6 SCIENTIFIC HIGHLIGHT OF THE MONTH

Hybrid functionals and GW approximation in the FLAPW method

Christoph Friedrich, Markus Betzinger, Martin Schlipf, and Stefan Blügel

Peter Grünberg Institut and Institute for Advanced Simulation,
Forschungszentrum Jülich and JARA, 52425 Jülich, Germany

Arno Schindlmayr

Department Physik, Universität Paderborn, 33095 Paderborn, Germany

Abstract

We present recent advances in numerical implementations of hybrid functionals and the GW approximation within the full-potential linearized augmented-plane-wave (FLAPW) method. The former is an approximation for the exchange-correlation contribution to the total energy functional in density-functional theory, and the latter is an approximation for the electronic self-energy in the framework of many-body perturbation theory. All implementations employ the mixed product basis, which has evolved into a versatile basis for the products of wave functions, describing the incoming and outgoing states of an electron that is scattered by interacting with another electron. It can thus be used for representing the non-local potential in hybrid functionals as well as the screened interaction and related quantities in GW calculations. In particular, the six-dimensional space integrals of the Hamiltonian exchange matrix elements (and exchange self-energy) decompose into sums over vector-matrix-vector products, which can be evaluated easily. The correlation part of the GW self-energy, which contains a time or frequency dependence, is calculated on the imaginary frequency axis with a subsequent analytic continuation to the real axis or, alternatively, by a direct frequency convolution of the Green function G and the dynamically screened Coulomb interaction W along a contour integration path that avoids the poles of the Green function. Hybrid-functional and GW calculations are notoriously computationally expensive. We present a number of tricks that reduce the computational cost considerably including the usage of spatial and time-reversal symmetries, modifications of the mixed product basis with the aim to optimize it for the correlation self-energy and another modification that makes the Coulomb matrix sparse, analytic expansions of the interaction potentials around the point of divergence at $\mathbf{k} = \mathbf{0}$, and a nested density and density-matrix convergence scheme for hybrid-functional calculations. We show CPU timings for prototype semiconductors and illustrative results for GdN and ZnO.

1 Introduction

Within the last decades density-functional theory (DFT) [1, 2] has evolved into a powerful tool for calculating the electronic ground-state properties of molecules and solids. It is usually applied within the Kohn-Sham (KS) formalism [3], in which the interacting many-electron system is mapped onto a system of fictitious noninteracting particles that, by construction, possesses the same particle density as that of the real system. All exchange and correlations effects of the many-electron system are incorporated into the so-called exchange-correlation (xc) energy functional, which is not known exactly and must be approximated in practice. The choice of the xc functional is the only practical approximation in this otherwise exact theory and determines the precision and efficiency of the numerical DFT calculations.

Fortunately, already simple approximations, such as the local-density (LDA) [4, 5] and generalized gradient approximation (GGA) [6, 7], give reliable results for a wide range of materials and properties. In these approximations the exchange-correlation energy density is given as a function of the local electron density (and density gradient in the case of the GGA) rather than as a functional, which enables a straightforward numerical implementation. However, the missing nonlocal dependence on the density leads to several shortcomings. First, the electrostatic interaction of the electron with the total electron charge, described by the Hartree potential, contains an unphysical interaction of the electron with itself, commonly referred to as self-interaction. This extra term should be compensated exactly by an identical term with opposite sign in the exchange potential, in the same manner as in Hartree-Fock theory. However, as the LDA and GGA exchange potentials are only approximate, this cancellation is incomplete and part of the self-interaction remains. This error leads, in particular, to an improper description of localized states, which appear too high in energy and tend to delocalize. This deficiency is particularly critical in systems whose electronic and magnetic properties are largely governed by the correlated motion of electrons in localized states. The rare-earth chalcogenides are among this class of materials, having incompletely filled f-electron shells. We will show results for one such material, GdN, below. Second, the LDA and GGA functionals do not give rise to a discontinuity of the xc potential with respect to changes in the particle number. Theoretically, the sum of the KS band gap and the discontinuity yields the real band gap [8, 9]. The KS band gap alone is known to underestimate experimental gaps by as much as 50%. This is often called the band-gap problem of LDA and GGA. Finally, the LDA and GGA functionals do not exhibit the correct asymptotic behavior of exchange and correlation effects between interacting but spatially separate parts of an electronic system.

During the last decade hybrid functionals, which combine a local or semi-local xc functional with nonlocal Hartree-Fock (HF) exchange, have been shown to overcome these deficiencies to a great extent [10, 11, 12, 13, 14]. Hybrid functionals are usually applied within the generalized Kohn-Sham (gKS) scheme [15], where the HF exchange term leads to a nonlocal exchange potential in the one-particle equations. The first hybrid functional, a half-and-half mixing of the LDA functional with HF exchange, was proposed by Becke in 1993 [16]. Since then various *ab initio* and semi-empirical hybrid functionals have been published [17, 10, 18, 19]. The hybrid functionals, on which we focus in this article, do not contain any empirical parameters and are thus *ab initio* hybrid functionals.

The presence of the nonlocal exchange potential requires the description of the scattering of two particles interacting with each other via the Coulomb interaction. This involves to treat the incoming and outgoing states of the two particles explicitly, which will lead us to introduce a special basis for the product of wave functions. Once such a basis formulation of the electron-electron interaction is found, the question arises whether one can take a wider perspective beyond the HF exchange term and build up a numerical framework for many-body perturbation theory [20]. We will see that this is, indeed, possible. In many-body perturbation theory, one starts with a noninteracting system and incorporates the electron-electron interaction by adiabatically switching on the interaction. In the limit of full interaction strength, one finds that physical quantities can be written systematically as summations over integrals – the so-called Feynman diagrams – that contain increasingly higher orders of the electron-electron interaction. In this way, one can in principle construct the electronic self-energy in a systematic way. The self-energy incorporates all many-body exchange and correlation effects and is, in this sense, the counterpart of the xc functional of DFT. A very successful approximation of the self-energy is given by the *GW* approximation [21], which contains the electronic exchange exactly while the screening is treated at the level of the random-phase approximation, where noninteracting electron-hole ring diagrams are summed to all orders. The *GW* approximation is mainly used to compute electronic excitation energies and lifetimes from first principles so that it complements DFT, which is in practice restricted to the electronic ground state.

Due to the complexity of hybrid-functional and *GW* calculations, most implementations [22, 23, 24] have been based on the pseudopotential plane-wave approach, which effectively restricts the range of materials that can be examined. In particular, materials containing transition metals and rare earths as well as oxides cannot be treated efficiently in this approach. Two early all-electron calculations using the *GW* approximation were carried out by Hamada *et al.* [25] for Si and by Aryasetiawan [26] for Ni, both within the LAPW method. However, only very recently were further all-electron implementations reported, based on the FLAPW [27, 28], the linearized muffin-tin orbital (LMTO) [29, 30, 31], the projector-augmented-wave (PAW) [32, 33, 34], and the Korringa-Kohn-Rostoker method [35] together with applications to a larger variety of systems. Apart from an early implementation within a Gaussian-type basis [36], the hybrid functionals were adopted to all-electron methods only in recent years. In 2005 Paier *et al.* [37] developed an implementation within the projector-augmented-wave (PAW) technique. In 2006 Novak *et al.* [38] proposed an approximate scheme within the FLAPW approach. There the nonlocal exchange term is evaluated only in individual atomic spheres and only for selected l channels. In 2010 we presented an efficient numerical implementation of hybrid functionals [39] within the FLAPW method, which is not subject to these constraints. The FLAPW method provides a highly accurate all-electron basis, with which the electron structure of a large variety of materials can be studied, including open systems with low symmetry, d- and f-electron systems as well as oxides. It treats core and valence electrons on an equal footing.

In the first Hartree-Fock implementation within the FLAPW method, Massidda *et al.* [40] employed an algorithm that is routinely used for solving the Poisson equation for the electronic and nuclear charge distribution, i.e., it generates the electrostatic potential [41]. This *Poisson solver* can also be used for the nonlocal exchange potential, because its matrix representation involves formally identical six-dimensional integrals over space. Instead of the real charge density

one then uses an artificial charge density formed by the product of two wave functions. Recently, this scheme was employed for an implementation of a Yukawa screened hybrid functional [42]. Unfortunately, although the algorithm is very fast itself, the *Poisson solver* must be called many, many times instead of just once when applied to the Hartree-Fock exchange potential, which makes this approach computationally very expensive. Furthermore, it can only be used to calculate the matrix elements of interaction potentials that possess a simple expansion in dyadic products of regular and irregular radial functions as well as spherical harmonics, e.g., the bare Coulomb and the screened Yukawa potential. The error function used in the HSE functional does not have this property. It also does not provide a framework for an implementation of the screened interaction within the random-phase approximation for the *GW* self-energy.

Here, we describe a numerical realization of hybrid functionals and the *GW* approximation within an auxiliary basis, the so-called mixed product basis, which is constructed from products of LAPW basis functions and consists of muffin-tin (MT) functions and interstitial plane waves [29]. This basis allows to describe the product formed by the incoming and outgoing states of an electron that scatters by interacting with another electron. For example, this leads mathematically to a decomposition of the state-dependent six-dimensional Hartree-Fock exchange integral into two three-dimensional and one state-independent six-dimensional integral. The latter gives rise to the Coulomb matrix represented in the mixed product basis. Thus, the matrix elements of the nonlocal exchange potential are evaluated as Brillouin-zone (BZ) sums over vector-matrix-vector products. It also allows to represent the polarization function, dielectric function, and screened interaction needed in *GW* calculations in an efficient way.

We have employed several tricks to speed up the expensive calculations of the nonlocal quantities. Among them, we employ a suitable unitary transformation, by which nearly all MT functions become multipole-free, which makes the Coulomb matrix sparse and reduces the computational effort considerably. Spatial and time-reversal symmetry is exploited to reduce the number of terms that have to be taken into account explicitly in the summations over the Brillouin zone. Symmetry can also be used to decide in advance which of the matrix elements will be nonzero and have to be calculated. Furthermore, several cutoffs, such as a band cutoff for the hybrid functionals as well as a Coulomb eigenvalue cutoff for the correlation part of the *GW* self-energy, are introduced to speed up the code. The long-range nature of the Coulomb interaction gives rise to a divergence of the Coulomb matrix at the center of the BZ leading to a divergent integrand in the exchange and correlation terms, as well as to an anisotropy of the screened interaction in the BZ center. The divergence of the Coulomb potential has to be taken into account in a numerically efficient way to obtain a favorable **k**-point convergence. Corrections beyond the divergent $1/k^2$ term obtained from **k·p** perturbation theory can improve the convergence further.

The article is organized as follows. Section 2 gives a brief introduction to the theory of hybrid functionals and the *GW* approximation. Our implementations are discussed in detail in Sec. 3. After some test calculations, we present results for prototype semiconductors and insulators in Sec. 4 and compare them with theoretical and experimental results from the literature. We also present illustrative results for GdN and ZnO. Section 5 gives a summary.

2 Theory

In the following, we will summarize the formulas that have to be implemented for the hybrid functionals and the GW approximation. Even though the underlying theories are rather different – DFT [1] and many-body perturbation theory [20] – the formulas have a similar structure, and we can use similar strategies for an implementation.

GW **approximation**   We start with the GW approximation for the electronic self-energy [21]. The self-energy describes the renormalization of an additional particle, which can be an electron or a hole, due to the interaction with the rest of the electron system. From the renormalized Green function, information about the eigenstates of the many-electron system can be obtained, e.g., the excitation energies and lifetimes, which are measured experimentally in photoelectron spectroscopy. The GW approximation incorporates the renormalization due to electron exchange exactly and a large part of the renormalization due to electron correlation, where screening effects are taken into account on the level of the random-phase approximation (RPA) [43]. In a many-electron system the bare Coulomb potential $v(|\mathbf{r} - \mathbf{r}'|) = 1/|\mathbf{r} - \mathbf{r}'|$ created by an electron at $\mathbf{r}$ is screened by the redistribution of the electrons nearby.[1] In the RPA this redistribution (excluding the feed-back potential change due to this charge redistribution, which will be included later by solving a Dyson-type equation) is described by the polarization function

$$P(\mathbf{r}, \mathbf{r}'; \omega) = -i \sum_\sigma \lim_{\eta \to 0^+} \int_{-\infty}^\infty d\omega' G^\sigma(\mathbf{r}, \mathbf{r}'; \omega') G^\sigma(\mathbf{r}', \mathbf{r}; \omega' - \omega) e^{i\eta\omega'}, \tag{1}$$

where σ is the electron spin. Diagrammatically, P is represented by a bubble formed by two Green function lines. For the non-interacting $G(\mathbf{r}, \mathbf{r}'; \omega)$ we employ the KS Green function, which can be written in terms of the KS wave functions $\varphi_{n\mathbf{k}}^\sigma(\mathbf{r})$ and energies $\epsilon_{n\mathbf{k}}^\sigma$ as $G^\sigma(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{n\mathbf{k}} \varphi_{n\mathbf{k}}^\sigma(\mathbf{r}) \varphi_{n\mathbf{k}}^{\sigma*}(\mathbf{r}') / [\omega - \epsilon_{n\mathbf{k}}^\sigma + i\eta \, \mathrm{sgn}(\epsilon_{n\mathbf{k}}^\sigma - \epsilon_F)]$ with the Fermi energy ϵ_F, where $\mathbf{k}$ and n are the Bloch vector and the band index, respectively. Inserting into Eq. (1) yields [43]

$$P(\mathbf{r}, \mathbf{r}'; \omega) = \lim_{\eta \to 0^+} \sum_\sigma \sum_{\mathbf{q},\mathbf{k}}^{\mathrm{BZ}} \sum_n^{\mathrm{occ}} \sum_{n'}^{\mathrm{unocc}} \varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r}) \varphi_{n'\mathbf{k}}^\sigma(\mathbf{r}) \varphi_{n\mathbf{q}}^\sigma(\mathbf{r}') \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}')$$
$$\times \left(\frac{1}{\omega + \epsilon_{n\mathbf{q}}^\sigma - \epsilon_{n'\mathbf{k}}^\sigma + i\eta} - \frac{1}{\omega - \epsilon_{n\mathbf{q}}^\sigma + \epsilon_{n'\mathbf{k}}^\sigma - i\eta} \right). \tag{2}$$

With the polarization function we obtain the dynamically screened interaction $W(\mathbf{r}, \mathbf{r}'; \omega)$ from the integral equation

$$W(\mathbf{r}, \mathbf{r}'; \omega) = v(|\mathbf{r} - \mathbf{r}'|) \tag{3}$$
$$+ \int v(|\mathbf{r} - \mathbf{r}''|) P(\mathbf{r}'', \mathbf{r}'''; \omega) W(\mathbf{r}''', \mathbf{r}'; \omega) \, d^3 r'' \, d^3 r''',$$

where the implicit infinite summation of this Dyson-type equation includes the feed-back effects mentioned above.

An expansion of the electronic self-energy $\Sigma_{\mathrm{xc}}^\sigma(\mathbf{r}, \mathbf{r}'; \omega)$ in terms of W up to linear order yields the GW approximation

$$\Sigma_{\mathrm{xc}}^\sigma(\mathbf{r}, \mathbf{r}'; \omega) = \frac{i}{2\pi} \lim_{\eta \to 0^+} \int G^\sigma(\mathbf{r}, \mathbf{r}'; \omega + \omega') W(\mathbf{r}, \mathbf{r}'; \omega') e^{i\eta\omega'} \, d\omega'. \tag{4}$$

---

[1]Here and in the following we use atomic units unless noted otherwise.

If we write the screened interaction as a sum of the bare interaction and a remainder,

$$W(\mathbf{r}, \mathbf{r}'; \omega) = v(|\mathbf{r} - \mathbf{r}'|) + W^{c}(\mathbf{r}, \mathbf{r}'; \omega), \tag{5}$$

the self-energy [Eq. (4)] decomposes into the terms

$$\Sigma_{x}^{\sigma}(\mathbf{r}, \mathbf{r}') = \frac{i}{2\pi} \lim_{\eta \to 0^{+}} \int G^{\sigma}(\mathbf{r}, \mathbf{r}'; \omega + \omega') v(|\mathbf{r} - \mathbf{r}'|) e^{i\eta\omega'} d\omega' \tag{6}$$

and

$$\Sigma_{c}^{\sigma}(\mathbf{r}, \mathbf{r}'; \omega) = \frac{i}{2\pi} \int G^{\sigma}(\mathbf{r}, \mathbf{r}'; \omega + \omega') W^{c}(\mathbf{r}, \mathbf{r}'; \omega') d\omega', \tag{7}$$

which are identified as the exchange and the correlation contributions to the electronic self-energy, respectively. We note that the exponential factor allows to close the integration path over the upper complex half plane in Eq. (6). As $W^{c}(\mathbf{r}, \mathbf{r}'; \omega)$ falls off quickly enough with increasing frequencies, we may take the limit $\eta \to 0$ before integrating in Eq. (7).

The excitation energies $E_{n\mathbf{q}}^{\sigma}$ of the $N+1$ and $N-1$ particle system (for unoccupied and occupied states n, respectively) form the poles of the renormalized Green function, which can be calculated as corrections on the KS energies within first-order perturbation theory, yielding the nonlinear equation

$$E_{n\mathbf{q}}^{\sigma} = \epsilon_{n\mathbf{q}}^{\sigma} + \langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma_{xc}^{\sigma}(E_{n\mathbf{q}}^{\sigma}) - v_{xc}^{\sigma} | \varphi_{n\mathbf{q}}^{\sigma} \rangle, \tag{8}$$

where $v_{xc}^{\sigma}(\mathbf{r})$ is the LDA exchange-correlation potential. While the expectation value of the latter and the Kohn-Sham energies are routinely calculated in the DFT code, we are left with the task of evaluating the expectation value $\langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma_{xc}^{\sigma}(E_{n\mathbf{q}}^{\sigma}) | \varphi_{n\mathbf{q}}^{\sigma} \rangle$, which decomposes into an exchange and a correlation term according to Eqs. (6) and (7). Finally, we must solve the nonlinear Eq. (8). Because of the exponential factor in Eq. (6), we can formally close the frequency integration contour with an infinite half circle over the positive complex plane without changing the value of the integral. This contour integral then equals the sum over the residues of the poles of the Green function. The expectation value of the exchange term with respect to a wave function $\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r})$ yields the well-known Hartree-Fock (HF) expression

$$\langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma_{x}^{\sigma} | \varphi_{n\mathbf{q}}^{\sigma} \rangle = - \sum_{\mathbf{k}}^{BZ} \sum_{n'}^{occ.} \int \varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r}) \varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r}) v(|\mathbf{r} - \mathbf{r}'|) \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}') \varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}') d^{3}r \, d^{3}r' \tag{9}$$

and from (7) we obtain

$$\langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma_{c}^{\sigma}(\omega) | \varphi_{n\mathbf{q}}^{\sigma} \rangle = \frac{i}{2\pi} \lim_{\eta \to 0^{+}} \sum_{\mathbf{k}}^{BZ} \sum_{n'}^{all} \int_{-\infty}^{\infty} d\omega' \tag{10}$$

$$\times \int \frac{\varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r}) \varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r}) W^{c}(\mathbf{r}, \mathbf{r}'; \omega') \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}') \varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}')}{\omega + \omega' - \epsilon_{n'\mathbf{k}}^{\sigma} + i\eta \, \text{sgn}(\epsilon_{n'\mathbf{k}}^{\sigma} - \epsilon_{F})} d^{3}r \, d^{3}r'.$$

In contrast to Eq. (9), the frequency integral in the latter expression cannot be replaced by a sum over residues because the poles of W^{c} are unknown.

