

Fully Automated Implementation of the Incremental Scheme: Application to CCSD Energies for Hydrocarbons and Transition Metal Compounds

Joachim Friedrich¹, Michael Hanrath¹, Michael Dolg¹

¹*Institute for Theoretical Chemistry, University of Cologne, Greinstr. 4, 50939 Cologne, Germany*

A general fully automated implementation of the incremental scheme for molecules and embedded clusters in the framework of the Coupled Cluster singles and doubles theory is presented. The code can be applied to arbitrary order of the incremental expansion and is parallelized in a master/slave structure. We found that the error in the total correlation energy is lower than 1 kcal/mol with respect to the canonical CCSD calculation if the incremental series is truncated in a proper way.

I. INTRODUCTION

During the last decade a lot of effort has been made in the development of local correlation methods. Ground state methods like LMP2[1–9], Laplace MP2[10], FMO-based MP2[11], NLSCC, LCCD, LCCSD or LCCSD(T)[12–18] were developed in several groups. An alternative approach which can be combined with most standard quantum chemical program packages without changing the correlation modules is the incremental scheme of Stoll [19–21] which is based on ideas of Nesbet [22]. Within this procedure the occupied orbitals are localized and grouped into local domains. In the next step correlation calculations are performed for all single domains, all pairs, etc. until the desired accuracy is reached. The correlation energy of the total system can be expanded as:

$$E_{corr} = \sum_i \Delta\varepsilon_i + \frac{1}{2!} \sum_{ij} \Delta\varepsilon_{ij} + \frac{1}{3!} \sum_{ijk} \Delta\varepsilon_{ijk} + \dots \quad (1)$$

$$\Delta\varepsilon_i = \varepsilon_i \quad \Delta\varepsilon_{ij} = \varepsilon_{ij} - \Delta\varepsilon_i - \Delta\varepsilon_j$$

where ε_i is the correlation energy of the subsystem i and ε_{ij} the correlation energy of the subsystem i and j together. The general increment is given as:

$$\begin{aligned} \Delta\varepsilon_{i_1 i_2 \dots i_t} = \varepsilon_{i_1 i_2 \dots i_t} - \sum_{I_{t-1}} \Delta\varepsilon_{I_{t-1}} - \sum_{I_{t-2}} \Delta\varepsilon_{I_{t-2}} - \dots \\ - \sum_{I_2} \Delta\varepsilon_{I_2} - \sum_{I_1} \Delta\varepsilon_{I_1} \end{aligned} \quad (2)$$

where the index I_{t-v} is defined as the elements of the power set of $\{i_1, i_2, \dots, i_t\}$ with the cardinality $t-v$ (v runs from 1 to $(t-1)$). For molecules the series in eq. 1 is always finite and exact, because the last term corresponds to a correction yielding the exact correlation energy. Since local orbitals usually decay very rapidly, we can expect that the series can be truncated at low order (e.g. second or third [19, 23–25]). A nice property of the incremental scheme is the possibility to extend it to the multi-reference case [26–30]. One aspect limiting the applications of the incremental scheme is that up to now a lot of handwork is required in order to obtain an energy within the framework of the incremental scheme.

Therefore it is the goal of this work is to generate a completely automatic procedure in order to use the incremental scheme as a black box method. Our present work relies on the use of standard correlation codes, i.e. the emphasis is on the potential accuracy of the approach rather than on computational savings.

II. THEORY

A. Notation

For compactness of notation we introduce the variable \mathbb{X} as summation index which runs over all members of the power set of the set of domains $\mathcal{P}(\mathbb{D})$ up to a given cardinality. Within this notation we

can write eq. 1 as:

$$E_{\text{corr}} = \sum_{\substack{\mathbb{X} \\ \mathbb{X} \in \mathcal{P}(\mathbb{D}) \wedge |\mathbb{X}| \leq \mathcal{O}}} \Delta \varepsilon_{\mathbb{X}}$$

$$\mathbb{D} := \text{set of domains} \tag{3}$$

$$\mathcal{P}(\mathbb{D}) := \text{power set of the set of domains}$$

$$\mathcal{O} := \text{order of the expansion}$$

B. Further simplifications

If two parts of a given domain are far apart, we can set the incremental energy to zero, since the parts can be treated as uncorrelated (e.g.: $\varepsilon_{ij} \approx \varepsilon_i + \varepsilon_j$).

$$\Delta \varepsilon_{ij} = \varepsilon_{ij} - \Delta \varepsilon_i - \Delta \varepsilon_j \approx \varepsilon_i + \varepsilon_j - \varepsilon_i - \varepsilon_j = 0 \tag{4}$$

This can easily be generalized to higher order increments. The implementation of such a truncation can be achieved by a distance threshold. Eq. 4 turns out to be very useful for production calculations.

