4}$$

where $g^{\lambda\nu}_{\kappa\mu} = \langle \kappa\mu || \lambda\nu \rangle$ are antisymmetrized two-electron repulsion integrals. The superscripts denote the order in perturbation assigned to the different constituents. The first-order part of the Fock operator is just the off-diagonal block that connects the occupied orbitals and the virtual orbitals in the formally complete virtual space. The reference determinant is thus an eigenfunction of $F^{(0)}$. The off-diagonal block of F that couples the virtual orbitals of the finite primary OBS and the complementary set are retained in the zeroth-order part. We note that this partitioning of the Hamiltonian is fully compatible with the one used to derive the CCSD(F12) model.[27] In F12 theory, the formally complete set of orbitals is either absorbed into analytical two-electron integrals or approximated with the aid of an additional CABS. The latter arises when an approximate resolution of the identity is inserted to avoid three-electron and higher-order integrals.[19,28] These additional CA orbitals, as already outlined in Sec. I, can also be used to improve the description of one-electron contributions to the total energy. In the following, we will drop the subtle distinction between the formally complete

complementary set $\{\alpha'\}$ and its auxiliary basis set representation $\{a'\}$, and just use the latter notation.

The first set of models that we want to consider employs a cluster operator with only single excitation operators, $T = T_1 + T_{1'}$, where

$$T_1 = t^i_a a^a_i \tag{5}$$

denotes the usual excitations into OBS virtual orbitals and

$$T_{1'} = t^i_{a'} a^{a'}_i \tag{6}$$

are excitations into the CABS virtual orbitals. Introducing the sets of de-excitation operators

$$\Lambda_1 = \Lambda^a_i a^i_a, \quad \Lambda_{1'} = \Lambda^{a'}_i a^i_{a'} \tag{7}$$

and the short hand $\tilde{H} = e^{-T_1}(H - E_0)e^{T_1}$, we can write the CCS Lagrange functional in the form

$$\mathcal{L} = E_0 + \langle 0|(1 + \Lambda_1 + \Lambda_{1'})e^{-T_{1'}}\tilde{H}e^{T_{1'}}|0\rangle. \tag{8}$$

Clearly, this energy functional is not fully equivalent to solving the Hartree–Fock equations in the larger basis, as the expression is based on pure excitation operators and a projection approach rather than employing orbital rotations and a variational procedure. The latter, however, is difficult to combine with a coupled-cluster model which is the goal in this work. The purpose of examining Eq. (8) is rather to quantify the projection error in comparison to the improvements in the one-electron energy obtained at various orders in $F^{(1)} + G^{(1)}$.

Starting from Eq. (8), we can define a set of approximations by truncating the expression to a given order in the perturbation. $T_{1'}$ and $\Lambda_{1'}$ obviously appear in the first order in the perturbation, whereas T_1 and Λ_1 are considered zeroth order. We will denote the resulting models nth order singles correction, abbreviated as Sn. The second-order model reads

$$E_{S2} = E_0 + \langle 0|\tilde{H} + [\tilde{H}, T_{1'}]|0\rangle, \tag{9}$$

$$0 = \langle 0|a^i_a(\tilde{H} + [\tilde{H}, T_{1'}])|0\rangle, \tag{10}$$

$$0 = \langle 0|a^i_{a'}(\tilde{H} + [F^{(0)}, T_1 + T_{1'}])|0\rangle. \tag{11}$$

Truncation at third, fourth, and fifth orders defines models S3, S4, and S5, respectively, where S5 is in fact the untruncated CCS model. The explicit expressions can be found in Appendix A. An alternative definition of the series is also possible, where T_1 is considered as first order in the perturbation, which in principle is the case for CCS. However, for any higher-order CC model, all T_1 terms will contribute in the zeroth-order model (i.e., the model within the OBS space), which would necessitate to distinguish between two sets of T_1 excitations if we want to truncate these consistently. In the present work, we will not make this distinction.

We notice that in order to solve for T_1 and $T_{1'}$, all these models require an iterative procedure. As alternative, two simplifications of the above defined S2 model exist in the literature. Effectively, $T_1 = 0$ at zeroth order if the reference determinant is a Hartree–Fock solution. If we insist on T_1

=0 at any order (by omitting all terms containing them) and truncate Eq. (8) to second order, we arrive at the following energy expression and amplitude equation:

$$E_{(S2^*)} = E_0 + \langle 0|[F^{(1)}, T_{1'}]|0\rangle = E_0 + f_i^{a'} t_{a'}^i, \tag{12}$$

$$0 = \langle 0|a_{a'}^i (F^{(1)} + [F^{(0)}, T_{1'}])|0\rangle = f_{a'}^i + f_{a'}^{b'} t_{b'}^i - f_j^i t_{a'}^j. \tag{13}$$

Assuming semicanonicalization of the CABS ($f_{a'}^{b'} = \epsilon_{a'} \delta_{a'}^{b'}$) we obtain an expression $\Delta E_{(S2^*)} = \Sigma_{a'i} |f_i^{a'}|^2 / (\epsilon_i - \epsilon_{a'})$, equivalent to the (nc) approximation discussed in Ref. 26. Here, we used the acronym (S2*), where the parentheses indicate a noniterative approach, and the star indicates the absence of couplings to T_1.

It is known,[14,26] however, that coupling to T_1 is rather important (see also below) and the simplest model to allow for that is to include a few terms linear in T_1. We then arrive at the model advocated by Werner and co-workers,[13,14] which is equivalent to Noga and Šimunek's (c) approximation[26] (note that there is no energy contribution from T_1 as long as the Brillouin theorem holds),

$$\begin{aligned} E_{(S2)} &= E_0 + \langle 0|[F^{(0)}, T_1] + [F^{(1)}, T_{1'}]|0\rangle \\ &= E_0 + f_i^a t_a^i + f_i^{a'} t_{a'}^i, \end{aligned} \tag{14}$$

$$\begin{aligned} 0 &= \langle 0|a_a^i (F^{(0)} + [F^{(0)}, T_1 + T_{1'}])|0\rangle \\ &= f_a^i + f_a^{b'} t_{b'}^i + f_a^b t_b^i - f_j^i t_a^j, \end{aligned} \tag{15}$$

$$\begin{aligned} 0 &= \langle 0|a_{a'}^i (F^{(1)} + [F^{(0)}, T_1 + T_{1'}])|0\rangle \\ &= f_{a'}^i + f_{a'}^b t_b^i + f_{a'}^{b'} t_{b'}^i - f_j^i t_{a'}^j. \end{aligned} \tag{16}$$

Model (S2) corresponds to the second-order energy obtained from a Møller–Plesset perturbation treatment of the CCS equations using the partitioning in Eq. (1). [Note that (S2*) can be obtained from (S2) by neglecting the contributions from Fock matrix elements $f_{a'}^b$.]

