Particle-particle ladder based basis-set corrections applied to atoms and molecules using coupled-cluster theory

Andreas Irmler and Andreas Grüneis*
Institute for Theoretical Physics, TU Wien,
Wiedner Hauptstraße 8-10/136, 1040 Vienna, Austria
(Dated: March 15, 2019)

We investigate the basis-set convergence of coupled cluster electronic correlation energies using a recently proposed finite basis-set correction technique. The correction is applied to atomic and molecular systems and is based on a diagrammatically decomposed coupled cluster singles and doubles correlation energy. Only the second-order energy and the particle-particle ladder term are corrected for their basis-set incompleteness error. We present absolute correlation energies for the H_2O and F_2 molecules. Furthermore atomization energies for a test set containing 49 molecules are investigated and the performance of the employed basis-set correction technique is compared to explicitly correlated methods. Our findings indicate that it is possible to achieve basis-set reductions that are comparable to state-of-the-art F12 theories. The employed technique can readily be transfered to other many-electron wavefunction methods without the need for three- and four-electron integrals.

Quantum chemical many-electron theories that employ anti-symmetrized one-particle functions (Slater determinants) as a basis for the many-electron wavefunction exhibit a frustratingly slow convergence of central quantities like the ground state energy to the complete basis-set limit. A large fraction of the computational cost involved in many-electron theory calculations of atoms, molecules and solids originates from the need to include large numbers of one-electron basis functions necessary to achieve the desired level of precision. Many techniques have been developed to accelerate the convergence to the complete basis-set limit including explicitly correlated methods, transcorrelated methods or simple yet less efficient basis-set extrapolation techniques [1–7].

It has been known since the early days of electronic structure theory that the slow convergence of the wavefunction expansion in Slater determinants is due to short-ranged interelectronic correlation. As the electrons coalesce, a derivative discontinuity or 'cusp' must arise, so that a divergence in the kinetic energy operator cancels an opposite one in the potential [8–11]. Explicitly correlated methods account for the cusp condition in an *a priori* manner and are commonly referred to as F12 theories, where F12 stands for a two-electron correlation factor that enables a compact expansion of the wavefunction at short interelectronic distances [1, 12, 13]. F12 theories introduce, however, the need for additional many-electron integrals such as three- and sometimes even four-electron integrals. Due to their large computational overhead, explicitly correlated methods are generally only beneficial for more complex parent methods. Furthermore they cause substantial additional effort in their computer code implementation. Nonetheless, thanks to their improved basis-set convergence and reliability, they have become an indispensable tool for quantum chemical calculations of large systems [1, 3, 14–20].

In this work we apply a recently proposed basis-set extrapolation technique for coupled cluster singles and doubles (CCSD) theory to atoms and molecules. The employed technique is based on a diagrammatic decomposition of the coupled cluster correlation energy and was first introduced to study the uniform electron gas [21]. An analysis of the decomposed correlation energy for large basis sets has shown that the slow convergence of the CCSD energy originates from the second-order correlation energy and the so-called particle-particle ladder term. Both contributions have been found to converge at the same rate to the complete basis-set (CBS) limit. Based on this observation, we have devised a basis-set correction that is based on re-scaling the partice-particle ladder (ppl) correlation energy contribution by a factor that is estimated from the corresponding basis-set incompleteness error of the MP2 correlation energy, which can be obtained in a computationally significantly cheaper manner. Previously this scheme has successfully been applied to estimate the CCSD correlation energy of simple semiconducting and insulating solids [21]. Here, we show that this method achieves a convergence rate that is comparable to state-of-the-art explicit correlation methods for total valence electron correlation energies of molecules and atomization energies for a large test set containing 49 molecules.

Theory. – Here, we employ a recently proposed method to correct for the basis-set incompleteness error of CCSD correlation energies. The central premise of this method is that the CCSD correlation energy,

$$E_c^{CCSD} = W_{ij}^{ab} T_{ij}^{ab}, \qquad (1)$$

can be decomposed into different diagrammatic contributions such that [21]

$$E_c^{CCSD} = E^{driver} + E^{ppl} + \underbrace{E^{phl} + E^{hhl} + E^{phr} + ...}_{=E^{rest}}, \quad (2)$$

where E^{driver} corresponds to the MP2 correlation energy

arXiv:1903.05559v2 [physics.chem-ph] 14 Mar 2019

TABLE I. Index notation for different orbital subspaces of the complete one-electron basis. The orbitals refer to the Hartree-Fock spatial orbital basis set over which the wavefunction amplitudes are defined.