C. Obtaining groups of occupied orbitals

The occupied orbitals are localized with a Foster-Boys procedure [31] using the algorithm of Edmiston and Ruedenberg [32]. For these localized orbitals we build the centers of charge due to the diagonal elements of the dipole integrals in MO-basis.

$$\phi_a \mapsto \vec{R}_a := \begin{pmatrix} \langle \phi_a | x | \phi_a \rangle \\ \langle \phi_a | y | \phi_a \rangle \\ \langle \phi_a | z | \phi_a \rangle \end{pmatrix} = \begin{pmatrix} x_a \\ y_a \\ z_a \end{pmatrix} \tag{5}$$

Using eq. 5 we are able to map our set of occupied orbitals \mathbb{O} to a set of vectors. From this set of vectors we build the distance matrix D of all vector pairs. In the next step we construct the connectivity matrix

C according to:

$$C_{ij} = \begin{cases} 10^8, & \text{if } D_{ij} \leq t \wedge \frac{f}{D_{ij}} \geq 10^8 \\ \frac{f}{D_{ij}}, & \text{if } D_{ij} \leq t \wedge \frac{f}{D_{ij}} < 10^8 \\ 0, & \text{if } D_{ij} > t \end{cases} \quad (6)$$

where t is a distance threshold and f is a constant stretching factor of 10000. The factor of 10^8 enters in order to avoid an overflow of integers. Sincemetis graph partitioning [33] needs the number of parts in advance, we introduce the `domain_size_parameter` to control this number. Due to the fast increase of the calculation time it is more convenient to control the size of the domains than a given number of domains. We calculate the `number_of_parts` parameter formetis according to:

$$\text{number_of_parts} = \frac{\text{nr_of_occupied_orbitals}}{\text{domain_size_parameter}} \quad (7)$$

Furthermore, since the graph partitioning requires integer variables we do a type cast from double to integer. The graph partitioning is done with the condition that the sum of the cut edge weights is minimal. According to the definition of our connectivity matrix C this forces close lying orbitals to be in one domain. At the end we obtain compact disjoint subsets of our localized occupied orbitals.

D. Excitation spaces for one-site domains

Our goal is to obtain a virtual space \mathbb{V}_{ϕ_a} for every occupied orbital ϕ_a :

$$\phi_a \mapsto \{\tilde{\phi}_i^{AO}\} \quad (8)$$

This is analogous to introducing excitation domains as in the well established local MPn methods of Pulay et al. [4] or Werner et al. [5]. Our virtual space is spanned by a set of projected atomic orbitals (PAO's) $\{\tilde{\phi}_i^{AO}\}$. The set of PAO's is constructed according to [4, 12]:

$$|\tilde{\phi}_i^{AO}\rangle = |\phi_i^{AO}\rangle - \sum_j^{\text{occ}} \langle \phi_j^{MO} | \phi_i^{AO} \rangle |\phi_j^{MO}\rangle \quad (9)$$

In local orbitals we recognize that an atomic orbital is important if it's center is close to the center of charge of the MO to be correlated. Guided by the decay of the localized occupied orbitals we can restrict

the excitation space of an occupied orbital according to:

$$\int (\phi_a^{MO} - \bar{\phi}_a^{MO})^2 d\tau \leq t_{dens} \quad (10)$$

We use the AO representation of $\bar{\phi}_a^{MO}$ to find the most important AO functions in ϕ_a^{MO} .