The other series of models that we want to discuss includes T_2 cluster operators

$$T_2 = \tfrac{1}{4} t_{ab}^{ij} a_{ij}^{ab} \tag{17}$$

and explicitly r_{12}-dependent geminal functions, introduced by a $T_{2'}$ operator,

$$T_{2'} = \tfrac{1}{2} \mathcal{R}_{\alpha' b}^{ij} a_{ij}^{\alpha' b} + \tfrac{1}{4} \mathcal{R}_{\alpha' \beta'}^{ij} a_{ij}^{\alpha' \beta'}. \tag{18}$$

The latter contains matrix elements over the correlation factor

$$\mathcal{R}_{pq}^{rs} = \mathcal{S}_{rs} \langle pq | \mathcal{Q}_{12} f(r_{12}) | rs \rangle, \tag{19}$$

where $f(r_{12}) = (-1/\gamma) e^{-\gamma r_{12}}$ and γ is a length scale parameter, as specified in the result section. The projector $\mathcal{Q}_{12}$ ensures strong orthogonality of the geminal functions to all configurations accessed by the operators T_1 and T_2, as well as $T_{1'}$. The operator $\mathcal{S}$ generates the proper prefactors for singlet and triplet pairs in order to ensure the cusp conditions for these two cases.[29,30]

For a more compact notation, we use the short hand $\bar{H} = e^{-T_1 - T_2} (H - E_0) e^{T_1 + T_2}$. At second order we obtain CCSD(F12)-S2, which is equivalent to the model discussed in Refs. 30 and 31. The energy and residual expressions read

$$\begin{aligned} E_{\text{F12-S2}} &= E_0 + \langle 0|\bar{H} + [\bar{H}, T_{1'} + T_{2'}]|0\rangle \\ &\quad + \langle 0|T_{2'}^\dagger, (\bar{H} + [F^{(0)}, T_{1'} + T_{2'}])|0\rangle, \end{aligned} \tag{20}$$

$$0 = \langle 0|a_a^i (\bar{H} + [\bar{H}, T_{1'} + T_{2'}])|0\rangle, \tag{21}$$

$$0 = \langle 0|a_{ab}^{ij} (\bar{H} + [\bar{H}, T_{1'} + T_{2'}])|0\rangle, \tag{22}$$

$$0 = \langle 0|a_{a'}^i (\bar{H} + [F^{(0)}, T_{1'} + T_{2'}])|0\rangle. \tag{23}$$

The explicit expressions are given in Appendix B. Analyzing these, one finds that all steps involving CABS indices scale with N^5 at most, in particular, these steps depend only linearly on the CABS size. The most expensive term arises from $\bar{g}_{ab}^{ic'} t_{c'}^j$, see Eq. (B7), which includes terms such as $g_{ab}^{dc'} t_d^i t_{c'}^j$. This term involves an initial OV^3X step, where O is the number of occupied, V is the number of virtual, and X is the number of CABS orbitals. In addition, two new N^6 scaling terms appear in doubles projection [the last two terms in Eq. (B7)] which, however, may be evaluated with at most $O^2 V^2 X$ cost by adding the $T_{1'}$ contributions to the standard intermediates of CCSD. The CCSD(F12) model contains terms that scale as $O^3 V^2 X$ in the iterations and the additional cost of the CCSD(F12)-S2 model is therefore relatively small.

In addition, we will consider a model, in which we truncate the $T_{1'}$ terms at third order. The truncation scheme for $T_{2'}$, i.e., the (F12) approximation, is retained, however. Denoting the above defined residuals for the F12-S2 model as $\Omega_{x,\text{F12-S2}}$, we can write the F12-S3 model equations as

$$\begin{aligned} E_{\text{F12-S3}} &= E_{\text{F12-S2}} + \tfrac{1}{2} \langle 0|[[H, T_{1'}], T_{1'}]|0\rangle \\ &\quad + \langle 0|T_{2'}^\dagger, ([\bar{H} - F^{(0)}, T_{1'}])|0\rangle, \end{aligned} \tag{24}$$

$$0 = \Omega_{1,\text{F12-S2}} + \langle 0|a_a^i ([[\bar{H}, T_{1'}], \tfrac{1}{2} T_{1'} + T_{2'}])|0\rangle, \tag{25}$$

$$0 = \Omega_{2,\text{F12-S2}} + \langle 0|a_{ab}^{ij} ([[\bar{H}, T_{1'}], \tfrac{1}{2} T_{1'} + T_{2'}])|0\rangle, \tag{26}$$

$$0 = \Omega_{1',\text{F12-S2}} + \langle 0|a_{a'}^i ([\bar{H} - F^{(0)}, T_{1'}])|0\rangle. \tag{27}$$

The major difference in the S2 correction is that now a number of terms appear in which $T_{1'}$ and $T_{2'}$ couple directly. From the explicit expressions in Appendix B, it becomes evident that this leads to quite a number of additional terms, of which a few also scale with the square power of the CABS size. The overall scaling of these terms is $OV^2 X^2$ i.e., N^5. Also, a few more N^6 scaling terms appear which, similar to the F12-S2 case, do not depend on X in the N^6 step.

TABLE II. Reference basis set limit CCSD energies from fc-CCSD(F12)/aug-cc-pV6Z(unc) calculations. Optimized ae-CCSD(T)/cc-pCVTZ structures were taken from Ref. 35.