	Occ. orbitals	Virt. orbitals	Complement
i, j, k, l, m, n	Yes	No	No
a, b, c, d	No	Yes	No
C	No	Yes	Yes
P, Q	Yes	Yes	Yes

$$E^{\text{driver}} = W_{ij}^{ab} \langle ab|ij \rangle \quad (3)$$

and the particle-particle ladder term is defined as

$$E^{\text{ppl}} = W_{ij}^{ab} \langle ab|cd \rangle T_{ij}^{cd}. \quad (4)$$

T_{ij}^{cd} is computed from the CCSD singles (t_i^a) and doubles (t_{ij}^{ab}) amplitudes such that $T_{ij}^{ab} = t_{ij}^{ab} + t_i^a t_j^b$. t_i^a and t_{ij}^{ab} are obtained by solving the corresponding amplitude equations[22, 23]. W_{ij}^{ab} is given by

$$W_{ij}^{ab} = \frac{2 \langle ij|ab \rangle - \langle ji|ab \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \quad (5)$$

The employed indices are explained and summarized in Table I. Einstein summation convention applies to repeated indices throughout this work.

We note that the employed decomposition of the CCSD correlation energy is achieved by replacing the doubles amplitudes in Eq. (1) with the corresponding right hand side of the amplitude equations using converged CCSD amplitudes. For the sake of brevity, we define E^{rest} such that it contains all remaining contributions to the CCSD correlation energy including terms such as the particle-hole ladder, hole-hole ladder and the particle-hole ring. We stress that the labelling of the contributions to the correlation energy on the right hand side of Eq. (2) is inspired by the corresponding terms in the amplitude equations and does not imply that the ppl term

includes particle-particle ladder contributions to the correlation energy only. The latter holds only in the case of $t_{ij}^{ab} = t^{(1)}{}_{ij}^{ab}$ where $t^{(1)}{}_{ij}^{ab} = \frac{\langle ab|ij \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$. We note that a similar labeling of the terms in the amplitude equations is used by Shepherd et al. in Ref. [24]. We refer the reader to Ref. [21] for a more detailed analysis of the different diagrammatic contributions to the CCSD correlation energy in periodic systems.

The MP2 correlation energy, E^{driver}, converges as $1/L^3$ using a correlation consistent atom-centered Gaussian basis set, where L refers to the cardinal number of the employed basis set. We will argue in the following that the first two terms on the right hand side of Eq. (2) exhibit the same convergence rate when approaching the complete basis set limit, constituting the dominant source of the basis set incompleteness error of the CCSD correlation energy. To better understand the convergence behaviour of E^{ppl}, we approximate the first-order amplitudes using the following expression that follows from F12 theory

$$t^{(1)}{}_{ij}^{ab} \approx \left\langle ab \left| (-1) \frac{t_{ij}}{\gamma} e^{-\gamma r_{12}} \right| ij \right\rangle. \quad (6)$$

t_{ij} are geminal amplitudes determined by the universal cusp conditions [8]. The expression above implies that the first-order doubles amplitudes can be approximated using an expression that is similar to the corresponding electron repulsion integral, where the Coulomb kernel has been replaced by the Slater-type correlation factor that depends on a parameter γ. Approximating the CCSD doubles amplitudes t_{ij}^{ab} on the right hand side of Eq. (4) by Eq. (6) and disregarding the contribution of single amplitudes, we can approximate the particle-particle ladder contribution to the correlation energy by

$$E^{\text{ppl}} \approx W_{ij}^{ab} \langle ab|cd \rangle \left\langle cd \left| (-1) \frac{t_{ij}}{\gamma} e^{-\gamma r_{12}} \right| ij \right\rangle \quad (7)$$

We now replace the summation over the virtual orbital indices c and d in Eq. (7) by the closure relation $|c\rangle\langle c| = |P\rangle\langle P| - |k\rangle\langle k|$, yielding the following approximate expression for E^{ppl}

$$E^{\text{ppl}} \approx W_{ij}^{ab}(-1)\frac{t_{ij}}{\gamma} \left[\left\langle ab \left| \frac{e^{-\gamma r_{12}}}{r_{12}} \right| ij \right\rangle + \underbrace{\langle ab|kl \rangle \langle kl|e^{-\gamma r_{12}}|ij \rangle - \langle ab|Pl \rangle \langle Pl|e^{-\gamma r_{12}}|ij \rangle}_{=-\langle ab|Cl \rangle \langle Cl|e^{-\gamma r_{12}}|ij \rangle} - \langle ab|kQ \rangle \langle kQ|e^{-\gamma r_{12}}|ij \rangle \right] \quad (8)$$

