THE JOURNAL OF CHEMICAL PHYSICS **130**, 154107 (2009)

Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency

Francesco Aquilante,[1,a)] Laura Gagliardi,[1,b)] Thomas Bondo Pedersen,[2,c)] and Roland Lindh[2,d)]
[1]*Department of Physical Chemistry, Université de Genève-Science II, 30 quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland*
[2]*Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden*

(Received 15 January 2009; accepted 17 March 2009; published online 17 April 2009)

Cholesky decomposition of the atomic two-electron integral matrix has recently been proposed as a procedure for automated generation of auxiliary basis sets for the density fitting approximation [F. Aquilante *et al.*, J. Chem. Phys. **127**, 114107 (2007)]. In order to increase computational performance while maintaining accuracy, we propose here to reduce the number of primitive Gaussian functions of the contracted auxiliary basis functions by means of a second Cholesky decomposition. Test calculations show that this procedure is most beneficial in conjunction with highly contracted atomic orbital basis sets such as atomic natural orbitals, and that the error resulting from the second decomposition is negligible. We also demonstrate theoretically as well as computationally that the locality of the fitting coefficients can be controlled by means of the decomposition threshold even with the long-ranged Coulomb metric. Cholesky decomposition-based auxiliary basis sets are thus ideally suited for local density fitting approximations. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3116784]

I. INTRODUCTION

Recent years have witnessed a growing interest in the use of *ab initio* and density functional theory (DFT) methods for the study of large molecules. Aided by increasing computer power, accurate electronic structure calculations can now be performed to study nanoscale molecular structures. General techniques to overcome the various bottlenecks in DFT calculations have been developed such as the fast multipole method,[1,2] fast quadratures,[3] and linear scaling self-consistent field optimizers.[4–6] Most of these results, together with the exploitation of the nearsightedness of the exact exchange,[7,8] have also made possible a linear scaling formulation of Hartree–Fock (HF) theory, the cornerstone of any single-reference correlation treatment. Finally, the advent of local correlation methods[9,10] and the atomic orbital (AO) reformulation of Møller–Plesset (MP) and coupled-cluster theories[11] have virtually completed this array of developments toward linear scaling implementations of the full arsenal of quantum chemistry.

A fundamental problem shared by all these techniques is that their efficiency is downgraded when atomic basis sets of higher quality are used. In such cases, the density fitting (DF) approximation for the electron repulsion integrals (ERIs) offers a reduction in the computational cost. The DF approximation was introduced in 1973 by Whitten[12] and Baerends *et al.*[13] The technique was investigated by Feyereisen and co-workers[14,15] in a number of papers where they used the name "resolution of identity" (RI) approximation. (We will exclusively use the term DF.) In the past decade, the impact of the DF approximation has grown in quantum chemistry. Next to the fast DF algorithms for evaluating the Coulomb Fock matrix,[16–19] screening techniques designed specifically for DF approximations allow asymptotic linear scaling evaluation of the exact exchange matrix[20–22] for HF and hybrid DFT. In addition, the use of DF approximations in correlated calculations has turned out to be extremely advantageous over conventional implementations.[23–25] The success of the DF techniques is partly due to the development of analytic gradients[26,27] and Hessians,[28,29] which have paved the way to geometry optimizations and vibrational frequency calculations for molecules with hundreds of atoms.

Some limitations of the DF approximation arise from the fact that until now the paradigm for designing the auxiliary basis sets has been that of preoptimization through data fitting. This approach has produced a number of auxiliary basis sets[16,26,30,31] designed to reproduce only subsets of the ERIs with sufficient accuracy, i.e., they are biased toward specific electronic structure methods. There are some shortcomings in adopting this standard DF paradigm: (1) it is impractical to design auxiliary basis sets for all atomic basis sets in use; in particular, the present authors are not aware of standard auxiliary basis sets to perform all-electron calculations on heavy elements, or to be used in connection with atomic natural orbital (ANO) type basis sets,[32] (2) different auxiliary basis sets have to be used for nonhybrid DFT, HF and hybrid DFT, and correlated calculations, and (3) there is no easy way to improve the accuracy of an auxiliary basis set except, of course, to generate a new set from scratch.

It is interesting to note that in the early days of DF,[14,15]

[a)] Electronic mail: francesco.aquilante@unige.ch.
[b)] Electronic mail: laura.gagliardi@unige.ch.
[c)] Electronic mail: thomas.pedersen@teokem.lu.se.
[d)] Electronic mail: roland.lindh@teokem.lu.se.

© 2009 American Institute of Physics

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

focus was on fitting the AO ERIs instead of individual energy contributions in a quantum chemical calculation. It was therefore natural to start from the AO product densities in order to define auxiliary basis sets. Although this concept was basically lost in the subsequent years, Ten-no and Iwata's RI with linear combination of atomic electron distributions method[33,34] contributed to a revived interest in finding more generic designs of auxiliary basis sets.[35] With this spirit and with the additional aim at rectifying the drawbacks of the standard auxiliary basis sets, three of the present authors proposed[63] to merge the DF approximation with the Cholesky decomposition (CD) approximation of Beebe and Linderberg.[36] The latter had been around for about 30 years as the forgotten twin brother of the DF method. Although CD presented some advantages over the DF approximation, strict error control in particular, it did not become popular.[37–39] The reasons for this include computational cost, lack of a general implementation, and the lack of an analytic derivative formulation. Since its first general implementation by Koch *et al.*,[40] the CD approximation has been used for efficient *ab initio* calculations,[41–61] and the first analytic derivative formulation is now available.[62] The key ingredient for the latter was to demonstrate that the CD approximation is a particular type of DF in which the auxiliary basis set is generated from the set of AO product densities. A pivotal difference between standard CD and DF approximations is the fact that the former contains, in general, bicentric auxiliary functions in addition to atom-centered ones. This characteristic is the reason behind the full error control of the approximation, but it has some drawbacks from the point of view of computational cost, as well as in terms of analytic formulation of geometric derivatives.[62] Various one-center forms of the standard CD approximation have therefore been introduced,[62,63] which retain essentially all of the error control property of the standard CD but considerably reduce the associated computational cost.[56]

In the present work, we will focus on the one-center CD approximation known as atomic Cholesky decomposition (aCD) for generating auxiliary basis sets from the AO basis set.[63] We will define a new family of aCD auxiliary basis sets expressed in terms of a much smaller number of primitive Gaussian basis functions. The accuracy of this new hierarchy of atomic compact CD (acCD) auxiliary basis sets is shown to be as good as that of the parent aCDs. We will also demonstrate that the aCD and acCD auxiliary basis sets introduce an inherent locality in the fitting coefficients even when a long-ranged metric, such as the Coulomb metric, is used. This property, which is not shared by existing standard auxiliary basis sets, opens new opportunities for faster and reduced-scaling DF-based algorithms in quantum chemistry.

