

Simplified CCSD(T)-F12 methods: Theory and benchmarks

Gerald Knizia, Thomas B. Adler, and Hans-Joachim Werner

Citation: *J. Chem. Phys.* **130**, 054104 (2009); doi: 10.1063/1.3054300

View online: <http://dx.doi.org/10.1063/1.3054300>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v130/i5>

Published by the [American Institute of Physics](#).

Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



**ALL THE PHYSICS
OUTSIDE OF
YOUR JOURNALS.**

www.physics today.org
**physics
today**

Simplified CCSD(T)-F12 methods: Theory and benchmarks

Gerald Knizia, Thomas B. Adler, and Hans-Joachim Werner^{a)}*Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany*

(Received 26 September 2008; accepted 2 December 2008; published online 2 February 2009)

The simple and efficient CCSD(T)-F12 x approximations ($x=a,b$) we proposed in a recent communication [T. B. Adler, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **127**, 221106 (2007)] are explained in more detail and extended to open-shell systems. Extensive benchmark calculations are presented, which demonstrate great improvements in basis set convergence for a wide variety of applications. These include reaction energies of both open- and closed-shell reactions, atomization energies, electron affinities, ionization potentials, equilibrium geometries, and harmonic vibrational frequencies. For all these quantities, results better than the AV5Z quality are obtained already with AVTZ basis sets, and usually AVDZ treatments reach at least the conventional AVQZ quality. For larger molecules, the additional cost for these improvements is only a few percent of the time for a standard CCSD(T) calculation. For the first time ever, total reaction energies with chemical accuracy are obtained using valence-double- ζ basis sets. © 2009 American Institute of Physics.

[DOI: [10.1063/1.3054300](https://doi.org/10.1063/1.3054300)]

I. INTRODUCTION

The CCSD(T) method is known to have an excellent intrinsic accuracy for a wide variety of applications in quantum chemistry. Usually energy differences are obtained, which reliably differ by no more than a few kJ/mol from full-configuration-interaction quality values.^{1,2} Unfortunately, the steep $O(N^7)$ cost scaling of CCSD(T), where N is a measure of the molecular size, combined with its strong basis set dependence, limits its applicability for obtaining highly accurate results to small molecules.

The problem of steep scaling with molecular size can be avoided by local approximations,^{3–10} while the basis set problem can be circumvented by the introduction of terms into the wave function *Ansatz* that depend explicitly on the interelectronic coordinates.¹¹ Following the groundbreaking work of Kutzelnigg and co-workers^{12–14} and several extensions of it,^{15–17} such explicitly correlated R12/F12 methods have found much attention in recent years. Important advances include the introduction of auxiliary basis sets¹⁸ and complementary auxiliary basis sets¹⁹ (CABSs) for resolution of the identity (RI) approximations, robust density fitting (DF) approximations for the numerous new integrals,^{20,21} nonlinear correlation factors,^{22–24} fixed amplitude *Ansätze*,²² excited state treatments,^{25,26} local treatments,^{27–30} systematic hierarchies of approximation for MP2-F12,³¹ as well as alternative *Ansätze* for MP2-R12.³² These advances greatly increased the understanding, efficiency, and applicability of F12 methods, which are now approaching a state of maturity in which nonexperts can apply them routinely. However, the largest part of prior work was dedicated to explicitly correlated methods at the MP2 level.

Although formally complete treatments of CCSD-R12 have already been given several years ago by Noga *et al.*,^{16,33} their original formulation has severe drawbacks. These limit

its practical applicability to systems where very large atomic orbital (AO) basis sets can be employed. Following this realization, Fliegl *et al.*^{34,35} introduced the simplified CC2(R12) and CCSD(R12) methods, which have recently been extended to CCSD(F12) by Tew *et al.*^{36,37} In these treatments, terms nonlinear in the F12-excitation operators are neglected, and certain other simplifying assumptions are made. Tew *et al.*^{36,37} stated that their CCSD(F12) method required about three times the computational resources of a conventional CCSD method. They demonstrated that results of quintuple-zeta quality can be obtained with triple-zeta basis sets. Implementations of full CCSD-F12^{38–40} and of nonlinear CCSD(R12) response properties⁴¹ have also been reported recently.

In this article we are concerned with the simplified CCSD-F12 x methods ($x=a,b$), which we proposed in a recent communication.⁴² Previous calculations⁴² of reaction energies (REs) suggested that these yield the same or even better accuracy than the CCSD(F12) method, although they are faster and much simpler. The iteration time in a CCSD-F12 x calculation is virtually the same as in a conventional CCSD calculation with the same basis set, and the only additional effort is to perform an initial MP2-F12 calculation. Since the latter method scales only with $\mathcal{O}(N^5)$, the time for this step quickly becomes negligible compared to the time for CCSD(T) when larger systems are considered. Due to the dramatically improved basis set convergence, computational savings of two to three orders of magnitude as compared to standard CCSD(T) calculations of similar accuracy can then be achieved. A recent development indicates that even greater enhancements are possible in a future local DF-LCCSD-F12 treatment: Suitable F12 corrections almost completely cancel the domain approximation error, the main error source in local methods.^{29,30}

We note that recently Valeev independently derived a method denoted as CCSD(2) $_{\overline{\text{R12}}}$, in which the F12 doubles

^{a)}Electronic mail: werner@theochem.uni-stuttgart.de.

basis set extension is treated as a perturbation on converged CCSD solutions.^{43–45} (For consistency with common notation, in the following we will use the acronym F12 instead of R12 to indicate that a nonlinear correlation factor has been used.) This approximation is closely related to our CCSD-F12b approximation. Apart from a different formal interpretation, the only major differences are that in our method the coupling of the conventional and explicitly correlated terms is treated self-consistently, rather than by perturbation theory with fixed CCSD amplitudes, and that we do not assume the simplifying extended Brillouin condition (EBC).

In Sec. II we review our current perspective on CCSD(T)-F12 methods in general. Based on that, we explain in detail which terms of the exact method are retained in our simplified CCSD(T)-F12a and -F12b approximations. In Sec. IV we present extensive benchmark calculations for both closed-shell and open-shell systems, which demonstrate that our method yields highly accurate results for a wide range of applications.

II. SIMPLIFIED CCSD-F12 THEORY

In the following, the indices i, j, \dots, p will denote occupied orbitals, a, b, \dots, d will denote virtual orbitals, and r, s, \dots, u will denote any orbitals representable in the AO basis. α, \dots, γ will denote the (orthonormal) orbitals of a formally complete virtual space. This can be partitioned into the virtual (external) orbital basis set $\{a, b, \dots\}$ and the CABS (Ref. 19) $\{x, y, \dots\}$, so that $|\alpha\rangle\langle\alpha|$ can be approximated by $|a\rangle\langle a| + |x\rangle\langle x|$ (here and in the following summation over repeated dummy indices is implied). The one-electron operators $\hat{o}_n = |i_n\rangle\langle i_n|$, $\hat{v}_n = |a_n\rangle\langle a_n|$ project onto the occupied and virtual orbital subspaces, and their subscripts refer to the electron coordinates on which they act.

For the sake of simplicity, all equations will be given only for the closed-shell case, but the extension to the spin-unrestricted open-shell case [using spin-restricted open-shell Hartree–Fock (RHF) reference functions] is straightforward. Some aspects of the open-shell implementation will be discussed in Sec. II E.

A. Form of the wave function

The wave function employed in the full CCSD-F12 approach has the form

$$|\Psi\rangle_{\text{CCSD-F12}} = \exp(\hat{T}_1 + \hat{T}_2)|\Phi\rangle, \quad (1)$$

where Φ is the Hartree–Fock reference function. The single and double excitation cluster operators \hat{T}_1 and \hat{T}_2 , respectively, are defined as

$$\hat{T}_1 = t_a^i \hat{E}_i^a, \quad (2)$$

$$\hat{T}_2 = \frac{1}{2} T_{ab}^{ij} \hat{E}_{ij}^{ab} + \frac{1}{2} T_{\alpha\beta}^{ij} \hat{E}_{ij}^{\alpha\beta}, \quad (3)$$

where \hat{E}_i^a and $\hat{E}_{ij}^{\alpha\beta} = \hat{E}_i^{\alpha} \hat{E}_j^{\beta}$ are the usual spin-free one- and two-electron excitation operators. The excitations into the standard virtual orbitals a, b in Eq. (2) and the first part of Eq. (3) are the same as those used in the conventional CCSD

theory. The additional amplitudes $T_{\alpha\beta}^{ij}$ are given implicitly by

$$T_{\alpha\beta}^{ij} = T_{mn}^{ij} \mathcal{F}_{\alpha\beta}^{mn}, \quad (4)$$

$$\mathcal{F}_{\alpha\beta}^{mn} = \langle mn | F_{12} \hat{Q}_{12} | \alpha\beta \rangle, \quad (5)$$

$$\hat{Q}_{12} = (1 - \hat{o}_1)(1 - \hat{o}_2)(1 - \hat{v}_1 \hat{v}_2), \quad (6)$$

where $F_{12} \equiv F(r_{12})$ is a suitable short-range correlation factor and T_{mn}^{ij} are the actual amplitudes used in the F12 treatment. The projector \hat{Q}_{12} is necessary to make the F12 configurations,

$$|\Phi_{ij}^{mn}\rangle = \mathcal{F}_{\alpha\beta}^{mn} \hat{E}_{ij}^{\alpha\beta} |\Phi\rangle, \quad (7)$$

orthogonal to the configurations in the molecular orbital (MO) space. Thus, $\mathcal{F}_{rs}^{mn} = 0$, and therefore $T_{rs}^{ij} = 0$.

The explicitly correlated terms improve the wave function's description of electrons coming close to each other. They augment the conventional CI expansion by additional functions, in which the orbital products $\varphi_i(\mathbf{r}_1)\varphi_j(\mathbf{r}_2)$ in Φ have been replaced by short-range pair correlation functions,

$$|u_{ij}(\mathbf{r}_1, \mathbf{r}_2)\rangle = T_{mn}^{ij} \hat{Q}_{12} F_{12}(r_{12}) |\varphi_m(\mathbf{r}_1)\varphi_n(\mathbf{r}_2)\rangle. \quad (8)$$

The terms for $mn=ij$ and $mn=ji$ are the most important ones. The function $|u_{ij}\rangle$ represents a negative short-range hole in the orbital product $|ij\rangle$. If added to the reference function, it directly suppresses the probability of finding the two $|ij\rangle$ electrons in a spatial configuration where they are close to each other. In contrast to orbital products, the short-range correlation factor $\hat{Q}_{12} F_{12} |\varphi_m \varphi_n\rangle$ can describe the wave function cusp for $r_{12} \rightarrow 0$ correctly. Both aspects counter major deficiencies of conventional wave function expansions in terms of Slater determinants.

In practice we use an exponential correlation factor fitted to a set of Gaussian geminals,

$$F(r_{12}) = -\frac{1}{\gamma} \exp(-\gamma r_{12}) \approx \sum_i c_i \exp(-\alpha_i r_{12}^2). \quad (9)$$

In the current work we use six Gaussians, and the coefficients are fitted as described in Ref. 31. This choice and other nonlinear factors have been investigated previously,^{22–24} and the simple Slater function was found to work very well. This Ansatz resembles analytic solutions of the hydrogen problem and parts of the helium problem, as already discussed by Hylleraas.¹¹

The well known short-range asymptotic behavior of the correlation cusp of the exact wave function^{46,47} can be used to great advantage in the F12 theory. As pointed out by Ten-no,⁴⁸ a wave function $T_{mn}^{ij} |\Phi_{ij}^{mn}\rangle$ [see Eq. (7)] with amplitudes T_{mn}^{ij} fixed to

$$T_{ii}^{ii} = t_s, \quad T_{ij}^{ij} = \frac{1}{2}(t_s + t_t), \quad T_{ji}^{ji} = \frac{1}{2}(t_s - t_t), \quad (10)$$

where $i \neq j$, $t_s = 1/2$, $t_t = 1/4$, and the remaining amplitudes are set to zero, fulfills these asymptotic conditions. The system- and position-dependent parameters Φ_0 and \mathbf{w} found in Eqs. (21) and (23) of Ref. 47 are given by the Hartree–Fock function in this case.

We use this fixed amplitude *Ansatz* in our F12 x approximations. It is free of geminal basis set superposition errors (BSSE), unitarily invariant, and size consistent. More details about the fixed amplitude *Ansatz* and its advantages and disadvantages compared with other possible *Ansätze* have been discussed elsewhere.⁴⁸⁻⁵¹

We note that due to the projector \hat{Q}_{12} in Eq. (8), only those parts of the F12 configurations are retained, which are not expressible by products of virtual orbitals. Thus, even if the F12 configurations are approximated by a fixed form in the intermediate r_{12} -region, the coarse form of the wave function is still treated accurately by CCSD because it is well described by the conventional contributions in its coupled cluster expansion.

B. The CCSD-F12 amplitude equations

A compact explicitly correlated form of the CCSD equations can be obtained by inserting the wave function (1) as defined in Sec. II A into the time-independent Schrödinger equation and projecting from the left with the contravariant configurations

$$|\tilde{\Phi}_i^a\rangle = \frac{1}{2}\hat{E}_i^a|\Phi\rangle, \quad (11)$$

$$|\tilde{\Phi}_{ij}^{ab}\rangle = \frac{1}{6}(2\hat{E}_{ij}^{ab} - \hat{E}_{ji}^{ab})|\Phi\rangle, \quad (12)$$

which have the convenient properties

$$\langle\tilde{\Phi}_i^a|\Psi_{\text{CCSD-F12}}\rangle = t_a^i, \quad (13)$$

$$\langle\tilde{\Phi}_{ij}^{ab}|\Psi_{\text{CCSD-F12}}\rangle = T_{ab}^{ij} + t_{ab}^{ij}. \quad (14)$$

This yields the energy and the singles and doubles residuals,

$$E = \langle\Phi|\hat{H}|\Psi_{\text{CCSD-F12}}\rangle, \quad (15)$$

$$R_a^i = \langle\tilde{\Phi}_i^a|\hat{H} - E|\Psi_{\text{CCSD-F12}}\rangle, \quad (16)$$

$$R_{ab}^{ij} = \langle\tilde{\Phi}_{ij}^{ab}|\hat{H} - E|\Psi_{\text{CCSD-F12}}\rangle. \quad (17)$$

We note that the dependence on the energy in the residual equations cancels out automatically, and this formulation is equivalent to other forms in which a similarity transformed Hamiltonian is used. The optimum amplitudes t_a^i and T_{ab}^{ij} are determined by the conditions $R_a^i=0$ and $R_{ab}^{ij}=0$, while the amplitudes of the explicitly correlated configurations are given by Eq. (10). Thus, the number of equations is the same as in the conventional CCSD theory, but the equations contain additional contributions from the explicitly correlated terms.

In some of the new terms in the doubles residuals R_{ab}^{ij} exact RIs such as

$$\hat{Q}_{12}|\alpha\beta\rangle\langle\alpha\beta| = \hat{Q}_{12} \quad (18)$$

can be used to eliminate the summations over the formally complete virtual orbital space. However, the terms $\hat{o}_1 + \hat{o}_2$ in \hat{Q}_{12} lead to many-electron integrals, which are too numerous and expensive to be evaluated exactly. In order to avoid these integrals, the projector \hat{Q}_{12} is approximated using RIs,^{18,31}

$$\hat{Q}_{12} \approx 1 - |rs\rangle\langle rs| - |xi\rangle\langle xi| - |ix\rangle\langle ix|. \quad (19)$$

The unit operator leads to new two-electron integrals, which often can be evaluated analytically. The remaining terms lead to sums of products of two-electron integrals. The dominant contributions are those arising from $1 - |rs\rangle\langle rs|$. These require no approximations. The last two terms in Eq. (19) are approximated by representing the CABS orbitals x in the union of the AO and an auxiliary RI basis.

Unfortunately, there are also contributions where neither the integrals arising from the unit operator in Eq. (19) can be computed analytically nor Eq. (18) be used. In these cases the projector has to be approximated as

$$\hat{Q}_{12} \approx |ax\rangle\langle ax| + |xa\rangle\langle xa| + |xy\rangle\langle xy|, \quad (20)$$

where x, y run over the finite CABS. This form is less preferable than Eq. (19) since the unit operator in Eq. (19) is not treated analytically but approximated by a double RI. More technical aspects of deriving practical working equations for MP2-F12 using this notation can be found in Refs. 31 and 50.

