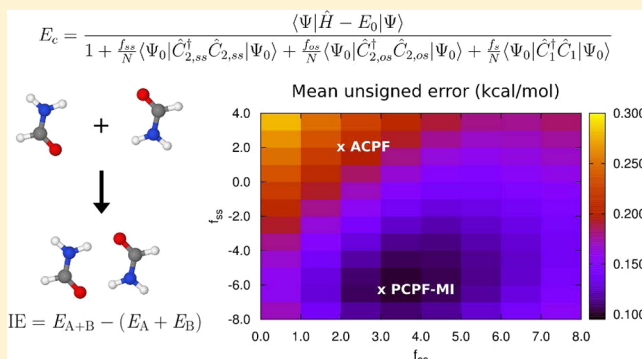# A Parametrized Coupled-Pair Functional for Molecular Interactions: PCPF-MI

Daniel R. Nascimento and A. Eugene DePrince, III*

Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32306-4390, United States

ABSTRACT: We present a parametrized coupled-pair functional, optimized to describe molecular interactions (PCPF-MI). The method has the same computational cost as singles and doubles configuration interaction, is size extensive, and yields energies that are stationary with respect to variations in all of its excitation coefficients. This last property facilitates the construction of density matrices and the evaluation of one- and two-electron properties. For the S22, HSG, S66, and A24 databases of van der Waals dimers, PCPF-MI computations yield interaction energies with mean unsigned errors of 0.326, 0.149, 0.214, and 0.044 kcal mol^{-1}, respectively, relative to benchmark computations. PCPF-MI interaction energies consistently improve upon those obtained from several other coupled-pair methods, including the averaged coupled-pair functional (ACPF), the coupled electron-pair approximation [CEPA(n), $n = 0,1,3$], and averaged quadratic coupled cluster (AQCC). Optimal parameters for spin-component-scaled (SCS) variants of each of these methods are also presented. SCS-CEPA(0), SCS-CEPA(1), SCS-ACPF, and SCS-PCPF-MI all perform similarly, with average errors for the four databases of roughly 0.1 kcal mol^{-1}. The PCPF-MI method, without additional SCS parametrization, has an average error of 0.192 kcal mol^{-1} over the four databases and, when compared to the SCS-parametrized coupled-pair methods, has the added benefit that the energy is stationary with respect to variations in all its excitation coefficients.

$$E_c = \frac{\langle\Psi|\hat{H} - E_0|\Psi\rangle}{1 + \frac{f_{ss}}{N}\langle\Psi_0|\hat{C}_{2,ss}^\dagger\hat{C}_{2,ss}|\Psi_0\rangle + \frac{f_{os}}{N}\langle\Psi_0|\hat{C}_{2,os}^\dagger\hat{C}_{2,os}|\Psi_0\rangle + \frac{f_s}{N}\langle\Psi_0|\hat{C}_1^\dagger\hat{C}_1|\Psi_0\rangle}$$

$$IE = E_{A+B} - (E_A + E_B)$$

1. INTRODUCTION

The accurate description of intermolecular interactions is a great challenge in computational chemistry, requiring a sophisticated treatment of electron correlation and the use of large (at least double-ζ quality) basis sets, augmented by diffuse functions. Coupled-cluster with single, double, and perturbative triple excitations [CCSD(T)],[1] regarded as the gold standard of quantum chemistry, has the potential to provide a quantitative description of these interactions. However, the steep computational scaling of CCSD(T) precludes its routine application to systems with more than two dozen heavy atoms. Hence, there is a rich literature describing reduced-cost modifications of, or alternatives to, CCSD(T) that aim to retain its predictive accuracy.

Spin-component-scaled (SCS) methods, inspired by the work of Grimme and co-workers,[2] achieve high accuracy by incorporating empirical parameters into the expression for the correlation energy obtained from a standard second-order perturbation theory (MP2) or singles and doubles coupled-cluster (CCSD)[3] computation. This class of methods includes SCS-MP2,[2] SCS-MP2 for molecular interactions [SCS(MI)-MP2],[4] SCS-CCSD,[5] SCS-CCSD for molecular interactions [SCS(MI)-CCSD],[6] and scaled-opposite-spin MP2 (SOS-MP2).[7] SCS(MI)-CCSD provides a particularly good description of noncovalent interactions; for the S22 set of van der Waals dimers,[8] SCS(MI)-CCSD/aug-cc-pVDZ computations provide interaction energies with average errors of only 0.05 kcal mol^{-1}, relative to those obtained at the CCSD(T)/

estimated complete-basis-set (CBS) limit level of theory.[6] Spin-component scaling has thus become a popular strategy to obtain near-CCSD(T)-quality interaction energies at CCSD or even MP2 computational cost.

Coupled-pair methods, which can be viewed as either simplifications to CCSD or size-extensive modifications of singles and doubles configuration interaction (CISD), provide a description of electron correlation that is often intermediate in quality between CCSD and CCSD(T)[9,10] at a computational cost equal to that of CISD. This class of methods includes the averaged coupled-pair functional (ACPF),[11] averaged quadratic coupled cluster (AQCC),[12] and the coupled electron pair approximation [CEPA(n), $n = 0,1,3$].[13−22] Coupled-pair methods are not often used to describe intermolecular interactions, perhaps due to a lack of systematic investigations of their accuracy for these systems. CEPA(1) has previously been shown to achieve a reasonably small mean unsigned error (MUE) of 0.24 kcal mol^{-1} for the S22 set, relative to benchmark computations.[23] In this paper, we develop a new, parametrized coupled-pair functional to describe molecular interactions, and we systematically assess its ability, and the ability of other coupled-pair methods, to provide high-quality noncovalent interaction energies.