II. THEORY

A. The DF approach

The DF approach aims at approximating the AO product densities with a linear expansion on a set of auxiliary basis functions $\chi_K = |K\rangle$: $\chi_\mu\chi_\nu = |\mu\nu\rangle \approx \Sigma_K C^K_{\mu\nu}\chi_K$. The expansion coefficients are determined by minimizing the "distance" (in Hilbert space) between the fitted and the actual product densities

$$\Delta_{\mu\nu} = \left(\chi_\mu\chi_\nu - \sum_K C^K_{\mu\nu}\chi_K \middle| \hat{g} \middle| \chi_\mu\chi_\nu - \sum_K C^K_{\mu\nu}\chi_K\right), \tag{1}$$

where $\hat{g}$ defines a given positive definite metric. Minimization of this function leads to the set of linear equations

$$\sum_I G_{IK} C^I_{\mu\nu} = h^K_{\mu\nu}, \tag{2}$$

where $h^K_{\mu\nu} = \langle \mu\nu | K \rangle_g$, $G_{IK} = \langle I | K \rangle_g$, and $\langle \cdot | \cdot \rangle_g = (\cdot | \hat{g} | \cdot)$. We now restrict the discussion to the case of the Coulomb metric $\hat{g} = r_{12}^{-1}$, as this is the most accurate choice for computing electronic properties.[15,64] The value of the function corresponding to the solution of Eq. (2) is given by

$$\Delta^{\min}_{\mu\nu} = (\mu\nu|\mu\nu) - \sum_{IK} C^I_{\mu\nu}(I|K)C^K_{\mu\nu}. \tag{3}$$

Thus, $\Delta^{\min}_{\mu\nu}$ measures the error in representing the diagonal element of the ERI matrix.

Equation (3) shows that in general an externally defined auxiliary basis set does not lead to the global minimum for the DF variational condition, Eq. (1). As $\Delta_{\mu\nu}$ is nonnegative, the global minimum is *zero* and is reached only when the auxiliary basis set spans the same space as the product densities $\chi_\mu\chi_\nu$. The ERI matrix is then represented exactly by the DF approach.

B. *Ab initio* generation of DF auxiliary basis sets

1. Atomic CD auxiliary basis sets

Standard auxiliary basis sets are chosen to be Gaussian functions centered on the nuclei. The number of auxiliary basis functions thus scales linearly with system size and the computed electronic energy is a continuous function of the nuclear positions. Auxiliary basis sets have been optimized to accurately fit some specific quantities (e.g., Coulomb energy, MP2 correlation energy correction, etc.) for a range of AO basis sets by minimization of the energy errors in a series of atomic calculations.[16,20,31,65] As the optimization is performed without any direct constraint regarding the representation of the integral diagonals as measured by Eq. (3), the resulting auxiliary basis sets cannot provide a uniform DF approximation to the ERI matrix.

On the other hand, if we consider an isolated atom, the CD of its ERI matrix[36] provides an upper bound for the value of $\Delta^{\min}_{\mu\nu}$, Eq. (3),

$$\Delta^{\min}_{\mu\nu} = (\mu\nu|\mu\nu) - \sum_J (L^J_{\mu\nu})^2 \leq \tau, \tag{4}$$

where $L^J_{\mu\nu}$ are the Cholesky vectors and τ is the decomposition threshold. Very recently[63] three of the present authors have shown that the CD is equivalent to a particular type of DF approximation to the ERIs in which the auxiliary basis set is selected from the full set of AO product densities. By using the auxiliary basis set defined through each aCD for a DF approximation to the molecular ERI matrix, one achieves strict error control on the one-center and two-center

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

"Coulomb" ERIs, namely, the integrals of the types $(AA|AA)$ and $(AA|BB)$, respectively. This can be understood by considering the fact that the DF approximation is an inner projection and therefore the following Cauchy–Schwarz inequality holds

$$|\Delta_{\mu\nu,\lambda\sigma}| = \left| (\mu\nu|\lambda\sigma) - \sum_{IK} C_{\mu\nu}^{I}(I|K)C_{\lambda\sigma}^{K} \right| \le \Delta_{\mu\nu}^{1/2}\Delta_{\lambda\sigma}^{1/2}. \quad (5)$$

(Henceforth, we drop the superscript min.) From Eqs. (4) and (5), we can see that the errors in the DF approximation to the ERIs of the types $(AA|AA)$ and $(AA|BB)$ are bound by the aCD threshold τ, whereas those of the type $(AB|**)$ may be affected by larger errors. Nonetheless, the robustness of the aCD-based DF approximation has been already established.[63] As the quality of the aCD auxiliary basis set is controlled by the aCD threshold, this approach defines a hierarchy of auxiliary basis sets for accurate DF approximations. Moreover, for a given decomposition threshold, the number of auxiliary functions grows linearly with the number of AO basis functions as a consequence of the linear dependence arising in the AO product space. From a practical point of view, the aCD auxiliary basis sets need not be stored in one of the usual basis set library formats, as they are constructed on the fly and, consequently, may be applied in conjunction with any atom and AO basis set. Finally, as they are derived *ab initio* by ensuring upper bounds on the DF error function of Eq. (3), the unbiased nature of the resulting DF approximation is guaranteed.[56,63]

The raw aCD auxiliary basis set does not have the structure of complete shells and the final auxiliary basis set is generated by adding auxiliary functions to get a complete shell structure. This gives rise to a somewhat larger auxiliary basis set as compared to, e.g., the so-called 1C-CD approximation using the same decomposition threshold.[63] However, the aCD auxiliary basis set is then spherically balanced.

2. Atomic compact CD auxiliary basis sets

By construction, the aCD auxiliary basis sets can contain all possible pairs of primitives, and it is indeed the case that even with a relatively high decomposition threshold the number of primitives can often be exceedingly large. Here we will show that it is possible to apply the same type of reasoning used to define the aCD auxiliary basis sets in order to reduce the size of the primitive basis required to expand a given aCD auxiliary basis set. This is achieved in two steps. The first step involves an elimination of redundant exponent pairs in an "angular free" CD. The matrix to decompose is given by

$$T_{\alpha\beta,\gamma\delta} = \sum_{ijkl} |(\chi_{\alpha i}\chi_{\beta j}|\chi_{\gamma k}\chi_{\delta l})|, \quad (6)$$

where χ are primitive Gaussians with α, β, γ, and δ as the exponent indices and i, j, k, and l as angular indices. The summation over angular index is restricted to those of the raw aCD basis set. The surviving exponent pairs after the CD of $\mathbf{T}$ are stored as individual single exponents $\widetilde{\alpha\beta}$. These exponents form the primitive set from which the contracted acCD auxiliary basis set is derived. The second step involves the derivation of the contraction coefficients of these nonredundant exponents $c_{\kappa\lambda,\widetilde{\gamma\delta}}$ to fit in a Coulombic metric to the original contracted integrals as computed from the primitive integrals of the set of redundant exponent pairs. That is, the coefficients are derived from the expression

$$T_{\widetilde{\alpha\beta},\kappa\lambda} = \sum_{\widetilde{\gamma\delta}} c_{\kappa\lambda,\widetilde{\gamma\delta}} T_{\widetilde{\alpha\beta},\widetilde{\gamma\delta}}. \quad (7)$$