Hybrid functionals The adiabatic-connection formula [44, 45, 46] establishes a connection between the noninteracting Kohn-Sham system with the fully interacting one by scaling the electron-electron interaction $\lambda v(r)$ from $\lambda = 0$ to $\lambda = 1$. In the weakly interacting limit, the

formula becomes identical to the HF exchange term, which motivated Becke [47, 16, 10] to replace a fraction a of the local exchange functional of DFT by the same fraction of nonlocal HF exchange energy

$$E_{\text{xc}}^{\text{HYB}} = E_{\text{xc}}^{\text{L}} + a \left(E_{\text{x}}^{\text{HF}} - E_{\text{x}}^{\text{L}} \right) , \tag{11}$$

where E_{xc}^{L} denotes the local xc functional and E_{xc}^{L} its exchange part.

$$\begin{aligned} E_{\text{xc}}^{\text{HF}} &= -\frac{1}{2} \sum_{\mathbf{q},\mathbf{k}}^{\text{BZ}} \sum_{n,n'}^{\text{occ.}} \int \varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r}) \varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r}) v(|\mathbf{r} - \mathbf{r}'|) \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}') \varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}') d^3r\, d^3r' \\ &= \frac{1}{2} \sum_{\mathbf{q}}^{\text{BZ}} \sum_{n}^{\text{occ.}} \langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma_{\text{x}}^{\sigma} | \varphi_{n\mathbf{q}}^{\sigma} \rangle \end{aligned} \tag{12}$$

is the HF exchange energy evaluated with the Kohn-Sham wave functions. The last equality shows that there is a direct relation to the expectation value of the exchange self-energy, Eq. (9). As the wave functions are functionals of the effective potential, which in turn is a functional of the density, $E_{\text{xc}}^{\text{HF}}$ is a true functional of the density, too.

By assuming a certain shape for the adiabatic-connection integrand, Perdew *et al.* [17] deduced a mixing parameter $a = 0.25$. With the Perdew-Burke-Ernzerhof (PBE) functional [6] for the local part, the resulting functional

$$E_{\text{xc}}^{\text{PBE0}} = E_{\text{xc}}^{\text{PBE}} + a \left(E_{\text{x}}^{\text{HF}} - E_{\text{xc}}^{\text{PBE}} \right) \tag{13}$$

is nowadays referred to as PBE0. For localized basis sets such as Gaussian-type functions, the long-range part of the nonlocal HF term is particularly expensive to evaluate. Therefore, Heyd *et al.* [48, 49] replaced it by a simple local functional, whereas the short-range part is still treated with a nonlocal potential. For this, they decomposed the bare Coulomb potential into a long-range (LR) and a short-range (SR) part

$$v(r) = \frac{\text{erf}(\kappa r)}{r} + \frac{\text{erfc}(\kappa r)}{r} = v^{\text{LR}}(r) + v^{\text{SR}}(r) , \tag{14}$$

where $\text{erf}(x)$ and $\text{erfc}(x) = 1 - \text{erf}(x)$ are the error function and its complement, respectively, and κ is an adjustable screening parameter. Later, this approach was demonstrated to even yield a better description of semiconductor band gaps than PBE0 [50]. The HSE hybrid functional is thus given by

$$E_{\text{xc}}^{\text{HSE}}(\kappa) = E_{\text{xc}}^{\text{PBE}} + a \left[E_{\text{x}}^{\text{HF,SR}}(\kappa) - E_{\text{xc}}^{\text{PBE,SR}}(\kappa) \right] , \tag{15}$$

where $E_{\text{x}}^{\text{HF,SR}}(\kappa)$ corresponds to Eq. (12) with the bare Coulomb potential $v(r)$ replaced by $v^{\text{SR}}(r)$. The numerical evaluation of $E_{\text{x}}^{\text{PBE,SR}}(\kappa)$, the local functional for the SR exchange according to the decomposition given in Eq. (14), is described in Refs. 48 and 51. The screening parameter was fitted to a set of benchmark data for molecules yielding $\kappa = 0.15$ [48]. Later Krukau *et al.* published a slightly different value, $\kappa = 0.11$, which was optimized for solids [52]. We employ the latter value in our implementation.

Hybrid functionals are usually applied within the generalized KS formalism [15], which differs from the KS formalism by the fact that the effective potential, in which the fictitious noninteracting electrons move, is not necessarily local. It may contain a nonlocal part in addition to the local part so that the wave functions of the particles are solutions of the differential equation

$$[-\frac{1}{2}\nabla^2 + V^{\sigma,\text{L}}(\mathbf{r})]\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}) + a \int V^{\sigma,\text{NL}}(\mathbf{r},\mathbf{r}')\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}') d^3r = \epsilon_{n\mathbf{q}}^{\sigma}\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}) , \tag{16}$$

where we have already included the mixing parameter a. The local part $V^{\sigma,\mathrm{L}}(\mathbf{r})$ consists of the contributions that are local *per se*, i.e., the external potential created by the nuclear charges, the Hartree potential (the electrostatic potential produced by the total electron charge density), as well as a contribution that derives from functional differentiation of the local parts of Eqs. (13) and (15) for the PBE0 and HSE functionals, respectively. The implementation of this local part of the xc potential requires only minor modifications of the DFT code, and we will focus on the nonlocal part $V^{\sigma,\mathrm{NL}}(\mathbf{r},\mathbf{r}')$ in the following, which derives from the nonlocal exchange energy functional $E_{\mathrm{x}}^{\mathrm{HF}}$. Leaving out the scaling factor a, its matrix representation in the basis of Kohn-Sham eigenstates and in reciprocal space is given by

$$
\begin{aligned}
V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL}}(\mathbf{q}) &= -\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{m}^{\mathrm{occ.}}\iint d^3r\,d^3r'\,\varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r})\varphi_{m\mathbf{k}}^{\sigma}(\mathbf{r})v(|\mathbf{r}-\mathbf{r}'|)\varphi_{m\mathbf{k}}^{\sigma*}(\mathbf{r}')\varphi_{n'\mathbf{q}}^{\sigma}(\mathbf{r}') \qquad (17)\\
&= \langle\varphi_{n\mathbf{q}}^{\sigma}|\Sigma_{\mathrm{x}}^{\sigma}|\varphi_{n'\mathbf{q}}^{\sigma}\rangle\,,
\end{aligned}
$$

the diagonal elements of which correspond to the expectation value of the exchange self-energy, Eq. (9). In a sense, the hybrid functionals can be regarded as alternative approximations to the electronic self-energy. In fact, the resulting one-particle energies are often in much better agreement to experiment than the usual Kohn-Sham energies within the LDA or GGA, as we will see in Sec. 4. Of course, Eqs. (13) and (15) miss the frequency dependence, which is present in the GW self-energy. On the other hand, the hybrid-functional calculations are carried out iteratively until self-consistency in the electron density is achieved, while GW calculations are usually performed only perturbatively as *one-shot* calculations.

The sum over the occupied states in Eq. (17) involves core and valence states. Core states are dispersionless, which can be shown to lead to particularly simple and computationally cheap expressions for their contribution to the exchange term [53]. The valence states, on the other hand, show a distinct $\mathbf{k}$ dependence, which must be taken into account properly. Here we employ the mixed product basis (MPB) [54, 29], which is introduced in the next section.

For the HSE functional, the bare interaction $v(r)$ in Eq. (17) would have to be replaced by the short-range interaction $v^{\mathrm{SR}}(r)$. (Then the last equality would turn into an inequality, of course). Yet in practice, we first evaluate the nonlocal potential in the form of Eq. (17) and subtract a corresponding potential with $v^{\mathrm{LR}}(r)[= v(r)-v^{\mathrm{SR}}(r)]$ afterwards, as will be explained in the next section.

3 Implementation

3.1 Basis sets – FLAPW and mixed product basis

Our implementation is based on the all-electron FLAPW method [55, 56, 57], in which space is partitioned into non-overlapping atom-centered muffin-tin (MT) spheres and the remaining interstitial region (IR). The fully relativistic Dirac equation is employed to determine the core states, which are confined to the spheres. The valence and conduction states are obtained by solving Eq. (16) in a basis representation, by which the one-particle differential equations turn into a generalized eigenvalue problem. The basis functions are defined differently in the two regions of space: plane waves $e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}$ with $|\mathbf{k}+\mathbf{G}| \leq G_{\mathrm{max}}$ in the IR and linear combinations of

$u_{lp}^{a\sigma}(r)Y_{lm}(\hat{\mathbf{r}})$ in the MT spheres, where r is measured from the sphere center, $u_{lp}^{a\sigma}(r)$ are numerical functions defined on a radial grid, $Y_{lm}(\mathbf{r})$ are spherical harmonics with angular-momentum quantum numbers $0 \leq l \leq l_{\max}$ and $|m| \leq l$, a labels the atom in the unit cell, and p is an index for different radial functions. $G_{\max}$ and $l_{\max}$ are cutoff parameters. The linear combinations are such that the basis functions and their radial derivatives are continuous at the MT sphere boundaries. The radial functions $u_{lp}^{a\sigma}(r)$ are constructed from the radial scalar-relativistic KS equation with the spherical part of the local effective potential. In this way, the basis functions are optimized for the given physical system. Simply speaking, the radial functions already represent approximate solutions of the one-particle KS equations.

Hybrid functionals and the *GW* approximation are methods that distinguish themselves from conventional DFT with local (LDA) or semilocal (GGA) functionals by the fact that individual two-particle scattering processes are described explicitly. In each scattering event the state of the incoming particle and the state that it scatters into form products of wave functions. This is apparent in the above formulas for the different physical quantities, such as the HF potential [Eq. (17)], the polarization function [Eq. (2)], the self-energy [Eqs. (6) and (7)], and so on. All these are given as six-dimensional integrals over products of wave functions. Thus, if the integral equations are rewritten in an auxiliary basis constructed for these product functions, they become matrix equations, which can easily be treated in a computer code using standard linear algebra libraries. In the context of the FLAPW method we have found the mixed product basis (MPB) [29] to be an optimal choice, which consists of two separate sets of functions that are defined only in one of the spatial regions, while being zero in the other. We will describe this basis in the following.

From the above basis representation inside the MT spheres, we would naively just take the product functions $U_{LP}^{a\sigma}(r) = u_{lp}^{a\sigma}(r)u_{l'p'}^{a\sigma}(r)$ as the radial part of a MPB function, while the spherical harmonics $Y_{lm}(\hat{\mathbf{r}})$ and $Y_{l'm'}(\hat{\mathbf{r}})$ multiply to linear combinations of spherical harmonics $Y_{LM}(\hat{\mathbf{r}})$ again with $|l - l'| \leq L \leq l + l'$ and $-L \leq M \leq L$. The index P counts all possible combinations of l, l', p, and p'. However, the set $\{U_{LP}^{a\sigma}(r)\}$ is very large and usually has a high degree of (near) linear dependence. An effective procedure to reduce the size of the set and to remove the linear dependences is to diagonalize the overlap matrix of $\{U_{LP}^{a\sigma}(r)\}$ and to retain only those eigenvectors whose eigenvalues exceed a specified threshold value [58]. By using both spin-up and spin-down products in the construction of the overlap matrix, we make the resulting basis spin independent. In practice, the basis set is reduced further by introducing a cutoff value $L_{\max}$ for the angular quantum number, which can be chosen much smaller than the theoretically exact limit of $2l_{\max}$, and by restricting the product functions to certain combinations of p and p'. On the other hand, it must be supplemented with a constant MT function for each atom in the unit cell, which is later needed to represent the eigenfunction that corresponds to the divergent eigenvalue of the Coulomb matrix in the limit $\mathbf{k} \to \mathbf{0}$. At $\mathbf{k} = \mathbf{0}$, this eigenfunction is the constant function. From the resulting MT functions $M_{LMP}^a(\mathbf{r}) = M_{LP}^a(r)Y_{LM}(\hat{\mathbf{r}})$, we formally construct Bloch functions.

In the IR, we have to form products of plane waves, which are plane waves $M_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r}) = e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}$ again. As above, the cutoff value $|\mathbf{k} + \mathbf{G}| \leq G'_{\max}$ can be chosen much smaller than the exact limit $2G_{\max}$ in practice. Together with the MT functions, we thus obtain the MPB $\{M_I^{\mathbf{k}}(\mathbf{r})\} = \{M_{LMP}^{a,\mathbf{k}}(\mathbf{r}), M_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r})\}$ for the representation of wave-function products. Since the interstitial plane

waves are not orthogonal to each other, we must introduce the biorthogonal set $\{\tilde{M}_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}'} M_{\mathbf{G}'}^{\mathbf{k}}(\mathbf{r}) O_{\mathbf{G}'\mathbf{G}}^{-1}(\mathbf{k})\}$, too, where $O_{\mathbf{G}'\mathbf{G}}(\mathbf{k})$ is the overlap matrix, which can be calculated analytically. (We note that the functions $M_{LMP}^{a,\mathbf{k}}(\mathbf{r})$ are orthonormal by construction.) In the Hilbert space spanned by the MPB, the two sets can be combined to yield the completeness relations

$$\sum_I |M_I^{\mathbf{k}}\rangle\langle\tilde{M}_I^{\mathbf{k}}| = \sum_I |\tilde{M}_I^{\mathbf{k}}\rangle\langle M_I^{\mathbf{k}}| = 1 \,, \tag{18}$$

which can be used to decouple the $\mathbf{r}$ and $\mathbf{r}'$ spaces in the nonlocal integrals. This technique is also called *resolution of the identity*. Sometimes, it is helpful to combine the sets $\{M_{LMP}^{a,\mathbf{k}}(\mathbf{r})\}$ and $\{M_{\mathbf{G}}^{\mathbf{k}}(\mathbf{r})\}$ to form a new set of functions that are continuous at the MT sphere boundaries. This basis set contains less functions without a loss of accuracy, which can increase the efficiency of the code. Furthermore, the elimination of the discontinuities can improve the numerical stability if matrices are diagonalized or inverted, e.g., the Coulomb and dielectric matrix in the GW approximation as well as the response matrix in the optimized effective potential method [59, 60].

3.2 MPB formulation

Let us first apply the MPB to the nonlocal potential Eq. (17) for the hybrid functionals. By placing the completeness relation Eq. (18) at both sides of $v(|\mathbf{r} - \mathbf{r}'|)$, the integrations over $\mathbf{r}$ and $\mathbf{r}'$ get decoupled, and the six-dimensional integral is replaced by a sum over vector-matrix-vector products

$$
\begin{aligned}
V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL}}(\mathbf{q}) &= -\sum_m^{\mathrm{occ.}} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{IJ} \langle\varphi_{n\mathbf{q}}^{\sigma}|\varphi_{m\mathbf{q}-\mathbf{k}}^{\sigma} M_{\mathbf{k}I}\rangle v_{IJ}(\mathbf{k})\langle M_{\mathbf{k}J}\varphi_{m\mathbf{q}-\mathbf{k}}^{\sigma}|\varphi_{n'\mathbf{q}}^{\sigma}\rangle \\
&= \langle\varphi_{n\mathbf{q}}^{\sigma}|\Sigma_{\mathrm{x}}^{\sigma}|\varphi_{n'\mathbf{q}}^{\sigma}\rangle
\end{aligned}
\tag{19}
$$

with the usual bra-ket notation $\langle f|g\rangle = \int d^3r\, f^*(\mathbf{r})g(\mathbf{r})$. The Coulomb matrix [61]

$$v_{IJ}(\mathbf{k}) = \iint d^3r\, d^3r'\, \tilde{M}_{\mathbf{k}I}^*(\mathbf{r})v(|\mathbf{r} - \mathbf{r}'|)\tilde{M}_{\mathbf{k}J}(\mathbf{r}') \,, \tag{20}$$

which still involves a six-dimensional integration, must be evaluated only once at the beginning of the self-consistency cycle. In the same manner, Eqs. (2), (3), (9), and (10) become algebraic matrix equations in the space of the MPB. Equation (3) becomes particularly simple if we perform a basis transformation $\{M_I^{\mathbf{k}}(\mathbf{r})\} \rightarrow \{E_\mu^{\mathbf{k}}(\mathbf{r})\}$ that diagonalizes the Coulomb matrix. We note that no approximation is involved at this stage. The new normalized basis functions are necessarily orthogonal, and we do not need a biorthogonal set. In this new basis, Eq. (3) becomes

$$W_{\mu\nu}(\mathbf{k},\omega) = v_\mu(\mathbf{k})\delta_{\mu\nu} + v_\mu(\mathbf{k})\sum_\gamma P_{\mu\gamma}(\mathbf{k},\omega)W_{\gamma\nu}(\mathbf{k},\omega) \,, \tag{21}$$

which by inversion yields

$$W_{\mu\nu}(\mathbf{k},\omega) = \varepsilon_{\mu\nu}^{-1}\sqrt{v_\mu(\mathbf{k})v_\nu(\mathbf{k})} \tag{22}$$

with the dielectric matrix

$$\varepsilon_{\mu\nu}(\mathbf{k},\omega) = \delta_{\mu\nu} - P_{\mu\nu}(\mathbf{k},\omega)\sqrt{v_\mu(\mathbf{k})v_\nu(\mathbf{k})} \,. \tag{23}$$

72

The last two equations do not contain any matrix multiplications, but only scalar products with $v_\mu(\mathbf{k})$, the eigenvalues of the Coulomb matrix [Eq. (20)]. In Sec. 3.4 we will come back to this special basis.

