This is a repository copy of *Optimized Basis Sets for the Environment in the Domain-Specific Basis Set Approach of the Incremental Scheme*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/97102/

Version: Accepted Version

Article:

Anacker, T., Hill, J. orcid.org/0000-0002-6457-5837 and Friedrich, J. (2016) Optimized Basis Sets for the Environment in the Domain-Specific Basis Set Approach of the Incremental Scheme. Journal of Physical Chemistry A, 120 (15). pp. 2443-2458. ISSN 1089-5639

https://doi.org/10.1021/acs.jpca.6b01097

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Optimized Basis Sets for the Environment in the Domain-Specific Basis Set Approach of the Incremental Scheme

Tony Anacker,[†] J. Grant Hill,[‡] and Joachim Friedrich[*,†]

[†]*Department of Theoretical Chemistry, Chemnitz University of Technology, Straße der Nationen 62, D-09111 Chemnitz, Germany*

[‡]*Department of Chemistry, University of Sheffield, Sheffield S3 7HF, UK*

E-mail: joachim.friedrich@chemie.tu-chemnitz.de

Phone: +49 (0)371 53139415

Abstract

Minimal basis sets, denoted DSBSenv, have been developed based on the segmented basis sets of Ahlrichs and co-workers for use as environmental basis set for the domain-specific basis set incremental scheme with the aim of decreasing the CPU requirements of the incremental scheme. The use of this minimal basis within explicitly correlated (F12) methods has been enabled by the optimization of matching auxiliary basis sets for use in density fitting of two-electron integrals and the resolution-of-the-identity. The accuracy of these auxiliary sets has been validated by calculations on a test set containing small- to medium-sized molecules. The errors due to density fitting are about two to four orders of magnitude smaller than the basis set incompleteness error of the DSBSenv orbital basis sets. Additional reductions in computational cost are tested with the reduced DSBSenv basis sets, where the highest angular momentum functions of the DSBSenv auxiliary basis sets have been removed. The optimized and reduced basis sets are used in the framework of the domain-specific basis set of the incremental scheme to decrease the computation time without significant loss of accuracy. The computation times and accuracy of the previously used environmental basis and that optimized in this work is validated with a test set of medium- to large-sized systems. The optimized and reduced DSBSenv basis sets decrease the CPU-time by about 15.4 % and 19.4 % compared to the old environmental basis and retains the accuracy in the absolute energy with a standard deviation of 0.99 and 1.06 kJ/mol, respectively.

Introduction

The application of highly accurate wave function based methods to large systems is a challenging research field in the last few years. A primary goal is to have an alternative method to density functional theory, which is systematically improvable and can be applied for even larger systems. Coupled-cluster with single, double and perturbative triple excitations [CCSD(T)] with an appropriate basis set is currently the "gold standard" of quantum chemistry.[1–4] Due to its unfavorable scaling with respect to system size, different strategies are frequently used to make CCSD(T) applicable. The first deals with the slow convergence of the coupled-cluster energies with respect to the one-particle basis set. To overcome this drawback complete basis set extrapolations,[5–7] interference-corrected coupled-cluster[8–11] or explicitly correlated methods[6,12–19] are frequently used strategies. The convergence with respect to the complete basis set (CBS) limit is improved and therefore smaller basis sets can be used to achieve accurate energies with affordable computational cost.

Another strategy to reduce the computational effort is to introduce local approximations in the wave function with the goal of reducing the scaling with system size. With the fundamental work of Pulay and Sæbø[20–24] many groups have developed local methods including Werner's local coupled-cluster methods,[17,18,25–33] Neese's pair natural orbitals (PNOs) coupled-cluster[34–44] and Yang's orbital specific virtuals (OSVs) coupled-cluster.[33,45–47] The incremental scheme proposed by Stoll[48–55] is another local method and is frequently used by Dolg,[56–60] Friedrich[61–66] and Paulus.[54,67–72] A fully automated implementation of this scheme for MP2 and CCSD(T) and the corresponding explicitly correlated methods were developed by Friedrich[73–80] and Dolg.[56,57] The incremental scheme very efficiently delivers highly accurate energies[81,82] and properties[66,83,84] for closed- and open-shell systems.[85,86]

In this work, the computational speed performance of the automated incremental scheme of Friedrich shall be improved. The main approximation for improving the computation times is the application of the domain-specific basis set. With this approach the virtual space of a domain is reduced in order to save computational time and is discussed in detail below. In previous work, the SV orbital basis set (OBS) of Ahlrichs[87] was used in conjunction with the SVP auxiliary basis sets (ABS)[88] in both the density fitting of two electron integrals in density-fitted MP2 (DF-MP2) and as the complementary basis for explicitly correlated methods (F12) within the complementary auxiliary basis set (CABS)[89] approach to the resolution-of-identity (RI) approximation of many electron integrals. This combination was used as an *ad hoc* guess to describe the environment of a domain in the domain-specific basis set. Therefore, the aim is to construct a minimal OBS along with specifically matched

ABS to improve the computational speed performance of the incremental scheme. Another main goal is to achieve negligible loss of accuracy in using the newly developed basis sets. The DSBSenv basis sets are optimized for the first three periods of the periodic table and these basis sets are then used in the incremental scheme with the domain-specific basis set. Moreover, the reduced DSBSenv basis sets remove the highest angular momentum functions in the ABS to further improve computational speed performance. We compare the old environmental SV/SVP basis sets with the newly optimized and reduced DSBSenv sets in respect to the computation times of the incremental scheme as well as the accuracy in the absolute correlation energy and relative energies.

Incremental Scheme

For an incremental calculation of the correlation energy, the system is divided into one-site domains consisting of disjoint sets of localized occupied orbitals.[61,62] The correlation energy of the total system is computed through a truncated many-body expansion, summing the correlation energy of these one-site domains and energy corrections for pairs, triples, etc. of these domains that approximately account for the non-additivity of the correlation energy. The incremental expansion of the correlation energy reads:[50,61]

$$E_{\mathrm{corr}} = \sum_{\substack{\mathbb{X} \\ \mathbb{X}\in\mathcal{P}(\mathbb{D})\wedge|\mathbb{X}|\leq n}} \Delta\varepsilon_{\mathbb{X}} \tag{1}$$

where $\mathcal{P}(\mathbb{D})$ is the power set of the set of domains $\mathbb{D}$. The restriction to the cardinality of the sets $\mathbb{X}$ truncates the incremental series at the desired order n and the increment $\Delta\varepsilon_{\mathbb{X}}$ is defined as

$$\Delta\varepsilon_{\mathbb{X}} = \varepsilon_{\mathbb{X}} - \sum_{\substack{\mathbb{Y} \\ \mathbb{Y}\in\mathcal{P}(\mathbb{X})\wedge|\mathbb{Y}|<|\mathbb{X}|}} \Delta\varepsilon_{\mathbb{Y}}. \tag{2}$$

$\varepsilon_{\mathbb{X}}$ is the correlation energy of the domain $\mathbb{X}$. Each domain is a set of localized occupied orbitals (LMOs), grouped using METIS graph partitioning[90] on the basis of charge-center distance criteria, and a subsequent pairwise refinement.[79]

The Domain-Specific Basis Set Approximation

To further improve the computational speed performance of the incremental scheme, for each incremental domain calculation the virtual space is reduced by using a domain-specific basis set, where the high quality basis set is used only at the atoms close to the charge centers of the domain and a small basis set is used at the remaining atoms.[63,73,76,91]

In this approach, we construct the domains from Boys-localized molecular orbitals obtained from a SCF calculation on the whole molecule in a small basis, which we are free to choose (we use the SV and the DSBSenv OBS developed in this work). The incremental correlation energy of a domain is computed after a Hartree-Fock (HF) calculation in the composed domain-specific basis set, and a unique mapping of the localized molecular orbitals in the domain-specific basis set to the small basis calculation. The computational time for the additional HF calculation is negligible compared to the CCSD(T) calculation for the domain. The environment of a domain is included in an incremental calculation to avoid artificial errors in the wave function. The reduction of the virtual space using a small environmental basis introduces errors in the incremental correlation energies, but we can effectively correct this using the incremental MP2 error. With the newly optimized DSBSenv basis sets, we use a minimal basis for the environment and want to investigate the influence of a larger virtual space reduction on the accuracy and computation time of the incremental scheme. In addition to improving computational speed performance, smaller environmental basis sets enable the use of high quality basis sets for the correlation treatment.

MP2 Error Correction

The domain-specific basis set approach introduces an error in the CCSD(T)(F12*)[14,15,92] energy. The leading contributions to these errors can be subtracted using the relatively cheap MP2 method, computed in the target basis

$$\begin{aligned} & E_{\text{corr}}^{\text{inc}}(\text{CCSD(T)(F12*)}|\text{MP2-F12}) \\ &\quad = E_{\text{corr}}^{\text{inc}}(\text{CCSD(T)(F12*)}) - \big(E_{\text{corr}}^{\text{inc}}(\text{MP2-F12}) - E_{\text{corr}}(\text{MP2-F12})\big) \end{aligned} \tag{3}$$

The MP2 error correction allows us to truncate the incremental expansion at second order for water clusters and at third order for organic systems, for both open-shell and closed-shell systems.[79,80,82,93]

Basis Set Development

In the following, we present the development of the DSBSenv OBS and ABS basis sets for the atoms of the first three periods of the periodic table, i.e. period 1 (H and He), period 2 (Li–Ne) and period 3 (Na–Ar).

The DSBSenv Orbital Basis Sets

The DSBSenv OBS was constructed by recontracting the def-SV basis set primitives of Ahlrichs and co-workers[87,94] to form a minimal basis with a segmented contraction scheme. For this purpose, we used the atomic orbital (AO) coefficients from symmetry adapted HF calculations on atomic ground-states carried out using the MCSCF program in MOLPRO.[95,96] For the H and He atoms, the $4s$ functions are contracted to $1s$ function, for the atoms Li–Ne $7s4p$ to $2s1p$, and for the atoms Na–Ar $10s7p$ to $3s2p$. The final composition of the basis sets are shown in Table 1, along with the composition of the matched auxiliary basis sets. The DSBSenv basis sets (including ABSs) can be found in the Supporting Information (SI).

Table 1: Composition of the DSBSenv orbital and the auxiliary MP2Fit and CABS basis sets for the elements H–Ar.

Element	OBS			MP2Fit			CABS
H,He	$(4s)$	$\rightarrow$	$[1s]$	$(4s3p)$	$\rightarrow$	$[3s2p]$	$(4s4p2d)$
Li,Be	$(7s4p)$	$\rightarrow$	$[2s1p]$	$(9s5p3d)$	$\rightarrow$	$[5s5p1d]$	$(6s6p5d3f)$
B–Ne	$(7s4p)$	$\rightarrow$	$[2s1p]$	$(8s6p5d)$	$\rightarrow$	$[5s5p4d]$	$(6s6p5d3f)$
Na,Mg	$(10s6p)$	$\rightarrow$	$[3s2p]$	$(10s8p5d)$	$\rightarrow$	$[5s5p2d]$	$(6s6p5d3f)$
Al–Ar	$(10s7p)$	$\rightarrow$	$[3s2p]$	$(10s7p7d)$	$\rightarrow$	$[5s5p5d]$	$(6s6p5d3f)$

The DSBSenv/MP2Fit Auxiliary Basis Sets

Auxiliary basis sets for use in density fitting of two-electron integrals (suffixed MP2Fit) matched to the DSBSenv basis have been optimized using a methodology based on that of Hättig and Weigend et al.[88,97] The functional δ_{DF} is minimized for neutral ground-state atoms using the analytical ABS gradients provided by the `ricc2` module[88,97–99] of TURBOMOLE:[100]

$$\delta_{\mathrm{DF}} = \frac{1}{4} \sum_{aibj} \frac{\left(\langle ab||ij\rangle_{\mathrm{DF}} - \langle ab||ij\rangle\right)^2}{\epsilon_a - \epsilon_i + \epsilon_b - \epsilon_j} \tag{4}$$

where the antisymmetrized two electron integrals are defined as $\langle ab||ij\rangle = (ai|bj) - (aj|bi)$. The integrals are in chemists notation, with i, j denoting occupied orbitals, a, b virtual orbitals and ϵ_x the HF orbital energies. Full details of the optimization can be found in the SI, but briefly the ABS uses the exponents of the def-SV(P)/MP2Fit ABS[88,97] primitives with a reoptimization of the contraction coefficients to produce a new, smaller contraction pattern.

