

Density Functional Theory for Molecular and Periodic Systems Using Density Fitting and Continuous Fast Multipole Methods

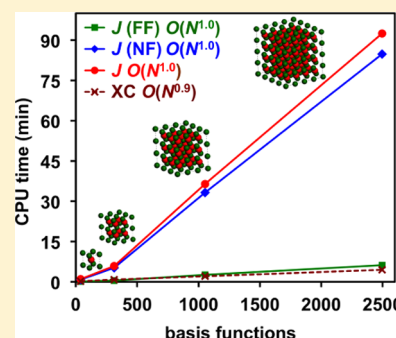
Roman Łazarski,[†] Asbjörn M. Burow,[‡] and Marek Sierka^{*,†}

[†]Otto-Schott-Institut für Materialforschung (OSIM), Friedrich-Schiller-Universität Jena, Löbdergraben 32, D-07743 Jena, Germany

[‡]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstrasse 7, D-81377 Munich, Germany

S Supporting Information

ABSTRACT: An implementation of Kohn–Sham density functional theory within the TURBOMOLE program package with Gaussian-type orbitals (GTO) as basis functions is reported that treats molecular and periodic systems of any dimensionality on an equal footing. Its key component is a combination of density fitting/resolution of identity (DF) approximation and continuous fast multipole method (CFMM) applied for the electronic Coulomb term. This DF-CFMM scheme operates entirely in the direct space and partitions Coulomb interactions into far-field part evaluated using multipole expansions and near-field contribution calculated employing density fitting. Computational efficiency and favorable scaling behavior of our implementation approaching $O(N)$ for the formation of Kohn–Sham matrix is demonstrated for various molecular and periodic systems including three-dimensional models with unit cells containing up to 640 atoms and 19072 GTO basis functions.



I. INTRODUCTION

The relatively low computational cost combined with accuracy makes density functional theory (DFT) methods important tools in contemporary computational chemistry and physics. DFT calculations are typically performed using a finite set of basis functions in order to solve algebraic rather than differential equations. The most common basis sets are plane waves (PW) and Gaussian-type orbitals (GTO).¹ PW are inherently periodic in three dimensions (3D) and require large simulation cells for calculations of lower dimensional systems—molecules (0D), chains and polymers (1D), or surfaces (2D)—in order to prevent overlapping of periodic images of the charge density. In contrast, GTO allow treating systems of any dimensionality on an equal footing without the need for constructing artificial 3D periodic models.^{2,3} GTO are also particularly well suited for DFT calculations on sparsely packed systems, such as zeolites and metal–organic frameworks.^{2,3}

When GTO are used, the computationally most demanding parts of DFT calculations are the electronic Coulomb (J) and exchange–correlation (E_{xc}) contributions. For the E_{xc} part several efficient numerical integration algorithms exist,^{4–7} among them our recent hierarchical scheme capable of achieving $O(N)$ scaling for systems of any dimensionality.⁸ Due to the long-range nature of electrostatic interactions an efficient evaluation of the Coulomb term is a much more challenging task, in particular under periodic boundary conditions.^{1,2,9} For GTO-based molecular and periodic DFT calculations the continuous fast multipole method (CFMM)^{2,10–12} is computationally one of the most efficient approaches. In CFMM, the Coulomb problem is partitioned into far-field (FF) and near-field (NF) portions. The FF part

which comprises the major part of interactions including the long-ranged crystal part is evaluated very efficiently using multipole expansions. The NF interactions are treated by a direct integration and represent computationally the most demanding part for both molecular and periodic systems.¹²

The computational cost of the direct integration can be significantly reduced by using the density fitting (DF) procedure, also known as resolution of identity (RI) approximation.^{13–16} In this method the electron density is approximated by a linear combination of atom-centered auxiliary basis functions. The corresponding expansion coefficients are determined by minimization of the difference between exact and auxiliary electron density in a particular metric. It has been shown that among different metric choices the use of the Coulomb one leads to a variational problem and the smallest fitting error.^{15,17} For molecular systems very efficient methods combining DF with multipole expansions¹⁸ or the continuous fast multipole method (CFMM)¹⁹ have been developed that allow one to reduce the asymptotic scaling to nearly $O(N)$. However, extension of DF approximation to periodic systems is not straightforward due to a slow decay behavior of the auxiliary density expansion coefficients and divergent terms in the Coulomb metric matrix under periodic boundary conditions (PBC).^{17,20} To circumvent these problems, several methods have been developed, including the use of non-Coulomb metrics,^{17,20} introducing basis sets comprising GTO and chargeless Poisson basis functions,²¹ combining electron and nuclear charge densities,^{22,23} and restricting the

Received: March 17, 2015

Published: May 27, 2015

fitting domain.²⁴ The Coulomb lattice sums within these methods are usually evaluated using the Ewald summation technique.²⁵ Recently, a fast and robust DF scheme for molecular and periodic systems operating entirely in direct space has been developed.⁹ Its core is the decomposition of auxiliary density into charged and chargeless parts. The use of the Coulomb metric under periodic boundary conditions constrains the charged part. The determination of the chargeless component is variational and involves only convergent Coulomb lattice sums.

Based on our previous work^{8,9} a full GTO-based DFT implementation is presented here that treats molecules as well as systems with 1D, 2D, and 3D periodicity on an equal footing. The novel aspect and the core of the implementation is a combination of DF and CFMM. Unlike other methods combining DF and multipole expansions,²⁶ this DF-CFMM scheme operates entirely in the direct space. The efficiency and scaling behavior of the program is evaluated for a variety of molecular and periodic systems. Our results show that the performance of the DF-CFMM scheme makes it competitive to DF algorithms based on plane waves,^{27,28} in particular for sparsely packed 3D, 2D, and 1D periodic systems.

II. GENERAL THEORY

A. DFT for Molecular and Periodic Systems. The general scheme of our GTO-based DFT implementation for periodic systems follows the standard procedures available in similar programs.^{1–3} Here, only the most important equations are given for the sake of completeness.

The translational symmetry of solids leads to Bloch orbitals $\psi_{p\sigma}^{\mathbf{k}}$ and one-particle energies $\epsilon_{p\sigma}^{\mathbf{k}}$ depending on the band index p , spin σ , and the wave vector \mathbf{k} within the Brillouin zone (BZ), which is the unit cell of reciprocal space. The orbitals

$$\psi_{p\sigma}^{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_{\text{UC}}}} \sum_{\mathbf{L}} e^{i\mathbf{k} \cdot \mathbf{L}} \sum_{\mu} C_{\mu p\sigma}^{\mathbf{k}} \mu_{\mathbf{L}}(\mathbf{r}) \quad (1)$$

are expanded in GTO basis functions $\mu(\mathbf{r}-\mathbf{R}_{\mu}-\mathbf{L}) \equiv \mu_{\mathbf{L}}(\mathbf{r})$ centered at atomic positions \mathbf{R}_{μ} in direct lattice cells \mathbf{L} over all N_{UC} unit cells. This results in the unrestricted Kohn–Sham equations,

$$\mathbf{F}_{\sigma}^{\mathbf{k}} \mathbf{C}_{\sigma}^{\mathbf{k}} = \mathbf{S}^{\mathbf{k}} \mathbf{C}_{\sigma}^{\mathbf{k}} \epsilon_{\sigma}^{\mathbf{k}} \quad (2)$$

which may be solved separately for each \mathbf{k} in the BZ. The same equations hold for the molecular case, where only $\mathbf{L} = \mathbf{k} = \mathbf{0}$ is a valid choice and N_{UC} is one. Equation 2 contains the reciprocal space Kohn–Sham (KS) and the overlap matrices $\mathbf{F}_{\sigma}^{\mathbf{k}}$ and $\mathbf{S}^{\mathbf{k}}$, respectively, obtained as Fourier transforms of real space matrices

$$F_{\mu\nu\sigma}^{\mathbf{k}} = \sum_{\mathbf{L}} e^{i\mathbf{k} \cdot \mathbf{L}} F_{\mu\nu\sigma}^{\mathbf{L}} \quad (3)$$

$$S_{\mu\nu}^{\mathbf{k}} = \sum_{\mathbf{L}} e^{i\mathbf{k} \cdot \mathbf{L}} S_{\mu\nu}^{\mathbf{L}} \quad (4)$$

The elements $F_{\mu\nu\sigma}^{\mathbf{L}}$ contain three contributions: elements $T_{\mu\nu}^{\mathbf{L}}$ of the kinetic energy matrix, elements $J_{\mu\nu}^{\mathbf{L}}$ of the Coulomb matrix, and elements $X_{\mu\nu\sigma}^{\mathbf{L}}$ of the exchange-correlation matrix,

$$F_{\mu\nu\sigma}^{\mathbf{L}} = T_{\mu\nu}^{\mathbf{L}} + J_{\mu\nu}^{\mathbf{L}} + X_{\mu\nu\sigma}^{\mathbf{L}} \quad (5)$$

The evaluation of $T_{\mu\nu}^{\mathbf{L}}$ is identical to the molecular case, and $X_{\mu\nu\sigma}^{\mathbf{L}}$ are calculated using a hierarchical integration scheme.⁸

The calculation of $J_{\mu\nu}^{\mathbf{L}}$ and $X_{\mu\nu\sigma}^{\mathbf{L}}$ requires the real space density matrix obtained by integration,

$$D_{\mu\nu\sigma}^{\mathbf{L}} = \frac{1}{V_{\mathbf{k}}} \int_{\text{BZ}} D_{\mu\nu\sigma}^{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{L}} d\mathbf{k} \quad (6)$$

of the reciprocal space density matrix

$$D_{\mu\nu\sigma}^{\mathbf{k}} = \sum_p f_{p\sigma}^{\mathbf{k}} (C_{\mu p\sigma}^{\mathbf{k}})^* C_{\nu p\sigma}^{\mathbf{k}} \quad (7)$$

over the BZ with volume $V_{\mathbf{k}}$. The orbital coefficients $C_{\mu p\sigma}^{\mathbf{k}}$ are taken from an initial guess or from the previous solution of eq 2. The occupation numbers $f_{p\sigma}^{\mathbf{k}}$ at zero temperature are either zero or one for virtual and occupied orbitals, respectively. The reciprocal space integral in eq 6 is evaluated numerically on a finite grid of \mathbf{k} -points.

