# MASTERARBEIT

zur Erlangung des akademischen Grades
Master of Science (M. Sc.)
im Fach Physik

# Implementation of Meta-GGA Functionals in the All-Electron Linearized Augmented Plane Wave Code exciting

Eingereicht am
Institut für Physik
der Humboldt-Universität zu Berlin am

14.10.2024

von

Elisa Marie Stephan

Gutachter/innen: Prof. Dr. Claudia Draxl
PD Dr. Denis Usvayat

# Abstract

Density Functional Theory (DFT) is a fundamental and powerful tool for electronic structure calculations. However, the accuracy and efficiency of DFT calculations strongly depend on the choice of the Density-Functional Approximation (DFA) used. Meta-Generalized Density Approximations (meta-GGAs), which incorporate the kinetic energy density, have gained increasing popularity because they maintain the computational efficiency of (semi-)local approximations such as Local Density Approximations (LDAs) and Generalized Gradient Approximations (GGAs), while improving the accuracy of various material properties, such as geometry, electronic structure, and binding energies, in both molecules and solids. This makes them a compelling alternative to LDAs and GGAs, or even to the more costly hybrid methods, especially when computational cost is a critical factor. In this work, meta-GGAs are implemented self-consistently in the all-electron (Linearized) Augumented Plane Wave + Local Orbital ((L)APW+LO) code `exciting`. The (L)APW+LO framework is known for its precision in solving the Kohn-Sham equations and is widely regarded as the gold standard for electronic structure calculations in DFT. The implementation covers both spin-unpolarized and spin-polarized cases, and the necessary formulas for both cases are derived and presented herein. The accuracy and correctness of the presented implementation are validated by calculating key material properties, such as band gaps, lattice constants, and magnetic moments, for a range of bulk materials using different meta-GGA functionals. These results are compared against literature results from established computational codes and put into context by comparing them to experimental data and higher-rung functionals like hybrids and lower-rung functionals like GGAs. Furthermore, the effectiveness of the implementation is also examined by evaluating the computational cost required for meta-GGA calculations, demonstrating that the inclusion of meta-GGAs only leads to a moderate increase in computational time of about a factor of 2-3 compared to GGA calculations. Together, these results demonstrate that the implementation presented here is both correct and efficient. They confirm that the meta-GGAs considered in this work not only improve upon GGAs in predicting lattice constants and band gaps, but also, in certain cases, yield results for these properties that are comparable to those obtained using much more computationally expensive hybrid functionals. Thus, the presented implementation of meta-GGAs lays the groundwork for the future application of meta-GGAs in the (L)APW+LO code `exciting`, allowing more accurate predictions with only a moderate increase in computational cost compared to lower-rung functionals like GGAs.

# Contents

# 1. Introduction

DFT [1, 2] has established itself as a widely used and powerful tool for electronic structure calculations. Although DFT is, in principle, an exact theory, its accuracy and computational efficiency are highly dependent on the exchange-correlation (xc) functional, $E_{\mathrm{xc}}[n]$. Since the exact form of $E_{\mathrm{xc}}[n]$ remains unknown, it must be approximated, which has led to the development of various xc functionals, commonly categorized within the so-called "Jacob's ladder" [3].
At the lowest rung of this ladder are the LDAs [3], which only depend on the local electron density $n(\mathbf{r})$. The GGAs [4], on the second rung, additionally incorporate the gradient of the density $\nabla n(\mathbf{r})$. On the third rung, the meta-GGAs [5] include the Kinetic Energy Density (KED) $\tau(\mathbf{r})$ or the Laplacian of the density $\nabla^2 n(\mathbf{r})$. Hybrid functionals [6], located on the fourth rung, mix in a portion of Hartree-Fock exact exchange. Each step up the ladder generally improves the accuracy of predictions, but this comes at the cost of increased computational demand and complexity [7].
Lower-rung (semi-)local functionals, such as LDAs, GGAs, and meta-GGAs, are widely used due to their balance between computational efficiency and accuracy. Among these, the Perdew-Burke-Ernzerhof (PBE) [8] functional has become a standard for electronic structure calculations. However, there has been a growing interest in meta-GGA functionals, particularly those that include the KED, as they often offer improved accuracy without the higher computational cost associated with hybrid functionals. For example, the Strongly Constrained and Appropriately Normed (SCAN) functional [9] has shown great performance in predicting several properties, including cohesive energies, formation energies, and geometries [10–12]. This makes the functional a compelling alternative to GGA functionals and hybrid methods, particularly when computational resources are limited. Meta-GGA functionals also have the advantage of being able to accurately describe both molecules and solids, something GGAs are not able to do [12, 13].
Additionally, when employed within the generalized Kohn-Sham (gKS) framework [14], orbital-dependent functionals like meta-GGAs can directly account for part of the xc derivative discontinuity $\Delta_{\mathrm{xc}}$ in the eigenvalues, resulting in more accurate band gap predictions [15]. This is a well-known limitation of LDAs and GGAs, which tend to strongly underestimate band gaps [16–18]. As a result, meta-GGAs, and especially those designed to exhibit a large xc derivative discontinuity $\Delta_{\mathrm{xc}}$, such as the TASK functional [19], can effectively address the common underestimation of band gaps seen with GGAs, yielding results that are closer to those of hybrid functionals [20, 21].
Thus, the primary goal of this work is to implement a self-consistent scheme for KED-dependent meta-GGAs within the gKS framework. To achieve this, the all-electron (L)APW+LO code `exciting` [22] is used. The (Linearized) Augumented Plane Wave ((L)APW) basis [23] is known for its accuracy when solving the Kohn-Sham (KS) equations, and in combination with Local Orbitals (LOs), it is considered the gold standard in DFT for electronic structure calculations. Moreover, as an all-electron method, `exciting` treats both valence and core electrons accurately without resorting to pseudopotentials [22]. The approach used for the implementation in this work follows a similar scheme to the one used in the `WIEN2k` code [24], which was presented in Ref. [7]. There, the meta-GGA potential is applied to the valence electrons only, and a GGA potential is

used for the underlying radial functions of the basis for both valence and core electrons.

The thesis is structured as follows. The theoretical background necessary to understand the implementation is provided in Chapter 2. This includes the fundamentals of DFT, the KS equations and generalized KS equations, spin-DFT, and the second-variational method. After this, Jacob's ladder of DFAs is explored, discussing the advantages and limitations of LDA, GGA, meta-GGA, and hybrid functionals. Additionally, the (L)APW+LO basis set is introduced, along with the basis used for the core states, which is obtained through the relativistic Dirac equations.
In Chapter 3, the necessary formulas to implement the KED, the key component of meta-GGA functionals, are derived for all basis types used in `exciting`. Following this, the necessary formulas, which include the xc potential, the Hamiltonian matrix elements, and the total meta-GGA energy, for both spin-unpolarized and spin-polarized cases are presented. In Chapter 4, the computational performance of the implementation is assessed by evaluating the run time for various calculations. Then, the implementation of the KED is tested, and finally, key properties - such as band gaps, lattice constants, and magnetic moments - are calculated with the new implementation across a range of meta-GGA functionals and materials, and compared with experimental data and results of other established codes. The thesis concludes with a summary of the main findings and an outlook on future directions for improving and expanding the implementation and use of meta-GGAs within `exciting`.

# 2. Theoretical Background

To evaluate the properties of a material, the electronic many-body wave function must be determined. This is done by solving the Schrödinger equation for a system containing $N$ electrons:

$$\hat{H}\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N), \tag{2.1}$$

with the many-body Hamiltonian defined as

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &- \sum_\alpha \frac{1}{2M_\alpha}\nabla_\alpha^2 + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|},\end{aligned} \tag{2.2}$$

where atomic units were used, such that $\hbar = m_e = 4\pi/\epsilon_0 = 1$. Here, $Z_\alpha$ and $M_\alpha$ are the atomic number and mass of the $\alpha$-th atomic ion respectively, $\mathbf{R}_\alpha$ is the atomic coordinate, and $\mathbf{r}_i$ is the coordinate of the $i$-th electron [25]. Due to the large mass difference between ions and electrons ($m_\mathrm{p}/m_\mathrm{e} \approx 1.8 \times 10^3$), ions can be considered fixed while solving the electronic Schrödinger equation. This simplification allows for the separation of nuclear and electronic motions, where the electronic Hamiltonian is solved for a static configuration of nuclei, effectively ignoring the kinetic energy of the nuclei. This is known as the Born-Oppenheimer approximation [26]. The electronic Hamiltonian is thus given by [25]:

$$\hat{H} = \sum_i \left[-\frac{1}{2}\nabla_i^2 + \sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|}\right] + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.3}$$

consisting of the kinetic energy of the electrons, the external electrostatic potential acting on the electrons $V_\mathrm{ext} = \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|}$, and the electron-electron interaction.

A direct solution of the Hamiltonian in Eq. (2.3) requires an enormous amount of computational storage. To store the three-dimensional many-body wave function for a system with $N$ electrons, the computational grid must be discretized with $M$ points in each spatial direction. This results in a total of $M^{3N}$ points. For example, even with a coarse grid of $M = 20$ points per direction, a system like oxygen with $N = 8$ electrons would require storing $20^{24}$ data points. In double floating precision, where each point requires 8 bytes, this results in $20^{24} \times 8$ bytes, or approximately $10^{17}$ petabytes, making the storage of the wave function both infeasible and impractical [27].

To overcome this challenge, DFT has emerged as a powerful and widely used approach that offers a balance between accuracy and computational efficiency. The following sections provide an overview of the fundamentals of DFT and its practical implementations, focusing on the role of xc functionals, in particular the meta-GGA functionals, and introducing the Linearized Augmented Plane Wave (LAPW) method.

## 2.1. Density-Functional-Theory

In DFT, the complexity of solving the many-body Schrödinger equation, at least for the ground state, is significantly reduced by using the ground state electron density $n(\mathbf{r})$ rather than the many-body wave function $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. The storage required for the electron density is $M^3$ data points, where $M$ is again the number of grid points in each spatial direction. Thus, the density requires $\frac{M^{3N}}{M^3} = M^{3(N-1)}$ times less storage than the wave function. Returning to the previous example with $M = 20$ and $N = 8$, this means the density would need approximately $10^{27}$ times less storage than the wave function. The approach of using the electron density instead of the wave function is grounded in the work of P. Hohenberg and W. Kohn from 1964 [1], who demonstrated that there is a functional mapping of the ground state density to the total energy of a many-body system. Their theorems laid the foundation for DFT, providing a framework in which the ground state energy is expressed as a functional of the electron density.
Building on this foundation, Kohn and Sham later derived the KS equations [2], which, when solved self-consistently, yield the exact ground state density. Beyond the standard KS formulation, the gKS method further provides a flexible mathematical framework that allows for the inclusion of a wider range of xc functionals, particularly those that involve orbital-dependent or non-local terms, such as meta-GGA or hybrid functionals. For spin-polarized systems, Spin-Density-Functional-Theory (SDFT), introduced by Barth and Hedin [28], incorporates spin degrees of freedom into the DFT framework. To efficiently solve spin-polarized calculations, the second-variational scheme is employed in `exciting`. In the following sections, these key theoretical concepts will be discussed in more detail.

### 2.1.1. Hohenberg-Kohn Theorems

The first Hohenberg-Kohn theorem states that the external potential $V_{\text{ext}}(\mathbf{r})$ is uniquely determined by the ground-state density $n(\mathbf{r})$ [1], assuming the system is in a non-degenerate ground state. Hohenberg and Kohn demonstrated that if two external potentials differ by more than a constant, they cannot lead to the same ground-state density. Therefore, $V_{\text{ext}}(\mathbf{r})$ can be written as a functional of the density:

$$V_{\text{ext}}(\mathbf{r}) = V_{\text{ext}}[n](\mathbf{r}). \tag{2.4}$$

Since the external potential determines the Hamiltonian, the many-body wave function becomes a functional of the density. As a result, observables such as the total energy $E[n]$ can be evaluated as functionals of $n(\mathbf{r})$.
The second Hohenberg-Kohn theorem states that there exists a universal functional of the electron density for the total energy $E[n]$, which is valid for any external potential:

$$E[n] = F[n] + \int d\mathbf{r}\, n(\mathbf{r})\, V_{\text{ext}}(\mathbf{r}), \tag{2.5}$$

where $F[n]$ is the universal functional that accounts for the internal energies (kinetic and electron-electron interaction) of the electron system, and $V_{\text{ext}}(\mathbf{r})$ is the external potential. Specifically, the theorem states that the ground state energy is obtained by minimizing the total energy functional under the constraint of electron number conservation, with the minimizing density being the ground state density.

### 2.1.2. The Kohn-Sham System

In 1965, Kohn and Sham [2] provided a way to determine the universal functional $F[n]$ by mapping the complex interacting many-body problem onto a simpler, non-interacting single-particle problem. They postulated that there exists an effective potential $V_{\text{eff}}(\mathbf{r})$ such that the ground-state density of a fictitious non-interacting system under this potential would match the ground-state density of the real interacting system. Kohn and Sham formulated the total energy of the interacting system as:

$$E[n] = T_{\text{s}}[n] + \int d\mathbf{r}\, n(\mathbf{r})\, V_{\text{ext}}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}' \, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{\text{xc}}[n], \tag{2.6}$$

where $T_{\text{s}}[n]$ is the kinetic energy of a non-interacting system, and $E_{\text{xc}}[n]$ is the xc energy, which accounts for the difference between the true interacting kinetic energy and $T_{\text{s}}[n]$, as well as exchange and correlation effects. Minimizing the functional under the constraint of electron number conservation leads to

$$\int d\mathbf{r} \left( \frac{\delta T_{\text{s}}[n]}{\delta n(\mathbf{r})} + V_{\text{eff}}(\mathbf{r}) \right) \delta n(\mathbf{r}) = 0, \tag{2.7}$$

with the effective potential $V_{\text{eff}}(\mathbf{r})$ given by:

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{\text{xc}}(\mathbf{r}). \tag{2.8}$$

Here, $V_{\text{H}}(\mathbf{r})$ is the Hartree potential and $V_{\text{xc}}(\mathbf{r})$ is the xc potential, respectively defined as:

$$V_{\text{H}}(\mathbf{r}) = \int d\mathbf{r}' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \quad V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})}. \tag{2.9}$$

Kohn and Sham argued that a non-interacting system subjected to an effective single-particle potential $V_{\text{eff}}(\mathbf{r})$ would yield the same variational equation (2.7). Consequently, they concluded that the ground-state density of the interacting system can be obtained by solving single-particle Schrödinger equations:

$$\left[ -\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}). \tag{2.10}$$

The ground-state density of the interacting system is then given by the sum over all the occupied single-particle states $(n, \mathbf{k})$:

$$n(\mathbf{r}) = \sum_{n\mathbf{k}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2. \tag{2.11}$$

The single-particle equations (2.10) are known as the KS equations and are solved iteratively within a Self-Consistent Field (SCF) cycle, as shown in Fig. 2.1. This iterative process is necessary because the potential depends on the electron density itself.
In the initial step ($i = 0$) of the SCF cycle, an initial guess for the electron density $n^0(\mathbf{r})$ is made, typically derived from atomic densities or a previously converged calculation. From this initial density, the total effective potential $V_{\text{eff}}(\mathbf{r})$ is then calculated. With the effective potential determined, the KS equations are solved to obtain the KS orbitals $\psi_{n\mathbf{k}}$ and KS eigenvalues $\varepsilon_{n\mathbf{k}}$. These orbitals are then used to evaluate the new electron density $n(\mathbf{r})$. Starting from the first iteration ($i \geq 1$), convergence is checked by comparing the

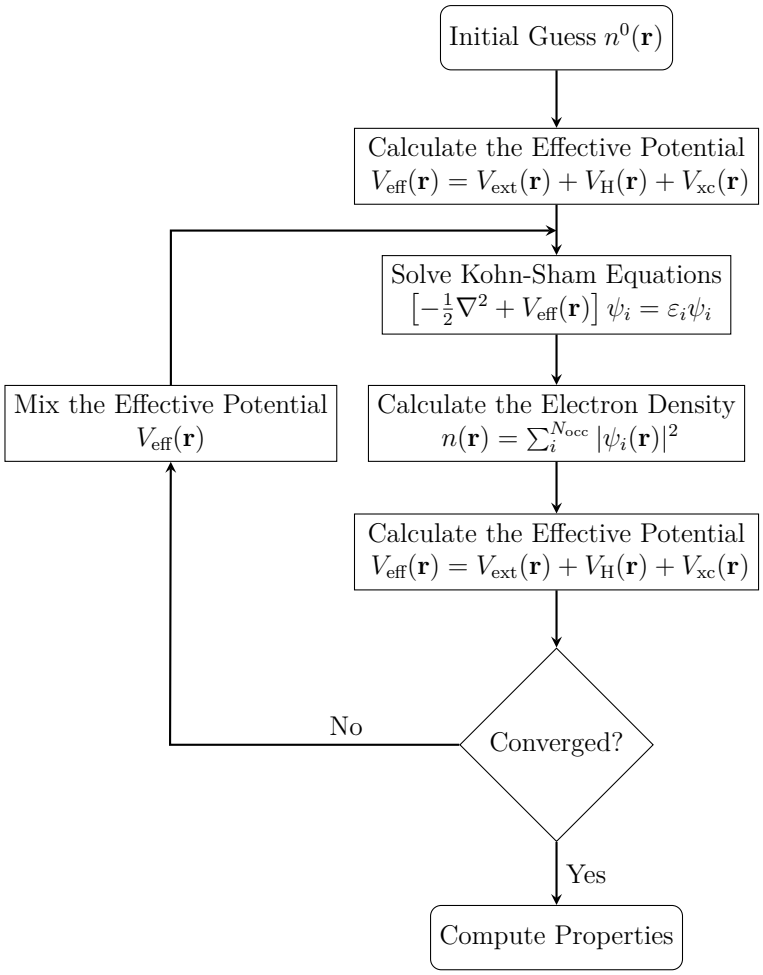



Figure 2.1.: Flowchart of the SCF cycle.

electron density or the effective potential of the current iteration $i$ with that of the previous iteration $i-1$. Convergence is achieved if $|n^i(\mathbf{r}) - n^{i-1}(\mathbf{r})| < \epsilon$ or $|V^i_{\text{eff}}(\mathbf{r}) - V^{i-1}_{\text{eff}}(\mathbf{r})| < \epsilon$, where $\epsilon$ is a predefined convergence criterion. Additionally, the convergence of the total energy can be checked. If the cycle has not converged, the process is repeated. To improve convergence stability, the densities or potentials from consecutive iterations are mixed using different mixing schemes. A basic approach is linear mixing, where the mixed potential, $V^{i+1}_{\text{eff, in}}$ for the next iteration $i+1$, is given by:

$$V^{i+1}_{\text{eff, in}} = (1-\alpha)V^i_{\text{eff, in}} + \alpha V^i_{\text{eff, out}}, \tag{2.12}$$

where $\alpha$ is the mixing parameter [23]. While linear mixing is simple, it often converges slowly, especially in complex systems. To overcome this, the multi-secant Broyden scheme [29] is commonly used, which improves convergence behavior by utilizing the history of previous iterations to construct a Jacobian for more efficient updates. The iterative procedure continues until the electron density or potential is self-consistent.
Theoretically, the ground-state density obtained from density functional theory is exact. However, the accuracy of the results depends on the xc functional $E_{\text{xc}}[n]$, whose exact form remains unknown. To address this, various approximations have been developed that balance computational efficiency with accuracy, including LDA, GGA, meta-GGA, and hybrid functionals, which will be further discussed in Sec. 2.2.

### 2.1.3. The Generalized Kohn-Sham System

While the Kohn-Sham scheme accurately reproduces the ground-state density of an interacting electron system using a fictitious system of non-interacting electrons, challenges arise when the xc functionals are orbital-dependent instead of purely density-dependent. This is the case for orbital-dependent meta-GGAs and hybrid functionals. For these functionals, evaluating the xc potential by taking the functional derivative with respect to the density, as defined in Eq. (2.9), would require the usage of the Optimized Effective Potential (OEP) method [30, 31], which involves solving a nontrivial integro-differential equation [32]. The OEP method then yields a local KS xc potential, but solving the equation is known to be computationally demanding and often unstable [33].
In 1996, Seidl et al. [14] proposed an alternative to the Kohn-Sham framework, where they showed that the non-interacting reference system introduced by Kohn and Sham is only one choice of many. They proposed an alternative framework where the reference system still uses a single Slater determinant of one-particle orbitals but includes some xc effects through a nonlocal, orbital-dependent potential. This approach maps the real interacting system to a partially interacting system rather than a fully non-interacting one. In the gKS scheme, an energy functional $S[\{\psi_i\}]$ of the $N$ orbitals that construct the Slater determinant is defined, with $i$ representing the combined index of $n$ and $\mathbf{k}$. From this, an energy functional $F^S[n]$ can be constructed, which is obtained from the orbitals that minimize $S$, while yielding the density $n(\mathbf{r})$:

$$F^S[n] = \min_{\{\psi_i\}\to n} S[\{\psi_i\}]. \tag{2.13}$$

The difference between the functional $F^S[n]$ and the previously defined universal Hohenberg-Kohn functional $F[n]$ can then be defined as:

$$F[n] = F^S[n] + R^S[n], \tag{2.14}$$

where $R^S[n]$ is the so-called remainder functional. The exact form of $R^S[n]$ is not known, but it can be approximated depending on how $S[\{\psi_i\}]$ is defined [14]. Using this, the total energy can be reformulated as:

$$E[n] = F^S[n] + R^S[n] + \int d\mathbf{r}\, n(\mathbf{r})\, V_{\text{ext}}(\mathbf{r}). \tag{2.15}$$

While $R^S[n]$ and the integral over the external potential depend only indirectly on the orbitals $\{\psi_i\}$, the total energy in the generalized Kohn-Sham scheme is minimized with respect to the orbitals, under the constraint that they yield the correct ground-state density. The resulting generalized KS equations are then given by [32]:

$$\left(\hat{O}^S[\{\psi_i\}] + V_R(\mathbf{r}) + V_{\text{ext}}(\mathbf{r})\right)\psi_j = \varepsilon_j \psi_j, \tag{2.16}$$

where $V_R(\mathbf{r}) = \frac{\delta R^S[n]}{\delta n(\mathbf{r})}$ is the so-called remainder potential and $\hat{O}^S[\{\psi_i\}]$ is, in general, a nonlocal operator.
In principle, both the KS and gKS mappings are exact [14]. However, while there is only one KS map, which is that of the non-interacting system, there are different possible gKS mappings depending on the choice of $S[\{\psi_i\}]$ [34]. For example, if one defines $S[\{\psi_i\}]$ as the non-interacting kinetic energy, then $\hat{O}^S[\{\psi_i\}]$ yields the kinetic operator already defined in the KS equations. The remainder functional can then be defined as the sum

of the Hartree energy and the xc energy. In this case, the KS mapping reduces to the standard KS scheme, and one can view the standard KS approach as a special case within the generalized gKS scheme [32, 34].
Additionally, any contributions from the true interacting kinetic energy can now be partially included in the operator $\hat{O}^S[\{\psi_i\}]$ [35]. As will be discussed in Sec. 2.2.5, this can lead to more realistic predictions of band gaps when they are directly evaluated from the eigenvalues of the system.
To conclude, the gKS scheme provides a framework to evaluate orbital-dependent functionals with respect to the orbitals, rather than the density, which is more efficient than using the standard scheme with the OEP method. Thus, we will explore the gKS mapping of orbital-dependent meta-GGAs in more detail in Sec. 3.2.1.

### 2.1.4. Spin-Density-Functional Theory

For spin-polarized systems, Barth and Hedin formulated the SDFT [28]. Spin-polarized systems can be classified into collinear and non-collinear systems. Many studied materials exhibit a collinear spin density, meaning that the magnetization of the system has a fixed spatial orientation along a particular axis, which is typically chosen as the z-axis. In such systems, the KS wave functions can be separated into spin-up and spin-down orbitals. Typical examples include conventional ferromagnets and antiferromagnets. In contrast, non-collinear magnetic systems have a magnetic orientation that can vary in space. In these systems, choosing spin as a quantum number is no longer a good choice, since the spin of the orbitals is no longer restricted to being either "up" or "down" [36]. An example of a non-collinear magnetic system is $\gamma$-Fe, where helical spin density waves or spin spirals form in the ground state [37].
The following section introduces SDFT for the general non-collinear case, which is equally valid for collinear systems in the collinear limit. In non-collinear cases, the KS orbitals are expressed as two-component spinors [28]:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \begin{pmatrix} \psi_{n\mathbf{k},\sigma}(\mathbf{r}) \\ \psi_{n\mathbf{k},\sigma'}(\mathbf{r}) \end{pmatrix}.$$

This leads to two coupled equations:

$$\sum_{\sigma'} \left[ -\frac{\nabla^2}{2}\delta_{\sigma\sigma'} + V^{\text{eff}}_{\sigma\sigma'}(\mathbf{r}) \right] \psi_{n\mathbf{k},\sigma'}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} \psi_{n\mathbf{k},\sigma}(\mathbf{r}), \tag{2.17}$$

where $V^{\text{eff}}_{\sigma\sigma'}$ is the effective potential consisting of the usual external, Hartree, and xc contributions:

$$V^{\text{eff}}_{\sigma\sigma'}(\mathbf{r}) = V^{\text{H}}(\mathbf{r})\delta_{\sigma\sigma'} + V^{\text{ext}}_{\sigma\sigma'}(\mathbf{r}) + V^{\text{xc}}_{\sigma\sigma'}(\mathbf{r}). \tag{2.18}$$

The kinetic energy and the Hartree potential contribute only to the diagonal elements of the Hamiltonian, whereas the external potential and the xc potential can have non-diagonal components. The xc potential is given by:

$$V^{\text{xc}}_{\sigma\sigma'} = \frac{\delta E_{\text{xc}}[\bar{n}]}{\delta n_{\sigma\sigma'}(\mathbf{r})}, \tag{2.19}$$

where the density is a $2 \times 2$ Hermitian matrix $\bar{n}(\mathbf{r})$, defined as [36]:

$$\bar{n}(\mathbf{r}) = \sum_{n\mathbf{k}} \psi^{\dagger}_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}) = \begin{pmatrix} n_{\sigma\sigma}(\mathbf{r}) & n_{\sigma\sigma'}(\mathbf{r}) \\ n_{\sigma'\sigma}(\mathbf{r}) & n_{\sigma'\sigma'}(\mathbf{r}) \end{pmatrix}. \tag{2.20}$$

Alternatively, the system can be described using a four-component vector $\mathbf{n}(\mathbf{r})$, which is defined as:

$$\mathbf{n}(\mathbf{r}) = \sum_{i=0}^{3} \mathrm{tr}\{\sigma_i \bar{n}(\mathbf{r})\}, \tag{2.21}$$

where $\sigma_i$ with $i = 1, 2, 3$ are the Pauli matrices and $\sigma_0$ is the identity matrix. This leads to the following expression:

$$\mathbf{n}(\mathbf{r}) = \begin{bmatrix} n_{\sigma\sigma}(\mathbf{r}) + n_{\sigma'\sigma'}(\mathbf{r}) \\ n_{\sigma\sigma'}(\mathbf{r}) + n_{\sigma'\sigma}(\mathbf{r}) \\ i\big(n_{\sigma\sigma'}(\mathbf{r}) - n_{\sigma'\sigma}(\mathbf{r})\big) \\ n_{\sigma\sigma}(\mathbf{r}) - n_{\sigma'\sigma'}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} n(\mathbf{r}) \\ m_x(\mathbf{r}) \\ m_y(\mathbf{r}) \\ m_z(\mathbf{r}) \end{bmatrix}, \tag{2.22}$$

where we now define the zeroth component of the vector as the scalar density $n(\mathbf{r})$, and the remaining components describe the direction and magnitude of the magnetization $\mathbf{m}(\mathbf{r})$. The total spin density matrix can thus be redefined in terms of the scalar density and magnetization as:

$$\bar{n}(\mathbf{r}) = \frac{1}{2}\sum_{j=0}^{3} \sigma_j n_j(\mathbf{r}) = \frac{1}{2}\left(\sigma_0\, n(\mathbf{r}) + \sum_{j=1}^{3} \sigma_j m_j(\mathbf{r})\right), \tag{2.23}$$

which results in

$$\bar{n}(\mathbf{r}) = \frac{1}{2}\begin{pmatrix} n(\mathbf{r}) + m_z(\mathbf{r}) & m_x(\mathbf{r}) - i m_y(\mathbf{r}) \\ m_x(\mathbf{r}) + i m_y(\mathbf{r}) & n(\mathbf{r}) - m_z(\mathbf{r}) \end{pmatrix}. \tag{2.24}$$

The KS equation defined in Eq. (2.17) can then be similarly rewritten as [36, 38]:

$$\left[-\frac{\nabla^2}{2} + \mathbf{V}_{\mathrm{eff}}(\mathbf{r}) \cdot \boldsymbol{\sigma}\right] \psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}), \tag{2.25}$$

where $\boldsymbol{\sigma}$ is again a four-component vector consisting of the identity matrix and the Pauli matrices, and $\mathbf{V}_{\mathrm{eff}}(\mathbf{r})$ is given by:

$$\mathbf{V}^{\mathrm{eff}}(\mathbf{r}) = \begin{bmatrix} V^{\mathrm{eff}}_{\sigma\sigma}(\mathbf{r}) + V^{\mathrm{eff}}_{\sigma'\sigma'}(\mathbf{r}) \\ V^{\mathrm{eff}}_{\sigma\sigma'}(\mathbf{r}) + V^{\mathrm{eff}}_{\sigma'\sigma}(\mathbf{r}) \\ i\big(V^{\mathrm{eff}}_{\sigma\sigma'}(\mathbf{r}) - V^{\mathrm{eff}}_{\sigma'\sigma}(\mathbf{r})\big) \\ V^{\mathrm{eff}}_{\sigma\sigma}(\mathbf{r}) - V^{\mathrm{eff}}_{\sigma'\sigma'}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} V^{\mathrm{eff}}(\mathbf{r}) \\ B^{\mathrm{eff}}_x(\mathbf{r}) \\ B^{\mathrm{eff}}_y(\mathbf{r}) \\ B^{\mathrm{eff}}_z(\mathbf{r}) \end{bmatrix}, \tag{2.26}$$

where $V^{\mathrm{eff}}(\mathbf{r})$ is the scalar effective potential and $\mathbf{B}^{\mathrm{eff}}(\mathbf{r})$ is the effective magnetic field. With the four-component vector defined in Eq. (2.26), the total $2\times2$ matrix $\bar{V}^{\mathrm{eff}}(\mathbf{r})$ in the KS-equation in Eq. (2.17) can then be expressed in terms of the four-component vector as:

$$\bar{V}^{\mathrm{eff}}(\mathbf{r}) = \frac{1}{2}\sum_{j=0}^{3} \sigma_j V^{\mathrm{eff}}_j(\mathbf{r}) = \frac{1}{2}\left(\sigma_0 V^{\mathrm{eff}}(\mathbf{r}) + \sum_{j=1}^{3} \sigma_j B^{\mathrm{eff}}_j(\mathbf{r})\right), \tag{2.27}$$

leading to a structure that is similar to the spin density matrix, defined in Eq. (2.24). The Hartree potential is scalar and only contributes to the scalar effective potential.

However, the xc potential contributes to both the scalar potential and the magnetic field and can be split into the scalar xc potential and the magnetic xc field, defined as:

$$V^{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n, \mathbf{m}]}{\delta n(\mathbf{r})}, \quad B_i^{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n, \mathbf{m}]}{\delta m_i(\mathbf{r})}, \quad i = \{1, 2, 3\}. \tag{2.28}$$

However, since the xc functionals $E_{\mathrm{xc}}[n, \mathbf{m}]$ now depend on both the scalar density $n(\mathbf{r})$ and the magnetization $\mathbf{m}(\mathbf{r})$, their construction becomes more complex. In collinear systems, where the magnetic field is aligned along the spin quantization axis $z$, the off-diagonal components of the spin density matrix are zero. It follows from Eq. (2.24) that the total density can then be constructed from the spin-up and spin-down densities [39]:

$$n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}), \tag{2.29}$$

given by

$$n_{\uparrow/\downarrow}(\mathbf{r}) = \frac{1}{2}\big(n(\mathbf{r}) \pm m_z(\mathbf{r})\big). \tag{2.30}$$

The collinear xc functionals can then be constructed to depend on the spin-up and spin-down densities instead of the scalar density and magnetization as follows [40]:

$$E_{\mathrm{xc}}[n, \mathbf{m}] \to E_{\mathrm{xc}}[n, m_z] = E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]. \tag{2.31}$$

In non-collinear systems, the off-diagonal components of the spin density matrix contribute, and consequently, the density cannot simply be described by spin-up and spin-down components. As a result, xc functionals designed for the collinear case cannot, in principle, be directly applied to the non-collinear case. However, the most common approach has been to still use the xc energy functionals from collinear SDFT and apply them to the non-collinear situation. This approach was pioneered by Kübler in 1988 [41], where he defined a local reference frame in which the total spin density matrix $\bar{n}$ is brought to a diagonal form composed of $n_\uparrow$ and $n_\downarrow$ [40]:

$$\bar{n}(\mathbf{r}) \Rightarrow \begin{pmatrix} n_\uparrow & 0 \\ 0 & n_\downarrow \end{pmatrix}, \tag{2.32}$$

where $n_{\uparrow/\downarrow}$ are the eigenvalues of the spin density matrix:

$$n_{\uparrow/\downarrow}(\mathbf{r}) = \frac{1}{2}\left(n(\mathbf{r}) \pm \sqrt{m_x^2(\mathbf{r}) + m_y^2(\mathbf{r}) + m_z^2(\mathbf{r})}\right). \tag{2.33}$$

Consequently, the collinear xc functionals can be used as:

$$E_{\mathrm{xc}}[n, \mathbf{m}] \to E_{\mathrm{xc}}[n, m_x, m_y, m_z] = E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]. \tag{2.34}$$

When defining $n_{\uparrow/\downarrow}$ in the non-collinear case, it is crucial that they recover the collinear case, which is clearly achieved here when $m_x = m_y = 0$. This ensures that the definition is applicable to both collinear and non-collinear cases.

### 2.1.5. Second-Variational Treatment

To treat spin-polarized systems, a two-step procedure is used in `exciting`. As derived in Eq. (2.27), the effective potential is divided into a scalar effective potential and an effective

magnetic field. Consequently, the Hamiltonian can be decomposed as follows [22]:

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{H}}_{\mathrm{SR}} + \hat{\boldsymbol{H}}_{\mathrm{B}}, \tag{2.35}$$

where $\hat{\boldsymbol{H}}_{\mathrm{SR}}$ represents the scalar (relativistic) Hamiltonian, consisting of the kinetic energy and the scalar potential $V^{\mathrm{eff}}(\mathbf{r})$, contributing only to the diagonal components of the total Hamiltonian. Relativistic effects can, for example, be accounted for by using the zeroth-order regular approximation (ZORA) [42, 43]. $\hat{\boldsymbol{H}}_{\mathrm{B}}$ is defined as:

$$\hat{\boldsymbol{H}}_{\mathrm{B}} = \boldsymbol{\sigma} \cdot (\mathbf{B}_{\mathrm{ext}} + \mathbf{B}_{\mathrm{xc}}), \tag{2.36}$$

consisting of the external magnetic field $\mathbf{B}_{\mathrm{ext}}$ and the exchange-correlation magnetic field $\mathbf{B}_{\mathrm{xc}}$. The external magnetic field, $\mathbf{B}_{\mathrm{ext}}$, perturbs the system from its initial paramagnetic state, breaking the original symmetry and helps develop a magnetic phase [22].
While the scalar (relativistic) Hamiltonian contributes only to the diagonal elements, the magnetic Hamiltonian couples the two spin components. Therefore, the spin-dependent Hamiltonian elements can be expressed as [44]:

$$\hat{H}_{\sigma\sigma'} = \delta_{\sigma\sigma'}\hat{H}_{\mathrm{SR},\sigma} + \hat{H}_{\mathrm{B},\sigma\sigma'}. \tag{2.37}$$

It is useful to avoid solving the full Hamiltonian with its full basis $N_b$ from the beginning, but instead to focus on the scalar (relativistic) part first. The direct solution of the full Hamiltonian for both spin components leads to a $2N_b \times 2N_b$ secular equation due to the spin. Since the diagonalization scales with $\mathcal{O}(N_b^3)$ [22], which is approximately eight times more expensive than the non-spin-polarized problem.
The second-variational approach allows us to first solve the scalar (relativistic) Hamiltonian $\hat{\boldsymbol{H}}_{\mathrm{SR}}$ in the so-called first-variational step and to then treat any non-collinear magnetic effects in the second step. First, the scalar (relativistic) Hamiltonian matrix is diagonalized to obtain the first-variational eigenvectors $\psi^{\mathrm{FV}}_{n\mathbf{k}}$. Then, a new spinor basis $\psi^{\mathrm{FV}}_{n\mathbf{k},\sigma}$ is constructed, built from the $N_f$ lowest-lying Kohn-Sham orbitals $\psi^{\mathrm{FV}}_{n\mathbf{k}}$ from the first-variational step. With the new spinor basis, the total Hamiltonian is then solved. Thus, the basis size in the second-variational step is reduced to $2N_f$, where the factor of 2 again accounts for the spin. This problem is computationally more efficient [22, 23]. Once the first-variational eigenvectors are obtained, the components of the total Hamiltonian are constructed as [44]:

$$\langle \psi^{\mathrm{FV}}_{i\mathbf{k},\sigma} | \, \hat{H}_{\sigma\sigma'} \, | \psi^{\mathrm{FV}}_{j\mathbf{k},\sigma'} \rangle = \delta_{\sigma\sigma'}\delta_{ij}\varepsilon^{\mathrm{FV}}_{i\mathbf{k},\sigma} + \langle \psi^{\mathrm{FV}}_{i\mathbf{k},\sigma} | \, \hat{H}_{\mathrm{B},\sigma\sigma'} \, | \psi^{\mathrm{FV}}_{j\mathbf{k},\sigma'} \rangle , \tag{2.38}$$

where the first term represents the contribution from the scalar (relativistic) Hamiltonian, and the second term includes the contributions from the magnetic Hamiltonian $\hat{\boldsymbol{H}}_{\mathrm{B}}$. Finally, the second-variational eigenvectors and eigenvalues are obtained by diagonalizing this Hamiltonian matrix, providing the complete solution to the system.

## 2.2. Exchange-Correlation Functionals

Apart from the Born-Oppenheimer approximation, no additional approximations have been used to derive the Kohn-Sham equations for both the spin-unpolarized and spin-polarized cases. If the exact xc potential $V_{\mathrm{xc}}(\mathbf{r})$ were known, the KS scheme would provide the exact ground-state energy and electron density. Therefore, the accuracy of DFT cal-

culations depends fundamentally on the DFA used. Identifying an accurate DFA remains one of the primary challenges in DFT. Perdew introduced the concept of the "Jacob's Ladder" [3] to categorize various DFAs based on their complexity, accuracy, and computational cost, organizing them into different rungs. This section discusses the functionals of the initial rungs of the ladder, including their construction and performance in solid-state calculations.

### 2.2.1. Local Density Approximation

The first and simplest rung of Jacob's Ladder is the LDA. The LDA is defined as follows [3, 45]:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{\mathrm{xc}}^{\mathrm{uniform}}(n(\mathbf{r})), \tag{2.39}$$

where $\varepsilon_{\mathrm{xc}}^{\mathrm{uniform}}(n)$ is the xc energy per electron of a homogeneous electron gas with density $n$. The exchange part is known analytically and given by:

$$\varepsilon_{\mathrm{x}}^{\mathrm{uniform}}(n(\mathbf{r})) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\mathbf{r})^{\frac{1}{3}}. \tag{2.40}$$

The correlation part is often estimated using Quantum Monte Carlo methods [46].
By construction, LDA is only valid for systems with uniform or very slowly-varying densities. However, it has been shown that LDA can predict some ground-state properties with surprising accuracy, even for non-homogeneous densities [27]. For instance, LDA can accurately predict lattice parameters for different classes of materials [11, 47]. For weakly correlated systems, such as metals, LDA describes structural and vibrational properties quite well [48]. Despite these strengths, LDA has notable limitations. It can incorrectly predict lattice types or the magnetic ordering of the ground-state. For example, for Fe, LDA predicts a ground state geometry of Face-Centered Cubic (FCC) instead of ferromagnetic Body-Centered Cubic (BCC) [49]. Additionally, LDA tends to overestimate cohesive energies [50] and often underestimates band gaps, sometimes predicting metallic instead of semiconducting behavior [16, 17].