$$\begin{aligned} \phi_a^{MO} &= \sum_i c_i \phi_{ai}^{AO}(\vec{r}, \vec{\lambda}) \\ \bar{\phi}_a^{MO} &= \sum_i c_{h_i} \phi_{ah_i}^{AO}(\vec{r}, \vec{\lambda}) \end{aligned} \quad (11)$$

where $\vec{\lambda}$ is the shift vector of the AO-function. Since we usually have a set of AO basis functions on the same center, we define the set $\mathbb{E}_{\vec{v}}$:

$$\mathbb{E}_{\vec{v}} = \{\phi^{AO}(\vec{r}, \vec{\lambda}) | \vec{\lambda} = \vec{v}\} \quad (12)$$

According to eq. 12 we can divide the set of AO-functions into disjoint subsets. We can find the important AO-functions in ϕ_a^{MO} by successively using the sets $\mathbb{E}_{\vec{v}}$ where \vec{v} is closest to the center of charge of ϕ_a^{MO} until eq. 10 is fulfilled. In order to get the ordering of the sets $\mathbb{E}_{\vec{v}}$ according to the distance to the center of charge of the a-th local MO we define:

$$|\vec{R}_a - \vec{v}_s| > |\vec{R}_a - \vec{v}_t| \text{ for } s > t \quad (13)$$

Now we find the smallest n where eq. 14 is fulfilled.

$$\int \left(\phi_a^{MO} - \sum_{i=1}^n \sum_{h \in \mathbb{E}_{\vec{v}_i}} \langle \phi_h^{AO} | \phi_a^{MO} \rangle \phi_h \right) d\tau \leq t_{dens} \quad (14)$$

With this procedure we obtain $\bar{\phi}_a^{MO}$. From eq. 9 we identify the mapping:

$$\phi_i^{AO} \mapsto \tilde{\phi}_i^{AO} \quad (15)$$

The representation of $\bar{\phi}_a^{MO}$ contains a set of AO-functions which can be mapped to their corresponding PAO's according to eq. 15. Thus we find a local excitation space \mathbb{V}_{ϕ_a} for the orbital ϕ_a^{MO} .

In order to obtain a local excitation space \mathbb{V}_i for our one-site domain i we have to unify the sets of PAO's which correspond to the occupied orbitals in the domain.

$$\mathbb{V}_i = \bigcup_{\phi_a \in \mathbb{O}_i} \mathbb{V}_{\phi_a} \quad (16)$$

E. Construction of the n-site domains

The n-site domains are constructed using simple set theory. $\mathbb{O}_{i_1 i_2}$ is defined as the unification of the set \mathbb{O}_{i_1} and \mathbb{O}_{i_2} . The same holds true for the virtual space $\mathbb{V}_{i_1 i_2}$. In general the n-site domains are constructed according to:

$$\mathbb{O}_{\mathbb{K}} = \bigcup_{\lambda \in \mathbb{K}} \mathbb{O}_{\lambda} \quad (17)$$

$$\mathbb{V}_{\mathbb{K}} = \bigcup_{\lambda \in \mathbb{K}} \mathbb{V}_{\lambda} \quad (18)$$

$\mathbb{K} = \{i_1, i_2, i_3, \dots, i_n\} :=$ subset of the domains

\mathbb{D} with the cardinality n

F. Obtaining correlation energies

In order to calculate the energies with a standard quantum chemistry package we have to account for the linear dependencies and the non-orthogonality in the PAO space. For this purpose we use a linear transformation which includes symmetric orthogonalization:

$$D^{-\frac{1}{2}} U^\dagger C^T S^{AO} C U D^{-\frac{1}{2}} = \mathbb{1} \quad (19)$$

with the MO coefficient matrix C , the overlap matrix in AO basis S^{AO} , the diagonal matrix $D = U^\dagger C^T S^{AO} C U$ and the Matrix \tilde{U} which diagonalizes $\tilde{S} = C^T S^{AO} C$. The matrix U is obtained by restricting \tilde{U} to those eigenvectors which correspond to an eigenvalue greater than 10^{-10} . Now we build the new MO matrix within the localized occupied orbitals and with the new orthogonalized linear independent PAO's. All occupied orbitals which are not in the set $\mathbb{O}_{\mathbb{K}}$ are frozen, and all virtual orbitals which are not in $\mathbb{V}_{\mathbb{K}}$ are deleted in the following CCSD calculation. Finally we obtain the total CCSD correlation energy according to eq. 1.

