The following article appeared in J. Chem. Phys. 142, 064111 (2015) and may be found at http://dx.doi.org/10.1063/1.4907591.

Accurate thermochemistry from explicitly correlated distinguishable cluster approximation

Daniel Kats, David Kreplin, Hans-Joachim Werner, and Frederick R. Manby

Citation: The Journal of Chemical Physics **142**, 064111 (2015); doi: 10.1063/1.4907591
View online: http://dx.doi.org/10.1063/1.4907591
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/142/6?ver=pdfcov
Published by the AIP Publishing

Accurate thermochemistry from explicitly correlated distinguishable cluster approximation

Daniel Kats,[1,a)] David Kreplin,[1] Hans-Joachim Werner,[1] and Frederick R. Manby[2]
[1]*Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany*
[2]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom*

(Received 25 November 2014; accepted 24 January 2015; published online 11 February 2015)

An explicitly correlated version of the distinguishable-cluster approximation is presented and extensively benchmarked. It is shown that the usual F12-type explicitly correlated approaches are applicable to distinguishable-cluster theory with single and double excitations, and the results show a significant improvement compared to coupled-cluster theory with singles and doubles for closed and open-shell systems. The resulting method can be applied in a black-box manner to systems with single- and multireference character. Most noticeably, optimized geometries are of coupled-cluster singles and doubles with perturbative triples quality or even better. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4907591]

I. INTRODUCTION

Accurate relative energies are an essential prerequisite for *ab-initio* thermochemical prediction, and systematic behaviour of the method is typically more important than the accuracy of absolute energies.

Coupled-cluster theory[1] is now widely accepted as the best choice for accurate calculations with predictive power, but one has to introduce triple excitations to achieve chemical accuracy. Coupled-cluster theory with single and double (CCSD) excitations[2] alone usually gives much poorer results and, although containing all possible Goldstone diagrams that correspond to singles and doubles, can be easily outperformed by other electron-pair theories.[3–11] After including triple excitations in a perturbative manner, the CCSD(T) method[12] is obtained, which shows a good balance between accuracy and cost. However, the (T) contribution is still too expensive for many applications, since the computation cost of evaluating it scales as $O(N^7)$ with the molecular size N. It would obviously be desirable to retain the $O(N^6)$ scaling of CCSD whilst achieving the accuracy of CCSD(T).

Recently, two of us introduced a new electron-pair method called the distinguishable cluster (DC) approximation.[11] In the DC approximation, the quadratic terms in the coupled-cluster doubles (CCD) amplitude equation are modified to remove exchange interactions *between* distinct 2-particle clusters. Such exchange terms should go to zero if the clusters were well separated and the orbitals were localizable on each fragment. But, if one uses restricted Hartree-Fock (RHF) theory as the reference, such exchange terms do not decay properly in the case of bond breaking, where only unrestricted orbitals are fully localizable. Therefore, in the distinguishable-cluster doubles (DCD) method, these terms are completely removed, and the remaining terms are adjusted (in a unique way) to retain exactness for two-electron systems and particle-hole symmetry of the equations.

The reasoning applies particularly to dissociating bonds, where DCD systematically out-performs CCD. In principle, the modifications could spoil the accuracy of CCSD calculations for single-reference cases, such as molecules close to their equilibrium geometries. However, our first benchmark calculations (using Brueckner DCD (BDCD) — the doubles theory with Brueckner orbitals[13] instead of single excitations) showed the opposite: reaction energies are predicted more accurately with BDCD than with Brueckner CCD. These findings have been confirmed also for other orbital-relaxation techniques,[14] which makes this method interesting for applications that are either too large to be treated by CCSD(T) or that have substantial multireference character, or both.

Another parameter that strongly influences the quality of wave-function based calculations is the size of the basis set. But recent progress in explicitly correlated methods[15–41] has made this problem much less severe, and nowadays one can closely approach the basis-set limit with a modest basis set of triple-zeta quality.

The combination of these two techniques — described in this paper — results in a black-box method that can quantitatively predict various experimental values with $O(N^6)$ scaling.

II. THEORY

As has been shown in Ref. 14, the DC method with singles and doubles (DCSD) can be straightforwardly obtained from the CCSD method by modifying terms quadratic in the doubles amplitudes. Some technical complications may arise depending on the factorization of the underlying CCSD equations. In the case of the factorization[42] used in MOLPRO,[43,44] where in some terms the singles amplitudes are incorporated into doubles amplitudes, one has to use, e.g., an unmodified α intermediate as in CCSD and correct this later by subtracting

a)Electronic mail: kats@theochem.uni-stuttgart.de

the quadratic term in the doubles amplitudes. The explicit amplitude equations based on the CCSD factorization used in MOLPRO[42] can be found in the Appendix.

The introduction of explicit correlation to the DCSD method is also a straightforward procedure when using the F12a and F12b approximations originally derived for CCSD,[28,39] since these consider only linear terms in CCD, which remain unchanged in DCD. Therefore, one can use the same F12-equations for DCSD-F12*x* as for CCSD-F12*x*.

We have implemented closed-shell and open-shell versions of DCSD-F12 in the MOLPRO package,[43,44] and in Sec. III, we will establish the applicability of these F12 approximations to the DCSD method and benchmark the quality of DCSD-F12 with various test systems.

III. BENCHMARK CALCULATIONS

In order to estimate the accuracy of DCSD-F12, we have recalculated all test cases from Ref. 39, i.e., reaction energies for closed- and open-shell molecules, atomization energies (AEs) and ionization energies, electron affinities, and spectroscopic constants. Additionally, we computed reaction-barrier heights for the DBH24 benchmark set,[45] interaction energies using the S22 benchmark set,[46] and optimized geometries of small molecules consisting of first-row[47] and second-row[48] elements. In order to test the sensitivity of DCSD-F12 to multireference character, we have calculated the potential-energy curves for N_2, C_2, and BN molecules and optimized the geometry of twisted ethene.

All calculations use a RHF reference. For all F12 calculations, we used the aug-cc-pVTZ basis set for first-row elements and aug-cc-pV(T+d)Z for second-row elements as was done in Ref. 39. As in the previous calculations for electron affinities,[39] we have further augmented the basis sets by one diffuse *s*-shell for hydrogen and one *s*- and one *p*-shell for other atoms. The exponents for these functions were obtained by dividing the lowest exponent of the corresponding set by 1.8. The complementary auxiliary basis set (CABS) singles correction of the HF energy was used in all calculations. The auxiliary aug-cc-pVTZ/MP2FIT basis set[49] was used for density fitting, and cc-pVTZ/JKFIT[50] for the resolution-of-the-identity approximation and for the calculation of the additional CABS blocks of Fock and exchange matrices. As has been shown in Ref. 39, these basis sets represent a good compromise between accuracy and efficiency. All computations are done with the frozen-core approximation apart from some geometry optimizations of ethene (see below). Calculations of open-shell systems are done using partially spin-restricted versions of CCSD,[51] CCSD(T),[52] and DCSD for all test cases apart from the DBH24 test set, where we have used the spin-unrestricted versions of correlation methods based on RHF orbitals.

From now on, we will omit the -F12 suffix for convenience and will only use the acronyms F12a and F12b where necessary.