The CCSD-F12 doubles residual [Eq. (17)] can be written in matrix form as

$$\mathbf{R}_{\text{CCSD-F12}}^{ij} = \mathbf{R}_{\text{MP2-F12}}^{ij} + \mathbf{K}(D^{ij}) + \mathbf{K}(T^{ij}) + \alpha_{ij,kl}\mathbf{D}^{kl} + \mathbf{G}^{ij} + \mathbf{G}^{ji\dagger}, \quad (21)$$

where here and in the following all quantities in bold face represent matrices in the basis of virtual orbitals (a, b). The general form of this expression and the meaning of the individual terms are given in Ref. 52, except that in that work the Fock-operator terms in the MP2 residual were included in the matrices \mathbf{G}^{ij} . In the following, we will consider the individual terms in some detail.

The first term, $\mathbf{R}_{\text{MP2-F12}}^{ij}$, is the MP2-F12 residual for the conventional amplitudes (*Ansatz* 3),

$$\mathbf{R}_{\text{MP2-F12}}^{ij} = \mathbf{K}^{ij} + \mathbf{fT}^{ij} + \mathbf{T}^{ij}\mathbf{f} - f_{ik}\mathbf{T}^{kj} - \mathbf{T}^{ik}f_{kj} + \mathbf{C}^{mn}T_{mn}^{ij}. \quad (22)$$

f_{ab} and f_{ij} are virtual-virtual and occupied-occupied blocks of the closed-shell Fock matrix, respectively, and $K_{ab}^{ij} = \langle ab|r_{12}^{-1}|ij\rangle$ are the usual exchange integrals. The only difference to the conventional MP2 residual (in the orbital invariant form) is the last term in Eq. (22), which describes the coupling to the explicitly correlated configurations. The coupling matrices \mathbf{C}^{ij} are defined as (using approximation 3C and CABS)

$$C_{ab}^{ij} = \langle ab|(\hat{f}_1 + \hat{f}_2)\hat{Q}_{12}F_{12}|ij\rangle \approx f_{ax}F_{xb}^{ij} + F_{ax}^{ij}f_{xb}, \quad (23)$$

with $F_{a\beta}^{ij} = \langle\alpha\beta|F_{12}|ij\rangle$. The matrices \mathbf{C}^{ij} vanish if the EBC is assumed to be valid ($f_{ax}=0$). This usually has only a small effect on relative MP2-F12 energies, but nevertheless we fully include the coupling in the present work.

The second term in Eq. (21), namely, the external exchange operator $\mathbf{K}(D^{ij})$, includes all contractions of the doubles amplitudes with integrals involving three or four virtual orbitals. These operators are defined as

$$[\mathbf{K}(\mathcal{D}^{ij})]_{ab} = \langle ab|r_{12}^{-1}|rs\rangle \mathcal{D}_{rs}^{ij}, \quad (24)$$

where \mathcal{D}^{ij} are composite amplitude matrices,

$$\mathcal{D}_{cd}^{ij} = T_{cd}^{ij} + t_c^i t_d^j, \quad (25)$$

$$\mathcal{D}_{kd}^{ij} = \delta_{ki} t_d^j, \quad \mathcal{D}_{ck}^{ij} = \delta_{jk} t_c^i,$$

$$\mathcal{D}_{kl}^{ij} = 0.$$

The virtual-occupied sub-blocks of $\mathbf{K}(\mathcal{D}^{ij})$ are defined analogously and are used in the singles residuals. We note that there are some further contractions of integrals involving three external orbitals with singles amplitudes, but these are unaffected by the explicitly correlated terms in the wave function.

The correction of the external exchange operators due to the explicitly correlated terms is described by the operators $\mathbf{K}(\mathcal{T}^{ij})$. Using Eqs. (4) and (18), these operators can be written as

$$[\mathbf{K}(\mathcal{T}^{ij})]_{ab} = V_{ab}^{mn} T_{mn}^{ij}, \quad (26)$$

where

$$V_{ab}^{mn} = \langle ab|r_{12}^{-1} \hat{Q}_{12} F_{12} |mn\rangle. \quad (27)$$

Using the RI approximation in Eq. (19), this can be evaluated as

$$V_{ab}^{mn} = \langle ab|r_{12}^{-1} F_{12} |mn\rangle - \langle ab|r_{12}^{-1} |rs\rangle \langle rs|F_{12} |mn\rangle \\ - \langle ab|r_{12}^{-1} |ix\rangle \langle ix|F_{12} |mn\rangle - \langle ab|r_{12}^{-1} |xi\rangle \langle xi|F_{12} |mn\rangle. \quad (28)$$

Note that in order to achieve a compact matrix formulation, where superscripts denote different matrices and subscripts their elements, the definition of V_{ab}^{mn} is different from that of V_{kl}^{ij} in our previous work (see also Sec. II D); i.e., the subscripts and superscripts have been exchanged.

The quantities $\alpha_{ij,kl}$ and \mathbf{G}^{ij} in Eq. (21) are intermediates depending on amplitudes and integrals with at most two external orbitals. For their full definitions we refer to Ref. 52 (closed shell) and Refs. 53 and 54 (open shell). If nonlinear terms in the residual are neglected [as in the coupled electron pair approximation, CEPA(0)], $\alpha_{ij,kl} = \langle ij|r_{12}^{-1} |kl\rangle$, and no corrections arise from the explicitly correlated configurations in this term. The matrices \mathbf{D}^{ij} correspond to the external-external block of \mathcal{D}^{ij} , i.e., $[\mathbf{D}^{ij}]_{ab} = \mathcal{D}_{ab}^{ij}$.

Omitting for simplicity the contributions of the single excitations and keeping only terms that are linear in the doubles amplitudes, the matrices \mathbf{G}^{ij} take the form

$$G_{ab}^{ij} = (2T_{ac}^{ik} - T_{ac}^{ki}) K_{cb}^{kj} - T_{ac}^{ik} J_{cb}^{kj} - J_{ac}^{kj} T_{cb}^{ik} \\ + (2T_{ax}^{ik} - T_{ax}^{ki}) K_{xb}^{kj} - T_{ax}^{ik} J_{xb}^{kj} - J_{ax}^{kj} T_{xb}^{ik}. \quad (29)$$

According to Eq. (4) the amplitudes T_{ax}^{ik} are

$$T_{ax}^{ik} = T_{mn}^{ik} F_{ax}^{mn} \approx T_{mn}^{ik} F_{ax}^{mn}. \quad (30)$$

In these terms it is not possible to use Eqs. (18) and (19), and therefore in Eq. (5) the projector must be approximated by Eq. (20). Large RI basis sets are required to make these approximations accurate and meaningful.

From Eq. (29) it is obvious that the matrices \mathbf{G}^{ij} could be evaluated exactly as in the standard CISD or CEPA residuals, except that the integral and amplitude matrices are augmented by the ax or xb blocks. This would increase the computational effort from $3N_o^3 N_v^3$ to $3N_o^3 N_v^2 (N_v + N_{CA})$, where N_o , N_v , and N_{CA} are the number of (correlated) occupied, virtual, and complementary auxiliary (CA) orbitals, respectively. Typically, $N_{CA} \approx 2N_v$, and therefore the computational effort for these terms would increase by a factor of 3. This is the case, for example, in the CCSD(F12) method,^{36,37} which includes such terms.

Similar considerations hold for all other terms in the full CCSD-F12 equations. The terms nonlinear in T_{mn}^{ij} even lead to multiple simultaneous RI expansions involving Eq. (20), and then also integrals of the type K_{xy}^{kl} are needed. The computational effort would then scale quadratically with $(N_v + N_{CA})$ and thus be approximately one order of magnitude larger than for a standard CCSD calculation with the same orbital basis.

Apart from the contributions in $\mathbf{R}_{MP2-F12}^{ij}$, the only terms where the use of Eq. (20) can be avoided are the $\mathbf{K}(\mathcal{T}^{ij})$ discussed above. Fortunately, these terms turn out to be the most important ones, because they are the only contributions in CCSD that involve contractions of doubles amplitudes with integrals over three or four external orbitals (such terms do not exist at the MP2 level). Unlike the occupied space, the external space shrinks in size when truncating the basis set. Therefore, a finite AO basis truncation affects these interactions more seriously than other terms, and the $\mathbf{K}(\mathcal{T}^{ij})$ -terms are the most important ones to include in an F12 treatment beyond MP2-F12.

C. The CCSD-F12a approximation

Taking these considerations into account, we arrived at approximation CCSD-F12a, in which we neglect all contributions of the explicitly correlated configurations to the doubles residual $\mathbf{R}_{CCSD-F12a}^{ij}$ except for $\mathbf{K}(\mathcal{T}^{ij})$ and the coupling matrices \mathbf{C}^{mn} in $\mathbf{R}_{MP2-F12}^{ij}$. Thus, the total correction to the canonical CCSD doubles residual can be written in the simple form

$$\Delta \mathbf{R}_{CCSD-F12a}^{ij} = [\mathbf{V}^{mn} + \mathbf{C}^{mn}] T_{mn}^{ij}, \quad (31)$$

where \mathbf{C}^{mn} and \mathbf{V}^{mn} are defined in Eqs. (23) and (28), respectively. \mathbf{C}_{ab}^{mn} is the same as in MP2-F12 theory and does not require any extra effort. However, the full evaluation of V_{ab}^{mn} requires new integrals $\langle ab|r_{12}^{-1} |xi\rangle$.

These $|ix\rangle\langle ix|$ contributions of \mathbf{V}^{mn} in Eq. (28) can be calculated without storing large intermediates. In a DF scheme, we have

$$\langle \mu\nu|r_{12}^{-1} |ix\rangle \langle ix|F_{12} |mn\rangle \approx [(\nu x|A) D_A^{\mu i}] F_{ix}^{mn}, \quad (32)$$

where A is a fitting basis index and $D_A^{\mu i} = J_{AB}^{-1}(\mu i|B)$ is the DF coefficient giving $|\mu i\rangle \approx D_A^{\mu i} |A\rangle$. Since Coulomb integrals are concerned, the fitting formula is robust already in this form.²¹ Eq. (32) can be calculated for each ν basis function shell independently: For each ν -shell, the integrals $(\nu x|A)$ are computed for all A and x and contracted first with $D_A^{\mu i}$ and then with F_{ix}^{mn} (which are precalculated and kept in

memory). Once all ν -shells are done, the final integral can be transformed to the desired basis and added to \mathbf{V}^{mn} . However, this evaluation requires significant computer time since two virtual indices and one complementary auxiliary (CA) index are involved. Numerical experiments (see Sec. IV E) have shown that these terms have a minor impact on relative energies; therefore we neglect them by default. Thus, unless otherwise noted, the projector in \mathbf{V}^{mn} is approximated as $1 - |rs\rangle\langle rs|$.

If the contributions of the CA orbitals are neglected as described, the matrix \mathbf{V}^{mn} can be evaluated simply as

$$V_{rs}^{mn} = W_{rs}^{mn} - K(F^{mn})_{rs}, \quad (33)$$

$$W_{rs}^{mn} = \langle rs|r_{12}^{-1}F_{12}|mn\rangle, \quad (34)$$

$$K(F^{mn})_{rs} = \langle rs|r_{12}^{-1}|tu\rangle F_{tu}^{mn}. \quad (35)$$

The operator $K(F^{mn})$ has the same form as the external exchange operators $K(\mathcal{D}^{ij})$, and therefore the total residual can be written as

$$\begin{aligned} \mathbf{R}_{\text{CCSD-F12a}}^{ij} = & \mathbf{R}_{\text{MP2}}^{ij} + \bar{\mathbf{C}}^{ij} + \bar{\mathbf{W}}^{ij} + \mathbf{K}(\mathcal{D}^{ij} - \bar{F}^{ij}) + \alpha_{ij,kl} \mathbf{D}^{kl} \\ & + \mathbf{G}^{ij} + \mathbf{G}^{j\ddagger i}, \end{aligned} \quad (36)$$

where $\bar{F}_{rs}^{ij} = F_{rs}^{mn} T_{mn}^{ij}$ (\bar{W}_{rs}^{ij} and \bar{C}_{ab}^{ij} are defined analogously). Note that in the fixed amplitude *Ansatz* we simply have

$$\bar{F}_{rs}^{ij} = \frac{3}{8} F_{rs}^{ij} + \frac{1}{8} F_{sr}^{ij}. \quad (37)$$

The internal-external blocks of the matrix

$$K(\mathcal{D}^{ij} + T^{ij})_{rs} = \bar{W}_{rs}^{ij} + K(\mathcal{D}^{ij} - \bar{F}^{ij})_{rs} \quad (38)$$

are also used in the singles residual, replacing the corresponding blocks of $K(\mathcal{D}^{ij})$ in the conventional CCSD [see Eq. (32) of Ref. 52].

The operators $K(\mathcal{D}^{ij} - \bar{F}^{ij})$ can be computed directly from the integrals in the AO basis,⁵² and the only change required in the corresponding subroutines is to replace \mathcal{D}^{ij} by $\mathcal{D}^{ij} - \bar{F}^{ij}$ and to add \bar{W}^{ij} to the final operators. Since the matrices \mathbf{F}^{mn} , \mathbf{W}^{mn} , and \mathbf{C}^{mn} are needed anyway in the MP2-F12, there is virtually no additional effort in the CCSD-F12a iterations as compared to the standard CCSD. The computational effort for the initial fixed amplitude MP2-F12 scales only as $\mathcal{O}(N^5)$, and for larger molecules it takes only a small fraction of the time of a CCSD(T) calculation. [For a full MP2-F12 with optimized amplitudes T_{mn}^{ij} the computational effort would scale as $\mathcal{O}(N^6)$ (with a *very* low prefactor), provided the F12 amplitude equations are solved iteratively. The commonly cited $\mathcal{O}(N^8)$ scaling of MP2-F12 only occurs if the equations are solved noniteratively with canonical orbitals.] For instance, in an application to the benzene dimer,⁵¹ the additional cost was only about 5% (using aug-cc-pVTZ basis sets). Conventional calculations of the same quality would require at least aug-cc-pV5Z basis sets and would take more than 100 times more CPU time (if the integrals could be stored on disk). Using current computational facilities, such calculations would be impossible for the benzene dimer, since the two-electron integrals could not be stored, and integral-direct calculations would be prohibitively expensive.

In summary, the two terms of Eq. (31) represent the coupling of the F12 configurations to the conventional doubles amplitudes at the full CCSD level for all terms that involve three or four external orbitals and to second-order perturbation theory in the remaining ones. It is expected that the second-order treatment of these remaining terms involving zero to two external orbitals is sufficient. This is suggested by the good performance of MP2-F12 methods employing the cruder EBC approximation, which entirely neglects the coupling between conventional and F12 configurations. Furthermore, the neglect of terms that would require the use of Eq. (20) is equivalent to the assumption that the RI basis is equal to the AO basis. This assumption is also made in MP2-F12 hybrid approximations,^{31,50,55} which have been found to work very well. A more complete treatment is not only computationally costly but also appears to be unnecessary.

D. The energy functional

The standard CCSD-F12 energy expression is given by

$$\begin{aligned} E = \langle \Phi | \hat{H} | \Psi_{\text{CCSD-F12}} \rangle = & (2D_{ab}^{ij} - D_{ab}^{ji}) K_{ab}^{ij} \\ & + (2T_{mn}^{ij} - T_{mn}^{ji}) V_{mn}^{ij}, \end{aligned} \quad (39)$$

$$V_{mn}^{ij} = \langle mn | F_{12} \hat{Q}_{12} r_{12}^{-1} | ij \rangle, \quad (40)$$

which differs from the MP2-F12 energy^{31,50} only by using D_{ab}^{ij} rather than T_{ab}^{ij} . This expression is exact for fully optimized amplitudes. However, as will be discussed in the following, for fixed amplitudes this energy expression depends sensitively on the shape of the correlation factor, i.e., the value of γ .⁴² This is related to a strong coupling between γ and the optimized amplitudes T_{mn}^{ij} .

