Computer Physics Communications 192 (2015) 60–69

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Hybrid functionals for large periodic systems in an all-electron, numeric atom-centered basis framework

Sergey V. Levchenko [a,*], Xinguo Ren [b], Jürgen Wieferink [a], Rainer Johanni [a], Patrick Rinke [a], Volker Blum [c], Matthias Scheffler [a]

[a] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany

[b] Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui, 230026, China

[c] Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA

ARTICLE INFO

Article history:
Received 16 February 2015
Accepted 21 February 2015
Available online 2 March 2015

Keywords:
Density-functional theory
Exact exchange
Hartree–Fock approximation
Hybrid functionals
Numeric atomic orbitals
All-electron
Linear scaling

ABSTRACT

We describe a framework to evaluate the Hartree–Fock exchange operator for periodic electronic-structure calculations based on general, localized atom-centered basis functions. The functionality is demonstrated by hybrid-functional calculations of properties for several semiconductors. In our implementation of the Fock operator, the Coulomb potential is treated *either* in reciprocal space *or* in real space, where the sparsity of the density matrix can be exploited for computational efficiency. Computational aspects, such as the rigorous avoidance of on-the-fly disk storage, and a load-balanced parallel implementation, are also discussed. We demonstrate linear scaling of our implementation with system size by calculating the electronic structure of a bulk semiconductor (GaAs) with up to 1,024 atoms per unit cell without compromising the accuracy.

1. Introduction

The Hartree–Fock method is the oldest viable electronic structure approach for practical computations that ensures the fermionic character of the electronic wave function explicitly. For accurate calculations, electron correlation must still be accounted for. In molecular calculations (quantum chemistry), so-called hybrid functionals based on a mixture of non-local exchange and semilocal density-functional theory have long been employed as a *de facto* standard (e.g., the B3LYP [1,2], PBEh [3] (also referred to as PBE0 [4] or PBE1PBE) functionals, or the HSE range-separated hybrid functional [5–7]). For solids, however, the non-local nature of the Coulomb operator has prevented a widespread use of Hartree–Fock-like exchange for a long time. A number of successful hybrid functional implementations for solids (periodic boundary conditions) have been reported in the past several years [8–20]. Aside from ever-improving computational resources, a main driving force is arguably the emergence of the HSE06 functional [5,6], which shows an appealing performance for lattice parameters, cohesive energies, and single-quasiparticle-like band structure properties of semiconductors and insulators. Thus, efficient implementations of hybrid functionals are highly desirable.

Aside from the "fundamental" choice of the right approximation for exchange and correlation, a primary challenge for theory is that of the accuracy of all *other* numerical approximations vs. the efficiency of the method. Efficient all-electron hybrid-functional implementations exist for the linearized augmented plane wave (LAPW) [18,19] and projector-augmented wave (PAW) [16,20] methods, but contain additional approximations for the core–valence electron interaction, and remain comparatively computationally heavy for large and/or low-symmetry structures. Much larger system sizes are attainable by methods which "pseudoize" away the core electrons [17,12,20] or rely on basis sets that allow analytic integration [8,15]. Even then, the attainable problem sizes (number of atoms or molecules in a supercell) in practice remain significantly below what is achieved by semilocal functionals [21,22].

The work presented here employs numeric atom-centered basis functions in an all-electron implementation of electronic-structure theory [23] for hybrid functionals with periodic boundary conditions. Depending on the choice of basis set, total-energy convergence vs. computational speed can be balanced from fast qualitative to meV-level converged accuracy as, e.g., demonstrated

* Corresponding author.
E-mail addresses: levchenko@fhi-berlin.mpg.de (S.V. Levchenko), renxg@ustc.edu.cn (X. Ren), rinke@fhi-berlin.mpg.de (P. Rinke), volker.blum@duke.edu (V. Blum), scheffler@fhi-berlin.mpg.de (M. Scheffler).

http://dx.doi.org/10.1016/j.cpc.2015.02.021
0010-4655/

for benchmark molecules by Delley [24]. Our own implementation, the Fritz Haber Institute *ab initio* molecular simulation (FHI-aims) code [23], features hierarchical numeric atom-centered basis sets for all elements (nuclear weight 1-102) that enables such accurate convergence for large systems and semilocal density-functional theory (DFT) in practice [21,25,26]. Molecules and atomic clusters can be treated with similar accuracy with hybrid functionals, using a "resolution of the identity" (RI) [27–29] approach (also referred to as "variational density fitting") tailored for numeric atom-centered basis functions [30]. Our implementation also allows for valence correlation-consistent atomic orbital basis sets [31] for extrapolating to the complete basis set limit, which is particularly important for perturbation-theory-based methods, such as the random-phase approximation or the *GW* approach. Recently, the RI approach has been also applied to periodic systems [32–35].

The goal of the present paper is to outline the methodological and algorithmic framework that allows us to significantly increase the computational performance of methods based on RI for periodic systems. As in the molecular case, we aim for a numerically accurate all-electron treatment that remains feasible up to relatively large system sizes (hundreds of atoms per unit cell), with low or no spatial symmetry. Based on this framework, we compare two different implementations: One that treats the Coulomb operator in reciprocal space, where it is diagonal, and another that treats the Coulomb operator in real space, where distance-based screening can be exploited to reduce computational time and memory use for larger systems without loss of accuracy. From a practical point of view, we are able to avoid any reliance on intermediate storage to disk (which would imply a very slow communication between the disk and the main memory) of any quantities and can exploit parallel computer architectures with reasonable efficiency for large problems. The performance of the method is verified for bulk semiconductor properties. In particular, calculations of GaAs supercells up to 1024 atoms demonstrate that our implementation scales linearly with the number of atoms, retaining the rigorous accuracy control.

The outline of the paper is as follows. Section 2 introduces the general formalism of the Hartree–Fock exchange in periodic systems. Section 3 describes the resolution of the identity approach and related approximations. Details of the reciprocal- and real-space implementations of the periodic Hartree–Fock exchange are given in Sections 4 and 5, respectively. The analysis of the accuracy and performance of the implementations is presented in Section 6. Finally, our conclusions are summarized in Section 7.

2. Exchange operator in periodic boundary conditions

The fundamental quantity to be expressed in the Hartree–Fock method is the exchange operator:

$$K^{\sigma}_{ij} = \sum_{kl} (ik|lj) D^{\sigma}_{kl}. \tag{1}$$

Here, σ represents the spin index, and D^{σ}_{kl} is the single-particle density matrix of the system. Indices i, j, k, and l refer to individual basis functions in the chosen basis set used to expand the single-particle states $\psi_{n\sigma}(\mathbf{r}) = \sum_{i=1}^{N_b} c^{i}_{n\sigma} \phi_i(\mathbf{r})$. D^{σ}_{kl} is defined as a sum over states n in each spin channel:

$$D^{\sigma}_{kl} = \sum_{n} f_{n\sigma} c^{k}_{n\sigma} c^{l*}_{n\sigma} \tag{2}$$

where $f_{n\sigma}$ are occupation numbers, $0 < f_{n\sigma} \leq 1$. For any given choice of basis functions $\{\phi_i(\mathbf{r})\}$, the key challenge is to represent (or, preferably, entirely circumvent the storage of) the two-electron, four-index *Coulomb integral* $(ij|kl)$:

$$(ij|kl) = \iint \frac{\phi^{*}_{i}(\mathbf{r}) \phi^{*}_{j}(\mathbf{r}) \phi_k(\mathbf{r}') \phi_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{3}$$

The exchange operator can be formally written in the coordinate representation,

$$\hat{K}^{\sigma} \psi_{n\sigma}(\mathbf{r}) = \sum_{m} f_{m\sigma} \int \frac{\psi^{*}_{m\sigma}(\mathbf{r}') \psi_{n\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \psi_{m\sigma}(\mathbf{r}). \tag{4}$$

$\psi_{n\sigma}(\mathbf{r})$ are eigenstates of the generalized Kohn–Sham hamiltonian

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v^{\sigma}_{\text{C}}(\mathbf{r}) + (1-\alpha) v^{\sigma}_{\text{X}}(\mathbf{r}) - \alpha \hat{K}^{\sigma} \right) \psi_{n\sigma}(\mathbf{r}) = \epsilon_{n\sigma} \psi_{n\sigma}(\mathbf{r}) \tag{5}$$

where $n(\mathbf{r}) = \sum_{n\sigma} f_{n\sigma} |\psi_{n\sigma}(\mathbf{r})|^2$ is the total electron density, $v_{\text{ext}}(\mathbf{r})$ is the external potential (including the potential due to atomic nuclei), $v^{\sigma}_{\text{X}}(\mathbf{r})$ and $v^{\sigma}_{\text{C}}(\mathbf{r})$ are the local exchange and correlation potentials, respectively, $\epsilon_{n\sigma}$ are the single-particle energies, and α is a mixing parameter. For hybrid functionals, the total-energy expression reads:

$$E^{\text{tot}} = \sum_{n\sigma} \epsilon_{n\sigma} - \sum_{\sigma} \left(\int ((1-\alpha) v^{\sigma}_{\text{X}}(\mathbf{r}) + v^{\sigma}_{\text{C}}(\mathbf{r})) n^{\sigma}(\mathbf{r}) d\mathbf{r} \right) - \frac{1}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E^{\text{loc}}_{\text{C}} + (1-\alpha) E^{\text{loc}}_{\text{X}} - \alpha E^{\text{HF}}_{\text{X}} \tag{6}$$