We stress that the particle-particle ladder contribution to the correlation energy is positive (for $t_{ij}^{ab} \approx t^{(1)}{}_{ij}^{ab}$), whereas the last three terms in the brackets on the right hand side of Eq. (8) yield a negative contribution. According to this, the leading contribution to the correla-

tion energy must be the first term in the brackets on the right hand side of the above expression. This term converges at the same rate to the complete basis set limit as E^{driver}. In particular the $\frac{e^{-\gamma r_{12}}}{r_{12}}$ kernel exhibits a singularity at $r_{12} = 0$ that causes an asymptotic basis
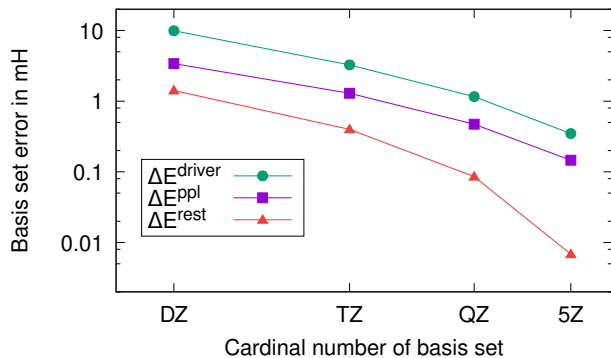
FIG. 1. Basis set error for the different channels in the He atom ($\Delta E^{\mathrm{driver}}(X)$, $\Delta E^{\mathrm{ppl}}(X)$ and $\Delta E^{\mathrm{rest}}(X)$) using the aug-cc-pVXZ basis sets. The reference energy is taken from aug-cc-pV6Z value.

set convergence rate which is identical to MP2 theory $(1/L^3)$. In short, we have shown above that E^{ppl} defined by Eq. (4) will converge to the CBS limit at the same rate as E^{driver} if the following approximations hold $t_{ij}^{ab} \approx t^{(1)}{}_{ij}^{ab} \approx \left\langle ab \left| (-1) \frac{t_{ij}}{\gamma} e^{-\gamma r_{12}} \right| ij \right\rangle$.

We now put the analysis carried out above to a numerical test for the He atom. Figure 1 depicts the convergence of the basis set error of the decomposed correlation energy contributions. The errors are retrieved as a function of the cardinal number of the employed aug-cc-pVXZ basis set. Basis set errors for the corresponding channels are defined by

$$\Delta E(X) = E(\mathrm{CBS}) - \mathrm{E}(X).$$

We refer to energies obtained in a given basis set X by $E^{\mathrm{driver}}(X)$, $E^{\mathrm{ppl}}(X)$ and $E^{\mathrm{rest}}(X)$. For the results shown in Figure 1, we estimate the CBS values using an aug-cc-pV6Z basis. We note that $\Delta E^{\mathrm{driver}}(X)$ and $\Delta E^{\mathrm{ppl}}(X)$ are parallel on a logarithmic scale, whereas $\Delta E^{\mathrm{rest}}(X)$ exhibits a faster convergence rate compared to the other terms. These results illustrate that the analysis for the uniform electron gas outlined in Ref. [21] and the approximation discussed above also hold for the He atom.

Based on the analysis above and in agreement with Ref. [21], it is reasonable to propose a finite basis set error correction to the CCSD correlation energy that requires the CBS limit estimate of the second-order correlation energy $E^{\mathrm{driver}}(\mathrm{CBS})$ only. We define the approximation to the CBS limit estimate of the CCSD correlation energy, CCSD-PPL, as

$$E_{\mathrm{c}}^{\mathrm{CCSD-PPL}} = E_{\mathrm{c}}^{\mathrm{CCSD}}(X) + \Delta E^{\mathrm{driver}}(X) + \Delta E^{\mathrm{ppl}}(X), \tag{9}$$

where we employ the following approximation to the particle-particle ladder correlation energy contribution

$$\Delta E^{\mathrm{ppl}}(X) = \frac{E^{\mathrm{driver}}(\mathrm{CBS})}{E^{\mathrm{driver}}(X)} E^{\mathrm{ppl}}(X) - E^{\mathrm{ppl}}(X) \tag{10}$$

The CBS limit of the second-order correlation energy $E^{\mathrm{driver}}(\mathrm{CBS})$ is approximated using a $[Q5]$ extrapolation throughout this work.

Computational details. – We have modified the coupled cluster code in the open-source quantum chemistry package PSI4 [25] such that the $E^{\mathrm{ppl}}(X)$ contribution to the CCSD correlation energy is computed separately at the end of each CCSD calculation using fully converged CCSD amplitudes. The required modifications to the exisiting CCSD code are minor and we expect that other computer code implementations of coupled cluster theory can also be modified in such a simple manner.