III. COMPUTATIONAL DETAILS

A. Geometries

The geometries were obtained by optimization with the RI-BP86/SVP method in the TURBOMOLE 5.6 [34] quantum chemistry package. Stationary points were characterized by force constant calculations.

B. Incremental Calculations

The implementation contains interfaces to the MOLPRO and DALTON [35] quantum chemistry packages in order to obtain the molecular orbital coefficient matrix, the overlap matrix in AO-basis and the dipole integrals in AO-basis from a previous SCF calculation. After extraction of this data a Foster-Boys localization [31] with unitary 2x2 rotations in the occupied space is performed. We use the Foster-Boys functional together with the transformation algorithm of Edmiston and Ruedenberg [32] using a threshold of 10^{-12} for $D_{\max}(u_i, u_j) - D(\varphi_i, \varphi_j)$ (for details see eq. 26 in [32] and eq. 15 in [36]).

Orbitals which are treated as frozen core in the calculation are excluded from the unitary transformations. In the next step we calculate the CCSD correlation energies of the domains with MOLPRO or DALTON. The DALTON calculations were done in order to check the convergence of the incremental series when the virtual space is truncated according to eq. 10. The MOLPRO calculations were done with the complete virtual space.

For n domains and the order \mathcal{O} , we include all $\binom{n}{\mathcal{O}}$ possible increments in the summation, in order to check the convergence behavior of the incremental series eq. 1. Since the number of calculations increases quite fast, we had to tighten the threshold for the SCF energies (GTHRESH, ENERGY=1.d-10) and the threshold for the CCSD energies (THRESH, ENERGY=1.00D-08, COEFF=1.00D-5).

C. Platform and Software

All calculations were done on a cluster of Pentium IV personal computers with a 32-bit architecture, 1.35 GB of random access memory (RAM) and 40 GB disk space.

By construction the incremental approach is inherently parallel with no significant communication. Our code is built upon a master/slave structure. The master keeps track of the calculations, sends the jobs to the slaves and calculates the incremental energy at the end. The slaves get the MO-Matrix and the number of frozen/deleted electrons via a wrapper class from the master. They perform independently the CCSD calculations and send the energy and the error information back to the master. This makes sure that all CCSD energies in the expansion are converged. If no more jobs are available a slave terminates.

IV. APPLICATIONS

We investigate the performance of the incremental scheme on a set of test molecules with different bonding situations (figures in [37]). For this purpose we choose hydrocarbons and complex compounds. We also included some challenging delocalized systems like naphthalene or the conjugated π -systems of an alkene or an alkyne. In order to monitor the convergence we evaluated the incremental expansion up to 4th order.

In this work we do not address the symmetry in our systems, because the Foster-Boys localization criterion may contradict with the symmetry of the molecule. The use of symmetry, also for periodic systems, will be studied in a future publication.

A. Hydrocarbon Compounds

Table I shows the convergence behavior of the incremental expansion for a set of hydrocarbon compounds. For molecules of this type we can truncate the expansion of the correlation energy at 3rd order while recovering the correlation energy almost exactly. An expansion up to second order increments is already a good guess for the correlation energy whereas the first order energy has a large error (for the chosen domain size). We emphasize that the error in the total correlation energy at 3rd order is lower than 1 kcal/mol for sigma bonded chains as well as for aromatic or conjugated compounds.

B. Transition metal complexes

Transition metal complex compounds are a very important class of compounds in organic and inorganic chemistry. A huge number of modern homogeneous catalysts belong to this type of molecules. Since these compounds usually assemble a nearly spherical shape they challenge local correlation methods. In table II we see that the convergence of the incremental series depends on the type of the ligands. For the molybdenum fluoro complex we obtain very fast convergence. In the other cases we need 4th order increments in order to achieve convergence to 0.1 % of the correlation energy. We note that in these cases the sum of the 4th order contributions is still quite large. The source of this behavior is that individual energy increments are still quite large at this order. In the case of $\text{Nb}_2\text{Cl}_{10}$, MoCl_6 and TiCp_2Cl_2 we find still large contributions of the adjacent 4th order increments ($\approx 10^{-4}$ Hartree). In MoF_6 for example the 4th order increments are in the order of 10^{-5} Hartree. Therefore we think that the large polarizability of the Cl^- ligands is responsible for the slow convergence behavior. This is supported by the fact, that the centers of charge in the MoF_6 complex are closer on each F^- and closer to the neighboring fluorine atoms than the corresponding centers of charge in MoCl_6 . The larger distance of the centers of charge on a Cl atom is the reason why it is easier to polarize Cl^- than F^- . This effect can be seen in the incremental energies of MoF_6 and MoCl_6 .