where $n^{\sigma}(\mathbf{r})$ is the electron density for spin-channel σ, $E^{\text{loc}}_{\text{X}}$ and $E^{\text{loc}}_{\text{C}}$ are the local exchange and correlation energies, and

$$E^{\text{HF}}_{\text{X}} = -\frac{1}{2} \sum_{kl\sigma} K^{\sigma}_{kl} D^{\sigma}_{kl} \tag{7}$$

is the Hartree–Fock exchange energy. For the PBE0 functional, $\alpha = 0.25$, $E^{\text{loc}}_{\text{X}} = E^{\text{PBE}}_{\text{X}}$, and $E^{\text{loc}}_{\text{C}} = E^{\text{PBE}}_{\text{C}}$ [3,4]. In the HSE functional, the exchange operator is replaced with a screened version by replacing the Coulomb interaction $1/r$ with a short-range interaction $\text{erfc}(\omega r)/r$, where ω is the screening parameter, and the missing long-range exchange piece is added by way of a separate semilocal expansion [5–7]:

$$E^{\text{HSE}}_{\text{X}}(\omega) = E^{\text{PBE}}_{\text{X}} - \alpha E^{\text{PBE,SR}}_{\text{X}}(\omega) + \alpha E^{\text{HF,SR}}_{\text{X}}(\omega) \tag{8}$$

where $E^{\text{HF,SR}}_{\text{X}}$ is the exchange energy calculated with the screened Coulomb interaction kernel, and $E^{\text{PBE,SR}}_{\text{X}}$ is the short-range PBE exchange energy calculated by rescaling the exchange hole by a short-range Coulomb screening factor. Note that Coulomb interaction is screened only in the *exchange* term, but not in the Hartree and correlation terms. As mentioned above, the screening parameter ω is not rigorously grounded by electronic structure theory. Instead, optimum ω and α values are sometimes determined by comparing to experiment or to accurate parameter-free methods, for example derived from many-body perturbation theory, for different system types [36–38]. For systems with translational symmetry, the above equations can be used with the following notational changes:

- The atomic-centered localized basis functions are repeated periodically in real space. The three vectors of the real-space lattice, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, also define a "reciprocal lattice" $\mathbf{b}_1 = 2\pi[\mathbf{a}_2, \mathbf{a}_3]/V$, $\mathbf{b}_2 = 2\pi[\mathbf{a}_3, \mathbf{a}_1]/V$, $\mathbf{b}_3 = 2\pi[\mathbf{a}_1, \mathbf{a}_2]/V$, where square brackets denote vector product, and $V = (\mathbf{a}_1, [\mathbf{a}_2, \mathbf{a}_3])$ is the scalar triple product.
- Bloch's theorem implies that the Kohn–Sham states of the periodic hamiltonian are labeled by a continuous vector-index $\mathbf{k}$ ($\mathbf{k}$-point) and a discrete index n (band index). The unique values of $\mathbf{k}$ can be restricted to only one unit cell of reciprocal space (by convention, the "first Brillouin zone" or Wigner–Seitz cell of the reciprocal lattice). Further, Eq. (5) can be solved independently for each $\mathbf{k}$-point. The single-particle states can be written as $\psi_{\mathbf{k},n\sigma}(\mathbf{r}) = \exp(i\mathbf{k}\mathbf{r}) u_{\mathbf{k},n\sigma}(\mathbf{r})$, where $u_{\mathbf{k},n\sigma}(\mathbf{r})$ is a function with the periodicity of the lattice. In general, $\psi_{\mathbf{k},n\sigma}(\mathbf{r})$ is complex-valued.

- For simplicity, we will define "extended" Bloch-like basis functions, by incorporating the Bloch phase factors right there: $\varphi_{\mathbf{ik}}(\mathbf{r}) = \sum_{\mathbf{R}} \phi_i(\mathbf{r} - \mathbf{R}) \exp(i\mathbf{kR})$, where $\mathbf{R}$ denotes real-space lattice vectors.

For practical calculations, the first Brillouin zone must be sampled on a discrete mesh. For instance, we can define such a mesh by dividing each reciprocal lattice vector into m subdivisions, which then define a grid that spans the first Brillouin zone. Due to the phase factors associated with the states $\psi_{\mathbf{k},n\sigma}(\mathbf{r})$, the resulting wave functions will be exactly periodic with a real-space period m. For instance, an $8 \times 8 \times 8$ $\mathbf{k}$-space grid subdivision corresponds to an exact periodicity across $8 \times 8 \times 8$ real-space unit cells (the so-called Born–von Karman cell). Because of the long-range nature of the Coulomb potential, this imposed periodicity requires a special treatment of the long-wavelength components of the Coulomb operator (as discussed further in the text).

3. Resolution of the identity

The challenge in the HF exchange calculations is to evaluate the two-electron Coulomb integrals as given in Eq. (3). Already in the 1950s [39], it was realized that, since two of the basis functions share the same spatial coordinate in the integral, the set of all basis function products $\{\phi_i(\mathbf{r})\phi_j(\mathbf{r})\}$ must be redundant as the basis set $\{\phi_i(\mathbf{r})\}$ approaches completeness. Thus, numerical efficiency can be notably enhanced when the basis function products are expressed as a linear combination of functions from a separate, auxiliary basis set $\{P_\mu(\mathbf{r})\}$. Even if the convergence requirements for this separate basis set are somewhat stringent if the Coulomb operator is to be expanded accurately, its size will still scale as $O(N)$, not $O(N^2)$ if N is a measure of the system size. This approach is known as "resolution of the identity" (RI), or "variational density fitting", in the literature [40–46]. In the following, we implicitly take into account that our localized basis functions (both main and auxiliary) are real-valued. Generally speaking, the RI approach consists in expanding the basis pair products in terms of the auxiliary basis set

$$\phi_i(\mathbf{r})\phi_j(\mathbf{r}) = \sum_\mu C_{ij}^\mu P_\mu(\mathbf{r}), \tag{9}$$

where C_{ij}^μ are the expansion coefficients to be specified later. This immediately leads to

$$(ij|kl) = \sum_{\mu\nu} C_{ij}^\mu V_{\mu\nu} C_{kl}^\nu, \tag{10}$$

where

$$V_{\mu\nu} = \iint \frac{P_\mu(\mathbf{r})P_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \tag{11}$$

We will discuss different versions of the RI below. They correspond to different strategies for determining the expansion coefficients C_{ij}^μ by minimizing errors due to the residual $\delta\rho_{ij}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_j(\mathbf{r}) - \sum_\mu C_{ij}^\mu P_\mu(\mathbf{r})$ in the RI expansion. A straightforward way to expand the four-index integrals is known as "RI-SVS", where the expansion coefficients are determined as

$$C_{ij}^\mu = \sum_\nu (ij\nu) S_{\nu\mu}^{-1}, \tag{12}$$

where $(ij\nu)$ and $S_{\nu\mu}$ are the three- and two-orbital *overlap* matrices,

$$(ij\nu) = \int \phi_i(\mathbf{r})\phi_j(\mathbf{r})P_\nu(\mathbf{r})d\mathbf{r}, \tag{13}$$

and

$$S_{\mu\nu} = \int P_\mu(\mathbf{r})P_\nu(\mathbf{r})d\mathbf{r}. \tag{14}$$

RI-SVS amounts to minimizing the norm of the residuals $\left(\delta\rho_{ij}\delta\rho_{ij}\right)$.

Another strategy, more accurate particularly for calculating the Coulomb integrals (see Eq. (3)), is an expansion that uses the Coulomb metric. This expansion minimizes the Coulomb repulsion between the residuals $\left(\delta\rho_{ij}|\delta\rho_{ij}\right)$, and eliminates the expansion error. The remaining quadratic error term goes to zero much faster as the auxiliary basis set size (expansion quality) increases. The corresponding expansion coefficients are:

$$C_{ij}^\mu = \sum_\nu (ij|\nu) V_{\nu\mu}^{-1}, \tag{15}$$

where

$$(ij|\nu) = \iint \frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r})P_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}', \tag{16}$$

and the V matrix is already defined in Eq. (11). The "RI-V" version of RI arguably reflects the most accurate RI standard in quantum chemistry today. Our own non-periodic implementation of the Coulomb operator [30] is also based on RI-V and encompasses hybrid functionals, perturbation expressions such as second-order Møller–Pesset perturbation theory (MP2), random-phase approximation (RPA), G_0W_0 or self-consistent GW [47,48] accurately for all practical purposes.

For extended (periodic) systems, straight RI-V does have one disadvantage. If the basis functions i and j are localized, the spatial extent of any given basis function product (i, j) is rigorously bounded to the volume where *both* are non-zero. In RI-V, this property is lost, as any basis function pair (i, j) localized at two specific atoms is now expanded into auxiliary basis functions μ centered at *any other* atom in the system. In an infinite periodic system, such a delocalization is obviously impractical. Based on simple completeness arguments, it should be equally well possible to restrict the auxiliary basis set into which a basis function pair (i, j) is expanded to only the two atoms at which those basis functions are centered. From a formal point of view, such a prescription destroys the rigorous error cancellation properties of the expansion, reintroducing a linear error of the resulting Coulomb matrix elements in the expansion. Nonetheless, in the limit of a large auxiliary basis set the expansion must still be complete. As a benefit, the locality of the two basis-function products in the original four-index Coulomb operator is now preserved in the RI expansion.