No.	Molecule		E_{HF}	ΔE_{CCSD}
1	CFN	Cyanogen fluoride	−191.786 911 5	−0.653 518 7
2	CFN	Isocyanogen fluoride	−191.670 179 9	−0.655 422 8
3	CF_2	Singlet difluoromethylene	−236.779 760 3	−0.728 870 4
6	CHF	Singlet fluoromethylene	−137.825 445 1	−0.453 129 7
9	CHN	Hydrogen cyanide	−92.915 237 7	−0.379 582 7
10	CHN	Hydrogen isocyanide	−92.899 960 3	−0.371 785 7
15	CH_2	Singlet methylene	−38.895 936 1	−0.175 522 6
20	CH_2O	Formaldehyde	−113.923 038 6	−0.445 932 5
21	CH_2O	Hydroxymethylene	−113.846 819 5	−0.440 517 0
30	CH_4	Methane	−40.217 045 8	−0.232 771 4
34	CO	Carbon monoxide	−112.790 251 8	−0.395 934 2
35	CO_2	Carbon dioxide	−187.724 439 5	−0.659 208 7
40	C_2H_2	Acetylene	−76.855 221 5	−0.346 637 0
73	FH	Hydrogen fluoride	−100.070 786 3	−0.313 886 3
74	FHO	Hypofluorous acid	−174.822 708 1	−0.579 010 8
76	FH_2N	Monofluoroamine	−155.040 971 9	−0.544 346 4
78	FNO	Nitrosyl fluoride	−228.730 468 6	−0.773 852 7
79	F_2	Difluorine	−198.773 196 3	−0.601 118 5
82	F_2O	Difluorine monoxide	−273.586 578 0	−0.865 134 5
85	HNO	Nitrosylhydride	−129.849 256 7	−0.488 156 6
91	H_2N_2	Diazene (*cis*)	−110.039 620 2	−0.459 575 9
92	H_2N_2	Diazene (*trans*)	−110.049 182 2	−0.458 847 8
93	H_2N_2	Diazene (*iso*)	−110.020 629 0	−0.448 996 8
95	H_2O	Water	−76.067 305 9	−0.297 967 1
96	H_2O_2	Hydrogen peroxide	−150.851 770 4	−0.562 166 7
97	NH_3	Ammonia	−56.224 790 4	−0.269 348 8
101	N_2	Dinitrogen	−108.992 513 4	−0.407 702 6
102	N_2O	Nitrous oxide	−183.765 661 0	−0.686 243 0
105	O_3	Ozone	−224.363 728 1	−0.825 830 5
106	H_2	Dihydrogen	−1.133 606 6	−0.040 834 6

III. TEST CALCULATIONS

The methods were implemented using automated derivation and string-based evaluation schemes, as implemented in the GECCO program, for previous applications see Refs. 17, 30, and 31. Sn and CCSD(F12)-Sn calculations were performed using the cc-pVXZ-F12 basis sets.[32] In our calculations the CABS is used for the approximate resolution of the identity (RI) and as the auxiliary basis for the single excitations. For CABS, we test the cc-pVXZ-F12-RI (Ref. 33) and the uncontracted cc-pVQZ-F12 basis sets against the uncontracted aug-cc-pV6Z basis set. We assess the performance of

our various coupled-cluster models by comparison to basis set limit HF and CCSD energies for a set of 30 small molecules, listed in Table II. These are a subset of the 106 molecules compiled by Bakowies[34] and used by Klopper *et al.*[21,35] to develop and assess explicitly correlated methods and basis sets. Optimized CCSD(T)/cc-pCVTZ structures (all electrons correlated) were taken from Ref. 35. Basis set limit energies were obtained from frozen core (fc) CCSD(F12) calculations using a partially decontracted aug-cc-pV6Z orbital basis and the TURBOMOLE program package.[36,37] The orbital, CABS, and density fitting basis sets were chosen identical to those used in Ref. 21 to determine

TABLE III. The FH molecule: deviation of Sn energies from the basis limit Hartree–Fock values for various choices of orbital and CABS, in kJ/mol pve.

Basis	CABS	SCF	(S2*)	(S2)	S2	S3	S4	S5
cc-pVDZ-F12	cc-pVDZ-F12-RI	3.730	2.621	2.582	2.584	2.493	2.493	2.493
	cc-pVQZ-F12(unc)	3.730	0.232	0.185	0.187	0.072	0.073	0.073
	aug-cc-pV6Z(unc)	3.730	0.156	0.109	0.110	−0.006	−0.005	−0.005
cc-pVTZ-F12	cc-pVTZ-F12-RI	0.910	0.586	0.562	0.562	0.539	0.539	0.539
	cc-pVQZ-F12(unc)	0.910	0.112	0.095	0.095	0.070	0.070	0.070
	aug-cc-pV6Z(unc)	0.910	0.044	0.027	0.027	0.001	0.001	0.001

TABLE IV. The reaction $2FH \rightarrow H_2 + F_2$: deviation of fc S$n$ energies from the basis limit Hartree–Fock value (+38.5 kJ/mol pve) for various choices of orbital and CABS, in kJ/mol pve.

Basis	CABS	SCF	(S2*)	(S2)	S2	S3	S4	S5
cc-pVDZ-F12	cc-pVDZ-F12-RI	0.125	0.024	−0.037	−0.036	−0.055	−0.055	−0.055
	cc-pVQZ-F12(unc)	0.125	0.063	0.000	0.001	−0.016	−0.016	−0.016
	aug-cc-pV6Z(unc)	0.125	0.076	0.013	0.014	−0.003	−0.003	−0.003
cc-pVTZ-F12	cc-pVTZ-F12-RI	−0.088	−0.015	−0.017	−0.017	−0.015	−0.015	−0.015
	cc-pVQZ-F12(unc)	−0.088	−0.011	−0.014	−0.014	−0.012	−0.012	−0.012
	aug-cc-pV6Z(unc)	−0.088	0.002	−0.002	−0.002	0.001	0.001	0.001

the basis set limit fc-MP2 correlation energies. The remaining basis set error in total energies is thus expected to be below 0.05 kJ/mol per valence electron (pve) and almost entirely from the correlation contribution.

A. The Sn series

To within a projection error, the Sn energies converge with increasing CABS and order n to the basis set limit HF energy. The relative importance of the model and the CABS basis is illustrated in Tables III and IV, where the deviations from the Hartree–Fock limit of Sn calculations with various OBS and CABS are reported for the FH molecule and its reaction to H_2 and F_2, respectively. We use this representative example for a detailed analysis preceding a subsequent discussion of performance over the test set. For the reaction (Table IV), we have used the fc approximation in the Sn calculations. To converge total energies to the HF limit, it is obviously necessary to perform all-electron calculations. For reaction energies, improvement of the core orbitals has very little effect, not more than 0.002 kJ/mol pve in our calculations.

The Sn energies converge very rapidly with the model n. For the FH molecule, Table III, the intrinsic error for the S2 model is 34 times smaller than that of uncorrected SCF, and progressing to S3 the error further reduces by a factor of 18.

At S3 the energy is essentially converged, leaving only the tiny projection error. The convergence with CABS is also rapid, provided that the correct functions are included in the CA basis. The cc-pVXZ-F12-RI basis sets do not appear to be well suited to the description of the HF state and the basis set error of a S5/cc-pVXZ-F12 calculation with cc-pVXZ-F12-RI as CABS is only two-thirds of that of a SCF/cc-pVXZ-F12 calculation, that is, the HF/cc-pVXZ-F12 energy. Using the uncontracted cc-pVQZ-F12 orbital basis as CABS reduces the HF/cc-pVDZ-F12 basis set error from 3.73 to 0.07 kJ/mol pve. The primary reason for the improvement of cc-pVQZ-F12(unc) over the cc-pVXZ-F12-RI CABS is that the number of s and p functions is substantially larger. The main difference between the cc-pVQZ-F12(unc) and aug-cc-pV6Z(unc) basis sets is the additional functions with high angular momentum quantum number. The effect of these functions is relatively small.