In the present work we have computed atomization energies for a molecular test set and compared to CBS limit estimates from Ref. [3]. The employed equilibrium geometries and reference energies have been taken from Ref. [3].

Results. – We first seek to assess the efficiency of CCSD-PPL for absolute valence electron correlation energies of the closed shell molecules H_2O and F_2. Table II summarizes the calculated valence electron correlation energies for aug-cc-pVXZ basis sets with $X = \mathrm{D, T, Q}$ and results from literature obtained using F12 approaches from Refs. [3, 26] that account explicitly for the cusp condition. The valence electron correlation energies shown in Table II demonstrate that CCSD-PPL converges to the CBS limit with respect to the employed basis set at a similar rate as state-of-the-art F12 theories. Compared to CCSD theory, CCSD-PPL allows for a reduction by approximately two cardinal numbers in the employed basis set size while capturing a similar fraction of the correlation energy.

As a further test, we have employed CCSD-PPL to compute atomization energies of a molecular test set containing 49 molecules. Figure 2 depicts the CBS error in the valence electron correlation energy contributions to the computed CCSD and CCSD-PPL atomization energies using an aug-cc-pVDZ (AVDZ), aug-cc-pVTZ (AVTZ) and aug-cc-pVQZ (AVQZ) basis set. We observe a monotonically decreasing error for increasing basis set size. A corresponding statistical analysis is summarized in Table III. We find that CCSD-PPL using AVDZ achieves on average the same level of precision as CCSD using AVQZ as denoted by the agreement of the mean absolute (MAD) and root mean squared (RMS) deviations. The same holds for the maximum (MAX) deviation to the CBS limit estimates, demonstrating that CCSD-PPL improves the basis set convergence in a robust manner for a wide range of systems. A comparison of CCSD-PPL to CCSD-F12a reveals that on average CCSD-F12a achieves a slightly better agreement with CBS limit values for AVDZ and AVTZ. For the AVQZ basis set, CCSD-PPL is closer to the CBS limit estimates than CCSD-F12a. However, we also note that CCSD-F12b yields atomization energies that are on average in an even better agreement with the CBS limit than CCSD-F12a and CCSD-PPL when employing AVTZ and AVQZ basis sets.

TABLE II. Valence electron correlation energies for H_2O and F_2 in mH. The CCSD CBS limit values are -297.9 mH and -601.17 mH, respectively [3].

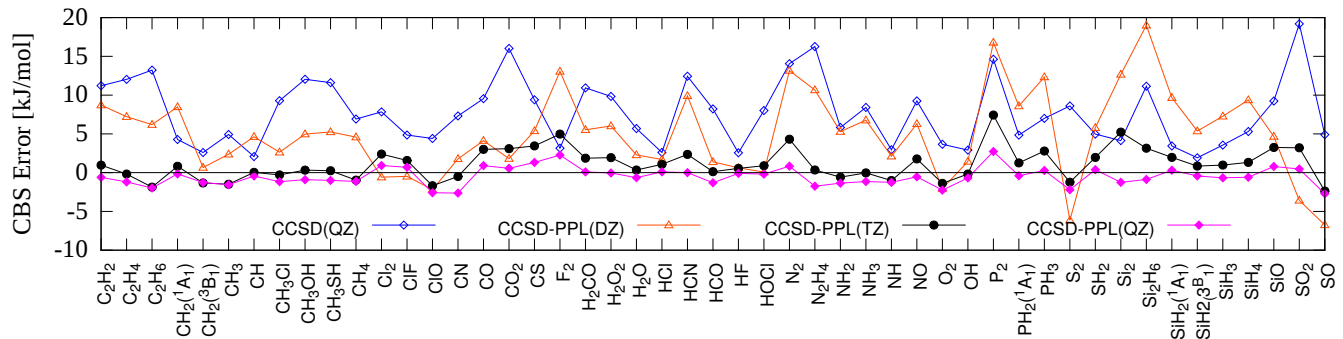|	H_2O			F_2		
	AVDZ	AVTZ	AVQZ	AVDZ	AVTZ	AVQZ
CCSD	-227.11	-273.05	-288.21	-435.39	-538.91	-575.10
CCSD-PPL	-291.44	-298.22	-299.50	-580.92	-598.32	-602.67
CCSD-F12a [3]	-293.22	-298.50	-299.37	-590.81	-599.65	-602.89
CCSD-F12 [26]	-289.86	-295.40	-297.23	-584.83	-594.98	-599.01

FIG. 2. Basis set incompleteness errors of computed valence electron correlation energy contributions to the atomization energies for a set of 49 molecules. The CCSD/CBS limit reference values have been taken from Ref. [3]. Errors are shown for CCSD and CCSD-PPL theory using aug-cc-pVDZ (DZ), aug-cc-pVTZ (TZ) and aug-cc-pVQZ (QZ) basis sets.