We now turn back to the hybrid functionals and, in particular, to the HSE functional, which is defined in terms of a short-range (or model screened) interaction v^{SR} defined in Eq. (14). The way of implementation seems obvious now: the bare Coulomb potential v in Eq. (19) for the PBE0 functional is simply replaced by v^{SR}. However, in this way we would loose a very favorable property of the bare Coulomb matrix, Eq. (20), namely that it can be made sparse by a simple unitary transformation of the MPB (see Sec. 3.4). The sparsity of $v_{IJ}(\mathbf{q})$ speeds up the matrix-vector multiplications in Eq. (19) considerably, while $v_{IJ}^{\mathrm{SR}}(\mathbf{q})$ does not have this mathematical property. Therefore, we evaluate Eq. (19) with the bare interaction as in the PBE0 case and subtract a corresponding expression, where we replace v by v^{LR}, the long-range interaction defined in Eq. (14). While $v(r)$ and $v^{\mathrm{SR}}(r)$ diverge at $r = 0$, $v^{\mathrm{LR}}(r)$ remains finite there and behaves like $v(r)$ for large r. So, it has a very smooth behavior for all distances and should thus be suitable to be described in reciprocal space. In fact, we find that its Fourier transform

$$v_{\mathbf{k}}^{\mathrm{LR}} = \frac{4\pi}{k^2} e^{-ik^2/(4\kappa^2)} \tag{24}$$

falls off rapidly with increasing k. The potential for the long-range interaction corresponding to Eq. (19) can thus be evaluated in reciprocal space

$$V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL,LR}}(\mathbf{q}) = -\sum_m^{\mathrm{occ.}} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{\mathbf{G}} \langle \varphi_{n\mathbf{q}}^\sigma | \varphi_{m\mathbf{q}-\mathbf{k}}^\sigma \frac{e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}}{\sqrt{V}} \rangle v_{\mathbf{k}+\mathbf{G}}^{\mathrm{LR}} \langle \frac{e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}}{\sqrt{V}} \varphi_{m\mathbf{q}-\mathbf{k}}^\sigma | \varphi_{n'\mathbf{q}}^\sigma \rangle \tag{25}$$

with the crystal volume V and

$$\langle \varphi_{n\mathbf{q}}^\sigma | \varphi_{m\mathbf{q}-\mathbf{k}}^\sigma \frac{e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}}{\sqrt{V}} \rangle = \sum_I \langle \varphi_{n\mathbf{q}}^\sigma | \varphi_{m\mathbf{q}-\mathbf{k}}^\sigma M_{\mathbf{k}I} \rangle \langle \tilde{M}_{\mathbf{k}I} | \frac{e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}}{\sqrt{V}} \rangle, \tag{26}$$

where the first integrals on the right-hand side are calculated routinely already for $V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL}}(\mathbf{q})$. The nonlocal short-range potential needed for HSE is then given by

$$V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL,SR}}(\mathbf{q}) = V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL}}(\mathbf{q}) - V_{\mathrm{x},nn'}^{\sigma,\mathrm{NL,LR}}(\mathbf{q}). \tag{27}$$

We note that any other form of the screened Coulomb interaction could easily be implemented at this stage. Since the long-range interaction is diagonal in reciprocal space, the evaluation of Eq. (25) takes in practical terms negligible time to compute. From the fact that the Fourier transform Eq. (24) approaches zero very quickly with k, it is clear that the results are easily converged up to machine precision, even if the Fourier coefficients in Eq. (26) fall off very slowly because of the rapidly varying all-electron wave functions. Since $v^{\mathrm{SR}}(r)$ can be understood as a model screened interaction, one might anticipate that a similar treatment of the long-range (or correlation) part W^{c} of the dynamically screened interaction in the GW approximation [Eq. (5)], can give rise to a similar speed-up of the computation time. We will demonstrate in Sec. 3.4 that this is, in fact, possible even without a projection onto plane waves.

3.3 Frequency dependence in GW

The GW approximation for the self-energy involves quantities that are time or frequency dependent such as the Green function, the polarization function, and the screened interaction. In this section, we show how this frequency dependence is treated in our code.

The polarization function describes the charge redistributions within a many-electron system due to a change in the effective potential. The electrons react quickly to this change but not instantaneous, making the polarization function dependent on the time that has passed since the perturbation or, after Fourier transformation, dependent on frequency. In the basis of eigenstates, the charge redistribution is realized by the formation of electron-hole pairs. Simply speaking, the expression in Eq. (2) can be understood as the summed response of a large number of oscillators with eigenfrequencies $\epsilon_{n'\mathbf{k}}^{\sigma} - \epsilon_{n\mathbf{q}}^{\sigma}$ (corresponding to the electron-hole pairs) to a perturbation that oscillates with frequency ω.

The response of the individual oscillators have a finite value over all frequencies, which complicates a direct summation over the Brillouin zone. It is more convenient to consider the representation $(\mathrm{Im}\, P)_{IJ}(\mathbf{k}, \omega)$ of the imaginary part $\mathrm{Im}\, P(\mathbf{r}, \mathbf{r}'; \omega)$ first, which is obtained by replacing expressions of the form $1/(a \pm i\eta)$ by $\mp \pi \delta(a)$. This simplifies the BZ summation significantly. Afterwards a Hilbert transformation yields the full polarization matrix $P_{IJ}(\mathbf{k}, \omega)$, where the frequency argument may be complex. In particular, this allows an evaluation on the imaginary-frequency axis, where the frequency-dependent quantities show a smooth behavior and can therefore be sampled and interpolated with few frequency points. As the bracket in Eq. (2) is real for frequencies on the imaginary axis, the corresponding matrix $P_{IJ}(\mathbf{k}, i\omega)$ with $\omega \in \mathbb{R}$ becomes Hermitian; it even becomes real symmetric if the system exhibits inversion symmetry and we use a symmetrized MPB as described in Sec. 3.4.

In contrast to the exchange self-energy, the frequency integral in Eq. (7) cannot be replaced by a sum over residues because the positions of the poles of $W_{\mu\nu}^{\mathrm{c}}(\mathbf{k}, \omega) = W_{\mu\nu}(\mathbf{k}, \omega) - \delta_{\mu\nu} v_{\mu}(\mathbf{k})$ in the complex-frequency plane are unknown. Therefore, the correlation self-energy still contains an explicit integration over frequencies. Unfortunately, the integrand has a lot of structure along the real frequency axis, which makes a direct evaluation difficult. There are two methods that avoid the integration over real frequencies and use the imaginary axis instead: analytic continuation [62] and contour integration [63, 64]. The former allows a faster and easier implementation, but contains a badly controlled fitting procedure, which can be tested with the more accurate contour-integration method. In the latter, the frequency integration path is chosen so as to avoid the real axis in the best possible way. It falls into two contributions: an integral that runs from $-i\infty$ to $+i\infty$ along the imaginary frequency axis and another integral that runs the perimeter of a rectangle only containing poles of the KS Green function. Since these poles are known, the latter integral is easily calculated as a sum over the residues. We have implemented both algorithms and find that they give similar results for the systems considered here. In the following, we hence discuss only the first approach in more detail.

The analytic-continuation method starts with the calculation of the self-energy for imaginary

frequencies

$$\langle\varphi_{n\mathbf{q}}^{\sigma}|\Sigma_{\mathrm{c}}^{\sigma}(i\omega)|\varphi_{n\mathbf{q}}^{\sigma}\rangle = -\frac{1}{2\pi}\lim_{\eta\to 0^+}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n'}^{\mathrm{all}}\int_{-\infty}^{\infty}d\omega' \tag{28}$$
$$\times\int\frac{\varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r})\varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r})W^{\mathrm{c}}(\mathbf{r},\mathbf{r}';i\omega')\varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}')\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}')}{i\omega+i\omega'-\epsilon_{n'\mathbf{k}}^{\sigma}}d^3r\,d^3r'\,.$$

on a mesh of points $\{i\omega_n;\,\omega_n\geq 0\}$. The frequency integration is along the imaginary-frequency axis, where the integrand is much smoother than on the real axis. In practice, we use a discrete and finite mesh for the imaginary frequencies, which is dense near $\omega = 0$ and contains typically 10-20 points. After evaluating Eq. (28), the self-energy is analytically continued to the whole complex frequency plane either by fitting to a model function given by a sum over N_{p} poles in the complex plane

$$f(\omega) = \sum_{\nu=1}^{N_{\mathrm{P}}}\frac{a_{\nu}}{\omega-\omega_{\nu}} \tag{29}$$

with complex fit parameters a_p and ω_p or by using a Padé approximant employing Thiele's reciprocal difference algorithm [65], which gives the approximant as a continued fraction. A subsequent search for the poles with the Newton method yields the self-energy again in the form of Eq. (29), usually with many more poles than is possible with the numerical fitting procedure. An accurate analytical continuation is characterized by the fact that all poles (or at least the dominant ones) are below the real frequency axis. We find that the Padé method is more precise than numerical fitting, but it can produce random outliers if a pole happens to appear too close to the real axis. The frequency convolution in Eq. (28) is either performed analytically with a Padé approximant again for the matrix elements of the screened interaction or by a cubic spline interpolation (see Ref. 54 for details).

After finding the parameters a_{ν} and ω_{ν}, the correlation self-energy is given by the analytical function

$$\langle\varphi_{n\mathbf{q}}^{\sigma}|\Sigma_{\mathrm{c}}^{\sigma}(\omega)|\varphi_{n\mathbf{q}}^{\sigma}\rangle = \begin{cases} f(\omega) & \text{for Re }\omega\geq 0 \\ f^*(\omega^*) & \text{for Re }\omega < 0 \end{cases}, \tag{30}$$

which allows to solve the nonlinear quasiparticle Eq. (8) to machine precision with the standard iterative Newton method and without any additional linearization of the self-energy.

3.4 Tricks to reduce computational cost

As already mentioned, the explicit orbital dependence and the nonlocality in the effective potential makes DFT calculations with hybrid functionals very expensive. For the same reason and also because of the additional frequency dependence, the same is true for *GW* calculations. In this section, we discuss several numerical tricks that can be applied to reduce the computational expense of these methods. We start with those that can be applied equally to both methods.

Sparsity of Coulomb matrix The Coulomb matrix in the MPB representation, Eq. (20), consists of four distinct blocks, the diagonal parts MT-MT and IPW-IPW as well as the two off-diagonal parts MT-IPW and IPW-MT. Their evaluation was discussed in detail in Ref. 61. The vector-matrix-vector products in Eq. (19), where the wave function products are regarded

as vectors in the MPB, must be evaluated for each combination of band indices n, n', and n'' as well as Bloch vectors $\mathbf{k}$ and $\mathbf{q}$. This easily amounts to billion matrix operations or more and constitutes the computationally most expensive step in the algorithm. Fortunately, we can accelerate these matrix operations considerably, by making the Coulomb matrix sparse by a suitable unitary transformation of the MPB MT functions. These transformations are done within the subspaces of each atom and LM channel, so it involves only few basis functions per subspace. According to Eq. (20), the matrix elements $v_{IJ}(\mathbf{q})$ can be regarded as the electrostatic energy of two charge distributions described by MPB functions. So, it makes sense to define a formal multipole moment for the MPB functions in the muffin-tin spheres

$$\mu_{LP}^a = \int_0^{S_a} M_{LP}^a(r)\, r^{L+2} dr\,. \tag{31}$$

Now, by forming linear combinations

$$M_{L1}^{a\prime}(r) = \frac{\mu_{L1}^a M_{L1}^a(r) + \mu_{Ln}^a M_{Ln}^a(r)}{\sqrt{\mu_{L1}^{a\,2} + \mu_{Ln}^{a\,2}}} \tag{32a}$$

$$M_{Ln}^{a\prime}(r) = \frac{\mu_{Ln}^a M_{L1}^a(r) - \mu_{L1}^a M_{Ln}^a(r)}{\sqrt{\mu_{L1}^{a\,2} + \mu_{Ln}^{a\,2}}}\,, \tag{32b}$$

we can make the multipole moment of the nth function vanish. By a successive application of this procedure in each subspace, only one of the resulting functions (the first) will have a nonzero multipole moment. The "charge distributions" represented by all other MT functions do not generate an electrostatic potential outside the MT spheres so that Coulomb matrix elements involving an element of this set can only be nonzero, if the other function is a MT function residing in the same MT sphere; all other matrix elements must vanish. This leads to a very sparse form of the Coulomb matrix, which drastically reduces the number of floating point operations needed in evaluating Eq. (19) and thus the computational cost.

Truncation of matrices The computational cost can be reduced significantly by introducing a cutoff for the basis representation of matrices without compromising the accuracy. Let us start with the nonlocal potential for the hybrid functionals. In Eq. (17) we have represented the exchange operator in terms of the wave functions rather than the LAPW basis functions. Although important, the exchange energy is a relatively small energy contribution compared to kinetic and potential energies. Therefore, we can afford to describe the nonlocal exchange potential in a subspace of wave functions up to a band cutoff $n_{\max}$, all other matrix elements are set to zero. We find that the results converge reasonably fast with the cutoff $n_{\max}$.

Another cutoff can be introduced in the calculation of the GW correlation self-energy. The contribution of Eq. (7) is smaller than the exchange self-energy, Eq. (6), which allows to use a reduced set $\{E_\mu^{\mathbf{k}}(\mathbf{r})\}$ for the correlation part. If we assume that the eigenvalues $v_\mu(\mathbf{k})$ are ordered according to decreasing size, then matrix elements $\varepsilon_{\mu\nu}(\mathbf{k}, \omega)$ and $W_{\mu\nu}(\mathbf{k}, \omega)$ with large indices will be relatively small, cf. Eqs. (22) and (23). We may then introduce a threshold value $v_{\min}$ for the eigenvalues and only retain the functions $E_\mu^{\mathbf{k}}(\mathbf{r})$ with $v_\mu(\mathbf{k}) \geq v_{\min}$. We thus restrict ourselves to the dominant part of the electron-electron interaction. The removal of basis functions with small eigenvalues can be viewed as an optimization step for the MPB, because

it reduces the matrix sizes and hence the computational cost. Note that with $v_{\min} = 0$, the full accuracy of the MPB is restored. In our implementation, this optimization of the MPB only affects the correlation self-energy while we calculate the exchange self-energy always with the full MPB.

Symmetry Spatial and time reversal symmetries are exploited to accelerate the code in three ways: (1) Inversion symmetry leads to real-valued quantities such as the Coulomb matrix $v_{IJ}(\mathbf{q})$, the polarization matrix $P_{IJ}(\mathbf{q}, i\omega)$ and the screened interaction $W_{IJ}(\mathbf{q}, i\omega)$ for $\omega \in \mathbb{R}$. (2) The $\mathbf{q}$ and $\mathbf{k}$ vector in Eq. (17) can be chosen from a reduced set of Bloch vectors, i.e., from the irreducible wedge of the BZ and from an *extended* irreducible BZ, respectively. And (3), if the wave functions $\varphi_{n\mathbf{k}}^{\sigma}$ and $\varphi_{n'\mathbf{k}}^{\sigma}$ in Eq. (17) fall into different irreducible representations, the corresponding exchange matrix element vanishes. This can be used as a criterion whether an element must be calculated or not.

In general, the matrix representation of real operators (e.g., the Coulomb interaction) in an arbitrary complex basis $\{f_\mu(\mathbf{r})\}$ is Hermitian. If the system exhibits inversion symmetry and the basis functions fulfill $f_\mu(-\mathbf{r}) = f_\mu^*(\mathbf{r})$, the matrices become real symmetric. Similarly, the representation of wave-function products $\langle M_{\mathbf{q}I} \varphi_{n\mathbf{k}}^{\sigma} | \varphi_{n'\mathbf{k+q}}^{\sigma} \rangle$, which appear in nearly all equations, are then real instead of complex. Of course, this reduces the computational cost considerably, in terms of both memory consumption and computation time. However, according to the current definition only the IPWs fulfill the above condition, while the MT functions do not. We hence apply a unitary transformation of the MT functions such that the condition is satisfied for the whole basis [66]. From the form of the condition it is clear that the transformation forms linear combinations of MT functions that reside in different, but symmetry-equivalent MT spheres, resulting in functions that resemble bonding and antibonding states. (If the atom is placed in the origin, linear combinations of different magnetic quantum numbers are formed.) In the following, it is understood that all quantities are represented in this symmetrized basis if inversion symmetry is present. We note that this symmetrization leaves the sparsity of the Coulomb matrix intact.

In a reciprocal space formulation, the nonlocal form of the interaction potentials, the *GW* self-energy, etc. gives rise to computationally expensive double BZ summations. For example, the nonlocal potential [Eq. (17)] at a point $\mathbf{q}$ does not depend only on quantities at that point in reciprocal space but also on the Coulomb matrix and wave functions at all other points in the BZ. Fortunately, if the system is spatially invariant under symmetry operations, we can reduce the computational expense considerably. Here, invariance means that a given symmetry operation Q commutes with the Hamiltonian H, i.e., $QH = HQ$, such that $Q\varphi$ is an eigenstate of H if φ is an eigenstate. Hence, the solutions of the one-particle equations at two different $\mathbf{q}$ points are equivalent, if the $\mathbf{q}$ vectors are related by a symmetry operation. This can be used to restrict the set of $\mathbf{q}$ points, at which the Hamiltonian must be diagonalized, to a smaller set, the so-called irreducible Brillouin zone (IBZ). Of course, this is nothing new, as the IBZ is routinely employed in computer codes that are used to treat systems with periodic boundary conditions. However, we can restrict the additional $\mathbf{k}$ summations in the calculation of the nonlocal quantities in a similar way, with the restriction that, due to the additional dependence on $\mathbf{k}$ and $\mathbf{q}-\mathbf{k}$, only those symmetry operations $Q_i^{\mathbf{q}}$ can be employed that leave the given $\mathbf{q}$

vector invariant, i.e., $Q_i^{\mathbf{q}}\mathbf{q} = \mathbf{q} + \mathbf{G}_i^{\mathbf{q}}$, where $\mathbf{G}_i^{\mathbf{q}}$ is a reciprocal lattice vector that folds the vector back into the BZ if $Q_i^{\mathbf{q}}\mathbf{q}$ is pointing out of it. This subset of operations $\{Q_i^{\mathbf{q}}\}$ is commonly called *little group* LG($\mathbf{q}$). In the same way as for the IBZ, the little group gives rise to a minimal set of inequivalent $\mathbf{k}$ points, which we denote by the extended IBZ [EIBZ($\mathbf{q}$)]. Depending on the particular quantity, the strategies to apply the EIBZ($\mathbf{q}$) are different. The basic strategy is to replace the summation over $\mathbf{k}$ by a summation over the EIBZ($\mathbf{q}$) and over the symmetry operations. The latter is performed such that it produces only little overhead so that the net computation time is much faster than if we summed over the whole BZ. In the following, we only give the final expressions and give additional explanations where necessary. For simplicity, non-symmorphic and time-reversal symmetry operations are omitted. The formulas can be generalized easily to take these operations into account as well (cmp. Refs. 39 and 54).

