Computer Physics Communications 184 (2013) 348–366

Contents lists available at SciVerse ScienceDirect

Computer Physics Communications

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

FHI-gap: A *GW* code based on the all-electron augmented plane wave method

Hong Jiang [a,b,*,1], Ricardo I. Gómez-Abal [a,1], Xin-Zheng Li [a], Christian Meisenbichler [c], Claudia Ambrosch-Draxl [c,2], Matthias Scheffler [a]

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

[b] *Beijing National Laboratory for Molecular Sciences, College of Chemistry, Peking University, 100871 Beijing, China*

[c] *Chair of Atomistic Modelling and Design of Materials, University of Leoben, Franz-Josef-Straße 18, A-8700, Austria*

ARTICLE INFO

Article history:
Received 14 May 2012
Received in revised form
31 August 2012
Accepted 10 September 2012
Available online 14 September 2012

Keywords:
Many-body perturbation theory
GW approach
LAPW method
Quasi-particles

ABSTRACT

The *GW* method has become the state-of-the-art approach for the first-principles description of the electronic quasi-particle band structure in crystalline solids. Most of the existing codes rely on pseudopotentials in which only valence electrons are treated explicitly. The pseudopotential method can be problematic for systems with localized *d*- or *f*-electrons, even for ground-state density-functional theory (DFT) calculations. The situation can become more severe in *GW* calculations, because pseudo-wavefunctions are used in the computation of the self-energy and the core–valence interaction is approximated at the DFT level. In this work, we present the package FHI-gap, an all-electron *GW* implementation based on the full-potential linearized augmented planewave plus local orbital (LAPW) method. The FHI-gap code can handle core, semicore, and valence states on the same footing, which allows for a correct treatment of core–valence interaction. Moreover, it does not rely on any pseudopotential or frozen-core approximation. It is, therefore, able to handle a wide range of materials, irrespective of their composition. Test calculations demonstrate the convergence behavior of the results with respect to various cut-off parameters. These include the size of the basis set that is used to expand the products of Kohn–Sham wavefunctions, the number of **k** points for the Brillouin zone integration, the number of frequency points for the integration over the imaginary axis, and the number of unoccupied states. At present, FHI-gap is linked to the WIEN2k code, and an implementation into the exciting code is in progress.

1. Introduction

The central quantity defining the electronic properties of solids is the electronic band structure, typically measured by direct and inverse photo-emission spectroscopy (PES/IPS) [1]. In the past decades, the most commonly used approach to theoretically describe the electronic structure (in particular electron density and total energy) has been the Kohn–Sham (KS) density-functional theory (DFT) in the local-density approximation (LDA) or generalized gradient approximation (GGA) [2,3]. However, comparing KS eigenvalues with experimental electronic band structures lacks rigorous justification, and, hence, may exhibit noticeable discrepancies. In particular, the band gaps of *sp* semiconductors obtained from LDA/GGA KS single-particles energies are systematically underestimated, with errors being of the order of 30%–100% [4]. A proper theoretical framework for the description of electronic band structures of extended systems is provided by many-body perturbation theory based on the interacting Green function $G(\mathbf{r}, \mathbf{r}'; \omega)$ [5], whose poles in the complex frequency plane determine the single-particle excitation energies of the system. All the information about the interaction among electrons is contained in the self-energy, which relates the interacting Green function to the non-interacting one via Dyson's equation. The simplest realistic approximation to the self-energy, treating the exchange interaction exactly and containing dynamical correlation effects, is the *GW* approximation, originally proposed by Hedin [6,7]. Under the assumption that the KS eigenvalues and eigenfunctions constitute a good zeroth-order approximation to the corresponding quasi-particle energies, the self-energy is calculated as a first-order correction to the

* Corresponding author at: Beijing National Laboratory for Molecular Sciences, College of Chemistry, Peking University, 100871 Beijing, China. Tel.: +86 10 62765970.
E-mail address: h.jiang@pku.edu.cn (H. Jiang).

1 Authors with equal contributions.

2 New address: Department of Physics, Humboldt-Universität zu Berlin, Newtonstraße 15, D-12489, Germany.

0010-4655/$ – see front matter
doi:10.1016/j.cpc.2012.09.018

KS eigenvalues. In this approach, termed G_0W_0, both the Green function, G, and the dynamically screened Coulomb potential, W, are calculated using KS eigenvalues and eigenfunctions. We note that the exact KS results may well be a good starting point as shown by the Slater–Janak transition state theorem [8]. However, approximate exchange–correlation functionals suffer from various shortcomings, as, e.g. electron self-interactions. The latter is also corrected by the G_0W_0 approach. If self-interaction or other errors are large, the starting point xc functional may not be good enough and must be replaced by a better one.

Since the seminal work by Hybertsen and Louie [9], and by Godby et al. [10], the G_0W_0 method with LDA as a starting point has established itself as a state-of-the-art tool for calculating the electronic band structure of materials, noticeably improving agreement with experiment over LDA/GGA (see, e.g., Refs. [4,11] and references therein). Often, however, a better starting point is necessary to achieve further improvement [12]. A majority of these calculations have been carried out within the pseudopotential approximation, in which the core–valence interaction is kept at the LDA level and pseudo-wavefunctions are used. The appearance of full-potential all-electron calculations starting in 2002 [13–15] showed a noticeable underestimation of band gaps, in clear contradiction with previous results. Discrepancies were traced back to the approximations implicit in the pseudopotential method [15,16]. These findings highlight the importance of an all-electron treatment not only as benchmark but also as a basis for the development of new electronic-structure methods.

In this paper, we present the implementation of the `FHI-gap` (*GW* with Augmented Planewaves) package, an all-electron *GW* code based on the full-potential linearized augmented planewave (LAPW) method, or its more recent version, the APW plus local-orbital (APW+lo) method. Since both, LAPW and APW+lo, are schemes to linearize the original APW method, we use the terminology LAPW synonymously. Our code has been developed as an add-on to the Wien2k code [17] (see Appendix E), but an according implementation into the `exciting` code [18] is currently in progress. The main features of the methodology used in our package are the following. (i) The products of Kohn–Sham wavefunctions are expanded into a dual basis, consisting of planewaves in the interstitial region and spherical harmonics inside the atom-centered non-overlapping spheres. (ii) The numerical Brillouin zone (BZ) integrations are carried out by a generalized tetrahedron method [19]. (iii) The Γ point ($\mathbf{q} = 0$), at which both bare and screened Coulomb interaction are singular, is included in the BZ integrations, and the divergences are accurately treated within $\mathbf{k} \cdot \mathbf{p}$ perturbation theory. (iv) frequency integrations are carried out on the imaginary axis. (v) In addition, the code is parallelized using the message-passing interface library.

An advantage of our code is the capability to explore *d*- and *f*-electron systems [20,21], materials traditionally categorized as strongly correlated. For such materials, a full-potential all-electron treatment is highly desirable. DFT with LDA or GGA exchange–correlation functionals fails dramatically for many such systems, and so may G_0W_0 carried out on top of LDA or GGA. Here, much (possibly all) of the problem may stem from the LDA/GGA starting point. As a first step towards establishing G_0W_0 for *d* and *f*-electron systems, we have implemented G_0W_0 based on LDA+U. This simple and effective approach has been applied to a series of prototype *d* and *f*-electron systems [20–22] and shown to overcome the major shortcomings of LDA/GGA.

In the following, we will give a short overview of the *GW* approach and the G_0W_0 approximation. This serves as a basis to describe the details of our implementation. We will sketch the basics of the LAPW method and the mixed basis to expand the *GW*-related non-local quantities. The final expressions, as implemented in the code, will be derived thereafter. Convergence tests with respect to several parameters are presented for Si to illustrate how different levels of accuracy can be achieved.

2. Theoretical background

2.1. The GW approach

The single-particle excitation energies of a many-electron system are given by the poles of the interacting single-particle Green function $G(\mathbf{r}, \mathbf{r}'; \omega)$. These can be obtained by solving the quasi-particle equation

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_H(\mathbf{r})\right]\Psi_{n\mathbf{k}}(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; E^{\text{qp}}_{n\mathbf{k}})\,\Psi_{n\mathbf{k}}(\mathbf{r}')d\mathbf{r}' = E^{\text{qp}}_{n\mathbf{k}}\,\Psi_{n\mathbf{k}}(\mathbf{r}), \tag{1}$$

where $V_{\text{ext}}(\mathbf{r})$ and $V_H(\mathbf{r})$ are the external (nuclear) and Hartree potential, respectively, $\Sigma(\mathbf{r}, \mathbf{r}'; \omega)$ is the self-energy, and $\Psi_{n\mathbf{k}}(\mathbf{r})$ is the quasi-particle wavefunction. $\mathbf{k}$ is the Bloch wavevector in the first BZ that accounts for the translational symmetry. Note that atomic units are used throughout the paper.

The self-energy contains all electron–electron interactions beyond the Hartree term, and is in general non-hermitian so that the corresponding eigenvalues, $E^{\text{qp}}_{n\mathbf{k}}$, are complex. Their real parts $\epsilon^{\text{qp}}_{n\mathbf{k}} \equiv \Re\left(E^{\text{qp}}_{n\mathbf{k}}\right)$ represent the single-particle excitation energies of the system, while their imaginary parts are related to the excitation's lifetime. In the *GW* approximation [6], only the first-order in the expansion of the self-energy in terms of the dynamically screened Coulomb potential is included:

$$\Sigma(\mathbf{r}, \mathbf{r}'; \omega) = \frac{i}{2\pi}\int G(\mathbf{r}, \mathbf{r}'; \omega + \omega')W(\mathbf{r}', \mathbf{r}; \omega')\,e^{i\omega'\eta}d\omega' \tag{2}$$

where η is an infinitesimal positive number. The dynamically screened potential $W(\mathbf{r}, \mathbf{r}'; \omega)$ is given by

$$W(\mathbf{r}, \mathbf{r}'; \omega) = v(\mathbf{r}, \mathbf{r}') + \iint v(\mathbf{r}, \mathbf{r}_1)\,P(\mathbf{r}_1, \mathbf{r}_2; \omega)\,W(\mathbf{r}_2, \mathbf{r}'; \omega)\,d\mathbf{r}_1\,d\mathbf{r}_2, \tag{3}$$

where $v(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$ denotes the bare Coulomb potential, and $P(\mathbf{r}, \mathbf{r}'; \omega)$ represents the polarizability in the random-phase approximation (RPA),

$$P(\mathbf{r}, \mathbf{r}'; \omega) = -\frac{i}{2\pi}\int G(\mathbf{r}, \mathbf{r}'; \omega + \omega')\,G(\mathbf{r}', \mathbf{r}; \omega')\,e^{i\omega'\eta}d\omega'. \tag{4}$$

2.2. The G_0W_0 approximation

The solution of the above system of Eqs. (1)–(4) requires, in principle, a self-consistent procedure. Assuming, however, that the respective Kohn–Sham quantities are good zeroth-order approximations to the quasi-particle counterparts, Hybertsen and Louie [9] proposed a scheme where the quasi-particle energy $\omega = \epsilon^{\rm qp}_{n\mathbf{k}}$ is obtained as a first-order correction to the Kohn–Sham eigenvalue $\epsilon_{n\mathbf{k}}$,

$$\epsilon^{\rm qp}_{n\mathbf{k}} = \epsilon_{n\mathbf{k}} + \left\langle \psi_{n\mathbf{k}}(\mathbf{r}) \mid \Re\left[\Sigma\left(\mathbf{r}, \mathbf{r}'; \epsilon^{\rm qp}_{n\mathbf{k}}\right)\right] - V^{\rm xc}(\mathbf{r})\,\delta\left(\mathbf{r}-\mathbf{r}'\right) \mid \psi_{n\mathbf{k}}\left(\mathbf{r}'\right)\right\rangle \tag{5}$$

where $V^{\rm xc}$ is the DFT exchange–correlation potential, and $\psi_{n\mathbf{k}}(\mathbf{r})$ are the Kohn–Sham eigenfunctions. The self-energy $\Sigma(\mathbf{r}, \mathbf{r}'; \omega)$ is given like in Eq. (2) but replacing G by the non-interacting single-particle Green function G_0 obtained from the Kohn–Sham states

$$G_0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{n\mathbf{k}} \frac{\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r}')}{\omega - \tilde{\epsilon}_{n\mathbf{k}}} \tag{6}$$

with $\tilde{\epsilon}_{n\mathbf{k}} \equiv \epsilon_{n\mathbf{k}} + i\eta\,\mathrm{sgn}(\epsilon_F - \epsilon_{n\mathbf{k}})$. The dynamically screened Coulomb potential $W_0(\mathbf{r}, \mathbf{r}'; \omega)$ is given by

$$W_0(\mathbf{r}, \mathbf{r}'; \omega) = \int \varepsilon^{-1}(\mathbf{r}, \mathbf{r}_1; \omega)\, v(\mathbf{r}_1, \mathbf{r}')\, d\mathbf{r}_1 \tag{7}$$

where $\varepsilon(\mathbf{r}, \mathbf{r}'; \omega)$ is the dielectric function calculated as

$$\varepsilon(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r}, \mathbf{r}') - \int v(\mathbf{r}, \mathbf{r}_1)\, P_0(\mathbf{r}_1, \mathbf{r}'; \omega)\, d\mathbf{r}_1. \tag{8}$$

$P_0(\mathbf{r}, \mathbf{r}'; \omega)$ is the RPA polarizability computed according to Eq. (4), but with G replaced by G_0. Carrying out the integration over ω' we obtain:

$$\begin{aligned} P_0(\mathbf{r}, \mathbf{r}'; \omega) &= 2\sum_{n\mathbf{k}, m\mathbf{k}'} f_{n\mathbf{k}}(1 - f_{m\mathbf{k}'})\psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{m\mathbf{k}'}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r}')\psi_{m\mathbf{k}'}(\mathbf{r}')\left\{\frac{1}{\omega - \epsilon_{m\mathbf{k}'} + \epsilon_{n\mathbf{k}} + i\eta} - \frac{1}{\omega + \epsilon_{m\mathbf{k}'} - \epsilon_{n\mathbf{k}} - i\eta}\right\} \\ &= \sum_{n,m}\sum_{\mathbf{k},\mathbf{q}} F_{nm}(\mathbf{k}, \mathbf{q}; \omega)\ \psi_{n\mathbf{k}}(\mathbf{r})\psi^*_{m\mathbf{k}-\mathbf{q}}(\mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r}')\psi_{m\mathbf{k}-\mathbf{q}}(\mathbf{r}') \end{aligned} \tag{9}$$

where $f_{n\mathbf{k}}$ denotes the occupation number of the state $n\mathbf{k}$, and the factor of 2 accounts for the spin degeneracy. Here, we have introduced $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ and the factor

$$F_{nm}(\mathbf{k}, \mathbf{q}; \omega) \equiv 2f_{n\mathbf{k}}(1 - f_{m\mathbf{k}-\mathbf{q}})\left\{\frac{1}{\omega - \epsilon_{m\mathbf{k}-\mathbf{q}} + \epsilon_{n\mathbf{k}} + i\eta} - \frac{1}{\omega + \epsilon_{m\mathbf{k}-\mathbf{q}} - \epsilon_{n\mathbf{k}} - i\eta}\right\} \tag{10}$$

depending on frequency and occupation.

