JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Analytical Gradients of Hartree−Fock Exchange with Density Fitting Approximations

Jonas Boström,[†] Francesco Aquilante,*[,‡,§] Thomas Bondo Pedersen,[∥] and Roland Lindh[§]

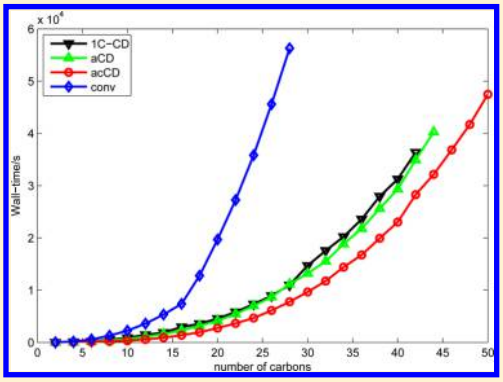
[†]Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O. Box 124 S-221 00 Lund, Sweden

[‡]Center for Biomolecular Nanotechnologies @UNILE, Italian Institute of Technology (IIT), Via Barsanti, I-73010 Arnesano (LE), Italy

[§]Department of Chemistry - Ångström, The Theoretical Chemistry Programme, Uppsala University, P.O. Box 518, SE-751 20 Uppsala, Sweden

[∥]Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Oslo, P.O. Box 1033 Blindern, N-0315 Oslo, Norway

ABSTRACT: We extend the local exchange (LK) algorithm [Aquilante, F.; Pedersen, T. B.; Lindh, R. *J. Chem. Phys.* **2007**, *126*, 194106] to the calculation of analytical gradients with density fitting. We discuss the features of the screening procedure and demonstrate the possible advantages of using this formulation, which is easily interfaced to a standard integral-direct gradient code. With auxiliary basis sets obtained from Cholesky decomposition of atomic or molecular integral blocks with a decomposition threshold of $10^{-4}E_h$, typical errors due to the density fitting in bond lengths, bond angles, and dihedral angles are 0.1 pm, 0.1°, and 0.5°, respectively. The overall speedup of geometry optimizations is about 1 order of magnitude for atomic natural-orbital-type basis sets but much less pronounced for correlation-consistent basis sets.



1. INTRODUCTION

However successful in describing the electronic structure of atoms and molecules, quantum chemical methods are confronted with a skepticism that derives from their high computational costs. One such bottleneck arises from the electron repulsion integrals (ERIs), as increasing the number of basis functions per atom leads to a quartic scaling of the number of integral evaluations. Much work is therefore put into devising approximations that avoid these expensive integral evaluations at a small loss of accuracy. For example, this can be achieved by Cholesky decomposition (CD) of the ERI matrix,[1,2] a method that exploits the existence of near and exact linear dependencies in the product space of the atomic orbitals (AOs). The use of this technique provides remarkably accurate results[2−5] while greatly reducing the prefactor associated with standard *ab initio* and density functional theory (DFT) methods due to the fact that the effective rank of the ERI matrix is much smaller than the total number of AO products. Moreover, the accuracy of the approximation can be systematically improved by lowering the decomposition threshold.

In the case of the density fitting (DF) approximation,[6−13] also known as resolution of the identity (RI), the same problem is attacked by introducing an auxiliary basis set that allows one to express four-index integrals in terms of two- and three-index integrals. The auxiliary basis sets are usually optimized by minimizing the error in specific energy contributions for individual atoms. Thus their accuracy is guaranteed only for a specific theoretical method,[14−20] and their use in a more general context is not recommended. One consequence is for example that such "standard" DFs, while extremely popular in applications of nonhybrid DFT models, are virtually nonexistent in the literature on applications of hybrid DFT, Hartree−Fock (HF), and multiconfigurational methods.

On the other hand, it has been shown that the generation of the auxiliary basis set in DF can be made without the atomic energy error minimization if one uses the so-called Cholesky basis obtained by CDs of each atomic sub-block of the integral matrix.[21,22] This type of *ab initio* DF combines the ease of computation of standard DF approximations with the unique features of molecular CD: accuracy control and unbiased nature.[4,5]

The increasing popularity[23−30] of Cholesky-based DF is also characterized by being widespread across many quantum chemical methods, including DFT and HF,[2,31,32] multiconfigurational second-order perturbation theory (CASSCF/CASPT2),[33−41] multireference configuration interaction,[42,43] Møller−Plesset second-order perturbation theory (MP2),[2,44,45] Coupled Cluster (CC) theory,[46−50] electron propagator methods,[51] symmetry-adapted perturbation theory,[52] fragment molecular orbital theory,[53] and quantum Monte Carlo.[54] The

Received: November 20, 2011
Published: November 26, 2012

development of Cholesky-based *ab initio* DF started as a side project while attempting a formal unification of the standard Cholesky method with the DF ideas.[55−58] The common identity of the two methods is now well understood and has been exploited to propose a formalism for computing analytical gradients and higher derivatives of integrals represented by Cholesky vectors. A few years ago, Aquilante et al.[56] reported an algorithm for computing analytical gradients in nonhybrid DFT models within the Cholesky framework.

One of the key elements of the success of Cholesky-based DF is represented by its combination with the local exchange (LK) screening algorithm,[31] based on localized Cholesky orbitals.[59] In fact, using the DF approximation to straightforwardly compute the exchange Fock matrix results in a quartic scaling algorithm.[16] Such a bottleneck in this type of self-consistent field (SCF) model renders the performance of DF-based algorithms very poor compared to the corresponding integral-direct implementations, except for calculations with few electrons and large basis sets. The idea put forward by Polly et al.[60] of using a local fitting auxiliary basis set for each orbital contribution to the exchange matrix is intended to solve the above "exchange problem" with a potentially linear-scaling algorithm. In this approach, a definition of local domains for the correct balance accuracy vs efficiency is not an easy task. Nonetheless, a wide range of local DF-based quantum mechanical models for HF and correlated calculations have been developed and used in recent years.[61−64] In contrast to the local DF schemes for HF exchange, the LK algorithm uses screening to reduce the computational scaling to quadratic with a low prefactor, thus providing considerable speedup for medium-sized molecules.

As we will show, the situation is even more involved when computing energy gradients in models that include HF exchange terms. Not only are we confronted with the quartic scaling in floating-point operations for the various tensor contractions but we also face the problem that, compared to an integral-direct algorithm, a considerable amount of input/output (I/O) overhead is involved. An existing implementation by Schütz et al.[65] circumvents such computational walls by employing local fitting domains for the HF exchange contributions to the gradients. The question then arises of whether or not an implementation of DF-based gradients for HF exchange that does not require fitting domains is meaningful for practical applications. In this article, we show that it is indeed possible to design such an implementation that allows geometry optimization on larger molecules. We achieve this goal by extending the LK algorithm to also cover analytical gradient calculations at the HF and hybrid DFT levels of theory. We demonstrate speedup and accuracy compared to integral-direct calculations.