These effects can be most easily demonstrated at the MP2-F12 level. In this case the Hylleraas functional provides a variational expression that is a strict upper bound to the exact MP2-F12 energy. In contrast to the linear energy expression (39), the Hylleraas functional depends only quadratically on errors of the first-order wave function. Such errors can be introduced by nonoptimized values of γ and/or T_{mn}^{ij} .

Figure 1 shows the dependence of the linear and quadratic energy expressions for $\gamma=1.0, 1.5, 2.0 a_0^{-1}$ as a function of a single fixed coefficient t using $t_s=t$ and $t_t=t/2$. As expected, the linear energy expression (open symbols, dashed lines) depends almost linearly on the coefficient t , while the quadratic Hylleraas functional (filled symbols and full lines) shows a much weaker dependence on t and has a minimum. The linear and quadratic energy curves cross very close to the minima. The positions of the minima depend on γ : The larger γ is, the larger the optimum amplitude t is. However, the energy of the minima depends only very slightly on γ ; the lowest energy is obtained for $\gamma=1.5 a_0^{-1}$, and in this case the optimum coefficient is $t=0.5$, i.e., the value that fulfills the cusp condition. Thus, it appears that nonoptimum γ -values can be compensated to a large extent by variation of the amplitudes, even if this violates the exact cusp condition. This is due to the fact that the region very

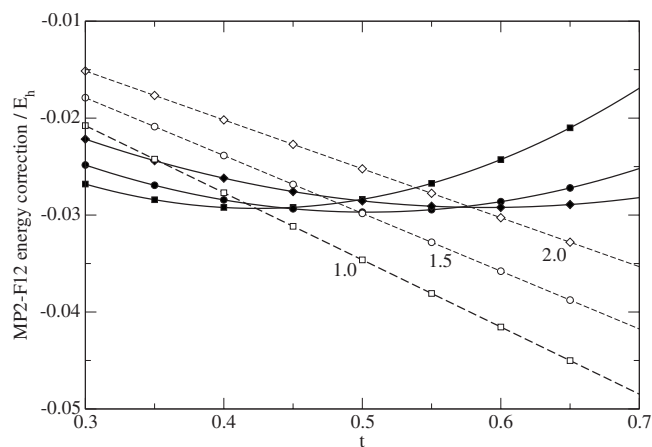


FIG. 1. Dependence of the MP2-F12 energy correction of H_2O on the choice of the fixed amplitudes $t_s=t$ and $t_i=t/2$ for $\gamma=1.0 a_0^{-1}$ (squares), $1.5 a_0^{-1}$ (circles), and $2.0 a_0^{-1}$ (diamonds); basis AVTZ. The full lines with full symbols represent the energy computed by the Hylleraas functional; the dashed lines with open symbols represent the linear energy approximation (see text).

close to the cusp does hardly contribute to the energy (since the volume element in the integration goes to zero). The region for medium values of r_{12} is most important.

Figure 1 also demonstrates that for fixed amplitudes (e.g., $t=0.5$) the dependence on γ is much stronger for the linear than for the quadratic energy expression. A plot of the γ -dependence for $t=0.5$ can be found in Ref. 42. Unfortunately, the γ -dependence depends on the molecule. Optimizing γ for each molecule individually would not only be inconvenient and expensive but would also violate size extensivity.

From these considerations it follows that with fixed amplitudes it is essential to use an energy functional that is approximately quadratic in the amplitudes and minimizes the γ -dependence. In principle, this could be achieved by using a Lagrangian formulation. But this would require obtaining the full CCSD-F12 residual R_{kl}^{ij} for the explicitly correlated amplitudes and solving the Λ -equations.⁵⁶ We have chosen a simpler approach: In analogy to the MP2 Hylleraas functional, we replace in Eq. (39) V_{kl}^{ij} by $V_{kl}^{ij}+R_{kl}^{ij}$, where R_{kl}^{ij} is the MP2-F12 residual (using canonical orbitals),

$$R_{kl}^{ij} = V_{kl}^{ij} + [B_{kl,mn} - (\varepsilon_i + \varepsilon_j)X_{kl,mn}]T_{mn}^{ij} + C_{ab}^{kl}T_{ab}^{ij}. \quad (41)$$

$B_{kl,mn}$ and $X_{kl,mn}$ are defined and computed as described in Refs. 31 and 50 using approximation 3C. In the last term, we use the converged CCSD amplitudes T_{ab}^{ij} .

One can consider including further terms in the residual R_{kl}^{ij} , which arise in the CCSD-F12 case. For the case of linearized CCD [CEPA(0)], the form of R_{kl}^{ij} is given in the Appendix. One possibility is to include only the terms arising from the external exchange operators $\mathbf{K}(\mathcal{D}^{ij})$, as for the residuals R_{ab}^{ij} . This leads to an additional energy correction⁴²

$$\Delta E = (2T_{mn}^{ij} - T_{nm}^{ij})\text{tr}[(\mathbf{W}^{mn} - \mathbf{K}(\mathbf{F}^{mn}))\mathbf{D}^{ij}]. \quad (42)$$

The resulting approximation has been denoted as CCSD-F12b. We found, however, that the inclusion of this term, which roughly doubles the coupling between the conventional and explicitly correlated amplitudes, does not lead to

an improvement of basis set convergence for small- and medium-sized basis sets. Usually, it leads to a substantial underestimation of the correlation energies, particularly for open-shell systems. On the other hand, our experience has shown that unlike F12a correlation energies, which often slightly overestimate the basis set limit, the absolute values of the F12b correlation energies converge systematically from below to the complete basis set (CBS) limit.

It is possible that consistently improved results combined with systematic CBS convergence can only be obtained if all of the remaining CEPA(0)-terms in the R_{kl}^{ij} residuals are included at least to linear order in the T_{mn}^{ij} -amplitudes. As long as no additional conventional/F12 couplings are introduced (which, as mentioned in Sec. II C, are unlikely to be necessary), this can again be done without impact on the conventional CCSD iterations. We reserve the name F12c for such a treatment.

In the perturbative CCSD(2)_{F12} approximation recently proposed by Valeev⁴³ a correction $2\Delta E$ is added to the conventional CCSD energy. In his case the amplitudes T_{mn}^{ij} are fully optimized at the MP2-F12/*C level, and \mathbf{T}^{ji} is used instead of \mathbf{D}^{ji} in the last term of Eq. (42). This method yields results that are rather similar to those of our CCSD-F12b method. Judging from recently published results on atomization energies (AEs),⁴⁵ it also underestimates the basis set limit more than MP2-F12 calculations do with the same basis set, just like our CCSD-F12b and the CCSD(F12) method of Tew *et al.*^{36,37}

E. Open-shell aspects

In the high-spin open-shell case, the reference function is built from a set of closed-shell orbitals, doubly occupied with both alpha (A)- and beta (B)-spin, and singly occupied A-spin active orbitals. The spin-free double excitation operator in Eq. (3) is replaced by a sum of spin-orbital operators representing the six spin excitations that commute with \hat{S}_z , $AA \rightarrow AA$, $BB \rightarrow BB$, $AB \rightarrow AB$, $AB \rightarrow BA$, $BA \rightarrow BA$, and $BA \rightarrow AB$. The latter four can be expressed by each other due to the anti-commutation relations of spin-orbital substitution operators. Thus only three sets (AA , BB , and $AB \rightarrow AB$) of spin-orbital amplitudes must be stored, and three sets of residual equations must be solved. All operators refer to the same set of spin-restricted spatial orbitals.

Efficient open-shell CCSD formulations in spin-unrestricted (RHF-UCCSD) and partially spin-restricted (RHF-RCCSD) forms have been given before.^{53,54} Corresponding perturbative triples corrections are also available.⁵⁷ A RMP2-F12 formulation compatible with these forms of CCSD was developed recently.⁵⁰ Most arguments of the previous sections hold unaltered for this open-shell case, and the equations can be transferred into this regime easily. (Note that Refs. 50 and 53 differ in the normalization of spin-orbital T -amplitudes.) But there are some aspects that need special consideration:

Diagonal amplitudes: In the open-shell case, an ambiguity in the choice of the AB -spin fixed amplitudes $t_{m_A n_B}^{i_A j_B}$ arises when i is an active orbital and j is a closed-shell orbital. In this case the direct diagonal excitation $t_{i_A j_B}^{i_A j_B}$ is unproblematic,

while the reversed diagonal excitation $t_{j_A i_B}^{i_A j_B}$ refers to the *unoccupied* spin orbital i_B . Such reversed excitations were previously not included in the wave function *Ansatz* (e.g., in Ref. 50). But both excitations are necessary to form the symmetric and antisymmetric spatial pair functions required to fulfill both singlet and triplet cusp conditions at the same time.

In this work we include such excitations: The F12 part of the cluster operator is [see Eqs. (3) and (4)]

$$\hat{T}_{\text{fix}} = \frac{1}{2} T_{mn}^{ij} \mathcal{F}_{\alpha\beta}^{mn} \hat{e}_{ij}^{\alpha\beta}, \quad (43)$$

where $\hat{e}_{ij}^{\alpha\beta}$ is a spin-orbital excitation operator and i and j run over occupied spin-orbitals as usual, while for *all* spins m and n run over *both* closed-shell and active orbitals – including *B*-spin, where active spatial orbitals are unoccupied. The spin orbital T_{mn}^{ij} of occupied spin-orbital pairs ij is fixed according to Eq. (10), where $t_s=0$ for *AA*- and *BB*-spins.

This *Ansatz* extension does not pose orthogonality problems because the projectors \hat{Q}_{12} still project out the overlap of F12 configurations with the conventional space. In principle one is free in the choice of the orbitals used for generating the F12-excitation manifold. Related *Ansatz* extensions were previously discussed by Neiss *et al.*²⁵ in the context of F12 treatments of excited states and by Bokhan *et al.*⁵⁸ in the context of cusp conditions.

Choice of \hat{Q}_{12} projectors: By default, we used spin-orbital projectors \hat{Q}_{12} . This means that the projectors $\hat{o} = |i\rangle\langle i|$ and $\hat{v} = |a\rangle\langle a|$ in $\hat{Q}_{12} = (1 - \hat{o}_1)(1 - \hat{o}_2)(1 - \hat{v}_1\hat{v}_2)$ are built from the occupied or unoccupied spin orbitals in the respective spin case (*A*, *B*).

Another possibility is to use projectors \hat{Q}_{12} with identical spatial parts for all spins. Then, for both *A*- and *B*-spins, i runs over both closed and active orbitals in $\hat{o} = |i\rangle\langle i|$ and a runs over pure external orbitals in $\hat{v} = |a\rangle\langle a|$. As a result, the F12 wave function is spin-free and the semi-internal contributions of all F12 pair functions are removed (i.e., the overlap with virtual spin-orbital pairs involving active spatial orbitals is projected out).

Our experiments indicated that the effect of these changes is only significant for double-zeta (aug-cc-pVDZ) basis sets. RMP2-F12 benchmarks showed better convergence when the semi-internal contributions were not removed.

Spin contamination: Although spin contamination does not directly contribute to the RMP2-F12 energy, the RMP2-F12 wave function is not completely spin-free. Its spin contaminations couple to the conventional CCSD part of the wave function in F12x schemes. We expect this effect to be small; therefore no additional spin restriction has been imposed on the F12 part of the RMP2-F12 wave function.

F. CABS single excitations

In MP2-F12 and CCSD-F12 calculations, the basis set errors of the Hartree–Fock energy are often larger than the errors of the correlation energy. The error of the HF contribution can be much reduced by including single excitations

into the CA orbital space and by computing the second-order energy contribution perturbatively.^{42,50} For closed-shell cases the CABS singles amplitudes are obtained by solving

$$f_{\alpha}^i = t_{\alpha}^k f_k^i - f_{\alpha}^{\beta} t_{\beta}^i, \quad (44)$$

where f_{α}^{β} are the matrix elements of the closed-shell Fock operator and the indices α, β run over the virtual and CA orbitals. Note that also the standard virtual orbitals need to be included since the couplings f_a^x are nonzero. The energy correction is then computed as

$$\Delta E_s = 2t_{\alpha}^i f_i^{\alpha}. \quad (45)$$

A spin-free extension for high-spin open-shell cases has been discussed in Ref. 50. This spin-free version was used exclusively in this work. Core orbitals were always omitted since they usually have a negligible effect on energy differences but are more sensitive to the choice of the RI basis and the CABS thresholds.

We consider the CABS singles correction as a correction to the HF energy and not as a correlation effect. The Fock matrix elements f_{α}^{β} and f_{α}^i are needed in MP2-F12 calculations anyway, and therefore the additional effort for this correction is negligible. It should be noted that this energy correction is entirely distinct from the CCSD calculations, and the singles amplitudes t_{β}^i used here are not the same as those in the CCSD.

G. Triple excitations

In our current implementation, the perturbative triples energy correction is computed exactly as in standard CCSD(T) calculations. Thus, the F12 treatment described in the previous sections has no direct effect on the triples energy. However, there is a small indirect effect, arising from the change in the doubles amplitudes T_{ab}^{ij} due to the coupling with the explicitly correlated terms (usually, this slightly reduces the triples energy correction). Even though the triples energy correction is fairly small, the basis set error of this correction can be larger than the basis set error of the CCSD-F12 energy and therefore can spoil the accuracy of the CCSD(T)-F12. Unfortunately, a direct inclusion of F12 terms in the triples is not straightforward and not yet available. Therefore, we have investigated the performance of a very simple basis set correction. For this correction, which we denote as (T*), the (T) triples are scaled by the factor $E_{\text{corr}}^{\text{MP2-F12}}/E_{\text{corr}}^{\text{MP2}}$. This assumes that basis set incompleteness affects the perturbative triples energy by the same factor as the MP2 correlation energy.

This correction may slightly break size consistency (but not size extensivity) if the ratios $E_{\text{corr}}^{\text{MP2-F12}}/E_{\text{corr}}^{\text{MP2}}$ are different for different dissociation products. However, this effect is minor, and even for the case of AEs, which one would expect to be the most susceptible to these kinds of errors, the correction improves the triples contributions (see Table II).

Like all F12 corrections, the (T*) approach will eventually converge to the correct CCSD(T) limit with increasing basis set size since F12 corrections go to zero if the orbital basis set becomes complete. The performance of the scaling correction will be investigated in Sec. IV B.

III. COMPUTATIONAL DETAILS

The methods described in Sec. II have been implemented into the CCSD(T) (Ref. 52) and RHF-UCCSD(T) and RHF-RCCSD(T) (Refs. 53 and 57) programs (spin-unrestricted and partially spin-restricted open-shell coupled clusters, respectively) of the MOLPRO quantum chemistry package.⁵⁹ RHF reference wave functions were used in all open- and closed-shell cases. In this work, we report on RHF-RCCSD as open-shell CCSD exclusively, but most open-shell calculations were performed with both RHF-RCCSD and RHF-UCCSD. We did not observe noteworthy differences in the basis set convergence between the two. All calculations used the frozen core approximation.

A. MP2-F12 calculations

The MP2-F12 (closed-shell) and RMP2-F12 (open-shell) calculations were performed using approximation 3C(FIX) as described in Refs. 31 and 50, respectively. In order to achieve a balanced description for closed-shell and open-shell cases, the approximations used in both cases must be equivalent; i.e., the open-shell method should exactly reproduce the results of the closed-shell program if applied to closed-shell systems. In the open-shell case it is convenient to use a projected zeroth-order Hamiltonian,⁵⁰ in which the contributions of Fock matrix elements f_{ai} and f_{xi} are projected out (the former vanishes in closed-shell cases, and the latter vanishes if the generalized Brillouin condition is assumed). These terms are then treated as a perturbation. It turns out that in the closed-shell case this is exactly equivalent to the MP2-F12/3C(+Z) method as described in Ref. 31 and the MP2-F12/C method of Kedžuch *et al.*⁶⁰ Apart from the RI approximations, this method does not make any approximations to the matrix elements. The +Z terms can be included without additional computational effort by replacing the exchange operator \hat{k} used in the $\bar{Y}_{kl,mn}$ terms [Eqs. (75) and (76) of Ref. 31] by $\hat{k} + \delta\hat{f}$, and their inclusion is now the default in our program. Unless otherwise noted, the γ -value in the correlation factor was chosen to be $1.0 a_0^{-1}$.