The accuracy of the optimized MP2Fit auxiliary basis sets has been investigated at the molecular level with a test set of 104 small- to medium-sized systems derived from Kritikou's extension of Weigend's test set.[101,102] The noble gas atoms have been removed from the test set as the minimal basis means they have no virtual orbitals to allow a correlation treatment. The 101 remaining systems of the test set do possess virtual orbitals due to the linear combination of AOs (LCAO). The statistics of the ΔDF error, defined as the absolute error in the MP2 correlation energy, for the test set are shown in Table 2, along with an estimate of the conventional MP2/DSBSenv BSIE. The latter is assessed by using MP2-F12/cc-pVQZ-F12 correlation energies as a simple estimate of the complete basis set limit. These F12 calculations used the (R)MP2-F12 method[103,104] with the 3C(FIX) *Ansatz*,[105] a geminal Slater exponent of 1.0 and correlation consistent basis sets optimized specifically for use with F12 methods.[106,107] It can be seen from Table 2 that the density fitting error is three to four orders of magnitude smaller than the conventional BSIE, indicating that the errors introduced in the density fitting are negligible. As it is envisaged that the DSBSenv basis will be used in explicitly correlated calculations, Table 2 also contains an estimate of the MP2-F12/DSBSenv BSIE, and it is evident that the BSIE is reduced by a factor of roughly 2–5 through the use of explicit correlation. A comparison of the MP2-F12 BSIE and ΔDF shows that the density fitting error is roughly three orders of magnitude smaller than this measure of BSIE and hence the MP2Fit ABS developed in this work should also be appropriate for density fitting in the explicitly correlated context.

Table 2: DSBSenv frozen-core correlation energy errors (μE$_h$, per correlated electron) relative to the MP2-F12/cc-pVQZ-F12 energy for the test set of 101 molecules. ΔDF evaluates the MP2 density fitting error using the optimized MP2Fit ABS and ΔRI gives the MP2-F12 error of the optimized OptRI ABS relative to the large even-tempered sets.

Error Type	μ	σ	MAX
BSIE (MP2)	27309.75	8619.67	40179.23
BSIE (MP2-F12)	6041.94	2725.67	25659.05
ΔDF (MP2Fit)	7.08	9.35	50.28
ΔRI (OptRI)	14.24	14.57	60.57

The DSBSenv/OptRI Complementary Auxiliary Basis Sets

RMP2-F12[103,104] calculations with the diagonal 3C(D) *Ansatz* and a geminal slater exponent of 1.4 were used to optimize the exponents of the ABS to be used in the CABS procedure.

For this purpose, either the Broyden-Fletcher-Goldfarb-Shanno (BFGS)[108] or Nelder-Mead simplex algorithm[109] were used. The optimizations were performed with a locally modified version of MOLPRO and use the diagonal elements of the intermediate **V** and **B** matrices, which appear in F12 theory.[103] The resulting ABSs follow the OptRI design philosophy of Peterson and co-workers, with the exponents optimized by minimizing the objective function δ_{RI},[110] where the superscript ref indicates a large, almost complete, RI basis.

$$\delta_{\mathrm{RI}} = \sum_{ij} \frac{\left(V_{ij,ij}^{\mathrm{RI}} - V_{ij,ij}^{\mathrm{RI_{ref}}}\right)^2}{\left|V_{ij,ij}^{\mathrm{RI_{ref}}}\right|} + \frac{\left(B_{ij,ij}^{\mathrm{RI}} - B_{ij,ij}^{\mathrm{RI_{ref}}}\right)^2}{\left|B_{ij,ij}^{\mathrm{RI_{ref}}}\right|} \tag{5}$$

δ_{RI} has units of energy and is always positive. The value ΔRI is also introduced to indicate the energetic error due to using an incomplete ABS in the RI.

$$\Delta\mathrm{RI} = \left|E_{\mathrm{corr}}^{\mathrm{OptRI}} - E_{\mathrm{corr}}^{\mathrm{ref}}\right| \tag{6}$$

Large even-tempered basis sets of Hill[107,111] were used as reference RI ABS. These uncontracted sets have $21s18p14d12f10g8h6i$ functions for H and He, $28s26p22d22f20g18h15i$ functions for Li–Ne, and $29s27p23d20f18g17h15i$ functions for Na–Ar. Additionally the aug-cc-pV5Z/MP2Fit[112] and cc-pV5Z/JKFit[101,113] ABS were used in the density fitting of the two-electron integrals, and the Fock and exchange matrices. Additional optimization details are included in the SI.

Table 3: RI errors in the frozen-core correlation energy (per correlated electron, relative to a large even-tempered reference set) at the MP2-F12 level for the atoms H–Ar. ΔRI illustrates the error in the absolute energy.

	ΔRI [μE_h]			δ_{RI} [nE_h]		
Element	μ	σ	MAX	μ	σ	MAX
H_2, He	1.67	1.92	3.03	0.33	0.03	0.35
Li–Ne	9.64	7.87	20.58	1.77	2.24	7.17
Na–Ar	9.55	8.99	22.03	6.16	4.15	10.69
H–Ar	8.72	8.10	22.03	3.56	3.88	10.69

The statistics of the δ_{RI} and ΔRI atomic errors are shown in Table 3, where it can be seen that the average energetic error per correlated electron is less then 10 μE_h. The period two and three elements have the same composition for the OptRI basis and therefore, the small increase in the statistical errors in the optimization criteria δ_{RI} are not entirely surprising. It is noted in passing that the DSBSenv/OptRI ABSs are almost equivalent in size to cc-pVDZ-F12/OptRI (the latter have fewer *spd* functions, but require higher angular momentum *g*-type functions). This is presumably due to the unified OBS and ABS needing to span a specific space in order to accurately reproduce the RI integrals, and in the present case the relatively large CABS compensates for the minimal nature of the OBS. The optimized OptRI ABS are also analyzed in terms of molecular error using the same test set as above for density fitting errors, with results displayed in Table 2. The error related to using the OptRI ABSs is once again insignificant compared to the MP2-F12 BSIE, as it is two to three orders of magnitude smaller. The reader is reminded at this stage that the main goal of this basis set development is to provide an efficient environmental basis set for the domain-specific basis set approach in order to increase the computational speed performance of the incremental scheme.

The reduced DSBSenv Auxiliary Basis Sets

The domain-specific basis set approach is the most important approximation within our incremental scheme to decrease the computation time and enable local correlation with large basis sets. Therefore, the use of minimal basis sets for the environment is preferable, but one has to take account of the errors introduced. The reduction of the virtual space using the DSBSenv OBS introduces a larger error than the density fitting. The DSBSenv minimal OBS is used for the environment of a domain, which is an insufficient description of the environment with regard to the large target basis used for the main part of a domain. This unbalanced basis leads to HF orbitals different to those of the target basis. Thus, the incremental scheme in combination with the domain-specific basis set approach has different HF orbital spaces for every domain calculation.[63] We can assess this error with the incremental MP2 error, as shown in previous studies.[79] It is expected that the density fitting in the environment has a negligible error compared to the use of a minimal basis (see Tab. 7). Therefore, we try to improve the computational speed performance of the domain-specific basis incremental scheme by removing the highest angular momentum functions from the MP2Fit and OptRI ABSs matched to the DSBSenv basis set. All other exponents and contraction coefficients remain fixed at their optimized values. These ABSs are referred to as rDSBSenv, which are used in conjunction with the DSBSenv OBS. The compositions of the MP2Fit and OptRI rDSBSenv ABSs are shown in Table 4.

Table 4: Composition of the auxiliary reduced DSBSenv (rDSBSenv) MP2Fit and CABS basis sets for the elements H–Ar. The rDSBSenv ABSs use the DSBSenv OBSs.

Element	MP2Fit			CABS
H,He	$(4s)$	$\rightarrow$	$[3s]$	$(4s4p)$
Li,Be	$(9s5p)$	$\rightarrow$	$[5s5p]$	$(6s6p5d)$
B–Ne	$(8s6p)$	$\rightarrow$	$[5s5p]$	$(6s6p5d)$
Na,Mg	$(10s8p)$	$\rightarrow$	$[5s5p]$	$(6s6p5d)$
Al–Ar	$(10s7p)$	$\rightarrow$	$[5s5p]$	$(6s6p5d)$

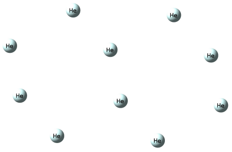
Application in the Incremental Scheme

The TURBOMOLE 6.6 program package[100] was used for all reference and incremental calculations to investigate the computation time and accuracy of the above optimized DSBSenv basis sets. For the density-fitted explicitly correlated calculations MP2-F12[114] and CCSD(T)(F12*),[14,15,92] the optimized cc-pVDZ-F12 basis sets[106,107] were used. All calculations used the frozen core approximation. Therefore, the $1s$ electrons of the Be–Al atoms, and the $1s2s2p$ electrons of the Si–Ar atoms are excluded from the correlation treatment. Third-order incremental MP2-F12 [i3MP2-F12], CCSD(T)(F12*) [i3CC(F12*)] and MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2][79,80,82,93] calculations were performed with our fully automated implementation.[73–80] For the small basis in the domain-specific basis approach,[91] we employed either the SV basis and the SV(P) auxiliary basis sets or the newly optimized DSBSenv basis sets for all atoms. The distance threshold used to specify the atoms belonging to a domain was $t_{\text{main}} = 3.0\,a_0$, as recommended in previous work.[80,93] The calculations employed an order-dependent distance threshold $t_{\text{dist}}(i)$, set using the truncation parameter $f = 30\,a_0$ as recommended previously.[79]

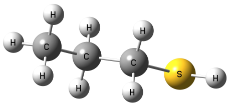
$$t_{\text{dist}}(i) = \frac{f}{(i-1)^2} \tag{7}$$

All incremental computations were performed on a cluster of 71 nodes connected by standard gigabit Ethernet. Each node is equipped with an Intel Xeon E3-1270 3.4 GHz Quad-core CPU, 8 GB RAM, and a single hard disk of 1 TB.

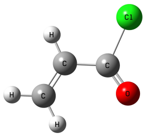
For the following investigations a test set was optimized which contains all elements of the first, second and third periods of the periodic table. The structures were optimized with the BP86/def2-TZVP[88,94] method or were taken from the literature.[79,115] The noble gas structures were optimized with a dispersion correction[116] (BP86-D3/def2-TZVPP). The test set contains 37 systems and is divided into two subsets. Subset I consists of neutral and ionic cluster structures (10 systems), for which the incremental scheme has an outstanding accuracy with the old environmental basis sets. For the noble gas structures, i.e. He_{10}, Ne_{10}, and Ar_{14}, we chose every atom as a domain. Subset II exists of diverse covalent bound systems (27 systems) with aromatic π systems as critical tests for the incremental scheme. Additionally, the test set can be divided into a subset containing only elements from periods one and two (period 2 subset) and a subset containing elements from periods one, two and three (period 3 subset). The structures investigated are displayed in Figure 1.