2. THEORY

We start by briefly presenting the idea used to formulate analytical gradients in Cholesky-based methods, which exploits the common origin of these methods with DF. Since DF is defined in an analytical form, expressions for the gradients of the energy are trivial to formulate. In what follows, indices $\mu,\nu,...$ are used to indicate AOs; $i,j,...$ refer to occupied molecular orbitals (MOs); and $I,J,...$ refer to auxiliary basis functions and Cholesky vectors, unless otherwise stated.

The expression for the electronic part of the molecular HF gradient reads[66]

$$E^{(1)} = \sum_{\mu\nu} D_{\mu\nu} h^{(1)}_{\mu\nu} + \sum_{\mu\nu\lambda\sigma} P_{\mu\nu\lambda\sigma} (\mu\nu|\lambda\sigma)^{(1)} - \sum_{ij} \varepsilon_{ij} S^{(1)}_{ij} \tag{1}$$

where the two-particle reduced density matrix (2-RDM) is defined in terms of the one-electron density matrix (D) as

$$P_{\mu\nu\lambda\sigma} = D_{\mu\nu} D_{\lambda\sigma} - \frac{1}{2} D_{\mu\lambda} D_{\nu\sigma} \tag{2}$$

and the last term arises due to the geometry dependence of the overlap matrix present in the Lagrangian (ε_{ij} being the matrix of multipliers associated with the MO orthonormality constraints). Given the separability of the HF 2-RDM, the two-electron terms are usually identified with the specific names of Coulomb and exchange contribution, respectively:

$$E^{(1)}_C = \sum_{\mu\nu\lambda\sigma} D_{\mu\nu} D_{\lambda\sigma} (\mu\nu|\lambda\sigma)^{(1)} \tag{3}$$

$$E^{(1)}_X = \sum_{\mu\nu\lambda\sigma} D_{\mu\lambda} D_{\nu\sigma} (\mu\nu|\lambda\sigma)^{(1)} \tag{4}$$

The ERI matrix can be approximated by Cholesky vectors ($L^J_{\mu\nu}$) according to[1,2]

$$(\mu\nu|\lambda\sigma) \approx \sum_J L^J_{\mu\nu} L^J_{\lambda\sigma} \tag{5}$$

The Cholesky vectors are obtained through a recursive numerical algorithm which is not a very suitable starting point for the analytical formulation of their gradients, $(\mu\nu|\lambda\sigma)^{(1)}$, needed in eq 1. The idea put forward by O'Neal and Simons[67] of performing a CD of the supermatrix that includes derivative integrals in addition to the ERIs has never been thoroughly tested. Instead one can follow a line of thought based on the fact that a Cholesky decomposition can be viewed as a density fitting employing a set of auxiliary product functions (h_J) defined by the decomposition procedure (Cholesky basis):[56−58]

$$\sum_J L^J_{\mu\nu} L^J_{\lambda\sigma} = \sum_{JK} C^J_{\mu\nu} G_{JK} C^K_{\lambda\sigma} \tag{6}$$

where

$$G_{JK} = (h_J|h_K) \tag{7}$$

$$\sum_J C^J_{\mu\nu} G_{JK} = (\mu\nu|h_K) \tag{8}$$

It is then possible to approximate the ERI derivative as

$$(\mu\nu|\lambda\sigma)^{(1)} \approx \sum_J C^J_{\mu\nu} (\lambda\sigma|h_J)^{(1)} + \sum_J (\mu\nu|h_J)^{(1)} C^J_{\lambda\sigma} - \sum_{JK} C^J_{\mu\nu} G^{(1)}_{JK} C^K_{\lambda\sigma} \tag{9}$$

where the fitting coefficients can be obtained by solving eq 8. We compute the fitting coefficients directly from the Cholesky (or DF) vectors **L**, as in ref 56 [eq A4], namely $\mathbf{C} = \mathbf{L}\mathbf{Z}^T$, with $\mathbf{G}^{-1} = \mathbf{Z}\mathbf{Z}^T$ being the inverse Cholesky decomposition of the metric **G**.

With eq 9 as a starting point, it is straightforward to derive the following expression for the Coulomb part of the total energy derivative:

dx.doi.org/10.1021/ct200836x | J. Chem. Theory Comput. 2013, 9, 204−212

$$E_C^{(1)} = 2\sum_J \sum_{\mu\nu} V^J (h_J|\mu\nu)^{(1)} D_{\mu\nu} - \sum_{JK} V^J G_{JK}^{(1)} V^K \tag{10}$$

where **D** is the one-electron density matrix in the AO basis and **V** is a contraction between the DF coefficients and the density

$$V^J = \sum_{\lambda\sigma} D_{\lambda\sigma} C_{\lambda\sigma}^J \tag{11}$$

Together with one-electron, overlap, nuclear repulsion and numerical quadrature terms, the above expression completes the contributions to the gradients in DFT calculations employing functionals without explicit HF exchange. In order to perform HF and hybrid DFT calculations, however, the expression for the exchange energy gradients must be taken into account. The latter can be written as

$$E_x^{(1)} = 2\sum_J \sum_{\mu\nu\lambda\sigma} D_{\mu\lambda} C_{\mu\nu}^J (h_J|\lambda\sigma)^{(1)} D_{\nu\sigma} - \sum_{JK} \sum_{\mu\nu\lambda\sigma} D_{\mu\lambda} C_{\mu\nu}^J G_{JK}^{(1)} D_{\nu\sigma} C_{\lambda\sigma}^K \tag{12}$$

which is much harder to handle since the indices of the densities do not match the indices of the DF coefficients, thus preventing a regrouping of the sums analogous to the Coulomb term. (See also ref 65 for analogous expressions.) To simplify eq 12, we utilize the fact that the SCF one-electron density matrix can be written as

$$D_{\mu\nu} = \sum_i X_{\mu i} X_{\nu i} \tag{13}$$

and introduce the MO transformed fitting coefficients

$$C_{ij}^J = \sum_{\mu\nu} X_{\mu i} C_{\mu\nu}^J X_{\nu j} \tag{14}$$

so that

$$E_x^{(1)} = 2\sum_J \sum_{ij} C_{ij}^J (h_J|ij)^{(1)} - \sum_{JK} G_{JK}^{(1)} \sum_{ij} C_{ij}^J C_{ij}^K = 2\sum_J \sum_{\mu\nu} B_{\mu\nu}^J (h_J|\mu\nu)^{(1)} - \sum_{JK} G_{JK}^{(1)} A_{JK} \tag{15}$$

with the following definitions:

$$A_{JK} = \sum_{ij} C_{ij}^J C_{ij}^K \text{ and } B_{\mu\nu}^J = \sum_{ij} X_{\mu i} C_{ij}^J X_{\nu j} \tag{16}$$

Equation 13 is fulfilled by an infinite number of matrices **X**, all related to each other by unitary transformations that connect different orbital representations of the occupied MO space. For reasons that will be clarified soon, localized MOs are to be preferred. More specifically, we employ Cholesky orbitals,[59] as they are sufficiently well-localized and easier to generate compared to other standard localized MO sets. (See refs 31 and 60 for a discussion of this aspect.)