Our incremental calculations on TiCp_2Cl_2 account for 99.99% of the correlation energy. This behavior does not change dramatically if different domains are chosen, as we can see from table III. From this table we can see that the convergence of the incremental expansion is good also for small domain sizes. We want to point out that it is possible to obtain accurate correlation energies before the spanned space of the orbitals in the domains approaches the full space.

Again we find that the first order approximation yields an inaccurate correlation energy (for the chosen domain size), the second order is a good guess and the third order has a small error, whereas we obtain excellent agreement at 4th order level.

C. Virtual space truncation

A necessary condition for a linear scaling correlation-method is the possibility to truncate the virtual space. Therefore we use DALTON to calculate the incremental series up to third order for the hexane molecule according to different values of the density threshold in eq. 10 (table IV). We found that the convergence of the incremental series is affected, but it still converges to the exact CCSD energy. We note that hexane is still a small molecule and we could not save significant parts of the virtual space. For big molecules we can expect a larger saving.

D. Neglecting small contributions

In order to check the approximation according to eq. 4 we summed the complete set of incremental energies in eq. 3 with different energy thresholds according to eq. 20.

$$E_{\text{corr}} = \sum_{\substack{\mathbb{X} \\ \mathbb{X} \in \mathcal{P}(\mathbb{D}) \wedge |\mathbb{X}| \leq \mathcal{O}}} \Delta \varepsilon_{\mathbb{X}} \quad (20)$$

$$|\Delta \varepsilon_{\mathbb{X}}| > E_{\text{thres}}$$

Where \mathbb{X} runs over all members of the power set of the set of the domains $\mathcal{P}(\mathbb{D})$, up to a certain cardinality as in eq. 3 (card=4 in table V). Table V demonstrates the performance of eq. 4. It is clear that we can neglect more increments in chain-type molecules than in sphere-shaped molecules, because the distances between certain domains are much larger in former. In naphthalene we have two effects, "nonlocal" orbitals and rather short distances between the fragments. It is therefore not possible to neglect a significant number of increments in this case. For a reasonable threshold of 10^{-5} Hartree we can save 13-83% of the calculations in quite small molecules. We note that the higher order increments are more expensive than the low order increments. Therefore we will save more than 13-83% of the total cpu time because the higher order increments are usually smaller than the low order increments.

V. CONCLUSIONS

Collecting the results of table I and table II, we find that the convergence of the incremental scheme is quite fast, also for some delocalized systems like aromatic compounds or conjugated polyenes. The convergence is somehow slower for compounds like $\text{Nb}_2\text{Cl}_{10}$ or TiCp_2Cl_2 . We emphasize that we still obtain 99.34% of total CCSD correlation energy at 3rd order level of increments in the worst case of this study. At 4th order we obtained only in the worst case an error of about 0.12%. This is a somewhat higher accuracy than Pflüger and Werner [43] got for their LCCSD|LMP2 scheme. For our alkane, alkene and alkyne-chains we get almost 100 % of the correlation energy at 3rd order level of increments, which is in excellent agreement with the exact CCSD correlation energy. If we compare this with the LCCSD scheme of Subotnik and Head-Gordon, we find that our method yields in this case a more accurate result, too.

The first order approximation is in all cases inaccurate, whereas the second order is usually a good guess for the correlation energy. The expansion up to 3rd order yields in standard organic molecules very accurate results, whereas it may have an error up to 7.6 kcal/mol in difficult cases.

According to table IV we find that it is possible to truncate the virtual space for the domains in the incremental expansion. This is a necessary condition to obtain a linear scaling incremental CCSD-method. Another one is a linear scaling MO-transformation which was already published by Werner and Schütz. The number of CCSD calculations will only increase linearly with the system size if eq. 4 is used and the time of the single CCSD calculations can be treated as constant if the virtual space is truncated. Therefore we conclude that the incremental expansion is a powerful scheme to obtain a parallel local Coupled Cluster code with low order scaling.