We have implemented such a localized resolution-of-identity expression (called "RI-LVL"). Details will be reported in a separate paper [49]; we here rely on the fact that our results for hybrid functionals, given below, show excellent convergence both compared to the literature and within themselves. We note that similar expansion recipes have been explored by others before [50–54]. In our localized RI, we *a priori* require that

$$C_{ij}^\mu = 0, \quad \text{for } \mu \notin \mathcal{P}(IJ), \tag{17}$$

where I, J label the atoms where basis functions i, j are centering, and $\mathcal{P}(IJ)$ denotes the set of auxiliary basis functions belonging to these two atoms. By minimizing the Coulomb self-repulsion $\left(\delta\rho_{ij}|\delta\rho_{ij}\right)$ enforcing condition (17), one arrives at the following "LVL" expansion coefficients

$$C_{ij}^\mu = \sum_{\nu\in\mathcal{P}(IJ)} (ij|\nu) L_{\nu\mu}^{IJ} \tag{18}$$

where $L_{\nu\mu}^{IJ} = \left(V^{IJ}\right)^{-1}$ with V^{IJ} being the V matrix (11) with non-zero entries only for $\mu \in \mathcal{P}(IJ)$. This is the form of the RI expansion that will be used for the remainder of this work.

4. Reciprocal-space formalism for the exchange operator

The Coulomb operator is non-local in real space (*i.e.*, it explicitly depends on $\mathbf{r}$ and $\mathbf{r}'$), but it is diagonal in reciprocal space. For a periodic system, the exact-exchange matrix in reciprocal space can be generally expressed using the RI approach

$$K_{ij}^{\sigma}(\mathbf{k}) = \frac{1}{N_{\mathbf{q}}}\sum_{kl,\mathbf{q}} D_{kl}^{\sigma}(\mathbf{q}) \iint \frac{\varphi_{i\mathbf{k}}^{*}(\mathbf{r})\varphi_{k\mathbf{q}}(\mathbf{r})\varphi_{l\mathbf{q}}^{*}(\mathbf{r}')\varphi_{j\mathbf{k}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$$
$$= \frac{1}{N_{\mathbf{q}}}\sum_{kl,\mathbf{q}} D_{kl}^{\sigma}(\mathbf{q}) \sum_{\mu\nu} C_{ik}^{\mu}(\mathbf{k},\mathbf{q}) V_{\mu\nu}(\mathbf{k}-\mathbf{q}) C_{lj}^{\nu}(\mathbf{q},\mathbf{k}), \qquad (19)$$

where $N_{\mathbf{q}}$ is the number of $\mathbf{k}$-points used to sample the first Brillouin zone, and $C_{ik}^{\nu}(\mathbf{k},\mathbf{q})$ are the RI expansion coefficients for a periodic system,

$$\varphi_{i\mathbf{k}}^{*}(\mathbf{r})\varphi_{k\mathbf{q}}(\mathbf{r}) = \sum_{\mu} C_{ik}^{\mu}(\mathbf{k},\mathbf{q}) P_{\mu}^{\mathbf{q}-\mathbf{k}}(\mathbf{r}), \qquad (20)$$

where both the single-particle and auxiliary basis functions now carry a $\mathbf{k}$ index referring to a vector in the first Brillouin zone. $P_{\mu}^{\mathbf{k}}(\mathbf{r})$ is the Bloch summation of auxiliary basis functions,

$$P_{\mu}^{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} P_{\mu}(\mathbf{r}-\mathbf{R}), \qquad (21)$$

and $D_{kl}^{\sigma}(\mathbf{q})$ is the density matrix that now also has a $\mathbf{q}$-dependence:

$$D_{kl}^{\sigma}(\mathbf{q}) = \sum_{m} f_{m\sigma}(\mathbf{q}) c_{m\sigma}^{k}(\mathbf{q}) c_{m\sigma}^{l*}(\mathbf{q}), \qquad (22)$$

where $f_{m\sigma}(\mathbf{q})$ are the occupation numbers at $\mathbf{k}$-point $\mathbf{q}$, and $c_{m\sigma}^{k}(\mathbf{q})$ are the Kohn–Sham eigenvectors.

We have implemented two different methods to treat the singularity of the Coulomb matrix $V_{\mu\nu}(\mathbf{k})$ at $\mathbf{k}=\mathbf{0}$: (i) the singularity-correction method of Gygi and Baldereschi [55], extended to sparse $\mathbf{k}$-point samplings [56], and (ii) a cutoff for the Coulomb operator beyond a certain distance (the cut-Coulomb operator) [57]. The singularity-corrected Coulomb matrix $\tilde{V}_{\mu\nu}$ calculated with either of the methods can be used for a compact real-space representation of the Coulomb matrix, as described in the next section (see Eq. (31)). In practice, the Coulomb matrix in reciprocal space is calculated using Ewald summation in case of the unmodified Coulomb operator, or discrete Fourier transformation by explicit real-space summation in case of cut-Coulomb or screened Coulomb potentials.

Now let us examine what simplification the localized RI ("LVL") can bring. For a periodic system, the RI-LVL approximation implies that in real space, one has

$$\phi_i(\mathbf{r}-\mathbf{R})\phi_k(\mathbf{r}-\mathbf{R}')$$
$$= \sum_{\mu}\left(C_{i(\mathbf{R})k(\mathbf{R}')}^{\mu(\mathbf{R})} P_{\mu}(\mathbf{r}-\mathbf{R}) + C_{i(\mathbf{R})k(\mathbf{R}')}^{\mu(\mathbf{R}')} P_{\mu}(\mathbf{r}-\mathbf{R}')\right), \qquad (23)$$

where $\mathbf{R}$, $\mathbf{R}'$ are Bravais lattice vectors, and we use $i(\mathbf{R})$ to emphasize that the basis function i belongs to the unit cell $\mathbf{R}$. Apparently Eq. (23) reflects the basic idea of RI-LVL: only the auxiliary functions belonging to the unit cells from which the single-particle basis functions originate are used in the expansion. In reciprocal space, we now have

$$\varphi_{i\mathbf{k}}^{*}(\mathbf{r})\varphi_{k\mathbf{q}}(\mathbf{r}) = \sum_{\mathbf{R},\mathbf{R}'} e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}'} \phi_i(\mathbf{r}-\mathbf{R})\phi_k(\mathbf{r}-\mathbf{R}')$$
$$= \sum_{\mathbf{R},\mathbf{R}'} e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}'} \sum_{\mu}\left(C_{i(\mathbf{R})k(\mathbf{R}')}^{\mu(\mathbf{R})} P_{\mu}(\mathbf{r}-\mathbf{R}) + C_{i(\mathbf{R})k(\mathbf{R}')}^{\mu(\mathbf{R}')} P_{\mu}(\mathbf{r}-\mathbf{R}')\right)$$
$$= \sum_{\mathbf{R},\mathbf{R}'} e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}'} \sum_{\mu}\left(C_{i(\mathbf{R}_0)k(\mathbf{R}'-\mathbf{R})}^{\mu(\mathbf{R}_0)} P_{\mu}(\mathbf{r}-\mathbf{R}) + C_{i(\mathbf{R}-\mathbf{R}')k(\mathbf{R}_0)}^{\mu(\mathbf{R}_0)} P_{\mu}(\mathbf{r}-\mathbf{R}')\right)$$
$$= \sum_{\mu}\left(C_{ik}^{\mu}(-\mathbf{k},\mathbf{R}_0) + C_{ki}^{\mu}(\mathbf{q},\mathbf{R}_0)\right) P_{\mu}^{\mathbf{q}-\mathbf{k}}(\mathbf{r}), \qquad (24)$$

where $\mathbf{R}_0 = \mathbf{0}$ denotes the unit cell at the origin, and the translational invariance of the expansion coefficient, *i.e.*, $C_{i(\mathbf{R})k(\mathbf{R}')}^{\mu(\mathbf{R}'')} = C_{i(\mathbf{R}-\mathbf{R}'')k(\mathbf{R}'-\mathbf{R}'')}^{\mu(\mathbf{R}_0)}$ is used above. The expansion coefficients in Eq. (24) $C_{ik}^{\mu}(\mathbf{k},\mathbf{R}_0)$ are represented in Fourier space

$$C_{ik}^{\mu}(\mathbf{k},\mathbf{R}_0) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} C_{i(\mathbf{R})k(\mathbf{R}_0)}^{\mu(\mathbf{R}_0)}. \qquad (25)$$

Comparing Eqs. (20) and (24) we arrive at the following key expression:

$$C_{ik}^{\mu}(\mathbf{k},\mathbf{q}) = C_{ik}^{\mu}(-\mathbf{k},\mathbf{R}_0) + C_{ki}^{\mu}(\mathbf{q},\mathbf{R}_0). \qquad (26)$$

This means that the RI-LVL scheme decouples the $\mathbf{k}$ and $\mathbf{q}$ dependence of the RI expansion coefficients $C_{ik}^{\mu}(\mathbf{k},\mathbf{q})$ in reciprocal space. In practice only one independent Bloch vector (or equivalently one Bravais vector in real space) needs to be kept in storage of the three-orbital expansion coefficients. This greatly reduces the memory cost and computational effort in HF-type calculations for periodic systems.