### 2.2.2. Generalized Gradient Approximation

To account for density inhomogeneity, GGAs extend the LDAs by introducing an additional dependence on the gradients of the density, making them semilocal functionals [4]:

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{\mathrm{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r})). \tag{2.41}$$

The exchange of GGAs is constructed as [4]:

$$E_{\mathrm{x}}^{\mathrm{GGA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{\mathrm{x}}^{\mathrm{uniform}}(n(\mathbf{r}))\, F_{\mathrm{x}}^{\mathrm{GGA}}(s(\mathbf{r})), \tag{2.42}$$

where $F_{\mathrm{x}}^{\mathrm{GGA}}(s)$ is the enhancement factor, which describes the deviation from $\varepsilon_{\mathrm{x}}^{\mathrm{uniform}}(n)$. From the gradient of the density $\nabla n$, the dimensionless variable $s = |\nabla n|/(2k_F n)$, known as the reduced density gradient, can be constructed, measuring the local density inhomogeneity [51].
The best-known and successful GGA functional is the PBE functional [8], which is widely

used in the physics community. Another well-known functional is PBEsol [52], a modification of PBE specifically designed to better satisfy physical constraints relevant to solids. PBE manages to reduce the overbinding seen in LDA but tends to overestimate lattice constants [53, 54]. This overestimation is improved by PBEsol [54]. However, for band gaps, GGA functionals show no significant improvement over LDA functionals [18].

### 2.2.3. Meta-Generalized Gradient Approximation

On the third rung of Jacob's Ladder are the meta-GGA functionals, which additionally account for the Laplacian of the electron density and/or the KED [5]:

$$E_{\mathrm{xc}}^{\mathrm{mGGA}} = \int d\mathbf{r}\, \varepsilon_{\mathrm{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r}), \nabla^2 n(\mathbf{r}), \tau(\mathbf{r})), \tag{2.43}$$

where $\tau(\mathbf{r}) = \frac{1}{2}\sum_i^{N_{\mathrm{occ}}} |\nabla\psi_i(\mathbf{r})|^2$ is the non-interacting KED. Many meta-GGA functionals do not take into account the Laplacian of the density $\nabla^2 n(\mathbf{r})$, partly due to the stronger sensitivity of $\nabla^2 n(\mathbf{r})$ to the integration grid, making most published meta-GGA functionals depend only on $\tau(\mathbf{r})$ [55].
The most prominent meta-GGA functional is the SCAN functional, proposed by Sun et al. [9]. SCAN satisfies all 17 known mathematical constraints that can be imposed on a meta-GGA functional, which are detailed in the supplementary information of Ref. [9]. Additionally, it is appropriately normed, ensuring accuracy for specific systems where exact results are known. The SCAN exchange energy can be expressed as:

$$E_{\mathrm{x}}^{\mathrm{SCAN}} = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{\mathrm{x}}^{\mathrm{uniform}}(n(\mathbf{r}))\, F_{\mathrm{x}}^{\mathrm{SCAN}}(s(\mathbf{r}), \alpha(\mathbf{r})), \tag{2.44}$$

where $F_{\mathrm{x}}^{\mathrm{SCAN}}(s, \alpha)$ is the enhancement factor, which depends on the density gradient $s$ and another dimensionless variable, the iso-orbital indicator $\alpha$.
The iso-orbital indicator $\alpha$ is related to the KED $\tau$ as follows. The KED can be expressed in terms of the density in two limits. For the one- and two-electron limit, $\tau$ reduces to the von Weizsäcker KED $\tau_W = \frac{|\nabla n|^2}{8n}$. For a uniform electron gas, the limit is $\tau_{\mathrm{uniform}} = \frac{3}{10}(3\pi^2)^{2/3} n^{5/3}$, where electrons are fully delocalized and orbitals overlap. From these quantities, the dimensionless variable $\alpha$ is defined as [56]:

$$\alpha = \frac{\tau - \tau_W}{\tau_{\mathrm{uniform}}}. \tag{2.45}$$

The dimensionless quantity $\alpha$ is thus able to capture different types of binding: for single bonds, $\tau$ equals the von Weizsäcker KED $\tau_W$, resulting in $\alpha = 0$. In the case of metallic bonds or slowly varying densities, where the electron density gradient $\nabla n$ is zero, $\tau \approx \tau_{\mathrm{uniform}}$, leading to $\alpha \approx 1$. For non-covalent bonds, the von Weizsäcker KED $\tau_W \approx 0$, making $\alpha \propto n/n^{5/3}$, thus $\alpha \gg 1$ [56]. Thus, SCAN is able to accurately treat different types of bonds simultaneously, something PBE is not able to do [12, 13].
Overall, SCAN has demonstrated improvements in accuracy, particularly for strongly bound compounds, when compared to other common functionals like PBE across several properties [10, 12]. Additionally, SCAN provides very accurate predictions of lattice constants [10, 11] and ground-state crystal structures of solids [10]. SCAN predicts formation energies for strongly bound compounds much more accurately than PBE, though it shows less accuracy for weakly bonded solids [10]. While SCAN still underestimates experimen-

tal band gaps like the other (semi)-local functionals, it offers slight improvements [10]. For magnetic moments, SCAN tends to overestimate moments in bulk ferromagnetic materials like Iron (Fe), Cobalt (Co), and Nickel (Ni) [57].
The SCAN functional also has limitations due to sharp oscillations in its functional derivatives, which can lead to numerical instabilities in meta-GGA calculations and result in a divergent xc potential [58]. These numerical sensitivities can require the use of very dense grids, which increases computational costs and can limit the practicality of SCAN functional calculations [59]. Due to this, Bartok and Yates [60] introduced regularized SCAN (rSCAN), which regularizes $\alpha$ to improve stability, though it compromises the accuracy of SCAN by violating scaling constraints, such as the correct uniform and non-uniform scaling properties of $\alpha$ [59, 61]. Furness et al. [59] developed the re-regularized SCAN (r2SCAN) to address the instability and restore accuracy by reducing the constraints violated by rSCAN. The performance of r2SCAN was studied thoroughly by Kingsbury et al. [62], where they found that r2SCAN predicts ground-state properties such as band gaps and formation energies with an accuracy comparable to that of SCAN, but also predicts slightly larger lattice constants. Overall, they showed that r2SCAN demonstrates a much more reliable convergence behavior, making it ideal for high-throughput calculations while overall retaining a similar accuracy to SCAN.
Another non-empirical meta-GGA is the Tao, Perdew, Staroverov, and Scuseria (TPSS) functional [63], with its exchange part defined as:

$$E_{\mathrm{x}}^{\mathrm{TPSS}} = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{\mathrm{x}}^{\mathrm{uniform}}(n(\mathbf{r}))\, F_{\mathrm{x}}^{\mathrm{TPSS}}(s(\mathbf{r})^2, z(\mathbf{r})). \tag{2.46}$$

Unlike SCAN, TPSS does not depend on the iso-orbital indicator $\alpha$, but instead on $z = \frac{\tau_W}{\tau}$. The issue with this iso-orbital indicator is that it can capture single bonds with $z = 1$, but it fails to differentiate between metallic and noncovalent bonds, where $z \approx 0$ [56]. Generally, TPSS does not offer a significant improvement over PBE. While it predicts slightly more accurate lattice constants, it is still less precise than the SCAN functional [7, 64]. Additionally, there is no notable improvement in the band gap compared to PBE [21]. However, TPSS generally needs fewer grid points to converge, which can be attributed to fewer peaks in the derivatives of $\varepsilon_{\mathrm{xc}}$ in comparison to the SCAN functional [65].

### 2.2.4. Hybrid Functionals

One of the reasons why the (semi-)local functionals described in the previous sections fail to accurately predict certain properties is that they do not include any nonlocal terms, preventing an accurate description of long-range xc effects. Furthermore, (semi-)local functionals suffer from the self-interaction error. This error arises because (semi-)local DFAs do not completely cancel out the unphysical interaction of an electron with itself introduced by the Hartree potential [66].
To address these limitations and advance up to the next rung of Jacob's ladder, hybrid functionals were introduced. These functionals mix a fraction of the nonlocal exact exchange from Hartree-Fock theory with the exchange term from a (semi-)local DFA. The inclusion of the Hartree-Fock exact exchange helps to reduce the self-interaction error by canceling part of the self-interaction and introduces nonlocality into the functional. A hybrid functional is expressed as [6]:

$$E_{\mathrm{xc}}^{\mathrm{hyb}} = \alpha E_{\mathrm{x}}^{\mathrm{NL}} + (1-\alpha) E_{\mathrm{x}}^{\mathrm{L}} + E_{\mathrm{c}}^{\mathrm{L}}, \tag{2.47}$$

where the correlation term $E_c^L$ from the (semi-)local functional is not mixed and the parameter $\alpha$ controls the mixing ratio between the nonlocal (NL) and local (L) exchange. A prominent example of a hybrid which is constructed like this is the PBE0 functional [67]. However, hybrids constructed in this way are computationally expensive because of the slowly decaying long-range (LR) part of the exchange interaction. To reduce the computational cost, screened hybrid functionals were developed, which replace the LR-exchange part with the respective part of the DFA [68]. One prominent example of this class is the Heyd Scuseria Ernzerhof (HSE) functional, generally defined as [69, 70]:

$$E_{xc}^{HSE} = E_c^{PBE} + \alpha E_x^{NL,\,SR}(\omega) + (1-\alpha)E_x^{PBE,\,SR}(\omega) + E_x^{PBE,\,LR}(\omega), \tag{2.48}$$

where $\omega$ is the screening parameter that controls the range separation between short-range (SR) and long-range (LR) interactions. There exist different versions of HSE, with different parameters. The well-known HSE06 functional has the parameters $\omega = 0.11\,\mathrm{Bohr}^{-1}$ and $\alpha = 0.25$ [71].
Hybrids have been shown to outperform most (semi-)local functionals for many properties, such as band gaps [18, 20, 21] and lattice constants [72]. However, despite their improved accuracy, hybrid functionals are significantly more expensive computationally compared to (semi-)local functionals. Reports show that hybrid functional calculations can be 10 to 100 times slower than (semi-)local calculations [21, 73].

### 2.2.5. The Band Gap Problem

Among the functionals introduced above, meta-GGA and hybrid functionals have been shown to improve upon the underestimation of the band gap in many materials. However, while the fictitious system introduced by Kohn and Sham provides a way to obtain the ground-state density of the real, interacting system, the question of how the obtained energies and thus the KS band gap relate to the real system remains. In the following sections, we will explore this relationship and discuss why functionals that are orbital-dependent, like the meta-GGAs and hybrid functionals, improve band gap predictions in comparison to the LDA and GGA functionals. The fundamental band gap $E_{gap}$ for a system with $N$ electrons is defined as [15]:

$$E_{gap}(N) = I(N) - A(N), \tag{2.49}$$

where $I(N)$ is the ionization potential and $A(N)$ is the electron affinity of the $N$-electron system. The KS band gap, in contrast, is given by the difference of the lowest unoccupied and highest occupied KS state:

$$E_{gap}^{KS} = \varepsilon_{LUMO}^{KS} - \varepsilon_{HOMO}^{KS}. \tag{2.50}$$

To analyze the relationship between the two expressions, Perdew et al. [74] considered a fractional electron number $N = N_0 + \delta$, where $0 < \delta < 1$ and stated that the total energy as a function of this fractional electron number consists of segments with straight lines between the values at the integer electron values. When crossing an integer electron number, the total energy exhibits a kink, leading to a discontinuous derivative. The derivative discontinuity with respect to the electron number then defines the fundamental

band gap:

$$E_{\text{gap}} = I(N) - A(N) = \left( \left.\frac{\partial E}{\partial N}\right|_{+} - \left.\frac{\partial E}{\partial N}\right|_{-} \right), \tag{2.51}$$

where + and − correspond to the right and left limits of the potential at integer electron numbers. The Hartree and external potential are continuous functionals of the density and thus the contribution to the band gap only stems from the non-interacting kinetic energy $T_s$ and the xc energy $E_{\text{xc}}$ [75] [76]:

$$E_{\text{gap}} = \left\{ \left.\frac{\delta T_s}{\delta n}\right|_{+} - \left.\frac{\delta T_s}{\delta n}\right|_{-} \right\} + \left\{ \left.\frac{\delta E_{\text{xc}}}{\delta n}\right|_{+} - \left.\frac{\delta E_{\text{xc}}}{\delta n}\right|_{-} \right\} = \Delta_{\text{KS}} + \Delta_{\text{xc}}. \tag{2.52}$$

The contribution from $T_s$ reduces to the band gap of the KS system, as defined in Eq. (2.50), leading to $E_{\text{gap}}^{\text{KS}} = \Delta_{\text{KS}}$ [15]. Consequently, one reason for underestimation of the band gaps is that the KS gap does not include the derivative discontinuity $\Delta_{\text{xc}}$. To compare the KS gap to the fundamental gap, $\Delta_{\text{xc}}$ must be evaluated and added to the KS gap $\Delta_{\text{KS}}$ [66]. However, for density-dependent functionals, like LDAs and GGAs, $\Delta_{\text{xc}}$ is zero, leading to a significant underestimation of the band gap.
In contrast, orbital-dependent functionals like meta-GGAs and hybrids can produce a non-zero value for the derivative discontinuity $\Delta_{\text{xc}}$ [15]. Within the gKS scheme, this discontinuity is even directly included in the gap, yielding $E_{\text{gap}}^{\text{gKS}} = \varepsilon_{\text{LUMO}}^{\text{gKS}} - \varepsilon_{\text{HOMO}}^{\text{gKS}} = E_{\text{gap}}$. This occurs because the orbital-dependent component of KED-dependent meta-GGA functionals leads to a non-multiplicative potential, contributing to the derivative discontinuity as $\Delta_{\text{xc}}^{\text{non-mult}}$. The band gap in this case can be expressed as [77]:

$$E_{\text{gap}}^{\text{gKS}} = E_{\text{gap}}^{\text{KS}} + \Delta_{\text{xc}}^{\text{mult}} + \Delta_{\text{xc}}^{\text{non-mult}}. \tag{2.53}$$

Whilst the multiplicative, density-dependent xc discontinuity $\Delta_{\text{xc}}^{\text{mult}}$ is zero, the xc derivative discontinuity obtained by the $\tau$-dependent part is given by [15]:

$$\Delta_{\text{xc}}^{\text{non-mult}} = \int d\mathbf{r}' \, \frac{\partial \varepsilon_{\text{xc}}(\mathbf{r}')}{\partial \tau} \left[ \left.\frac{\delta \tau(\mathbf{r}')}{\delta n(\mathbf{r})}\right|_{+} - \left.\frac{\delta \tau(\mathbf{r}')}{\delta n(\mathbf{r})}\right|_{-} \right]. \tag{2.54}$$

However, $\Delta_{\text{xc}}^{\text{non-mult}}$ is often relatively small for these functionals, resulting in minimal improvement in the band gap [15]. For instance, the TPSS functional does not significantly enhance the band gap compared to PBE [21], which is expected due to its small xc-derivative discontinuity [78].
To understand how a larger portion of the derivative discontinuity can be included, T. Aschebrock et al. [19] proposed that the first factor of Eq. (2.54) can be averaged over an energetically relevant region, such that:

$$\Delta_{\text{xc}} \approx \overline{\frac{\partial \varepsilon_{\text{xc}}}{\partial \tau}} \left[ \left.\frac{\delta T_s}{\delta n}\right|_{+} - \left.\frac{\delta T_s}{\delta n}\right|_{-} \right] = \overline{\frac{\partial \varepsilon_{\text{xc}}}{\partial \tau}} \Delta_{\text{KS}}. \tag{2.55}$$

To obtain a positive xc derivative discontinuity, the following condition,

$$\frac{\partial \varepsilon_{\text{x}}}{\partial \tau} > 0, \tag{2.56}$$

has to be enforced, firstly only considering exchange as the contribution as the exchange

is much larger [15]. It can be shown that this condition is equal to [19]:

$$\frac{\partial F_{\mathrm{x}}}{\partial \alpha} < 0, \tag{2.57}$$

where $F_{\mathrm{x}}$ is the exchange enhancement factor. The validity of averaging the derivative is discussed further in Ref. [15]. However, in general, this provides a way to estimate the contribution of the xc derivative discontinuity to the band gap by simply considering the enhancement factor. It has been shown in Ref. [79], that the slope of the enhancement factor $F_{\mathrm{x}}$ for SCAN is larger than TPSS, leading to a larger xc derivative discontinuity and consequently also an expected better performance in the prediction of band gaps.
The constructed functional Thilo Aschebrock and Stephan Kümmel (TASK) [19] even exhibits a stronger slope. This is because it was constructed non-empirically by reconstructing the SCAN exchange functional to specifically enhance the xc derivative discontinuity and thus enlarge the total band gap. A more thorough analysis of this can also be found in Ref. [15]. The TASK exchange is combined with the LDA correlation functional PW92 [80], as for now only the exchange functional has been addressed by the authors in Ref. [19]. As a result of developing the TASK functional specifically to increase the xc derivative discontinuity, it has been identified as one of the most accurate meta-GGA functionals for predicting the band gap of solids [21]. In a study of 2D materials [20], TASK and particularly its modified version mTASK [81], modified for low-dimensional materials, have shown a performance comparable to the hybrid functional HSE06 [69, 70]. Given that meta-GGA functionals are generally less computationally expensive than hybrid calculations, TASK offers a promising balance between accuracy and computational cost. However, TASK significantly overestimates lattice constants [7], indicating that it cannot serve as a general-purpose functional like SCAN.
Similarly to meta-GGA functionals, hybrid functionals also exhibit a non-zero xc derivative discontinuity when implemented in the gKS scheme [15, 35]. This discontinuity also arises through the explicit dependence on the orbitals through the nonlocal exact exchange.

## 2.3. (L)APW Method

The (L)APW method solves the KS equations using a basis set expansion:

$$\psi_\mu(\mathbf{r}) = \sum_j C_j^\mu \phi_j(\mathbf{r}), \tag{2.58}$$

where $\phi_j(\mathbf{r})$ are the basis functions and $C_j^\mu$ are the expansion coefficients for band index and wave vector $\mu = \{n, \mathbf{k}\}$. This transforms the KS equations (2.10) into an eigenvalue problem:

$$\sum_j \hat{H}_{ij} C_j^\mu = \varepsilon_\mu \sum_j S_{ij} C_j^\mu, \tag{2.59}$$

where $\hat{H}_{ij}$ and $S_{ij}$ are the Hamiltonian and overlap matrices, respectively:

$$\hat{H}_{ij} = \langle \phi_i | \hat{H} | \phi_j \rangle, \quad S_{ij} = \langle \phi_i | \phi_j \rangle. \tag{2.60}$$

In the (L)APW method, the unit cell is partitioned into two regions: the Interstitial Region (IR) region and the Muffin-Tin (MT) regions. In the MT regions, non-overlapping

spheres of radius $R_\alpha^{\mathrm{MT}}$ are centered around each atom $\alpha$ at position $\mathbf{R}_\alpha$. Within these spheres, atomic-like functions describe the rapid oscillations of the wave function near the nuclei. In the interstitial region, where electrons behave more like free electrons, the wave function varies more slowly and can be accurately described by plane waves. Unless specified otherwise, the theoretical framework discussed in the following sections is primarily based on the concepts and methodologies outlined in [23].
The most general form of the (L)APW basis function is defined as follows:

$$\phi_{\mathbf{G}+\mathbf{k}}^{\mathrm{(L)APW}}(\mathbf{r}) = \begin{cases} \sum_{\ell m\xi} A_{\ell m\xi,\alpha}^{\mathbf{G}+\mathbf{k}}\, u_{\ell\xi}^{\alpha}(r_\alpha, \varepsilon_\ell)\, Y_{\ell m}(\hat{\mathbf{r}}_\alpha), & \mathbf{r} \in \mathrm{MT}_\alpha, \\ \frac{1}{\sqrt{\Omega}}\, e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}}, & \mathbf{r} \in \mathrm{I}. \end{cases} \tag{2.61}$$

The interstitial basis functions are represented by plane waves with a cut-off of $G_{\max}$. Although $G_{\max}$ defines the size of the interstitial basis, reducing the muffin-tin radius $R_\alpha^{\mathrm{MT}}$ for an atom $\alpha$ results in an interstitial region where the wave functions vary more rapidly. This, in turn, requires a different cut-off to achieve the same level of convergence. Therefore, a more practical convergence parameter, $R_{\mathrm{MT}}G_{\max}$, which is the product of the smallest muffin-tin radius in the material and the plane wave cut-off, is commonly used to define the cut-off for the interstitial basis. This parameter must be converged to ensure that the interstitial region is well-represented by the basis functions.
Within the muffin-tin spheres, the basis functions are expanded in terms of spherical harmonics $Y_{\ell m}(\hat{\mathbf{r}}_\alpha)$, where $\mathbf{r}_\alpha = \mathbf{r} - \mathbf{R}_\alpha$, and $\ell$ and $m$ are the angular and magnetic quantum numbers, respectively. Here, the basis size is controlled by the cut-off parameter $\ell_{\max}$. The parameter $\xi$ defines the linearization order, which determines the order of the energy derivative of the radial function $u_{\ell\xi}^{\alpha}$.
In the pure Augmented Plane Wave (APW) method, originally proposed by Slater [82], only the radial function $u_\ell^\alpha$ contributes to the basis function. Consequently, the basis function in the muffin-tin region is given by

$$\phi_{\mathbf{G}+\mathbf{k}}^{\mathrm{APW}}(\mathbf{r}) = \sum_{\ell m} A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}\, u_\ell^\alpha(r_\alpha, \varepsilon_\ell)\, Y_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{2.62}$$

The radial function $u_\ell^\alpha$ is the solution to the radial Schrödinger equation for atom $\alpha$ at energy $\varepsilon_\ell$:

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{2r^2} + V_{\mathrm{sph}}(r) - \varepsilon_\ell\right] r u_\ell^\alpha(r) = 0. \tag{2.63}$$

Since the spheres are centered around the nuclei, the potential $V_{\mathrm{sph}}(r)$ within these regions is assumed to be spherically symmetric.
The coefficients $A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ are determined to ensure the continuity of the APW wave function $\phi_{\mathbf{G}+\mathbf{k}}(\mathbf{r})$ at the MT boundary $R_\alpha^{\mathrm{MT}}$. This is achieved by expanding the plane waves from the interstitial region in spherical harmonics using the Rayleigh expansion. By requiring that the plane wave expansion equals the $(\ell m)$-contribution of the spherical harmonics at the sphere boundary, the $A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ are given by:

$$A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}} = \frac{4\pi i^\ell}{\sqrt{\Omega}\, u_\ell^\alpha(R_\alpha^{\mathrm{MT}}, \varepsilon_\ell)}\, j_\ell(|\mathbf{G}+\mathbf{k}|R_\alpha^{\mathrm{MT}})\, Y_{\ell m}^*(\widehat{\mathbf{G}+\mathbf{k}}), \tag{2.64}$$

where $j_\ell$ are the spherical Bessel functions. An issue with the APW method arises when the denominator in Eq. (2.64) approaches zero, which is known as the asymptote problem, leading to numerical difficulties. Furthermore, $\psi_{n\mathbf{k}}$ is only accurate if the energies $\varepsilon_\ell$

match the actual band energies, making the eigenvalue problem non-linear in energy. One approach to address the non-linearity in the APW method is to use fixed trial energies $E_\ell$ to construct the basis. The secular equation is then solved for the band energies $\varepsilon_\ell$. If the trial energies do not match the resulting $\varepsilon_\ell$, the process is repeated with updated trial energies until convergence, making the APW method inefficient [66].
To overcome the nonlinearity issue of the APW method, Andersen developed the LAPW method in 1975 [83]. In this method, the basis in the MT is made up of a linear combination of the radial function $u_\ell^\alpha(r_\alpha, E_\ell)$ and its energy derivative $\dot{u}_\ell^\alpha(r_\alpha, E_\ell)$ at a fixed energy value $E_\ell$, called the linearization energy. Thus, the basis function in the muffin-tin is given by

$$\phi_{\mathbf{G}+\mathbf{k}}^{\mathrm{LAPW}}(\mathbf{r}) = \sum_{\ell m} \left\{ A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}\, u_\ell^\alpha(r_\alpha, E_\ell) + B_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}\, \dot{u}_\ell^\alpha(r_\alpha, E_\ell) \right\} Y_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{2.65}$$

Here, $A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ and $B_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ are the matching coefficients to ensure not only the continuity of $\phi_{\mathbf{G}+\mathbf{k}}$ but also of its derivative at the MT boundary. This requires higher plane wave cutoffs to achieve a certain level of convergence. It can however be shown, that the asymptote problem does not occur in the LAPW method, as the denominator in $A_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ and $B_{\ell m\alpha}^{\mathbf{G}+\mathbf{k}}$ does not vanish. The LAPW method provides more variational freedom compared to the APW method. This is because, for a fixed trial energy $E_\ell$ differing from the band energy $\varepsilon_\ell$, a linear combination expanded around the band energy $\varepsilon_\ell$ can reproduce the radial function at the band energy up to an error of order $\mathcal{O}((\varepsilon_\ell - E_\ell)^2)$:

$$u_\ell^\alpha(r_\alpha, \varepsilon_\ell) = u_\ell^\alpha(r_\alpha, E_\ell) + (\varepsilon_\ell - E_\ell)\, \dot{u}_\ell^\alpha(r_\alpha, E_\ell) + \mathcal{O}\left((\varepsilon_\ell - E_\ell)^2\right), \tag{2.66}$$

with $\dot{u}_\ell^\alpha(r_\alpha, E_\ell) = \frac{\partial u_\ell^\alpha(r_\alpha, E_\ell)}{\partial \varepsilon_\ell}$. While this modified basis introduces more flexibility, it is still crucial that the linearization energies $E_\ell$ are chosen close to the band energies to minimize the linearization error of order $\mathcal{O}((\varepsilon_\ell - E_\ell)^2)$ introduced by the LAPW method.

### 2.3.1. Local Orbitals

The LAPW basis is primarily used to describe valence states, which extend beyond the boundary of the muffin-tin sphere. In contrast, core states, which are low in energy and tightly bound to the nucleus, are fully contained within the muffin-tin sphere. However, there exist intermediate states that are called semi-core states. These are typically high-lying in energy, at least compared to the deep-lying core states, and like valence states, they are not completely confined to the muffin-tin sphere and can leak out into the interstitial region. Challenges arise when the fixed energy parameter $E_\ell$ that is used to describe the high-lying valence state is then used to describe semi-core states with the same angular momentum quantum number $\ell$, but with a different principal quantum number $n$. This is because $E_\ell$ may not sufficiently represent the energy characteristics of the semi-core state. To solve this, LOs were introduced, which are defined as follows [22]:

$$\phi_L^{\mathrm{LO}}(\mathbf{r}) = \begin{cases} \delta_{\alpha\alpha_L}\delta_{\ell\ell_L}\delta_{mm_L} \left( a_L^{\mathrm{LO}} u_\ell^\alpha(r_\alpha, E_\ell) + b_L^{\mathrm{LO}} \dot{u}_\ell^\alpha(r_\alpha, E_\ell) \right) Y_{\ell_L m_L}(\hat{\mathbf{r}}_\alpha), & \mathbf{r} \in \mathrm{MT}_\alpha, \\ 0, & \mathbf{r} \in \mathrm{IR}. \end{cases} \tag{2.67}$$

A local orbital is entirely localized within its muffin-tin sphere, with its value going to zero at the boundary. Consequently, the basis does not require matching to the plane waves in the interstitial region, making the coefficients $a_L^{\mathrm{LO}}$ and $b_L^{\mathrm{LO}}$ independent of $\mathbf{G}$

and **k**. These coefficients are determined only by the requirements that the local orbital is normalized and that it vanishes in value at the muffin-tin boundary.
Another approach to linearize the original APW method was proposed by Sjöstedt et al. in 2000 [84]. Known as the Augmented Plane Wave + Local Orbital (APW+LO) method, it enhances the flexibility of the APW basis by incorporating an additional local orbital. This local orbital is evaluated at the same linearization energy $E_l$ as the main APW basis function, limiting the energy to one $E_l$ per quantum number $n$. The coefficients $a_L^{\mathrm{LO}}$ and $b_L^{\mathrm{LO}}$ are determined through normalization of the local orbital and by ensuring it is zero at the muffin-tin boundary. By relaxing the requirement that the first derivative of the APW must match at the boundary, it has been shown that a lower cut-off for the plane wave basis is needed [84, 85].
However, both the LAPW and APW+LO methods are subject to linearization errors due to the fixed linearization energy. Consequently, the accuracy of the calculation can depend on the chosen energies and the muffin-tin radius, as it defines the region where the (L)APW basis describes the wave function. Ideally, the results should be as independent as possible of both parameters. One approach to achieve this is to choose a small muffin-tin radius, minimizing the region described by the (L)APW. The downside is that this requires a higher plane wave basis set cut-off, as the plane waves must describe a more oscillatory behavior, making the calculation more computationally expensive [85].
An alternative to minimize the linearization error and make calculations less dependent on the muffin-tin radius involves the addition of High Derivative Local Orbitals (HDLO) to both valence and semi-core states. HDLOs are constructed from higher derivatives of local orbitals, such as second energy derivatives $\ddot{u}_\ell^\alpha(r_\alpha, E_\ell)$. Incorporating HDLOs allows for a larger muffin-tin radius, as the wave function within the muffin-tin is better described, thereby enabling a smaller plane wave basis set cut-off [85].

### 2.3.2. Core States

Core electrons, which are fully confined within the MT sphere, are treated fully relativistically by solving the Dirac equation for a spherically symmetric potential. It is valid to treat the core states separately from the valence states, as the radial functions $u_\ell$ and $\dot{u}_\ell$ of the (L)APW basis are orthogonal to any other radial function $u_\ell$ obtained at a different energy $E_\ell$ that drops to zero at the MT boundary, as is the case for core electrons. Consequently, the (L)APW basis cannot accurately describe the core states, justifying their separate treatment [23].
The Dirac Hamiltonian commutes with $\hat{J}^2$ and $\hat{J}_z$, where $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ represents the total angular momentum. The eigenstates of this system can thus be constructed from the spherical harmonics $Y_{\ell m}(\theta, \phi)$, which are the eigenstates of $\hat{L}^2$ and $\hat{L}_z$, and the two-component spinor $\chi_\sigma$, which is the eigenstate of $\hat{S}^2$ and $\hat{S}_z$ [86]. The eigenstates are represented by the spin spherical harmonics, $\Omega_{JSLM}(\theta, \phi)$, with $S = \frac{1}{2}$, defined as follows:

$$\Omega_{LJM}(\theta, \phi) = \sum_{m_s = \pm\frac{1}{2}} C^{JM}_{L,M-m_s,\frac{1}{2},m_s} Y_{L,M-m_s}(\theta, \phi)\chi_{m_s}, \tag{2.68}$$

where $C^{JM_J}_{LMSM_S}$ are the Clebsch-Gordan coefficients and $\chi_\sigma$ for $m_s = \pm\frac{1}{2}$ is given by:

$$\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.69}$$

The quantum numbers $\{L, J, M\}$ are not independent because $J = L \pm \frac{1}{2}$. The spin spherical harmonics $\Omega_{LJM}(\hat{\mathbf{r}})$ can be expressed as:

$$\Omega_{L,\pm\frac{1}{2},M}(\hat{\mathbf{r}}) = \begin{pmatrix} \sqrt{\frac{L\pm M+\frac{1}{2}}{2L+1}} Y_{L,M\mp\frac{1}{2}}(\hat{\mathbf{r}}) \\ \sqrt{\frac{L\mp M+\frac{1}{2}}{2L+1}} Y_{L,M\pm\frac{1}{2}}(\hat{\mathbf{r}}) \end{pmatrix}. \tag{2.70}$$

One can further introduce the quantum number $\kappa$ defined as

$$\kappa = \begin{cases} -L-1 & \text{for } J = L+\frac{1}{2} \\ L & \text{for } J = L-\frac{1}{2} \end{cases}. \tag{2.71}$$

Using this definition, the solution to the Dirac equation can be expressed as [87]:

$$\psi^{\alpha}_{\kappa M}(\mathbf{r}) = \begin{pmatrix} g^{\alpha}_{\kappa}(r)\Omega_{\kappa M}(\hat{\mathbf{r}}) \\ -if^{\alpha}_{\kappa}(r)\Omega_{-\kappa M}(\hat{\mathbf{r}}) \end{pmatrix}. \tag{2.72}$$

The core wave functions $\psi^{\alpha}_{\kappa M}$ are therefore four-dimensional Dirac spinors with so-called large and small components. The radial functions $g^{\alpha}_{\kappa}(r)$ for the large component and $f^{\alpha}_{\kappa}(r)$ for the small component are obtained by solving the coupled radial Dirac equations [87]:

$$\frac{\partial g^{\alpha}_{\kappa}}{\partial r} = \frac{1}{c}\left(V^{\alpha}_{\text{eff}} - \epsilon_{\alpha,\kappa}\right) f^{\alpha}_{\kappa} + \left(\kappa - \frac{1}{r}\right) g^{\alpha}_{\kappa}, \tag{2.73}$$

$$\frac{\partial f^{\alpha}_{\kappa}}{\partial r} = -\kappa\frac{1}{r} f^{\alpha}_{\kappa} + 2c\left[1 + \frac{1}{2c^2}\left(\epsilon_{\alpha,\kappa} - V^{\alpha}_{\text{eff}}\right)\right] g^{\alpha}_{\kappa}. \tag{2.74}$$

# 3. Implementation of Meta-GGA Functionals

In this chapter, the necessary formulas for implementing meta-GGA functionals in the `exciting` code are derived. The added key component in meta-GGA calculations is the inclusion of the KED, which must be handled separately for valence electrons in both the interstitial and muffin-tin regions, and for the core electrons within the muffin-tin region. Following the implementation of the KED, the focus shifts to the derivation of the meta-GGA potential and the KS equations for spin-unpolarized systems. Next, we extend the derivations to the spin-polarized case, where different treatments are required due to the complexity introduced by the spin. Finally, this chapter concludes with the derivation of the total meta-GGA energy expression, covering all necessary steps to perform a SCF cycle using meta-GGA functionals in `exciting`.

## 3.1. Kinetic Energy Density

The KED $\tau(\mathbf{r})$ is defined as a function that integrates to the total non-interacting kinetic energy [88]:

$$T_s = \int d\mathbf{r}\, \tau(\mathbf{r}). \tag{3.1}$$

This definition, however, allows for different expressions and local values of the KED. Within the KS framework, the positive-definite KED is used, defined as:

$$\tau^{\mathrm{PD}}(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla\psi_i(\mathbf{r})|^2. \tag{3.2}$$

Another valid definition is the following KED,

$$\tau^{\mathrm{L}}(\mathbf{r}) = -\frac{1}{2}\sum_i \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r}), \tag{3.3}$$

which follows from the kinetic energy operator used in the KS equations. These two definitions are related through

$$\tau^{\mathrm{L}}(\mathbf{r}) = \tau^{\mathrm{PD}}(\mathbf{r}) - \frac{1}{4}\nabla^2 n(\mathbf{r}). \tag{3.4}$$

Both definitions of the KED integrate to the same non-interacting kinetic energy, as $\nabla^2 n(\mathbf{r})$ integrates to zero [89]. The positive-definite definition is generally preferred, as it is positive everywhere and can be considered more physical. For example, in the case of a single-orbital system, it leads to the Weizsäcker term [89, 90], $\tau_W$, which was already defined in Sec. 2.2.3 and is a component of the iso-orbital indicator implemented in many meta-GGA functionals. Furthermore, the alternative definition, which requires the use of the Laplacian of the orbitals, may be numerically less stable [90]. As a result, many implementations prefer to use the positive-definite definition [7, 91, 92]. In the case of relativistic wave functions, such as those for core electrons solved using the Dirac equation,

the relativistic definition of the KED does not satisfy the non-negativity condition $\alpha \geq 0$ required by the iso-orbital indicator. This failure can lead to inaccuracies in xc functionals that depend on the iso-orbital indicator [7, 93]. Therefore, the positive-definite KED defined in Eq. (3.2) is used for the core electrons to ensure that $\alpha \geq 0$ is satisfied.
The total KED is then the sum of the contributions from the valence and core electrons [7]:

$$\tau(\mathbf{r}) = \tau^{\text{val}}(\mathbf{r}) + \tau^{\text{core}}(\mathbf{r}). \tag{3.5}$$

For valence electrons, the positive-definite KED is given by

$$\tau^{\text{val}}(\mathbf{r}) = \frac{1}{2}\sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \nabla\psi^*_{n\mathbf{k}}(\mathbf{r}) \cdot \nabla\psi_{n\mathbf{k}}(\mathbf{r}), \tag{3.6}$$

where $w_{\mathbf{k}}$ is the $\mathbf{k}$-point weight and $f_{n\mathbf{k}}$ is the occupation number. The KED for valence electrons is expressed differently in the interstitial and muffin-tin regions:

$$\tau^{\text{val}}(\mathbf{r}) = \begin{cases} \sum_{\alpha\ell m} \tau^{\alpha}_{\ell m}(r_\alpha) R_{\ell m}(\hat{\mathbf{r}}_\alpha), & \mathbf{r} \in \text{MT}, \\ \sum_{\mathbf{G}} \tau_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}, & \mathbf{r} \in \text{I}. \end{cases} \tag{3.7}$$

The scalar KED for core electrons is evaluated by taking the gradient of the orbitals obtained through the Dirac equation as:

$$\tau^{\text{core}}(\mathbf{r}) = \frac{1}{2}\sum_{\alpha\kappa M} \nabla\psi^{\alpha *}_{\kappa M}(\mathbf{r}) \cdot \nabla\psi^{\alpha}_{\kappa M}(\mathbf{r}). \tag{3.8}$$

In the following sections, explicit expressions for the KED as implemented in the `exciting` code for both core and valence electrons will be derived, and strategies to reduce computational expense where possible will be discussed.

### 3.1.1. Valence States in the Interstitial Region

In the interstitial region, the wave function is expanded in terms of plane waves

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} C^{n\mathbf{k}}_{\mathbf{G}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}, \tag{3.9}$$

which yields the following gradient:

$$\nabla\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} i(\mathbf{G}+\mathbf{k}) C^{n\mathbf{k}}_{\mathbf{G}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}. \tag{3.10}$$

Inserting Eq. (3.10) into the positive-definite KED definition from Eq. (3.2) yields:

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{n\mathbf{k}} \frac{w_{\mathbf{k}} f_{n\mathbf{k}}}{\Omega} \sum_{\mathbf{G}'}\sum_{\mathbf{G}''} C^{n\mathbf{k}*}_{\mathbf{G}'} C^{n\mathbf{k}}_{\mathbf{G}''} \left[(\mathbf{G}'+\mathbf{k})\cdot(\mathbf{G}''+\mathbf{k})\right] e^{i(\mathbf{G}''-\mathbf{G}')\cdot\mathbf{r}}. \tag{3.11}$$

Defining $\mathbf{G} = \mathbf{G}'' - \mathbf{G}'$, the plane wave coefficients for the KED are given by:

$$\tau_{\mathbf{G}} = \frac{1}{2}\sum_{n\mathbf{k}} \frac{w_{\mathbf{k}} f_{n\mathbf{k}}}{\Omega} \sum_{\mathbf{G}'} C^{n\mathbf{k}*}_{\mathbf{G}'} C^{n\mathbf{k}}_{\mathbf{G}+\mathbf{G}'} \left[(\mathbf{G}'+\mathbf{k})\cdot(\mathbf{G}+\mathbf{G}'+\mathbf{k})\right]. \tag{3.12}$$

This shows that the computation of the KED in real space results from a convolution in **G**-space. The computational cost of this convolution scales as $\mathcal{O}(n^2)$. Given that the cutoff for the **G**-vector is $G_{\max}$, the computational effort scales with $\mathcal{O}\left((G_{\max})^6\right)$. To improve efficiency, the KED is calculated using the Fast Fourier Transform (FFT). The gradient of the wave function from Eq. (3.10) for each state is first transformed to real space using FFT, squared on the real space mesh, and then summed over all states. Finally, the result is converted back to **G**-space using FFT. This approach significantly reduces computational demands, scaling as $\mathcal{O}(G_{\max}^3 \ln(G_{\max}^3))$, and offers a more efficient implementation of the KED than the direct evaluation of the KED in **G**-space [23, 94].