Comparing the (S2*), (S2), and S2 models, we see that the (S2) and S2 energies are almost identical. Obviously, (S2) includes all important second-order terms. The (S2*) energies are somewhat inferior, which indicates that the Fock matrix elements $f_a^{p'}$ should not be neglected.

The rapid convergence with the model and the size of CABS is also evident for the reaction energy, Table IV. We notice, however, that the cc-pVXZ-F12-RI CA basis per-

TABLE V. Statistical measures for the basis set error pve for 25 reaction energies computed using various fc-CCSn models (kJ/mol).

Basis/auxbasis	Model	MAD	rms	Maximum	
cc-pVDZ-F12/—	HF	0.164	0.183	0.292	96. H_2O_2
cc-pVTZ-F12/—	HF	0.035	0.035	0.115	97. NH_3
cc-pVQZ-F12/—	HF	0.004	0.005	0.015	97. NH_3
cc-pVDZ-F12/cc-pVDZ-F12-RI	(S2*)	0.019	0.022	0.043	73. FH
	(S2)	0.018	0.017	0.045	102. N_2O
	S2	0.016	0.016	0.042	102. N_2O
	S3	0.017	0.018	0.055	73. FH
cc-pVDZ-F12/cc-pVQZ-F12(unc)	(S2*)	0.015	0.020	0.063	3. CF_2
	(S2)	0.006	0.007	0.019	93. H_2N_2
	S2	0.007	0.007	0.019	93. H_2N_2
	S3	0.006	0.007	0.016	73. FH
cc-pVTZ-F12/cc-pVTZ-F12-RI	(S2*)	0.009	0.008	0.029	97. NH_3
	(S2)	0.007	0.006	0.021	97. NH_3
	S2	0.006	0.006	0.021	97. NH_3

TABLE VI. Set of 25 reactions (heats of formation with respect to H_2, CO, CO_2, N_2, and F_2) considered in this work.

1	$NCF+CO_2 \rightarrow 2CO+\frac{1}{2}N_2+\frac{1}{2}F_2$
2	$CNF+CO_2 \rightarrow 2CO+\frac{1}{2}N_2+\frac{1}{2}F_2$
3	$CF_2+CO_2 \rightarrow 2CO+F_2$
6	$CHF+CO_2 \rightarrow 2CO+\frac{1}{2}H_2+\frac{1}{2}F_2$
9	$NCH+CO_2 \rightarrow 2CO+\frac{1}{2}N_2+\frac{1}{2}H_2$
10	$CNH+CO_2 \rightarrow 2CO+\frac{1}{2}N_2+\frac{1}{2}H_2$
15	$CH_2+CO_2 \rightarrow 2CO+H_2$
20	$CH_2O \rightarrow CO+H_2$
21	$HCOH \rightarrow CO+H_2$
30	$CH_4+CO_2 \rightarrow 2CO+2H_2$
40	$C_2H_2+2CO_2 \rightarrow 4CO+H_2$
73	$FH \rightarrow \frac{1}{2}F_2+\frac{1}{2}H_2$
74	$FHO+CO \rightarrow CO_2+\frac{1}{2}F_2+\frac{1}{2}H_2$
76	$FH_2N \rightarrow \frac{1}{2}F_2+\frac{1}{2}N_2+H_2$
78	$FNO+CO \rightarrow CO_2+\frac{1}{2}F_2+\frac{1}{2}N_2$
82	$F_2O+CO \rightarrow CO_2+F_2$
91	$H_2N_2(cis) \rightarrow H_2+N_2$
92	$H_2N_2(trans) \rightarrow H_2+N_2$
93	$H_2N_2(iso) \rightarrow H_2+N_2$
95	$H_2O+CO \rightarrow CO_2+H_2$
96	$H_2O_2+2CO \rightarrow H_2+2CO_2$
85	$HNO+CO \rightarrow CO_2+\frac{1}{2}H_2+\frac{1}{2}N_2$
97	$NH_3 \rightarrow \frac{1}{2}N_2+\frac{3}{2}H_2$
102	$N_2O+CO \rightarrow CO_2+N_2$
105	$O_3+3CO \rightarrow 3CO_2$

forms much better for relative energies than for total energies. The dominant source of error for total energies is the description of the core orbitals, which is unimportant for reaction energies. We note that a recent study of Werner and co-workers[38] resulted in similar findings. Nevertheless, the cc-pVXZ-F12-RI CABS would probably benefit from additional s and p functions in the valence region for use in CABS singles methods, which would also improve the RI in F12 methods. We also notice that the inferior performance of the (S2*) method as compared to the (S2) and S2 methods is far more significant for relative energies than for total energies.

In Table V we present statistics for the basis set error from fc-Sn calculations for 25 reaction energies (see Table VI), compared to the HF basis set limit. The mean absolute deviation (MAD) of 0.164 kJ/mol pve for the HF/cc-pVDZ-F12 basis set error is reduced by a factor of approximately 10 by the fc-(S2*) method for both the cc-pVDZ-F12-RI and

cc-pVQZ-F12(unc) CABS. For the cc-pVDZ-F12-RI CABS, the basis set error is not reduced further when progressing to the S2 and S3 models, suggesting that the error is limited by the quality of the CABS. For the cc-pVQZ-F12(unc) CABS, the error reduces by a factor of 2 by moving to the (S2) or S2 models. The remaining error is essentially that of the cc-pVQZ-F12 basis and progressing to S3 does not improve matters.

To summarize, the Sn models converge rapidly and systematically to the HF energy of the combined orbital and CA basis, to within a tiny projection error. The dominant contributions are already contained in (S2) and contributions beyond S3 can safely be neglected for total energies. For relative energies, contributions beyond (S2) are inconsequential. The quality of the CABS must be sufficiently high for the implicit accuracy of the (S2), S2, and S3 models to be realized. To this end, we suggest that the cc-pVXZ-F12-RI basis sets should be reoptimized for use with CABS singles approaches.