The MO transformed fitting coefficients are calculated by transforming the Cholesky vectors to the MO basis in a manner analogous to eq 14 followed by multiplication with the inverse Cholesky factors. (We later refer to these two parts of the algorithm as "step 1" and "step 2", respectively.) Since the number of occupied orbitals is much smaller than the total number of orbitals in almost any calculation of interest, this quantity is easier to store on disk than the original AO basis vectors. The **A** matrix is evaluated ("step 3") and also stored on disk. Triangular forms are employed whenever possible. The **B** vectors are then constructed on-the-fly ("step 4") and immediately contracted with the three-center integrals as they are needed. The evaluation of the integral derivatives constitutes "step 5" of our algorithm.

The above algorithm has a rather high computational scaling, proportional to O^2M^2 for the construction of the **A** matrix and proportional to $O^2NM + ON^2M$ for the transformation of Cholesky vectors and backtransformation of **C** vectors, where O is the number of occupied orbitals, N is the number of AO basis functions, and M is the number of auxiliary basis functions. The problem at hand is very similar to that of calculating the exchange contribution to the Fock matrix. For the latter, tremendous improvements are obtained by using the LK screening,[31] and therefore we have chosen to investigate a similar strategy for the gradients code. The details of the LK screening are presented and extensively discussed in in the original paper,[31] and here we will review only its basic ingredients.

The starting point of the LK algorithm is the following inequality defining an upper bound to the contribution of molecular orbital i to the exchange matrix elements:

$$\begin{aligned} |K_{\mu\nu}^i| &= |(\mu i|\nu i)| \\ &\le \sum_{\lambda\sigma} |X_{\lambda i}||(\mu\lambda|\nu\sigma)||X_{\sigma i}| \\ &\le \Big(\sum_\lambda |X_{\lambda i}|(\mu\lambda|\mu\lambda)^{1/2}\Big)\Big(\sum_\sigma |X_{\sigma i}|(\nu\sigma|\nu\sigma)^{1/2}\Big) \\ &= Y_\mu^i Y_\nu^i \end{aligned} \tag{17}$$

where the ith vector $\mathbf{Y}^i$ will feature sparsity if the MO coefficient vector $\boldsymbol{X}^i$ corresponds to a localized (Cholesky) orbital. Whenever the contribution to the exchange Fock matrix from a certain number of DF or Cholesky vectors has been computed, the above inequality can be written in terms of an updated (more sparse) integral diagonal. In LK, this is exploited by computing the **Y** vectors using the updated diagonal elements, so that the number of significant elements of the **Y** vectors will dynamically decrease. Finally, upper bounds based on Frobenius norms are employed to reduce the scaling (from cubic to quadratic) of the first MO transformation of the DF or Cholesky vectors.

In the present algorithm for gradients, we replace the above inequality with an analogous one, related directly to the exchange energy contributions:

$$\begin{aligned} |(ij|ij)| &\le \sum_{\mu\nu\lambda\sigma} |X_{\mu i}||X_{\nu j}||(\mu\nu|\lambda\sigma)||X_{\lambda i}||X_{\sigma j}| \\ &\le \Big(\sum_{\mu\nu} |X_{\mu i}||X_{\nu j}|(\mu\nu|\mu\nu)^{1/2}\Big)\Big(\sum_{\lambda\sigma} |X_{\lambda i}||X_{\sigma j}|(\lambda\sigma|\lambda\sigma)^{1/2}\Big) \\ &= Y_{ij}^2 \end{aligned} \tag{18}$$

The **Y** vectors now allow us to identify the pairs of occupied orbitals that contribute significantly to the energy gradient. The original LK algorithm is deployed to compute the half-transformed vectors $C_{\mu j}^K$, from which the final C_{ij}^K vectors are computed only for the interacting pairs, but stored on disk in full.

In order to integrate an algorithm for the DF gradients into an existing integral-direct gradient code, the simplest solution is to construct effective density matrices for the two-center and

Table 1. Baker Test Set

no.	name	no.	name	no.	name
1	water	11	disilyl ether	21	ACHTAR10
2	ammonia	12	1,3,5-trisilacyclohexane	22	ACANIL01
3	ethane	13	benzaldehyde	23	benzidine
4	acetylene	14	1,3-diflourobenzene	24	pterin
5	allene	15	1,3,5-triflourobenzene	25	difuropyrazine
6	hydroxysulfane	16	neopentane	26	mesityl oxide
7	benzene	17	furan	27	histidine
8	methylaminei	18	napthalene	28	dimethylpentane
9	ethanol	19	1,5-diflouronapthalene	29	caffeine
10	acetone	20	2-hydroxybicyclopentane	30	menthone

three-center contributions. In the HF case, for the three-center term such a matrix is given by

$$P^K_{\mu\nu} = V^K D_{\mu\nu} - \frac{1}{2} B^K_{\mu\nu} \tag{19}$$

whereas the effective density for the two-center term is given by

$$P_{JK} = V^J V^K - \frac{1}{2} A_{JK} \tag{20}$$

These formulas should be compared to the definition, eq 2, of the 2-RDM in integral-direct HF gradient calculations. Despite the formal similarities between eqs 19, 20, and 2, there is a striking difference between them: in the integral-direct case, the matrix elements are retrieved from the one-particle density matrix at virtually no cost. Can the gain due to the reduced number of integral derivatives needed in DF pay off the computational cost and I/O overhead involved in the evaluation of the effective two-particle density matrix? As will appear from our test calculations, the construction of the **A** matrix in eq 20 can be performed with the straightforward quartic scaling direct matrix−matrix multiplication, since the small computational prefactor that we observed for this step makes it of little relevance for the total timings. For eq 19, we can start by exploiting the fact that the AO indices $(\mu\nu)$ must belong to the basis functions that (a) contribute to the occupied orbital space and (b) have significant mutual spatial overlap. In practice, we include those orbital products $\mu\nu$ that satisfy the inequality

$$|D_{\mu\nu}|(\mu\nu|\mu\nu) \geq \tau_0 \tag{21}$$

with τ_0 set to 10^{-9} au in the present implementation. We observe that with this preliminary screening, the computational cost of computing the B terms should become proportional to O^3M, and correspondingly OM is the asymptotic scaling for the calculation of the three-center integral derivatives to be contracted with the matrix elements of eq 19. We will refer to this type of algorithm for DF gradients as the "*B-unscreened*" algorithm. On the other hand, it will be clear from our results that an additional screening is necessary in order to reduce even further the computational cost of the construction of the B terms and make this step less expensive than the calculation of the two- and three-index integral derivatives. This second screening, specific to the B terms, is based on the quantity $Y_{ij}[\mu\nu] = |X_{\mu i}||X_{\nu j}|(\mu\nu|\mu\nu)^{1/2}$ derived from eq 18. A given occupied orbital j is included in the list of the $n_o \ll O$ contributing orbitals to the B term of the shell pair μ,ν if $Y_{ij}[\mu\nu] \geq \tau$. The threshold τ is set in relation to the decomposition threshold, and it is found that a value of τ equal to the decomposition threshold does not downgrade the accuracy, while producing a significant reduction in the cost for computing the B terms. Clearly, as the occupied orbitals are localized, with the same argument used in discussing eq 18 we expect that only a constant (size-independent) number of i,j pairs will contribute for each given μ,ν shell pair. In practice, we use the full list of contributing orbitals n_o, therefore including $n_o(n_o + 1)/2$ orbital pairs for each μ,ν. In this way, by changing the storage of the C^K_{ij} vectors to the above subset of orbital pairs, we can use standard optimized BLAS utilities for performing the required matrix−matrix multiplications. The algorithm that uses this type of screening will be referred to as "LK-based". We will demonstrate that this type of LK screening produces a significant computational gain in the evaluation of the effective densities entering the expression for the exchange-HF DF gradients.