Within a self-consistent-field (SCF) cycle, the calculation of all contributions to the KS and overlap matrices is first completed in direct space and then eq 2 is solved in reciprocal space. The last step is accomplished within a loop over a grid of \mathbf{k} -points, where the direct space matrices are transformed into the orthonormal orbital basis on a single \mathbf{k} -point. In this way, a regular matrix eigenvalue problem is obtained in eq 2, which is solved by diagonalization of the KS matrix. The elements $D_{\mu\nu\sigma}^{\mathbf{k}}$ in eq 7 are formed from the eigenvectors $C_{\mu p\sigma}^{\mathbf{k}}$ of eq 2 and are subsequently transformed on the fly to yield the direct space density matrix according to eq 6. In each SCF cycle, the total energy per unit cell, E , is calculated as

$$E = \sum_{\mu\nu\mathbf{L}} \sum_{\sigma} D_{\mu\nu\sigma}^{\mathbf{L}} T_{\mu\nu}^{\mathbf{L}} + E_{\text{xc}} + J \quad (8)$$

where E_{xc} is the exchange-correlation energy evaluated as in ref 8 and J is the Coulomb energy calculated as described in the next section.

An important feature of the real space KS matrix elements $F_{\mu\nu\sigma}^{\mathbf{L}}$ is their exponential decay with increasing separation of the basis functions μ and ν . Combined with the direct space symmetry relation of the matrices, $M = F, T, J, X_{\sigma}$, and D_{σ} defined previously,

$$M_{\nu\mu}^{-\mathbf{L}} = M_{\mu\nu}^{\mathbf{L}} \quad (9)$$

the exponential decay allows for a very efficient sparse storage of the real space matrices.⁸ There is further a reciprocal space symmetry relation

$$M_{\mu\nu}^{-\mathbf{k}} = (M_{\mu\nu}^{\mathbf{k}})^* \quad (10)$$

which shows that all direct space matrix elements must be real, if the BZ contains for each \mathbf{k} the corresponding $-\mathbf{k}$ and real GTO basis functions are used.

B. Density Fitting Scheme. A full description of the method is given in ref 9, and here we provide only the basic summary. The total crystal electron density ρ^{cryst} is a (infinite) sum of local densities $\rho_{\mathbf{L}}$ centered in cells translated by \mathbf{L} ,

$$\rho^{\text{cryst}} = \sum_{\mathbf{L}} \rho_{\mathbf{L}} \quad (11)$$

with

$$\rho_{\mathbf{L}} = \sum_{\mu\nu\mathbf{L}'} \sum_{\sigma} D_{\mu\nu\sigma}^{\mathbf{L}'} \mu_{\mathbf{L}} \nu_{\mathbf{L}'} \quad (12)$$

where the shorthand notation $\nu_{\mathbf{L}}(\mathbf{r}-\mathbf{L}') \equiv \nu_{\mathbf{L}\mathbf{L}'}$ is used and the subscript $\mathbf{0}$ is omitted for $\mathbf{L} = \mathbf{0}$. Atom-centered GTO auxiliary basis functions that are elements of vector $\boldsymbol{\alpha}$ are denoted as α .

The total crystal electron density ρ^{cryst} is approximated by auxiliary crystal electron density $\tilde{\rho}^{\text{cryst}}$,

$$\rho^{\text{cryst}} \approx \tilde{\rho}^{\text{cryst}} = \sum_{\mathbf{L}} \tilde{\rho}_{\mathbf{L}} \quad (13)$$

composed of unit cell auxiliary densities

$$\tilde{\rho}_{\mathbf{L}} = \sum_{\alpha} \mathbf{c}^T \boldsymbol{\alpha}_{\mathbf{L}} \quad (14)$$

The vector of expansion coefficients \mathbf{c} is independent of the lattice vectors \mathbf{L} since it is identical in every unit cell. It is determined by minimizing the Coulomb repulsion D of the residual density $\delta\rho = \rho - \tilde{\rho}$

$$D = \iint \delta\rho(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \sum_{\mathbf{L}} \delta\rho_{\mathbf{L}}(\mathbf{r}') \, d\mathbf{r}' = (\delta\rho|\delta\rho_{\mathbf{L}}) = (\rho - \tilde{\rho}|\rho_{\mathbf{L}} - \tilde{\rho}_{\mathbf{L}}) \quad (15)$$

Unless explicitly stated, in eq 15 and throughout this work, a direct space lattice sum over lattice vectors \mathbf{L} is understood on the right-hand side of Coulomb integrals.

In periodic systems D is finite only if $\delta\rho$ is chargeless, i.e.,

$$\int \delta\rho(\mathbf{r}) \, d\mathbf{r} = 0 \quad \Rightarrow \quad \int \tilde{\rho}(\mathbf{r}) \, d\mathbf{r} = N_{\text{el}} \quad (16)$$

Thus, the charge of $\tilde{\rho}$ is naturally constrained to the number of electrons, N_{el} . In order to explicitly enforce this condition $\tilde{\rho}$ is decomposed into charged and chargeless components, $\tilde{\rho}_{\parallel}$ and $\tilde{\rho}_{\perp}$, respectively,

$$\tilde{\rho} = \tilde{\rho}_{\parallel} + \tilde{\rho}_{\perp} = \mathbf{c}_{\parallel}^T \boldsymbol{\alpha} + \mathbf{c}_{\perp}^T \boldsymbol{\alpha} \quad (17)$$

with

$$\int \tilde{\rho}_{\parallel}(\mathbf{r}) \, d\mathbf{r} = N_{\text{el}} \quad \text{and} \quad \int \tilde{\rho}_{\perp}(\mathbf{r}) \, d\mathbf{r} = 0 \quad (18)$$

The vectors of expansion coefficients \mathbf{c}_{\parallel} and \mathbf{c}_{\perp} are orthogonal to each other and can be expressed with the help of projection matrices

$$\mathbf{P}_{\parallel} = \mathbf{n}\mathbf{n}^T \quad \text{and} \quad \mathbf{P}_{\perp} = \mathbf{1} - \mathbf{n}\mathbf{n}^T \quad (19)$$

as

$$\mathbf{c}_{\parallel} = \mathbf{P}_{\parallel}\mathbf{c} \quad \text{and} \quad \mathbf{c}_{\perp} = \mathbf{P}_{\perp}\mathbf{c} \quad (20)$$

The vector \mathbf{n} in eq 19 is a normalized charge vector of the auxiliary basis with elements

$$n_{\alpha} = \frac{1}{|\mathbf{q}|} q_{\alpha}, \quad \mathbf{q} = (q_1, q_2, \dots), \quad q_{\alpha} = \int \alpha(\mathbf{r}) \, d\mathbf{r} \quad (21)$$

The same projection matrices are used to define the vectors of charged and chargeless auxiliary basis functions $\boldsymbol{\alpha}_{\parallel}$ and $\boldsymbol{\alpha}_{\perp}$, respectively, as $\boldsymbol{\alpha}_{\parallel} = \mathbf{P}_{\parallel}\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_{\perp} = \mathbf{P}_{\perp}\boldsymbol{\alpha}$. Due to the mentioned charge constraint of $\tilde{\rho}$ the expansion coefficients \mathbf{c}_{\parallel} of the charged component $\tilde{\rho}_{\parallel}$ take the form

$$\mathbf{c}_{\parallel} = \frac{N_{\text{el}}}{|\mathbf{q}|} \mathbf{n} \quad (22)$$

The complementary chargeless part $\tilde{\rho}_{\perp}$ is determined by minimization of D (eq 15) with respect to all elements of \mathbf{c}_{\perp} . This leads to a system of linear equations

$$(\mathbf{V}_{\perp} + \mathbf{P}_{\parallel})\mathbf{c}_{\perp} = \boldsymbol{\xi}_{\perp} \quad (23)$$

where \mathbf{V}_{\perp} is the projected Coulomb metric matrix

$$\mathbf{V}_{\perp} = \mathbf{P}_{\perp}\mathbf{V}\mathbf{P}_{\perp} = (\alpha_{\perp}|\alpha_{\perp}^T) \quad (24)$$

and vector $\boldsymbol{\xi}_{\perp}$ is given by

$$\boldsymbol{\xi}_{\perp} = (\alpha_{\perp}|\rho_{\mathbf{L}} - \tilde{\rho}_{\parallel\mathbf{L}}) \quad (25)$$

An important property of this DF scheme is that all lattice sums in eqs 23–25 are convergent since they employ exclusively chargeless quantities. Contrary to other approaches for DF in extended systems,^{22,23} our formulation does not use the nuclear charge distribution to obtain such convergent lattice sums.

The final set of expansion coefficients of the auxiliary density is obtained as $\mathbf{c} = \mathbf{c}_{\parallel} + \mathbf{c}_{\perp}$. The elements $J_{\mu\nu}^{\mathbf{L}}$ of the real space Coulomb matrix are calculated as

$$J_{\mu\nu}^{\mathbf{L}} = (\mu\nu_{\mathbf{L}}|\tilde{\rho}_{\mathbf{L}} - \rho_{\mathbf{nL}}) \quad (26)$$

where $\rho_{\mathbf{n}}$ denotes the unit cell nuclear charge distribution. They are used to calculate the total Coulomb energy including the nuclear contribution as

$$J = \sum_{\mu\nu\mathbf{L}} D_{\mu\nu}^{\mathbf{L}} J_{\mu\nu}^{\mathbf{L}} - \frac{1}{2} (\tilde{\rho} + \rho_{\mathbf{n}}|\tilde{\rho}_{\mathbf{L}} - \rho_{\mathbf{nL}}) \quad (27)$$

C. DF-Accelerated Continuous Fast Multipole Method. The evaluation of Coulomb lattice sums appearing in eqs 25–27 is performed entirely in the direct space employing CFMM.^{2,11,29–31} We term this combination the DF-accelerated CFMM (DF-CFMM). This section provides basic definitions and description of the algorithm.