B. The CCSD(F12)-Sn models

The $T_{1'}$ amplitudes entering the CCSD-Sn models affect the description of both the HF state and electron correlation. The difference between Sn and HF is clearly a pure one-electron effect. The correlation energy may be defined as the difference between CCSD-Sn and Sn (note that this reduces to the usual definition of correlation energy for zeroth order in $T_{1'}$). In Table VII we present the deviation of fc-CCSD(F12)-Sn energies from the basis set limit for FH, listing total and correlation contributions. The SP ansatz[26] is used with the recommended STG exponents for the cc-pVXZ-F12 basis sets. The corresponding values for the reaction $2FH \rightarrow H_2+F_2$ are given in Table VIII.

For the FH molecule, the fc-CCSD(F12)/cc-pVDZ-F12 energy deviates by 7.18 kJ/mol pve from the basis set limit CCSD energy, 3.45 kJ/mol of this is from the correlation energy and 3.73 kJ/mol due to the HF energy. The total basis set error is reduced to 4.03 kJ/mol pve when including single excitations into the aug-cc-pV6Z(unc) CABS to third order and this residual error can be decomposed into 3.47 kJ/mol pve correlation contribution and 0.56 kJ/mol pve HF error due almost entirely to the core orbitals. In contrast to the smooth convergence of the HF contribution, the error in the CCSD(F12)-Sn energies fluctuates with n due to the coupling, or lack thereof, of the CABS singles with the CCSD

TABLE VII. The FH molecule: deviation of fc CCSD(F12)-Sn energies from the basis set limit CCSD values for various choices of orbital and CABS, in kJ/mol pve.

Basis	CABS	Total energy					Correlation contribution		
		CCSD(F12)	(S2*)	(S2)	S2	S3	CCSD(F12)	S2	S3
cc-pVDZ-F12	cc-pVDZ-F12-RI	7.23	6.13	6.09	6.15	6.01	3.50	3.56	3.50
	cc-pVQZ-F12(unc)	7.23	4.28	4.23	4.28	4.13	3.50	3.55	3.51
	aug-cc-pV6Z(unc)	7.18	4.17	4.13	4.18	4.03	3.45	3.50	3.47
cc-pVTZ-F12	cc-pVTZ-F12-RI	1.81	1.49	1.47	1.49	1.48	0.90	0.93	0.94
	cc-pVQZ-F12(unc)	1.84	1.11	1.09	1.11	1.10	0.93	0.96	0.97
	aug-cc-pV6Z(unc)	1.78	0.98	0.96	0.99	0.98	0.87	0.89	0.91

TABLE VIII. The reaction $2FH \rightarrow H_2 + F_2$: deviation of fc CCSD(F12)-Sn energies from the basis set limit CCSD values for various choices of orbital and CABS, in kJ/mol pve.

Basis	CABS	Total energy					Correlation contribution		
		CCSD(F12)	(S2*)	(S2)	S2	S3	CCSD(F12)	S2	S3
cc-pVDZ-F12	cc-pVDZ-F12-RI	0.219	0.118	0.057	0.096	0.086	0.094	0.132	0.141
	cc-pVQZ-F12(unc)	0.287	0.225	0.162	0.214	0.213	0.162	0.213	0.229
	aug-cc-pV6Z(unc)	0.303	0.254	0.191	0.243	0.243	0.178	0.229	0.245
cc-pVTZ-F12	cc-pVTZ-F12-RI	−0.059	0.013	0.011	0.021	0.013	0.029	0.038	0.028
	cc-pVQZ-F12(unc)	−0.060	0.017	0.013	0.024	0.017	0.027	0.039	0.028
	aug-cc-pV6Z(unc)	−0.046	0.043	0.040	0.051	0.044	0.042	0.053	0.042

singles. In fact, the coupling present in the S2 and S3 models increases the error in correlation energy. Although the effect of CABS singles on the correlation energy is two orders of magnitude smaller than the overall reduction in the HF basis set error, it is on the same order as the difference between the Sn models and thus has a non-negligible effect on relative energies. This is evident from Table VIII, where, for this example, the coupling increases the basis set error.

In Table IX we present statistics for the basis set error for our set of 25 reaction energies using the CCSD(F12)-Sn models and various orbital and CA basis sets. The errors are an order of magnitude larger than those associated with the HF contribution (cf. Table V). All CABS singles models studied practically eliminate the HF basis set error, provided that the CABS basis is properly chosen. The difference between the performance of the CCSD(F12)-Sn models results from the differing correlation contributions. The corresponding basis set errors for the correlation contribution to the reaction energies are listed in Table X.

For the cc-pVDZ-F12 basis and both considered auxiliary basis sets, we indeed find—for the present set of examples—a slight improvement in the correlation energy due to the S2 model. For the cc-pVTZ-F12 basis, the S2 model does not improve the correlation energy, but with a MAD of less than 0.003 kJ/mol pve from the CCSD(F12)

correlation energy, one cannot reasonably claim that it worsens the correlation energy either, which means that S2 is an equivalent replacement for (S2) whenever the focus is on response properties. Inspecting Table VIII one finds that the deviations due to less complete CA basis sets are larger by nearly an order of magnitude. The MADs in the correlation energy errors for the S3 model, however, are inferior to those of S2, although the maximum error (which occurs for FH) is reduced. In order to investigate whether the poorer performance of the S3 correlation energies is due to an unconverged CABS, we calculated the MAD between reaction energies with either cc-pVDZ-F12-RI or the uncontracted cc-pVQZ-F12 basis as CABS and cc-pVDZ-F12 as OBS. While we see a slight increase going from CCSD(F12) (MAD =0.01 kJ/mol pve) to CCSD(F12)-S2 (MAD=0.03 kJ/mol pve), the result for CCSD(F12)-S3 is exactly the same as for CCSD(F12)-S2. In other words, there is a slightly increased sensitivity toward CABS for the Sn corrected models, but both S2 and S3 seem equally sensitive. Thus, the slightly worse results of S3 for the correlation contribution must have a different origin, most likely due to an imbalance between $T_{1'}$ (up to third order) and $T_{2'}$ (up to second order).

IV. CONCLUSIONS

We investigated a series of coupled-cluster theory based models that make use of single excitations into an auxiliary

TABLE IX. Statistical measures for the fc-CCSD basis set error of 25 reaction energies in J/mol pve. $\gamma=0.9$ for DZ and $\gamma=1.0$ for TZ.