To improve the description of noncovalent interactions afforded by coupled-pair methods, one could follow the route

Table 1. Values of Δ_2^{ss}, Δ_2^{os}, and Δ_1 that Define Several Coupled-Pair Methods

method	Δ_2^{ss}	Δ_2^{os}	Δ_1
CEPA(0)	0	0	0
CISD	E_c	E_c	E_c
AQCC	$((4N-6)/(N(N-1)))E_c$	$((4N-6)/(N(N-1)))E_c$	$((4N-6)/(N(N-1)))E_c$
ACPF	$(2/N)E_c$	$(2/N)E_c$	$(2/N)E_c$
PCPF-MI	$(f_{ss}/N)E_c$	$(f_{os}/N)E_c$	$(f_s/N)E_c$

of Grimme and others, determining optimal SCS parameters for each of these methods, but we would prefer to modify the methods in a way that would leave some of their more desirable features intact. For some coupled-pair methods, the correlation energy can be expressed as a functional, and the correlation energy and excitation coefficients are obtained by minimizing this functional with respect to variations in the coefficients. The stationary nature of these methods greatly simplifies the construction of density matrices and the evaluation of geometric derivatives or one-electron properties. In this paper, we incorporate same-spin and opposite-spin parameters directly into a coupled-pair energy functional, thus preserving the stationary nature of the method. The resulting approach, which we call a parametrized coupled-pair functional for molecular interactions (PCPF-MI), provides a better description of noncovalent interactions than any of these other coupled-pair methods as compared to benchmark interaction energies. Conventional spin-component-scaled parameters can, of course, be optimized for the PCPF-MI, but the resulting SCS-PCPF-MI energies are not stationary with respect to variations in the wave function parameters, and that particular style of parametrization runs contrary to the spirit of the method. Nonetheless, we also present optimal SCS parameters for PCPF-MI and all of the other coupled-pair methods considered herein.

2. THEORY

Coupled-pair methods can be motivated as size-extensive extensions of truncated configuration interaction or as linear approximations to coupled cluster theory, and the working equations of the methods appear very similar to those of CISD or linearized CCSD. The CISD wave function (in intermediate normalization) takes the form

$$|\Psi\rangle = (1 + \hat{C}_1 + \hat{C}_2)|\Psi_0\rangle \tag{1}$$

where $|\Psi_0\rangle$ represents a closed-shell restricted Hartree–Fock reference determinant, and $\hat{C}_1$ and $\hat{C}_2$ are single and double excitation operators, respectively:

$$\hat{C}_1 = \sum_{ia} c_i^a a^\dagger i \tag{2}$$

$$\hat{C}_2 = \frac{1}{4} \sum_{ij,ab} c_{ij}^{ab} a^\dagger b^\dagger ji \tag{3}$$

The symbols c_i^a and c_{ij}^{ab} represent the corresponding single and double excitation coefficients, and the indices i and j (a and b) represent spin–orbitals that are occupied (virtual) in the reference wave function. The CISD correlation energy can be obtained by minimizing the energy functional

$$E_c = \frac{\langle\Psi|\hat{H} - E_0|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{4}$$

with respect to the coefficients c_i^a and c_{ij}^{ab}. Here, E_c represents the correlation energy, and E_0 represents the energy of the reference configuration.

In the ACPF and AQCC methods, size-extensivity is restored by modifying the denominator of eq 4 to give

$$E_c = \frac{\langle\Psi|\hat{H} - E_0|\Psi\rangle}{1 + \frac{f}{N}(\langle\Psi_0|\hat{C}_2^\dagger \hat{C}_2|\Psi_0\rangle + \langle\Psi_0|\hat{C}_1^\dagger \hat{C}_1|\Psi_0\rangle)} \tag{5}$$

where f/N is a renormalization factor that depends on the number of electrons N in the system. We note that energy given by eq 5 is only strictly size extensive for systems of identical molecules. The ACPF and AQCC methods are obtained by choosing $f = 2$ or $((4N-6)/(N-1))$, respectively. If we decompose the overlap of the doubly excited functions, $\langle\Psi_0|\hat{C}_2^\dagger \hat{C}_2|\Psi_0\rangle$, into opposite-spin and same-spin components, we can increase the flexibility of the method by assigning a different renormalization factor to each component. This more general energy functional is now

$$E_c = \frac{\langle\Psi|\hat{H} - E_0|\Psi\rangle}{1 + \frac{f_{ss}}{N}\langle\Psi_0|\hat{C}_{2,ss}^\dagger \hat{C}_{2,ss}|\Psi_0\rangle + \frac{f_{os}}{N}\langle\Psi_0|\hat{C}_{2,os}^\dagger \hat{C}_{2,os}|\Psi_0\rangle + \frac{f_s}{N}\langle\Psi_0|\hat{C}_1^\dagger \hat{C}_1|\Psi_0\rangle} \tag{6}$$

The same-spin, opposite-spin, and singles renormalization factors are denoted by f_{ss}/N, f_{os}/N, and f_s/N, respectively, and the opposite-spin and same-spin double-excitation operators are defined as

$$\hat{C}_{2,os} = \sum_{ij,ab} c_{i\bar{j}}^{a\bar{b}} a^\dagger \bar{b}^\dagger \bar{j} i \tag{7}$$

$$\hat{C}_{2,ss} = \frac{1}{4} \sum_{ij,ab} (c_{ij}^{ab} a^\dagger b^\dagger ji + c_{\bar{i}\bar{j}}^{\bar{a}\bar{b}} \bar{a}^\dagger \bar{b}^\dagger \bar{j} \bar{i}) \tag{8}$$

In eqs 7 and 8, an overbar denotes β spin. We consider only closed-shell restricted wave functions, so $c_{ij}^{ab} = c_{i\bar{j}}^{a\bar{b}}$.