Well-Separated Charge Distributions. The definition of well-separated charge distributions plays a central role in the DF-CFMM. Two charge distributions centered at \mathbf{P} and \mathbf{Q} with extents (radii) r_P and r_Q , respectively, are well-separated if the distance between the centers is larger than the sum of their extents; i.e.,

$$|\mathbf{P} - \mathbf{Q}| \geq r_P + r_Q \quad (28)$$

The extents in eq 28 ensure that for well-separated distributions the error of the Coulomb interaction introduced by the multipole approximation is smaller than a given threshold ε .¹⁸ For a primitive auxiliary basis function with an exponent ζ_{ak} the extent is given by

$$r_k = \sqrt{\frac{-\ln \varepsilon + \frac{1}{2} \ln(\zeta_{ak})}{\zeta_{ak}}} \quad (29)$$

For a contracted auxiliary basis function α containing k primitive Gaussian functions the extent is

$$r_{\alpha} = \max_k \{r_k\} \quad (30)$$

and the center is simply the atom at which α is centered. For a product $\mu_i\nu_j$ of two primitive basis functions with exponents $\zeta_{\mu i}$ and $\zeta_{\nu j}$ centered at \mathbf{R}_{μ} and $\mathbf{R}_{\nu\mathbf{L}} = \mathbf{R}_{\nu} + \mathbf{L}$, respectively, the extent is given as

$$r_{ij\mathbf{L}} = \sqrt{\frac{-\ln \varepsilon + \ln S_{ij}^{\mathbf{L}} + \frac{1}{2} \ln(\zeta_{\mu i} + \zeta_{\nu j})}{\zeta_{\mu i} + \zeta_{\nu j}}} \quad (31)$$

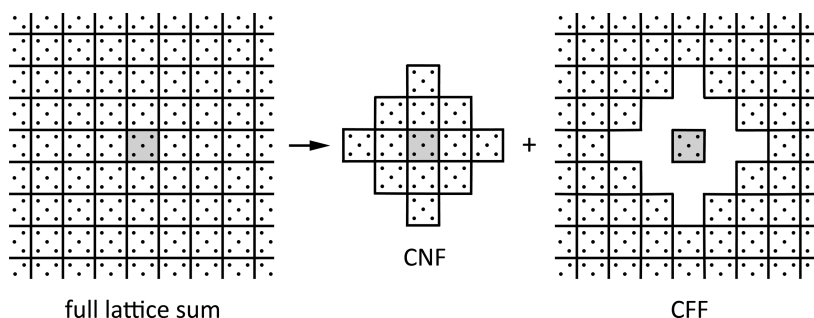


Figure 1. Partitioning of a Coulomb lattice sum into the CNF and CFF contributions. The central unit cell with $\mathbf{L} = \mathbf{0}$ is marked in gray.

with S_{ij}^L as the overlap integral evaluated neglecting angular parts of the basis functions.¹⁸ The center of the product is

$$\mathbf{R}_{ij\mathbf{L}} = \frac{\zeta_{\mu i} \mathbf{R}_{\mu} + \zeta_{\nu j} \mathbf{R}_{\nu}}{\zeta_{\mu i} + \zeta_{\nu j}} \quad (32)$$

For a product of two contracted basis functions μ and ν , a sphere is defined with the extent $r_{\mu\nu\mathbf{L}}$ enclosing all primitive products $\mu_i\nu_j\mathbf{L}$. The center $\mathbf{R}_{\mu\nu\mathbf{L}}$ of the sphere is the center of the charge distribution associated with the product $\mu\nu\mathbf{L}$.

Partitioning of Coulomb Lattice Sums. The Coulomb lattice sums in eqs 25–27 are partitioned into the crystal near-field (CNF) and far-field (CFF) contributions as depicted in Figure 1. In a full lattice sum of the form $(\rho_1 | \rho_2\mathbf{L})$ the distribution ρ_1 in the central unit cell ($\mathbf{L} = \mathbf{0}$) interacts with an infinite number of distributions $\rho_2\mathbf{L}$, i.e., ρ_2 translated by all possible \mathbf{L} . The distributions ρ_1 and ρ_2 can contain point charges (nuclei) as well as continuous charge distributions (shell pairs $\mu\nu\mathbf{L}$ or/and auxiliary basis functions α).

The CNF contribution consists of ρ_1 interacting with $\rho_2\mathbf{L}$ for a finite number of \mathbf{L} corresponding to the neighboring unit cells. The remaining part of the lattice sum defines the CFF; i.e.,

$$\sum_{\mathbf{L}} (\rho_1 | \rho_2\mathbf{L}) = \sum_{\mathbf{L} \in \text{CNF}} (\rho_1 | \rho_2\mathbf{L}) + \sum_{\mathbf{L} \in \text{CFF}} (\rho_1 | \rho_2\mathbf{L}) \quad (33)$$

A lattice vector \mathbf{L} belongs to the CFF if it fulfills two conditions: (i) all charge distributions forming ρ_1 are well-separated from those of $\rho_2\mathbf{L}$, and (ii) $|\mathbf{L}| \geq w_s \times (r_1^{\max} + r_2^{\max})$, where r_1^{\max} and r_2^{\max} are maximum distances between the unit cell center and all charge centers comprising ρ_1 and ρ_2 , respectively, and w_s is the well-separateness criterion. Note that our definition of w_s is different from that used by other authors, e.g., in refs 10 and 30.

Crystal Far-Field Contribution. The infinite lattice sums of the CFF part for systems of 1D, 2D, and 3D periodicity with arbitrary lattice vectors are evaluated using multipole expansions and recurrence relations derived by Kudin and Scuseria¹¹ for calculation of the lattice sum S of so-called “external to local” translation operators. S is calculated only once per SCF run and stored since it depends only on the size and shape of the unit cell. Its calculation is a very efficient $O(1)$ operation. Evaluation of the CFF contribution S_{CFF} to the lattice sum $(\rho_1 | \rho_2\mathbf{L})$ involves contraction

$$S_{\text{CFF}} = \omega^{\rho_1} \otimes \mu^{\rho_2} = \sum_{l=0}^{L_{\max}} \sum_{m=-l}^l \omega_{lm}^{\rho_1} \mu_{lm}^{\rho_2} \quad (34)$$

of the multipole moments $\omega_{lm}^{\rho_1}$ of ρ_1 about the unit cell center with the corresponding Taylor moments $\mu_{lm}^{\rho_2}$ of $\rho_2\mathbf{L}$ for $\mathbf{L} \in \text{CFF}$

obtained from S and $\omega_{lm}^{\rho_2}$ as described in ref 11. For systems with 3D periodicity and unit cells with nonvanishing dipole moment additional correction terms are necessary in order to obtain convergent lattice sums.^{29,30} Our dipole correction scheme places fictitious point charges on the unit cell face centers in order to cancel the dipole moment of the unit cell, in analogy to the method proposed by Kudin and Scuseria.^{29,30}

Crystal Near-Field Contribution. The CNF contribution of the lattice sum $(\rho_1 | \rho_2\mathbf{L})$ is evaluated employing CFMM that combines high computational efficiency and favorable scaling approaching $O(N)$. The key feature of the CFMM is an octree-based hierarchical spatial decomposition of the charge distributions ρ_1 and ρ_2 . This allows partitioning of Coulomb interactions between the two distributions into far-field (FF) and near-field (NF) contributions. The NF part is usually evaluated by means of direct integration. The FF part comprises the majority of interactions and is calculated very efficiently using a hierarchy of multipole expansions. Our CFMM implementation closely follows the algorithm described in refs 2, 10, and 29–31. The main differences of our approach are (i) the NF contribution to the Coulomb matrix elements is evaluated employing density fitting and (ii) octree construction and the spatial decomposition are applied directly to ρ_1 and $\rho_2\mathbf{L}$ only for $\mathbf{L} = \mathbf{0}$. The interactions between ρ_1 and $\rho_2\mathbf{L}$ for the remaining $\mathbf{L} \in \text{CNF}$ are calculated using replicas of the octree.

The construction of the octree is based on the fractional tiers scheme of White and Head-Gordon.³² It achieves an optimum balance between the NF and FF work by allowing for an arbitrary number of lowest level boxes. In contrast, standard boxing schemes restrict the number of lowest level boxes to be a power of eight. In short, a cubic parent box enclosing all distribution centers of ρ_1 and ρ_2 is constructed that is large enough to yield a predefined number n_{targ} of distribution centers per lowest level box. The parent box is successively subdivided in half along all Cartesian axes yielding the octree. In the next step all charge distributions comprising ρ_1 and ρ_2 are sorted into boxes. This is achieved by assigning the center of each distribution (nuclear point charge, auxiliary basis function, or shell pair) of ρ_1 and ρ_2 to the smallest box with the edge length $L_i \geq 2r_i/(w_s - 1)$, where r_i is the distribution extent. This condition ensures that on each tree level the distributions from well-separated boxes do not overlap. Nuclear point charges are sorted into the lowest level boxes by their coordinates.

The CNF contribution of the lattice sum $(\rho_1 | \rho_2\mathbf{L})$ is calculated using the following scheme, depicted schematically in Figure 2. The necessary definitions of multipole and Taylor expansions as well as translation operators are given in ref 31.

Step 1: Formation and translation of multipole expansions.

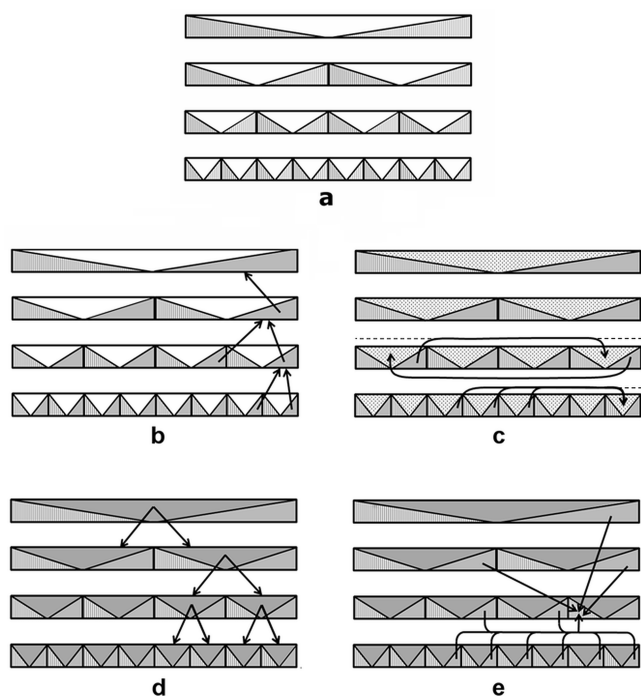


Figure 2. One-dimensional scheme representation of DF-CFMM with $ws = 3$ in the CNF part of the lattice sum ($\rho_1 | \rho_{2L}$): (a) formation of partial multipole expansions of ρ_1 (left lower triangle in each box) and ρ_2 (right lower triangle), (b) upward translation and summation of partial multipole expansions of ρ_2 , (c) formation of Taylor expansions (upper triangle) from partial multipole expansions in well-separated boxes whose parents are not well-separated (dashed lines represent contributions from periodic replicas of the octree), (d) downward shift and summation of Taylor expansions, and (e) calculation of the NF contribution between non-well-separated boxes at all levels and between levels.

This step does not involve any periodic images of ρ_2 and is identical for molecular and periodic systems. For each occupied octree box at all levels multipole expansions about the center of the box are formed for charge distributions comprising ρ_1 and ρ_2 that belong to the given box. These multipole expansions are calculated using the method described in ref 18. The expansions are stored separately for ρ_1 and ρ_2 (see Figure 2a). Multipole expansions of ρ_2 are translated upward applying the “external to external” translation operator³¹ A and accumulated starting from the lowest level of the octree (Figure 2b). After this step each box contains partial multipole expansion of charge distributions of ρ_2 contained in the box itself and in its child boxes.

Step 2: Formation and translation of Taylor expansions.