Method	MAD	rms	Maximum	
cc-pVDZ-F12/cc-pVQZ-F12(unc)				
CCSD(F12)	0.182	0.208	0.399	93. H_2N_2
CCSD(F12)-(S2)	0.104	0.094	0.308	93. H_2N_2
CCSD(F12)-S2	0.085	0.080	0.243	93. H_2N_2
CCSD(F12)-S3	0.101	0.116	0.213	73. FH
cc-pVDZ-F12/cc-pVDZ-F12-RI				
CCSD(F12)	0.186	0.213	0.406	2. CFN
CCSD(F12)-(S2)	0.116	0.101	0.336	93. H_2N_2
CCSD(F12)-S2	0.095	0.082	0.262	93. H_2N_2
CCSD(F12)-S3	0.118	0.131	0.235	2. CFN
cc-pVTZ-F12/cc-pVTZ-F12-RI				
CCSD(F12)	0.039	0.035	0.116	97. NH_3
CCSD(F12)-(S2)	0.014	0.012	0.060	93. H_2N_2
CCSD(F12)-S2	0.017	0.014	0.063	93. H_2N_2

TABLE X. Statistical measures for the basis set error in the correlation contribution to 25 fc-CCSD reaction energies in kJ/mol pve. $\gamma=0.9$ for DZ and $\gamma=1.0$ for TZ.

Method	MAD	rms	Maximum	
cc-pVDZ-F12/cc-pVQZ-F12(unc)				
CCSD(F12)	0.102	0.094	0.290	93. H_2N_2
CCSD(F12)-S2	0.084	0.081	0.229	93. H_2N_2
CCSD(F12)-S3	0.105	0.122	0.229	73. FH
cc-pVDZ-F12/cc-pVDZ-F12-RI				
CCSD(F12)	0.104	0.092	0.291	93. H_2N_2
CCSD(F12)-S2	0.084	0.074	0.220	93. H_2N_2
CCSD(F12)-S3	0.111	0.126	0.223	2. CFN
cc-pVTZ-F12/cc-pVTZ-F12-RI				
CCSD(F12)	0.010	0.012	0.045	93. H_2N_2
CCSD(F12)-S2	0.013	0.014	0.048	93. H_2N_2

basis, which is chosen complementary to the underlying smaller basis set for which the Hartree–Fock equations have been solved. Based on a partitioning of the Hamiltonian that is equivalent to that used in the definition of the approximate explicitly correlated coupled-cluster model CCSD(F12),[27] we defined a series of iterative models termed Sn (n =2,3,4,5). The approach extends recent work by Werner and co-workers[13,14] and Noga and Šimunek,[26] who investigated two related noniterative models, here referred to as (S2) and (S2*). In the latter model the coupling between virtual orbitals and the CA basis is omitted. If solved simultaneously with the CCSD(F12) equations, the proposed extensions are termed CCSD(F12)-Sn.

We also present a set of benchmark results for the HF and CCSD basis set limit, which were used to evaluate the above models. Due to the projective nature of the coupled-cluster type equations, the Sn models feature a tiny projection error which amounts to less than 0.01 kJ/mol pve. To within this projection error, we find a nice convergence of the error in the total one-electron energy. For relative energies, the convergence is less pronounced, rather we observe that the improvement is mainly determined by the completeness of the CA basis. In this respect, the newly developed Petersen CA basis sets[33] seem not optimal for the Sn corrections. Although they perform reasonably for relative energies, they are not suited for obtaining corrections for total energies. Our results suggest that a few additional functions in the s and p set of these CA basis sets would improve performance for both total and relative energies.

The noniterative HF correction of Adler et al.,[13] here denoted (S2), performs basically as well as our second-order model S2. Also, we agree with Noga and Šimunek[26] that the (S2*) model is inferior to (S2). Compared to the remaining correlation error, both S2 and (S2) are sufficiently accurate.

While the CABS singles in the (S2) correction do not couple to the CCSD(F12) equations, this is the case in the CCSD(F12)-S2 model. Our benchmark results indicate that this coupling slightly improves the one-electron contributions to the correlation energy, when smaller basis sets are used. Moving to the CCSD(F12)-S3 model seems not to pay, according to our test set. Despite the significant increase in computational cost, the results are inferior to the S2 models.

For most cases the performance of CCSD(F12)-S2 is comparable to its noniterative variant, i.e., the (S2) energy correction added to the CCSD(F12) energy. In terms of computation time, CCSD(F12)-S2 is not competitive. Its main advantage, however, is that a well defined energy Lagrangian exists and that it thus lends itself easily to the formulation of response theory, as demonstrated previously.[30,31]

ACKNOWLEDGMENTS

APPENDIX A: EXPLICIT EXPRESSIONS FOR THE Sn MODELS

We use the following intermediates for brevity: The T_1-transformed Hamiltonian $e^{-T_1}He^{T_1}$ can be partitioned into two types of effective one-electron operators with matrix elements

$$\tilde{f}'^p_q = (e^{-T_1}Fe^{T_1})^p_q \tag{A1}$$

and

$$\tilde{f}''^p_q = (e^{-T_1}Ge^{T_1})^p_q, \tag{A2}$$

and a two-electron operator with matrix elements

$$\tilde{g}^{pr}_{qs} = (e^{-T_1}Ge^{T_1})^{pr}_{qs}. \tag{A3}$$

The distinction between $\tilde{F}'$ and $\tilde{F}''$ is necessary, as the second one is already first order according to our partitioning scheme, Eq. (1). Whenever this distinction is not necessary, we will use $\tilde{F}=\tilde{F}'+\tilde{F}''$. More explicitly, the matrix elements of the Fock-type operators read

$$\tilde{f}'^i_j = f^i_j + f^i_j t^i_c, \tag{A4}$$

$$\tilde{f}'^a_i = f^a_i, \tag{A5}$$

$$\tilde{f}'^i_a = f^i_a + f^c_a t^i_c - f^i_k t^k_a - f^c_k t^k_a t^i_c, \tag{A6}$$

$$\tilde{f}'^a_b = f^a_b - f^a_k t^k_b, \tag{A7}$$

and

$$\tilde{f}''^i_j = g^{ic}_{jk} t^k_c + g^{dc}_{jk} t^k_c t^i_d, \tag{A8}$$

$$\tilde{f}''^a_i = g^{ac}_{ik} t^k_c, \tag{A9}$$

$$\tilde{f}''^i_a = g^{ic}_{ak} t^k_c + g^{dc}_{ak} t^k_c t^i_d - g^{ic}_{lk} t^k_c t^l_a - g^{cd}_{kl} t^k_c t^l_a t^i_d, \tag{A10}$$

$$\tilde{f}''^a_b = g^{ac}_{bk} t^k_c - g^{ac}_{lk} t^k_c t^l_b. \tag{A11}$$

As a generalization of the above definition, the virtual indices of the matrix elements may also refer to CABS orbitals. However, only the similarity transformation with T_1 (not $T_{1'}$) is considered, thus, e.g., the third and fourth terms in Eq. (A6) will not be present in the definition of $\tilde{f}'^i_{a'}$.

