The Journal of Chemical Physics

Implementation of the full explicitly correlated coupled-cluster singles and doubles model CCSD-F12 with optimally reduced auxiliary basis dependence

Cite as: J. Chem. Phys. **129**, 201103 (2008); https://doi.org/10.1063/1.3028546
Submitted: 10 October 2008 . Accepted: 30 October 2008 . Published Online: 26 November 2008

Andreas Köhn, Gareth W. Richings, and David P. Tew

J. Chem. Phys. **129**, 201103 (2008); https://doi.org/10.1063/1.3028546 **129**, 201103

THE JOURNAL OF CHEMICAL PHYSICS **129**, 201103 (2008)

Implementation of the full explicitly correlated coupled-cluster singles and doubles model CCSD-F12 with optimally reduced auxiliary basis dependence

Andreas Köhn,[1,a)] Gareth W. Richings,[1] and David P. Tew[2]
[1] *Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany*
[2] *Lehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie, Universität Karlsruhe (TH), D-76128 Karlsruhe, Germany*

(Received 10 October 2008; accepted 30 October 2008; published online 26 November 2008)

An implementation of the full explicitly correlated coupled-cluster singles and doubles model CCSD-F12 using a single Slater-type geminal has been obtained with the aid of automated term generation and evaluation techniques. In contrast to a previously reported computer code [T. Shiozaki *et al.*, J. Chem. Phys. **129**, 071101 (2008)], our implementation features a reduced dependence on the auxiliary basis set due to the use of a reformulated evaluation of the so-called Z-intermediate rather than straight forward insertion of an auxiliary basis expansion, which allows an unambiguous comparison to more approximate CCSD-F12 models. First benchmark results for total correlation energies and reaction energies indicate an excellent performance of the much cheaper CCSD(F12) model. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3028546]

Explicitly correlated F12 methods have been developed to overcome the slow convergence of the correlation energy with basis set size observed for orbital based expansions.[1–4] In combination with coupled-cluster methods, highly accurate results can be obtained with relatively small basis sets.[5–10]

Explicitly correlated coupled-cluster theory was pioneered by Noga *et al.*[6] and Noga and Kutzelnigg[7] and extended to F12-theory only recently.[8] The implementation was restricted to the standard approximation, which greatly simplified the equations. Within the CABS approach (*vide infra*), a fully hand-coded implementation does not seem feasible, and only implementations of approximate models have been reported so far.[9–11]

Here, we report an implementation of the full CCSD-F12 model, which was made possible by the use of automated term generation and evaluation techniques. The underlying code, named GECCO, consists of a symbolic algebra part and a string-based general contraction kernel for the numerical evaluation of the resulting expressions. Details of the program will be presented elsewhere.[12] While work on this project was in progress, an implementation of CCSD-F12 by another group was published,[13] which was also based on an automated approach, in this case automatic code generation. The presented numerical results, however, featured a peculiar increased auxiliary basis set dependence of the CCSD-F12 energy, as compared to truncated models as CCSD(F12).[9] In the present communication, we will investigate this behavior and show how to avoid this unpleasant feature. This will enable us to present an unambiguous comparison to the CCSD(F12) model and a linearized CCSD-F12 model.

The CCSD-F12 wave function is obtained by extension of the usual cluster operator, with an R12-dependent double excitation operator $T_{2'}$, resulting in

$$|\Psi_{\text{CCSD-F12}}\rangle = e^{T_1+T_2+T_{2'}}|0\rangle, \tag{1}$$

where $|0\rangle$ is the reference state and the cluster operators are defined as

$$T_1 = t^i_a a^a_i, \quad T_2 = \tfrac{1}{4} t^{ij}_{ab} a^{ab}_{ij}, \quad T_{2'} = \tfrac{1}{8} c^{ij}_{kl} R^{kl}_{\alpha\beta} a^{\alpha\beta}_{ij}. \tag{2}$$

We adhere to the usual convention that indices $i,j,\ldots$ denote occupied orbitals, $a,b,\ldots$ virtual orbitals (in a finite, computationally tractable basis), and $\alpha,\beta,\ldots$ refer to virtual orbitals in a formally complete basis. The symbol

$$R^{kl}_{\alpha\beta} = \langle \alpha\beta | Q_{12} f(r_{12}) | kl \rangle \tag{3}$$

denotes matrix elements over the correlation factor $f(r_{12})$, which is usually chosen to be a Slater-type geminal, see below. Q_{12} is the strong orthogonality projector.