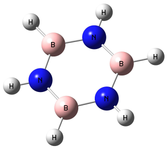
I, period 2



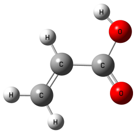
II, period 3



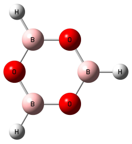
II, period 3



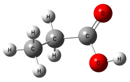
II, period 2



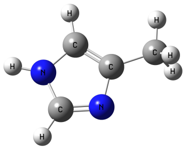
II, period 2



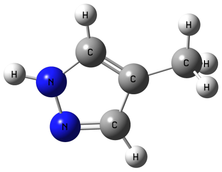
II, period 2



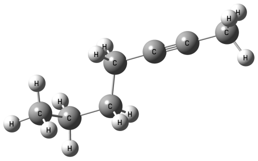
II, period 2



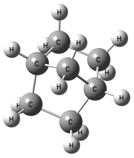
II, period 2



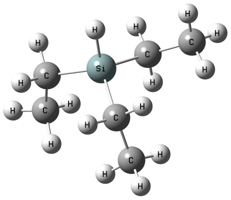
II, period 2



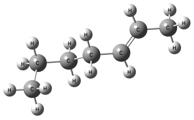
II, period 2



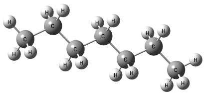
II, period 2



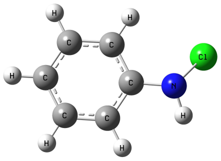
II, period 3



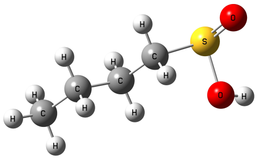
II, period 2



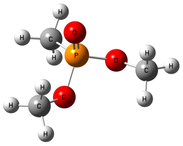
II, period 2



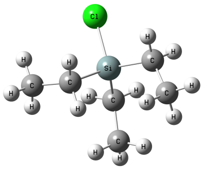
II, period 3



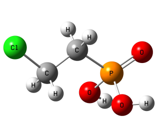
II, period 3



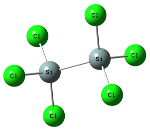
II, period 3



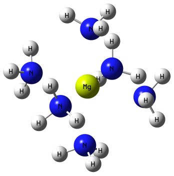
II, period 3



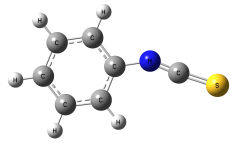
II, period 3



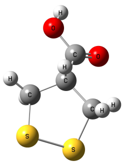
II, period 3



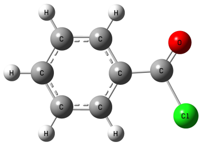
I, period 3



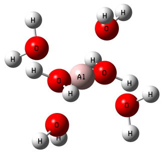
II, period 3



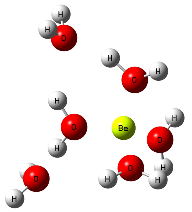
II, period 3



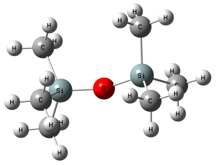
II, period 3



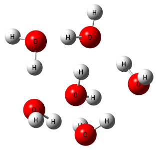
I, period 3



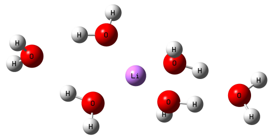
I, period 2



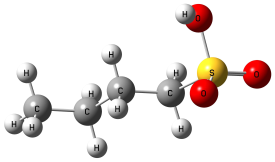
II, period 3



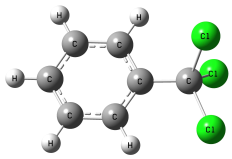
I, period 2



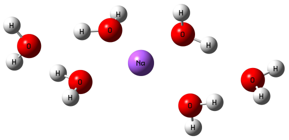
I, period 2



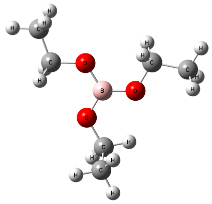
II, period 3



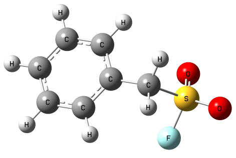
II, period 3



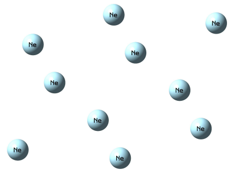
I, period 3



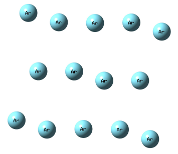
II, period 2



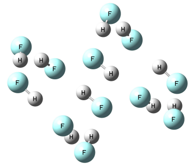
II, period 3



I, period 2



I, period 3



I, period 2

Figure 1: Structures of the systems in the test set. The structures are ordered according to increasing correlation energy. Additionally their sub groups are mentioned: neutral and ionic clusters (I), organic and organic aromatic systems (II), containing elements only from periods one and two (period 2), and containing elements from periods one, two and three (period 3).

Computation Time

For all incremental calculations, the thresholds given above are used. The reference calculations were done on one of the above mentioned nodes. Only the calculations for the $(HF)_{12}$ and Ar_{14} structures needed to be done on a larger node with 12 cores and 64 GB RAM. All incremental calculations were done with one core for the incremental server and 10 clients with 4 cores to carry out the calculations. This ensures a comparable communication between the server and the clients. The advantage of our incremental scheme is the high parallelization of the clients. The clients can use the parallelization of the TURBOMOLE suite to increase the speed performance and additionally one can use as many clients as calculations needed to be done to decrease the wall-time, i.e. the wall-time is as long as the longest incremental calculation. In this work, we concentrate on the better comparable CPU-time, which delivers us the real time savings in the incremental calculations.

In the framework of the domain-specific basis set, the small environmental basis set is used for an automatic generation of the one-site domains. The orbitals of the domain in the domain-specific basis set are identified with the charge centers of the localized molecular orbitals (LMOs). For this purpose, we use the Boys localization in combination with our template localization for a unique mapping.[91] The second orthogonal transformation is necessary, since the Boys localization has more than one maximum, e.g. for aromatic π systems. Due to the re-contraction of the basis functions of the new DSBSenv OBS, the LMOs can be different to the ones of the SV OBS, which can lead to a different number of one-site domains as well as different one-site domains. In a few cases, i.e. the 4-methyl-pyrazol and 5-methyl-imidazol, the DSBSenv OBS led to more domains than the SV OBS using the default parameters. In such cases, we defined the domains to be equal by hand, in order to retain a one to one comparability.

The test set contains structures with different atoms and spatial distribution. This introduces many parameters which makes a systematically analysis very difficult. Therefore, we give exemplary structures with the lowest and the highest time savings when the

Table 5: CPU-times for the third-order incremental MP2-F12 corrected CCSD(T)(F12*) energies. The structures with the lowest (CH_2=CHCOCl) and the highest ($HSiEt_3$) time savings, and the largest structure ($(HF)_{12}$) of the test set are listed as well as the number of domains (#dom). The percentage time savings reflect the time savings using the DSBSenv or rDSBSenv basis sets instead of the SV basis sets. In the last line the arithmetic mean μ of the time savings for the whole test set are added. In general, the time savings increase with the system size.

		CPU-time [h]				time savings [%]	
	#dom.	ref.	SV	DSBSenv	rDSBSenv	DSBSenv	rDSBSenv
CH_2=CHCOCl	4	1.1	2.3	2.2	2.2	2.5	2.1
$HSiEt_3$	6	27.7	67.0	41.2	46.6	38.5	30.5
$(HF)_{12}$	12	537.3[a]	93.4	58.4	44.6	37.0	52.2
mean savings						15.4	19.4

[a] The canonical reference calculation was performed on a node with 12 cores and 64 GB RAM, since it was not computable on the other nodes.

smaller DSBSenv basis sets are used instead of the old SV/SV(P) basis sets in Table 5. The CH_2=CHCOCl structure is partitioned into 4 domains for the incremental correlation treatment. For such small-sized molecules, we do not expect large time savings (2.5 or 2.1 %). Also the smaller rDSBSenv ABSs have nearly no influence on the computation time. The difference of 0.4 % is due to the implementation of the incremental scheme. There are different procedures which needs to fulfill convergence criteria, e.g. the template localization. The template localization has no parallel implementation and is done for every domain calculation. This can lead to different number of cycles until a given threshold is reached. The $HSiEt_3$ structure is divided into 6 domains and has the largest percentage time savings with 38.5 % for the full calculation. As discussed before, we used the incremental scheme with the default parameters as black box method. In this case, the screening of small increments with the order-dependent distance threshold[79] neglects more incremental calculations for the old SV environmental basis sets than the new DSBSenv ones, but the latter still has a much better speed performance. Without screening of small increments, the $HSiEt_3$ structure is calculated six first-order, 15 second-order and 20 third-order incremental domains. Using screening, three and one third-order domains are screened for the SV and DSBSenv basis sets, respectively. It should be mentioned that the third-order domains have the highest computational demand, as they are merged from the first-order domains. This illustrates that a straightforward discussion of the computation time improvement is quite difficult, since we do not want to make further assumptions for the incremental scheme. The $(HF)_{12}$ structure has the largest correlation energy of the test set. The screening of small increments neglects the same calculations for both environmental OBS. In this case we save 35 h and 48.8 h CPU-time for the DSBSenv and rDSBSenv basis sets, respectively. Moreover, the canonical calculation needed to be done on a larger node with 12 cores and 64 GB RAM. The larger amount of memory should make the calculation additionally much faster. Comparing the CPU-times, we can see that both incremental calculations are much faster than the reference calculation on nodes with less memory. In summary, the newly optimized DSBSenv OBSs and ABSs as well as the rDSBSenv ABSs improve the speed performance in all test cases without significant loss of accuracy (see discussion below). The He_{10} system is the only exception, but due to the fast canonical CCSD(T)(F12*)/cc-pVDZ-F12 calculation (264 seconds CPU-time), the application of the incremental scheme to this system is not required using the cc-pVDZ-F12 basis. The application will be more interesting for basis sets with higher ζ level, i.e. cc-pVXZ-F12 with X = T, Q. For the investigated test set, we save an average of 15.4 % or 19.4 % of the incremental calculation time when using the DSBSenv or rDSBSenv basis sets instead of the SV OBSs combined with SVP ABSs.

The analysis focuses on the computation time using the cc-pVDZ-F12 basis set as target basis. We expect the time savings for this basis to be very small, since the canonical implementation is quite fast. Nevertheless, to obtain energies close to the CCSD(T) CBS limit, one has to use the cc-pVTZ-F12 or cc-pVQZ-F12 basis sets in combination with the CCSD(T)(F12*) method. These calculations are much more time consuming and not feasible for large systems. We use the ${Na(H_2O)_6}^+$ structure to illustrate the increasing CPU- and wall-times for the CCSD(T)(F12*) method using the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets.

Table 6: Computation times for the $Na(H_2O)_6{}^+$ structure with different methods using the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets. The use of the optimized and reduced DSBSenv basis sets improve the speed performance of the incremental scheme with respect to the SV basis sets. Using the incremental scheme for calculations with larger basis sets (higher ζ level in the correlation consistent basis sets), the speed performance increases for the new DSBSenv basis sets and the calculations are much faster than the canonical ones.

	cc-pVDZ-F12			cc-pVTZ-F12		
Method	CPU-time	wall-time	#proc	CPU-time	wall-time	#proc
RHF	0.1 h	0.5 min	12	0.8 h	0.1 h	12
DF-MP2-F12	4 h	19.9 min	12	37.2 h	3.1 h	12
DF-CCSD(T)(F12*)	46.6 h	233.2 min	12	404.1 h	33.7 h	12
i3CC(F12*)|MP2/SV	10.1 h	18.4 min	41	37.4 h	0.6 h	161
i3CC(F12*)|MP2/DSBSenv	9.7 h	16.4 min	41	31.6 h	0.5 h	161
i3CC(F12*)|MP2/rDSBSenv	7.6 h	13.8 min	41	28.9 h	0.5 h	161

For the $Na(H_2O)_6{}^+$ structure, the CPU-time increase from 46.6 h to 404.1 h (about a factor of 9) for the CCSD(T)(F12*) method, when the cc-pVTZ-F12 basis is used instead of cc-pVDZ-F12 (see Table 6). In contrast to this, the CPU-time of the incremental methods increase about a factor of 3 to 4. This illustrates that the incremental scheme makes the CCSD(T)(F12*) method applicable to larger basis sets and achieve a higher computational speed performance. Additionally, the new environmental DSBSenv basis sets lead to a larger decrease of the CPU-times in the cc-pVTZ-F12 basis. Moreover, using the reduced rDSBSenv ABSs saves 3 h of CPU-time compared to the optimized ones. The incremental calculations using the cc-pVTZ-F12 as target basis were carried out on 41 nodes (1 for the server and 40 for the clients), so that every domain calculation is performed in parallel. This massive parallelization of the calculations reduces the wall-time from 33.7 h for the canonical calculation to 28 min using the reduced DSBSenv basis sets.