The energy expressions for S2 and S3 are

$$E_{S2} = E_0 + f^a_i t^i_a + f^{a'}_i t^i_{a'} + \tfrac{1}{2}g^{ab}_{ij}t^i_a t^j_b + g^{ab'}_{ij} t^i_a t^j_{b'}, \tag{A12}$$

$$E_{S3} = E_{S2} + \tfrac{1}{2}g^{a'b'}_{ij} t^i_{a'} t^j_{b'}. \tag{A13}$$

For S4 and S5 the energy expression is the same as for S3. Likewise, the OBS singles residual only differs for S2 and S3,

$$[\Omega_{S2}]^i_a = \tilde{f}^i_a + \tilde{f}'^i_a t^i_{c'} + \tilde{g}'^i_{ka} t^k_{c'}, \tag{A14}$$

$$[\Omega_{S3}]_a^i = [\Omega_{S2}]_a^i + \widetilde{g}_{ka}^{c'd'} t_{c'}^k t_{d'}^i, \tag{A15}$$

while again for S4 and S5 the same expression is obtained as for S3. All models differ, however, for the CABS singles residual,

$$[\Omega_{S2}]_{a'}^i = \widetilde{f}_{a'}^i - \widetilde{f}_{a'}^{\prime i} + \widetilde{f}_{c'}^{\prime a'} t_{c'}^i, \tag{A16}$$

$$[\Omega_{S3}]_{a'}^i = [\Omega_{S2}]_{a'}^i - \widetilde{f}_k^{\prime \prime i} t_{a'}^k + + \widetilde{f}_{c'}^{\prime a'} t_{c'}^i + \widetilde{g}_{ak}^{ic'} t_{c'}^k, \tag{A17}$$

$$[\Omega_{S4}]_{a'}^i = [\Omega_{S3}]_{a'}^i - (f_l^{c'} t_c^i + \widetilde{g}_{kl}^{c'i} t_{c'}^k) t_{a'}^l + g_{ka'}^{c'd'} t_{c'}^k t_{d'}^i$$
$$- \widetilde{f}_k^{\prime \prime c'} t_{a'}^k t_{c'}^i, \tag{A18}$$

$$[\Omega_{S5}]_{a'}^i = [\Omega_{S4}]_{a'}^i - g_{kl}^{c'd'} t_{a'}^k t_{c'}^i t_{d'}^l. \tag{A19}$$

APPENDIX B: DEFINITION OF THE CCSD(F12)-S2 AND CCSD(F12)-S3 MODELS

We will only discuss the additional terms due to the S2 and S3 corrections; the explicit expressions for CCSD(F12) can be found in literature, e.g., Refs. 10 and 12. In addition to the above defined intermediates, the following definitions are necessary:

$$V_{pq}^{rs} = \mathcal{S}_{rs}\langle pq | r_{12}^{-1} \mathcal{Q}_{12} f(r_{12}) | rs \rangle, \tag{B1}$$

$$V_q^p = \sum_j V_{qj}^{pj}, \quad [V^\dagger]_q^p = \sum_j V_{pj}^{qj}. \tag{B2}$$

The V intermediate is one of several special intermediates that appear in the F12 theory.[28] For a more compact notation, we will use the following T_1-transformed matrix elements of the transposed V operator

$$\widetilde{[V^\dagger]}_i^{a'} = [V^\dagger]_i^{a'} + [V^\dagger]_{ik}^{a'} t_c^k + [\mathcal{R}^\dagger]_{ik}^{d'b} g_{d'l}^{a'k} t_b^l$$
$$- [\mathcal{R}^\dagger]_{ik}^{d'c} g_{d'l}^{a'e} t_e^k t_c^l. \tag{B3}$$

We recall that both $T_{1'}$ and $T_{2'}$ are considered first order. With that, the CCSD(F12)-S2 energy becomes

$$E_{F12\text{-}S2} = E_{F12} + f_i^{a'} t_{a'}^i + g_{ij}^{ab'} t_a^i t_{b'}^j + \frac{1}{2} \mathcal{R}_{ij}^{ab'} f_k^c t_{b'}^k t_{ac}^{ij}. \tag{B4}$$

Due to the Brillouin condition, the last term will usually not contribute. The equations read

$$[\Omega_{F12\text{-}S2}]_a^i = [\Omega_{F12}]_a^i + \widetilde{f}_a^{c'} t_{c'}^i + \widetilde{g}_{ak}^{ic'} t_{c'}^k + g_{kl}^{c'd} t_{c'}^k t_{ad}^{il}$$
$$- \frac{1}{2} g_{kl}^{c'd} t_{c'}^i t_{ad}^{kl}, \tag{B5}$$

$$[\Omega_{F12\text{-}S2}]_{a'}^i = \widetilde{f}_{a'}^i - \widetilde{f}_k^{\prime i} t_{a'}^k + f_{a'}^{c'} t_{c'}^i + \frac{1}{2} g_{ka'}^{cd} t_{cd}^{ki} + f_k^c R_{ca'}^{ki}$$
$$+ f_k^{c'} \mathcal{R}_{c'a'}^{ki}, \tag{B6}$$

$$[\Omega_{F12\text{-}S2}]_{ab}^{ij} = [\Omega_{F12}]_{ab}^{ij} + P(i|j)$$
$$\times \left(\widetilde{g}_{ab}^{ic'} t_{c'}^j - \widetilde{f}_k^{c'} t_{c'}^j t_{ab}^{ik} - \widetilde{g}_{kl}^{c'j} t_{c'}^k t_{ab}^{il} + \frac{1}{2} \widetilde{g}_{kl}^{c'j} t_{c'}^i t_{ab}^{kl} \right)$$
$$+ P(a|b) \widetilde{g}_{kb}^{c'd} t_{c'}^k t_{ad}^{ij} + P(i|j)P(a|b) \widetilde{g}_{bl}^{c'd} t_{c'}^i t_{ad}^{il}. \tag{B7}$$

Here, we used the permutation operators $P(p|q)$ which generate all possible permutations of indices left and right of the bar, accompanied with the proper sign change for parity, e.g., $P(i|j) \widetilde{g}_{ab}^{ic'} t_{c'}^j = \widetilde{g}_{ab}^{ic'} t_{c'}^j - \widetilde{g}_{ab}^{jc'} t_{c'}^i$.