Since the incremental scheme is independent of the correlation method, it can be applied to almost all quantum chemical methods. Another aspect is the extension of the presented scheme to multi-reference methods in order to calculate excited states, as well as the calculation of the correlation energy in crystalline solids. In the latter case we need a correct treatment of the symmetry for the local occupied orbitals, since errors due to approximate symmetric Foster-Boys orbitals will increase very rapidly with the order of the incremental expansion.

VI. ACKNOWLEDGMENTS

Financial support of the German Science Foundation (DFG) through the priority programme 1145 is gratefully acknowledged.

- [1] P. Pulay, S. Saebø, *Theor. Chim. Acta.* **69**, 357 (1986).
- [2] S. Saebø, P. Pulay, *J. Chem. Phys.* **86** (2), 914 (1987).
- [3] S. Saebø, P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).
- [4] J. W. Boughton, P. Pulay, *J. Comput. Chem.* **14**, 736 (1993).
- [5] M. Schütz, G. Hetzer, H. J. Werner, *J. Chem. Phys.* **111** (13), 5691 (1999).
- [6] P. E. Maslen, M. Head-Gordon, *Chem. Phys. Lett.* **283**, 102 (1998).
- [7] P. E. Maslen, M. Head-Gordon, *J. Chem. Phys.* **109**, 7093 (1998).
- [8] P. E. Maslen, M. S. Lee, M. Head-Gordon, *Chem. Phys. Lett.* **319**, 205 (2000).
- [9] M. S. Lee, P. E. Maslen, M. Head-Gordon, *J. Chem. Phys.* **112**, 3592 (2000).
- [10] P. Y. Ayala, G. E. Scuseria, *J. Chem. Phys.* **110**, 3660 (1999).
- [11] D. G. Federov, K. Kitaura, *J. Chem. Phys.* **121**, 2483 (2004).
- [12] C. Hampel, H.-J. Werner, *J. Chem. Phys.* **104**, 6286 (1996).
- [13] M. Schütz, *J. Chem. Phys.* **113** (22), 9986 (2000).
- [14] M. Schütz, H.-J. Werner, *J. Chem. Phys.* **114** (2), 661 (2001).
- [15] N. Flocke, R. J. Bartlett, *J. Chem. Phys.* **121**, 10935 (2004).
- [16] J. E. Subotnik, M. Head-Gordon, *J. Chem. Phys.* **123**, 64108 (2005).
- [17] J. E. Subotnik, A. Sodt, M. Head-Gordon, *J. Chem. Phys.* **125**, 74116 (2006).
- [18] A. Auer, M. Nooijen, *J. Chem. Phys.* **125**, 24104 (2006).
- [19] H. Stoll, *Chem. Phys. Lett.* **91**, 548 (1992).
- [20] H. Stoll, *Phys. Rev. B: Condens. Matter Mater. Phys.* **46**, 6700 (1992).
- [21] H. Stoll, *J. Chem. Phys.* **97**, 8449 (1992).
- [22] R. K. Nesbet, *Adv. Chem. Phys.* **14**, 1 (1969).