Evaluation of the Accuracy in the Absolute Energy

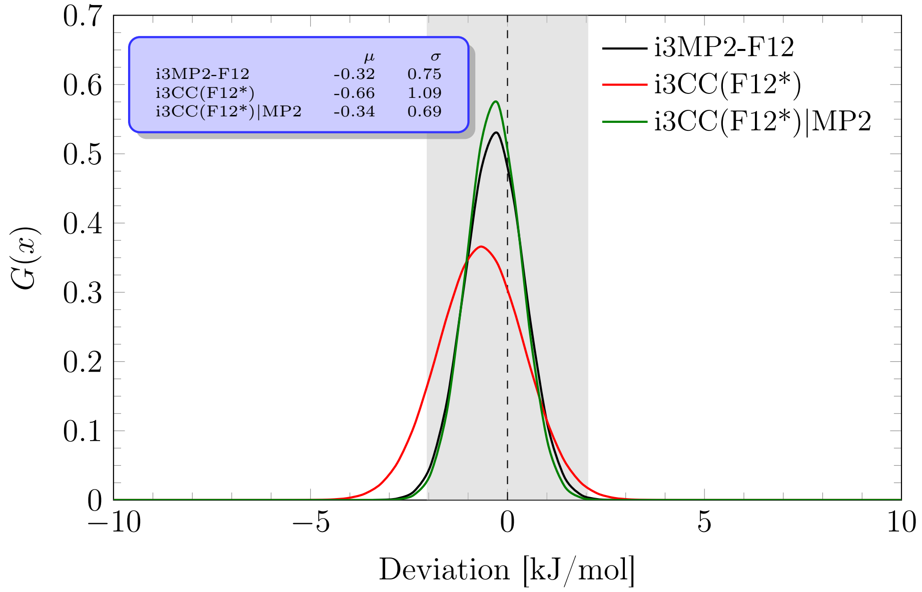
In the following section the influence of the DSBSenv basis sets as environmental basis in the framework of the domain-specific basis set is analyzed within the errors in the absolute and reaction energies.

Table 7: **Errors in the absolute correlation energy [kJ/mol] for the third-order incremental MP2-F12 [i3MP2-F12], CCSD(T)(F12*) [i3CC(F12*)] and MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] methods for the test set.**

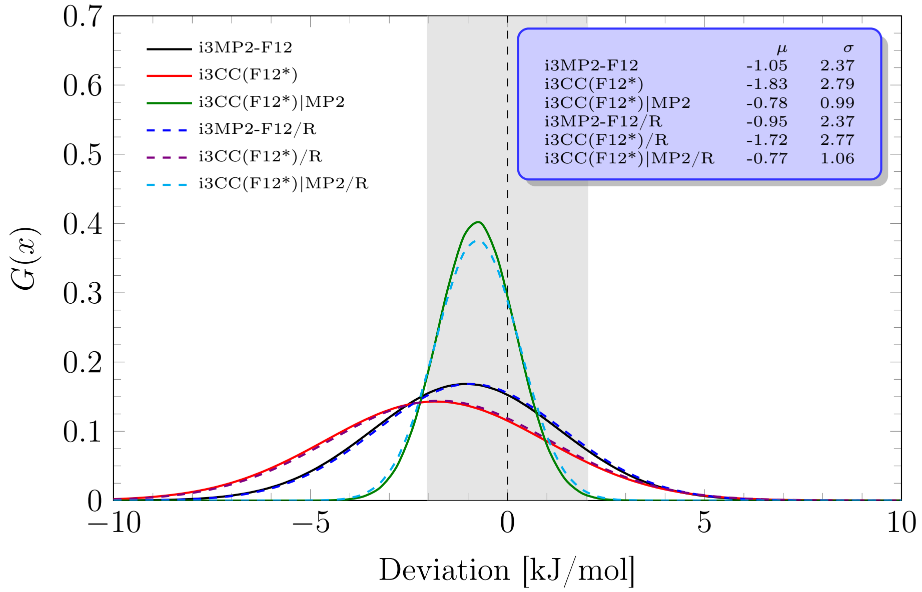
					subset I				
		SV			DSBSenv			reduced DSBSenv	
system	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2
$Al(H_2O)_6{}^{3+}$	-0.82	-1.51	-0.69	-1.93	-4.01	-2.08	-1.61	-3.69	-2.08
$(Ar)_{14}$	-0.16	-3.15	-2.98	-0.17	-1.10	-0.93	-0.22	0.15	0.37
$Be(H_2O)_6{}^{2+}$	0.25	0.20	-0.05	0.52	0.34	-0.18	0.52	0.34	-0.18
$(H_2O)_6$	0.01	-0.02	-0.03	-0.05	-0.05	0.01	-0.05	-0.05	0.01
$(He)_{10}$	0.02	0.02	0.00	0.02	0.02	0.00	0.01	0.02	0.00
$(HF)_{12}$	0.20	-0.06	-0.26	-0.17	-0.44	-0.27	-0.14	-0.41	-0.27
$Li(H_2O)_6{}^{+}$	0.26	0.27	0.01	0.27	0.34	0.06	0.28	0.34	0.07
$Mg(NH_3)_6{}^{2+}$	-0.85	-1.51	-0.66	-1.72	-2.37	-0.65	-1.54	-2.20	-0.66
$Na(H_2O)_6{}^{+}$	0.19	0.35	0.16	0.21	0.42	0.20	0.22	0.43	0.21
$(Ne)_{10}$	-0.02	-0.04	-0.02	-0.02	-0.04	-0.02	-0.02	-0.04	-0.02
μ	-0.09	-0.54	-0.45	-0.30	-0.69	-0.38	-0.26	-0.51	-0.26
σ	0.41	1.14	0.94	0.83	1.44	0.69	0.73	1.35	0.70
MAX	0.85	3.15	2.98	1.93	4.01	2.08	1.61	3.69	2.08
					subset II				
		SV			DSBSenv			reduced DSBSenv	
system	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2	i3MP2-F12	i3CC(F12*)	i3CC(F12*)|MP2
4-methyl−pyrazol	-0.17	-1.27	-1.10	-6.53	-7.45	-0.91	-6.08	-7.15	-1.07
5-methyl−imidazol	0.74	0.70	-0.04	1.37	1.36	-0.02	1.40	1.37	-0.03
$B_3H_6N_3$	-0.01	-0.06	-0.04	-0.42	-0.43	-0.01	-0.42	-0.43	-0.01
$B_3O_3H_3$	0.04	0.02	-0.02	-0.92	-1.21	-0.29	-0.89	-1.19	-0.30
$BOEt_3$	-0.58	-1.20	-0.62	-2.13	-2.87	-0.73	-2.12	-2.86	-0.73
C_2H_5COOH	0.01	-0.05	-0.05	0.57	0.36	-0.21	0.59	0.37	-0.21
$C_3H_4S_2COOH$	-0.81	-1.62	-0.81	-7.72	-9.33	-1.60	-7.70	-9.32	-1.63
C_3H_7SH	-0.25	-1.13	-0.87	-0.01	-0.79	-0.77	-0.01	-0.78	-0.77
$C_4H_9SO_2H$	-3.00	-4.19	-1.19	-5.30	-6.00	-0.70	-5.22	-5.94	-0.72
$C_4H_9SO_3H$	-1.28	-2.02	-0.74	-2.28	-3.20	-0.92	-2.21	-3.13	-0.92
$CH_2CHCOCl$	0.34	0.15	-0.19	1.02	0.53	-0.49	1.02	0.53	-0.49
$CH_2CHCOOH$	0.18	0.09	-0.09	0.70	0.43	-0.28	0.70	0.43	-0.28
$CH_3{-}C{\equiv}C{-}C_4H_9$	-0.66	-0.82	-0.16	-2.04	-2.54	-0.50	-2.03	-2.53	-0.50
$CH_3{-}CH{=}CH{-}C_4H_9$	-0.17	-0.41	-0.24	-2.00	-2.29	-0.29	-2.00	-2.29	-0.29
$ClEtPO_3H_2$	0.45	-0.05	-0.50	0.20	-0.36	-0.56	0.23	-0.33	-0.56
$ClSiEt_3$	-1.37	-1.42	-0.04	-2.33	-2.56	-0.24	-2.31	-2.55	-0.24
$(Me_3Si)_2O$	-0.38	-0.73	-0.35	-0.58	-1.00	-0.42	-0.57	-0.99	-0.42
C_7H_{16}	-0.01	-0.07	-0.06	-1.68	-2.31	-0.63	-1.67	-2.31	-0.64
$HSiEt_3$	-0.84	-0.92	-0.08	-0.14	-0.67	-0.54	-0.13	-0.67	-0.54
norbornan	-1.80	-1.15	0.65	-4.44	-4.02	0.42	-4.42	-4.02	0.40
$OPMe(OMe)_2$	0.05	-0.02	-0.06	4.12	4.50	0.38	4.14	4.52	0.38
$PhCCl_3$	-0.68	0.21	0.89	2.01	-0.15	-2.16	2.33	0.03	-2.29
PhCOCl	-1.01	-2.74	-1.73	-3.28	-6.69	-3.42	-3.18	-6.63	-3.45
PhNCS	1.10	0.06	-1.04	-3.09	-6.04	-2.95	-2.75	-5.82	-3.07
PhNHCl	0.29	0.80	0.51	-0.81	-4.28	-3.47	0.19	-3.64	-3.84
$PhSO_2F$	-0.41	0.01	0.41	2.51	0.40	-2.11	3.05	0.76	-2.29
Si_2Cl_6	-0.70	-1.27	-0.57	-2.65	-4.22	-1.57	-2.59	-4.14	-1.55
μ	-0.40	-0.71	-0.30	-1.33	-2.25	-0.93	-1.21	-2.17	-0.97
σ	0.83	1.09	0.59	2.70	3.06	1.06	2.71	3.03	1.12
MAX	3.00	4.19	1.73	7.72	9.33	3.47	7.70	9.32	3.84

The incremental errors in the absolute energy are listed in Table 7 for the whole test set. It is expected that the use of the DSBSenv basis sets as environmental basis in the domain-specific basis set introduce larger errors in the i3MP2-F12 and i3CC(F12*) energies. For the cluster subset (subset I) the error only increase slightly, as the standard deviation of the i3MP2-F12 method increase from 0.41 to 0.83 or 0.73 kJ/mol changing the SV/SVP basis sets with the DSBSenv or rDSBSenv basis sets. In the same manner, the i3CC(F12*) method has a small increase in error. The MP2-F12 error correction leads to an improvement of all errors and statistical values. Although the statistical errors of the i3MP2-F12 and i3CC(F12*) methods are increased using the optimized and reduced DSBSenv basis sets, the i3CC(F12*)|MP2 method has nearly the same accuracy within the absolute energy for the three environmental basis sets. Thus, the small DSBSenv basis sets decrease the computation time of the incremental scheme without significant loss of accuracy for cluster structures and the reduced DSBSenv ABS have a negligible influence on the accuracy.
For subset II the change of the *ad hoc* environmental basis to the DSBSenv basis sets also increases the errors of the i3MP2-F12 and i3CC(F12*) methods. The influence of the environmental basis for organic systems, especially aromatic π systems can be seen, e.g. for the PhNCS system where the error of the i3CC(F12*) method is increased from 0.06 to -6.04 kJ/mol. The MP2-F12 correction is important to assess the local error and achieve chemical accuracy for such systems.