TABLE III. Statistical analysis of errors depicted in Fig. 2.

	MAD	RMS	MAX
CCSD/AVQZ	7.740	8.816	19.192
CCSD-PPL/AVDZ	5.886	7.291	19.008
CCSD-PPL/AVTZ	1.735	2.296	7.425
CCSD-PPL/AVQZ	1.015	1.264	2.728
CCSD-F12a/AVDZ[3]	5.972	7.031	18.202
CCSD-F12a/AVTZ[3]	1.534	1.859	4.102
CCSD-F12a/AVQZ[3]	1.910	2.167	4.654
CCSD-F12b/AVDZ[3]	8.390	10.230	27.434
CCSD-F12b/AVTZ[3]	1.717	2.144	6.138
CCSD-F12b/AVQZ[3]	0.554	0.701	1.714

Concluding Remarks. – We have applied a recently introduced finite basis-set correction method for coupled cluster singles and doubles theory (CCSD-PPL) to atoms and molecules. The method is based on a decomposed correlation energy and corrects for the basis-set incompleteness error in the second-order and the particle-particle ladder term of the CCSD energy only. Compared to CCSD, CCSD-PPL allows for a reduction of atom centered correlation consistent Gaussian basis sets by approximately two cardinal numbers, while capturing a similar fraction of the CBS limit correlation energy. We have applied CCSD-PPL to a test set containing 49 molecules, demonstrating a similar level of precision on approach to the CBS limit values for atomization energies as state-of-the-art F12 theories. We stress that it is possible to transfer the outlined method to other widely-used many-electron theories that allow for a decomposition of the electronic correlation energy such as full configuration interaction and related theories. Finally, we note that the employed decomposition of the electronic correlation energy will potentially also be useful for the further development and improvement of correlation factors used in F12 theories.

[1] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[2] A. Grüneis, S. Hirata, Y.-y. Ohnishi, and S. Ten-no, J.

Chem. Phys. **146**, 080901 (2017).

[3] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).

[4] S. F. Boys, N. C. Handy, and J. W. Linnett, Proc. Roy. Soc. London Ser. A. **309**, 209 (1969).

[5] S. F. Boys, N. C. Handy, and J. W. Linnett, Proc. Roy. Soc. London Ser. A. **310**, 43 (1969).

[6] S. Ten-no, Chem. Phys. Lett. **330**, 169 (2000).

[7] J. J. Shepherd, A. Grüneis, G. H. Booth, G. Kresse, and A. Alavi, Phys. Rev. B **86**, 035111 (2012).

[8] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[9] R. T. Pack and W. Byers Brown, J. Chem. Phys. **45**, 556 (1966).

[10] A. Carlsson and N. Ashcroft, Phys. Rev. B **25**, 3474 (1982).

[11] X.-Y. Pan and V. Sahni, J. Chem. Phys. **119**, 7083 (2003).

[12] H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).

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

[14] C. Hättig, D. P. Tew, and A. Köhn, J. Chem. Phys. **132**, 231102 (2010).

[15] L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).

[16] D. P. Tew, C. Hättig, R. A. Bachorz, and W. Klopper, in *Recent Progress in Coupled Cluster Methods – Theory and Applications*, edited by P. Čársky, J. Paldus, and J. Pittner (Springer, Dordrecht, Heidelberg, London,

[17] H.-J. Werner, T. B. Adler, G. Knizia, and F. R. Manby, in *Recent Progress in Coupled Cluster Methods – Theory and Applications*, edited by P. Čársky, J. Paldus, and J. Pittner (Springer, Dordrecht, Heidelberg, London, New York, 2010) pp. 573–620.

[18] S. Ten-no and J. Noga, WIREs Comput. Mol. Sci. **2**, 114 (2012).

[19] S. Ten-no, Theor. Chem. Acc. **131**, 1070 (2012).

[20] T. Shiozaki and S. Hirata, J. Chem. Phys. **132**, 151101 (2010).

[21] A. Irmler, F. Hummel, and A. Grüneis, arXiv:1903.05458 (2019).

[22] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[23] J. Cizek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).

[24] J. J. Shepherd, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **140**, 124102 (2014).

[25] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).

[26] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **129**, 071101 (2008).

New York, 2010) pp. 535–572.