### 3.1.2. Valence States in the Muffin-Tin Region

As outlined in Sec. 2.3, the wave function within the muffin-tin region for atom $\alpha$ is given by the (L)APW + LO basis:

$$\psi_{n\mathbf{k}}^{\alpha}(\mathbf{r}) = \sum_{\mathbf{G}+\mathbf{k}} C_{\mathbf{G}+\mathbf{k}}^{n\mathbf{k}\alpha} \phi_{\mathbf{G}+\mathbf{k}}^{\alpha}(\mathbf{r}) + \sum_{L} C_{L}^{n\mathbf{k}\alpha}\, \phi_{L}^{\alpha}(\mathbf{r}), \tag{3.13}$$

where the $\phi_{\mathbf{G}+\mathbf{k}}^{\alpha}(\mathbf{r})$ represents the (L)APW basis and $\phi_{L}^{\alpha}(\mathbf{r})$ the local orbitals. To derive the coefficients $\tau_{\ell m}^{\alpha}(r_\alpha)$ of Eq. (3.7), it is convenient to rewrite the wave function in a more compact form as

$$\psi_{n\mathbf{k}}^{\alpha}(\mathbf{r}) = \sum_{\lambda} C_{\lambda}^{n\mathbf{k}\alpha}\, \phi_{\lambda}^{\alpha}(\mathbf{r}) = \sum_{\lambda} C_{\lambda}^{n\mathbf{k}\alpha}\, g_{\lambda}^{\alpha}(r_\alpha)\; Y_{\ell_\lambda m_\lambda}(\hat{\mathbf{r}}_\alpha), \tag{3.14}$$

where $\lambda$ is a combined index consisting of $\{\ell, m, \xi\}$ for the (L)APWs and $\{L\}$ for the local orbitals. The eigenvectors in this notation are then defined as follows

$$C_{\lambda}^{n\mathbf{k}\alpha} = \begin{cases} \sum_{\mathbf{G}+\mathbf{k}} C_{\mathbf{G}+\mathbf{k}}^{n\mathbf{k}}\, A_{\ell m\,\xi,\alpha}^{\mathbf{G}+\mathbf{k}}, & \lambda \rightarrow \text{ (L)APW}, \\ C_{L}^{n\mathbf{k}}\, \delta_{\alpha\alpha_L}, & \lambda \rightarrow \text{ LO}, \end{cases} \tag{3.15}$$

and the radial functions are given by

$$g_{\lambda}^{\alpha}(r_\alpha) = \begin{cases} u_{\ell\xi}^{\alpha}(r_\alpha), & \lambda \rightarrow \text{ (L)APW}, \\ f_{L}^{\alpha}(r_\alpha), & \lambda \rightarrow \text{ LO}, \end{cases} \tag{3.16}$$

where $f_{L}^{\alpha}(r_\alpha)$ is the sum of the radial functions with their coefficients of the local orbital, as defined in Eq. (2.67). The gradient of a product of a radial function and a spherical harmonic can be expressed analytically in terms of vector spherical harmonics $\mathbf{Y}_{JM_J}^{L}$ [95]. By applying the analytical formula onto the basis $\phi_{\lambda}^{\alpha}(\mathbf{r})$, the following expression can be derived

$$\begin{aligned} \nabla\phi_{\lambda}^{\alpha}(\mathbf{r}) &= \nabla\left(g_{\lambda}^{\alpha}(r_\alpha)\, Y_{\ell_\lambda m_\lambda}(\hat{\mathbf{r}}_\alpha)\right) \\ &= g_{\lambda}^{\alpha-}(r_\alpha)\, \mathbf{Y}_{\ell_\lambda m_\lambda}^{\ell_\lambda - 1}(\hat{\mathbf{r}}_\alpha) + g_{\lambda}^{\alpha+}(r_\alpha)\, \mathbf{Y}_{\ell_\lambda m_\lambda}^{\ell_\lambda + 1}(\hat{\mathbf{r}}_\alpha), \end{aligned} \tag{3.17}$$

with

$$g_{\lambda}^{\alpha-}(r_\alpha) = \left[\frac{\ell_\lambda}{2\ell_\lambda + 1}\right]^{\frac{1}{2}} \left(\frac{d}{dr} - \frac{\ell_\lambda + 1}{r}\right) g_{\lambda}^{\alpha}(r_\alpha), \tag{3.18}$$

$$g_\lambda^{\alpha+}(r_\alpha) = -\left[\frac{\ell_\lambda + 1}{2\ell_\lambda + 1}\right]^{\frac{1}{2}} \left(\frac{d}{dr} + \frac{\ell_\lambda}{r}\right) g_\lambda^\alpha(r_\alpha). \tag{3.19}$$

The vector spherical harmonics are expanded in terms of scalar spherical harmonics [95],

$$\mathbf{Y}_{JM_J}^L = \sum_\gamma \sum_{M=-L}^{L} C_{LM1\gamma}^{JM_J} Y_{LM} \hat{\mathbf{e}}_\gamma, \tag{3.20}$$

where $C_{LM1\gamma}^{JM_J}$ are the Clebsch-Gordan coefficients, $Y_{LM}$ are the spherical harmonics and $\hat{\mathbf{e}}_\gamma$ is the spherical basis vector in direction $\gamma$, defined as

$$\hat{\mathbf{e}}_1 = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x + i\hat{\mathbf{e}}_y), \qquad \hat{\mathbf{e}}_0 = \hat{\mathbf{e}}_z, \qquad \hat{\mathbf{e}}_{-1} = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x - i\hat{\mathbf{e}}_y). \tag{3.21}$$

Then the gradient for the total wave function in terms of spherical harmonics in the direction $\gamma$ is given by

$$\begin{aligned}\partial_\gamma \Psi_{n\mathbf{k}}^\alpha(\mathbf{r}) = \sum_\lambda C_\lambda^{n\mathbf{k}\alpha} \Bigg[ & g_\lambda^{\alpha+}(r_\alpha) \sum_{m=-(\ell_\lambda+1)}^{\ell_\lambda+1} C_{\ell_\lambda+1,m,1,\gamma}^{\ell_\lambda,m_\lambda} Y_{\ell_\lambda+1,m}(\hat{\mathbf{r}}_\alpha) \\ & + g_\lambda^{\alpha-}(r_\alpha) \sum_{m=-(\ell_\lambda-1)}^{\ell_\lambda-1} C_{\ell_\lambda-1,m,1,\gamma}^{\ell_\lambda,m_\lambda} Y_{\ell_\lambda-1,m}(\hat{\mathbf{r}}_\alpha) \Bigg].\end{aligned} \tag{3.22}$$

With the gradient of the wave function now fully defined, its product can be computed, resulting in the coefficients of the KED $\tau_{\ell m}^\alpha$:

$$\tau_{\ell m}^\alpha = \frac{1}{2} \left\{ \sum_{\lambda_1\lambda_2} D_{\lambda_1\lambda_2}^\alpha \sum_\pm \sum_{\pm'} P_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}}^{\pm\pm'} g_{\lambda_1}^{\alpha\pm} g_{\lambda_2}^{\alpha\pm'} \right\}, \tag{3.23}$$

where

$$P_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}}^{\pm\pm'} = \sum_{\substack{m'= \\ -(\ell_{\lambda_1}\pm 1)}}^{\ell_{\lambda_1}\pm 1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}\pm' 1)}}^{\ell_{\lambda_2}\pm' 1} \sum_\gamma \left( C_{\ell_{\lambda_1}\pm 1,m',1,\gamma}^{\ell_{\lambda_1},m_{\lambda_1}} \right)^* C_{\ell_{\lambda_2}\pm' 1,m'',1,\gamma}^{\ell_{\lambda_2},m_{\lambda_2}} G_{\ell_{\lambda_1}\pm 1,m',\ell_{\lambda_2}\pm' 1,m''}^{\ell m}. \tag{3.24}$$

Here, $G_{\ell_{\lambda_1}\pm 1,m',\ell_{\lambda_2}\pm' 1,m''}^{\ell m}$ are the Gaunt coefficients, defined as the angular integral of a product of three spherical harmonics

$$G_{\ell_1 m_1 \ell_2 m_2}^{\ell m} = \int Y_{\ell_1 m_1}^* R_{\ell m} Y_{\ell_2 m_2} \, d\Omega. \tag{3.25}$$

For many $(\ell m)$-pairs, the Gaunt integral vanishes for specific combinations of $(\ell_1 m_1)$ and $(\ell_2 m_2)$ [95], so it only needs to be computed for the non-vanishing combinations [22]. The matrix $D_{\lambda_1\lambda_2}^\alpha$ is the product of the eigenvectors $C_\lambda^{n\mathbf{k}\alpha}$:

$$D_{\lambda_1\lambda_2}^\alpha = \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \left(C_{\lambda_1}^{n\mathbf{k}\alpha}\right)^* C_{\lambda_2}^{n\mathbf{k}\alpha}. \tag{3.26}$$

A more detailed derivation of the coefficients $\tau_{\ell m}^\alpha$ can be found in the Appendix A.3.

The total KED in real space for valence electrons is then given by

$$\tau^{\text{val}}(\mathbf{r}) = \sum_{\alpha\ell m} \tau^{\alpha}_{\ell m}(r_\alpha) R_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{3.27}$$

The KED is recalculated at every SCF step. However, for an efficient implementation, the different components are handled separately. While the eigenvectors $C^{n\mathbf{k}\alpha}_{\lambda}$ and the radial functions $g^{\alpha\pm}_{\lambda}$ are updated at every iteration, only the eigenvectors are dependent on the state $(n, \mathbf{k})$. Consequently, the summation over states is necessary only for these eigenvectors, allowing results to be efficiently stored in the matrix $D^{\alpha}_{\lambda_1\lambda_2}$. Furthermore, since the Gaunt coefficients and Clebsch-Gordan coefficients remain invariant across states and SCF iterations, they are computed once and reused, to reduce the computational effort at each step.

### 3.1.3. Core States

The scalar KED for core electrons is given by:

$$\tau^{\text{core}}(\mathbf{r}) = \frac{1}{2} \sum_{\alpha\kappa M} \nabla\psi^{\alpha *}_{\kappa M}(\mathbf{r}) \cdot \nabla\psi^{\alpha}_{\kappa M}(\mathbf{r}), \tag{3.28}$$

where $\psi^{\alpha}_{\kappa M}(\mathbf{r})$ is the relativistic four-component core wave function obtained from the Dirac equation, as described in Sec. 2.3.2. Despite being a four-component solution, each component still consists of a product of a radial function and spherical harmonics. Consequently, the core KED can also be expressed in terms of spherical harmonics as:

$$\tau^{\text{core}}(\mathbf{r}) = \sum_{\alpha\ell m} \tau^{\alpha}_{\ell m}(r_\alpha) R_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{3.29}$$

Furthermore, to calculate $\tau^{\text{core}}(\mathbf{r})$, the same analytical formula used for the evaluation of the valence KED in the muffin-tin from Eq. (3.17) can be applied here again. The gradient of the wave function $\psi^{\alpha}_{i,\kappa M}(\mathbf{r})$ for either the major or minor component $i = \{1, 2\}$ can thus be expressed as:

$$\nabla\psi^{\alpha}_{i,\kappa M}(\mathbf{r}) = \sum_{m_s=\pm\frac{1}{2}} C^{JM}_{L,M-m_s,\frac{1}{2},m_s} \left[ u^{\alpha+}_{i,\kappa} \sum_{m_L=-(L+1)}^{L+1} \sum_{\gamma} C^{L,M-m_s}_{L+1,m_L,1,\gamma} Y_{L+1,m_L} \hat{\mathbf{e}}_\gamma \right.$$
$$\left. + u^{\alpha-}_{i,\kappa} \sum_{m_L=-(L-1)}^{L-1} \sum_{\gamma} C^{L,M-m_s}_{L-1,m_L,1,\gamma} Y_{L-1,m_L} \hat{\mathbf{e}}_\gamma \right] \chi_{m_s},$$

where $u^{\alpha-}_{i,\kappa}$ and $u^{\alpha+}_{i,\kappa}$ are defined similarly to Eq. (3.18) and Eq. (3.19), respectively, and $u^{\alpha}_{i,\kappa}$ represents the radial function of the major or minor component $i$.
Following the steps outlined in Appendix A.3 for the valence muffin-tin basis, the gradient is taken for each of the components, and from that, the KED is constructed by taking the dot product of the gradients of the wave function and its complex conjugate. This leads to the following expression for the major and minor components,

$$\tau^{\alpha}_{i,\ell m} = \frac{1}{2} \sum_{m_s=\pm\frac{1}{2}} \left( C^{JM}_{L,M-m_s,\frac{1}{2},m_s} \right)^* C^{JM}_{L,M-m_s,\frac{1}{2},m_s} \sum_{\pm} \sum_{\pm'} P^{\pm\pm'}_{\ell m;L,M-m_s} u^{\alpha\pm *}_{i,\kappa} u^{\alpha\pm'}_{i,\kappa}, \tag{3.30}$$

where $P^{\pm\pm'}_{\ell m;L,M-m_s}$ is given by

$$P^{\pm\pm'}_{\ell m;L,M-m_s} = \sum_{\substack{m'_L= \\ -(L\pm1)}}^{L\pm1} \sum_{\substack{m''_L= \\ -(L\pm1)}}^{L\pm1} \sum_{\gamma} \left(C^{L,M-m_s}_{L\pm1,m'_L,1,\gamma}\right)^* C^{L,M-m_s}_{L\pm'1,m''_L,1,\gamma} G^{\ell m}_{L\pm1,m'_L,L\pm'1,m''_L}, \tag{3.31}$$

consisting of a product of Clebsch-Gordan and Gaunt coefficients. The total core KED is then given by

$$\tau^{\text{core}}(\mathbf{r}) = \sum_{\alpha\ell m} \sum_{i} \tau^{\alpha}_{i,\ell m}(r_\alpha) R_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{3.32}$$

The calculation of the core KED is expected to be significantly faster compared to the muffin-tin valence KED, as it involves summing exclusively over the core states, which are defined through the quantum numbers $\{L, J, M\}$ or $\{\kappa, M\}$, rather than being expanded in spherical harmonics up to a defined $\ell_{\max}$.

### 3.1.4. Spin-Polarized Kinetic Energy Density

Up until now, the derivation of KED has been limited to spin-unpolarized systems. In both spin-unpolarized and spin-polarized systems, the core states are directly solved via the Dirac equation. Consequently, the KED of the core states remains unaffected by the second-variational treatment, which is applied only to the valence states. Because of the two-component spinor wave functions, the spin KED matrix is given by:

$$\bar{\tau}(\mathbf{r}) = \frac{1}{2} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \nabla \psi^{\dagger}_{n\mathbf{k}} \cdot \nabla \psi_{n\mathbf{k}} = \begin{pmatrix} \tau_{\sigma\sigma} & \tau_{\sigma\sigma'} \\ \tau_{\sigma'\sigma} & \tau_{\sigma'\sigma'} \end{pmatrix}. \tag{3.33}$$

Therefore, the spin-dependent elements $\tau_{\sigma\sigma'}$ must be calculated. Since spin-polarized systems are solved using the second-variational scheme, the second-variational wave function needs to be defined first. If we write the first-variational wave function in the muffin-tin region for atom $\alpha$ in its most general form for state $(j, \mathbf{k})$ [44],

$$\psi^{\alpha,\text{FV}}_{j\mathbf{k},\sigma}(\mathbf{r}) = \sum_{\lambda} C^{j\mathbf{k}\alpha,\text{FV}}_{\lambda\sigma} \phi^{\alpha}_{\lambda}(\mathbf{r}), \tag{3.34}$$

the second-variational spin-dependent wave function is then constructed as

$$\psi^{\alpha,\text{SV}}_{n\mathbf{k},\sigma}(\mathbf{r}) = \sum_{j} C^{n\mathbf{k}\alpha,\text{SV}}_{j\sigma} \psi^{\alpha,\text{FV}}_{j\mathbf{k},\sigma}(\mathbf{r}) = \sum_{j} C^{n\mathbf{k}\alpha,\text{SV}}_{j\sigma} \sum_{\lambda} C^{j\mathbf{k}\alpha,\text{FV}}_{\lambda\sigma} \phi^{\alpha}_{\lambda}(\mathbf{r}), \tag{3.35}$$

where $j$ runs over all occupied and some unoccupied states from the first-variational step [44]. The valence KED in the muffin-tin for the atom $\alpha$ is then defined as

$$\tau^{\alpha}(\mathbf{r}) = \frac{1}{2} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \nabla \psi^{\alpha,\text{SV}\dagger}_{n\mathbf{k}} \cdot \nabla \psi^{\alpha,\text{SV}}_{n\mathbf{k}}, \tag{3.36}$$

resulting in the spin KED matrix defined as:

$$\tau^{\alpha}(\mathbf{r}) = \frac{1}{2} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \begin{pmatrix} \nabla \psi^{\alpha,\text{SV}*}_{n\mathbf{k},\sigma} \cdot \nabla \psi^{\alpha,\text{SV}}_{n\mathbf{k},\sigma} & \nabla \psi^{\alpha,\text{SV}*}_{n\mathbf{k},\sigma} \cdot \nabla \psi^{\alpha,\text{SV}}_{n\mathbf{k},\sigma'} \\ \nabla \psi^{\alpha,\text{SV}*}_{n\mathbf{k},\sigma'} \cdot \nabla \psi^{\alpha,\text{SV}}_{n\mathbf{k},\sigma} & \nabla \psi^{\alpha,\text{SV}*}_{n\mathbf{k},\sigma'} \cdot \nabla \psi^{\alpha,\text{SV}}_{n\mathbf{k},\sigma'} \end{pmatrix}. \tag{3.37}$$

The product of the gradients of the wave functions can be generalized to:

$$\nabla\psi^{\alpha,\mathrm{SV}*}_{n\mathbf{k},\sigma}\cdot\nabla\psi^{\alpha,\mathrm{SV}}_{n\mathbf{k},\sigma'}=\sum_{\lambda_1\lambda_2}\left(\sum_j C^{n\mathbf{k}\alpha,\mathrm{SV}}_{j\sigma}C^{j\mathbf{k}\alpha,\mathrm{FV}}_{\lambda_1\sigma}\right)^*\left(\sum_i C^{n\mathbf{k}\alpha,\mathrm{SV}}_{i\sigma'}C^{i\mathbf{k}\alpha,\mathrm{FV}}_{\lambda_2\sigma'}\right)\nabla\phi^*_{\lambda_1}\cdot\nabla\phi_{\lambda_2}. \tag{3.38}$$

Here, the product of the gradients of the basis functions $\nabla\phi_\lambda$ can clearly be treated in the same way as in the spin-unpolarized case since they are spin-independent. The product of the first- and second-variational eigenvectors is spin-dependent, allowing us to define a $(n,\mathbf{k})$-dependent matrix, similar to Eq. (3.26):

$$D^\alpha_{\lambda_1\lambda_2,\sigma\sigma'}=\sum_{n\mathbf{k}}w_{\mathbf{k}}f_{n\mathbf{k}}\left(\sum_j C^{n\mathbf{k}\alpha,\mathrm{SV}}_{j\sigma}C^{j\mathbf{k}\alpha,\mathrm{FV}}_{\lambda_1\sigma}\right)^*\left(\sum_i C^{n\mathbf{k}\alpha,\mathrm{SV}}_{i\sigma'}C^{i\mathbf{k}\alpha,\mathrm{FV}}_{\lambda_2\sigma'}\right). \tag{3.39}$$

This leads to the following expression for the spin-dependent coefficients of the KED in the muffin-tin region:

$$\tau^\alpha_{\ell m,\sigma\sigma'}=\frac{1}{2}\left\{\sum_{\lambda_1\lambda_2}D^\alpha_{\lambda_1\lambda_2,\sigma\sigma'}\sum_{\pm}\sum_{\pm'}P^{\pm\pm'}_{\ell m;\ell_{\lambda_1}m_{\lambda_1}\ell_{\lambda_2}m_{\lambda_2}}g^{\alpha\pm}_{\lambda_1}g^{\alpha\pm'}_{\lambda_2}\right\}. \tag{3.40}$$

This expression differs from the spin-unpolarized definition only through the spin dependence in the eigenvectors within $D^\alpha_{\lambda_1\lambda_2,\sigma\sigma'}$.
Similarly, the second-variational wave functions in the interstitial region can be defined as:

$$\psi^{\mathrm{SV}}_{n\mathbf{k},\sigma}(\mathbf{r})=\sum_j C^{n\mathbf{k},\mathrm{SV}}_{j\sigma}\psi^{\mathrm{FV}}_{j\mathbf{k},\sigma}(\mathbf{r})=\sum_j C^{n\mathbf{k},\mathrm{SV}}_{j\sigma}\frac{1}{\sqrt{\Omega}}\sum_{\mathbf{G}}C^{j\mathbf{k},\mathrm{FV}}_{\mathbf{G}\sigma}e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{3.41}$$

where $j$ again runs over a number of specified states from the first-variational step. The KED in the interstitial region is thus given by

$$\tau_{\sigma\sigma'}(\mathbf{r})=\frac{1}{2}\sum_{n\mathbf{k}}\frac{w_{\mathbf{k}}f_{n\mathbf{k}}}{\Omega}\sum_{ij}C^{n\mathbf{k},\mathrm{SV}*}_{j\sigma}C^{n\mathbf{k},\mathrm{SV}}_{i\sigma'}\sum_{\mathbf{G}'\mathbf{G}''}C^{jk,\mathrm{FV}*}_{\mathbf{G}'\sigma}C^{ik,\mathrm{FV}}_{\mathbf{G}''\sigma'}\left[(\mathbf{G}'+\mathbf{k})\cdot(\mathbf{G}''+\mathbf{k})\right]e^{i(\mathbf{G}''-\mathbf{G}')\cdot\mathbf{r}}. \tag{3.42}$$

Defining $\mathbf{G}=\mathbf{G}''-\mathbf{G}'$, we again obtain the plane wave coefficients similar to the spin-unpolarized case:

$$\tau_{\mathbf{G},\sigma\sigma'}=\frac{1}{2}\sum_{n\mathbf{k}}\frac{w_{\mathbf{k}}f_{n\mathbf{k}}}{\Omega}\sum_{ij}C^{n\mathbf{k},\mathrm{SV}*}_{j\sigma}C^{n\mathbf{k},\mathrm{SV}}_{i\sigma'}\sum_{\mathbf{G}'}C^{jk,\mathrm{FV}*}_{\mathbf{G}'\sigma}C^{ik,\mathrm{FV}}_{\mathbf{G}+\mathbf{G}'\sigma'}\left[(\mathbf{G}'+\mathbf{k})\cdot(\mathbf{G}+\mathbf{G}'+\mathbf{k})\right]. \tag{3.43}$$

In practice, similar to the spin-unpolarized case, the gradient of the wave function is first constructed in reciprocal space. It is then transformed to real space using a FFT to compute the product of the gradients. This product is then converted back to reciprocal space to build the components of the spin KED matrix, reducing the computational cost.

## 3.2. Spin-Unpolarized Systems

With the KED derived, the focus shifts to the spin-unpolarized case. Section 3.2.1 begins with the derivation of the meta-GGA xc potential and its implementation in `exciting`. Following this, Sec. 3.2.2 covers the derivation of the Hamiltonian matrix elements for both the muffin-tin and interstitial regions. In Sec. 3.2.3, the influence of the LAPW basis on the meta-GGA SCF cycle is analyzed, along with the necessary adjustments. Although these formulas are developed for the spin-unpolarized case, they provide a foundation for extending the implementation to spin-polarized systems in later sections.

### 3.2.1. Exchange-Correlation Potential

In standard Kohn-Sham theory, as discussed in Sec. 2.1.2, the xc potential is evaluated by taking the functional derivative of the xc energy with respect to the density. However, because meta-GGAs introduce an additional dependency with the KED $\tau(\mathbf{r})$, the functionals become implicitly dependent on the density and explicitly dependent on the orbitals. As a result, evaluating the total functional derivative with respect to the density is less straightforward. Specifically, the functional derivative with respect to the density yields [33]:

$$V_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r})} = \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n(\mathbf{r})} - \nabla \cdot \left( \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \nabla n(\mathbf{r})} \right) + \int d\mathbf{r}' \, \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau(\mathbf{r}')} \frac{\delta \tau(\mathbf{r}')}{\delta n(\mathbf{r})}. \tag{3.44}$$

While the density-dependent parts of Eq. (3.44) are relatively simple to evaluate, the solution of the $\tau$-dependent integral requires the use of the OEP method [30, 31]. However, the solution of the integral using OEP can be computationally unstable and demanding [33], which is why it is common practice to approximate the OEP integral using the Krieger-Li-Iafrate (KLI) [96] approximation [33].
To circumvent the need for an approximation altogether, many instead choose to calculate the xc potential within the gKS scheme instead of the standard KS scheme, as this offers several advantages. First of all, there is no need for an approximation as with the OEP method. Additionally, as discussed in the previous Sec. 2.2.5, the gKS scheme directly captures the xc derivative discontinuity $\Delta_{\mathrm{xc}}$ and includes the xc derivative discontinuity in the solution of the gKS eigenvalues. In contrast, as the OEP scheme stays within the standard KS scheme, to gain an improvement in the prediction of the band gap, one has to additionally compute the contribution of the derivative discontinuity in the OEP frame. Furthermore, an OEP implementation requires the evaluation of $\nabla \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau(\mathbf{r})}$, which can cause numerical instabilities [33]. On the other hand, as will be discussed in the following sections, the gKS scheme only requires the evaluation of $\frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau(\mathbf{r})}$, which is expected to be numerically more stable, as it avoids the need to evaluate the gradient.
Within the gKS scheme, the energy functional $S[\{\psi_i\}]$ depends on the one-particle orbitals $\psi_i$, which make up the Slater determinant $\Psi$. The functional can then be defined as [97]:

$$S[\Psi] = \langle \Psi | \hat{T} | \Psi \rangle + E_{\mathrm{xc}}^{\mathrm{mGGA}}[\{\psi_i\}], \tag{3.45}$$

where $\hat{T}$ is the non-interacting kinetic energy operator and $E_{\mathrm{xc}}^{\mathrm{mGGA}}$ depends on the one-particle orbitals through the KED $\tau(\mathbf{r})$ and the density $n(\mathbf{r})$. The potential $\hat{V}_{\mathrm{xc}}^{\mathrm{mGGA}}$ is then obtained by taking the functional derivative with respect to the orbitals $\psi_i^*$ instead

of the density $n(\mathbf{r})$ [98, 99]:

$$\frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta \psi_i^*(\mathbf{r})} = \int d\mathbf{r}' \, \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r}')} \frac{\delta n(\mathbf{r}')}{\delta \psi_i^*(\mathbf{r})} = \frac{\delta E_{\mathrm{xc}}^{\mathrm{mGGA}}}{\delta n(\mathbf{r})} \psi_i(\mathbf{r}) = \hat{V}_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r}) \psi_i(\mathbf{r}), \tag{3.46}$$

which, if one follows the functional derivative scheme from Eq. (A.2) - (A.7), results in the following potential:

$$\hat{V}_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r}) = V_{\mathrm{xc}}^{\mathrm{mult}}(\mathbf{r}) + \hat{V}_{\mathrm{xc},\tau}^{\text{non-mult}}(\mathbf{r}). \tag{3.47}$$

Here, $V_{\mathrm{xc}}^{\mathrm{mult}}(\mathbf{r})$ has a form equivalent to the GGA potential and is multiplicative:

$$V_{\mathrm{xc}}^{\mathrm{mult}}(\mathbf{r}) = \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n(\mathbf{r})} - \nabla \cdot \left( \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \nabla n(\mathbf{r})} \right), \tag{3.48}$$

while $\hat{V}_{\mathrm{xc},\tau}^{\text{non-mult}}(\mathbf{r})$ is non-multiplicative:

$$\hat{V}_{\mathrm{xc},\tau}^{\text{non-mult}}(\mathbf{r}) = -\frac{1}{2} \nabla \cdot \left( \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau(\mathbf{r})} \nabla \cdot \right). \tag{3.49}$$

Approximating the Hartree potential as the remainder potential in the gKS scheme, the meta-GGA gKS equation is then given by [97]:

$$\left[ -\frac{1}{2}\nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + \hat{V}_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}). \tag{3.50}$$

In practice, the xc potential is computed using the `Libxc` library [55] for both the muffin-tin and interstitial region. Meta-GGA xc functionals in the `Libxc` library are defined using the following generalized variables:

$$E_{\mathrm{xc}}^{\mathrm{mGGA}} = \int d\mathbf{r} \, \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}(n(\mathbf{r}), \gamma(\mathbf{r}), \tau(\mathbf{r})), \tag{3.51}$$

where $\gamma(\mathbf{r}) = \nabla n(\mathbf{r}) \cdot \nabla n(\mathbf{r})$ is the contracted density gradient. Consequently, the derivation of the multiplicative potential leads to a slightly different formula:

$$\begin{aligned} V_{\mathrm{xc}}^{\mathrm{mult}}(\mathbf{r}) &= \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n(\mathbf{r})} - \nabla \cdot \left( \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma(\mathbf{r})} \frac{\partial \gamma(\mathbf{r})}{\partial \nabla n(\mathbf{r})} \right) \\ &= \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n(\mathbf{r})} - 2 \left[ \nabla \left( \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma(\mathbf{r})} \right) \cdot \nabla n(\mathbf{r}) + \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma(\mathbf{r})} \nabla^2 n(\mathbf{r}) \right]. \end{aligned} \tag{3.52}$$

The density $n(\mathbf{r})$, the contracted density gradient $\gamma(\mathbf{r})$, and the KED $\tau(\mathbf{r})$ need to be provided to the `Libxc` library as inputs. The library then returns $\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}$ and all required partial derivatives, from which both the multiplicative and non-multiplicative potentials can then be constructed.

### 3.2.2. Hamiltonian

Following the derivation of the xc potential $\hat{V}_{\text{xc}}^{\text{mGGA}}(\mathbf{r})$ in the previous section, the Hamiltonian matrix elements specific to the meta-GGA case will be derived in the following. The total Hamiltonian is defined as:

$$\hat{H} = -\frac{1}{2}\nabla^2 + \hat{V}_{\text{eff}}(\mathbf{r}). \tag{3.53}$$

Since the matrix elements associated with the kinetic energy operator are well-established and unaffected by the meta-GGA potential, they are not derived here. The total effective potential can be decomposed as follows:

$$\hat{V}_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}^{\text{mult}}(\mathbf{r}) + \hat{V}_{\tau}^{\text{non-mult}}(\mathbf{r}), \tag{3.54}$$

where $V_{\text{eff}}^{\text{mult}}(\mathbf{r})$ represents the multiplicative part, which includes the external, Hartree, and multiplicative density-dependent meta-GGA potentials. The non-multiplicative component, $\hat{V}_{\tau}^{\text{non-mult}}(\mathbf{r})$, accounts for the $\tau$-dependent meta-GGA xc potential. For the multiplicative part of the potential, the Hamiltonian matrix elements are given by:

$$\langle\phi_i|V_{\text{eff}}^{\text{mult}}|\phi_j\rangle = \int_{\Omega} d\mathbf{r}\, \phi_i^*(\mathbf{r})\, V_{\text{eff}}^{\text{mult}}(\mathbf{r})\, \phi_j(\mathbf{r}) = \left(\sum_{\alpha}\int_{S_\alpha} d\mathbf{r} + \int_{I} d\mathbf{r}\right) \phi_i^*(\mathbf{r})\, V_{\text{eff}}^{\text{mult}}(\mathbf{r})\, \phi_j(\mathbf{r}), \tag{3.55}$$

where the unit cell has been further divided into the interstitial region $I$ and the muffin-tin spheres $S_\alpha$ to account for the different basis sets in (L)APW. The derivation of the matrix elements corresponding to the multiplicative potential for the muffin-tin and interstitial basis is provided in Appendix A.4.
For the non-multiplicative potential, the integral can be re-expressed as follows:

$$\begin{aligned}\langle\phi_i|V_{\tau}^{\text{non-mult}}|\phi_j\rangle &= -\frac{1}{2}\int_{\Omega} d\mathbf{r}\, \phi_i^*(\mathbf{r})\, \nabla\cdot\left(\frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}\nabla\phi_j(\mathbf{r})\right) \\ &= \frac{1}{2}\int_{\Omega} d\mathbf{r}\, \frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}\nabla\phi_i^*(\mathbf{r})\cdot\nabla\phi_j(\mathbf{r}) - \frac{1}{2}\int_{\Omega} d\mathbf{r}\, \nabla\cdot\left(\frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\right).\end{aligned} \tag{3.56}$$

The first term is further divided into contributions from the MT and interstitial regions:

$$\frac{1}{2}\int_{\Omega} d\mathbf{r}\, \frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}\nabla\phi_i^*(\mathbf{r})\cdot\nabla\phi_j(\mathbf{r}) = \frac{1}{2}\left(\sum_{\alpha}\int_{S_\alpha} d\mathbf{r} + \int_{I} d\mathbf{r}\right) \frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}\nabla\phi_i^*(\mathbf{r})\cdot\nabla\phi_j(\mathbf{r}), \tag{3.57}$$

and clearly no longer requires the evaluation of the gradient of $\frac{\partial\varepsilon_{\text{xc}}^{\text{mGGA}}}{\partial\tau(\mathbf{r})}$, making the gKS scheme numerically more stable.
In the muffin-tin region, the (L)APW+LO basis functions can be rewritten in a more compact form as follows:

$$\phi_{\mu\mathbf{k}}^{\alpha}(\mathbf{r}) = \sum_{\lambda} T_{\lambda\mu}^{\alpha}\, g_{\lambda}^{\alpha}(r_\alpha)\; Y_{\ell_\lambda m_\lambda}(\hat{\mathbf{r}}_\alpha), \tag{3.58}$$

where $\lambda$ is a combined index that includes $\{\ell, m, \xi\}$ for the (L)APWs, and $\{L\}$ for the

local orbitals. The coefficients $T^{\alpha}_{\lambda\mu}$ are defined as:

$$T^{\alpha}_{\lambda\mu} = \begin{cases} A^{\mathbf{G}+\mathbf{k}}_{\ell m\,\xi,\alpha}, & \lambda \rightarrow \text{ (L)APW}, \\ \delta_{\alpha\alpha_L}, & \lambda \rightarrow \text{ LO}, \end{cases} \tag{3.59}$$

where $\mu$ corresponds to $\mathbf{G}+\mathbf{k}$ when associated with an (L)APW, and to $L$ when associated with a local orbital. Consequently, the size of the basis set, and thus the Hamiltonian matrix, is given by the sum of the (L)APW basis functions and the local orbitals. The radial functions $g^{\alpha}_{\lambda}(r_\alpha)$ are defined as in Eq. (3.16).
To evaluate the gradient of the basis functions $\nabla\phi^{\alpha}_{\mu\mathbf{k}}(\mathbf{r})$, the analytical expression for the gradient of a function expanded in spherical harmonics, as introduced in Eq. (3.17), is applied here again. Defining $v_\tau(\mathbf{r}) = \frac{\partial \varepsilon^{\text{mGGA}}_{\text{xc}}}{\partial \tau(\mathbf{r})}$, the $\tau$-dependent contribution within the muffin-tin region is also expanded in spherical harmonics as:

$$v^{\alpha}_{\tau}(\mathbf{r}) = \sum_{\ell m} v^{\alpha}_{\tau,\ell m}(r_\alpha)\, Y_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{3.60}$$

With this expansion, the matrix elements for the first part of the non-multiplicative potential in the muffin-tin $\alpha$ are given by:

$$\begin{aligned} V^{\alpha}_{\tau,\mu\nu} &= \langle \nabla\phi^{\alpha}_{\mu\mathbf{k}} \mid v^{\alpha}_{\tau} \mid \nabla\phi^{\alpha}_{\nu\mathbf{k}} \rangle \\ &= \int_{S_\alpha} d\mathbf{r}\, \nabla\phi^{\alpha *}_{\mu\mathbf{k}}(\mathbf{r}) \sum_{\ell m} v^{\alpha}_{\tau,\ell m}(r_\alpha)\, Y_{\ell m}(\hat{\mathbf{r}}_\alpha)\, \nabla\phi^{\alpha}_{\nu\mathbf{k}}(\mathbf{r}) \\ &= \sum_{\lambda_1\lambda_2} C^{\alpha *}_{\lambda_1\mu} C^{\alpha}_{\lambda_2\nu} \sum_{\ell m} R^{\alpha\pm\pm'}_{\ell m;\lambda_1\lambda_2}\, P^{\pm\pm'}_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}}. \end{aligned} \tag{3.61}$$

Here,

$$R^{\alpha\pm\pm'}_{\ell m;\lambda_1\lambda_2} = \int_0^{R_{\text{MT}_\alpha}} g^{\alpha\pm}_{\lambda_1}(r)\, v^{\alpha}_{\tau,\ell m}(r_\alpha)\, g^{\alpha\pm'}_{\lambda_2}(r)\, r^2\, dr, \tag{3.62}$$

is the radial integral over the product of the radial functions $g^{\alpha\pm}_{\lambda}(r)$, as defined in Eq. (3.18) and Eq. (3.19), and the radial function of the non-multiplicative potential $v^{\alpha}_{\tau}(r_\alpha)$. The spherical part is given by

$$P^{\pm\pm'}_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}} = \sum_{\substack{m'= \\ -(\ell_{\lambda_1}\pm 1)}}^{\ell_{\lambda_1}\pm 1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}\pm' 1)}}^{\ell_{\lambda_2}\pm' 1} \sum_{\gamma} C^{\ell_{\lambda_1}, m_{\lambda_1} *}_{\ell_{\lambda_1}\pm 1, m', 1, \gamma}\, C^{\ell_{\lambda_2}, m_{\lambda_2}}_{\ell_{\lambda_2}\pm' 1, m'', 1, \gamma}\, \langle Y_{\ell_{\lambda_1}\pm 1, m'} | R_{\ell m} | Y_{\ell_{\lambda_2}\pm' 1, m''} \rangle, \tag{3.63}$$

a product of an integral over three spherical harmonics, known as the Gaunt coefficients, and Clebsch-Gordan coefficients, which arise through the vector spherical harmonics as defined in Eq. (3.20). Since the radial integral is $\mathbf{k}$-independent, it is evaluated only once per SCF iteration to improve computational efficiency. The Gaunt and Clebsch-Gordan coefficients are precomputed, allowing the spherical integral to be efficiently evaluated as well. These integrals are then multiplied by the $\mathbf{k}$-dependent coefficients for each $\mathbf{k}$-point.
In the interstitial region, the basis functions are given by

$$\phi_{\mathbf{G}+\mathbf{k}} = \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}, \tag{3.64}$$

and the potential is similarly expanded in plane waves as

$$v_\tau^{\mathrm{IR}}(\mathbf{r}) = \sum_{\mathbf{G}} v_\tau^{\mathrm{IR}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{3.65}$$

Thus, the following expression is derived for the matrix elements of the non-multiplicative potential in the interstitial region:

$$\begin{aligned} v^{\mathrm{IR}}_{\tau,\mathbf{G_1G_2}} &= \langle \nabla\phi_{\mathbf{G_1}+\mathbf{k}} \mid v_\tau^{\mathrm{IR}} \mid \nabla\phi_{\mathbf{G_2}+\mathbf{k}} \rangle \\ &= \int_{\mathrm{IR}} d\mathbf{r}\, \nabla\phi^*_{\mathbf{G_1}+\mathbf{k}}(\mathbf{r}) \cdot v_\tau^{\mathrm{IR}}(\mathbf{r})\, \nabla\phi_{\mathbf{G_2}+\mathbf{k}}(\mathbf{r}) \\ &= \sum_{\mathbf{G}} v_\tau^{\mathrm{IR}}(\mathbf{G})\, (\mathbf{G}_1+\mathbf{k}) \cdot (\mathbf{G}_2+\mathbf{k})\, \frac{1}{\Omega} \int_\Omega e^{-i(\mathbf{G}_1-\mathbf{G}_2-\mathbf{G})\cdot\mathbf{r}}\, \Theta_{\mathrm{IR}}(\mathbf{r})\, d\mathbf{r} \\ &= \sum_{\mathbf{G}} v_\tau^{\mathrm{IR}}(\mathbf{G})\, (\mathbf{G}_1+\mathbf{k}) \cdot (\mathbf{G}_2+\mathbf{k})\, \Theta_{\mathrm{IR}}(\mathbf{G}_1-\mathbf{G}_2-\mathbf{G}) \\ &= v^{\mathrm{IR}}_{\tau,\Theta}(\mathbf{G}_1-\mathbf{G}_2)\, (\mathbf{G}_1+\mathbf{k}) \cdot (\mathbf{G}_2+\mathbf{k}). \end{aligned} \tag{3.66}$$

The characteristic function $\Theta_{\mathrm{IR}}(\mathbf{r})$ is used to ensure that the integral over the unit cell volume $\Omega$ contributes only where $\mathbf{r} \in \mathrm{IR}$, as $\Theta_{\mathrm{IR}}(\mathbf{r})$ is defined to be zero outside the interstitial region. In practice, the matrix elements in the interstitial region are evaluated similarly to the KED by multiplying the interstitial potential $v_\tau^{\mathrm{IR}}(\mathbf{G})$ and the characteristic function in real space, avoiding the computational cost of convolution in reciprocal space. To accurately capture high-frequency components associated with large $\mathbf{G}$-vectors, the potential is transformed to real space using a grid with double the original density. After multiplying $v_\tau^{\mathrm{IR}}(\mathbf{r}) \cdot \Theta_{\mathrm{IR}}(\mathbf{r})$, the product is transformed back to reciprocal space using the FFT and truncated back to the original density.
As a final step, the second term in Eq. (3.56) must be evaluated to obtain the total matrix elements for the non-multiplicative potential. Gauss' theorem is applied to convert the volume integral into a surface integral over the muffin-tin boundary, which can be further divided into contributions from the muffin-tin and surface regions:

$$\begin{aligned} -\frac{1}{2}\int_\Omega \nabla\cdot\Big(v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\Big)\,d\mathbf{r} = -\frac{1}{2}\Big[\sum_\alpha \int_{\partial S_\alpha} v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\cdot d\mathbf{S}_{\mathrm{MT}_\alpha} \\ + \int_{\partial S_I} v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\cdot d\mathbf{S}_I\Big], \end{aligned} \tag{3.67}$$

where $d\mathbf{S}_{\mathrm{MT}_\alpha} = \mathbf{n}\, dS_{\mathrm{MT}_\alpha}$, with $\mathbf{n}$ being the outward-pointing unit normal vector on the surface $\partial S_\alpha$. In contrast, $d\mathbf{S}_I$ points outward from the interstitial surface, such that $d\mathbf{S}_I = -d\mathbf{S}_{\mathrm{MT}_\alpha}$. Thus, we can write

$$-\frac{1}{2}\int_\Omega \nabla\cdot\Big(v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\Big)\,d\mathbf{r} = -\frac{1}{2}\sum_\alpha \int_{\partial S_\alpha} \Big[v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\Big|_{\mathrm{MT}} \tag{3.68}$$

$$-v_\tau(\mathbf{r})\,\phi_i^*(\mathbf{r})\,\nabla\phi_j(\mathbf{r})\Big|_I\Big]\cdot d\mathbf{S}_{\mathrm{MT}_\alpha}, \tag{3.69}$$

where the orbitals are evaluated using either the muffin-tin basis $\big|_{\mathrm{MT}}$ or the interstitial basis $\big|_I$ to account for the contributions from the respective sides of the surface.
The respective surface terms evaluated for the muffin-tin and interstitial regions are generally non-zero. However, when using the LAPW method, the first derivative of the radial

function is matched at the muffin-tin boundary, ensuring a continuous derivative. This matching causes the surface terms of both regions to cancel each other. In contrast, with the APW method, continuity is not maintained because the radial functions of the basis functions are matched only in value, not in slope. Consequently, the gradients of the basis functions and the potential $v_\tau(\mathbf{r})$ are discontinuous, leading to a non-zero surface term. The authors of Ref. [7] reported that including a surface term in their implementation of the meta-GGA potential in the `WIEN2k` code led to instabilities. They also found that the surface term had a negligible influence on the outcomes of their calculations, which prompted them to discard the surface term in their final work. In this thesis, the surface term is similarly excluded. To ensure the accuracy of all calculations, only the LAPW basis is used for further computations.