Figure 2 shows the normal distributions of the old SV and the new DSBSenv environmental basis sets for the full test set. Comparing the i3MP2-F12 normal distributions, the error increases by about a magnitude of 1.6 kJ/mol in the standard deviation when using the smaller DSBSenv basis sets. Also the arithmetic mean errors μ shift slightly from the origin. It can be concluded that the smaller DSBSenv basis introduces a larger error in the incremental correlation energies, but in combination with the MP2-F12 correction the loss of accuracy is negligible compared to the old environmental basis. Moreover, the removal of the highest angular momentum functions of the DSBSenv ABSs introduces a negligible error,



(a) SV



(b) DSBSenv & rDSBSenv (R)

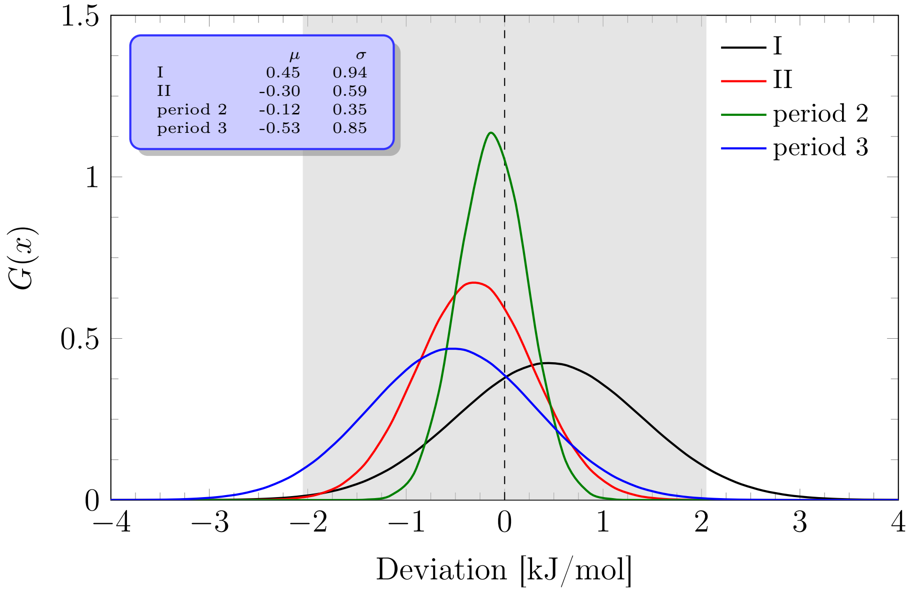
Figure 2: Normal distributions of the incremental error in the correlation energy [kJ/mol] for the test set. The deviations of the third-order incremental MP2-F12 [i3MP2-F12], the incremental CCSD(T)(F12*) [i3CC(F12*)] and incremental MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] are shown for the different environmental basis sets, SV (a) and the two DSBSenv (b). The R illustrates the normal distributions of the reduced DSBSenv ABSs. The gray highlighted area illustrates the 1 kcal/mol accuracy interval.

while increasing speed performance. In Figure 2b, the normal distributions of the reduced DSBSenv and DSBSenv basis sets are almost identical. For i3CC(F12*)|MP2 method, the difference in the statistical errors are well below 1 kcal/mol.

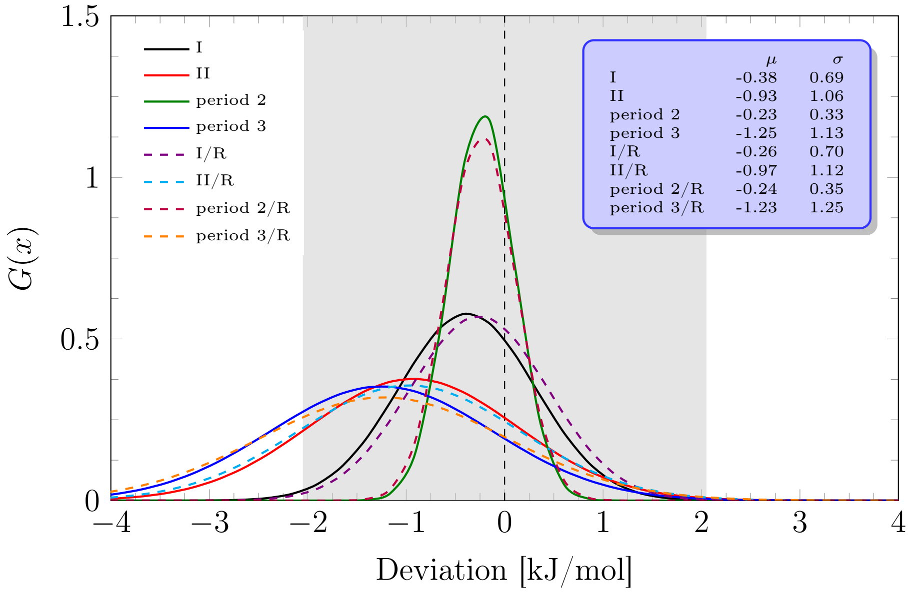
The diversity of the test set shows that the environmental DSBSenv basis sets work well for all different types of structures. The statistical errors are analyzed within the sub groups and the elements contained in the structures. Therefore, the normal distributions of the error in the absolute correlation energy for the subsets are illustrated in Figure 3. The statistical errors of the DSBSenv and the reduced DSBSenv are nearly the same. This illustrates that the reduction of the ABSs used for the environment of a domain introduces only a small error compared to the use of a minimal OBS. For all environmental basis sets, the standard deviations for the period 2 and period 3 subsets are nearly the same. The accuracy for organic systems (subset II) decreased when using the DSBSenv basis sets. For the cluster subset (subset I), the standard deviation of the DSBSenv improves about 0.3 kJ/mol with respect to the old SV environmental basis sets.

It is also important to compare the local error due to the domain-specific basis set incremental scheme with the intrinsic error of the method used. Therefore, the statistical errors of the incremental correlation energies for the CCSD(T)(F12*)/cc-pVDZ-F12 method are given in kJ/mol per correlated electron for the full test set and the subsets in Table 8.

Comparing the statistical errors of the subsets in Table 8, we see that for all environmental basis sets, the statistical errors are smaller for subset II than subset I and for the period 2 subset than the period 3 subset. Thus, the DSBSenv minimal basis sets for the environment of a domain have the same behavior as the SV basis sets for different types of systems. Also we can see that for the cluster subset (subset I) and the period 2 subset, the new environmental DSBSenv basis sets have nearly the same standard deviations as the old SV basis sets, and differ only by 0.05 kJ/mol per correlated electron. For the subset II and the period 3 subset, the DSBSenv basis sets have larger standard deviations. A possible reason for this is due to the optimized DSBSenv/OptRI ABSs. For the explicitly correlated



(a) SV



(b) DSBSenv & rDSBSenv (R)

Figure 3: Normal distributions of the errors in the absolute energies [kJ/mol] using the incremental MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] method with the different small environmental basis sets, SV (a), DSBSenv and reduced DSBSenv (b). The normal distributions are given for the subset I (I), subset II (II), the period 2 subset and the period 3 subset. The error for the reduced DSBSenv auxiliary basis sets is negligible, since the Gaussians are very close to those of the DSBSenv.

Table 8: Incremental errors in the absolute correlation energy of the MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] method [kJ/mol per correlated electron] to the canonical CCSD(T)(F12*)/cc-pVDZ-F12 energies using different basis sets for the environment of the domain-specific basis. The errors are analyzed concerning their subsets and atom types of the structures. The intrinsic error of the CCSD(T)(F12*)/cc-pVDZ-F12 method to the CCSD(T)/CBS limit is 0.130 kJ/mol per correlated electron.[14,15,117,118]

	SV			DSBSenv			reduced DSBSenv		
Test Set	μ	σ	MAX	μ	σ	MAX	μ	σ	MAX
subset I	-0.006	0.009	0.027	-0.007	0.014	0.043	-0.006	0.014	0.043
subset II	-0.008	0.014	0.038	-0.021	0.023	0.083	-0.022	0.025	0.091
period 2 subset	-0.003	0.010	0.034	-0.006	0.009	0.029	-0.006	0.009	0.033
period 3 subset	-0.011	0.015	0.038	-0.027	0.025	0.083	-0.027	0.027	0.091
full test set	-0.007	0.013	0.038	-0.017	0.022	0.083	-0.017	0.023	0.091

methods, the union of OBS and corresponding CABS is built for the calculation. The optimized DSBSenv/OptRI ABS have the same number of functions for period two and three elements. This leads to slightly larger RI errors (δ_{RI}) for the period three elements (see Table 3). All statistical errors due to the local approximations of the incremental scheme are smaller than the BSIE of the CCSD(T)(F12*)/cc-pVDZ-F12 method. A standard deviation of 0.130 kJ/mol per correlated electron relative to the CCSD(T)/CBS energies was found for this method.[14,15,117,118] In summary, the smaller environmental DSBSenv basis sets introduce a negligible error to the canonical correlation energy and these errors are only slightly different to the old environmental SV basis sets, which were used as starting guess for the application of the domain-specific basis set approach.[91] The local error is much smaller than the BSIE of the CCSD(T)(F12*)/cc-pVDZ-F12 method for both environmental basis sets.

Evaluation of the Accuracy in the Relative Energies

In previous work, the incremental scheme shows high accuracy in relative energies, resulting from the small error in the absolute energies.[79] Therefore, the change in accuracy for different environmental basis sets has been tested. The reactions are divided into two sets, one set contains structures from the subset I and the second set contains only structures of subset II. The reactions of subset I are:

$$Al^{3+} + 6H_2O \longrightarrow Al(H_2O)_6{}^{3+} \tag{8}$$

$$Be^{2+} + 6H_2O \longrightarrow Be(H_2O)_6{}^{2+} \tag{9}$$

$$Li^{+} + 6H_2O \longrightarrow Li(H_2O)_6{}^{+} \tag{10}$$

$$Na^{+} + 6H_2O \longrightarrow Na(H_2O)_6{}^{+} \tag{11}$$

$$6H_2O \longrightarrow (H_2O)_6 \tag{12}$$

$$12HF \longrightarrow (HF)_{12} \tag{13}$$

$$Mg^{2+} + 6NH_3 \longrightarrow Mg(NH_3)_6{}^{2+} \tag{14}$$

$$10He \longrightarrow (He)_{10} \tag{15}$$

$$10Ne \longrightarrow (Ne)_{10} \tag{16}$$

$$14Ar \longrightarrow (Ar)_{14} \tag{17}$$

The set of reactions that contain structures of subset II is:

$$B(OH)_3 + 3EtOH \longrightarrow B(OEt)_3 + 3H_2O \tag{18}$$

$$3\,B_2H_6 + 6NH_3 \longrightarrow 2B_3H_6N_3 + 12H_2 \tag{19}$$

$$HSiEt_3 + Cl_2 \longrightarrow ClSiEt_3 + HCl \tag{20}$$

$$H_2C{=}CHCOOH + PhCCl_3 \longrightarrow H_2C{=}CHCOCl + PhCOCl + HCl \tag{21}$$

$$H_3C{-}C{\equiv}C{-}C_4H_9 + H_2 \longrightarrow H_3C{-}CH{=}CH{-}C_4H_9 \tag{22}$$

$$H_3C{-}CH{=}CH{-}C_4H_9 + H_2 \longrightarrow C_7H_{16} \tag{23}$$

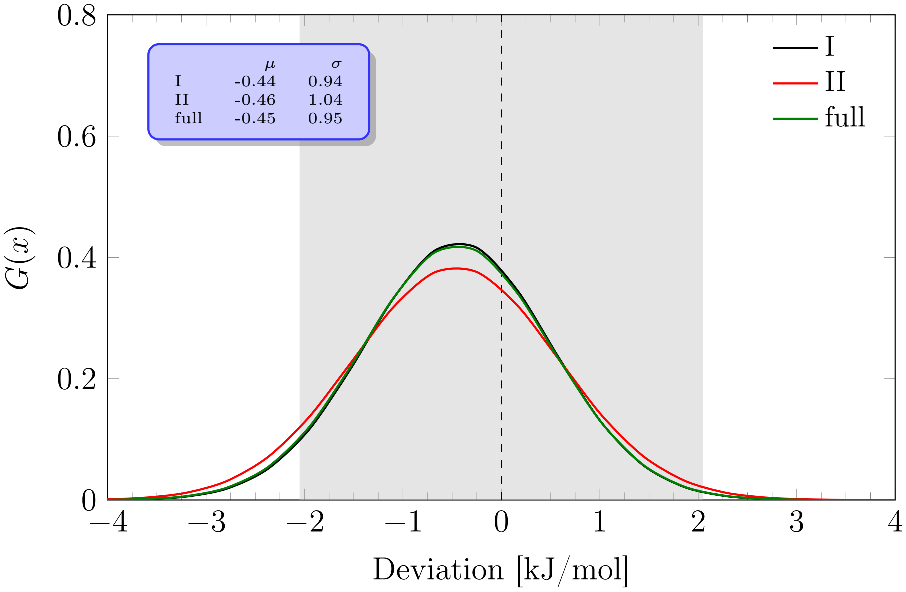
$$H_3C{-}C{\equiv}C{-}C_4H_9 + 2H_2 \longrightarrow C_7H_{16} \tag{24}$$