3. TEST CALCULATIONS

The algorithm for calculating exchange contributions to HF and DFT gradients presented above has been implemented in a development version of the quantum chemical software package MOLCAS,[68] which has been used to obtain all results reported in this section. We start by asserting the accuracy of the algorithm, followed by an investigation of its computational efficiency.

3.1. Accuracy. For testing the accuracy of our implementation, we have performed geometry optimizations on the Baker test suite,[69] which is listed in Table 1 for later reference. All the results presented in this section refer to the *B*-unscreened algorithm. However, test calculations (see also Table 4) show that no loss of accuracy is to be expected from the use of the LK-type screening. We exclude the histidine molecule (no. 27) due to convergence issues, which we observe both in DF and in integral-direct reference calculations (i.e., the problem is not related to the DF algorithm presented above). We performed geometry optimizations at the HF, hybrid DFT (B3LYP), and nonhybrid DFT (BLYP) levels of theory. The latter were performed only for comparison purposes, to isolate the effect of HF exchange on the accuracy of the gradient calculations. The ANO-RCC-VDZP basis set[70] was used in all calculations. Three types of auxiliary basis sets were used: one-center CD (1C-CD),[21] atomic CD (aCD),[21] and atomic compact CD (acCD),[22] all obtained with the decomposition threshold set to $10^{-4}E_h$.

Table 2 reports maximum absolute errors in bond distances, bond angles, and dihedral angles for the structures obtained using the DF algorithm compared to the structures computed using the integral-direct algorithm. For the present purposes, bond distances refer to all interatomic distances less than 2.0 Å, and bond angles and dihedral angles are obtained among these

dx.doi.org/10.1021/ct200836x | J. Chem. Theory Comput. 2013, 9, 204−212

Table 2. Maximum Absolute Errors in Bond Distance (in pm), Bond Angles (degrees), Dihedral Angles (degrees), and Total Energies (kcal/mol)

method	auxiliary basis	bond distance	bond angle	dihedral angle	energy
HF	acD	0.0494	0.06	0.49	0.377
	acCD	0.0496	0.09	0.45	0.421
	1C-CD	0.0683	0.08	5.90	0.689
B3LYP	aCD	0.8651	1.28	5.74	0.186
	acCD	0.8306	1.27	5.66	0.187
	1C-CD	0.8355	1.26	5.57	0.187
BLYP	aCD	0.0428	0.12	0.54	0.025
	acCD	0.0652	0.15	0.88	0.022
	1C-CD	0.0598	0.15	0.75	0.015

"bonded" atoms. On average, 18 338 distances, 935 bond angles, and 1364 dihedral angles are considered, with a slight variation depending on quantum chemical method.

For HF and BLYP, the errors in bond distance and bond angle are below 0.1 pm and 0.1°, respectively, whereas dihedral angles are accurate to within 1.0°. These deviations are insignificant in most applications. We do, however, observe errors 1 order of magnitude larger than these for HF with the 1C-CD auxiliary basis and for B3LYP. Table 2 also reports maximum absolute energy errors. These are obtained by comparing total energies computed with integral-direct techniques for the optimized DF and reference structures. The spread in the total energy errors is less than one might have expected based on the errors in structural parameters, suggesting that the structural errors might arise from internal motions associated with shallow minima on the potential energy surface.

As an alternative structural comparison, we have attempted to superimpose the DF structures on the reference ones. This is done by minimizing the root-mean-square deviation between the Cartesian coordinates of the DF structure and the corresponding reference structure through rigid rotation and translation of the former. The minimized root-mean-square deviation is then divided by the largest distance between an atom and the center of mass of the reference structure to yield a dimensionless normalized root-mean-square deviation, which we shall simply refer to as RMS. These RMS values are plotted in Figure 1. We note in passing that there is no particular auxiliary basis set that always performs the best or worst of those tested here. The results clearly show that the errors reported in Table 2 originate from only a few molecules, primarily disilyl ether (molecule no. 11), neopentane (16), and menthone (30). The main sources of error for disilyl ether and neopentane are dihedral angles corresponding to the rotation of silyl and methyl groups, respectively. These internal rotations are characterized by low barriers, and even very small errors on the potential energy surface may lead to the unusually large errors in dihedral angles that we observe for these two molecules. For menthone, the largest error is observed for the dihedral angle defined by the carbonyl oxygen atom and the isopropyl carbon atom attached to the ring. Unlike the methyl and silyl rotations, this is not a nearly free internal motion. Omitting the menthone molecule altogether and all dihedral angles involving hydrogen atoms, we get the results shown in Table 3. Bond distances are now all below 0.1 pm, bond angles below 0.15°, and dihedral angles below 0.5°.

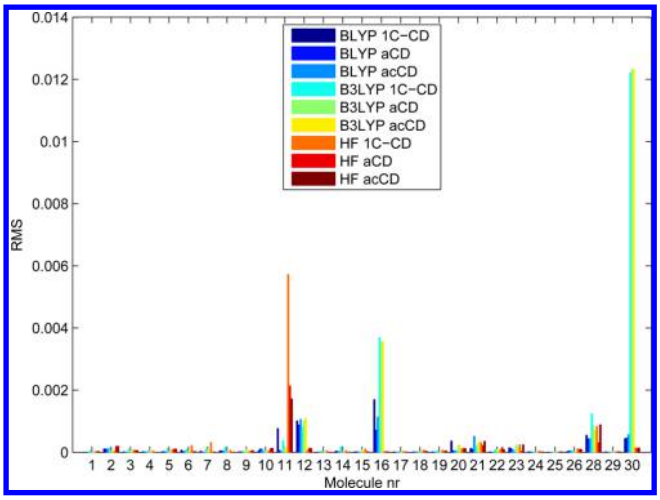


Figure 1. RMS of Cartesian distance when the structures are rotated and translated to maximize overlap.