B. CABS treatment

All RI approximations in the MP2-F12 calculations employed orthogonal CA orbitals, which are represented in the union of the orbital and RI basis sets. The CA orbitals $|x\rangle$ were constructed by a two-step process: First, the overlap matrix of the RI functions is diagonalized, and RI functions are orthonormalized among themselves. Second, the MOs are projected out from this orthonormal set. The overlap matrix of the remaining projected basis functions is diagonalized, and orthonormal projected basis functions are obtained. In both steps, eigenvectors corresponding to eigenvalues smaller than a threshold THR are deleted. The threshold was taken to be $\text{THR} = \max(10^{-9}, s_{\max} \cdot 10^{-9})$, where s_{\max} is the largest eigenvalue of the corresponding overlap matrix. (In MOLPRO, this is specified as `explicit,ortho_cabs=1,thrcabs=1d-9,thrcabs_rel=1d-9`.) Using such small thresholds, in most cases no functions are removed at all, and therefore the results are unaffected by the thresholds and the details of this

procedure. Despite the small thresholds, the results appeared to be numerically stable, even with large basis sets.

The CA orbitals thus obtained can be written as

$$|x\rangle = |\chi_\mu\rangle C_{\mu x} + |\chi_{\mu'}\rangle \tilde{C}_{\mu' x}, \quad (46)$$

where $C_{\mu x}$ and $\tilde{C}_{\mu' x}$ are the coefficients of the AO and RI basis functions, respectively. A typical matrix element is

$$f_{\mu x} F_{x\nu}^{ij} = \bar{f}_{\mu\sigma} F_{\sigma\nu}^{ij} + \tilde{f}_{\mu\sigma'} F_{\sigma'\nu}^{ij}, \quad (47)$$

with transformation matrices

$$\bar{f}_{\mu\sigma} = f_{\mu\rho} C_{\rho x} C_{\sigma x} + f_{\mu\rho'} \tilde{C}_{\rho' x} C_{\sigma x}, \quad (48)$$

$$\tilde{f}_{\mu\sigma'} = f_{\mu\rho} C_{\rho x} \tilde{C}_{\sigma' x} + f_{\mu\rho'} \tilde{C}_{\rho' x} \tilde{C}_{\sigma' x}. \quad (49)$$

We found that the numerical accuracy depends sensitively on the order in which the summations in Eqs. (48) and (49) are performed. In the method as described in Ref. 31 the summation over x was carried out first, and then the resulting matrices were multiplied with the Fock (or exchange) matrices. Meanwhile it turned out, however, that much better numerical accuracy is achieved if one first multiplies the Fock matrices with the transformation matrices $C_{\rho x}$ ($\tilde{C}_{\rho' x}$) and then multiplies them with the second transformation matrix. Alternatively, one can first transform the Fock and F^{ij} matrices to the orthogonal MO/CA basis and then perform the summation over x .

C. Basis sets

As orbital basis sets, we used the aug-cc-pVnZ basis sets⁶¹ for first-row elements and the aug-cc-pV(n+d)Z basis sets⁶² for second-row elements. For the sake of brevity, these basis sets will be denoted as “AVnZ” in the remainder of this article. In order to estimate the basis set limits, conventional CCSD(T)/AV5Z and CCSD(T)/AV6Z doubles and triples correlation energies were extrapolated using the formula^{63,64} $E_n = E_{\text{CBS}} + a/n^3$; the corresponding results are denoted as CCSD(T)/CBS[56]. Hartree–Fock energies and the direct singles correlation contributions in open-shell CCSD (t_{ad}^{ia} for spins A, B) were not extrapolated but taken directly from the larger basis set. For some of the larger molecules, the AV6Z calculations were not feasible. In these cases the AVQZ and AV5Z basis sets have been used in the extrapolations (CCSD/CBS[45]).

In the F12 calculations of electron affinities (EAs), the AVnZ orbital basis sets were augmented by one diffuse s -shell for hydrogen and one diffuse s -shell and one p -shell for other atoms. For AVQZ, additionally one diffuse p -shell was added for hydrogen and one diffuse d -shell for other atoms. The (nonoptimized) exponents of these functions were obtained by dividing the lowest exponent of the non-augmented sets by 1.8. The resulting basis sets, denoted as “A2VnZ,” are the same as that already used in previous benchmark calculations with the RMP2-F12 method.⁵⁰ It was found that the additional diffuse functions significantly improve the accuracy of the computed EAs with the RMP2-F12 method.

TABLE I. Comparison of approximate and exact CCSD-F12 valence correlation energies for H₂O and F₂ (in mH). The CCSD/CBS[56] values are -297.9 and -601.17 mH, respectively.

Method	H ₂ O			F ₂		
	AVDZ	AVTZ	AVQZ	AVDZ	AVTZ	AVQZ
CCSD	-227.11	-273.05	-288.21	-435.39	-538.91	-575.10
CCSD-F12a(FIX)	-293.22	-298.50	-299.37	-590.81	-599.65	-602.89
CCSD-F12b(FIX)	-287.27	-294.33	-296.88	-579.81	-591.43	-597.71
CCSD-F12a(FIX) ^a	-293.95	-298.65	-299.40	-592.81	-600.02	-602.99
CCSD-F12b(FIX) ^a	-288.25	-294.59	-296.95	-582.34	-592.07	-597.90
CCSD-F12a(FULL) ^{a, b}	-295.17	-298.80	-299.55	-595.46	-601.41	-603.80
CCSD-F12b(FULL) ^{a, b}	-288.79	-294.65	-297.10	-584.40	-593.67	-598.90
CCSD(2) _{F12} ^c	-290.72	-294.75	-296.94	-590.95	-593.49	-598.15
CCSD(2) _{F12} ^d	-291.51	-295.42	-297.40	-592.20	-594.54	-598.94
CCSD(F12) ^d	-290.33	-295.81	-297.54	-585.07	-595.52	-599.69
CCSD-F12 ^d	-289.86	-295.40	-297.23	-584.83	-594.98	-599.01

^aIncluding the $|\dot{i}\dot{x}\rangle\langle\dot{i}\dot{x}| + |\dot{x}\dot{i}\rangle\langle\dot{x}\dot{i}|$ contributions in Eq. (28).

^bUsing full MP2-F12 T_{mn}^{ij} -amplitudes.

^cMethod as described in Ref. 43, computed with our program.

^dValues taken from Ref. 39.

The MO/CABS and CABS/CABS blocks of the Fock and exchange matrices used in the MP2-F12 calculations were computed using DF approximations; the cc-pVnZ/JKFIT auxiliary basis sets of Weigend⁶⁵ were used for this purpose. In the case of EAs, they were augmented by one additional diffuse shell per angular momentum, as described in Ref. 50.

The MO/MO blocks of the Fock matrices used in the MP2-F12 and CCSD residual equations were not density fitted but taken directly from the RHF calculation. In the MP2-F12 part, all other integral quantities were obtained by robust DF using the aug-cc-pVnZ/MP2FIT basis sets of Weigend *et al.*⁶⁶ For the RI approximations, the cc-pVnZ/JKFIT basis sets were used.

Except for double- ζ calculations, the cardinal numbers n of the RI, JKFIT, and MP2FIT basis sets were chosen to be equal to that of the orbital basis sets (e.g., VTZ/JKFIT and AVTZ/MP2FIT were used for AVTZ orbital calculations). In AVDZ/A2VDZ calculations the same auxiliary basis sets were used as in the AVTZ/A2VTZ calculations because no VDZ/JKFIT basis sets are available at this moment. The RHF and CCSD(T) calculations did not employ any DF or other approximations.

IV. RESULTS

Most results will be presented in statistical or graphical form. Individual data for some representative benchmarks are provided in the supplementary material.⁶⁷ Unless otherwise noted, all values in tables and figures refer to differences of total energies, including the RHF contribution and the CABS singles correction.

A. Comparison with exact CCSD-F12 correlation energies

Very recently Shiozaki *et al.*³⁹ presented exact CCSD-F12 calculations for Ne, H₂O, and F₂. In this section we compare the exact CCSD-F12 correlation energies with values from our CCSD-F12 x approximations and values of the other approximate schemes CCSD(F12) of Fliegl *et al.*³⁴ and CCSD(2)_{F12} of Valeev and Crawford.⁴⁵

The same geometries and the 19s14p10d8f6g4h2i/9s6p4d3f2g RI basis set as defined in Refs. 39 and 43 were used. The exponents and contraction coefficients of the Gaussian geminals [see Eq. (9)] were taken from Ref. 43 and correspond to $\gamma = 1.5$. For the DF of the integrals in the MP2-F12, we used the V5Z/JKFIT basis for the Fock and exchange matrices and the AV5Z/MP2FIT basis for all other integrals.

The results for H₂O and F₂ are presented in Table I. We reimplemented the CCSD(2)_{F12} method according to Ref. 43. This method differs from the one used in Ref. 39 by using a different energy functional.⁴⁵ A more detailed account of these methods and the uncommon form of EBC used in them is given in the supplementary material.⁶⁷

The results for the CCSD-F12a and CCSD-F12b methods show the same trends as already noticed in our preliminary communication.⁴² The CCSD-F12a correlation energies overestimate the exact CCSD-F12 ones for the same basis set, but the values for the AVDZ and AVTZ basis sets are closer to the basis set limit than for any other approximation. Clearly, this is the effect of an error compensation. But, as will be shown in Secs. IV C and IV D, this is very systematic, and therefore energy differences computed with CCSD-F12a and AVDZ or AVTZ basis sets turn out to be amazingly accurate. On the other hand, CCSD-F12b underestimates the correlation energies. For comparison with the other methods, we also added CCSD-F12 x values using fully optimized

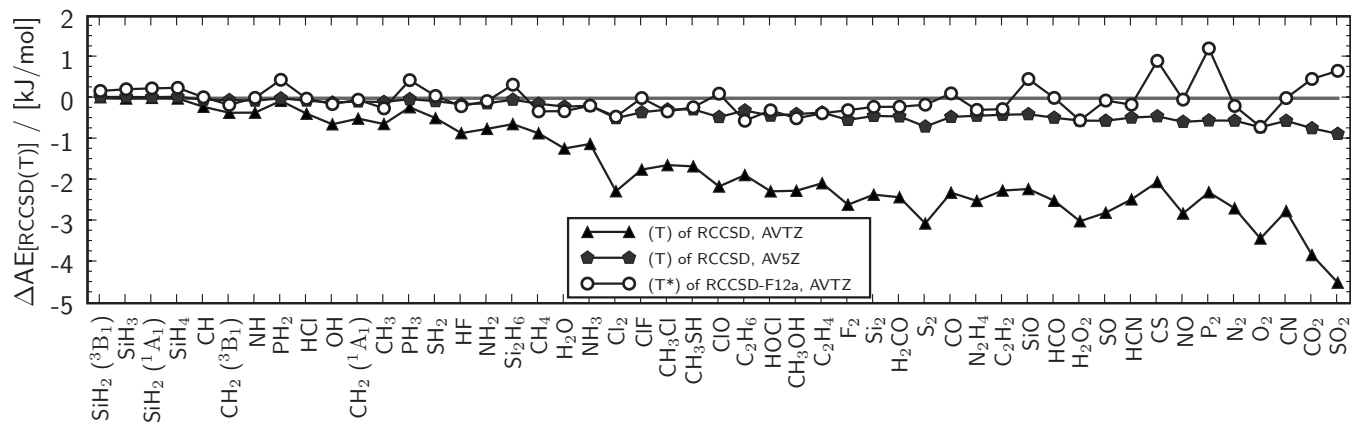


FIG. 2. Deviations of isolated (T) and (T*) energy contributions from conventional CCSD(T)/CBS[56] results for AEs, in kJ/mol.

MP2-F12 T_{mn}^{ij} -amplitudes instead of fixed diagonal ones. If these optimized amplitudes are used, CCSD-F12b values show good agreement with the ones from exact CCSD-F12, which also uses full T_{mn}^{ij} . Additionally, the tables show the effect of neglecting the $|ix\rangle\langle ix| + |xi\rangle\langle xi|$ contributions in Eq. (28). This effect is small and quickly decreases with increasing basis set size.

In conclusion, it appears that the differences between the exact CCSD-F12 correlation energies and those of all the more approximate methods are significantly smaller than the remaining basis set incompleteness errors. Therefore the approximate methods should be sufficient for practical applications, especially if the difference in computational cost is taken into account. The most important question regarding these methods is how accurately they predict energy differences and other molecular properties. In Secs. IV C–IV F we will present extensive benchmark data to answer this question.

B. Contributions of triple excitations

Figure 2 demonstrates that the basis set error of the isolated triples energy contribution to atomization energies (AEs) computed with the AVTZ basis sets can amount to several kJ/mol. The scaling correction described in Sec. II G significantly reduces the deviation from the basis set limit in

most cases. This is also found for other properties such as reaction energies (REs), electron affinities (EAs), and ionization potentials (IPs), as summarized statistically in Table II. The concrete systems that enter in these statistics are described in Secs. IV C and IV D.

C. Reaction energies

In order to evaluate the practical performance of our simplified CCSD(T)-F12x corrections, we calculated REs of a large set of trial reactions and compared the results with conventional CCSD(T)/CBS estimates. Many reactions are taken from a benchmark set developed previously.⁶⁸ Further reactions, which can be formed with additional molecules from the AE/IP/EA sets described below, were added in the current work.

All molecular geometries were optimized at the MP2/RMP2 level with AVTZ basis sets. The geometries and reference values can be found in the supplementary material.⁶⁷

In total, there are 54 reactions involving exclusively closed-shell molecules, and 50 further reactions involving also open-shell molecules and atoms. The reference RHF, CCSD, and CCSD(T) values for the closed- and open-shell reactions are listed in Tables III and IV, respectively. We

TABLE II. RMS deviations of isolated (T) contributions from conventional CCSD(T)/CBS[56] results for closed- and open-shell reaction energies (REc and REo, in kJ/mol), AEs (in kJ/mol), IPs (in meV), and EAs (in meV). Only reactions with CBS[56] results were included in the statistics. All listed (T) corrections apart from CBS values were calculated with CCSD-F12a, $\gamma=1.0 a_0^{-1}$.

Method/basis	REc	REo	AEs	IPs	EAs
Conv/CBS[23]	0.713	0.611	0.604	2.15	2.14
Conv/CBS[34]	0.144	0.136	0.130	0.64	1.52
Conv/CBS[45]	0.032	0.062	0.092	0.24	0.44
(T)/AVDZ	2.683	5.930	8.767	33.40	42.09
(T*)/AVDZ	1.375	3.779	3.671	24.34	26.19
(T)/AVTZ	1.292	1.670	2.470	10.51	13.07
(T*)/AVTZ	0.564	0.822	0.368	4.97	4.63
(T)/AVQZ	0.604	0.731	1.066	4.58	5.78
(T*)/AVQZ	0.227	0.380	0.150	1.67	1.87

TABLE III. Closed-shell reactions and reference values (kJ/mol).