A. Reaction energies

Calculations of reaction energies represent a typical application for which accurate energetics are required. We have calculated energies of 54 closed-shell and 47 open-shell reactions. The corresponding reactions and molecular geometries can be found in the supplementary material of Ref. 39. In the following, we compare results from DCSD calculations with CCSD and CCSD(T).

In Figure 1 a graphical representation of deviations of CCSD-F12a and DCSD-F12a results from CCSD(T)-F12a results is given. From these plots, it is evident that in most cases the DCSD results are closer to the reference values than the CCSD results. This is even more pronounced in the case of the open-shell reactions. Note that larger absolute deviations from reference energies of the open-shell reactions reflect the higher absolute values of these reaction energies compared to the closed-shell ones.

The statistical measures, i.e., mean absolute deviations (MAD), root mean square deviations (RMSD), and maximal deviation (MaxD), for errors in reaction energies for F12a and F12b approximations are given in Table I. As already seen in the previous publications,[11,14] the RMSD of DCSD is around 60% of that of CCSD for the closed-shell reactions. For the open-shell reactions, the improvement is even larger and the RMSD is roughly halved on going from CCSD to DCSD. The difference between the F12a and F12b results is rather small, but one can see that the discrepancy from CCSD(T) is slightly larger in the case of F12b.

Overall, one can conclude that the F12*x* explicitly correlated corrections are valid for the DCSD method, do not corrupt its accuracy, and do not deteriorate on going to open-shell systems.

B. Atomization energies, ionization potentials (IPs), and electron affinities (EAs)

In order to test sensitivity of DCSD to highly energetic processes and to variable number of electrons, we have calculated atomization energies, IPs, and EAs for the G2 test set of molecules,[53] which was also utilized previously in Ref. 39.

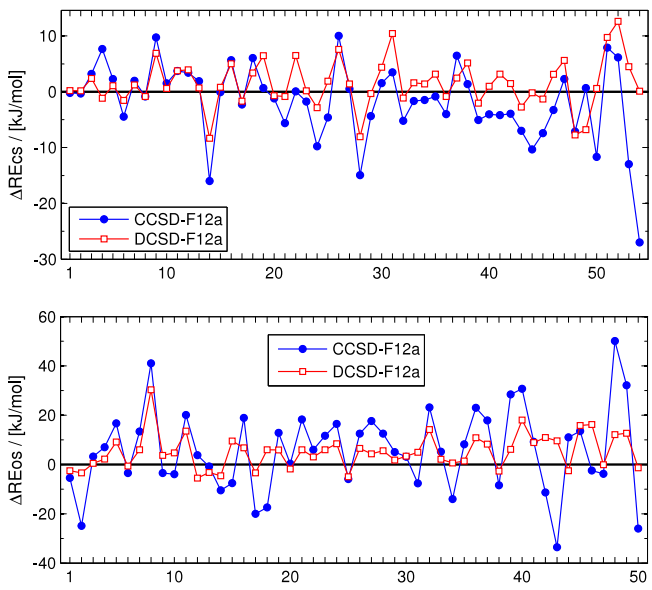

FIG. 1. Reaction energies: deviations of CCSD-F12a and DCSD-F12a results from the corresponding CCSD(T)-F12a results. Top: closed-shell reactions. Bottom: open-shell reactions. On the x-coordinate: reaction numbers from Ref. 39.

TABLE I. Reaction energies: statistical analysis of the absolute deviations of the CCSD and DCSD values from CCSD(T) values in kJ/mol. Note that since the F12x approximations differ only in the way the energy is calculated, the energy differences $E_{\mathrm{CCSD-F12a}} - E_{\mathrm{CCSD(T)-F12a}}$ and $E_{\mathrm{CCSD-F12b}} - E_{\mathrm{CCSD(T)-F12b}}$ are exactly equal, but this is not the case for $E_{\mathrm{DCSD-F12x}} - E_{\mathrm{CCSD(T)-F12x}}$.

	Closed-shell			Open-shell		
	MAD	RMSD	MaxD	MAD	RMSD	MaxD
CCSD-F12x	5.04	7.00	27.9	14.08	17.68	50.1
DCSD-F12a	3.22	4.40	12.6	6.66	8.67	30.3
DCSD-F12b	3.22	4.41	12.6	6.73	8.75	30.4

We have used the RMP2-optimized geometries for the molecules and ions from the supplementary material of Ref. 39. As already described at the beginning of Sec. III, the basis set for molecules from the EA benchmark calculation was expanded by diffuse s- and p-shell functions to improve accuracy.

We again use the CCSD(T) results as our reference values and evaluate the accuracy of the DCSD (and CCSD) methods relative to them. The statistical analysis of the deviations is shown in Table II, and the plots of individual deviations (in the same order as Table VIII in Ref. 39) are presented in Figure 2.

The accuracy of AEs from CCSD calculations varies strongly depending on the fragment sizes, i.e., for molecules containing hydrogen atoms it usually gives quite accurate results, but for molecules consisting of more than one heavy atom the results are much worse. This dependency is much less pronounced in the DCSD calculations, although one can still see the same pattern in the results. It is obvious already from the plot that the DCSD results are much better than the CCSD ones, and the statistical analysis shows the considerable improvement: errors in DCSD AEs are around 2–3 times smaller than in CCSD. The difference in the quality of DCSD-F12a and DCSD-F12b results is again very small.

The computed IPs and EAs have values up to 17 eV and up to 4 eV, respectively; therefore, the errors of less than 50 meV can be considered as relatively small. Again, the DCSD errors are around 2.5 times smaller than CCSD errors. Interestingly, the results of DCSD show a systematic error of ca. 25 meV for IPs as well as EAs, which gives a hint that there is a small dependency of the DCSD vs. CCSD(T) discrepancy on the number of electrons, i.e., CCSD(T) gives systematically somewhat larger absolute values of correlation energy than DCSD with the larger number of electrons, according to

$$\begin{aligned}\Delta X &= (E_{n_{el}} - E_{n_{el}+1})^{\mathrm{DCSD}} - (E_{n_{el}} - E_{n_{el}+1})^{\mathrm{CCSD(T)}} \\ &= \Delta E^{\mathrm{DCSD}}_{n_{el}} - \Delta E^{\mathrm{DCSD}}_{n_{el}+1}. \end{aligned} \quad (1)$$

Here, $E_{n_{el}}$ represents the energy calculated for a n_{el}-electron molecule, and ΔX stands for ΔIP and ΔEA. A possible source of this shift is the neglect of the exchange terms in DCSD. After adding a constant shift of 25 meV to the DCSD values, the RMSDs become only 19 meV and 17 meV for IPs and EAs, respectively.

To summarize, DCSD shows remarkably good results even for these complicated cases of open-shell systems and ions and is more accurate than CCSD by a factor of around 2.5.

C. Interaction energies

Previous results on neon dimer dissociation[14] have shown that the DCSD method can improve on CCSD results even for this very delicate system. In this section, we will investigate the accuracy of DCSD for the computation of counter-poise corrected interaction energies by applying it to the S22 benchmark set,[46] which contains 22 molecular dimers representing various non-covalent interactions. Molecular systems from this test set have been utilized previously for benchmarking CCSD(T)-F12x methods,[54–56] where it was shown that the explicitly correlated treatment works very well for interaction energies.