### 3.2.3. Treatment of the Basis

When implementing the meta-GGA within the gKS scheme, a drawback arises with the (L)APW+LO method. This issue becomes evident when examining the setup of the radial functions $u_\ell^\alpha(r)$ of the basis functions $\phi_i(\mathbf{r})$ in the muffin-tin region. As discussed in Sec. 2.3, these radial functions are obtained by directly solving the radial Schrödinger equation through integration. For this, the radial Schrödinger equation is transformed into two coupled first-order differential equations, which are then solved for two grid points $r_{i+1}$ and $r_i$. If the interval between $r_{i+1}$ and $r_i$ is too large, it introduces significant errors, requiring the solution over many grid points [22]. In `exciting`, the Bulirsch-Stoer algorithm [100] is used to enhance the efficiency and accuracy of this integration process. The algorithm begins by dividing the integration interval into smaller subintervals and solving the differential equations over these subintervals. This is done for an increasing number of subintervals, after which this sequence of different subintervals is used to extrapolate to an infinite number of subintervals, reducing numerical errors. The potential $V_{\text{eff}}(r)$ must be evaluated at any point within $0 < r < R_{\text{MT}}$. When necessary, cubic spline interpolation is used to determine $V_0(r)$ at points not directly on the grid [22].
In the case of meta-GGA functionals, the potential includes the non-multiplicative potential:

$$\hat{V}_\tau^{\text{non-mult}}(\mathbf{r}) = -\frac{1}{2}\nabla \cdot \left( \frac{\partial \varepsilon_{xc}^{\text{mGGA}}}{\partial \tau(\mathbf{r})} \nabla \cdot \right), \tag{3.70}$$

where the $\nabla\cdot$ operator acts directly on the yet unknown radial functions $u_\ell^\alpha(r)$. As a result, the potential cannot be fully determined without knowing the radial functions. Consequently, the Bulirsch-Stoer algorithm cannot be directly applied, as it relies on the fact that the potential is known beforehand.
There have been attempts to address this issue. For instance, a cubic spline solver method that solves the radial Schrödinger equation with a gKS meta-GGA potential was presented by Holtzwarth et al. [101]. This method, however, solves the radial equation by diagonalizing an algebraic eigenvalue problem. In contrast, (L)APW-based methods calculate radial functions at chosen energy parameters close to band energies, not at atom-like eigenvalues that would result from the algebraic solution [102].
The authors of Ref. [7], who implemented self-consistent meta-GGAs in the `WIEN2k` code, decided to instead use a GGA potential to solve the radial Schrödinger equation and obtain the radial functions $u_\ell^\alpha(r)$. The appropriate GGA potential was determined using the variational principle in Ref. [103], where the band gap was calculated non-self-consistently using a meta-GGA functional by evaluating the meta-GGA total energy. The orbitals

used to construct the meta-GGA total energy were obtained from a self-consistent GGA calculation. This procedure was repeated for various GGA functionals to identify the one that minimized the non-self-consistent meta-GGA total energy, thus determining the optimal GGA functional for constructing the most accurate meta-GGA total energy.
The authors of Ref. [7] then used the optimal GGA functionals identified in Ref. [103] to set up the basis in the self-consistent meta-GGA scheme. They validated this approach by showing that the resulting converged properties, such as band gaps, lattice constants, and magnetic moments, were close to those obtained with the `VASP` code [104]. In addition to that, they also compared results obtained with the variationally optimal GGA functional to those obtained when using the PBE functional and concluded that for most cases, using the PBE functional for the basis setup gives satisfactory results as well. Consequently, in this implementation of self-consistent meta-GGA, the same scheme is applied. It will later be verified that using either the variationally optimal GGA potential determined in Ref. [103] or the PBE functional leads to similar results as in the `WIEN2k` code [24].
However, using GGA functionals to obtain the radial functions of the muffin-tin basis results in a SCF cycle that differs from the standard KS cycle, which was introduced in Fig. 2.1. Therefore, the necessary modifications to the meta-GGA SCF cycle are presented in Fig. 3.1.

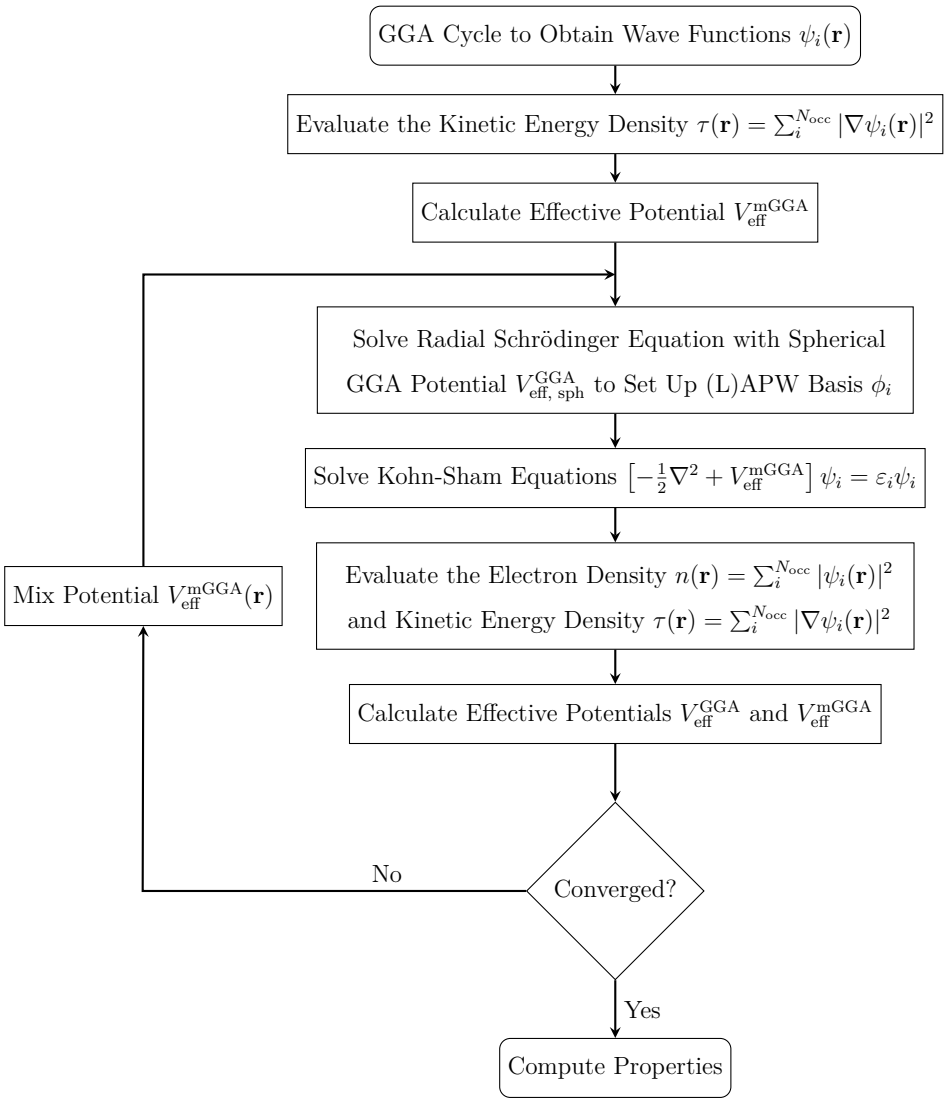



Figure 3.1.: Flowchart of the SCF cycle for a meta-GGA calculation in `exciting`.

Initially, a full standard KS SCF cycle is completed using a GGA functional. This step is necessary to obtain the initial wave functions required to evaluate the KED $\tau(\mathbf{r})$. The lack of wave functions in the first iteration is not an issue for GGA functionals, as the initial density can be constructed from atomic densities to provide a first guess. From the resulting KED, the effective meta-GGA potential is constructed. The radial Schrödinger equation is then solved using the spherical part of the GGA potential. The radial functions obtained by solving the radial Schrödinger equation are used to set up the full (L)APW+LO basis and construct the Hamiltonian and overlap matrices to solve the KS equations. From the solutions of the KS equations, the new density, KED and the potentials can be recalculated. The convergence criterion can be applied to the difference in the density, KED, or the potential. In the calculations presented in this work, only potential mixing is applied. This approach has been shown to be more stable, and it requires mixing only one variable instead of two, as in the case of density and KED. Since the potential determines the KS orbitals, which are then used to construct both the density and KED, mixing the potential effectively ensures that both the density and KED are indirectly mixed toward their self-consistent values.
So far, only the valence states have been discussed. Core states are solved by numerical integration of the Dirac equation. Due to the non-multiplicative component of the meta-GGA potential, the authors of Ref. [7] also use the optimal GGA potential to solve the Dirac equation. However, this approach implies that the scheme is not fully self-consistent for meta-GGA, as the core states are treated only within the GGA potential [105].

## 3.3. Spin-Polarized Systems

Having established the self-consistent scheme and formulas for the spin-unpolarized case, the following section extends the derivation of the xc potential and the setup of the Hamiltonian to the spin-polarized case.
The meta-GGA functionals in the spin-polarized case are defined as:

$$E_{\mathrm{xc}}^{\mathrm{mGGA}} = \int d\mathbf{r}\, \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}(n_\uparrow, n_\downarrow, \gamma_{\uparrow\uparrow}, \gamma_{\downarrow\downarrow}, \gamma_{\uparrow\downarrow}, \tau_\uparrow, \tau_\downarrow). \tag{3.71}$$

To derive the xc potential $V_{\mathrm{xc}}(\mathbf{r})$ and the magnetic xc field $\mathbf{B}_{\mathrm{xc}}(\mathbf{r})$, we will now define the components of the KED $\tau_{\uparrow/\downarrow}$ for the general non-collinear case. Analogous to the four-component vector $\mathbf{n}(\mathbf{r})$, introduced in Sec. 2.1.4, a four-component vector $\boldsymbol{\tau}(\mathbf{r})$ can be defined [39]:

$$\boldsymbol{\tau}(\mathbf{r}) = \sum_{i=0}^{3} \mathrm{tr}\{\sigma_i \bar{\tau}(\mathbf{r})\}, \tag{3.72}$$

with the spin KED matrix $\bar{\tau}(\mathbf{r})$ given by

$$\bar{\tau}(\mathbf{r}) = \frac{1}{2} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \nabla \psi_{n\mathbf{k}}^\dagger \cdot \nabla \psi_{n\mathbf{k}} = \begin{pmatrix} \tau_{\sigma\sigma} & \tau_{\sigma\sigma'} \\ \tau_{\sigma'\sigma} & \tau_{\sigma'\sigma'} \end{pmatrix}. \tag{3.73}$$

This definition yields:

$$\boldsymbol{\tau}(\mathbf{r}) = \begin{bmatrix} \tau_{\sigma\sigma}(\mathbf{r}) + \tau_{\sigma'\sigma'}(\mathbf{r}) \\ \tau_{\sigma\sigma'}(\mathbf{r}) + \tau_{\sigma'\sigma}(\mathbf{r}) \\ i(\tau_{\sigma\sigma'}(\mathbf{r}) - \tau_{\sigma'\sigma}(\mathbf{r})) \\ \tau_{\sigma\sigma}(\mathbf{r}) - \tau_{\sigma'\sigma'}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \tau(\mathbf{r}) \\ u_x(\mathbf{r}) \\ u_y(\mathbf{r}) \\ u_z(\mathbf{r}) \end{bmatrix}, \tag{3.74}$$

which defines the scalar KED $\tau(\mathbf{r})$ and the components of the spin KED $u_i(\mathbf{r})$. From this, $\bar{\tau}(\mathbf{r})$ can be reformulated in terms of the four-component vector as [39]

$$\bar{\tau}(\mathbf{r}) = \frac{1}{2}\sum_{i=0}^{3}\sigma_i\tau_i(\mathbf{r}) = \frac{1}{2}\begin{pmatrix} \tau + u_z & u_x - iu_y \\ u_x + iu_y & \tau - u_z \end{pmatrix}. \tag{3.75}$$

Next, we introduce the generalized variables $\tau_{\uparrow/\downarrow}$ for the non-collinear case, following Kübler's scheme. To transform the spin KED matrix to $\tau_{\uparrow/\downarrow}$, we transform the matrix in Eq. (3.75) accordingly. Several definitions for these generalized variables have been proposed. The most straightforward definition is [39]:

$$\tau_{\uparrow/\downarrow}(\mathbf{r}) = \frac{1}{2}\left(\tau(\mathbf{r}) \pm \hat{\mathbf{m}}(\mathbf{r}) \cdot \mathbf{u}(\mathbf{r})\right), \tag{3.76}$$

where $\mathbf{u}(\mathbf{r})$ is projected onto the same quantization axis as $\mathbf{m}(\mathbf{r})$ to ensure proper treatment in the collinear limit. Similarly, the contracted density $\gamma_{\sigma\sigma'}(\mathbf{r})$ can be defined as

$$\gamma_{\sigma\sigma'}(\mathbf{r}) = \nabla n_\sigma(\mathbf{r}) \cdot \nabla n_{\sigma'}(\mathbf{r}), \tag{3.77}$$

where $\sigma$ and $\sigma'$ can be $\uparrow$ or $\downarrow$. The drawback of this approach is that when substituting the density, $n_{\uparrow/\downarrow} = \frac{1}{2}(n \pm |\mathbf{m}|)$, into Eq. (3.77), only the longitudinal variation of $\mathbf{m}$ is considered. This is because $\gamma_{\sigma\sigma'}$ then only includes terms where the change in magnetization $\nabla\mathbf{m}$ is projected onto the direction of the magnetization $\hat{\mathbf{m}}$ [106]. As a result, the energy is insensitive to changes in the direction of $\mathbf{m}$, causing $\mathbf{B}_{\text{xc}}$ to remain parallel to the magnetization $\mathbf{m}$. This limitation prevents the capture of local torque moments. In general, the generalization of variables in the non-collinear case must adhere to the global torque theorem, which states that the magnetic field does not exert any net torque on the total system [106]:

$$0 = \int d\mathbf{r}\, \mathbf{m}(\mathbf{r}) \times \mathbf{B}_{\text{xc}}(\mathbf{r}). \tag{3.78}$$

However, local torques can still arise, and in fact, they are essential for accurately capturing spin dynamics and time-dependent behavior of the magnetization [107].
To address this limitation, Scalmani and Frisch [108, 109] proposed an alternative method for defining functional variables that incorporates both longitudinal and transverse changes in magnetization. This approach allows the xc functional to capture local torques while satisfying the global zero torque theorem [110], and most importantly, has demonstrated improved numerical stability [111]. Scalmani and Frisch define the variables up to the meta-GGA case as

$$\begin{aligned}
n_{\uparrow/\downarrow} &= \frac{1}{2}\left(n \pm \sqrt{\mathbf{m}\cdot\mathbf{m}}\right), \\
\gamma_{\uparrow\uparrow/\downarrow\downarrow} &= \frac{1}{4}\nabla n\cdot\nabla n + \frac{1}{4}\nabla\mathbf{m}\cdot\circ\nabla\mathbf{m} \pm \frac{f_\nabla}{2}\sqrt{\nabla n\cdot\nabla\mathbf{m}\circ\nabla\mathbf{m}\cdot\nabla n}, \\
\gamma_{\uparrow\downarrow} &= \frac{1}{4}\nabla n\cdot\nabla n - \frac{1}{4}\nabla\mathbf{m}\cdot\circ\nabla\mathbf{m}, \\
\tau_{\uparrow/\downarrow} &= \frac{1}{2}\left(\tau \pm f_u\sqrt{\mathbf{u}\circ\mathbf{u}}\right), \\
f_\nabla &= \operatorname{sgn}(\nabla n\cdot\nabla\mathbf{m}\circ\mathbf{m}), \\
f_u &= \operatorname{sgn}(\mathbf{u}\circ\mathbf{m}).
\end{aligned} \tag{3.79}$$

In this formulation the dot "$\cdot$" contracts indices of the gradient operator $\nabla$ and the circle

"$\circ$" contracts the indices of the three-dimensional vector fields $\mathbf{m}$ and $\mathbf{u}$.
For very small magnetization values $\mathbf{m} < 10^{-12}$, the transformations in Eq. (3.79) can cause numerical instabilities in the xc potential due to the appearance of the magnetization length $\sqrt{\mathbf{m} \cdot \mathbf{m}}$ in the denominator during functional differentiation. To mitigate this, alternative scalar transformations are introduced to ensure stable convergence for small $\mathbf{m}$ [112]:

$$\begin{aligned} \sqrt{\mathbf{m} \cdot \mathbf{m}} &\to m_s = \frac{1}{3} \sum_i m_i, \\ \sqrt{\nabla n \cdot \nabla \mathbf{m} \circ \nabla \mathbf{m} \cdot \nabla n} &\to \nabla n \cdot \nabla m_s, \end{aligned} \tag{3.80}$$

and equally for the meta-GGA extension,

$$\sqrt{\mathbf{u} \cdot \mathbf{u}} \to u_s = \frac{1}{3} \sum_i u_i. \tag{3.81}$$

One drawback of this method, pointed out by Vingale et al. [106], is its equal treatment of transverse and longitudinal components. To address this, they introduced a formalism that weighs the transversal and longitudinal components differently. Nevertheless, the formalism proposed by Scalmani and Frisch has been shown to produce satisfactory results for systems with weak or significantly inhomogeneous spin polarization [106, 108, 109].
As discussed, meta-GGA functionals can pose numerical challenges. To address these, we adopt the Scalmani and Frisch formalism for the meta-GGA extension, which offers improved numerical stability [111]. In the following sections, we will derive the xc potential using this formalism and evaluate the Hamiltonian matrix elements for the spin-polarized case within the second-variational scheme.

### 3.3.1. Exchange-Correlation Potential

With the definition of the generalized variables in Eq. (3.79), the total xc potential can be evaluated. As derived earlier, the potential is given by:

$$\bar{V}_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}}{\delta n} \sigma_0 + \sum_{i=1,2,3} \frac{\delta E_{\mathrm{xc}}}{\delta m_i} \sigma_i. \tag{3.82}$$

However, as in the spin-unpolarized case, the meta-GGA potential is only implicitly dependent on the density $n(\mathbf{r})$ and magnetization $\mathbf{m}(\mathbf{r})$, making it difficult to evaluate the functional derivative of the xc energy with respect to them. For computational codes that expand the wave functions in a basis set, such as $\psi_{i,\sigma} = \sum_\mu C^i_{\mu\sigma} \phi_\mu$, a density matrix is defined as [113]:

$$D^{\sigma\sigma'}_{\mu\nu} = \sum_i C^{i*}_{\mu\sigma} C^i_{\nu\sigma'}, \tag{3.83}$$

as already derived in Eq. (3.26) and Eq. (3.39) for the KED.
The spin-dependent electron density $n_{\sigma\sigma'}$ can be expressed in terms of the density matrix as:

$$n_{\sigma\sigma'} = \sum_{\mu\nu} D^{\sigma\sigma'}_{\mu\nu} \phi^*_\mu(\mathbf{r}) \, \phi_\nu(\mathbf{r}), \tag{3.84}$$

and following the generalization of the variables for the general non-collinear case in

Eq. (3.79), the density $n(\mathbf{r})$ can be written as:

$$n(\mathbf{r}) = \sum_{\mu\nu} D^0_{\mu\nu}\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu\nu}\left(D^{\sigma\sigma}_{\mu\nu} + D^{\sigma'\sigma'}_{\mu\nu}\right)\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}). \tag{3.85}$$

Similarly, the scalar KED and spin KED can be rewritten in terms of $D^{\sigma\sigma'}_{\mu\nu}$ as follows:

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu} D^0_{\mu\nu}\nabla\phi^*_\mu(\mathbf{r})\nabla\phi_\nu(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu}\left(D^{\sigma\sigma}_{\mu\nu} + D^{\sigma'\sigma'}_{\mu\nu}\right)\nabla\phi^*_\mu(\mathbf{r})\cdot\nabla\phi_\nu(\mathbf{r}), \tag{3.86}$$

$$u_x(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu} D^x_{\mu\nu}\nabla\phi^*_\mu(\mathbf{r})\nabla\phi_\nu(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu}\left(D^{\sigma\sigma'}_{\mu\nu} + D^{\sigma'\sigma}_{\mu\nu}\right)\nabla\phi^*_\mu(\mathbf{r})\cdot\nabla\phi_\nu(\mathbf{r}), \tag{3.87}$$

$$u_y(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu} D^y_{\mu\nu}\nabla\phi^*_\mu(\mathbf{r})\nabla\phi_\nu(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu} i\left(D^{\sigma\sigma'}_{\mu\nu} - D^{\sigma'\sigma}_{\mu\nu}\right)\nabla\phi^*_\mu(\mathbf{r})\cdot\nabla\phi_\nu(\mathbf{r}), \tag{3.88}$$

$$u_z(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu} D^z_{\mu\nu}\nabla\phi^*_\mu(\mathbf{r})\nabla\phi_\nu(\mathbf{r}) = \frac{1}{2}\sum_{\mu\nu}\left(D^{\sigma\sigma}_{\mu\nu} - D^{\sigma'\sigma'}_{\mu\nu}\right)\nabla\phi^*_\mu(\mathbf{r})\cdot\nabla\phi_\nu(\mathbf{r}). \tag{3.89}$$

The magnetization $\mathbf{m}(\mathbf{r})$ can be rewritten similarly (see Appendix A.2).
The matrix elements of the xc potential for the Hamiltonian are then evaluated by taking the functional derivative with respect to the density matrix $D^i_{\mu\nu}$ as follows [112],

$$(V_{\mathrm{xc}})^i_{\mu\nu} = \sum_{pq}\int d\mathbf{r}\,\frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial U_p}\frac{\partial U_p}{\partial V_q}\frac{\partial V_q}{\partial D^i_{\mu\nu}}. \tag{3.90}$$

Here, $i = \{0,1,2,3\}$ represents all components of scalar density and the magnetization $\mathbf{m}$, as well as the scalar KED and the spin KED $\mathbf{u}$. The variables are $\{U\} = \{n^\uparrow, n^\downarrow, \gamma^{\uparrow\uparrow}, \gamma^{\downarrow\downarrow}, \gamma^{\uparrow\downarrow}, \tau^\uparrow, \tau^\downarrow\}$, as defined by Scalmani and Frisch in Eq. (3.79) and $\{V\} = \{n, \mathbf{m}, \tau, \mathbf{u}\}$. The partial derivatives of $\{V_q\}$ can be obtained by differentiating equations (3.85) - (3.89). The partial derivatives of $\{U_p\}$ and the resulting potential matrix elements are provided in the Appendix A.2.
In `exciting`, the xc potential is evaluated first, and later the matrix elements are calculated as:

$$(V_{\mathrm{xc}})_{\mu\nu} = \int d\mathbf{r}\,\phi^*_\mu(\mathbf{r})\,V_{\mathrm{xc}}\,\phi_\nu(\mathbf{r}). \tag{3.91}$$

The expressions for the xc potential and xc magnetic potential can be derived from the matrix elements $(V_{\mathrm{xc}})^0_{\mu\nu}$ (see Eq. (A.19)) and $(B_{\mathrm{xc}})^i_{\mu\nu}$ (see Eq. (A.24)), respectively. Making use of partial integration, the scalar potential is given by

$$\begin{aligned}(V_{\mathrm{xc}})^0 = {} & \frac{1}{2}\left(\frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\uparrow} + \frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\downarrow}\right) \\ & - \frac{1}{2}\nabla\left[\left(\frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\gamma_{\uparrow\uparrow}} + \frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\gamma_{\downarrow\downarrow}} + \frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\gamma_{\uparrow\downarrow}}\right)\nabla n\right] \\ & - \frac{f_\nabla}{2}\nabla\left[\left(\frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\gamma_{\uparrow\uparrow}} - \frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\gamma_{\downarrow\downarrow}}\right)\frac{\nabla n\cdot\nabla\mathbf{m}\circ\nabla\mathbf{m}}{\Gamma}\right] \\ & - \frac{1}{4}\nabla\left[\left(\frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\tau_\uparrow} + \frac{\partial\varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial\tau_\downarrow}\right)\nabla\cdot\right],\end{aligned} \tag{3.92}$$

and each component $i$ of the xc magnetic potential is given by

$$\begin{aligned}(B_{\mathrm{xc}})^i = &\frac{1}{2}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\downarrow}\right)\frac{m^i}{|\mathbf{m}|} \\ &- \frac{1}{2}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\downarrow}}\right)\nabla m^i\right] \\ &- \frac{f_\nabla}{2}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}}\right)\frac{f_\nabla}{2}\frac{\nabla n\cdot\nabla\mathbf{m}\cdot\nabla n}{\Gamma}\right] \\ &- \frac{f_{\mathbf{u}}}{4}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\frac{u^i}{|\mathbf{u}|}\nabla\cdot\right].\end{aligned} \tag{3.93}$$

As previously mentioned, in regions where the magnetization is small, the generalized variables are redefined. The corresponding modifications to the xc potential and xc magnetic field, implemented to handle these regions, are given in the Appendix A.2.

### 3.3.2. Hamiltonian

Following the derivation of the potential and its matrix elements in the previous section, this section discusses the specific setup of the matrix elements for the spin-polarized case. Since the spin-polarized case is solved using the second-variational scheme, the scalar (relativistic) Hamiltonian in the first-variational scheme is set up first. As in the spin-unpolarized case, the potential is split into a multiplicative and a non-multiplicative part, such that the non-multiplicative potential is given by its $\tau$-dependent part, and the remaining density-dependent part of the xc potential contributes to the multiplicative potential $V_{\mathrm{eff}}^{\mathrm{mult}}$. To evaluate the multiplicative potential in the muffin-tin and interstitial regions, we apply the same scheme derived in the Appendix A.4. Similarly, for the non-multiplicative part the derivation given in Eq. (3.61) for the muffin-tin and in Eq. (3.66) for the interstitial region holds.
After evaluating the total scalar (relativistic) Hamiltonian of the first variation, the resulting first-variational eigenvectors $\psi_{n\mathbf{k}}^{\mathrm{FV}}$ are used to calculate the magnetic Hamiltonian matrix elements for the second-variational step, which includes the external magnetic field $\mathbf{B}_{\mathrm{ext}}$ and the xc magnetic field $\mathbf{B}_{\mathrm{xc}}$. The total effective magnetic field is again divided into multiplicative and non-multiplicative parts. Thus, in the muffin-tin region for atom $\alpha$, the Hamiltonian matrix for the component $i$ of the xc magnetic field is given by:

$$\begin{aligned}(B_{\mathrm{eff}}^\alpha)_{mn,\sigma\sigma'}^i &= \langle\psi_{m\mathbf{k},\sigma}^{\alpha,\mathrm{FV}}|\,(B_{\mathrm{eff}}^\alpha)^i\,|\psi_{n\mathbf{k},\sigma'}^{\alpha,\mathrm{FV}}\rangle \\ &= \sum_{\mu\nu} C_{\mu\sigma}^{m\mathbf{k}\alpha,\mathrm{FV}^*} C_{\nu\sigma'}^{n\mathbf{k}\alpha,\mathrm{FV}} \int_{S_\alpha} \phi_{\mu\mathbf{k}}^{\alpha*}\,(B_{\mathrm{eff}}^\alpha)^i\,\phi_{\nu\mathbf{k}}\,d\mathbf{r} \\ &= \sum_{\mu\nu} C_{\mu\sigma}^{m\mathbf{k}\alpha,\mathrm{FV}^*} C_{\nu\sigma'}^{n\mathbf{k}\alpha,\mathrm{FV}} \int_{S_\alpha} \phi_{\mu\mathbf{k}}^{\alpha*}\left[\left(B_{\mathrm{eff}}^{\alpha,\mathrm{mult}}\right)^i + \left(B_{\mathrm{xc}}^{\alpha,\mathrm{nonmult}}\right)^i\right]\phi_{\nu\mathbf{k}}\,d\mathbf{r} \\ &= \sum_{\mu\nu} C_{\mu\sigma}^{m\mathbf{k}\alpha,\mathrm{FV}^*} C_{\nu\sigma'}^{n\mathbf{k}\alpha,\mathrm{FV}} \left[\left(B_{\mathrm{eff}}^{\alpha,\mathrm{mult}}\right)_{\mu\nu}^i + \left(B_{\mathrm{xc}}^{\alpha,\mathrm{nonmult}}\right)_{\mu\nu}^i\right],\end{aligned} \tag{3.94}$$

where the basis functions $\phi_{\mu\mathbf{k}}$ are defined as in Eq. (3.58) to account for both the (L)APW basis and local orbitals. The matrix elements $\left(B_{\mathrm{eff}}^{\alpha,\mathrm{mult}}\right)_{\mu\nu}^i$ and $\left(B_{\mathrm{xc}}^{\alpha,\mathrm{nonmult}}\right)_{\mu\nu}^i$ are obtained by following Eq. (A.45) and Eq. (3.61), respectively.

For the interstitial region, the same treatment is done, with the integral given by

$$\begin{aligned}
\left(B_{\text{eff}}\right)^i_{mn,\sigma\sigma'} &= \langle \psi^{\text{FV}}_{m\mathbf{k},\sigma} | \left(B_{\text{xc}}\right)^i | \psi^{\text{FV}}_{n\mathbf{k},\sigma'} \rangle \\
&= \sum_{\mathbf{G}_1\mathbf{G}_2} C^{m\mathbf{k},\text{FV}^*}_{\mathbf{G}_1\sigma} C^{n\mathbf{k},\text{FV}}_{\mathbf{G}_2\sigma'} \frac{1}{\Omega} \int_I \phi^*_{\mathbf{G}_1+\mathbf{k}} \left(B_{\text{xc}}\right)^i \phi_{\mathbf{G}_2+\mathbf{k}} d\mathbf{r} \\
&= \sum_{\mathbf{G}_1\mathbf{G}_2} C^{m\mathbf{k},\text{FV}^*}_{\mathbf{G}_1\sigma} C^{n\mathbf{k},\text{FV}}_{\mathbf{G}_2\sigma'} \frac{1}{\Omega} \int_I d\mathbf{r} \phi^*_{\mathbf{G}_1+\mathbf{k}} \left[ \left(B^{\text{mult}}_{\text{eff}}\right)^i + \left(B^{\text{nonmult}}_{\text{xc}}\right)^i \right] \phi_{\mathbf{G}_2+\mathbf{k}} \\
&= \sum_{\mathbf{G}_1\mathbf{G}_2} C^{m\mathbf{k},\text{FV}^*}_{\mathbf{G}_1\sigma} C^{n\mathbf{k},\text{FV}}_{\mathbf{G}_2\sigma'} \left[ \left(B^{\text{mult}}_{\text{eff}}\right)^i_{\mathbf{G}_1\mathbf{G}_2} + \left(B^{\text{nonmult}}_{\text{xc}}\right)^i_{\mathbf{G}_1\mathbf{G}_2} \right].
\end{aligned} \tag{3.95}$$

The interstitial matrix elements are similarly obtained by following Eq. (A.48) for the multiplicative and Eq. (3.66) for the non-multiplicative part. With the integrals evaluated for both the muffin-tin and interstitial regions and for all components $i$, the total effective magnetic field matrix $\bar{B}_{\text{eff}}(\mathbf{r})$ can be set up as

$$\bar{B}_{\text{eff}}(\mathbf{r}) = \frac{1}{2} \begin{pmatrix} B^z_{\text{eff}}(\mathbf{r}) & B^x_{\text{eff}}(\mathbf{r}) - iB^y_{\text{eff}}(\mathbf{r}) \\ B^x_{\text{eff}}(\mathbf{r}) + iB^y_{\text{eff}}(\mathbf{r}) & -B^z_{\text{eff}}(\mathbf{r}) \end{pmatrix}, \tag{3.96}$$

to complete the total setup of the Hamiltonian. The Hamiltonian is then diagonalized to obtain the full solution.

## 3.4. Total Energy

The total energy is given by [114]:

$$E[n] = T_s[n] + U[n] + E_{\text{xc}}[n], \tag{3.97}$$

where $T_s[n]$ is the kinetic energy of the non-interacting electrons defined as [22]

$$T_s[n] = \sum_{i\mathbf{k}} f_{i\mathbf{k}} \varepsilon_{i\mathbf{k}} - \tilde{E}_{\text{eff}}, \tag{3.98}$$

where the sum $i$ runs over both valence and core electrons. The effective energy $\tilde{E}_{\text{eff}}$ is made up of integrals over the Coulomb potential $V_{\text{C}}(\mathbf{r})$, the xc potential $V_{\text{xc}}(\mathbf{r})$, the xc magnetic field $\mathbf{B}_{\text{xc}}(\mathbf{r})$ and the external magnetic field $\mathbf{B}_{\text{ext}}(\mathbf{r})$:

$$\tilde{E}_{\text{eff}} = \int_\Omega V_{\text{C}}(\mathbf{r})\, n(\mathbf{r})\, d\mathbf{r} + \int_\Omega V_{\text{xc}}(\mathbf{r})\, n(\mathbf{r})\, d\mathbf{r} + \int_\Omega \left[\mathbf{B}_{\text{xc}}(\mathbf{r}) + \mathbf{B}_{\text{ext}}(\mathbf{r})\right] \cdot \mathbf{m}(\mathbf{r})\, d\mathbf{r}. \tag{3.99}$$

The interaction energy $U[n]$ consists of electron-electron, electron-nuclear, and nuclear-nuclear interactions [114]. It can be expressed in terms of the Coulomb and Madelung potentials as follows:

$$U[n] = \frac{1}{2} \int_\Omega V_{\text{C}}(\mathbf{r}) n(\mathbf{r})\, d\mathbf{r} - \frac{1}{2} \sum_\alpha Z_\alpha V^\alpha_{\mathbf{M}}(\mathbf{R}_\alpha). \tag{3.100}$$

In the case of the gKS scheme, the xc potential $V_{\mathrm{xc}}(\mathbf{r})$ can be separated into contributions from core and valence electrons, as they are treated by different potentials [7]:

$$\int_\Omega V_{\mathrm{xc}}(\mathbf{r})\, n(\mathbf{r})\, d\mathbf{r} = \int_\Omega V_{\mathrm{xc}}^{\mathrm{GGA}}(\mathbf{r})\, n^{\mathrm{core}}(\mathbf{r})\, d\mathbf{r} + \int_\Omega V_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r}. \tag{3.101}$$

The meta-GGA potential $V_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r})$ can then further be split up into its multiplicative and non-multiplicative part,

$$\int_\Omega V_{\mathrm{xc}}^{\mathrm{mGGA}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r} = \int_\Omega V_{\mathrm{xc}}^{\mathrm{mult}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r} + \int_\Omega V_{\mathrm{xc}}^{\text{non-mult}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r}. \tag{3.102}$$

The integral over the non-multiplicative potential $V_{\mathrm{xc}}^{\text{non-mult}}(\mathbf{r})$ can be expressed as:

$$\begin{aligned}\int_\Omega V_{\mathrm{xc}}^{\text{non-mult}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r} &= -\frac{1}{2}\sum_i \int_\Omega \nabla\cdot\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau}\nabla\psi_i\right)\psi_i^*\, d\mathbf{r} \\ &= \frac{1}{2}\sum_i \int_\Omega \left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau}\nabla\psi_i^*\cdot\nabla\psi_i\right) d\mathbf{r} - \frac{1}{2}\sum_i \int_{\partial S}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau}\psi_i^*\nabla\psi_i\right)\cdot d\mathbf{S},\end{aligned} \tag{3.103}$$

where, by applying Gauss' theorem, a surface term arises, as it is also the case for the Hamiltonian matrix elements. Similar to the treatment of the Hamiltonian matrix elements, the surface term is neglected in the present implementation. This results in the following final expression for the total xc potential energy for the valence electrons:

$$\int_\Omega V_{\mathrm{xc}}^{\text{non-mult}}(\mathbf{r})\, n^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r} = \int_\Omega \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau}\,\tau^{\mathrm{val}}(\mathbf{r})\, d\mathbf{r}. \tag{3.104}$$

In the case of spin-polarized systems, the xc magnetic field $\mathbf{B}_{\mathrm{xc}}$ needs to be evaluated additionally. Similar to the spin-unpolarized case, the xc magnetic field is decomposed into a multiplicative and a non-multiplicative part. The non-multiplicative part for the component $i$ can then be evaluated by considering the contribution of the KED and using the definition of the magnetization $m^i = \sum_j \psi_j^\dagger \sigma_i \psi_j$:

$$\int_\Omega B_{\mathrm{xc}}^{\text{non-mult},i}(\mathbf{r})\, m^i(\mathbf{r})\, d\mathbf{r} = -\sum_j \int_\Omega \frac{f_{\mathbf{u}}}{4}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\frac{u^i}{|\mathbf{u}|}\nabla\psi_j^\dagger\right]\sigma_i\,\psi_j\, d\mathbf{r}. \tag{3.105}$$

After partial integration, and with the definition of the spin KED, $u^i = \frac{1}{2}\sum_j \nabla\psi_j^\dagger\sigma_i\nabla\psi_j$, we obtain the final expression:

$$\begin{aligned}\int_\Omega B_{\mathrm{xc}}^{\text{non-mult},i}(\mathbf{r})\, m^i(\mathbf{r})\, d\mathbf{r} &= \sum_j \int_\Omega \frac{f_{\mathbf{u}}}{2}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\frac{u^i}{|\mathbf{u}|}\,\frac{1}{2}\nabla\psi_j^\dagger\sigma_i\nabla\psi_j\, d\mathbf{r} \\ &= \int_\Omega \frac{f_{\mathbf{u}}}{2}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\frac{u^i}{|\mathbf{u}|}\, u^i\, d\mathbf{r},\end{aligned} \tag{3.106}$$

where the surface term has also been neglected.