Table 3. Maximum Absolute Errors in Bond Distance (in pm), Bond Angles (degrees), and Dihedral Angles (degrees)[a]

method	auxiliary basis	bond distance	bond angle	dihedral angle
HF	acD	0.0494	0.06	0.11
	acCD	0.0496	0.09	0.44
	1C-CD	0.0683	0.08	0.41
B3LYP	aCD	0.0954	0.13	0.35
	acCD	0.0799	0.13	0.42
	1C-CD	0.0780	0.10	0.44
BLYP	aCD	0.0428	0.12	0.32
	acCD	0.0652	0.15	0.50
	1C-CD	0.0598	0.15	0.46

[a]Menthone and dihedral angles involving hydrogen atoms are omitted.

3.2. Performance. To test the computational performance of the DF algorithm we started by looking at a set of conjugated dienes, C_nH_{n+2}, n = 2, 4, 6, ..., 50, with alternating double and single bonds in the carbon chain. The geometries for ethene, 1,3-butadiene, 1,3,5-hexatriene, and 1,3,5,7-octatriene were taken from the test suite of Schreiber et al.,[71] and the remaining geometries were extrapolated utilizing that the inner part of the chain shows a high degree of periodicity. The reason for choosing a set of similar molecules with different sizes is that it is easier to observe trends for the overall scaling. Since the principal difference between HF and hybrid DFT gradient calculations is the derivate of the energy with respect to the grid, a computational step independent of the DF approximation, only HF calculations are reported. These were carried out using the ANO-RCC-VDZP basis set[70] and a Cholesky threshold of $10^{-4}E_h$. In some cases, we have investigated the influence of different choices of localized orbitals to the overall speed of the algorithm and found that the Cholesky localization did not lower performances in a significative way (<10%) compared to more robust orbital localizations, for example compared to the use of the Pipek−Mezey scheme.[72] Integral-direct calculations were performed for comparison. Wall times and CPU times for the gradient calculation were recorded from tests performed using a single core on a Xeon E5520 (2.26 GHz, quadcore, 3 GB RAM/core).

In Figure 2, we report the wall times for computing the integral derivatives only. The DF curves refer to the *B*-

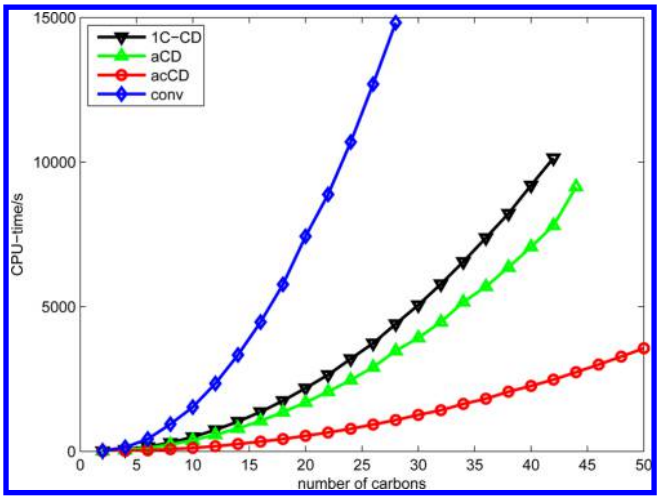


Figure 2. CPU time for computing the integral derivatives in integral-direct ("conv") and in the various DF approximations with the *B-unscreened* algorithm (ANO-RCC-VDZP basis set[70] and Cholesky threshold of $10^{-4}E_h$).

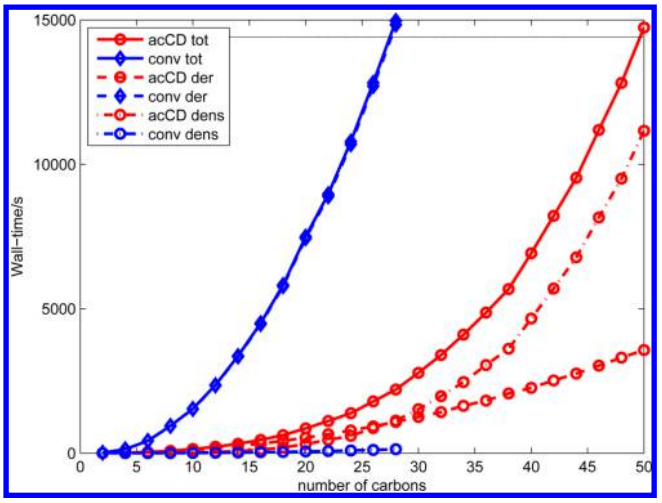


Figure 3. Wall time for computing HF gradients: conventional vs *B-unscreened* DF algorithm. Labels: "dens" refers to calculating the two- and three-index densities. "der" stands for the evaluation of the integral derivatives. "tot" indicates the total computing time for a gradient evaluation (ANO-RCC-VDZP basis set[70] and Cholesky threshold of $10^{-4}E_h$).

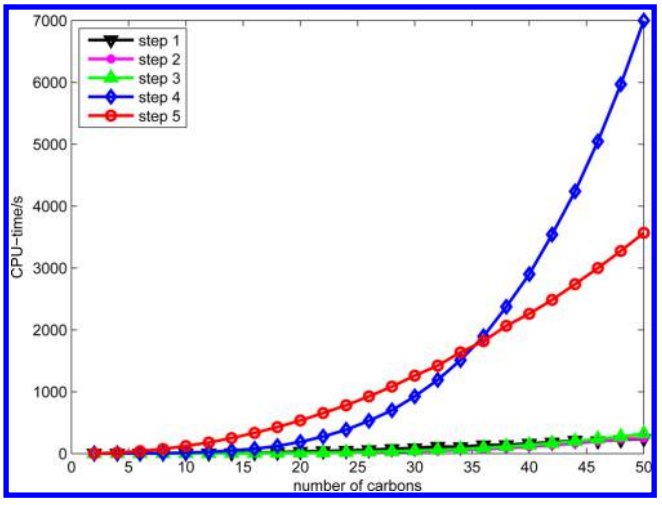


Figure 4. CPU time spent in the different computationally intensive steps of the *B-unscreened* DF gradient algorithm. The Coulomb term in the effective 2-RDM is computed at negligible costs. (Definition of the various steps: (1) MO-transformation of the Cholesky/DF vectors, (2) multiplication of the MO-basis vectors with the inverse Cholesky factors of the metric matrix, (3) construction of the **A** matrix, (4) generation of the **B** vectors, (5) evaluation of the integral derivatives.)