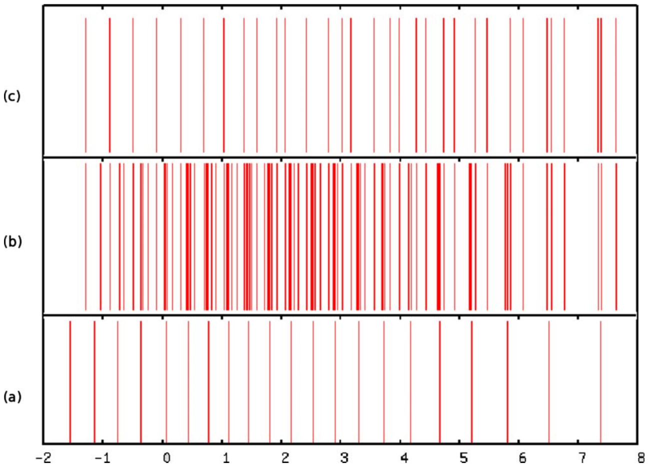

FIG. 1. (Color online) Tc atom: Progression of the primitive set in (a) ANO-RCC s-shell and corresponding (b) aCD-4 and (c) acCD-4 basis sets. The decimal logarithm of the Gaussian exponent is reported on the horizontal axes.

Figure 1 gives an idea of the degree of reduction in the size of the primitive basis for these acCD auxiliary basis sets compared to the corresponding aCD ones. We notice how the selection of the reduced primitive basis set acts in the sense of replacing the large number of "near degenerate" exponents in the crowded regions with one or few effective exponents. On the other hand, the progressions of exponents for aCD and acCD span the same range, as one would require. We shall demonstrate in what follows that despite the significant reduction in the size of the primitive basis, the acCD auxiliary basis sets exhibit virtually the same accuracy as the corresponding aCD ones.

III. RESULTS AND DISCUSSION

The acCD basis sets defined above constitute a hierarchy of auxiliary basis sets for DF approximations in exactly the same way as the original aCD. It is clear that the acCD and aCD sets must show the same accuracy as the threshold for the CD of the $\mathbf{T}$ matrix of Eq. (6) approaches zero. In all examples discussed here, we have set the threshold for the CD of the $\mathbf{T}$ matrix to be one order of magnitude tighter than the corresponding aCD threshold. We will then document the accuracy of the acCD auxiliary basis sets only for relatively loose decomposition thresholds by comparison with the aCD results and in some cases with conventional integral calculations. A more thorough investigation of accuracy can be found in Ref. 66. The remaining part of this section will focus on computational efficiency, and in particular on analyzing an interesting locality of DF approximations based on various CD formulations.

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

TABLE I. Accuracy of the acCD-4 B3LYP energy (in units of CD threshold $\tau=10^{-4}E_h$) with respect to conventional and aCD-4 reference values. N_b is the number of AO basis functions, M_c is the number of contracted auxiliary basis functions, $\Delta_{\max}$ is the largest error in the diagonal integrals, and $\zeta\%$ is the percentage of the number of primitive auxiliary functions in acCD-4 compared to aCD-4.

Basis set	Molecule[a]	N_b	M_c	$(E_{\text{acCD-4}}-E_{\text{conv}})/\tau$	$(E_{\text{acCD-4}}-E_{\text{aCD-4}})/\tau$	$\Delta_{\max}/\tau$	$\zeta\%$
ANO-VDZP	$(3\text{-btph})_1$	190	917	−0.510	−0.077	2.14	35.6
	$(3\text{-btph})_2$	370	1804	−0.964	−0.157	2.12	35.4
	$(3\text{-btph})_3$	550	2691	−1.025	−0.217	1.95	35.3
	$(3\text{-btph})_4$	730	3578	−1.136	−0.293	2.01	35.3
	$(3\text{-btph})_8$	1450	7126		−0.561	2.08	35.3
	$(3\text{-btph})_{10}$	1810	8900		−0.637	2.05	35.3
	DDA	896	4762	18.449	−0.341	2.33	32.3
	DTTC	464	1858	−18.356	0.616	1.03	27.7
	DTTC[b]	356	1426	−294.371	0.618	2.29	23.6
ANO-VTZP	$(3\text{-btph})_1$	442	3329	0.033	0.008	1.25	41.2
	$(3\text{-btph})_2$	856	6512	1.060	0.022	1.11	41.0
	$(3\text{-btph})_3$	1270	9695		0.052	1.20	41.0

[a] 3-btph=3-butylthiophene, DDA=dinuclear diazene iron complex, DTTC=$Tc_2Cl_4(PMe)_3$.

[b] Polarization (*p*-functions) removed from ANO-VDZP basis set on all hydrogen atoms.

All calculations are performed with a development version of the MOLCAS quantum chemistry software.[67,68] To expedite the evaluation of the exchange Fock matrix, we employ the local-K (LK) screening[69] with localized Cholesky molecular orbitals.[70]

A. Accuracy

The aCD basis sets analyzed in the present work have previously[56,63] been associated with the acronym aCD-n^*, where n is related to the atomic CD threshold as $\tau=10^{-n}$. In the present work, however, we will use the unstarred notation, which in previous papers[56,63] signified removal of the highest angular momentum auxiliary functions (i.e., we keep the highest angular momentum functions in this work). In Refs. 56 and 63 we assessed the accuracy of the aCD-based DF approximation. Here we will briefly focus on the accuracy of the acCD basis sets. A thorough benchmarking of all the CD-based *ab initio* DF approximations can be found in Ref. 66.

We have summarized in Tables I and II a set of results on the accuracy of acCD-4. As explained before, there is no need to investigate the accuracy of these auxiliary basis sets for tighter decomposition thresholds. The choice $\tau=10^{-4}$ implies an accuracy comparable to standard auxiliary basis sets.[63,66] The test systems are primarily polymers of 3-butylthiophene (see Fig. 2), which are of interest as donors in organic tandem solar cells.[71] In addition, we have also verified the accuracy of the acCD-4 auxiliary basis set on systems containing transition metal atoms [the dinuclear diazene (DDA) compound in Ref. 72] and heavy elements (the ditechnetium tetrachloride trimethylphosphine compound synthetized by Sattelberger.[73]

Table I shows that the error in the computed B3LYP energy introduced by the DF approximation that uses acCD-4 is typically on the order of 10^{-4}. There is a scaling of the error with the size (number of electrons) of the system but it is not severe: in going from 1 to 4 units of 3-butylthiophenes, the absolute deviation in total energy increases only by about a factor of 2. The only odd result in

TABLE II. Accuracy of the acCD-4 HF and MP2 energy (in units of $10^{-4}E_h$) with respect to conventional integral-direct and aCD-4 reference values. Values in parenthesis represent the MP2 energy deviation. N_b is the number of AO basis functions, M_c the number of contracted auxiliary basis functions, $\Delta_{\max}$ is the largest error in the diagonal integrals, and $\zeta\%$ is the percentage of the number of primitive auxiliary functions in acCD-4 compared to aCD-4.