It is common and convenient to decompose the self-energy into exchange and correlation terms by defining

$$W^c_0(\mathbf{r}, \mathbf{r}'; \omega) = W_0(\mathbf{r}, \mathbf{r}'; \omega) - v(\mathbf{r}, \mathbf{r}'). \tag{11}$$

The exchange part, which is just given by the Hartree–Fock exchange potential, reads

$$\begin{aligned} \Sigma^x(\mathbf{r}, \mathbf{r}') &= \frac{i}{2\pi}\int G_0(\mathbf{r}, \mathbf{r}'; \omega')v(\mathbf{r}', \mathbf{r})e^{i\omega'\eta}d\omega' \\ &= -\sum_{n\mathbf{k}} f_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})v(\mathbf{r}', \mathbf{r})\psi^*_{n\mathbf{k}}(\mathbf{r}'). \end{aligned} \tag{12}$$

The GW correlation self-energy is obtained from a frequency integral as

$$\Sigma^c(\mathbf{r}, \mathbf{r}'; \omega) = \frac{i}{2\pi}\int G_0(\mathbf{r}, \mathbf{r}'; \omega + \omega')W^c_0(\mathbf{r}', \mathbf{r}; \omega')\, d\omega'. \tag{13}$$

2.3. The G_0W_0 equations in matrix form

Let us start with a few words on the notation: all wavefunctions as well as the basis functions are defined such that they are normalized within the unit cell with volume Ω. When a function $f^{\mathbf{k}}(\mathbf{r})$, characterized by the wavevector $\mathbf{k}$, is involved in integrals over the whole crystal space, $V \equiv N_c\Omega$, with N_c being the number of unit cells, we will use its Bloch form

$$f^{\mathbf{k}}(\mathbf{r}) \to \frac{1}{\sqrt{N_c}}\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} f^{\mathbf{k}}(\mathbf{r} + \mathbf{R}), \tag{14}$$

where $\mathbf{R}$ are Bravais lattice vectors.

To convert the G_0W_0 equations into matrix form, we need a basis set, $\chi_i^{\mathbf{q}}(\mathbf{r})$, that can describe products of two Kohn–Sham wavefunctions accurately. In such representation,

$$\psi_{n\mathbf{k}}(\mathbf{r})\,\psi^*_{m\mathbf{k}-\mathbf{q}}(\mathbf{r}) = \sum_i M^i_{nm}(\mathbf{k},\mathbf{q})\,\chi_i^{\mathbf{q}}(\mathbf{r}), \tag{15}$$

$M^i_{nm}(\mathbf{k},\mathbf{q})$ are expansion coefficients given by

$$M^i_{nm}(\mathbf{k},\mathbf{q}) \equiv \int_\Omega \left[\chi_i^{\mathbf{q}}(\mathbf{r})\,\psi_{m\mathbf{k}-\mathbf{q}}(\mathbf{r})\right]^* \psi_{n\mathbf{k}}(\mathbf{r})d\mathbf{r}. \tag{16}$$

We will denote this basis the *product basis* to distinguish it from the basis that is used to expand single-particle orbitals. All two-particle functions, O, including the polarizability, $P_0(\mathbf{r},\mathbf{r}';\omega)$, the bare Coulomb interaction, $v(\mathbf{r},\mathbf{r}')$, the dielectric function, $\varepsilon(\mathbf{r},\mathbf{r}';\omega)$, and the screened Coulomb interaction, $W(\mathbf{r},\mathbf{r}';\omega)$, will be expanded into this product basis according to

$$O(\mathbf{r},\mathbf{r}') = \sum_{\mathbf{q}}^{BZ}\sum_{i,j} O_{ij}(\mathbf{q})\,\chi_i^{\mathbf{q}}(\mathbf{r})\left[\chi_j^{\mathbf{q}}(\mathbf{r}')\right]^*. \tag{17}$$

The diagonal elements of the bare (Hartree–Fock) exchange self-energy in the product basis read:

$$\begin{aligned}\Sigma^x_{n\mathbf{k}} &\equiv \int_V d\mathbf{r}\int_V \mathbf{r}'\,[\psi_{n\mathbf{k}}(\mathbf{r})]^*\,\Sigma^x(\mathbf{r},\mathbf{r}')\,\psi_{n\mathbf{k}}(\mathbf{r}')\\ &= -\sum_{m\mathbf{k}'} f_{m\mathbf{k}'}\int_V d\mathbf{r}\int_V d\mathbf{r}'\,[\psi_{n\mathbf{k}}(\mathbf{r})]^*\,\psi_{m\mathbf{k}'}(\mathbf{r})\,v(\mathbf{r},\mathbf{r}')\,\psi_{n\mathbf{k}}(\mathbf{r}')\left[\psi_{m\mathbf{k}'}(\mathbf{r}')\right]^*,\\ &= -\frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\sum_{i,j} v_{ij}(\mathbf{q})\sum_m f_{m\mathbf{k}-\mathbf{q}}\left[M^i_{nm}(\mathbf{k},\mathbf{q})\right]^* M^j_{nm}(\mathbf{k},\mathbf{q}),\end{aligned} \tag{18}$$

where $v_{ij}(\mathbf{q})$ is the bare Coulomb potential in the product basis representation. The matrix elements of the polarizability (Eq. (9)), $P_{ij}(\mathbf{q},\omega)$, can be written as

$$\begin{aligned}P_{ij}(\mathbf{q},\omega) &\equiv \int_V\int_V \left[\chi_i^{\mathbf{q}}(\mathbf{r})\right]^* P(\mathbf{r},\mathbf{r}';\omega)\,\chi_j^{\mathbf{q}}(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'\\ &= \frac{1}{N_c}\sum_{\mathbf{k}}^{BZ}\sum_{n,m} F_{nm}(\mathbf{k},\mathbf{q};\omega)\,M^i_{nm}(\mathbf{k},\mathbf{q})\left[M^j_{nm}(\mathbf{k},\mathbf{q})\right]^*.\end{aligned} \tag{19}$$

Due to the singularity of the bare Coulomb interaction in reciprocal space as $\mathbf{q}$ goes to zero, the dielectric function at $\mathbf{q}\to 0$ requires a special treatment. Mathematically, it is more convenient to use the symmetrized form of the dielectric function [23], which, in matrix form, can be obtained from $P_{ij}(\mathbf{q},\omega)$ by

$$\varepsilon_{ij}(\mathbf{q},\omega) = \delta_{ij} - \sum_{lm} v_{il}^{\frac{1}{2}}(\mathbf{q})P_{lm}(\mathbf{q},\omega)v_{mj}^{\frac{1}{2}}(\mathbf{q}). \tag{20}$$

The correlation term of the screened Coulomb interaction (Eq. (7)) can then be calculated through

$$W^c_{ij}(\mathbf{q},\omega) = \sum_{lm} v_{il}^{\frac{1}{2}}(\mathbf{q})\left[\varepsilon^{-1}_{lm}(\mathbf{q},\omega) - \delta_{lm}\right]v_{mj}^{\frac{1}{2}}. \tag{21}$$

The diagonal matrix elements of the correlation self-energy can be written as

$$\begin{aligned}\Sigma^c_{n\mathbf{k}}(\omega) &\equiv \int_V d\mathbf{r}\int_V \mathbf{r}'\,[\psi_{n\mathbf{k}}(\mathbf{r})]^*\,\Sigma^c(\mathbf{r},\mathbf{r}';\omega)\,\psi_{n\mathbf{k}}(\mathbf{r}')\\ &= \sum_{m\mathbf{k}'}\int d\mathbf{r}\int d\mathbf{r}'\frac{i}{2\pi}\int d\omega'\frac{1}{\omega+\omega'-\tilde{\epsilon}_{m\mathbf{k}'}}\,[\psi_{n\mathbf{k}}(\mathbf{r})]^*\,\psi_{m\mathbf{k}'}(\mathbf{r})\,W^c_0(\mathbf{r},\mathbf{r}';\omega')\,\psi_{n\mathbf{k}}(\mathbf{r}')\left[\psi_{m\mathbf{k}'}(\mathbf{r}')\right]^*\\ &= \frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\sum_m \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega'\,\frac{X_{nm}(\mathbf{k},\mathbf{q};\omega')}{\omega+\omega'-\tilde{\epsilon}_{m\mathbf{k}-\mathbf{q}}}\end{aligned} \tag{22}$$

where the auxiliary quantity $X_{nm}(\mathbf{k},\mathbf{q};\omega)$ is defined as

$$X_{nm}(\mathbf{k},\mathbf{q};\omega) \equiv \sum_{ij}\left[M^i_{nm}(\mathbf{k},\mathbf{q})\right]^* W^c_{ij}(\mathbf{q},\omega)M^j_{nm}(\mathbf{k},\mathbf{q}). \tag{23}$$

As we can see from Eqs. (21) and (23), M^i_{nm} often appears together with $v_{ij}^{\frac{1}{2}}$. We therefore introduce

$$\tilde{M}^i_{nm}(\mathbf{k},\mathbf{q}) \equiv \sum_p v_{ip}^{\frac{1}{2}}(\mathbf{q})M^p_{nm}(\mathbf{k},\mathbf{q}). \tag{24}$$

Using this contracted quantity gives

$$\varepsilon_{ij}(\mathbf{q},\omega) = \delta_{ij} - \frac{1}{N_c}\sum_{\mathbf{k}}^{BZ}\sum_{n,m} F_{nm}(\mathbf{k},\mathbf{q};\omega)\,\tilde{M}^{i}_{nm}(\mathbf{k},\mathbf{q})\left[\tilde{M}^{j}_{nm}(\mathbf{k},\mathbf{q})\right]^{*} \tag{25}$$

and

$$X_{nm}(\mathbf{k},\mathbf{q};\omega) \equiv \sum_{ij}\left[\tilde{M}^{i}_{nm}(\mathbf{k},\mathbf{q})\right]^{*}\left[\varepsilon^{-1}_{ij}(\mathbf{q},\omega) - \delta_{ij}\right]\tilde{M}^{j}_{nm}(\mathbf{k},\mathbf{q}). \tag{26}$$

The exchange part of the self-energy now reads

$$\Sigma^{x}_{n\mathbf{k}} = -\frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\sum_{i,j}\sum_{m}^{\mathrm{occ}}\left[\tilde{M}^{i}_{nm}(\mathbf{k},\mathbf{q})\right]^{*}\tilde{M}^{j}_{nm}(\mathbf{k},\mathbf{q}). \tag{27}$$

In this way, an explicit construction of the polarizability and screened Coulomb interaction matrix can be avoided.

The G_0W_0 quasi-particle energies are then given by

$$\epsilon^{\mathrm{qp}}_{n\mathbf{k}} = \epsilon_{n\mathbf{k}} + Z_{n\mathbf{k}}\left[\Re\,\Sigma^{c}_{n\mathbf{k}}(\epsilon_{n\mathbf{k}}) + \Sigma^{x}_{n\mathbf{k}} - V^{\mathrm{xc}}_{n\mathbf{k}}\right] \tag{28}$$

where $V^{\mathrm{xc}}_{n\mathbf{k}}$ are the diagonal matrix elements of the exchange–correlation potential that is employed in the single-particle KS Hamiltonian, and $Z_{n\mathbf{k}}$ is the so-called QP renormalization factor arising from the energy dependence of the correlation self-energy,

$$Z_{n\mathbf{k}} = \left[1 - \left(\frac{\partial}{\partial\omega}\Re\,\Sigma^{c}_{n\mathbf{k}}(\omega)\right)_{\epsilon_{n\mathbf{k}}}\right]^{-1}. \tag{29}$$

3. Basics of LAPW

The basic idea behind the LAPW method is that, close to nuclei, potential and wavefunctions are similar to those in the free atoms, and slowly varying in between the atoms. Hence, Slater [24] proposed to partition the unit cell into non-overlapping muffin-tin (MT) spheres, centered around each atom (indexed by α, positioned at $\mathbf{r}_\alpha$), and the interstitial region (I). Correspondingly, the LAPW basis consists of two types of functions inside the two types of volumes [25]:

$$\phi^{\mathbf{k}}_{\mathbf{G}}(\mathbf{r}) = \begin{cases} \displaystyle\sum_{lm}\left[A_{\alpha lm}(\mathbf{k}+\mathbf{G})u_{\alpha l}(r^{\alpha},E_l) + B_{\alpha lm}(\mathbf{k}+\mathbf{G})\dot{u}_{\alpha l}(r^{\alpha},E_l)\right]Y_{lm}(\hat{\mathbf{r}}^{\alpha}) & r^{\alpha} < R^{\alpha}_{\mathrm{MT}}, \\ \dfrac{1}{\sqrt{\Omega}}e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} & \mathbf{r}\in I, \end{cases} \tag{30}$$

with $\mathbf{r}^{\alpha} \equiv \mathbf{r} - \mathbf{r}_\alpha$ and R^{α}_{MT} being the radius of α-th MT sphere. $u_l(r^{\alpha},\epsilon)$ are the solutions of the radial Schrödinger equation in the spherical potential of the respective MT sphere, taken at a fixed reference energy E_l and $\dot{u}_{\alpha l}(r^{\alpha},E_l)$ is its energy derivative. The augmentation coefficients, $A_{\alpha lm}(\mathbf{k}+\mathbf{G})$ and $B_{\alpha lm}(\mathbf{k}+\mathbf{G})$, are determined from the continuity of the basis functions and their first derivatives at the sphere boundary.