Expanding the CC equations in terms of the elements of $\bar{H} = e^{-T} H e^{T}$, where $T = T_1 + T_2$, the CCSD-F12 energy reads

$$E = \langle 0 | \bar{H} + [\bar{H}, T_{2'}] | 0 \rangle, \tag{4}$$

and the equations become

$$0 = \left\langle {}^{a}_{i} \right| \bar{H} + [\bar{H}, T_{2'}] | 0 \rangle, \tag{5}$$

$$0 = \left\langle {}^{ab}_{ij} \right| \bar{H} + [\bar{H}, T_{2'}] + \tfrac{1}{2}[[\bar{H}, T_{2'}], T_{2'}] | 0 \rangle, \tag{6}$$

$$0 = \left\langle {}^{kl}_{ij} \right| \bar{H} + [\bar{H}, T_{2'}] + \tfrac{1}{2}[[\bar{H}, T_{2'}], T_{2'}] | 0 \rangle, \tag{7}$$

where $\left\langle {}^{kl}_{ij} \right| = \langle 0 | a^{ij}_{\alpha\beta} R^{\alpha\beta}_{kl}$. The well-established CCSD(F12) method of Fliegl *et al.*[9] reduces the effort of solving these equations by truncation of Eqs. (6) and (7) to

---

[a] Electronic mail: andreas.koehn@uni-mainz.de.

0021-9606/2008/129(20)/201103/5/$23.00

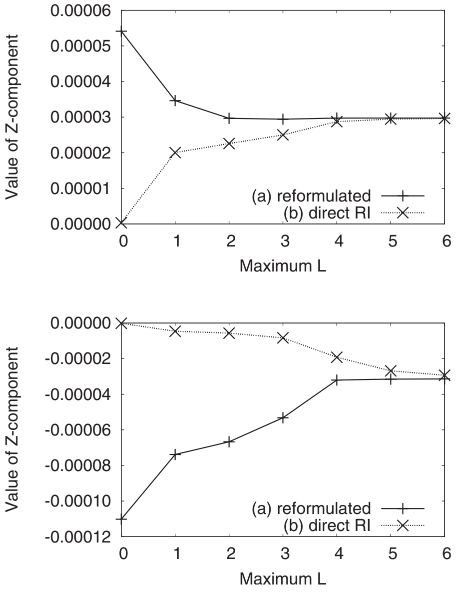

FIG. 1. Neon. Convergence of elements $Z^{23,3}_{23,3}$ (upper panel) and $Z^{2\bar{3},3}_{3\bar{3},2}$ (lower panel) with auxiliary basis set size. The orbital basis is aug-cc-pVTZ, as auxiliary basis subsets of the K4 basis set are used.

$$0 = \left\langle {}^{ab}_{ij} \middle| \bar{H} + [\bar{H}, T_{2'}] \middle| 0 \right\rangle, \tag{8}$$

$$0 = \left\langle {}^{kl}_{ij} \middle| \bar{H} + [F, T_{2'}] \middle| 0 \right\rangle, \tag{9}$$

where F is the Fock operator.

Removal of all terms within a particular commutator ensures that the remaining equations are size consistent.[9] With this in mind, we propose an alternative, truncated CCSD-F12 method. The rationale behind the method is similar to that behind the CCSD(F12) approach. It is assumed that commutator terms involving the R12-dependent excitations will be relatively small, hence the double commutators involving these excitations should be almost negligible. We thus propose the CCSD-F12(lin) method whereby only terms quadratic in $T_{2'}$ (third order in the fluctuation potential) are ignored. Therefore the equations we use are Eqs. (4), (5), and (8) as well as

$$0 = \left\langle {}^{kl}_{ij} \middle| \bar{H} + [\bar{H}, T_{2'}] \middle| 0 \right\rangle. \tag{10}$$