The excitation coefficients are obtained by enforcing the stationary condition $(\partial E/\partial c) = 0$, which results in the following set of coupled equations:

$$\langle\Psi_{ij}^{ab}|\hat{H} - E_0 - \Delta_2^{ss}|\Psi\rangle = 0 \tag{9}$$

$$\langle\Psi_{i\bar{j}}^{a\bar{b}}|\hat{H} - E_0 - \Delta_2^{os}|\Psi\rangle = 0 \tag{10}$$

$$\langle\Psi_i^a|\hat{H} - E_0 - \Delta_1|\Psi\rangle = 0 \tag{11}$$

Here, the symbols $\langle\Psi_{ij}^{ab}|$, $\langle\Psi_{i\bar{j}}^{a\bar{b}}|$, and $\langle\Psi_i^a|$ represent doubly or singly substituted configurations, and different choices for Δ_2^{ss}, Δ_2^{os}, and Δ_1 define various coupled-pair approximations. The values for Δ_2^{ss}, Δ_2^{os}, and Δ_1 that define some of the coupled-pair methods studied here are provided in Table 1. The optimal scaling factors, f_{os}, f_{ss}, and f_s for PCPF-MI have yet to be specified, so we optimize them to minimize the error in the interaction energies for small noncovalently bound dimers
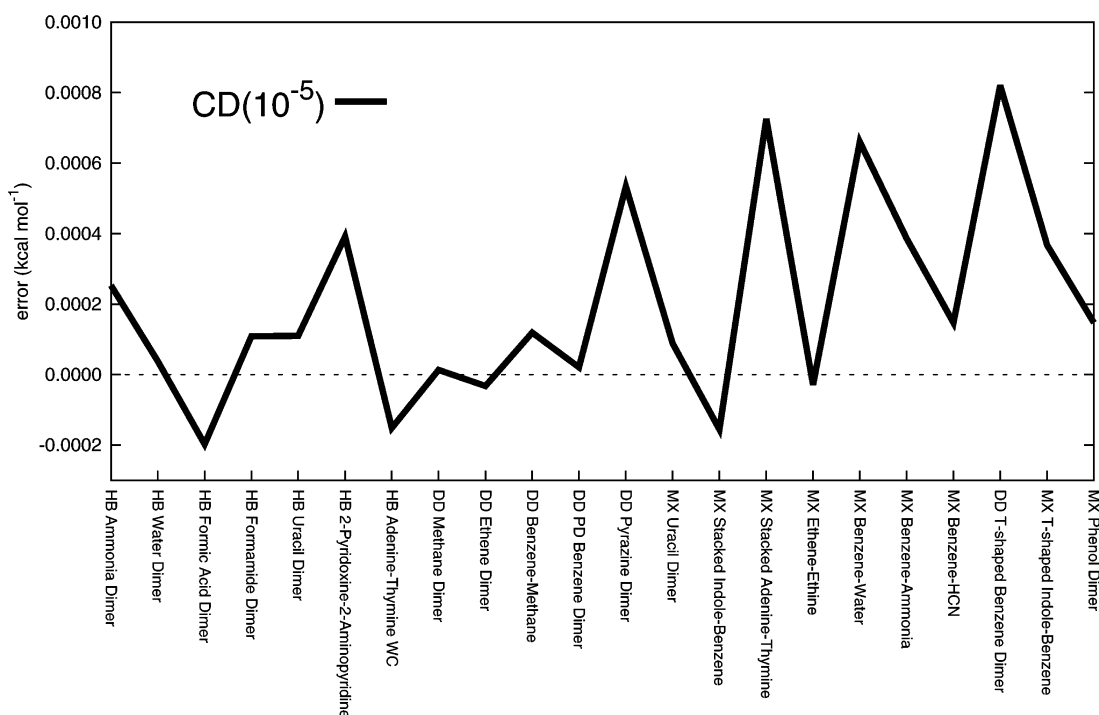
Figure 1. Errors in the interaction energies (kcal mol^{-1}) computed at the CD-CCSD(T)/aug-cc-pVDZ level of theory relative to those obtained by canonical CCSD(T)/aug-cc-pVDZ methods utilizing conventional four-index integrals. Interaction types are denoted as hydrogen-bonded (HB), dispersion-dominated (DD), and mixed interactions (MX).

relative to those obtained from CCSD(T) computations. Other CEPA variants [CEPA(n), $n = 1,2,3$] can be implemented using eqs 9−11 by choosing orbital-dependent values for Δ_2^{os}, Δ_2^{ss}, and Δ_1.[13−22] However, energies obtained using orbital-dependent shifts are not orbital invariant, and the coefficients do not satisfy the stationary condition because the energy can no longer be expressed as a functional. We note here that there exist several variational methods that exhibit similar numerical behavior to variants of CEPA that use orbital-dependent shifts; these methods include the coupled pair functional,[24] variational CEPA,[25] and the parametric two-electron reduced-density-matrix (2-RDM) method.[26−28]

3. COMPUTATIONAL DETAILS

The CEPA/ACPF/AQCC/PCPF-MI methods were implemented in a development version of the Psi4 electronic structure package.[29] Unless otherwise noted, all computations utilize Cholesky-decomposed (CD) two-electron integrals in the Hartree−Fock and correlated portions of the algorithm. The CD integrals were generated by the lib3index library in the Psi4 package with a very conservative[30] error tolerance of 10^{-5}. We account for basis set superposition error by the counterpoise method.[31] The computer implementation of the coupled-pair methods was validated against the CEPA(n) (n = 0,1,3), ACPF, and AQCC implementations in the Molpro electronic structure package.[32] In all tests against Molpro, we used a CD threshold of 10^{-12} to match Molpro results obtained with conventional four-index integrals. All computations were performed within the frozen core approximation.

To obtain estimated CBS-limit interaction energies for the coupled-pair methods, we employ a focal-point analysis[33,34] that leads to an estimate of the CBS-limit energy of

$$E_X^{CBS} \approx E_{MP2}^{CBS} + \delta_{MP2}^X \tag{12}$$

$$\delta_{MP2}^X = [E_X^{small} - E_{MP2}^{small}] \tag{13}$$

Here, X denotes a given coupled-pair or coupled-cluster method. We choose the "small" basis set to be either the aug-cc-pVDZ or aug-cc-pVTZ basis, and the CBS-limit MP2 energy is estimated from a two-point Helgaker extrapolation scheme[35] using the aug-cc-pVTZ and aug-cc-pVQZ basis sets, as implemented in Psi4.

4. RESULTS AND DISCUSSION

Approximate factorizations of the electron repulsion integral tensor are well-established in quantum chemistry. Systematic studies of the accuracy of CD and density fitting approximations in the context of noncovalent interactions have been presented elsewhere,[30] so here we only consider the error arising from the CD approximation for interaction energies computed at the CCSD(T) level of theory within the aug-cc-pVDZ basis set.[36] Figure 1 illustrates the error in the interaction energies for the set of 22 van der Waals dimers that comprise the S22 set.[8] The MUE for CD-CCSD(T) is only 0.0002 kcal mol^{-1}, which is insignificant relative to errors due to the incomplete descriptions of electron correlation and the single particle basis set. Hence, errors resulting from the use of the CD approximation are considered negligible when comparing the present results to existing benchmark interaction energies. Unless otherwise noted, the CD approximation with a threshold of 10^{-5} is utilized for all CCSD(T) and coupled-pair computations in the remainder of this manuscript.