# 4. Validation and Results

Having successfully implemented the meta-GGA functionals, the next step is to validate the implementation. The following sections first examine the computational cost and time of a meta-GGA calculation, followed by an assessment of the correctness of the implementation through tests of the KED, and by calculating and comparing properties such as band gaps, lattice constants, and magnetic moments to theoretical and experimental literature results across different meta-GGA functionals and materials.

## 4.1. Computational Performance

In the following, the computational performance of the meta-GGA implementation is evaluated by analyzing the convergence behavior and the computational time for the SCAN, r2SCAN, TASK, and TPSS functionals across different materials. To this end, the number of SCF iterations required for spin-unpolarized convergence are evaluated and summarized in Fig. 4.1.

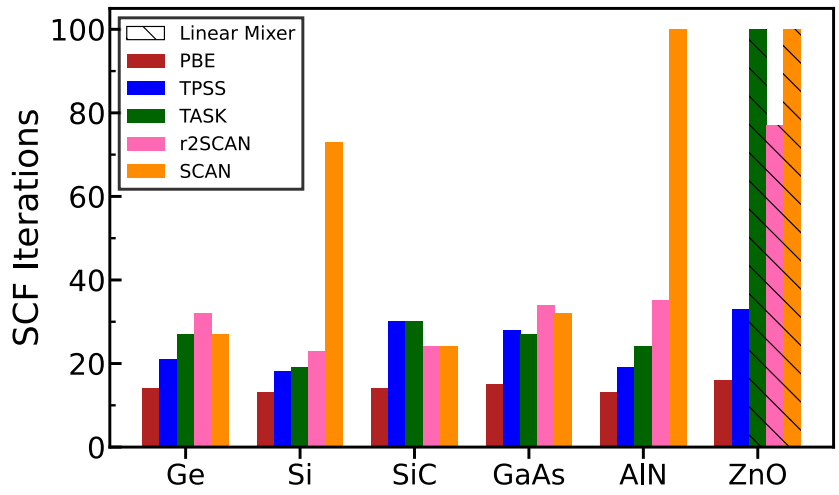
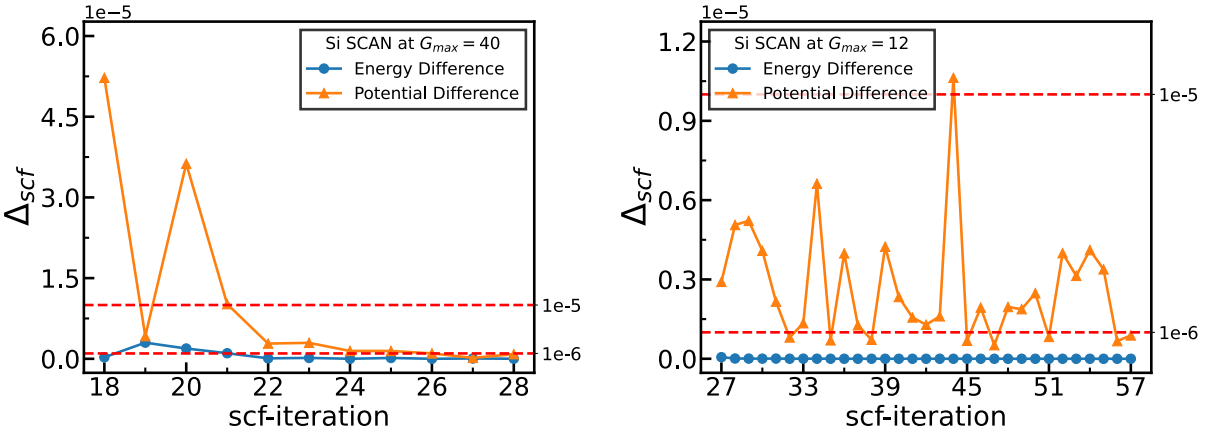



Figure 4.1.: Number of SCF iterations required for different materials with different functionals with $G_{\max} = 12\,\text{Bohr}^{-1}$ starting from atomic densities. The default convergence criteria are $\Delta_{\text{energy}} = 1 \times 10^{-6}$ Ha for the energy and $\Delta_{\text{pot}} = 1 \times 10^{-6}$ Ha for the potential. Values equal to 100 indicate that convergence was not achieved unless the potential convergence criterion was reduced to $\Delta_{\text{pot}} = 1 \times 10^{-5}$ Ha. Unless specified, the standard multi-secant Broyden mixer is used for potential mixing.

From the results, it is evident that meta-GGA functionals require more iterations than PBE. Among them, TPSS is generally the most stable, with only a moderate increase in iterations across all materials. As noted in Sec. 2.2, SCAN is prone to numerical instabilities, leading to unpredictable behavior and significantly more SCF iterations, which is reflected in the results, especially for Silicon (Si), Aluminium Nitride (AlN), and Zinc Oxide (ZnO). For ZnO with SCAN, r2SCAN, and TASK the linear mixer was

required to reach convergence. Additionally, SCAN and TASK only converged when the convergence criterion for the potential was relaxed. However, r2SCAN, designed to address SCAN's instabilities, achieved convergence with the default criterion for ZnO and all other cases. The cases of Si and ZnO will be discussed in more detail below.
For Si, to reach the default convergence criterion of $1 \times 10^{-6}$ Ha over two consecutive steps, a high cut-off of $G_{\max} = 40$ Bohr$^{-1}$ is needed, see Fig. 4.2a. This cut-off controls the Fourier coefficients used to construct the density, potential, and KED. A higher $G_{\max}$ provides a more accurate representation of the density and potential, capturing oscillations better, though at higher computational cost. Fig. 4.2b shows that a low $G_{\max}$ requires significantly more SCF iterations with SCAN and is not able to reach convergence systematically. However, with a relaxed convergence criterion of $1 \times 10^{-5}$ Ha for the potential, fewer iterations are needed, even with a $G_{\max}$ of 12 Bohr$^{-1}$.
In comparison, the PBE functional requires fewer iterations than for Si, converging in just 13 iterations. The TPSS and TASK functionals also show no major convergence issues, requiring 18 and 19 iterations, respectively. r2SCAN converges in 32 iterations, thus offering a more efficient convergence than SCAN, making it a favorable choice for Si.



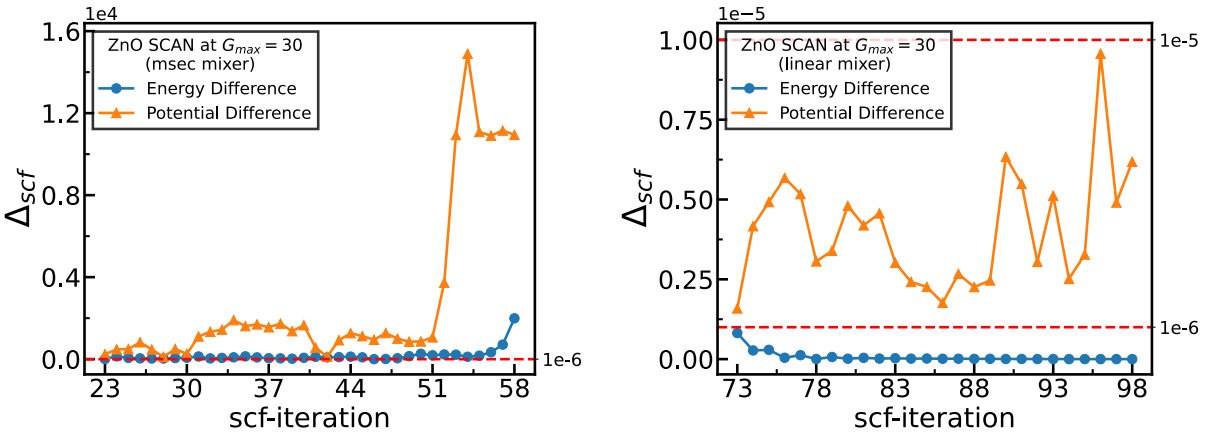

(a) Convergence behaviour with $G_{\max} = 40$ Bohr$^{-1}$.

(b) Convergence behaviour with $G_{\max} = 12$ Bohr$^{-1}$.

Figure 4.2.: Convergence behaviour of Si with the SCAN functional.

For ZnO, the convergence issues with SCAN are more pronounced. With $G_{\max} = 30$ Bohr$^{-1}$, the multi-secant Broyden mixer fails, as shown in Fig. 4.3a, where the potential diverges at higher SCF steps. Switching to the linear mixer, which provides slower but more stable convergence, helps mitigate this. However, even with higher $G_{\max}$ values, the SCAN potential fails to meet the $1 \times 10^{-6}$ Ha convergence criterion, as seen in Fig. 4.3b. Lowering the criterion to $1 \times 10^{-5}$ Ha for the potential helps to achieve convergence.
The divergence in ZnO might be attributed to several factors. In the beginning, the key reason could be its ionic nature, leading to a poor description of the atomic distribution in the first iterations. In such cases, it may be beneficial to first fully converge the calculation using PBE before switching to SCAN. While a linear mixer is still required in this case, we are able to observe a faster convergence with the SCAN functional when starting from converged PBE potential and density rather than atomic densities.
In comparison to SCAN, TPSS requires 33 iterations, while TASK, like SCAN, needs a relaxed convergence criterion and the linear mixer most likely due to its similarity to SCAN. The r2SCAN functional requires 77 iterations with the linear mixer when

starting from atomic densities without lowering the convergence criterion. Starting from a converged PBE calculation reduces the iterations needed to 52.
We conclude that the TPSS functional is the most stable among the meta-GGAs considered here. TASK and SCAN occasionally exhibit severe convergence problems with divergent potentials and it can be expected that with more difficult materials, such as 2D materials, where vacuum plays an important role, more issues with divergent potentials can occur. A clear alternative to SCAN is r2SCAN, which shows more consistent convergence in the cases studied here.




(a) Convergence behavior with $G_{\max} = 30$ Bohr$^{-1}$ using the multi-secant Broyden mixer.

(b) Convergence behavior with $G_{\max} = 30$ Bohr$^{-1}$ using the linear mixer.

Figure 4.3.: Convergence behaviour of ZnO with the SCAN functional.

The SCF iterations discussed here were based on calculations starting from atomic densities. It may be worth investigating whether meta-GGA calculations could benefit from fewer iterations when incorporating previous GGA iterations or starting from a pre-converged GGA solution. We find, however, that in all cases considered here, starting from a converged PBE calculation does not significantly reduce the required SCF iterations. Specifically, for Si and AlN with the SCAN functional, the extended convergence time is mainly due to the potential exhibiting too strong oscillations close to the convergence criterion. At this stage, it is reasonable to argue that PBE orbitals no longer have a significant influence on the potential. While starting from PBE orbitals can be more effective for ionic systems, such as ZnO, where it leads to faster convergence, in SCAN@PBE, the convergence threshold for the potential still needs to be lowered with ZnO. In contrast, r2SCAN@PBE does not require such an adjustment with ZnO.
To fully assess the computational cost, we further examine the time required per iteration. To this end, we evaluate the time per SCF iteration for four functionals for Si: PBE, and SCAN, TPSS, and TASK. To assess the time required for each iteration, we used the Intel VTune Profiler, performing each calculation 10 times to determine the mean value and minimize measurement errors. All calculations were run serially on a single core of an Intel(R) Core(TM) Processor (Ivybridge) at 3.2 GHz. The results are presented in Fig. 4.1.
The significant differences in time primarily stem from three factors: the evaluation of the kinetic energy density (KED), the potential (V), and the Hamiltonian setup and diagonalization (H). Clearly, the KED is not evaluated in PBE calculations, and the KED time does not vary significantly across meta-GGA functionals since it is independent of

the specific functional.
For the potential (V), the increased time in meta-GGA functionals arises from the need to evaluate both the meta-GGA and GGA potentials. Additionally, the $\tau$-dependent part, which accounts for the kinetic energy density, is evaluated alongside the standard GGA components in meta-GGA calculations. Here, the TASK functional shows a slightly lower evaluation time compared to SCAN and TPSS, likely due to its simpler LDA based correlation. Overall, we observe that the evaluation of the potential is approximately 2 to 3 times more expensive in meta-GGA calculations, consistent with findings in Ref. [7]. The Hamiltonian setup is another major factor. The inclusion of $\tau$-dependent integrals increases the computational cost for meta-GGA functionals. The cost arises from the need to evaluate gradients of the basis functions across three dimensions, as derived in Sec. 3.2.2, requiring three separate integrals per gradient component. The diagonalization step of the Hamiltonian should, in principle, remain consistent across all functionals, as its timing primarily depends on the matrix size, which is the same for all calculations.

| Method | V | H | KED | Iterations |
|---|---|---|---|---|
| PBE | 0.40 | 6.57 | 0.00 | 12 |
| SCAN | 1.07 | 10.06 | 0.93 | 25 |
| TPSS | 1.21 | 9.86 | 0.87 | 18 |
| TASK | 0.93 | 9.89 | 0.89 | 19 |

Table 4.1.: Time per iteration (in seconds) needed for the evaluation of the potential (V), Hamiltonian setup and diagonalization (H), and kinetic energy density (KED) for Si using different functionals. Additionally, the number of iterations required to converge calculations with a convergence criterion of $1 \times 10^{-5}$ Ha for the potential and $1 \times 10^{-6}$ Ha for the energy (starting from an atomic superposition) is shown. All computational settings are consistent across the different functionals.

The increase in computational time per iteration is mainly due to these factors, although additional elements contribute. These include the symmetrization of the KED, initialization of meta-GGA-related variables, and potential mixing, which involves both multiplicative and non-multiplicative potentials. Together, these factors make meta-GGA calculations approximately 1.5 times more expensive per iteration in the case of Si compared to GGA calculations. Moreover, meta-GGA calculations require more iterations to achieve convergence, which however depends on the material and functional considered. In summary, the materials discussed here take about 2-3 times longer to evaluate.
To compare the computational cost of hybrid functionals with meta-GGA calculations, we examine PBE0, HSE06, and the SCAN functional. Table 4.2 shows that hybrid functionals require significantly more computational time, with HSE06 being 170 times slower than SCAN and PBE0 around 80 times slower. Though PBE0 appears faster, it generally requires more **k**-points due to long-range exact exchange, whereas HSE06 requires fewer **k**-points for convergence because of its semi-local treatment. Furthermore, both hybrid functionals require the inclusion of extra unoccupied states for proper convergence, which further adds to their computational expense when performing convergence calculations. This makes meta-GGA functionals like SCAN a more efficient choice when balancing accuracy and computational resources.

| Method | Time (hh:mm:ss) |
|---|---|
| SCAN | 00:01:44 |
| HSE06 | 04:55:30 |
| PBE0 | 02:17:19 |

Table 4.2.: Mean computation time (hh:mm:ss) over 10 runs for the hybrid functionals HSE06 and PBE0, as well as the meta-GGA functional SCAN. All calculations were performed with identical settings, using OpenMP threading with 10 threads on a single node.

## 4.2. Kinetic Energy Density

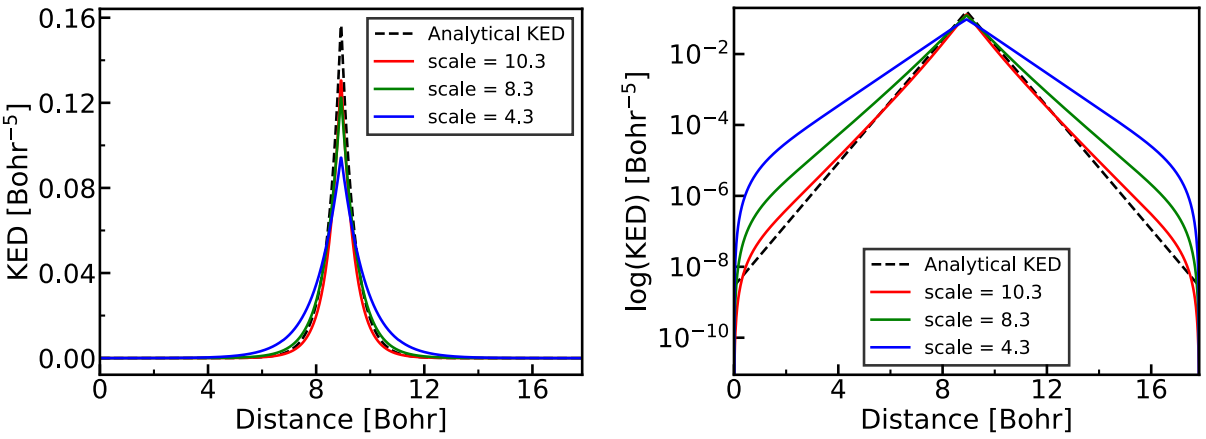
To verify the KED implementation from Sec. 4.2, we conduct two tests. First, we compute the KED for an isolated hydrogen atom and compare it to the analytical solution. The analytical solution for the hydrogen wave function is given by [115]:

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}, \tag{4.1}$$

from which the KED can be derived as:

$$\tau(r) = \frac{1}{2\pi a_0^3} e^{-2r/a_0}, \tag{4.2}$$

where $r$ is the radial distance from the nucleus, and $a_0$ is the Bohr radius, which is set to 1 in atomic units. Fig. 4.4 compares the analytical KED solution with the numerical results for different unit cell volumes. The unit cell volume is increased by adjusting the `scale` parameter to approximate the isolated atom. As can be seen, increasing `scale` improves the agreement, showing that the numerical approach converges to the analytical result. It should be noted, however, that the KED for the hydrogen atom only captures




(a) KED plotted on a linear scale.

(b) KED plotted on a logarithmic scale.

Figure 4.4.: KED of the hydrogen atom for different unit cell sizes, scaled by the factors of `scale` = 3.3, 8.3, and 10.3, to model the isolated atom. The numerical results are compared to the analytical solution.

the 1s state, which is a valence state. Consequently, this analysis does not provide any

insight into the implementation of the core states.
As defined in Sec. 4.2, the KED integrates to yield the total non-interacting kinetic energy:

$$T_s = \int_\Omega d\mathbf{r}\, \tau(\mathbf{r}). \tag{4.3}$$

Thus, as a second test, the values of the integrated KED can be compared to the values of the kinetic energy obtained by the eigenvalues $\varepsilon_{i\mathbf{k}}$ of the KS equation, as defined in Eq. (3.98). However, this test is also limited in that any integral over spherical harmonics captures only the $\ell = 0, m = 0$ component, since it is the only one that does not vanish. Specifically, the integral in the muffin-tin region

$$T_s = \sum_\alpha \int_{\mathrm{MT}_\alpha} d\mathbf{r} \sum_{\ell m} \tau^\alpha_{\ell m}(r) R_{\ell m}(\hat{\mathbf{r}}) = \sum_\alpha 4\pi \int_0^{R_{\mathrm{MT}_\alpha}} dr\, r^2\, \tau^\alpha_{00}(r), \tag{4.4}$$

shows that only the spherical component of the MT implementation is tested, since $\int d\Omega\, R_{\ell m}$ equals $\sqrt{4\pi}$ for $\ell = 0, m = 0$ and vanishes for all other $\ell$ and $m$.
Furthermore, the KED for core electrons is evaluated non-relativistically, so eigenvalues from the relativistic Dirac equation cannot be directly compared with the integrated core KED. To make a valid comparison, we focus on the valence contributions, comparing the valence KED and $T_{s,\mathrm{val}}$. Additionally, to ensure the KED integrates to the correct non-interacting kinetic energy, relativistic effects in the Hamiltonian must be disabled when evaluating the valence electrons, as the implemented KED is non-relativistic.

| Material | $\Delta_{\mathrm{rel,\ KE}}$ |
|---|---|
| He | $1.33 \times 10^{-12}$ |
| Li(s) | $2.73 \times 10^{-9}$ |
| C | $1.91 \times 10^{-13}$ |
| Si | $6.94 \times 10^{-13}$ |
| BN(s) | $4.48 \times 10^{-8}$ |
| LiF | $9.15 \times 10^{-12}$ |
| GaAs | $8.37 \times 10^{-12}$ |

Table 4.3.: Relative differences $\Delta_{\mathrm{rel,\ KE}}$ of the kinetic energy $T_{s,\mathrm{val}}$ and integrated non-relativistic kinetic energy density $\int \tau^{\mathrm{val}}$ for valence states for different materials. For Li and BN, the spin-polarized KED is evaluated.

The relative differences $\Delta_{\mathrm{rel,\ KE}}$ across different materials, shown in Table 4.3, are consistently small, indicating good agreement between $T_{s,\mathrm{val}}$ and the integrated KED. In spin-polarized cases, the differences are slightly larger, likely due to the complexity of the second-variational scheme. Nevertheless, the small differences observed in both the spin-polarized and spin-unpolarized cases validate the accuracy of the implemented KED.

## 4.3. Band Gaps

To validate the correctness of the implementation of meta-GGAs in `exciting`, this section focuses on converging and computing the band gaps of several semiconductors and on comparing those results with other theoretical band gaps found in literature, as well as experimental band gaps.
The implementation in `exciting` is tested using a small subset of semiconductors from Ref. [7], where the implementation of meta-GGAs in `WIEN2k` [24] was validated against results obtained with the `VASP` code [104]. The dataset for their study originates from Ref. [18], where calculations using different functionals across various rungs of the Jacob's ladder were compared using the `VASP` code. This database was later extended in Ref. [21] by adding more results obtained with meta-GGA functionals. The dataset consists of 472 materials, each with defined experimental band gaps and crystal structures, available in the supplementary information of Ref. [18]. For consistent comparison, the same experimental geometries are used for the smaller subset analyzed in this work, listed in Table 4.4, along with the expected points of the Conduction Band Minimum (CBM) and Valence Band Maximum (VBM) for the band gap and the experimental band gap. As discussed in Sec. 2.2.5, the band gap in the gKS scheme is obtained by taking the

| Solid | Structure type | $a_0$ [Å] | $c_0$ [Å] | CBM | VBM | exp. gap [eV] |
|---|---|---|---|---|---|---|
| Si | A4 (diamond) | 5.44 | - | X′ | Γ | 1.17 |
| Ge | A4 (diamond) | 5.67 | - | L | Γ | 0.74 |
| SiC | B3 (zincblende) | 4.35 | - | X | Γ | 2.42 |
| GaAs | B3 (zincblende) | 5.75 | - | Γ | Γ | 1.52 |
| AlN | B4 (wurtzite) | 3.13 | 5.02 | Γ | Γ | 6.19 |
| ZnO | B4 (wurtzite) | 3.24 | 5.22 | Γ | Γ | 3.44 |

Table 4.4.: Summary of materials, their structure type, experimental lattice constants, and experimental band gaps. The lattice constants $a_0$ and $c_0$ are given in Å, where $c_0$ is provided for non-cubic materials. CBM denotes the Conduction Band Minimum and VBM denotes the Valence Band Maximum. Note: X′ denotes a point slightly off X along the $\Gamma - \mathrm{X}$ path. The experimental band gaps are taken from the supplementary information of [18].

difference between the eigenvalues at the CBM point, or Lowest Unoccupied Molecular Orbital (LUMO), and the VBM point, or Highest Occupied Molecular Orbital (HOMO),

$$E_{\mathrm{gap}}^{\mathrm{gKS}} = \varepsilon_{\mathrm{CBM}}^{\mathrm{gKS}} - \varepsilon_{\mathrm{VBM}}^{\mathrm{gKS}}. \tag{4.5}$$

The calculated band gaps with `exciting` for the materials listed in Table 4.4 are presented in Table 4.5. The SCAN, TPSS, and TASK functionals were used for each material. These results are compared to those from the `VASP` code, which uses a plane wave basis with pseudopotentials or the Projector Augmented-Wave (PAW) method, and the `WIEN2k` code, also an all-electron APW+LO code. When comparing results obtained with `exciting` to those from other codes, discrepancies in the results can occur. Although `WIEN2k` is also an (L)APW code like `exciting`, and good agreement is generally expected, differences in parameter settings, such as the local orbital configuration and muffin-tin radii, can lead to small variations in the results. Despite these potential differences, Table 4.5 shows that the band gaps obtained with `exciting` are comparable to those obtained with `WIEN2k`

| Solid | SCAN | $\Delta_{\texttt{WIEN2k}}$ | $\Delta_{\texttt{VASP}}$ | TPSS | $\Delta_{\texttt{WIEN2k}}$ | $\Delta_{\texttt{VASP}}$ | TASK | $\Delta_{\texttt{WIEN2k}}$ | $\Delta_{\texttt{VASP}}$ |
|---|---|---|---|---|---|---|---|---|---|
| Si | 0.84 | 0.01 | −0.03 | 0.66 | 0.00 | −0.04 | 1.01 | 0.00 | −0.01 |
| Ge | 0.18 | 0.00 | 0.01 | 0.19 | 0.01 | 0.05 | 0.86 | −0.02 | −0.01 |
| SiC | 1.70 | 0.02 | −0.01 | 1.32 | −0.06 | 0.00 | 1.99 | 0.05 | 0.02 |
| GaAs | 0.80 | 0.03 | 0.00 | 0.68 | 0.00 | 0.00 | 1.73 | 0.00 | 0.06 |
| AlN | 4.81 | 0.02 | 0.02 | 4.15 | 0.01 | 0.01 | 5.84 | 0.07 | −0.01 |
| ZnO | 1.16 | −0.03 | 0.02 | 0.78 | 0.02 | 0.05 | 2.25 | −0.01 | 0.17 |

Table 4.5.: Band gaps (in eV) calculated with different meta-GGA functionals using the `exciting` code at experimental lattice parameters. The results are compared to those obtained with the `WIEN2k` and `VASP` codes. The differences between the `exciting` code and the respective codes ($\Delta_{\texttt{WIEN2k}} = E_g^{\texttt{exciting}} - E_g^{\texttt{WIEN2k}}$ and $\Delta_{\texttt{VASP}} = E_g^{\texttt{exciting}} - E_g^{\texttt{VASP}}$) are shown. Band gap data is taken from Refs. [7, 18, 21].

and `VASP`, with almost all differences well below 0.1 eV. The largest discrepancy is found for the TASK functional with ZnO, where a difference of about 0.17 eV compared to the `VASP` code is observed.
Additionally, as discussed in Sec. 4.1, the r2SCAN functional improves the convergence behavior of the SCAN functional, which is why we also evaluate the band gaps obtained with r2SCAN, which are presented in Table 4.6 and compare them to the SCAN functional. We observe that in most cases the differences between the results of the SCAN and r2SCAN functionals are within 0.1 eV, except for Germanium (Ge) and Gallium Arsenide (GaAs), where a difference of 0.3 eV and 0.24 eV is observed. In Ref. [59], it was reported that the typical difference between the SCAN and r2SCAN functionals was found to be around ±0.15 eV for band gaps, which is consistent with the mean difference of approximately 0.13 eV observed here.

| Solid | SCAN | r2SCAN |
|---|---|---|
| Si | 0.84 | 0.76 |
| Ge | 0.18 | 0.48 |
| SiC | 1.70 | 1.72 |
| GaAs | 0.80 | 1.04 |
| AlN | 4.80 | 4.82 |
| ZnO | 1.16 | 1.25 |

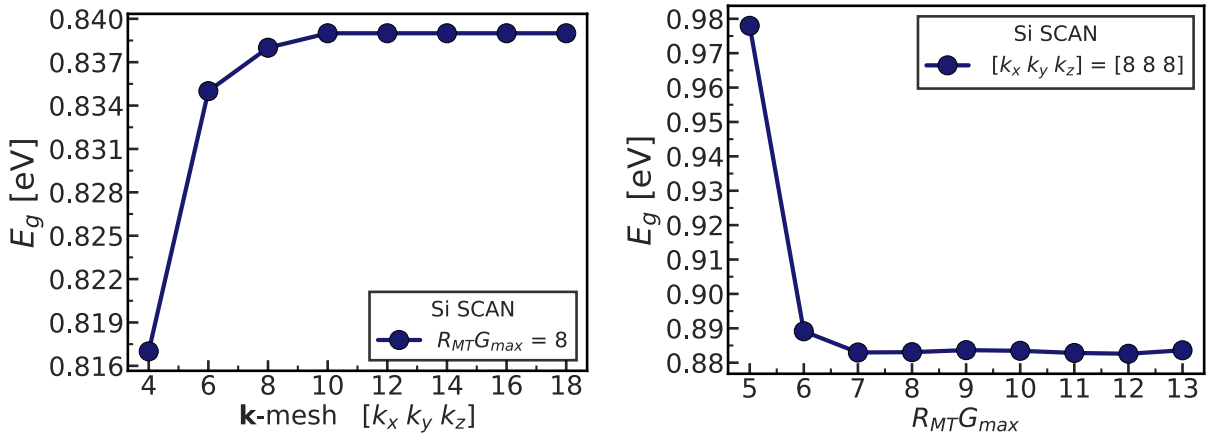
Table 4.6.: Band gaps (in eV) obtained with SCAN and r2SCAN using `exciting`.

To obtain reliable band gap values, it is essential to always converge the size of the interstitial basis, $R_{\mathrm{MT}}G_{\max}$, as defined in Sec. 2.3, as well as the $\mathbf{k}$-grid. This is because, for the calculation of quantities like the density, it is necessary to integrate over the Brillouin Zone (BZ). In practice, these integrals are evaluated by summing over a finite number of $\mathbf{k}$-points [23]:

$$\frac{1}{\Omega_{\mathrm{BZ}}} \int_{\mathrm{BZ}} d\mathbf{k} \to \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} .$$

However, replacing an integral with a sum requires convergence of the number of $\mathbf{k}$-points $[k_x, k_y, k_z]$ to ensure an accurate representation of the calculated properties. For

the semiconductors considered here, coarser meshes are typically sufficient because of their localized states and gap around their Fermi level. In this study, the band gap for all materials is converged to within at least 0.01 eV for both the **k**-grid and $R_{\mathrm{MT}}G_{\mathrm{max}}$. Additionally, setting up a good basis within the muffin-tin region is crucial. As discussed in Sec. 2.3.1, this requires the addition of several local orbitals for each atom $\alpha$, to achieve an accurate description of the wave function in the muffin-tin region.
In the following sections, the convergence behavior with respect to the parameters $R_{\mathrm{MT}}G_{\mathrm{max}}$ and the **k**-mesh of a few materials will be further discussed. The respective convergence plots for all materials can be found in Figs. B.1 - B.6.
The first material discussed is Si, a semiconductor with an experimental indirect band gap of 1.17 eV along the $\Gamma - \mathrm{X}$ points. Since the CBM of Si is not directly at a high symmetry point, a higher density of **k**-points is required to capture the indirect band gap accurately when evaluating the CBM directly from the **k**-mesh. This is because there can be variations in the specific **k**-point that is evaluated for the CBM for different **k**-point densities, making convergence analysis unstable. To address this, it is either useful to examine the band gap at consistent high symmetry points, such as the $\Gamma - \Gamma$ gap, which are always included in the **k**-mesh, or by sampling more eigenvalues along the $\Gamma - \mathrm{X}$ path by doing a band structure calculation. This provides a reliable convergence study for eigenvalues, as shown in Fig. 4.5a, accurately demonstrating how the **k**-point density is converged. As shown in Fig. 4.5b, a smooth convergence is equally achieved for Si using




(a) The band gap $E_g$ is evaluated from the band structure path along the $\Gamma - \mathrm{X}$ path, providing accurate convergence results.

(b) Convergence of the indirect Silicon band gap along $\Gamma - X$ with increasing basis set size $R_{\mathrm{MT}}G_{\mathrm{max}}$ at a constant **k**-mesh of [8 8 8].

Figure 4.5.: Convergence behavior of the Si band gap w.r.t. the **k**-mesh and $R_{\mathrm{MT}}G_{\mathrm{max}}$

the SCAN functional at a constant **k**-mesh, indicating that the band gap value converges as $R_{\mathrm{MT}}G_{\mathrm{max}}$ increases. This behavior is consistent across the other functionals, TPSS and TASK, see Fig. B.1.
The converged results for Si are close to either the `WIEN2k` or `VASP` results for all three functionals, as can be seen in Table 4.5. The differences in the band gaps between the functionals are also consistent with expectations. The TASK functional yields the largest band gap for Si, while the TPSS functional results in the smallest band gap, very close to PBE values found in literature of 0.58 eV [7], confirming that TPSS does not significantly improve the underestimation of the band gap that is inherent to the PBE functional.
For Ge, we similarly observe that different functionals lead to significantly varied band

gaps. Ge is characterized by a small indirect band gap at the $\Gamma - \mathrm{L}$ point, with an experimental value of approximately 0.74 eV. It also has a not much larger direct gap at the $\Gamma - \Gamma$ point of approximately 0.9 eV [116]. PBE generally underestimates the band gap strongly with values around $0.06 - 0.11$ eV [7, 18]. In contrast, the meta-GGAs evaluated in this study show an improved performance in predicting the band gap. The TASK functional, in particular, predicts a gap of 0.86 eV and r2SCAN a value of 0.48 eV. The SCAN and TPSS functionals, however, predict much smaller gaps of 0.18 eV and 0.19 eV, respectively.

Further analysis reveals that for the SCAN functional, convergence with respect to the basis size $R_{\mathrm{MT}}G_{\max}$ at smaller basis set sizes, $R_{\mathrm{MT}}G_{\max} = 6, 7$, the calculation incorrectly predicts a small non-zero Density of States (DOS) at the Fermi level, which implies the material is metallic. This issue is also seen when using the TPSS functional. To address this, the broadening parameter $\sigma$, linked to the artificial temperature $k_B T$, is adjusted. A small $\sigma$ smooths the occupancy of electronic states near the Fermi level by replacing the sharp Dirac delta function with a Gaussian function. This approach is beneficial for metals, but for materials like Ge, with small band gaps, refining $\sigma$ narrows the Gaussian, improving accuracy around the Fermi level. By refining $\sigma$ and increasing **k**-point density, a zero DOS at the Fermi level is consistently achieved, leading to more reliable band gap predictions (see Fig. B.3d).

Furthermore, when analyzing the band structure of Ge using the SCAN, r2SCAN, TPSS, and TASK functionals in Fig. 4.6, we observe that only the r2SCAN and TASK functionals correctly predict the order of the gaps around the Fermi level. Specifically, both functionals accurately identify the $\Gamma - \mathrm{L}$ gap as smaller than the $\Gamma - \Gamma$ gap, which aligns with the known properties of Ge [116]. In contrast, both the TPSS and SCAN functionals incorrectly predict the smallest gap at the $\Gamma - \Gamma$ point. SCAN and TPSS predict a $\Gamma - \Gamma$ gap of 0.18 eV and 0.19 eV, respectively, while the $\Gamma - \mathrm{L}$ gap is larger, with TPSS predicting 0.26 eV and SCAN predicting 0.31 eV. The r2SCAN functional improves the predictions by providing a more accurate gap hierarchy, with a $\Gamma - \mathrm{L}$ gap of 0.478 eV, slightly smaller than the $\Gamma - \Gamma$ gap of 0.482 eV. Similarly, the TASK functional accurately predicts a $\Gamma - \mathrm{L}$ gap of 0.86 eV, which is smaller than the predicted $\Gamma - \Gamma$ gap of 0.91 eV. This demonstrates that both r2SCAN and TASK provide a more accurate representation of the band structure of Ge.

The results for Silicon Carbide (SiC) and GaAs show good agreement with those from other codes, with detailed convergence plots in Fig. B.2 and B.4. For AlN and ZnO, both wurtzite structures with direct band gaps at $\Gamma - \Gamma$, convergence requires a non-uniform **k**-mesh to account for the non-cubic symmetry. Consequently, in the convergence plots for ZnO and AlN, convergence is shown with respect to the total number of **k**-points ($k_x \cdot k_y \cdot k_z$) rather than an equal **k**-mesh, to better reflect the hexagonal structures. Both materials show smooth convergence with respect to **k**-points and $R_{\mathrm{MT}}G_{\max}$, as shown in Fig. B.5 and Fig. B.6. Overall, we observe no increase in the needed **k**-points or $R_{\mathrm{MT}}G_{\max}$ values to converge meta-GGA band gaps to 0.01 eV.

Finally, it should be noted that in our implementation, as in `WIEN2k`, a GGA potential is used to calculate the radial functions for both valence and core states. The authors of Ref. [7] compared band gaps using both the so-defined optimal GGA and PBE for the basis setup, as discussed in Sec. 3.2.3. They found small differences (typically $0.0 - 0.02$ eV) for the TASK, SCAN, and TPSS functionals when using PBE instead of the optimal GGA. Thus, it was concluded that PBE is also suitable for the basis setup, and consequently the calculations presented here were performed using the PBE functional.

To analyze the influence of the functional used for the basis, the calculations for all other

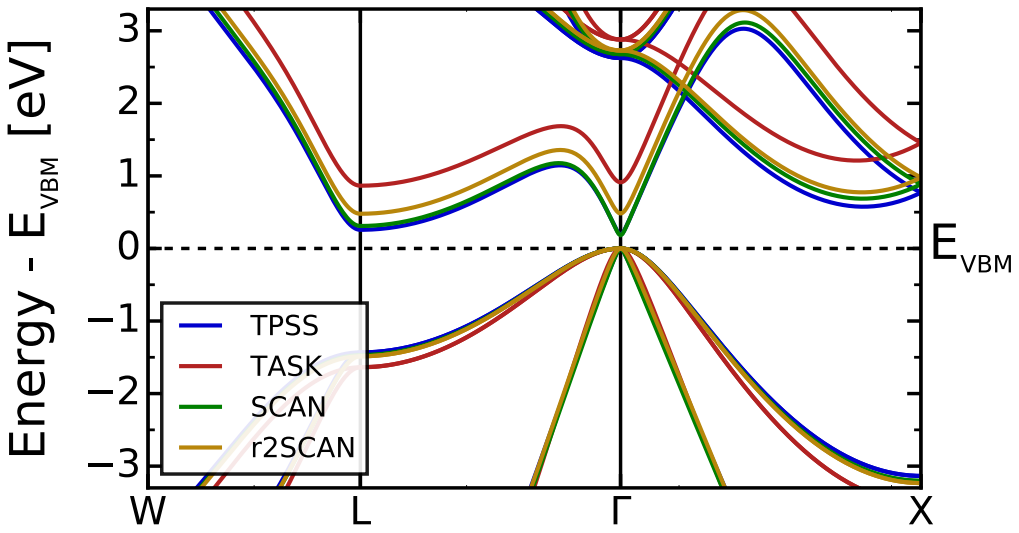



Figure 4.6.: Bandstructure of Ge for the SCAN, r2SCAN, TPSS and TASK functional along the $W - L - \Gamma - X$ path.

materials are repeated with `exciting` using the optimal functional for each respective meta-GGA functional. The differences are shown in Table 4.7. As can be seen, there is no significant discrepancy between the results when using the PBE and optimal functionals for the chosen materials. The largest differences between results using the PBE functional and the optimal GGA functional were found with the TASK functional. This is most likely due to the fact that the optimal GGA functional for SCAN and TPSS is the RPBE functional [117], a revised version of PBE, which explains the minimal differences observed. However, for the TASK functional, the optimal functional is the HCTH/407 functional [118], which resulted in more significant differences, such as a discrepancy of 0.22 eV for ZnO in Ref. [7].

| Solid | $\Delta_{\text{SCAN}}$ | $\Delta_{\text{TPSS}}$ | $\Delta_{\text{TASK}}$ |
|---|---|---|---|
| Si | 0.00 | 0.02 | 0.00 |
| Ge | 0.00 | 0.00 | 0.02 |
| SiC | 0.00 | 0.00 | 0.00 |
| GaAs | 0.01 | 0.00 | −0.01 |
| AlN | 0.00 | 0.00 | 0.00 |
| ZnO | 0.01 | 0.00 | 0.04 |

Table 4.7.: Differences in the meta-GGA band gaps (in eV) due to the chosen GGA potential (PBE or the optimal one) to evaluate the radial basis functions. The optimal potentials are RPBE for TPSS and SCAN, and HCTH/407 for TASK. The differences are given as $\Delta_{\text{Functional}} = E_g^{\text{optimalGGA}} - E_g^{\text{PBE}}$.