Three-electron and higher order integrals are avoided by inserting the approximate resolution of the identity (RI), using the union of the orbital basis set (OBS) and an additional auxiliary basis set (ABS). This procedure is known as the complementary auxiliary basis set (CABS approach[14]). In doing so, care must be taken to avoid RIs at places that introduce slowly convergent partial wave expansions. Thus, it is in many cases necessary to consider reformulated intermediates as already discussed in the initial work on MP2-R12 (Ref. 2) and CCSD-R12 (Ref. 7) theories. Beyond the intermediates needed in MP2-F12 theory (a good overview can be found in Ref. 3), which are also sufficient for implementing the CCSD(F12) model, for CCSD-F12, only two additional cases need special consideration, namely, the P and the Z intermediates,

$$P^{ij}_{kl} = R^{\alpha\beta}_{kl} g^{\gamma\delta}_{\alpha\beta} R^{ij}_{\gamma\delta}, \tag{11}$$

$$Z^{ij;p}_{kl;m} = R^{\alpha\beta}_{kl} g^{\gamma p}_{\alpha m} R^{ij}_{\gamma\beta}. \tag{12}$$

All other terms involving the geminal operator R can be evaluated by directly approximating the complementary space by the CABS, as the corresponding partial wave expansions are either rapidly convergent or finite.[7,8]

Reformulated expressions for P and Z within the CABS approach have been given by Noga *et al.*,[8] but numerical results were only reported within the so-called standard approximation. Unfortunately, the formula quoted for Z in Ref. 8 seems to contain an error, so we give a corrected version here as follows:

$$Z^{ij;p}_{kl;m} = {}^{C}Z^{ij;p}_{kl;m} - {}^{X}Z^{ij;p}_{kl;m}, \tag{13}$$

$$\begin{aligned} {}^{C}Z^{ij;p}_{kl;m} = {} & \tfrac{1}{2}\Big[g^{p';p}_{k;m} (R^2)^{ij}_{p'l} - g^{p';p}_{k;l} (R^2)^{ij}_{p'm} + (R^2)^{p'j}_{kl} g^{i;p}_{p';m} \\ & - (R^2)^{p'i}_{kl} g^{j;p}_{p';m} \Big] - R^{qs}_{kl} g^{r;p}_{q;m} R^{ij}_{rs} - R^{q''o}_{kl} g^{r'';p}_{q'';m} R^{ij}_{r''o} \\ & - R^{q''s}_{kl} g^{r;p}_{q'';m} R^{ij}_{rs} - R^{qs}_{kl} g^{r'';p}_{q;m} R^{ij}_{r''s} - R^{os''}_{kl} g^{r'';p}_{o;m} R^{ij}_{r''s''} \\ & - R^{q''s''}_{kl} g^{o;p}_{q'';m} R^{ij}_{os''} - R^{os''}_{kl} g^{n;p}_{o;m} R^{ij}_{ns''} - R^{os''}_{kl} g^{a;p}_{o;m} R^{ij}_{as''} \\ & - R^{as''}_{kl} g^{o;p}_{a;m} R^{ij}_{os''}, \end{aligned} \tag{14}$$

$$\begin{aligned} {}^{X}Z^{ij;p}_{kl;m} = {} & R^{p''q''}_{kl} g^{p;r''}_{p'';m} R^{ij}_{r''q''} + R^{aq''}_{kl} g^{p;r''}_{a;m} R^{ij}_{r''q''} + R^{aq''}_{kl} g^{p;b}_{a;m} R^{ij}_{bq''} \\ & + R^{p''q''}_{kl} g^{p;b}_{p'';m} R^{ij}_{bq''} + R^{p''c}_{kl} g^{p;r''}_{p'';m} R^{ij}_{r''c}. \end{aligned} \tag{15}$$

The notation for the two-electron repulsion integrals $g^{p;r}_{q;s}$ indicates that the pairs pr and qs are not antisymmetrized, which allows splitting into Coulomb (C) and exchange (X) contributions to Z. Here p , q are general orbitals in the finite basis, p'' , q'' are auxiliary orbitals complementary to the finite basis, and $p' = p \oplus p''$.