The LAPW functions can be supplemented by the so-called local orbitals (LO), defined as [26]:

$$\phi^{\mathrm{LO}}_{lm}(\mathbf{r}) = \begin{cases} \left[A^{\mathrm{LO}}_{\alpha lm}u_{\alpha l}(r^{\alpha},E_l) + B^{\mathrm{LO}}_{\alpha lm}\dot{u}_{\alpha l}(r^{\alpha},E_l) + C^{\mathrm{LO}}_{\alpha lm}u_{\alpha l}(r^{\alpha},E^{(2)}_l)\right]Y_{lm}(\hat{r}^{\alpha}) & r^{\alpha} < R^{\alpha}_{\mathrm{MT}} \\ 0 & \mathbf{r}\in I. \end{cases} \tag{31}$$

Here $u_{\alpha l}(r,E^{(2)}_l)$ is a radial wavefunction evaluated at a different linearization energy, $E^{(2)}_l$, enabling the proper treatment of a semicore state. The coefficients $A^{\mathrm{LO}}_{\alpha lm}$, $B^{\mathrm{LO}}_{\alpha lm}$, and $C^{\mathrm{LO}}_{\alpha lm}$ are determined by the requirement that $\phi^{\mathrm{LO}}_{lm}(\mathbf{r})$ is normalized and is zero in value and slope at the MT sphere boundary.

Eqs. (30) and (31) contain the basis functions for the APW+lo method [27] as special cases. In this scheme, $B_{\alpha lm}(\mathbf{k}+\mathbf{G}) = 0$, reducing an LAPW to an APW basis function, and for the relevant valence states these functions are supplemented by a local orbital with $C^{\mathrm{LO}}_{\alpha lm} = 0$, taken at the same trial energy E_l as the corresponding APW.

One can write down the basis set of all APW-related methods in its most general form as

$$\phi^{\mathbf{k}}_{\mathbf{G}}(\mathbf{r}) = \begin{cases} \displaystyle\sum_{\zeta lm} A_{\alpha\zeta lm}(\mathbf{k}+\mathbf{G})u_{\alpha\zeta l}(r^{\alpha})Y_{lm}(\hat{r}^{\alpha}) & r^{\alpha} < R^{\alpha}_{\mathrm{MT}}, \\ \dfrac{\theta^{\mathrm{LO}}_{\mathbf{G}}}{\sqrt{\Omega}}e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} & \mathbf{r}\in I. \end{cases} \tag{32}$$

$\theta^{\mathrm{LO}}_{\mathbf{G}}$ is equal to one for a normal APW or LAPW function, but zero for a local orbital. The summation in the first term over lm is reduced to a single term when ϕ corresponds to a LO function. The index ζ has been introduced to denote different possible radial functions, including energy derivatives as appearing in Eq. (30). This notation also comprises higher-order basis functions as implemented in the `exciting` code [28,18].

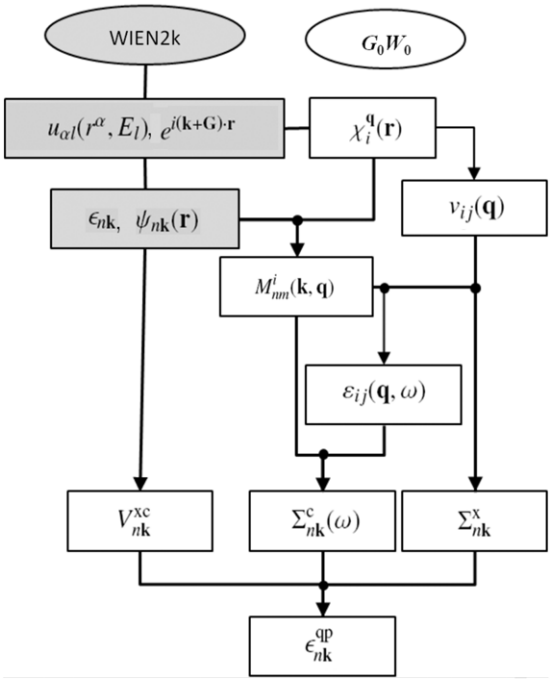

Fig. 1. Flowchart of the FHI-gap code. The gray boxes indicate quantities obtained from the ground-state calculation.

4. Implementation

In this section, we present the implementation of the FHI-gap package. The basic structure of the code is illustrated by the flowchart in Fig. 1. Each block represents an important quantity, and the lines connecting different blocks illustrate the inter-dependence of these quantities. The blocks marked by the gray background display results obtained from the ground-state package. In particular, KS eigenvalues and eigenfunctions (together with the basis functions), corresponding to a converged DFT potential, are the starting point for the G_0W_0 run.

We will present below details about the implementation of all the quantities, which are relevant for the G_0W_0 calculations. Before that, however, we want to briefly address partial self-consistency, which has been discussed in the literature by many authors [4,29,30], and often leads to further improvements. In such schemes, the diagonal correlation self-energy, Eq. (22), is recalculated by quasi-particle energies with fixed W_0, hence it is called GW_0, or the *energy-only* GW_0 [29,30]. The GW_0 quasi-particle energies can be easily obtained with little computational overhead, since the auxiliary matrix $X_{nm}(\mathbf{k}, \mathbf{q}; \omega)$ as defined in Eq. (23), which is one of the most time-consuming quantities, is fixed here and therefore needs to be calculated only once. This feature is also implemented in our code.

4.1. The product basis functions

The LAPW basis is a mixed one, where the basis functions have different character in the MT region and the interstitial. To represent products of two Kohn–Sham wavefunctions accurately, a basis with similar features should be used. In this section, we present the procedure to construct an optimal product basis within the LAPW framework.

The product of two spherical harmonics can be expanded in spherical harmonics using Clebsch–Gordan coefficients, and the product of two planewaves is naturally a planewave. Hence a *mixed basis set* making use of the space partitioning into muffin-tin spheres and interstitial region is straightforward. Analogous to the suggestion of Kotani and van Schilfgaarde [13], we use an optimized set of functions consisting of planewaves in the interstitial region and a spherical harmonics expansion within the MT spheres. The latter are similar to the product basis originally proposed by Aryasetiawan and Gunnarsson [31].

Inside the MT sphere of atom α, we define our basis functions as

$$\gamma_{\alpha NLM}(\mathbf{r}^\alpha) = \upsilon_{\alpha NL}(r^\alpha)\, Y_{LM}\left(\hat{\mathbf{r}}^\alpha\right). \tag{33}$$

An optimal set of radial functions $\upsilon_{\alpha NL}(r^\alpha)$ is constructed from the products of the radial wavefunctions (see Eq. (30)) by the following procedure.

- Following Ref. [32], we take only $u_{\alpha l}(r^\alpha)$'s with $l \leq l^{\mathrm{MB}}_{\mathrm{max}}$, and $\dot{u}_{\alpha l}(r)$'s are not taken into account. A careful analysis regarding the impact of this approximation on the self-energy of various materials is under way.
- For each L in $\upsilon_{\alpha NL}(r^\alpha)$, we take all the products of two radial functions $u_{\alpha l}(r^\alpha)u_{\alpha l'}(r^\alpha)$ which fulfill the triangular condition $|l - l'| \leq L \leq l + l'$.
- We calculate the overlap matrix between these product functions:

$$\mathbb{O}_{ll';l_1l'_1} = \int_0^{R^\alpha_{\mathrm{MT}}} u_{\alpha l}(r^\alpha)u_{\alpha l'}(r^\alpha)u_{\alpha l_1}(r^\alpha)u_{\alpha l'_1}(r^\alpha)(r^\alpha)^2 dr^\alpha. \tag{34}$$

- The matrix $\mathbb{O}_{(ll');(l_1l_1')}$ is diagonalized, yielding the corresponding set of eigenvalues, λ_N^{MB}, and eigenvectors, $\{c_{ll',N}\}$.
- Eigenvectors corresponding to eigenvalues smaller than a certain tolerance, $\lambda_{\mathrm{min}}^{\mathrm{MB}}$, are assumed to be linearly dependent and discarded; $\lambda_{\mathrm{min}}^{\mathrm{MB}}$ is typically chosen to be 10^{-4} [33].
- The remaining eigenvectors, after normalization, form the radial basis set:

$$\upsilon_{\alpha NL}(r^\alpha) = \sum_{ll'} c_{ll',N} u_l(r^\alpha) u_{l'}(r^\alpha). \tag{35}$$

So defined, the functions $\{\gamma_{\alpha NLM}\}$ constitute an orthonormal basis. Taking into account the crystalline translational symmetry, we obtain the optimal product basis functions belonging to wavevector $\mathbf{q}$ for the MT region

$$\gamma_{\alpha NLM}^{\mathbf{q}}(\mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}_\alpha} \upsilon_{\alpha NL}(r^\alpha) Y_{LM}(\hat{\mathbf{r}}^\alpha). \tag{36}$$

In the interstitial region, the product basis is constructed from planewaves. These interstitial planewaves (IPW's) are not orthogonal, and the overlap of two IPW's is given by

$$\mathbb{O}_{\mathbf{G}\mathbf{G}'} = \frac{1}{\Omega}\int_\Omega \theta_I(\mathbf{r})\, e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}}\, d^3r \equiv \frac{1}{\Omega}\mathcal{I}_{\mathbf{G}-\mathbf{G}'} \tag{37}$$

where $\theta_I(\mathbf{r})$ is one in the interstitial, and vanishes otherwise. $\mathcal{I}_{\mathbf{G}}$ is calculated as

$$\mathcal{I}_{\mathbf{G}} = \begin{cases} \Omega - \sum_\alpha V_{\mathrm{MT}}^\alpha & \mathbf{G} = 0 \\ -3\sum_\alpha V_{\mathrm{MT}}^\alpha\, e^{i\mathbf{G}\cdot\mathbf{r}_\alpha} \dfrac{\sin(GR_{\mathrm{MT}}^\alpha) - GR_{\mathrm{MT}}^\alpha \cos(GR_{\mathrm{MT}}^\alpha)}{(GR_{\mathrm{MT}}^\alpha)^3} & \mathbf{G} \neq 0 \end{cases} \tag{38}$$

with V_{MT}^α being the volume of MT sphere α. To obtain a set of orthogonal wavefunctions, we diagonalize the overlap matrix by solving the eigenvalue equation

$$\sum_{\mathbf{G}'} \mathbb{O}_{\mathbf{G}\mathbf{G}'} S_{\mathbf{G}'i} = \lambda_i^I S_{\mathbf{G}i}. \tag{39}$$

The orthonormal basis set in the interstitial region is then defined by

$$P_i^{\mathbf{q}}(\mathbf{r}) \equiv \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} \tilde{S}_{\mathbf{G}i} e^{i(\mathbf{G}+\mathbf{q})\cdot\mathbf{r}} \theta_I(\mathbf{r}) \tag{40}$$

with

$$\tilde{S}_{\mathbf{G}i} \equiv \frac{S_{\mathbf{G}i}}{\sqrt{\lambda_i^I}}. \tag{41}$$

The planewave expansion is truncated at a certain $G_{\mathrm{max}}^{\mathrm{MB}}$, which is the parameter determining the quality of the mixed basis in the interstitial region.

To summarize, our orthonormal mixed basis set is given by

$$\left\{\chi_j^{\mathbf{q}}(\mathbf{r})\right\} \equiv \left\{\gamma_{\alpha NLM}^{\mathbf{q}}(\mathbf{r}), P_i^{\mathbf{q}}(\mathbf{r})\right\}, \tag{42}$$

whose quality is mainly determined by the cut-off parameters $l_{\mathrm{max}}^{\mathrm{MB}}$ and $G_{\mathrm{max}}^{\mathrm{MB}}$.

This mixed basis is optimal to represent the products of Kohn–Sham wavefunctions. On the other hand, planewaves are simpler to use. In particular, the bare Coulomb interaction $v(\mathbf{r}, \mathbf{r}')$ is diagonal in the planewave representation. To construct the matrix elements in the mixed basis representation, it is often necessary to take advantage of the simplicity of the planewave representation. For that purpose, we use the overlap integrals,

$$\mathcal{W}_{\mathbf{G}}^i(\mathbf{q}) \equiv \frac{1}{\sqrt{\Omega}} \int_\Omega \left[\chi_i^{\mathbf{q}}(\mathbf{r})\right]^* e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} d\mathbf{r}, \tag{43}$$

between these two basis sets frequently. For $\chi_i^{\mathbf{q}}(\mathbf{r}) \in$ the MT region, *i.e.*, $\chi_i^{\mathbf{q}}(\mathbf{r}) \to \gamma_{\alpha NLM}^{\mathbf{q}}(\mathbf{r})$, this expression becomes

$$\mathcal{W}_{\mathbf{G}}^i(\mathbf{q}) = \frac{4\pi}{\sqrt{\Omega}} e^{i\mathbf{G}\cdot\mathbf{r}_\alpha} i^L Y_{LM}^*(\widehat{\mathbf{q}+\mathbf{G}}) \mathbb{J}_{\alpha NL}(|\mathbf{G}+\mathbf{q}|) \tag{44}$$

with

$$\mathbb{J}_{\alpha NL}(|\mathbf{G}+\mathbf{q}|) \equiv \int_0^{R_{\mathrm{MT}}^\alpha} j_L(|\mathbf{G}+\mathbf{q}|\, r)\, \upsilon_{\alpha NL}(r) r^2 dr. \tag{45}$$

For $\chi_i^{\mathbf{q}}(\mathbf{r}) \in I$, *i.e.*, $\chi_i^{\mathbf{q}}(\mathbf{r}) \to P_i^{\mathbf{q}}(\mathbf{r})$, we obtain

$$\mathcal{W}_{\mathbf{G}}^i(\mathbf{q}) \equiv \frac{1}{\sqrt{\Omega}} \int_\Omega \left[P_i^{\mathbf{q}}(\mathbf{r}_1)\right]^* e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}_1} d\mathbf{r}_1 = \frac{1}{\Omega} \sum_{\mathbf{G}'} \mathcal{I}_{\mathbf{G}-\mathbf{G}'} \tilde{S}_{\mathbf{G}'i}^*. \tag{46}$$

4.2. Evaluation of $M^i_{nm}(\mathbf{k},\mathbf{q})$

The expansion coefficients for the product of two Kohn–Sham eigenvectors, $M^i_{nm}(\mathbf{k},\mathbf{q})$, as defined in Eq. (16), are used in both the construction of the polarizability matrix (Eq. (19)) and the evaluation of the self-energy matrix elements (Eqs. (18) and (22)) and therefore play a central role in the implementation of our G_0W_0 approach. In this section, we present detailed expressions for the evaluation of $M^i_{nm}(\mathbf{k},\mathbf{q})$.