In this step the partial multipole expansion of ρ_2 about the center of a box i_L (i.e., box i shifted by a lattice vector $L \in \text{CNF}$) is transformed into the Taylor expansion about the center of box j . This transformation uses the “external to local” operator³¹ B and is performed only for pairs of boxes at the same level of the octree that are well-separated while their parents are not (cf. Figure 2c). Two boxes i_L and j with edge lengths L_i and L_j , respectively, are called well-separated if the distance d_{ij} between their centers satisfies the condition

$$d_{ij} \geq_{ws} \frac{(L_i + L_j)}{2} \quad (35)$$

Since the partial multipole expansions along the octree are formed only for ρ_{2L} with $L = 0$ (step 1), this procedure formally involves replica of the octree shifted by $L \in \text{CNF}$. Using the replica allows for a significant speedup of this “external to local” transformation step by first calculating the sum of operators B for all $L \in \text{CNF}$ that fulfill the condition given by eq 35. This sum is then applied to transform partial multipole expansion of ρ_2 about the center of box i into the Taylor expansion about the center of box j for all well-separated i_L in only one step. This is the main difference to the implementation of Kudin et al.^{11,29,30} who sort charge distributions for all $L \in \text{CNF}$ into one bigger octree.

The local Taylor expansions are passed down the tree using the operator C .³¹ This pass starts at the highest level and translates Taylor expansions of each parent box to the centers of the child boxes. The expansions within child boxes are summed up, and the procedure is repeated until the lowest level of the octree is reached (Figure 2d).

Step 3: Far-field contribution.

After steps 1 and 2 each box within the octree contains partial Taylor expansions of ρ_{2L} from all well-separated boxes i_L for $L \in \text{CNF}$. These expansions are combined with partial multipole expansions of ρ_1 yielding the FF portion of the CNF contribution to lattice sum ($\rho_1 | \rho_{2L}$).

Step 4: Near-field contribution.

In this step the remaining interactions between non-well-separated boxes and within boxes are evaluated by direct integration at all octree levels. This is the most time-consuming step of the algorithm. Note that the NF part includes interactions between boxes at different levels, as depicted in Figure 2e. The resulting NF contribution is combined with the FF part obtained in step 3 yielding the CNF part of the lattice sum ($\rho_1 | \rho_{2L}$).

III. IMPLEMENTATION DETAILS

The methodology described in section II along with the DF-CFMM algorithm have been implemented within the TURBOMOLE³³ program package. In order to save computational time several quantities are precalculated and stored before entering the main SCF loop (i.e., iterative building of the KS matrix and solution of eq 2). The most important is the list of significant basis function products $\mu\nu_L$ which along with the symmetry relation given by eq 9 allows for efficient sparse storage of all matrices in eq 5. This is because the elements of these matrices have significant contribution only for non-vanishing $\mu\nu_L$. The list is created based on the overlap criterion neglecting angular parts of the basis functions. Due to the exponential decay of $\mu\nu_L$ with increasing separation between centers of μ and ν_L the length of this list shows asymptotic $O(N)$ scaling. For local exchange-correlation density functional approximations, the real space density matrix elements, eq 6, may also be stored only for the significant $\mu\nu_L$ products.

The lattice sums in eqs 25–27 involve combinations of different types of charge distributions. It is efficient and yet simple to use the same common set of lattice vectors for partitioning all lattice sums into CNF and CFF portions (cf. eq 33). Similarly, the products $\mu\nu_L$, auxiliary basis functions α , and nuclear point charges are all sorted into one common octree. The direct space elements $S_{\mu\nu}^L$ and $T_{\mu\nu}^L$ of the overlap and kinetic energy matrices, respectively, as well as the combined Coulomb metric matrix $V = V_{\perp} + P_{\parallel}$ along with its Cholesky factorization needed for the solution of eq 23 are evaluated and stored prior to the main SCF loop. The procedure used to

calculate \mathbf{V} is the same as that used in ref 9. Sparse matrix storage is used for all direct space matrices in eq 5.

The procedures used within the main SCF loop are analogous to other implementations.^{2,3} The main difference is the application of DF-CFMM for the evaluation of the Coulomb matrix elements $J_{\mu\nu}^L$ and the Coulomb energy J . The procedure implemented consists of the following steps:

- Calculation of the vector ξ_{\perp} , eq 25. First, the vector $\xi = (\alpha | \rho_L - \tilde{\rho}_{\parallel L})$ is calculated using CFMM. ξ_{\perp} is obtained applying charge projection $\xi_{\perp} = \mathbf{P}_{\perp} \xi$ with the projection matrix \mathbf{P}_{\perp} given in eq 19.

- Solution of eq 23 for the chargeless component c_{\perp} of the auxiliary density using precalculated Cholesky factorization of $\mathbf{V}_{\perp} + \mathbf{P}_{\perp}$.

- Assembly of the final auxiliary density $\tilde{\rho}$ according to eqs 17 and 22.

- Calculation of the Coulomb matrix elements $J_{\mu\nu}^L$ according to eq 26 using CFMM.

- Calculation of the total Coulomb energy J . The second term in eq 27 is evaluated using CFMM. The CNF part of this term is simplified as discussed in ref 9.

The integral in eq 6 is evaluated approximately on grids of sampling points \mathbf{k}_α . Uniform grids are used yielding the same weight V_k/N_k with the total number of k -points, N_k , for each \mathbf{k}_α so that eq 6 may be written as

$$D_{\mu\nu\sigma}^L \approx \frac{1}{N_k} \sum_{\alpha}^{BZ} e^{i\mathbf{k}_\alpha^T \mathbf{L}} D_{\mu\nu\sigma}^{\mathbf{k}_\alpha} \quad (36)$$

In the current implementation, the reciprocal density matrix $D_{\mu\nu\sigma}^{\mathbf{k}_\alpha}$ defined in eq 7 is calculated for occupation numbers $f_{p\sigma}^{\mathbf{k}_\alpha}$ which are constant within each band throughout the BZ. This includes insulators with fully occupied and empty bands. Open shell defect states with a single one-particle energy in k -space can also be investigated. In general, our methods may be augmented with schemes for semiconductors and metals found in refs 34 and 35 and citations therein.

For simplicity, we employ uniform grids centered at the Γ -point using a single odd number k of k -points in each periodic direction such that $N_k = kN_d$, where N_d is the number of periodic directions. Similar to Monkhorst and Pack,³⁶ the positions of the k -points are defined by lattice points \mathbf{G}_α in the k -fold supercell $\mathbf{S}^{(k)}$ of reciprocal space shrunk by the factor $1/k$,

$$\mathbf{k}_\alpha = \frac{1}{k} \mathbf{G}_\alpha \quad \forall \mathbf{G}_\alpha \in \mathbf{S}^{(k)} \quad (37)$$

This definition yields a Γ -centered mesh of points $\mathbf{k}_\alpha = (k_1, k_2, k_3)$ with components

$$k_j = i/k \quad \text{with } i = -\frac{k-1}{2}, \dots, \frac{k-1}{2} \quad (38)$$

in crystal coordinates. Inserting the sampled wave vectors \mathbf{k}_α eq 37, into a plane wave

$$e^{i\mathbf{k}_\alpha^T \mathbf{L}} = e^{i\mathbf{G}_\alpha^T \mathbf{L}/k} \quad (39)$$

yields all plane waves belonging to the chosen k -point grid that are periodic on the lattice of direct space supercell translations $k \cdot \mathbf{L}$, since $e^{i\mathbf{G}^T \mathbf{L}}$ is one by definition. Therefore, the orbitals and the density matrix in eqs 1 and 6 are periodic on the k -fold supercell in direct space.³⁷ The same periodicity is obtained by a Γ -point calculation on the k -fold supercell. In general, all n -

fold supercells with equivalent shape yield the same normalized energy E/n^{N_d} in the SCF calculation, if k is selected such that nk is a constant.

Using the k -points defined previously, the direct space density matrix in eq 36 is obtained from the sparse KS matrix in direct space, eq 5, by the following steps within a loop over k -points $\mathbf{k}=\mathbf{k}_\alpha$:

- Real sparse matrices $F_{\mu\nu\sigma}^L$ and $S_{\mu\nu}^L$ are transformed to complex Hermitian square matrices $\mathbf{F}_\sigma^k \equiv \{F_{\mu\nu\sigma}^k\}$ and $\mathbf{S}^k \equiv \{S_{\mu\nu}^k\}$ according to eqs 3 and 4.

- Reorthonormalized orbitals $\mathbf{C}_\sigma^k = \mathbf{C}_\sigma^L \mathbf{\Lambda}^{-1}$ are formed with upper triangular matrix $\mathbf{\Lambda}$ obtained from the Cholesky factorization $(\mathbf{C}_\sigma^k)^\dagger \mathbf{S}^k \mathbf{C}_\sigma^k = \mathbf{\Lambda}^\dagger \mathbf{\Lambda}$ using complex orbital coefficients $\mathbf{C}_\sigma^k \equiv \{C_{\mu p \sigma}^k\}$ from the previous SCF cycle.

- KS matrix is formed in orbital basis, $\mathbf{F}_\sigma^k = (\mathbf{C}_\sigma^k)^\dagger \mathbf{F}_\sigma^L \mathbf{C}_\sigma^k$ and diagonalized to obtain new orbitals $\tilde{\mathbf{C}}_\sigma^k$ ordered by ascending orbital energies $\epsilon_{p\sigma}^k$.

- Rescaled orbitals $\tilde{\mathbf{C}}_\sigma^k$ are formed by multiplying each p th column in \mathbf{C}_σ^k with the square root of the occupation number, $(f_{p\sigma}^k)^{1/2}$, and neglecting zero columns.

- The real part of the reciprocal density matrix is formed, eq 7, as $\mathbf{D}_{\sigma,R}^k = \tilde{\mathbf{C}}_{\sigma,R}^k (\tilde{\mathbf{C}}_{\sigma,R}^k)^T + \tilde{\mathbf{C}}_{\sigma,I}^k (\tilde{\mathbf{C}}_{\sigma,I}^k)^T$, where R and I indicate real and imaginary parts.

- For $\mathbf{k} = \mathbf{0}$ matrix $\mathbf{D}_{\sigma,R}^k$ is saved for DIIS procedure in the next SCF cycle.

- Rescaled real part of orbitals $\tilde{\mathbf{C}}_{\sigma,R}^k$ is formed by multiplying each p th column in $\mathbf{C}_{\sigma,R}^k$ with the occupation number $f_{p\sigma}^k$ and neglecting zero columns.

- The imaginary part of reciprocal density matrix, eq 7, $\mathbf{D}_{\sigma,I}^k = \tilde{\mathbf{C}}_{\sigma,R}^k (\tilde{\mathbf{C}}_{\sigma,I}^k)^T + \tilde{\mathbf{C}}_{\sigma,I}^k (\tilde{\mathbf{C}}_{\sigma,R}^k)^T$ is formed.