In the recent work of Shiozaki *et al.*,[13] only the P intermediate was considered (in addition to the intermediates known from MP2-F12 theory), whereas terms that formally involve the Z intermediate were all treated by direct insertion of the RI. This corresponds to evaluating the Z intermediate according to Eq. (12) with the complete basis approximated by the union of the orbital basis and the auxiliary basis. The consequences of this approximation (which we will refer to as "Z with direct RI") will be discussed below.

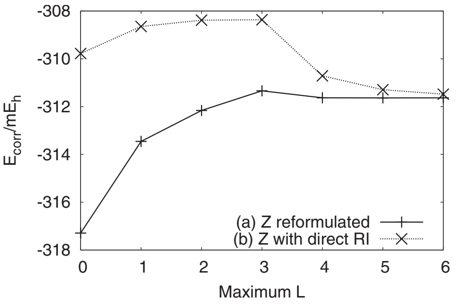

FIG. 2. Neon. Convergence of CCSD-F12 correlation energy with auxiliary basis set size. The orbital basis is aug-cc-pVTZ, as auxiliary basis subsets of the K4 basis set are used.

TABLE I. CCSD valence correlation energies (E_{corr}) of Ne in mE_H. Ansatz 3 in approximation B with a Slater-type geminal with $\gamma=1.3$ has been used.

		Z using Eqs. (13)–(15)		Z with direct RI
OBS/CABS	CCSD(F12)	CCSD-F12(lin)	CCSD-F12	CCSD-F12
aug-cc-pVDZ/K2	−305.469	−306.675	−306.526	−305.814
aug-cc-pVDZ/K3	−305.533	−306.721	−306.572	−306.303
aug-cc-pVDZ/K4	−305.538	−306.727	−306.578	−306.190
aug-cc-pVTZ/K2	−311.434	−311.646	−311.630	−311.044
aug-cc-pVTZ/K3	−311.450	−311.640	−311.624	−311.397
aug-cc-pVTZ/K4	−311.465	−311.651	−311.634	−311.478
aug-cc-pVQZ/K2	−314.327	−314.258	−314.255	−313.758
aug-cc-pVQZ/K3	−314.372	−314.258	−314.256	−314.060
aug-cc-pVQZ/K4	−314.381	−314.264	−314.262	−314.114
aug-cc-pV5Z/K2	−315.144	−315.128	−315.128	−314.640
aug-cc-pV5Z/K3	−315.246	−315.134	−315.134	−314.961
aug-cc-pV5Z/K4	−315.246	−315.131	−315.131	−314.988
aug-cc-pV6Z/K2	−315.517	−315.454	−315.454	−315.282
aug-cc-pV6Z/K3	−315.502	−315.452	−315.452	−315.235
aug-cc-pV6Z/K4	−315.524	−315.454	−315.454	−315.315

All numerical values presented in this contribution have been obtained using an automated approach with our newly developed program GECCO.[12] All special F12-integrals are extracted from a local version of DALTON[15] in atomic-orbital form and transformed as appropriate. A Slater-type geminal $\exp(-\gamma r_{12})$ was used as implemented in DALTON where a set of six Gaussian-type functions are fitted to the required function (using the appropriately scaled exponents and the coefficients given in Ref. 16). For Q_{12}, the modified ansatz 2,[14] also referred to as ansatz 3, was used and the special F12-intermediates were evaluated according to approximation B.[17] In this work, neither the extended nor generalized Brillouin condition (GBC) (Ref. 2) are assumed. All calculations will be carried out by use of the frozen-core approximation.

In order to examine closely the behavior of the CCSD-F12 method for different formulations of the Z intermediate, it is informative to compare the convergence of correlation energies with the systematic improvement of the ABS. In this respect, the frozen-core CCSD-F12/3B energies of the neon atom were calculated, using, as OBS, the aug-cc-pVTZ basis set along with subsets of a $19s14p10d8f6g4h2i$ basis[17] as ABS (the K4 set of Ref. 13), i.e., the ABS was systematically improved by adding the basis functions of higher angular momentum quantum number to the minimal set of all s-type functions.