- [23] A. Shukla, M. Dolg, P. Fulde, H. Stoll, Phys. Rev. B **60**, 5211 (1999).
- [24] A. B. Mukhopadhyay, M. Dolg, J. Chem. Phys. **120**, 8734 (2004).
- [25] H. Stoll, B. Paulus, P. Fulde, J. Chem. Phys. **123**, 144108 (2005).
- [26] H. Stoll, Ann. Phys. **5**, 355 (1996).
- [27] M. Mödel, M. Dolg, P. Fulde, J. Chem. Phys. **106**, 1836 (1997).
- [28] B. Paulus, Chem. Phys. Lett. **371**, 7 (2003).
- [29] V. Bezugly, U. Birkenheuer, Chem. Phys. Lett. **399**, 57 (2004).
- [30] V. Bezugly, *Wavefunction-based method for excited- state electron correlations in periodic systems*, Doktorarbeit, Dresden (2004).
- [31] J. M. Foster, S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).
- [32] C. Edmiston, K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
- [33] G. Karypis, V. Kumar, SIAM J. Sci. Comput. **20** (1), 359 (1998).
- [34] R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, H. Horn, C. Huber, U. Huniar, C. Kölmel, M. Kollwitz, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, *Turbomole 5*, Institut für Physikal. Chemie, Universität Karlsruhe, Kaiserstr. 12, D-76128 Karlsruhe (2002).
- [35] *DALTON, a molecular electronic structure program, Release 2.0 (2005)*, see <http://www.kjemi.uio.no/software/dalton/dalton.html>.
- [36] J. Pipek, P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
- [37] See EPAPS Document No. for (brief description). This document can be reached through a direct link in the online article's HTML reference section or via the EPAPS homepage (<http://www.aip.org/pubser/epaps.html>).
- [38] W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
- [39] P. C. Hariharan, J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
- [40] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).
- [41] M. Dolg, U. Wedig, H. Stoll, H. Preuss, J. Chem. Phys. **86**, 866 (1987).
- [42] D. Andrae, U. Häussermann, M. Dolg, H. Stoll, H. Preuss, Theor. Chim. Acta. **77**, 123 (1990).
- [43] H.-J. Werner, K. Pflüger, <http://www.theochem.uni-stuttgart.de/forschung/preprints.html>.

system	order i	i -th order correction	$E_{\text{corr}}(i)$	error	% E_{corr}
		[au]	[au]	[kcal/mol]	
pentane ^a	1	-0.661436	-0.661436	89.89	82.20
	2	-0.145661	-0.807097	-1.52	100.30
	3	0.002399	-0.804697	-0.01	100.00
	4	0.000020	-0.804677	0.00	100.00
full CCSD			-0.804677		
decane ^a	1	-1.257413	-1.257413	204.12	79.45
	2	-0.331197	-1.588610	-3.71	100.37
	3	0.005823	-1.582787	-0.05	100.01
	4	0.000084	-1.582704	0.00	100.00
full CCSD			-1.582702		
alkene ^a	1	-1.052959	-1.052959	229.75	74.20
	2	-0.374002	-1.426960	-4.94	100.55
	3	0.007689	-1.419271	-0.12	100.01
	4	0.000203	-1.419069	0.01	100.00
full CCSD			-1.419088		
alkyne ^a	1	-1.399036	-1.399036	395.63	68.93
	2	-0.650050	-2.049085	-12.28	100.96
	3	0.018755	-2.030331	-0.52	100.04
	4	0.000753	-2.029578	-0.04	100.00
full CCSD			-2.029509		
naphthalene ^a	1	-0.906710	-0.906710	278.87	67.11
	2	-0.461387	-1.368097	-10.65	101.26
	3	0.016755	-1.351341	-0.14	100.02
	4	0.000467	-1.350875	0.15	99.98
full CCSD			-1.351119		

TABLE I: Comparison of the incremental energies with the full CCSD calculations for the a series of hydrocarbons. Structures can be found in [37].

^a domain_size_parameter=3. All calculations in the 6-31G** basis set of Pople [38, 39].

system	order i	i -th order correction [au]	$E_{\text{corr}}(i)$ [au]	error [kcal/mol]	% E_{corr}
MoF ₆ ^{a†}	1	-1.214064	-1.214064	24.62	96.87
	2	-0.043326	-1.257390	-2.57	100.33
	3	0.004799	-1.252591	0.44	99.94
	4	-0.000805	-1.253396	-0.06	100.01
exact CCSD			-1.253299		
MoCl ₆ ^{a†}	1	-0.983501	-0.983501	46.86	92.94
	2	-0.090874	-1.074375	-10.17	101.53
	3	0.021664	-1.052711	3.43	99.48
	4	-0.006767	-1.059478	-0.82	100.12
exact CCSD			-1.058176		
TiCp ₂ Cl ₂ ^{a#}	1	-1.360373	-1.360373	305.25	73.66
	2	-0.541651	-1.902023	-34.64	102.99
	3	0.067304	-1.834719	7.60	99.34
	4	-0.011983	-1.846702	0.08	99.99
exact CCSD			-1.846825		
Nb ₂ Cl ₁₀ ^{b#}	1	-1.648100	-1.648100	78.78	92.92
	2	-0.153063	-1.801164	-17.27	101.55
	3	0.035460	-1.765703	4.98	99.55
	4	-0.009367	-1.775070	-0.90	100.08
exact CCSD			-1.773638		