5. Real-space formalism for the exchange operator

While algorithmically transparent and relatively easy to implement, the reciprocal-space approach has the disadvantage that the $C_{ik}^{\mu}(-\mathbf{k},\mathbf{R}_0)$ matrix is not sparse. It can be either stored or recalculated on the fly by Fourier transform of the real-space RI coefficients, and in each case there will be a time overhead. Moreover, the possibility of a distance-based screening, based on the locality of the basis functions, is lost.

As an alternative to the reciprocal-space formulation, one may also formulate the exchange operator according to Eq. (1) in real space. In this case, the number of arithmetic operations required to calculate the Coulomb integrals (see Eq. (3)) will be reduced due to the following two properties:

(1) the overlap between real-space basis functions (i,k) becomes negligible with distance, as does the overlap between (l,j).
(2) The density matrix D_{kl}^{σ} decays with increasing distance between the basis function pairs (i,k) and (l,j).

These properties result in a significant sparsity of the four-center Coulomb integrals in real space, and pave the way to an *a priori* screening of near-zero integrals.

It is this property that would be destroyed by traditional RI-V, as then each basis function product (i,k), even if limited in range on its own, would first become delocalized over the expanding auxiliary basis functions μ. By choosing our RI expansion to be localized at only the atoms at which i and k are centered, we retain the same localization as i and j, and therefore preserve the sparsity properties of Eq. (1).

The real-space version of the exchange operator is implemented as follows. The $\mathbf{k}$-dependent exchange operator in Eq. (1) can be expressed as [58]:

$$K_{ij}^{\sigma}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\mathbf{R}} X_{ij}^{\sigma}(\mathbf{R}), \qquad (27)$$

where

$$X_{ij}^{\sigma}(\mathbf{R}) = \sum_{kl}\sum_{\mathbf{R}'} D_{kl}^{\sigma}(\mathbf{R}') \times \sum_{\mathbf{R}''}\int d\mathbf{r}d\mathbf{r}' \frac{\phi_i(\mathbf{r})\phi_k(\mathbf{r}-\mathbf{R}'')\phi_j(\mathbf{r}'-\mathbf{R})\phi_l(\mathbf{r}'-\mathbf{R}'-\mathbf{R}'')}{|\mathbf{r}-\mathbf{r}'|} \quad (28)$$

and

$$D_{kl}^{\sigma}(\mathbf{R}') = \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{k}} D_{kl}^{\sigma}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{R}'}. \quad (29)$$

The translational symmetry is taken into account by considering basis functions ϕ_i in Eq. (28) only within the reference unit cell ($\mathbf{R} = \mathbf{0}$). The meaning of the different lattice vectors entering Eq. (28) is illustrated by Fig. 1. Using the resolution of identity, the real-space exchange matrix can be written as:

$$X_{ij}^{\sigma}(\mathbf{R}) = \sum_{k\mathbf{R}'}\sum_{l\mathbf{R}''}\sum_{\mu\mathbf{Q}'}\sum_{\nu\mathbf{Q}''} C_{ik(\mathbf{R}')}^{\mu(\mathbf{Q}')} V_{\mu\nu(\mathbf{R}+\mathbf{Q}''-\mathbf{Q}')} C_{jl(\mathbf{R}'')}^{\nu(\mathbf{Q}'')} \times D_{kl}^{\sigma}(\mathbf{R}+\mathbf{R}''-\mathbf{R}'). \quad (30)$$

where $\mathbf{Q}'$ and $\mathbf{Q}''$ are lattice vectors. For brevity, we omit index $\mathbf{0}$ for basis functions belonging to the reference unit cell. In case of our localized RI, the index μ of an auxiliary basis function in the above equation is restricted to either the same atom to which basis function ϕ_i belongs (and then $\mathbf{Q}' = \mathbf{0}$), or the atom of basis function ϕ_k (then $\mathbf{Q}' = \mathbf{R}'$). Similarly, $\mathbf{Q}''$ can only be either $\mathbf{0}$ or $\mathbf{R}''$, for a given $\mathbf{R}''$.

In the above formulation, lattice vectors $\mathbf{R}'$ and $\mathbf{R}''$ that can contribute to the sums are restricted by basis function overlaps, while the vector $\mathbf{R}$ is restricted only by the decay of the density matrix (and the very slow $1/|\mathbf{R}|$ decay of the Coulomb matrix). It has been shown previously [59] that the density matrix decays exponentially in insulators, but in metals the decay is only polynomial. Thus, a naive implementation would have to deal with a large number of lattice vectors for which the Coulomb and exchange matrices would have to be calculated and stored (or re-calculated). A more robust and consistent implementation can be obtained if one imposes Born–von Karman (BvK) periodic boundary conditions. For a given finite mesh of $\mathbf{k}$-points, the lattice-vector-dependent density matrix calculated using Eq. (29) will automatically satisfy BvK conditions. The BvK-periodic Coulomb matrix can be obtained in a similar way, *i.e.*, by the reverse Fourier transformation of the singularity-corrected reciprocal-space Coulomb matrix introduced in the previous section:

$$V_{\mu\nu(\mathbf{R})} = \frac{1}{N_{\mathbf{k}}}\sum_{\mathbf{k}} \tilde{V}_{\mu\nu}(\mathbf{k}) \exp(-i\mathbf{k}\cdot\mathbf{R}), \quad (31)$$

where lattice vectors $\mathbf{R}$ are restricted to the BvK cell. An approximate but more efficient method for calculating the BvK-periodic Coulomb matrix is to use the cut-Coulomb operator

$$V_{\mathrm{cut}}(\mathbf{r}) = \frac{1}{2|\mathbf{r}|}\mathrm{erfc}\left[\eta(|\mathbf{r}| - r_{\mathrm{cut}})\right], \quad (32)$$

where r_{cut} determines the radius beyond which the cut-Coulomb operator decays to zero, and η is the inverse decay width. The BvK-periodic Coulomb matrix is then constructed by first summing up the cut-Coulomb potentials in real space with appropriate phases calculated for the specified $\mathbf{k}$-point mesh:

$$\tilde{V}_{\mu\nu}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \int P_{\mu}(\mathbf{r}) V_{\mathrm{cut}}(|\mathbf{r}-\mathbf{r}'|) P_{\nu}(\mathbf{r}'-\mathbf{R}) d\mathbf{r}d\mathbf{r}',$$

and then Fourier-transformed back to real space according to Eq. (31). This method is a natural choice in case of range-separated hybrid functionals.

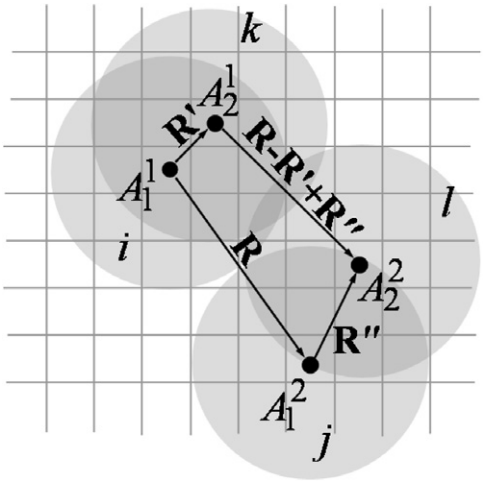

Fig. 1. Four atom centers and the lattice vectors used for grouping of four-center integrals in Eq. (30). The grid lines represent borders of the unit cells, atom centers are shown as black dots, and lattice vectors are shown by arrows. The superscript in A_i^j denotes the number of pair, while the subscript denotes atoms within the pairs.

The real-space Coulomb matrix obtained in either of these ways has the advantage that the Coulomb interaction with auxiliary basis functions beyond the BvK cell is folded in, since no explicit lattice-vector-dependent cut-offs are applied when calculating $\tilde{V}_{\mu\nu}(\mathbf{k})$.

Note that the vectors $\mathbf{R}'$ and $\mathbf{R}''$ in Eq. (30) are restricted solely by the overlap of basis functions, and, therefore, are allowed to go beyond the BvK cell, which can happen in case of sparse $\mathbf{k}$-point samplings. In this case, we simply take into account BvK periodicity of the Coulomb and density matrices, *e.g.*, $V_{\mu\nu(\mathbf{R})} = V_{\mu\nu(\mathbf{R}+\mathbf{R}^{\mathrm{BvK}})}$, where $\mathbf{R}^{\mathrm{BvK}}$ is any lattice vector of the BvK superlattice.

Due to the complex algebraic structure of Eq. (30) (it is a product of four multi-index matrices), its actual implementation involves multiple choices in the order, grouping, and distribution of operations, with the effect of each choice on the overall performance being hard to predict. Below we outline our algorithmic solution, which showed the best performance among several implementations that we have tried.