Beginning with the convergence of selected elements of Z calculated by the two methods in Fig. 1, it is apparent that a much faster convergence is obtained using the reformulated Z, Eqs. (13)–(15). In case of $Z^{23;3}_{23;3}$ (i.e., all labels referring to α-spin), both approaches are tightly converged for the full ABS (upper panel of Fig. 1). While the direct RI approach needs h-functions at least, the reformulated expression is already converged for d-functions. The situation for $Z^{2\bar{3};3}_{3\bar{3};2}$ is somewhat different (Fig. 1 lower panel). Up to $L=3$, Eqs. (13)–(15) seem to yield an inferior approximation but once that g-functions are included in the ABS, the result is converged. Obviously a term with a partial wave-expansion breaking off at finite order dominates in this case. The direct RI approximation to the matrix element, however, shows a slowly convergent behavior and even for the full ABS (including up to i-functions) no tight convergence is obtained.

The convergence behavior of the Z matrix elements are directly reflected in that of the correlation energy, Fig. 2. With Eqs. (13)–(15), the correlation energy is converged to within 2 μE_H of the full K4 ABS energy when using the ABS truncated at the g-type functions. At the same point, the energy obtained with direct RI is only within 800 μE_H of the full ABS value. Upon adding further higher angular momentum functions, a very slow L^{-3} convergence behavior of the direct RI based approach for Z can be seen, which suggests that it will be hard to obtain an accurate representation of the Z-contribution in this way. Moreover, as the full CCSD-F12 model contains several terms that scale with the square of the auxiliary basis size (one term even scales cubically), obtaining high accuracy with a small ABS is advantageous.

Calculations on the neon atom and on the water molecule have also been performed with a selection of orbital (chosen from the sets aug-cc-pVXZ with X=D, T, Q, 5, and 6) and auxiliary basis sets (K2, K3, and for Ne only K4,[17,18] the nomenclature follows Ref. 13) in order to assess the performance of the different formulations for the Z intermediate. In addition, results from calculations using the two approximate CCSD-F12 models are presented, which quantify the errors made in (a) neglecting terms nonlinear in $T_{2'}$ [CCSD-F12(lin)] and (b) additionally approximating $[\bar{H},T_{2'}]$ by $[F,T_{2'}]$, [CCSD(F12)]. Results are collected in Tables I and II. Please note that at difference to the Dalton implementation, the CCSD(F12) energies are evaluated without assuming the GBC. This was done as the GBC is not an essential approximation and does not significantly change the operation count of that model.

The first point to note is that the more accurate evaluation of the Z-dependent terms leads in all cases to a much

TABLE II. CCSD valence correlation energies (E_{corr}) of H_2O in mE_H. Ansatz 3 in approximation B with a Slater-type geminal with $\gamma=1.3$ has been used.

		Z using Eqs. (13)–(15)		*Z* with direct RI
OBS/CABS	CCSD(F12)	CCSD-F12(lin)	CCSD-F12	CCSD-F12
aug-cc-pVDZ/K2	−290.998	−291.306	−291.196	−290.664
aug-cc-pVDZ/K3	−291.005	−291.300	−291.190	−290.994
aug-cc-pVTZ/K2	−296.089	−295.962	−295.953	−295.582
aug-cc-pVTZ/K3	−296.097	−295.959	−295.950	−295.778
aug-cc-pVQZ/K2	−297.803	−297.654	−297.653	−297.385
aug-cc-pVQZ/K3	−297.820	−298.655	−297.654	−297.520

reduced auxiliary basis set dependence, which is now comparable to that of CCSD(F12). For both neon and water, the inaccuracy introduced by the direct RI for the *Z*-dependent terms leads to deviations on the order of 200 μE_H for aug-cc-pVTZ orbital basis sets with a K3 auxiliary basis. Noticeable effects persist even for large OBSs, >100 μE_H for the CCSD-F12 calculations on neon atom with an aug-cc-pV6Z orbital basis and a K4 auxiliary basis. Of course, the error due to an incomplete orbital basis is considerably larger, in particular for the smaller basis sets, e.g., the deviation between the aug-cc-pVTZ and the aug-cc-pV6Z basis amounts to approximately 3800 μE_H. On the other hand, the deviations due to the different treatment of the *Z* intermediate are on the same order as those between the CCSD(F12) and the full CCSD-F12 model.