The f_{ss} and f_{os} parameters of the PCPF-MI method were optimized by minimizing the MUE in the interaction energies of the 11 smallest members of the S22 set (denoted S11 and presented in Figure 2) relative to the interaction energies obtained at the CCSD(T) level of theory. Here, both PCPF-MI and CCSD(T) computations were performed within the aug-
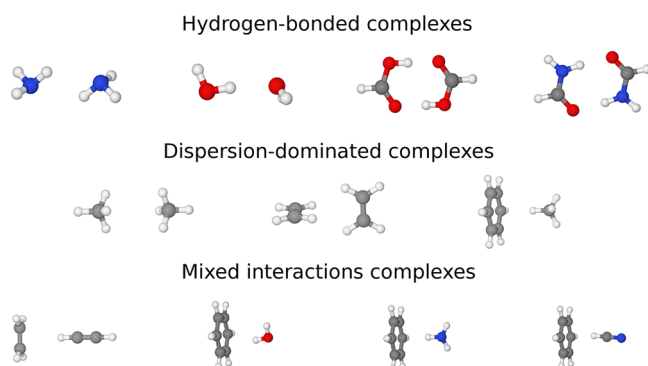
Figure 2. S11 set of van der Waals dimers used to parametrize the PCPF-MI method.

cc-pVDZ basis set. This optimization is more complicated than the SCS parametrization of the SCS-MP2/SCS-CCSD methods in that each component of the counterpoise-corrected interaction energy must be minimized for each possible set of parameters. Interaction energies were computed for the S11 set at the PCPF-MI/aug-cc-pVDZ level of theory using values of f_{os} in the range from 0.0 to 8.0 and f_{ss} in the range from −8.0 to 4.0. For this optimization, the singles factor is held fixed at the value used in the ACPF method (f_s = 2). PCPF-MI interaction energies were then compared to those obtained at the CCSD(T)/aug-cc-pVDZ level of theory, and the MUE for the PCPF-MI is provided in Figure 3. Figure 3a depicts a coarse-grain parameter sweep, where f_{os} and f_{ss} were varied by
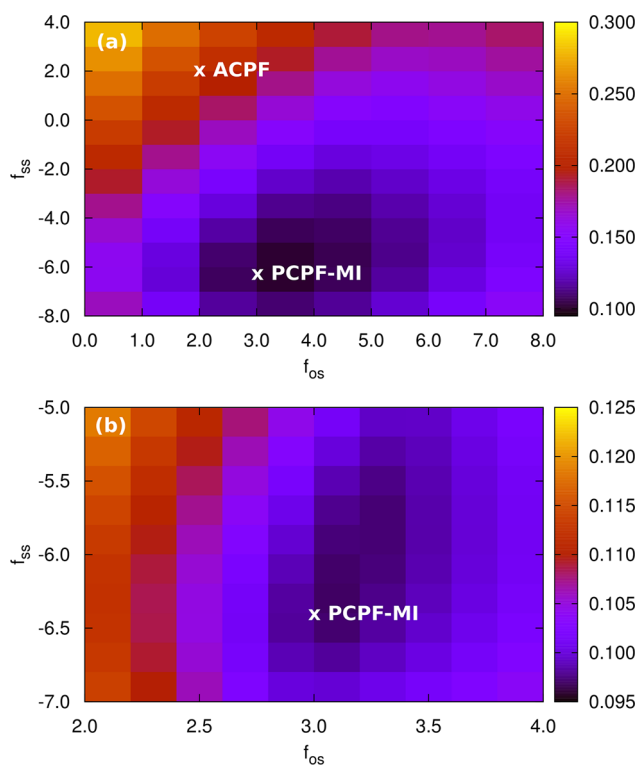

Figure 3. Mean unsigned error (kcal mol^{-1}) in the interaction energies for the S11 set of van der Waals dimers computed at the PCPF-MI/aug-cc-pVDZ level of theory, relative to interaction energies obtained at the CCSD(T)/aug-cc-pVDZ level of theory in the region (a) $f_{ss} \in \{-8.0$ to $4.0\}$, $f_{os} \in \{0.0$ to $8.0\}$ and (b) $f_{ss} \in \{-7.0$ to $-5.0\}$, $f_{os} \in \{2.0$ to $4.0\}$.

1.0. Figure 3b shows a refined search near the minimum in the MUE where f_{os} and f_{ss} were varied by 0.2. The minimum in the MUE for the PCPF-MI method in Figure 3b is only 0.096 kcal mol^{-1} and occurs with f_{os} = 3.0 and f_{ss} = −6.4. These values will define the f_{os} and f_{ss} parameters for the remainder of this manuscript. The ACPF method can be located in Figure 3a at $f_{os} = f_{ss} = 2$, and the ACPF MUE for the S11 set is 0.216 kcal mol^{-1}.

Note that the sign of the optimal f_{ss} is negative; in the context of conventional SCS parametrization, it is understood that scaling parameters should take only positive values, as the same-spin and opposite-spin contributions to the correlation energy, by definition, must be negative.[4] In the present scheme, however, we are not scaling the same-spin and opposite-spin components of the energy directly; we are scaling contributions to the normalization denominator. Using the present parametrization and minimizing the energy given by the resulting functional, a negative value for f_{ss} can never result in a positive correlation energy. Allowing the parameters to assume negative values merely increases the flexibility of the energy functional.