The reaction energies for the canonical CCSD(T)(F12*)/cc-pVDZ-F12 method are presented in Table 9, along with the errors in the i3CC(F12*)|MP2 reaction energies using the SV, DSBSenv and reduced DSBSenv basis sets as environmental basis in the framework of the domain-specific basis set. In general, all environmental basis sets have small errors and all errors are within chemical accuracy of 1 kcal/mol relative to the reference energies. The largest deviations are in reaction (8) with the $Al(H_2O)_6{}^{3+}$ system and reaction 17 containing the $(Ar)_{14}$ system. The standard deviations are 0.95, 0.64 and 0.63 kJ/mol for the SV, optimized and reduced DSBSenv basis sets, respectively. For this test set of reactions, the statistical errors in the reaction energies are nearly the same for all of the environmental basis sets, and both the optimized DSBSenv and reduced DSBSenv basis sets slightly improve the accuracy in addition to the efficiency. Furthermore, the incremental error is negligible compared to the intrinsic error for CCSD(T)(F12*)/cc-pVDZ-F12 reaction energies.

Further investigations of the errors have been performed by analyzing the subsets of the reactions. The first subset I (containing ionic and neutral cluster structures) is where the incremental scheme is expected to perform with an outstanding accuracy as documented in previous work.[80,81] The second subset contains organic systems, especially aromatic π systems, which are a challenging task for the incremental scheme. Figure 4 illustrates the error of the i3CC(F12*)|MP2 reaction energies relative to the canonical calculations on the CCSD(T)(F12*)/cc-pVDZ-F12 level. It can be seen that the optimized and reduced DSBSenv basis sets have a slightly better accuracy, since the standard deviations for the two subsets and the full set of reactions are smaller than those of the SV basis sets.

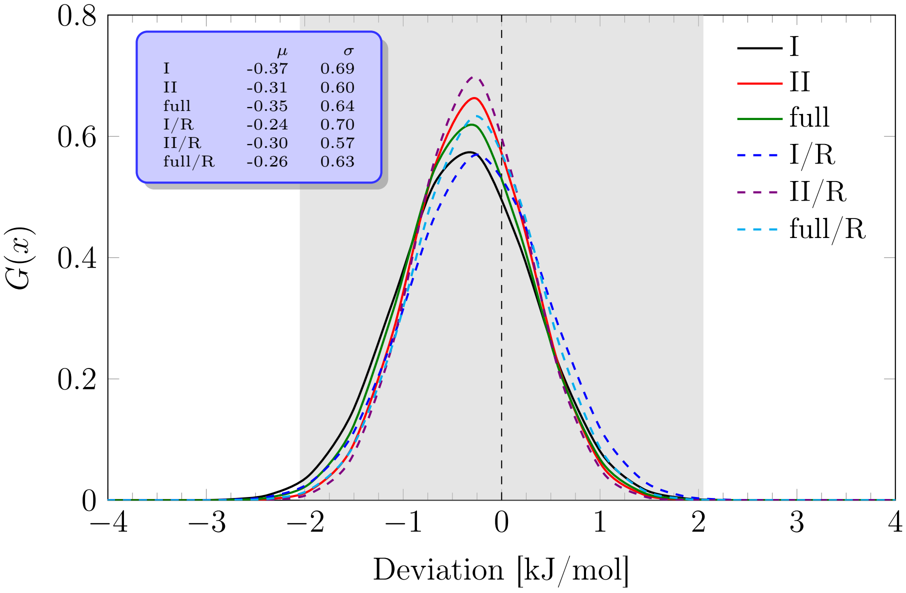
Local Treatment for Large Systems

Within the old environmental basis it was possible to calculate the incremental CCSD(T) correlation energy for the $Al(H_2O)_{25}{}^{3+}$ system at a quintuple-ζ quality.[81] The acronym aCVXZ' basis represents the cc-pwCVXZ basis for Al, cc-pVXZ for H, and aug-cc-pVXZ



	μ	σ
I	-0.44	0.94
II	-0.46	1.04
full	-0.45	0.95

(a) SV



	μ	σ
I	-0.37	0.69
II	-0.31	0.60
full	-0.35	0.64
I/R	-0.24	0.70
II/R	-0.30	0.57
full/R	-0.26	0.63

(b) DSBSenv & rDSBSenv (R)

Figure 4: Normal distributions of the errors in the reaction energies [kJ/mol] using the incremental MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] method with two different small environmental basis sets, SV (a) and DSBSenv and reduced DSBSenv (b). The normal distributions are given for reactions, containing structures of subset I (I), subset II (II) or all reactions (full). The shape of the Gaussian distributions indicates chemical accuracy.

Table 9: Reaction energies of all test reactions using the canonical CCSD(T)(F12*)/cc-pVDZ-F12 method ($\Delta E_{\mathbf{ref}}$) [kJ/mol] and the incremental errors of the MP2-F12 corrected CCSD(T)(F12*) [i3CC(F12*)|MP2] method with the SV ($\Delta\Delta E_{\mathbf{inc}}^{\mathbf{SV}}$), the DSBSenv ($\Delta\Delta E_{\mathbf{inc}}^{\mathbf{DSBSenv}}$) and the reduced DSBSenv ($\Delta\Delta E_{\mathbf{inc}}^{\mathbf{rDSBSenv}}$) basis as environmental basis in the framework of the domain-specific basis set.

Reaction	ΔE_{ref}	$\Delta\Delta E_{\mathrm{inc}}^{\mathrm{SV}}$	$\Delta\Delta E_{\mathrm{inc}}^{\mathrm{DSBSenv}}$	$\Delta\Delta E_{\mathrm{inc}}^{\mathrm{rDSBSenv}}$
Eq. (8)	-2938.04	-0.69	-2.08	-2.08
Eq. (9)	-1870.95	-0.05	-0.18	-0.18
Eq. (10)	-560.32	0.01	0.06	0.07
Eq. (11)	-173.89	-0.03	0.01	0.01
Eq. (12)	-435.90	0.16	0.20	0.21
Eq. (13)	-404.97	-0.11	-0.12	-0.12
Eq. (14)	-1437.18	-0.66	-0.65	-0.66
Eq. (15)	-0.46	0.00	0.00	0.00
Eq. (16)	-2.57	-0.02	-0.02	-0.02
Eq. (17)	-53.35	-2.98	-0.93	0.37
Eq. (18)	-16.63	-0.62	-0.73	-0.73
Eq. (19)	-652.82	-0.09	-0.02	-0.02
Eq. (20)	-296.45	0.03	0.30	0.30
Eq. (21)	-28.61	-2.73	-1.47	-1.37
Eq. (22)	-187.13	-0.25	0.04	0.04
Eq. (23)	-148.68	0.35	-0.17	-0.17
Eq. (24)	-335.81	0.10	-0.13	-0.13
μ		-0.45	-0.35	-0.26
σ		0.95	0.64	0.63
MAX		2.98	2.08	2.08

for O (X=D,T,Q,5). The incremental scheme enabled the 4–5 complete basis set extrapolation [CBS(45)] on CCSD(T) level using the extrapolation of Helgaker for the correlation energy.[2,5] For this large system, we want to illustrate the increased computational speed performance and the accuracy with the system size using the reduced DSBSenv basis sets. The second-order incremental scheme with the MP2 corrected CCSD(T) [i2CC|MP2] method as well as the domain based local pair natural orbital CCSD(T) of Neese including the T_0 approximation which is default in ORCA[119] [DLPNO-CCSD(T_0)][39] was used to calculate the interaction energy with respect to

$$\mathrm{Al}^{3+} + 25\mathrm{H_2O} \longrightarrow \mathrm{Al(H_2O)_{25}}^{3+} \tag{25}$$

Using DLPNO-CCSD(T), it is required to set tight thresholds for the PNOs describing non-covalent interactions. Therefore, both the normalPNO and tightPNO thresholds were used.[119] Table 10 details the correlation contribution to the interaction energy. On the above mentioned hardware, it was not possible to calculate the DLPNO-CCSD(T_0) correlation energy with higher level than triple-ζ quality. The correlation energy with the DLPNO-CCSD(T_0)/aCVTZ' method and normalPNO and tightPNO thresholds could be computed on a node with 64 GB RAM and a RAID with 4 times 1 TB disks for a fast data transfer.

Table 10: Correlation contribution to the interaction energies [kJ/mol] for the $\mathbf{Al(H_2O)_{25}}^{3+}$ system using different post-HF methods. The domain based local pair natural orbital CCSD(T) [DLPNO-CCSD(T_0)] with normal (normalPNO) and tight (tightPNO) thresholds, the second-order incremental MP2 corrected CCSD(T) [i2CC|MP2] and density fitted MP2 [DF-MP2], the site-site functional counterpoise correction[120] on MP2 level [DF-MP2+SSFC], and explicitly correlated MP2 [MP2-F12] methods were used. The CBS(45) extrapolated energies for the DF-MP2 and i2CC|MP2 methods are added as benchmark.

	DLPNO-CCSD(T_0)		i2CC|MP2		DF-MP2		MP2-F12	
Basis	normalPNO	tightPNO	SV	rDSBSenv		+SSFC		
aCVDZ'	-451.4	-475.6	-506.1	-504.2	-518.8	-357.4	cc-pVDZ-F12	-532.6
aCVTZ'	-476.3	-497.6	-532.0	-529.7	-534.3	-447.9	cc-pVTZ-F12	-518.1
aCVQZ'	-	-	-531.8	-530.2	-532.8	-488.8		
aCV5Z'	-	-	-528.9	-526.6	-530.2	-		
CBS(23)	-486.8	-506.9	-543.0	-540.5	-540.8	-486.0		
CBS(34)	-	-	-531.6	-530.5	-531.8	-518.7		
CBS(45)	-	-	-525.8	-522.9	-527.4	-		

The i2CC|MP2 energies with the two different environmental basis sets differ by around 2 kJ/mol for both the aCVDZ' and aCVTZ' basis, indicating that the reduced DSBSenv basis sets introduce practically negligible errors. The use of the new environmental basis sets reduces the CPU-time for the aCVDZ' calculation by 66% (from 417.4 to 140.4 h), and aCVTZ' by 63% (from 864.8 to 320.7 h). It can be seen from Table 10 that the incremental scheme recovers significantly more correlation energy than DLPNO-CCSD(T_0) for this system. Furthermore, with a CPU-time of 13.4 days, the incremental scheme with the reduced DSBSenv basis sets is faster than tightPNO DLPNO-CCSD(T_0) (19.4 days) for the aCVTZ' basis. Reducing the DLPNO-CCSD(T_0) thresholds to normalPNO settings results in a CPU-time of 2.3 days, but this introduces an additional error of around 20 kJ/mol, consistent with the recommendation of only using tightPNO settings for such a system.[119] Due to the inherent parallelization of the incremental scheme (using 1 processor for the server and 100 as clients) the wall-time for the aCVTZ' calculation is only 5.8 h, rising to just 3.3 days for the large aCV5Z' calculation.