TABLE II: Comparison of the incremental energies with the full CCSD calculations for a series of transition metal complexes. Structures can be found in [37].^a domain_size_parameter=3, ^b domain_size_parameter=4, [†] 6-31G* basis set of Pople [38, 39], [#] cc-pVDZ basis set of Dunning [40], Mo, Ti, Nb, ECP28MWB [41, 42]

order i	i-th order correction [au]	$E_{\text{corr}}(i)$ [au]	error [kcal/mol]	% E_{corr}
domain_size_parameter=3				
connectivity parameter=2.5				
9 domains				
1	-1.360373	-1.360373	305.25	73.66
2	-0.541651	-1.902023	-34.64	102.99
3	0.067304	-1.834719	7.60	99.34
4	-0.011983	-1.846702	0.08	99.99
domain_size_parameter=4				
connectivity parameter=3.5				
8 domains				
1	-1.421872	-1.421872	266.66	76.99
2	-0.473666	-1.895537	-30.57	102.64
3	0.059733	-1.835805	6.92	99.40
4	-0.010099	-1.845904	0.58	99.95
domain_size_parameter=6				
connectivity parameter=3.5				
5 domains				
1	-1.524918	-1.524918	202.00	82.57
2	-0.346342	-1.871260	-15.33	101.32
3	0.028312	-1.842948	2.43	99.79
4	-0.004008	-1.846956	-0.08	100.01
exact CCSD		-1.846825		

TABLE III: Comparison of the convergence for different domain_size_parameter and connectivity parameter values in the case of TiCp_2Cl_2 .

density threshold	order i	i -th order correction [au]	$E_{\text{corr}}(i)$ [au]	error [kcal/mol]	% E_{corr}
0.1	1	-0.706723	-0.706723	110.00	80.13
	2	-0.189472	-0.896195	-8.90	101.61
	3	0.013845	-0.882350	-0.21	100.04
0.01	1	-0.711767	-0.711767	106.84	80.70
	2	-0.180256	-0.892022	-6.28	101.13
	3	0.015149	-0.876873	3.23	99.42
0.001	1	-0.713198	-0.713198	105.94	80.86
	2	-0.174295	-0.887493	-3.43	100.62
	3	0.005505	-0.881989	0.02	100.00
0	1	-0.713208	-0.713208	105.93	80.86
	2	-0.174253	-0.887460	-3.41	100.62
	3	0.005430	-0.882031	-0.01	100.00
	4	0.000013	-0.882018	0.00	100.00
exact CCSD			-0.882020		

TABLE IV: Effect of eq. 10 on the convergence of the incremental scheme applied to hexane. The calculations were performed in the CCSD/6-31G* basis of Pople [38, 39].

energy threshold	E_{corr}	error	number of calculations	
[au]	[au]	[kcal/mol]		
alkyne	total number of calculations		561	
	10^{-8}	-2.029578	-0.04	555
	10^{-7}	-2.029578	-0.04	506
	10^{-6}	-2.029565	-0.03	299
	10^{-5}	-2.029656	-0.09	135
	10^{-4}	-2.030857	-0.85	71
	decane	total number of calculations		385
10^{-8}		-1.582704	0.00	348
10^{-7}		-1.582700	0.00	223
10^{-6}		-1.582681	0.01	124
10^{-5}		-1.582699	0.00	65
10^{-4}		-1.583047	-0.22	45
pentaene		total number of calculations		162
	10^{-8}	-1.419069	0.01	162
	10^{-7}	-1.419069	0.01	157
	10^{-6}	-1.419067	0.01	144
	10^{-5}	-1.419134	-0.03	81
	10^{-4}	-1.419772	-0.43	43
	naphthalene	total number of calculations		98
10^{-8}		-1.350875	0.15	98
10^{-7}		-1.350875	0.15	98
10^{-6}		-1.350875	0.15	98
10^{-5}		-1.350931	0.12	85
10^{-4}		-1.351735	-0.39	51

TABLE V: Performance of the approximation in eq. 4 for a set of molecules using different energy thresholds for the summation according to eq. 20.