Basis set	Molecule[a]	N_b	M_c	$(E_{\text{acCD-4}}-E_{\text{conv}})/\tau$	$(E_{\text{acCD-4}}-E_{\text{aCD-4}})/\tau$	$\Delta_{\max}/\tau$	$\zeta\%$
SVP	$(3\text{-btph})_1$	190	864	2.7 (3.3)	−0.13 (−0.13)	8.41	81.8
	$(3\text{-btph})_2$	370	1698	8.9 (9.6)	−0.21 (−0.21)	8.52	81.7
	$(3\text{-btph})_3$	550	2532	16.5 (17.6)	−0.33 (−0.32)	7.93	81.7
	$(3\text{-btph})_4$	730	3366	26.8 (29.2)	−0.43 (−0.42)	8.09	81.7
	$(3\text{-btph})_5$	910	4200	39.8 (42.7)	−0.58 (−0.57)	8.17	81.7
cc-pVDZ	$(3\text{-btph})_1$	190	768	−2.5 (−2.9)	−0.41 (−0.41)	7.86	58.5
	$(3\text{-btph})_2$	370	1508	−7.5 (−7.9)	−0.79 (−0.78)	8.14	58.2
	$(3\text{-btph})_3$	550	2248	−15.3 (−16.1)	−1.09 (−1.07)	8.02	58.2
ANO-VDZP	$(3\text{-btph})_1$	190	917	1.1 (1.5)	−0.008 (0.02)	2.14	35.6
	$(3\text{-btph})_2$	370	1804	2.3 (2.7)	−0.008 (0.05)	2.12	35.4
	$(3\text{-btph})_3$	550	2691	3.4 (4.0)	−0.009 (0.08)	1.95	35.3

[a] 3-btph=3-butylthiophene.

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

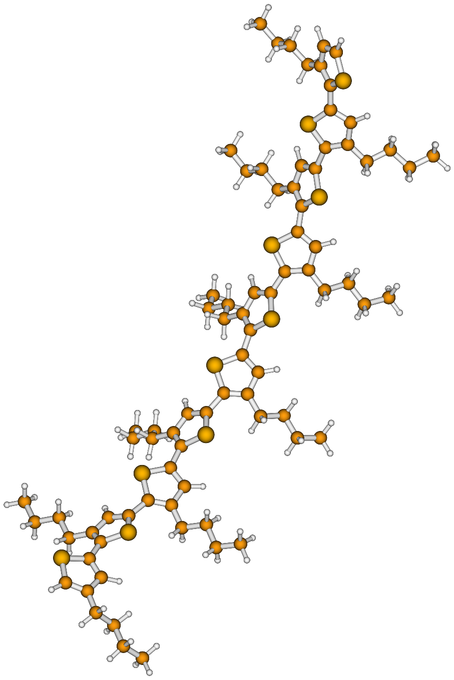

FIG. 2. (Color online) Schematic representation of the structure of 10-unit 3-butylthiophene polymer.

Table I is observed for the energy of the ditechnetium compound computed by replacing the ANO-VDZP basis set on hydrogen atoms with the ANO-VDZ basis set: the inaccuracy of the DF approximation is an order of magnitude larger than for the full ANO-VDZP calculation. The ANO-VDZ basis set for hydrogen contains only (two) *s*-functions, thus it lacks the single polarization *p*-function of ANO-VDZP. The acCD and aCD auxiliary basis sets for hydrogen ANO-VDZ basis set are therefore composed of only *s*-functions. This is the source of the manifested inaccuracy in total energies whenever using auxiliary basis sets generated from AO basis sets containing only *s*-functions. We do not currently know if this is a real problem in computing energy differences, and we simply suggest to avoid using AO basis sets of this type. We also note in passing that standard auxiliary basis sets circumvent this problem by using a somewhat larger (polarized) set of auxiliary functions for, e.g., double-ζ AO basis sets on hydrogen atoms.

More important for the purpose of the present work is the fact that the error introduced by the acCD approximation with respect to the corresponding aCD is negligible, typically an order of magnitude smaller than the intrinsic error of the parent aCD set. The acCD error becomes smaller when we use a basis set of higher quality, e.g., in going from ANO-VDZP to ANO-VTZP, and similarly the accuracy of the DF approximation is generally improved. The reduction in the number of primitive auxiliary functions depends on the nature of the AO basis sets: the largest gain is obtained for generally contracted basis sets such as the ANOs, whereas segmented basis sets exhibit less linear dependence in the primitive product basis. A small increase in the number of primitives for the acCD basis sets is observed in going from double to triple-zeta quality AO basis sets. Finally, the unbiased nature of the aCD and acCD auxiliary basis sets is demonstrated by the fact that the accuracy of the DF approximation is nearly constant across different methods, as shown in Table II, for the errors in HF and MP2 energies. The errors for the B3LYP calculations in Table I are smaller than those corresponding to HF and MP2 with basis sets of similar degree of polarization and reported in Table II. This difference is not attributable to any bias toward DFT but is peculiar to the ANOs, as these are by construction AO basis sets of higher quality than segmented basis sets with same degree of polarization. The intrinsic high quality of ANOs is then reflected in the corresponding CD-based auxiliary basis sets, giving rise to a higher accuracy of the DF approximation. With the same ANO-VDZP basis set, in fact, we have basically the same accuracy for B3LYP in Table I and for HF and MP2 in Table II; thus the auxiliary basis set produces an unbiased DF approximation to the ERIs.

Before leaving this section on accuracy, we would like to point out a natural way to achieve an increased level of error control in the acCD-based DF approximations to the ERIs. As mentioned, Eq. (5) contains both the key to the accuracy of the acCD and aCD and also the reason for the limited control of the error in the DF approximation. The latter consists of the lack of strict error control of the two-center diagonal ERIs when employing only one-center atomic auxiliary basis functions: $\Delta_{\max}$, the largest error in the diagonal integrals, is therefore not reduced significantly by lowering the aCD threshold.[63] We can easily be convinced that the best auxiliary basis set for a DF approximation of the two-center product densities can be derived by CD of the "exchange" integral matrix $(AB|AB)$ for each interacting atom pair AB (i.e., with overlapping AOs). This approach would allow for full error control in the ERI representation, as achieved by standard full CD, but at a much more competitive cost. Although very appealing for energy calculations, introducing a geometry dependence in the auxiliary basis set is a source of difficulties for derivative calculations. We intend to report soon on our work in seeking for general solutions to further improve the robustness and applicability of this type of *ab initio* DF approximations.

B. Efficiency

We start by the obvious statement that a lower number of primitive functions in the acCDs compared to aCDs implies that the computational cost of evaluating the two- and three-center integrals needed to set up the DF equations is reduced. All subsequent steps are completely unaffected by the difference in the size of the primitive basis. Consequently, there is no guarantee that significant speedups will be observed for the overall calculation. As an example, we report in Table III the timings for computing the DF vectors (see Ref. 56 for the definition) for the DDA compound in Ref. 72. The reduction in timings for the acCD is significant only for the calculation performed with the ANO-VDZP basis set. In this case the acCD basis set requires only about 30% of the aCD primitive basis but shows no loss in accuracy, as measured by the largest error in the diagonal integrals $\Delta_{\max}$. It should also be noted that the speedup of any of the DF-based calculations compared to conventional integral-direct algorithms is sub-

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

TABLE III. DDA iron complex. Timings for the evaluation of the DF vectors (see Ref. 56 for definitions) with different AO and corresponding auxiliary basis sets. (BS1 indicates ANO-VDZP on all atoms and BS2 indicates cc-pVTZ on Fe, S, and N and cc-pVDZ on C and H.) The last column reports the largest error in the diagonal integrals. N_b is the number of AO basis functions; M_c and M_p are the numbers of contracted and primitive auxiliary basis functions, respectively.