It is convenient to group the RI coefficients C in Eq. (30) by atoms to which the auxiliary basis functions belong. For clarity, the four atoms and the corresponding lattice vectors involved in each single term in Eq. (30), are depicted in Fig. 1, along with their notations. Thus, the exchange matrix for each pair of atoms $A_1^1A_1^2$ can be written as:

$$X_{i\in A_1^1 j\in A_1^2}^{\sigma}(\mathbf{R}) = \sum_{A_2^1(\mathbf{R}')}\sum_{k\in A_2^1(\mathbf{R}')}\sum_{\nu\in A_1^2} F_{ik}^{\nu(\mathbf{R})} E_{jk(\mathbf{R}-\mathbf{R}')}^{\nu\sigma} + \sum_{A_2^2(\mathbf{R}'')}\sum_{l,\nu\in A_2^2(\mathbf{R}'')} \left(2G_{il}^{\nu\sigma}(\mathbf{R}+\mathbf{R}'') + H_{il}^{\nu\sigma}(\mathbf{R}+\mathbf{R}'')\right) C_{lj(-\mathbf{R}'')}^{\nu}, \quad (33)$$

where

$$E_{jk(\mathbf{R})}^{\nu\sigma} = \sum_{l\mathbf{R}''} C_{jl(\mathbf{R}'')}^{\nu} D_{kl}^{\sigma}(\mathbf{R}+\mathbf{R}''), \quad (34)$$

$$F_{ik(\mathbf{R}')}^{\nu(\mathbf{R})} = \sum_{\mu\in A(i)} C_{ik(\mathbf{R}')}^{\mu} V_{\mu\nu(\mathbf{R})}, \quad (35)$$

$$G_{il}^{\nu\sigma}(\mathbf{R}) = \sum_{\mu\in A(i)} \left(E_{il(-\mathbf{R})}^{\nu\sigma}\right)^{*} V_{\mu\nu(\mathbf{R})}, \quad (36)$$

$$H_{il}^{\nu\sigma}(\mathbf{R}) = \sum_{k\mathbf{R}'} F_{ki(-\mathbf{R}')}^{\nu(\mathbf{R}-\mathbf{R}')} D_{kl}^{\sigma}(\mathbf{R}-\mathbf{R}'). \quad (37)$$

$A(i)$ denotes the atom on which basis function i resides, "$*$" denotes complex conjugation, $\mu \in A(i)$ signifies that the auxiliary function $P_\mu(\mathbf{r})$ is centered on atom $A(i)$, and $k \in A(i)$ signifies that the atomic basis function $\phi_k(\mathbf{r})$ is centered on atom $A(i)$. The choice of these intermediate matrices is dictated by the effort to minimize the number of matrix multiplications. In the above equation we make use of the translational symmetry of the RI coefficients, namely $C^{\mu(\mathbf{R}')}_{i(\mathbf{0})k(\mathbf{R}')} = C^{\mu(\mathbf{0})}_{k(\mathbf{0})i(-\mathbf{R}')}$.

For brevity, we introduce the following short-hand notations that are used throughout the text below. The auxiliary basis functions $P_\mu(\mathbf{r}+\mathbf{R})$ are labeled by their greek-letter index and the lattice vector: $\mu(\mathbf{R})$ or simply μ for $\mathbf{R} = \mathbf{0}$. Similarly, the atomic basis functions $\phi_i(\mathbf{r}+\mathbf{R})$ are labeled as $i(\mathbf{R})$ or simply i for $\mathbf{R} = \mathbf{0}$.

The matrices E and F are sparse. The indices i and $k(\mathbf{R}')$ of matrix $F^{\nu(\mathbf{R})}_{ik(\mathbf{R}')}$ are bounded by the overlap of corresponding basis functions. In addition, in cases when auxiliary basis functions $\nu(\mathbf{R})$ do not overlap with i and $k(\mathbf{R}')$ atomic orbitals, the corresponding matrix elements are determined by the multipole moments of $\nu(\mathbf{R})$. Thus, we can make an efficient use of the idea by Betzinger et al. [18] to perform a simple unitary transform of the auxiliary basis functions on each atom and in each lm channel, such that the multipole moments vanish for all but one function. We also extended this method to the functionals with screened Coulomb interaction (in particular, the HSE functional), in which case the multipole moment of all but one function can be made small, although not exactly zero. Indices j and $k(\mathbf{R})$ in $E^{\nu\sigma}_{jk(\mathbf{R})}$ are bounded by the decay of the density matrix with increasing distance between the corresponding functions.

A key ingredient of an efficient implementation of non-local exchange is screening out small elements of the exchange matrix without actual calculation. The above representation and the sparsity patterns allow for such an efficient *a priori* screening. We introduce a single parameter ε, and discard all exchange matrix elements below ε. *A priori* screening of the exchange matrix is based on estimates of the matrix elements for each of the three terms in Eq. (33). Since the implementation of these estimates is closely tied to the algorithm of the exchange matrix calculation, we combine its description with a rough sketch of the algorithm. The same screening procedure can be applied also to non-periodic systems (clusters), simply by setting all lattice vectors in the above expressions to $\mathbf{0}$.

The outer-most loop is over basis functions on atom A^1_2 (see Fig. 1 for atom notations). In practice, the loop is performed over all basis functions in the central unit cell ($\mathbf{R}' = \mathbf{0}$), and for each basis function the atom to which it belongs is treated as atom A^1_2, with additional loops over the lattice vector $\mathbf{R}'$ when needed. First, we define the following quantities:

$$\tilde{C}^{A(i)}_{ij(\mathbf{R}')} = \sqrt{\sum_{\mu\in A(i)} \left(C^{\mu}_{ij(\mathbf{R}')}\right)^2},$$

$$\tilde{C}_{A_1A_2(\mathbf{R}')} = \max_{i\in A_1, j\in A_2} \tilde{C}^{A_1}_{ij(\mathbf{R}')},$$

$$V_{A_1A_2(\mathbf{R})} = \sqrt{\sum_{\mu\in A_1,\nu\in A_2} (V_{\mu\nu(\mathbf{R})})^2},$$

$$D_{kA}(\mathbf{R}) = \max_{\sigma, l\in A} |D^{\sigma}_{kl}(\mathbf{R})|,$$

$$\tilde{D}_{A_1A_2}(\mathbf{R}) = \max_{k\in A_1} D_{kA_2}(\mathbf{R}),$$

$$s^{A_3}_{A_1A_2}(\mathbf{R}') = \max_{\mathbf{R}} \left(V_{A_2A_3(\mathbf{R})} \tilde{D}_{A_1A_3}(\mathbf{R}+\mathbf{R}')\right),$$

$$O_{A_1A_2}(\mathbf{R}') = \max_{A_3} \left(s^{A_3}_{A_1A_2}(\mathbf{R}') \max_{A_4,\mathbf{R}''} \tilde{C}_{A_3A_4(\mathbf{R}'')}\right).$$

For each atom A^1_1 and lattice vector $\mathbf{R}'$, we select basis function pairs $ik(\mathbf{R}')$, $i \in A^1_1$ (k is the index of the basis function in the outermost loop), that satisfy the following criterion:

$$\tilde{C}^{A^1_1}_{ik(\mathbf{R}')} \max\left[\max_{A^2_1,\mathbf{R}}\left[V_{A^1_1A^2_1(\mathbf{R}+\mathbf{R}')} \max_{A^2_2,\mathbf{R}''} \left(D_{kA^2_2}(\mathbf{R}+\mathbf{R}'')\tilde{C}_{A^2_1A^2_2(\mathbf{R}'')}\right)\right], O_{A^1_1A^1_1}(\mathbf{0})\right] > \varepsilon.$$

Analogously, basis pairs $ki(-\mathbf{R}')$ are considered only if

$$\tilde{C}^{A^1_2}_{ki(-\mathbf{R}')} O_{A^1_1A^1_2}(\mathbf{R}') > \varepsilon.$$

The coefficients $C^{\mu}_{ik(\mathbf{R}')}$ for the selected subset of the basis function pairs are then distributed among processors and used for the evaluation of the intermediate matrix products.

As a next step, we loop over atoms A^2_1. A subset (within the BvK supercell) of lattice vectors $\mathbf{R}$ for which the density matrix $D^{\sigma}_{kl}(\mathbf{R}+\mathbf{R}'')$ needs to be calculated to obtain the matrix $E^{\nu\sigma}_{jk(\mathbf{R})}$ is determined by the following condition

$$D_{kA^2_2}(\mathbf{R}+\mathbf{R}'')\tilde{C}_{A^2_1A^2_2(\mathbf{R}'')} \max_{A^1_1,\mathbf{R}'} \left(V_{A^1_1A^2_1(\mathbf{R}+\mathbf{R}')}\tilde{C}_{A^1_1A^1_2(\mathbf{R}')}\right) > \varepsilon,$$

tested for every atom pair $A^2_1A^2_2(\mathbf{R}'')$.

Withing the same loop over atom A^2_1, we identify, for each lattice vector $\mathbf{R}$ within the subset determined above, atom A^1_1, and lattice vector $\mathbf{R}'$, basis function pairs $ik(\mathbf{R}')$ ($i \in A^1_1$) which need to be considered in calculation of the first and third terms ($E\cdot F$ and $H\cdot C$) in Eq. (33). For the $E\cdot F$ term, the basis function pairs should satisfy the condition

$$\tilde{C}^{A^1_1}_{ik(\mathbf{R}')} V_{A^1_1A^2_1(\mathbf{R})} E_{A^2_1k(\mathbf{R}-\mathbf{R}')} > \varepsilon,$$

where

$$E_{A^2_1k(\mathbf{R}-\mathbf{R}')} = \max_{\sigma} \sqrt{\max_{j\in A^2_1} \sum_{\nu\in A^2_1} \left(E^{\nu\sigma}_{jk(\mathbf{R}-\mathbf{R}')}\right)^2}.$$

For the $H\cdot C$ term, the basis function pairs are selected according to

$$\tilde{C}^{A^1_1}_{ik(\mathbf{R}')} V_{A^1_1A^2_1(\mathbf{R})} \tilde{D}_{A^1_1A^2_1}(\mathbf{R}) \max_{A^2_2,\mathbf{R}''} \tilde{C}_{A^2_1A^2_2(\mathbf{R}'')} > \varepsilon.$$

Finally, the basis function pairs $ki(-\mathbf{R}')$ ($i \in A^1_1$) needed for the second term ($G\cdot C$) in Eq. (33) are selected according to

$$\tilde{C}^{A^1_2}_{ki(-\mathbf{R}')} V_{A^1_2A^2_1(\mathbf{R})} \tilde{D}_{A^1_1A^2_1}(\mathbf{R}+\mathbf{R}') \max_{A^2_2,\mathbf{R}''} \tilde{C}_{A^2_1A^2_2(\mathbf{R}'')} > \varepsilon.$$

Note that, while the $F\cdot E$ term (the first term in Eq. (33)) contributes to the $X^{\sigma}_{i\in A^1_1 j\in A^2_1}$ elements of the exchange matrix, the matrices H and G are indexed in such a way that the $H\cdot C$ and $G\cdot C$ terms contribute to the $X^{\sigma}_{k\in A^1_2 l\in A^2_2}$ elements. This allows to reduce the communication between CPUs and to avoid redundant operations.