Reaction	Total	RHF	Δ CCSD	$\Delta(T)$
1 $\text{CH}_3\text{COOH} + \text{NH}_3 \rightarrow \text{CH}_3\text{CONH}_2 + \text{H}_2\text{O}$	-0.13 ^a	2.73	-3.05	0.19
2 $\text{HCOOH} + \text{NH}_3 \rightarrow \text{HCONH}_2 + \text{H}_2\text{O}$	-4.47 ^a	-2.90	-1.90	0.32
3 $\text{CH}_3\text{ONO} \rightarrow \text{CH}_3\text{NO}_2$	-9.08 ^a	3.35	-9.21	-3.22
4 $\text{H}_2\text{SO}_3 \rightarrow \text{SO}_2 + \text{H}_2\text{O}$	-9.12 ^a	11.38	-12.56	-7.95
5 $\text{HCONH}_2 + \text{CH}_3\text{OH} \rightarrow \text{HCOOCH}_3 + \text{NH}_3$	-14.32 ^a	-6.79	-5.22	-2.31
6 $\text{H}_2\text{C}=\text{C}=\text{O} + \text{HCHO} \rightarrow \text{C}_2\text{H}_4\text{O} + \text{CO}$	-16.97 ^a	-24.51	2.98	4.56
7 $\text{HCOOH} + \text{CH}_3\text{OH} \rightarrow \text{HCOOCH}_3 + \text{H}_2\text{O}$	-18.79 ^a	-9.69	-7.11	-1.99
8 $\text{CO} + \text{H}_2 \rightarrow \text{HCHO}$	-21.64	1.06	-23.43	0.73
9 $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{H}_2$	-26.88	0.24	-16.96	-10.16
10 $\text{CH}_3\text{OH} + \text{HCl} \rightarrow \text{CH}_3\text{Cl} + \text{H}_2\text{O}$	-33.61	-25.11	-6.66	-1.83
11 $\text{H}_2\text{O} + \text{CO} \rightarrow \text{HCOOH}$	-37.74	-7.59	-26.09	-4.05
12 $\text{CO} + \text{NH}_3 \rightarrow \text{HCONH}_2$	-41.92 ^a	-10.49	-27.71	-3.73
13 $\text{CH}_3\text{OH} + \text{H}_2\text{S} \rightarrow \text{CH}_3\text{SH} + \text{H}_2\text{O}$	-45.59	-32.17	-11.29	-2.13
14 $\text{CS}_2 + 2\text{H}_2\text{O} \rightarrow \text{CO}_2 + 2\text{H}_2\text{S}$	-47.36	-122.28	57.28	17.65
15 $\text{Furane} + \text{NH}_3 \rightarrow \text{Pyrrole} + \text{H}_2\text{O}$	-54.43 ^a	-46.09	-8.29	-0.05
16 $\text{CO} + \text{CH}_3\text{OH} \rightarrow \text{HCOOCH}_3$	-56.24 ^a	-17.28	-32.92	-6.03
17 $\text{C}_2\text{H}_6 + \text{H}_2 \rightarrow 2\text{CH}_4$	-75.69	-88.77	10.64	2.44
18 $\text{Furane} + \text{H}_2\text{S} \rightarrow \text{Thiophene} + \text{H}_2\text{O}$	-80.87 ^a	-49.65	-24.44	-6.77
19 $\text{CH}_3\text{CN} + \text{H}_2\text{O} \rightarrow \text{CH}_3\text{CONH}_2$	-81.10 ^a	-80.31	-0.20	-0.59
20 $\text{HNCO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{NH}_3$	-85.93	-96.84	9.60	1.30
21 $\text{HNCO} + \text{NH}_3 \rightarrow \text{NH}_2\text{CONH}_2$	-86.28 ^a	-88.07	-3.88	5.67
22 $\text{HCN} + \text{H}_2\text{O} \rightarrow \text{HCONH}_2$	-92.02 ^a	-98.95	6.85	0.08
23 $\text{CH}_4 + \text{Cl}_2 \rightarrow \text{CH}_3\text{Cl} + \text{HCl}$	-98.17	-110.01	8.81	3.03
24 $\text{COCl}_2 + 2\text{NH}_3 \rightarrow \text{NH}_2\text{CONH}_2 + 2\text{HCl}$	-102.72 ^a	-147.80	33.97	11.11
25 $\text{Cl}_2 + \text{F}_2 \rightarrow 2\text{ClF}$	-113.25	-142.44	22.96	6.22
26 $\text{CO} + \text{Cl}_2 \rightarrow \text{COCl}_2$	-114.79	-56.53	-48.17	-10.09
27 $\text{CO}_2 + 3\text{H}_2 \rightarrow \text{CH}_3\text{OH} + \text{H}_2\text{O}$	-117.85	-118.69	-14.66	15.49
28 $\text{C}_2\text{H}_4 + \text{Cl}_2 \rightarrow \text{C}_2\text{H}_3\text{Cl} + \text{HCl}$	-117.85 ^a	-117.09	-1.24	0.48
29 $\text{HCHO} + \text{H}_2 \rightarrow \text{CH}_3\text{OH}$	-123.09	-119.50	-8.19	4.61
30 $\text{C}_2\text{H}_2 + \text{HCl} \rightarrow \text{C}_2\text{H}_3\text{Cl}$	-132.91 ^a	-120.03	-11.00	-1.89
31 $\text{HCN} + \text{NH}_3 + \text{CO} + \text{H}_2\text{O} \rightarrow 2\text{HCONH}_2$	-133.94 ^a	-109.43	-20.85	-3.65
32 $\text{CO} + 2\text{H}_2 \rightarrow \text{CH}_3\text{OH}$	-144.73	-118.45	-31.62	5.34
33 $\text{H}_2\text{C}=\text{C}=\text{O} + \text{H}_2\text{O} \rightarrow \text{CH}_3\text{COOH}$	-156.66 ^a	-152.44	-5.74	1.52
34 $\text{H}_2\text{C}=\text{C}=\text{O} + \text{NH}_3 \rightarrow \text{CH}_3\text{CONH}_2$	-156.79 ^a	-149.71	-8.79	1.71
35 $\text{C}_2\text{H}_2 + \text{H}_2\text{O} \rightarrow \text{CH}_3\text{CHO}$	-161.15 ^a	-169.99	7.75	1.09
36 $\text{C}_2\text{H}_4 + \text{H}_2 \rightarrow \text{C}_2\text{H}_6$	-165.48	-163.86	-5.92	4.30
37 $\text{SO}_3 + \text{CO} \rightarrow \text{SO}_2 + \text{CO}_2$	-180.95	-159.44	-15.12	-6.39
38 $\text{C}_2\text{H}_2 + \text{HCN} \rightarrow \text{C}_2\text{H}_3\text{CN}$	-185.16 ^a	-179.87	-4.07	-1.22
39 $\text{H}_2 + \text{Cl}_2 \rightarrow 2 \text{HCl}$	-190.59	-213.35	15.98	6.78
40 $\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$	-206.34	-216.39	5.68	4.37
41 $\text{SO}_2 + \text{H}_2\text{O}_2 \rightarrow \text{SO}_3 + \text{H}_2\text{O}$	-210.97	-231.70	16.38	4.34
42 $\text{C}_2\text{H}_4 + \text{H}_2\text{O}_2 \rightarrow \text{C}_2\text{H}_4\text{O} + \text{H}_2\text{O}$	-216.13 ^a	-223.12	2.77	4.22
43 $\text{CO} + 3 \text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-270.76	-246.90	-31.12	7.26
44 $\text{HCN} + 3 \text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_3$	-321.02	-335.39	3.30	11.06
45 $\text{H}_2\text{O}_2 + \text{H}_2 \rightarrow 2 \text{H}_2\text{O}$	-365.04	-391.37	18.22	8.11
46 $\text{CH}_3\text{CHO} + \text{H}_2\text{O}_2 \rightarrow \text{CH}_3\text{COOH} + \text{H}_2\text{O}$	-379.35 ^a	-398.48	15.40	3.74
47 $\text{CO} + \text{H}_2\text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O}$	-391.92	-391.13	1.26	-2.05
48 $2\text{NH}_3 + 3\text{Cl}_2 \rightarrow \text{N}_2 + 6\text{HCl}$	-407.43	-482.76	63.68	11.65
49 $3\text{N}_2\text{H}_4 \rightarrow 4\text{NH}_3 + \text{N}_2$	-439.08	-470.67	32.16	-0.57
50 $\text{H}_2 + \text{F}_2 \rightarrow 2\text{HF}$	-564.97	-610.74	32.89	12.88
51 $\text{CH}_3\text{NH}_2 + 3\text{H}_2\text{O}_2 \rightarrow \text{CH}_3\text{NO}_2 + 4\text{H}_2\text{O}$	-601.08 ^a	-603.36	9.61	-7.33
52 $\text{NH}_3 + 4\text{H}_2\text{O}_2 \rightarrow \text{HNO}_3 + 5\text{H}_2\text{O}$	-746.89 ^a	-759.28	17.75	-5.35
53 $\text{CH}_4 + 4\text{H}_2\text{O}_2 \rightarrow \text{CO}_2 + 6\text{H}_2\text{O}$	-1216.29	-1318.35	87.04	15.02
54 $2\text{NH}_3 + 3\text{F}_2 \rightarrow \text{N}_2 + 6\text{HF}$	-1530.57	-1674.93	114.42	29.93

^aRow was calculated from CCSD(T)/CBS[45] values instead of CBS[56].

TABLE IV. Open-shell reactions and reference values (kJ/mol).

Reaction	Total	RHF	Δ CCSD	$\Delta(T)$
1 HCl+H \rightarrow Cl+H ₂	-8.61	-12.52	-2.01	5.92
2 H ₂ O+F ₂ \rightarrow 2HF+O	-49.30	-292.86	216.01	27.55
3 CH ₄ +OH \rightarrow CH ₃ +H ₂ O	-55.30	-9.73	-41.98	-3.60
4 OH+H ₂ \rightarrow H ₂ O+H	-67.86	-27.53	-32.67	-7.66
5 CO+OH \rightarrow CO ₂ +H	-94.74	-27.29	-49.63	-17.82
6 CH ₃ +Cl ₂ \rightarrow CH ₃ Cl+Cl	-119.33	-140.33	16.12	4.88
7 S+2HCl \rightarrow H ₂ S+Cl ₂	-119.81	5.02	-109.11	-15.72
8 2NO+O ₂ \rightarrow 2NO ₂	-124.15	55.87	-138.00	-42.02
9 N+O ₂ \rightarrow NO+O	-141.31	-146.48	1.12	4.05
10 4HCl+O ₂ \rightarrow 2H ₂ O+2Cl ₂	-156.20	-129.15	-28.67	1.62
11 NO+H ₂ O ₂ \rightarrow NO ₂ +H ₂ O	-158.43	-85.52	-52.42	-20.49
12 2NO \rightarrow N ₂ +O ₂	-173.94	-118.61	-50.92	-4.41
13 2H ₂ O ₂ \rightarrow 2H ₂ O+O ₂	-192.71	-226.90	33.16	1.03
14 Cl ₂ +H \rightarrow HCl+Cl	-199.20	-225.87	13.97	12.70
15 2SO ₂ +O ₂ \rightarrow 2SO ₃	-229.24	-236.50	-0.40	7.66
16 Cl+OH \rightarrow HOCl	-247.13	-72.33	-153.94	-20.86
17 H ₂ S+F ₂ \rightarrow S+2HF	-254.57	-402.41	126.02	21.82
18 4NO ₂ +O ₂ +2H ₂ O \rightarrow 4HNO ₃	-262.02 ^a	-218.35	-61.08	17.41
19 2NH ₂ \rightarrow N ₂ H ₄	-305.87	-162.50	-129.32	-14.06
20 NO+N \rightarrow O+N ₂	-315.25	-265.09	-49.80	-0.36
21 O+2HCl \rightarrow H ₂ O+Cl ₂	-325.09	-104.53	-199.10	-21.45
22 2SiH ₃ \rightarrow Si ₂ H ₆	-333.37	-246.87	-79.75	-6.75
23 CH ₃ +SH \rightarrow CH ₃ SH	-334.38	-223.96	-97.72	-12.70
24 SO ₂ +O \rightarrow SO ₃	-361.61	-158.21	-184.96	-18.43
25 CS+O \rightarrow CO+S	-366.54	-320.32	-52.20	5.98
26 CH ₃ +Cl \rightarrow CH ₃ Cl	-369.61	-252.85	-103.03	-13.73
27 CH ₃ OH+O \rightarrow HCHO+H ₂ O	-392.59	-198.38	-174.94	-19.28
28 CH ₃ +OH \rightarrow CH ₃ OH	-412.47	-267.79	-131.04	-13.65
29 NH+H \rightarrow NH ₂	-415.53	-285.49	-124.54	-5.50
30 Si+2H ₂ \rightarrow SiH ₄	-443.20	-385.34	-54.73	-3.13
31 2C ₂ H ₄ +O ₂ \rightarrow 2CH ₃ CHO	-447.96 ^a	-463.08	6.54	8.58
32 CS+S \rightarrow CS ₂	-451.20	-296.57	-129.58	-25.04
33 NH ₂ +H \rightarrow NH ₃	-482.55	-361.36	-115.51	-5.67
34 2H ₂ +O ₂ \rightarrow 2H ₂ O	-537.38	-555.84	3.28	15.19
35 CO ₂ +C \rightarrow 2CO	-540.25	-438.83	-92.35	-9.08
36 CO+O \rightarrow CO ₂	-542.55	-317.64	-200.09	-24.82
37 C+H ₂ O \rightarrow CO+H ₂	-567.13	-438.59	-109.31	-19.24
38 N ₂ H ₄ +O ₂ \rightarrow N ₂ +2H ₂ O	-574.18	-607.88	24.50	9.20
39 2NH \rightarrow N ₂ +H ₂	-715.63	-434.59	-250.00	-31.04
40 C+S ₂ \rightarrow CS ₂	-734.97	-527.19	-176.11	-31.66
41 2CO+2NO \rightarrow N ₂ +2CO ₂	-765.07	-673.97	-81.56	-9.54
42 CH ₄ +2O ₂ \rightarrow CO ₂ +2H ₂ O	-830.87	-864.55	20.72	12.96
43 4NH ₃ +5O ₂ \rightarrow 4NO+6H ₂ O	-935.57	-1115.75	143.20	36.98
44 2NH ₃ +2NO+O \rightarrow 2N ₂ +3H ₂ O	-1062.65	-835.05	-215.01	-12.59
45 C+O ₂ \rightarrow CO ₂	-1131.39	-994.19	-122.99	-14.21
46 CS ₂ +3O ₂ \rightarrow CO ₂ +2SO ₂	-1144.66	-1097.09	-51.07	3.50
47 CH ₄ +4NO \rightarrow 2N ₂ +CO ₂ +2H ₂ O	-1178.75	-1101.76	-81.13	4.13
48 CH ₄ +NH ₃ +3O \rightarrow HCN+3H ₂ O	-1226.00	-618.26	-552.68	-55.06
49 2C+H ₂ \rightarrow C ₂ H ₂	-1228.27	-901.96	-291.26	-35.06
50 4NH ₃ +3O ₂ \rightarrow 2N ₂ +6H ₂ O	-1283.44	-1352.96	41.35	28.16

^aRow was calculated from CCSD(T)/CBS[45] values instead of CBS[56].

note that on the average the open-shell reactions studied here have larger REs than the closed-shell reactions. In fact, there are only 14 open-shell reactions with less than 200 kJ/mol RE, while 39 closed-shell reactions fall into this category. Therefore, larger mean absolute errors can be anticipated in the open-shell case.