The molecular complexes in S22 can be divided into three groups according to the type of intermolecular interaction: hydrogen-bonded complexes, complexes with a predominant dispersion contribution, and mixed complexes with a similar amount of electrostatic and dispersion contributions. Large molecular complexes in this set, e.g., the adenine-thymine complex, contain up to 30 atoms, and thus a calculation of their interaction energies using all three methods would be very expensive. Therefore, we decided to compute only the four smallest complexes of each group. The most expensive calculation was the computation of the interaction energy of benzene dimer in C_{2h} symmetry containing 24 atoms.

As in Secs. III A and III B, we use CCSD(T) as the reference. One can see from the deviation plots in Figure 3 that DCSD makes a modest improvement over CCSD for these

TABLE II. Statistical analysis of the absolute deviations from CCSD(T) of atomization energies, ionization potentials, and electron affinities.

	AE (kJ/mol)			IP (meV)			EA (meV)		
	MAD	RMSD	MaxD	MAD	RMSD	MaxD	MAD	RMSD	MaxD
CCSD-F12x	21.84	26.32	62.6	60.5	71.8	178.6	82.4	95.9	175.5
DCSD-F12a	8.35	9.88	23.3	26.6	32.2	69.7	35.8	38.2	59.2
DCSD-F12b	8.45	10.00	23.6	26.9	32.4	68.9	36.9	39.4	59.1

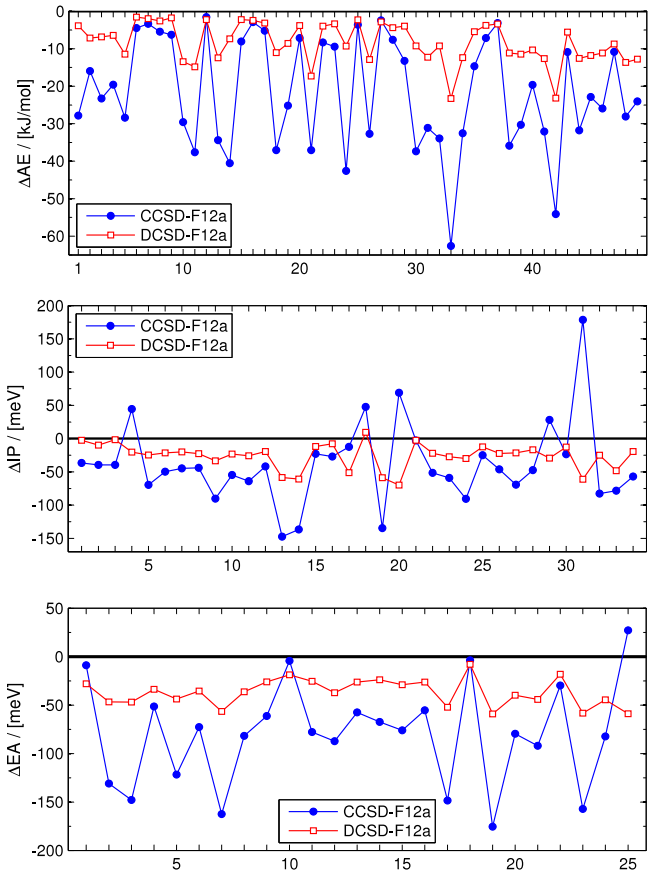

FIG. 2. Deviations of CCSD-F12a and DCSD-F12a from CCSD(T)-F12a results. Top: atomization energies; middle: ionization potentials; and bottom: electron affinities.

interaction energies. In particular, interaction energies with substantial dispersion contributions come out better with the DC approximation, with errors nearly halved relative to CCSD. The improvement of electrostatically dominated based interactions is smaller, and in the case of formic-acid dimer the CCSD interaction energy is slightly closer to the CCSD(T) results.

Overall, one can say that the DCSD results are more accurate than the CCSD results, although not as dramatically as for the test sets in Secs. III A and III B. This observation is also confirmed by the statistical analysis (Table III). Judging from the large calculations performed in this section, the DCSD method is applicable to large molecules, since the quality of calculations remains constant with the molecular size as expected for a size-extensive method.

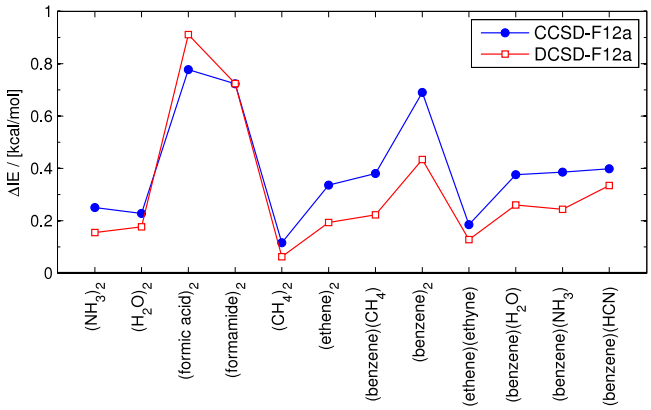

FIG. 3. Deviation of interaction energies computed with CCSD-F12a and DCSD-F12a from CCSD(T)-F12a values in kcal/mol.

TABLE III. S22 interaction energy benchmark: statistical analysis of the absolute deviations of CCSD and DCSD values from CCSD(T) ones (in kcal/mol).

	MAD	RMSD	MaxD
CCSD-F12*x*	0.40	0.45	0.78
DCSD-F12a	0.32	0.40	0.91
DCSD-F12b	0.32	0.40	0.91

D. Spectroscopic constants

As shown in Secs. III A–III C, the DCSD results are noticeably closer to CCSD(T) than are the CCSD results. In Secs. III D–III G, we will compare all three methods with reference values from experimental data and highly accurate theoretical calculations.

The simplest way of comparing computational results to experiment is to calculate spectroscopic constants of diatomic molecules. We have used the same set of molecules as in Ref. 39, i.e., closed-shell molecules HF, N_2, CO, BF, F_2, and C_2 ($^1\Sigma^+$), and open-shell molecules OH ($^2\Pi$), NH ($^3\Sigma^-$), CH ($^2\Pi$), CN ($^2\Sigma^+$), NO ($^2\Pi$), O_2 ($^3\Sigma_g^-$), and CF ($^2\Pi$).

The energy values have been computed at an equidistant grid of eleven points within a distance of ±0.15 a_0 of the optimized bond length, and the equilibrium distance r_e and harmonic frequency ω_e have been determined from an eighth-order polynomial fit of the energy values. The computed values are compared with experimental data obtained from Ref. 57, and the deviations from the experimental data are presented in Figure 4. The statistical analysis of the absolute deviations is given in Table IV.

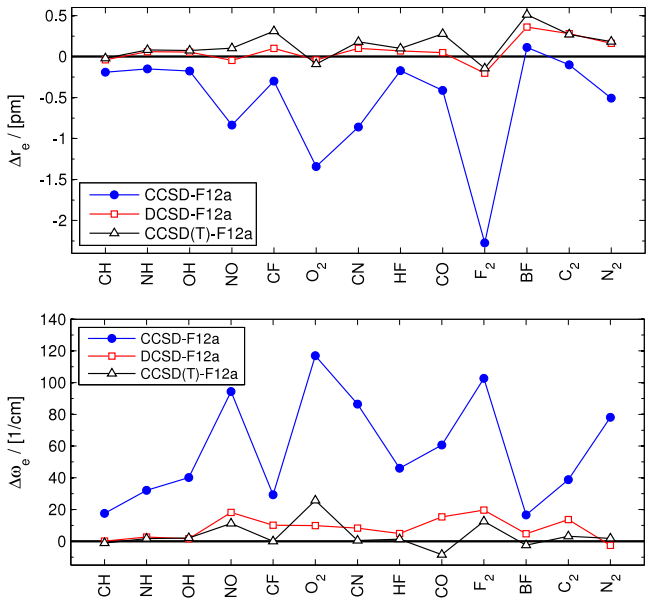

FIG. 4. Deviations from experimental reference values[57] of spectroscopic constants. Top: equilibrium distances. Bottom: harmonic frequencies.