Conclusions

In this work DSBSenv minimal basis sets and matching auxiliary basis sets have been developed for H–Ar for use as the environmental basis in the framework of the domain-specific basis set approach of the incremental scheme. These basis sets increase the computational speed performance of the incremental scheme and achieve a comparable accuracy to the previous *ad hoc* choice of environmental basis. Using the DSBSenv/MP2Fit ABS, the density fitting for the evaluation of the two-electron integrals are three to four orders of magnitude smaller than the BSIE of the DSBSenv OBS. Also the CABS approach to the RI with the DSBSenv/OptRI ABS introduces errors about two to three orders of magnitude smaller than the BSIE.

In combination with the domain-specific basis set approach, the environmental DSBSenv basis sets save 2.5–38.5 % CPU-time compared to the old def-SV OBS combined with the def-SV(P) ABS. Exchanging the environmental SV with the DSBSenv OBS, the standard deviation for the incremental MP2-F12 corrected CCSD(T)(F12*) method of the error in the absolute correlation energy increase slightly from 0.69 to 0.99 kJ/mol. Moreover, the DSBSenv basis sets have a slightly better accuracy for reaction energies, as the standard deviation decreases from 0.95 to 0.64 kJ/mol. To further decrease the computation time, the highest angular momentum functions are removed from every DSBSenv ABS. This reduced DSBSenv leads to significantly lower CPU-times for the test set, saving 19.4 % compared to the old environmental basis, with standard deviations of the error in the absolute and reaction energies of 1.06 and 0.63 kJ/mol, respectively. These values are close to those of the fully-optimized DSBSenv basis sets, thus it is recommended to use the reduced DSBSenv as environmental basis for the domain-specific basis incremental scheme.

The use of the reduced DSBSenv basis sets in the second-order domain-specific basis incremental scheme reduces the CPU-time for the $Al(H_2O)_{25}{}^{3+}$ system from 17.4 to 5.8 days in the aCVDZ' basis and from 36.0 to 13.4 days in the aCVTZ' basis. For this system, the incremental scheme is more accurate and faster than the tightPNO DLPNO-CCSD(T_0) method using basis sets with higher level than double ζ quality.

Acknowledgement

JF, JGH and TA thank the DAAD for financial support and facilitating the collaboration, D. P. Tew for the modifications to the `ricc2` module and Marcel John for providing the noble gas structures.

Supporting Information Available

The geometries of the investigated structures and their energies, further details of the ABS optimizations and all basis sets (in TURBOMOLE format) are contained in the Supporting Information. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A 5th-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(2) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley, Chichester, 2000.

(3) Ramabhadran, R. O.; Raghavachari, K. Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy. *J. Chem. Theory Comput.* **2013**, *9*, 3986–3994.

(4) Rezac, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the "Gold Standard," CCSD(T) at the Complete Basis Set Limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.

(5) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N-2, and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(6) Liakos, D. G.; Izsák, R.; Valeev, E. F.; Neese, F. What is the most efficient way to reach the canonical MP2 basis set limit? *Mol. Phys.* **2013**, *111*, 2653–2662.

(7) Heßelmann, A.; Korona, T. Intermolecular symmetry-adapted perturbation theory study of large organic complexes. *J. Chem. Phys.* **2014**, *141*, 094107.

(8) Vogiatzis, K. D.; Klopper, W. Accurate non-covalent interactions with basis-set corrections from interference-corrected perturbation theory: comparison with the S22B database. *Mol. Phys.* **2013**, *111*, 2299–2305.

(9) Vogiatzis, K. D.; Haunschild, R.; Klopper, W. Accurate atomization energies from combining coupled-cluster computations with interference-corrected explicitly correlated second-order perturbation theory. *Theo. Chem. Acc.* **2014**, *133*, 1446.

(10) Ahnen, S.; Hehn, A.-S.; Vogiatzis, K. D.; Trachsel, M. A.; Leutwyler, S.; Klopper, W. Accurate computations of the structures and binding energies of the imidazolebenzene and pyrrolebenzene complexes. *Chem. Phys.* **2014**, *441*, 17–22.

(11) Vogiatzis, K. D.; Klopper, W.; Friedrich, J. Non-covalent Interactions of CO2 with Functional Groups of Metal-Organic Frameworks from a CCSD(T) Scheme Applicable to Large Systems. *J. Chem. Theory Comput.* **2015**, *11*, 1574–1584.

(12) Klopper, W.; Samson, C. C. M. Explicitly correlated second-order Moller-Plesset methods with auxiliary basis sets. *J. Chem. Phys.* **2002**, *116*, 6397–6410.

(13) Klopper, W.; Bachorz, R. A.; Hättig, C.; Tew, D. P. Accurate computational thermochemistry from explicitly correlated coupled-cluster theory. *Theor. Chem. Acc.* **2010**, *126*, 289–304.

(14) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-zeta quality coupled-cluster correlation energies with triple-zeta basis sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921–1930.

(15) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Comment on Quintuple-zeta quality coupled-cluster correlation energies with triple-zeta basis sets by D. P. Tew, W. Klopper, C. Neiss and C. Hättig, Phys. Chem. Chem. Phys., 2007, 9, 1921 [erratum]. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6325–6327.

(16) Tew, D. P.; Klopper, W.; Hättig, C. Diagonal orbital-invariant explicitly-correlated coupled-cluster method. *Chem. Phys. Lett.* **2008**, *452*, 326–332.

(17) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(18) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(19) Feller, D.; Peterson, K. A.; Hill, J. G. On the effectiveness of CCSD(T) complete basis set extrapolations for atomization energies. *J. Chem. Phys.* **2011**, *135*, 044102.

(20) Pulay, P. Localizability of Dynamic Electron Correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(21) Sæbø, S.; Pulay, P. Local Configuration-interaction - An Efficient Approach For Larger Molecules. *Chem. Phys. Lett.* **1985**, *113*, 13–18.

(22) Pulay, P.; Sæbø, S. Orbital-invariant Formulation and 2nd-order Gradient Evaluation In Moller-plesset Perturbation-theory. *Theor. Chim. Acta* **1986**, *69*, 357–368.

(23) Sæbø, S.; Pulay, P. 4th-order Moller-plessett Perturbation-theory In the Local Correlation Treatment .1. Method. *J. Chem. Phys.* **1987**, *86*, 914–922.

(24) Sæbø, S.; Pulay, P. Local Treatment of Electron Correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

(25) Hampel, C.; Werner, H.-J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(26) Mata, R. A.; Werner, H.-J.; Schütz, M. Correlation regions within a localized molecular orbital approach. *J. Chem. Phys.* **2008**, *128*, 144106.

(27) Adler, T. B.; Werner, H.-J. Local explicitly correlated coupled-cluster methods: Efficient removal of the basis set incompleteness and domain errors. *J. Chem. Phys.* **2009**, *130*, 241101.

(28) Adler, T. B.; Werner, H.-J. An explicitly correlated local coupled cluster method for calculations of large molecules close to the basis set limit. *J. Chem. Phys.* **2011**, *135*, 144117.

(29) Werner, H.-J.; Schütz, M. An efficient local coupled cluster method for accurate thermochemistry of large systems. *J. Chem. Phys.* **2011**, *135*, 144116.

(30) Krause, C.; Werner, H. J. Comparison of explicitly correlated local coupled-cluster methods with various choices of virtual orbitals. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7591–7604.

(31) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(32) Schütz, M. Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T). *J. Chem. Phys.* **2000**, *113*, 9986–10001.

(33) Schütz, M.; Yang, J.; Chan, G. K. L.; Manby, F. R.; Werner, H. J. The orbital-specific virtual local triples correction: OSV-L(T). *J. Chem. Phys.* **2013**, *138*, 054109.

(34) Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108.

(35) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

(36) Liakos, D. G.; Hansen, A.; Neese, F. Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods. *J. Chem. Theory Comput.* **2011**, *7*, 76–87.

(37) Hansen, A.; Liakos, D. G.; Neese, F. Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals. *J. Chem. Phys.* **2011**, *135*, 214102.

(38) Huntington, L. M. J.; Hansen, A.; Neese, F.; Nooijen, M. Accurate thermochemistry from a parameterized coupled-cluster singles and doubles model and a local pair natural orbital based implementation for applications to larger systems. *J. Chem. Phys.* **2012**, *136*, 064101.

(39) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

(40) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.

(41) Tew, D. P.; Helmich, B.; Hättig, C. Local explicitly correlated second-order Moller-Plesset perturbation theory with pair natural orbitals. *J. Chem. Phys.* **2011**, *135*, 074107.

(42) Hättig, C.; Tew, D. P.; Helmich, B. Local explicitly correlated second- and third-order Moller-Plesset perturbation theory with pair natural orbitals. *J. Chem. Phys.* **2012**, *136*, 204105.

(43) Tew, D. P.; Hättig, C. Pair natural orbitals in explicitly correlated second-order moller-plesset theory. *Int. J. Quant. Chem.* **2013**, *113*, 224–229.

(44) Schmitz, G.; Hättig, C.; Tew, D. P. Explicitly correlated PNO-MP2 and PNO-CCSD and their application to the S66 set and large molecular systems. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167–22178.

(45) Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K. L. Tensor factorizations of local second-order Moller-Plesset theory. *J. Chem. Phys.* **2011**, *134*, 044123.

(46) Yang, J.; Chan, G. K. L.; Manby, F. R.; Schütz, M.; Werner, H. J. The orbital-specific-virtual local coupled cluster singles and doubles method. *J. Chem. Phys.* **2012**, *136*, 144105.

(47) Kurashige, Y.; Yang, J.; Chan, G. K. L.; Manby, F. R. Optimization of orbital-specific virtuals in local Moller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *136*, 124106.

(48) Stoll, H. Correlation-energy of Diamond. *Phys. Rev. B* **1992**, *46*, 6700–6704.

(49) Stoll, H. On the Correlation-energy of Graphite. *J. Chem. Phys.* **1992**, *97*, 8449–8454.

(50) Stoll, H. The Correlation-energy of Crystalline Silicon. *Chem. Phys. Lett.* **1992**, *191*, 548–552.

(51) Stoll, H.; Paulus, B.; Fulde, P. On the accuracy of correlation-energy expansions in terms of local increments. *J. Chem. Phys.* **2005**, *123*, 144108.

(52) Stoll, H. Can incremental expansions cope with high-order coupled-cluster contributions? *Mol. Phys.* **2010**, *108*, 243–248.

(53) Mata, R. A.; Stoll, H. An incremental correlation approach to excited state energies based on natural transition/localized orbitals. *J. Chem. Phys.* **2011**, *134*, 034122.

(54) Müller, C.; Usvyat, D.; Stoll, H. Local correlation methods for solids: Comparison of incremental and periodic correlation calculations for the argon fcc crystal. *Phys. Rev. B* **2011**, *83*, 245136.

(55) de Lara-Castells, M. P.; Stoll, H.; Civalleri, B.; Causà, M.; Voloshina, E.; Mitrushchenkov, A. O.; Pi, M. Communication: A combined periodic density functional and incremental wave-function-based approach for the dispersion-accounting

time-resolved dynamics of 4He nanodroplets on surfaces: 4He/graphene. *J. Chem. Phys.* **2014**, *141*, 151102.