- Real-valued direct space density matrix is accumulated, $D_{\mu\nu\sigma}^L \leftarrow (1/N_k) [\cos(\mathbf{k}^T \mathbf{L}) D_{\mu\nu\sigma,R}^k - \sin(\mathbf{k}^T \mathbf{L}) D_{\mu\nu\sigma,I}^k]$

Symmetric real and antisymmetric imaginary parts of complex Hermitian matrices are stored and used separately in real-valued packed form when possible, i.e., as vectors containing only matrix elements M_{ij} with $i \geq j$. Before the k -point loop, the KS matrix is extrapolated by linear combination of the sparse direct space KS matrices from the current and previous SCF cycles. The linear combination coefficients are determined from $\mathbf{F}_\sigma^{k=0}$, $\mathbf{D}_{\sigma,R}^{k=0}$, and $\mathbf{S}^{k=0}$ according to the DIIS scheme.^{38–41}

All steps of the DF-CFMM scheme as well as the SCF procedure described earlier are parallelized for shared-memory machines using the OpenMP application programming interface.⁴²

IV. COMPUTATIONAL DETAILS

A. Model Systems. The performance and scaling behavior of our implementation is evaluated for different sets of molecular and periodic systems with examples shown in Figure 3.

The molecular models are similar to those used by Sodt et al.¹⁹ and consist of (i) alkane chains, (ii) H-terminated graphite sheets, and (iii) H-terminated diamond chunks. The alkane chains have the composition $C_n H_{2n+2}$, with $n = 10, 20, 40, 60, 120, 160$, and 200 and the longest chain uses 5010 basis and 17024 auxiliary basis functions. The graphite sheets, $C_{16} H_{10}$, $C_{76} H_{22}$, $C_{102} H_{26}$, $C_{184} H_{34}$, and $C_{210} H_{38}$, contain up to 3340 basis and 13266 auxiliary basis functions. The diamond chunks have compositions $C_{11} H_{18}$, $C_{87} H_{64}$, $C_{168} H_{130}$, and $C_{246} H_{184}$. The largest chunk uses 4610 basis and 17214 auxiliary basis functions.

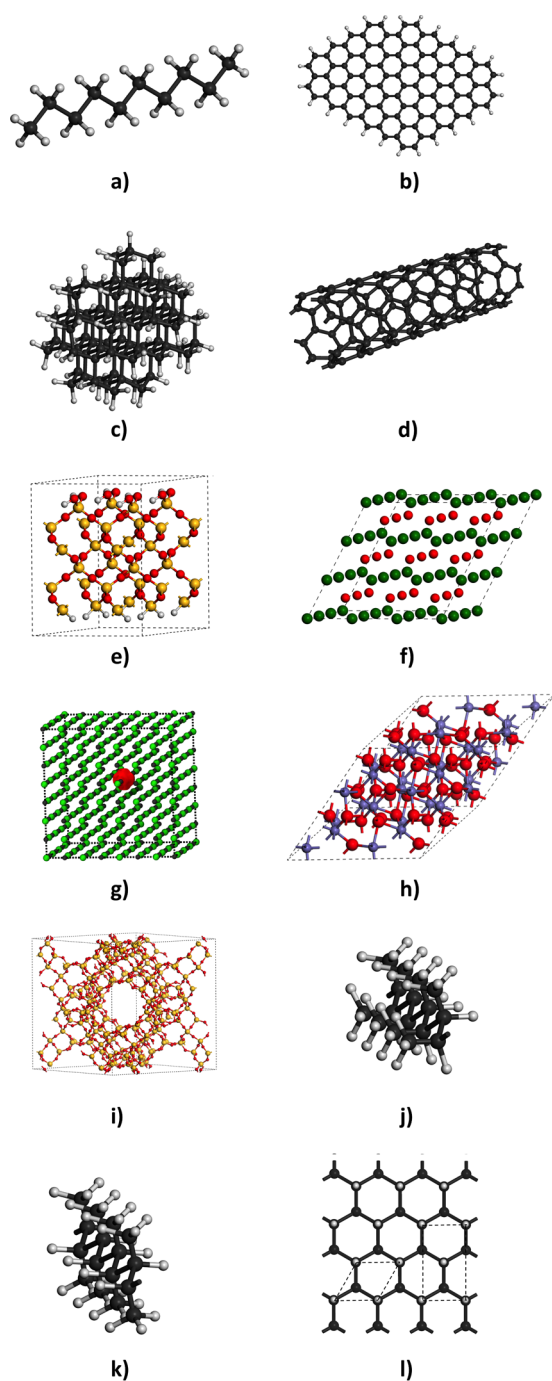


Figure 3. Examples of test molecular and periodic systems: (a) $C_{10}H_{22}$ alkane chain, (b) H-terminated $C_{102}H_{26}$ graphite sheet, (c) H-terminated $C_{87}H_{84}$ diamond chunk, (d) 1D quadrupled supercell of (4,4) carbon nanotube $(C_{16})_n$, (e) 2D 2×2 supercell of hydroxylated α -quartz (0001) $(Si_6O_{12}H_4)_n$ surface, (f) 3D $3 \times 3 \times 3$ supercell of MgO, (g) 3D $4 \times 4 \times 4$ supercell of NaCl $(Na_4Cl_4)_n$ containing a single Cl vacancy along with spin density isosurface, (h) 3D $2 \times 2 \times 2$ supercell of hematite $(Fe_4O_6)_n$, (i) 3D all-silica faujasite $(Si_{192}O_{384})_n$, (j) 1D *cis*- and (k) *trans*-methylated $(C_2H_2(CH_3)_2)_n$ alkane isomers, and (l) 2D primitive and orthogonal unit cells (dashed lines) of $(CH)_n$ graphane.

For 1D periodic calculations a series of single-walled carbon (4,4) armchair nanotubes (SWNT) is constructed. The unit cell of the smallest SWNT with length of 2.46 Å in periodic direction consists of 16 C atoms. The series uses doubled,

tripled, and quadrupled supercells. The largest supercell contains 960 basis and 3904 auxiliary basis functions.

Two-dimensional periodic calculations are performed for models of the hydroxylated α -quartz (0001) surface. The hexagonal surface unit cell (lattice vector $a = 5.01$ Å) of the smallest model has the composition $Si_6O_{12}H_4$. The series consists of (1×1) , (2×2) , (3×3) , and (4×4) surface supercells with the largest model using 5024 basis and 18528 auxiliary basis functions.

The 3D periodic model systems consist of (i) MgO supercells, (ii) supercells of NaCl with a chlorine vacancy (F-center), (iii) hematite supercells with ferromagnetic spin configuration, and (iv) an all-silica faujasite zeolite model. A series of MgO $(1 \times 1 \times 1)$, $(2 \times 2 \times 2)$, $(3 \times 3 \times 3)$, and $(4 \times 4 \times 4)$ supercells is constructed starting from the primitive unit cell containing one MgO unit ($a = b = c = 2.98$ Å and $\alpha = \beta = \gamma = 60^\circ$). The largest supercell uses 2496 basis and 8320 auxiliary basis functions. The cubic unit cell of the smallest NaCl model (lattice vector $a = 5.62$ Å) has the composition Na_4Cl_4 . The series uses $(1 \times 1 \times 1)$, $(2 \times 2 \times 2)$, $(3 \times 3 \times 3)$, and $(4 \times 4 \times 4)$ supercells with the largest model containing 11008 basis and 33792 auxiliary basis functions. For each size the corresponding model of F-center is constructed by a removal of one Cl atom. In case of the $(4 \times 4 \times 4)$ NaCl supercell this results in a system comprising 10985 basis and 33729 auxiliary basis functions. For hematite a series of $(1 \times 1 \times 1)$, $(2 \times 2 \times 2)$, $(3 \times 3 \times 3)$, $(3 \times 3 \times 3)$, and $(4 \times 4 \times 4)$ supercells is constructed starting from the primitive unit cell ($a = b = c = 5.43$ Å and $\alpha = \beta = \gamma = 55.2^\circ$) with the composition Fe_4O_6 . The largest hematite model uses 19072 basis and 53632 auxiliary basis functions. For evaluation of the parallel performance of our implementation, an all-silica faujasite zeolite model constructed from the experimental crystal structure⁴³ is used. Its cubic unit cell with $a = 25.0$ Å shown in Figure 3 contains 576 atoms and 9408 basis and 35520 auxiliary basis functions.

For a better comparison between systems of different sizes the performance and scaling behavior of our implementation is evaluated using single point energy calculations at the Γ -point only. The energy convergence with the number of k-points is examined using two periodic isomeric alkane chains with unit cell composition $C_2H_2(CH_3)_2$ and lattice constant $a = 2.54$ Å containing methyl groups on one side (*cis*) or on opposite sides (*trans*) of the plane formed by the carbon zigzag backbone (see Figure 3). The hexagonal graphane sheet shown in Figure 3 is also tested for different unit cell definitions. The primitive cell is defined with $a = b = 2.55$ Å and C–C and C–H bond distances of 1.54 and 1.12 Å, respectively. This model contains 6-fold carbon rings in chair conformation. Finally, the $Fm\bar{3}m$ structure of magnesium oxide, MgO, and cubic diamond ($Fd\bar{3}m$) are investigated with Mg–O and C–C bond distances of 2.11 and 1.54 Å, respectively.^{44,45} The corresponding unit cells are defined with $a = 4.28$ Å for MgO and $a = 3.57$ Å for diamond. In the case of diamond, additional calculations using primitive rhombohedral cell defined with $a = b = c = 2.52$ Å and $\alpha = \beta = \gamma = 60^\circ$ are conducted.

B. Methods and Basis Sets. All calculations employ the Becke–Perdew (BP86) exchange-correlation functional.^{46,47} The performance and scaling behavior is investigated using the split-valence plus polarization (def2-SVP) basis^{48,49} and auxiliary basis⁵⁰ sets. For 3D periodic dense models (MgO, NaCl with F-center, and hematite) the pob-TZVP⁵¹ basis sets along with appropriate auxiliary basis sets⁵⁰ are employed to avoid SCF convergence problems caused by small orbital

exponents. The calculations are performed on the 2.40 GHz Intel Xeon E5-2695v2 CPU. Parallel efficiency is evaluated using up to 24 CPU cores. Timings for molecular systems (alkane chains, graphite sheets, and diamond chunks) are obtained using a single CPU core. Calculations for the remaining systems are performed using 4 CPU cores. Calculations on the F-center and hematite models use spin unrestricted KS formalism. For remaining systems spin restricted calculations are performed. Unless stated otherwise numerical integration of the exchange-correlation term uses grids of size 3.^{5,8} DF-CFMM uses well-separateness criterion $w_s = 3$, the order of multipole expansions $L = 20$, and the threshold $\varepsilon = 1 \times 10^{-9}$ in eqs 29 and 31. For all investigated systems these settings yield energies that differ from reference values ($w_s = 4$, $L = 30$, $\varepsilon = 1 \times 10^{-11}$) by less than $10^{-6} E_h$.