unscreened algorithm. We observe that the reduced number of integral derivatives to be computed and the absence of four-center terms result in a substantial computational gain for all three types of DF approximations. The simpler structure of the acCD auxiliary basis set compared to the aCD and 1C-CD ones (in terms of number of primitives and shell-pair completion, respectively) explains the much better performance in the evaluation of the derivative integrals. The advantage of compacting the set of primitive Gaussians in the auxiliary basis set is also clear from this plot. At the same time, one observes the drawback in using 1C-CD. The lack of shell-completion in 1C-CD forces the integral program to compute a large number of unused derivative integrals. This is due to the fact that the CD procedure generates vectors with no regard for the shell-pair structure of products of AOs. Although this has no consequences in energy calculations, where instead it ensures a given accuracy with a minimal number of vectors,[57] it makes the calculation of integral derivatives inefficient. In fact, in Figure 2 the 1C-CD curve accounts for the evaluation of the integral-derivatives required according to eq 9 if the full set of one-center AO products was used as an auxiliary basis set. Although 1C-CD uses only a fraction of that, there is no simple way to exploit this sparsity because of the shell-pair driven structure of the integral derivatives code. This becomes less of a problem when tighter CD thresholds are used. However, computing the integral derivatives is not the most time-consuming step in the *B-unscreened* DF algorithm. Indeed, as seen in Figure 3 the calculation of the effective density is more expensive than computing the derivatives of the two- and three-index integrals for larger systems. As expected, differences in timings between 1C-CD, acCD, and aCD are not very significant when restricted to the overhead in the construction of the effective density matrices. We report a more detailed account of this overhead in Figure 4 for the case of acCD. Notice that the evaluation of integral derivatives, a step that scales quadratically, carries a large computational prefactor that makes it the time-determining step for a range of molecular sizes. Nonetheless, the construction of the *B* term is clearly dominating for larger systems. It may seem quite surprising that two apparently very similar tasks, the formation of the C^K_{ij} from the AO vectors and their backtransformation to $B^K_{\mu\nu}$, could have such dramatically different computational costs ("step 1" vs "step 4," respectively). The answer lies in the efficiency of the original LK algorithm that is used in "step 1": the first half of the backtransformation is performed on the $C^K_{\mu j}$ vectors with the *i,j* indices stored in full, contrary to the corresponding step in the LK where only the significant occupied orbital pairs are processed—and that is the source of the striking difference between "step 1" and "step 4" in Figure 4. We have been able to remedy this unsatisfactory performance of the *B-unscreened* DF gradient algorithm by introducing the screening defined by the quantity $Y_{ij}[\mu\nu]$, as previously discussed. The results of Figure 5 show that with the introduction of the LK-based screening, "step 4" is no longer the bottleneck of the calculation, as it becomes negligible compared to the evaluation of the integral derivatives. These results were obtained with a choice of the

dx.doi.org/10.1021/ct200836x | J. Chem. Theory Comput. 2013, 9, 204–212

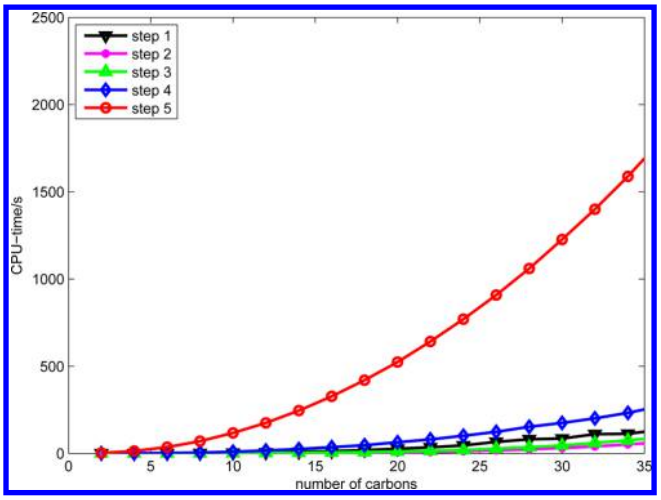


Figure 5. CPU time spent in the different computationally intensive steps of the LK-based DF gradient algorithm ($\tau = 10^{-5}$). The Coulomb term in the effective 2-RDM is computed at negligible costs. (Definition of the various steps: (1) MO-transformation of the Cholesky/DF vectors, (2) multiplication of the MO-basis vectors with the inverse Cholesky factors of the metric matrix, (3) construction of the **A** matrix, (4) generation of the **B** vectors, (5) evaluation of the integral derivatives.).

screening threshold of $\tau = 10^{-5}$, that is 10 times smaller than the decomposition threshold. This choice of the threshold is in line with our experience in testing the LK-based algorithm, as confirmed also by the results collected in Table 4.

Table 4. Accuracy of the *B-unscreened* and LK-based Algorithms (ANO-RCC-VDZP Basis Set[70] and Cholesky Threshold of $10^{-4}E_h$)[a]

			MAE (LK-based)	
parameter	MAE (*B-unscreened*)		$\tau = 10^{-5}$	$\tau = 10^{-4}$
next energy	HF	1.7×10^{-5}	1.8×10^{-4}	4.4×10^{-3}
	B3LYP	3.7×10^{-7}	1.3×10^{-6}	1.4×10^{-7}
RMSD		2.5×10^{-3}	1.7×10^{-3}	8.5×10^{-3}
		1.6×10^{-3}	2.0×10^{-3}	7.0×10^{-4}
bond length		4.8×10^{-5}	1.7×10^{-4}	6.9×10^{-4}
		8.3×10^{-5}	1.1×10^{-4}	1.4×10^{-4}
bond angle		0.01	0.02	0.07
		0.02	0.03	0.03
dihedral		0.21	0.13	0.62
		0.09	0.14	0.08

[a]Results from geometry optimization of molecule no. 23 of our test set. Max absolute errors (MAE) are computed with respect to the corresponding integral-direct calculation.

So far, the discussion has been based on test calculations performed with a general contracted ANO basis set. The latter contains a large number of primitive functions which might make the calculation of derivatives fairly expensive compared to the densities and favor our conclusions. In Tables 5 and 6, we list results for a molecule in the series of conjugated dienes with 24 carbon atoms using cc-pVDZ and cc-pVTZ. We note that the results are not as good as for the ANO series but still slightly faster than the conventional approach for cc-pVDZ and with about a 50% improvement for cc-pVTZ using a screening threshold of 10^{-5}. Going to higher screening thresholds gives a rather small gain, whereas the loss of accuracy amounts to roughly the same as the increase in value of the threshold.

We have performed a similar calculation on a Buckminster fullerene, Table 7, for which timings might look disappointing at first sight. However, in more practical terms one needs to perform an SCF energy calculation and other preliminary tasks (e.g., related to ERI evaluation) in order to get to the point where gradients are computed. For the same example, computing the gradients constitutes a minor portion of the whole calculation. When also taking the scf and integral setup

Table 5. CPU and Wall Times for Gradients Calculations on the $C_{24}H_{54}$ Molecule Using Different Screening Thresholds[a]

	screening thresholds/E_h				
	0.0	10^{-5}	10^{-4}	10^{-3}	conv
	CPU times (s):				
derivatives	48	47	47	46	262
densities	516	146	102	76	17
scr. overhead	276	49	24	9	
total	567	196	152	126	279
	wall times (s):				
derivatives	49	48	48	47	264
densities	593	209	155	135	17
scr. overhead	275	50	25	9	
total	647	261	207	187	281

[a]Basis set: cc-pVDZ. Auxiliary basis set: acCD-4.