We first apply the extended IBZ to the polarization function defined in Eq. (2) and represented in the MPB. The final expression is

$$P_{IJ}(\mathbf{k},\omega) = \sum_{i=1}^{N_A^{\mathbf{k}}} \sum_{\sigma} \sum_{\mathbf{q}}^{\text{EIBZ}(\mathbf{k})} \frac{1}{N_{\text{LG}(\mathbf{k},\mathbf{q})}} \sum_{n}^{\text{occ}} \sum_{n'}^{\text{unocc}} \langle Q_i^{\mathbf{k}} \tilde{M}_I^{\mathbf{k}} \varphi_{n\mathbf{q}}^{\sigma} | \varphi_{n'\mathbf{q}+\mathbf{k}}^{\sigma} \rangle \langle \varphi_{n'\mathbf{q}+\mathbf{k}}^{\sigma} | \varphi_{n\mathbf{q}}^{\sigma} Q_i^{\mathbf{k}} \tilde{M}_J^{\mathbf{k}} \rangle (\ldots) ,$$

(33)

where the factor $1/N_{\text{LG}(\mathbf{k},\mathbf{q})}$ avoids double counting of $\mathbf{q}$ points. $N_{\text{LG}(\mathbf{k},\mathbf{q})}$ is the number of symmetry operations that are members of both LG($\mathbf{k}$) and LG($\mathbf{q}$). The application of a symmetry operation to a MPB function can be written as a linear combination of MPB basis functions so that the sum over the symmetry operations can be performed at the very end after summing over the bands, the $\mathbf{q}$ points, and the electron spin. We note that this is also possible with the set $\{E_\mu^{\mathbf{k}}(\mathbf{r})\}$ instead of $\{M_I^{\mathbf{k}}(\mathbf{r})\}$. In a similar way, we can accelerate the evaluation of the expectation values of Σ_x^{σ} and $\Sigma_c^{\sigma}(\omega)$. To this end, we write Eqs. (6) and (7) in a general form with a (possibly frequency-dependent) function $f(\mathbf{r}, \mathbf{r}')$, which fulfills all symmetry properties of the system. Using the great orthogonality theorem of group theory [67, 54], we can then write

$$\begin{aligned}
\langle \varphi_{n\mathbf{q}}^{\sigma} | \Sigma^{\sigma} | \varphi_{n\mathbf{q}}^{\sigma} \rangle &= \sum_{\mathbf{k}}^{\text{BZ}} \sum_{n'} \iint d^3r\, d^3r'\, \varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r})\, \varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r})\, \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}')\, \varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r}')\, f(\mathbf{r},\mathbf{r}') \\
&= \sum_{\mathbf{k}}^{\text{EIBZ}(\mathbf{q})} \frac{N_{\text{LG}(\mathbf{q})}}{N_{\text{LG}(\mathbf{k},\mathbf{q})}(n_2 - n_1 + 1)} \sum_{m=n_1}^{n_2} \sum_{n'} \iint d^3r\, d^3r'\, \varphi_{m\mathbf{q}}^{\sigma*}(\mathbf{r})\, \varphi_{n'\mathbf{k}}^{\sigma}(\mathbf{r}) \\
&\quad \times \varphi_{n'\mathbf{k}}^{\sigma*}(\mathbf{r}')\, \varphi_{m\mathbf{q}}^{\sigma}(\mathbf{r}')\, f(\mathbf{r},\mathbf{r}')
\end{aligned}$$

(34)

where the wave functions $\{\varphi_{m\mathbf{q}}^{\sigma}(\mathbf{r}), n_1 \leq m \leq n_2\}$ form an eigenspace with $n_1 \leq n \leq n_2$ and $N_{\text{LG}(\mathbf{q})}$ is the number of elements in LG($\mathbf{q}$). Instead of summing over the symmetry operations, we average here over the degenerate states of which $\varphi_{n\mathbf{q}}^{\sigma}(\mathbf{r})$ is a member. This is the method of choice if only few diagonal elements must be calculated. For the hybrid functionals, however, we have to calculate a whole matrix according to Eq. (17). Here, we use a different strategy, which is akin to Eq. (33): we calculate a first matrix by summing only over the EIBZ($\mathbf{q}$) and then let all symmetry operations of LG($\mathbf{q}$) act on this matrix and sum up, which adds the contribution of all other $\mathbf{k}$ points. This time, the symmetry operations act on a matrix represented in the LAPW basis, which is particularly simple because it can be written as a one-to-one mapping,

which takes negligible computation time. With the definitions of above,

$$
\begin{aligned}
V_{\mathrm{x},\mathbf{GG}'}^{\mathrm{NL},\sigma}(\mathbf{q}) \; = \; & -\sum_i^{\mathrm{LG}(\mathbf{q})}\sum_{\mathbf{k}}^{\mathrm{EIBZ}(\mathbf{q})}\frac{1}{N_{\mathrm{LG}(\mathbf{k},\mathbf{q})}}\sum_n^{\mathrm{occ.}}\iint \chi_{\mathbf{q}(P_i^{\mathbf{q}}\mathbf{G}+\mathbf{G}_i^{\mathbf{q}})}^{\sigma*}(\mathbf{r})\varphi_{n\mathbf{k}}^{\sigma}(\mathbf{r})v(|\mathbf{r}-\mathbf{r}'|) \\
& \times \varphi_{n\mathbf{k}}^{\sigma*}(\mathbf{r}')\chi_{\mathbf{q}(P_i^{\mathbf{q}}\mathbf{G}'+\mathbf{G}_i^{\mathbf{q}})}^{\sigma}(\mathbf{r}')\,d^3r\,d^3r' \,.
\end{aligned}
\tag{35}
$$

Local orbitals transform in a similar way.

Finally, we can employ the great orthogonality theorem of group theory to reduce the cost of evaluating the exchange matrix in Eq. (17). This theorem demands that the matrix elements can only be nonzero, if the wave functions are members of the same irreducible representation. Since the irreducible representations are not available in our DFT code (and their evaluation in each iteration would be computationally expensive), we exploit the fact that the great orthogonality theorem applies to *any* operator that has the full symmetry of the system. A suitable operator is given by the MT step function $\Theta^{\mathrm{MT}}(\mathbf{r})$, which is one in the MT spheres and zero elsewhere. The calculation of its matrix elements $\langle \varphi_{n\mathbf{k}}^{\sigma}|\Theta^{\mathrm{MT}}|\varphi_{n'\mathbf{k}}^{\sigma}\rangle$ is elementary and takes negligible CPU time. These matrix elements can be used as a *probe* in the sense that we only calculate those exchange matrix elements for which the corresponding Θ^{MT} matrix elements are nonzero.

Treatment of Γ point   The Γ point, i.e, $\mathbf{k}=\mathbf{0}$, is a special point in Eqs. (19), (22), and (23), because the long-range nature of the Coulomb interaction makes the Coulomb matrix $v_{IJ}(\mathbf{k})$ and also the screened interaction $W_{\mu\nu}(\mathbf{k},\omega)$ diverge in the limit $\mathbf{k}\to\mathbf{0}$. A proper treatment of this divergence is crucial to obtain a fast and stable convergence with respect to the $\mathbf{k}$-point sampling.

We start with Eq. (19). The integrand contains the Coulomb matrix, whose leading element is proportional to $1/k^2$ in the limit $\mathbf{k}\to\mathbf{0}$. Therefore, the three-dimensional integration over the BZ yields a finite value. We find that the contribution of the divergence is numerically important so that we cannot simply exclude the problematic Γ point from the $\mathbf{k}$-point set. Instead, the Coulomb matrix $v_{IJ}(\mathbf{k})$ is decomposed into a divergent and a nondivergent part [61]

$$
v_{IJ}(\mathbf{k}) = \frac{4\pi}{V}\frac{1}{k^2}\langle \tilde{M}_I^{\mathbf{k}}|e^{i\mathbf{k}\cdot\mathbf{r}}\rangle\langle e^{i\mathbf{k}\cdot\mathbf{r}}|\tilde{M}_J^{\mathbf{k}}\rangle + v_{IJ}'(\mathbf{k}) \,.
\tag{36}
$$

The long-wavelength limit of the non-divergent second term replaces the matrix $v_{IJ}(\mathbf{0})$ in Eq. (19), which allows to perform the $\mathbf{k}$ summation numerically. The divergent first term (since the MPB contains the constant function, this term is given exactly, and we may calculate with $e^{i\mathbf{k}\cdot\mathbf{r}}$) gives rise to a contribution

$$
V_{\mathrm{x},nn'}^{\mathrm{NL},\sigma}(\mathbf{q})\Big|_{\mathrm{div}} = -\frac{1}{2\pi^2}\left(\sum_{n''}^{\mathrm{occ.}}\int_{\mathrm{BZ}}\langle\varphi_{n\mathbf{q}}^{\sigma}|\varphi_{n''\mathbf{q}-\mathbf{k}}^{\sigma}e^{i\mathbf{k}\cdot\mathbf{r}}\rangle\frac{1}{k^2}\langle e^{i\mathbf{k}\cdot\mathbf{r}}\varphi_{n''\mathbf{q}-\mathbf{k}}^{\sigma}|\varphi_{n'\mathbf{q}}^{\sigma}\rangle d^3k - \mathrm{d.c.}\right),
\tag{37}
$$

where d.c. is a double-counting correction to be defined later. For the important region close to $\mathbf{k}=\mathbf{0}$ we can replace $\langle\cdot|\cdot\rangle$ by $\delta_{nn''}$ and $\delta_{n'n''}$, respectively, and leave out higher-order corrections [39] for simplicity. Even though this simplifies the integral to $\int_{\mathrm{BZ}}d^3k/k^2$ the $\mathbf{k}$ integration can still not be done analytically because the integration region is bound by the complicated BZ boundary. Therefore, we employ a method that is similar to the one proposed by Massidda *et*

al. in Ref. 40. We replace $1/k^2$ by the function

$$F(\mathbf{k}) = \sum_{\mathbf{G}} \frac{e^{-\beta|\mathbf{k}+\mathbf{G}|^2}}{|\mathbf{k}+\mathbf{G}|^2} , \tag{38}$$

which is defined in the whole reciprocal space. It contains as many terms as the crystal contains BZs. Therefore, we can write

$$V_{\mathrm{x},nn'}^{\mathrm{NL},\sigma}(\mathbf{q})\Big|_{\mathrm{div}} = -\delta_{nn'} f_{n\mathbf{q}}^{\sigma} \left(\frac{1}{2\pi^2} \int \frac{e^{-\beta|\mathbf{k}|^2}}{k^2} d^3k - \frac{1}{N_{\mathbf{k}}\Omega} \sum_{k\neq 0} \frac{e^{-\beta|\mathbf{k}|^2}}{k^2} \right) , \tag{39}$$

where the parameter $\beta > 0$ ensures that the integral over the whole reciprocal space is finite and Ω is the unit-cell volume. In contrast to Massidda *et al.*, we choose this parameter as small as possible so that the function remains close to $1/k^2$. (This also ensures that calculations of the same system with different choices of the unit cell yield identical values.) Here, we have defined the double-counting term. $N_{\mathbf{k}}$ denotes the number of $\mathbf{k}$ points and $f_{n\mathbf{k}}^{\sigma}$ is the occupation number. The contribution of large $\mathbf{k}$ vectors tend to cancel in the difference of Eq. (39) so that we may introduce a reciprocal cutoff radius k_0 and finally obtain

$$V_{\mathrm{x},nn'}^{\mathrm{NL},\sigma}(\mathbf{q})\Big|_{\mathrm{div}} = -\delta_{nn'} f_{n\mathbf{q}}^{\sigma} \left(\frac{1}{\sqrt{\pi\beta}} \mathrm{erf}\left(\sqrt{\beta}k_0\right) - \frac{1}{N_{\mathbf{k}}\Omega} \sum_{0<k\leq k_0} \frac{e^{-\beta|\mathbf{k}|^2}}{k^2} \right) . \tag{40}$$

By setting $e^{-\beta k_0^2} = \beta$, we are left with β as the only convergence parameter, which just has to be chosen small enough; $\beta = 0.005$ is a good choice.

Fig. 1 demonstrates the performance of the Γ-point treatment for the exchange energy $E_{\mathrm{x}}^{\mathrm{NL}} = 2\sum_{n\mathbf{q}}^{\mathrm{occ.}} V_{\mathrm{x},nn}^{\mathrm{NL}}(\mathbf{q})$ of NaCl. While the contributions from the divergent term, Eq. (40), and the remainder separately converge poorly with respect to the $\mathbf{k}$-point sampling, their sum nearly look constant on the energy scale of the diagram. The $\mathbf{k}$-point convergence can be improved further by taking corrections at $\mathbf{k} = \mathbf{0}$ into account that arise from multiplying $1/k^2$ with second-order terms of $\langle\cdot|\cdot\rangle\langle\cdot|\cdot\rangle$ [39].

We note that the Fourier transform in Eq. (24) diverges as $1/k^2$ in the limit $\mathbf{k} \to \mathbf{0}$. The same divergence is found for the bare Coulomb potential [39, 61], such that the $1/k^2$ terms cancel in the difference. The remainder is finite and is given by

$$\lim_{\mathbf{k}\to\mathbf{0}} \frac{4\pi}{k^2} \left(1 - e^{-k^2/(4\kappa^2)}\right) = \frac{\pi}{\kappa^2} . \tag{41}$$

In Fig. 2, we show that this nondivergent behavior of the short-range interaction gives rise to a favorable $\mathbf{k}$-point convergence, here for the example of the silicon band gap. Within HSE the values converge nearly as fast as in PBE, whereas the PBE0 functional requires larger $\mathbf{k}$-point meshes. This was already observed in Ref. 68.

In the calculation of the dielectric matrix [Eq. (23)], we multiply the polarization function with the Coulomb matrix, and in the limit $\mathbf{k} \to \mathbf{0}$ we must again deal with the k^{-2} Coulomb divergence. Its treatment is simplified by the basis transformation $\{M_I^{\mathbf{k}}(\mathbf{r})\} \to \{E_\mu^{\mathbf{k}}(\mathbf{r})\}$ introduced above because this transformation confines the divergence of the Coulomb matrix to a single eigenvalue $v_1(\mathbf{k}) \sim 4\pi/k^2$ [61]. The corresponding eigenfunction $E_1^{\mathbf{k}}(\mathbf{r}) \sim e^{i\mathbf{k}\cdot\mathbf{r}}/\sqrt{V}$ is known analytically. The situation is then analogous to a pure plane-wave basis, but we emphasize that,
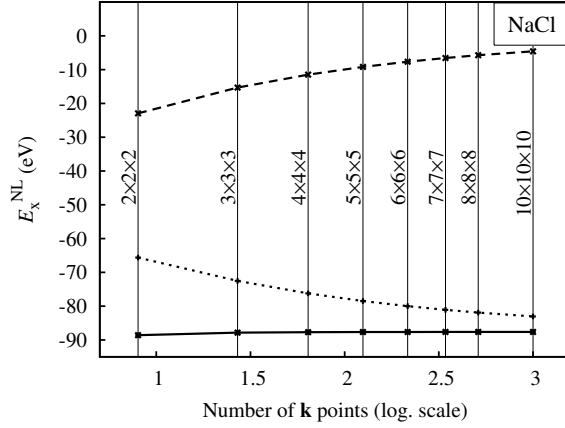
Figure 1: Exchange energy as a function of the **k**-point mesh for NaCl. The dashed and dotted curves correspond to the contributions of the divergent term Eq. (40) and the numerical sum, respectively. Their sum shown by the solid curve converges very quickly. [39]
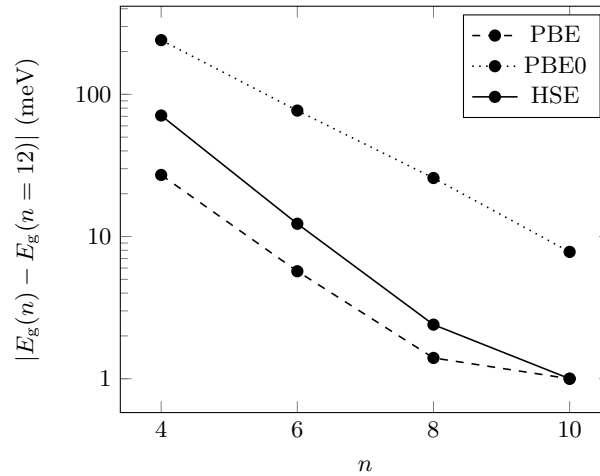

Figure 2: Convergence of the silicon band gap for the functionals PBE (dashed), PBE0 (dotted), and HSE (solid) with respect to the **k**-point mesh $(n \times n \times n)$. [69]

since $\{E_\mu^{\mathbf{k}}(\mathbf{r})\}$ has been transformed from the MPB, we still retain the all-electron description without a projection onto plane waves. Using $\mathbf{k} \cdot \mathbf{p}$ perturbation theory, we can show that the head ($\mu = \nu = 1$) and wing elements ($\mu \neq \nu = 1$ and $1 = \mu \neq \nu$) of the polarization matrix $P_{\mu\nu}(\mathbf{k}, \omega)$ [Eq. (33)] behave as k^2 and k, respectively, so that the resulting dielectric matrix $\varepsilon(\mathbf{k}, \omega)$ remains finite albeit angular dependent at $\mathbf{k} = \mathbf{0}$ [70, 71]. This form of the dielectric matrix gives rise to angular and energy dependent head and wing elements that diverge as k^{-2} and k^{-1}, respectively, in the $\mathbf{k} \to \mathbf{0}$ limit of the screened interaction $W_{\mu\nu}(\mathbf{k}, \omega)$ [Eq. (22)], which consists of a constant and an angular dependent part. Without a derivation we only give the final expression here

$$
W_{\mu\nu}(\mathbf{k}, i\omega) \sim \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \tilde{W}_{22}(\omega) & \cdots & \tilde{W}_{2n}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \tilde{W}_{n2}(\omega) & \cdots & \tilde{W}_{nn}(\omega) \end{pmatrix}
$$

$$
+ \frac{1}{\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{L}(\omega)\hat{\mathbf{k}}} \begin{pmatrix} 4\pi/k^2 & \hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_2(\omega)/k & \cdots & \hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_n(\omega)/k \\ \hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_2^*(\omega)/k & \left|\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_2(\omega)\right|^2 & \cdots & \left[\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_2^*(\omega)\right]\left[\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_n(\omega)\right] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_n^*(\omega)/k & \left[\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_n^*(\omega)\right]\left[\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_2(\omega)\right] & \cdots & \left|\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{w}_n(\omega)\right|^2 \end{pmatrix} ,
$$

(42)

were $\mathbf{L}(\omega)$ is a 3×3 matrix, $\mathbf{w}_\mu(\omega)$ are three-dimensional vectors, and $\hat{\mathbf{k}}$ is a unit Bloch vector denoting the direction from which the limit $\mathbf{k} \to \mathbf{0}$ is taken. The appearance of the divergent wing elements and the angular dependence makes the evaluation of the correlation self-energy more complicated than in the case of the exchange self-energy. For the angular dependence, we choose an expansion in terms of spherical harmonics. (This expansion contains infinitely many terms if $\mathbf{L}(\omega)$ is not a multiple of the unit matrix, but the expansion usually converges rapidly. Furthermore, it can be calculated analytically by solving a system of linear equations.) Once the expansion is found, angular integration over the point $\mathbf{k} = \mathbf{0}$ is straightforward, $\int Y_{lm}(\hat{\mathbf{k}})d\hat{\mathbf{k}} = \delta_{l0}$. The numerically most important contribution derives from the k^{-2} divergence, which is treated with the same strategy as in the case of the exchange self-energy. For further details we refer the reader to our original publication [54].