For ZnO, it is noteworthy that the band gap result from `VASP` (2.08 eV) is significantly smaller than the band gaps obtained by us using `exciting` (2.25 eV) and the `WIEN2k` code (2.26 eV) when using PBE as a basis for the radial functions. Using the optimal GGA functional HCTH/407 for ZnO in TASK calculations, we observe a discrepancy of 0.04 eV in the case of `exciting`, see Table 4.7. In the case of `WIEN2k`, the calculated value is 2.04 eV, resulting in a larger discrepancy of 0.22 eV when compared to the result

using PBE, though it now closely matches the result obtained with `VASP`. Although we did not reverify the `VASP` or `WIEN2k` results, comparing the theoretical results with the experimental band gap for ZnO, which is 3.44 eV, indicates that the larger band gaps obtained using `exciting` and `WIEN2k` are closer to the experimental value. Considering that no additional theoretical data on TASK with ZnO was found, and given that our results match those of `WIEN2k` and `VASP` in all other cases, it can be concluded that the implementation in `exciting` is valid and produces accurate band gaps for meta-GGA calculations.
To assess the performance of meta-GGAs in predicting band gaps, we compare the theoretical results with experimental data in Fig. 4.7, alongside values obtained with PBE and HSE06. As expected, TPSS consistently underestimates the band gaps, offering no improvement over PBE. In contrast, SCAN, r2SCAN, and TASK show better agreement with experimental data, with TASK even slightly overestimating the gaps in two cases and generally outperforming SCAN.

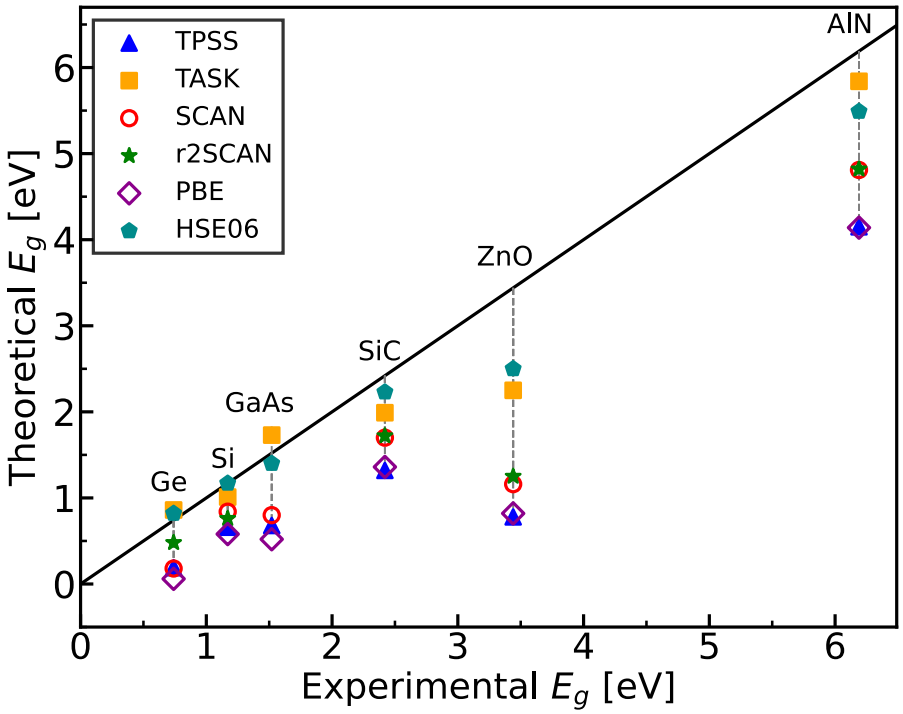



Figure 4.7.: Comparison of experimental band gaps (in eV) with theoretical values obtained using different functionals. The theoretical band gaps are computed using the meta-GGA functionals TPSS, r2SCAN, SCAN, and TASK with the `exciting` code. PBE data is taken from Ref. [7], HSE06 data from Ref. [119], and experimental band gaps from Ref. [18].

The comparison shows that meta-GGAs generally predict band gaps more accurately than PBE, but are less accurate than HSE06, which provides more accurate results for most materials. An exception is the TASK functional, which yields results closest to HSE06 and even outperforms it for the wide band gap material AlN.
As discussed in Sec. 4.1, HSE06 requires approximately 170 times more computational time, and PBE0 about 80 times more, compared to SCAN. While hybrid calculations are manageable for smaller systems, their computational cost scales quickly with system size [120]. Thus, when balancing the trade-off between accuracy and computational efficiency, TASK is found to be a viable option among the discussed functionals.

## 4.4. Lattice Constants

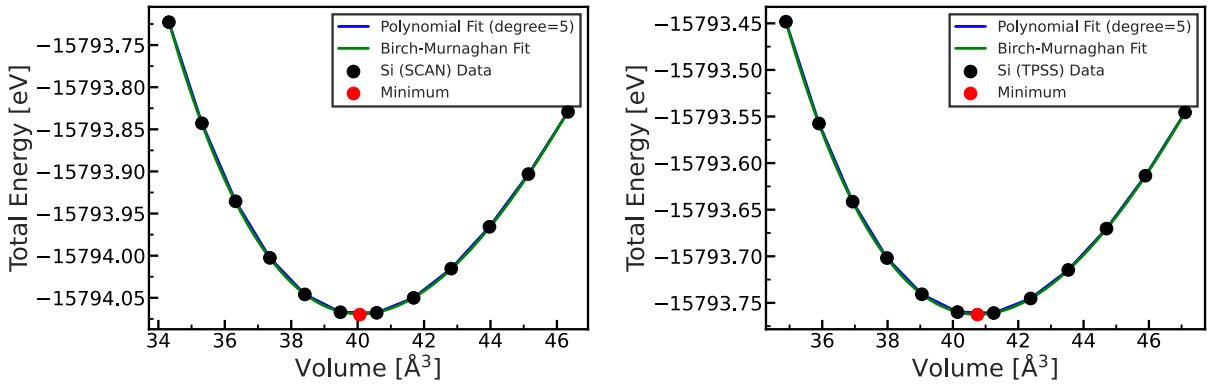
In this section, to further validate the implementation of meta-GGAs in `exciting`, specifically the total energy as defined in Sec. 3.4, equilibrium lattice constants $a_0$ are computed for Si, Ge, SiC, and GaAs. Equilibrium lattice constants can be evaluated either self-consistently or non-self-consistently. The latter involves using GGA or LDA orbitals to converge the Kohn-Sham SCF cycle, then evaluating the meta-GGA total energy across different volumes. In this work, lattice constants are determined self-consistently, using the converged $R_{\mathrm{MT}}G_{\max}$ and $\mathbf{k}$-mesh values from Sec. 4.3.
The equilibrium lattice constants are found by fitting the total energy $E(V)$ as a function of volume to both the Birch-Murnaghan Equation of State (EOS) [121] and a polynomial fit, as shown for Si in Fig. 4.8a. Both methods yield consistent lattice constants up to the third decimal place, with the Birch-Murnaghan EOS providing a more physically motivated fit. The Birch-Murnaghan EOS expresses the energy as a function of the volume:

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B_1 + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2 \left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right\},$$

where $V_0$ and $E_0$ are the equilibrium unit cell volume and energy, respectively, $B_0$ is the bulk modulus, and $B_1$ is the derivative of the bulk modulus with respect to pressure. The lattice constant can then be evaluated from the volume $V_0$ that minimizes the energy by taking into account the symmetry of the system. Given the cubic structure of the studied materials, only a single lattice constant $a_0$ needs to be computed for each material, which is given by:

$$a_0 = \sqrt[3]{N_{\mathrm{cubic}}V_0},$$

where $N_{\mathrm{cubic}}$ is the number of atoms in the cubic unit cell. For all materials evaluated here, $N_{\mathrm{cubic}} = 4$. The $E(V)$-fitted plots for Ge, SiC, and GaAs can be found in the Appendix in Fig. B.8.




(a) $E(V)$ curve fitted for Si using SCAN. (b) $E(V)$ curve fitted for Si using TPSS.

Figure 4.8.: Fitting of the $E(V)$ curves for Si using the SCAN and TPSS functionals. Fitted with the Birch-Murnaghan EOS and a polynomial fit of order 5.

We compare our results for the SCAN and TPSS functionals to those found in the literature, obtained with different methods and codes, as shown in Tables 4.8 and 4.9. Overall, the results obtained with `exciting` show good agreement with other codes such as `WIEN2k`, `VASP` and `FHI-Aims`, which utilizes atom-centered numerical basis sets [122];

| Solid | exciting | WIEN2k[1] | VASP[2] | FHI-Aims[3] |
|---|---|---|---|---|
| Si | 5.432 | 5.436 | 5.428 | 5.433 |
| Ge | 5.687 | - | 5.685 | 5.669 |
| SiC | 4.355 | - | 4.352 | 4.353 |
| GaAs | 5.658 | 5.658 | 5.665 | 5.656 |

Table 4.8.: Lattice constants (in Å) obtained using the SCAN functional with different computational codes.

[1] Ref. [7]
[2] Ref. [126]
[3] Ref. [127]

Gaussian, which employs Gaussian-Type Orbitals (GTOs) [123]; and BAND, which makes use of localized atomic orbitals [124, 125].
The results for the SCAN functional in Table 4.8 show that the lattice constants calculated using exciting for Si, SiC and GaAs are in close agreement with the values obtained using WIEN2k, VASP and FHI-Aims, with deviations well within 0.01 Å. For Ge, exciting is also in good agreement with VASP, but FHI-Aims predicts a lower value, showing a larger deviation of 0.018 Å from exciting.
For the TPSS functional, see Table 4.9, the lattice constants computed using exciting generally also show good agreement with the other codes, with most deviations also within 0.01 Å. In the cases of Si and SiC, Gaussian predicts slightly higher lattice constants, deviating by 0.015 Å and 0.02 Å, respectively, from exciting. For GaAs, we find that BAND predicts a slightly lower value, deviating by 0.02 Å. Nevertheless, exciting remains closely aligned with the other codes for these materials.

| Solid | exciting | WIEN2k[1] | VASP[2] | Gaussian[3] | BAND[4] |
|---|---|---|---|---|---|
| Si | 5.462 | 5.458 | 5.453 | 5.477 | 5.452 |
| Ge | 5.732 | - | 5.729 | 5.731 | 5.723 |
| SiC | 4.372 | - | 4.366 | 4.392 | 4.366 |
| GaAs | 5.722 | 5.711 | 5.718 | 5.716 | 5.702 |

Table 4.9.: Lattice constants (in Å) obtained using the TPSS functional with different computational codes.

[1] Ref. [7]
[2] Ref. [128]
[3] Ref. [129]
[4] Ref. [130]

To put the meta-GGA results from exciting into context, we compare them to theoretical literature values obtained using the PBE and HSE06 functionals, as well as Zero-Point Anharmonic Expansion (ZPAE)-corrected experimental lattice constants. Experimental lattice constants at zero temperature are influenced by zero-point vibration energies, which are absent in 0 K DFT calculations, affecting the absolute energy and EOS, as phonon frequencies decrease with volume [72]. Removing the ZPAE contribution from the experimental data allows for a more accurate comparison with theoretical lattice constants. To visualize the performance of each functional compared to the experimental data, we plot the deviations of the calculated lattice constants from the experimental

data in Fig. 4.9. As expected, PBE consistently overestimates the experimental lattice constants. The meta-GGA functionals show improvements over PBE, although TPSS, while more accurate, still strongly overestimates the values. As discussed in Sec. 2.2.3, SCAN stands out as one of the most accurate meta-GGA for predicting lattice constants of solids. This is evident in the results, where SCAN and HSE06 predict values closest to the experimental data. We also evaluated lattice constants using the more stable r2SCAN functional, which yields values similar to the SCAN functional but generally produces larger constants, consistent with findings in Ref. [62].
Among the meta-GGAs and materials discussed, SCAN is the most reliable functional. For SCF convergence issues, TPSS and r2SCAN serve as viable alternatives. Additionally, lattice constants can be calculated non-self-consistently, as shown in Ref. [7], where self-consistent and non-self-consistent results using GGA orbitals showed minimal differences for both the SCAN and TPSS functionals.
In summary, the observed deviations from other computational codes are minimal and do not impact the overall reliability of the results. When compared to experimental results, the `exciting` calculations reproduce the findings reported in the literature: TPSS tends to overestimate lattice constants, while SCAN provides more accurate values. This consistency across different functionals and materials further validates the correctness of the implementation of meta-GGAs in `exciting`.

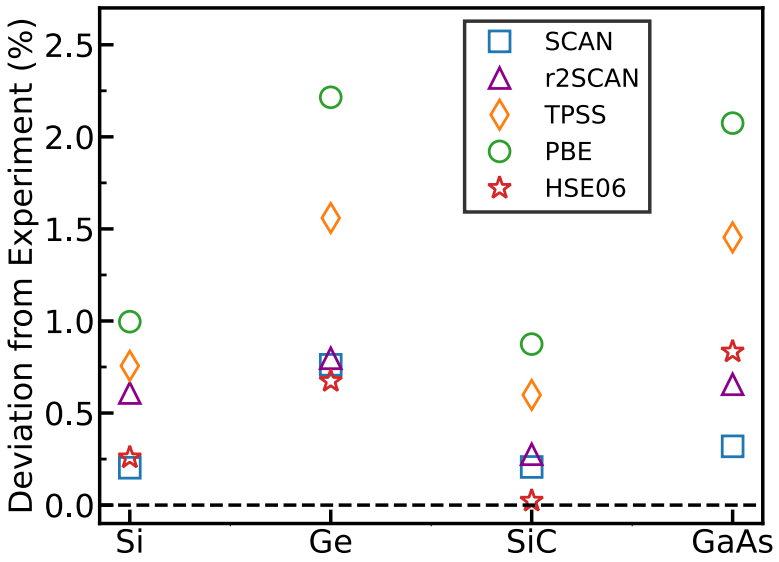



Figure 4.9.: Deviation (in %) of the calculated theoretical lattice constants $a_0$ from the ZPAE-corrected experimental values. Theoretical lattice constants are obtained using the TPSS, SCAN, and r2SCAN functionals with the `exciting` code. PBE data is taken from Ref. [53], and HSE06 and experimental data is from Ref. [72].

## 4.5. Magnetic Moments

To validate the spin-polarized implementation outlined in Sec. 3.3, the magnetic moments of the ferromagnetic solids Ni, Fe, and Co are calculated and compared to both `WIEN2k` results and experimental data. To ensure an accurate comparison with the `WIEN2k` results from Ref. [7], the calculations are performed at the same experimental geometries, which are summarized in Table 4.10.

| Material | Structure Type | $a_0$ | $c_0$ | exp. $\mu_s$ |
|---|---|---|---|---|
| Ni | A1 (fcc) | 3.523 | - | 0.52-0.55 |
| Fe | A2 (bcc) | 2.867 | - | 1.98-2.08 |
| Co | A3 (hcp) | 2.507 | 4.070 | 1.52-1.62 |

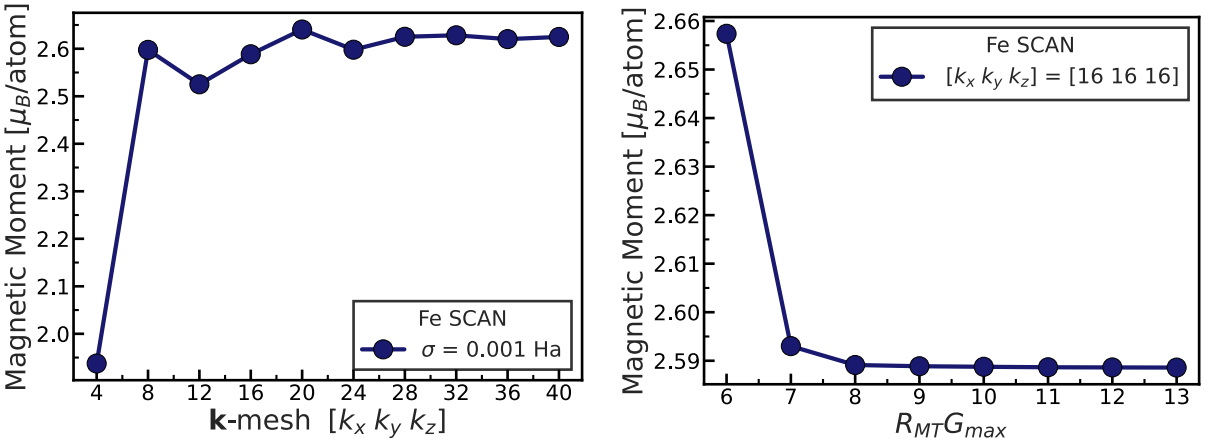
Table 4.10.: Summary of materials, their structure type, experimental lattice constants, and experimental magnetic moments $\mu_s$. The lattice constants $a_0$ and $c_0$ are given in Å, where $c_0$ is provided for non-cubic materials. The experimental magnetic moments are taken from Ref. [57] and are given in $\mu_B$/atom.

To converge magnetic moments, it is necessary to converge the **k**-mesh, $R_{\mathrm{MT}}G_{\max}$, and the number of states used in the second-variational basis, which is done with the `nempty` parameter in `exciting`. In principle, this parameter determines the eigenstates required beyond charge neutrality. The inclusion of these empty states is crucial in metals to accurately capture the Fermi level, where the number of states below the Fermi energy is not known beforehand. Furthermore, it determines the basis set size of the second-variational basis set in spin-polarized calculations. Then, the number of empty states depends on the total basis set size of the (L)APWs. Therefore, we first converge $R_{\mathrm{MT}}G_{\max}$, followed by the number of empty eigenstates, see Fig. 4.10b and 4.11.
In metals, it is crucial to accurately describe the DOS at the Fermi level due to partially occupied bands. This creates a discontinuity in the occupation number $f_{n\mathbf{k}}$, which can cause SCF instability if the **k**-point grid is too sparse. To address this, smearing methods are used, replacing the occupancy $f_{n\mathbf{k}}$ with a smooth function near the Fermi level. Common smearing functions include Gaussian and Fermi-Dirac distributions, with broadening controlled by $\sigma = k_B T$. A smaller $\sigma$ improves accuracy but requires a denser **k**-point mesh, while a larger $\sigma$ reduces the **k**-point density at the cost of accuracy.
Although a full convergence study of $\sigma$ is important for predicting magnetic moments, we did not perform an extensive convergence here, focusing instead on validating the implementation. However, such tests should ideally be done. For the materials considered in this study, we chose $\sigma$ values between 0.005 Ha and 0.001 Ha. For Ni and Co, larger **k**-point meshes were required for convergence, which is why we use $\sigma = 0.005$ Ha and $\sigma = 0.003$ Ha, respectively. For Fe, convergence is achieved more easily, so we use $\sigma = 0.001$ Ha, as shown in Fig. 4.10a. All other convergence plots here can be found in the Appendix B.3.
The converged results for magnetic moments converged within $0.01\,\mu_B$/atom for TPSS and SCAN are listed in Table 4.11. As can be seen, the results of the `WIEN2k` code and the `exciting` code are in good agreement.
For Co, the TPSS functional shows smooth convergence, and the resulting magnetic moment aligns well with the `WIEN2k` result. However, using the SCAN functional presents significant challenges, with unpredictable convergence behavior. A calculation may con-

| Material | SCAN | $\Delta_{\texttt{WIEN2k}}$ | TPSS | $\Delta_{\texttt{WIEN2k}}$ |
|---|---|---|---|---|
| Fe | 2.61 | 0.01 | 2.23 | 0.00 |
| Ni | 0.76 | −0.01 | 0.66 | 0.00 |
| Co | 1.76 | −0.01 | 1.65 | 0.00 |

Table 4.11.: Converged magnetic moments (in $\mu_B$/atom) for Ni, Fe, and Co using the SCAN and TPSS functionals obtained with the `exciting` code. The differences for the magnetic moments obtained with the `WIEN2k` code ($\Delta_{\texttt{WIEN2k}} = \mu_s^{\texttt{exciting}} - \mu_s^{\texttt{WIEN2k}}$) are also presented, with values taken from [7].

verge for a specific **k**-mesh, but fail with a finer or coarser mesh. This issue is not unique to Co and is also observed for Fe and Ni, suggesting it is inherent to SCAN. While convergence studies are possible, they are inconsistent, with some **k**-point configurations converging and others not. As a result, some convergence plots have missing data points where the convergence criteria were not met. In general, we find that increasing $G_{\max}$ and $\ell_{\max}$ to describe the potential and KED more accurately in both the interstitial and MT regions helps mitigate convergence issues.




(a) Convergence w.r.t. the **k**-mesh with Gaussian smearing with $\sigma = 0.001$ Ha.

(b) Convergence w.r.t. the basis size $R_{\mathrm{MT}}G_{\max}$ at constant **k**-mesh = [16 16 16].

Figure 4.10.: Convergence of the magnetic moment (in $\mu_B$/atom) of Fe SCAN.

To further improve the convergence issues encountered here, alternative mixing schemes could be used. For example, the authors of Ref. [131] have shown that convergence problems, especially in spin-polarized calculations, can be improved by using the Kerker mixer [132], which reduces convergence problems caused by charge sloshing [133]. The study also addressed convergence problems due to noise in the exchange-correlation potential by testing different mixers, such as the Pulay mixer [134, 135] and a modified version of it. They showed that these can improve convergence in the presence of oscillating potentials. Therefore, the convergence problems encountered here with the SCAN functional could be improved by using one of these mixers, and investigating these options could improve the reliability of meta-GGA spin-polarized calculations. Thus, before performing more extensive calculations with meta-GGAs in the spin-polarized case, it may be beneficial to first improve the convergence behavior.
To further put the results obtained with TPSS and SCAN in context by comparing them

with experimental values, we find that TPSS slightly overestimates the magnetic moments, while SCAN overestimates them even more. These findings are consistent with those reported in Ref. [7, 57], where it was shown that SCAN produces magnetic moments significantly higher than those obtained with LDA and PBE, both of which already overestimate magnetic moments for ferromagnetic elementary solids. TPSS yields magnetic moments similar to those obtained by PBE. Although the meta-GGAs do not provide satisfactory results for ferromagnetic materials, that does not mean that meta-GGAs in general are not applicable to magnetic systems. For example, for cuprate superconductors, SCAN is able to predict the most accurate magnetic properties compared to other DFAs, such as GGAs and hybrid functionals [136]. Thus, the study here mostly serves the validation of the spin-polarized calculation.

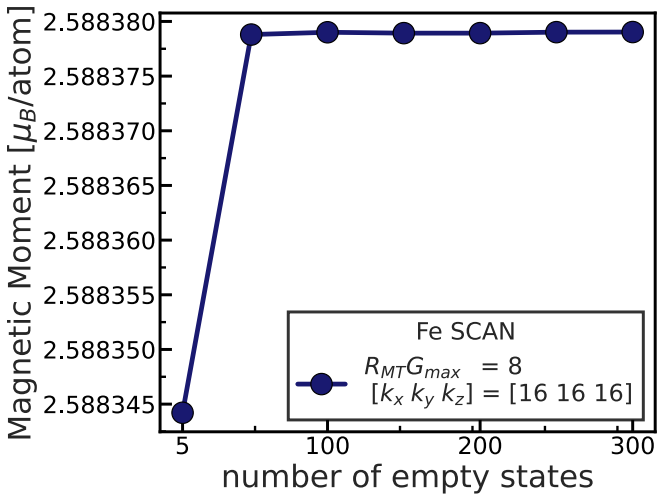



Figure 4.11.: Convergence of the magnetic moment (in $\mu_B$/atom) of Fe SCAN w.r.t. the empty states.

As a final note, in evaluating the magnetic moments of ferromagnetic materials, only the collinear part of the implementation has been tested. Specifically, we always apply an external field $\mathbf{B}_{\text{ext}}$ along the spin quantization axis $z$, which is gradually reduced during each SCF iteration to find the ground state. Thus, to fully test the implementation, further testing would be required for the general non-collinear case. However, since the implemented non-collinear formalism reduces to the collinear limit, and the magnetic moments are accurately predicted here, this suggests that the implementation is most likely correct.

# 5. Conclusion and Outlook

In this work, we have successfully extended the capabilities of the all-electron (L)APW+LO code `exciting` by implementing self-consistent meta-GGAs for both spin-unpolarized and spin-polarized systems. By integrating the `Libxc` library for the evaluation of the meta-GGAs potentials, we have further expanded the applicability of meta-GGA calculations in `exciting` by providing access to a wide range of KED-dependent meta-GGA functionals. The validity of the implementation was verified by comparisons with established computational codes for fundamental properties such as band gaps, lattice constants, and magnetic moments. The meta-GGA functionals considered in this work - SCAN, r2SCAN, TPSS, and TASK - consistently produced results in close agreement with other codes. This confirms the correctness of the approach and implementation of meta-GGAs within `exciting`.
Furthermore, we demonstrated that the meta-GGA implementation is computationally efficient. Our meta-GGA calculations require only about 2-3 times the computational effort of standard GGA functionals, consistent with other timing results found in the literature, making them a practical choice for achieving improved accuracy without significant additional cost. In contrast, hybrid functionals within the `exciting` code take at least 80 to 170 times longer than the meta-GGAs considered here.
Moreover, our comparative analysis between the different meta-GGA functionals provides guidance on selecting the most appropriate meta-GGA functional based on the property of interest - at least among the functionals and properties considered here. For band gap prediction, the TASK functional proved to be the best, approaching the accuracy of the HSE06 hybrid functional. This makes TASK particularly suitable for electronic structure calculations where an accurate band gap estimation is crucial. For lattice constants, both SCAN and r2SCAN showed good performance, closely matching experimental data and hybrid functional results, thus serving as reliable choices for structural predictions. For the magnetic moments of the ferromagnetic elementary solids Fe, Ni, and Co, none of the meta-GGA functionals considered (SCAN and TPSS) provided an improvement over GGA functionals such as PBE.
Finally, the usage of GGA potentials to obtain the radial functions for the valence and core electrons implies that we do not have a fully self-consistent meta-GGA implementation, since the core states are evaluated only through the GGA potential. However, we have shown that using PBE to set up the radial functions does not significantly affect our results, as all calculated properties remain in close agreement with those from other codes. Therefore, we conclude that it is acceptable to use the PBE functional for the setup of the radial functions.

While the successful implementation of meta-GGA functionals within the all-electron (L)APW+LO code `exciting` opens up numerous possibilities for future applications in `exciting`, there are also areas in the implementation that may require further work to overcome the current limitations of the implementation.
Most importantly, during our analysis, it became evident that meta-GGA calculations, particularly in spin-polarized systems, can suffer from convergence issues. Addressing these challenges could involve exploring alternative mixing schemes, such as the Kerker

mixer. Improving the convergence behavior would make meta-GGA calculations more reliable, broaden their applicability to a wider range of complex materials and systems in `exciting`, and enable a more high-throughput approach.
In addition, achieving a fully self-consistent all-electron meta-GGA implementation would include the accurate treatment of the core electrons with a meta-GGA functional as well. This would require the implementation of a (relativistic) radial solver capable of handling non-multiplicative potentials, as meta-GGA potentials are orbital-dependent and cannot be simply integrated into the existing framework for solving for the radial functions.
Finally, another technical aspect is the implementation of surface terms. Although these terms are expected to be negligible and small, to ensure a completely accurate treatment of the gradients of orbitals when using APW basis functions, they should, in principle, be implemented.
Despite these technical aspects that can still be improved, the implementation has already demonstrated accurate results, thereby making it suitable for immediate use and exploration of various applications with meta-GGA functionals.

# Bibliography


[1] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas", Phys. Rev. **136**, B864–B871 (1964).

[2] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects", Phys. Rev. **140**, A1133–A1138 (1965).

[3] J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy", AIP Conference Proceedings **577**, 1–20 (2001).

[4] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

[5] J. P. Perdew, S. Kurth, A. š. Zupan, and P. Blaha, "Accurate density functional with correct formal properties: a step beyond the generalized gradient approximation", Phys. Rev. Lett. **82**, 2544–2547 (1999).

[6] M. Betzinger, C. Friedrich, and S. Blügel, "Hybrid functionals within the all-electron flapw method: implementation and applications of pbe0", Phys. Rev. B **81**, 195117 (2010).

[7] J. Doumont, F. Tran, and P. Blaha, "Implementation of self-consistent mgga functionals in augmented plane wave based methods", Phys. Rev. B **105**, 195138 (2022).

[8] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

[9] J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional", Phys. Rev. Lett. **115**, 036402 (2015).

[10] E. B. Isaacs and C. Wolverton, "Performance of the strongly constrained and appropriately normed density functional for solid-state materials", Phys. Rev. Mater. **2**, 063801 (2018).

[11] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids", New Journal of Physics **20**, 063020 (2018).

[12] F. Tran, J. Stelzl, and P. Blaha, "Rungs 1 to 4 of DFT Jacob's ladder: Extensive test on the lattice constant, bulk modulus, and cohesive energy of solids", The Journal of Chemical Physics **144**, 204120 (2016).

[13] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, "Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional", Nature Chemistry **8**, 831–836 (2016).

[14] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, "Generalized kohn-sham schemes and the band-gap problem", Phys. Rev. B **53**, 3764–3774 (1996).

[15] T. Lebeda, T. Aschebrock, J. Sun, L. Leppert, and S. Kümmel, "Right band gaps for the right reason at low computational cost with a meta-gga", Phys. Rev. Mater. **7**, 093803 (2023).

[16] R. W. Godby, M. Schlüter, and L. J. Sham, "Trends in self-energy operators and their corresponding exchange-correlation potentials", Phys. Rev. B **36**, 6497–6500 (1987).

[17] F. Tran and P. Blaha, "Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential", Phys. Rev. Lett. **102**, 226401 (2009).

[18] P. Borlido, T. Aull, A. W. Huran, F. Tran, M. A. L. Marques, and S. Botti, "Large-scale benchmark of exchange–correlation functionals for the determination of electronic band gaps of solids", Journal of Chemical Theory and Computation **15**, PMID: 31306006, 5069–5079 (2019).

[19] T. Aschebrock and S. Kümmel, "Ultranonlocality and accurate band gaps from a meta-generalized gradient approximation", Phys. Rev. Res. **1**, 033082 (2019).

[20] F. Tran, J. Doumont, L. Kalantari, P. Blaha, T. Rauch, P. Borlido, S. Botti, M. A. L. Marques, A. Patra, S. Jana, and P. Samal, "Bandgap of two-dimensional materials: Thorough assessment of modern exchange–correlation functionals", The Journal of Chemical Physics **155**, 104103 (2021).

[21] P. Borlido, J. Schmidt, A. W. Huran, F. Tran, M. A. L. Marques, and S. Botti, "Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning", npj Computational Materials **6**, 96 (2020).

[22] A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner, and C. Draxl, "Exciting: a full-potential all-electron package implementing density-functional theory and many-body perturbation theory", Journal of Physics: Condensed Matter **26**, 363202 (2014).

[23] D. J. Singh and L. Nordström, *Planewaves, pseudopotentials, and the lapw method*, 2nd ed., eBook ISBN: 978-0-387-29684-5, Hardcover ISBN: 978-0-387-28780-5, Softcover ISBN: 978-1-4419-3954-8 (Springer New York, NY, Oct. 2006), pp. XIII, 134.

[24] P. Blaha, K. Schwarz, F. Tran, R. Laskowski, G. K. H. Madsen, and L. D. Marks, "WIEN2k: An APW+lo program for calculating the properties of solids", The Journal of Chemical Physics **152**, 074101 (2020).

[25] R. M. Martin, *Electronic structure: basic theory and practical methods*, 2nd ed. (Cambridge University Press, 2020).

[26] M. Born and W. Heisenberg, "Zur quantentheorie der molekeln", Annalen der Physik **389**, 457–484.

[27] W. Kohn, "Nobel lecture: electronic structure of matter—wave functions and density functionals", Rev. Mod. Phys. **71**, 1253–1266 (1999).

[28] U. von Barth and L. Hedin, "A local exchange-correlation potential for the spin polarized case. i", Journal of Physics C: Solid State Physics **5**, 1629 (1972).

[29] L. D. Marks and D. R. Luke, "Robust mixing for ab initio quantum mechanical calculations", Phys. Rev. B **78**, 075114 (2008).

[30] R. T. Sharp and G. K. Horton, "A variational approach to the unipotential many-electron problem", Phys. Rev. **90**, 317–317 (1953).

[31] J. D. Talman and W. F. Shadwick, "Optimized effective atomic central potential", Phys. Rev. A **14**, 36–40 (1976).

[32] R. Garrick, A. Natan, T. Gould, and L. Kronik, "Exact generalized kohn-sham theory for hybrid functionals", Phys. Rev. X **10**, 021040 (2020).

[33] Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, "More realistic band gaps from meta-generalized gradient approximations: only in a generalized kohn-sham scheme", Phys. Rev. B **93**, 205205 (2016).

[34] R. Baer and L. Kronik, "Time-dependent generalized kohn–sham theory", The European Physical Journal B **91**, 170 (2018).

[35] L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, "Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals", Journal of Chemical Theory and Computation **8**, PMID: 26593646, 1515–1531 (2012).

[36] C. A. Ullrich, "Density-functional theory for systems with noncollinear spin: orbital-dependent exchange-correlation functionals and their application to the hubbard dimer", Phys. Rev. B **98**, 035140 (2018).

[37] E. Sjöstedt and L. Nordström, "Noncollinear full-potential studies of $\gamma-$Fe", Phys. Rev. B **66**, 014447 (2002).

[38] A. Eiguren and C. Ambrosch-Draxl, "Spin polarization and relativistic electronic structure of the 1×1 h/w(110) surface", New Journal of Physics **11**, 013056 (2009).

[39] J. E. Peralta, G. E. Scuseria, and M. J. Frisch, "Noncollinear magnetism in density functional calculations", Phys. Rev. B **75**, 125119 (2007).

[40] I. W. Bulik, G. Scalmani, M. J. Frisch, and G. E. Scuseria, "Noncollinear density functional theory having proper invariance and local torque properties", Phys. Rev. B **87**, 035117 (2013).

[41] J. Kubler, K. .-. Hock, J. Sticht, and A. R. Williams, "Density functional theory of non-collinear magnetism", Journal of Physics F: Metal Physics **18**, 469 (1988).

[42] E. v. Lenthe, E. J. Baerends, and J. G. Snijders, "Relativistic regular two-component Hamiltonians", The Journal of Chemical Physics **99**, 4597–4610 (1993).

[43] E. van Lenthe, E. J. Baerends, and J. G. Snijders, "Relativistic total energy using regular approximations", The Journal of Chemical Physics **101**, 9783–9792 (1994).

[44] C. Vona, S. Lubeck, H. Kleine, A. Gulans, and C. Draxl, "Accurate and efficient treatment of spin-orbit coupling via second variation employing local orbitals", Phys. Rev. B **108**, 235161 (2023).

[45] P. Haas, F. Tran, P. Blaha, K. Schwarz, and R. Laskowski, "Insight into the performance of gga functionals for solid-state calculations", Phys. Rev. B **80**, 195109 (2009).

[46] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method", Phys. Rev. Lett. **45**, 566–569 (1980).

[47] M. Fuchs, M. Bockstedte, E. Pehlke, and M. Scheffler, "Pseudopotential study of binding properties of solids within generalized gradient approximations: the role of core-valence exchange correlation", Phys. Rev. B **57**, 2134–2145 (1998).

[48] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory", Rev. Mod. Phys. **73**, 515–562 (2001).

[49] T. C. Leung, C. T. Chan, and B. N. Harmon, "Ground-state properties of fe, co, ni, and their monoxides: results of the generalized gradient approximation", Phys. Rev. B **44**, 2923–2927 (1991).

[50] P. H. T. Philipsen and E. J. Baerends, "Cohesive energy of 3d transition metals: density functional theory atomic and bulk calculations", Phys. Rev. B **54**, 5326–5333 (1996).

[51] J. W. Furness and J. Sun, "Enhancing the efficiency of density functionals with an improved iso-orbital indicator", Phys. Rev. B **99**, 041119 (2019).

[52] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces", Phys. Rev. Lett. **100**, 136406 (2008).

[53] P. Haas, F. Tran, and P. Blaha, "Calculation of the lattice constant of solids with semilocal functionals", Phys. Rev. B **79**, 085104 (2009).

[54] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids", New Journal of Physics **20**, 063020 (2018).

[55] S. Lehtola, C. Steigemann, M. J. Oliveira, and M. A. Marques, "Recent developments in libxc — a comprehensive library of functionals for density functional theory", SoftwareX **7**, 1–5 (2018).

[56] J. Sun, B. Xiao, Y. Fang, R. Haunschild, P. Hao, A. Ruzsinszky, G. I. Csonka, G. E. Scuseria, and J. P. Perdew, "Density functionals that recognize covalent, metallic, and weak bonds", Phys. Rev. Lett. **111**, 106401 (2013).

[57] F. Tran, G. Baudesson, J. Carrete, G. K. H. Madsen, P. Blaha, K. Schwarz, and D. J. Singh, "Shortcomings of meta-gga functionals when describing magnetism", Phys. Rev. B **102**, 024407 (2020).

[58] J. W. Furness and J. Sun, "Enhancing the efficiency of density functionals with an improved iso-orbital indicator", Phys. Rev. B **99**, 041119 (2019).

[59] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, "Accurate and numerically efficient r2scan meta-generalized gradient approximation", The Journal of Physical Chemistry Letters **11**, PMID: 32876454, 8208–8215 (2020).

[60] A. P. Bartók and J. R. Yates, "Regularized SCAN functional", The Journal of Chemical Physics **150**, 161101 (2019).

[61] D. Mejía-Rodríguez and S. B. Trickey, "Comment on "Regularized SCAN functional" [J. Chem. Phys. 150, 161101 (2019)]", The Journal of Chemical Physics **151**, 207101 (2019).

[62] R. Kingsbury, A. S. Gupta, C. J. Bartel, J. M. Munro, S. Dwaraknath, M. Horton, and K. A. Persson, "Performance comparison of $r^2$SCAN and scan metagga density functionals for solid materials via an automated, high-throughput computational workflow", Phys. Rev. Mater. **6**, 013801 (2022).

[63] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: nonempirical meta–generalized gradient approximation designed for molecules and solids", Phys. Rev. Lett. **91**, 146401 (2003).

[64] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids", New Journal of Physics **20**, 063020 (2018).

[65] Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, "More realistic band gaps from meta-generalized gradient approximations: only in a generalized kohn-sham scheme", Phys. Rev. B **93**, 205205 (2016).

[66] S. Kümmel and L. Kronik, "Orbital-dependent density functionals: theory and applications", Rev. Mod. Phys. **80**, 3–60 (2008).

[67] J. P. Perdew, M. Ernzerhof, and K. Burke, "Rationale for mixing exact exchange with density functional approximations", The Journal of Chemical Physics **105**, 9982–9985 (1996).

[68] C. Franchini, "Hybrid functionals applied to perovskites", Journal of Physics: Condensed Matter **26**, Published 29 May 2014, 253202 (2014).

[69] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential", The Journal of Chemical Physics **118**, 8207–8215 (2003).

[70] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]", The Journal of Chemical Physics **124**, 219906 (2006).

[71] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the exchange screening parameter on the performance of screened hybrid functionals", The Journal of Chemical Physics **125**, 224106 (2006).

[72] L. Schimka, J. Harl, and G. Kresse, "Improved hybrid functional for solids: The HSEsol functional", The Journal of Chemical Physics **134**, 024116 (2011).

[73] F. Furche and J. P. Perdew, "The performance of semilocal and hybrid density functionals in 3d transition-metal chemistry", The Journal of Chemical Physics **124**, 044103 (2006).

[74] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, "Density-functional theory for fractional particle number: derivative discontinuities of the energy", Phys. Rev. Lett. **49**, 1691–1694 (1982).

[75] J. P. Perdew and M. Levy, "Physical content of the exact kohn-sham orbital energies: band gaps and derivative discontinuities", Phys. Rev. Lett. **51**, 1884–1887 (1983).

[76] L. J. Sham and M. Schlüter, "Density-functional theory of the energy gap", Phys. Rev. Lett. **51**, 1888–1891 (1983).

[77] M. Grüning, A. Marini, and A. Rubio, "Effect of spatial nonlocality on the density functional band gap", Phys. Rev. B **74**, 161103 (2006).

[78] F. G. Eich and M. Hellgren, "Derivative discontinuity and exchange-correlation potential of meta-GGAs in density-functional theory", The Journal of Chemical Physics **141**, 224107 (2014).