Table 6. CPU and Wall Times for Gradients Calculations on the $C_{24}H_{54}$ Molecule Using Different Screening Thresholds[a]

	screening threshold/E_h				
	0.0	10^{-5}	10^{-4}	10^{-3}	conv
	CPU times (s):				
derivatives	377	371	374	371	2808
densities	3900	1082	859	689	259
scr. overhead	1598	204	95	31	
total	4297	1471	1251	1077	3068
	wall times (s):				
derivatives	378	377	374	371	2817
densities	4572	1624	1329	1099	262
scr. overhead	1606	206	98	34	
total	4994	2049	1749	1518	3080

[a]Basis set: cc-pVTZ. Auxiliary basis set: acCD-4.

Table 7. CPU and Wall Times for Gradients Calculations on the C_{60} Molecule Using the Screening Threshold 10^{-5} E_h[a]

	acCD-4	conv
	CPU times (s):	
derivatives	536	5377
densities	6087	1222
scr. overhead	2879	
total	6650	6736
	wall times (s):	
derivatives	538	5394
densities	8630	1235
scr. overhead	2887	
total	9224	6800
SCF energy	19200	40260

[a]Wall-times for the SCF calculation of the energy are reported for comparison. Basis set: cc-pVDZ.

dx.doi.org/10.1021/ct200836x | J. Chem. Theory Comput. 2013, 9, 204–212

into account, the acCD calculation required a wall time of just under 8 h, whereas the conventional calculation took just over 13 h. With this in mind, and since SCF calculations are carried out several times in a full geometry optimization, our results suggest that there is certainly a wide target range of applicability of our DF-based algorithm.

4. CONCLUSIONS AND OUTLOOK

Computing the HF exchange matrix is the bottleneck for all types of DF-based SCF models. The so-called LK screening algorithm[31] represents a way to overcome this scaling bottleneck. For geometry optimization with SCF models that include HF exchange, the evaluation of such contributions to the energy gradients becomes even more demanding. Besides the quartic scaling of the floating point operations, the intense I/O requirements have probably been the reason for the absence of any such implementation in existing quantum chemistry software, the only exception being the implementation described in Schütz et al.[65] At variance with the implementation presented here, the latter is based on the definition of local fitting domains, a potentially linear-scaling solution to the "exchange problem."

We have here extended the LK algorithm to the calculation of analytical gradients with DF, necessary for SCF models that include HF exchange. We have shown that our implementation can be used for SCF and hybrid DFT with an insignificant loss of accuracy. In our test calculations using CD-based auxiliary basis sets with a threshold of $10^{-4}E_h$, bond lengths are typically reproduced within 0.1 pm, bond angles within 0.1°, and dihedral angles within 0.5° compared to results obtained with an integral-direct implementation of the same SCF models. In a few cases, we have observed errors up to about 6° in dihedral angles and 0.9 pm in bond distances. These errors are of course ascribable to deficiencies in the auxiliary basis set and can be reduced by lowering the CD threshold and by increasing the AO basis set quality. A more thorough benchmark study that includes lower CD thresholds as well as larger AO basis sets, both of which lead to more accurate CD-based auxiliary basis sets, is the subject of a forthcoming publication. The improved performance is obtained even if the present algorithm uses the structure of an existing standard integral-direct code. We have chosen this strategy to make it as easy as possible to modify existing codes. Finally, we stress that this work serves as an important step toward the implementation of analytical gradients for DF-based CASSCF.

As for the conventional gradients algorithms, the bottleneck of the calculation remains the integral derivative evaluation. With a screening of the three-index quantities contributing to an effective density matrix, we have avoided introducing any further bottleneck to the calculation. The algorithm shows quadratic scaling, and for ANO basis sets we report speedups of more than 1 order of magnitude compared to integral-direct gradient evaluation in medium-sized molecules (up to 50 atoms). For segmented basis sets (e.g., cc-pVXZ family), the performance improvement over integral-direct is much less pronounced, depending on the choice of the screening threshold. For acceptable accuracies, the corresponding speedup in the gradient calculation is only at most 2-fold for the systems investigated, and for C_{60} we even observe a slowdown of 35%. Note, however, that the overall geometry optimization is still substantially faster than the conventional calculation due to savings in the energy evaluation. In the future, we plan to exploit the inherent locality of the CD-based auxiliary basis sets,[57] to improve on the performance of the present algorithm and move toward linear scaling. Numerical experiments with these ideas constitute work in progress in our laboratories.

AUTHOR INFORMATION

Corresponding Author

*E-mail: francesco.aquilante@gmail.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Funding from the Swedish Research Council (VR) is gratefully acknowledged. T.B.P. acknowledges support from the CoE Centre for Theoretical and Computational Chemistry through grant no. 179568/V30.

REFERENCES

(1) Beebe, N. H. F.; Linderberg, J. *Int. J. Quantum Chem.* **1977**, *12*, 683.
(2) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. *J. Chem. Phys.* **2003**, *118*, 9481.
(3) Pedersen, T. B.; Sánchez de Merás, A. M. J.; Koch, H. *J. Chem. Phys.* **2004**, *120*, 8887.
(4) Boström, J.; Aquilante, F.; Pedersen, T. B.; Lindh, R. *J. Chem. Theory Comput.* **2009**, *5*, 1545.
(5) Boström, J.; Delcey, M.; Aquilante, F.; Pedersen, T. B.; Serrano-Andrés, L.; Lindh, R. *J. Chem. Theory Comput.* **2010**, *6*, 747.
(6) Whitten, J. L. *J. Chem. Phys.* **1973**, *58*, 4496.
(7) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys.* **1973**, *2*, 41.
(8) Sambe, H.; Felton, R. H. *J. Chem. Phys.* **1975**, *62*, 1122.
(9) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396.
(10) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993.
(11) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.
(12) Vahtras, O.; Almlöf, J.; Feyereisen, M. *Chem. Phys. Lett.* **1993**, *213*, 514.
(13) Manby, F. R.; Knowles, P. J.; Lloyd, A. W. *J. Chem. Phys.* **2001**, *115*, 9144.
(14) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chim. Acta* **1997**, *97*, 119.
(15) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.
(16) Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285.
(17) Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167.
(18) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.
(19) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.
(20) Weigend, F.; Kattannek, M.; Ahlrichs, R. *J. Chem. Phys.* **2009**, *130*, 164106.
(21) Aquilante, F.; Lindh, R.; Pedersen, T. B. *J. Chem. Phys.* **2007**, *127*, 114107.
(22) Aquilante, F.; Pedersen, T. B.; Gagliardi, L.; Lindh, R. *J. Chem. Phys.* **2009**, *130*, 154107.
(23) Söderhjelm, P.; Aquilante, F.; Ryde, U. *J. Phys. Chem. B* **2009**, *113*, 11085.
(24) Pierloot, K.; Zhao, H.; Vancoillie, S. *Inorg. Chem.* **2010**, *49*, 10316.
(25) Vancoillie, S.; Zhao, H.; Radon, M.; Pierloot, K. *J. Chem. Theory Comput.* **2010**, *6*, 576.
(26) El-Khoury, P. Z.; George, L.; Kalume, A.; Schapiro, I.; Olivucci, M.; Tarnovsky, A. N.; Reid, S. A. *Chem. Phys. Lett.* **2010**, *496*, 68.
(27) Riley, K. E.; Pitoňák, M.; Černý, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66.
(28) de Graaf, C.; Sousa, C. *Chem.—Eur. J.* **2010**, *16*, 4550.