			M_p		CPU time/min		$\Delta_{\max}$	
	N_b	M_c	aCD-4	acCD-4	aCD-4	acCD-4	aCD-4	acCD-4
BS1	896	4762	101 026	32 700	265	167	2.3×10^{-4}	2.3×10^{-4}
BS2	1156	6066	36 054	16 858	456	455	7.3×10^{-4}	7.3×10^{-4}

stantial; as an example, for the B3LYP/ANO-VTZP calculations in Table I, we register a wall clock time speedup of nearly a factor of 22 already for the single unit of 3-butylthiophene.

These aspects of the efficiency introduced by the acCD auxiliary basis sets compared to aCD will not be investigated further here. Instead, we focus on an important advantage of acCD and aCD over standard auxiliary basis sets: the fitting coefficients show an inherent locality, which can be controlled by the decomposition threshold, even with the Coulomb metric. The use of the Coulomb metric in DF approximations is the choice of reference, as it is unanimously recognized as the most accurate one.[14,15] However, this choice introduces a slow decay behavior of the fitting coefficients C of Eq. (2). Jung *et al.*[64] demonstrated that the coefficients fall off as the inverse distance between the auxiliary function and the fitted on-site product density when the Coulomb metric is employed. Using a simple model system, Jung *et al.*[64] explained this behavior by contaminants in the fitting coefficients arising from the incompleteness of the on-site auxiliary basis set, i.e., off-site auxiliary basis functions are needed to produce an accurate fit. This would discourage any attempt to use screening of the fitting coefficient to reduce the cost of the linear algebra involving these quantities and leads to divergences when trying to extend the DF approximation to systems with periodic boundary conditions (bulk material). The solution proposed by Jung *et al.*[64] is to introduce the attenuated Coulomb metric for the fitting in order to maintain the accuracy of the Coulomb metric at short distances, while allowing for a fast decay of the fitting coefficients. The long-ranged behavior of the fitting coefficients is also the source of another limitation of present DF approximations, namely, the so-called "extensivity errors" arising in DF-based explicitly correlated theories, at least when slowly decaying correlation factors are used.[74] Last but not least, the presence of contaminants gives rise to nonvanishing geometrical derivatives for one-center integrals, an unphysical characteristic of DF approximations which has not been analyzed yet but which could have interesting implications.

With this state of affairs, it is certainly of interest to discover that by employing CD-based auxiliary basis sets instead of externally defined ones, these problems are significantly reduced. To clarify the matter, we need to specify that the fitting coefficients of interest here are those of the on-site type $C^{K}_{\mu_A\nu_A}$, i.e., those corresponding to one-center product densities. If computed employing a short-ranged metric such as the overlap metric, these matrices will have non-negligible elements only for auxiliary basis functions K centered on A and atoms in the proximity of A. This is not the case when working in the Coulomb metric, as the matrices tend to be rather dense even for relatively large systems. In the supplementary material,[75] we report a pictorial analysis of the locality of the fitting coefficients for different auxiliary basis sets. In particular, we observe that by increasing the accuracy of the acCD, i.e., lowering the decomposition threshold, we gain a substantial degree of locality.

Before going into the details of the mathematical justification of this result, it is instructive to look at the physical reason behind it. Standard auxiliary basis sets have been optimized by fixing the number of functions and trying to reproduce some specific energies in atomic calculations. There is therefore no guarantee that such basis set is sufficiently accurate to also represent locally (on each atom) the various product densities. The DF procedure recognizes such deficiencies of the auxiliary basis set and compensates them through the metric, i.e., through very delocalized fitting coefficients. The acCD, as well as any other type of CD-based auxiliary basis set, does not have such strong need to borrow extra functions from neighboring and distant atoms. These *ab initio* auxiliary basis sets are constructed to be sufficient for describing the AO product densities on each atom, and they approach completeness by means of lowering the CD threshold. We should also stress another point. As the ultimate source of the long-ranged behavior of the fitting coefficients is the Coulomb metric, we cannot completely change this decay behavior by simply changing the nature of the auxiliary basis set, but it is also clear that the effects of these tails can be reduced to a level where they are practically negligible.

We now attempt a more rigorous proof of the behavior observed for the fitting coefficients of CD-based auxiliary basis sets. Let us consider the effect on the fitting of the product densities of atom A in the presence of auxiliary functions on atom B. We use the notation (C^a_{AA}, C^b_{AA}) to distinguish the component along the intrinsic auxiliary functions (C^a_{AA}) from the contaminants (C^b_{AA}), where the lower case letters indicate auxiliary basis functions centered on the corresponding atom. The DF equations in this case read

$$(C^a_{AA} \;\; C^b_{AA})\begin{pmatrix} G_{aa} & G_{ab} \\ G_{ba} & G_{bb} \end{pmatrix} = (h^a_{AA} \;\; h^b_{AA}), \tag{8}$$

which may be solved to give

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

$$C^a_{AA} = C^{a(0)}_{AA} - C^b_{AA} G_{ba} G^{-1}_{aa} \tag{9}$$

$$C^b_{AA} = [h^b_{AA} - C^{a(0)}_{AA} G_{ab}](G_{bb} - G_{ba} G^{-1}_{aa} G_{ab})^{-1}, \tag{10}$$

where

$$C^{a(0)}_{AA} = h^a_{AA} G^{-1}_{aa} \tag{11}$$

is the solution in the absence of auxiliary functions on atom B. Equations (9) and (10) are valid for any type of auxiliary basis set and any choice of metric. We note that they reproduce the solutions derived by Jung *et al.*[64] for a simple model system. For large distances, the contaminant C^b_{AA} decays as r^{-1} with the Coulomb metric.[64] To overcome this long-range problem, Jung *et al.*[64] investigated the overlap and the attenuated Coulomb metrics. At large distances, these metrics imply $h^b_{AA} \to 0$, $G_{ab} \to 0$, and $G_{ba} \to 0$ at a rate faster than r^{-1}. Consequently, in this limit, $C^b_{AA} \to 0$ and $C^a_{AA} \to C^{a(0)}_{AA}$.