Turning to the performance of the approximate models, we note that starting from triple-zeta quality orbital basis sets, the contributions from terms nonlinear in $T_{2'}$ become negligible, as seen from the close resemblance of the CCSD-F12 and CCSD-F12(lin) numbers. The results for the CCSD(F12) are excellent as well, the deviations amount to fractions of mE_H only.

These observations carry over to reaction energies. As an example, we give the water gas shift reaction, $H_2+CO_2 \rightarrow H_2O+CO$. We use an aug-cc-pVTZ orbital basis set and an uncontracted aug-cc-pV5Z auxiliary basis set with a Slater-type geminal of exponent $\gamma=1.3$ and approximation B for the F12-intermediates. Geometries were taken from Ref. 19. The Hartree–Fock contribution is +0.65 kJ mol^{-1}. The frozen-core correlation contribution is −17.06 kJ mol^{-1} using CCSD(F12), −17.00 kJ mol^{-1} using CCSD-F12(lin), and −16.99 kJ mol^{-1} using the full CCSD-F12 model with the reformulated evaluation of the *Z*-intermediate. Again, CCSD-F12 and CCSD-F12(lin) are in virtually perfect agreement, and the computationally much cheaper CCSD(F12) model performs excellent again, with a deviation below 0.1 kJ mol^{-1} to the full CCSD-F12 result. Using direct RI for the *Z*-intermediate the result for CCSD-F12 is −16.77 kJ mol^{-1}, which is a significantly larger error than introduced by the CCSD(F12) approximation. We note that the differences due to changing to approximation C are below 0.001 kJ mol^{-1} and the variation due to choosing a different $\gamma=1.0$ amounts to 0.04 kJ mol^{-1}.

In conclusion, we have shown that using automated implementation techniques, the evaluation of highly complex wave function models has become a feasible task. In this we agree with Shiozaki *et al.*,[13] who have recently achieved the same goal using a code generation approach. We have presented results for the completely untruncated CCSD-F12 model using the CABS approach to avoid many-electron integrals. In contrast to Ref. 13, our implementation avoids any RI approximations with slowly convergent partial wave expansions, namely the correctly reformulated expression for the *Z*-intermediate has been used. This turns out to be an important factor in maximizing the stability of the correlation energy to changes in the ABS. Therefore this nontruncated CCSD-F12 implementation can be confidently used for benchmarking purposes.

Preliminary benchmark results indicate that CCSD(F12) performs excellently compared to full CCSD-F12, with errors well below 1 mE_H for neon atom and water molecule absolute correlation energies, and below 0.1 kJ mol^{-1} for the correlation contribution to the water gas shift reaction.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Grant Nos. KO2337/2 and TE644/1). The authors would like to acknowledge Wim Klopper for assistance with the DALTON integral package and for helpful discussions.

[1] W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985).
[2] W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).
[3] H. J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[4] W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).
[5] D. P. Tew, W. M. Klopper, C. Neiss, and C. Hättig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).
[6] J. Noga, W. Kutzelnigg, and W. Klopper, Chem. Phys. Lett. **199**, 497 (1992).
[7] J. Noga and W. Kutzelnigg, J. Chem. Phys. **101**, 7738 (1994).
[8] J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, J. Chem. Phys. **128**, 174103 (2008).
[9] H. Fliegl, W. M. Klopper, and C. Hättig, J. Chem. Phys. **122**, 084107 (2005).
[10] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[11] E. F. Valeev and T. D. Crawford, J. Chem. Phys. **128**, 244113 (2008).
[12] A. Köhn and G. Richings (unpublished).
[13] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **129**, 071101 (2008).

[14] E. F. Valeev, Chem. Phys. Lett. **395**, 190 (2004).
[15] DALTON, a molecular electronic structure program, Release 2.0 (2005), see http://www.kjemi.uio.no/software/dalton/dalton.html.
[16] D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[17] W. Klopper and C. C. M. Samson, J. Chem. Phys. **116**, 6397 (2002).
[18] W. Klopper, J. Chem. Phys. **102**, 6168 (1995).
[19] D. P. Tew, W. M. Klopper, and C. Hättig, Chem. Phys. Lett. **452**, 326 (2008).