Additional computations to parametrize the singles normalization factor were also performed. The MUE in Figure 4
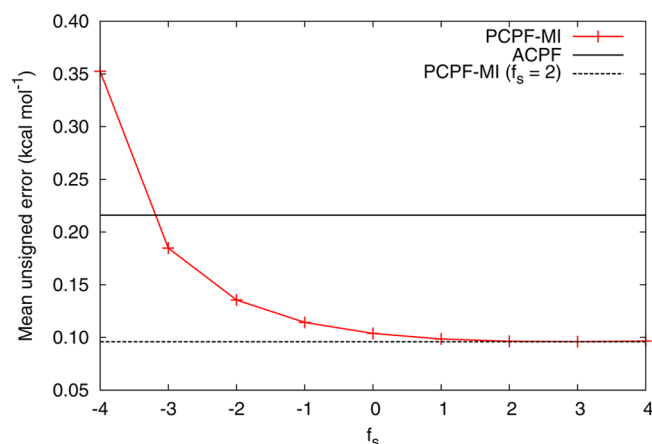

Figure 4. Mean unsigned error (kcal mol^{-1}) in the interaction energies for the S11 set of van der Waals dimers computed at the PCPF-MI/aug-cc-pVDZ level of theory, relative to interaction energies obtained at the CCSD(T)/aug-cc-pVDZ level of theory as a function of the parameter f_s.

indicates that the accuracy of PCPF-MI for the S11 set cannot be improved much by varying f_s in the range from −4.0 to 4.0 while holding f_{os} and f_{ss} fixed at 3.0 and −6.4, respectively. The error is essentially flat near f_s = 2, which is the value of f_s in the ACPF method. Hence, PCPF-MI will retain the singles factor defined for ACPF. We assess in Figure 5 the basis-set-dependence of the optimized PCPF-MI f_{os} factor. For the S11 set within the aug-cc-pVTZ basis set, f_{os} was varied in the vicinity of the optimal aug-cc-pVDZ basis value (3.0). While the MUE in the aug-cc-pVTZ basis is slightly larger than that in the aug-cc-pVDZ basis, the minimum occurs at the same value for f_{os}. We therefore conclude that the parameters optimized for the aug-cc-pVDZ basis are insensitive to the choice of basis.

The error statistics for each of the coupled-pair methods discussed above are illustrated in Table 2. We consider the full S22 database, as well as the HSG,[37] S66,[38] and A24[39] databases. We use the revised benchmark interaction energies of ref 40 for the S22 and HSG sets. Estimated CBS-limit coupled-pair interaction energies are obtained from coupled-pair energies
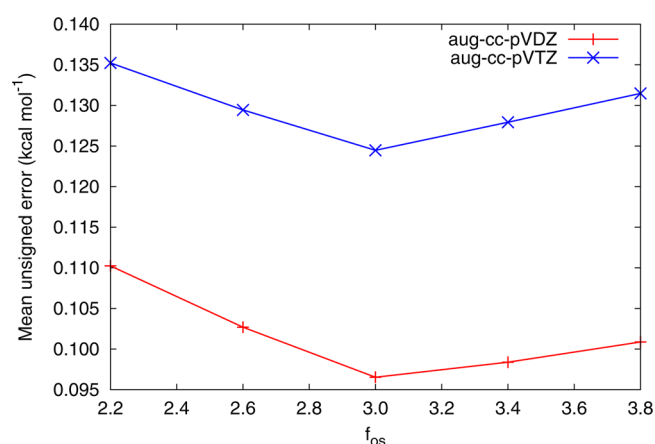
Figure 5. Mean unsigned error in PCPF-MI interaction energies relative to CCSD(T) for the S11 set within the aug-cc-pVDZ and aug-cc-pVTZ basis.

obtained within the aug-cc-pVDZ basis set, combined with the δ_{MP2}^X correction described in Section 3. Some coupled-pair methods perform reasonably well, with mean unsigned errors of less than 0.5 kcal mol^{-1}. CEPA(0), CEPA(1), and ACPF clearly outperform CCSD in terms of MUEs. AQCC and CEPA(3) perform worse with average MUEs over the four databases near 1 kcal mol^{-1}. The PCPF-MI parametrization substantially improves the quality of the interaction energies; the PCPF-MI MUE over the four databases is less than half of that from the ACPF. The full set of 133 interaction energies is more general than the S11 training set; it is reassuring that the PCPF-MI MUEs for the HSG, S66, and A24 sets are smaller than that for the S22 set, despite the fact that S22 contains the training set. Table 2 also includes the standard deviation in the energy errors and the maximum errors for each method. With the exception of the standard deviation for the HSG set, the PCPF-MI method exhibits the best errors statistics accross each metric and data set. Of all coupled-pair methods considered, PCPF-MI is the only method with a maximum error of less than 1 kcal mol^{-1} for all 133 interaction energies.

SCS parametrizations in the usual sense (where post-optimization same-spin and opposite-spin correlation energies are scaled) can be obtained for any method whose correlation energy can be decomposed into same-spin and opposite-spin components. Table 3 presents mean unsigned errors for SCS variants of all coupled-pair methods considered in this manuscript. SCS parameters were optimized for the revised S22 database[40] and tested against the revised HSG,[40] S66, and A24 databases. The results suggest that SCS parametrization is the great equalizer among coupled-pair methods, and four SCS methods emerge as accurate alternatives to CCSD(T): SCS-CEPA(0), SCS-CEPA(1), SCS-ACPF, and SCS-PCPF-MI. Among these methods, SCS-CEPA(0) and SCS-ACPF exhibit MUEs over the four test sets of roughly 0.11 kcal mol^{-1}, while the MUEs for SCS-CEPA(1) and SCS-PCPF-MI are about 0.10 kcal mol^{-1}. Maximum errors are also quite small for these four SCS methods; on average, they exhibit maximum errors over the four databases of about 0.3 kcal mol^{-1}. Alternatively, some methods perform quite poorly, even with spin-component-scaling. SCS-CEPA(3) exhibits an average maximum error over the four databases of more than 1 kcal mol^{-1}. We have also included results for SCS(MI)-CCSD, which we have reparameterized for this work against the revised S22 database. The δ_{MP2}^X correction used herein is slightly different than that