(29) Morgado, C. A.; Jurečka, P.; Svozil, D.; Hobza, P.; Šponer, J. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3522.
(30) Sauri, V.; Serrano-Andrés, L.; Moughal Shahi, A. R.; Vancoillie, S.; Pierloot, K.; Gagliardi, L. *J. Chem. Theory Comput.* **2011**, *7*, 153.
(31) Aquilante, F.; Pedersen, T. B.; Lindh, R. *J. Chem. Phys.* **2007**, *126*, 194106.
(32) Boman, L.; Koch, H.; Sánchez de Merás, A. *J. Chem. Phys.* **2008**, *129*, 134107.
(33) Aquilante, F.; Malmqvist, P.-Å.; Pedersen, T. B.; Ghosh, A.; Roos, B. O. *J. Chem. Theory. Comput.* **2008**, *4*, 694−702.
(34) Aquilante, F.; Pedersen, T. B.; Roos, B. O.; Sánchez de Merás, A.; Koch, H. *J. Chem. Phys.* **2008**, *129*, 024113.
(35) Aquilante, F.; Todorova, T. K.; Pedersen, T. B.; Gagliardi, L.; Roos, B. O. *J. Chem. Phys.* **2009**, *131*, 034113.
(36) Roca-Sanjuán, D.; Aquilante, F.; Lindh, R. *WIREs Comput. Mol. Sci.* **2012**, *2*, 585−603.
(37) Macchia, G. L.; Aquilante, F.; Veryazov, V.; Roos, B. O.; Gagliardi, L. *Inorg. Chem.* **2008**, *47*, 11455.
(38) Huber, S. M.; Shahi, A. R. M.; Aquilante, F.; Cramer, C. J.; Gagliardi, L. *J. Chem. Theory Comput.* **2009**, *5*, 2967.
(39) Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10964.
(40) La Macchia, G.; Li Manni, G.; Todorova, T. K.; Brynda, M.; Aquilante, F.; Roos, B. O.; Gagliardi, L. *Inorg. Chem.* **2010**, *49*, 5216.
(41) Li Manni, G.; Aquilante, F.; Gagliardi, L. *J. Chem. Phys.* **2011**, *134*, 034114.
(42) Chwee, T. S.; Carter, E. A. *J. Chem. Phys.* **2010**, *132*, 074104.
(43) Chwee, T. S.; Carter, E. A. *Mol. Phys.* **2010**, *108*, 2519.
(44) Aquilante, F.; Pedersen, T. B. *Chem. Phys. Lett.* **2007**, *449*, 354−357.
(45) Vysotskiy, V.; Cederbaum, L. S. *J. Chem. Theory Comput.* **2011**, *7*, 320.
(46) Pedersen, T. B.; Koch, H.; Boman, L.; Sánchez de Merás, A. M. J. *Chem. Phys. Lett.* **2004**, *393*, 319−326.
(47) Pedersen, T. B.; Kongsted, J.; Crawford, T. D. *Chirality* **2009**, *21*, S68.
(48) Pitonák, M.; Aquilante, F.; Hobza, P.; Neogrády, P.; Noga, J.; Urban, M. *Collect. Czech. Chem. Commun.* **2011**, *76*, 713.
(49) Pitoňák, M.; Heßelmann, A. *J. Chem. Theory Comput.* **2010**, *6*, 168.
(50) Pitoňák, M.; Neogrády, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1369.
(51) Vysotskiy, V. P.; Cederbaum, L. S. *J. Chem. Phys.* **2010**, *132*, 044110.
(52) Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 184111.
(53) Okiyama, Y.; Nakano, T.; Yamashita, K.; Mochizuki, Y.; Taguchi, N.; Tanaka, S. *Chem. Phys. Lett.* **2010**, *490*, 84.
(54) Purwanto, W.; Krakauer, H.; Virgus, Y.; Zhang, S. *J. Chem. Phys.* **2011**, *135*, 164105.
(55) Aquilante, F. New Approaches to Large-Scale Electronic Structure Calculations. Ph.D. thesis, Lund University, Lund, Sweden, 2007; ISBN 978-91-7422-169-5.
(56) Aquilante, F.; Lindh, R.; Pedersen, T. B. *J. Chem. Phys.* **2008**, *129*, 034106.
(57) Pedersen, T. B.; Aquilante, F.; Lindh, R. *Theor. Chem. Acc.* **2009**, *124*, 1.
(58) Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; Sánchez de Merás, A.; Pedersen, T. B. In *Linear-Scaling Techniques in Computational Chemistry and Physics. Methods and Applications*; Zalesny, R., Papadopoulos, M. G., Mezey, P. G., Leszczynski, J., Eds.; Challenges and Advances in Computational Chemistry and Physics; Springer: New York, 2011; Vol. *13*, Chapter 13, pp 301−343.
(59) Aquilante, F.; Pedersen, T. B.; Sánchez de Merás, A.; Koch, H. *J. Chem. Phys.* **2006**, *125*, 174101.
(60) Polly, R.; Werner, H. J.; Manby, F. R.; Knowles, P. J. *Mol. Phys.* **2004**, *102*, 2311.
(61) Werner, H. J.; Manby, F. R. *J. Chem. Phys.* **2006**, *124*, 054114.
(62) Usvyat, D.; Maschio, L.; Manby, F.; Casassa, S.; Schütz, M.; Pisani, C. *Phys. Rev. B* **2007**, *76*, 075102.
(63) Goll, E.; Leininger, T.; Manby, F. R.; Mitrushchenkov, A.; Werner, H.-J.; Stoll, H. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3353.
(64) Loibl, S.; Manby, F. R.; Schütz, M. *Mol. Phys.* **2010**, *108*, 477.
(65) Schütz, M.; Lindh, R.; Manby, F. R. *J. Chem. Phys.* **2004**, *121*, 737.
(66) Roos, B. O.; Widmark, P.-O. *European Summer School in Quantum Chemistry*; Lund University: Lund, Sweden, 2000; Vol. *I-III*.
(67) O'Neal, D. W.; Simons, J. *Int. J. Quantum Chem.* **1989**, *36*, 673.
(68) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.
(69) Baker, J.; Chan, F. *J. Comput. Chem.* **1996**, *17*, 888.
(70) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. *J. Phys. Chem. A* **2004**, *108*, 2851.
(71) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110.
(72) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916.