TABLE IV. Statistical analysis of the absolute deviations of equilibrium distance r_e and harmonic frequency ω_e from experimental data.[57]

	r_e (pm)			ω_e (cm^{-1})		
	MAD	RMSD	MaxD	MAD	RMSD	MaxD
CCSD-F12a	0.57	0.84	2.27	58.4	66.8	117.0
CCSD(T)-F12a	0.18	0.22	0.51	5.5	9.0	25.7
DCSD-F12a	0.12	0.16	0.36	8.6	10.6	19.6
CCSD-F12b	0.59	0.85	2.32	59.9	68.0	117.3
CCSD(T)-F12b	0.17	0.20	0.46	6.4	9.4	26.0
DCSD-F12b	0.11	0.15	0.31	10.2	12.2	21.8

The equilibrium distances computed using CCSD(T) and DCSD are very similar, and are much closer to the experimental values than the values from CCSD calculations. Interestingly, the DCSD graph at the top of Figure 4 has approximately the same form as the CCSD(T) one. The results for harmonic frequencies from CCSD(T) and DCSD are also close to each other and much better than the CCSD frequencies. Even for difficult cases like F_2 or O_2, for which CCSD gives very large errors, DCSD results are of the same quality as CCSD(T) or even better.

The high quality of the DCSD results is also evident from the statistical analysis. The errors of CCSD are more than 5 times larger than DCSD errors. Surprisingly, DCSD yields more accurate results for the equilibrium distances than the CCSD(T) method. The reason is that CCSD(T) usually overestimates the bond lengths, while CCSD underestimates them. DCSD results are close to CCSD(T) ones but usually slightly lower, coming into a better agreement with experiment.

In the case of harmonic frequencies, where both CCSD and CCSD(T) overestimate the results, DCSD is somewhat worse than CCSD(T). F12b results of DCSD and CCSD(T) are slightly better than corresponding F12a ones for equilibrium distances, but worse for frequencies. Calculations with the quadruple-zeta basis sets showed similar results, underlining that we are close to CBS limit.

To summarize, the DCSD method is much more accurate for spectroscopic constant calculations than the CCSD method, and produces results of a similar quality as CCSD(T). DCSD geometrical parameters look very promising, making the approach an interesting candidate for geometry optimizations, as discussed in Sec. III E.

E. Geometry optimization

Motivated by the very promising results for diatomics, we performed numerical geometry optimizations on small molecules consisting of first-[47] and second-[48] row elements. The results are represented graphically in Figure 5 (bond lengths) and Figure 6 (angles), and the statistical measures are compiled in Table V. We have excluded the H_2O_2 molecule from the statistical analysis, since no method under consideration was able to provide reasonable results for the HOO-angle (all methods showed errors of more than 3°), and it was assumed before[10,47] that the experimental values are not accurate enough.

It is evident that DCSD outperforms CCSD and reaches or surpasses the accuracy of CCSD(T). Especially for bond

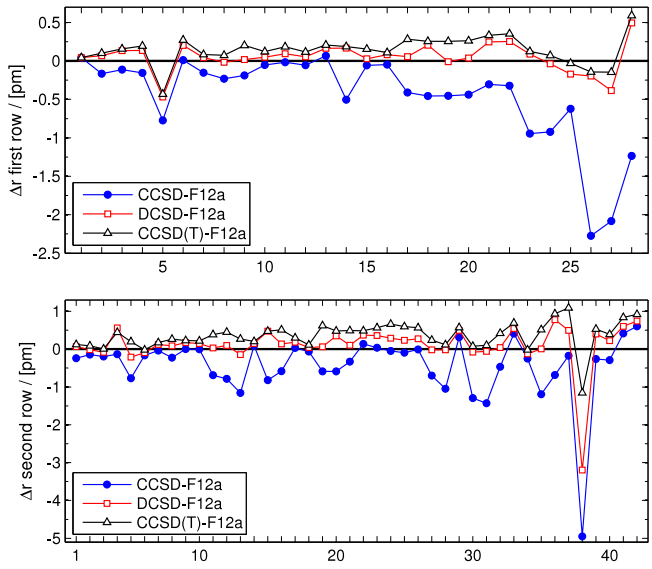

FIG. 5. Deviations of optimized bond lengths from experimental values (in the ascending order of experimental bond lengths). Top: first row benchmark set.[47] Bottom: second row benchmark set.[48]

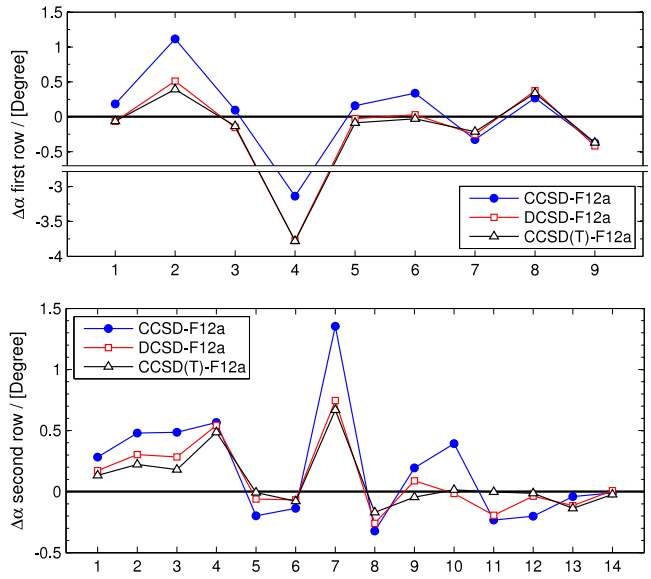

FIG. 6. Deviations of optimized bond angles from experimental values (in the ascending order of experimental angles). Top: first row benchmark set.[47] Bottom: second row benchmark set.[48]

TABLE V. Statistical analysis of optimized geometries using F12a aug-cc-pVTZ calculations versus experimental values.

	Bond lengths (pm)			Angles (deg)		
	MAD	RMSD	MaxD	MAD	RMSD	MaxD
First-row molecules without H_2O_2[47]						
CCSD	0.43	0.71	2.28	0.36	0.47	1.12
CCSD(T)	0.17	0.19	0.35	0.20	0.25	0.39
DCSD	0.12	0.15	0.39	0.23	0.29	0.51
Second-row molecules[48]						
CCSD	0.55	0.97	4.95	0.34	0.46	1.24
CCSD(T)	0.41	0.47	1.15	0.17	0.24	0.57
DCSD	0.32	0.58	3.20	0.22	0.29	0.62
Without NCl outlier						
CCSD	0.44	0.61	1.51	0.34	0.46	1.24
CCSD(T)	0.39	0.44	0.90	0.17	0.24	0.57
DCSD	0.25	0.30	0.60	0.22	0.29	0.62

lengths it usually gives better results than CCSD(T). Excluding the cases of H_2, for which all three methods are equivalent, and H_2O_2 for reasons noted above, DCSD agrees better with experiment than CCSD(T) for 85%–90% of the bonds.