Let us now turn to the case of a metallic system, where in addition to the interband transitions with $n \neq n'$ there is a contribution from intraband transitions across the Fermi surface. These transitions occur within one electron band, i.e., $n = n'$, and give rise to a contribution to the head element of the polarization function, the so-called Drude term [72]

$$
P_{11}^{\mathrm{D}}(\mathbf{k}, i\omega) \sim -\frac{k^2}{4\pi} \frac{\omega_{\mathrm{pl}}^2}{\omega(\omega + i\eta)} ,
$$

(43)

where ω_{pl} is the plasma frequency obtained by an integration over the Fermi surface. The main contribution to the screened interaction is given by the *bare* Drude term

$$
W_{11}^{\mathrm{c,D}}(\mathbf{k}, i\omega) \sim -\frac{4\pi}{k^2} \frac{\omega_{\mathrm{pl}}^2}{\omega^2 + \omega_{\mathrm{p}}^2} .
$$

(44)

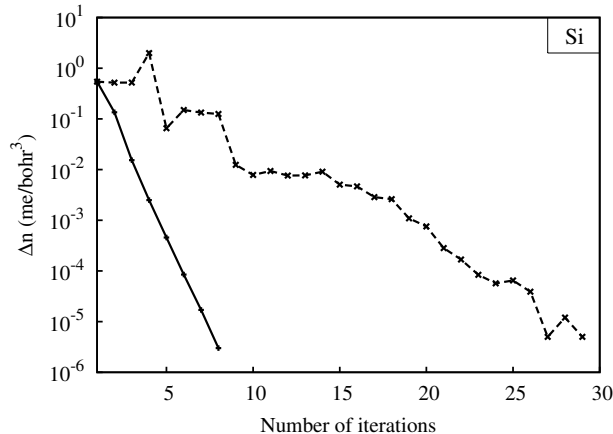This expression can be convoluted with the Green function in Eq. (28) analytically.

Figure 3: Convergence behavior of the electron density for Si in a self-consistent-field cycle. The solid and dashed curves correspond to calculations with and without the nested density convergence scheme. [39]

Nested density convergence   For hybrid functional calculations one must solve the one-particle equation [Eq. (16)] self-consistently, meaning that the input density that generates the local and nonlocal potential coincides with the output density formed by the wave functions that are solutions of the one-particle equation. It is known that already in conventional DFT calculations with purely local functionals a density mixing scheme must be employed that combines the current density with several previous densities, as for example in the standard simple-mixing and Broyden-mixing schemes [73, 74]. However, in addition to the local effective potential, Eq. (16) contains a nonlocal potential that depends on the density matrix, for which no simple mixing procedure is available. Indeed, a standard density-only mixing (without a mixing of the density matrix) leads to poor convergence: 27 iterations for Si, as illustrated in Fig. 3, and more than 200 iterations for $SrTiO_3$ are necessary. Rather than constructing a complicated mixing scheme for the density matrix, we choose a pragmatic approach: The self-consistency cycle is divided into an outer loop, where the density and the density matrix is updated, and an inner loop, where only the density is converged while the density matrix is kept fixed. With this nested iterative procedure the *outer* loop converges after eight steps for Si, see Fig. 3, and after only twelve steps for $SrTiO_3$. One iteration of the *inner* loop lasts only 1.0 s for Si and 8.3 s for $SrTiO_3$ on a single Intel Xeon X5355 at 2.66 GHz (Cache 4 MB) using a 4×4×4 **k**-point set. This is small compared to the cost for the construction of the nonlocal potential in the *outer* loop, which takes 11.9 s for Si and 573.1 s for $SrTiO_3$.

4   Results

We have implemented the hybrid functionals PBE0 and HSE in the FLEUR program package [75] and the GW approximation as the SPEX code [54]. In addition to the radial functions $u_{lp}^{a\sigma}(r)$ with $p = 0, 1$, corresponding to the solution of the radial Kohn-Sham equation with the local PBE potential and its energy derivative, respectively, we employ local orbitals with $p \geq 2$ to describe semi-core states as well as high-lying conduction states if necessary [76, 77]. In the

hybrid-functional calculations presented here, the core states are taken from a preceding PBE calculation and kept fixed during the self-consistent-field cycle.

In the following sections we discuss the convergence of the MPB (Sec. 4.1), results for simple materials, comparison to other codes, and timings (Sec. 4.2), as well as results for GdN (Sec. 4.3), and ZnO (Sec. 4.4).

4.1 Test calculations

In this section we analyze the performance of the MPB, in particular its convergence behavior with respect to the parameters $L_{\max}$, $G'_{\max}$, and $n_{\max}$ for single-particle excitation energies and total energy differences. We only show results for the PBE0 functional. A similar performance of the MPB is found for the HSE functional and the GW approximation. Figures 4(a) and (b) show excitation energies of Si ($\Gamma_{25'v} \rightarrow \Gamma_{15c}$ and $\Gamma_{25'v} \rightarrow X_{1c}$) and SrTiO$_3$ ($\Gamma_{15v} \rightarrow \Gamma_{25'c}$ and R$_{15'v} \rightarrow \Gamma_{25'c}$) obtained from the self-consistent PBE0 solution as functions of the MPB parameters. The diagrams show that the convergence of these transition energies to within $0.01\,\text{eV}$ is achieved for $G'_{\max} = 2.0\,\text{bohr}^{-1}$ and $G'_{\max} = 2.7\,\text{bohr}^{-1}$ for Si and SrTiO$_3$, respectively. This is not only below the exact limit $G'_{\max} = 2G_{\max}$ for the wave-function products, it is even below the reciprocal cutoff radius $G_{\max}$ for the wave functions themselves ($G_{\max} = 3.6\,\text{bohr}^{-1}$ for Si and $G_{\max} = 4.3\,\text{bohr}^{-1}$ for SrTiO$_3$). Similarly, we find an angular-momentum cutoff value of $L_{\max} = 4$ to be sufficient, while a proper matching of the wave functions at the MT boundaries requires a much larger value of $l_{\max} = 8$. Furthermore, the figures show that the number of bands, in which the exchange potential is represented, can be cut at $n_{\max} = 100$ for Si and at $n_{\max} = 250$ for SrTiO$_3$, corresponding to about 50 bands per atom.

Figure 4(c) shows that the total energy difference between the diamond and wurtzite structure of Si converges even faster than the transition energies above. With a reciprocal cutoff radius of $G'_{\max} = 2.25\,\text{bohr}^{-1}$, an angular-momentum cutoff of $L_{\max} = 4$, and 20 bands per atom we achieve an accuracy of $1\,\text{meV}$, which is well below the error resulting from the BZ discretization. With the finer 8×8×8 **k**-point mesh, the total-energy difference is converged to within $2\,\text{meV}$ and amounts to $112\,\text{meV}$, i.e., the diamond structure is lower in energy than the wurtzite structure. This is very close to the value of $92\,\text{meV}$ obtained with the PBE functional.

In summary, we find that $G'_{\max}$ can be chosen universally smaller than $G_{\max}$, $G'_{\max} = 0.75\,G_{\max}$ as a rule of thumb, while the cutoff parameter $L_{\max}$ is more material-specific. If the material contains occupied d or f electrons, we have to use $L_{\max} = 5$ or even $L_{\max} = 6$. Also, the band cutoff $n_{\max}$ is material-specific.

In GW calculations we also have to deal with a band cutoff $n_{\max}$ that determines the number of bands that are taken into account in the calculation of the polarization function (2) and the correlation self-energy (10), because the number of available eigenstates is limited by the size of the basis set. Furthermore, the computational cost depends strongly on the number of unoccupied states. It is well known that a proper convergence of the GW quasiparticle energies requires very many unoccupied states [78, 79, 80, 81]. While for SrTiO$_3$, a relatively modest number of 200 states is sufficient, the band gaps of Si, in particular, the indirect one, are more difficult to converge. We will see in Sec. 4.4 that ZnO is an extreme case in this respect. A
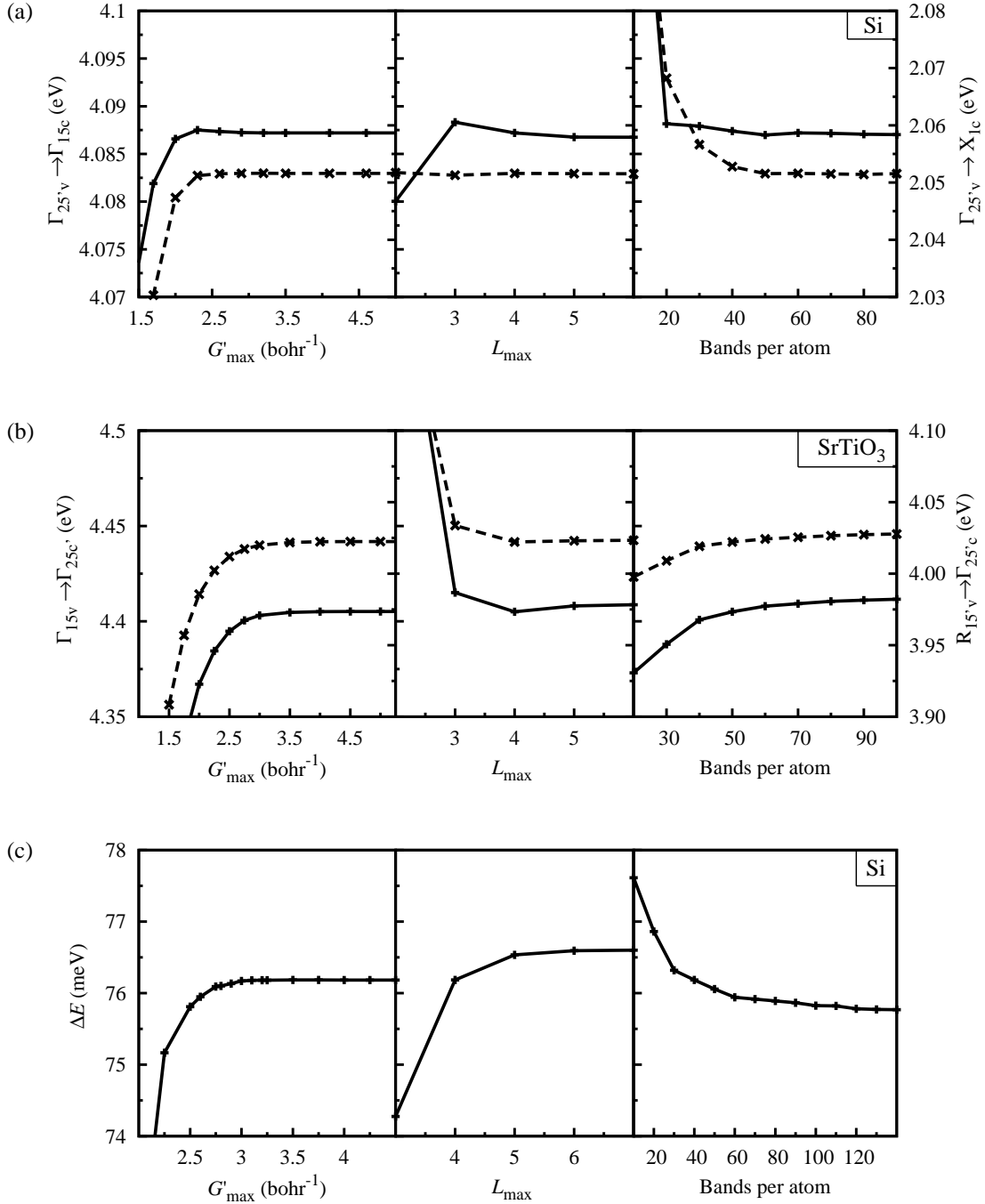
Figure 4: Convergence of single-particle transition energies for (a) Si and (b) SrTiO$_3$ with respect to the MPB parameters (solid line: left scale, dashed line: right scale). (c) Convergence of the total energy difference between the diamond and wurtzite phase of Si with respect to the same parameters. For these test calculations we have employed a $4\times4\times4$ **k**-point set. See the tables for fully converged results and the text for details. [39]
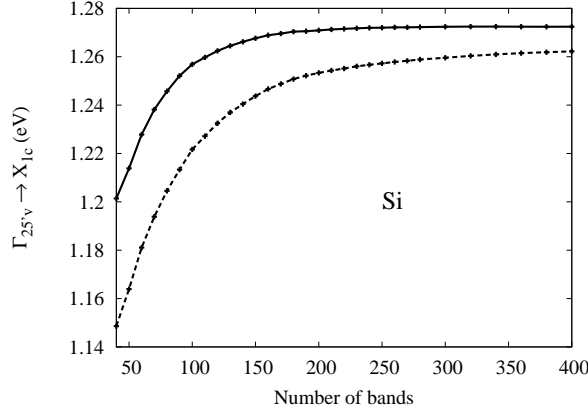
Figure 5: Convergence of the $\Gamma_{25'v} \to X_{1c}$ gap of Si with respect to the number of bands used in the construction of the polarization function as well as the self-energy (solid line). The dashed line shows the effect of the extrapolar correction. [54]
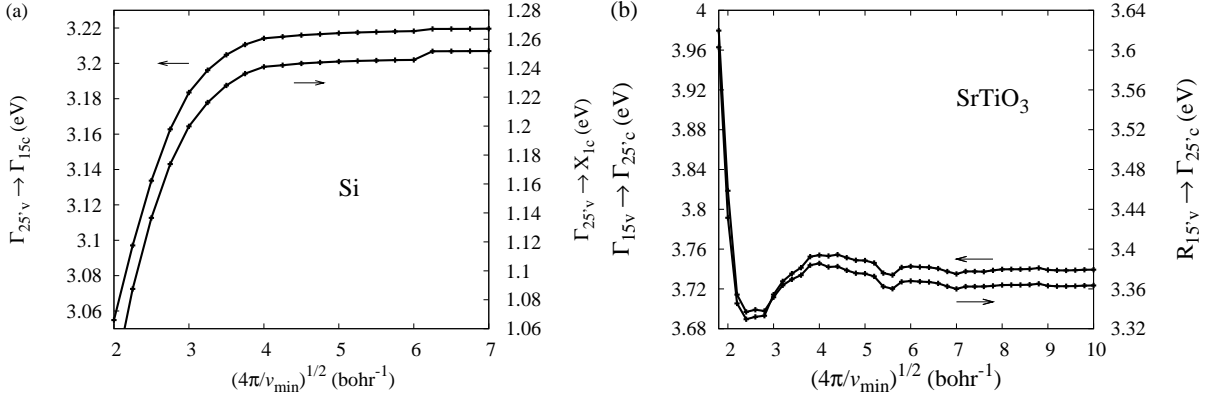

Figure 6: Convergence of single-particle transition energies of (a) Si and (b) $SrTiO_3$ as a function of $\sqrt{4\pi/v_{min}}$. [54]

possible solution to this problem is the extrapolar correction introduced by Bruneval and Gonze [82], in which the states not explicitly taken into account (i.e., $n' > n_{max}$) are placed on a fixed energy that is taken as a parameter. In this way, the frequency-dependent term becomes state independent, which allows to apply the completeness relation $\sum_{n'>n_{max}} |n'\rangle\langle n'| = 1 - \sum_{n'\leq n_{max}} |n'\rangle\langle n'|$ to eliminate the states $n' > n_{max}$. Figure 5 shows the normal convergence with respect to the number of states and also the improvement from the extrapolar approximation. However, we have found mixed success with this method. In some cases it works very well, in other cases it is less successful. Another drawback is that the determination of the additional fixed energy requires further test calculations. Therefore, all other results presented here were obtained with the conventional summation. We note that the LAPW basis is a relatively small and accurate basis for the occupied states. In order to get enough unoccupied states for GW calculations it is therefore often necessary to extend the LAPW basis by increasing the reciprocal cutoff radius G_{max} and introducing additional local orbitals.
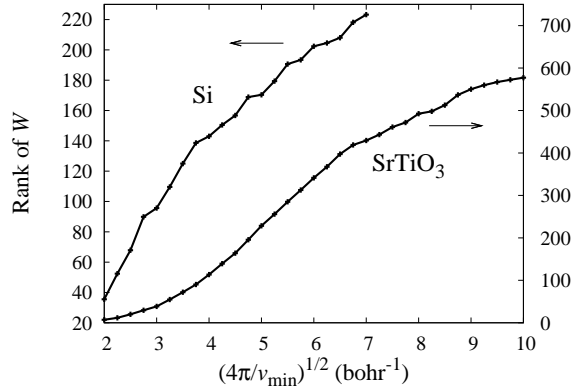
Figure 7: Rank of the matrix W for the screened interaction, i.e., the number of basis functions after the optimization, for Si and $SrTiO_3$ as a function of $\sqrt{4\pi/v_{min}}$. [54]

Geometry	Atoms	**k** mesh	Symmetry	v_{min}	CPU time
1×1×1	2	4×4×4	No	—	5 min 52 s
			Only IS	—	3 min 34 s
			Only IBZ	—	30 s
			Yes	—	11 s
			Yes	0.65	5 s
2×2×2	16	2×2×2	Yes	0.65	14 min 15 s
4×4×4	128	1×1×1	Yes	0.65	34 h 11 min

Table 1: Computational time for the calculation of quasiparticle shifts of diamond in the conventional (1×1×1) and supercell (2×2×2 and 4×4×4) geometries. For the former we also show corresponding timings without use of symmetry ("No") and with restricted use ("Only IS" and "Only IBZ") as well as the effect of the MPB optimization with a threshold value v_{min}. [54]

The computationally most expensive step in GW calculations is the evaluation of the correlation part of the self-energy, which also involves the calculation of the screened interaction W. This step can be accelerated by introducing a threshold value v_{min} for the Coulomb eigenvalues as explained in Sec. 3.4. In Fig. 6, we show the convergence of the single-particle transition energies of Si and $SrTiO_3$ with respect to $\sqrt{4\pi/v_{min}}$. The latter quantity is chosen to relate the threshold value to a reciprocal cutoff radius such as G_{max}. (In a pure plane-wave representation the minimum Coulomb eigenvalue would be $v_{min} = 4\pi/G_{max}^2$.) We observe that the energies are converged to within 0.01 eV at around $\sqrt{4\pi/v_{min}} = 3.5\,bohr^{-1}$ or $v_{min} = 1.9\,ha$. This reduces the rank of the matrix W (see Fig. 7) from 392 (the full MPB) to around 75 and the computation time from 140 to 42s on an Intel Xeon (2.66 GHz, 4 MB cache) work station. (The step-like feature at around 5.3 bohr^{-1} is caused by a sudden increase of the basis set, because there are several Coulomb eigenfunctions with the same eigenvalue of $4\pi/5.3^2\,ha$.) A similar improvement is found for $SrTiO_3$.

In order to test the efficiency of the code, we show the computational time for calculating quasiparticle shifts for diamond in the conventional unit cell (1×1×1) containing two atoms as

well as in 2×2×2 and 4×4×4 supercell geometries containing 16 and 128 atoms, respectively. The BZ is sampled with **k**-point sets having the same **k**-point density for all calculations, i.e., a 4×4×4, 2×2×2, and 1×1×1 set, respectively. The computation times on a single CPU are given in Table 1. While the *GW* calculation takes only 5 s for the conventional unit cell, even the treatment of supercells containing 16 and 128 atoms only consumes affordable 0.24 and 34.2 h of computation time, respectively. We also demonstrate the efficiency gain achieved by exploiting the symmetry and by using a threshold parameter $v_{\min}$. The different entries are as follows: "Only IS" refers to the symmetrization of the MPB due to inversion symmetry, which gives rise to real-valued quantities; "Only IBZ", only the IBZ for **q**, but not the EIBZ(**q**) for **k** is used; "Yes", spatial and time-reversal symmetries are fully exploited; the next row shows the effect of introducing a Coulomb eigenvalue threshold $v_{\min}$. With this last step the computation time has been reduced to only 5 s, gaining an overall factor of 70 with respect to the first row.