[79] T. Aschebrock, T. Lebeda, M. Brütting, R. Richter, I. Schelter, and S. Kümmel, "Exact exchange-like electric response from a meta-generalized gradient approximation: A semilocal realization of ultranonlocality", The Journal of Chemical Physics **159**, 234107 (2023).

[80] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", Phys. Rev. B **45**, 13244–13249 (1992).

[81] B. Neupane, H. Tang, N. K. Nepal, S. Adhikari, and A. Ruzsinszky, "Opening band gaps of low-dimensional materials at the meta-gga level of density functional approximations", Phys. Rev. Mater. **5**, 063803 (2021).

[82] J. C. Slater, "Wave functions in a periodic potential", Phys. Rev. **51**, 846–851 (1937).

[83] O. K. Andersen, "Linear methods in band theory", Phys. Rev. B **12**, 3060–3083 (1975).

[84] E. Sjöstedt, L. Nordström, and D. Singh, "An alternative way of linearizing the augmented plane-wave method", Solid state communications **114**, 15–20 (2000).

[85] G. Michalicek, M. Betzinger, C. Friedrich, and S. Blügel, "Elimination of the linearization error and improved basis-set convergence within the flapw method", Computer Physics Communications **184**, 2670–2679 (2013).

[86] C. Vorwerk, C. Cocchi, and C. Draxl, "Addressing electron-hole correlation in core excitations of solids: an all-electron many-body approach from first principles", Phys. Rev. B **95**, 155121 (2017).

[87] D. D. Koelling and B. N. Harmon, "A technique for relativistic spin-polarised calculations", Journal of Physics C: Solid State Physics **10**, 3107 (1977).

[88] J. P. Perdew and L. A. Constantin, "Laplacian-level density functionals for the kinetic energy density and exchange-correlation energy", Phys. Rev. B **75**, 155109 (2007).

[89] H. I. Francisco, J. Carmona-Espíndola, and J. L. Gázquez, "Analysis of the kinetic energy functional in the generalized gradient approximation", The Journal of Chemical Physics **154**, 084107 (2021).

[90] F. Della Sala, E. Fabiano, and L. A. Constantin, "Kinetic-energy-density dependent semilocal exchange-correlation functionals", International Journal of Quantum Chemistry **116**, 1641–1694 (2016).

[91] H. Levämäki and L. Vitos, "Electron localization function implementation in the exact muffin-tin orbitals method", Phys. Rev. B **103**, 035118 (2021).

[92] L.-H. Ye, "Computation of the kohn-sham orbital kinetic energy density in the full-potential linearized augmented plane-wave method", Phys. Rev. B **91**, 075101 (2015).

[93] T. M. Maier, Y. Ikabata, and H. Nakai, "Restoring the iso-orbital limit of the kinetic energy density in relativistic density functional theory", The Journal of Chemical Physics **151**, 174114 (2019).

[94] S. Blügel and G. Bihlmayer, "Full-potential linearized augmented planewave method", in *Computational nanoscience: do it yourself!*, Vol. 31, edited by J. Grotendorst, S. Blügel, and D. Marx, NIC Series (John von Neumann Institute for Computing, Jülich, 2006), p. 85.

[95] G. Arfken, G. Arfken, H. Weber, and F. Harris, *Mathematical methods for physicists: a comprehensive guide* (Elsevier Science, 2013).

[96] J. B. Krieger, Y. Li, and G. J. Iafrate, "Construction and application of an accurate local spin-polarized kohn-sham potential with integer discontinuity: exchange-only theory", Phys. Rev. A **45**, 101–126 (1992).

[97] R. Richter, T. Aschebrock, I. Schelter, and S. Kümmel, "Meta-generalized gradient approximations in time dependent generalized Kohn–Sham theory: Importance of the current density correction", The Journal of Chemical Physics **159**, 124117 (2023).

[98] A. V. Arbuznikov and M. Kaupp, "The self-consistent implementation of exchange-correlation functionals depending on the local kinetic energy density", Chemical Physics Letters **381**, 495–504 (2003).

[99] J. C. Womack, N. Mardirossian, M. Head-Gordon, and C.-K. Skylaris, "Self-consistent implementation of meta-GGA functionals for the ONETEP linear-scaling electronic structure package", The Journal of Chemical Physics **145**, 204114 (2016).

[100] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: the art of scientific computing*, 3rd (Cambridge University Press, Cambridge, 2007), pp. 921–924.

[101] N. A. W. Holzwarth, M. Torrent, J.-B. Charraud, and M. Côté, "Cubic spline solver for generalized density functional treatments of atoms and generation of atomic datasets for use with exchange-correlation functionals including meta-gga", Phys. Rev. B **105**, 125144 (2022).

[102] J. Doumont, F. Tran, and P. Blaha, "Erratum: implementation of self-consistent mgga functionals in augmented plane wave based methods [phys. rev. b 105, 195138 (2022)]", Phys. Rev. B **106**, 159901 (2022).

[103] F. Tran, J. Doumont, P. Blaha, M. A. L. Marques, S. Botti, and A. P. Bartók, "On the calculation of the bandgap of periodic solids with MGGA functionals using the total energy", The Journal of Chemical Physics **151**, 161102 (2019).

[104] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set", Phys. Rev. B **54**, 11169–11186 (1996).

[105] S. Lehtola, "Meta-gga density functional calculations on atoms with spherically symmetric densities in the finite element formalism", Journal of Chemical Theory and Computation **19**, PMID: 37084260, 2502–2517 (2023).

[106] F. G. Eich, S. Pittalis, and G. Vignale, "Transverse and longitudinal gradients of the spin magnetization in spin-density-functional theory", Phys. Rev. B **88**, 245102 (2013).

[107] A. Petrone, D. B. Williams-Young, S. Sun, T. F. Stetina, and X. Li, "An efficient implementation of two-component relativistic density functional theory with torque-free auxiliary variables", The European Physical Journal B **91**, 169 (2018).

[108] G. Scalmani and M. J. Frisch, "A new approach to noncollinear spin density functional theory beyond the local density approximation", Journal of Chemical Theory and Computation **8**, PMID: 26588951, 2193–2196 (2012).

[109] I. W. Bulik, G. Scalmani, M. J. Frisch, and G. E. Scuseria, "Noncollinear density functional theory having proper invariance and local torque properties", Phys. Rev. B **87**, 035117 (2013).

[110] F. G. Eich and E. K. U. Gross, "Transverse spin-gradient functional for noncollinear spin-density-functional theory", Phys. Rev. Lett. **111**, 156401 (2013).

[111] J. K. Desmarais, S. Komorovsky, J.-P. Flament, and A. Erba, "Spin–orbit coupling from a two-component self-consistent approach. II. Non-collinear density functional theories", The Journal of Chemical Physics **154**, 204110 (2021).

[112] F. Egidi, S. Sun, J. J. Goings, G. Scalmani, M. J. Frisch, and X. Li, "Two-component noncollinear time-dependent spin density functional theory for excited state calculations", Journal of Chemical Theory and Computation **13**, PMID: 28481542, 2591–2603 (2017).

[113] J. A. Pople, P. M. Gill, and B. G. Johnson, "Kohn—sham density-functional theory within a finite basis set", Chemical Physics Letters **199**, 557–560 (1992).

[114] M. Weinert, E. Wimmer, and A. J. Freeman, "Total-energy all-electron density functional method for bulk solids and surfaces", Phys. Rev. B **26**, 4571–4578 (1982).

[115] W. Demtröder, *Atoms, molecules and photons: an introduction to atomic-, molecular- and quantum physics*, 3rd ed., Graduate Texts in Physics, Published: 20 February 2019 (hardcover), 09 February 2019 (eBook) (Springer Berlin, Heidelberg, 2019) Chap. Relevant Chapter, pp. XVIII, 551.

[116] O. Madelung, *Semiconductors: data handbook* (Springer-Verlag Berlin Heidelberg, 2004).

[117] B. Hammer, L. B. Hansen, and J. K. Nørskov, "Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals", Phys. Rev. B **59**, 7413–7421 (1999).

[118] A. D. Boese and N. C. Handy, "A new parametrization of exchange–correlation generalized gradient approximation functionals", The Journal of Chemical Physics **114**, 5497–5503 (2001).

[119] F. Tran and P. Blaha, "Importance of the kinetic energy density for band gap calculations in solids with density functional theory", The Journal of Physical Chemistry A **121**, PMID: 28402113, 3318–3325 (2017).

[120] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids", New Journal of Physics **20**, 063020 (2018).

[121] F. Birch, "Finite elastic strain of cubic crystals", Physical Review **71**, 809–824 (1947).

[122] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, "Ab initio molecular simulations with numeric atom-centered orbitals", Computer Physics Communications **180**, 2175–2196 (2009).

[123] M. J. Frisch et al., *GAUSSIAN Development Version, Revision B.01*, Development Version, Pittsburgh, PA, 2003.

[124] G. te Velde and E. J. Baerends, "Precise density-functional method for periodic structures", Phys. Rev. B **44**, 7888–7903 (1991).

[125] SCM, Theoretical Chemistry, *BAND 2024.1*, `http://www.scm.com`, Vrije Universiteit, Amsterdam, The Netherlands, 2024.

[126] D. Mejia-Rodriguez and S. B. Trickey, "Deorbitalized meta-gga exchange-correlation functionals in solids", Phys. Rev. B **98**, 115161 (2018).

[127] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids", New Journal of Physics **20**, 063020 (2018).

[128] J. Sun, M. Marsman, G. I. Csonka, A. Ruzsinszky, P. Hao, Y.-S. Kim, G. Kresse, and J. P. Perdew, "Self-consistent meta-generalized gradient approximation within the projector-augmented-wave method", Phys. Rev. B **84**, 035117 (2011).

[129] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, "Tests of a ladder of density functionals for bulk solids and surfaces", Phys. Rev. B **69**, 075102 (2004).

[130] P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew, "Lattice constants from semilocal density functionals with zero-point phonon correction", Phys. Rev. B **85**, 014111 (2012).

[131] J. Kim, A. Gulans, and C. Draxl, "Robust mixing in self-consistent linearized augmented planewave calculations", Electronic Structure **2**, 037001 (2020).

[132] G. P. Kerker, "Efficient iteration scheme for self-consistent pseudopotential calculations", Phys. Rev. B **23**, 3082–3084 (1981).

[133] N. D. Woods, M. C. Payne, and P. J. Hasnip, "Computing the self-consistent field in kohn–sham density functional theory", Journal of Physics: Condensed Matter **31**, 453001 (2019).

[134] P. Pulay, "Improved scf convergence acceleration", Journal of Computational Chemistry **3**, 556–560 (1982).

[135] P. Pulay, "Convergence acceleration of iterative sequences. the case of scf iteration", Chemical Physics Letters **73**, 393–398 (1980).

[136] K. Pokharel, C. Lane, J. W. Furness, A. Faghaninia, Y. Gao, D. Seo, J. J. Kas, F. D. Vila, J. J. Rehr, G. A. Glatzmaier, A. A. Kananenka, and J. Sun, "Sensitivity of the electronic and magnetic structures of cuprate superconductors to density functional approximations", npj Computational Materials **8**, 31 (2022).

[137] F. Della Sala, E. Fabiano, and L. A. Constantin, "Kinetic-energy-density dependent semilocal exchange-correlation functionals", International Journal of Quantum Chemistry **116**, 1641–1694 (2016).

# A. Derivations

## A.1. Functional Derivative of the Exchange-Correlation Functional

Let us consider the calculation of the functional derivative with respect to the orbital $\psi_i^*$ of an exchange-correlation functional of the form [137]

$$E_{\mathrm{xc}} = \int d\mathbf{r}\, \varepsilon_{\mathrm{xc}}\left(n(\mathbf{r}), \nabla n(\mathbf{r}), \tau(\mathbf{r})\right) \tag{A.1}$$

where $n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$, and $\tau(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla\psi_i(\mathbf{r})|^2$ is the kinetic energy density. The functional derivative of $E_{\mathrm{xc}}$ can be derived by considering the variation of the orbital $\delta\psi_i^*$,

$$\int d\mathbf{r}\, \frac{\delta E_{\mathrm{xc}}}{\delta\psi_i^*}\, \delta\psi_i^* = \frac{d}{d\lambda} E_{\mathrm{xc}}[\psi_i^* + \lambda\delta\psi_i^*]\bigg|_{\lambda=0}. \tag{A.2}$$

The transformation $\psi_i^* \to \psi_i^* + \lambda\delta\psi_i^*$ results in the following variations:

$$\begin{aligned}
\tau(\mathbf{r}) &\to \tau(\mathbf{r}) + \lambda \int d\mathbf{r}'\, \frac{\delta\tau(\mathbf{r})}{\delta\psi_i^*(\mathbf{r}')}\, \delta\psi_i^*(\mathbf{r}') \\
&= \tau(\mathbf{r}) + \lambda \int d\mathbf{r}'\, \frac{1}{2}\nabla\psi_i(\mathbf{r}) \cdot \nabla\delta(\mathbf{r}-\mathbf{r}')\, \delta\psi_i^*(\mathbf{r}') \\
&= \tau(\mathbf{r}) + \lambda\frac{1}{2}\nabla\psi_i(\mathbf{r}) \cdot \nabla\delta\psi_i^*(\mathbf{r}).
\end{aligned} \tag{A.3}$$

Similarly, for the density components, we find:

$$n(\mathbf{r}) \to n(\mathbf{r}) + \lambda\psi_i(\mathbf{r})\, \delta\psi_i^*(\mathbf{r}), \tag{A.4}$$

$$\nabla n(\mathbf{r}) \to \nabla n(\mathbf{r}) + \lambda\nabla\psi_i(\mathbf{r})\, \delta\psi_i^*(\mathbf{r}). \tag{A.5}$$

The right-hand side of (A.2) becomes

$$\begin{aligned}
\frac{d}{d\lambda} E_{\mathrm{xc}}[n + \lambda\delta n]\bigg|_{\lambda=0} &= \int d\mathbf{r} \left[\frac{\partial\varepsilon_{\mathrm{xc}}}{\partial n(\mathbf{r})}\psi_i(\mathbf{r})\, \delta\psi_i^*(\mathbf{r}) + \frac{\partial\varepsilon_{\mathrm{xc}}}{\partial\nabla n(\mathbf{r})} \cdot \nabla\psi_i(\mathbf{r})\, \delta\psi_i^*(\mathbf{r})\right] \\
&\quad + \int d\mathbf{r}\, \frac{1}{2}\frac{\partial\varepsilon_{\mathrm{xc}}}{\partial\tau(\mathbf{r})}\nabla\psi_i(\mathbf{r}) \cdot \nabla\, \delta\psi_i^*(\mathbf{r}) \\
&= \int d\mathbf{r} \left[\frac{\partial\varepsilon_{\mathrm{xc}}}{\partial n(\mathbf{r})}\psi_i(\mathbf{r}) + \frac{\partial\varepsilon_{\mathrm{xc}}}{\partial\nabla n(\mathbf{r})} \cdot \nabla\psi_i(\mathbf{r})\right. \\
&\quad \left. -\frac{1}{2}\nabla \cdot \left(\frac{\partial\varepsilon_{\mathrm{xc}}}{\partial\tau(\mathbf{r})}\nabla\psi_i(\mathbf{r})\right)\right] \delta\psi_i^*(\mathbf{r}).
\end{aligned} \tag{A.6}$$

Finally, comparing this with the left side of Eq. (A.2), the functional derivative is given by:

$$\frac{\delta E_{\mathrm{xc}}}{\delta \psi_i^*} = \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n(\mathbf{r})}\psi_i(\mathbf{r}) - \nabla \cdot \left(\frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \nabla n(\mathbf{r})}\right)\psi_i(\mathbf{r}) - \frac{1}{2}\nabla \cdot \left(\frac{\partial \varepsilon_{\mathrm{xc}}}{\partial \tau(\mathbf{r})}\nabla\psi_i(\mathbf{r})\right). \tag{A.7}$$

## A.2. Non-Collinear Functional Variables and Derivatives

The spin magnetisation related variables are rewritten in terms of the density matrix

$$D_{\mu\nu}^{\sigma\sigma'} = \sum_i C_{\mu\sigma}^{i*} C_{\nu\sigma}^{i} \tag{A.8}$$

as follows:

$$n(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}^{0} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu\nu} \left(D_{\mu\nu}^{\alpha\alpha} + D_{\mu\nu}^{\sigma'\sigma'}\right)\phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{A.9}$$

$$m_x(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}^{x} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu\nu} \left(D_{\mu\nu}^{\sigma\sigma'} + D_{\mu\nu}^{\sigma'\sigma}\right)\phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{A.10}$$

$$m_y(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}^{y} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu\nu} i\left(D_{\mu\nu}^{\sigma\sigma'} - D_{\mu\nu}^{\sigma'\sigma}\right)\phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{A.11}$$

$$m_z(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}^{z} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu\nu} \left(D_{\mu\nu}^{\sigma\sigma} - D_{\mu\nu}^{\sigma'\sigma'}\right)\phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{A.12}$$

The xc potential matrix elements are evaluated by taking the functional derivative with respect to the density matrix as follows

$$(V_{\mathrm{xc}})_{\mu\nu}^{i} = \sum_{pq} \int d\mathbf{r} \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial U_p} \frac{\partial U_p}{\partial V_q} \frac{\partial V_q}{\partial D_{\mu\nu}^{i}}. \tag{A.13}$$

where $\{U\} = \{n^\uparrow, n^\downarrow, \gamma^{\uparrow\uparrow}, \gamma^{\downarrow\downarrow}, \gamma^{\uparrow\downarrow}, \tau^\uparrow, \tau^\downarrow\}$ as defined by Scalmani and Frisch (3.79) and $\{V\} = \{n, \mathbf{m}, \tau, \mathbf{u}\}$. Using the chain rule, the following expression is derived for the xc potential matrix $(V_{\mathrm{xc}})_{\mu\nu}^{0}$:

$$\begin{aligned}
(V_{\mathrm{xc}})_{\mu v}^{0} = &\int d\mathbf{r} \left(\frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\uparrow}\frac{\partial n_\uparrow}{\partial n} + \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\downarrow}\frac{\partial n_\downarrow}{\partial n}\right)\frac{\partial n}{\partial D_{\mu\nu}^{0}} \\
&+ \int d\mathbf{r} \left(\frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma_{\uparrow\uparrow}}\frac{\partial \gamma_{\uparrow\uparrow}}{\partial \nabla n} + \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma_{\downarrow\downarrow}}\frac{\partial \gamma_{\downarrow\downarrow}}{\partial \nabla n} + \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \gamma_{\uparrow\downarrow}}\frac{\partial \gamma_{\uparrow\downarrow}}{\partial \nabla n}\right)\frac{\partial \nabla n}{\partial D_{\mu\nu}^{0}} \\
&+ \int d\mathbf{r} \left(\frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau_\uparrow}\frac{\tau_\uparrow}{\tau} + \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial \tau_\downarrow}\frac{\tau_\downarrow}{\tau}\right)\frac{\partial \tau}{\partial D_{\mu\nu}^{0}}
\end{aligned} \tag{A.14}$$

The needed derivatives can be evaluated straightforwardly:

$$\frac{\partial n_{\uparrow/\downarrow}}{\partial n} = \frac{1}{2} \tag{A.15}$$

$$\frac{\partial \gamma_{\uparrow\downarrow}}{\partial \nabla n} = \frac{1}{2}\nabla n \tag{A.16}$$

$$\frac{\partial \gamma_{\uparrow\uparrow/\downarrow\downarrow}}{\partial \nabla n} = \frac{1}{2}\nabla n \pm \frac{1}{2}\frac{\nabla \mathbf{m} \circ (\nabla \mathbf{m} \cdot \nabla n)}{\Gamma} \tag{A.17}$$

$$\frac{\partial \tau_{\uparrow/\downarrow}}{\tau} = \frac{1}{2} \tag{A.18}$$

where $\Gamma = \sqrt{(\nabla n \cdot \nabla \mathbf{m}) \circ (\nabla n \cdot \nabla \mathbf{m})}$. The resulting matrix elements are then given by:

$$\begin{aligned}
(V_{\mathrm{xc}})^0_{\mu\nu} = &\frac{1}{2}\int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\uparrow} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\downarrow} \right) \phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \\
&+ \frac{1}{2}\int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\uparrow\uparrow}} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\downarrow\downarrow}} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\uparrow\downarrow}} \right) \nabla n \cdot \nabla \left[ \phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \right] \\
&+ \int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\uparrow\uparrow}} - \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\downarrow\downarrow}} \right) \frac{f_\nabla}{2} \frac{\nabla n \cdot \nabla \mathbf{m} \circ \nabla \mathbf{m}}{\Gamma} \nabla \left[ \phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \right] \\
&+ \int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \tau_\uparrow} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \tau_\downarrow} \right) \frac{1}{4} \nabla \phi^*_\mu(\mathbf{r}) \nabla \phi_\nu(\mathbf{r})
\end{aligned}$$

Similarly, to evaluate the expression for the xc magnetic potential field matrix $(B_{\mathrm{xc}})^i_{\mu\nu}$ the chain rule is taken:

$$\begin{aligned}
(B_{\mathrm{xc}})^i_{\mu\nu} = &\int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\uparrow} \frac{\partial n_\uparrow}{\partial m^i} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial n_\downarrow} \frac{\partial n_\downarrow}{\partial m^i} \right) \frac{\partial m^i}{\partial D^i_{\mu\nu}} \\
&+ \int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\uparrow\uparrow}} \frac{\partial \gamma_{\uparrow\uparrow}}{\partial \nabla m^i} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\downarrow\downarrow}} \frac{\partial \gamma_{\downarrow\downarrow}}{\partial \nabla m^i} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \gamma_{\uparrow\downarrow}} \frac{\partial \gamma_{\uparrow\downarrow}}{\partial \nabla m^i} \right) \frac{\partial \nabla m^i}{\partial D^i_{\mu\nu}} \\
&+ \int d\mathbf{r} \left( \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \tau_\uparrow} \frac{\partial \tau_\uparrow}{\partial u^i} + \frac{\partial \varepsilon^{\mathrm{mGGA}}_{\mathrm{xc}}}{\partial \tau_\downarrow} \frac{\partial \tau_\downarrow}{\partial u^i} \right) \frac{\partial u^i}{\partial D^i_{\mu\nu}}
\end{aligned} \tag{A.19}$$

The derivatives are given by:

$$\frac{n_{\uparrow/\downarrow}}{\partial m} = \pm\frac{1}{2}\frac{\mathbf{m}}{\sqrt{\mathbf{m} \circ \mathbf{m}}} \tag{A.20}$$

$$\frac{\partial \gamma_{\uparrow\downarrow}}{\partial \nabla \mathbf{m}} = -\frac{1}{2}\nabla \mathbf{m} \tag{A.21}$$

$$\frac{\partial \gamma_{\uparrow\uparrow/\downarrow\downarrow}}{\partial \nabla \mathbf{m}} = \frac{1}{2}\nabla \mathbf{m} \pm \frac{f_\nabla}{2}\frac{(\nabla n \cdot \nabla \mathbf{m}) \circ \nabla n}{\Gamma} \tag{A.22}$$

$$\frac{\partial \tau_{\uparrow/\downarrow}}{\mathbf{u}} = \pm\frac{f_\tau}{2}\frac{\mathbf{u}}{\sqrt{\mathbf{u} \circ \mathbf{u}}} \tag{A.23}$$

This yields the following matrix elements:

$$\begin{aligned}
(B_{\mathrm{xc}})^i_{\mu\nu} = &\frac{1}{2}\int d\mathbf{r}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\downarrow}\right)\frac{m^i}{|\mathbf{m}|}\left[\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\right] \\
&+\frac{1}{2}\int d\mathbf{r}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\downarrow}}\right)\nabla m^i\cdot\nabla\left[\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\right] \\
&+\int d\mathbf{r}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}}\right)\frac{f_\nabla}{2}\frac{\nabla n\cdot\nabla\mathbf{m}\cdot\nabla n}{\Gamma}\nabla\left[\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\right] \\
&+\int d\mathbf{r}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\frac{f_{\mathbf{u}}}{4}\frac{u^i}{|\mathbf{u}|}\nabla\phi^*_\mu(\mathbf{r})\nabla\phi_\nu(\mathbf{r}).
\end{aligned} \tag{A.24}$$

Finally, for regions where $\mathbf{m}(\mathbf{r}) < 10^{-12}$ the transformations

$$\begin{aligned}
\sqrt{\mathbf{m}\cdot\mathbf{m}} &\to m_s(\mathbf{r}) = \frac{1}{3}\sum_i m_i(\mathbf{r}), \\
\sqrt{\nabla n\cdot\nabla\mathbf{m}\circ\nabla\mathbf{m}\cdot\nabla n} &\to \nabla n\cdot\nabla m_s \\
\sqrt{\mathbf{u}\cdot\mathbf{u}} &\to m_s(\mathbf{r}) = \frac{1}{3}\sum_i u_i(\mathbf{r})
\end{aligned} \tag{A.25}$$

are done, to ensure a proper convergence behaviour. Thus the formulas for the xc potential change to

$$\begin{aligned}
(V_{\mathrm{xc}})^0 = &\frac{1}{2}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\uparrow} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\downarrow}\right) \\
&-\frac{1}{2}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\downarrow}}\right)\nabla n\right] \\
&-\frac{f_\nabla}{2}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}}\right)\nabla m_s\right] \\
&-\frac{1}{4}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\nabla\cdot\right],
\end{aligned} \tag{A.26}$$

and equally the xc magnetic field changes to

$$\begin{aligned}
(B_{\mathrm{xc}})^u = &\frac{1}{6}\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial n_\downarrow}\right) \\
&-\frac{1}{2}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} + \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\downarrow}}\right)\nabla m^u\right] \\
&-\frac{f_\nabla}{6}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\uparrow\uparrow}} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\gamma_{\downarrow\downarrow}}\right)\nabla n\right] \\
&-\frac{f_{\mathbf{u}}}{12}\nabla\left[\left(\frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\uparrow} - \frac{\partial\varepsilon_{\mathrm{xc}}^{\mathrm{mGGA}}}{\partial\tau_\downarrow}\right)\nabla\cdot\right].
\end{aligned} \tag{A.27}$$

## A.3. Derivation of the KED in the Muffin-Tin Region

To derive the KED coefficients, the following expression needs to be evaluated [95]

$$\begin{aligned}\nabla\phi_\lambda^\alpha(\mathbf{r}) &= \nabla\left(g_\lambda^\alpha(r_\alpha)Y_{\ell_\lambda m_\lambda}(\hat{\mathbf{r}}_\alpha)\right)\\ &= g_\lambda^{\alpha-}(r_\alpha)\mathbf{Y}_{\ell_\lambda m_\lambda}^{\ell_\lambda-1}(\hat{\mathbf{r}}_\alpha) + g_\lambda^{\alpha+}(r_\alpha)\mathbf{Y}_{\ell_\lambda m_\lambda}^{\ell_\lambda+1}(\hat{\mathbf{r}}_\alpha)\end{aligned} \tag{A.28}$$

to obtain the gradient of the occupied orbitals.
The radial functions $g_\lambda^{\alpha-}(r_\alpha)$ and $g_\lambda^{\alpha+}(r_\alpha)$ are defined as

$$g_\lambda^{\alpha-}(r_\alpha) = \left[\frac{\ell_\lambda}{2l_\lambda+1}\right]^{\frac{1}{2}}\left(\frac{d}{dr} - \frac{\ell_\lambda+1}{r}\right)g_\lambda^\alpha(r_\alpha) \tag{A.29}$$

$$g_\lambda^{\alpha+}(r_\alpha) = -\left[\frac{\ell_\lambda+1}{2\ell_\lambda+1}\right]^{\frac{1}{2}}\left(\frac{d}{dr} + \frac{\ell_\lambda}{r}\right)g_\lambda^\alpha(r_\alpha), \tag{A.30}$$

and the vector spherical harmonics $\mathbf{Y}_{\ell_\lambda,m_\lambda}^{\ell_\lambda\pm1}(\hat{\mathbf{r}})$ are expanded in terms of scalar spherical harmonics, defined as:

$$\mathbf{Y}_{\ell_\lambda m_\lambda}^{\ell_\lambda\pm1}(\hat{\mathbf{r}}_\alpha) = \sum_{\gamma}^{-1,0,1}\sum_{m=-(\ell_\lambda\pm1)}^{\ell_\lambda\pm1} C_{\ell_{\lambda_1}\pm1,m,1,\gamma}^{\ell_\lambda,m_\lambda}Y_{\ell_\lambda\pm1m}\hat{\mathbf{e}}_\gamma. \tag{A.31}$$

The coefficients $C_{\ell_{\lambda_1}\pm1,m,1,\gamma}^{\ell_\lambda,m_\lambda}$ are the Clebsch-Gordan coefficients. The spherical basis is given by:

$$\hat{\mathbf{e}}_1 = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x + i\hat{\mathbf{e}}_y), \qquad \hat{\mathbf{e}}_0 = \hat{\mathbf{e}}_z, \qquad \hat{\mathbf{e}}_{-1} = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x - i\hat{\mathbf{e}}_y). \tag{A.32}$$

Then the gradient for the total wave function in terms of the scalar spherical harmonics in the direction $\gamma$ is given by

$$\begin{aligned}\partial_\gamma\Psi_{n\mathbf{k}}^\alpha(\mathbf{r}) = \sum_\lambda C_\lambda^{n\mathbf{k}\alpha}\Bigg[&g_\lambda^{\alpha+}(r_\alpha)\sum_{m=-(\ell_\lambda+1)}^{\ell_\lambda+1} C_{\ell_\lambda+1,m,1,\gamma}^{\ell_\lambda,m_\lambda}Y_{\ell_\lambda+1,m}(\hat{\mathbf{r}}_\alpha)\\ &+ g_\lambda^{\alpha-}(r_\alpha)\sum_{m=-(\ell_\lambda-1)}^{\ell_\lambda-1} C_{\ell_\lambda-1,m,1,\gamma}^{\ell_\lambda,m_\lambda}Y_{\ell_\lambda-1,m}(\hat{\mathbf{r}}_\alpha)\Bigg].\end{aligned} \tag{A.33}$$

To obtain the KED the product of the gradient of the wave function needs to be evaluated:

$$\nabla\psi^{\alpha*}_{n\mathbf{k}}\cdot\nabla\psi^{\alpha}_{n\mathbf{k}} = \sum_{\lambda_1\lambda_2} C^{n\mathbf{k}\alpha *}_{\lambda_1} C^{n\mathbf{k}\alpha}_{\lambda_2} \Bigg\{ \tag{A.34}$$

$$\left[g^{\alpha+}_{\lambda_1} g^{\alpha+}_{\lambda_2} \sum_{\substack{m'= \\ -(\ell_{\lambda_1}+1)}}^{\ell_{\lambda_1}+1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}+1)}}^{\ell_{\lambda_2}+1} \sum_{\gamma} \left(C^{\ell_{\lambda_1},m_{\lambda_1}}_{\ell_{\lambda_1}+1,m',1,\gamma}\right)^* C^{\ell_{\lambda_2},m_{\lambda_2}}_{\ell_{\lambda_1}+1,m'',1,\gamma} Y^*_{\ell_{\lambda_1}+1,m'} Y_{\ell_{\lambda_2}+1,m''}\right] \tag{A.35}$$

$$+ \left[g^{\alpha+}_{\lambda_1} g^{\alpha-}_{\lambda_2} \sum_{\substack{m'= \\ -(\ell_{\lambda_1}+1)}}^{\ell_{\lambda_1}+1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}-1)}}^{\ell_{\lambda_2}-1} \sum_{\gamma} \left(C^{\ell_{\lambda_1},m_{\lambda_1}}_{\ell_{\lambda_1}+1,m',1,\gamma}\right)^* C^{\ell_{\lambda_2},m_{\lambda_2}}_{\ell_{\lambda_2}-1,m'',1,\gamma} Y^*_{\ell_{\lambda_1}+1,m'} Y_{\ell_{\lambda_2}-1,m''}\right] \tag{A.36}$$

$$+ \left[g^{\alpha-}_{\lambda_1} g^{\alpha+}_{\lambda_2} \sum_{\substack{m'= \\ -(\ell_{\lambda_1}-1)}}^{\ell_{\lambda_1}-1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}+1)}}^{\ell_{\lambda_2}+1} \sum_{\gamma} \left(C^{\ell_{\lambda_1},m_{\lambda_1}}_{\ell_{\lambda_1}-1,m',1,\gamma}\right)^* C^{\ell_{\lambda_2},m_{\lambda_2}}_{\ell_{\lambda_2}+1,m'',1,\gamma} Y^*_{\ell_{\lambda_1}-1,m'} Y_{\ell_{\lambda_2}+1,m''}\right] \tag{A.37}$$

$$+ \left[g^{\alpha-}_{\lambda_1} g^{\alpha-}_{\lambda_2} \sum_{\substack{m'= \\ -(\ell_{\lambda_1}-1)}}^{\ell_{\lambda_1}-1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}-1)}}^{\ell_{\lambda_2}-1} \sum_{\gamma} \left(C^{\ell_{\lambda_1},m_{\lambda_1}}_{\ell_{\lambda_1}-1,m',1,\gamma}\right)^* C^{\ell_{\lambda_2},m_{\lambda_2}}_{\ell_{\lambda_2}-1,m'',1,\gamma} Y^*_{\ell_{\lambda_1}-1,m'} Y_{\ell_{\lambda_2}-1,m''}\right] \Bigg\}, \tag{A.38}$$

and plugged into the definition of the positive-definite KED, see Eq. 3.6. On the other hand, we want to obtain the coefficients $\tau^{\alpha}_{\ell m}(r_\alpha)$ of the KED, which are defined through the expansion:

$$\tau(\mathbf{r}) = \sum_{\alpha\ell m} \tau^{\alpha}_{\ell m}(r_\alpha) R_{\ell m}(\hat{\mathbf{r}}_\alpha). \tag{A.39}$$

For this, both sides of the equation (A.39) and (A.34) are multiplied by $R^*_{\ell m}$ and we integrate over the whole unit cell. Due to the orthogonality condition $\langle R_{\ell m} \mid R_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'}$ [95], the coefficients $\tau^{\alpha}_{\ell m}(r_\alpha)$ can then be determined from (A.39). For the right hand-side (A.34) an integral over three spherical harmonics is obtained, which are given by the Gaunt coefficients [95]:

$$G^{\ell m}_{\ell_1 m_1 \ell_2 m_2} = \langle Y_{\ell_1 m_1} \mid R_{\ell m} \mid Y_{\ell_2 m_2}\rangle = \int Y^*_{\ell_1 m_1}\, R_{\ell m}\, Y_{\ell_2 m_2}\, d\Omega. \tag{A.40}$$

Consequently, by defining

$$P^{\pm\pm'}_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}} = \sum_{\substack{m'= \\ -(\ell_{\lambda_1}\pm 1)}}^{\ell_{\lambda_1}\pm 1} \sum_{\substack{m''= \\ -(\ell_{\lambda_2}\pm' 1)}}^{\ell_{\lambda_2}\pm' 1} \sum_{\gamma} \left(C^{\ell_{\lambda_1},m_{\lambda_1}}_{\ell_{\lambda_1}\pm 1,m',1,\gamma}\right)^* C^{\ell_{\lambda_2},m_{\lambda_2}}_{\ell_{\lambda_2}\pm 1,m'',1,\gamma} \langle Y_{\ell_{\lambda_1}\pm 1,m'} | R_{\ell m} | Y_{\ell_{\lambda_2}\pm' 1 m''}\rangle\,, \tag{A.41}$$

the coefficients for each muffin-tin sphere $\alpha$ are given by

$$\tau^{\alpha}_{\ell m} = \frac{1}{2} \sum_{\lambda_1\lambda_2} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \left(C^{n\mathbf{k}\alpha}_{\lambda_1}\right)^* C^{n\mathbf{k}\alpha}_{\lambda_2} \sum_{\pm} \sum_{\pm'} P^{\pm\pm'}_{\ell m;\ell_{\lambda_1} m_{\lambda_1} \ell_{\lambda_2} m_{\lambda_2}} g^{\alpha\pm}_{\lambda_1} g^{\alpha\pm'}_{\lambda_2}. \tag{A.42}$$

Additionally, defining the matrix

$$D^{\alpha}_{\lambda_1\lambda_2} = \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \left(C^{n\mathbf{k}\alpha}_{\lambda_1}\right)^* C^{n\mathbf{k}\alpha}_{\lambda_2}, \tag{A.43}$$

leads to the final implemented expression for the KED:

$$\tau^{\alpha}_{\ell m} = \frac{1}{2}\left\{\sum_{\lambda_1\lambda_2} D^{\alpha}_{\lambda_1\lambda_2} \sum_{\pm}\sum_{\pm'} P^{\pm\pm'}_{\ell m;\ell_{\lambda_1} m_{\lambda_1}\ell_{\lambda_2} m_{\lambda_2}} g^{\alpha\pm}_{\lambda_1} g^{\alpha\pm'}_{\lambda_2}\right\}. \tag{A.44}$$

## A.4. Derivation of the Matrix Elements

In the following, the matrix elements for any multiplicative potential $V_{\mathrm{xc}}(\mathbf{r})$ within the LAPW basis are derived, considering both the muffin-tin and interstitial regions. The basis functions used are defined in Section 3.2.2, where the matrix elements for the non-multiplicative potential are also derived. For the muffin-tin region $\alpha$, the matrix elements can be expressed as follows:

$$\begin{aligned}
V^{\alpha}_{\mathrm{xc},\mu\nu} &= \langle \phi^{\alpha}_{\mu\mathbf{k}} \mid V^{\alpha}_{\mathrm{xc}} \mid \phi^{\alpha}_{\nu\mathbf{k}} \rangle \\
&= \int_{S_\alpha} d\mathbf{r}\, \phi^{\alpha *}_{\mu\mathbf{k}}(\mathbf{r}) \sum_{\ell m} V^{\alpha}_{\mathrm{xc},\ell m}(r_\alpha) Y_{\ell m}(\hat{\mathbf{r}}_\alpha)\, \phi^{\alpha}_{\nu\mathbf{k}}(\mathbf{r}) \\
&= \sum_{\lambda_1\lambda_2} C^{\alpha *}_{\lambda_1\mu} C^{\alpha}_{\lambda_2\nu} \sum_{\ell m} \int_{S_\alpha} d\mathbf{r}\, g^{\alpha}_{\lambda_1}(r_\alpha)\, v^{\alpha}_{\mathrm{xc},\ell m}(r_\alpha)\, g^{\alpha}_{\lambda_2}(r_\alpha)\, Y^*_{\ell_{\lambda_1} m_{\lambda_1}}(\hat{\mathbf{r}}_\alpha) Y_{\ell m}(\hat{\mathbf{r}}_\alpha) Y_{\ell_{\lambda_2} m_{\lambda_2}}(\hat{\mathbf{r}}_\alpha) \\
&= \sum_{\lambda_1\lambda_2} C^{\alpha *}_{\lambda_1\mu} C^{\alpha}_{\lambda_2\nu} \sum_{\ell m} R^{\alpha}_{\ell m;\lambda_1\lambda_2} G_{\ell m;\ell_{\lambda_1} m_{\lambda_1}\ell_{\lambda_2} m_{\lambda_2}},
\end{aligned} \tag{A.45}$$

where the summation over $\{\ell m\}$ is independent of any $\mathbf{k}$-points. Here,

$$R^{\alpha}_{\ell m;\lambda_1\lambda_2} = \int_0^{R_{MT_\alpha}} g^{\alpha}_{\lambda_1}(r)\, v^{\alpha}_{\mathrm{xc},\ell m}(r_\alpha)\, g^{\alpha}_{\lambda_2}(r)\, r^2\, dr \tag{A.46}$$

is the integral over the radial functions, and

$$G_{\ell m;\ell_{\lambda_1} m_{\lambda_1}\ell_{\lambda_2} m_{\lambda_2}} = \langle Y_{\ell_{\lambda_1} m_{\lambda_1}} \mid Y_{\ell m} \mid Y_{\ell_{\lambda_2} m_{\lambda_2}} \rangle \tag{A.47}$$

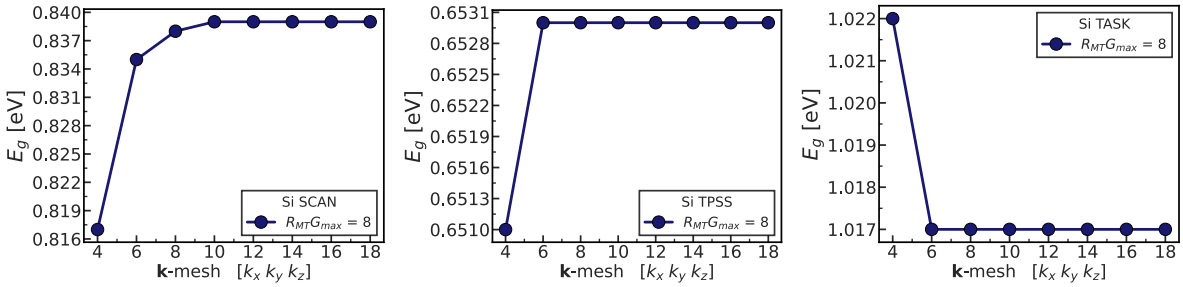
is the spherical integral over three spherical harmonics, known as the Gaunt coefficients. To evaluate the matrix elements in the interstitial region, the following expression is derived:

$$\begin{aligned}
V^{\mathrm{IR}}_{\mathrm{xc},\mathbf{G_1}\mathbf{G_2}} &= \langle \phi_{\mathbf{G_1}+\mathbf{k}} \mid V^{\mathrm{IR}}_{\mathrm{xc}} \mid \phi_{\mathbf{G_2}+\mathbf{k}} \rangle \\
&= \int_{\mathrm{IR}} d\mathbf{r}\, \phi^*_{\mathbf{G_1}+\mathbf{k}}(\mathbf{r})\, V^{\mathrm{IR}}_{\mathrm{xc}}(\mathbf{r})\, \phi_{\mathbf{G_2}+\mathbf{k}}(\mathbf{r}) \\
&= \sum_{\mathbf{G}} v^{\mathrm{IR}}_{\mathrm{xc}}(\mathbf{G}) \frac{1}{\Omega} \int_\Omega d\mathbf{r}\, e^{-i(\mathbf{G}_1+\mathbf{k})\cdot\mathbf{r}}\, e^{i\mathbf{G}\cdot\mathbf{r}}\, e^{i(\mathbf{G}_2+\mathbf{k})\cdot\mathbf{r}}\, \Theta_{\mathrm{IR}}(\mathbf{r}) \\
&= \sum_{\mathbf{G}} v^{\mathrm{IR}}_{\mathrm{xc}}(\mathbf{G}) \frac{1}{\Omega} \int_\Omega d\mathbf{r}\, e^{-i(\mathbf{G}_1-\mathbf{G}_2-\mathbf{G})\cdot\mathbf{r}}\, \Theta_{\mathrm{IR}}(\mathbf{r}) \\
&= \sum_{\mathbf{G}} v^{\mathrm{IR}}_{\mathrm{xc}}(\mathbf{G})\, \Theta_{\mathrm{IR}}(\mathbf{G}_1 - \mathbf{G}_2 - \mathbf{G}) \\
&= v^{\mathrm{IR}}_{\mathrm{xc},\theta}(\mathbf{G}_1 - \mathbf{G}_2).
\end{aligned} \tag{A.48}$$

In practice, the product of the characteristic function $\Theta(\mathbf{G})$ and the potential is multiplied in real space to avoid the computational cost of a convolution, as outlined in Sections A.3 and 3.2.2.