Test calculations for the accuracy of k-point sampling employ the same basis sets, except for the *cis*- and *trans*-alkane models which use the triple- ζ plus polarization basis sets (def2-TZVP).^{48,49} The total density of states (TDOS) is calculated as superposition of Gaussian distributions centered at the one-particle energies of the sampled k-points. The Gaussian distributions are used as approximates to the δ distributions, which are exact in the limiting case of high-density k-point grids.⁵² The Fermi level is placed in the middle of the gap, which is the limit of the chemical potential as temperature tends to zero.⁵²

V. PERFORMANCE AND ACCURACY

A. Performance. Performance of our implementation for the formation of the KS matrix is evaluated for a series of model systems described in section A with examples depicted in Figure 3. CPU times per SCF iteration for the Coulomb term, its NF and FF parts, and the exchange-correlation contribution (matrix and energy) are shown in Figures 4–6. The scaling exponents are determined using logarithmic fit for the largest models in each series. The overlap and kinetic energy matrices are precalculated before entering the SCF loop and stored in a sparse format. Their evaluation time is negligible. The computational cost of the remaining SCF procedures, such as formation of the density matrix, orbitals orthonormalization, and KS matrix diagonalization, is not included since even for the largest test systems the computational cost of the KS matrix formation is the dominant one as shown in Table 1. For the largest model, a $(4 \times 4 \times 4)$ hematite supercell with 19072 basis functions, it takes about 73% of the total CPU time per SCF iteration. An exception is the alkane series, for which the computational cost of KS matrix formation is comparable to its diagonalization (cf. Table 1). We note that for even larger systems the diagonalization of the KS matrix will become the dominant step due to its cubic scaling. However, the availability of an efficient procedure for the formation of the KS matrix opens a possibility for future implementation of low-scaling methods that avoid matrix diagonalization.⁵³

As shown in Figure 4a the evaluation of the KS matrix for alkane chains achieves approximately linear scaling. For all molecules in the series the CPU times for calculation of the Coulomb and exchange-correlation contributions are similar. Alkanes with over 3000 basis functions exhibit sublinear scaling for the Coulomb term with dominant contribution from the NF part. CPU times for the evaluation of the FF part are less than 5 s for the largest alkane in the series.

Figure 4b shows CPU times per SCF iteration for the series of H-terminated graphite sheets. For small models the scaling

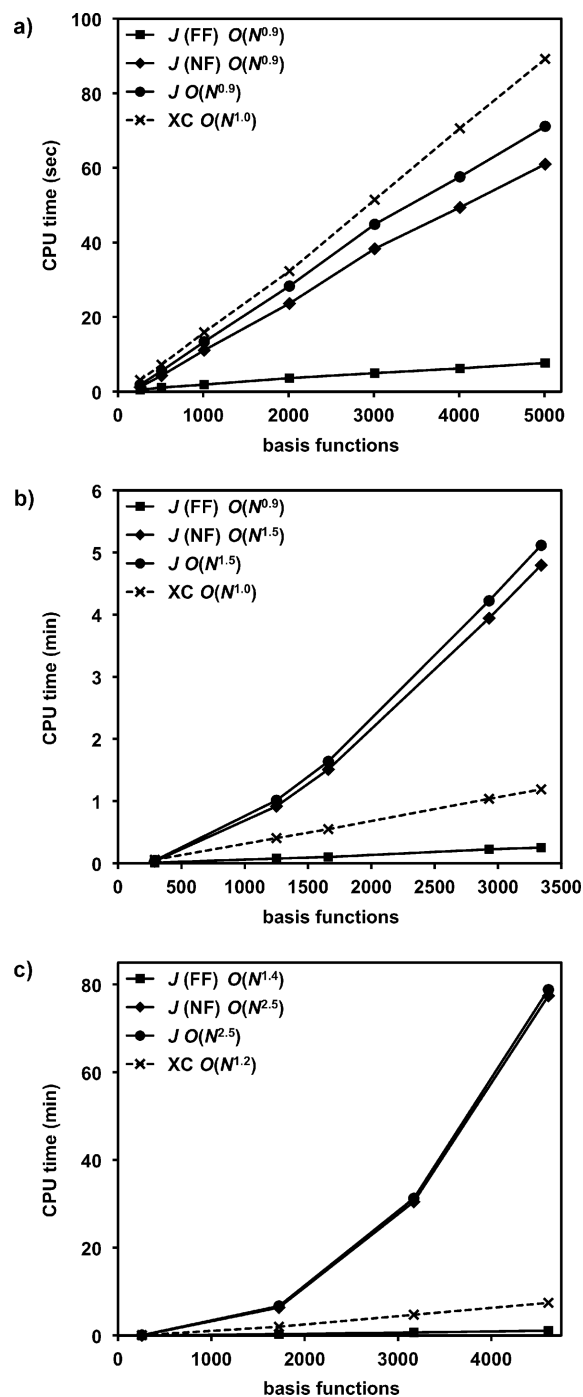


Figure 4. CPU timing per SCF iteration for calculation of the FF and NF Coulomb (J) contributions, the total Coulomb term (matrix and energy), and the exchange-correlation (XC) term for (a) alkane chains, (b) graphite sheets, and (c) diamond pieces.

of the evaluation of the Coulomb term is approximately quadratic and decreases to 1.5 for larger sheets. The computational cost of the NF part is dominant whereas the FF contribution takes only a small portion of the CPU time. The exchange-correlation term shows perfect linear scaling with the system size and is evaluated significantly faster than the Coulomb contribution.

As shown in Figure 4c the scaling behavior of the formation of the KS matrix for the series of diamond pieces is less favorable than for other molecular systems. CPU times are

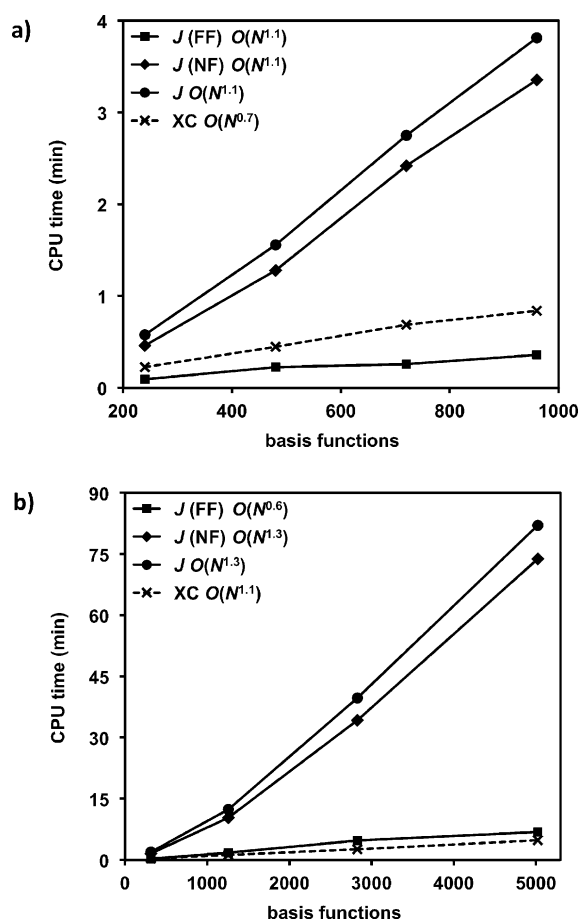


Figure 5. CPU timing per SCF iteration for calculation of the FF and NF parts of the CNF Coulomb (J) contribution, the total Coulomb term (matrix and energy), and the exchange-correlation (XC) term for (a) 1D carbon nanotubes and (b) 2D hydroxylated α -quartz (0001) surfaces.

dominated by the NF part of the Coulomb term, with scaling exponents decreasing from 2.7 to 2.5 with increasing system size. The computational cost of the Coulomb FF contribution and the exchange-correlation part is small with scaling exponents of 1.4 and 1.2, respectively.

Figure 5a shows CPU times per SCF iteration for the series of 1D periodic carbon nanotubes. The times for the formation of the KS matrix are dominated by the Coulomb NF contribution. Calculation of both the NF and FF parts shows nearly $O(N)$ complexity. Formation of the exchange-correlation term achieves a sublinear scaling.

The series of 2D periodic hydroxylated α -quartz (0001) surface models shows similar behavior with the major part of computational time for the formation of the KS matrix due to the Coulomb NF contribution with a scaling exponent of 1.3 (cf. Figure 5b). Evaluation of the exchange-correlation and Coulomb FF contributions takes only a small portion of the CPU time with the later showing a sublinear scaling behavior.

The evaluation of the Coulomb NF part is also the dominant step for the series of 3D periodic models of magnesium oxide (cf. Figure 6a). Similar to other periodic test systems, the calculation of the FF and exchange-correlation parts takes only a little CPU time. The computational effort for the formation of the KS matrix shows nearly perfect linear scaling behavior with the system size.

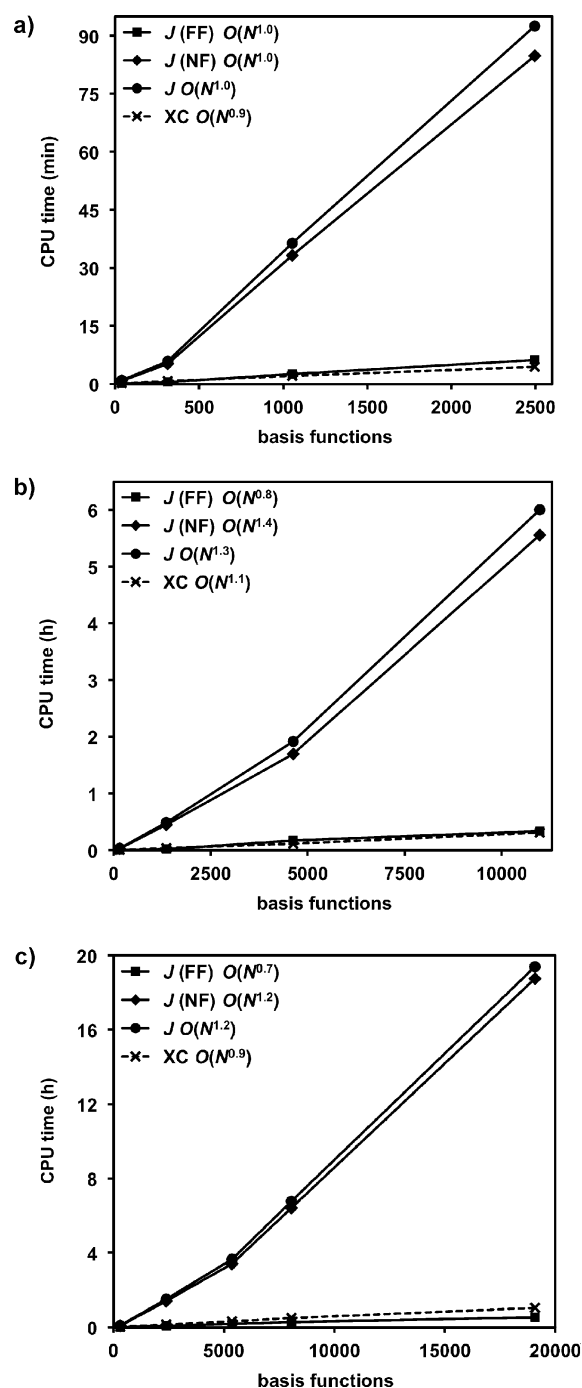


Figure 6. CPU timing per SCF iteration for calculation of the FF and NF parts of the CNF Coulomb (J) contribution, the total Coulomb term (matrix and energy), and the exchange-correlation (XC) term for supercells of 3D systems of (a) MgO, (b) NaCl containing the F-center, and (c) hematite.