In Fig. 3 the deviations of CCSD(T)-F12a/AVTZ results from CBS estimates are compared with conventional

CCSD(T)/AVTZ and CCSD(T)/AV5Z results. For most of the closed-shell reactions (upper panel), the basis set dependence is rather weak, the root mean square (RMS) errors being only 6.9 kJ/mol already for conventional CCSD(T)/AVTZ calculations. For the open-shell reactions (lower panel), the basis set dependence is much stronger. Conventional CCSD(T)/AVTZ here has a RMS deviation of 19.1 kJ/mol. In both cases, the CCSD(T)-F12a results are uni-

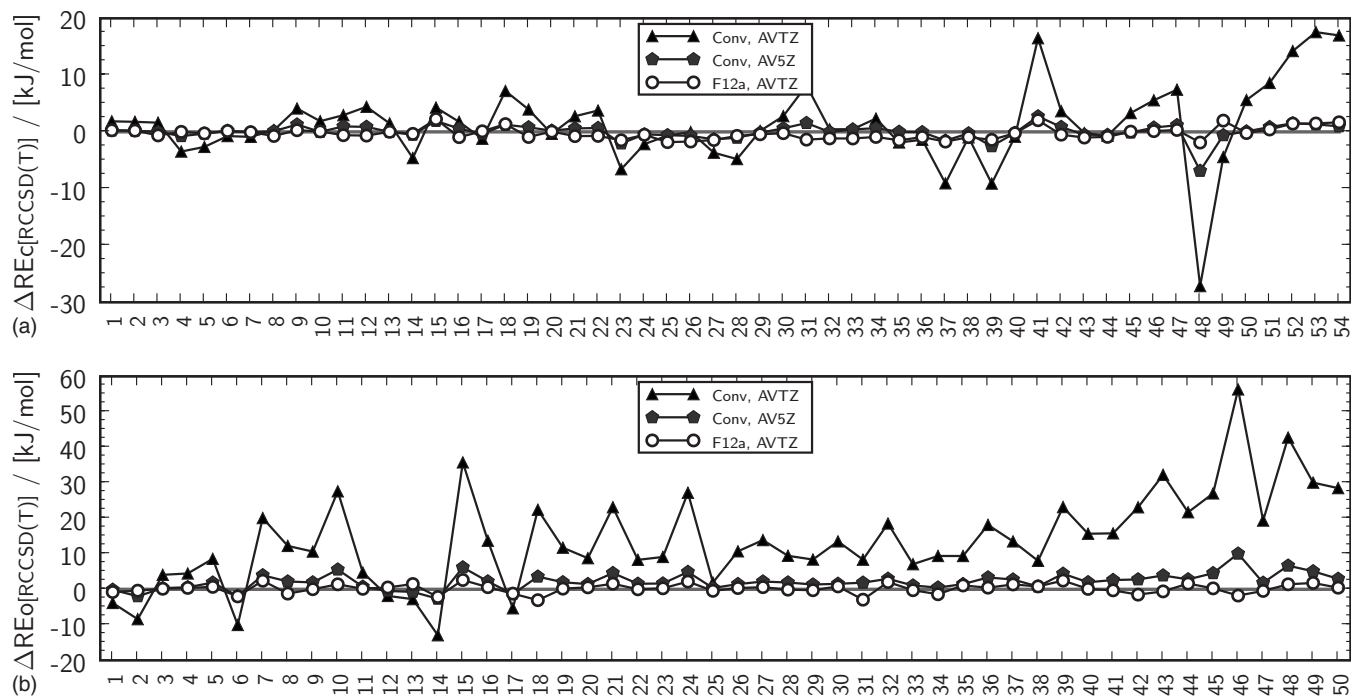


FIG. 3. Deviations of CCSD(T)-F12a and conventional CCSD(T)/AVTZ and CCSD(T)/AV5Z REs from CCSD(T)/CBS results. Upper panel: Closed-shell reactions; lower panel: Open-shell reactions.

formly accurate, and there are no outliers. The RMS errors are only 1.0 and 1.3 kJ/mol for the closed- and open-shell reactions, respectively.

A statistical analysis of the RHF, CCSD, and CCSD-F12a/b treatments for all the closed- and open-shell reactions

is presented in Table V. Mean absolute deviations (MADs), RMS deviations, and maximum deviations (MAX) from CBS estimates are shown. For F12/AVQZ values, only reactions for which AV6Z calculations were feasible enter in the statistics because the CBS[45] accuracy is not sufficient to

TABLE V. Statistical analysis of deviations of REs from conventional RHF/CBS or CCSD/CBS values (see text). Rows marked by an asterisk include only reactions with CBS[56] reference values.

Method/basis	Closed shell (kJ/mol)			Open shell (kJ/mol)		
	MAD	RMS	MAX	MAD	RMS	MAX
RHF/AVDZ	8.999	13.451	55.111	15.540	26.878	144.085
RHF/AVTZ	1.406	2.066	6.516	2.304	3.608	17.273
RHF/AVQZ	0.249	0.491	1.846	0.517	1.173	6.720
RHF/AV5Z*	0.085	0.168	0.572	0.137	0.354	2.021
RHF+singlets/AVDZ	1.409	1.786	6.026	1.651	2.218	6.160
RHF+singlets/AVTZ	0.378	0.477	1.608	0.398	0.536	2.021
RHF+singlets/AVQZ*	0.124	0.179	0.651	0.131	0.173	0.488
CCSD/AVDZ	10.761	18.743	99.302	39.774	51.966	197.439
CCSD/AVTZ	4.410	6.802	23.061	15.060	18.645	57.157
CCSD/AVQZ	1.561	2.462	11.696	5.322	6.708	23.679
CCSD/AV5Z	0.791	1.229	5.922	2.461	3.087	10.388
CCSD/CBS[23]	4.287	6.422	20.436	9.259	12.825	50.624
CCSD/CBS[34]	1.239	1.810	7.130	1.488	2.064	7.105
CCSD/CBS[45]*	0.341	0.468	1.394	0.440	0.605	1.683
CCSD-F12a/AVDZ	2.104	2.677	6.811	3.806	4.683	14.040
CCSD-F12a/AVTZ	0.922	1.275	4.625	0.941	1.286	3.551
CCSD-F12a/AVQZ*	0.395	0.507	1.145	0.974	1.241	4.387
CCSD-F12b/AVDZ	1.657	2.343	7.177	3.855	5.002	17.692
CCSD-F12b/AVTZ	0.820	1.180	5.313	1.228	1.809	7.347
CCSD-F12b/AVQZ*	0.457	0.585	1.304	0.554	0.675	1.735

TABLE VI. Statistical analysis of deviations of REs from conventional CCSD(T)/CBS results. Rows marked by an asterisk include only reactions with CBS[56] reference values.

Method/basis	Closed shell (kJ/mol)			Open shell (kJ/mol)		
	MAD	RMS	MAX	MAD	RMS	MAX
CCSD(T)/AVTZ	4.544	6.933	27.153	15.670	19.142	56.310
CCSD(T)/AVQZ	1.646	2.660	13.430	5.541	6.901	23.060
CCSD(T)/AV5Z	0.832	1.332	6.863	2.551	3.171	10.040
CCSD(T)/CBS[23]	4.357	6.501	19.766	9.131	12.769	51.248
CCSD(T)/CBS[34]	1.309	1.903	7.343	1.563	2.138	6.652
CCSD(T)/CBS[45]*	0.354	0.492	1.502	0.472	0.639	1.778
CCSD(T)-F12a/AVDZ	2.140	2.982	10.223	3.202	4.441	16.716
CCSD(T)-F12a/AVTZ	0.833	1.042	2.246	1.033	1.325	3.062
CCSD(T)-F12a/AVQZ*	0.565	0.700	1.813	0.824	1.112	2.805
CCSD(T*)-F12a/AVDZ	2.643	3.418	8.620	3.492	4.577	15.961
CCSD(T*)-F12a/AVTZ	0.900	1.250	4.740	1.108	1.469	4.817
CCSD(T*)-F12a/AVQZ*	0.392	0.530	1.304	1.134	1.451	4.318
CCSD(T)-F12b/AVDZ	1.983	2.974	13.849	3.877	5.209	17.820
CCSD(T)-F12b/AVTZ	0.953	1.259	4.727	1.757	2.220	5.593
CCSD(T)-F12b/AVQZ*	0.815	1.038	3.063	0.629	0.837	2.112
CCSD(T*)-F12b/AVDZ	2.318	3.161	8.655	3.699	4.619	11.583
CCSD(T*)-F12b/AVTZ	0.938	1.285	5.428	1.233	1.717	5.531
CCSD(T*)-F12b/AVQZ*	0.508	0.670	1.744	0.689	0.861	2.566

judge such high-level F12 treatments. “RHF+singles” refers to conventional RHF, to which the CABS relaxation correction described in Sec. II F has been added.

This CABS singles correction clearly reduces the Hartree–Fock errors to levels at which they are not problematic any more. As evidenced by the RHF/AVDZ row, without this correction no sensible total REs could be obtained in any isolated AVDZ treatment due to the RHF errors alone. Even the AVTZ results would be seriously affected by Hartree–Fock errors. However, the corrected RHF+singles/AVDZ and RHF+singles/AVTZ rows show basis set limit deviations that are small compared to those of the correlation contributions.

On the correlation side, both the F12a and F12b correlation treatments deliver an excellent performance. While for the closed-shell molecules both F12a and F12b perform equally well, for the open-shell reactions F12a is ahead when used with AVDZ and AVTZ basis sets. This difference in the performance of the F12a and F12b methods for difficult open-shell cases was also observed in other cases (see Sec. IV D).

The statistical analysis of F12x energy differences involving either direct (T) triples contributions or the scaled (T*) triples treatment (see Sec. II G) is seen in Table VI. While Table II demonstrated that the isolated triples contributions are recovered more accurately using (T*), Table VI shows that this improvement does not always lead to better total energy differences. This suggests that a part of the good CCSD(T)-F12a/AVTZ performance stems from a fortuitous cancellation of the overshooting F12a correlation energies and the undershooting noncorrected triples contributions.

F12b values, which do not overshoot in the correlation energy, do not show this behavior and usually get better when (T*) is applied instead of simply adding (T).

For both, closed- and open-shell reactions, both with F12a and F12b, an explicitly correlated CCSD-F12x or CCSD(T*)-F12x calculation with AVnZ basis sets yields a performance comparable to conventional AVnZ/AV(n+1)Z extrapolated CBS estimates. This comes at the reduced cost of *only* the smaller calculation. Compared to the special case of the CBS[23] extrapolation – which generally leads to rather poor results – F12/AVDZ calculations even offer greatly increased accuracy. In the case of the AVQZ basis set, it is not clear if the CBS[45] or the CCSD(T*)-F12b results are more accurate since the errors of the CBS[56] extrapolation may be of the same magnitude as the differences between the two.

These conclusions still hold when the more involved extrapolation formulas of Schwenke⁶⁹ are used. In fact, in our benchmark set there are only few systems where using either extrapolation scheme makes a noteworthy difference. A table using Schwenke extrapolation reference values can be found in the supplementary material.

The 15 reactions calculated by Tew *et al.* in Ref. 37 are a subset of the current closed-shell reaction set. (This is also the same set of reactions we investigated in Ref. 42. The values reported in that article slightly differ from the current ones because now other reference values are used, the CABS singles correction does not include core orbitals anymore, and the CABS for MP2-F12 is constructed with different thresholds.) A statistical comparison of the CCSD(T)(F12) results reported in that reference with our F12x methods is shown in Table VII. (F12)- c_{ij}^{kl} and (F12)- c refer to the

TABLE VII. Comparison of correlation contributions (kJ/mol) of CCSD(T)-F12a and -F12b with the CCSD(T)(F12) values of Ref. 37. Statistics refer to 15 reactions. See text for details.

Method	γ (a_0^{-1})	MAD	RMS	MAX
F12a	1.0	0.55	0.74	1.88
F12b	1.0	0.47	0.62	1.32
(F12)- c_{ij}^{kl}	1.0	0.75	0.97	2.00
(F12)- c	1.0	1.43	1.85	5.00
F12a	1.3	0.78	1.07	2.21
F12b	1.3	0.71	0.90	1.67
(F12)- c_{ij}^{kl}	1.3	0.72	0.97	2.16
(F12)- c	1.3	0.96	1.27	3.14
(F12)- c	1.8	0.97	1.21	2.80

CCSD(F12) method applied with fully optimized amplitudes for the explicitly correlated part and with the fixed amplitude approximation (which we also employ), respectively. The statistics are calculated from the basis set deviations and reference values for CCSD and (T) correlation contributions reported in Ref. 37, which have been added together. The reference values are CBS[56] for CCSD and CBS[45] for (T). An exception was made for the $C_2H_2+H_2O \rightarrow CH_3CHO$ reaction because the CCSD/CBS[56] reference value of 3.06 kJ/mol reported in Table II of Ref. 37 is apparently in error. For this reaction we used our CCSD/CBS[45] value of 7.75 kJ/mol, which is also much more consistent with both our and their F12 calculations. Our F12 x methods were applied with $V(n+1)Z$ auxiliary basis sets and have been compared to the same reference values.

As a result we can see that for this particular set of reactions our F12 x methods are slightly more accurate than the (F12) methods. Surprisingly we also see that the dependence on the geminal exponent of both approaches is apparently the reverse of each other – a finding that should be investigated in future work.

D. Atomization energies, ionization potentials, and electron affinities

In addition to the REs, we benchmarked AEs, IPs, and EAs as representatives of certain classes of worst-case scenarios. AEs are particularly difficult because of the huge BSSE and because the electronic structure of atoms and molecules is very different, thus making error cancellations unlikely. EAs and IPs are also tough because the number of electrons in the calculated species differs. For the additional/missing electron, there is nothing that could cancel an error in its description, and even highly systematic errors related to the interaction with the remaining electrons would inevitably show up in benchmark results. The same AEs/IPs/EAs are calculated as in the original G2 benchmark set,⁷⁰ with the exception of species involving alkali or alkali-earth elements. The latter had to be excluded because suitable JKFIT or augmented orbital basis sets for them are not available at the moment.

TABLE VIII. Systems in the benchmark set. EAs and IPs are adiabatic. RMP2/AVTZ geometries were used.

Electron affinities: NO, CH ₃ , NH, O ₂ , CH ₂ , P, NH ₂ , PH, SiH ₂ , PO, CH, PH ₂ , SiH, C, SiH ₃ , Si, O, S ₂ , OH, S, SH, Cl ₂ , F, Cl, CN
Ionization potentials: Al, Si, B, S ₂ , PH ₃ , PH ₂ → PH ₂ ^(1A₁) , PH, S, NH ₃ , SH, SH ₂ → SH ₂ ^(2B₁) , P, C ₂ H ₄ , P ₂ , SiH ₄ , C, CS, Cl ₂ , C ₂ H ₂ , O ₂ , ClF, SH ₂ → SH ₂ ^(2A₁) , HCl, H ₂ O, CH ₄ , Cl, OH, O, CO, N, N ₂ → N ₂ ^(2Σ_g) , HF, N ₂ → N ₂ ^(2Π_u) , F
Atomization energies: F ₂ , Cl ₂ , ClO, ClF, Si ₂ , NH, CH, HCl, OH, S ₂ , P ₂ , SiH ₂ ^(3B₁) , SO, O ₂ , HF, SiH ₂ ^(1A₁) , PH ₂ , NO, HOCl, CH ₂ ^(1A₁) , CS, SH ₂ , NH ₂ , CN, CH ₂ ^(3B₁) , SiO, SiH ₃ , PH ₃ , H ₂ O, N ₂ , CO, H ₂ O ₂ , SO ₂ , HCO, NH ₃ , CH ₃ , SiH ₄ , HCN, H ₂ CO, CH ₃ Cl, C ₂ H ₂ , CO ₂ , CH ₄ , N ₂ H ₄ , CH ₃ SH, CH ₃ OH, Si ₂ H ₆ , C ₂ H ₄ , C ₂ H ₆

Table VIII lists the benchmarked systems. The geometries (RMP2/AVTZ optimized) and reference values can be found in the supplementary material. The deviations relative to conventional CCSD(T)/CBS[56] will be shown.

Figure 4 compares the basis set deviations of AEs, IPs, and EAs obtained with conventional AVTZ and AV5Z calculations with those of CCSD(T)-F12a/AVTZ calculations. The systems are ordered by increasing absolute reference values. Again we find that the CCSD(T)-F12a method with AVTZ basis sets yields more accurate results than conventional CCSD(T) calculations with quintuple-zeta (AV5Z) basis sets. Additionally, even for these difficult tasks the accuracy is very uniform. The RMS basis set deviation of F12a/AVTZ for AEs is only 1.3 kJ/mol.

Table IX contains a statistical analysis of the RHF, CCSD, and CCSD-F12a/b treatments for all AEs, IPs, and EAs. Like in the REs, we see that the perturbative RHF corrections are important for AVDZ and AVTZ basis sets. This holds especially for AEs, in which the RHF correction reduces the maximum error from 11.4 kJ/mol with RHF/AVTZ (which is much larger than the error of the F12/AVTZ correlation treatment) to mere 1.1 kJ/mol when the correction is applied.

On the correlation side, both the F12a and F12b correlation treatments show a robust performance in absolute terms. For small- to medium-sized basis sets, particularly the important AVTZ set, F12a clearly outperforms F12b. These CCSD-F12a RMS deviations are close to the ones obtained with a full DF-RMP2-F12/3C(FIX) treatment⁵⁰ (compared with RMP2/CBS[56]), which for these kinds of tasks can be regarded as a measure of the theoretical optimum performance of F12 treatments. This is because the RMP2-F12 method does not involve any approximations other than DF and the RI, which are both very accurate and well controllable.

For large basis sets F12a shows a troubling behavior. This is related to the incomplete energy functional, which leads to an overshooting of correlation energies, as already mentioned in Sec. IV A. This is exemplified by the slight *increase* in the F12a RMS deviation for AEs when going from AVTZ to AVQZ basis sets (see Table IX). F12b, however, converges systematically. For AVQZ and larger basis sets, F12b is therefore the preferred choice.