# B. Plots

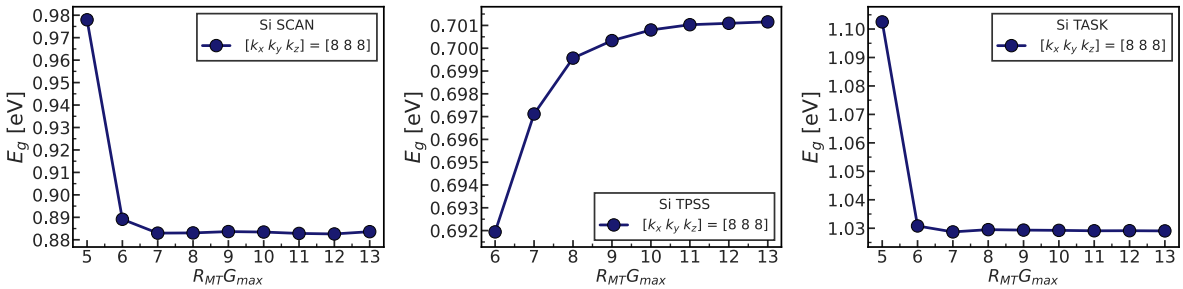
## B.1. Convergence Plots: Band Gap




(a) Convergence behavior of the Si SCAN band gap w.r.t. the **k**-mesh.

(b) Convergence behavior of the Si TPSS band gap w.r.t. the **k**-mesh.

(c) Convergence behavior of the Si TASK band gap w.r.t. the **k**-mesh.

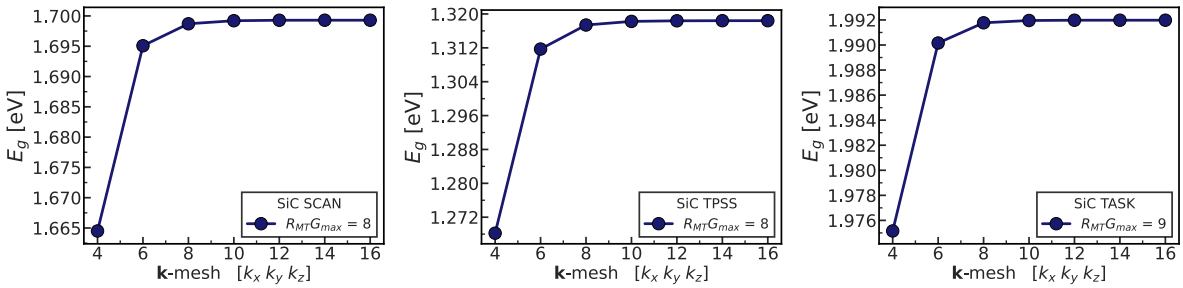



(d) Convergence behavior of the Si SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

(e) Convergence behavior of the Si TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

(f) Convergence behavior of the Si TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

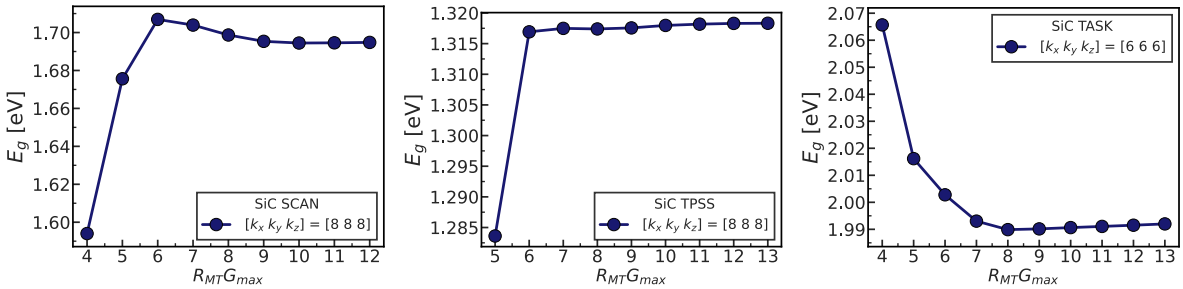
Figure B.1.: Convergence behavior of the Si band gap.




(a) Convergence behavior of the SiC SCAN band gap w.r.t. the **k**-mesh.

(b) Convergence behavior of the SiC TPSS band gap w.r.t. the **k**-mesh.

(c) Convergence behavior of the SiC TASK band gap w.r.t. the **k**-mesh.

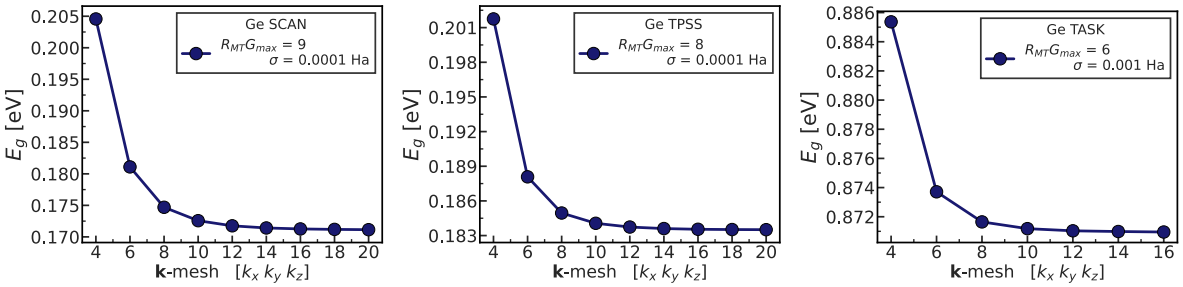



(d) Convergence behavior of the SiC SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

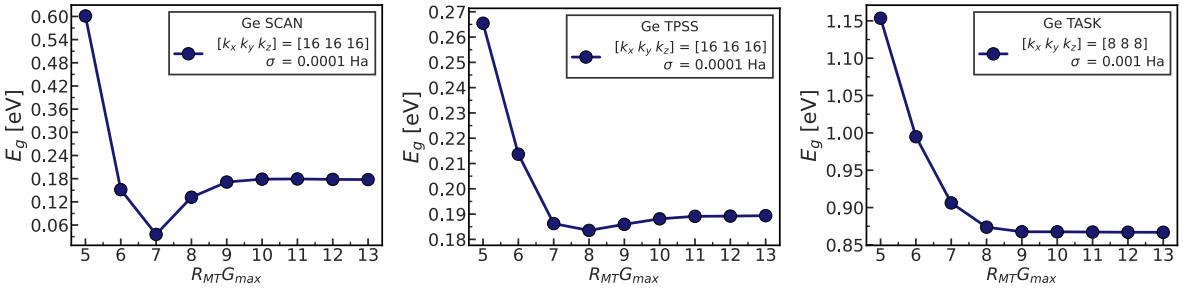
(e) Convergence behavior of the SiC TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

(f) Convergence behavior of the SiC TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

Figure B.2.: Convergence behavior of the SiC band gap.

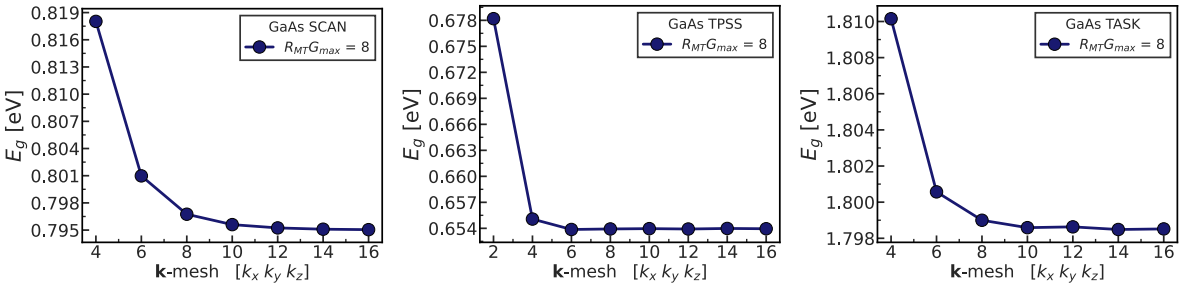



(a) Convergence behavior of the Ge SCAN band gap w.r.t. the **k**-mesh.
(b) Convergence behavior of the Ge TPSS band gap w.r.t. the **k**-mesh.
(c) Convergence behavior of the Ge TASK band gap w.r.t. the **k**-mesh.

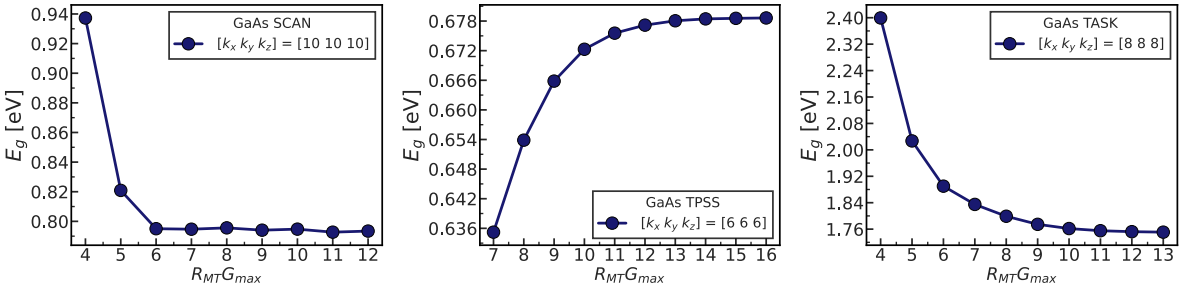



(d) Convergence behavior of the Ge SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(e) Convergence behavior of the Ge TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(f) Convergence behavior of the Ge TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

Figure B.3.: Convergence behavior of the Ge band gap.

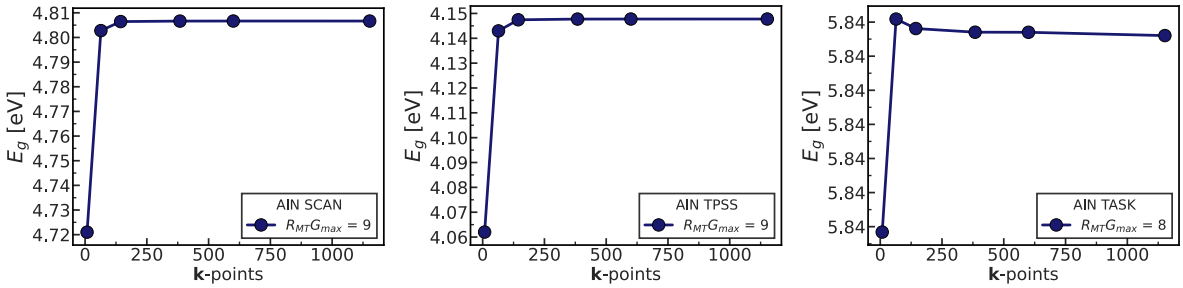



(a) Convergence behavior of the GaAs SCAN band gap w.r.t. the **k**-mesh.
(b) Convergence behavior of the GaAs TPSS band gap w.r.t. the **k**-mesh.
(c) Convergence behavior of the GaAs TASK band gap w.r.t. the **k**-mesh.

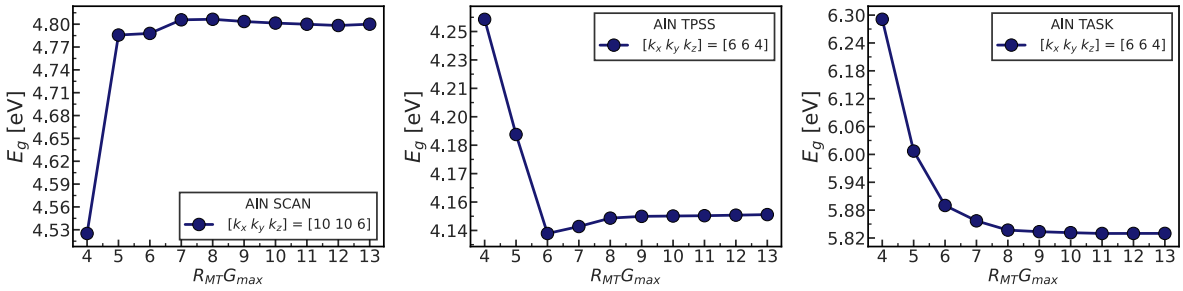



(d) Convergence behavior of the GaAs SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(e) Convergence behavior of the GaAs TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(f) Convergence behavior of the GaAs TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

Figure B.4.: Convergence behavior of the GaAs band gap.

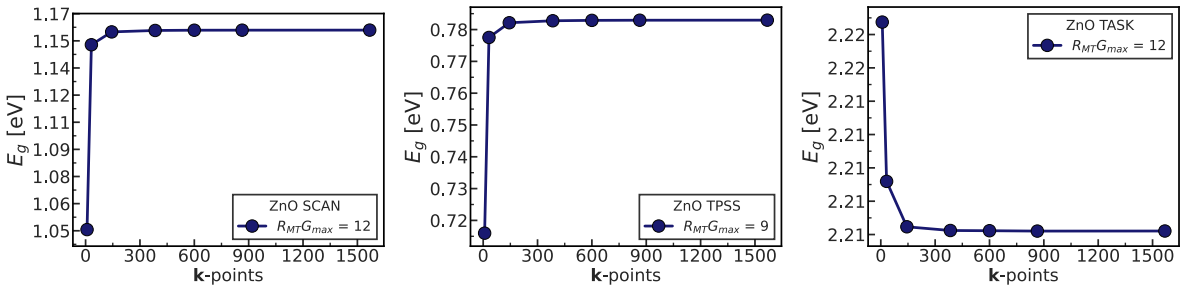



(a) Convergence behavior of the AlN SCAN band gap w.r.t. the k-mesh.
(b) Convergence behavior of the AlN TPSS band gap w.r.t. the k-mesh.
(c) Convergence behavior of the AlN TASK band gap w.r.t. the k-mesh.

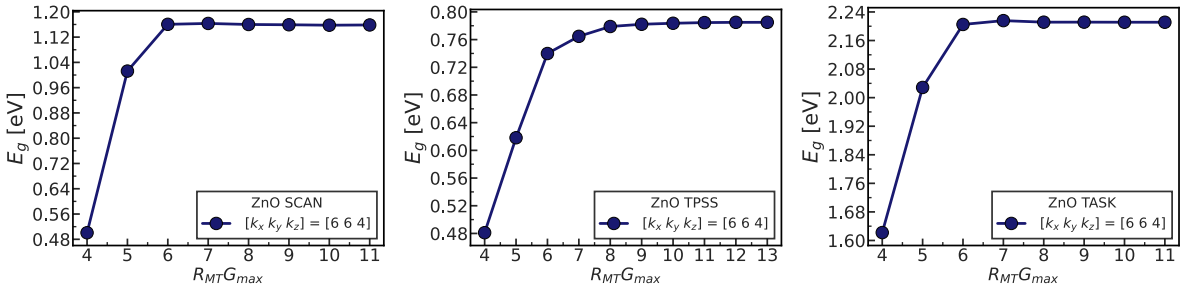



(d) Convergence behavior of the AlN SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(e) Convergence behavior of the AlN TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(f) Convergence behavior of the AlN TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

Figure B.5.: Convergence behavior of the AlN band gap.



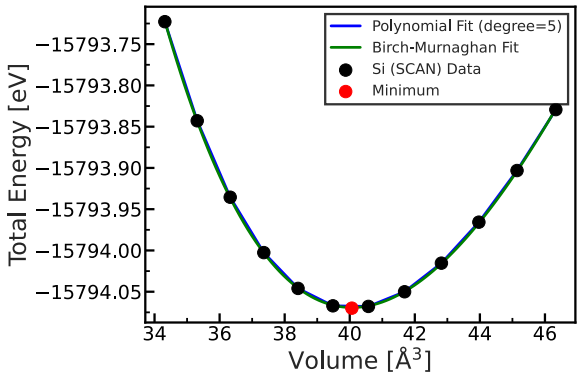

(a) Convergence behavior of the ZnO SCAN band gap w.r.t. the k-mesh.
(b) Convergence behavior of the ZnO TPSS band gap w.r.t. the k-mesh.
(c) Convergence behavior of the ZnO TASK band gap w.r.t. the k-mesh.



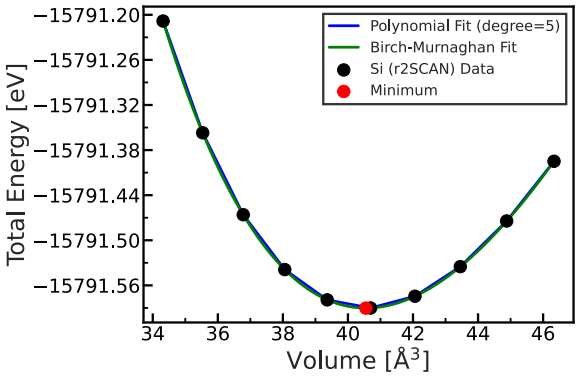

(d) Convergence behavior of the ZnO SCAN band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(e) Convergence behavior of the ZnO TPSS band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.
(f) Convergence behavior of the ZnO TASK band gap w.r.t. $R_{\mathrm{MT}}G_{\mathrm{max}}$.

Figure B.6.: Convergence behavior of the ZnO band gap.

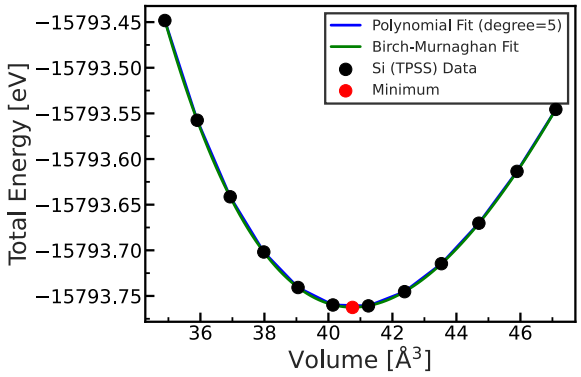
## B.2. Volume Optimization



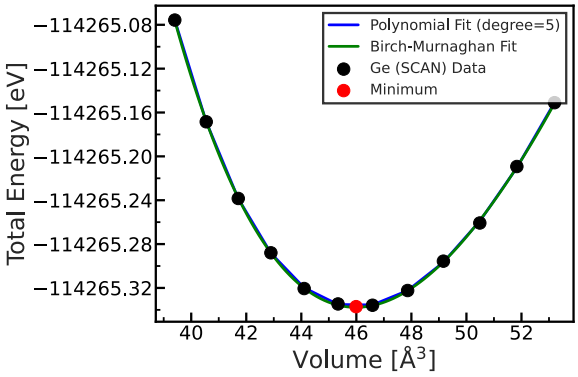

(a) Volume optimization curve $E(V)$ for Si using the SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.432$ Å.



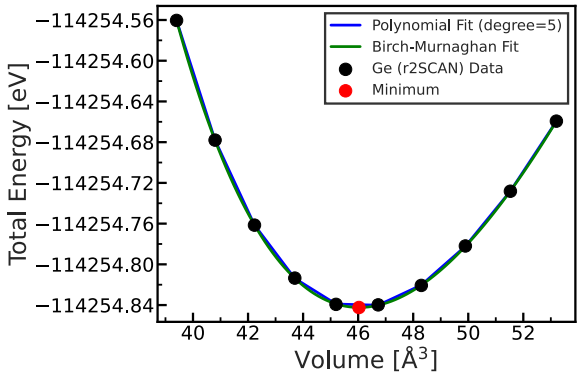

(b) Volume optimization curve $E(V)$ for Si using the r2SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.454$ Å.



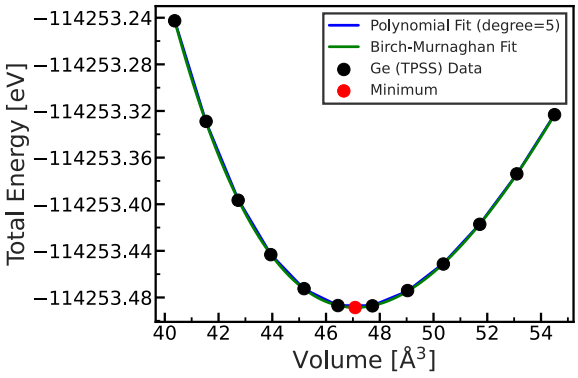

(c) Volume optimization curve $E(V)$ for Si using the TPSS functional. The fits yield an equilibrium lattice constant of $a_0 = 5.462$ Å.



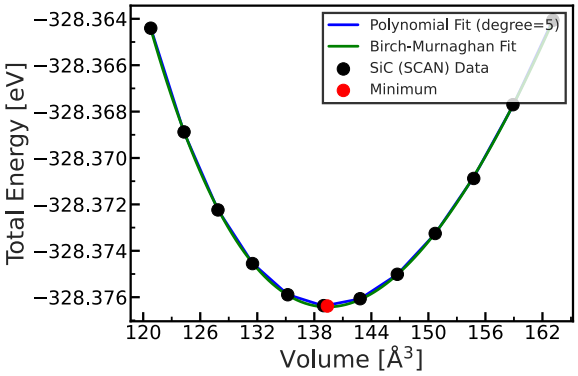

(d) Volume optimization curve $E(V)$ for Ge using the SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.687$ Å.



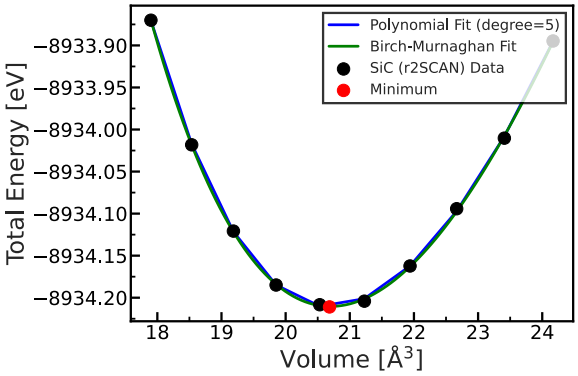

(e) Volume optimization curve $E(V)$ for Ge using the r2SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.689$ Å.



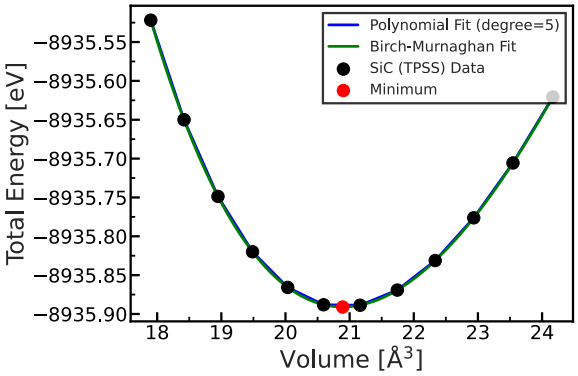

(f) Volume optimization curve $E(V)$ for Ge using the TPSS functional. The fits yield an equilibrium lattice constant of $a_0 = 5.732$ Å.



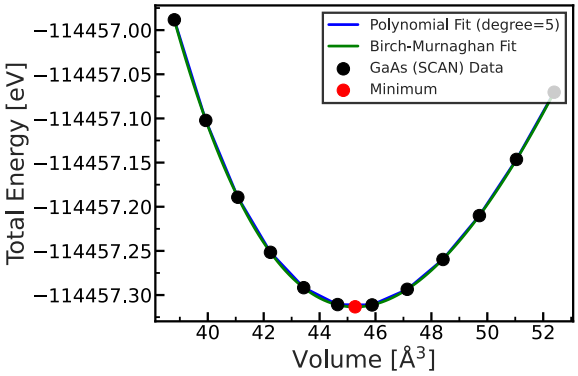

(g) Volume optimization curve $E(V)$ for SiC using the SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 4.355$ Å.



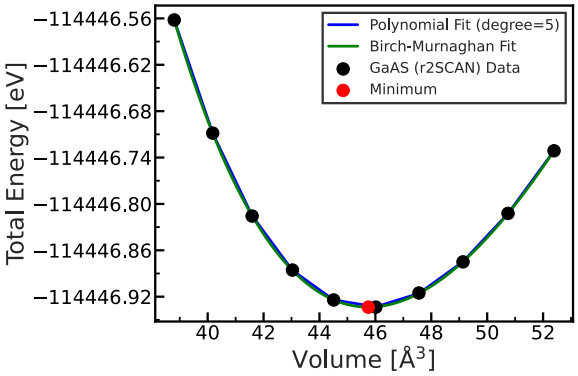

(h) Volume optimization curve $E(V)$ for SiC using the r2SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 4.359$ Å.



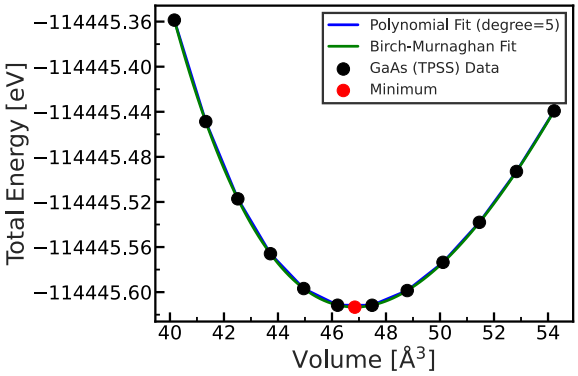

(i) Volume optimization curve $E(V)$ for SiC using the TPSS functional. The fits yield an equilibrium lattice constant of $a_0 = 4.372$ Å.




(j) Volume optimization curve $E(V)$ for GaAs using the SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.722$ Å.



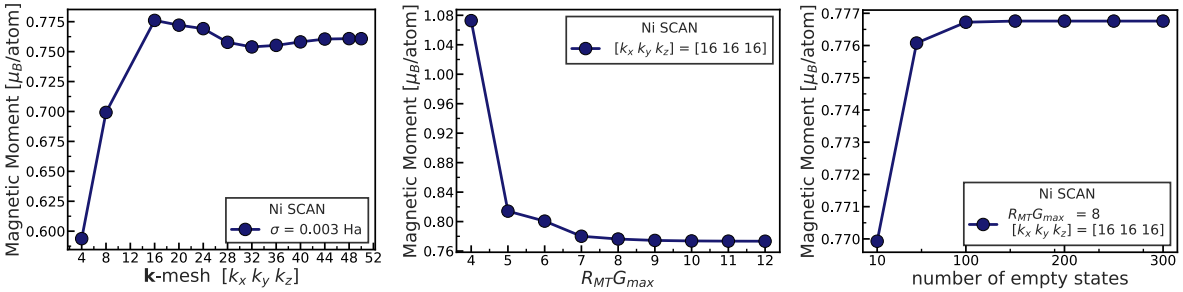

(k) Volume optimization curve $E(V)$ for GaAs using the r2SCAN functional. The fits yield an equilibrium lattice constant of $a_0 = 5.677$ Å.




(l) Volume optimization curve $E(V)$ for GaAs using the TPSS functional. The fits yield an equilibrium lattice constant of $a_0 = 5.640$ Å.

Figure B.8.: Volume Optimization curve $E(V)$ for different materials. A polynomial and Birch-Murnaghan fit are done to determine the equilibrium lattice constant $a_0$.

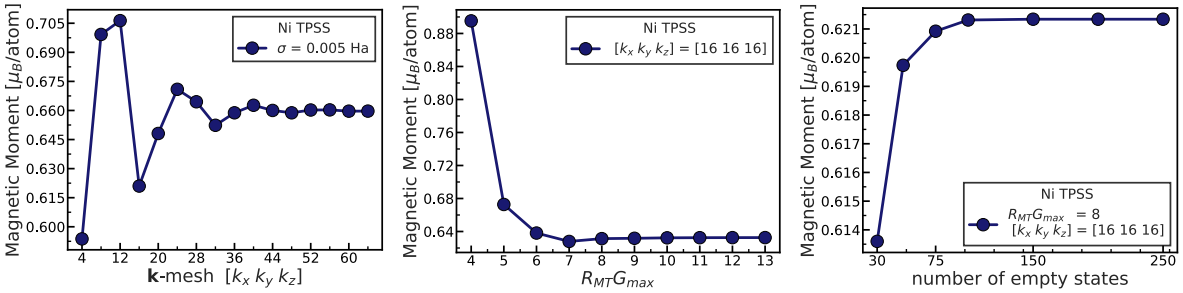
## B.3. Convergence Plots: Magnetic Moments




(a) Convergence of Ni SCAN w.r.t. the **k**-mesh with Gaussian smearing with $\sigma = 0.003$.

(b) Convergence of Ni SCAN w.r.t. the basis size $R_{\mathrm{MT}}G_{\max}$ at constant **k**-mesh = [16 16 16].

(c) Convergence of Ni SCAN w.r.t. the number of empty eigenstates.

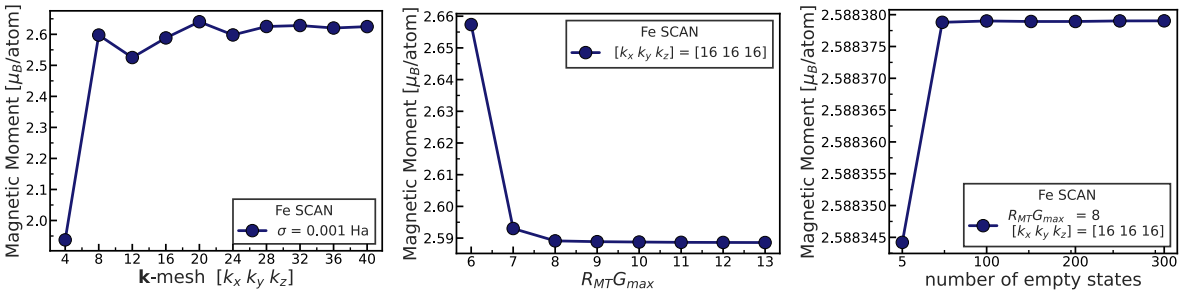



(d) Convergence of Ni TPSS w.r.t. the **k**-mesh with Gaussian smearing with $\sigma = 0.005$.

(e) Convergence of Ni TPSS w.r.t. the basis size $R_{\mathrm{MT}}G_{\max}$ at constant **k**-mesh = [16 16 16].

(f) Convergence of Ni TPSS w.r.t. the number of empty eigenstates.

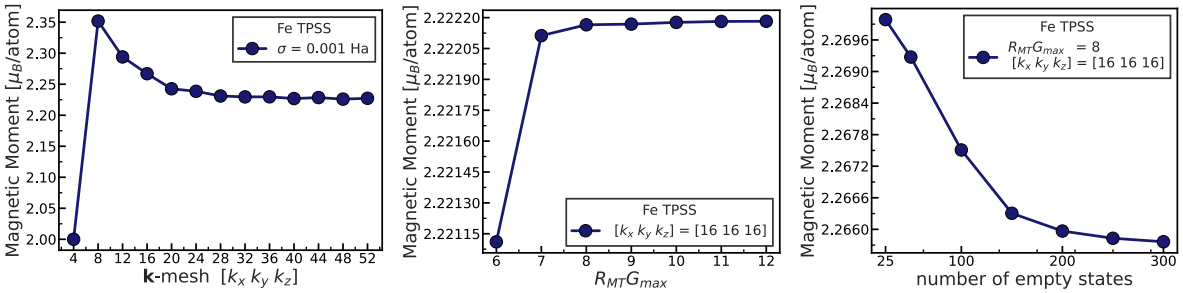
Figure B.9.: Convergence behavior of the magnetic moment for Ni.




(a) Convergence of Fe SCAN w.r.t. the **k**-mesh with Gaussian smearing with $\sigma = 0.001$.

(b) Convergence of Fe SCAN w.r.t. the basis size $R_{\mathrm{MT}}G_{\max}$ at constant **k**-mesh = [16 16 16].

(c) Convergence of Fe SCAN w.r.t. the number of empty eigenstates

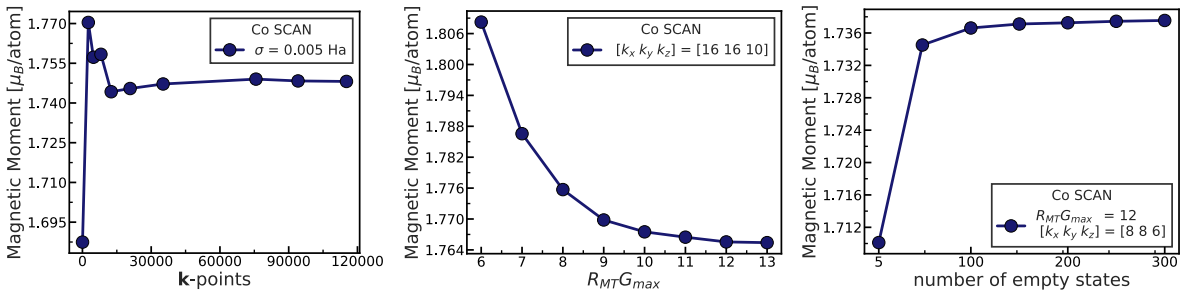



(d) Convergence of Fe TPSS w.r.t. the **k**-mesh with Gaussian smearing with $\sigma = 0.001$.

(e) Convergence of Fe TPSS w.r.t. the basis size $R_{\mathrm{MT}}G_{\max}$ at constant **k**-mesh = [16 16 16].

(f) Convergence of Fe TPSS w.r.t. the number of empty eigenstates.

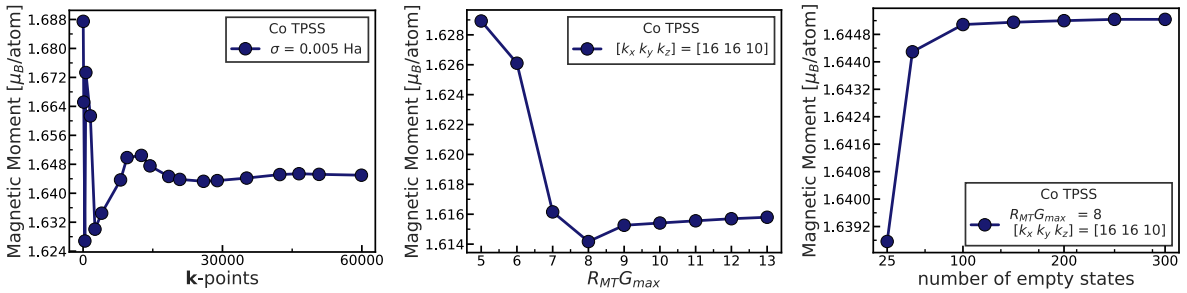
Figure B.10.: Convergence behavior of the magnetic moment for Fe.




(a) Convergence of Co SCAN w.r.t. the **k**-points with Gaussian smearing with $\sigma = 0.005$.

(b) Convergence of Co SCAN w.r.t. the basis size $R_{\mathrm{MT}}G_{\mathrm{max}}$ at constant **k**-mesh = [16 16 10].

(c) Convergence of Co TPSS w.r.t. the number of empty eigenstates.




(d) Convergence of Co TPSS w.r.t. the **k**-points with Gaussian smearing with $\sigma = 0.005$.

(e) Convergence of Co TPSS w.r.t. the basis size $R_{\mathrm{MT}}G_{\mathrm{max}}$ at constant **k**-mesh = [16 16 10].

(f) Convergence of Co TPSS w.r.t. the number of empty eigenstates.

Figure B.11.: Convergence behavior of the magnetic moment for Co.

# List of Figures

# List of Tables

# List of Abbreviations

**(L)APW** (Linearized) Augumented Plane Wave

**(L)APW+LO** (Linearized) Augumented Plane Wave + Local Orbital

**AlN** Aluminium Nitride

**APW** Augmented Plane Wave

**APW+LO** Augmented Plane Wave + Local Orbital

**BCC** Body-Centered Cubic

**BZ** Brillouin Zone

**CBM** Conduction Band Minimum

**Co** Cobalt

**DFA** Density-Functional Approximation

**DFT** Density Functional Theory

**DOS** Density of States

**EOS** Equation of State

**FCC** Face-Centered Cubic

**Fe** Iron

**FFT** Fast Fourier Transform

**GaAs** Gallium Arsenide

**Ge** Germanium

**GGA** Generalized Gradient Approximation

**gKS** generalized Kohn-Sham

**GTO** Gaussian-Type Orbital

**HDLO** High Derivative Local Orbitals

**HOMO** Highest Occupied Molecular Orbital

**HSE** Heyd Scuseria Ernzerhof

**IR** Interstitial Region

**KED** Kinetic Energy Density

**KS** Kohn-Sham

**LAPW** Linearized Augmented Plane Wave

**LDA** Local Density Approximation

**LO** Local Orbital

**LUMO** Lowest Unoccupied Molecular Orbital

**meta-GGA** meta-Generalized Density Approximation

**MT** Muffin-Tin

**Ni** Nickel

**OEP** Optimized Effective Potential

**PAW** Projector Augmented-Wave

**PBE** Perdew-Burke-Ernzerhof

**r2SCAN** re-regularized SCAN

**rSCAN** regularized SCAN

**SCAN** Strongly Constrained and Appropriately Normed

**SCF** Self-Consistent Field

**SDFT** Spin-Density-Functional-Theory

**Si** Silicon

**SiC** Silicon Carbide

**TASK** Thilo Aschebrock and Stephan Kümmel

**TPSS** Tao, Perdew, Staroverov, and Scuseria

**VBM** Valence Band Maximum

**xc** exchange-correlation

**ZnO** Zinc Oxide

**ZPAE** Zero-Point Anharmonic Expansion

**Eigenständigkeitserklärung**

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen, Bilder sowie die Nutzung von Künstlicher Intelligenz für die Erstellung von Texten und Abbildungen, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.

Berlin, 14.10.2024

Elisa Marie Stephan