The only outlier where DCSD is much worse than CCSD(T) is the N–Cl bond of the ClNO molecule, but even there, the DCSD results are much closer to the experimental values than the CCSD results. Note that all methods under consideration have the maximal deviation for this bond; therefore, we show also statistical results without this outlier in Table V. After exclusion of NCl bond, the RMSD of DCSD is one and a half times smaller than that of CCSD(T). Moreover, DCSD bond angles are nearly as accurate as results from the much more expensive CCSD(T) calculations.

One can conclude that the DCSD method is especially good for geometry optimizations, challenging the accuracy of CCSD(T) with lower computational cost scaling. These results are very promising considering the fact that the DCSD nuclear gradients are much less expensive than the CCSD(T) ones. They are also easier to implement analytically, and analytical DCSD gradients are already available in the development version of MOLPRO. Gradients of the explicitly correlated versions are under development.

F. Reaction barriers

In this section, we will investigate the accuracy of DCSD for the computation of reaction-barrier heights. We will use the DBH24 benchmark set of reaction barriers[45] for which accurate theoretical reference values calculated using the W4 *ab initio* computational thermochemistry protocol exist.[58] This test set contains 24 reaction barriers of 12 reactions (forward and reverse directions). There are 22 unique reaction barriers because of the symmetry of two reactions, but we will handle all 24 reaction barriers independently to be consistent with the previous calculations on this test set. The DBH24 benchmark set consist of first- and second-row heavy-atom and hydrogen transfer, nucleophilic substitution, and unimolecular and recombination reactions.

Our benchmark calculations with explicitly correlated CCSD, DCSD, and CCSD(T) methods are done using unrestricted versions of the methods based on a restricted HF reference[51] as was done previously for CCSD(T)-F12x.[59] We have used original geometry data of molecules and transition states from Ref. 45, which have also been used in the previous calculations.

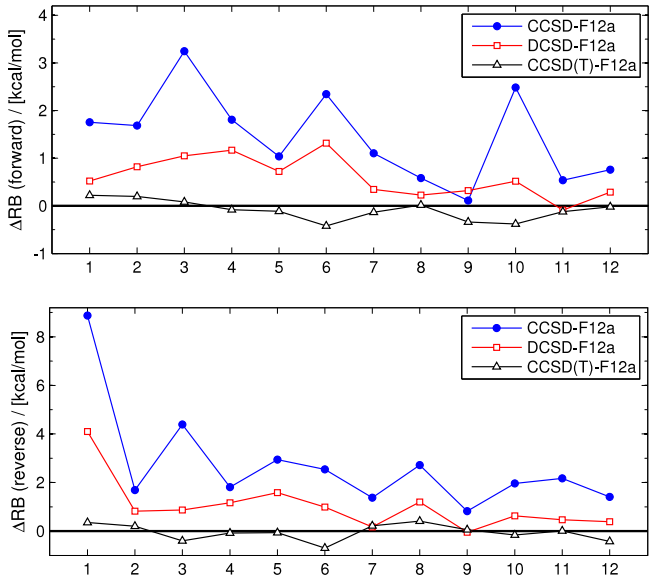

FIG. 7. Deviation of reaction-barrier heights computed with CCSD-F12a, DCSD-F12a, and CCSD(T)-F12a from W4 values[58] in kcal/mol for reactions in forward (top) and reverse (bottom) directions. The order of reactions is equal to the order in Table II of Ref. 45.

Plots of deviations from the W4 data are presented in Figure 7 and statistical analysis is shown in Table VI. The DCSD methods are more than twice as accurate for computations of reaction barriers as the CCSD methods for both F12a and F12b versions (although the accuracy is still far from the CCSD(T) calculations).

G. Systems with multireference character

The DC approximation has shown remarkable stability in cases where the conventional single-reference methods completely fail, i.e., in cases with strong correlation as encountered, e.g., in multiple-bond breaking processes.[11] It has been shown that this ability is not coupled to a particular choice of orbital relaxation, and DCSD is also capable of treating multireference problems.[14] In this section, we benchmark DCSD-F12x methods for systems with pronounced multireference character, i.e., dissociating molecules and ethene in a twisted geometry.

TABLE VI. Reaction barrier: statistical analysis of the absolute deviations of CCSD, CCSD(T), and DCSD values from W4 values[58] in kcal/mol.

	MAD	RMSD	MaxD
CCSD-F12a	2.09	2.70	8.88
CCSD(T)-F12a	0.22	0.28	0.70
DCSD-F12a	0.83	1.15	4.10
CCSD-F12b	2.12	2.72	8.78
CCSD(T)-F12b	0.22	0.27	0.58
DCSD-F12b	0.88	1.19	4.03

In the derivation of the F12x methods, some perturbation-based approximations are used, e.g., by retaining only linear terms of the CCD residual, which is related to linearized CCD or third-order variational or unitary coupled-cluster theory.[60–62] For this reason, it is not immediately obvious that the F12 approximations will work well on the onset of strong correlation.

In Figure 8, we present potential energy curves of the nitrogen molecule calculated with explicitly correlated CCSD, CCSD(T), and DCSD using aug-cc-pVTZ basis, and with plain DCSD using aug-cc-pV5Z basis. Evidently, the form of the F12 curves resembles results from the previous calculations without explicit correlation, with DCSD-F12a yielding a qualitatively correct potential energy curve. It should be noted, however, that as already shown in previous publications,[11,14] the dissociation energy is about 17% too large. Comparing DCSD-F12a values with large-basis DCSD values, one can conclude that the F12a approximation works well even in the strong-correlation region, and the DCSD-F12a method can be applied for studying problems with considerable multireference character, at least qualitatively. (The DCSD-F12b curve is very close to DCSD-F12a and is therefore not shown in the plot.)

In Figures 9 and 10, potential energy curves for the C_2 and BN molecules are plotted. These molecules are often used to test the accuracy of various methods and represent a difficult case for CCSD(T) even at the equilibrium distance.[63–67] The dissociation curves look similar to the N_2 curve; although in the case of BN $X^3\Pi$ dissociation, we encountered convergence problems with DCSD. Presumably, it is related to the wrong behaviour of DCSD in the case of carbon monoxide and could be fixed by using more flexible orbital-relaxation techniques,[14] i.e., Brueckner or optimized orbitals. Implementation of explicit correlation for these methods would require F12-integral calculations and transformations in each iteration, which would constitute a bottleneck of the calculation, but the transformation step can be sped up using semi-direct integral transformations.[68] Additionally, we have optimized the two states for each molecule and compiled the results in Table VII. Most noticeably, CCSD(T) predicts a wrong ground state for the BN molecule and CCSD — for the C_2 molecule. DCSD gives the correct answer for both systems (although the absolute deviations of the energy splittings from the reference values are similar to CCSD(T) ones), and the equilibrium bond lengths from DCSD are on a par with or better than those from CCSD(T).