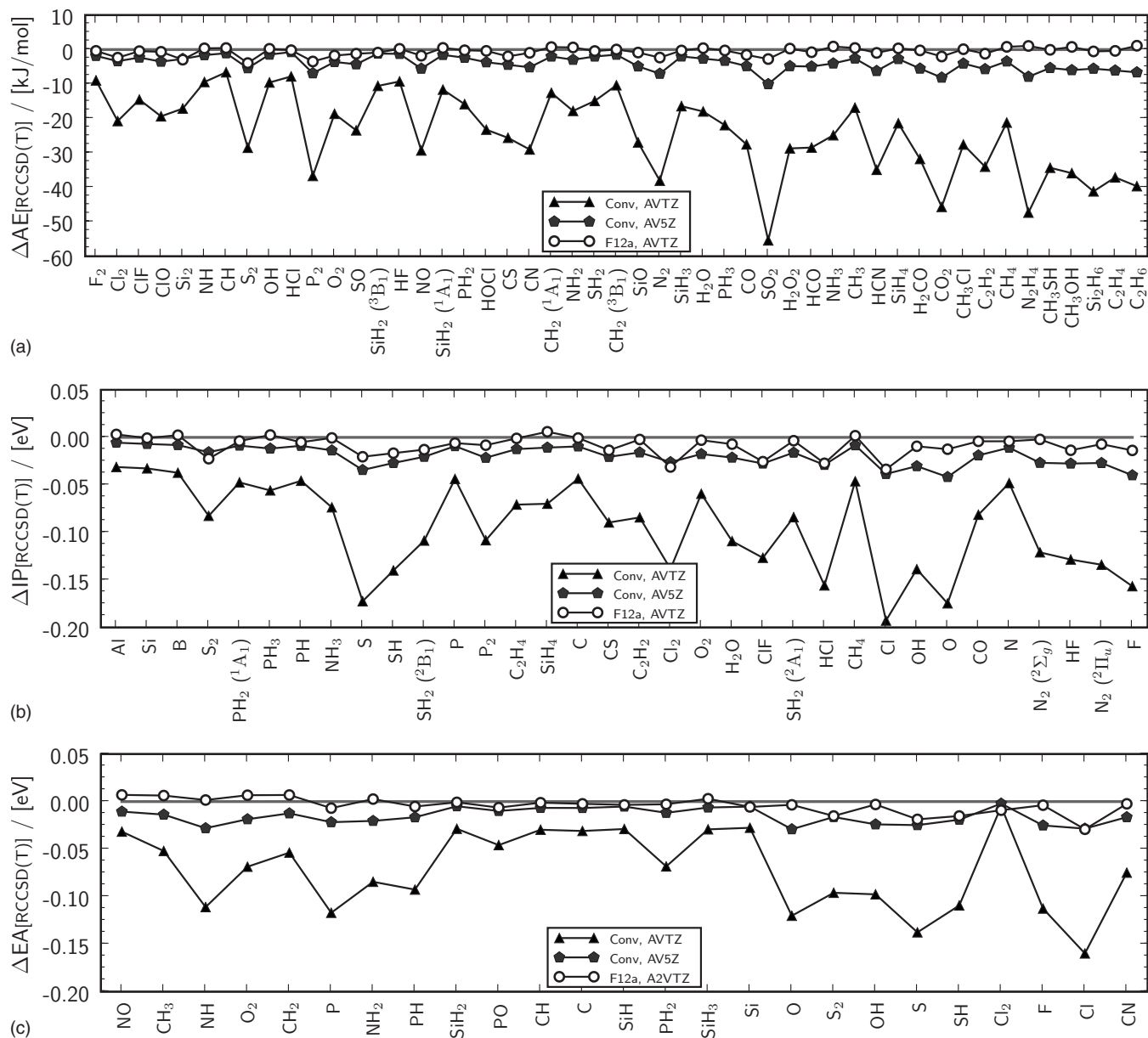


FIG. 4. Deviations of CCSD(T)-F12a and conventional CCSD(T)/AVTZ and CCSD(T)/AV5Z from CCSD(T)/CBS[56]. Upper panel: AEs; middle panel: IPs; lower panel: EAs.

The statistical analysis of F12x energy differences including the triples contributions is presented in Table X. Apart from the F12a/AVQZ combination – which should not be used, as pointed out above – the (T*) approach now generally improves also the total energy differences. In particular, the scaling correction leads to rather significant improvements of the F12b/AVQZ results.

E. Impact of additional calculation parameters

Additional statistics characterizing the dependence of the results on the geminal exponent γ , the size of the auxiliary basis sets, and the neglected $|ix\rangle\langle ix| + |xi\rangle\langle xi|$ projector contributions in Eq. (28) can be found in the supplementary material.⁶⁷ In short, the following results can be extracted:

For F12a and F12b, $\gamma=1.0$ is a good choice for both AVDZ and AVTZ basis sets, and the dependence on the actual γ -value is rather weak. F12b/AVQZ calculations may

profit significantly from larger exponents such as 1.2–1.6, especially for AEs, IPs, and EAs. Note that the optimum γ -values are supposed to get harder with more complete orbital basis sets. This is because the softer parts of the F12 configuration tails can then be represented well within the conventional orbital expansion, and the F12 contributions are only needed to correct for deficiencies near the wave function cusp.

The neglect of the $|ix\rangle\langle ix| + |xi\rangle\langle xi|$ projector in Eq. (28) is justified. Its inclusion leads to very small changes in the statistics, and for AVTZ and larger basis sets the effect is hardly noticeable at all.

Regarding the auxiliary basis sets, we found that the combinations described in Sec. III C are usually sufficient. But sometimes improvements in the statistics are observed when the cardinal numbers of the auxiliary basis sets are increased by 1. For example, the RMS deviation of CCSD-

TABLE IX. Statistical analysis of deviations of AEs, IPs, and EAs from conventional RHF/AV6Z or CCSD/CBS[56].

Method/basis	AEs (kJ/mol)			IPs (meV)			EAs (meV)		
	MAD	RMS	MAX	MAD	RMS	MAX	MAD	RMS	MAX
RHF/AVDZ	20.659	24.390	84.600	45.87	65.14	220.36	42.22	76.22	325.53
RHF/AVTZ	2.904	3.467	11.374	8.22	11.07	28.52	8.10	11.80	35.09
RHF/AVQZ	0.580	0.823	3.912	1.63	2.27	7.66	2.45	3.58	12.78
RHF/AV5Z	0.161	0.251	1.213	0.40	0.59	1.58	0.68	0.98	3.29
RHF+Singlets/AVDZ	1.304	1.639	4.343	6.55	8.16	20.80	5.63	8.04	29.16
RHF+Singlets/AVTZ	0.212	0.287	1.100	2.12	2.76	7.71	1.37	1.80	5.00
RHF+Singlets/AVQZ	0.093	0.127	0.403	0.83	1.04	2.21	0.72	1.01	2.57
CCSD/AVDZ	72.493	80.491	153.635	226.55	245.72	471.99	160.50	177.76	355.81
CCSD/AVTZ	22.540	24.944	50.879	87.33	98.10	180.02	61.34	72.18	146.25
CCSD/AVQZ	7.876	8.841	20.373	35.64	40.10	77.11	24.75	28.00	49.38
CCSD/AV5Z	3.692	4.168	9.164	18.09	20.54	40.00	12.76	14.64	24.82
CCSD/CBS[23]	8.578	11.075	37.614	24.44	31.37	80.18	22.27	30.35	95.61
CCSD/CBS[34]	1.429	1.694	4.206	6.46	8.99	24.32	8.91	12.54	37.58
CCSD/CBS[45]	0.282	0.390	1.080	1.24	1.53	3.19	1.39	1.73	3.76
CCSD-F12a/AVDZ	5.972	7.031	18.202	39.02	51.96	143.45	26.79	34.93	98.90
CCSD-F12a/AVTZ	1.534	1.859	4.102	7.70	9.23	23.97	10.28	11.76	18.14
CCSD-F12a/AVQZ	1.910	2.167	4.654	8.73	9.10	14.35	10.37	11.38	18.08
CCSD-F12b/AVDZ	8.390	10.230	27.434	60.47	70.84	176.81	39.24	50.22	132.47
CCSD-F12b/AVTZ	1.717	2.144	6.138	17.12	21.80	44.69	11.92	14.83	30.35
CCSD-F12b/AVQZ	0.554	0.701	1.814	4.24	5.20	12.07	3.15	3.76	8.13

F12a/AVTZ, $\gamma=1.0$ of closed-shell reactions [reduced set with the 28 reactions for which CBS[56] results are available] drops from 1.51 to 1.20 kJ/mol. However, these improvements are not usually worth the increase in calculation time – especially if small molecules are calculated for which the additional cost of the initial MP2-F12 is not negligible.

F. Spectroscopic constants of diatomic molecules

In order to investigate how the CCSD(T)-F12x corrections influence potential energy surfaces, we benchmarked the basis set convergence of the spectroscopic constants of several diatomic molecules. These are the closed-shell mol-

TABLE X. Statistical analysis of deviations of AEs, IPs, and EAs from conventional CCSD(T)/CBS[56].

Method/basis	AEs (kJ/mol)			IPs (meV)			EAs (meV)		
	MAD	RMS	MAX	MAD	RMS	MAX	MAD	RMS	MAX
CCSD(T)/AVTZ	24.213	26.764	55.378	95.08	105.50	192.79	72.51	83.12	159.80
CCSD(T)/AVQZ	8.564	9.609	22.279	38.82	43.20	80.52	29.53	33.17	56.52
CCSD(T)/AV5Z	4.005	4.521	10.032	19.63	22.05	41.69	15.34	17.39	29.10
CCSD(T)/CBS[23]	8.317	10.845	37.227	23.86	30.85	82.44	22.62	30.40	94.17
CCSD(T)/CBS[34]	1.490	1.770	4.320	6.55	8.87	23.59	9.14	12.46	36.65
CCSD(T)/CBS[45]	0.329	0.447	1.175	1.38	1.69	3.71	1.65	2.09	4.42
CCSD(T)-F12a/AVDZ	7.771	9.454	18.923	67.34	76.19	174.92	50.14	58.25	129.96
CCSD(T)-F12a/AVTZ	0.967	1.326	3.920	9.64	13.32	33.62	6.75	9.29	29.15
CCSD(T)-F12a/AVQZ	1.126	1.363	3.491	4.69	5.02	8.94	5.70	6.89	12.73
CCSD(T*)-F12a/AVDZ	5.806	6.870	14.315	59.62	70.11	161.81	35.89	47.61	110.18
CCSD(T*)-F12a/AVTZ	1.496	1.769	3.817	7.79	10.96	32.80	8.36	9.71	17.24
CCSD(T*)-F12a/AVQZ	1.969	2.243	4.738	7.49	7.95	13.72	8.92	10.33	18.19
CCSD(T)-F12b/AVDZ	12.757	14.877	31.808	89.22	97.41	208.28	77.78	84.40	163.53
CCSD(T)-F12b/AVTZ	3.572	3.967	8.936	26.29	30.27	60.00	22.61	25.72	46.81
CCSD(T)-F12b/AVQZ	0.702	0.972	3.562	7.34	8.76	17.81	6.85	8.20	17.87
CCSD(T*)-F12b/AVDZ	9.380	11.408	25.945	81.50	90.40	195.17	63.03	70.32	143.75
CCSD(T*)-F12b/AVTZ	1.717	2.104	5.807	21.01	25.47	53.18	14.60	17.83	35.41
CCSD(T*)-F12b/AVQZ	0.579	0.731	2.148	4.84	6.24	12.82	3.81	4.76	12.41

TABLE XI. RMS deviations of equilibrium distances r_e and harmonic frequencies ω_e from CCSD(T)/CBS[56].

Basis	r_e (pm)			ω_e (cm ⁻¹)		
	Conv	F12a	F12b	Conv	F12a	F12b
AVDZ	2.49	0.20	0.15	75.7	5.7	5.0
AVTZ	0.59	0.07	0.05	21.0	2.6	1.5
AVQZ	0.19	0.013	0.008	6.5	0.7	1.2
AV5Z	0.08	0.008	0.006	2.9	0.8	0.6

ecules HF, N₂, CO, BF, F₂, and C₂ (¹Σ⁺) and the open-shell molecules OH (²Π), NH (³Σ⁻), CH (²Π), CN (²Σ⁺), NO (²Π), O₂ (³Σ_g⁻), and CF (²Π).

For each molecule, a grid of ten energy points within ±0.15 a_0 of the equilibrium distance was calculated, and the equilibrium harmonic frequency ω_e and equilibrium distance r_e were determined from an eighth-order polynomial fit. F12 methods have been applied with $\gamma=1.0, 1.2, 1.4, 1.6 a_0^{-1}$ for AVDZ, AVTZ, AVQZ, and AV5Z, respectively. The number of CABS functions was frozen to the value obtained by the procedure described in Sec. III B at the minimum distance. This had only an effect on AV5Z calculations; in other calculations no functions needed to be deleted.

The F12 corrections greatly accelerate the convergence of equilibrium bond lengths. Table XI lists the RMS deviations from CCSD(T)/CBS[56]. In fact, the convergence is so fast that even with F12/AVQZ it is not clear whether the CBS[56] used as a reference is more accurate. For comparison: For this set of molecules the RMS deviation of the reference r_e with respect to experimental values⁷¹ is 0.17 pm. With F12 methods, this intrinsic accuracy can already be reached with AVDZ or AVTZ basis sets.

Similar conclusions hold for the convergence of harmonic frequencies ω_e . Again, both CCSD(T)-F12a and -F12b show a rapid convergence, and using F12 methods, even with valence-double- ζ basis sets accurate frequencies can be calculated. This rapid convergence even holds for difficult cases, such as the triply bonded N₂, C₂, and CN, which contain low lying excited states, and F₂, which is unbound at the Hartree-Fock level.

The basis set limit CCSD(T) frequencies themselves have a RMS deviation of 10.3 cm⁻¹ from experimental values.⁷¹ This is consistent with the previous findings of Ruden *et al.*⁷² and Tew *et al.*,⁷³ who also discussed which further effects need to be considered to obtain more accurate values.

All calculations included raw (T) triples. Surprisingly, applying either (T^{*})-scaling of the triples or adding extrapolated AVnZ/AV(n-1)Z-triples instead of the raw triples substantially increased the observed deviation from CBS[56].

V. SUMMARY OF THE RESULTS

For routine ground state calculations, the results of CCSD(T)-F12a calculations employing the AVTZ basis set reliably reach conventional AV5Z quality. Remaining deviations from CBS[56] are usually smaller than for CBS[34],

particularly for situations that challenge conventional basis set expansions. In contrast to conventional methods, the dependence of F12 errors on the molecular size and the size of the quantity being measured is low.

Results involving AVDZ basis sets usually reach conventional AVQZ quality. However, this depends sensitively on the errors of the AVDZ-(T) triples and the perturbative RHF corrections. While in the present benchmarks no obvious outliers were found, we cannot exclude that some exist in other applications or with other molecules. More balanced F12 basis sets may be helpful in this regard.⁷⁴ It should be noted that the diffuse basis functions in the augmented correlation consistent basis sets are important both for obtaining accurate HF values as well as for the F12 correlation treatment. Non-augmented basis sets should therefore not be used in F12 calculations (an exception are non-augmented basis sets for H atoms, which may often be sufficient).

F12a produces relative energies of exceptional quality for orbital basis sets up to AVTZ. But starting with AVQZ, the convergence speed becomes slow, while F12b converges smoothly to the basis set limit. Therefore F12b is preferable for calculations with AVQZ or larger basis sets, aiming at sub-kJ/mol accuracy.