An alternative solution is to make the quantity in square brackets in Eq. (10) vanish. This can be done in a conceptually simple way: make sure that the auxiliary basis on each center is sufficient to fit the product densities on the same center. In this case there will be no need to "borrow" auxiliary functions from other centers. Now, consider the absolute value of an element of the matrix in square brackets,

$$\left| (\mu_A \nu_A | K_B) - \sum_{I_A} C^{I_A(0)}_{\mu_A \nu_A} (I_A | K_B) \right|$$
$$= |(d_{\mu_A \nu_A} | K_B)| \le \sqrt{(d_{\mu_A \nu_A} | d_{\mu_A \nu_A})(K_B | K_B)}, \tag{12}$$

where $|d_{\mu_A \nu_A}) = |\mu_A \nu_A) - \Sigma_{I_A} C^{I_A(0)}_{\mu_A \nu_A} | I_A)$. While standard DF approximations do not ensure a value for $(d_{\mu_A \nu_A} | d_{\mu_A \nu_A})$ sufficiently close to zero (global minimum), the CD-based sets are special in the sense that they define an upper bound to this quantity. As the CD is performed for each atomic ERI block, $(d_{\mu_A \nu_A} | d_{\mu_A \nu_A}) \le \tau$, and hence

$$\left| (\mu_A \nu_A | K_B) - \sum_{I_A} C^{I_A(0)}_{\mu_A \nu_A} (I_A | K_B) \right| \le \sqrt{\tau (K_B | K_B)}. \tag{13}$$

By decreasing the decomposition threshold, the contaminant can be made arbitrarily small and the intrinsic part of the fitting coefficients approaches the value of the isolated atom, Eq. (11). For an exact decomposition ($\tau = 0$), the contaminant vanishes regardless of the distance between A and B.

As for CD-based auxiliary basis sets the integral matrices $\mathbf{G}$ and $\mathbf{h}$ are subsets of the ERI matrix, in the limit of $\tau \to 0$ discussed above, we can then write

$$h^b_{AA} \approx C^{a(0)}_{AA} G_{ab} C^b_b + \Delta_{AA,b}. \tag{14}$$

Moreover, $C^b_b = h^b_b G^{-1}_{bb} = G_{bb} G^{-1}_{bb} = I$, the identity matrix, and so we specialize Eq. (10) to

$$C^b_{AA} \approx \Delta_{AA,b} (G_{bb} - G_{ba} G^{-1}_{aa} G_{ab})^{-1}. \tag{15}$$

Equation (15) relates directly the contaminant component of CD-based DF coefficients to the corresponding value of the remainder Δ of Eq. (5). In particular, we notice that the absolute value of $\Delta_{AA,b}$ is bound by the decomposition threshold τ for any CD-based DF approximation. We should point out that in deriving Eq. (15) we have assumed that the acCD basis is a subset of the original AO product densities. It must be noted, however, that there are some manipulations (shell completion, fitting to a smaller set of primitives, and fitting of contractions) that render this assumption not completely legitimate for aCD and acCD. It is however valid in the $\tau \to 0$ limit and all our results show trends in perfect agreement with Eq. (15).

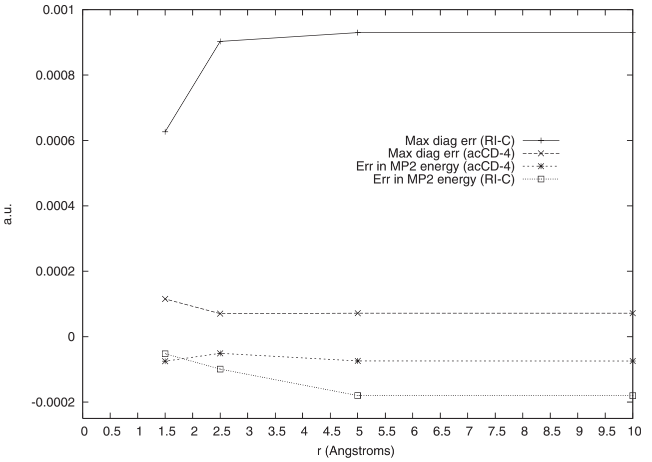

FIG. 3. Benzene dimer: maximum error in the diagonal integrals and error in the computed MP2 energy correction at various distances (cc-pVDZ basis set; 1*s* orbitals on the carbon atoms left uncorrelated). The values for the maximum error in the diagonals have been scaled by a factor of 0.1 for a better visualization of the curves.

We emphasize that this type of locality achieved not only in the long-range regime but at any range of distances increases with the accuracy of the CD-based auxiliary basis set. Figure 3 displays the geometry dependence of the largest error in the diagonal integrals for DF approximation based on RI-C (Refs. 31, 76, and 77) and acCD-4. The latter shows a less pronounced dependence on the geometry of the dimer, especially in the repulsive region. In any case, the maximum diagonal error for RI-C is an order of magnitude larger than for acCD-4, and this happens while the locality of the fitting coefficient is almost invariably much higher for acCD-4 than for RI-C. The corresponding error in computing the MP2 correction to the HF energy is also typically smaller for acCD-4 than for RI-C. The latter has been optimized specifically to minimize this error and therefore gives a higher accuracy than suggested by the largest error in the diagonal ERIs. However, we note that the accuracy of RI-C is more geometry dependent than that of acCD-4 in accordance with the behavior of the diagonal error statistics.

The largest error in diagonal integrals with different auxiliary basis sets is shown for the benzene dimer in Fig. 4. For RI-C, the largest error occurs for a one-center diagonal, implying a rather large contaminant in the fitting coefficients. For acCD, on the other hand, the one-center diagonal errors decrease with the decomposition threshold and therefore the contaminant approaches zero. The largest errors occur for the two-center diagonals for acCD and the value is almost independent of decomposition threshold as expected.

In Fig. 5, the decay behavior of the fitting coefficients for aCDs and RI-C is analyzed in more detail for the benzene-He dimer. The AO basis on He is cc-pVTZ, whereas cc-pVDZ is used for the benzene molecule. The value of the largest contaminant fitting coefficient on He atom due to the

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

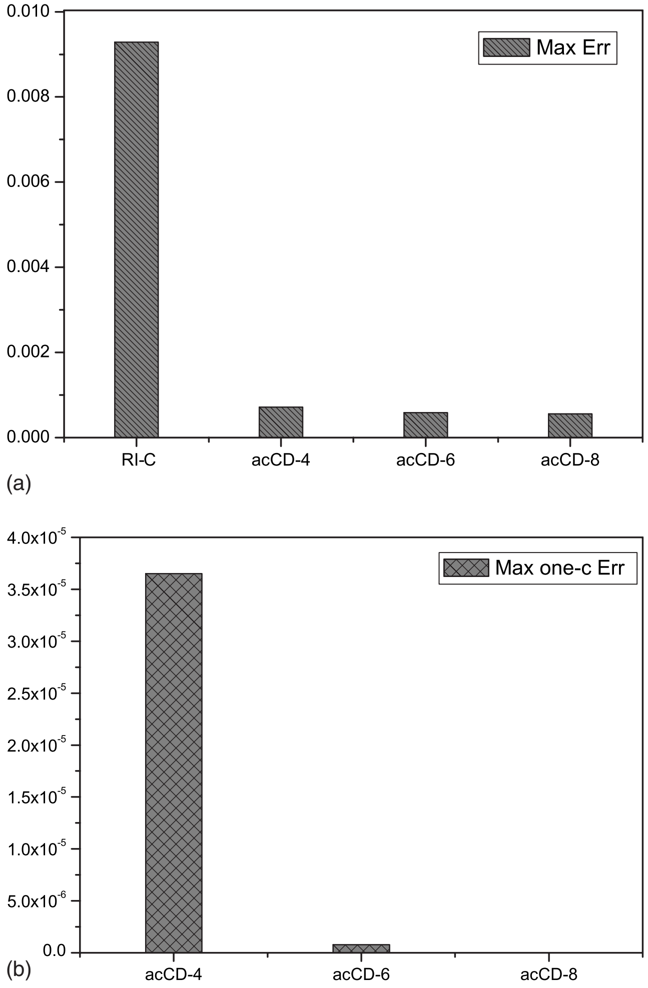

FIG. 4. Benzene dimer: maximum error in the (a) diagonal and (b) one-center diagonal integrals at $r=5.0$ Å (cc-pVDZ basis set). The values for RI-C and acCD-8 in (b) are 9.29×10^{-3} and 8.69×10^{-9}, respectively. Notice that for RI-C the values in (a) and (b) coincide.