4.2 Simple materials

In this section we present HSE, PBE0, and *GW* calculations for a prototypical set of semi-conductors and insulators and compare the results with previous works from the literature [68, 52, 34, 83]. All calculations are performed at the experimental lattice constant taken from Ref. 84 with an 8×8×8 **k**-point mesh, except for the PBE0 calculations, which are performed with a 12×12×12 mesh. We focus in particular on direct and indirect band transitions. These are calculated as the energy differences of the highest occupied and the lowest unoccupied eigen-states at the corresponding points in the BZ.

In Table 2, our PBE0 and HSE results for the $\Gamma \to \Gamma$, $\Gamma \to X$, and $\Gamma \to L$ transition energies are compared with those obtained by the projector-augmented-wave (PAW) method [68] and experimental data. Overall, we observe excellent agreement between the calculated values, except for the PBE0 values for systems with wider band gaps, where we find larger discrepancies. This is a basis-set effect caused by the fact that only two radial functions per lm channel have been used for each MT sphere. For the PBE and HSE calculations, which were performed later, we have overcome this slight inaccuracy by including local orbitals (in addition to those that are already in use for the semicore states). In all cases the admixture of the nonlocal exchange potential leads to an increase of the transition energies in such a way that they come close to the measured values. For the materials shown in Table 2 there is still a slight underestimation of the band gaps for large-gap insulators in PBE0 and HSE as well as an overestimation for semiconductors in PBE0.

Although the hybrid functionals thus seem to give an overall better description of excitation energies than conventional DFT functionals such as LDA or PBE, one should not forget that the eigenvalues obtained within a (generalized) Kohn-Sham formalism are merely mathematical Lagrangian parameters without a direct physical meaning. A theoretically better founded theory for excitation (or quasiparticle) energies is provided by the *GW* approximation for the electronic self-energy. Of course, one could interpret the hybrid functionals as approximate self-energies and, thus, the KS eigenvalues as quasiparticle energies. However, the definition of the hybrid functionals is somewhat arbitrary: the mixing parameter $a = 0.25$ is estimated, the parameter κ of the decomposition [Eq. (14)] is fitted to experiment, the mathematical decomposition itself is

		This work [39, 69]			PAW [68]			Expt.
		PBE	PBE0	HSE	PBE	PBE0	HSE	
GaAs	$\Gamma \to \Gamma$	0.54	2.02	1.43	0.56	2.01	1.45	1.52 [85], 1.63 [86]
	$\Gamma \to X$	1.47	2.69	2.06	1.46	2.67	2.02	1.90 [85], 2.01, 2.18 [86]
	$\Gamma \to L$	1.01	2.38	1.78	1.02	2.37	1.76	1.74 [85], 1.84, 1.85 [86]
Si	$\Gamma \to \Gamma$	2.56	3.96	3.32	2.57	3.97	3.32	3.05 [87], 3.35 [88], 3.4 [86]
	$\Gamma \to X$	0.71	1.93	1.29	0.71	1.93	1.29	1.13 [88], 1.25 [87]
	$\Gamma \to L$	1.54	2.87	2.24	1.54	2.88	2.24	2.06 [89], 2.40 [86]
C	$\Gamma \to \Gamma$	5.60	7.74	6.98	5.59	7.69	6.97	7.3 [85]
	$\Gamma \to X$	4.75	6.69	5.90	4.76	6.66	5.91	—
	$\Gamma \to L$	8.46	10.88	10.02	8.46	10.77	10.02	—
MgO	$\Gamma \to \Gamma$	4.77	7.31	6.49	4.75	7.24	6.50	7.7 [90]
	$\Gamma \to X$	9.14	11.63	10.86	9.15	11.67	10.92	—
	$\Gamma \to L$	7.93	10.51	9.69	7.91	10.38	9.64	—
NaCl	$\Gamma \to \Gamma$	5.20	7.13	6.57	5.20	7.26	6.55	8.5 [91]
	$\Gamma \to X$	7.58	9.59	9.05	7.60	9.66	8.95	—
	$\Gamma \to L$	7.30	9.33	8.66	7.32	9.41	8.67	—
Ar	$\Gamma \to \Gamma$	8.70	11.15	10.36	8.68	11.09	10.34	14.2 [92]

Table 2: Kohn-Sham transition energies in eV obtained with the functionals PBE, PBE0, and HSE at experimental lattice constants compared with values from PAW calculations and experiment. An 8×8×8 (12×12×12) **k**-point mesh was employed for the PBE and HSE (PBE0) values. The deviation of the PBE0 values from the PAW ones for the large band-gap insulators is a basis-set effect (see text). (Compiled from Refs. 39 and 69.)

	This work [54]		LMTO [29]		PAW [83]		Expt.
	LDA	*GW*	LDA	*GW*	LDA	*GW*	
Ge	0.02	0.75	−0.08	0.57	—	—	0.74 [93]
Si	0.62	1.11	0.46	0.90	0.62	1.12	1.17 [86]
GaAs	0.29	1.31	0.33	1.31	0.49	1.30	1.63 [86]
CdS	1.17	2.18	—	—	1.14	2.06	2.58 [94]
GaN	1.67	2.83	1.81	3.03	1.62	2.80	3.27 [95]
$SrTiO_3$	1.80	3.36	—	—	—	—	3.25 [96]
$BaTiO_3$	2.18	3.18	—	—	—	—	3.3 [97]
CaSe	2.04	3.63	—	—	—	—	3.85 [98]
C	4.15	5.62	4.11	5.49	4.12	5.50	5.48 [86]
BN	4.35	6.20	—	—	4.45	6.10	5.97 [99]
MgO	4.64	7.17	4.85	6.77	4.76	7.25	7.83 [100]
NaCl	4.90	7.53	—	—	—	—	8.5 [101]

Table 3: Fundamental *GW* band gaps for a variety of semiconductors and insulators compared with experimental and theoretical values from the literature. We also indicate the LDA eigenvalue gaps. All values are in electron volts. [54]

non-unique, the choice of the local xc potential is arbitrary, and so on. One should also mention that the meaning of the local xc potential is different from that of a self-energy. Furthermore, retardation effects in the screening process are completely neglected.

In contrast, quasiparticle energies calculated with the *GW* self-energy have the clear physical interpretation of being excitation energies of the many-body system. Within this theory retardation and lifetime effects are taken into account without resorting to adjustable parameters. In Table 3, we list the fundamental LDA and *GW* band gaps for a variety of semiconductors and insulators, together with experimental and other theoretical values for comparison. The latter are calculated with the LMTO (Ref. 29) and the PAW method (Ref. 83). Additional local orbitals are included to improve the description of high-lying unoccupied states (also cmp. Sec. 4.4). Local orbitals are also used to describe the shell below the valence band, e.g., the cation $2s$ and $2p$ states of MgO and NaCl. Overall our LDA and *GW* values agree very well with those of Ref. 83, but somewhat less so with the older Ref. 29. (We note that there are many more convergence parameters in *GW* calculations as there are in PBE0 or HSE calculations, which explains the larger deviations here compared with Table 2.) As expected, the *GW* self-energy corrects the underestimation of LDA in such a way that the results come very close to the measured values. When comparing with Table 2, we see that the *GW* values are at least as close to experiment as the hybrid-functional results and even better for the large-gap insulators.

However, there is still a slight underestimation in most cases. It has been suggested that a self-consistent scheme could improve the *GW* values further [83, 102]. The starting point is then optimized in such a way that the resulting one-particle orbitals are as close as possible to the quasiparticle wave functions; in particular, closer than those from standard local or semilocal functionals such as LDA. In this way, the first-order perturbative correction [Eq. (8)], where the quasiparticle wave functions are approximated by the one-particle orbitals, is better justified.

| | This work [69] | | PAW [68] | Gaussian [52] | Expt. [84] |
Functional	PBE	HSE	HSE	HSE	—
GaAs	5.743	5.660	5.687	5.710	5.648
Si	5.472	5.441	5.435	5.451	5.430
C	3.571	3.549	3.549	3.557	3.567
MgO	4.265	4.217	4.210	4.222	4.207
NaCl	5.703	5.627	5.659	5.645	5.595

Table 4: Optimized lattice constants in Å obtained with the PBE and the HSE functional. An $8 \times 8 \times 8$ **k**-point mesh was employed. Results are compared to experimental results and theoretical calculations. [69]

| | This work [69] | | PAW [68] | Expt. [84] |
Functional	PBE	HSE	HSE	—
GaAs	64.5	79.2	70.9	75.6
Si	88.9	98.0	97.7	99.2
C	433	467	467	443
MgO	153	177	169	165
NaCl	21.3	28.8	24.5	26.6

Table 5: Bulk moduli in GPa obtained with the PBE and the HSE functional. An $8 \times 8 \times 8$ **k**-point mesh was employed. Results are compared to experimental results and calculations using the HSE functional within a PAW method. [69]

However, self-consistent *GW* calculations are computationally very expensive. When compared with the electronic self-energy, the most obvious source of errors in LDA or GGA is the missing self-interaction correction, which influences the shape of the KS wave functions. Therefore, better results might alternatively be obtained if one uses a functional that treats electronic exchange more accurately, e.g., the hybrid functionals [103]. These combined approaches go beyond the scope of the present article. Nevertheless, we note that the numerical procedure for the *GW* approximation presented here is independent of the starting point and could also be applied within a self-consistent scheme or to functionals containing an exact exchange term.

The main virtue of DFT is, of course, its ability to yield accurate total energies rather than the excitation energies discussed above. In Tables 4 and 5 we present equilibrium lattice constants and bulk moduli derived from HSE total energies for different lattice constants and compare the results with other implementations based on plane-wave (PAW) [68] and Gaussian basis sets [52]. The results of all three methods agree very well. Except for diamond, the HSE functional yields lattice constants and bulk moduli in much better agreement with experiment than the semilocal PBE functional, which tends to overestimate the former and underestimate the latter.

	HSE		LSDA+U			B3LYP	Expt.
	This work		[2]	[3]	[4]	[5]	
Latt. const. (Å)	(4.988)	4.967	4.92	5.08	(4.988)	5.05	4.988 [120]
Bulk mod. (GPa)		164		150		159	192 [121]
Mag. mom. (μ_B)	6.99[1]	6.99		6.93[1]		7.0	6.88 [104]
$E_g^{\text{dir.}}(X)$ (FM)	0.90[1]	0.85	−0.16		0.91[1]	1.18	0.90 [122]
$E_g^{\text{dir.}}(X)$ (PM)	1.17[1]	1.11	0.10	0.98[1]	1.30[1]	1.77	1.31 [122]
$E_g^{\text{indir.}}(\Gamma \rightarrow X)$ (FM)	0.01[1]	−0.06	−0.45	0.14[1]	0.43[1]	0.72	
$E_g^{\text{indir.}}(\Gamma \rightarrow X)$ (PM)	0.90[1]	0.85	−0.13	0.69[1]	0.98[1]	1.47	
$\epsilon_{4f}^{\uparrow} - \epsilon_F$	−6.00[1]	−6.00	−7.8	−8.1[1]		−6.3	−7.8 [105]
$\epsilon_{4f}^{\uparrow} - \epsilon_F$	6.05[1]	6.05	6.6	5.0[1]	4.8[1]	5.5	5.5–6.1 [6]

[1] At the experimental lattice constant.

[2] [108, 109, 110]; U optimized for Gd bulk [123].

[3] [118]; U chosen to reproduce the experimental direct gap of paramagnetic GdN [124].

[4] [122]; U chosen to reproduce the experimental direct gap of paramagnetic GdN.

[5] [119]; insulating solution.

[6] [125]; measured for GdX (X =P, As, Sb, and Bi).

Table 6: Comparison of our HSE results [69] for GdN with those from LSDA+U and experiment. For the calculation of the ferro- and paramagnetic states (FM, $T < T_C$; PM, $T > T_C$) see text. Energies are given in eV. The calculations were performed with a $8\times8\times8$ **k**-point set. [69]

4.3 GdN – a HSE study

In this section we present results for the ground-state properties and the band structure of GdN calculated with the HSE functional. Gadolinium nitride (GdN) is widely studied owing to the ferromagnetic order, large magnetic moment of $6.88\,\mu_B$ per Gd atom [104] and its large magnetoresistive effect [105], which makes the material interesting for technological applications. The mechanism of the ferromagnetic order is still under debate. Various types are being discussed, such as carrier mediated [106, 107] and superexchange mechanisms [108, 109, 110]. Another point of debate are the electronic properties. It was experimentally demonstrated to be a low carrier semimetal in single crystals [111] and insulating in thin films [112]. There are also several recent reports of thin films of GdN having a degenerately doped semiconducting [113, 114, 115] or a metallic ground state [115] based on the resistivity data measured at low temperatures. Results from LSDA+U calculations are inconclusive. While the linearized muffin-tin orbital (LMTO) approach yields a narrow-gap semiconductor as ground state [116, 117, 118, 106], one obtains a transition from a half-metallic to a semiconducting ground state under strain within the FLAPW method [108, 109, 110]. Two different solutions close in energy were obtained in an investigation using the hybrid functional B3LYP [119]. Both solutions were half-metallic, one in the majority spin channel, the other one in the minority spin channel.

GdN crystallizes in the rocksalt structure with a room-temperature lattice constant of $a_{\text{GdN}} = 4.988$ Å[121]. Half-filling of the Gd $4f$ states gives rise to ferromagnetic ordering in the electronic ground state at 0 K. The rest of the valence and conduction bands are formed by the N $2s$ and
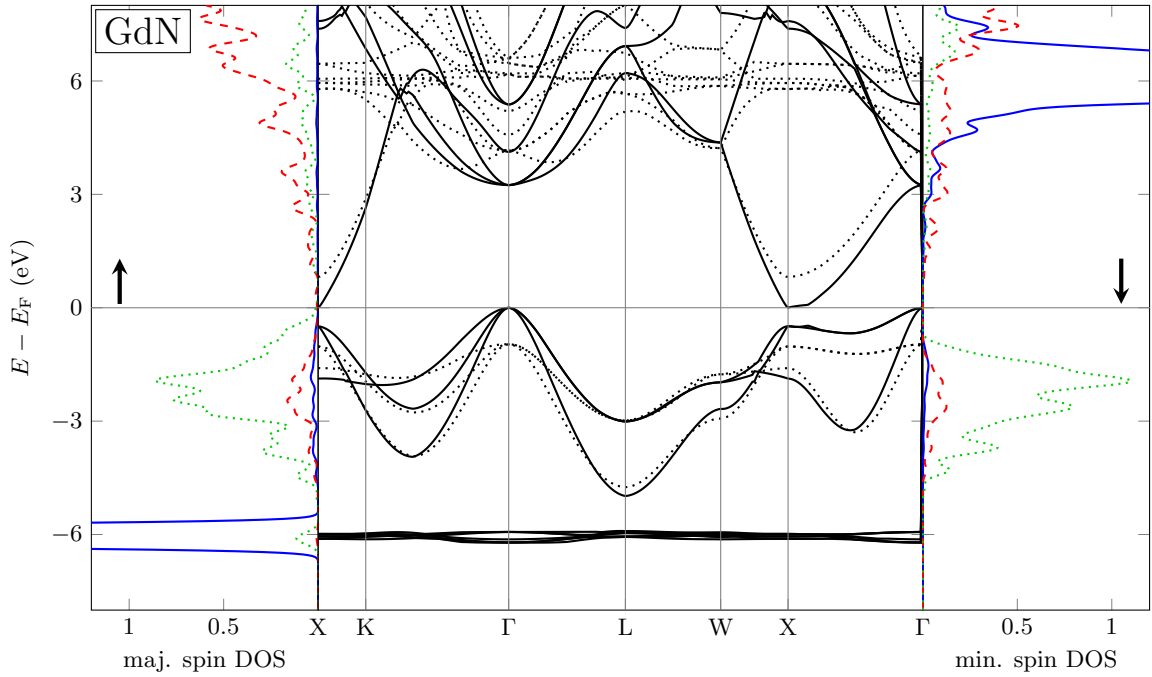
Figure 8: Band structure and density of states (DOS; in states per eV) of GdN at the experimental lattice constant. The majority and minority bands are plotted as solid and dotted lines, respectively. In the DOS, the solid blue line shows the Gd $4f$ states, the red dashed line the Gd $5d$ states, and the green dotted line the N $2p$ states. [69]

$2p$ as well as Gd $5d$ and $6s$ states, respectively. In order to compare our theoretical results with experimental measurements obtained at room temperature, where GdN is in the paramagnetic state, we assume that the exchange splitting and the local magnetic $4f$ moment calculated for the ferromagnetic state at 0 K persist in the paramagnetic state. However, the long-range ordering is lost, and the moments fluctuate resulting in a vanishing macroscopic magnetization. The fluctuating magnetic moments produce a random exchange potential, which effectively leads to electronic single-particle energies that can be approximated by averaging over the corresponding spin-up and spin-down energies of the ferromagnetic phase [118].

In Table 6 we compare the structural and electronic properties calculated with the HSE functional to some of the available experimental [104, 120, 121, 122] and theoretical data, obtained with the hybrid B3LYP functional [119] and within the LSDA+U approach [108, 118, 122]. Our HSE calculations yield a lattice parameter of 4.967 Åin very close agreement to the experiment, while B3LYP and the LSDA+U approach overestimate the value by about 2%. We note that thermal expansion could account for the remaining difference to the experimental lattice parameter, which was determined at the room temperature (whereas the theoretical result corresponds to 0 K). Assuming linear expansion between 0 K and room temperature (293 K) with the thermal expansion coefficient of isostructural and isovalent EuO ($\approx 13 \times 10^{-6}$K^{-1}; the coefficient of GdN is unknown) [126], one would extrapolate to 4.969 Åat 0 K, which is, indeed, very close to our optimized lattice constant.