For each atom A^2_2 and lattice vector $\mathbf{R}''$, we also determine for which lattice vectors $\mathbf{R}$ the last two terms in Eq. (33) have to be calculated from matrices G and H, based on the following criterion

$$\tilde{C}_{A^2_1A^2_2(\mathbf{R}'')} \max_{j\in A^2_1,\sigma} \sqrt{\sum_{\nu\in A^2_1} \left(H^{\nu\sigma}_{kj}(\mathbf{R}) + 2G^{\nu\sigma}_{kj}(\mathbf{R})\right)^2} > \varepsilon.$$

Although the sparsity patterns described above allow for more refined screening, we find the above screening scheme conveniently compatible with such grouping of matrix multiplications that the highly optimized dense matrix multiplications can be used without the need for an overwhelming number of indexing arrays.

We aim at an implementation that does not require any storage of large matrices to disk. Such a storage would make calculations unacceptably slow for any periodic system of practical interest. A

disk-free implementation, on the other hand, must deal with the necessity to repeat some operations. In particular, we need to recalculate the product of the Coulomb matrix and the RI coefficients (the F matrix above, Eq. (35)). This product does not change during the SCF cycle, and could be calculated only once. It is also a three-index matrix (if one associates the lattice vector indices with the basis function index in a supercell). However, due to the non-local nature of the Coulomb interaction, we cannot efficiently benefit from the sparsity of the matrices, which will most likely make the storage of the complete product in memory unfeasible even if it is distributed.

Ideally, we want to distribute the storage and operations on multiple CPUs so that: (i) memory load is distributed evenly between different CPUs, (ii) the communication between CPUs is minimized, and (iii) computational load (number of operations) is distributed evenly between CPUs. We find that it is very non-trivial to fulfill all three conditions simultaneously for an arbitrary system. Nevertheless, computational effort for a wide range of sizes (from few- to few-hundred atoms per unit cell) with the current implementation on average scales linearly with the system size (see the example in the next section).

In cases of a large number of **k**-points and intermediate unit cell sizes (bulk metallic systems), the Coulomb matrix would dominate the memory consumption for storage, followed by the density matrix, the exchange matrix, and the RI coefficients C. We find that the on-the-fly calculation of the Coulomb matrix using a multipole expansion is not fast enough to compete with fetching the matrix from memory, which has to be done multiple times to avoid additional storage, computations, and communication. However, we do benefit from the multipole expansion during the initial calculation of the matrix. We also make full use of the sparsity of the Coulomb matrix mentioned above. Namely, we store only elements with absolute values above a small number η. In fact, we find that the results are rather insensitive to the value of η in the range 10^{-12}–10^{-4}, due to the on-site transformation that removes the intermediate values of the Coulomb matrix [18] (see text above). The sparse matrix $V_{\mu\nu(\mathbf{R})}$ is computed and stored for each atom pair A_1A_2 ($\mu \in A_1$, $\nu \in A_2$) within one unit cell, and each lattice vector **R** within the BvK supercell. It is distributed among CPUs first over atoms A_2, but if the number of CPUs is larger than the number of atoms, it is also distributed over lattice vectors **R**. The distribution by atoms A_2 is used for the density matrix $D^{\sigma}_{kl}(\mathbf{R})$ ($k \in A_2$) and for the real-space exchange matrix $X^{\sigma}_{ij}(\mathbf{R})$ ($j \in A_2$).

For systems with large numbers of atoms and basis functions the RI coefficients $C^{\mu(\mathbf{Q}')}_{ik(\mathbf{R}')}$ can be expected to dominate the storage. However, due to the use of localized RI (LVL), the size of the C matrix is significantly reduced. The index i is distributed according to the atom distribution used for the Coulomb matrix storage.

The matrix operations are distributed such that communication between CPUs is minimized. For each basis function k, RI coefficients $C^{\mu(\mathbf{Q}')}_{ik(\mathbf{R}')}$ that passed the screening criteria described above, are first communicated to all CPUs. Then we loop over local atoms and local unit cells within BvK supercell on each CPU, and calculate all local contributions to the exchange matrix. The only additional global communication takes place at the end of all operations for given k, namely, to synchronize the exchange matrix.

6. Performance

6.1. Benchmark systems

We demonstrate the accuracy of our approach by comparing to benchmark periodic HSE06 results in the literature. Specifically, we focus on the well studied semiconductors Si, GaAs, and GaSb.

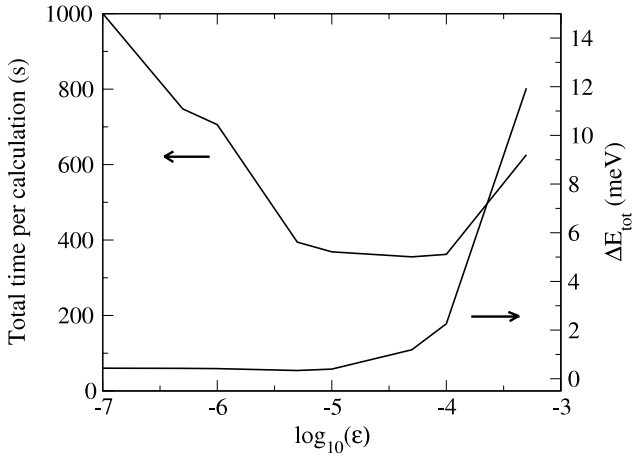

Fig. 2. Dependence of the total energy (right axis) and computational time (left axis) on the screening parameter ε. The energy zero corresponds to the total energy obtained with the **k**-space implementation. The system is zinc-blende GaAs (2 atoms in the unit cell), the method is HSE06 ($\omega = 0.11$ bohr^{-1}). The total computational time increases for larger ε because the number of scf cycles needed for convergence also increases. Each calculation was distributed over 96 CPUs.

Table 1 contains the calculated and experimental lattice parameters, bulk moduli, cohesive energies, and selected electronic properties for each material. The settings used in FHI-aims are "tight" and an $8 \times 8 \times 8$ **k**-space grid. All parameters are calculated for the diamond/zinc-blende crystal structure.

6.2. Real-space vs. reciprocal space implementation

In this subsection, we compare the relative accuracy and efficiency of our own implementation. We focus on total energy deviations between the **k**-space implementation and the real-space implementation. The **k**-space implementation has no adjustable parameters, and can thus serve as a reference. In the real-space implementation, the key parameter to be examined is ε, whose use is described in detail in Section 5.

The dependence of the accuracy and computational time on the screening parameter ε for the primitive unit cell of GaAs is shown in Fig. 2.

The results clearly demonstrate that the *a priori* four-center integral screening can significantly reduce the computational time without loss of accuracy. However, this behavior may be system-dependent, and we recommend to test it for every system or class of similar systems.

6.3. Computational scaling

Scaling of the computational time with the number of CPUs for the real-space implementation is shown in Fig. 3 for different unit cell sizes. For the unit cell of the largest size (1024 atoms), only $\mathbf{k} = \mathbf{0}$ is used. Other sizes are obtained by the successive division of the unit cell by a factor of two along one, two, and then three directions, with the number of **k**-points being doubled along corresponding directions in the reciprocal space. Based on the results shown in Fig. 2, the value of the screening parameter ε is set to 10^{-5}. As can be seen from the figure, the computational time consistently improves with increasing the number of CPUs, up to the maximum tested number (8192 cores).

Finally, we demonstrate the linear scaling of the total computational time with the system size for the real-space implementation, using GaAs as an example. Fig. 4 shows total time per iteration of the self-consistency cycle as a function of the number of atoms in the unit cell.

Table 1
Calculated and experimental lattice constants (a), bulk moduli (B_0), cohesive (atomization) energies (E_{coh}), band gap at the Γ point ($E^{\Gamma}_{\mathrm{gap}}$), and the gap between lowest conduction band at X point and the valence band at the Γ point ($E_{\mathrm{X_{1c}}}$). The first row for each material shows results obtained in this work. The method is HSE06 ($\omega = 0.11$ bohr^{-1}), the crystal structures are diamond for Si, and zinc-blende for GaAs and GaSb. The relativity is taken into account via scaled zeroth-order regular approximation [60]. The spin–orbit coupling is not included in our calculations. The experimental numbers for Si and GaAs are the same as cited by Paier et al. [61]. Values in parentheses are calculated at the experimental lattice constant.