Table 2. Error Statistics for Interaction Energies (kcal mol^{-1}) Obtained from Several Coupled-Pair Methods Relative to Benchmark Computations

database	method[a]	MUE[b]	STDDEV[c]	max E[d]
S22	CCSD(T)	0.081	0.092	0.291
	CCSD	0.835	0.610	2.370
	CEPA(0)	0.555	0.377	1.233
	CEPA(1)	0.674	0.519	1.973
	CEPA(3)	1.346	0.852	3.105
	ACPF	0.644	0.489	1.830
	AQCC	1.125	0.620	2.707
	PCPF-MI	0.326	0.273	0.978
HSG	CCSD(T)	0.024	0.036	0.156
	CCSD	0.475	0.199	1.104
	CEPA(0)	0.202	0.143	0.584
	CEPA(1)	0.313	0.124	0.649
	CEPA(3)	0.719	0.230	1.427
	ACPF	0.274	0.118	0.609
	AQCC	1.207	0.506	2.784
	PCPF-MI	0.149	0.119	0.353
S66	CCSD(T)	0.075	0.070	0.317
	CCSD	0.655	0.290	1.634
	CEPA(0)	0.360	0.198	0.938
	CEPA(1)	0.485	0.235	1.272
	CEPA(3)	0.973	0.492	2.879
	ACPF	0.468	0.226	1.167
	AQCC	0.943	0.293	1.902
	PCPF-MI	0.214	0.121	0.552
A24	CCSD(T)	0.029	0.029	0.086
	CCSD	0.186	0.079	0.392
	CEPA(0)	0.083	0.061	0.244
	CEPA(1)	0.142	0.060	0.252
	CEPA(3)	0.585	0.571	2.201
	ACPF	0.121	0.063	0.295
	AQCC	0.518	0.145	0.743
	PCPF-MI	0.044	0.040	0.173
average	CCSD(T)	0.060	0.061	0.246
	CCSD	0.572	0.290	1.448
	CEPA(0)	0.317	0.194	0.806
	CEPA(1)	0.427	0.233	1.106
	CEPA(3)	0.925	0.524	2.565
	ACPF	0.404	0.223	1.031
	AQCC	0.938	0.354	1.965
	PCPF-MI	0.192	0.131	0.523

[a]Estimated CBS-limit interaction energies were obtained from aug-cc-pVDZ coupled-pair computations, combined with the δ_{MP2}^X correction described in Section 3. Results are presented for the S22, HSG, S66, and HSG sets of van der Waals dimers. [b]Mean unsigned error. [c]Standard deviation in the MUE. [d]Maximum error.

used in ref 6, so a fair comparison to the present SCS coupled-pair methods necessitated the reparametrization. SCS(MI)-CCSD provides a considerably better description of non-covalent interactions (~0.07 kcal mol^{-1} MUE for all four databases), albeit at a slightly increased computational cost. SCS-PCPF-MI errors for the A24 set are *larger* than the errors for the underlying PCPF-MI method, implying some degree of overparametrization. The novelty and utility of PCPF-MI lies in its stationary nature, and conventional SCS parametrization destroys this property. Thus, SCS(MI)-CCSD may be a better choice when stationarity is not required.

Table 3. Error Statistics for Interaction Energies (kcal mol^{-1}) Obtained from Several Spin-Component-Scaled Coupled-Pair Methods Relative to Benchmark Computations

database	method[a]	OS	SS	MUE[b]	STDDEV[c]	max E[d]
S22	CCSD	1.02	1.35	0.050	0.034	0.122
	CEPA(0)	0.95	1.23	0.122	0.098	0.324
	CEPA(1)	1.07	1.25	0.064	0.046	0.213
	CEPA(3)	1.00	1.53	0.276	0.229	0.923
	ACPF	1.05	1.25	0.071	0.052	0.175
	AQCC	1.12	1.41	0.162	0.126	0.439
	PCPF-MI	1.10	1.08	0.085	0.087	0.312
HSG	CCSD			0.076	0.044	0.169
	CEPA(0)			0.128	0.054	0.223
	CEPA(1)			0.149	0.057	0.252
	CEPA(3)			0.133	0.106	0.472
	ACPF			0.173	0.073	0.280
	AQCC			0.519	0.305	1.380
	PCPF-MI			0.113	0.067	0.245
S66	CCSD			0.076	0.058	0.263
	CEPA(0)			0.113	0.090	0.402
	CEPA(1)			0.118	0.102	0.448
	CEPA(3)			0.227	0.215	0.988
	ACPF			0.116	0.099	0.421
	AQCC			0.165	0.128	0.504
	PCPF-MI			0.106	0.112	0.392
A24	CCSD			0.056	0.030	0.135
	CEPA(0)			0.082	0.054	0.232
	CEPA(1)			0.051	0.031	0.128
	CEPA(3)			0.321	0.527	1.883
	ACPF			0.085	0.060	0.239
	AQCC			0.263	0.121	0.583
	PCPF-MI			0.097	0.069	0.285
average	CCSD			0.068	0.047	0.202
	CEPA(0)			0.111	0.079	0.330
	CEPA(1)			0.102	0.073	0.320
	CEPA(3)			0.237	0.256	1.057
	ACPF			0.112	0.080	0.325
	AQCC			0.238	0.154	0.646
	PCPF-MI			0.102	0.093	0.336

[a]Estimated CBS-limit interaction energies were obtained from aug-cc-pVDZ coupled-pair computations, combined with the δ^{X}_{MP2} correction described in Section 3. Results are presented for the S22, HSG, S66, and A24 sets of van der Waals dimers. Optimal SCS parameters are also provided. [b]Mean unsigned error. [c]Standard deviation in the MUE. [d]Maximum error.

Figure 6 provides the MUE for the PCPF-MI method decomposed by interaction type (hydrogen-bonded, dispersion-dominated, and mixed interactions). It is immediately clear that the PCPF-MI method provides a better description of hydrogen-bonding interactions than dispersion-dominated interactions. This result is not surprising; the S11 training set contains only three dispersion-dominated complexes. SCS parametrization improves the quality of the interaction energies for dispersion-bound complexes and lowers the total error over the four databases to 0.102 kcal mol^{-1}. Again, however, SCS-PCPF-MI is to some degree an overparametrized method, as is evidenced by the slight increase in errors for the A24 database.