Figure 8 shows the spin-resolved band structure and the spin- and orbital-resolved density of

states calculated at the experimental lattice constant. The majority and minority Gd $4f$ states are located at $-6\,\text{eV}$ and $+6\,\text{eV}$, respectively. GdN is in narrow-gap semiconducting ground state with an almost vanishing indirect band gap ($\Gamma \rightarrow X$) in the majority spin ($0.01\,\text{eV}$), while the minority bands exhibit a robust gap of $1.5\,\text{eV}$. Thus, GdN is just at the border between an insulator and a semimetal. Its electronic properties might change easily depending on the experimental circumstances, which explains the contradicting experimental reports. Upon decreasing the lattice constant isotropically by just 0.02 Å(to the theoretically optimized value), we observe a transition to a half-metallic state or, more precisely, to a semimetallic state just for the majority states, we formally get a negative band-gap value (see Table 6). Duan *et al.* [108] also found this transition but at a much larger lattice constant of 5.63 Å. The proximity of semiconducting and semimetallic states suggest that growth conditions or the lattice expansion upon temperature changes play a decisive role in the transport properties of GdN. Furthermore, upon doping at a small concentration, e.g., with Eu, one might be able to achieve significant charge currents with 100% spin polarization. When the temperature is raised across the Curie temperature, the spin polarization of the charge current disappears and the band gap opens due to the averaged-out exchange potential, giving rise to a possible change of the conductivity by orders or magnitude. A coupling of the electrons to the fluctuating $4f$ moments might also play a role [127]. The band transition energies calculated with our method for the ferromagnetic ($T < T_\text{C}$) and the paramagnetic ($T > T_\text{C}$) state are listed in Table 6. They compare well with the experimental data where available. We note the aforementioned increase of the band gap, both upon increasing the lattice parameter and upon the transition to the paramagnetic state. The partial compensation of the self-interaction error in HSE leads to a pronounced shift of the localized Gd $4f$ majority band to larger binding energies compared to PBE. While the PBE functional yields a much too shallow f majority band, located at $3.1\,\text{eV}$ below the Fermi energy, this band appears at a binding energy of $6.0\,\text{eV}$ in HSE, much closer to its experimentally measured position at $7.8\,\text{eV}$ [105]. We note a very good agreement with the insulating B3LYP result, where the position of the $4f$ peak is located at $6.3\,\text{eV}$ [119].

The ground state of GdN is ferromagnetic (FM), with an experimental Curie temperature of $58\,\text{K}$ [121] and a magnetic moment of $6.88\,\mu_\text{B}$ per Gd atom [104] determined from the saturation magnetization at $1.2\,\text{K}$. The calculated ground state is ferromagnetic, too, with a total magnetic moment of $7\,\mu_\text{B}$, of which $6.99\,\mu_\text{B}$ originates from the Gd muffin-tin sphere. From total-energy differences between ferromagnetic and antiferromagnetic spin configurations, we derive the exchange coupling constants for a Heisenberg spin Hamiltonian, from which the Curie temperature is determined. This is a critical test for the quality of total energies that can be expected from the HSE functional. The classical spin Heisenberg Hamiltonian is given by

$$\mathcal{H} = -\frac{1}{2}\sum_i \mathbf{S}_i \left(J_1 \sum_{j=\text{nn}} \mathbf{S}_j + J_2 \sum_{j=\text{nnn}} \mathbf{S}_j \right) \tag{45}$$

including the nearest-neighbors (nn), and the next-nearest-neighbors (nnn) interaction where J_1 and J_2 are the respective coupling constants with normalized spin vectors $\mathbf{S}_i$ and $\mathbf{S}_j$. The coupling constants are extracted from the differences of the total energies of the ferromagnetic (FM) and two types of antiferromagnetic (AFM) configurations characterized by the planes of ferromagnetically ordered moments that are antiferromagnetically stacked along the crystallo-

	ΔE_I	ΔE_{II}	J_1	J_2	$T_{\mathrm{C}}^{\mathrm{MFA}}$	$T_{\mathrm{C}}^{\mathrm{MFA}}$
This work [69]	8.8	7.6	1.09	0.17	55	42
Duan *et al.* [109]	6.7	4.2	0.84	−0.14	36	26
Mitra *et al.* [106]	3.4	0.4	0.42	−0.36	11	5

Table 7: Total-energies differences for different magnetic configurations [Eq. (46)], the Heisenberg coupling constants, and the corresponding Curie temperatures within the mean-field [Eq. (48)] and the random-phase approximation [Eq. (49)]. Energies and coupling constants are given in meV and the Curie temperatures in K. [69]

graphic [001] or [111] directions (AFM-I and AFM-II, respectively) [109]. For the calculation of the AFM-I (AFM-II) phase, we use an $1 \times 1 \times 2$ ($\sqrt[3]{2} \times \sqrt[3]{2} \times \sqrt[3]{2}$) unit cell. All the calculations are performed at the experimental lattice constant. The Heisenberg coupling constants J_1 and J_2 are obtained from the relations

$$\Delta E_I = E_{\mathrm{AFM,I}} - E_{\mathrm{FM,I}} = 8J_1 \tag{46}$$

$$\Delta E_{II} = E_{\mathrm{AFM,II}} - E_{\mathrm{FM,II}} = 6J_1 - 6J_2 \,. \tag{47}$$

The resulting coupling constants shown in Table 7 are both positive, confirming the ferromagnetic nature of the ground state. Compared with previous studies using the LSDA+U method within the FLAPW [109] and LMTO basis [106], our values obtained with HSE yield significantly higher coupling constants and, as a consequence, a higher Curie temperature (T_{C}), in closer agreement with experiment. The Curie temperature was estimated within two methods: the mean-field approximation (MFA)

$$T_{\mathrm{C}}^{\mathrm{MFA}} = \frac{1}{3k_{\mathrm{B}}} \left(12J_1 + 6J_2 \right) , \tag{48}$$

yielding $T_{\mathrm{C}}^{\mathrm{MFA}} = 55\,\mathrm{K}$, and the random-phase approximation as described in Refs. 128 and 129

$$T_{\mathrm{C}}^{\mathrm{RPA}} = \frac{1}{3k_{\mathrm{B}}} \left[\int_{\mathrm{BZ}} \frac{d^3q}{J(\mathbf{0}) - J(\mathbf{q})} \right]^{-1} , \tag{49}$$

where we evaluate the integral on a discrete mesh of $\mathbf{q}$ points within the Brillouin zone. $J(\mathbf{q})$ is the Fourier transform of the exchange coupling constants defined as

$$J(\mathbf{q}) = \sum_{\mathrm{nn}} J_1 e^{i\mathbf{q}\cdot\mathbf{R}_{\mathrm{nn}}} + \sum_{\mathrm{nnn}} J_2 e^{i\mathbf{q}\cdot\mathbf{R}_{\mathrm{nnn}}} , \tag{50}$$

where $\mathbf{R}_{\mathrm{nn}}$ and $\mathbf{R}_{\mathrm{nnn}}$ are the positions of the nearest and the next nearest neighbors, respectively. The resulting $T_{\mathrm{C}}^{\mathrm{RPA}} = 42\,\mathrm{K}$ is smaller than the mean-field estimate, which is not surprising as it is known that the mean-field approximation usually overestimates the Curie temperature. The two theoretical values are in very good agreement with the experimental situation, e.g., $T_{\mathrm{C}} = 68$, 69, 58, or 37 K as reported by Granville *et al.* [113], Khazen *et al.* [130], Leuenberger *et al.* [105], and Yoshitomi *et al.* [131], which vary in value also depending on film thickness, strain, grain size, stoichiometry, and N vacancies [132, 133].

4.4 ZnO – an extreme case for GW calculations

In a publication Shih *et al.* [80] presented a new calculation for the band gap of ZnO in the wurtzite structure employing the GW approximation [21] for the electronic self-energy in a conventional *one-shot* approach [63, 23] corresponding to Eq. (8). All previous calculations [134, 135, 28, 83] invoking the one-shot LDA+GW approach had shown that the band gap of wurtzite ZnO is underestimated with respect to the experimental value by more than 1 eV. They fall in the range 2.12–2.6 eV, while the experimental gap amounts to 3.6 eV, [136] after correction for vibrational effects. This large underestimation is untypical for GW calculations of sp-bound systems.

Shih *et al.* investigated the convergence of the correlation self-energy and the dielectric matrix with respect to the number of bands. They performed calculations with up to 3000 bands corresponding to a maximal band energy of 67 Ry as well as a cutoff for the dielectric matrix of up to 80 Ry and showed that the resulting GW band gaps, 3.4 eV for LDA+GW (and 3.6 eV for LDA+U+GW), turned out to be in very good agreement with experiment. These new results for ZnO are in striking contrast to the other GW calculations and seem to indicate that the convergence with respect to the unoccupied states is even more problematic than previously thought. Of course, several points of criticism can be raised. First, the calculation relies on the pseudopotential approximation, in which the numerically important core-valence exchange is neglected, the wave functions are pseudized, and high-lying states are inaccurately described. Second, the study was based on a plasmon-pole model for the dielectric function, which only approximates the full expression of the RPA.

Therefore, we reexamined [81] the one-shot GW band gap of wurtzite ZnO with our own code, which neither employs the pseudopotential approximation nor a plasmon-pole model for the dielectric matrix. We employ the contour integration for the frequency convolution in Eq. (7), and Eq. (8) is solved on an energy mesh around $\epsilon_{n\mathbf{q}}^{\sigma}$ with a cubic spline interpolation between the mesh points. We carefully converged the number of empty bands for the calculation of both the polarization function and the correlation self-energy as well as the MPB parameters. While the ground-state electron density was converged with a standard LAPW basis with moderate cutoff parameters ($l_{\max} = 8$, $G_{\max} = 4.3\,a_0^{-1}$, where a_0 is the Bohr radius, and including second-derivative local orbitals [79]), we had to employ much larger cutoffs to generate enough wave functions for the GW calculation: $l_{\max} = 12$ and $G_{\max} = 8.0\,a_0^{-1}$. Furthermore, in order to avoid linearization errors in the MT part of the LAPW basis [79, 77], we added local orbitals [137, 76] (LOs) with different angular momentum quantum numbers and energy parameters distributed over the relevant energy range: 292 LOs for Zn (for l = 0-6) and 186 for O (for l = 0-5). We also treat the $3s$ and $3p$ semicore states of Zn explicitly with LOs. We found that relatively modest MPB parameters were sufficient: $L_{\max} = 4$ and $G'_{\max} = 2.4\,a_0^{-1}$. However, we had to include a large number of radial functions in the MT spheres, which after optimization led to 177 MT functions for Zn and 127 for O. Obviously, the rapid variations close to the atomic nuclei must be accurately described. Within the all-electron MPB this is possible with a comparatively small number of MT functions, while in a pure plane-wave approach a very large number of plane waves is necessary to resolve the variations sufficiently. This explains the finding of Shih *et al.* that the dielectric matrix must be converged to very large energy cutoffs. The
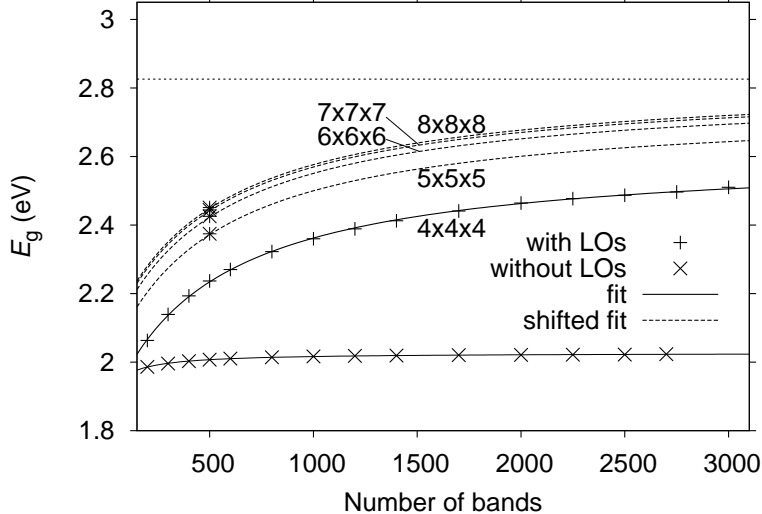
Figure 9: Band convergence of the quasiparticle band gap of ZnO employing a 4×4×4 **k**-point set and calculated with (pluses) and without local orbitals (LOs) (crosses) for high-lying states. The solid lines show hyperbolic fits. We also indicate results with finer **k**-point samplings (stars) calculated with LOs and 500 bands. The dashed lines show the hyperbolic fit shifted to align with these results. The fit asymptote for the 8×8×8 **k**-point set at 2.83 eV (dotted line) is considered the best estimate for the all-electron one-shot *GW* band gap. [81]

total number of MPB functions in the calculations is less than 700 per **k** point. This number is further reduced to around 490 by constructing linear combinations that are continuous in value and radial derivative at the MT sphere boundaries, cmp. Sec. 3.1 and Ref. 39.

Figure 9 shows the quasiparticle band gap of ZnO as a function of the number of states included in the calculation of the polarization function and the correlation self-energy. The calculations were performed with a 4×4×4 **k**-point sampling of the Brillouin zone. There is a large difference between calculations with (pluses) and without the LOs for unoccupied states (crosses), which shows the importance of eliminating the linearization error of the conventional LAPW basis. As the linearization error becomes larger for higher and higher bands, it is not surprising that the difference between the convergence curves grows toward increasing numbers of bands. We find an asymptotic difference of 0.5 eV. The calculations without LOs for unoccupied states show a false convergence behavior, as they seem to converge with a few hundred bands, but toward a much too small value. A similar false convergence behavior was found in Ref. 80 for calculations with underconverged dielectric functions.

As Fig. 9 shows, the calculations with eliminated linearization error (pluses) take far more bands to converge. In fact, even with 3000 bands the band gap is still not completely converged. Therefore, we fitted the values with the hyperbolic function

$$f(N) = \frac{a}{N - N_0} + b \tag{51}$$

where a, b, and N_0 are fit parameters. It is surprising how closely the fitted curve (solid line) follows the calculated data points. This makes us confident in taking the fit asymptote b as the band gap extrapolated to infinite bands. Furthermore, we have recalculated the band gap

with finer **k**-point meshes (up to 8×8×8) and 500 bands (crosses). The dashed lines show correspondingly shifted hyperbolic fits. The asymptote of the fit corresponding to an 8×8×8 **k**-point sampling is found at 2.83 eV, which we take as the final best estimate for the all-electron one-shot GW band gap.

This band gap is 0.4–0.9 eV larger than the previously reported values. Both the large number of bands that are needed for proper convergence and the elimination of the linearization error, which has not been undertaken in the previous all-electron studies, are responsible for this large difference. Our value is still about 0.4 eV smaller than the band gap of Shih *et al.*, though. In fact, a certain discrepancy should be expected because of the pseudopotential approximation and the plasmon-pole model for the dielectric function used in Ref. 80. The pseudopotential approximation not only neglects the important contribution of core-valence exchange. It also yields accurate wave functions only in the vicinity of the atomic electron energies, but not for high-lying states. This error is very similar in spirit to the linearization error of the LAPW basis and is also of the same magnitude [79]. Thus, it should have an impact on the GW results comparable in size to the linearization error of the LAPW basis.

With the LDA band gap of only 0.73 eV the quasiparticle correction amounts to more than 2 eV. It can be expected that a treatment beyond the one-shot approach, for example, by taking into account offdiagonal elements of the self-energy, by updating the Green function, or by including vertex corrections, will further increase the value and, thus, bring it even closer to the experimental value. As was shown in Ref. 80, already using LDA+U instead of LDA as the mean-field starting point, which corrects the $2p$-$3d$ hybridization, gives an upward correction of 0.2 eV in the resulting GW band gap. Self-consistency can be simulated in a simplified way by aligning the chemical potential of the LDA starting point in such a way that the ionisation potential (position of the highest occupied state) of the starting point becomes identical to that after the GW self-energy correction [138]. In practice, this means that the LDA eigenvalues are all shifted by a constant energy determined by the condition that the self-energy correction for the highest occupied state vanishes. While the LDA band gap trivially remains unchanged, it does influence the GW band gap. In the case of ZnO we observe a sizable upward correction by about 0.2 eV toward the experimental value.

5 Conclusions

We have presented details of numerical realizations of the hybrid functionals PBE0 and HSE as well as the GW approximation for the electronic self-energy within the all-electron FLAPW method as realized in the FLEUR [75] and SPEX [54] codes. All implementations are based on the all-electron mixed product basis (MPB), which is specifically designed for the representation of wave-function products and retains the full accuracy of the all-electron framework. For example, the six-dimensional integrals for the nonlocal exchange potential (or self-energy) decompose into vector-matrix-vector products in this basis, where the Coulomb matrix must be calculated only once at the beginning of the self-consistent-field cycle. As hybrid-functional and GW calculations are notoriously expensive in terms of computation time and memory, we presented ways to accelerate the calculations considerably. As a demonstration, we showed a GW calculation

for 128 carbon atoms in a diamond supercell. If inversion symmetry is present, the MPB can be defined in such a way that the Coulomb matrix, the screened interaction (for imaginary frequencies), wave-function products, etc. become real-valued. Spatial and time-reversal symmetries are further exploited (1) to restrict the $\mathbf{k}$-point summation for the nonlocal quantities to an irreducible wedge of the BZ and (2) to identify those exchange matrix elements in advance that are zero and need not to be calculated. For the exchange term, the computational cost is considerably reduced by a suitable unitary transformation of the MPB, which makes the Coulomb matrix sparse. The divergence of the bare and the screened interaction potential in the limit $\mathbf{k} \to \mathbf{0}$ is treated analytically, while higher-order corrections can be derived with the help of $\mathbf{k} \cdot \mathbf{p}$ perturbation theory. The anisotropy of the RPA screening at $\mathbf{k} = \mathbf{0}$ is fully taken into account. This procedure gives rise to a fast $\mathbf{k}$-point convergence, which is particularly important for the evaluation of the nonlocal quantities in the hybrid functionals and the GW method. The short-range nonlocal exchange term for the HSE functional is evaluated by subtracting the corresponding long-range part of exchange from the bare exchange, where we exploit the rapidly converging Fourier series of the long-range interaction potential. This procedure allows constructing the nonlocal HSE potential from PBE0 up to machine precision at a negligible computational cost. We note that this approach is quite general and not restricted to the error function used in the HSE functional. For the correlation part of the GW self-energy we can apply an optimization of the MPB that involves a basis transformation to the eigenvectors of the Coulomb matrix. By neglecting eigenvectors with eigenvalues below a certain threshold value we only retain the dominant part of the bare electron-electron interaction. The threshold value then becomes a convergence parameter. This optimization reduces the matrix sizes of response quantities such as the screened interaction, again giving rise to a speed up of the calculation. As a demonstration, we showed CPU timings of calculations, where up to 128 carbon atoms were treated explicitly. Furthermore, we have demonstrated that while a direct iteration of the one-particle equation employing hybrid functionals needs extensively many steps to converge due to the nonlocal exchange potential, a nested density-only and density-matrix iteration scheme accelerates the convergence of the self-consistent-field cycle considerably.

We have analyzed the convergence of the interband transition and total energies with respect to the parameters of the MPB for the example of the PBE0 hybrid functional and found that parameters well below the exact limit with respect to the LAPW basis are sufficient. This makes the MPB an efficient and versatile basis for the calculation of nonlocal two-particle quantities within the FLAPW method.

We have presented results for prototype semiconductors and insulators calculated with the hybrid functional PBE0 and HSE as well as with the GW approximation. The resulting gap energies, lattice constants, and bulk moduli for a variety of semiconductors and insulators are consistently closer to experimental measurements than within conventional local xc functionals and compare very well with recent theoretical results obtained with PAW and Gaussian-based methods. We have confirmed the finding of Paier *et al.* [68] that the $\mathbf{k}$-point convergence within HSE is comparable to the conventional local PBE functional, whereas in PBE0 much larger $\mathbf{k}$-point meshes are necessary.