Material	a (Å)	B_0 (GPa)	E_{coh} (eV)	$E^{\Gamma}_{\mathrm{gap}}$ (eV)	$E_{\mathrm{X_{1c}}}$ (eV)
Si	5.446	97.5	4.540	3.33 (3.33)	1.34 (1.33)
Ref. [61]	5.435	97.7	4.582	(3.32)	(1.29)
exp.	5.430[a]	99.2[b]	4.63[c]	3.34-3.36[d]	1.13[d,e]
				3.05[f]	1.25[f]
GaAs	5.695	71.1	3.181	1.12 (1.34)	2.09 (2.06)
Ref. [61]	5.687	70.9	3.149	(1.45)	(2.02)
exp.	5.648	78.9[g]	3.26[c]	1.52[g]	1.98[g]
GaSb	6.157	53.1	2.859	0.58 (0.87)	1.43 (1.40)
Ref. [62]	6.16			0.65	
exp.	6.086[h]	55[i]	3.03	0.81[j,k]	
		55.7[l]			

[a] Ref. [63].
[b] Ref. [64].
[c] Ref. [65].
[d] Ref. [66].
[e] Ref. [67].
[f] Ref. [68].
[g] Ref. [69].
[h] Ref. [70].
[i] Ref. [71].
[j] Ref. [72].
[k] Ref. [73].
[l] Ref. [74].

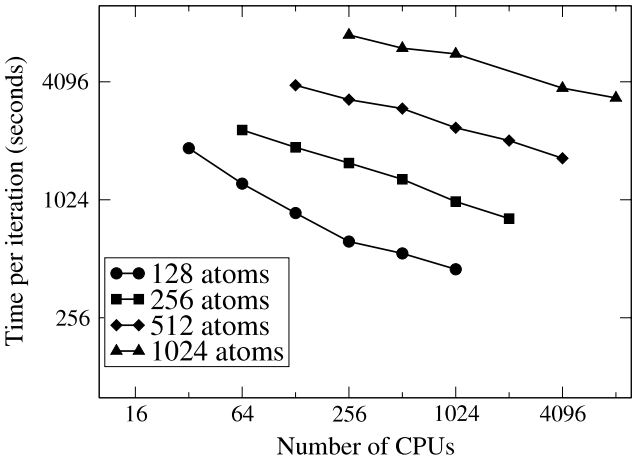

Fig. 3. Scaling of computational time (total time per calculation divided by the number of iterations) with the number of CPUs for different system sizes. The system is zinc-blende GaAs, the method is HSE06 ($\omega = 0.11$ bohr^{-1}).

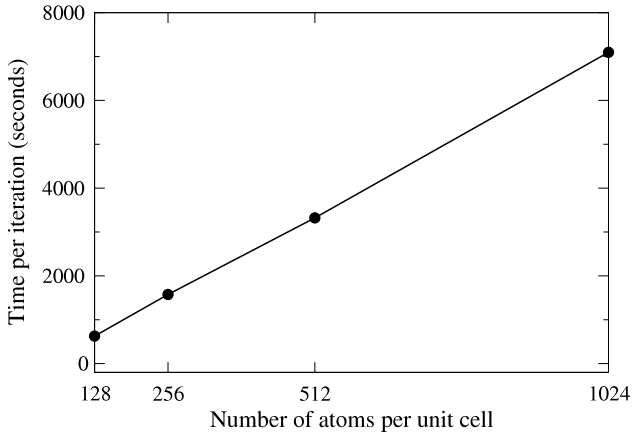

Fig. 4. Scaling of computational time (total time per calculation divided by the number of iterations) with the system size. The system is zinc-blende GaAs, the method is HSE06 ($\omega = 0.11$ bohr^{-1}). Each calculation was distributed over 256 CPUs.

7. Conclusions

We have presented an all-electron implementation of the Hartree–Fock ("exact") exchange for periodic systems using numeric atomic orbitals as basis functions. The resolution of the identity (variational density fitting) is used to avoid storage of the four-center electron repulsion integrals. A localized version of the RI approach, RI-LVL, allows to further reduce memory consumption and computational time. In conjunction with an *a priori* screening of the four-center electron repulsion integrals, based on the locality of the basis functions, the use of RI-LVL leads to a linear scaling of the computational time with system size (number of atoms per unit cell). The numerical accuracy of the implementation has been tested by comparison with the theoretical and experimental results for Si, GaAs, and GaSb reported in literature.

Acknowledgments

SVL is grateful for the financial support by the Cluster of Excellence "Unifying Concepts in Catalysis" (EXC 314; www.unicat.tu-berlin.de; funded by the Deutsche Forschungsgemeinschaft) and the Alexander von Humboldt Foundation.

References

[1] K. Kim, K. Jordan, Comparison of density functional and MP2 calculations on the water monomer and dimer, J. Phys. Chem. 98 (1994) 10089–10094.
[2] P. Stephens, F. Devlin, C. Chabalowski, M. Frisch, *Ab initio* calculation of vibrational absorption and circular dichroism spectra using density functional force fields, J. Phys. Chem. 98 (1994) 11623–11627.
[3] M. Ernzerhof, G. Scuseria, Assessment of the Perdew–Burke–Ernzerhof exchange–correlation functional, J. Chem. Phys. 110 (1999) 5029–5036.
[4] C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: the PBE0 model, J. Chem. Phys. 110 (1999) 6158–6170.