Estimated CBS-limit coupled-pair interaction energies were also obtained from results obtained within the aug-cc-pVTZ basis set and are compared to benchmark interaction energies from the S11 and A24 databases in Table 4. Again, we use a
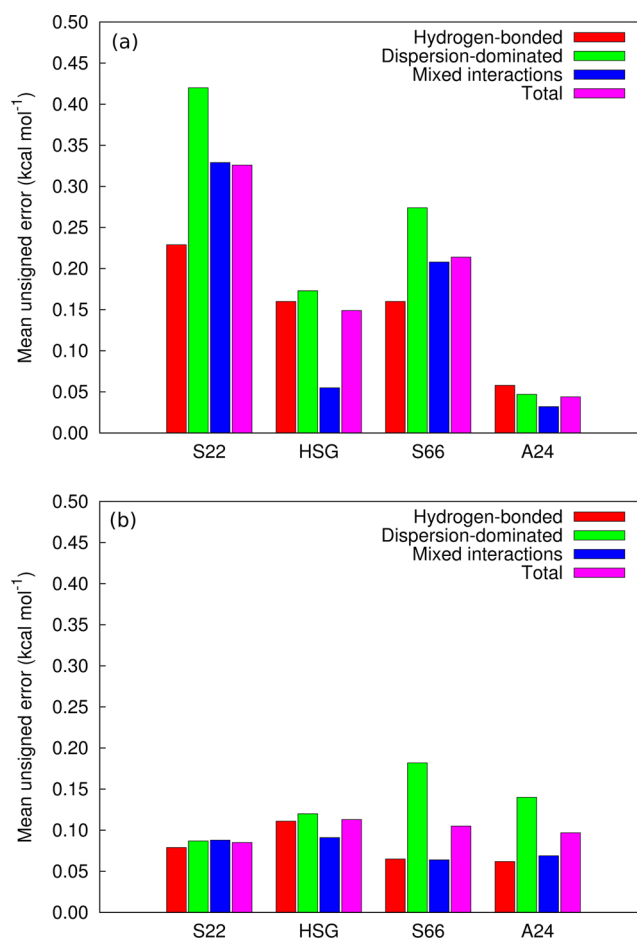

Figure 6. Mean unsigned error (kcal mol^{-1}) for (a) PCPF-MI and (b) SCS-PCPF-MI interaction energies for the S22, HSG, S66, and A24 databases. Errors are decomposed by binding type and are relative to benchmark interaction energies. Estimated CBS-limit interaction energies were obtained from aug-cc-pVDZ PCPF-MI computations, combined with the δ^{X}_{MP2} correction described in Section 3.

δ^{X}_{MP2} correction to the coupled-pair energies, as described in Section 3. Since the PCPF-MI method was optimized using the S11 set of van der Waals dimers represented by the aug-cc-pVDZ basis, we expect this method to perform better than all other coupled-pair methods. As has been consistently observed, PCPF-MI is the coupled-pair method that most accurately reproduces benchmark interaction energies, with an MUE value of 0.125 kcal mol^{-1}. This value is not much larger than the MUE observed for the S11 set in the aug-cc-pVDZ basis set: 0.096 kcal mol^{-1}. It is important to note that PCPF-MI also yields the most accurate interaction energies for the A24 set, with an MUE value of 0.038 kcal mol^{-1}. Spin-component-scaled interaction energies are in better agreement with benchmark values; MUE values for SCS(MI)-CCSD, SCS-CEPA(0), SCS-CEPA(1), SCS-ACPF, and SCS-PCPF-MI are all less than or close to 0.1 kcal mol^{-1}. SCS-CEPA(0) and SCS-CEPA(1) performs the best here, with MUEs of only 0.070 and 0.063 kcal mol^{-1} for the S11 and A24 sets, respectively.

5. CONCLUSIONS

A new coupled-pair functional, parametrized to describe molecular interactions, was presented. The method offers a description of nonbonded interactions that is superior to that of other coupled-pair methods of comparable cost. For the S22,

Table 4. Error Statistics for Interaction Energies (kcal mol^{-1}) Obtained from Several Coupled-Pair and Spin-Component-Scaled Coupled-Pair Methods Relative to Benchmark Computations

	method[a]	spin-component scaled?					
		no			yes		
		MUE[b]	STDDEV[c]	max E[d]	MUE	STDDEV	max E
S11	CCSD(T)	0.014	0.015	0.050	−	−	−
	CCSD	0.382	0.226	0.835	0.096	0.071	0.251
	CEPA(0)	0.289	0.243	0.744	0.070	0.062	0.192
	CEPA(1)	0.272	0.119	0.455	0.097	0.143	0.470
	CEPA(3)	1.056	0.625	2.138	0.473	0.450	1.235
	ACPF	0.268	0.165	0.546	0.104	0.079	0.280
	AQCC	0.735	0.162	0.979	0.279	0.123	0.449
	PCPF-MI	0.125	0.135	0.453	0.098	0.070	0.262
A24	CCSD(T)	0.018	0.012	0.045	−	−	−
	CCSD	0.209	0.093	0.414	0.118	0.284	0.143
	CEPA(0)	0.084	0.047	0.202	0.078	0.062	0.268
	CEPA(1)	0.167	0.091	0.366	0.063	0.043	0.152
	CEPA(3)	0.673	0.658	2.530	0.368	0.621	2.195
	ACPF	0.142	0.070	0.323	0.073	0.049	0.169
	AQCC	0.642	0.141	0.914	0.357	0.124	0.605
	PCPF-MI	0.038	0.032	0.110	0.093	0.058	0.216

[a]Results are presented for the S11 and A24 sets of van der Waals dimers. Estimated CBS-limit interaction energies were obtained from aug-cc-pVTZ coupled-pair computations, combined with the δ_{MP2}^X correction described in Section 3. [b]Mean unsigned error. [c]Standard deviation in the MUE. [d]Maximum error.

HSG, S66, and A24 databases of van der Waals dimers, PCPF-MI/aug-cc-pVDZ combined with a δ_{MP2}^X correction yields estimated CBS-limit interaction energies with MUEs of only 0.326, 0.149, 0.214, and 0.044 kcal mol^{-1}, respectively, relative to those obtained from benchmark computations. PCFP-MI/aug-cc-pVTZ+δ_{MP2}^X yields lower MUEs than all other coupled-pair methods considered for the S11 and A24 benchmark databases.