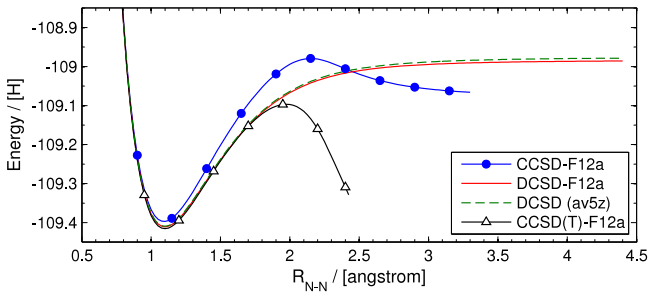

FIG. 8. N_2 dissociation computed with CCSD-F12a, DCSD-F12a, and CCSD(T)-F12a using the aug-cc-pVTZ basis set, and DCSD using aug-cc-pV5Z basis set.

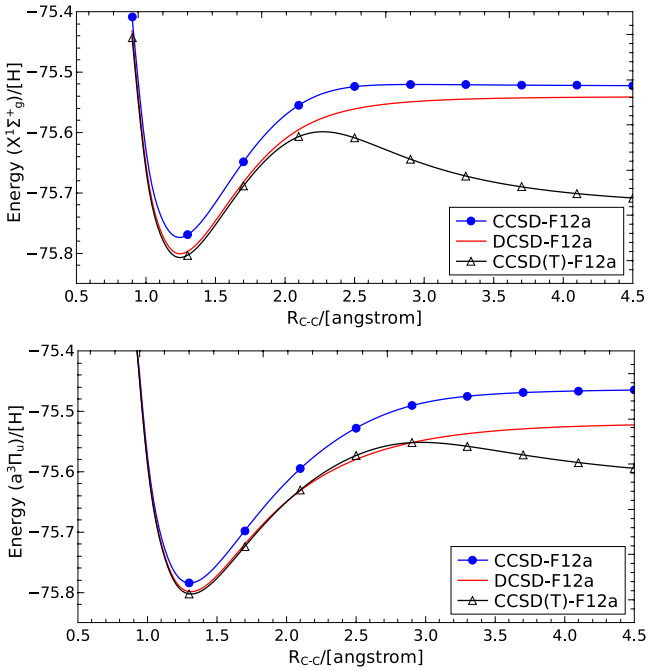

FIG. 9. C_2 $X^1\Sigma_g^+$ (top) and $a^3\Pi_u$ (bottom) potential energy functions computed with CCSD-F12a, DCSD-F12a, and CCSD(T)-F12a using the aug-cc-pVTZ basis set.

Twisted ethene is another system often used as a test for methods capable of treating multireference problems.[71–73] Since DCSD shows great stability in strong-correlation regimes and produces excellent optimized geometries, we will now test its ability to reproduce the correct geometry of ethene by rotating along the C–C axis and relaxing all other coordinates (r_{CC}, r_{CH}, and ∠HCH).

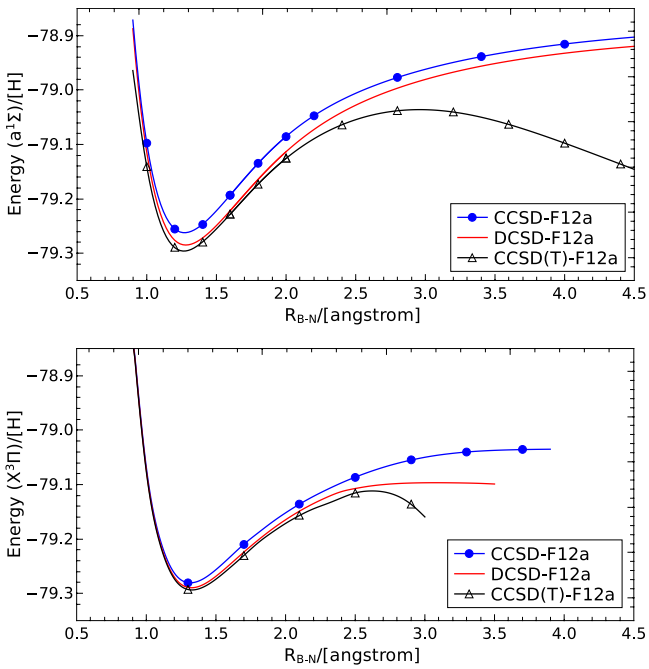

FIG. 10. BN $a^1\Sigma^+$ (top) and $X^3\Pi$ (bottom) potential energy functions computed with CCSD-F12a, DCSD-F12a, and CCSD(T)-F12a using the aug-cc-pVTZ basis set.

TABLE VII. Equilibrium bond-lengths for $^1\Sigma^+$ and $^3\Pi$ states, and the energy splitting between them, of BN and C_2, computed using F12a methods and the aug-cc-pVTZ basis set.

	Bond length (Å)		
	$^1\Sigma^+$	$^3\Pi$	$E_3 - E_1$ (cm^{-1})
		BN	
CCSD	1.271	1.317	−4026
CCSD(T)	1.267	1.328	500
DCSD	1.280	1.325	−1044
icMRCI + Q/CBS[a]	1.277	1.328	−305
CCSDTQP/CBS[b]			−282
Expt.[c]	1.283	1.329	
		C_2	
CCSD	1.242	1.306	−2328
CCSD(T)	1.245	1.314	956
DCSD	1.245	1.313	421
icMRCI+Q/CBS[a]	1.246	1.316	443
CCSDTQP/CBS[b]			548
Expt.[d]	1.243	1.312	716

[a]Estimated CBS internally contracted MRCI + Q results from Ref. 64.
[b]Estimated CCSDTQP/CBS results from Ref. 67.
[c]Experimental values from Refs. 69 and 70.
[d]Experimental values from Ref. 57.

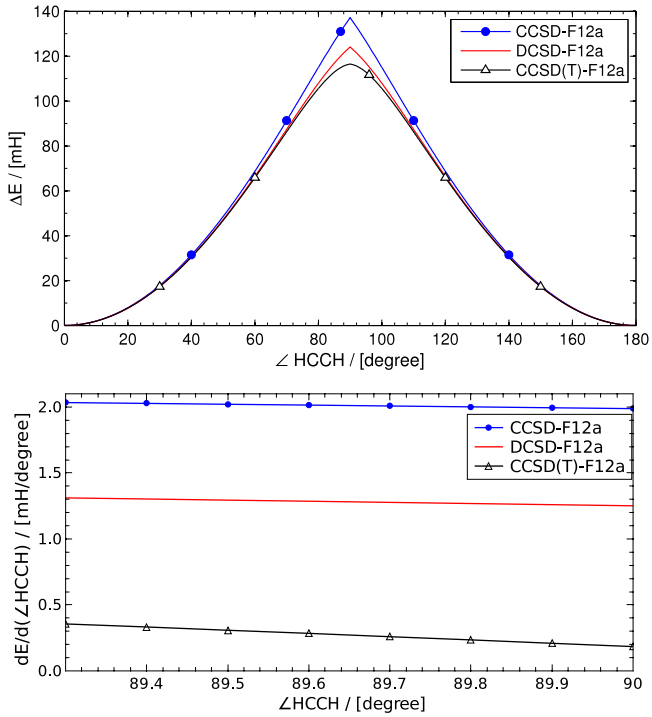

FIG. 11. Top: energy curves for the ethene twisting calculated with explicitly correlated methods in aug-cc-pVTZ basis set with the frozen core approximation. Bottom: energy-derivative curves calculated with third-order backward finite differences.