[5] J. Heyd, G. Scuseria, M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, J. Chem. Phys. 118 (2003) 8207–8215.
[6] J. Heyd, G. Scuseria, M. Ernzerhof, Erratum: "Hybrid functionals based on a screened coulomb potential" [J. Chem. Pys. 118, 8207 (2003)], J. Chem. Phys. 124 (2006) 219906.
[7] A.V. Krukau, O.A. Vydrov, A.F. Izmaylov, G.E. Scuseria, Influence of the exchange screening parameter on the performance of screened hybrid functionals, J. Chem. Phys. 125 (2006) 224106.
[8] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. Sounders, C. Zicovich-Wilson, CRYSTAL: a computational tool for the *ab initio* study of the electronic properties of crystals, Z. Kristalogr. 220 (2005) 571–573.
[9] R. Dovesi, V. Sounders, C. Roetti, R. Orlando, C. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. Harrison, I. Bush, P. D'Arco, M. Llunell, CRYSTAL09 user's manual, University of Torino, Torino, 2009.
[10] E. Bylaska, K. Tsemekhman, S.B. Baden, J. Weare, H. Jonsson, Parallel implementation of γ-point pseudopotential plane-wave DFT with exact exchange, J. Comput. Chem. 32 (2011) 54–69.
[11] H. Shang, Z. Li, J. Yang, Implementation of screened hybrid density functional for periodic systems with numerical atomic orbitals: basis function fitting and integral screening, J. Chem. Phys. 135 (2011) 034110.
[12] M. Guidon, J. Hutter, J. VandeVondele, Robust periodic Hartree–Fock exchange for large-scale simulations using Gaussian basis sets, J. Chem. Theory Comput. 5 (2009) 3010–3021.
[13] C. Tymczak, V. Weber, E. Schwegler, M. Challacombe, Linear scaling computation of the Fock matrix. VIII. Periodic boundaries for exact exchange at the γ point, J. Chem. Phys. 122 (2005) 124105.
[14] J. Fernández, C. Tablero, P. Wahnón, Development and implementation of the exact exchange method for semiconductors using a localized basis set, Comput. Mater. Sci. 28 (2003) 274–286.
[15] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ã. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09 Revision A.1, Gaussian Inc., Wallingford, CT, 2009.
[16] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. Gerber, J. Ángyán, Screened hybrid density functionals applied to solids, J. Chem. Phys. 124 (2006) 154709.
[17] X. Wu, A. Selloni, R. Car, Order-n implementation of exact exchange in extended insulating systems, Phys. Rev. B 79 (2009) 085102.
[18] M. Betzinger, C. Friedrich, S. Blügel, Hybrid functionals within the all-electron FLAPW method: implementation and applications of PBE0, Phys. Rev. B 81 (2010) 195117.
[19] M. Schlipf, M. Betzinger, C. Friedrich, M. Ležaić, S. Blügel, HSE hybrid functional within the FLAPW method and its application to GdN, Phys. Rev. B 84 (2011) 125142.
[20] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A.D. Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, J. Phys.: Condens. Matter 21 (2009) 395502.
[21] P. Havu, V. Blum, V. Havu, P. Rinke, M. Scheffler, Large-scale surface reconstruction energetics of Pt(100) and Au(100) by all-electron density functional theory, Phys. Rev. B 82 (2010) 161418(R).
[22] D. Bowler, T. Miyazaki, Calculations for millions of atoms with density functional theory: linear scaling shows its potential, J. Phys.: Condens. Matter 22 (2010) 074207.
[23] V. Blum, F. Hanke, R. Gehrke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, Ab-initio molecular simulations with numeric atom-centered orbitals, Comput. Phys. Comm. 180 (2009) 2175.
[24] B. Delley, An all-electron numerical method for solving the local density functional for polyatomic molecules, J. Chem. Phys. 92 (1990) 508–517.
[25] M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, M. Scheffler, Secondary structure of Ac-Ala$_n$-LysH$^+$ polyalanine peptides (n = 5, 10, 15) in vacuo: Helical or not? J. Phys. Chem. Lett. 1 (2010) 3465–3470.
[26] N. Artrith, T. Morawietz, J. Behler, High-dimensional neural-network potentials for multicomponent systems: applications to zinc oxide, Phys. Rev. B 83 (2011) 153101.
[27] C. Van Alsenoy, *Ab initio* calculations on large molecules: the multiplicative integral approximation, J. Comput. Chem. 9 (1988) 620–626.
[28] Vahtras, J. Almlöf, M. Feyereisen, Integral approximations for LCAO-SCF calculations, Chem. Phys. Lett. 213 (1993) 514.
[29] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary basis sets to approximate Coulomb potentials, Chem. Phys. Lett. 240 (1995) 283–290.
[30] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, M. Scheffler, New J. Phys. 14 (2012) 053020.
[31] I.Y. Zhang, X. Ren, P. Rinke, V. Blum, M. Scheffler, Numeric atom-centered-orbital basis sets with valence-correlation consistency from H to Ar, New J. Phys. 15 (2013) 123033.
[32] L. Maschio, D. Usvyat, F. Manby, S. Casassa, C. Pisani, M. Schütz, Fast local-MP2 method with density-fitting for crystals. I. Theory and algorithms, Phys. Rev. B 76 (2007) 075101.
[33] D. Usvyat, L. Maschio, F. Manby, S. Casassa, M. Schütz, C. Pisani, Fast local-MP2 method with density-fitting for crystals. II. Test calculations and application to the carbon dioxide crystal, Phys. Rev. B 76 (2007) 075102.
[34] M. Lorenz, D. Usvyat, M. Schütz, Local *ab initio* methods for calculating optical band gaps in periodic systems. I. Periodic density fitted local configuration interaction singles method for polymers, J. Chem. Phys. 134 (2011) 094101.
[35] M. Lorenz, L. Maschio, M. Schütz, D. Usvyat, Local *ab initio* methods for calculating optical bandgaps in periodic systems. II. Periodic density fitted local configuration interaction singles method for solids, J. Chem. Phys. 137 (2012) 204119.
[36] V. Atalla, M. Yoon, F. Caruso, P. Rinke, M. Scheffler, Hybrid density functional theory meets quasiparticle calculations: a consistent electronic structure approach, Phys. Rev. B 88 (2013) 165122.
[37] T. Körzdörfer, R. Parrish, N. Marom, J. Sears, C. Sherrill, J.-L. Brédas, Assessment of the performance of tuned range-separated hybrid density functionals in predicting accurate quasiparticle spectra, Phys. Rev. B 86 (2012) 205110.
[38] T. Körzdörfer, N. Marom, Strategy for finding a reliable starting point for g_0w_0 demonstrated for molecules, Phys. Rev. B 86 (2012) 041110(R).
[39] S.F. Boys, I. Shavitt, Univeristy of Wisconsin Rept. WIS-AF-13.
[40] J.L. Whitten, Coulomb potential energy integrals and approximations, J. Chem. Phys. 58 (1973) 4496.
[41] B.I. Dunlap, J.W.D. Connolly, J.R. Sabin, On some approximations of xα method, J. Chem. Phys. 71 (1979) 3396.
[42] B. Dunlap, Robust variational fitting: Gáspár's variational exchange can accurately be treated analytically, THEOCHEM 502 (2000) 221.
[43] B. Dunlap, N. Rösch, S. Trickey, Variational fitting methods for electronic structure calculations, Mol. Phys. 108 (2010) 3176.
[44] O. Vahtras, J. Almlöf, M.W. Feyereisen, Integral approximations for LCAO-SCF calculations, Chem. Phys. Lett. 213 (1993) 514.
[45] M. Feyereisen, G. Fitzgerald, A. Komornicki, Use of approximate integrals in ab initio theory, an application in MP2 energy calculations, Chem. Phys. Lett. 208 (1993) 359.
[46] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, RI-MP2: optimized auxiliary basis sets and demonstration of efficiency, Chem. Phys. Lett. 294 (1998) 143.
[47] F. Caruso, P. Rinke, X. Ren, M. Scheffler, A. Rubio, Unified description of ground and excited states of finite systems: the self-consistent GW approach, Phys. Rev. B 86 (2012) 081102(R).
[48] F. Caruso, P. Rinke, X. Ren, A. Rubio, M. Scheffler, Self-consistent GW: all-electron implementation with localized basis functions, Phys. Rev. B 88 (2013) 075105.
[49] A. Ihrig, J. Wieferink, I. Zhang, V. Blum, X. Ren, P. Rinke, M. Scheffler, to be published.
[50] F. Billingsley, J. Bloor, Limited expansion of diatomic overlap (LEDO): a near-accurate approximate *ab initio* LCAO MO method. I. Theory and preliminary investigations, J. Chem. Phys. 55 (1971) 5178–5190.
[51] J. Whitten, Coulombic potential energy integrals and approximations, J. Chem. Phys. 58 (1973) 4496–4501.
[52] A. Sodt, J.E. Subotnik, M. Head-Gordon, Linear scaling density fitting, J. Chem. Phys. 125 (2006) 194109.
[53] A. Sodt, M. Head-Gordon, Hartree–Fock exchange computed using the atomic resolution of the identity approximation, J. Chem. Phys. 128 (2008) 104106.
[54] P. Merlot, T. Kjærgaard, T. Helgaker, R. Lindh, F. Aquilante, S. Reine, T.B. Pedersen, Attractive electron–electron interactions within robust local fitting approximations, J. Comput. Chem. 34 (2013) 1486.
[55] F. Gygi, A. Baldereschi, Self-consistent Hartree–Fock and screened-exchange calculations in solids: application to silicon, Phys. Rev. B 34 (1986) 4405–4408.
[56] P. Broqvist, A. Alkauskas, A. Pasquarello, Hybrid-functional calculations with plane-wave basis sets: effect of singularity correction on total energies, energy eigenvalues, and defect energy levels, Phys. Rev. B 80 (2009) 085114.
[57] J. Spencer, A. Alavi, Efficient calculation of the exact exchange energy in periodic systems using a truncated Coulomb potential, Phys. Rev. B 77 (2008) 193110.
[58] R. Dovesi, R. Orlando, C. Roetti, C. Pisani, V. Saunders, The periodic Hartree–Fock method and its implementation in the Crystal code, Phys. Status Solidi B 217 (2000) 63–88.
[59] S. Goedecker, Linear scaling electronic structure methods, Rev. Modern Phys. 71 (1999) 1085–1123.
[60] E. van Lenthe, E. Baerends, J. Snijders, Relativistic total energy using regular approximations, J. Chem. Phys. 101 (1994) 9783.
[61] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. Gerber, J. Ángyán, Erratum: screened hybrid density functionals applied to solids [j. chem. phys. 124, 154709 (2006)], J. Chem. Phys. 125 (2006) 249901.
[62] V. Virkkala, V. Havu, F. Tuomisto, M.J. Puska, Hybrid functional study of band structures of GaAs$_{1-x}$N$_x$ and GaSb$_{1-x}$N$_x$ alloys, Phys. Rev. B 85 (2012) 085132.
[63] C. Kittel, Introduction to Solid State Physics, Wiley, New York, 1996.
[64] K.-H. Hellwege (Ed.), Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology, in: New Series, Group III, vol. 17, Springer, Berlin, 1966, Pt. a.
[65] M.W. Chase Jr. (Ed.), NIST-JANAF Thermochemical Tables, AIP, New York, 1998.
[66] M. Walkowsky, R. Braunstein, Interband transitions and exciton effects in semiconductors, Phys. Rev. B 5 (1972) 497–509.
[67] R. Zucca, J. Walter, Y. Shen, M. Cohen, Wavelength modulation spectra of GaAs and silicon, Solid State Commun. 8 (1970) 627–632.
[68] J. Ortega, F. Himpsel, Inverse-photoemission study of Ge(100), Si(100), and GaAs(100): bulk bands and surface states, Phys. Rev. B 47 (1993) 2130–2137.

[69] J.S. Blakemore, Semiconducting and other major properties of gallium arsenide, J. Appl. Phys. 53 (1982) R123.
[70] V.T. Bublik, J. Wilke, A.T. Pereversev, Die Natur der festen lösungen der überschüssigen Komponenten in unlegiertem galliumantimonid, Phys. Status Solidi (a) 73 (1982) K271.
[71] V.A. Sidorov, V.V. Brazhkin, L.G. Khvostantsev, A.G. Lyapin, A.V. Sapelkin, O.B. Tsiok, Nature of semiconductor-to-metal transition and volume properties of bulk tetrahedral amorphous GaSb and GaSb–Ge semiconductors under high pressure, Phys. Rev. Lett. 73 (1994) 3262.
[72] M.-C. Wu, C.-C. Chen, Photoluminescence of high-quality GaSb grown from Ga- and Sb-rich solutions by liquid-phase epitaxy, J. Appl. Phys. 72 (1992) 4275.
[73] C. Ghezzi, R. Magnanini, A. Parisini, B. Rotelli, L. Tarricone, A. Bosacchi, S. Franchi, Optical absorption near the fundamental absorption edge in GaSb, Phys. Rev. B 52 (1995) 1463.
[74] V.V. Brazhkin, A.G. Lyapin, V.A. Goncharova, O.V. Stal'gorova, S.V. Popova, Elastic softness of amorphous tetrahedrally bonded GaSb and $Ge_{20.27}GaSb_{0.73}$ semiconductors, Phys. Rev. B 56 (1997) 990.