Similarly, the timings for the evaluation of the KS matrix in the case of the series of F-center models are dominated by the Coulomb NF contribution with scaling exponent of 1.4 (Figure 6b). The computational cost of the exchange-correlation term and the Coulomb FF part is negligible. In both systems calculation of the Coulomb FF contribution achieves a sublinear scaling.

Figure 6c shows CPU times per SCF iteration for the series of hematite supercells. CPU times are dominated by the NF part of the Coulomb term, with scaling exponent of 1.2. For the

Table 1. Performance Results for the Largest Model Systems Studied: Number of Atoms (N_{at}), Number of Basis (N_{bf}) and Auxiliary (N_{aux}) Basis Functions, CPU Times (min) per SCF Iteration for Calculation of the Coulomb (t_{f}) and Exchange-Correlation (t_{xc}) Terms as Well as the Whole SCF Iteration (t_{SCF})

system	N_{at}	N_{bf}	N_{aux}	t_{f}	t_{xc}	t_{SCF}
alkane chain $\text{C}_{200}\text{H}_{402}$	602	5010	17024	1.2	1.5	4.7
graphite sheet $\text{C}_{210}\text{H}_{38}$	248	3340	13266	5.1	1.2	6.9
diamond chunk $\text{C}_{246}\text{H}_{184}$	430	4610	17214	78.8	7.5	87.9
1D carbon nanotube $(\text{C}_{64})_n$	64	960	3904	3.8	0.8	4.7
2D α -quartz surface $(\text{Si}_{96}\text{O}_{192}\text{H}_{64})_n$	352	5024	18528	82.1	4.8	89.4
3D magnesium oxide $(\text{Mg}_{64}\text{O}_{64})_n$	128	2496	8320	92.5	4.5	97.5
3D NaCl with F-center $(\text{Na}_{256}\text{Cl}_{1255})_n$	511	10985	33729	360.1	18.4	449.6
3D hematite $(\text{Fe}_{256}\text{O}_{384})_n$	640	19072	53632	1163.3	62.6	1681.8
3D faujasite $(\text{Si}_{192}\text{O}_{384})_n$	576	9408	35520	144.6	4.9	171

Table 2. Three-Dimensional Periodic Faujasite Zeolite Model (576 Atoms and 9408 Basis Functions): Wall Times (min) per SCF Iteration for Parallel Computing of the Near-Field (t_{f}^{NF}) and Far-Field (t_{f}^{FF}) Contributions to the Coulomb Matrix (J), Equation 26, and the ξ_{\perp} Vector ($t_{\xi_{\perp}}^{\text{NF}}$ and $t_{\xi_{\perp}}^{\text{FF}}$), Equation 25^a

CPU cores	t_{f}^{NF}	t_{f}^{FF}	$t_{\xi_{\perp}}^{\text{NF}}$	$t_{\xi_{\perp}}^{\text{FF}}$	t_{f}	t_{xc}	E , %
1	66.5	6.5	62.0	6.5	144.6	4.9	
2	37.6	3.2	31.8	3.2	77.4	2.6	93
4	18.8	1.6	15.9	1.6	38.7	1.3	93
6	13.5	1.1	10.7	1.1	26.9	0.9	90
12	8.6	0.5	5.4	0.5	15.3	0.5	78
24	3.6	0.3	2.8	0.3	7.1	0.2	85

^aFor comparison total wall times for calculation of the Coulomb t_{f} and exchange-correlation t_{xc} terms along with the overall parallel efficiency E are given.

Table 3. Relative BP86 Single-Point Energies $E(\text{cis}) - E(\text{trans})$ (kJ/mol) of Periodic Alkanes (see Figure 3) for k points

k	SVP	TZVP
1	62.5928	66.2368
5	3.5273	6.7358
9	3.4277	6.5991
13	3.4606	6.6424
17	3.4618	6.6441
101	3.4618	6.6441

FF part of the Coulomb contribution and the exchange-correlation term a sublinear scaling behavior is achieved with scaling exponents of 0.7 and 0.9, respectively.

In summary, for all systems the formation of the KS matrix is dominated by the Coulomb NF contribution. For molecular systems it shows an increase of computational complexity with increasing “dimensionality”—from linear alkane models to bulky diamond pieces. For systems comprising about 3000 basis functions the CPU time for the formation of the KS matrix grows from less than 1 to approximately 30 min for alkane and diamond chunk, respectively. Similarly, the scaling exponents increase from 1.0 to 2.5. The main reason for this behavior is a steep increase of the computational cost of the NF Coulomb contribution due to an increasing number of overlapping charge distributions (basis functions pairs and auxiliary basis functions). This effect is more important for compact systems and diminishes with increasing system size as demonstrated by the decreasing scaling exponents for H-

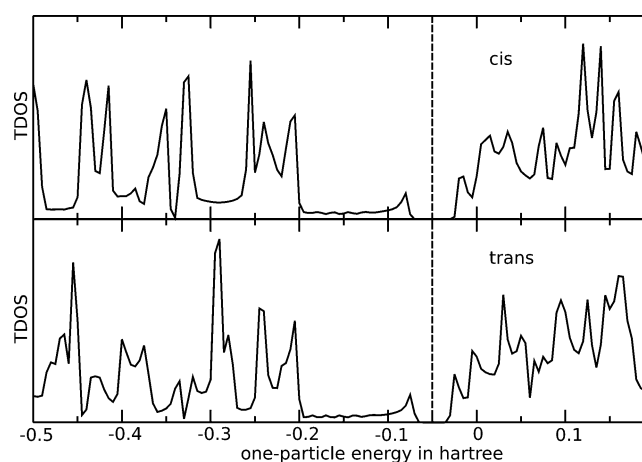


Figure 7. TDOS of *cis*- and *trans*-alkanes (shown in Figure 3) with Fermi level (dashed line) obtained from BP86/TZVP calculations.

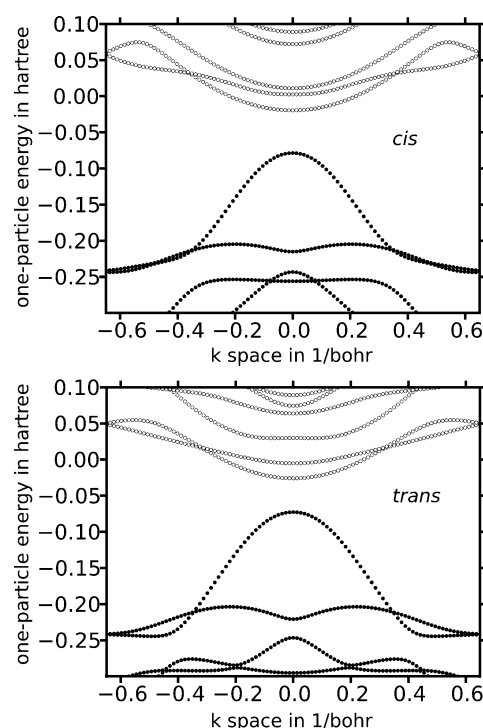


Figure 8. Valence (dark) and conduction bands (light) of *cis*- and *trans*-alkanes (shown in Figure 3) obtained from BP86/TZVP calculations.

Table 4. Normalized BP86/SVP Single-Point Energies in E_h , Relative to $E_0 = -154 E_h$, of Graphane for $(k \times k)$ Grids, Primitive Cell (1×1) , Its Supercells (3×3) and (5×5) , and Orthogonal Cell (ortho; See Figure 3)^a

k	(1×1)	(3×3)	(5×5)	ortho
1	-0.0512169	-0.7293277	-0.7502349	-0.4607915
3	-0.7293277	-0.7513577	-0.7513664	-0.7438141
5	-0.7502349	-0.7513664	-0.7513664	-0.7509877
9	-0.7513576	-0.7513664	:	-0.7513635
15	-0.7513663	:	:	-0.7513664

^aAdditional factor 2 included in the normalized energies of cells (1×1) , (3×3) , and (5×5) .

terminated graphite sheets and diamond chunks. In the case of alkane chains the formation of the KS matrix is particularly efficient due to the linear structure resulting in a small number of overlapping charge distributions. The sublinear scaling behavior observed for alkane chains comprising over 3000 basis functions is most probably caused by CPU cache effects.

In periodic systems the computational time for the formation of the KS matrix grows with increasing periodicity from 1D (carbon nanotubes) to 3D (MgO, F-center models and hematite supercells). For about 1000 basis functions it takes about 4 min for carbon nanotube, 10 min for hydroxylated α -quartz (0001) surface, and over a half-hour for 3D systems. Similar to the molecular case this effect arises from a larger number of near-field Coulomb interactions in denser systems. The scaling exponents for all periodic model systems are close to 1.0, with the highest value of 1.4 for F-center models. For 2D and 3D systems the Coulomb FF and the exchange-correlation contributions are evaluated within virtually negligible CPU time.

The parallel efficiency of our implementation is examined using the 3D all-silica faujasite model (Figure 3i) with the unit cell containing 576 atoms and 9408 basis functions. Table 2 shows wall times for the evaluation of the Coulomb and exchange-correlation contributions on up to 24 CPU cores. The overall parallel performance is very good with efficiency of 85% for 24 cores. Parallel calculation of the KS matrix using 24 cores takes only about 7.5 min.