Using the expansion of the Kohn–Sham wavefunctions in terms of LAPW basis functions, $\phi^{\mathbf{k}}_{\mathbf{G}}(\mathbf{r})$ (Eq. (32)),

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n\mathbf{k},\mathbf{G}}\phi^{\mathbf{k}}_{\mathbf{G}}(\mathbf{r}) \tag{47}$$

$\psi_{n\mathbf{k}}(\mathbf{r})$ can be written as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \begin{cases} \sum_{\zeta lm} \mathcal{A}_{n\mathbf{k},\alpha\zeta lm}\, u_{\alpha\zeta l}(r^\alpha)\, Y_{lm}(\hat{\mathbf{r}}^\alpha) & r^\alpha < R^\alpha_{\mathrm{MT}}, \\ \dfrac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} \theta^{\mathrm{LO}}_{\mathbf{G}}\, C_{n\mathbf{k},\mathbf{G}}\, e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} & \mathbf{r}\in I \end{cases} \tag{48}$$

with

$$\mathcal{A}_{n\mathbf{k},\alpha\zeta lm} \equiv \sum_{\mathbf{G}} C_{n\mathbf{k},\mathbf{G}}\, A_{\alpha\zeta lm}(\mathbf{k}+\mathbf{G}). \tag{49}$$

The calculation of the matrix elements obviously differs for the two situations that $\chi^{\mathbf{q}}_i$ corresponds to a function in the muffin-tin sphere or to an interstitial planewave. For a basis function in the muffin-tin region, $\chi^{\mathbf{q}}_i \to \gamma^{\mathbf{q}}_{\alpha NLM}(\mathbf{r})$, one obtains with Eqs. (36) and (48):

$$M^i_{nm}(\mathbf{k},\mathbf{q}) = e^{-i\mathbf{q}\cdot\mathbf{r}_\alpha} \sum_{l_1m_1}\sum_{l_2m_2}\left[\mathcal{G}^{l_1m_1}_{LM;l_2m_2}\right]^* \sum_{\zeta_1\zeta_2} \mathcal{A}_{n\mathbf{k},\alpha\zeta_1l_1m_1}\mathcal{A}^*_{m\mathbf{k}-\mathbf{q},\alpha\zeta_2l_2m_2}\langle NL|\zeta_1l_1,\zeta_2l_2\rangle_\alpha. \tag{50}$$

Here $\mathcal{G}^{lm}_{l_1m_1;l_2m_2}$ are the corresponding Gaunt coefficients, defined as

$$\mathcal{G}^{lm}_{l_1m_1;l_2m_2} \equiv \int d\hat{\mathbf{r}}\left[Y_{lm}(\hat{\mathbf{r}})\right]^* Y_{l_1m_1}(\hat{\mathbf{r}})Y_{l_2m_2}(\hat{\mathbf{r}}), \tag{51}$$

where the notation

$$\langle NL|\zeta_1l_1,\zeta_2l_2\rangle_\alpha \equiv \int_0^{R^\alpha_{MT}} v_{\alpha NL}(r)\, u_{\alpha\zeta_1l_1}(r)\, u_{\alpha\zeta_2l_2}(r)\, r^2dr \tag{52}$$

has been introduced. When $\chi^{\mathbf{q}}_i(\mathbf{r})$ belongs to the interstitial region, $\chi^{\mathbf{q}}_i \to P^{\mathbf{q}}_i(\mathbf{r})$, as given by Eq. (40), the corresponding expression reads:

$$M^i_{nm}(\mathbf{k},\mathbf{q}) \equiv \frac{1}{\Omega^{\frac{3}{2}}}\sum_{\mathbf{G}\mathbf{G}'} C_{n\mathbf{k},\mathbf{G}}\left[C_{m\mathbf{k}-\mathbf{q},\mathbf{G}'}\right]^* \sum_{\mathbf{G}_1} \mathcal{I}_{\mathbf{G}-\mathbf{G}'-\mathbf{G}_1}\tilde{S}^*_{\mathbf{G}_1 i}. \tag{53}$$

4.3. The bare Coulomb potential

For the matrix elements of the bare Coulomb potential in the mixed basis,

$$\begin{aligned} v_{ij}(\mathbf{q}) &\equiv \int_V\int_V \left[\chi^{\mathbf{q}}_i(\mathbf{r})\right]^* v(\mathbf{r},\mathbf{r}')\chi^{\mathbf{q}}_j(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &= \int_\Omega\int_\Omega \left[\chi^{\mathbf{q}}_i(\mathbf{r})\right]^* \sum_{\mathbf{R}} v_c(\mathbf{r},\mathbf{r}'-\mathbf{R})\, e^{-i\mathbf{q}\cdot\mathbf{R}}\chi^{\mathbf{q}}_j(\mathbf{r}')\, d\mathbf{r}'d\mathbf{r}, \end{aligned} \tag{54}$$

we need to distinguish between three different cases according to the different domains of the mixed basis functions. For a detailed derivation of the formulas, see Ref. [33]. Note that the bare Coulomb interaction in reciprocal space is singular at the Γ point ($\mathbf{q}\to 0$) and hence requires special treatment, which is described in Appendix A.

4.3.1. Case 1: $\chi^{\mathbf{q}}_i(\mathbf{r})$ and $\chi^{\mathbf{q}}_j(\mathbf{r})$ with $\mathbf{r}\in$ MT spheres

When $\mathbf{r}$ for both functions is inside MT spheres, we have to insert $\chi^{\mathbf{q}}_i(\mathbf{r}) = \gamma^{\mathbf{q}}_{\alpha NLM}(\mathbf{r})$ and $\chi^{\mathbf{q}}_j(\mathbf{r}) = \gamma^{\mathbf{q}}_{\alpha'N'L'M'}(\mathbf{r})$ into Eq. (54). Making use of the Laplace expansion for the Coulomb potential in terms of spherical harmonics and the addition theorem for lattice harmonics [34], we obtain

$$v_{ij}(\mathbf{q}) = (-1)^M \mathbb{S}^{\alpha\alpha'}_{LM;L'M'}(\mathbf{q})\mathbb{R}_{\alpha NL}\mathbb{R}_{\alpha'N'L'} + \frac{4\pi}{2L+1}\mathbb{D}^{\alpha L}_{NN'}\delta_{LL'}\delta_{MM'}\delta_{\alpha\alpha'}. \tag{55}$$

The radial integrals, $\mathbb{R}_{\alpha NL}$ and $\mathbb{D}^{\alpha L}_{NN'}$ are

$$\mathbb{R}_{\alpha NL} \equiv \int_0^{R^\alpha_{\rm MT}} \upsilon_{\alpha NL}(r)\, r^{L+2} dr \tag{56}$$

$$\mathbb{D}^{\alpha L}_{NN'} \equiv \int_0^{R^\alpha_{\rm MT}} dr_1 \int_0^{R^\alpha_{\rm MT}} dr_2\, r_1^2 r_2^2\, \upsilon_{\alpha NL}(r_1) \frac{r_<^L}{r_>^{L+1}} \upsilon_{\alpha N'L}(r_2), \tag{57}$$

where $r_< = \min(r_1, r_2)$ and $r_> = \max(r_1, r_2)$. The lattice structure constants $\mathbb{S}^{\alpha\alpha'}_{l'm';lm}(\mathbf{q})$ are given by

$$\mathbb{S}^{\alpha\alpha'}_{l'm';lm}(\mathbf{q}) = \tilde{g}_{l'm';lm}\, \Xi^{\alpha\alpha'}_{l'+l,m'+m}(-\mathbf{q}). \tag{58}$$

Thereby we have introduced the abbreviation

$$\tilde{g}_{lm;l'm'} = (-1)^l (4\pi)^{\frac{3}{2}} \sqrt{\frac{(l+l'+m+m')!(l+l'-m-m')!}{(2l+1)(2l'+1)[2(l+l')+1](l+m)!(l-m)!(l'+m')!(l'-m')!}} \tag{59}$$

and the lattice sums

$$\Xi^{\alpha\alpha'}_{lm}(\mathbf{q}) = \sum_{\mathbf{R}} \frac{e^{i\mathbf{q}\cdot\mathbf{R}_{\alpha\alpha'}}}{R^{l+1}_{\alpha\alpha'}} Y_{lm}(\hat{R}_{\alpha\alpha'}) \tag{60}$$

with $\mathbf{R}_{\alpha\alpha'} = \mathbf{R} + \mathbf{r}_\alpha - \mathbf{r}_{\alpha'}$. The calculation of $\Xi^{\alpha\alpha'}_{\lambda\mu}(\mathbf{q})$ is carried out following Refs. [35,36].

4.3.2. Case 2: $\chi_i^{\mathbf{q}}(\mathbf{r})$ and $\chi_j^{\mathbf{q}}(\mathbf{r})$ with $\mathbf{r} \in I$

Introducing the Fourier expansion of the Coulomb interaction

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{V} \sum_{\mathbf{qG}} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}_1} \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}_2} \tag{61}$$

and the planewave representation for the basis functions, *i.e.*, $\chi_i^{\mathbf{q}}(\mathbf{r}) = P_i^{\mathbf{q}}(\mathbf{r})$ and $\chi_j^{\mathbf{q}}(\mathbf{r}) = P_j^{\mathbf{q}}(\mathbf{r})$, Eq. (54) becomes

$$\begin{aligned} v_{ij}(\mathbf{q}) &= \int_V \int_V \left[P_i^{\mathbf{q}}(\mathbf{r}_1)\right]^* \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} P_j^{\mathbf{q}}(\mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \int_V \int_V \left[P_i^{\mathbf{q}}(\mathbf{r}_1)\right]^* \frac{1}{V} \sum_{\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}_1} \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}_2} P_j^{\mathbf{q}}(\mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \sum_{\mathbf{G}} \left(\frac{1}{\sqrt{\Omega}} \int_\Omega \left[P_i^{\mathbf{q}}(\mathbf{r}_1)\right]^* e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}_1} d\mathbf{r}_1 \right) \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} \left(\frac{1}{\sqrt{\Omega}} \int_\Omega e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}_2} \tilde{P}_j^{\mathbf{q}}(\mathbf{r}_2)\, d\mathbf{r}_2 \right) \\ &= \sum_{\mathbf{G}} w^i_{\mathbf{G}}(\mathbf{q}) \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} \left[w^j_{\mathbf{G}}(\mathbf{q})\right]^*. \end{aligned} \tag{62}$$

4.3.3. Case 3: $\chi_i^{\mathbf{q}}(\mathbf{r})$ with $\mathbf{r} \in I$ and $\chi_j^{\mathbf{q}}(\mathbf{r})$ with $\mathbf{r} \in$ MT sphere

In this case, we need to consider $\chi_i^{\mathbf{q}}(\mathbf{r}) = P_i^{\mathbf{q}}(\mathbf{r})$ and $\chi_j^{\mathbf{q}}(\mathbf{r}) = \gamma^{\mathbf{q}}_{\alpha NLM}(\mathbf{r})$. Starting again from the Fourier expansion of the Coulomb potential, we obtain

$$v_{ij}(\mathbf{q}) = \frac{(4\pi)^2}{\sqrt{\Omega}} \sum_{\mathbf{G}} \frac{1}{|\mathbf{q}+\mathbf{G}|^2} \tilde{S}^*_{\mathbf{G},i}\, i^L Y_{LM}\left(\widehat{\mathbf{G}+\mathbf{q}}\right) \mathbb{J}_{\alpha NL}(|\mathbf{G}+\mathbf{q}|) \tag{63}$$

where $\tilde{S}_{\mathbf{G}i}$ is defined in Eq. (41), and $\mathbb{J}_{\alpha NL}$ is given in Eq. (45).

4.4. The dielectric function

The dielectric matrix can be constructed straightforwardly from $\tilde{M}^i_{nm}(\mathbf{k}, \mathbf{q})$ according to Eq. (25) except for $\mathbf{q} \to 0$ due to the singularity of the bare Coulomb interaction. To treat this special case properly, we start from the representation of the symmetrized dielectric function in Fourier space,

$$\varepsilon_{\mathbf{GG}'}(\mathbf{q}, \omega) = \delta_{\mathbf{GG}'} - \frac{4\pi}{|\mathbf{q}+\mathbf{G}|\,|\mathbf{q}+\mathbf{G}'|} P_{\mathbf{GG}'}(\mathbf{q}, \omega). \tag{64}$$

It can be seen that possible divergences are located in the head, ε_{00}, and the wings, $\varepsilon_{0\mathbf{G}'}$ and $\varepsilon_{\mathbf{G}0}$. Using $\mathbf{k}\cdot\mathbf{p}$ perturbation theory, the limit for $\mathbf{q} \to 0$ can be calculated analytically [37]. Denoting

$$\hat{\mathbf{q}}_0 \equiv \lim_{\mathbf{q}\to 0} \frac{\mathbf{q}}{|\mathbf{q}|} \tag{65}$$

as the direction along which the limit of $\mathbf{q} \to 0$ is taken, the final expressions are

$$\varepsilon_{00}(0,\omega) \equiv H(\omega) = 1 - \frac{4\pi}{N_c \Omega} \sum_{\mathbf{k}} \left\{ \sum_n \delta(\epsilon_F - \epsilon_{n\mathbf{k}}) \frac{\left(\mathbf{p}_{nn\mathbf{k}} \cdot \hat{\mathbf{q}}_0\right)^2}{\omega^2} + \sum_{n \neq n'} F_{nn'}(\mathbf{k}, 0; \omega) \left| \frac{\mathbf{p}_{n'n\mathbf{k}} \cdot \hat{\mathbf{q}}_0}{\epsilon_{n'\mathbf{k}} - \epsilon_{n\mathbf{k}}} \right|^2 \right\} \tag{66}$$

for the head and

$$\varepsilon_{0\mathbf{G}}(0,\omega) = \frac{1}{N_c \Omega^{1/2}} \frac{4\pi}{|\mathbf{G}|} \sum_{\mathbf{k}} \sum_{n' \neq n} F_{nn'}(\mathbf{k}, 0; \omega) \frac{\mathbf{p}^*_{nn'\mathbf{k}} \cdot \hat{\mathbf{q}}_0}{\epsilon_{n'\mathbf{k}} - \epsilon_{n\mathbf{k}}} \left[M^{\mathbf{G}}_{nn'}(\mathbf{k}, 0) \right]^* \tag{67}$$

for the wings. $\mathbf{p}_{nn'\mathbf{k}}$ are the matrix elements of the momentum operator in the basis of KS eigenvectors, $\mathbf{p}_{nn'\mathbf{k}} \equiv \langle \psi_{n\mathbf{k}} | \mathbf{p} | \psi_{n'\mathbf{k}} \rangle$.