The calculations were done without the frozen-core approximation using the cc-pVTZ basis set in order to be consistent with values from Ref. 73. Additionally, we have performed F12 calculations using the frozen-core approximation and the aug-cc-pVTZ basis.

The optimized parameters are listed in Table VIII. Geometries optimized at the DCSD level using the cc-pVTZ basis are in a close agreement with CCSDT results for equilibrium and transition state, whilst the CCSD geometries and the CCSD(T) twisted geometry differ noticeably from the CCSDT results. The frozen-core calculations are very close to experimental values for DCSD and CCSD(T); therefore, one can assume that the frozen-core DCSD-F12a twisted geometry is also very accurate.

The energy curve of DCSD has an unphysical cusp for ∠HCCH = 90°, but it is much smaller than the cusp of the CCSD curve (cf. Figure 11). Interestingly, the CCSD(T)-F12 method has a much smaller cusp for a twisted geometry, although it clearly is unable to describe the electronic structure of twisted ethene correctly.

IV. CONCLUSION

We have extended our DCSD method to include explicit correlation corrections in the form of the F12a and F12b approximations originally developed for CCSD. Extensive benchmarks have been performed for reaction energies for closed- and open-shell processes, atomization energies, ionization potentials, electron affinities, interaction energies, reaction-

TABLE VIII. Optimized parameters (in Å and degrees) for equilibrium and twisted ethene geometries using aug-cc-pVTZ basis set and F12a methods.

	Equilibrium			Twisted		
	r_{CC}	r_{CH}	∠HCH	r_{CC}	r_{CH}	∠HCH
			cc-pVTZ			
CCSD	1.3271	1.0772	117.08	1.4351	1.0825	116.12
CCSD(T)	1.3330	1.0788	117.20	1.4753	1.0804	117.67
CCSDT[a]	1.3329	1.0788	117.18	1.4513	1.0826	116.76
DCSD	1.3317	1.0783	117.16	1.4507	1.0824	116.74
			F12a, aug-cc-pVTZ, frozen core			
CCSD	1.3280	1.0806	117.05	1.4360	1.0862	116.08
CCSD(T)	1.3339	1.0824	117.16	1.4761	1.0841	117.58
DCSD	1.3325	1.0817	117.12	1.4516	1.0860	116.68
Expt.[b]	1.3342	1.0812	117.37	...	...	...

[a]CCSDT results are taken from Ref. 73.
[b]Experimental values taken from Ref. 47.

barrier heights, spectroscopic constants, geometry optimizations, and systems with multireference character.

The DCSD method showed great improvement over CCSD results in all tested cases with improvements up to five times in accuracy (for spectroscopic constants). Especially big improvements were encountered for optimized geometries, where DCSD even outperformed CCSD(T) for more than 85% of the bond lengths in our test sets. We have shown that the F12x approximations when applied to DCSD can be also used for systems with a large amount of static correlation, and geometry optimizations of twisted ethene yielded geometries similar to much more demanding highly accurate CCSDT calculations.

The F12a version usually showed better results for basis sets of triple-zeta quality, which is consistent with the previous benchmarks on F12x methods.

In order to reduce the scaling of this method, one can use local approximations developed for coupled-cluster theory.[74] The resulting methods would additionally benefit from the reduced number of quadratic terms in the amplitude equations, since the remaining terms can be efficiently implemented using density-fitting techniques.[75]

Overall, one can conclude that DCSD-F12 is a very promising method for geometry optimizations of single- and multireference systems, but also for computation of thermochemical energies and other energy-based properties.

ACKNOWLEDGMENTS

Financial support from the ERC (Advanced Grant ASES 320723) is acknowledged.

APPENDIX: DCSD-AMPLITUDE EQUATIONS

$$\begin{aligned} R_{ab}^{ij} &= (ai|bj) + (ac|bd)D_{cd}^{ij} \\ &\quad + \alpha_{ijkl}D_{ab}^{kl} - (kc|ld)T_{cd}^{ij}T_{ab}^{kl} \\ &\quad + \mathcal{P}(ia;jb)\left\{x_{ac}T_{cb}^{ij} - x_{ki}T_{ab}^{kj} + (ai|bc)t_c^j\right. \\ &\quad - \left[(ac|kd)D_{cd}^{ij} + (ac|kj)t_c^i + k_{aikj}\right]t_b^k \\ &\quad \left. + \tilde{T}_{ac}^{ik}Y_{cb}^{kj} - \frac{1}{2}T_{ac}^{ki}Z_{cb}^{kj} - T_{ac}^{kj}Z_{cb}^{ki}\right\}, \end{aligned} \tag{A1}$$

$$\begin{aligned} r_a^i &= f_{ai} + \breve{f}_{ac}t_c^i - \breve{x}_{ki}t_a^k - [2(lc|ki) - (kc|li)]T_{ac}^{kl} \\ &\quad + [2(ai|kc) - (ki|ac)]t_c^k \\ &\quad + \left[f_{kc} + L_{cd}^{kl}t_d^l\right]\tilde{T}_{ca}^{ki} + (ac|kd)\tilde{D}_{cd}^{ik}, \end{aligned} \tag{A2}$$

with

$$\begin{aligned} \alpha_{ijkl} &= (ki|lj) + (kc|ld)D_{cd}^{ij} + t_c^i(kc|lj) + t_c^j(lc|ki), \\ \bar{f}_{ac} &= f_{ac} - \frac{1}{2}\tilde{T}_{ad}^{kl}(ld|kc), \\ \bar{f}_{ki} &= f_{ki} + \frac{1}{2}\tilde{T}_{cd}^{il}(ld|kc), \\ x_{ac} &= \bar{f}_{ac} - \left[f_{kc} + L_{cd}^{kl}t_d^l\right]t_a^k + [2(ac|kd) - (kc|ad)]t_d^k, \\ x_{ki} &= \bar{f}_{ki} + f_{kc}t_c^i + t_c^l l_{lcki}, \\ \breve{f}_{ac} &= \bar{f}_{ac} - \frac{1}{2}\tilde{T}_{ad}^{kl}(ld|kc), \\ \breve{x}_{ki} &= x_{ki} + \frac{1}{2}\tilde{T}_{cd}^{il}(ld|kc), \\ Y_{cb}^{kj} &= (kc|bj) + (kc|bd)t_d^j - \frac{1}{2}\left[(kj|bc) + (bc|kd)t_d^j\right. \\ &\quad - \left.(kc|ld)\tilde{T}_{db}^{lj} + l_{kclj}t_b^l\right], \\ Z_{cb}^{kj} &= (kj|bc) + (bc|kd)t_d^j - k_{clkj}t_b^l, \\ k_{aikj} &= (ai|kj) + (ai|kc)t_c^j, \\ l_{lcki} &= 2(lc|ki) - (kc|li) + L_{dc}^{kl}t_d^i, \\ L_{cd}^{kl} &= 2(kc|ld) - (kd|lc), \\ D_{ab}^{ij} &= T_{ab}^{ij} + t_a^i t_b^j, \\ \tilde{T}_{ab}^{ij} &= 2T_{ab}^{ij} - T_{ab}^{ij}, \\ \mathcal{P}(ia;jb)X_{ab}^{ij} &= X_{ab}^{ij} + X_{ba}^{ji}. \end{aligned} \tag{A3}$$

f_{pq} denotes Fock matrix, $(pq|rs)$ — two-electron repulsion integrals in the chemical notation, and T_{ab}^{ij} and t_a^i — distinguishable cluster amplitudes. As usual, $i,j,k,\ldots$, $a,b,c,\ldots$, and $p,q,r,\ldots$ indices denote the occupied, virtual, and general orbitals, respectively, and we assume summation over repeated indices.