Finally, the equation for the CCSD(F12)-S3 energy is

$$E_{F12\text{-}S3} = E_{F12\text{-}S2} + \frac{1}{2} g_{ij}^{a'b'} t_{a'}^i t_{b'}^j$$
$$- \left(\frac{1}{2}[\mathcal{R}^\dagger]_{ij}^{a'b} g_{kb}^{ij} + \frac{1}{2}[\mathcal{R}^\dagger]_{ij}^{a'b'} g_{kb'}^{ij} + \widetilde{[V^\dagger]}_k^{a'} \right) t_{a'}^k$$
$$- [\mathcal{R}^\dagger]_{ij}^{ab'} (g_{kl}^{cj} t_{b'}^l + g_{kb'}^{cd'} t_{d'}^j) t_{ac}^{ik} + \frac{1}{2}[\mathcal{R}^\dagger]_{ij}^{ab'} (\widetilde{f}_k^{\prime \prime c} t_{b'}^k$$
$$+ g_{b'k}^{cd'} t_{d'}^k) t_{ac}^{ij} + \frac{1}{4}([\mathcal{R}^\dagger]_{ij}^{ab'} g_{ka}^{cd} t_{b'}^k$$
$$- [\mathcal{R}^\dagger]_{ij}^{a'b'} g_{ka}^{cd} t_{b'}^k) t_{cd}^{ij}. \tag{B8}$$

For the amplitude equations, we obtain

$$[\Omega_{F12\text{-}S3}]_a^i = [\Omega_{F12\text{-}S2}]_a^i, \tag{B9}$$

$$[\Omega_{F12\text{-}S3}]_{a'}^i = [\Omega_{F12\text{-}S2}]_{a'}^i + V_{a'}^i - \frac{1}{2} \widetilde{g}_{kl}^{id} \mathcal{R}_{a'd}^{kl} - \frac{1}{2} \widetilde{g}_{kl}^{id'} \mathcal{R}_{a'd'}^{kl}$$
$$+ \widetilde{f}_k^{c'} \mathcal{R}_{a'c'}^{ik} - \widetilde{f}_k^{\prime \prime c} \mathcal{R}_{a'c}^{ik} - \widetilde{f}_k^{\prime \prime i} t_{a'}^k + f_{a'}^{c'} t_{c'}^i$$
$$+ \widetilde{g}_{a'k}^{ic'} t_{c'}^k + \frac{1}{2} g_{kl}^{cd} t_{a'}^l t_{cd}^{ki}, \tag{B10}$$

$$[\Omega_{F12\text{-}S3}]_{ab}^{ij} = [\Omega_{F12\text{-}S2}]_{ab}^{ij} + P(i|j)P(a|b)(\widetilde{g}_{kb}^{c'd'} t_{d'}^j \mathcal{R}_{ac'}^{ik}$$
$$+ \widetilde{g}_{kb}^{c'd'} t_{c'}^j \mathcal{R}_{ad'}^{ik}) + P(i|j)\left(\widetilde{g}_{ab}^{c'd'} t_{c'}^i t_{d'}^j \right.$$
$$\left. - g_{kl}^{c'd'} t_{c'}^k t_{d'}^j t_{ab}^{il} + \frac{1}{4} g_{kl}^{c'd'} t_{c'}^i t_{d'}^j t_{ab}^{kl} \right). \tag{B11}$$

[1] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[2] A. J. May, E. F. Valeev, R. Polly, and F. R. Manby, Phys. Chem. Chem. Phys. **7**, 2710 (2005).
[3] D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[4] W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).
[5] H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[6] G. Knizia and H.-J. Werner, J. Chem. Phys. **128**, 154103 (2008).
[7] S. Höfener, F. A. Bischoff, A. Glöß, and W. Klopper, Phys. Chem. Chem. Phys. **10**, 3390 (2008).
[8] F. A. Bischoff, S. Höfener, A. Glöß, and W. Klopper, Theor. Chem. Acc. **121**, 11 (2008).
[9] S. Ten-no, Chem. Phys. Lett. **447**, 175 (2007).
[10] D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).
[11] D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, Phys. Chem. Chem. Phys. **10**, 6325 (2008).
[12] D. P. Tew, W. Klopper, and C. Hättig, Chem. Phys. Lett. **452**, 326 (2008).
[13] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[14] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[15] D. Bokhan, S. Bernadotte, and S. Ten-no, Chem. Phys. Lett. **469**, 214 (2009).

[16] E. F. Valeev and T. D. Crawford, J. Chem. Phys. **128**, 244113 (2008).

[17] A. Köhn, J. Chem. Phys. **130**, 131101 (2009).

[18] W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985).

[19] W. Klopper and W. Kutzelnigg, Chem. Phys. Lett. **134**, 17 (1987).

[20] S. Höfener, D. P. Tew, W. Klopper, and T. Helgaker, Chem. Phys. **356**, 25 (2009).

[21] F. A. Bischoff, S. Wolfsegger, D. P. Tew, and W. Klopper, Mol. Phys. **107**, 963 (2009).

[22] R. Jurgens-Lutovsky and J. Almlöf, Chem. Phys. Lett. **178**, 451 (1991).

[23] K. Wolinski and P. Pulay, J. Chem. Phys. **118**, 9497 (2003).

[24] J. Deng, A. Gilbert, and P. Gill, J. Chem. Phys. **130**, 231101 (2009).

[25] J. Noga, S. Kedžuch, and J. Šimunek, J. Chem. Phys. **127**, 034106 (2007).

[26] J. Noga and J. Šimunek, Chem. Phys. **356**, 1 (2009).

[27] H. Fliegl, W. Klopper, and C. Hättig, J. Chem. Phys. **122**, 084107 (2005).

[28] W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).

[29] S. Ten-no, J. Chem. Phys. **121**, 117 (2004).

[30] A. Köhn, J. Chem. Phys. **130**, 104104 (2009).

[31] M. Hanauer and A. Köhn, J. Chem. Phys. **131**, 124118 (2009).

[32] K. A. Peterson, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **128**, 084102 (2008).

[33] K. E. Yousaf and K. A. Peterson, J. Chem. Phys. **129**, 184108 (2008).

[34] D. Bakowies, J. Chem. Phys. **127**, 084105 (2007).

[35] W. Klopper, B. Ruscic, D. P. Tew, F. A. Bischoff, and S. Wolfsegger, Chem. Phys. **356**, 14 (2009).

[36] TURBOMOLE development version, 2009, for further information see http://www.turbomole.com.

[37] R. A. Bachorz, "Implementation and application of the explicitly correlated coupled-cluster method in Turbomole," Ph.D. thesis, Universität Karlsruhe, 2009.

[38] *Recent Progress in Coupled Cluster Methods: Theory and Applications*, edited by P. Cársky, J. Pittner, and J. Paldus (Springer, New York, in press).