VI. CONCLUSIONS

The results presented in this paper clearly demonstrate that our explicitly correlated coupled cluster approximations are ready for use in practice. In fact, they are already available for public use in the latest release of the MOLPRO package of *ab initio* programs.⁵⁹ The presented benchmark calculations convincingly show that the CCSD(T)-F12a method with triple-zeta basis sets yields results very close to the basis set limit. Generally, the F12 treatment reduces the basis set errors of all considered properties (AEs, REs, EAs, IPs, equilibrium structures, and vibrational frequencies) typically by one order of magnitude. Results of similar quality were recently also obtained for intermolecular interaction energies,⁵¹ as well as for equilibrium geometries and anharmonic vibrational frequencies of larger molecules.⁷⁵ In many cases the remaining basis set errors are comparable to or smaller than the intrinsic errors of the CCSD(T) method. Compared to standard CCSD(T) calculations with aug-cc-pV5Z or aug-cc-pV6Z basis sets, which would be required to achieve comparable accuracy, the computational effort is reduced by two orders of magnitude.

The great gains in accuracy obtained for such a wide variety of applications, combined with the minimal increase in computation time compared with conventional CCSD(T) calculations with the same basis set, suggest that the simplified F12 corrections should be used by default in quantum chemical applications of coupled cluster methods. Since DF and RI basis sets suitable for calculations with correlation consistent orbital basis sets are available and chosen automatically in the MOLPRO program, the user has nothing more to do than to change the command CCSD(T) to CCSD(T)-F12 in order to run an explicitly correlated F12 calculation. By default the program then computes both the F12a and the F12b energies.

The F12 treatment reduces the basis set error in calculations with medium size (augmented triple-zeta) basis sets to the same size as other small errors in high-accuracy quantum chemical treatments. The good balance of the remaining errors due to different sources (perturbative treatment of triple excitations and neglect of higher excitations, core-valence correlation, relativistic effects, nonadiabatic effects, etc.) makes it unlikely that a more complete treatment of the CCSD-F12-Lagrangian would significantly improve the overall accuracy of total energies and energy differences, unless also these remaining effects are explicitly accounted for in separate calculations.

One shortcoming of the present method is the missing direct F12 treatment of triple excitations. Furthermore, the performance of explicitly correlated methods for core-core and core-valence correlations, particularly when transition metals and heavier elements are present, still needs to be evaluated systematically; first tests are very promising. The development of analytical CCSD(T)-F12 gradient techniques would be of great help for accurate geometry optimizations and the prediction of vibrational frequencies. Last but not least, local approximations, which have already been shown to work very well for LMP2-F12,^{29,30} should be extended to LCCSD(T)-F12, in order to avoid the steep $\mathcal{O}(N^7)$ scaling of the CCSD(T) method and make it possible to treat much larger systems. Our future efforts will be concentrated on these aspects.

ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft in the priority program SPP 1145 (“Modern and universal first-principles methods for many-electron systems in chemistry and physics”). H.-J.W. is also grateful for generous support by the Fonds der Chemischen Industrie. We thank Fred Manby for providing the integral routines used in the density fitted F12 programs.

APPENDIX: DERIVATION FOR THE CASE OF CEPA(0)-DOUBLES

For illustrative purposes we present some arguments of Sec. II in more detail for CEPA(0)-doubles. This CEPA(0)D can be obtained formally from the CCSD method by dropping all terms which are non-linear in the cluster operator and furthermore neglecting all single excitation terms. The amplitude equations then become

$$\mathbf{R}^{ij} = \mathbf{R}_{\text{MP2}}^{ij} + \mathbf{K}(\mathbf{T}^{ij}) + \mathbf{K}_{kl}^{ij} \mathbf{T}^{kl} + \mathbf{G}^{ij} + \mathbf{G}^{ji\dagger}, \quad (\text{A1})$$

$$[\mathbf{K}(\mathbf{T}^{ij})]_{\alpha\beta} = \langle \alpha\beta | r_{12}^{-1} | \gamma\delta \rangle T_{\gamma\delta}^{ij}, \quad (\text{A2})$$

$$\mathbf{R}_{\text{MP2}}^{ij} = \mathbf{K}^{ij} + \mathbf{f} \mathbf{T}^{ij} + \mathbf{T}^{ij} \mathbf{f} - \mathbf{T}^{ik} f_{kj} - f_{ik} \mathbf{T}^{kj}, \quad (\text{A3})$$

$$\mathbf{G}^{ij} = \tilde{\mathbf{T}}^{ik} \mathbf{K}^{kj} - \mathbf{T}^{ik} \mathbf{J}^{kj} - \mathbf{J}^{kj} \mathbf{T}^{ik}, \quad (\text{A4})$$

$$\tilde{\mathbf{T}}^{ik} = 2\mathbf{T}^{ik} - \mathbf{T}^{ik\dagger}. \quad (\text{A5})$$

Here matrices have the dimension of the virtual space, which is assumed to be complete (indices α, β, \dots), and the integral

matrices are given by $K_{\alpha\beta}^{ij} = \langle \alpha\beta | r_{12}^{-1} | ij \rangle$ and $J_{\alpha\beta}^{ij} = \langle \alpha i | r_{12}^{-1} | \beta j \rangle$ as usual.

We consider the amplitude equations for the explicitly correlated configurations. For this, we form

$$R_{mn}^{ij} = \mathcal{F}_{\alpha\beta}^{mn} R_{\alpha\beta}^{ij}, \quad (\text{A6})$$

with $R_{\alpha\beta}^{ij}$ from Eq. (A1) and

$$T_{\alpha\beta}^{ij} = \delta_{\alpha}^i \delta_{\beta}^j T_{ab}^{ij} + T_{\alpha\beta}^{ij} \quad (\text{A7})$$

from Eq. (3). We examine only non-MP2 terms since the MP2 terms have been discussed elsewhere.^{31,50} Due to the terms $\mathbf{K}(\mathbf{T}^{ij})$ and $\mathbf{K}_{kl}^{ij} \mathbf{T}^{kl}$, we obtain

$$R_{mn}^{ij} = \dots + H_{mn}^{kl} T_{kl}^{ij} + V_{cd}^{mn} T_{cd}^{ij} + K_{kl}^{ij} T_{op}^{kl} X_{mn}^{op} + \dots, \quad (\text{A8})$$

$$H_{mn}^{kl} = \mathcal{F}_{\alpha\beta}^{kl} \langle \alpha\beta | r_{12}^{-1} | \gamma\delta \rangle \mathcal{F}_{\gamma\delta}^{mn} = \langle kl | F_{12} \hat{Q}_{12} r_{12}^{-1} \hat{Q}_{12} F_{12} | mn \rangle, \quad (\text{A9})$$

$$V_{cd}^{ij} = \langle ij | F_{12} \hat{Q}_{12} r_{12}^{-1} | cd \rangle. \quad (\text{A10})$$

While complex to calculate accurately and efficiently, these terms are tractable, and the dominant contributions can be evaluated exactly. The evaluation of V_{cd}^{ij} is described in Eq. (28) and Sec. II C. H_{mn}^{kl} can be evaluated from the similar quantities V_{rs}^{kl} , V_{mx}^{kl} , and $\langle kl | F_{12} \hat{Q}_{12} (F_{12} r_{12}^{-1}) | mn \rangle$ by inserting Eq. (19) for the right-hand \hat{Q}_{12} projector in Eq. (A9).

The terms stemming from $\mathcal{F}_{\alpha\beta}^{mn} G_{\alpha\beta}^{ij}$ are more troublesome. For example, the $\mathbf{T}^{ik} \mathbf{J}^{kj}$ term of Eq. (A4) turns into

$$\begin{aligned} \mathcal{F}_{\alpha\beta}^{mn} T_{\alpha\gamma}^{ik} J_{\gamma\beta}^{kj} &= \langle mn | F_{12} \hat{Q}_{12} | \alpha x \rangle T_{ac}^{ik} J_{cx}^{kj} \\ &+ \langle mn | F_{12} \hat{Q}_{12} | \alpha\beta \rangle T_{op}^{ik} \langle op | F_{12} \hat{Q}_{12} | \alpha\gamma \rangle J_{\gamma\beta}^{kj}. \end{aligned} \quad (\text{A11})$$

The second summand of this expression is problematic. Since none of the tensors share more than one external index, no exact RI can be used and all external indices need to be resolved simultaneously with approximate RIs in the combined MO+CABS basis. Furthermore, the approximate $G_{\alpha\beta}^{ij} = \dots - T_{\alpha\gamma}^{ik} J_{\gamma\beta}^{kj}$ must be explicitly formed as first contraction because \mathcal{F} shares only one index with either T or J . This also holds for the other contributions to G . As a result, in the G -terms nothing is saved compared with actually doing the CEPA calculation with the combined MO+CABS bases *as orbital basis*. If the T_{mn}^{ij} -amplitudes were optimized, this would have to be done in each CCSD iteration, leading to a strongly increased iteration time.

¹ A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, *J. Chem. Phys.* **125**, 144108 (2006).

² A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 11599 (2004).

³ P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).

⁴ S. Saebø and P. Pulay, *Chem. Phys. Lett.* **113**, 13 (1985).

⁵ S. Saebø and P. Pulay, *J. Chem. Phys.* **86**, 914 (1987).

⁶ C. Hampel and H.-J. Werner, *J. Chem. Phys.* **104**, 6286 (1996).

⁷ M. Schütz, G. Hetzer, and H.-J. Werner, *J. Chem. Phys.* **111**, 5691 (1999).

⁸ M. Schütz and H.-J. Werner, *Chem. Phys. Lett.* **318**, 370 (2000).

⁹ M. Schütz, *J. Chem. Phys.* **113**, 9986 (2000).

¹⁰ M. Schütz and H.-J. Werner, *J. Chem. Phys.* **114**, 661 (2001).

- ¹¹E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).
- ¹²W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- ¹³V. Termath, W. Klopper, and W. Kutzelnigg, *J. Chem. Phys.* **94**, 2002 (1991).
- ¹⁴W. Klopper and W. Kutzelnigg, *J. Chem. Phys.* **94**, 2020 (1991).
- ¹⁵J. Noga, W. Kutzelnigg, and W. Klopper, *Chem. Phys. Lett.* **199**, 497 (1992).
- ¹⁶J. Noga and W. Kutzelnigg, *J. Chem. Phys.* **101**, 7738 (1994).
- ¹⁷R. J. Gdanitz, *J. Chem. Phys.* **109**, 9795 (1998).
- ¹⁸W. Klopper and C. C. M. Samson, *J. Chem. Phys.* **116**, 6397 (2002).
- ¹⁹E. F. Valeev, *Chem. Phys. Lett.* **395**, 190 (2004).
- ²⁰S. Ten-no and F. R. Manby, *J. Chem. Phys.* **119**, 5358 (2003).
- ²¹F. R. Manby, *J. Chem. Phys.* **119**, 4607 (2003).
- ²²S. Ten-no, *Chem. Phys. Lett.* **398**, 56 (2004).
- ²³A. J. May, E. Valeev, R. Polly, and F. R. Manby, *Phys. Chem. Chem. Phys.* **7**, 2710 (2005).
- ²⁴D. P. Tew and W. Klopper, *J. Chem. Phys.* **123**, 074101 (2005).
- ²⁵C. Neiss, C. Hättig, and W. Klopper, *J. Chem. Phys.* **125**, 064111 (2006).
- ²⁶H. Fliegl, C. Hättig, and W. Klopper, *J. Chem. Phys.* **124**, 044112 (2006).
- ²⁷H.-J. Werner and F. R. Manby, *J. Chem. Phys.* **124**, 054114 (2006).
- ²⁸F. R. Manby, H.-J. Werner, T. B. Adler, and A. J. May, *J. Chem. Phys.* **124**, 094103 (2006).
- ²⁹H.-J. Werner, *J. Chem. Phys.* **129**, 101103 (2008).
- ³⁰T. B. Adler, F. R. Manby, and H.-J. Werner, *J. Chem. Phys.* **130**, 054106 (2009).
- ³¹H.-J. Werner, T. B. Adler, and F. R. Manby, *J. Chem. Phys.* **126**, 164102 (2007).
- ³²J. Noga, S. Kedžuch, and J. Šimunek, *J. Chem. Phys.* **127**, 034106 (2007).
- ³³J. Noga and P. Valiron, *Chem. Phys. Lett.* **324**, 166 (2000).
- ³⁴H. Fliegl, W. Klopper, and C. Hättig, *J. Chem. Phys.* **122**, 084107 (2005).
- ³⁵H. Fliegl, C. Hättig, and W. Klopper, *Int. J. Quantum Chem.* **106**, 2306 (2006).
- ³⁶D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Phys. Chem. Chem. Phys.* **9**, 1921 (2007).
- ³⁷D. P. Tew, W. Klopper, and C. Hättig, *Chem. Phys. Lett.* **452**, 326 (2008).
- ³⁸J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, *J. Chem. Phys.* **128**, 174103 (2008).
- ³⁹T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *J. Chem. Phys.* **129**, 071101 (2008).
- ⁴⁰T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3358 (2008).
- ⁴¹C. Neiss and C. Hättig, *J. Chem. Phys.* **126**, 154101 (2007).
- ⁴²T. B. Adler, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **127**, 221106 (2007).
- ⁴³E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 106 (2008).
- ⁴⁴M. Torheyden and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008).
- ⁴⁵E. F. Valeev and T. D. Crawford, *J. Chem. Phys.* **128**, 244113 (2008).
- ⁴⁶T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- ⁴⁷R. T. Pack and W. Byers Brown, *J. Chem. Phys.* **45**, 556 (1966).
- ⁴⁸S. Ten-no, Abstr. Pap. - Am. Chem. Soc. **228**, 275 (2004).
- ⁴⁹D. P. Tew and W. Klopper, *J. Chem. Phys.* **125**, 094302 (2006).
- ⁵⁰G. Knizia and H.-J. Werner, *J. Chem. Phys.* **128**, 154103 (2008).
- ⁵¹O. Marchetti and H.-J. Werner, *Phys. Chem. Chem. Phys.* **10**, 3400 (2008).
- ⁵²C. Hampel, K. A. Peterson, and H.-J. Werner, *Chem. Phys. Lett.* **190**, 1 (1992).
- ⁵³P. J. Knowles, C. Hampel, and H.-J. Werner, *J. Chem. Phys.* **99**, 5219 (1993).
- ⁵⁴P. J. Knowles, C. Hampel, and H.-J. Werner, *J. Chem. Phys.* **112**, 3106 (2000).
- ⁵⁵W. Klopper, *J. Chem. Phys.* **120**, 10890 (2004).
- ⁵⁶T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, Chichester, 2000).
- ⁵⁷M. J. O. Deegan and P. J. Knowles, *Chem. Phys. Lett.* **227**, 321 (1994).
- ⁵⁸D. Bokhan, S. Ten-no, and J. Noga, *Phys. Chem. Chem. Phys.* **10**, 3320 (2008).
- ⁵⁹H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2008.1, a package of *ab initio* programs 2008, see <http://www.molpro.net>.
- ⁶⁰S. Kedžuch, M. Milko, and J. Noga, *Int. J. Quantum Chem.* **105**, 929 (2005).
- ⁶¹R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ⁶²T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).
- ⁶³T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- ⁶⁴K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, *J. Chem. Phys.* **112**, 9229 (2000).
- ⁶⁵F. Weigend, *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).
- ⁶⁶F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).
- ⁶⁷See EPAPS Document No. E-JCPA6-130-016902 for geometries, reference values, results using different values of γ and full inclusion of the $|ix\rangle\langle ix| + |xi\rangle\langle xi|$ projector contributions in Eq. (28), larger RI basis sets, different extrapolation schemes, and a comparison with exact CCSD-F12 correlation energies and other methods. For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.
- ⁶⁸K. Pflüger, H.-J. Werner, A. Schäfer, and M. Schütz (unpublished).
- ⁶⁹D. W. Schwenke, *J. Chem. Phys.* **122**, 014107 (2005).
- ⁷⁰L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991).
- ⁷¹K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules*, Molecular Spectra and Molecular Structure Vol. IV (Van Nostrand Reinhold, New York, 1979).
- ⁷²T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, *J. Chem. Phys.* **121**, 5874 (2004).
- ⁷³D. P. Tew, W. Klopper, M. Heckert, and J. Gauss, *J. Phys. Chem. A* **111**, 11242 (2007).
- ⁷⁴K. A. Peterson, T. B. Adler, and H.-J. Werner, *J. Chem. Phys.* **128**, 084102 (2008).
- ⁷⁵G. Rauhut, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **130**, 054105 (2009).