B. k-Point Sampling. The first test for the wave vector implementation starts with the two periodic isomeric alkane chains shown in Figure 3. The relative energies of these structures are presented in Table 3 demonstrating rapid energy convergence with the number of k-points for double- and triple- ζ basis sets. The TDOS and the band structures near the Fermi energy of the converged triple- ζ calculations are shown

Table 6. Single-Point BP86/SVP Energies in E_h , Relative to $E_0 = -304 E_h$, of Diamond for Rhombohedral and Cubic Cell and k k-Points in Each Periodic Direction

DFT grid	k	rhombohedral ^a	cubic ^b	abs. deviation
3	1	+3.6209449	-0.1564608	
	3	-0.6402795	-0.7221706	
	5	-0.7163772	-0.7242943	
	9	-0.7209168	-0.7243362	
	15	-0.7209493	-0.7243363	
	21	-0.7209494	-0.7243363	3.4×10^{-3}
5	21	-0.7210545	-0.7215739	5.2×10^{-4}
7	21	-0.7210127	-0.7210678	5.5×10^{-5}

^aEnergies include factor 4. ^bNormalized energy of the 3-fold cubic supercell is $-304.7221706 E_h$ with Γ -point only.

in Figures 7 and 8 predicting isolators with direct KS band gaps at the Γ -point of 1.8 and 1.4 eV for the *cis*- and *trans*-methylated alkanes, respectively.

Next, energies from supercell and k-point calculations are compared to assess the numerical accuracy of our electronic structure approach. As explained in section III, electronic structure calculations of the same type nk yield equal normalized energies. This relation is demonstrated on the graphane sheet shown in Figure 3 by the energies listed in Table 4. For example, the Γ -point calculation of the 5-fold supercell and the (5×5) k-point calculation using the primitive reference cell have the same energy of $-154.7502349 E_h$. The converged energies differ at most by $4 \times 10^{-8} E_h$ independent of the unit cell definition.

Similar tests are conducted with magnesium oxide and diamond cells; the results are given in Tables 5 and 6. The energies for different MgO cells are virtually converged within $nk = 11$ and calculations of the same type nk have the same energy within $10^{-7} E_h$. For diamond, the cubic cell computed with $k = 3$ agrees with the energy from a Γ -point calculation of the 3-fold supercell; however, energies of rhombohedral and cubic cells converged with the number of k-points differing by $3.4 \times 10^{-3} E_h$ when DFT grids of size 3 were used. The diamond example documents inaccuracy issues well-known for the numerical integration of the exchange-correlation term.⁵⁴ Integration errors of the same magnitude are also obtained from comparable molecular calculations (see Supporting Information). The deviation between rhombohedral and cubic cell energies decreases systematically for larger DFT grids.

Table 5. Normalized BP86/TZVP Single-Point Energies in E_h , Relative to E_0 , of MgO for k k-Points in Each Periodic Direction and n -Fold Cells

1D, $E_0 = -1101$				2D, $E_0 = -1102$				3D, $E_0 = -1102$			
k	$n = 1$	k	$n = 3$	k	$n = 1$	k	$n = 3$	k	$n = 1$	k	$n = 3$
1	-0.8828346	1	-0.9320707	1	+0.0512312	1	-0.0855391	1	+0.0749854	1	-0.1735513
3	-0.9320707	3	-0.9325479	3	-0.0855391	3	-0.0860178	3	-0.1735513	3	-0.1739073
5	-0.9325122	5	-0.9325484	5	-0.0859945	5	-0.0860180	5	-0.1738899	5	-0.1739078
7	-0.9325441		$n = 5$	7	-0.0860160		$n = 5$	7	-0.1739054	:	:
9	-0.9325479	1	-0.9325122	9	-0.0860178	1	-0.0859944	9	-0.1739074		
11	-0.9325483	3	-0.9325484	11	-0.0860180	3	-0.0860180	11	-0.1739077		
13	-0.9325484		$n = 13$	13	-0.0860180		$n = 9$	13	-0.1739078		
101	-0.9325484	1	-0.9325484	15	-0.0860180	1	-0.0860178	15	-0.1739078		

VI. SUMMARY

An implementation of Kohn–Sham density functional theory with Gaussian-type orbitals as basis functions is reported that treats molecular and periodic systems of any dimensionality on an equal footing. Its core is a combination of density fitting (DF) and the continuous fast multipole method (CFMM) applied for evaluation of the Coulomb term. In this DF-CFMM approach the Coulomb interactions are partitioned into near- and far-field contributions. The far-field part is evaluated very efficiently using a hierarchy of multipole expansions whereas the calculation of the near-field contribution employs the DF scheme. For systems with periodic boundary conditions both contributions are evaluated entirely in direct space. The DF procedure is based on the decomposition of auxiliary density into charged and chargeless parts. In periodic systems the charged component is constrained by using the Coulomb metric. The chargeless part is determined variationally and involves only convergent lattice sums. For periodic systems a k-point sampling scheme has been implemented allowing for consistent results across different definitions of unit cells. A high efficiency of our implementation and favorable scaling behavior for the formation of the KS matrix approaching $O(N)$ has been demonstrated for various molecular and periodic systems including a 3D hematite model containing 640 atoms and 19072 basis and 53632 auxiliary basis functions. The favorable computational performance of the DF-CFMM scheme makes it competitive to DF algorithms based on plane waves,^{27,28} in particular for sparsely packed 3D, 2D, and 1D periodic systems.

Finally, our implementation opens a way for the development of efficient GTO-based code for quantum chemical methods in periodic systems. Currently we are expanding our code to include energy gradients, exact exchange and hybrid functionals.

■ ASSOCIATED CONTENT

Supporting Information

Text discussing the accuracy assessment of exchange-correlation energies and figures showing diamond cubic cell with crystal axes a , b , and c and the $C_{156}O_{64}H_{70}$ cluster. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00252.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: marek.sierka@uni-jena.de.

Funding

Financial support from the Fonds der Chemischen Industrie and Turbomole GmbH is gratefully acknowledged.

Notes

The authors declare the following competing financial interests: M.S. is a cofounder and CEO of Turbomole GmbH, which may potentially benefit from the research results.

■ REFERENCES

- (1) Evarestov, R. A. *Quantum Chemistry of Solids*; Springer: Berlin, 2012; Vol. 153.
- (2) Kudin, K. N.; Scuseria, G. E. *Phys. Rev. B* **2000**, *61*, 16440–16453.
- (3) Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalieri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; de la Pierre, M.; D'Arco, P.; Noël, Y.; Causà, M.; Rérat, M.; Kirtman, B. *Int. J. Quantum Chem.* **2014**, *114*, 1287–1317.

- (4) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *Chem. Phys. Lett.* **1996**, *257*, 213–223.
- (5) Treutler, O.; Ahlrichs, R. *J. Chem. Phys.* **1995**, *102*, 346–354.
- (6) Towler, M. D.; Zupan, A.; Causà, M. *Comput. Phys. Commun.* **1996**, *98*, 181–205.
- (7) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.
- (8) Burow, A. M.; Sierka, M. *J. Chem. Theory Comput.* **2011**, *7*, 3097–3104.
- (9) Burow, A. M.; Sierka, M.; Mohamed, F. *J. Chem. Phys.* **2009**, *131*, No. 214101.
- (10) White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M. *Chem. Phys. Lett.* **1994**, *230*, 8–16.
- (11) Kudin, K. N.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *121*, 2886–2890.
- (12) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51–53.
- (13) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys.* **1973**, *2*, 41–51.
- (14) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396–3402.
- (15) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
- (16) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652–660.
- (17) Jung, Y.; Sodt, A.; Gill, P. M. W.; Head-Gordon, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692–6697.
- (18) Sierka, M.; Hogeckamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136–9148.
- (19) Sodt, A.; Subotnik, J. E.; Head-Gordon, M. *J. Chem. Phys.* **2006**, *125*, 194109.
- (20) Varga, Š. *Int. J. Quantum Chem.* **2008**, *108*, 1518–1527.
- (21) Maschio, L.; Usyat, D. *Phys. Rev. B* **2008**, *78*, No. 073102.
- (22) Jaffe, J. E.; Hess, A. C. *J. Chem. Phys.* **1996**, *105*, 10983–10998.
- (23) Dunlap, B. I.; Rösch, N.; Trickey, S. B. *Mol. Phys.* **2010**, *108*, 3167–3180.
- (24) Izmaylov, A. F.; Scuseria, G. E. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3421–3429.
- (25) Ewald, P. P. *Ann. Phys.* **1921**, *369*, 253–287.
- (26) Franchini, M.; Philipsen, P. H. T.; van Lethe, E.; Visscher, L. *J. Chem. Theory Comput.* **2014**, *10*, 1994–2004.
- (27) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- (28) Pulay, P. *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*; Springer: Dordrecht, The Netherlands, 2011; Chapter 1, pp 142–143.
- (29) Kudin, K. N.; Scuseria, G. E. *Chem. Phys. Lett.* **1998**, *289*, 611–616.
- (30) Kudin, K. N.; Scuseria, G. E. *Chem. Phys. Lett.* **1998**, *283*, 61–68.
- (31) White, C. A.; Head-Gordon, M. *J. Chem. Phys.* **1994**, *101*, 6593–6605.
- (32) White, C. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1996**, *257*, 647–650.
- (33) *TURBOMOLE developer version*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, and TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- (34) Pisani, C.; Dovesi, R. *Int. J. Quantum Chem.* **1980**, *17*, 501–516.
- (35) Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- (36) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188–5192.
- (37) Sundararaman, R.; Arias, T. A. *Phys. Rev. B* **2013**, *87*, No. 165122.
- (38) Pulay, P. *Chem. Phys. Lett.* **1980**, *73*, 393–398.
- (39) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556–560.
- (40) Kudin, K. N. *Linear Scaling Density Functional Theory with Gaussian Orbitals and Periodic Boundary Conditions*. Ph.D. thesis, Rice University, Houston, TX, USA, 2000.

- (41) Cancès, E.; Defranceschi, M.; Kutzelnigg, W.; Le, C.; Maday, Y. In *Special Volume: Computational Chemistry*, Handbook of Numerical Analysis series; Le Bris, C., Ciarlet, P. G., Eds.; Elsevier: Amsterdam, The Netherlands, 2003; Vol. X, pp 121–122.
- (42) OpenMP Architecture Review Board. *OpenMP Application Program Interface*, Version 4.0, July 2013. <http://www.openmp.org>.
- (43) Baerlocher, C.; Meier, W. M.; Olson, D. H. *Atlas of Zeolite Framework Types*, 6th ed.; Elsevier: Amsterdam, The Netherlands, 2007.
- (44) Hazen, R. M. *Am. Mineral.* **1976**, *61*, 266–271.
- (45) Wyckoff, R. W. G. *Crystal structures*, 2nd ed.; Interscience: New York, 1963; Vol. 1.
- (46) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (47) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (48) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (49) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (50) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (51) Peintinger, M. F.; Oliveira, D. V.; Bredow, T. *J. Comput. Chem.* **2013**, *34*, 451–459.
- (52) Ashcroft, N. W.; Mermin, N. D. *Solid State Physics*; Thomson Learning: Stamford, CT, USA, 1976; Chapter 8, pp 142–143.
- (53) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085–1123.
- (54) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.