[1]J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[2]G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[3]W. Meyer, Int. J. Quantum Chem. **5**, 341 (1971).
[4]J. Paldus, J. Čížek, and M. Takahashi, Phys. Rev. A **30**, 2193 (1984).
[5]P. Piecuch and J. Paldus, Int. J. Quantum Chem. **40**, 9 (1991).
[6]P. Piecuch, R. Tobola, and J. Paldus, Phys. Rev. A **54**, 1210 (1996).
[7]M. Nooijen and R. J. Le Roy, J. Mol. Struct.: THEOCHEM **768**, 25 (2006).
[8]F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).
[9]L. M. J. Huntington and M. Nooijen, J. Chem. Phys. **133**, 184109 (2010).
[10]L. M. J. Huntington, A. Hansen, F. Neese, and M. Nooijen, J. Chem. Phys. **136**, 064101 (2012).
[11]D. Kats and F. R. Manby, J. Chem. Phys. **139**, 021102 (2013).
[12]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[13]R. K. Nesbet, Phys. Rev. **109**, 1632 (1958).
[14]D. Kats, J. Chem. Phys. **141**, 061101 (2014).
[15]W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985).
[16]W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).
[17]W. Klopper and C. C. M. Samson, J. Chem. Phys. **116**, 6397 (2002).
[18]F. R. Manby, J. Chem. Phys. **119**, 4607 (2003).
[19]S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[20]S. Ten-no, J. Chem. Phys. **121**, 117 (2004).
[21]E. F. Valeev, Chem. Phys. Lett. **395**, 190 (2004).
[22]D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[23]S. Kedžuch, M. Milko, and J. Noga, Int. J. Quantum Chem. **105**, 929 (2005).
[24]H. Fliegl, W. Klopper, and C. Hättig, J. Chem. Phys. **122**, 084107 (2005).
[25]H. Fliegl, C. Hättig, and W. Klopper, Int. J. Quantum Chem. **106**, 2306 (2006).
[26]H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[27]J. Noga, S. Kedžuch, and J. Šimunek, J. Chem. Phys. **127**, 034106 (2007).
[28]T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[29]D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).
[30]G. Knizia and H.-J. Werner, J. Chem. Phys. **128**, 154103 (2008).
[31]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **129**, 071101 (2008).
[32]T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, Phys. Chem. Chem. Phys. **10**, 3358 (2008).
[33]J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, J. Chem. Phys. **128**, 174103 (2008).
[34]D. P. Tew, W. Klopper, and C. Hättig, Chem. Phys. Lett. **452**, 326 (2008).
[35]E. F. Valeev, Phys. Chem. Chem. Phys. **10**, 106 (2008).
[36]E. F. Valeev and T. D. Crawford, J. Chem. Phys. **128**, 244113 (2008).
[37]M. Torheyden and E. F. Valeev, Phys. Chem. Chem. Phys. **10**, 3410 (2008).
[38]D. Bokhan, S. Ten-no, and J. Noga, Phys. Chem. Chem. Phys. **10**, 3320 (2008).
[39]G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[40]H.-J. Werner, G. Knizia, and F. R. Manby, Mol. Phys. **109**, 407 (2011).

[41]D. Usvyat, J. Chem. Phys. **139**, 194101 (2013).
[42]C. Hampel, K. A. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).
[43]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.
[44]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).
[45]J. Zheng, Y. Zhao, and D. G. Truhlar, J. Chem. Theory Comput. **3**, 569 (2007).
[46]P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[47]K. L. Bak, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, and J. F. Stanton, J. Chem. Phys. **114**, 6548 (2001).
[48]S. Coriani, D. Marchesan, J. Gauss, C. Hättig, T. Helgaker, and P. Jørgensen, J. Chem. Phys. **123**, 184107 (2005).
[49]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[50]F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
[51]P. J. Knowles, C. Hampel, and H.-J. Werner, J. Chem. Phys. **99**, 5219 (1993).
[52]M. J. O. Deegan and P. J. Knowles, Chem. Phys. Lett. **227**, 321 (1994).
[53]L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
[54]O. Marchetti and H.-J. Werner, J. Phys. Chem. A **113**, 11580 (2009).
[55]M. S. Marshall and C. D. Sherrill, J. Chem. Theory Comput. **7**, 3978 (2011).
[56]M. S. Marshall, L. A. Burns, and C. D. Sherrill, J. Chem. Phys. **135**, 194102 (2011).
[57]K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure* (Springer, US, 1979).
[58]A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz, and J. M. L. Martin, J. Phys. Chem. A **112**, 12868 (2008).
[59]G. Knizia, "Explicitly correlated quantum chemistry methods for high-spin open-shell molecules," Ph.D. thesis (Universität Stuttgart, 2010).
[60]R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).
[61]D. Kats, D. Usvyat, and M. Schütz, Phys. Rev. A **83**, 062503 (2011).
[62]G. Wälz, D. Kats, D. Usvyat, T. Korona, and M. Schütz, Phys. Rev. A **86**, 052519 (2012).
[63]J. M. L. Martin, T. J. Lee, G. E. Scuseria, and P. R. Taylor, J. Chem. Phys. **97**, 6549 (1992).
[64]K. A. Peterson, J. Chem. Phys. **102**, 262 (1995).
[65]X. Li and J. Paldus, Chem. Phys. Lett. **431**, 179 (2006).
[66]A. Karton and J. M. L. Martin, J. Chem. Phys. **125**, 144313 (2006).
[67]J. M. Martin, Mol. Phys. **112**, 785 (2014).
[68]D. Kats, J. Chem. Phys. **141**, 244101 (2014).
[69]H. Bredohl, I. Dubois, Y. Houbrechts, and P. Nzohabonayo, J. Phys. B: At. Mol. Phys. **17**, 95 (1984).
[70]H. Bredohl, I. Dubois, Y. Houbrechts, and P. Nzohabonayo, J. Mol. Spectrosc. **112**, 430 (1985).
[71]V. Molina, M. Merchán, B. O. Roos, and P.-Å. Malmqvist, Phys. Chem. Chem. Phys. **2**, 2211 (2000).
[72]M. Barbatti, J. Paier, and H. Lischka, J. Chem. Phys. **121**, 11614 (2004).
[73]A. Melnichuk and R. J. Bartlett, J. Chem. Phys. **140**, 064113 (2014).
[74]T. Korona, D. Kats, M. Schütz, T. B. Adler, Y. Liu, and H.-J. Werner, in *Linear-Scaling Techniques in Computational Chemistry and Physics*, edited by R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski (Springer, Dordrecht, Netherlands, 2011), p. 345.
[75]D. Kats and F. R. Manby, J. Chem. Phys. **138**, 144101 (2013).