(56) Zhang, J.; Dolg, M. Third-Order Incremental Dual-Basis Set Zero-Buffer Approach: An Accurate and Efficient Way To Obtain CCSD and CCSD(T) Energies. *J. Chem. Theory Comput.* **2013**, *9*, 2992–3003.

(57) Zhang, J.; Dolg, M. Approaching the complete basis set limit of CCSD(T) for large systems by the third-order incremental dual-basis set zero-buffer F12 method. *J. Chem. Phys.* **2014**, *140*, 044114.

(58) Zhang, J.; Heinz, N.; Dolg, M. Understanding Lanthanoid(III) Hydration Structure and Kinetics by Insights from Energies and Wave functions. *Inorg. Chem.* **2014**, *53*, 7700–7708.

(59) Zhang, J.; Dolg, M. Dispersion Interaction Stabilizes Sterically Hindered Double Fullerenes. *Chem. Eur. J.* **2014**, *20*, 13909–13912.

(60) Heinz, N.; Zhang, J.; Dolg, M. Actinoid(III) Hydration-First Principle Gibbs Energies of Hydration Using High Level Correlation Methods. *J. Chem. Theory Comput.* **2014**, *10*, 5593–5598.

(61) Friedrich, J.; Hanrath, M.; Dolg, M. Fully automated implementation of the incremental scheme: Application to CCSD energies for hydrocarbons and transition metal compounds. *J. Chem. Phys.* **2007**, *126*, 154110.

(62) Friedrich, J.; Hanrath, M.; Dolg, M. Error analysis of incremental electron correlation calculations and applications to clusters and potential energy surfaces. *Chem. Phys.* **2007**, *338*, 33–43.

(63) Friedrich, J.; Dolg, M. Implementation and performance of a domain-specific basis

set incremental approach for correlation energies: Applications to hydrocarbons and a glycine oligomer. *J. Chem. Phys.* **2008**, *129*, 244105.

(64) Friedrich, J.; Hanrath, M.; Dolg, M. Evaluation of incremental correlation energies for open-shell systems: Application to the intermediates of the 4-exo cyclization, Arduengo carbenes and an anionic water cluster. *J. Phys. Chem. A* **2008**, *112*, 8762–8766.

(65) Friedrich, J.; Hanrath, M.; Dolg, M. Using symmetry in the framework of the incremental scheme: Molecular applications. *Chem. Phys.* **2008**, *346*, 266–274.

(66) Friedrich, J.; Coriani, S.; Helgaker, T.; Dolg, M. Implementation of the incremental scheme for one-electron first-order properties in coupled-cluster theory. *J. Chem. Phys.* **2009**, *131*, 154102.

(67) Paulus, B.; Rosciszewski, K.; Stoll, H.; Birkenheuer, U. Ab initio incremental correlation treatment with non-orthogonal localized orbitals. *Phys. Chem. Chem. Phys.* **2003**, *5*, 5523–5529.

(68) Paulus, B. The method of increments - a wavefunction-based ab initio correlation method for solids. *Phys. Rep.* **2006**, *428*, 1–52.

(69) Müller, C.; Paulus, B. Wavefunction-based electron correlation methods for solids. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7605–7614.

(70) Voloshina, E.; Paulus, B. First Multireference Correlation Treatment of Bulk Metals. *J. Chem. Theory Comput.* **2014**, *10*, 1698–1706.

(71) Steenbergen, K. G.; Gaston, N.; Müller, C.; Paulus, B. Method of increments for the halogen molecular crystals: Cl, Br, and I. *J. Chem. Phys.* **2014**, *141*, 124707.

(72) Hammerschmidt, L.; Maschio, L.; Müller, C.; Paulus, B. Electron Correlation at the MgF2(110) Surface: A Comparison of Incremental and Local Correlation Methods. *J. Chem. Theory Comput.* **2015**, *11*, 252–259.

(73) Friedrich, J.; Dolg, M. Fully Automated Incremental Evaluation of MP2 and CCSD(T) Energies: Application to Water Clusters. *J. Chem. Theory Comput.* **2009**, *5*, 287–294.

(74) Friedrich, J.; Walczak, K.; Dolg, M. Evaluation of core and core-valence correlation contributions using the incremental scheme. *Chem. Phys.* **2009**, *356*, 47–53.

(75) Walczak, K.; Friedrich, J.; Dolg, M. On the incremental evaluation of BSSE-free interaction energies. *Chem. Phys.* **2009**, *365*, 38–43.

(76) Friedrich, J.; Hanrath, M.; Dolg, M. Fully Automated Implementation of the Incremental Scheme for Correlation Energies. *Z. Phys. Chem.* **2010**, *224*, 513–525.

(77) Friedrich, J.; Tew, D. P.; Klopper, W.; Dolg, M. Automated incremental scheme for explicitly correlated methods. *J. Chem. Phys.* **2010**, *132*, 164114.

(78) Walczak, K.; Friedrich, J.; Dolg, M. Fully automated incremental evaluation of MP2 and CCSD(T) core, core-valence and valence correlation energies. *Chem. Phys.* **2010**, *376*, 36–45.

(79) Friedrich, J.; Hänchen, J. Incremental CCSD(T)(F12*)|MP2: A Black Box Method To Obtain Highly Accurate Reaction Energies. *J. Chem. Theory Comput.* **2013**, *9*, 5381–5394.

(80) Friedrich, J.; Walczak, K. Incremental CCSD(T)(F12)|MP2-F12 - A Method to Obtain Highly Accurate CCSD(T) Energies for Large Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 408–417.

(81) Anacker, T.; Friedrich, J. Highly accurate incremental CCSD(T) calculations on aqua- and amine-complexes. *Mol. Phys.* **2013**, *111*, 1161–1172.

(82) Anacker, T.; Friedrich, J. New Accurate Benchmark Energies for Large Water Clusters: DFT Is Better Than Expected. *J. Comput. Chem.* **2014**, *35*, 634–643.

(83) Yang, J.; Dolg, M. Evaluation of electronic correlation contributions for optical tensors of large systems using the incremental scheme. *J. Chem. Phys.* **2007**, *127*, 084108.

(84) Friedrich, J.; McAlexander, H. R.; Kumar, A.; Crawford, T. D. Incremental evaluation of coupled cluster dipole polarizabilities. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14284–14296.

(85) Anacker, T.; Tew, D. P.; Friedrich, J. First UHF Implementation of the Incremental Scheme for Open-Shell Systems. *J. Chem. Theory Comput.* **2015**, just accepted, DOI:10.1021/acs.jctc.5b00933.

(86) Zhang, J.; Dolg, M. Third-Order Incremental Dual-Basis Set Zero-Buffer Approach for Large High-Spin Open-Shell Systems. *J. Chem. Theory Comput.* **2015**, *11*, 962–968.

(87) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian-basis Sets For Atoms Li To Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(88) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(89) Valeev, E. F. Improving on the resolution of the identity in linear R12 ab initio theories. *Chem. Phys. Lett.* **2004**, *395*, 190–195.

(90) Karypis, G.; Kumar, V. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.* **1998**, *20*, 359–392.

(91) Friedrich, J. Localized Orbitals for Incremental Evaluations of the Correlation Energy within the Domain-Specific Basis Set Approach. *J. Chem. Theory Comput.* **2010**, *6*, 1834–1842.

(92) Hättig, C.; Tew, D. P.; Köhn, A. Communications: Accurate and efficient approximations to explicitly correlated coupled-cluster singles and doubles, CCSD-F12. *J. Chem. Phys.* **2010**, *132*, 231102.

(93) Friedrich, J. Incremental Scheme for Intermolecular Interactions: Benchmarking the Accuracy and the Efficiency. *J. Chem. Theory Comput.* **2012**, *8*, 1597–1607.

(94) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(95) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; *et. al*, MOLPRO, version 2012.1, a package of *ab initio* programs. 2012.

(96) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.

(97) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple-[small zeta] basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(98) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(99) Hättig, C.; Hellweg, A.; Köhn, A. Distributed memory parallel implementation of energies and gradients for second-order Moller-Plesset perturbation theory with the resolution-of-the-identity approximation. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1159–1169.

(100) TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007.

(101) Weigend, F. Hartree-Fock exchange fitting basis sets for H to Rn. *J. Comput. Chem.* **2008**, *29*, 167–175.

(102) Kritikou, S.; Hill, J. G. Auxiliary Basis Sets for Density Fitting in Explicitly Correlated Calculations: The Atoms H–Ar. *J. Chem. Theory Comput.* **2015**, *11*, 5269–5276.

(103) Werner, H. J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, 164102.

(104) Knizia, G.; Werner, H. J. Explicitly correlated RMP2 for high-spin open-shell reference states. *J. Chem. Phys.* **2008**, *128*, 154103.

(105) Ten-no, S. Initiation of explicitly correlated Slater-type geminal theory. *Chem. Phys. Lett.* **2004**, *398*, 56–61.

(106) Peterson, K. A.; Adler, T. B.; Werner, H. J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B-Ne, and Al-Ar. *J. Chem. Phys.* **2008**, *128*, 084102.

(107) Hill, J. G.; Peterson, K. A. Correlation consistent basis sets for explicitly correlated wavefunctions: valence and core-valence basis sets for Li, Be, Na, and Mg. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10460–10468.

(108) Fletcher, R. *Practical Methods of Optimization*, 2nd ed.; Wiley: Chichester, 1987.

(109) Nelder, J. A.; Mead, R. A Simplex-method For Function Minimization. *Comput. J.* **1965**, *7*, 308–313.

(110) Yousaf, K. E.; Peterson, K. A. Optimized auxiliary basis sets for explicitly correlated methods. *J. Chem. Phys.* **2008**, *129*, 184108.

(111) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H. J. Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets. *J. Chem. Phys.* **2009**, *131*, 194105.

(112) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(113) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

(114) Bachorz, R. A.; Bischoff, F. A.; Glöß, A.; Hättig, C.; Höfener, S.; Klopper, W.; Tew, D. P. Software News and Updates The MP2-F12 Method in the TURBOMOLE Program Package. *J. Comput. Chem.* **2011**, *32*, 2492–2513.

(115) Friedrich, J.; Perlt, E.; Roatsch, M.; Spickermann, C.; Kirchner, B. Coupled Cluster in Condensed Phase. Part I: Static Quantum Chemical Calculations of Hydrogen Fluoride Clusters. *J. Chem. Theory Comput.* **2011**, *7*, 843–851.

(116) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

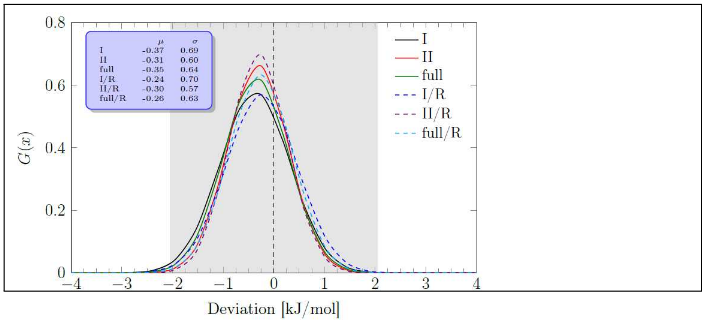
(117) Köhn, A. Explicitly correlated coupled-cluster theory using cusp conditions. II. Treatment of connected triple excitations. *J. Chem. Phys.* **2010**, *133*, 174118.

(118) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4–74.

(119) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

(120) Wells, B. H.; Wilson, S. Vanderwaals Interaction Potentials - Many-body Basis Set Superposition Effects. *Chem. Phys. Lett.* **1983**, *101*, 429–434.

Graphical TOC Entry



Normal distributions of the errors in the reaction energies [kJ/mol] using the incremental MP2-F12 corrected CCSD(T)(F12*) method.