PCPF-MI does not necessarily provide a balanced description of different interaction types; the method provides a good description of hydrogen-bonded interactions, while errors in dispersion-dominated complexes are considerably larger. The description of dispersion-bound complexes is substantially improved by conventional SCS parametrization. The SCS-PCPF-MI MUE for all interaction types over the four databases is only 0.102 kcal mol^{-1}, relative to benchmark interaction energies. The description of dispersion also could be improved through the use of an expanded test set in which dispersion-bound dimers are well-represented and by considering a more flexible energy functional. The coupled pair functional[24] and parametric 2-RDM methods[26−28] both restore size extensivity to CISD by introducing a topological factor into the normalization of the wave function. In the case of the parametric 2-RDM method, the factors are chosen by considering two-particle N-representability conditions.[41] Alternatively, one could choose factors based on an optimization similar to that presented herein. However, with increasing flexibility in the energy functional comes a commensurate increase in the complexity of the parameter space.

PCPF-MI exhibits the useful feature that, upon minimization of the correlation energy, it can provide unrelaxed one- and two-electron density matrices (1-RDM and 2-RDM, respectively) without the need to solve any additional equations, as would be the case in coupled-cluster theory. This feature is desirable in the context of noncovalent interactions, where it is well-known that the interaction energy converges very slowly with the size of the one-particle basis set. In such cases,

explicitly correlated methods such as the popular CCSD(T**)-F12a/b[42,43] or related dispersion-weighted (DW) CCSD-(T**)-F12[44] methods are particularly powerful, as they provide large-basis accuracy with modest basis sets. These methods provide interaction energies with MUE values of less than 0.1 kcal mol^{-1} for the S22 set, relative to the revised S22B benchmark results of Marshall et al.[40] The perturbative explicitly correlated approximation of Valeev and co-workers, denoted as $[2]_{R12}$,[45,46] is an interesting alternative to more familiar F12 methods in that it is defined in terms of the 1-RDM and 2-RDM. The $[2]_{R12}$ approximation is generally applicable to any method for which 1-RDM and 2-RDM are known and is thus an attractive target for methods such as the PCPF-MI, which automatically yield the 2-RDM. Future work will consider the quality of PCPF-MI$[2]_{R12}$ interaction energies as well as those from other coupled-pair methods that can provide the unrelaxed 2-RDM without any additional work.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: deprince@chem.fsu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(2) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095−9102.

(3) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910−1918.

(4) Distasio, R. A., JR.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073−1083.

(5) Takatani, T.; Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2008**, *128*, 124111.

(6) Pitoňák, M.; Řezáč; Hobza, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9611−9614.

(7) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793−9802.

(8) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985−1993.

(9) Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. *Acc. Chem. Res.* **2009**, *42*, 641−648.

(10) Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, 114108.

(11) Gdanitz, R. J.; Ahlrichs, R. *Chem. Phys. Lett.* **1988**, *143*, 413−420.

(12) Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481−488.

(13) Ahlrichs, R.; Driessler, F. *Theor. Chim. Acta* **1975**, *36*, 275−287.

(14) Meyer, W. *Theor. Chim. Acta* **1974**, *35*, 277−292.

(15) Ahlrichs, R.; Lischka, H.; Zurawski, B.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *63*, 4685−4694.

(16) Ahlrichs, R.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1225−1234.

(17) Ahlrichs, R.; Keil, F.; Lischka, H.; Kutzelnigg, W.; Staemmler, V. *J. Chem. Phys.* **1975**, *63*, 455−463.

(18) Ahlrichs, R.; Driessler, F.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1235−1247.

(19) Meyer, W.; Rosmus, P. *J. Chem. Phys.* **1975**, *63*, 2356−2375.

(20) Meyer, W. *J. Chem. Phys.* **1973**, *58*, 1017−1035.

(21) Dykstra, C. E.; Schaefer, H. F., III; Meyer, W. *J. Chem. Phys.* **1976**, *65*, 2740−2750.

(22) Meyer, W. *Int. J. Quantum Chem.* **1971**, *5*, 341−348.

(23) Liakos, D. G.; Hansen, A.; Neese, F. *J. Chem. Theory Comput.* **2011**, *7*, 76−87.

(24) Ahlrichs, R.; Scharf, P.; Ehrhardt, C. *J. Chem. Phys.* **1985**, *82*, 890−898.

(25) Kollmar, C.; Neese, F. *Mol. Phys.* **2010**, *108*, 2449−2458.

(26) Kollmar, C. *J. Chem. Phys.* **2006**, *125*, 084108.

(27) DePrince, A. E.; Mazziotti, D. A. *Phys. Rev. A* **2007**, *76*, 042501.

(28) Mazziotti, D. A. *Phys. Rev. A* **2010**, *81*, 062515.

(29) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *WIREs: Comput. Mol. Sci.* **2012**, *2*, 556−565.

(30) DePrince, A. E.; Sherrill, C. D. *J. Chem. Theory Comput.* **2013**, *9*, 2687−2696.

(31) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(32) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *WIREs: Comput. Mol. Sci.* **2012**, *2*, 242−253.

(33) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638−4650.

(34) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 9751−9764.

(35) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437−446.

(36) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(37) Faver, J. C.; Benson, M. L.; He, X.; Roberts, B. P.; Wang, B.; Marshall, M. S.; Kennedy, M. R.; Sherrill, C. D.; Merz, K. M. *J. Chem. Theory Comput.* **2011**, *7*, 790−797.

(38) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.

(39) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151−2155.

(40) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *135*, 194102.

(41) Garrod, C.; Percus, J. K. *J. Math. Phys.* **1964**, *5*, 1756−1776.

(42) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.

(43) Knizia, G.; Adler, T.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.

(44) Marshall, M. S.; Sherrill, C. D. *J. Chem. Theory Comput.* **2011**, *7*, 3978−3982.

(45) Torheyden, M.; Valeev, E. F. *J. Chem. Phys.* **2009**, *131*, 171103.

(46) Kong, L.; Valeev, E. F. *J. Chem. Phys.* **2011**, *135*, 214105.