auxiliary basis functions on benzene is shown in Fig. 5 as a function of the distance between the two subsystems. We observe the same slow decay of the contaminants with distance as reported by Jung *et al.*[64] but with the significant difference that with aCD, one can drastically reduce (at any distance) the magnitude of the contaminants by lowering the decomposition threshold. We see for instance that aCD-8 gives already a substantial reduction (about four orders of magnitude) of the contaminants, as expected from Eq. (15) and the diagonal statistics of Fig. 4(b).

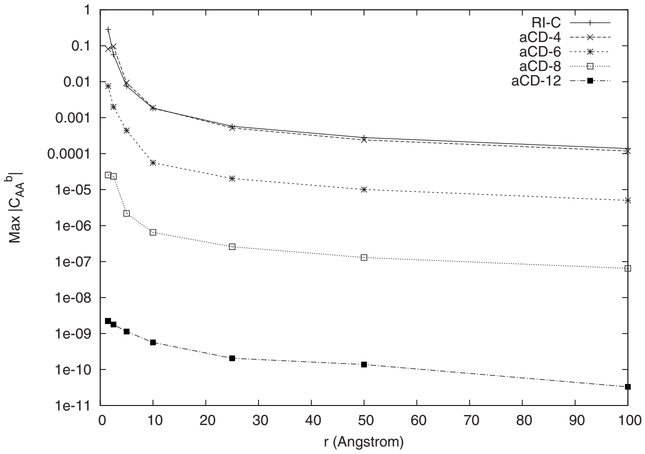

FIG. 5. He–C_6H_6 system: decay of the largest contaminant of the fitting coefficients on He atom as a function of the distance r between He atom and benzene molecule. The AO basis set used for He is cc-pVTZ, whereas a cc-pVDZ basis set is used for the atoms of the benzene molecule. Logarithmic scale is employed for the vertical axes.

We conclude by pointing out that although the data presented here are limited to a few molecular systems, the results are general, as the locality of this type of auxiliary basis sets stems from the way in which they are constructed and not on the nature of the AO basis set (e.g., if it includes diffuse functions or high angular moment components). The decay behavior of the fitting coefficients may be improved further by using a short-ranged metric, and it is reasonable to expect that the CD-based auxiliary basis sets are accurate also in this case. Finally, we have mentioned that employing auxiliary basis sets from CDs of the exchange integral matrix $(AB|AB)$ for each interacting atom pair AB allows DF approximation to the two-center product densities with full error control. The contaminants of $C_{\mu_A\nu_B}$ from a third atom can be avoided if the threshold for the CD of the $(AB|AB)$ matrix is sufficiently tight.

IV. CONCLUSIONS

The aCD approximation[63] to the AO ERIs can be used to generate auxiliary basis sets for DF approximations. This family of *ab initio* auxiliary basis sets is of particular interest because the accuracy of the subsequent DF approximation is unbiased toward any quantum chemical method, is nearly independent of geometry, and the error can be smoothly controlled by varying the decomposition threshold. A drawback of the aCDs compared to standard auxiliary basis sets is the large number of primitive functions. We have investigated here the possibility to reduce this number by exploiting the linear dependence in an atomic ERI matrix constructed in the primitive basis. This is a natural extension of the principle underlying the aCD auxiliary basis sets. Our results indicate that these new auxiliary basis sets, named acCD, achieve virtually the same level of accuracy as the corresponding aCD sets, but require only a fraction of the primitive space needed by the latter. In general, acCDs should be preferred to aCDs for standard decomposition thresholds ($\sim10^{-4}$), but not for very high accuracy, where the reduction in computational cost is negligible.

The present study has also highlighted a hitherto unexplored property of the CD-based auxiliary basis sets. This concerns the locality of the fitting coefficients obtained in the subsequent DF approximation, in particular, those coefficients corresponding to one-center AO product densities. It is shown that the locality is a result of the nature of the CD-based auxiliary basis sets and the need for non-Coulomb metrics is reduced if not entirely removed. The degree of locality can simply be increased by reducing the decomposition threshold, thus increasing the completeness of the auxiliary basis set. We envisage a great potential for reduced-scaling DF-based formulation of quantum chemical methods

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

by exploiting this property. Thus these *ab initio* DF schemes may rectify a number of situations where standard auxiliary basis sets cannot be used in connection with long-ranged metrics.

ACKNOWLEDGMENTS

Part of this work was carried out under the HPC-EUROPA project (Grant No. RII3-CT-2003-506079) with the support of the European Community-Research Infrastructure Action of the FP6 “Structuring the European Research Area” Programme. Funding from the Swiss National Science Foundation (SNF) (Grant No. 200020-120007), and Swedish Research Council (VR) are also acknowledged.

[1] M. C. Strain, G. E. Scuseria, and M. J. Frisch, Science **271**, 51 (1996).

[2] Y. Shao, C. A. White, and M. Head-Gordon, J. Chem. Phys. **114**, 6572 (2001).

[3] R. Lindh, P.-Å. Malmqvist, and L. Gagliardi, Theor. Chem. Acc. **106**, 178 (2001).

[4] J. M. Millam and G. E. Scuseria, J. Chem. Phys. **106**, 5569 (1997).

[5] W. Liang, Y. Zhao, and M. Head-Gordon, J. Chem. Phys. **123**, 194106 (2005).

[6] S. Høst, J. Olsen, B. Jansik, L. Thøgersen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **129**, 124106 (2008).

[7] E. Schwegler, M. Challacombe, and M. Head-Gordon, J. Chem. Phys. **106**, 9708 (1997).

[8] C. Ochsenfeld, C. A. White, and M. Head-Gordon, J. Chem. Phys. **109**, 1663 (1998).

[9] C. Hampel and H. J. Werner, J. Chem. Phys. **104**, 6286 (1996).

[10] M. Schutz and H. J. Werner, J. Chem. Phys. **114**, 661 (2001).

[11] G. E. Scuseria and P. Y. Ayala, J. Chem. Phys. **111**, 8330 (1999).

[12] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).

[13] E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).