We have calculated the geometric, electronic, and magnetic properties of the rare-earth compound GdN within the HSE functional. There is an ongoing discussion whether the ground

state is insulating or metallic. In fact, within the HSE functional the ground state is very close to a phase transition: we observe a tiny indirect band gap for the experimental lattice constant at room temperature, which vanishes at the theoretically optimized 0 K lattice constant – the compound becomes half-metallic. The experimentally known band transitions are in good agreement with our theoretical results. Furthermore, we have calculated the coupling constants for the Heisenberg spin Hamiltonian from total-energy differences of ferromagnetic and antiferromagnetic configurations. The resulting Curie temperature of 42 K evaluated in the random-phase approximation is in good agreement with the experimental values varying between 37 and 69 K and gives confidence in the energetics obtained by HSE for different magnetic phases. We note that the HSE functional does not contain an adjustable parameter such as the Hubbard U parameter in LSDA+U calculations.

The band convergence is a serious issue in GW calculations and must be thoroughly dealt with. We have presented results for wurtzite ZnO, which is an extreme case in this respect, and could confirm the main result of a recent GW calculation [80] using the plane-wave pseudopotential approach: the GW band gap of ZnO shows a very slow convergence with respect to the number of states used to construct the polarization function and the correlation self-energy. Furthermore, when high-lying bands are included in the calculation, the linearization error of all-electron approaches becomes another important issue. As we have shown, it can be eliminated systematically within the FLAPW method by augmenting the basis in the MT spheres with LOs (see Ref. 81 for details). In the case of ZnO this procedure yields a correction of about 0.5 eV and brings the calculated band gap (2.83 eV) much closer to experiment than in previous studies.

The numerical approach presented in this article can also be applied to other methods. Recently, we have employed the MPB to solve the optimized-effective-potential (OEP) equation, which yields a local instead of a nonlocal effective potential from the exact Hartree-Fock exchange term [59]. Here, additional contributions from radial Sternheimer solutions can improve the precision and convergence of the density response function significantly [60]. Additional contributions from electron correlation can be considered by using the RPA total energy within the adiabatic-connection fluctuation-dissipation theory, either by a one-shot correction of a preceding DFT calculation or self-consistently by solving the optimized-effective-potential equation. The extension of the numerical procedure to off-diagonal elements of the GW self-energy is straightforward. The quasiparticle equation can then be solved beyond perturbation theory, which also makes self-consistent calculations possible, e.g., employing the quasiparticle self-consistent scheme [102]. When strongly correlated electron systems are studied, one often uses methods that rely on the Hubbard U parameter, which is a measure of the strength of the effective electron-electron interaction in the localized states, such as the LSDA+U method and the dynamical mean-field theory. We have extended our RPA implementation to the constrained RPA (cRPA) approach, with which the Hubbard U parameter defined in a basis of Wannier functions [139] expressed in terms of FLAPW basis functions [140, 141] can be calculated from first principles [142]. A Wannier basis can also be used to represent the four-point T matrix, which describes the correlated motion of an electron-hole pair in terms of ladder diagrams. If the electron and hole are of opposite spin, the two-particle propagation function contains information about the collective spin excitations [143]. These spin waves or magnons can also contribute to the renormalization of quasiparticles. The corresponding self-energy is then approximated by a GT diagram [144].

Acknowledgment

References

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] C. Fiolhais, F. Noguiera, and M. A. L. Marques, ed., *A Primer in Density Functional Theory*, vol. 620 of *Lecture Notes in Physics* (Springer, Heidelberg, 2003).

[3] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[4] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[5] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[6] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[7] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).

[8] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[9] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[10] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[11] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[12] J. Muscat, A. Wander, and N. M. Harrison, Chem. Phys. Letters **342**, 397 (2001).

[13] C. Franchini, V. Bayer, R. Podloucky, J. Paier, and G. Kresse, Phys. Rev. B **72**, 045132 (2005).

[14] T. Bredow and A. R. Gerson, Phys. Rev. B **61**, 5194 (2000).

[15] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).

[16] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).

[17] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[18] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[19] J. Tao, S. Tretiak, and J. Zhu, J. Chem. Phys. **128**, 084110 (2008).

[20] G. D. Mahan, *Many-particle physics* (Plenum, New York, 1990).

[21] L. Hedin, Phys. Rev. **139**, A796 (1965).

[22] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[23] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[24] S. Chawla and G. A. Voth, J. Chem. Phys. **108**, 4697 (1998).

[25] N. Hamada, M. Hwang, and A. J. Freeman, Phys. Rev. B **41**, 3620 (1990).

[26] F. Aryasetiawan, Phys. Rev. B **46**, 13051 (1992).

[27] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).

[28] M. Usuda, N. Hamada, T. Kotani, and M. van Schilfgaarde, Phys. Rev. B **66**, 125101 (2002).

[29] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).

[30] A. Yamasaki and T. Fujiwara, Phys. Rev. B **66**, 245108 (2002).

[31] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).

[32] B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).

[33] S. Lebègue, B. Arnaud, M. Alouani, and P. E. Bloechl, Phys. Rev. B **67**, 155208 (2003).

[34] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).

[35] A. Ernst, M. Lüders, P. Bruno, W. M. Temmerman, and Z. Szotek, unpublished.

[36] V.R. Saunders, R. Dovesi, C. Roetti, M. Causa, N.M. Harrison, R. Orlando, and C.M. Zicovich-Wilson, *CRYSTAL98 User's Manual* (University of Torino, 1998).

[37] J. Paier, R. Hirschl, M. Marsman, and G. Kresse, J. Chem. Phys. **122**, 234102 (2005).

[38] P. Novak, J. Kunes, L. Chaput, and W. E. Pickett, Phys. Stat. Sol. B **243**, 563 (2006).

[39] M. Betzinger, C. Friedrich, and S. Blügel, Phys. Rev. B **81**, 195117 (2010).

[40] S. Massidda, M. Posternak, and A. Baldereschi, Phys. Rev. B **48**, 5058 (1993).

[41] M. Weinert, J. Math. Phys. **22**, 2433 (1981).

[42] F. Tran and P. Blaha, Phys. Rev. B **83**, 235118 (2011).

[43] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[44] J. Harris and R. O. Jones, Journal of Physics F: Metal Physics **4**, 1170 (1974).

[45] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[46] D. C. Langreth and J. P. Perdew, Phys. Rev. B **15**, 2884 (1977).

[47] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).

[48] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[49] B. G. Janesko, T. M. Henderson, and G. E. Scuseria, Phys. Chem. Chem. Phys. **11**, 443 (2009).

[50] E. N. Brothers, A. F. Izmaylov, J. O. Normand, V. Barone, and G. E. Scuseria, J. Chem. Phys. **129**, 011102 (pages 4) (2008).

[51] J. Heyd and G. E. Scuseria, J. Chem. Phys. **120**, 7274 (2004).

[52] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. **125**, 224106 (pages 5) (2006).

[53] L. Dagens and F. Perrot, Phys. Rev. B **5**, 641 (1972).

[54] C. Friedrich, S. Blügel, and A. Schindlmayr, Phys. Rev. B **81**, 125102 (2010).

[55] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).

[56] M. Weinert, E. Wimmer, and A. J. Freeman, Phys. Rev. B **26**, 4571 (1982).

[57] H. J. F. Jansen and A. J. Freeman, Phys. Rev. B **30**, 561 (1984).

[58] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. B **49**, 16214 (1994).

[59] M. Betzinger, C. Friedrich, S. Blügel, and A. Görling, Phys. Rev. B **83**, 045105 (2011).

[60] M. Betzinger, C. Friedrich, and S. Blügel, Phys. Rev. B (in preparation).

[61] C. Friedrich, A. Schindlmayr, and S. Blügel, Comput. Phys. Commun. **180**, 347 (2009).

[62] H. N. Rojas, R. W. Godby, and R. J. Needs, Phys. Rev. Lett. **74**, 1827 (1995).

[63] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[64] V. I. Anisimov, ed., *Electronic Structure Calculations in Advances in Condensed Matter Science* (Gordon and Breach, New York, 2000).

[65] G. A. Baker, Jr., *Essentials of Padé Approximants* (Academic Press, New York, 1975).

[66] M. Betzinger, *Efficient implementation of the non-local exchange potential within the flapw method*, Diplomarbeit, RWTH Aachen, Institut für Festkörperphysik, Forschungszentrum Jülich (2007).

[67] S. K. Kim, *Group Theoretical Methods and Applications to Molecules and Crystals* (Cambridge University Press, Cambridge, 1999).

[68] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, J. Chem. Phys. **124**, 154709 (pages 13) (2006); *ibid.* **124**, 249901 (2006).

[69] M. Schlipf, M. Betzinger, C. Friedrich, M. Ležaić, and S. Blügel, Phys. Rev. B **84**, 125142 (2011).

[70] S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).

[71] C. Freysoldt, P. Eggert, P. Rinke, A. Schindlmayr, R. W. Godby, and M. Scheffler, Comput. Phys. Commun. **176**, 1 (2007).

[72] P. Ziesche and G. Lehmann, ed., *Ergebnisse in der Elektronentheorie der Metalle* (Akademie/Springer Verlag Berlin, 1983).

[73] C. G. Broyden, Math. Comput. **19**, 577 (1965).

[74] C. G. Broyden, Math. Comput. **21**, 368 (1966).

[75] http://www.flapw.de.

[76] D. Singh, Phys. Rev. B **43**, 6388 (1991).

[77] E. E. Krasovskii, A. N. Yaresko, and V. N. Antonov, J. Electron Spectrosc. Relat. Phenom. **68**, 157 (1994).

[78] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, Phys. Rev. B **69**, 125212 (2004).

[79] C. Friedrich, A. Schindlmayr, S. Blügel, and T. Kotani, Phys. Rev. B **74**, 045104 (2006).

[80] B.-C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **105**, 146401 (2010).

[81] C. Friedrich, M. C. Müller, and S. Blügel, Phys. Rev. B **83**, 081101(R) (2011); ibid. **84**, 039906(E) (2011).

[82] F. Bruneval and X. Gonze, Phys. Rev. B **78**, 085125 (2008).

[83] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).

[84] J. Heyd and G. E. Scuseria, J. Chem. Phys. **121**, 1187 (2004).

[85] K.-H. Hellwege, O. Madelung, M. Schulz, and H. Weiss, eds., *Numerical Data and Functional Relationships in Science and Technology*, vol. 17 & 22 of *Landolt-Börnstein - Group III Condensed Matter* (Springer-Verlag, New York, 1982).

[86] T. C. Chiang and F. J. Himpsel, *Electronic Structure of Solids: Photoemission Spectra and Related Data*, vol. 23 a of *Landolt-Börnstein - Group III Condensed Matter Numerical Data and Functional Relationships in Science and Technology* (Springer Verlag, Berlin, 1989).

[87] J. E. Ortega and F. J. Himpsel, Phys. Rev. B **47**, 2130 (1993).

[88] M. Welkowsky and R. Braunstein, Phys. Rev. B **5**, 497 (1972).

[89] R. Hulthn and N. G. Nilsson, Solid State Commun. **18**, 1341 (1976).

[90] S. Adachi, *Optical Properties of Crystalline and Amorphous Semiconductors: Numerical Data and Graphical Information* (Kluwer Academic, Dordrecht, 1999).

[91] R. T. Poole, J. Liesegang, R. C. G. Leckey, and J. G. Jenkin, Phys. Rev. B **11**, 5190 (1975).

[92] R. J. Magyar, A. Fleszar, and E. K. U. Gross, Phys. Rev. B **69**, 045111 (2004).

[93] G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts, Phys. Rev. **108**, 1377 (1957).

[94] K. O. Magnusson, U. O. Karlsson, D. Straub, S. A. Flodström, and F. J. Himpsel, Phys. Rev. B **36**, 6566 (1987).

[95] H. Okumura, S. Yoshida, and T. Okahisa, Appl. Phys. Lett. **64**, 2997 (1994).

[96] K. van Benthem, C. Elsässer, and R. H. French, J. Appl. Phys. **90**, 6156 (2001).

[97] L. T. Hudson, R. L. Kurtz, S. W. Robey, D. Temple, and R. L. Stockbauer, Phys. Rev. B **47**, 1174 (1993).

[98] Y. Kaneko and T. Koda, J. Cryst. Growth **86**, 72 (1990).

[99] K. Watanabe, T. Taniguchi, and H. Kanda, Nature Mater. **3**, 404 (2004).

[100] R. C. Whited, C. J. Flaten, and W. C. Walker, Solid State Commun. **13**, 1903 (1973).

[101] R. T. Poole, J. G. Jenkin, J. Liesegang, and R. C. G. Leckey, Phys. Rev. B **11**, 5179 (1975).

[102] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, Phys. Rev. B **76**, 165106 (2007).

[103] F. Bechstedt, F. Fuchs, and G. Kresse, Phys. Status Solidi B **246**, 1877 (2009).

[104] D. X. Li, Y. Haga, H. Shida, and T. Suzuki, Physica B: Condensed Matter **199-200**, 631 (1994).

[105] F. Leuenberger, A. Parge, W. Felsch, K. Fauth, and M. Hessler, Phys. Rev. B **72**, 014427 (2005).

[106] C. Mitra and W. R. L. Lambrecht, Phys. Rev. B **78**, 134421 (2008).

[107] A. Sharma and W. Nolting, Phys. Rev. B **81**, 125303 (2010).

[108] C.-G. Duan, R. F. Sabiryanov, J. Liu, W. N. Mei, P. A. Dowben, and J. R. Hardy, Phys. Rev. Lett. **94**, 237201 (2005).

[109] C.-G. Duan, R. F. Sabiryanov, W. N. Mei, P. A. Dowben, S. S. Jaswal, and E. Y. Tsymbal, Appl. Phys. Lett. **88**, 182505 (pages 3) (2006).

[110] C.-G. Duan, R. F. Sabirianov, W. N. Mei, P. A. Dowben, S. S. Jaswal, and E. Y. Tsymbal, J. Phys.: Condens. Matter **19**, 315220 (2007).

[111] P. Wachter and E. Kaldis, Solid State Commun. **34**, 241 (1980).

[112] J. Q. Xiao and C. L. Chien, Phys. Rev. Lett. **76**, 1727 (1996).

[113] S. Granville, B. J. Ruck, F. Budde, A. Koo, D. J. Pringle, F. Kuchler, A. R. H. Preston, D. H. Housden, N. Lund, A. Bittar, et al., Phys. Rev. B **73**, 235335 (2006).

[114] B. M. Ludbrook, I. L. Farrell, M. Kuebel, B. J. Ruck, A. R. H. Preston, H. J. Trodahl, L. Ranno, R. J. Reeves, and S. M. Durbin, J. Appl. Phys. **106**, 063910 (pages 4) (2009).

[115] M. Scarpulla, C. Gallinat, S. Mack, J. Speck, and A. Gossard, J. Cryst. Growth **311**, 1239 (2009).

[116] W. R. L. Lambrecht, Phys. Rev. B **62**, 13538 (2000).

[117] P. Larson and W. R. L. Lambrecht, Phys. Rev. B **74**, 085108 (2006).

[118] P. Larson, W. R. L. Lambrecht, A. Chantis, and M. van Schilfgaarde, Phys. Rev. B **75**, 045114 (2007).

[119] K. Doll, J. Phys.: Condens. Matter **20**, 075214 (2008).

[120] R. C. Brown and N. J. Clark, J. Inorg. Nucl. Chem. **36**, 2507 (1974).

[121] U. Rössler and D. Strauch, *Group IV Elements, IV-IV and III-V Compounds. Part a - Lattice Properties*, vol. 41A1a of *Landolt-Börnstein - Group III Condensed Matter* (Springer-Verlag, Berlin, 2001).

[122] H. J. Trodahl, A. R. H. Preston, J. Zhong, B. J. Ruck, N. M. Strickland, C. Mitra, and W. R. L. Lambrecht, Phys. Rev. B **76**, 085211 (2007).

[123] B. N. Harmon, V. P. Antropov, A. I. Liechtenstein, I. V. Solovyev, and V. I. Anisimov, J. Phys. Chem. Solids **56**, 1521 (1995); proceedings of the 1994 Conference on Magneto-optic Materials.

[124] F. Hullinger, *Handbook on the Physics and Chemistry of Rare Earths*, vol. 4 (North-Holland Physics Publishing, New York, 1979).

[125] H. Yamada, T. Fukawa, T. Muro, Y. Tanaka, S. Imada, S. Suga, D.-X. Li, and T. Suzuki, J. Phys. Soc. Jpn. **65**, 1000 (1996).

[126] F. Lévy, Phys. Kondens. Mater. **10**, 71 (1969).

[127] M. Arnold and J. Kroha, Phys. Rev. Lett. **100**, 046404 (2008).

[128] N. N. Bogoliubov and S. V. Tyablikov, Dokl. Akad. Nauk SSSR **126**, 53 (1959).

[129] U. Nowak, in *Magnetism goes Nano*, edited by S. Blügel, T. Brückel, and C. M. Schneider (Forschungszentrum Jülich, Jülich, 2005), vol. 26 of *Matter and Materials*, p. A3.1.

[130] K. Khazen, H. J. von Bardeleben, J. L. Cantin, A. Bittar, S. Granville, H. J. Trodahl, and B. J. Ruck, Phys. Rev. B **74**, 245330 (2006).

[131] H. Yoshitomi, S. Kitayama, T. Kita, O. Wada, M. Fujisawa, H. Ohta, and T. Sakurai, Phys. Rev. B **83**, 155202 (2011).

[132] K. Senapati, T. Fix, M. E. Vickers, M. G. Blamire, and Z. H. Barber, Phys. Rev. B **83**, 014403 (2011).

[133] A. Punya, T. Cheiwchanchamnangij, A. Thiess, and W. R. L. Lambrecht, MRS Proceedings **1290** (2011).

[134] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).

[135] P. Gori, M. Rakel, C. Cobet, W. Richter, N. Esser, A. Hoffmann, R. Del Sole, A. Cricenti, and O. Pulci, Phys. Rev. B **81**, 125207 (2010).

[136] S. Tsoi, X. Lu, A. K. Ramdas, H. Alawadhi, M. Grimsditch, M. Cardona, and R. Lauck, Phys. Rev. B **74**, 165203 (2006).

[137] E. Sjöstedt, L. Nordström, and D. J. Singh, Solid State Commun. **114**, 15 (2000).

[138] T. J. Pollehn, A. Schindlmayr, and R. W. Godby, J. Phys.: Condens. Matter **10**, 1273 (1998).

[139] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[140] F. Freimuth, Y. Mokrousov, D. Wortmann, S. Heinze, and S. Blügel, Phys. Rev. B **78**, 035120 (2008).

[141] J. Kuneš, R. Arita, P. Wissgott, A. Toschi, H. Ikeda, and K. Held, Computer Physics Communications **181**, 1888 (2010).

[142] E. Şaşıoğlu, C. Friedrich, and S. Blügel, Phys. Rev. B **83**, 121101 (2011).

[143] E. Şaşıoğlu, A. Schindlmayr, C. Friedrich, F. Freimuth, and S. Blügel, Phys. Rev. B **81**, 054434 (2010).

[144] M. C. Müller, *Electron-magnon interaction in the GT approximation*, Master's thesis, Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich (2011), in preparation.