In the mixed basis representation, the bare Coulomb interaction is not diagonal. As a result, the head and wings are no longer restricted to a particular region of the dielectric matrix, hence all components, corresponding to head, wing, and body (denoted by H, W, and B respectively) appear in general matrix form

$$\varepsilon_{ij}(0,\omega) = \varepsilon^H_{ij}(0,\omega) + \varepsilon^W_{ij}(0,\omega) + \varepsilon^B_{ij}(0,\omega). \tag{68}$$

Using the separation for $v^{\frac{1}{2}}_{ij}(\mathbf{q})$

$$v^{\frac{1}{2}}_{ij}(\mathbf{q} \to 0) = \frac{v^{s\frac{1}{2}}_{ij}}{q} + \tilde{v}^{\frac{1}{2}}_{ij} \tag{69}$$

with

$$v^{s\frac{1}{2}}_{ij} = \sqrt{4\pi}\, \mathcal{W}^i_0(0) \mathcal{W}^{j*}_0(0), \tag{70}$$

as outlined in the Appendix for $v_{ij}(\mathbf{q})$ (Eq. (A.4)), one obtains for the three terms:

$$\begin{aligned}
\varepsilon^H_{ij}(0,\omega) &= H(\omega) \mathcal{W}^i_0(0) \mathcal{W}^{j*}_0(0) \\
\varepsilon^W_{ij}(0,\omega) &= \left(-\sqrt{\frac{4\pi}{\Omega}} \right) \frac{1}{N_c} \sum_{\mathbf{k}} \sum_{n' \neq n} F_{nn'\mathbf{k}}(0,\omega) \left\{ \frac{\mathbf{p}^*_{nn'\mathbf{k}} \cdot \hat{\mathbf{q}}_0}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \mathcal{W}^i_0 \left[\tilde{M}^j_{nn'}(\mathbf{k}, 0) \right]^* + \frac{\mathbf{p}_{nn'\mathbf{k}} \cdot \hat{\mathbf{q}}_0}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \tilde{M}^i_{nn'}(\mathbf{k}, 0) \mathcal{W}^{j*}_0 \right\} \\
\varepsilon^B_{ij}(0,\omega) &= \delta_{ij} - \mathcal{W}^i_0(0) \mathcal{W}^{j*}_0(0) - \sum_{\mathbf{k}} \sum_{n,n'} F_{nn'}(\mathbf{k}, 0; \omega)\, \tilde{M}^i_{nn'}(\mathbf{k}, 0) \left[\tilde{M}^j_{nn'}(\mathbf{k}, 0) \right]^*
\end{aligned} \tag{71}$$

where we have introduced matrix elements $\tilde{M}^j_{nn'}$ that contract $\tilde{v}^{\frac{1}{2}}_{ij}$ with $M^i_{nn'}(\mathbf{k}, 0)$ in the same way as in Eq. (24), *i.e.*, $\tilde{M}^i_{nm}(\mathbf{k}, \mathbf{q}) \equiv \sum_p \tilde{v}^{\frac{1}{2}\, ip}(\mathbf{q}) M^p_{nm}(\mathbf{k}, \mathbf{q})$.

A significant improvement of the efficiency can be achieved by using the eigenvectors of the bare Coulomb matrix, v, as a secondary basis to represent the dielectric function and the subsequent evaluation of the screened Coulomb interaction, W, as described by Friedrich et al. [38]. Denoting the μ-th eigenvector of v at $\mathbf{q}$ as $\left\{ U^{\mathbf{q}}_{i\mu} \right\}$, we define the new basis, henceforth called the "v-diagonal basis", as

$$|\chi^{\mathbf{q}}_\mu\rangle = \sum_i |\chi^{\mathbf{q}}_i\rangle U^{\mathbf{q}}_{i\mu}. \tag{72}$$

For $\mathbf{q} \neq 0$, the new basis is just a unitary transformation of the original mixed basis, and therefore the two basis sets are mathematically equivalent. For $\mathbf{q} = 0$, a special approach is needed for the treatment of the singularity. As shown in Appendix A, the bare Coulomb matrix at $\mathbf{q} = 0$ can be decomposed into a singular term and a regular term, both being a full matrix in the mixed basis representation. In this case, the natural choice is to use the eigenvectors of the regular part, $\left\{ \tilde{v}_{ij} \right\}$, as the new basis. The singular part can be obtained by a unitary transform

$$v^s_{\mu\nu} = \sum_{ij} U^{0*}_{i\mu} v^s_{ij} U^0_{j\nu}. \tag{73}$$

On the other hand, considering that the singularity in the planewave representation occurs only at $\mathbf{G} = 0$, $v^s_{\mathbf{G}\mathbf{G}'} = 4\pi \delta_{\mathbf{G}\mathbf{G}'} \delta_{\mathbf{G}0}$, we replace $v^s_{\mu\nu}$ by its exact limit, $v^s_{\mu\nu} = 4\pi \delta_{\mu 0} \delta_{\nu 0}$. In this new basis, the dielectric function is represented as

$$\varepsilon(\omega) = \begin{pmatrix} H(\omega) & \mathbf{w}^\dagger(\omega) \\ \mathbf{w}(\omega) & \mathbf{B}(\omega) \end{pmatrix} \tag{74}$$

where $H(\omega)$ is defined in Eq. (66), and

$$\begin{aligned}
w_\mu(\omega) &= \left(-\sqrt{\frac{4\pi}{\Omega}} \right) N_c^{-1} \sum_{\mathbf{k}}^{BZ} \sum_{n' \neq n} F_{nn'\mathbf{k}}(0,\omega) \frac{\mathbf{p}_{nn'\mathbf{k}} \cdot \hat{\mathbf{q}}_0}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \tilde{M}^\mu_{nn'}(\mathbf{k}, 0) \\
B_{\mu\nu}(\omega) &= \delta_{\mu\nu} - N_c^{-1} \sum_{\mathbf{k}}^{BZ} \sum_{n,n'} F_{nn'\mathbf{k}}(0,\omega) \tilde{M}^\mu_{nn'}(\mathbf{k}, 0) \left[\tilde{M}^\nu_{nn'}(\mathbf{k}, 0) \right]^*
\end{aligned} \tag{75}$$

with $\mu, \nu \neq 0$. The corresponding inverse dielectric matrix now reads (the frequency is dropped to simplify the notation)

$$\begin{aligned}
\varepsilon_{00}^{-1} &= \left[H - \sum_{\mu\nu} w_\mu^\dagger B_{\mu\nu}^{-1} w_\nu \right]^{-1} \\
\varepsilon_{\mu 0}^{-1} &= -\left(\sum_\nu B_{\mu\nu}^{-1} w_\nu \right) \varepsilon_{00}^{-1} \\
\varepsilon_{\mu\nu}^{-1} &= B_{\mu\nu}^{-1} + \varepsilon_{\mu 0}^{-1} \left(\varepsilon_{00}^{-1}\right)^{-1} \varepsilon_{0\nu}^{-1}.
\end{aligned} \tag{76}$$

There are several significant advantages in using the v-diagonal basis. First of all, the matrix multiplication involving the bare Coulomb matrix, which is diagonal in the new basis, is simplified. Second, the size of the basis can be reduced by keeping only the eigenvectors of v with a corresponding eigenvalue larger than a certain criterion. Most important, the treatment of the Γ point singularity in the v-diagonal basis is significantly simpler. Since the contribution related to the singularity at $\mathbf{q} = 0$ is now confined only to ε_{00} (the head), $\varepsilon_{\mu 0}$ and $\varepsilon_{0\nu}$ (the wings), the dielectric anisotropy, *i.e.*, the fact that the dielectric function at the limit of $\mathbf{q} \to 0$ depends on the direction, $\hat{\mathbf{q}}_0$, along which the limit is taken, is treated in a more rigorous way using the technique developed by Freysoldt et al. [39].

4.5. The self-energy

In this section, we discuss the evaluation of the diagonal elements of the self-energy (Eqs. (18) and (22)). There are mainly two technical issues that require careful treatment: the Brillouin zone integration, in particular, the treatment of singular term, and the frequency dependence.

We start with the Brillouin zone integration. The diagonal elements of the exchange and correlation self-energies can be generally evaluated as

$$\Sigma_{n\mathbf{k}}^{\mathrm{x/c}} = N_c^{-1} \sum_{\mathbf{q}}^{BZ} \Upsilon_{n\mathbf{k}}^{\mathrm{x/c}}(\mathbf{q}). \tag{77}$$

In the v-diagonal basis, we have

$$\begin{aligned}
\Upsilon_{n\mathbf{k}}^{x}(\mathbf{q}) &= -\sum_\mu \sum_m f_{m\mathbf{k}-\mathbf{q}} \left[\tilde{M}_{nm}^{\mu}(\mathbf{k}, \mathbf{q}) \right]^* \tilde{M}_{nm}^{\mu}(\mathbf{k}, \mathbf{q}) \\
\Upsilon_{n\mathbf{k}}^{c}(\mathbf{q}, \omega) &= \sum_m \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{X_{nm}(\mathbf{k}, \mathbf{q}, \omega')}{\omega + \omega' - \tilde{\epsilon}_{m\mathbf{k}-\mathbf{q}}}
\end{aligned} \tag{78}$$

with

$$X_{nm}(\mathbf{k}, \mathbf{q}, \omega') = \sum_{\mu\nu} \left[\tilde{M}_{nm}^{\mu}(\mathbf{k}, \mathbf{q}) \right]^* \left(\varepsilon_{\mu\nu}^{-1} - \delta_{\mu\nu} \right) \tilde{M}_{nm}^{\nu}(\mathbf{k}, \mathbf{q}). \tag{79}$$

In the $\mathbf{q} \to 0$ limit, the singular terms in $\Upsilon_{n\mathbf{k}}^{\mathrm{x/c}}(\mathbf{q})$ can be separated as follows

$$\Upsilon_{n\mathbf{k}}^{\mathrm{x/c}}(\mathbf{q} \to 0) = \frac{\Upsilon_{n\mathbf{k}}^{\mathrm{x/cs2}}}{q^2} + \frac{\Upsilon_{n\mathbf{k}}^{\mathrm{x/cs1}}}{q} + \tilde{\Upsilon}_{n\mathbf{k}}^{\mathrm{x/c}}. \tag{80}$$

The singularity can then be integrated using the standard analytical technique as described in Appendix B in more detail. By using

$$\begin{aligned}
M_{nm}^{\mu=0}(\mathbf{k}, \mathbf{q} = 0) &= \int_\Omega d\mathbf{r} \left[\chi_{\mu=0}^{\mathbf{q}=0}(\mathbf{r}) \right]^* \psi_{n\mathbf{k}} \left[\psi_{m\mathbf{k}} \right]^* \\
&= \Omega^{-1/2} \int_\Omega d\mathbf{r} \psi_{n\mathbf{k}}(\mathbf{r}) \left[\psi_{m\mathbf{k}}(\mathbf{r}) \right]^* = \Omega^{-1/2} \delta_{nm},
\end{aligned} \tag{81}$$

we obtain the explicit expressions for $\Upsilon_{n\mathbf{k}}^{\mathrm{x/cs1/2}}$ within the v-diagonal basis

$$\begin{aligned}
\Upsilon_{n\mathbf{k}}^{\mathrm{xs2}} &= \frac{4\pi}{\Omega} f_{n\mathbf{k}} \\
\Upsilon_{n\mathbf{k}}^{\mathrm{xs1}} &= 0 \\
\Upsilon_{n\mathbf{k}}^{\mathrm{cs1/2}} &= \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{X_{n\mathbf{k}}^{s1/2}(\omega')}{\omega + \omega' - \tilde{\epsilon}_{m\mathbf{k}-\mathbf{q}}},
\end{aligned} \tag{82}$$

with

$$\begin{aligned}
X_{n\mathbf{k}}^{s2}(\omega) &= \frac{4\pi}{\Omega} \left[\varepsilon_{00}^{-1}(0, \omega) - 1 \right] \\
X_{n\mathbf{k}}^{s1}(\omega) &= \sqrt{\frac{4\pi}{\Omega}} \sum_{\mu \neq 0} \left\{ \varepsilon_{0\mu}^{-1}(0, \omega) \tilde{M}_{nn}^{\mu}(\mathbf{k}, 0) + \varepsilon_{\mu 0}^{-1}(0, \omega) \left[\tilde{M}_{nn}^{\mu}(\mathbf{k}, 0) \right]^* \right\}.
\end{aligned} \tag{83}$$

The correlation self-energy involves a convolution integral over frequency (Eq. (22)). Due to the poles of both the Green function and W^c, infinitesimally close to the real axis, this integral is difficult to converge numerically, *i.e.*, it requires a large number of frequencies.

In our implementation, we compute the polarizability, the screened Coulomb potential, and the self-energy on the imaginary frequency axis, $\omega = iu$. Making use of the inversion symmetry of W^c on the imaginary frequency axis, $W^c_{ij}(\mathbf{q}, iu) = W^c_{ij}(\mathbf{q}, -iu)$, we obtain

$$\Sigma^c_{n\mathbf{k}}(iu) = \frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\sum_m \int_0^\infty \frac{du'}{2\pi} X_{nm}(\mathbf{k},\mathbf{q};iu')\frac{2\left(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu\right)}{u'^2+\left(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu\right)^2}. \tag{84}$$

When $\epsilon_{m\mathbf{k}-\mathbf{q}}$ is small, the integrand in Eq. (84) is peaked around $u' = u$, and therefore a direct numerical integration becomes unstable. This problem can be circumvented by adding and subtracting the term [40]

$$\int_0^\infty \frac{du'}{2\pi} X_{nm}(\mathbf{k},\mathbf{q};iu)\frac{(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu)}{u'^2+\left(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu\right)^2} = \frac{1}{2}\mathrm{sgn}(\epsilon_{m\mathbf{k}-\mathbf{q}})\, X_{nm}(\mathbf{k},\mathbf{q};iu) \tag{85}$$

such that we arrive at

$$\Sigma^c_{n\mathbf{k}}(iu) = \frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\sum_m \left\{\int_0^\infty \frac{du'}{2\pi}\left[X_{nm}(\mathbf{k},\mathbf{q};iu') - X_{nm}(\mathbf{k},\mathbf{q};iu)\right]\frac{2\left(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu\right)}{u'^2+\left(\epsilon_{m\mathbf{k}-\mathbf{q}}-iu\right)^2} + \frac{1}{2}\mathrm{sgn}(\epsilon_{m\mathbf{k}-\mathbf{q}})\, X_{nm}(\mathbf{k},\mathbf{q};iu)\right\}. \tag{86}$$

The integrand is now a smooth function for any $\epsilon_{m\mathbf{k}-\mathbf{q}}$ and therefore a standard Gaussian quadrature may be used. In practice, we apply a double Gauss–Legendre quadrature [10,41], in which the semi-infinite integral is divided into two intervals, $[0, \omega_0]$ and $[\omega_0, \infty)$, and the integration in each interval is carried out by standard Gauss–Legendre quadrature. As shown in the next section (see also Ref. [41]), the integration over frequency can be converged with a rather small number of points when choosing appropriate ω_0.

After the matrix elements of the correlation self-energy along the imaginary axis are calculated according to Eq. (86), they are fitted by a function with N_p poles [42]

$$\Sigma^c_{n\mathbf{k}}(iu) = \sum_p^{N_p}\frac{a_{p;n\mathbf{k}}}{iu - b_{p;n\mathbf{k}}} \tag{87}$$

where $a_{p;n\mathbf{k}}$ and $b_{p;n\mathbf{k}}$ are fitting parameters. Eq. (87) is then analytically continued onto the real frequency axis. We call this approach the multi-pole fitting scheme.

Another way to obtain the correlation self-energies on the real frequency is by the Pade-approximant method [43,44] (see Appendix C), which is also implemented in our code. In general, the two approaches (with $N_p = 2$ in the multi-pole fitting scheme) give very similar results for *sp*-semiconductors and insulators. On the other hand, the Pade approach may in some cases lead to unphysical QP renormalization factors (Eq. (29)) for states far away from the Fermi energy.

4.6. G_0W_0 with input from LDA + U

In this section, we provide some technical details about the implementation of G_0W_0 based on the LDA + U [45] single-particle Hamiltonian. Here LDA + U is used as a synonym to refer to the Hubbard U corrected LDA or GGA in general. In the LDA + U method, the total energy is not only an explicit functional of Kohn–Sham orbitals but also of the local on-site density matrix $n^{\alpha l}_{mm'}$ for the orbital l of atom α (usually the d orbital of transition metal atoms and the f orbital of lanthanides or actinides):

$$E^{\mathrm{LDA}+U}_{\mathrm{tot}} = E^{\mathrm{LDA}}_{\mathrm{tot}}[\rho(\mathbf{r})] + \Delta E^U[n^{\alpha l}_{mm'}]. \tag{88}$$

The corresponding KS single-particle equation becomes:

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + V^{\mathrm{LDA}}_{\mathrm{xc}}(\mathbf{r})\right]\psi_{n\mathbf{k}}(\mathbf{r}) + \sum_{m,m'} v^{\alpha l}_{mm'}\frac{1}{f_{n\mathbf{k}}}\frac{\delta n^{\alpha l}_{mm'}}{\delta\psi^*_{n\mathbf{k}}(\mathbf{r})} = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}) \tag{89}$$

with $v^{\alpha l}_{mm'} \equiv \delta\Delta E^U/\delta n^{\alpha l}_{mm'}$. Note that the spin index has been dropped for brevity. Using Eq. (48), the local density matrix corresponding to the l-th orbital of atom α can be written as [46]

$$\begin{aligned} n^{\alpha l}_{mm'} &= \sum_{n\mathbf{k}} f_{n\mathbf{k}}\sum_{\zeta,\zeta'}\mathcal{A}_{n\mathbf{k},\alpha\zeta lm} I^{\alpha l}_{\zeta,\zeta'}\left[\mathcal{A}_{n\mathbf{k},\alpha\zeta' lm'}\right]^* \\ &= \sum_{n\mathbf{k}} f_{n\mathbf{k}}\sum_{\zeta,\zeta'}\left\langle\psi_{n\mathbf{k}}|u_{\alpha\zeta l}Y_{lm'}\right\rangle\left[I^{\alpha l}\right]^{-1}_{\zeta\zeta'}\left\langle u_{\alpha l\zeta'}Y_{lm}|\psi_{n\mathbf{k}}\right\rangle \end{aligned} \tag{90}$$

with

$$I^{al}_{\zeta\zeta'} \equiv \int_0^\infty dr\, r^2\, u_{\alpha\zeta l}(r)\, u_{\alpha\zeta' l}(r). \tag{91}$$

The KS equation in the LDA + U framework then reads

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + V^{\mathrm{LDA}}_{\mathrm{xc}}(\mathbf{r}) + \delta\hat{V}^U\right]\psi_{n\mathbf{k}}(\mathbf{r}) = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}) \tag{92}$$

where

$$\delta\hat{V}^U = \sum_{m,m'} v^{\alpha l}_{mm'} \sum_{\zeta\zeta'} |u_{\alpha\zeta l} Y_{lm'}\rangle \left[I^{\alpha l}\right]^{-1}_{\zeta\zeta'} \langle u_{\alpha l\zeta'} Y_{lm}|. \tag{93}$$

Formally, the only difference between LDA-based and LDA + U-based G_0W_0 is the contribution from Eq. (93),

$$\begin{aligned}\delta V^U_{n\mathbf{k}} &\equiv \langle\psi_{n\mathbf{k}}|\,\delta\hat{V}^U\,|\psi_{n\mathbf{k}}\rangle \\ &= \sum_{m,m'} v^{\alpha l}_{mm'} \sum_{\zeta,\zeta'} \left\langle\psi_{n\mathbf{k}}|u_{\alpha l\zeta} Y_{lm'}\right\rangle \left[I^{\alpha l}\right]^{-1}_{\zeta\zeta'} \left\langle u_{\alpha l\zeta'} Y_{lm}|\psi_{n\mathbf{k}}\right\rangle \\ &= \sum_{m,m'} v^{\alpha l}_{mm'} \sum_{\zeta,\zeta'} \mathcal{A}_{n\mathbf{k},\alpha\zeta lm} \left[I^{\alpha l}\right]_{\zeta\zeta'} \left[\mathcal{A}_{n\mathbf{k},\alpha\zeta' lm'}\right]^* .\end{aligned} \tag{94}$$

Therefore, the G_0W_0 quasi-particle energy based on LDA + U reads:

$$\epsilon^{\mathrm{qp}}_{n\mathbf{k}} = \epsilon_{n\mathbf{k}} + \Re\left[\langle\psi_{n\mathbf{k}}|\,\Sigma(\epsilon^{\mathrm{qp}}_{n\mathbf{k}}) - V^{\mathrm{LDA}}_{\mathrm{xc}} - \delta V^U\,|\psi_{n\mathbf{k}}\rangle\right]. \tag{95}$$

5. Convergence tests

The advantage of the LAPW method is that it does not rely on any shape approximation for the potential and density, and therefore for ground state calculations it is numerically "exact" as long as the *convergence parameters* are well controlled. Concerning the evaluation of the self-energy, the numerical accuracy within the LAPW framework is certainly more problematic. It depends on:

1. the quality of the mixed basis set,
2. the size of the LAPW basis set, as determined by the parameter $RG_{\max}$,
3. the number of unoccupied states included in the calculation of the correlation term, as determined by the energy cutoff $\epsilon_{\max}$
4. the number of **k**- and **q**-points used for the Brillouin-zone integration, and
5. the number of frequencies, N_ω, in the calculation of the self-energy.

The convergence with respect to these parameters is essential for the reliability of the obtained results. In addition, the precision of the Kohn–Sham results used as input also requires attention, since they could indirectly affect the G_0W_0 output. In particular, the accuracy of the LAPW basis to represent the Kohn–Sham orbitals is crucial. It is mainly controlled by a quantity, usually termed '$RK_{\max}$' or '$RG_{\max}$', which is the product of the (smallest) MT radius, min R^α_{MT}, and the largest G vector of the interstitial planewaves, $G_{\max}$. It not only determines the quality of the ground-state, but also that of the unoccupied orbitals. On the other hand, the capability of the LAPW basis to represent Kohn–Sham wavefunctions within the MT spheres is limited by its linearization error. The latter is negligible for occupied states, but can become severe for unoccupied states when the orbitals have more nodes inside the MT sphere. How strongly G_0W_0 results are affected by the linearization error of the LAPW basis is still unclear. In this work, we use the "standard" LAPW basis.

In the following, we discuss the convergence of G_0W_0 results with respect to the various parameters mentioned above using Si as the example. The main quantity to monitor this convergence is the band gap between Γ and X, which is 1.13 eV with an estimated error of $\lesssim$0.02 eV when using the following parameters: $Q \equiv G^{\mathrm{MB}}_{\max}/G_{\max} = 1.0$, $l^{\mathrm{MB}}_{\max} = 3$, $RG_{\max} = 8.0$, $\epsilon_{\max} = 204$ eV, $\omega_0 = 13.6$ eV, $N_\omega = 16$, and $N_k = 6^3$. Unless otherwise stated, the results discussed in this section are obtained using these parameters, except the parameter under consideration which is varied.

5.1. Quality of the mixed basis set

The quality of the mixed basis, $\chi^{\mathbf{q}}_j$, is controlled by two parameters as described in Section 4.1: in the interstitial region, it is determined by the cut-off for the interstitial planewaves, $G^{\mathrm{MB}}_{\max}$. It can be characterized by the quantity $Q \equiv G^{\mathrm{MB}}_{\max}/G_{\max}$, where $G_{\max}$ is the corresponding cut-off of the LAPW basis functions. Within the muffin-tin spheres, the decisive parameter is $l^{\mathrm{MB}}_{\max}$, the maximal angular-momentum quantum number of the radial functions used to construct the mixed-basis. Fig. 2 shows the convergence of the Si $\Gamma - X$ band gap with respect to $l^{\mathrm{MB}}_{\max}$ and Q, respectively. For the former (left panel), Q is fixed to 1.0 and for the latter (right panel) $l^{\mathrm{MB}}_{\max} = 3$ is used.

In Si, $l^{\mathrm{MB}}_{\max} = 2$ is enough to ensure an accuracy of ∼0.02 eV. With respect to Q, one would expect that $Q \sim 2.0$ is needed to obtain converged G_0W_0 results, since the product of two Kohn–Sham wavefunctions contains the contributions of interstitial planewaves with G values up to $2G_{\max}$. In reality, however, a much smaller Q can be used. In fact, already $Q = 0.8$ gives reasonably converged band gaps with an accuracy of ∼0.01 eV. Note, however, that such findings should not be generalized, and convergence tests need to be performed for every material.

5.2. LAPW basis set size

As mentioned above, we will not address the possible error due to the linearization of the basis functions inside the MT spheres. Beyond this issue, $RG_{\max}$ plays a two-fold role: it determines the size of the LAPW basis and, therefore, also that of the mixed basis. It also determines the maximum number of unoccupied states that can be taken into account in the summation of Eqs. (19) and (22) (see below). To check the influence of the former, Fig. 3 shows the convergence of the $\Gamma - X$ gap with respect to $RG_{\max}$ with a fixed energy cutoff for the unoccupied states $\epsilon_{\max} = 82$ eV. We can see that the value converges quite quickly with respect to $RG_{\max}$. Using $RG_{\max} = 7.0$ ensures an accuracy of $\sim$ 0.01 eV.

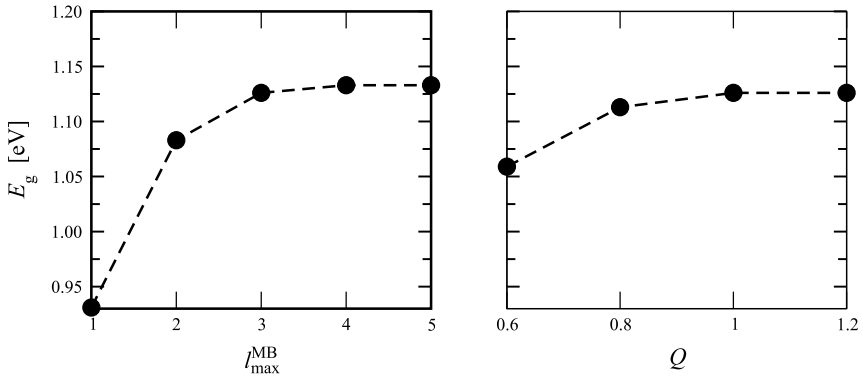

Fig. 2. Convergence behavior of the $\Gamma - X$ band gap of Si with respect to l_{max}^{MB} (left), the parameter determining the quality of the mixed basis inside the muffin-tin spheres, and Q (right), the parameter determining the quality of the mixed basis in the interstitial region.

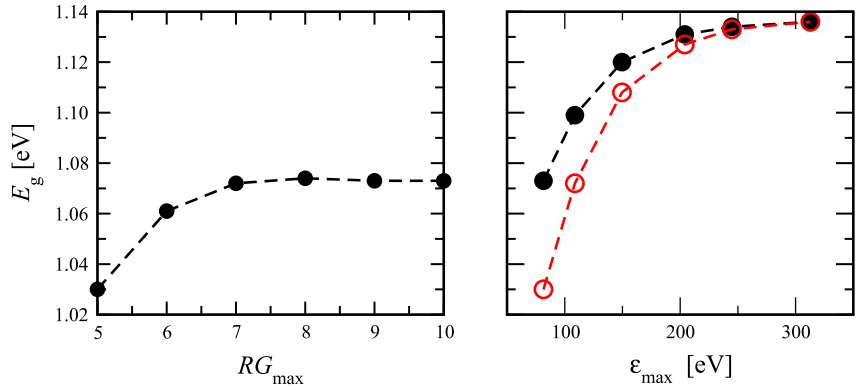

Fig. 3. Left: Convergence of the $\Gamma - X$ band gap for Si with respect to RG_{max}. The values are computed with a fixed $\epsilon_{max} = 82$ eV (an energy cutoff corresponding to $RG_{max} = 5.0$) for different RG_{max} and thus are not fully converged with respect to ϵ_{max}. Right: Convergence with respect to ϵ_{max}. Filled circles represent data obtained with a fixed RG_{max} of 10.0, while open symbols represent results obtained with varying RG_{max} and using all available unoccupied states, *i.e.*, setting $\epsilon_{max} = \epsilon_{max}^{PW}$.

5.3. Number of unoccupied states

We now study the convergence of the G_0W_0 band gap with respect to the number of unoccupied states involved in the G_0W_0 calculation. The number of available unoccupied states, ϵ_{max}^{PW}, is determined by the planewave cutoff of the LAPW basis functions. To check the convergence with respect to the number of unoccupied states, we choose a large $RG_{max} = 10.0$ which allows to access a large number of empty states. As shown in the left panel of Fig. 3, the G_0W_0 $\Gamma - X$ band gap of Si increases as a function of ϵ_{max}, and becomes saturated when $\epsilon_{max} >\sim 200$ eV. For comparison, we also show the results for the case in which RG_{max} is varied, and all available unoccupied states are used ($\epsilon_{max} = \epsilon_{max}^{PW}$). The two sets of data differ significantly in the regime of small cutoffs, but they naturally converge to the same value as ϵ_{max} increases.

5.4. Frequency integration

As described in Section 4.5, the semi-infinite integral is performed along the imaginary frequency axis (Eq. (86)) by a double Gauss–Legendre quadrature scheme, in which the integration region $[0, \infty)$ is split into two intervals separated by ω_0. The numerical accuracy is mainly determined by the total number of discrete imaginary frequency points, N_ω, but the parameter ω_0 can also have significant influence on the accuracy.

Fig. 4 shows the convergence of the $\Gamma - X$ band gaps of Si with respect to N_ω obtained with different ω_0 using the 2-pole fitting scheme for the analytic continuation. We consider three choices of ω_0, namely 1.36, 13.6, and 27.2 eV, which covers a broad energy range and allows for evaluating the sensitivity of the results. The convergence of the band gap with respect to N_ω is poor when $\omega_0 = 1.36$ eV. Choosing $\omega_0 = 13.6$ or 27.2 eV, the results obtained with $N_\omega = 16$ is already quite close to the fully converged value. In addition, we note that the band gaps converged with respect to N_ω for different ω_0 are nearly identical except that the one with $\omega_0 = 1.36$ eV differs by about 0.01 eV from the others.

5.5. Number of **k**-points

Finally, we test the convergence of the $\Gamma - X$ gap with respect to the number of **k** points, N_k, used for BZ integrations. Considering that G_0W_0 calculations scale quadratically with N_k, it is highly desirable to use an as small as possible **k**-point grid. Generally, the number of **k**-points needed for a converged LDA calculation may be different from that required for a G_0W_0 calculation. Since LDA requires much less computational effort than G_0W_0, it is a common practice to run the G_0W_0 calculation using a smaller number of **k**-points which is based on a well-converged LDA calculation, performed on a fine **k**-mesh. This LDA calculation not only provides Kohn–Sham single-particle energies and wave-functions, but also the exchange–correlation potential, V^{xc}. We find that, in some materials, the G_0W_0 gap converges faster with

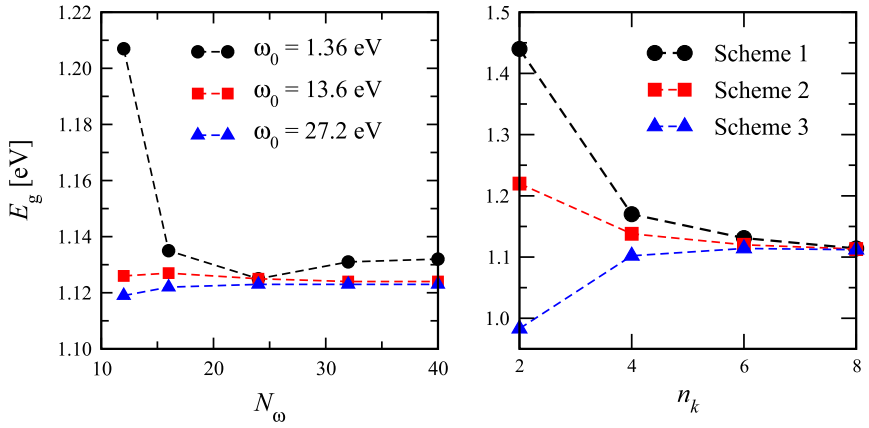

Fig. 4. Left: Convergence behavior of the $\Gamma - X$ gap of Si with respect to the number of frequency points, N_ω. Right: Convergence with respect to $n_k \equiv N_k^{1/3}$ using different schemes (see the text in Section 5.5).

respect to N_k when V^{xc} is recalculated with the same **k**-mesh as used for G_0W_0 calculations. This may be likely due to error cancellation between V^{xc} and Σ^{xc}, but it is not clear whether this is a general feature of G_0W_0 calculations.

In the right panel of Fig. 4, we systematically investigate how the $\Gamma - X$ gap behaves as a function of N_k by comparing the following three schemes.

- Scheme 1: LDA calculations are performed on a fine $10 \times 10 \times 10$ **k**-mesh, and V^{xc} obtained with this fine **k**-mesh is directly used in the G_0W_0 calculation.
- Scheme 2: LDA calculations are carried out on a fine $10 \times 10 \times 10$ **k**-mesh, but V^{xc} is recalculated with the same **k**-mesh as used for the G_0W_0 calculation.
- Scheme 3: The same **k**-mesh is used for LDA and G_0W_0 calculations.

Comparing Schemes 1–3, we see that using the same **k**-grid for Σ^{xc} and V^{xc} tends to speed up the convergence with respect to N_k. The $\Gamma - X$ gap calculated with a $6 \times 6 \times 6$ mesh differs from that obtained with an $8 \times 8 \times 8$ mesh by less than 0.01 eV in Schemes 2 and 3. The difference is slightly larger in Scheme 1 (∼0.02 eV).

6. Concluding remarks

In this work we have presented the implementation of an all-electron G_0W_0 code, `FHI-gap`, which is based on the LAPW method. `FHI-gap` also allows for calculations in combination with the LDA + U approach. This way, it has been used for exploring d and f electron systems such as transition metal and lanthanide oxides [20,21].

One issue we did not address in this paper are the linearization errors of the LAPW basis and their relevance for G_0W_0 calculations. In the standard LAPW basis, the linearization energies are chosen in the respective valence band regime, and local orbitals are introduced to describe semicore states or to improve the flexibility of the basis in case of large band widths. This procedure has proven to be very accurate for ground-state calculations. In contrast, *GW* calculations require the knowledge of all states including highly-lying unoccupied states, for which the standard LAPW basis is not very accurate. This issue was first recognized by Friedrich et al. in 2006 [47], who, however, found that for systems like Si the linearization error for G_0W_0 calculations accounts for only a small change of the band gap (∼0.03 eV) and is therefore negligible. The situation may, however, be more severe for other materials, as recently found for ZnO [48], where the addition of 292 and 186 LO's at high energies for Zn and O, respectively, increased the G_0W_0 band gap by nearly 0.5 eV. Such calculations cannot be carried out with `FHI-gap` at present, as only one LO per l quantum number can be used in the WIEN2k package. In any case, a systematic approach to reliably describe unoccupied states up to very high energies in a controllable way is still lacking, and at the same time the summation over these bands is computationally extremely expensive. In this context, recently proposed numerical techniques that eliminate unoccupied states in *GW* calculations appear attractive [49,50]. Such approaches, however, so far rely on the use of the pseudopotential planewave method. Similar techniques for the all-electron full-potential LAPW framework need to be developed.

Finally, we would like to give an outlook on envisaged extensions of the code. Important developments in the field include (i) more efficient algorithms and techniques [51,49] that give hope that the G_0W_0 method will be feasible for real materials; (ii) the so-called quasi-particle self-consistent *GW* schemes (QSGW) [52–55] that can get rid of the starting-point dependence; (iii) the proposal of combining the *GW* method with dynamical mean-field theory (DMFT) to treat strong correlations [56,57]. Some of these new features, including different variants of the QSGW approach, are currently implemented as well as the incorporation of spin–orbit coupling, and the application of the constrained random-phase approximation [58].

The `FHI-gap` code will be made available to the WIEN2k users free of charge and can be downloaded from the WIEN2k website [59]. Further information can be obtained from the webpage of the code [60].

Acknowledgments

This work was partly funded by the European Union through the NANOQUANTA network, contract no. NMP4-CT-2004-500198, and the European Theoretical Spectroscopy Facility, e-Infrastructure grant no. 211956. Financial support from the Austrian Science Fund (projects no. P16227 and S9714) and the National Natural Science Foundation of China (project no. 20973009) is also acknowledged.

Appendix A. The singularity of the bare Coulomb interaction at the Γ point

The bare Coulomb interaction $v(\mathbf{r}, \mathbf{r}') \equiv 1/|\mathbf{r}-\mathbf{r}'|$ is singular in the reciprocal space at the Γ point ($\mathbf{q} \to 0$). This singularity is integrable but requires special treatment. The divergence of the Coulomb matrix $v_{ij}(\mathbf{q})$ as $\mathbf{q} \to 0$ can be easily seen by taking a planewave expansion:

$$v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2}\delta_{\mathbf{G}\mathbf{G}'}. \tag{A.1}$$

Evidently, $v_{00} \to \infty$ as $\mathbf{q} \to 0$, since the limit corresponds to the potential generated by a constant, finite charge density, infinitely extended in space. The advantage of the planewave expansion is that one can clearly separate the divergent terms by writing

$$v_{\mathbf{G}\mathbf{G}'}(\mathbf{q} \to 0) = \frac{v^s_{\mathbf{G}\mathbf{G}'}}{q^2} + \tilde{v}_{\mathbf{G}\mathbf{G}'} \tag{A.2}$$

with

$$\begin{aligned} v^s_{\mathbf{G}\mathbf{G}'} &= 4\pi\,\delta_{\mathbf{G}\mathbf{G}'}\delta_{\mathbf{G}0} \\ \tilde{v}_{\mathbf{G}\mathbf{G}'} &= \frac{4\pi}{G^2}\delta_{\mathbf{G}\mathbf{G}'}(1-\delta_{\mathbf{G}0}). \end{aligned} \tag{A.3}$$

The same separation can be used when expanding the bare Coulomb potential in the mixed basis,

$$v_{ij}(\mathbf{q} \to 0) = \frac{v^s_{ij}}{q^2} + \tilde{v}_{ij}. \tag{A.4}$$

The singular term can be obtained from Eq. (A.3) transforming to the mixed basis:

$$v^s_{ij} = 4\pi\ \mathcal{W}^i_0(0)\ \mathcal{W}^{j*}_0(0). \tag{A.5}$$

The calculation of the regular term $\tilde{v}_{ij}$ is trivial in cases when at least one of the basis functions stems from the interstitial region (Eq. (62) or Eq. (63)). We just need to remove the $\mathbf{G} = 0$ component in the corresponding summation. In the case that both χ_i and χ_j live in the MT region, *i.e.*, $\chi_i = \gamma_{\alpha NLM}$ and $\chi_j = \gamma_{\alpha' N'L'M'}$, the lattice sum appearing in Eq. (55) diverges for $\mathbf{q} = 0$ (see e.g. Eq. (60)). Since the singularity only applies to the case $L = L' = 0$, we can carry out a planewave expansion for these terms, removing the contribution from $\mathbf{G} = 0$. That leads for $L = L' = 0$ to

$$\tilde{v}_{ij} = \sum_{\mathbf{G}\neq 0} \mathcal{W}^i_{\mathbf{G}}(0)\frac{4\pi}{|\mathbf{G}|^2}\left[\mathcal{W}^j_{\mathbf{G}}(0)\right]^*. \tag{A.6}$$

Appendix B. Brillouin zone integration for the self-energy

Calculating the matrix elements of the self-energy (Eqs. (18) and (22)) requires an integration in $\mathbf{q}$ over the Brillouin zone. To simplify the notation, we consider a general function

$$\Sigma = \frac{1}{N_c}\sum_{\mathbf{q}}^{BZ}\Upsilon(\mathbf{q}) = \frac{\Omega}{(2\pi)^3}\int_{\mathrm{BZ}} d\mathbf{q}\,\Upsilon(\mathbf{q}) \tag{B.1}$$

where the integrand $\Upsilon(\mathbf{q})$ in the $\mathbf{q} \to 0$ limit can be decomposed as

$$\Upsilon(\mathbf{q} \to 0) = \frac{\Upsilon^{s2}}{q^2} + \frac{\Upsilon^{s1}}{q} + \tilde{\Upsilon}(q) \tag{B.2}$$

with $\tilde{\Upsilon}(q)$ being the regularized form of $\Upsilon(\mathbf{q})$.

The singularity at the Γ point appearing in Eq. (B.2) is integrable, but a direct numerical integration will converge very slowly. Following [61], we introduce two auxiliary functions that show similar singularities:

$$\begin{aligned} F_1(\mathbf{q}) &= \sum_{\mathbf{G}}\frac{e^{-\beta|\mathbf{q}+\mathbf{G}|}}{|\mathbf{q}+\mathbf{G}|} \\ F_2(\mathbf{q}) &= \sum_{\mathbf{G}}\frac{e^{-\beta|\mathbf{q}+\mathbf{G}|^2}}{|\mathbf{q}+\mathbf{G}|^2}. \end{aligned} \tag{B.3}$$

One can see that

$$\begin{aligned} F_1(\mathbf{q} \to 0) &= \frac{1}{q} + \tilde{F}_1(\mathbf{q}) \\ F_2(\mathbf{q} \to 0) &= \frac{1}{q^2} + \tilde{F}_2(\mathbf{q}) \end{aligned} \tag{B.4}$$

where $\tilde{F}_1(\mathbf{q})$ and $\tilde{F}_2(\mathbf{q})$ are regularized forms of $F_1(\mathbf{q})$ and $F_2(\mathbf{q})$, respectively, obtained by neglecting $\mathbf{G} = 0$ in the summation on the right-hand side. After some algebra, the integration of $\Upsilon(\mathbf{q})$ over $\mathbf{q}$ can be written as

$$\Sigma = \frac{1}{N_c}\sum_{\mathbf{q}} \Upsilon(\mathbf{q}) = C_{s1}\Upsilon^{s1} + C_{s2}\Upsilon^{s2} + N_c^{-1}\sum_{\mathbf{q}} \tilde{\Upsilon}(\mathbf{q}) \tag{B.5}$$

with

$$\begin{aligned} C_{s1} &= \frac{\Omega}{(2\pi)^2\beta} - N_c^{-1}\sum_{\mathbf{q}} \tilde{F}_1(\mathbf{q}) \\ C_{s2} &= \frac{\Omega}{(2\pi)^2}\sqrt{\frac{\pi}{\beta}} - N_c^{-1}\sum_{\mathbf{q}} \tilde{F}_2(\mathbf{q}). \end{aligned} \tag{B.6}$$

β is a parameter chosen such that the width of the Gaussian is comparable to the Brillouin zone diameter R_{BZ}. In this work, we adopt the parameter $\beta = \left(\Omega/6\pi^2\right)^{1/3}$ which is obtained by requiring $\beta R_{BZ} = 1$ with $\frac{4\pi}{3}R_{BZ}^3 = (2\pi)^3/\Omega$. The expressions of Υ^{s1} and Υ^{s2} for the exchange and correlation self-energies can be easily obtained from Eqs. (18), (21), (22) and (A.4).

Since $\tilde{\Upsilon}(\mathbf{q})$ in Eq. (B.5) is a well-behaved function of $\mathbf{q}$ in the whole BZ, the integration over $\mathbf{q}$ can be carried out by standard techniques like the special-$\mathbf{k}$-points method [62] or the tetrahedron method [63,64]. The latter is used in our code [19].

Appendix C. The Pade approximant method for analytic continuation

A function $f(z)$, whose values at N discrete points $\{z_n | n = 1 \ldots N\}$ are given $f(z_n) = f_n$, can be fitted to an N-point Pade approximant

$$P_N(z) = \frac{A_N(z)}{B_N(z)}. \tag{C.1}$$

The Pade approximant has the feature that it is the "best" approximation of a function by a rational function of given order in the sense that $P_N(z)$ agrees with $f(z)$ at $z = 0$ up to the N-th order derivative

$$\begin{aligned} &f(0) = P_N(0) \\ &f'(0) = P_N'(0) \\ &\ldots \\ &f^{(N)}(0) = P_N^{(N)}(0). \end{aligned} \tag{C.2}$$

$A_N(z)$ and $B_N(z)$ are complex polynomials of order $N/2$ and $N/2-1$, respectively, for even N. For odd N, both are of order $(N-1)/2$. In this work we always use even N. Therefore the number of poles that are represented by the N-point Pade approximant is equal to $N/2$. Based on Thiele's reciprocal difference method [43,44], $A_N(z)$ and $B_N(z)$ can be calculated recursively according to

$$A_n(z) = A_{n-1}(z) + (z - z_{n-1})\, a_n\, A_{n-2}(z) \tag{C.3}$$

$$B_n(z) = B_{n-1}(z) + (z - z_{n-1})\, a_n\, B_{n-2}(z) \tag{C.4}$$

with $A_0 = 0$, $A_1 = a_1$, and $B_0 = B_1 = 1$. The coefficients a_n are also calculated recursively by

$$a_n = g_n(z_n) \tag{C.5}$$

$$g_1(z_n) = f_n \tag{C.6}$$

$$g_p(z) = \frac{g_{p-1}(z_{p-1}) - g_{p-1}(z)}{(z - z_{p-1})g_{p-1}(z)} \quad p \geq 2. \tag{C.7}$$

Appendix D. Parallelization

G_0W_0 calculations are both CPU-time and memory intensive, and therefore parallelization is particularly important. The parallelization of `FHI-gap` is based on the message-passing interface (MPI), with the aim to reduce CPU time as well as memory requirements for each process. The latter becomes extremely relevant for large systems.

As obvious from Eqs. (18) and (22), the calculation of the self-energy includes the summation of independent contributions from different $\mathbf{q}$ vectors, hence suggesting the simplest way of parallelization. This can be realized with hardly any overhead. This procedure is perfectly suitable for small systems where hundreds of $\mathbf{q}$ points are considered. For larger systems, however, with typically more than ten atoms in the unit cell, usually a small number of $\mathbf{q}$ points is needed to obtain converged results. In this case, the $\mathbf{q}$ parallelization alone is not enough to take full advantage of the evermore powerful high-performance computing facilities. We therefore have introduced new parallelization mechanisms to allow for the use of a large number of processors.

To be more specific, the total number of processes is divided into groups. They are called *columns* in the code, and their number is denoted as n_{col}. Processes in different groups perform calculations for different $\mathbf{q}$ points. Processes within a "column" are further divided into "rows", the number of which is denoted as n_{row}. Parallelization over the rows is achieved as follows: the construction of $M_{nm}^i(\mathbf{k}, \mathbf{q})$ is parallelized over unoccupied bands (indexed m) as the number of unoccupied bands is usually quite large; for the calculations of $P_{ij}(\mathbf{q}, \omega)$, $\varepsilon_{ij}(\mathbf{q}, \omega)$, and $W_{ij}^c(\mathbf{q}, \omega)$, the parallelization is performed over frequency points. In terms of CPU time and memory consumption, the parallelization over *columns* scales almost linearly, but the memory usage is not reduced. On the other hand, the parallelization over *rows* can reduce the memory allocation down to $1/n_{\text{row}}$, but the gain in CPU time is less efficient, because a certain amount of communication is needed, and some quantities have to be recalculated in the different processes.

Appendix E. Interface with WIEN2k

The FHI-gap code is currently interfaced to WIEN2k by several C-shell scripts, which can be used for preparing the input files needed for *GW* calculations (preprocessing) and analyzing the *GW* results afterwards (post-processing). Here we give a brief description of their usage. More detailed documents on the compilation and use of FHI-gap can be found in the package together with the source code.

Preprocessing. After a self-consistent field (SCF) calculation using WIEN2k [17], all data files required for the subsequent G_0W_0 calculation are prepared by the C-shell script gap_init. It creates Kohn–Sham wave-functions and energies on a new **k**-mesh (usually smaller than the one used in the SCF calculation), providing energies up to $\epsilon_{\max}$. It also generates the master input file for the *GW* calculation, *case.ingw*, which is a free-formatted text file that assigns default values to all important parameters [65]. Note that these usually work for most cases, but careful convergence tests, as demonstrated in Section 5, should be conducted to obtain reliable results.

A sample *case.ingw* reads as follows:

```
Task = "gw"             # Option for task
Restart = F             # Option for whether restarting from a previous calculation
nspin = 1               # 1 for spin-unpolarized and 2 for spin-polarized calculations
nvel = 8.0              # number of valence electrons
ComplexVector = F       # T for systems with complex Kohn-Sham vectors
SymVector = F           # whether to use Kohn--Sham eigenvectors in the irreducible Brillouin zone
barcevtol = 0.0         # tolerance to reduce the bare Coulomb matrix eigenvectors for the correlation self-energies
                        # for sp systems, barcevtol=0.6 is usually quite safe.
emingw = -2.0           # emingw and emaxgw (in Ry) to control the range of bands
emaxgw =  2.0           # for which GW correction are calculated. Only states
                        # between E_Fermi+emingw and E_Fermi+emaxgw are calculated
%SelfEnergy             # options for correlation self-energy
  2  |  0   |   1       # npol |  iopes  |  iopac
%                       # npol: Number of poles
                        # iopes: 0/1/2/3 - without or with iteration
                        # iopsac:0/1     - Pade's approximation / multipole fitting
%FreqGrid               # Frequency grid parameters
3 | 16 | 0.42 |0.0 | 0  # iopfreq | nomeg | omegmax | omegmin | nomeg_blk
%                       # iopfreq= 1 (equally spaced), 2 (Gauss--Laguerre), 3 (double Gauss--Legendre)
%MixBasis               # Mixed basis parameters
1.0 | 3 | 1.E-4         # Q, l_max^MB, lamda_max^MB
%
```

Post-processing: The main results of a G_0W_0 calculation by running the FHI-gap code are a set of quasi-particle energies on an equally spaced **k**-mesh. To calculate densities of states or plot the band structure diagram along high symmetry directions in the BZ, one usually needs QP energies on a fine **k**-mesh. Calculating QP energies at an arbitrary **k** is possible, but computationally expensive. Currently, the FHI-gap code uses the Fourier interpolation approach [66] to obtain QP energies on **k**-points that are not included in the original **k**-mesh. For that purpose, the C-shell script gap_analy can be used.

More detailed information on the compilation and use of the code can be obtained from the webpage of the FHI-gap code [60].

References

[1] S. Hüfner, Photoelectron Spectroscopy: Prinples and Applications, third ed., Springer, Berlin, 2003.
[2] P. Hohenberg, W. Kohn, Phys. Rev. B 136 (1964) 864.
[3] W. Kohn, L.J. Sham, Phys. Rev. A 140 (1965) 1133.
[4] F. Aryasetiawan, O. Gunnarsson, Rep. Progr. Phys. 61 (1998) 237.
[5] A.L. Fetter, J.D. Walecka, Quantum Theory of Many-Particle Systems, McGraw-Hill, New York, 1971.
[6] L. Hedin, Phys. Rev. 139 (1965) A796.
[7] L. Hedin, B.I. Lundqvist, Solid State Phys. 23 (1969) 1.
[8] J.F. Janak, Phys. Rev. B 18 (1978) 7165.
[9] M.S. Hybertsen, S.G. Louie, Phys. Rev. B 34 (1986) 5390.
[10] R.W. Godby, M. Schlüter, L.J. Sham, Phys. Rev. B 37 (1988) 10159.
[11] P. Rinke, A. Qteish, J. Neugebauer, M. Scheffler, Phys. Status Solidi (b) 245 (2008) 929.
[12] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, M. Scheffler, New J. Phys. 7 (2005) 126.
[13] T. Kotani, M. van Schilfgaarde, Solid State Commun. 121 (2002) 461.
[14] W. Ku, A.G. Eguiluz, Phys. Rev. Lett. 89 (2002) 126401.
[15] R. Gomez-Abal, X. Li, M. Scheffler, C. Ambrosch-Draxl, Phys. Rev. Lett. 101 (2008) 106404.
[16] X.-Z. Li, R. Gomez-Abal, H. Jiang, C. Ambrosch-Draxl, M. Scheffler, New J. Phys. 14 (2012) 023006.
[17] P. Blaha, K. Schwarz, G.K.H. Madsen, D. Kvasnicka, J. Luitz, WIEN2k, an augmented plane wave + local orbitals program for calculating crystal properties, Karlheinz Schwarz, Techn. Universität Wien, Austria, 2001.
[18] C. Ambrosch-Draxl, S. Sagmeister, C. Meisenbichler, J. Spitaler, (in preparation).
[19] R.I. Gomez-Abal, X. Li, C. Ambrosch-Draxl, M. Scheffler, Extended linear tetrahedron method for the calculation of q-dependent dynamical response functions (in preparation).
[20] H. Jiang, R. Gomez-Abal, P. Rinke, M. Scheffler, Phys. Rev. Lett. 102 (2009) 126403.
[21] H. Jiang, R.I. Gomez-Abal, P. Rinke, M. Scheffler, Phys. Rev. B 82 (2010) 045108.
[22] H. Jiang, R. Gomez-Abal, P. Rinke, M. Scheffler, Phys. Rev. B 81 (2010) 085119.
[23] A. Baldereschi, E. Tosatti, Solid State Commun. 29 (1979) 131.
[24] J.C. Slater, Phys. Rev. 51 (1937) 846.
[25] O.K. Andersen, Phys. Rev. B 12 (1975) 3060.
[26] D.J. Singh, Phys. Rev. B 43 (1991) 6388.
[27] E. Sjöstedt, L. Nordström, D.J. Singh, Solid State Commun. 114 (2000) 15.
[28] S. Sagmeister, C. Ambrosch-Draxl, Phys. Chem. Chem. Phys. 11 (2009) 4451.
[29] M. Shishkin, G. Kresse, Phys. Rev. B 75 (2007) 235102.
[30] H. Jiang, Acta Phys.—Chim. Sin. 26 (2010) 1017.
[31] F. Aryasetiawan, Phys. Rev. B 46 (1992) 13051.
[32] F. Aryasetiawan, O. Gunnarsson, Phys. Rev. B 49 (1994) 16214.

[33] X. Li, All-electron G_0W_0 code based on FP-(L)APW+lo and applications, Ph.D. Thesis, Free University of Berlin, 2008.
[34] M.E. Rose, J. Math. Phys. 37 (1958) 215.
[35] B. Nijboer, F.D. Wette, Physica 23 (1957) 309.
[36] N. Karasawa, W.A. Goddard III, J. Phys. Chem. 93 (1989) 7320.
[37] C. Ambrosch-Draxl, J.O. Sofo, Comput. Phys. Comm. 175 (2006) 1.
[38] C. Friedrich, A. Schindlmayr, S. Blügel, Comput. Phys. Comm. 180 (2009) 347.
[39] C. Freysoldt, et al., Comput. Phys. Comm. 176 (2007) 1.
[40] F. Aryasetiawan, The GW approximation and vertex corrections, in: V.I. Anisimov (Ed.), Strong Coulomb Correlations in Electronic Structure Calculations: Beyond the Local Density Approximation, Gordon and Breach Science Publishers, 2000, p. 1.
[41] H. Jiang, E. Engel, J. Chem. Phys. 127 (2007) 184108.
[42] H.N. Rojas, R.W. Godby, R.J. Needs, Phys. Rev. Lett. 74 (1995) 1827.
[43] H.J. Vidberg, J.W. Serene, J. Low Temp. Phys. 29 (1977) 179.
[44] K.-H. Lee, K.J. Chang, Phys. Rev. B 54 (1996) R8285.
[45] V.I. Anisimov, F. Aryasetiawan, A.I. Lichtenstein, J. Phys.: Condens. Matter. 9 (1997) 767.
[46] A.B. Shick, A.I. Liechtenstein, W.E. Pickett, Phys. Rev. B 60 (1999) 10763.
[47] C. Friedrich, A. Schindlmayr, S. Blögel, T. Kotani, Phys. Rev. B 74 (2006) 045104.
[48] C. Friedrich, M.C. Müller, S. Blügel, Phys. Rev. B 83 (2011) 081101(R).
[49] P. Umari, G. Stenuit, S. Baroni, Phys. Rev. B 79 (2009) 201104(R).
[50] J.A. Berger, L. Reining, F. Sottile, Phys. Rev. B 82 (2010) 041103(R).
[51] L. Caramella, G. Onida, F. Finocchi, L. Reining, F. Sottile, Phys. Rev. B 75 (2007) 205405.
[52] S.V. Faleev, M. van Schilfgaarde, T. Kotani, Phys. Rev. Lett. 93 (2004) 126406.
[53] F. Bruneval, N. Vast, L. Reining, Phys. Rev. B 74 (2006) 045102.
[54] M. Shishkin, M. Marsman, G. Kresse, Phys. Rev. Lett. 99 (2007) 246403.
[55] R. Sakuma, T. Miyake, F. Aryasetiawan, Phys. Rev. B 78 (2008) 075106.
[56] S. Biermann, F. Aryasetiawan, A. Georges, Phys. Rev. Lett. 90 (2003) 086402.
[57] P. Sun, G. Kotliar, Phys. Rev. Lett. 92 (2004) 196402.
[58] F. Aryasetiawan, K. Karlsson, O. Jepsen, U. Schönberger, Phys. Rev. B 74 (2006) 125106.
[59] http://www.wien2k.at/.
[60] http://www.chem.pku.edu.cn/jianghgroup/codes/fhi-gap.html.
[61] S. Massidda, M. Posternak, A. Baldereschi, Phys. Rev. B 48 (1993) 5058.
[62] H.J. Monkhorst, J.D. Pack, Phys. Rev. B 13 (1976) 5188.
[63] G. Lehman, M. Taut, Phys. Status Solidi (b) 54 (1972) 469.
[64] P.E. Blöchl, O. Jepsen, O.K. Andersen, Phys. Rev. B 49 (1994) 16223.
[65] The input parsing is handled by using the parser library obtained from the octopus package. http://www.tddft.org.
[66] W.E. Pickett, H. Krakauer, P.B. Allen, Phys. Rev. B 38 (1988) 2721.