[14] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[15] O. Vahtras, J. Almlöf, and M. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).

[16] K. Eichkorn, O. Treutler, H. Öm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **240**, 283 (1995).

[17] F. R. Manby, P. J. Knowles, and A. W. Lloyd, J. Chem. Phys. **115**, 9144 (2001).

[18] M. Sierka, A. Hogekamp, and R. Ahlrichs, J. Chem. Phys. **118**, 9136 (2003).

[19] A. Sodt, J. E. Subotnik, and M. Head-Gordon, J. Chem. Phys. **125**, 194109 (2006).

[20] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).

[21] R. Polly, H. J. Werner, F. R. Manby, and P. J. Knowles, Mol. Phys. **102**, 2311 (2004).

[22] F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007).

[23] A. P. Rendell and T. J. Lee, J. Chem. Phys. **101**, 400 (1994).

[24] C. Hättig and F. Weigend, J. Chem. Phys. **113**, 5154 (2000).

[25] H. J. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).

[26] F. Weigend and M. Häser, Theor. Chim. Acta **97**, 331 (1997).

[27] D. Rappoport and F. Furche, J. Chem. Phys. **122**, 064105 (2005).

[28] P. Deglmann, F. Furche, and R. Ahlrichs, Chem. Phys. Lett. **362**, 511 (2002).

[29] P. Deglmann, K. May, F. Furche, and R. Ahlrichs, Chem. Phys. Lett. **384**, 103 (2004).

[30] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, Theor. Chim. Acta **97**, 119 (1997).

[31] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).

[32] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark, Chem. Phys. Lett. **409**, 295 (2005).

[33] S. Ten-no and S. Iwata, Chem. Phys. Lett. **240**, 578 (1995).

[34] S. Ten-no and S. Iwata, J. Chem. Phys. **105**, 3604 (1996).

[35] R. Yang, A. P. Rendell, and M. J. Frisch, J. Chem. Phys. **127**, 074102 (2007).

[36] N. H. F. Beebe and J. Linderberg, Int. J. Quantum Chem. **12**, 683 (1977).

[37] I. Røeggen and E. Wisløff-Nielsen, Chem. Phys. Lett. **132**, 154 (1986).

[38] D. W. O’Neal and J. Simons, Int. J. Quantum Chem. **36**, 673 (1989).

[39] S. Wilson, Comput. Phys. Commun. **58**, 71 (1990).

[40] H. Koch, A. Sánchez de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).

[41] T. B. Pedersen, A. M. J. Sánchez de Merás, and H. Koch, J. Chem. Phys. **120**, 8887 (2004).

[42] T. B. Pedersen, H. Koch, L. Boman, and A. M. J. Sánchez de Merás, Chem. Phys. Lett. **393**, 319 (2004).

[43] I. García Cuesta, T. B. Pedersen, H. Koch, and A. M. J. Sánchez de Merás, Chem. Phys. Lett. **390**, 170 (2004).

[44] I. García Cuesta, T. B. Pedersen, H. Koch, and A. Sánchez de Merás, ChemPhysChem **7**, 2503 (2006).

[45] I. Røeggen, J. Chem. Phys. **124**, 184502 (2006).

[46] I. Røeggen, J. Chem. Phys. **126**, 204303 (2007).

[47] A. Öhrn and F. Aquilante, Phys. Chem. Chem. Phys. **9**, 470 (2007).

[48] B. Fernández, T. B. Pedersen, A. Sánchez de Merás, and H. Koch, Chem. Phys. Lett. **441**, 332 (2007).

[49] F. Aquilante and T. B. Pedersen, Chem. Phys. Lett. **449**, 354 (2007).

[50] D. Yamaki, H. Koch, and S. Ten-no, J. Chem. Phys. **127**, 144104 (2007).

[51] I. García Cuesta, J. S. Marín, T. B. Pedersen, H. Koch, and A. M. J. Sánchez de Merás, Phys. Chem. Chem. Phys. **10**, 361 (2008).

[52] F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, and B. O. Roos, J. Chem. Theory Comput. **4**, 694 (2008).

[53] K. Pierloot and S. Vancoillie, J. Chem. Phys. **128**, 034104 (2008).

[54] M. Radon and K. Pierloot, J. Phys. Chem. A **112**, 11824 (2008).

[55] I. Røeggen and T. Johansen, J. Chem. Phys. **128**, 194107 (2008).

[56] F. Aquilante, T. B. Pedersen, B. O. Roos, A. Sánchez de Merás, and H. Koch, J. Chem. Phys. **129**, 024113 (2008).

[57] L. Boman, H. Koch, and A. Sánchez de Merás, J. Chem. Phys. **129**, 134107 (2008).

[58] M. Pitoňák, P. Neogrády, J. Řezáč, P. Jurečka, M. Urban, and P. Hobza, J. Chem. Theory Comput. **4**, 1829 (2008).

[59] M. Pitoňák, K. E. Riley, P. Neogrády, and P. Hobza, ChemPhysChem **9**, 1636 (2008).

[60] K. Pluháčková, S. Grimme, and P. Hobza, J. Phys. Chem. A **112**, 12469 (2008).

[61] T. B. Pedersen, J. Kongsted, T. D. Crawford, and K. Ruud, J. Chem. Phys. **130**, 034310 (2009).

[62] F. Aquilante, R. Lindh, and T. B. Pedersen, J. Chem. Phys. **129**, 034106 (2008).

[63] F. Aquilante, R. Lindh, and T. B. Pedersen, J. Chem. Phys. **127**, 114107 (2007).

[64] Y. Jung, A. Sodt, P. M. W. Gill, and M. Head-Gordon, Proc. Natl. Acad. Sci. U.S.A. **102**, 6692 (2005).

[65] D. E. Bernholdt and R. J. Harrison, J. Chem. Phys. **109**, 1593 (1998).

[66] J. Boström, F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Theory Comput. (to be published).

[67] MOLCAS 7, University of Lund, Sweden, http://www.teokem.lu.se/molcas/, 2007.

[68] G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, and L. Seijo, Comput. Mater. Sci. **28**, 222 (2003).

[69] F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007).

[70] F. Aquilante, T. B. Pedersen, A. Sánchez de Merás, and H. Koch, J. Chem. Phys. **125**, 174101 (2006).

[71] G. Dennler, M. C. Scharber, T. Ameri, P. Denk, K. Forberich, C. Waldauf, and C. J. Brabec, Adv. Mater. (Weinheim, Ger.) **20**, 579 (2008).

[72] M. Reiher, O. Salomon, D. Sellmann, and B. A. Hess, Chem.-Eur. J. **7**, 5195 (2001).

[73] A. Sattelberger, personal communication (2008).

[74] H. J. Werner and F. R. Manby, J. Chem. Phys. **124**, 054114 (2006).

[75] See EPAPS Document No. E-JCPSA6-130-012916 for a graphical analysis of the sparsity of the fitting coefficient matrices. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.

[76] C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).

[77] J. G. Hill and J. A. Platts, J. Chem. Phys. **128**, 044104 (2008).

Downloaded 21 Apr 2009 to 129.194.8.73. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp