JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Mortensen, Jens Jørgen; Larsen, Ask Hjorth; Kuisma, Mikael; Ivanov, Aleksei V.; Taghizadeh, Alireza; Peterson, Andrew; Haldar, Anubhab; Dohn, Asmus Ougaard; Schäfer, Christian; Jónsson, Elvar Örn; Hermes, Eric D.; Nilsson, Fredrik Andreas; Kastlunger, Georg; Levi, Gianluca; Jónsson, Hannes; Häkkinen, Hannu; Fojt, Jakub; Kangsabanik, Jiban; Sødequist, Joachim; Lehtomäki, Jouko; Heske, Julian;

Title: GPAW : An open Python package for electronic structure calculations

Year: 2024

Version: Published version

Copyright: © 2024 Author(s)

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Mortensen, J. J., Larsen, A. H., Kuisma, M., Ivanov, A. V., Taghizadeh, A., Peterson, A., Haldar, A., Dohn, A. O., Schäfer, C., Jónsson, E. Ö., Hermes, E. D., Nilsson, F. A., Kastlunger, G., Levi, G., Jónsson, H., Häkkinen, H., Fojt, J., Kangsabanik, J., Sødequist, J., . . . Thygesen, K. S. (2024). GPAW : An open Python package for electronic structure calculations. Journal of Chemical Physics, 160(9), Article 092503. https://doi.org/10.1063/5.0182685

The Journal of Chemical Physics

RESEARCH ARTICLE | MARCH 07 2024

GPAW: An open Python package for electronic structure calculations

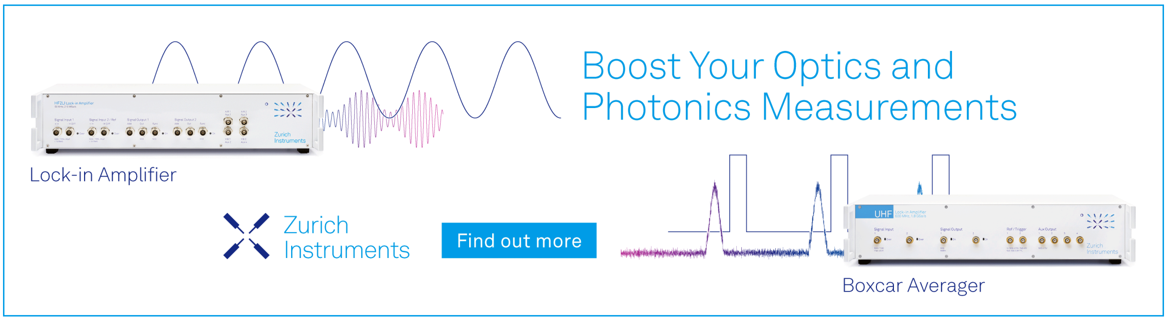
Jens Jørgen Mortensen ✉ ; Ask Hjorth Larsen ; Mikael Kuisma ; Aleksei V. Ivanov ; Alireza Taghizadeh ; Andrew Peterson ; Anubhab Haldar ; Asmus Ougaard Dohn ; Christian Schäfer ; Elvar Örn Jónsson; Eric D. Hermes ; Fredrik Andreas Nilsson ; Georg Kastlunger ; Gianluca Levi ; Hannes Jónsson ; Hannu Häkkinen ; Jakub Fojt ; Jiban Kangsabanik ; Joachim Sødequist ; Jouko Lehtomäki; Julian Heske ; Jussi Enkovaara ; Kirsten Trøstrup Winther ; Marcin Dulak ; Marko M. Melander ; Martin Ovesen ; Martti Louhivuori ; Michael Walter ; Morten Gjerding ; Olga Lopez-Acevedo ; Paul Erhart ; Robert Warmbier ; Rolf Würdemann ; Sami Kaappa ; Simone Latini; Tara Maria Boland ; Thomas Bligaard; Thorbjørn Skovhus ; Toma Susi ; Tristan Maxson ; Tuomas Rossi ; Xi Chen ; Yorick Leonard A. Schmerwitz ; Jakob Schiøtz ; Thomas Olsen ; Karsten Wedel Jacobsen ; Kristian Sommer Thygesen

Check for updates

J. Chem. Phys. 160, 092503 (2024)
https://doi.org/10.1063/5.0182685

View Online
Export Citation
CrossMark

12 March 2024 13:19:04


Boost Your Optics and Photonics Measurements
Lock-in Amplifier
Zurich Instruments
Find out more
Boxcar Averager

AIP Publishing

The Journal
of Chemical Physics

SOFTWARE

pubs.aip.org/aip/jcp

GPAW: An open Python package for electronic structure calculations

Cite as: J. Chem. Phys. 160, 092503 (2024); doi: 10.1063/5.0182685
Submitted: 20 October 2023 • Accepted: 15 January 2024 •
Published Online: 7 March 2024

Jens Jørgen Mortensen,[1,a)] Ask Hjorth Larsen,[1] Mikael Kuisma,[1] Aleksei V. Ivanov,[2] Alireza Taghizadeh,[1] Andrew Peterson,[3] Anubhab Haldar,[4] Asmus Ougaard Dohn,[5] Christian Schäfer,[6] Elvar Örn Jónsson,[7] Eric D. Hermes,[8] Fredrik Andreas Nilsson,[1] Georg Kastlunger,[9] Gianluca Levi,[7] Hannes Jónsson,[7] Hannu Häkkinen,[10] Jakub Fojt,[6] Jiban Kangsabanik,[1] Joachim Sødequist,[1] Jouko Lehtomäki,[11] Julian Heske,[1] Jussi Enkovaara,[12] Kirsten Trøstrup Winther,[13] Marcin Dulak,[1] Marko M. Melander,[14] Martin Ovesen,[1] Martti Louhivuori,[12] Michael Walter,[15] Morten Gjerding,[1] Olga Lopez-Acevedo,[16] Paul Erhart,[6] Robert Warmbier,[17] Rolf Würdemann,[18] Sami Kaappa,[19] Simone Latini,[20] Tara Maria Boland,[1] Thomas Bligaard,[21] Thorbjørn Skovhus,[1] Toma Susi,[22] Tristan Maxson,[23] Tuomas Rossi,[12] Xi Chen,[24] Yorick Leonard A. Schmerwitz,[7] Jakob Schiøtz,[1] Thomas Olsen,[1] Karsten Wedel Jacobsen,[1] and Kristian Sommer Thygesen[1]

AFFILIATIONS

[1] CAMD, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Riverlane Ltd., St Andrews House, 59 St Andrews Street, Cambridge CB2 3BZ, United Kingdom
[3] School of Engineering, Brown University, Providence, Rhode Island 02912, USA
[4] Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts 02215, USA
[5] Department of Physics, Technical University of Denmark, 2800 Lyngby, Denmark and Science Institute and Faculty of Physical Sciences, VR-III, University of Iceland, Reykjavík 107, Iceland
[6] Department of Physics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[7] Science Institute and Faculty of Physical Sciences, University of Iceland, VR-III, 107 Reykjavík, Iceland
[8] Quantum-Si, 29 Business Park Drive, Branford, Connecticut 06405, USA
[9] CatTheory, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[10] Departments of Physics and Chemistry, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland
[11] Department of Applied Physics, Aalto University, P.O. Box 11100, 00076 Aalto, Finland
[12] CSC–IT Center for Science Ltd., P.O. Box 405, FI-02101 Espoo, Finland
[13] SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA
[14] Department of Chemistry, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland
[15] FIT Freiburg Centre for Interactive Materials and Bioinspired Technologies, University of Freiburg, Georges-Köhler-Allee 105, 79110 Freiburg, Germany
[16] Biophysics of Tropical Diseases, Max Planck Tandem Group, University of Antioquia UdeA, 050010 Medellin, Colombia
[17] School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, 1 Jan Smuts Avenue, 2001 Johannesburg, South Africa
[18] Freiburger Materialforschungszentrum, Universität Freiburg, Stefan-Meier-Straße 21, D-79104 Freiburg, Germany
[19] Computational Physics Laboratory, Tampere University, P.O. Box 692, FI-33014 Tampere, Finland
[20] Nanomade, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[21] Department of Energy Conversion and Storage, Technical University of Denmark, DK-2800 Lyngby, Denmark

12 March 2024 13:19:04

© Author(s) 2024

[22]Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria
[23]Department of Chemical and Biological Engineering, The University of Alabama, Tuscaloosa, Alabama 35487, USA
[24]School of Physical Science and Technology, Lanzhou University, Lanzhou, Gansu 730000, China

[a)]**Author to whom correspondence should be addressed:** jjmo@dtu.dk

ABSTRACT

We review the GPAW open-source Python package for electronic structure calculations. GPAW is based on the projector-augmented wave method and can solve the self-consistent density functional theory (DFT) equations using three different wave-function representations, namely real-space grids, plane waves, and numerical atomic orbitals. The three representations are complementary and mutually independent and can be connected by transformations via the real-space grid. This multi-basis feature renders GPAW highly versatile and unique among similar codes. By virtue of its modular structure, the GPAW code constitutes an ideal platform for the implementation of new features and methodologies. Moreover, it is well integrated with the Atomic Simulation Environment (ASE), providing a flexible and dynamic user interface. In addition to ground-state DFT calculations, GPAW supports many-body GW band structures, optical excitations from the Bethe–Salpeter Equation, variational calculations of excited states in molecules and solids via direct optimization, and real-time propagation of the Kohn–Sham equations within time-dependent DFT. A range of more advanced methods to describe magnetic excitations and non-collinear magnetism in solids are also now available. In addition, GPAW can calculate non-linear optical tensors of solids, charged crystal point defects, and much more. Recently, support for graphics processing unit (GPU) acceleration has been achieved with minor modifications to the GPAW code thanks to the CuPy library. We end the review with an outlook, describing some future plans for GPAW.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0182685

12 March 2024 13:19:04

I. INTRODUCTION

The electronic-structure (ES) problem, i.e., the solution to the time-independent Schrödinger equation for a collection of electrons and atomic nuclei, forms the starting point for the quantum-mechanical treatment of matter. Indeed, all chemical and physical properties of any substance (solid, molecule, surface, etc.) can, in principle, be obtained from the energies and wave functions that constitute the solution. A pioneering step towards solving the many-body ES problem was the formulation and formal proof of density functional theory (DFT) by Hohenberg and Kohn in 1964[1] and a practical scheme for its solution by Kohn and Sham in 1965.[2] At present, most codes solving the ES problem from first principles are based on DFT. Such codes are extremely powerful and allow one to determine the atomic structure of solids and molecules containing hundreds of atoms with a relative error below 1%.[3–5] Once the atomic structure of the compound has been solved, its properties (electronic, magnetic, optical, topological, etc.) can, in principle, be determined. The evaluation of properties often involves theories beyond the formal DFT framework to account for effects such as temperature and lattice vibrations,[6,7] many-body interactions in excited states,[8,9] or time dependence.[10,11] As such, first-principles atomistic calculations often involve two successive phases: the solution of the ground-state ES problem (including ion dynamics) and the subsequent evaluation of physical properties. This review is structured accordingly, as Secs. III and IV deal with the first phase while Secs. V–IX are devoted to the second.

In recent years, the scientific significance of ES codes has shifted from a useful tool to describe and understand matter at the atomic scale to an independent driver of the discovery and development of new materials.[12–16] This change in scope has been fueled by the exponential increase in computer power accompanied by improved numerical algorithms[17,18] as well as the use of workflow management software for high-throughput computations[19–22] and the adoption of machine-learning techniques to leverage the rapidly growing data generated by ES codes.[23–25] In parallel with these capacity-extending developments, continuous progress in the fundamental description of exchange–correlation effects has advanced the predictive power of ES calculations to a level where they rival experiments in terms of accuracy for many important properties.[26–31]

The grid-based projector-augmented wave (GPAW) code was originally intended as a Python-based multigrid solver of the basic DFT equations within the projector-augmented wave (PAW) formalism.[32] The name GPAW, accordingly, was an abbreviation for "grid-based projector-augmented wave." At present, other choices than regular grids for representations of the wave functions exist in GPAW, but the name has stuck. During the years 2005–2010, GPAW evolved to a full-blown DFT package[33] supporting most of the functionality expected from a modern ES code, in addition to a few more specialized features, including real-time propagation of wave functions[34] and an alternative basis of numerical atomic orbitals [referred to as the linear combination of atomic orbitals (LCAOs) basis][35] to supplement the real-space grid. In 2011, a plane-wave (PW) basis set was also implemented. At present, the possibility to use three different types of basis sets and even combine them within a single run remains a unique feature of GPAW, rendering the code very versatile.

The implementation of the PW basis set laid the groundwork for GPAW's linear-response module, which at present supports the calculation of linear response functions,[36] total energies from the adiabatic connection fluctuation–dissipation theorem,[29,37] the GW self-energy method for quasiparticle band structures,[38] the

© Author(s) 2024

Bethe–Salpeter Equation (BSE) for optical excitations,[39] and more. The code also supports a wide range of features related to the calculation of magnetism and spin–orbit effects. Examples include spin-spiral calculations using the generalized Bloch theorem,[40] external magnetic fields, orbital magnetization, magnetic anisotropy,[41] adiabatic magnon dispersions from the magnetic force theorem,[42] and dynamic magnetic response from time-dependent density functional theory (TDDFT).[43] For solids, the **k**-space Berry phases can be computed directly from the Bloch orbitals and may be used to obtain spontaneous polarization,[44] Born effective charges, piezoelectric response tensors,[45] and various indices characterizing the band topology.[46]

In addition, GPAW can compute the localization matrices forming the basis for the construction of Wannier functions with, e.g., the Atomic Simulation Environment (ASE)[47] or Wannier90.[48] Electrostatic corrections to the formation energies of charged point defects in insulators are implemented, as are calculations of the hyperfine coupling and zero-field splitting for localized electron spins. GPAW also offers the possibility to perform time-independent, variational calculations of localized electronic excitations in, e.g., molecules or crystal point defects, using direct orbital optimization strategies implemented for all three types of basis sets.[49–51] This provides an efficient and robust alternative to traditional "ΔSCF" approaches. GPAW can also be used to describe ultrafast electron dynamics within time-dependent density functional theory (TDDFT), with wave functions represented either on a real-space grid[34] or in the LCAO basis.[52] The latter can provide a significant speed-up due to the relatively small size of the basis.[53–55] The LCAO representation also forms the basis for the calculation of electron–phonon couplings as well as non-linear optical spectra such as Raman scattering[56] (which can alternatively be obtained in the PW mode as a finite difference of the dielectric tensor), second-harmonics generation,[57] and shift currents[58] using higher-order perturbation theory.

II. WHY GPAW?

A. User's perspective

There are dozens of electronic-structure codes available for interested users. The codes differ in their license (in particular, whether open or proprietary), the underlying programming language (e.g., Fortran, C, and Python), their treatment of core electrons (all-electron vs pseudopotentials), the employed representations of the wave functions (plane waves, atom-centered orbitals, and real-space grids), and the beyond-DFT features they support. Why should one choose GPAW?

In this section, we describe some of the features that make GPAW interesting from the point of view of a common user who wants to perform electronic-structure calculations. Section II B focuses on its possibilities for more advanced users, who may want to modify the code or implement completely new functionalities.

A first point to note is that GPAW is written almost exclusively in Python and is directly integrated with the Atomic Simulation Environment. This integration with ASE makes the setup, control, and analysis of calculations easy and flexible. The programming language is of course a key issue for developers, but the common user also benefits from Python and the ASE/GPAW integration. A typical stand-alone program only offers a fixed (though, of course, possibly large) set of tasks that it can perform, while Python scripting allows for a more flexible use of the code. This could, for example, mean combining several different GPAW calculations in new ways. Another advantage is that "inner parts" of the code like the density or the Kohn–Sham eigenvalues are directly accessible in a structured format within Python for further analysis. It is even possible to "open up the main loop" of GPAW and have access to, inspect, and also modify key quantities during program execution (see Fig. 1).

As already mentioned in the introduction, GPAW distinguishes itself from other available ES codes by supporting three different ways of representing the wave functions. The most commonly used basis set is plane waves (PW), which is appropriate for small or medium-sized systems where high precision is required. Convergence is easily and systematically controlled by tuning the cutoff energy. A large number of advanced features and "beyond-DFT" methods are available in the PW mode. These include the calculation of hybrid functionals, RPA total energies, linear-response TDDFT, and many-body perturbation theory techniques like GW and the Bethe–Salpeter equations. The new graphics processing unit (GPU) implementation also uses the PW mode.

The wave functions can alternatively be represented on real-space grids, which was the original approach in GPAW. The implementation of this so-called finite-difference (FD) mode relies on multi-grid solutions of the Poisson and Kohn–Sham equations. The FD mode allows for more flexible boundary conditions than the PW mode, which is restricted to periodic supercells. The boundary conditions may, for example, be taken to reflect the charge distribution in the unit cell. Calculations in the FD mode can be systematically converged through lowering the grid spacing, but the approach to full convergence is slower than in the PW mode. The FD mode is particularly well suited for large systems because the wave-function representation allows for large-scale parallelization through real-space decomposition. Furthermore, it is possible to

```
import os
import psutil
process = psutil.Process(os.getpid())
atoms = ...
calc = ...
for ctx in calc.icalculate(atoms):
    # Inside SCF loop
    mem = process.memory_info().rss
    # Write memory usage to log-file:
    ctx.log(f'MEM: {ctx.niter} {mem}')
    if ctx.niter == 15:
        # Stop after 15 iterations
        break
```

FIG. 1. The variable `calc` is the ground-state DFT calculator object, and its `icalculate` method yields a context object at every self-consistent field (SCF) step. As seen, one can use this in a for-loop to implement special logic for the termination of the SCF iterations or for diagnostics. In this example, the memory usage is written to the log-file for the first 15 SCF iterations.

12 March 2024 13:19:04

© Author(s) 2024

perform time-propagation TDDFT, including Ehrenfest dynamics, in this mode.

The third representation of the wave functions is a basis of numerical atom-centered orbitals in the linear combination of atomic orbitals (LCAOs) mode. The size of the basis set can be varied through the inclusion of more angular momentum channels, additional orbitals within a channel, or polarization functions. GPAW comes with a standard set of orbitals, but a basis-set generator is included with the code so that users may construct different basis sets depending on their needs and requirements. The LCAO mode is generally less accurate than the PW and FD modes, but it allows for the treatment of considerably larger systems—more than ten thousand atoms. It is also possible to study electron dynamics through the fast implementation of time-propagation DFT, and Ehrenfest dynamics is under development.

As explained, the different modes have different virtues and limitations, and it can, therefore, be an advantage to apply several modes to a project. For larger systems, it is, for example, possible to divide structure optimization into two steps. First, an optimization is performed with the fast LCAO basis, leading to an approximately correct structure. This is then followed by an optimization in either the PW or FD mode, which now requires much fewer steps because of the good initial configuration. Due to the ASE/Python interface, this combined calculation can easily be performed within a single script.

Since GPAW was originally created with the FD mode only, and the LCAO mode was added next, some features have been implemented for only those modes. Examples are real-time TDDFT (see Sec. VII) and electron–phonon coupling (see Sec. VI G). Conversely, some new features only work for the PW mode, which was added after the real-space modes. Examples are RPA total energies (see Sec. VI C) and the calculation of the stress tensor. To summarize, given the limitations just mentioned, users should most of the time use PW or LCAO mode, and the choice will depend on the accuracy needed and the resources available. The GPAW webpage has a table showing up-to-date information about which features work with which modes.[59]

B. Developer's perspective

The GPAW source code is written in the Python and C languages and is hosted on GitLab,[60] licensed under the GNU General Public License v3.0. This ensures the transparency of all features and allows developers to fully customize their experience and contribute new features to the community.

An advantage of having Python code is that the Python script you write to carry out your calculations will have access to everything inside a GPAW calculation. An example showing the power and flexibility this affords is the possibility of having user-code inserted inside the self-consistent field (SCF) loop, as demonstrated in Fig. 1.

At the time of this writing (July 2023), GPAW has two versions of the ground-state DFT code in the main branch of the code. There is the older version that has grown organically since the birth of GPAW: it has many features but also a lot of technical debt that makes it harder to maintain and less ideal to build new features on top of. The newer ground-state code addresses these issues by having a better overall design.

The new design greatly improves the ease of implementing new features. The goal is to have the new code be feature-complete so that it can pass the complete test suite and then delete the old code once that is achieved. At the moment, we recommend that all production calculations be performed with the old code and that work on new features be performed on top of the new code, even though certain features are not yet production-ready. Three new features, not present in the old code base, have already been implemented based on the new code: GPU implementation of PW mode calculations (see Sec. III B 9), reuse of the wave functions after unit-cell changes during cell optimization, and spin-spiral calculations (see Sec. V E).

GPAW uses pytest[61] for its test suite, which currently consists of ∼1600 unit and integration tests (see Table I). A subset of those tests runs as part of GitLab's continuous integration (CI), thereby checking the correctness of every code change. Unfortunately, the full test suite is too time-consuming to run as part of CI, so we run that nightly both in serial as well as in parallel using MPI.

Many of the code examples in GPAW's documentation, exercises, and tutorials[62] require resources [time and number of central processing units (CPUs)] beyond what would make sense to run as part of the pytest test suite. For that, we use MyQueue[21] to submit those scripts as jobs to a local supercomputer every weekend. At the moment, this amounts to ∼5200 core-hours of calculations.

As can be seen from Table I, the majority of the code is written in Python, which is an interpreted language that is easy to read, write, and debug.

Interpreter-executed code will not run as efficiently as code that is compiled into native machine code. It is, therefore, important to make sure that the places in the code where most of the time is spent (hot spots) are in native machine code and not in the interpreter. GPAW achieves this by implementing the hot spots in C-code with Python wrappers that can be called from the Python code. Examples of such computationally intensive tasks are applying a finite-difference stencil to a uniform grid, interpolating from one uniform grid to another, or calculating overlaps between projector functions and wave functions. In addition, we have Python interfaces to the numerical libraries FFTW,[63] ScaLAPACK,[64] ELPA,[17] BLAS, Libxc,[65,66] libvdwxc,[67] and MPI. Finally, GPAW makes heavy use of the NumPy[68] and Scipy[69] Python packages. NumPy provides us with the `numpy.ndarray` data type, which is an *N*-dimensional array that we use for storing wave functions, electron densities, potentials, matrices like the overlap matrix or the LCAO wave function coefficients, and much more. The use of NumPy arrays allows us to use the many sub-modules of SciPy to manipulate data. This also gives us an efficient memory layout, allowing us to simply pass a pointer to the memory whenever we need to call the C-code from the Python code.

TABLE I. Number of files and number of lines of code in the git repository of GPAW. The Python source-code files are split into three parts: the actual code, the test suite, and code examples in the documentation.

Type of file	Files	Lines
Python (the code)	513	146 604
C	80	19 719
Python (test suite)	681	47 147
Python (documentation)	744	32 014

12 March 2024 13:19:04

© Author(s) 2024

With this strategy, we can get away with having most of the code written in a relatively slow interpreted language and still have most of the time spent in highly efficient C-code or optimized numerical libraries.

The advantage of the original FD-mode, where there are no Fourier transforms of the wave functions, is that the algorithms should parallelize well for large systems. In practice, it has turned out that FD-mode has a number of disadvantages: (1) Due to integrals over the unit cell being performed as sums over grid-points, there will be a small periodic energy variation as you translate atoms, and the period of the variation will be equal to the grid-spacing used (the so-called egg-box error); (2) The system sizes that are typically most interesting for applications of DFT are too small for the parallel scalability to be the decisive advantage; (3) The memory used to store the wave functions on uniform grids in real-space is significant. In contrast, the PW mode has practically no egg-box error, is very efficient for the most typical system sizes, and often uses a factor of 10 less memory compared to an FD mode calculation of similar accuracy. The main advantages of LCAO mode are low memory usage and high efficiency for large systems; for small unit cells with many **k**-points, the PW mode is the most efficient. One disadvantage of LCAO-mode is egg-box errors: LCAO-mode uses the same uniform grids as used in FD-mode for integration of matrix elements like $\langle \Phi_\mu | \tilde{v} | \Phi_\nu \rangle$ and, therefore, has similar egg-box energy variation. A second disadvantage of LCAO-mode is that, as for any localized basis-set, reaching the complete basis-set limit is more involved compared to PW and FD modes. This can have severe consequences, even for ground-state calculations of difficult systems such as Cr_2, for example.[70] In PW or FD modes, the complete basis set limit is easy to reach by simply increasing the number of plane-waves or grid-points, respectively, which leads to smooth convergence.[32] At the moment, we only provide double-ζ polarized (DZP) basis sets, and going beyond DZP is left for users to do themselves.

III. GROUND-STATE DFT

A. Projector augmented-wave method

The diverging Coulomb potential causes rapid oscillations in electronic wave functions near the nuclei, and special care is required to be able to work with smooth wave functions. The projector augmented-wave (PAW) method by Blöchl[71] is a widely adopted generalization of pseudopotential methods, utilizing their strength of smooth pseudo-wave functions while retaining a mapping from all-electron wave functions ($|\psi_n\rangle$) to pseudo-wave functions ($|\tilde{\psi}_n\rangle$).

The crux is to define a linear transformation $\hat{T}$ from pseudo to all-electron space,

$$\hat{T}|\tilde{\psi}\rangle = |\psi\rangle, \tag{1}$$

where

$$\hat{T} = 1 + \sum_a \sum_i \left(|\phi_i^a\rangle - |\tilde{\phi}_i^a\rangle \right) \langle \tilde{p}_i^a |. \tag{2}$$

Here, $\tilde{p}_i^a(\mathbf{r})$, $\phi_i^a(\mathbf{r})$ and $\tilde{\phi}_i^a(\mathbf{r})$ are called projectors, partial waves, and pseudo-partial waves, respectively. The pseudo-partial waves and projectors are chosen to be biorthogonal, $\int d\mathbf{r}\, \tilde{\phi}_i^a(\mathbf{r}) \tilde{p}_j^a(\mathbf{r}) = \delta_{ij}$, allowing for an approximate closure relation inside the augmentation sphere

$$\sum_i \tilde{\phi}_i^a(\mathbf{r}) \tilde{p}_i^a(\mathbf{r}') \approx \delta(\mathbf{r} - \mathbf{r}'), \tag{3}$$

which is utilized heavily to obtain an efficient, but all-electron, picture. In addition to biorthogonality, the pseudo and all-electron partial waves are chosen to be equal outside the PAW augmentation sphere cutoff radius $\tilde{\phi}_i(\mathbf{r}) = \phi_i(\mathbf{r})$, $r > r_c$.

The basic recipe for converting an operator $\hat{O}$ is $\langle \psi_n | \hat{O} | \psi_n \rangle = \langle \tilde{\psi}_n | \underbrace{\hat{T}^\dagger \hat{O} \hat{T}}_{\hat{\tilde{O}}} | \tilde{\psi}_n \rangle$. For example, the all-electron Kohn–Sham equations

$$\hat{H}[n(\mathbf{r})] |\psi_n\rangle = \epsilon_n |\psi_n\rangle, \tag{4}$$

where $\hat{H}$ is the single-particle all-electron Kohn–Sham Hamiltonian operator, can be transformed to their PAW counterparts,

$$\hat{T}^\dagger \hat{H}[n(\mathbf{r})] \hat{T} |\tilde{\psi}_n\rangle = \epsilon_n \hat{T}^\dagger \hat{T} |\tilde{\psi}_n\rangle. \tag{5}$$

We have used Eq. (1) and also multiplied with $\hat{T}^\dagger$ from left to make its dual space the pseudo one, i.e., $\langle \tilde{\psi} | \in \mathcal{H}^*$ can act from left. This results in a PAW Hamiltonian and in PAW overlap operators as follows:

$$\begin{aligned} \hat{\tilde{H}} &= \hat{T}^\dagger \hat{H}[n(\mathbf{r})] \hat{T} \\ &= -\frac{1}{2}\nabla^2 + \tilde{v}_{\mathrm{KS}}(\mathbf{r}) + \sum_a \sum_{ii'} |\tilde{p}_i^a\rangle \Delta H_{ii'}^a \langle \tilde{p}_{i'}^a |, \end{aligned} \tag{6}$$

$$\hat{\tilde{S}} = \hat{T}^\dagger \hat{T} = 1 + \sum_a \sum_{ii'} |\tilde{p}_i^a\rangle \Delta S_{ii'}^a \langle \tilde{p}_{i'}^a |. \tag{7}$$

Terms such as $\Delta H_{ii'}^a$ and $\Delta S_{ii'}^a$ represent so-called PAW corrections. In each part of the description that handles a particular kind of operator, such as kinetic energy, spin-operators, or electrostatic potential, the respective PAW corrections must be calculated. The most crucial ones are precalculated, such as overlap, kinetic energy, and Coulomb, and stored in the "setup" file, which also stores the partial waves and projectors. As an example, the overlap PAW corrections are precalculated to setup as follows:

$$\Delta S_{ii'}^a = \int d\mathbf{r}\, \phi_i^a(\mathbf{r}) \phi_{i'}^a(\mathbf{r}) - \int d\mathbf{r} \tilde{\phi}_i^a(\mathbf{r}) \tilde{\phi}_{i'}^a(\mathbf{r}). \tag{8}$$

We further define the atomic density matrices as

$$D_{\sigma ii'}^a = \sum_n f_n \langle \tilde{\psi}_{\sigma n} | \tilde{p}_i^a \rangle \langle \tilde{p}_{i'}^a | \tilde{\psi}_{\sigma n} \rangle. \tag{9}$$

The atomic density matrix contains all the information required to construct PAW corrections to any local all-electron expectation value,

$$\langle O \rangle = \sum_n \langle \tilde{\psi}_n | \tilde{O} | \tilde{\psi}_n \rangle + \sum_a \sum_{ii'} D_{ii'}^a \Delta O_{i'i}^a. \tag{10}$$

The all-electron atomic density can be constructed as

$$n_\sigma^a(\mathbf{r}) = \sum_{ii'} D_{\sigma ii'}^a \phi_i^{a*}(\mathbf{r}) \phi_{i'}^a(\mathbf{r}) + n_{\mathrm{core}}^a(\mathbf{r}), \tag{11}$$

and the corresponding equation for pseudo-densities holds with $n \to \tilde{n}$ and $\phi \to \tilde{\phi}$.

12 March 2024 13:19:04

© Author(s) 2024

Since the exchange–correlation (xc) potential is non-linear, the PAW corrections must be evaluated explicitly. The xc PAW corrections are performed by constructing the atomic all-electron and pseudo-electron densities as given by Eq. (11):

$$\Delta H^a_{\mathrm{xc},ii'} = \int d\mathbf{r}\phi^a_i(\mathbf{r})v_{\mathrm{xc}}[n^a](\mathbf{r})\phi^a_{i'}(\mathbf{r}) - \int d\mathbf{r}\tilde{\phi}^a_i(\mathbf{r})v_{\mathrm{xc}}[\tilde{n}^a](\mathbf{r})\tilde{\phi}^a_{i'}(\mathbf{r}). \tag{12}$$

This integral is numerically evaluated in the Cartesian product of a Lebedev angular grid and a non-uniform radial grid (a denser mesh closer to the nucleus) for each atom.

B. Numerical implementation

1. Wave-function representations

GPAW supports three representations for smooth wave functions. The plane wave (PW)

$$\tilde{\psi}_{\mathbf{k}n}(r) = \sum_{\mathbf{G}} C_{n\mathbf{k}\mathbf{G}}e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{13}$$

and linear combination of atomic orbitals (LCAOs)

$$\tilde{\psi}_{\mathbf{k}n}(r) = \sum_{\mu} C_{n\mathbf{k}\mu}\Phi_{\mathbf{k}\mu}(\mathbf{r}), \tag{14}$$

representations rely on basis functions. The finite difference (FD) mode relies on a representation of the kinetic energy operator on a uniform Cartesian grid.

2. All-electron quantities

The beauty of the PAW method is that you never *need* to transform the pseudo-wave functions to all-electron wave functions, but you *can* do it if you want to. GPAW has tools for interpolating the pseudo-wave functions to a fine real-space grid and adding the PAW corrections. A fine real-space grid is needed to properly represent the cusp and all the oscillations necessary for the wave function to be orthogonal to all the frozen core states.

GPAW also has tools for calculating the all-electron electrostatic potential. This is useful for transmission electron microscopy (TEM) simulations.[72] Most TEM simulations have relied on the so-called independent atom model (IAM), where the specimen potential is described as a superposition of isolated atomic potentials. While this is often sufficient, there is increasing interest in understanding the influence of valence bonding.[73] This can be investigated by a TEM simulation code such as *ab*TEM,[74] which can directly use *ab initio* scattering potentials from GPAW.

3. Solving the Kohn–Sham equation

The default method for solving the Kohn–Sham equation for PW and FD modes is to do iterative diagonalization combined with density mixing; for LCAO mode, we do a full diagonalization of the Hamiltonian. Alternatively, one can do direct minimization as described in Sec. III B 5.

For PW and FD modes, we need an initial guess of the wave functions. For this, we calculate the effective potential from a superposition of atomic densities and diagonalize an LCAO Hamiltonian in a small basis set consisting of all the pseudo-partial waves corresponding to bound atomic states.

Each step in the self-consistent field (SCF) loop consists of the following operations: (1) diagonalization of the Hamiltonian in the subspace of the current wave functions (skipped for LCAO); (2) one or more steps through the iterative eigensolver (except for LCAO, where a full diagonalization is performed); (3) update of eigenvalues and occupation numbers; and (4) density mixing and symmetrization. See previous work[32,33] and Ref. 75 for details.

GPAW has two kinds of Poisson equation solvers: direct solvers based on Fourier transforms or Fourier-sine transforms, and iterative multi-grid solvers (Jacobi or Gauss–Seidel). The default is to use a direct solver, whereas iterative solvers may be chosen for larger systems where they can be more efficient.

For 0-, 1-, and 2-dimensional systems, the default boundary conditions are to have the potential go to zero on the cell boundaries. This becomes a problem for systems involving large dipole moments. The potential due to the dipole is long-ranged and, therefore, the converged potential requires large vacuum sizes. For molecules (0D systems), the boundary conditions can be improved by adding multipole moment corrections to the density so that the corresponding multipoles of the density vanish. The potential of these corrections is added to the obtained potential. The same trick is used to handle charged systems. For slabs (2D systems), a dipole layer can be added to account for differences in the work functions on the two sides of the slab.

Methods for calculating occupation numbers are the Fermi–Dirac, Marzari–Vanderbilt,[76] and Methfessel–Paxton distributions, as well as the tetrahedron method and the improved tetrahedron method.[77]

4. Updating wavefunctions in dynamics

Simulations commonly move the atoms without changing other parameters. If an atom moves only slightly, we would expect most of the charge in its immediate vicinity to move along with it. We use this to compute an improved guess for the wave functions in the next self-consistency loop with FD or PW mode, where the eigensolver is iterative.

Near the atoms, the dual basis of pseudo-partial waves and projectors is almost complete, i.e.,

$$\sum_i |\phi^a_i\rangle\langle\tilde{p}^a_i| = 1 \quad (\text{near atom } a). \tag{15}$$

If an atom moves by $\Delta\mathbf{R}^a$, the wave functions $\tilde{\psi}_n(\mathbf{r})$ are updated by rigidly moving the projection $\sum_i \tilde{\phi}^a_i(\mathbf{r})\langle\tilde{p}^a_i|\tilde{\psi}_n\rangle$ along with it, i.e.,

$$\tilde{\psi}^{\mathrm{new}}_n(\mathbf{r}) = \tilde{\psi}_n(\mathbf{r}) + \sum_{ai} \tilde{\phi}^a_i(\mathbf{r} - \Delta\mathbf{R}^a)\langle\tilde{p}^a_i|\tilde{\psi}_n\rangle - \sum_{ai} \tilde{\phi}^a_i(\mathbf{r})\langle\tilde{p}^a_i|\tilde{\psi}_n\rangle. \tag{16}$$

As the partial waves on different atoms are not orthonormal, this expression generally "double-counts" contributions, resulting in wave functions that are to some extent unphysical. Nevertheless, we have found that this simple method achieves a significant speedup (~15% in realistic structure optimizations) compared to not updating the wave functions.

The method could be further improved by using the LCAO basis set and the overlap matrix to prevent double-counting.

5. Direct minimization

Direct orbital minimization[78–81] is a robust alternative to the conventional eigensolver and density mixing routines. The orbitals

12 March 2024 13:19:04

© Author(s) 2024

can be expressed as a unitary transformation of a set of reference, or auxiliary, orbitals $\mathbf{\Psi}_0$,

$$\mathbf{\Psi} = \mathbf{U}\mathbf{\Psi}_0. \tag{17}$$

In the direct minimization (DM) method implemented in GPAW,[50,82] the unitary matrix $\mathbf{U}$ is parametrized as an exponential transformation, i.e., $\mathbf{U} = e^{\mathbf{A}}$, where $\mathbf{A}$ is an anti-Hermitian matrix $(\mathbf{A} = -\mathbf{A}^{\dagger})$. The energy can be considered a functional of both $\mathbf{A}$ and $\mathbf{\Psi}_0$,

$$E[\mathbf{\Psi}] = \mathcal{F}[\mathbf{A}, \mathbf{\Psi}_0]. \tag{18}$$

Therefore, in general, the optimal orbitals corresponding to the minimum of the energy functional can be found in a dual-loop procedure. First, the energy is minimized with respect to the elements of $\mathbf{A}$, and second, the functional $\mathcal{L}[\mathbf{\Psi}_0] = \min_{\mathbf{A}} \mathcal{F}[\mathbf{A}, \mathbf{\Psi}_0]$ is minimized with respect to $\mathbf{\Psi}_0$,

$$\min_{\mathbf{\Psi}} E[\mathbf{\Psi}] = \min_{\mathbf{\Psi}_0} \min_{\mathbf{A}} \mathcal{F}[\mathbf{A}, \mathbf{\Psi}_0]. \tag{19}$$

Since anti-Hermitian matrices form a linear space, the inner loop minimization can use well-established local minimization strategies such as efficient quasi-Newton methods with inexact line search, e.g., the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm. The outer loop minimization follows the gradient $\partial \mathcal{L}/\partial \mathbf{\Psi}_0$ projected on the tangent space at $\mathbf{\Psi}_0$.

The GPAW formulation of DM is applicable to all representations of the orbitals available in GPAW, as well as Kohn–Sham (unitary invariant) and orbital density-dependent (nonunitary invariant) energy functionals, and can be used for both finite and extended systems. In LCAO calculations, the reference orbitals are expressed as a linear combination of the atomic basis functions $\mathbf{\Phi}$, $\mathbf{\Psi}_0 = \mathbf{\Phi}\mathbf{C}_0$, where the matrix of coefficients $\mathbf{C}_0$ is fixed. Therefore, only a minimization with respect to the elements of matrix $\mathbf{A}$ is required.[82] For plane wave and real-space grid representations, a minimization in the space tangent to the reference orbitals is sufficient if the functional is unitary invariant. Otherwise, if the functional is nonunitary invariant, such as when self-interaction correction is used (see Sec. III C 5), an inner loop minimization in the occupied–occupied block of the matrix $\mathbf{A}$ is performed to make the energy stationary with respect to the unitary transformation of the occupied orbitals.[50]

The DM method avoids diagonalization of the Hamiltonian at each step, and as a result, it usually involves a smaller computational effort. The DM method has also been shown to be more robust than conventional eigensolvers and density mixing in calculations of molecules and extended systems.[82] However, the current implementation does not support a finite-temperature distribution of occupation numbers and thus can only be used for systems with a finite bandgap.

6. Convergence criteria

The modular architecture of GPAW allows the user to have precise control over how the SCF loop decides that the electronic structure has converged to sufficient precision. GPAW contains simple keywords for common convergence criteria, such as “energy,” “forces,” (electron) “density,” and “work function,” which are sufficient for the most common use cases.

Internally, all convergence criteria are Python objects that are instances of convergence classes. For each step through the SCF loop, each convergence object is called. When any convergence object is called, it is passed a context object that contains the current state of the calculation, such as the wavefunctions and the Hamiltonian. The criterion can thus pull relevant data from the calculation to decide if it is convergent. Because the convergence criterion itself is an object, it can store information such as previous values of the energy for comparison to the new value. When all convergence criteria report that they are converged, the calculation as a whole is considered to be converged and terminates.

This modular nature gives the user full control over how each convergence criterion operates. For example, the user can easily ask the energy criterion to check the differences in the last four values of the energy rather than the last three. If a convergence criterion itself is expensive to calculate, it can make sense to not check it until the rest of the convergence criteria are met. This can be accomplished by activating an internal “calculate_last” flag within the convergence criterion.

Users can easily add their own custom convergence criteria to the SCF loop. If a user would like to use a criterion not included by default with GPAW, it is straightforward to write a new criterion as a Python class and pass this to the convergence dictionary of GPAW. For example, if one wanted to be sure the bandgap of a semiconductor was converged, the criterion could check the bandgap at each iteration, compare it to stored values from previous iterations, and report that the calculation is converged when the peak-to-peak variance among the last n iterations is below a given threshold.

7. PAW datasets and pseudopotentials

GPAW can read PAW datasets from (possibly compressed) XML files following the PAW-XML specification.[83] Dataset files for most of the periodic table can be downloaded from the GPAW webpage or installed with the `gpaw install-data` command-line tool. The datasets are available for the local-density approximation (LDA), Perdew–Burke–Ernzerhof (PBE),[3] revPBE,[346] RPBE,[345] and GLLBSC[97] functionals. The electronic structure code Abinit[84] also reads the PAW-XML format, allowing GPAW and Abinit to share PAW dataset collections such as the Jollet–Torrent–Holzwarth collection.[85]

Specialized datasets can be generated with the `gpaw dataset` command-line tool. This allows one to tweak the properties of a dataset. Some examples could be: (1) add more/less semi-core states; (2) increase/decrease the augmentation sphere radius to make the pseudo-wave functions more/less smooth; (3) add/remove projector functions and corresponding pseudo and all-electron partial waves; or (4) base the PAW dataset on a different XC-functional. These changes will affect the accuracy and cost of the calculations.

GPAW is also able to use norm-conserving pseudopotentials (NCPP) such as Hartwigsen-Goedecker-Hutter (HGH)[86] and pseudopotentials in the Unified Pseudopotential Format (UPF), such as SG15.[87] Non-local NCPPs can be considered an approximation to PAW: in the PAW description, the non-local part of the Hamiltonian [the term containing ΔH_{ij}^a in Eq. (6)] will adapt to the environment, whereas for a NCPP, ΔH_{ij}^a will be diagonal and have a fixed value taken from a reference atom. Because of norm-conservation, NCPPs will have $\Delta S_{ij}^a = 0$.

12 March 2024 13:19:04

© Author(s) 2024

8. Parallelization

GPAW can parallelize over various degrees of freedom, depending on the type of calculation, and it implements multiple algorithms for achieving good parallel performance and scalability. In calculations involving **k**-points, parallelization over them is typically the most efficient, as little communication is needed during the summation of wave functions to calculate the density and any derived quantities. As the number of **k**-points is often limited, especially in large systems, parallelization is also possible over real-space grids in the FD and LCAO modes, as well as over plane waves in the PW mode. All modes also support parallelization over electronic bands, which is particularly efficient for real-time TDDFT where the time-propagation of each band can be carried out independently. Additional parallelization possibilities exist depending on calculations, such as over electron–hole pairs in linear-respose TDDFT calculations.

Parallelization is performed mainly with MPI. In the FD and LCAO modes, it is possible to use OpenMP within shared-memory nodes, which can improve performance when the number of CPU cores per node is large. Dense linear algebra, such as matrix diagonalization and Cholesky decomposition, can be carried out with the parallel ScaLAPACK or ELPA libraries. This applies to both the direct diagonalization in the LCAO mode as well as to subspace diagonalizations in the iterative Davidson and Residual Minimization Method with Direct Inversion in the Iterative Subspace (RMM-DIIS) methods in the FD and PW modes.

For ground-state calculations, GPAW will divide the cores into three MPI communicators: **k**-points, bands, and domain. When parallelizing over **k**-points and/or bands, all the cores of the **k**-point and/or band communicators will have a copy of the density (possibly distributed over the domain communicator). GPAW has the option to redistribute the density from the domain-communicator to all cores so that operations such as evaluating the XC-energy and solving the Poisson equation can be performed more efficiently.

For each **k**-point, all the cores in the band and domain communicators will cooperate on calculating the matrix elements of the Hamiltonian and the overlap operators. Dense-matrix linear-algebra operations on those matrices can be performed on a single core (most efficient for small systems) or with ScaLAPACK, where the matrices are distributed over either all or some of the cores from the pool of cores in the band and domain communicators.

One drawback of Python is that in large parallel calculations, its import mechanism may incur a heavy load on the filesystem because all the parallel processes are trying to read the same Python files. GPAW tries to alleviate this by using a special "broadcast imports" mechanism: during initial module imports, only a single process loads the modules; afterward, MPI is used to broadcast the data to all the processes.

Parallel scalability depends strongly on the calculation mode and the system. The FD mode offers the best scalability for high core counts, as only nearest-neighbor communication is needed across domains over a domain decomposition. In PW mode, the limiting factor is all-to-all communication in parallelization over plane waves. In LCAO mode, communications arise from multi-center integrals of basis functions across domains. At best, GPAW scales to tens or hundreds of nodes in supercomputers.

9. GPU implementation

The GPU implementation of GPAW works both on NVIDIA and Advanced Micro Devices (AMD) GPUs, targeting either CUDA or Heterogeneous-Compute Interface for Portability (HIP) backends, respectively. GPAW uses a combination of manually written GPU kernels, external GPU libraries (such as cuBLAS/hipBLAS), and CuPy.[88] CuPy offers an easy-to-use Python interface for GPUs centered around a NumPy-like GPU array and makes many hardware details completely transparent to the end-user.

In the manually written GPU kernels, both GPU backends (CUDA and HIP) are targeted using a header-only porting approach[89] in which generic GPU identifiers are translated to vendor-specific identifiers at compile time. For example, to allocate GPU memory, the identifier `gpuMalloc` is used in the code, which is then translated either to `cudaMalloc` or `hipMalloc` depending on which GPU backend is targeted. This allows us to avoid unnecessary code duplication and still target multiple hardware platforms natively.

An earlier GPU implementation of GPAW[90,91] served as the starting point for the recent work on a new GPU code based on the rewritten ground-state code. The objects that store quantities like $\tilde{\psi}_n(\mathbf{r})$, $\tilde{p}_i^a(\mathbf{r})$, $\langle \tilde{p}_i^a | \tilde{\psi}_n \rangle$, $D_{ii'}^a$, $\tilde{n}(\mathbf{r})$ and $\tilde{v}(\mathbf{r})$ use Numpy arrays for the CPU code and CuPy arrays when running on a GPU. At the moment, GPUs can be used for total-energy calculations with LDA/generalized gradient approximation (GGA) in the PW mode.

Parallelization to multiple GPUs is performed using MPI. Each MPI rank is assigned a single GPU, and communication between the GPUs is handled by MPI. Support for GPU-aware MPI makes it possible to do direct GPU-to-GPU communication without unnecessary memory copies between the GPU device and the host CPU.

C. XC-functionals

Exchange–correlation (XC) functionals provide a mapping between the interacting and non-interacting systems of electrons. In Kohn–Sham DFT, the density is built from a set of occupied non-interacting single-particle orbitals ψ_i: $n = \sum_i f_i |\psi_i|^2$, with f_i denoting the occupation numbers, leading to the same density as the interacting system. The total energy in DFT is expressed as a sum of the density functionals for the different contributions,

$$E_{\text{tot}}[n] = T_{\text{S}}[n] + V_{\text{ext}}[n] + U_{\text{H}}[n] + E_{\text{xc}}[n], \tag{20}$$

where $T_{\text{S}}[n]$ denotes the kinetic energy of the non-interacting system, $V_{\text{ext}}[n]$ the energy of the density in the external potential, $U_{\text{H}}[n]$ the classical Coulomb energy of the density with itself, and $E_{\text{xc}}[n]$ the so-called exchange–correlation energy, which collects all energy contributions missing in the prior terms and therefore provides a mapping between the interacting and the non-interacting system of electrons. While the first three terms can be calculated exactly, even the form of E_{xc} is unknown, and although proven to exist and be exact in principle, it has to be approximated in practice. A huge number of approaches belonging to several families exist.[1,2,92] Several of these approximations are available in GPAW, and an overview is given in the following.

12 March 2024 13:19:04

© Author(s) 2024

1. Libxc and libvdwxc

The *libxc* library[66] provides implementations of several (semi-)local variants of the XC functional, given by the LDA, GGA, and meta-GGA (MGGA)[91] families. These are available in GPAW by a combination of their names from *libxc*, e.g., "GGA_X_PBE +GGA_C_PBE" for PBE.[3] At the moment, using MGGA functionals from *libxc* in GPAW requires *libxc* to be compiled with --disable-fhc. Most MGGA functionals in *libxc* do not depend on the laplacian of the density—those that do are not supported.

Additionally, GPAW provides its own implementation of several (semi-)local functionals called by their short names, e.g., TPSS,[93] PBE,[3] and LDA, the latter with the correlation of Perdew and Wang.[94] Several hybrids (see below) are implemented in GPAW with the support of the *libxc* library for their local parts.

For fully non-local van der Waals functionals, like the van der Waals (vdW)-DF functional,[31] GPAW uses the efficient fast Fourier-transform convolution algorithm by Román-Pérez and Soler[95] as implemented in the *libvdwxc* library.[67]

All LDA, GGA, and common MGGA functionals, such as Strongly Constrained and Appropriately Normed semilocal density functional (SCAN)[28] and TPSS, from libxc can be used in GPAW.

2. GLLB-sc

GPAW has an implementation of the solid-state modification of the Gritsenko–van Leeuwen–van Lenthe–Baerends exchange–correlation model potential (GLLB-sc).[96,97] This potential has been shown to improve the description of the bandgap[16] and the *d*-electron states in noble metals.[52,98–100]

3. Hubbard U

DFT+U calculations using a Hubbard-like correction can be performed in GPAW to improve the treatment of Coulombic interactions of localized electrons. This correction is most commonly applied to the valence orbitals of transition metals to assist with obtaining experimental bandgaps of oxides,[101] which may otherwise be underestimated.[102,103] In addition, formation energies and magnetic states are often improved due to a more correct description of the electronic structure. It may also be applied to main-group elements such as N, O, P, or S, but this is less commonly performed.[104]

In the formalism chosen in GPAW,[105] one uses a single U^{eff} parameter rather than separate U and J parameters for on-site Coulombic and on-site exchange, respectively.[106] The correction influences the calculation by applying an energy penalty to the system

$$U^{\mathrm{eff}} = U - J, \tag{21}$$

$$E_{\mathrm{DFT+U}} = E_{\mathrm{DFT}} + \sum_{a,i} \frac{U_i^{\mathrm{eff}}}{2} \mathrm{Tr}\left(\rho_i^a - \rho_i^a \rho_i^a\right), \tag{22}$$

where the sum runs over the atoms a and orbitals i, for which the correction should be applied. E_{DFT} is calculated by a standard GPAW calculation and is corrected to $E_{\mathrm{DFT+U}}$ by penalizing the energy such that fully occupied or fully unoccupied orbitals are stabilized. The magnitude of the correction depends on U_i^{eff}, and the atomic orbital occupation matrix (ρ_i^a) controls which orbitals contribute to the correction based on their occupation. In principle, any orbital on any element that is partially occupied can be corrected.

GPAW supports the normalization of the occupation matrix, accounting for the truncated portion of the wave function outside the augmentation sphere. To maintain consistency with other codes that do not support normalization, this normalization can be disabled, but large disagreements are expected when applied to p orbitals. GPAW is one of the few codes that currently supports multiple simultaneous corrections to orbitals on the same atom; this is useful when two types of orbitals, such as p and d orbitals, are nearly degenerate but both are partially occupied.

There is no U^{eff} that is strictly correct, but methods such as RPA[107] or linear response[108,109] can allow for a value to be calculated from first principles. More commonly, U^{eff} is chosen semi-empirically to fit experimental properties such as formation energy,[110] bandgap[111] or, more recently, machine learning predictions.[112]

4. Hybrids

Hybrid functionals, especially range-separated functionals (see below), correct problems present in calculations utilizing (semi-)local functionals such as the wrong asymptotic behavior of the effective potential leading to an improper description of Rydberg excitations,[113] the improper calculation of the total energy against fractional charges leading to the charge delocalization error,[114] and the wrong description of (long-range) charge transfer excitations due to the locality of the exchange hole.[115–117]

The exchange–correlation energy E_{xc} can be split into the contributions from exchange, E_{x}, and from correlation, E_{c}.[2,118] Hybrid functionals combine the exchange from the (semi-)local functionals with the exchange from Hartree–Fock theory (HF). Global hybrids such as PBE0[119] mix the exchange from DFT with the exchange from HF by a global fixed amount. Range-separated functionals (RSF) add a separation function ω_{RSF}[120,121] to express the Coulomb kernel in the exchange integrals as

$$\frac{1}{r_{12}} = \underbrace{\frac{1 - [\alpha + \beta(1 - \omega_{\mathrm{RSF}}(\gamma, r_{12}))]}{r_{12}}}_{\mathrm{SR}} + \underbrace{\frac{[\alpha + \beta(1 - \omega_{\mathrm{RSF}}(\gamma, r_{12}))]}{r_{12}}}_{\mathrm{LR}}, \tag{23}$$

where $r_{12} = |\vec{r}_1 - \vec{r}_2|$. Here, α and β are mixing parameters for global and range-separated mixing, respectively. ω_{RSF} is a soft function with values ranging from one for $r_{12} = 0$ to zero for $r_{12} \to \infty$, where the decay is controlled by the parameter γ. Long-range-corrected (LC/LCY) RSFs such as LCY-PBE[122] use (semi-)local approximations for short-range (SR) interaction and apply HF exchange for long-range (LR) interaction. Short-range-corrected RSFs such as Heyd-Scuseria-Ernzerhof (HSE)[5] reverse this approach. The parameter γ is either fixed or can be varied to match the criteria of the ideal functional, e.g., that the energy of the highest occupied molecular orbital matches the ionization potential.[118,123]

Details on the FD-mode implementation of long-range corrected RSF can be found in Refs. 124 and 125. In general, the

12 March 2024 13:19:04

© Author(s) 2024

FD-mode implementation of hybrids is limited to molecules, and forces have not been implemented. The PW-mode implementation of hybrids handles **k**-points, exploits symmetries, and can calculate forces.

5. SIC

A fully self-consistent and variational implementation of the Perdew–Zunger self-interaction correction (PZ-SIC)[126] is available in GPAW. It corrects the various problems with (semi-)local Kohn–Sham (KS) functionals mentioned earlier in the context of hybrid functionals. Atomic forces are available in the GPAW implementation with all three types of basis sets. A corrected KS functional has the form

$$E^{\mathrm{SIC}}[n_1,\ldots,n_n] = E^{\mathrm{KS}}[n] - \alpha\sum_i^{\mathrm{occ}}\left(U_{\mathrm{H}}[n_i] + E_{\mathrm{xc}}[n_i,0]\right), \tag{24}$$

where the self-Coulomb and self-XC of each occupied orbital density are subtracted from the Kohn–Sham energy functional. Due to the explicit dependence on the orbital densities, the corrected energy functional is not invariant under a unitary transformation among occupied orbitals and, thereby, not a KS functional. As a result, the minimization of E^{SIC} requires special direct minimization techniques (see Sec. III B 5) and delivers a specific set of (typically localized) optimal orbitals. The calculations should be carried out using complex orbitals.[127–129] The full PZ-SIC has been shown to give an over-correction to binding energy as well as bandgaps, and improved results for these properties are obtained by scaling the SIC by $\alpha = 1/2$,[130,131] while the long-range form of the effective potential, necessary for Rydberg state calculations, requires the full correction.[132]

PZ-SIC has been shown to give accurate results in cases where commonly used KS functionals fail. This includes, for example, the Mn dimer, where the PBE functional gives qualitatively incorrect results while the corrected functional gives close agreement with high-level quantum chemistry results as well as experimental measurements.[133] Another example is the defect state of a substitutional Al atom in α-quartz.[134] In addition, PZ-SIC has been shown to improve the values of the excitation energy of molecules obtained in variational calculations of excited states[50,135] (see Sec. VIII B).

6. BEEF

A great strength of Kohn–Sham DFT and its extensions is that a reasonably high accuracy of physical, material, and chemical properties can be obtained at a relatively moderate computational cost. DFT is thus often used to simulate materials, reactions, and properties where, de facto, there is no “better” alternative. Even though more accurate electronic structure methods in principle might exist for a given application, the poor scaling of systematically more accurate methods often makes these computationally infeasible at the given system size that is being studied using DFT. Therefore, one is often in the situation where the accuracy of a given DFT calculation of some materials or chemical properties cannot be verified against, e.g., a more accurate solution of the Schrödinger equation, even on the biggest available supercomputers.

On the other hand, the wealth of available XC functionals naturally allows one to look at how sensitive a DFT result is to the choice of functional, and often accuracy is, therefore, being judged primarily by applying a small set of different XC functions, especially if no accurate benchmark theoretical simulation or experimental measurement is available. A challenge, however, is that the different available functionals are often known to be particularly good at simulating certain properties and poor at others. It is thus not at all clear how much one should trust a given functional for a given simulated property. The Bayesian Error Estimation (BEE) class of functionals[136] attempts to develop a practical framework for establishing an error estimate from a selected set of functional “ingredients.”

Assume that the XC model $M(a)$ is a function of a set of parameters, a, that can be varied freely. If a benchmark dataset, D, of highly accurate properties established from experiments or higher-accuracy electronic structure simulations is available, we can attempt to identify the ensemble of models given by some distribution function, $P(M(a))$, such that the most likely model in the ensemble, $M(a_0)$, makes accurate predictions for the benchmark dataset, while the spread of the ensemble reproduces the spread between the predictions of the most likely model and the benchmark data.

Bayes’ theorem provides a natural framework to search for the ensemble distribution. If a joint distribution between $M(a)$ and D exists, which we would assume, then Bayes’ theorm gives

$$P(M(a)|D) \propto P(D|M(a))P_0(M(a)), \tag{25}$$

where $P_0(M(a))$ is the prior expectation of the distribution of models before looking at the data, D, and $P(D|M(a))$ is the likelihood of seeing the data given the model. To achieve a useful ensemble, much care has to be put into finding a large enough set of varied and accurate data for different materials and chemical properties, and much care has to be applied as to how the ensemble is regularized to avoid overfitting.[137,138]

In all the BEE functionals (BEEF), choices are made such that ultimately the XC functional is linear in a and such that the distribution of a ends up following a multidimensional normal distribution given by a regularized covariance matrix Γ,

$$P(a) \propto \exp\left(-(a-a_0)\Gamma(a-a_0)\right), \tag{26}$$

where Γ has been scaled in such a way that the ensemble reproduces the observed standard deviation between $M(a_0)$ and D.

Figure 2 shows an ensemble of exchange enhancement factors $F_x(s)$ from the BEEF-vdW functional, where $s = |\nabla n|/(2k_F n)$ is the reduced density gradient. The approach has given rise to several functionals, BEEF-vdW,[137] meta-BEEF (mBEEF),[138] and mBEEF-vdW,[139] which all include error estimation and are readily available in GPAW. Through the ASE interface, one can, for example, utilize the error ensembles to establish error estimates on Python-implemented models that are using DFT simulations for their parametrization. An example of this is the application of error estimates to adsorption scaling relations, microkinetics, and material selection.[140]

One risk to the approach of establishing error estimates from a small selected “ensemble” of XC-functionals is clearly that if the simulated property in question is poorly described by all functionals in the ensemble, then the error estimate might also become poor. This could, for example, be the case if one tried to establish an error estimate for bandgaps in oxides or van der Waals bonding of adsorbates on surfaces based on an ensemble of GGA XC-functionals, since no GGA functional may be accurate for simulating either property.

12 March 2024 13:19:04

© Author(s) 2024

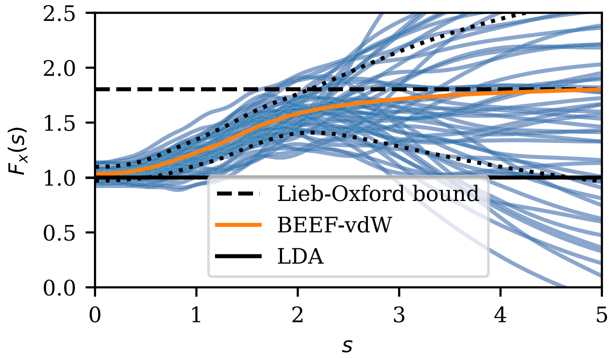


FIG. 2. Bayesian ensemble of XC functionals around BEEF–vdW. The orange solid line is the BEEF–vdW exchange enhancement factor, while the blue lines depict $F_x(s)$ for 50 samples of the randomly generated ensemble. Dotted black lines mark the exchange model perturbations that yield DFT results ±1 standard deviation away from BEEF–vdW results.

IV. ION DYNAMICS

GPAW can be employed as a “black-box” calculator, supplying energies and forces to other programs such as, e.g., ASE, which then perform optimization of ground state geometries and reaction paths or carry out molecular dynamics. In fact, this is a key design principle behind GPAW. Methodological developments and general implementations that are not directly dependent on fast access to detailed electronic structure information should preferably be implemented externally by GPAW. This led to the maximal simplicity of the GPAW code itself while also allowing for the external code to be utilized and tested with other electronic structure codes and atomistic potentials. A key to the high efficiency of GPAW simulations involving ionic displacements is the versatile implementation of restraints in ASE. Here, many types of restraints are readily accessible, from the simple removal of degrees of freedom to more exotic restraints allowing for rigid molecule dynamics,[141] harmonic restoring forces,[142] spacegroup preservation, and combined ionic-unit cell dynamics.[143] Many algorithms are available for various structure optimization and molecular dynamics tasks.

A. Structure relaxation

Local structure optimization in GPAW is typically achieved through the use of an optimizer from ASE. Here, a larger range of standard optimizers are available, such as quasi-Newton algorithms including BFGS[144–147] and limited-memory BFGS,[148] or Newtonian dynamics-based algorithms such as MDMin and FIRE.[149] The fact that the optimizers have been implemented externally in GPAW provides benefits in terms of simple procedures for restarting long simulations and monitoring their convergence. Some optimizers from SciPy[69] are also available through the open-source ASE package, which provides a simple way to interface any optimizer in the SciPy format to GPAW. Preconditioning is implemented in an accessible way,[150] which often leads to significant performance improvements.

Of the classical optimization methods, the quasi-Newton algorithms are often highly competitive. Here, one builds up information on the Hessian or inverse Hessian matrix from calculated forces, ultimately leading to an accurate harmonic model of the potential energy surface in the vicinity of the local minimum. Such algorithms can, however, have problems both dealing with anharmonicity in the potential energy surface and with any noise in the electronic structure simulations. It often makes sense instead to fit a Gaussian process to the calculated energies and forces and minimize the potential within this model. This is implemented as the so-called GPmin method in ASE and often converges on the order of three times faster than the best quasi-Newton optimizers.[151]

B. Reaction paths and barriers

Reliable calculations of energy barriers are of key importance for determining the rate of atomistic processes. In many quantum-chemical codes utilizing accurate atom-centered basis functions, this is achieved using analytical second derivatives in the direct search for first-order saddle points. This approach is less useful in the plane-wave based or grid-based modes of GPAW. The Dimer method[152] is implemented in ASE and can be used with GPAW, but often one would like to have an overview of the entire reaction path of an atomic-scale process to verify that the correct energy barrier for a process has been determined and to obtain a full overview of the atomistic mechanism. For this purpose, the Nudged Elastic Band (NEB) method is typically employed through the ASE interface to GPAW. Both the original method[153] and a range of later improvements are available.[154–159] Special care has to be taken in selecting the optimizer for a NEB algorithm, as this choice can have a drastic influence on the convergence rate. For optimization of the reaction path, drastic performance improvements can be obtained by carrying out the optimization in a surrogate machine-learning model fitted to the potential energy surface.[160–162] GPAW has been used to drive NEB calculations on both surface systems[163–166] as well as in molecules.[166,167]

C. Global structure optimization

GPAW is integrated with various tools for global optimization of structures, compositions, as well as material morphologies. Some quite generally applicable global optimization tools available are basin hopping,[168] minima hopping,[169] and genetic algorithms.[170] Some of the most powerful global optimization problems addressed using GPAW rely on machine-learning accelerated global-optimization strategies. These strategies have, for example, been applied to surfaces, clusters,[171] and crystal structures generally.[172] In other machine-learning accelerated global-optimization routines, GPAW was used to generate initial databases of surface and bulk systems and for later model validation in a strategy that uses Gaussian processes to generate surrogate potential-energy surfaces. These were then explored with Bayesian-optimization techniques, achieving speed-ups over more conventional methods of several orders of magnitude in finding the optimal structures of the systems under investigation.[173–175] The method has been augmented by introducing extra (hyper)dimensions by interpolating between chemical elements, which speeds up the efficiency of the global search.[176] GPAW has been integrated with a Covariance Matrix Adaptation Evolution Strategy (CMA-ES) framework, providing energies and forces, generating training data, and evaluating CMA-ES candidate structures.[177]

12 March 2024 13:19:04

D. Molecular dynamics and QM/MM

Ab initio molecular dynamics (MD) can be performed and analyzed through the ASE interface to GPAW. This includes standard simulations such as constant-NVE, constant-NVT, and constant-NPT simulations. There is access to, e.g., Langevin-, Andersen-, Nose–Hoover-, and Berendsen-dynamics. Due to the externalization of the dynamics, the development of novel algorithms and analysis tools becomes facile. Some examples of the use of GPAW span *ab initio* MD studies that have explored the liquid structure of water[178] and the water/Au(111) electrochemical interface.[179]

Furthermore, GPAW is capable of working with external electrostatic potential terms from user-supplied point charge values and positions, enabling Quantum Mechanical (QM)/Molecular Mechanical (MM) simulations. Since GPAW is from the outset designed around highly efficient grid operations, the computational overhead of evaluating this potential as well as the resulting forces on the MM point charges from the QM density is kept low.[180] GPAW has been central in a range of studies of ion dynamics in solution, both in and out of equilibrium. By using the molecular-dynamics functionality of ASE, researchers have performed QM/MM MD simulations of excited-state bond formation in photocatalyst model systems[181–184] and electron transfer, as well as coupled solute–solvent conformational dynamics in photosensitizer systems.[167,185] Work on polarizable embedding QM/MM within GPAW is ongoing.[186,187]

V. MAGNETISM AND SPIN

Many important technological applications utilize magnetic order or the manipulation of spin in materials. GPAW has a wide range of functionalities that facilitate the analysis of magnetic properties. This includes calculations with non-collinear spin, the inclusion of external magnetic fields, spin–orbit coupling, and spin spiral calculations within the generalized Bloch theorem. The implementation of these features is described below, while additional methods for calculating magnetic excitations are described in Sec. VI D.

A. Spin–orbit coupling

The spin–orbit coupling is typically completely dominated by regions close to the nuclei, where the electrostatic potential becomes strong. Within the PAW augmentation sphere of atom a, the spin-σ component of orbital n is given by

$$\psi^a_{n\sigma}(\mathbf{r}) = \sum_i \langle \tilde{p}^a_i | \tilde{\psi}_{n\sigma} \rangle \phi^a_i(\mathbf{r}). \tag{27}$$

Assuming a spherically symmetric potential, the spin–orbit Hamiltonian for atom a is then written as

$$\hat{H}^a_{\mathrm{SOC}} = \frac{\mu_{\mathrm{B}}}{\hbar m_e e c^2} \frac{1}{r} \frac{\mathrm{d}V^a_r}{\mathrm{d}r} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \tag{28}$$

where V^a_r is the radial electrostatic potential of atom a. We evaluate V^a_r as the spherical part of the XC- and Hartree-potential from the local expansion of the density given by Eq. (11). Since the partial waves ϕ^a_i are eigenstates of the scalar-relativistic Hamiltonian, these are independent of spin, and it is straightforward to evaluate the action of $\hat{\mathbf{L}}$ on them.

The eigenenergies can be accurately calculated in a non-selfconsistent treatment of the spin–orbit coupling and may be obtained by diagonalizing the full Hamiltonian on a basis of scalar-relativistic orbitals,[188]

$$H_{mn} = \varepsilon^0_m \delta_{mn} + \sum_{aij\sigma\sigma'} \langle \tilde{\psi}^0_{m\sigma} | \tilde{p}^a_i \rangle \langle \phi^a_i | \hat{H}^a_{\mathrm{SOC},\sigma\sigma'} | \phi^a_j \rangle \langle \tilde{p}^a_j | \tilde{\psi}^0_{n\sigma'} \rangle. \tag{29}$$

Here, ε^0_m and ψ^0_m represent the scalar-relativistic eigenenergies and eigenstates, respectively. This constitutes a fast post-processing step for any scalar-relativistic calculation and only requires the projector overlaps $\langle \tilde{p}^a_j | \tilde{\psi}^0_{n\sigma'} \rangle$. It should be noted that this approach in principle requires convergence with respect to the number of scalar-relativistic states included in the basis, but the eigenvalues typically converge rather rapidly with respect to the basis. In Fig. 3, we show the band structure of a WS_2 monolayer obtained from PBE with non-selfconsistent spin–orbit coupling. The W atoms introduce strong spin–orbit coupling in this material, and the valence band is split by 0.45 eV at the K point. The spin degeneracy is retained along the $\Gamma - \mathrm{M}$ line, which is left invariant by two non-commuting mirror symmetries.

For magnetic materials, the non-selfconsistent treatment of spin–orbit coupling is convenient for evaluating the magnetic anisotropy. The magnetic force theorem[189] implies that rotating the magnetic moments away from the ground state configuration yields a contribution to the energy, which is well approximated by the change in Kohn–Sham eigenvalues. The change in energy for a given orientation of the magnetization density can thus be obtained as

$$\Delta E^{\theta,\varphi} = \sum_n f^{\theta,\varphi}_n \varepsilon^{\theta,\varphi}_n, \tag{30}$$

where $\varepsilon^{\theta,\varphi}_n$ are the eigenvalues obtained from diagonalizing Eq. (29) with the spins rotated in a direction defined by the angles (θ, φ) and $f^{\theta,\varphi}_n$ are the associated occupation numbers. This corresponds to rotating the xc-magnetic field (defined below), which will lead to different eigenvalues when spin–orbit coupling is included. In two-dimensional magnets, an easy-axis anisotropy is decisive for magnetic order,[41,190,191] and Eq. (30) is easily applied to high-throughput computations of magnetic properties.[192,193]

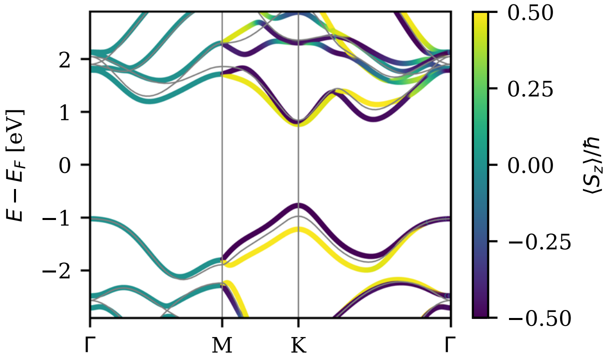


FIG. 3. Band structure of the WS_2 monolayer obtained from PBE with non-selfconsistent spin–orbit coupling. The colors indicate the expectation value of S_z for each state. The gray lines show the band structure without spin–orbit coupling.

12 March 2024 13:19:04

© Author(s) 2024

B. Self-consistent non-collinear magnetism

The Kohn–Sham framework for treating non-collinear magnetism was developed by Barth and Hedin[194] and involves the spin density matrix

$$\rho_{\sigma\sigma'}(\mathbf{r}) = \sum_n f_n \psi^*_{n\sigma}(\mathbf{r})\psi_{n\sigma'}(\mathbf{r}), \tag{31}$$

as the basic variable. The electronic density and magnetization are then given by $n = \mathrm{Tr}[\rho]$ and $\mathbf{m} = \mathrm{Tr}[\boldsymbol{\sigma}\rho]$, respectively. The XC-potential acquires four components, which are the functional derivatives of the XC-energy with respect to the density matrix and may be represented as a 2 × 2 matrix acting on spinor Kohn–Sham states. These can be expressed in terms of density and magnetization, which lead to the XC-part of the Kohn–Sham Hamiltonian

$$H_{\mathrm{xc}} = v_{\mathrm{xc}} + \mathbf{B}_{\mathrm{xc}} \cdot \boldsymbol{\sigma}, \tag{32}$$

where the scalar potential and XC-magnetic field are given by

$$v_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}, \qquad \mathbf{B}_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta \mathbf{m}(\mathbf{r})}. \tag{33}$$

In GPAW, the self-consistent treatment of non-collinear spin is implemented within LDA, where $\mathbf{B}_{\mathrm{xc}}$ is approximated as

$$\mathbf{B}^{\mathrm{LDA}}_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E^{\mathrm{LDA}}_{\mathrm{xc}}}{\delta m(\mathbf{r})} \hat{\mathbf{m}}(\mathbf{r}). \tag{34}$$

Here, $m(\mathbf{r}) = |\mathbf{m}(\mathbf{r})|$ is the magnitude, and $\hat{\mathbf{m}}(\mathbf{r})$ is the direction of the magnetization. The generalization of Eq. (34) to GGAs is plagued by formal as well as numerical problems,[195,196] and the self-consistent solution of the non-collinear Kohn–Sham equations is presently restricted to LDA. It should be emphasized that the PAW formalism allows for a fully non-collinear treatment that does not rely on intra-atomic collinearity, which is often imposed in other electronic structure packages.

Spin–orbit coupling may be included by adding Eq. (28) to the Kohn–Sham Hamiltonian, and this constitutes a fully self-consistent framework for spin–orbit calculations within LDA.

C. Orbital magnetization

Current *ab initio* methods for determining the orbital magnetization of a material involve either the modern theory, i.e., a Berry-phase formula, or the atom-centered approximation (ACA), where contributions to the expectation value of the angular momentum operator are restricted within atom-centered muffin-tin (MT) spheres with specified cutoff radii. The PAW formulation of the wave functions allows for an approximation similar to the MT-ACA, where only the PAW expansion of the wave functions is assumed to contribute significantly to the expectation value of the angular momentum. The orbital magnetic moments can then be calculated through

$$\mathbf{m}^a_{\mathrm{orb}} = -\frac{e}{2m_e} \sum_a \sum_{ii'} \sum_\sigma D^a_{\sigma ii'} \langle \phi^a_i | \hat{\mathbf{L}} | \phi^a_{i'} \rangle, \tag{35}$$

where $D^a_{\sigma ii'}$ are elements of the atomic density matrix as defined in Eq. (9), $\phi^a_i(\mathbf{r})$ are the *bound* all-electron partial waves of atom a, and $\hat{\mathbf{L}}$ is the angular-momentum operator.

TABLE II. Calculated and measured values for the orbital magnetization in units of μ_B per atom.

Crystal easy axis	bcc-Fe [001]	fcc-Ni [111]	fcc-Co [111]	hcp-Co [001]
PAW-ACA	0.0611	0.0546	0.0845	0.0886
MT-ACA[197]	0.0433	0.0511	0.0634	0.0868
Modern theory[197]	0.0658	0.0519	0.0756	0.0957
Experiment[198]	0.081	0.053	0.120	0.133

We note that this expression does not entail cutoff radii for the atomic contributions to the orbital magnetic moments; instead, the entire all-electron partial wave is included, although the all-electron atomic expansion is only formally exact inside the PAW spheres. Additionally, this means that the unbounded, i.e., non-normalizable, all-electron partial waves must be excluded in Eq. (35).

A prerequisite for nonzero orbital magnetization is broken time-reversal symmetry, as represented by a complex Hamiltonian that is not unitarily equivalent to its real counterpart. Practically, this means that a finite orbital magnetization requires magnetic order and either the inclusion of spin–orbit coupling or non-coplanar spin textures. In GPAW, the spin–orbit interaction can be included either self-consistently in a non-collinear calculation or as a non-self-consistent post-processing step following a collinear calculation. The orbital magnetization has been calculated using Eq. (35) for the simple ferromagnets bcc-Fe, fcc-Ni, fcc-Co, and hcp-Co, where the spin–orbit interaction is included non-self-consistently. The PAW-ACA results are displayed in Table II along with MT-ACA results, modern theory results, and measurements from experiments, demonstrating mainly that the PAW-ACA can be an improvement over the MT-ACA and secondly that there is a decent agreement between the PAW-ACA and the modern theory for these systems.

D. Constant B-field

The coupling of the electronic spin magnetic moments to a constant external magnetic field **B** can be included by the addition of a Zeeman term in the Kohn–Sham Hamiltonian

$$\Delta H_{\mathrm{B}} = \mathbf{B} \cdot \boldsymbol{\sigma}. \tag{36}$$

As an example, we can consider the spin-flop transition in $\mathrm{Cr_2O_3}$. The ground state is anti-ferromagnetic with a weak anisotropy that prefers alignment of the spins with the z-axis, and if a magnetic field is applied along the z-direction, it becomes favorable to align the spins along a perpendicular direction (with a small ferromagnetic component) at the critical field where the Zeeman energy overcomes the anisotropy. This is clearly a spin–orbit effect, and one has to carry out the calculations in a fully non-collinear framework with self-consistent spin–orbit coupling. In Fig. 4, we show the energy of the two alignments of spin as a function of the external magnetic field obtained with LDA+U ($U = 2$ eV). The minimum energy configuration changes from S_z to S_x alignment at 6 T, i.e., the spin flips at this critical field. This is in excellent agreement with the experimental value of Ref. 199. The critical field is, however, rather strongly dependent on the chosen value of U and increases by a factor of two

12 March 2024 13:19:04

© Author(s) 2024

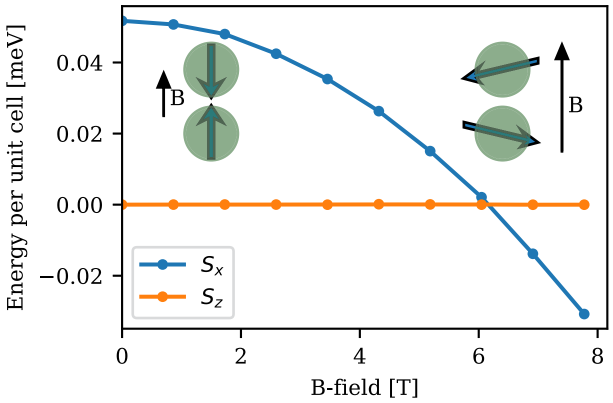


FIG. 4. Spin-flop transition in Cr_2O_3. The alignment of the spins with respect to the magnetic field is sketched for small and large magnitudes of the field. The canting (the small ferromagnetic component after the spin flop) is exaggerated for visualization. The actual canting is roughly 1°.

in the absence of U. The magnetic anisotropy (per unit cell) can be read off from the energy difference at $B = 0$ T.

E. Spin spirals

The ground-state magnetic structure of frustrated and/or chiral magnets is often non-collinear and may be incommensurate with the chemical unit cell. In the classical isotropic Heisenberg model, the energy is always minimized by a planar spin spiral.[200] The energy of a general planar spin spiral can be evaluated efficiently within the chemical unit cell using the Generalized Bloch's theorem (GBT).[201] In the GBT implementation of GPAW,[40] there is no restriction on interatomic collinearity and, therefore, it encodes a rotation of the all-electron magnetization density by an angle $\varphi = \mathbf{q} \cdot \mathbf{R}_i$ upon translations of a lattice vector $\mathbf{R}_i$. The Kohn–Sham equations can then be solved self-consistently for a fixed wave vector $\mathbf{q}$,

$$\hat{H}_{\mathbf{q}}(\mathbf{k})|u_{\mathbf{q},n\mathbf{k}}\rangle = \varepsilon_{\mathbf{q},n\mathbf{k}}|u_{\mathbf{q},n\mathbf{k}}\rangle, \tag{37}$$

using only the periodic part of the generalized Bloch orbitals, $|\psi_{\mathbf{q},n\mathbf{k}}\rangle = U_{\mathbf{q}}^{\dagger}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}|u_{\mathbf{q},n\mathbf{k}}\rangle$, where $U_{\mathbf{q}}$ denotes the spin rotation around the z-axis:

$$U_{\mathbf{q}}(\mathbf{r}) = \begin{pmatrix} e^{i\mathbf{q}\cdot\mathbf{r}/2} & 0 \\ 0 & e^{-i\mathbf{q}\cdot\mathbf{r}/2} \end{pmatrix}. \tag{38}$$

The generalized Bloch Hamiltonian without spin–orbit coupling is given by

$$\hat{H}_{\mathbf{q}}(\mathbf{k}) = e^{-i\mathbf{k}\cdot\mathbf{r}}U_{\mathbf{q}}(\mathbf{r})\hat{H}U_{\mathbf{q}}^{\dagger}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{39}$$

and once Eq. (37) has been solved self-consistently for a given $\mathbf{q}$, the corresponding spin-spiral energy is evaluated as usual, $E_{\mathrm{SS}}(\mathbf{q}) = E_{\mathbf{q}}^{\mathrm{DFT}}[n, \mathbf{m}]$.

Using the GPAW functionality to compute the spin-spiral energy as a function of $\mathbf{q}$, one can then search for the ground-state wave vector $\mathbf{Q}$ that minimizes the energy. In Fig. 5, we show that the monolayer $NiBr_2$ has an incommensurate spin-spiral ground state

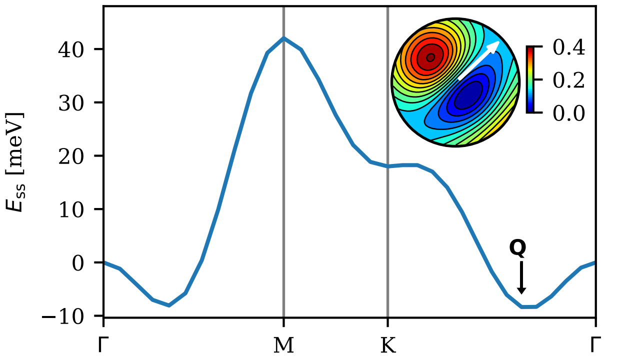


FIG. 5. LSDA spin-spiral spectrum of the $NiBr_2$ monolayer (structure taken from the C2DB[15]). The ground state has an incommensurate wave vector $\mathbf{Q} \simeq [0.1, 0.1, 0]$. The magnetic moment displays only weak longitudinal fluctuations, and the bandgap remains finite for all wave vectors $\mathbf{q}$, indicating that a Heisenberg model description of the material would be appropriate. The inset shows the spin–orbit correction to the spin-spiral energies (in meV) as a function of the normal vector $\mathbf{n}$ of the planar spin spiral. $\mathbf{n}$ is depicted in terms of its stereographic projection in the upper hemisphere above the monolayer plane. The spiral plane is found to be orthogonal to $\mathbf{Q}$ and tilted slightly with respect to the out-of-plane direction.

and that the local magnetic moment on the Ni atom depends only weakly on the wave vector $\mathbf{q}$.

Furthermore, the orientation of the ground state spin spiral can be obtained by including spin–orbit coupling non-self-consistently in the projected spin–orbit approximation[202] and search for the orientation that minimizes the energy. The ordering vector $\mathbf{Q}$ and the normal vector $\mathbf{n}$ of the spiral plane thus constitute a complete specification of the magnetic ground state within the class of single-$\mathbf{q}$ states. The normal vector is of particular interest since spin spirals may lead to spontaneous breaking of the crystal symmetry, and the normal vector largely determines the direction of magnetically induced spontaneous polarization.[40] In Fig. 5, we show that $\mathbf{n}$ is perpendicular to the wavevector $\mathbf{Q}$ in the monolayer $NiBr_2$ ground state, corresponding to a cycloidal spin spiral.

VI. RESPONSE FUNCTIONS AND EXCITATIONS

Linear response functions are the bread and butter of condensed matter physics. Their applications include the description of dielectric screening, optical and electron energy loss spectra, many-body excitations, and ground state correlation energies. In this section, we describe the methods available in GPAW for calculating electronic response functions as well as GW quasiparticle band structures and optical excitations from the Bethe–Salpeter Equation (BSE). In addition to the electronic response function, GPAW can also calculate the transverse magnetic susceptibility, which holds information about the magnetic excitations, e.g., magnons, and can be used to derive parameters for classical Heisenberg spin models and to estimate magnetic transition temperatures. Finally, we present methods to calculate Raman spectra of solids and quadratic optical response tensors for describing second harmonics generation and the Pockels electro-optical effect.

12 March 2024 13:19:04

© Author(s) 2024

A. Linear-response TDDFT

To linear order, the change in electron density induced by a time-dependent external (scalar) potential $V_{\mathrm{ext}}(\mathbf{r},t)$ is governed by the electronic susceptibility, χ,

$$\delta n(\mathbf{r},t) = \int_{-\infty}^{\infty} dt' \int d\mathbf{r}'\, \chi(\mathbf{r},\mathbf{r}',t-t')V_{\mathrm{ext}}(\mathbf{r}',t'). \tag{40}$$

The susceptibility is itself given by the Kubo formula[203]

$$\chi(\mathbf{r},\mathbf{r}',t-t') = -\frac{i}{\hbar}\theta(t-t')\langle[\hat{n}_0(\mathbf{r},t),\hat{n}_0(\mathbf{r}',t')]\rangle_0, \tag{41}$$

where the expectation value is taken with respect to the ground state at zero temperature and the density operators are cast in the interaction picture.

In the noninteracting Kohn–Sham system, the susceptibility can be evaluated explicitly from the Kohn–Sham orbitals, eigenvalues, and occupations,

$$\begin{aligned}\chi_0(\mathbf{r},\mathbf{r}',\omega) = \lim_{\eta\to 0^+} 2\sum_{n\mathbf{k}}\sum_{m\mathbf{k}'}(f_{n\mathbf{k}} - f_{m\mathbf{k}'})\\ \times \frac{\psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{m\mathbf{k}'}(\mathbf{r})\psi^*_{m\mathbf{k}'}(\mathbf{r}')\psi_{n\mathbf{k}}(\mathbf{r}')}{\hbar\omega - (\epsilon_{m\mathbf{k}'} - \epsilon_{n\mathbf{k}}) + i\hbar\eta}.\end{aligned} \tag{42}$$

Based on the Kohn–Sham susceptibility, χ_0, the many-body susceptibility can be calculated via a Dyson-like equation[204]

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',\omega) = \chi_0(\mathbf{r},\mathbf{r}',\omega) + \iint d\mathbf{r}_1 d\mathbf{r}_2\, \chi_0(\mathbf{r},\mathbf{r}_1,\omega)\\ \times K_{\mathrm{Hxc}}(\mathbf{r}_1,\mathbf{r}_2,\omega)\chi(\mathbf{r}_2,\mathbf{r}',\omega).\end{aligned} \tag{43}$$

Here, electronic interactions are accounted for via the Hartree-XC kernel, which is defined in terms of the effective potentials of time-dependent DFT (TDDFT),[205]

$$K_{\mathrm{Hxc}}(\mathbf{r},\mathbf{r}',t-t') = v_c(\mathbf{r}-\mathbf{r}') + f_{\mathrm{xc}}(\mathbf{r},\mathbf{r}',t-t'), \tag{44}$$

where v_c is the Coulomb interaction and f_{xc} is the XC-kernel defined as

$$f_{\mathrm{xc}}(\mathbf{r},\mathbf{r}',t-t') = \frac{\delta v_{\mathrm{xc}}(\mathbf{r},t)}{\delta n(\mathbf{r}',t')}. \tag{45}$$

In prototypical linear-response TDDFT (LR-TDDFT) calculations, the exchange–correlation part of the kernel is either neglected [leading to the random phase approximation (RPA)] or approximated via the adiabatic local density approximation (ALDA),

$$v_{\mathrm{xc}}(\mathbf{r},t) \simeq \left.\frac{\partial[\epsilon_{\mathrm{xc}}(n)n]}{\partial n}\right|_{n(\mathbf{r},t)}, \tag{46}$$

where $\epsilon_{\mathrm{xc}}(n)$ denotes the XC-energy per electron of the homogeneous electron gas of density n. The ALDA is in fact rather restricted, as it only ensures a correct description of the kernel for metals in the long wavelength limit. This implies that it cannot account for excitons in extended systems and, furthermore, it leads to a divergent XC-hole. The latter problem can be resolved by a simple renormalization of the ALDA that regulates the ontop XC-hole and drastically improves the description of local correlations over the ALDA (and RPA).[37]

1. Implementation for periodic systems

In crystalline systems, the susceptibility is periodic with respect to translations on the Bravais lattice, $\chi(\mathbf{r}+\mathbf{R},\mathbf{r}'+\mathbf{R},\omega) = \chi(\mathbf{r},\mathbf{r}',\omega)$. Consequently, the susceptibility can be Fourier-transformed according to

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \iint \frac{d\mathbf{r}d\mathbf{r}'}{\Omega} e^{-i(\mathbf{G}+\mathbf{q})\cdot\mathbf{r}}\chi(\mathbf{r},\mathbf{r}',\omega)e^{i(\mathbf{G}'+\mathbf{q})\cdot\mathbf{r}'}, \tag{47}$$

translating the Dyson Eq. (43) into a plane-wave matrix equation, which is diagonal in the wave vector $\mathbf{q}$ and can be inverted numerically.

By using a plane-wave representation, the crux of the implementation then becomes calculating the reciprocal-space pair densities

$$n_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{q}) = \langle\psi_{n\mathbf{k}}|e^{-i(\mathbf{G}+\mathbf{q})\cdot\mathbf{r}}|\psi_{m\mathbf{k}+\mathbf{q}}\rangle, \tag{48}$$

and to Fourier transform the XC-kernel

$$f_{\mathrm{LDA}}(\mathbf{G}) = \int_{\Omega_{\mathrm{cell}}} d\mathbf{r}\, e^{-i\mathbf{G}\cdot\mathbf{r}} \left.\frac{\partial^2[\epsilon_{\mathrm{xc}}(n)n]}{\partial n^2}\right|_{n(\mathbf{r})}. \tag{49}$$

Crucially, both can be evaluated by adding a PAW correction to the analogous pseudo-quantity, which itself can be evaluated by means of a fast Fourier transform (FFT). Implementational details hereto are reported in Refs. 36 and 43.

The Hartree contribution to the kernel (44) is simply given by the bare Coulomb interaction (here given in atomic units)

$$v_{c,\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \frac{4\pi}{|\mathbf{G}+\mathbf{q}|^2}\delta_{\mathbf{G},\mathbf{G}'}, \tag{50}$$

and can be evaluated analytically when $\mathbf{q}$ is finite. However, in the optical limit, the $\mathbf{G} = \mathbf{0}$ component diverges, and one needs to be careful when inverting the Dyson Eq. (43). In GPAW, this is handled by expanding the pair densities within *kp*-perturbation theory,

$$\langle\psi_{n\mathbf{k}}|e^{-i(\mathbf{G}+\mathbf{q})\cdot\mathbf{r}}|\psi_{m\mathbf{k}+\mathbf{q}}\rangle_{\mathbf{q}\to\mathbf{0},\mathbf{G}=\mathbf{0}} = -i\mathbf{q}\cdot\frac{\langle\psi_{n\mathbf{k}}|\nabla|\psi_{m\mathbf{k}}\rangle}{\epsilon_{m\mathbf{k}} - \epsilon_{n\mathbf{k}}}. \tag{51}$$

In the expansion, the diverging Coulomb term is exactly canceled by the $\mathbf{q}$-dependence of the pair densities so that the product $\chi_0 v_c$ remains finite. For additional details, see Ref. 36.

2. Spectral representation

GPAW offers two different ways of dealing with the frequency dependence of the Kohn–Sham susceptibility (42). One is to evaluate the expression explicitly for the frequencies of interest. This is advantageous if one is interested in a few specific frequencies. The other is to evaluate the associated spectral function

$$\begin{aligned}S_{0,\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \frac{2}{\Omega}\sum_{\mathbf{k}}\sum_{n,m}(f_{n\mathbf{k}} - f_{m\mathbf{k}+\mathbf{q}})\\ \times n_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{q})n_{m\mathbf{k}+\mathbf{q},n\mathbf{k}}(-\mathbf{G}'-\mathbf{q})\\ \times \delta(\hbar\omega - [\epsilon_{m\mathbf{k}+\mathbf{q}} - \epsilon_{n\mathbf{k}}]),\end{aligned} \tag{52}$$

12 March 2024 13:19:04

© Author(s) 2024

from which the susceptibility can be obtained via a Hilbert transform,

$$\chi_{0,\mathbf{GG}'}(\mathbf{q},\omega) = \lim_{\eta\to 0^+}\int_{-\infty}^{\infty} d\omega' \, \frac{S_{0,\mathbf{GG}'}(\mathbf{q},\omega')}{\omega' - \omega + i\eta}. \tag{53}$$

To converge the Hilbert transform, the spectral function is evaluated on a nonlinear frequency grid spanning the range of eigenenergy differences included in Eq. (52). Although the calculation becomes more memory-intensive as a result, it is usually faster to compute χ_0 via its spectral function. Furthermore, since S_0 is linearly interpolatable, the tetrahedron method can be employed to improve convergence with respect to **k**-points.

B. Dielectric function

The longitudinal part of the dielectric tensor is related to the susceptibility χ as

$$\varepsilon^{-1}_{\mathbf{GG}'}(\mathbf{q},\omega) = \delta_{\mathbf{GG}'} + v_{c,\mathbf{GG}}(\mathbf{q})\chi_{\mathbf{GG}'}(\mathbf{q},\omega). \tag{54}$$

From the dielectric matrix, the macroscopic dielectric function, including local field corrections, is given by

$$\varepsilon_M(\mathbf{q},\omega) = \frac{1}{\varepsilon^{-1}_{00}(\mathbf{q},\omega)}, \tag{55}$$

from which it is straightforward to extract the optical absorption spectrum

$$\mathrm{ABS} = \mathrm{Im}\ \varepsilon_M(\mathbf{q}\to 0,\omega), \tag{56}$$

as well as the electron energy-loss spectrum

$$\mathrm{EELS} = -\mathrm{Im}\,\frac{1}{\varepsilon_M(\mathbf{q},\omega)}. \tag{57}$$

It is also possible to define a symmetrized version of the dielectric matrix as

$$\tilde{\varepsilon}_{\mathbf{GG}'}(\mathbf{q},\omega) = v_{\mathbf{G}}^{-1/2}(\mathbf{q})\varepsilon_{\mathbf{GG}'}(\omega,\mathbf{q})v_{\mathbf{G}'}^{1/2}(\mathbf{q}). \tag{58}$$

The susceptibility can be evaluated at the RPA level [by setting V_{xc} in Eq. (43) to zero] or using one of the XC-kernels implemented in GPAW such as the local ALDA, the non-local renormalized ALDA (rALDA),[37] or the bootstrap kernel.[206]

The evaluation χ_0 is computationally demanding since it involves an integration over the Brillouin zone (BZ) as well as a summation over occupied and unoccupied states. The **k**-point convergence can be increased substantially with the tetrahedron method.[207] Contrary to simple point integration, where the δ-function in Eq. (52) is replaced by a smeared-out Lorentzian, the tetrahedron method utilizes linear interpolation of the eigenvalues and matrix elements.

In Fig. 6, we compare the dielectric function computed using the two different integration methods for two prototypical cases: (semiconducting) Si and (metallic) Ag. Since Si is semiconducting, there are no low-energy excitations, and consequently, the imaginary part of the dielectric function is zero, while the real part is flat for low frequencies. The point integration and tetrahedron integration yield the same value for $\omega \to 0$, but the tetrahedron integration avoids the unphysical oscillations of ε at higher frequencies exhibited by the point integration due to the finite **k**-point sampling. For metals, the **k**-point convergence is even slower, and the difference between point integration and tetrahedron integration is thus even more pronounced for Ag. By increasing the **k**-point sampling, the point integration results will eventually approach the results obtained with the tetrahedron method, but at a much higher computational cost.

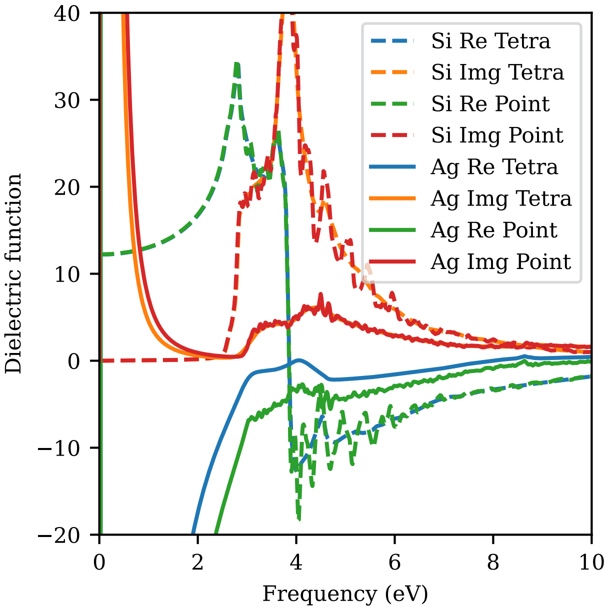


FIG. 6. Real and imaginary parts of the dielectric function for Ag (solid) and Si (dashed) using the tetrahedron and point integration methods.

1. *Screening in low-dimensional systems*

In a three dimensional (3D) bulk crystal, the macroscopic dielectric function is related to the macroscopic polarizability α_M of the material as

$$\varepsilon_M(\mathbf{q},\omega) = 1 + 4\pi\alpha_M(\mathbf{q},\omega). \tag{59}$$

Since ε_M is related to the macroscopic average of the induced potential, it will depend on the unit cell in low-dimensional systems and tend to unity when increasing the cell in the non-periodic direction. In contrast, it is straightforward to generalize the definition of polarizability so that it measures the induced dipole moment per length or area rather than volume. The d-dimensional polarizability is thus defined as $\alpha^d_M = \Omega^d\alpha_M$, where Ω^d is the cell volume with the periodic directions projected out. For example, for a 2D material, Ω^d is simply the length of the unit cell in the non-periodic direction. To improve convergence with respect to the size of the unit cell, the Coulomb kernel is truncated using the reciprocal-space method introduced in Ref. 208. This enables efficient calculations of the dielectric properties of low-dimensional materials.[209,210] This does, however, imply

12 March 2024 13:19:04

that the polarizability cannot be evaluated directly from the dielectric constant (which is just one), but one may obtain it directly from the susceptibility as

$$\alpha_M = -\lim_{\mathbf{q}\to 0} \frac{\Omega^d}{q^2} \chi_{00}(\mathbf{q}). \tag{60}$$

C. Adiabatic-connection fluctuation–dissipation theorem

The adiabatic-connection fluctuation–dissipation theorem (ACFDT) is a highly promising approach for constructing accurate correlation functionals with non-local characteristics. Unlike regular XC-functionals, the ACFDT correlation functionals do not rely on error cancellation between exchange and correlation. Instead, the ACFDT provides an exact theory for the electronic correlation energy E_c in terms of the interacting electronic susceptibility, which can be combined with the exact exchange,

$$E_c = -\int_0^\infty \frac{d\omega}{2\pi} \int_0^1 d\lambda\, \mathrm{Tr}\left[v_c \chi^\lambda(\omega) - v_c \chi^{\mathrm{KS}}(\omega)\right]. \tag{61}$$

The response function can be expressed in terms of the Kohn–Sham response function and the exchange–correlation kernel f_{xc}

$$\chi^\lambda(\omega) = \chi^{\mathrm{KS}}(\omega) + \chi^{\mathrm{KS}}(\omega)[\lambda v_c + f_{\mathrm{xc}}(\omega)]\chi^\lambda(\omega). \tag{62}$$

The random phase approximation (RPA) can be derived from ACFDT if the XC-kernel f_{xc} is neglected. RPA has the strength of capturing non-local correlation effects and provides high accuracy across different bonding types, including van der Waals interactions.[29,211–213]

Simple exchange–correlation kernels can also be incorporated into the response function, such as the adiabatic LDA (ALDA) kernel. However, the locality of adiabatic kernels leads to divergent characteristics of the pair-correlation function.[214] This issue can be overcome by a renormalization (r) scheme,[37] which is implemented in GPAW as rALDA. This class of renormalized kernels provides a significantly better description of the short-range correlations and, hence, also yields highly accurate total correlation energies.[215–218]

D. Magnetic response

The LR-TDDFT framework described in Sec. VI A can be generalized to include spin degrees of freedom.[43,219] Using the four-component density as the basic variable, $n^\mu(\mathbf{r})$ (where $\mu \in \{0, x, y, z\}$), one can define the four-component susceptibility tensor,

$$\chi^{\mu\nu}(\mathbf{r}, \mathbf{r}', t - t') = -\frac{i}{\hbar}\theta(t - t')\langle[\hat{n}_0^\mu(\mathbf{r}, t), \hat{n}_0^\nu(\mathbf{r}', t')]\rangle_0. \tag{63}$$

In a similar fashion to Eq. (43), the many-body $\chi^{\mu\nu}$ can be calculated from the corresponding susceptibility tensor of the Kohn–Sham system, $\chi^{\mu\nu}_{\mathrm{KS}}$. In the most general case, the Dyson equation for $\chi^{\mu\nu}$ is a matrix equation in the μ and ν indices, explicitly coupling the charge and spin degrees of freedom. However, for collinear magnetic systems, the transverse components of the susceptibility decouple from the rest in the absence of spin–orbit coupling.

1. Transverse magnetic susceptibility

By taking the spins to be polarized along the z-direction, the transverse magnetic susceptibility of collinear nonrelativistic systems may be expressed in terms of the spin-raising and spin-lowering density operators $\hat{n}^+(\mathbf{r}) = (\hat{n}^x(\mathbf{r}) + i\hat{n}^y(\mathbf{r}))/2 = \hat{\psi}_\uparrow^\dagger(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r})$ and $\hat{n}^-(\mathbf{r}) = (\hat{n}^x(\mathbf{r}) - i\hat{n}^y(\mathbf{r}))/2 = \hat{\psi}_\downarrow^\dagger(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r})$. In the ALDA, the Dyson equation for χ^{+-} then becomes a scalar one,[43]

$$\begin{aligned}\chi^{+-}(\mathbf{r}, \mathbf{r}', \omega) = \chi^{+-}_{\mathrm{KS}}(\mathbf{r}, \mathbf{r}', \omega) + \int d\mathbf{r}_1 \\ \times \chi^{+-}_{\mathrm{KS}}(\mathbf{r}, \mathbf{r}_1, \omega) f^{-+}_{\mathrm{LDA}}(\mathbf{r}_1)\chi^{+-}(\mathbf{r}_1, \mathbf{r}', \omega),\end{aligned} \tag{64}$$

where the transverse XC-kernel is given by $f^{-+}_{\mathrm{LDA}} = 2W_{\mathrm{LDA}}(\mathbf{r})/n^z(\mathbf{r})$. The susceptibilities χ^{+-} and χ^{+-}_{KS} are themselves defined via the Kubo formula (63), where the latter can be evaluated explicitly in complete analogy to Eq. (42),

$$\begin{aligned}\chi^{+-}_{\mathrm{KS}}(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta\to 0^+} \sum_{n\mathbf{k}} \sum_{m\mathbf{k}'} \left(f_{n\mathbf{k}\uparrow} - f_{m\mathbf{k}'\downarrow}\right) \\ \times \frac{\psi^*_{n\mathbf{k}\uparrow}(\mathbf{r})\psi_{m\mathbf{k}'\downarrow}(\mathbf{r})\psi^*_{m\mathbf{k}'\downarrow}(\mathbf{r}')\psi_{n\mathbf{k}\uparrow}(\mathbf{r}')}{\hbar\omega - (\epsilon_{m\mathbf{k}'\downarrow} - \epsilon_{n\mathbf{k}\uparrow}) + i\hbar\eta}.\end{aligned} \tag{65}$$

In GPAW, the transverse magnetic susceptibility is calculated using a plane-wave basis as described in Sec. VI A 1, with the notable exception that no special care is needed to treat the optical limit since the Hartree kernel plays no role in the Dyson Eq. (64). In terms of the temporal representation, it is at the time of this writing only possible to do a literal evaluation of Eq. (65) at the frequencies of interest. For metals, this means that η has to be left as a finite broadening parameter, which has to be carefully chosen depending on the **k**-point sampling (see Ref. 43).

2. The spectrum of transverse magnetic excitations

Based on the transverse magnetic susceptibility, one may calculate the corresponding spectral function[43]

$$S^{+-}(\mathbf{r}, \mathbf{r}', \omega) = -\frac{1}{2\pi i}\left[\chi^{+-}(\mathbf{r}, \mathbf{r}', \omega) - \chi^{-+}(\mathbf{r}', \mathbf{r}, -\omega)\right], \tag{66}$$

which directly governs the energy dissipation in a collinear magnet relating to induced changes in the spin projection along the z-axis S_z. In particular, one can decompose the spectrum into contributions from majority and minority spin excitations, $S^{+-}(\mathbf{r}, \mathbf{r}', \omega) = A^{+-}(\mathbf{r}, \mathbf{r}', \omega) - A^{-+}(\mathbf{r}', \mathbf{r}, -\omega)$, that is, into spectral functions for the excited states where S_z has been lowered or raised by one unit of spin angular momentum, respectively,

$$A^{+-}(\mathbf{r}, \mathbf{r}', \omega) = \sum_{\alpha>0} \langle 0|\hat{n}^+(\mathbf{r})|\alpha\rangle\langle\alpha|\hat{n}^-(\mathbf{r}')|0\rangle\delta(\hbar\omega - [E_\alpha - E_0]). \tag{67}$$

Here, α iterates the system eigenstates, with $\alpha = 0$ denoting the ground state.

For collinear magnetic systems, the spectrum S^{+-} is composed of two types of excitations: collective spin-wave excitations (referred to as magnons) and excitations in the Stoner-pair continuum (electron–hole pairs of opposite spin). Since GPAW employs a plane-wave representation of the spectrum, $S^{+-}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega)$, one can directly compare the calculational output to the inelastic neutron scattering cross-section measured in experiments.[220]

12 March 2024 13:19:04

© Author(s) 2024

In particular, one can extract the magnon dispersion directly by identifying the position of the peaks in the spectrum diagonal (see Fig. 7). In this way, GPAW allows the user to study various magnon phenomena in magnetic systems of arbitrary collinear order, such as nonanalytic dispersion effects in itinerant ferromagnets,[43] correlation-driven magnetic phase transitions,[221] and the emergence of distinct collective modes inside the Stoner continuum of an antiferromagnet.[222]

Additionally, one can analyze the spectrum more intricately by extracting the majority and minority eigenmodes from S^{+-} as the positive and negative eigenvalues, respectively. Contrary to the analysis of the plane-wave diagonal, this makes it possible to completely separate the analysis of each individual magnon lineshape as well as the many-body Stoner continuum (see Fig. 7).

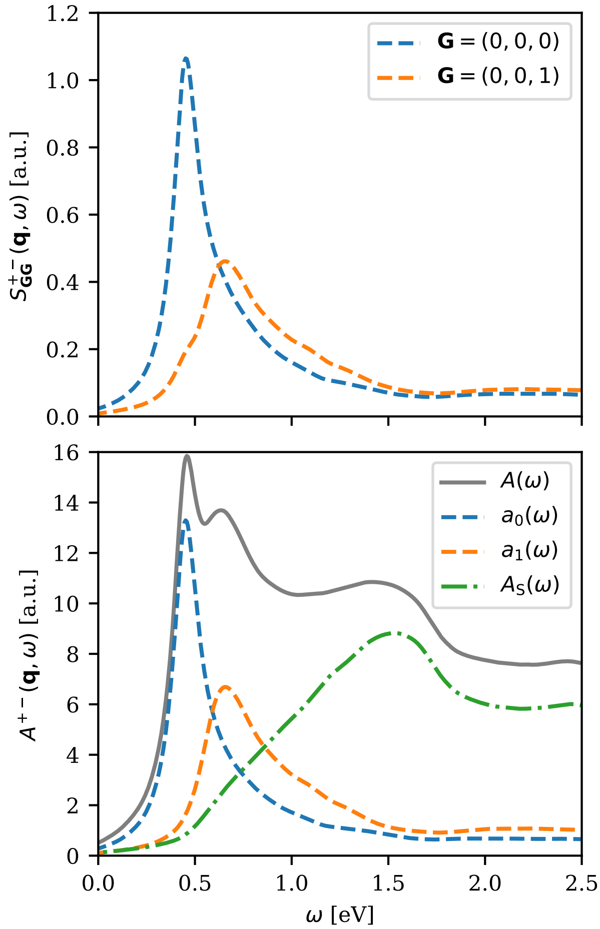


FIG. 7. Spectrum of transverse magnetic excitations for ferromagnetic hcp-Co evaluated at $\mathbf{q} = 5\mathbf{q}_{\rm M}/6$. The spectrum was calculated using eight empty-shell bands per atom, a plane-wave cutoff of 800 eV, a $(60, 60, 36)$ **k**-point mesh, and a spectral broadening of $\eta = 50$ meV. In the upper panel, the spectrum diagonal is depicted for the first and second Brillouin zones out of the hexagonal plane, from which the acoustic and optical magnon frequencies can be respectively extracted. In the lower panel, the spectrum of majority excitations is shown. The full spectral weight $A(\omega)$ is calculated as the sum of all positive eigenvalues of S^{+-}, the acoustic and optical mode lineshapes $a_0(\omega)$ and $a_1(\omega)$ are obtained via the two largest eigenvalues (which are significantly larger than the rest), and the Stoner spectrum is extracted as the difference $A_S(\omega) = A(\omega) - a_0(\omega) - a_1(\omega)$.

3. Linear response MFT

Not only can the transverse magnetic susceptibility be used to study magnetic excitations in a literal sense, but one can also use it to map the spin degrees of freedom to a classical Heisenberg model,

$$E_{\rm H} = -\frac{1}{2}\sum_{i,j}\sum_{a,b} J_{ij}^{ab}\,\mathbf{u}_{ia}\cdot\mathbf{u}_{jb}, \tag{68}$$

where i, j and a, b denote the indices of the Bravais lattice and the magnetic sublattice, respectively, $\mathbf{u}_{ia}$ being the direction of spin-polarization of the given magnetic site. Based on the magnetic force theorem (MFT), the LSDA Heisenberg exchange parameters J_{ij}^{ab} can be calculated from the reactive part of the static Kohn–Sham susceptibility[42] and the effective magnetic field $B_{\rm xc}(\mathbf{r}) = \delta E_{\rm xc}/\delta m(\mathbf{r})$, using the well-known Liechtenstein MFT formula[189]

$$J_{ij}^{ab} = -2\int_{\Omega_{ia}} d\mathbf{r}\int_{\Omega_{jb}} d\mathbf{r}'\, B_{\rm xc}(\mathbf{r})\chi'^{+-}_{\rm KS}(\mathbf{r},\mathbf{r}')B_{\rm xc}(\mathbf{r}'). \tag{69}$$

Here, Ω_{ia} denotes the site volume, which effectively defines the Heisenberg model (68).

Using GPAW's plane-wave representation of $\chi'^{+-}_{\rm KS}$, one can directly compute the lattice Fourier-transformed exchange parameters,[42]

$$\bar{J}^{ab}(\mathbf{q}) = \sum_i J_{0i}^{ab} e^{i\mathbf{q}\cdot\mathbf{R}_i}, \tag{70a}$$

$$= -\frac{2}{\Omega_{\rm cell}} B_{\rm xc}^{\dagger} K^{a\dagger}(\mathbf{q})\chi'^{+-}_{\rm KS}(\mathbf{q})K^{b}(\mathbf{q})B_{\rm xc}, \tag{70b}$$

where the right-hand side of the second equality is written in a plane-wave basis and K^a denotes the sublattice site-kernel,

$$K^{a}_{\mathbf{GG}'}(\mathbf{q}) = \frac{1}{\Omega_{\rm cell}}\int_{\Omega_{0a}} d\mathbf{r}\, e^{-i(\mathbf{G}-\mathbf{G}'+\mathbf{q})\cdot\mathbf{r}}. \tag{71}$$

Since an *a priori* definition for magnetic site volumes does not exist, GPAW supplies functionality to calculate exchange parameters based on spherical, cylindrical, and/or parallelepipedic site configurations of variable size.

Upon calculation of the exchange parameters $\bar{J}^{ab}(\mathbf{q})$, it is straightforward to compute the magnon dispersion within the classical Heisenberg model using linear spin-wave theory and to estimate thermal quantities such as the Curie temperature (see, e.g., Ref. 42). In Fig. 8, the MFT magnon dispersion of hcp-Co is compared to the majority magnon spectrum calculated within LR-TDDFT. For Co, the two are in excellent agreement except for the dispersion of the optical magnon mode along the K − M − Γ high-symmetry path, where MFT underestimates the magnon frequency and neglects the fine structure of the spectrum. The fine structure of the ALDA spectrum appears due to the overlap between the magnon mode and the Stoner continuum. This gives rise to so-called Kohn anomalies (nonanalytical points in the magnon dispersion), which are a trademark of itinerant electron magnetism. Since itinerancy is largely ignored in a localized spin model such as (68), one cannot generally expect to capture such effects.

12 March 2024 13:19:04

© Author(s) 2024

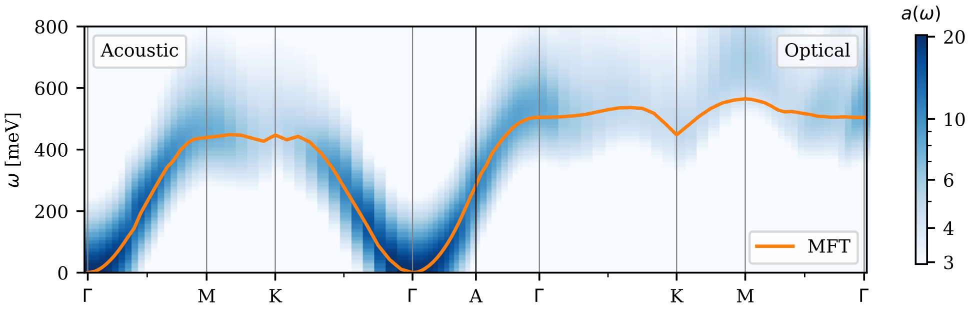


FIG. 8. Magnon spectrum of ferromagnetic hcp-Co calculated using ALDA LR-TDDFT (shown as a heat map) compared to the spin-wave dispersion of Liechtenstein MFT. The acoustic magnon mode $a_0(\omega)$ is shown to the left of the A-point, while the optical magnon mode $a_1(\omega)$ is shown to the right. Note that the two modes are degenerate at both the A and K high-symmetry points. The calculations were carried out using eight empty-shell bands per atom, a plane-wave cutoff of 800 eV, and a $(36, 36, 24)$ **k**-point mesh. A finite value of $\eta = 100$ meV was used to broaden the ALDA spectrum. For the MFT calculations, the dispersion was calculated using linear spin-wave theory based on a Heisenberg model of closed-packed spherical sites centered at each of the Co atoms.

E. GW approximation

GPAW supports standard G_0W_0 quasiparticle (QP) calculations based on a first-order perturbative treatment of the linearized QP equation[9]

$$E^{\mathrm{QP}}_{n\mathbf{k}\sigma} = \varepsilon_{n\mathbf{k}\sigma} + Z_{n\mathbf{k}\sigma}\langle\psi_{n\mathbf{k}\sigma}|\Sigma_{\mathrm{GW}}(\varepsilon_{n\mathbf{k}\sigma}) + v_{\mathrm{x}} - v_{\mathrm{xc}}|\psi_{n\mathbf{k}\sigma}\rangle, \tag{72}$$

where $\varepsilon_{n\mathbf{k}\sigma}$ and $\psi_{n\mathbf{k}\sigma}$ are Kohn–Sham eigenvalues and wave functions, and v_{xc} and v_{x} are the local XC-potential and non-local exchange potential, respectively. Σ_{GW} is the (dynamical part of) the GW self-energy whose frequency-dependence is accounted for in first order by the renormalization factor $Z_{n\mathbf{k}\sigma} = (1 - \mathrm{Re}\,\Sigma'_{\mathrm{GW}}(\varepsilon_{n\mathbf{k}\sigma}))^{-1}$. As indicated by the σ-index, spin-polarized G_0W_0 calculations are supported.

The GW self-energy is calculated on a plane-wave basis using full frequency integration along the real axis to evaluate the convolution between G and W.[38] Compared to alternative schemes employing contour deformation techniques or analytical continuation,[223–225] this approach is time-consuming but numerically accurate and can provide the full spectral function. A highly efficient and accurate evaluation of the self-energy based on a multipole expansion of the screened interaction, W,[226] is currently being implemented, and a GPU version of the full GW code is under development.

An important technical issue concerns the treatment of the head and wings of W ($\mathbf{G} = 0$ and/or $\mathbf{G}' = 0$, respectively) in the $\mathbf{q} \to 0$ limit. The divergence of the Coulomb interaction appears both in the evaluation of the inverse dielectric matrix $\varepsilon^{-1}_{\mathbf{GG}'}$ and in the subsequent evaluation of the screened interaction

$$W_{\mathbf{GG}'}(\mathbf{q},\omega) = \sum_{\mathbf{G}''} \varepsilon^{-1}_{\mathbf{GG}''}(\mathbf{q},\omega)v_{c,\mathbf{G}''\mathbf{G}'}(\mathbf{q}).$$

For 3D bulk crystals, GPAW obtains these components by evaluating $\varepsilon^{-1}_{\mathbf{GG}'}$ on a dense **k**-grid centered at $\mathbf{k} = 0$, while $v_{c,\mathbf{GG}'}$ can be integrated numerically or analytically around the Γ-point.

For low-dimensional structures, in particular atomically thin 2D semiconductors, GPAW can use a truncated Coulomb interaction to avoid interactions between periodical images when evaluating W.[208] It has been shown that the use of a truncated Coulomb kernel leads to slower **k**-point convergence.[209] To mitigate this drawback, a special 2D treatment of $W(q)$ at $q = 0$ can be applied to significantly improve the **k**-point convergence.[227] A detailed account of the GW implementation in GPAW can be found in Ref. 38.

Figure 9 shows two matrix elements of the dynamical GW self-energy for the valence and conduction band states at the gamma point of bulk silicon. As can be seen, the agreement with the corresponding quantities obtained with the Yambo GW code[228] is striking.

F. Bethe–Salpeter equation (BSE)

In addition to the LR-TDDFT discussed in Sec. VI A, the interacting response function may be approximated by solving the Bethe–Salpeter equation (BSE).[11] In particular, for a certain wave vector **q**, one may obtain the two-particle excitations by diagonalizing the Hamiltonian

$$\begin{aligned} H_{\substack{\mathbf{k}_2 m_1 m_2\\ \mathbf{k}_4 m_3 m_4}}(\mathbf{q}) &= \delta_{m_1 m_3}\delta_{m_2 m_4}\delta_{\mathbf{k}_2\mathbf{k}_4}(\varepsilon_{m_1\mathbf{k}_2+\mathbf{q}} - \varepsilon_{m_2\mathbf{k}_2}) \\ &\quad - (f_{m_1\mathbf{k}_2+\mathbf{q}} - f_{m_2\mathbf{k}_2})K_{\substack{\mathbf{k}_2 m_1 m_2\\ \mathbf{k}_4 m_3 m_4}}(\mathbf{q}), \end{aligned} \tag{73}$$

where $\varepsilon_{\mathbf{k}m}$ are the Kohn–Sham eigenvalues and $f_{\mathbf{k}m}$ are the associated occupation numbers. The kernel is defined by $K = v_c - W$ with

$$v_c = \langle m_2\mathbf{k}_2, m_3\mathbf{k}_4 + \mathbf{q}|\hat{v}_c|m_1\mathbf{k}_2 + \mathbf{q}, m_4\mathbf{k}_4\rangle, \tag{74}$$

$$W = \langle m_2\mathbf{k}_2, m_3\mathbf{k}_4 + \mathbf{q}|\hat{W}|m_4\mathbf{k}_4, m_1\mathbf{k}_2 + \mathbf{q}\rangle, \tag{75}$$

where $\hat{W}$ is the static screened Coulomb interaction defined in Sec. VI E. The matrix elements of the kernel are evaluated on a plane-wave basis, where they are easily expressed in terms of the pair densities (48) and the reciprocal-space representation of the dielectric matrix (54).

In the Tamm-Dancoff approximation, one neglects the coupling of the resonant and anti-resonant blocks of the BSE Hamiltonian (those corresponding to positive and negative transition energies, respectively). This makes the BSE Hamiltonian Hermitian.[11] The interacting retarded response function may then be written as

$$\chi^{\mathbf{GG}'}(\mathbf{q},\omega) = \frac{1}{\Omega}\sum_{\lambda}\frac{B_\lambda(\mathbf{q},\mathbf{G})C^*_\lambda(\mathbf{q},\mathbf{G}')}{\omega - E_\lambda(\mathbf{q}) + i\eta}, \tag{76}$$

12 March 2024 13:19:04

© Author(s) 2024

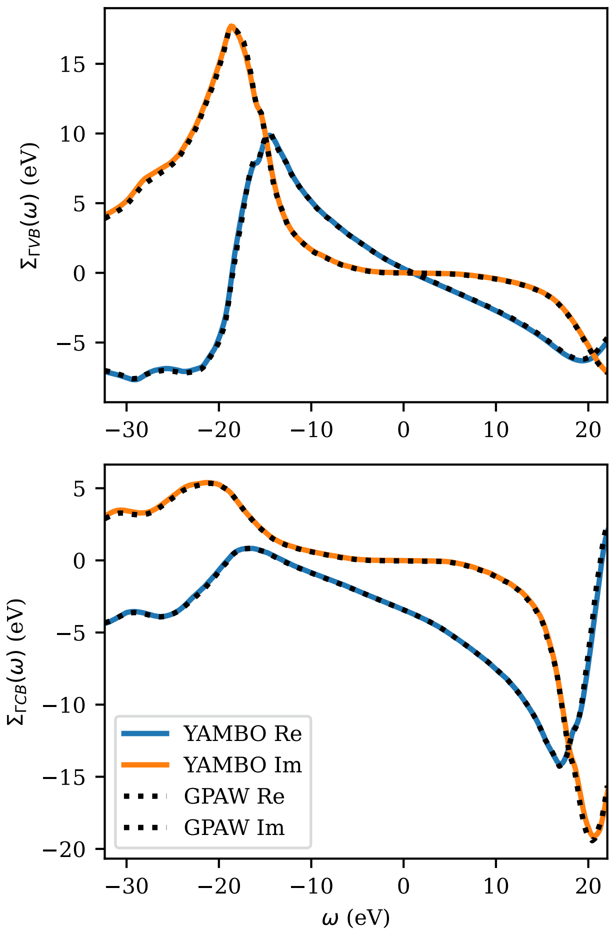


FIG. 9. The real (Re) and imaginary (Im) parts of the self-energy matrix elements at the Gamma point for valence (top) and conduction (bottom) bands evaluated with Yambo and GPAW for bulk silicon. Both codes are using full frequency integration with a broadening of 0.1 eV. Yambo is using norm-conserving pseudo-potentials, and GPAW is its standard PAW setup. The k-point grid was 12 × 12 × 12, the plane wave cutoff was 200 eV, and the number of bands was 200 for both codes. The results are virtually indistinguishable.

where

$$B_\lambda(\mathbf{q},\mathbf{G}) = \sum_{\mathbf{k}m_1m_2} n_{m_2\mathbf{k}m_1\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{q})A^\lambda_{m_1m_2\mathbf{k}}(\mathbf{q}), \tag{77}$$

$$C_\lambda(\mathbf{q},\mathbf{G}) = \sum_{\mathbf{k}m_1m_2} n_{m_2\mathbf{k}m_1\mathbf{k}+\mathbf{q}}(\mathbf{G}+\mathbf{q})A^\lambda_{m_1m_2\mathbf{k}}(\mathbf{q})(f_{m_1\mathbf{k}+\mathbf{q}} - f_{m_2\mathbf{k}}), \tag{78}$$

and $E_\lambda(\mathbf{q})$ denotes the eigenvalue of the Hamiltonian (73) corresponding to the eigenvector $A^\lambda_{m_1m_2\mathbf{k}}(\mathbf{q})$.

In GPAW, the construction of the BSE Hamiltonian (73) proceeds in two steps.[39] First, the static screened interaction is calculated at all inequivalent **q**-points in a plane-wave basis, and the kernel is then subsequently expressed in a basis of two-particle KS states. The first step is efficiently parallelized over either states or **k**-points, and the second step is parallelized over pair densities. The Hamiltonian elements are thus distributed over all CPUs, and the diagonalization is carried out using ScaLAPACK such that the full Hamiltonian is never collected on a single CPU. The dimension of the BSE Hamiltonian and memory requirements are, therefore, only limited by the number of CPUs used for the calculation. We note that the implementation is not limited to the Tamm–Dancoff approximation, but calculations become more demanding without it. The response function may be calculated for spin-paired as well as spin-polarized systems, and spin–orbit coupling can be included non-selfconsistently.[45,229] In low-dimensional systems, it is important to eliminate spurious screening from periodic images of the structure, which is accomplished with the truncated Coulomb interaction in Ref. 208.

The most important application of BSE is arguably the calculation of optical absorption spectra of solids, where BSE provides an accurate account of the excitonic effects that are not captured by semi-local approximations for K_{Hxc} (44). In 2D systems, the excitonic effects are particularly pronounced due to inefficient screening,[209,229] and in Fig. 10, we show the 2D polarizability of WS_2 calculated from BSE with $\mathbf{q} = \mathbf{0}$. Comparing with Fig. 3, it is observed that the absorption edge is expected to be located at the *K* point, where spin–orbit coupling splits the highest valence band by 0.45 eV. This splitting is seen as two excitonic peaks below the bandgap, which is interpreted as distinct excitons originating from the highest and next-highest valence bands (the splitting of the lowest conduction band is negligible in this regard). For comparison, we also show the RPA polarizabilty, obtained with the BSE module by neglecting the screened interaction in the kernel, and this shows the expected absorption edge at the bandgap. This yields identical results to the Dyson equation approach of Sec. VI A but has the advantage that the eigenvalues and weights of Eq. (76) are obtained directly, such that the artificial broadening η may be varied without additional computational cost. The eigenstate decomposition also allows one to access "dark states," and the BSE calculation reveals two (one for each valley) triplet-like excitons that are situated 70 meV below the lowest bright exciton in Fig. 10.

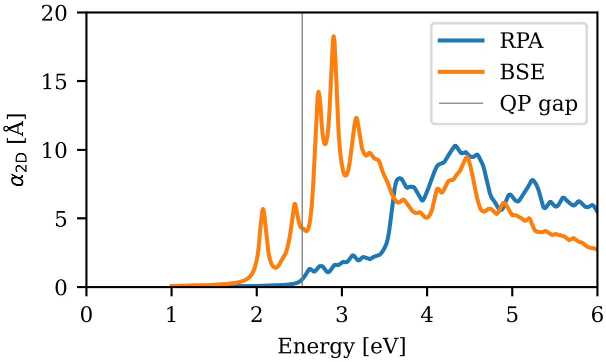


FIG. 10. 2D polarizability of WS_2 calculated from the BSE and the RPA. For this calculation, we included spin–orbit coupling and used the 2D Coulomb truncation to eliminate screening from periodic images. A gamma-centered uniform **k**-point grid of 48 × 48 was applied, and eight valence states and eight conduction states (shifted by 1 eV to match the GW bandgap[45]) were included in the Tamm–Dancoff approximation. This yields a BSE Hamiltonian of size $N \times N$ with $N = 147\,456$, which is easily diagonalized with ScaLAPACK on 240 CPUs.

12 March 2024 13:19:04

© Author(s) 2024

In addition to optical properties, GPAW allows for solving the BSE at finite $\mathbf{q}$, which can be used to obtain plasmon dispersion relations from the electron energy loss spectrum (EELS) (57) and magnon dispersions from the transverse magnetic susceptibility of Sec. VI D.[230]

G. Electron–phonon coupling

The electron–phonon coupling is the origin of several important material properties, ranging from electrical and thermal conductivity to superconductivity. In addition, it provides access to the deformation potential, which can be used to obtain transport properties for electrons and holes in solids.[231]

The first-order electron–phonon coupling matrix $g^{\nu}_{mn}(\mathbf{k},\mathbf{q})$ measures the strength of the coupling between a phonon branch ν with wave vector $\mathbf{q}$ and frequency ω_ν and the electronic states $m(\mathbf{k}+\mathbf{q})$ and $n(\mathbf{k})$,[232,233]

$$g^{\nu}_{mn}(\mathbf{k},\mathbf{q}) = \sqrt{\frac{\hbar}{2m_0\omega_\nu}}M^{\nu}_{mn}(\mathbf{k},\mathbf{q}), \tag{79}$$

with

$$M^{\nu}_{mn}(\mathbf{k},\mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}}|\nabla_u v^{\mathrm{KS}}\cdot \mathbf{e}_\nu|\psi_{n\mathbf{k}}\rangle. \tag{80}$$

Here, m_0 is the sum of the masses of all the atoms in the unit cell, ∇_u denotes the gradient with respect to atomic displacements, and $\mathbf{e}_\nu$ projects the gradient in the direction of the phonon displacements. In the case of the three translational modes at $|\mathbf{q}| = 0$, the matrix element g^{ν}_{mn} vanishes as a consequence of the acoustic sum rule.[233]

In GPAW, $\nabla_u v^{\mathrm{KS}}(\mathbf{r})$ is determined using a finite-difference method with a supercell description of the system. This step can be performed in any of the wave function representations available in GPAW. The derivative is then projected onto a set of atomic orbitals from an LCAO basis ϕ_{NM}, where N denotes the cell index and M the orbital index,

$$\mathbf{g}^{sc}_{\substack{NM\\N'M'}} = FT\left[\langle \phi_{NM}(\mathbf{k})|\nabla_u v^{\mathrm{KS}}|\phi_{N'M'}(\mathbf{k})\rangle\right]. \tag{81}$$

The Fourier transform from the Bloch to the real space representation makes it possible to compute M^{ν}_{mn} for arbitrary $\mathbf{q}$. Finally, the electron–phonon coupling matrix is obtained by projecting the matrix corresponding to the supercell into the primitive unit cell bands m, n and phonon modes ν,

$$M^{\nu}_{mn}(\mathbf{k},\mathbf{q}) = \sum_{\substack{NM\\N'M'}} C^{*}_{mM}C_{nM'}\mathbf{g}^{sc}_{\substack{NM\\N'M'}}\cdot \mathbf{u}_{q\nu}e^{i[(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_N-\mathbf{k}\cdot\mathbf{R}'_N]}, \tag{82}$$

where C_{nM} are the LCAO coefficients and $\mathbf{u}_{q\nu}$ are mass-scaled phonon displacement vectors.

H. Raman spectrum

The Raman effect describes inelastic light scattering, where vibrational modes are excited within the material. Resonant and non-resonant Raman spectra of finite systems such as molecules can be calculated in various approximations[234] using the corresponding interfaces in ASE. The Stokes Raman intensity is then written as

$$I(\omega) = I_0\sum_\nu \frac{n_\nu+1}{\omega_\nu}\Big|\sum_{\alpha,\beta} u^{\alpha}_{\mathrm{in}}R^{\nu}_{\alpha\beta}u^{\beta}_{\mathrm{out}}\Big|^2\delta(\omega-\omega_\nu), \tag{83}$$

where ν denotes phonon mode at $\mathbf{q} = 0$ with a frequency of ω_ν and n_ν is the corresponding Bose–Einstein distribution. Furthermore, u^{α}_{in} and u^{β}_{out} are the polarization vectors of the incoming and outgoing light, and $R^{\nu}_{\alpha\beta}$ denotes the Raman tensor for phonon mode ν.

The predominant approach for calculating $R^{\nu}_{\alpha\beta}$ involves the use of the Kramers–Heisenberg–Dirac (KHD) method. Within the KHD framework, the Raman tensor is determined by taking the derivative (utilizing a finite-difference method) of the electric polarizability concerning the vibrational normal modes. Alternatively, one can employ time-dependent third-order perturbation theory to compute Raman tensors. These two approaches are equivalent when local field effects are negligible.[56] However, each approach comes with its own set of advantages and drawbacks. The KHD method is computationally more efficient but is limited to computing first-order Raman processes. The perturbative approach can be extended to higher-order Raman processes involving multiple phonons, but it is more computationally demanding, necessitating a greater number of bands and a finer k-mesh grid to achieve convergence. The perturbative approach has been implemented in GPAW and is elaborated below, while the KHD method has been implemented in the ASR package,[22] utilizing GPAW as the computational backend.

In the perturbative approach, the Raman tensor $R^{\nu}_{\alpha\beta}$ is given by[56,235]

$$\begin{aligned}R^{\nu}_{\alpha\beta} \equiv \sum_{ijmn\mathbf{k}}\Bigg[&\frac{p^{\alpha}_{ij}(M^{\nu}_{jm}\delta_{in}-M^{\nu}_{ni}\delta_{jm})p^{\beta}_{mn}}{(\hbar\omega_{\mathrm{in}}-\varepsilon_{ji})(\hbar\omega_{\mathrm{out}}-\varepsilon_{mn})}+\frac{p^{\alpha}_{ij}(p^{\beta}_{jm}\delta_{in}-p^{\beta}_{ni}\delta_{jm})M^{\nu}_{mn}}{(\hbar\omega_{\mathrm{in}}-\varepsilon_{ji})(\hbar\omega_\nu-\varepsilon_{mn})}\\&+\frac{p^{\beta}_{ij}(M^{\nu}_{jm}\delta_{in}-M^{\nu}_{ni}\delta_{jm})p^{\alpha}_{mn}}{(-\hbar\omega_{\mathrm{out}}-\varepsilon_{ji})(-\hbar\omega_{\mathrm{in}}-\varepsilon_{mn})}+\frac{p^{\beta}_{ij}(p^{\alpha}_{jm}\delta_{in}-p^{\alpha}_{ni}\delta_{jm})M^{\nu}_{mn}}{(-\hbar\omega_{\mathrm{out}}-\varepsilon_{ji})(\hbar\omega_\nu-\varepsilon_{mn})}\\&+\frac{M^{\nu}_{ij}(p^{\alpha}_{jm}\delta_{in}-p^{\alpha}_{ni}\delta_{jm})p^{\beta}_{mn}}{(-\hbar\omega_\nu-\varepsilon_{ji})(\hbar\omega_{\mathrm{out}}-\varepsilon_{mn})}+\frac{M^{\nu}_{ij}(p^{\beta}_{jm}\delta_{in}-p^{\beta}_{ni}\delta_{jm})p^{\alpha}_{mn}}{(-\hbar\omega_\nu-\varepsilon_{ji})(-\hbar\omega_{\mathrm{in}}-\varepsilon_{mn})}\Bigg]\\&\times f_i(1-f_j)f_n(1-f_m),\end{aligned} \tag{84}$$

where the first term is referred to as the resonant part and the remaining terms represent different time orderings of the interaction in terms of Feynman diagrams. $p^{\alpha}_{nm} = \langle\psi_{n\mathbf{k}}|\hat{p}^{\alpha}|\psi_{m\mathbf{k}}\rangle$ is the momentum matrix element between electronic bands m and n, with transition energy $\varepsilon_{nm} = E_n - E_m$ in polarization direction α, and M^{ν}_{nm} is the electron–phonon coupling strength in the optical limit $|\mathbf{q}| = 0$ as defined in Eq. (82).

Figure 11 shows the polarization-resolved Raman spectrum of bulk $\mathrm{MoS_2}$ in the 2H phase as computed with a laser frequency of $\omega_{\mathrm{in}} = 488$ nm. This example uses only the resonant term, as the other contributions are small in this case. These calculations used a 700 eV plane wave cutoff and a $2 \times 2 \times 2\mathbf{k}$-point mesh in a $3 \times 3 \times 2$ supercell. Force constants and potential changes were calculated simultaneously. The present implementation uses the ASE phonon module as the backend to perform the finite difference displacement and to obtain the phonon eigenvalues and eigenvectors. The convergence of the supercell size for phonons and potential changes must be checked carefully by the users for each given system to achieve the desired precision of line positions and intensities. As Raman spectrum calculations only use $\mathbf{q} = 0$ phonons and electron–phonon matrices, relatively small supercells often yield satisfactory results.

12 March 2024 13:19:04

© Author(s) 2024

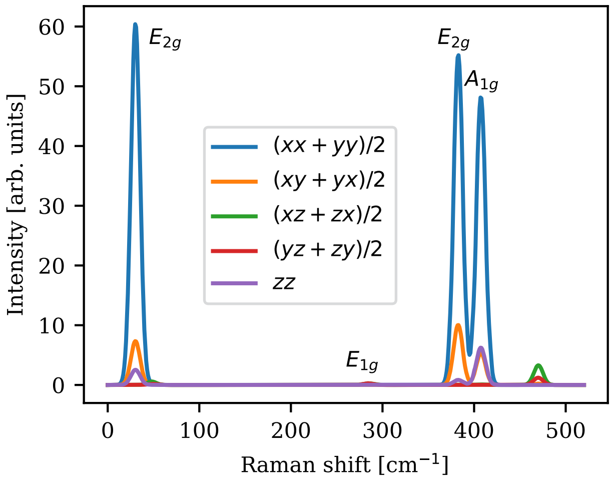


FIG. 11. Polarization-resolved Raman spectrum of bulk MoS$_2$ in the 2H phase at ω = 488 nm. Phonons and potential changes were computed using a 700 eV plane wave cutoff and a 2 × 2 × 2**k**-point mesh in a 3 × 3 × 2 supercell. Each peak has been labeled according to its irreducible representation.

It is worth noting that we have conducted a comparison of the calculated spectra using the ASR package and observed a high level of agreement for several materials, for example, MoS$_2$. The results obtained from both methods closely align with each other in terms of peak positions and dominant peaks, and minor disagreements between the two methods can be attributed to differences in implementation details and the distinct approximations employed by each. Specifically, within the ASR package, we utilized the phonopy package[236] to compute phonon frequencies and eigenvectors, whereas in GPAW, we directly computed phonon frequencies and eigenvectors using ASE's phonon module. Furthermore, in the ASR implementation, we rigorously enforced the symmetry of the polarizability tensor, and the ASR results, therefore, typically exhibit a more accurate adherence to the required symmetry of the Raman tensor compared to the GPAW implementation.

I. Quadratic optical response functions

The nonlinear optical response of materials can be obtained by going beyond first-order perturbation theory. Presently, the GPAW implementation is restricted to second-order response within the dipole approximation and without the inclusion of local field effects. We apply the independent particle approximation, which cannot capture collective behavior such as excitonic effects.[57] A spatially homogeneous incident electric field can be written in terms of its Fourier components as

$$\boldsymbol{\mathcal{E}}(t) = \sum_{\alpha,\omega_1} \mathcal{E}_\alpha(\omega_1)\mathbf{e}_\alpha e^{-i\omega_1 t}, \tag{85}$$

where ω_1 runs over positive and negative frequencies, $\mathbf{e}_\alpha$ denotes the unit vector along the α-direction, and $\mathcal{E}_\alpha(\omega_1)$ is the electric field at frequency ω_1. The induced quadratic polarization density $\mathcal{P}^{(2)}_\gamma(t)$ can then be expressed as

$$\mathcal{P}^{(2)}_\gamma(t) = \epsilon_0 \sum_{\omega_1\omega_2} \sum_{\alpha\beta} \chi^{(2)}_{\gamma\alpha\beta}(\omega_1,\omega_2)\,\mathcal{E}_\alpha(\omega_1)\,\mathcal{E}_\beta(\omega_2)\,e^{-i(\omega_1+\omega_2)t}, \tag{86}$$

where $\chi^{(2)}_{\gamma\alpha\beta}$ is the rank-3 quadratic susceptibility tensor. Due to intrinsic permutation symmetry, i.e., $\chi^{(2)}_{\gamma\alpha\beta}(\omega_1,\omega_2) = \chi^{(2)}_{\gamma\beta\alpha}(\omega_2,\omega_1)$, $\chi^{(2)}_{\gamma\alpha\beta}$ has at most 18 independent elements, which may be further reduced by the Neumann principle and point group symmetries.[237] We note that the corresponding quadratic conductivity tensor is readily derived from $\chi^{(2)}_{\gamma\alpha\beta}$ as $\sigma^{(2)}_{\gamma\alpha\beta}(\omega_1,\omega_2) = -i\epsilon_0\chi^{(2)}_{\gamma\alpha\beta}(\omega_1,\omega_2)(\omega_1 + \omega_2)$ due to the relationship between current density and polarization density.[238]

Among the various response functions that can be calculated from $\chi^{(2)}_{\gamma\alpha\beta}(\omega_1,\omega_2)$, we have implemented second-harmonics generation (SHG) and the shift current tensor. The implementation currently requires time-reversal symmetry, which limits the application to non-magnetic systems. For SHG, the susceptibility tensor is separated into a pure interband term $\chi^{(2e)}_{\gamma\alpha\beta}$ and a mixed term $\chi^{(2i)}_{\gamma\alpha\beta}$ that read

$$\chi^{(2e)}_{\gamma\alpha\beta} \equiv C_0 {\sum_{\mathbf{k},nml}}' \frac{\mathrm{Re}\{r^\gamma_{nm}(r^\alpha_{ml}r^\beta_{\mathrm{ln}} + r^\beta_{ml}r^\alpha_{\mathrm{ln}})\}}{2(\varepsilon_{\mathrm{ln}} - \varepsilon_{ml})} \times \left(\frac{2f_{nm}}{2\hbar\omega - \varepsilon_{mn}} - \frac{f_{nl}}{\hbar\omega - \varepsilon_{\mathrm{ln}}} + \frac{f_{ml}}{\hbar\omega - \varepsilon_{ml}}\right), \tag{87}$$

$$\begin{aligned}\chi^{(2i)}_{\gamma\alpha\beta} \equiv iC_0 {\sum_{\mathbf{k},nm}}' f_{nm}\Bigg[&\frac{2\,\mathrm{Im}\,\{r^\gamma_{nm}(r^\alpha_{mn;\beta} + r^\beta_{mn;\alpha})\}}{\varepsilon_{mn}(2\hbar\omega - \varepsilon_{mn})} \\ &+ \frac{\mathrm{Im}\,\{r^\alpha_{mn}r^\gamma_{nm;\beta} + r^\beta_{mn}r^\gamma_{nm;\alpha}\}}{\varepsilon_{mn}(\hbar\omega - \varepsilon_{mn})} + \frac{\mathrm{Im}\,\{r^\gamma_{nm}(r^\alpha_{mn}\Delta^\beta_{mn} + r^\beta_{mn}\Delta^\alpha_{mn})\}}{\varepsilon^2_{mn}} \\ &\times\left(\frac{1}{\hbar\omega - \varepsilon_{mn}} - \frac{4}{2\hbar\omega - \varepsilon_{mn}}\right) - \frac{\mathrm{Im}\,\{r^\alpha_{mn}r^\beta_{nm;\gamma} + r^\beta_{mn}r^\alpha_{nm;\gamma}\}}{2\varepsilon_{mn}(\hbar\omega - \varepsilon_{mn})}\Bigg].\end{aligned} \tag{88}$$

Here, $C_0 \equiv e^3/(2\epsilon_0 V)$, where V is the crystal volume, $\Delta^\alpha_{mn} \equiv (p^\alpha_{mm} - p^\alpha_{nn})/m$ denotes the velocity difference between bands n and m, and r^α_{nm} is the interband $(n \neq m)$ position matrix element, obtained from $imr^\alpha_{nm} = \hbar p^\alpha_{nm}/\varepsilon_{nm}$. All energies, occupations, and matrix elements in the preceding expressions depend on the **k**-vector. In addition, the summation over **k** implies an integral over the first BZ, i.e., $(2\pi)^3\sum_{\mathbf{k}} \to V\int_{\mathrm{BZ}}\mathrm{d}^3\mathbf{k}$. The primed summation signs indicate the omission of terms with two or more identical indices. Finally, the generalized derivative $r^\beta_{nm;\alpha}$ (for $n \neq m$) is evaluated using the sum rule[239]

$$r^\beta_{nm;\alpha} = \hbar\frac{r^\alpha_{nm}\Delta^\beta_{mn} + r^\beta_{nm}\Delta^\alpha_{mn}}{\varepsilon_{nm}} + \frac{i\hbar}{\varepsilon_{nm}}\sum_{l\neq n,m}(\varepsilon_{lm}r^\alpha_{nl}r^\beta_{lm} - \varepsilon_{nl}r^\beta_{nl}r^\alpha_{lm}). \tag{89}$$

Here, infinite sums have been substituted with finite sums over a limited, yet sizable, set of bands. It is important to emphasize that both sides of the equation are dependent on the **k**-vector, and the summation on the right-hand side pertains exclusively to bands. It should be mentioned that another implementation for the quadratic susceptibility tensor in the velocity gauge is also available in the code

12 March 2024 13:19:04

© Author(s) 2024

but is not documented here. For sufficiently many bands, the results of the two implementations are identical.[238]

Regarding the shift current, where a DC current is induced in response to an incident AC field, one needs to compute the quadratic conductivity tensor $\sigma^{(2)}_{\gamma\alpha\beta}$,

$$\sigma^{(2)}_{\gamma\alpha\beta} = \frac{\pi e^3}{\hbar^2 V} {\sum}'_{\mathbf{k},nm} f_{nm}\, \mathrm{Im}\,\{r^{\alpha}_{mn} r^{\beta}_{nm;\gamma} + r^{\beta}_{mn} r^{\beta}_{nm;\gamma}\} \delta(\hbar\omega - \varepsilon_{mn}). \quad (90)$$

In practice, the delta function is replaced by a Lorentzian with a finite broadening η. To avoid numerical instabilities, the implementation of Eqs. (87)–(90) uses a tolerance, such that terms are neglected if the associated energy differences or differences in Fermi levels are smaller than a tolerance. The default value of the tolerance is 10^{-6} eV and 10^{-4} for the energy and Fermi level differences, respectively.

VII. REAL-TIME TDDFT

The real-time TDDFT (RT-TDDFT) scheme, also known as time-propagation TDDFT, is implemented in FD[34] and LCAO[52] modes. It requires non-periodic boundary conditions but is not restricted to the linear regime and can be applied to model molecules in strong fields. The method may be combined with hybrid quantum–classical modeling to simulate the dynamical interaction between molecules and plasmon resonances at metal surfaces.[240] The LCAO-RT-TDDFT is the more recent implementation, supporting versatile analyses and enabling the modeling of large systems thanks to its efficiency.[52–54,241–247] We focus on the capabilities of the LCAO version in this section, but some of the described functionalities are also available in FD mode.

The time-dependent KS equation in the PAW formalism is

$$i\hat{S}\frac{\mathrm{d}}{\mathrm{d}t}|\tilde{\psi}_n(t)\rangle = \left(\hat{H}[n(\mathbf{r})] + \hat{v}(t)\right)|\tilde{\psi}_n(t)\rangle, \quad (91)$$

where the Kohn–Sham Hamiltonian $\hat{H}[n(\mathbf{r})]$ is implicitly dependent on time through the time-dependent density and $\hat{v}(t)$ is an explicitly time-dependent external potential. We have additionally assumed that the overlap matrix $\hat{S}$ is independent of time, i.e., there are no ion dynamics.

Starting from the ground state, Eq. (91) is propagated numerically forward. After each step, a new density is computed, and $\hat{H}[n(\mathbf{r})]$ is updated accordingly. The user can freely define the external potential, and implemented standard potentials include the delta-kick $\hat{v}(t) = \hat{x}\delta(t)$ and a Gaussian pulse $\hat{v}(t) = \hat{x}\,\exp\left[-\sigma^2(t-t_0)^2/2\right]\cos\left[\omega(t-t_0)\right]$, where $\hat{x}$ is the dipole operator in the direction x.

During the propagation, different time-dependent variables can be recorded, and after the propagation, they can be post-processed into quantities of physical and chemical interest. As a basic example, the time-dependent dipole moment recorded for a delta-kick perturbation can be Fourier-transformed to yield the photoabsorption spectrum.[248] Observables are recorded by attaching observers to the calculation, and implemented observers include writers for the dipole moment, magnetic moment, KS density matrix in frequency space, and wave functions.

RT-TDDFT calculations can be started from the ground state or continued from the last state of a previous time propagation, and the time-limiter feature allows one to limit jobs to a predefined amount of wall time. Together with the continuation capability, this facilitates time propagation in short chunks, efficiently using shared high-performance resources.

In the LCAO-RT-TDDFT implementation, Eq. (91) is cast into a matrix equation and solved with ScaLAPACK.[52] The intermediate step of updating the Hartree and XC potentials is performed on the real-space grid.

A. Kohn–Sham decomposition

The time-dependent KS density matrix can be written as

$$\rho_{nn'}(t) = \sum_m f_m \times \int \left(\psi^{(0)}_n(\mathbf{r})\right)^* \psi_m(\mathbf{r},t)\mathrm{d}\mathbf{r} \int \psi^*_m(\mathbf{r}',t)\psi^{(0)}_{n'}(\mathbf{r}')\mathrm{d}\mathbf{r}', \quad (92)$$

where $\psi^*_m(\mathbf{r}',t)$ are the time-dependent KS orbitals with ground state occupation factors f_m. The KS density matrix is a central quantity enabling the computation of observables and may be evaluated efficiently in the LCAO mode.[54]

The Fourier transform of the induced density matrix $\delta\rho_{nn'}(\omega) = \mathcal{F}[\rho_{nn'}(t) - \rho_{nn'}(0)]$ can be built on the fly during time propagation through the density-matrix observer. Details on the implementation are described in Ref. 54. The KS density matrix in frequency space is related to the Casida eigenvectors and gives similar information as the solution of the Casida equation.[54] Observables such as the polarizability can be decomposed into a sum over the electron–hole part of the KS density matrix $\rho_{nn'}$, where $f_n > f_{n'}$. This enables illustrative analyses, e.g., by visualizing $\rho_{nn'}(\omega)$ on energy axes as a transition contribution map,[249] from which the nature of the localized surface plasmon resonance can be understood (Fig. 12; see Ref. 54 for detailed discussion).

B. Hot-carrier analysis

The KS density matrix is a practical starting point for analyzing hot-carrier generation in plasmonic nanostructures. In the regime of weak perturbations, the KS density matrix resulting from arbitrary pulses can be obtained from delta-kick calculations by a

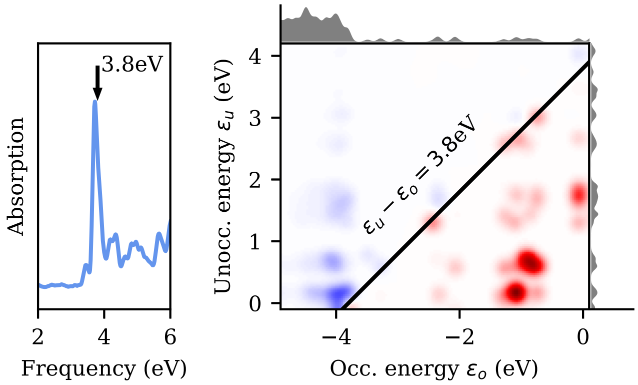


FIG. 12. Photoabsorption spectrum for an Ag_{147} icosahedral nanoparticle and the transition contribution map at 3.8 eV. The map reveals that transitions between KS states near the Fermi level contribute constructively to the plasmon resonance, while transitions from occupied states at the *d*-band edge contribute destructively (screening).

12 March 2024 13:19:04

© Author(s) 2024

post-processing routine.[245,246] By decomposing the matrix into different spatial and energetic contributions, hot-carrier generation in nanostructures[245,247] and across nanoparticle-semiconductor[243] and nanoparticle-molecule[242,246] interfaces can be studied. Computational codes and workflows for such hot-carrier analyses are provided in Refs. 245 and 246.

C. Circular dichroism for molecules

Electronic circular dichroism (ECD) is a powerful spectroscopic method for investigating chiral properties at the molecular level. The quantity that characterizes the ECD is the rotatory strength, which is defined through the magnetic dipole moment $m(\omega)$ as

$$R(\omega) = \frac{1}{\pi\kappa}\mathrm{Re}\left[\sum_{\alpha} m_{\alpha}^{(\alpha)}(\omega)\right], \tag{93}$$

where the index α enumerates Cartesian directions, the α superscript in parenthesis indicates the δ-kick direction, and κ is the strength of the kick. To resolve $R(\omega)$, one needs to perform the δ-kick in all three Cartesian directions using a perturbing electric field $\mathbf{E}^{\alpha}(t) = \kappa\hat{\boldsymbol{\alpha}}\delta(t)$. The frequency components of the dipole moment $m_i^{(i)}(\omega)$ are calculated by Fourier transforming $m_{\alpha}^{(\alpha)}(t)$, which is recorded during the propagation. Finally, the time-dependent magnetic dipole moment is obtained as the expectation value of the operator $\hat{\mathbf{m}} = -\frac{i}{2c}\hat{\mathbf{r}} \times \hat{\nabla}$,

$$\mathbf{m}(t) = \sum_{n} f_n \int \psi_n^{*}(\mathbf{r},t)\, \hat{\mathbf{m}}\, \psi_n(\mathbf{r},t)\, \mathrm{d}\mathbf{r}, \tag{94}$$

where f_n is the occupation number of KS orbital n and $\psi_n(\mathbf{r},t)$ is the time-evolved KS state. The current GPAW implementation supports both the FD and LCAO modes, and the computational efficiency of the LCAO mode enables the calculation of the ECD of nanoscale metal-organic nanoclusters. More details on the implementation can be found in Ref. 55.

D. Radiation reaction potential

Plasmonic or collective molecular excitations are strongly susceptible to any kind of optical interaction. Induced currents will couple via Maxwell's equation of motion to the optical environment and result in radiative losses, i.e., decay towards the ground state. It is possible to solve the Maxwell problem formally by obtaining Green's tensor $\mathbf{G}_{\perp}(\omega)$. The dipolar interaction between the electric field and the electronic dipole can be absorbed into a local potential $v_{\mathrm{rr}}(\mathbf{r},t) = e\mathbf{r}\cdot\mathbf{E}_{r,\perp}(t) = e\mathbf{r}\cdot\left[\mathcal{F}_t^{-1}(i\mu_0\omega\mathbf{G}_{\perp}(\omega)) * \partial_t\mathbf{R}(t)\right]$ suitable for Kohn–Sham TDDFT, where $\mathbf{R}(t)$ is the total electronic dipole moment. A detailed discussion can be found in Ref. 250.

For many simple structures, such as free-space, one-dimensional waveguides, or dielectric spheres, $\mathbf{G}_{\perp}(\omega)$ is analytically known, and radiative losses can then be included in TDDFT without additional computational cost. The tutorials on the GPAW webpage[62] include an example of one-dimensional waveguides for which the user can specify the cross-sectional area and the polarization of the propagating modes. Extending the functionality of the radiation-reaction potential to 3D free-space and the collective interaction of large ensembles in Fabry–Pérot cavities from first principles[251] is essential for the understanding of polaritonic chemistry. This functionality is currently under development.

E. Ehrenfest dynamics

Molecular dynamics (MD) simulations usually rely on the Born–Oppenheimer approximation, where the electronic system is assumed to react so much faster than the ionic system that it reaches its ground state at each time step. Therefore, forces for the dynamics are calculated from the DFT ground-state density. While this approximation is sufficiently valid in most situations, there are cases where the explicit dynamics of the electronic system can affect the molecular dynamics or the movement of the atoms can affect averaged spectral properties. These cases can be handled using so-called Ehrenfest dynamics, i.e., time-dependent density functional theory molecular dynamics (TDDFT/MD).

Ehrenfest dynamics is implemented in the FD mode.[252] A description of the theory and a tutorial is available on the GPAW webpage.[62] This functionality has been used to model the electronic stopping of ions including core-electron excitations,[253] study charge transfer at hybrid interfaces in the presence of water,[254] simulate the coherent diffraction of neutral atom beams from graphene,[255] model the dependence of carbon bond breaking under Ar^{+}-ion irradiation on *sp* hybridization,[256] and reveal charge-transfer dynamics at electrified sulfur cathodes.[257] An LCAO implementation, inspired by recent work in the Siesta code,[258,259] is currently under development.

VIII. EXCITED-STATE DFT METHODS

A. Improved virtual orbitals

The linear response TDDFT approach generally provides reasonably accurate excitation energies for low-lying valence excited states, where the orbitals associated with the holes and excited electrons overlap significantly. However, it tends to fail for excitations involving spatial rearrangement of the electrons, such as charge transfer,[115,260] Rydberg,[261] and doubly excited states.[262]

Some of these problems can be alleviated by using range separated functionals (see Sec. III C 4). However, these functionals come with a significantly increased computational cost due to the evaluation of exchange integrals. Moreover, due to the missing cancellation of Coulomb and exchange terms for canonical unoccupied orbitals within Hartree–Fock theory, one obtains spurious unoccupied orbitals. This leads to a slow convergence of linear-response TDDFT calculations with respect to the number of unoccupied orbitals when hybrid and range-separated functionals are used.[124,125]

Substantial improvement in convergence with respect to unoccupied orbitals can be obtained using improved virtual orbitals, as devised by Huzinaga and Arnau.[263,264] In this approach, a modified Fock operator is used for the unoccupied orbitals, which mimics the interaction between a hole arbitrarily chosen among the occupied ground-state orbitals and the excited electron. This leads to faster convergence in linear-response TDDFT calculations with hybrid and range-separated functionals and also makes it possible to evaluate excited state properties. For example, the energetics of long-range charge transfer can be obtained by means of ground state calculations because the difference between the energy of an

12 March 2024 13:19:04

© Author(s) 2024

improved virtual orbital and a hole tends to approximate the excitation energy. The improved virtual orbitals approach is available in GPAW, and details on the implementation are described in Ref. 125.

B. Variational excited-state calculations

GPAW also offers the possibility to perform excited-state calculations using an alternative time-independent density functional approach[265] (sometimes referred to as the "ΔSCF" method), which does not suffer from the limitations of linear-response TDDFT mentioned in Sec. VIII A. The method involves variational optimization of the orbitals corresponding to a specific excited state by optimizing the electronic energy to a stationary point other than the ground state. The computational effort is similar to that of ground-state calculations, and the variational optimization guarantees that the Hellmann–Feynman theorem is fulfilled. Therefore, all the ground-state machinery available in GPAW to evaluate atomic forces can be used for geometry optimization and for simulating the dynamics of atoms in the excited state. Furthermore, coupling this time-independent, variational approach for excited-state calculations with external MM potentials (see Sec. IV) does not involve additional implementation efforts compared to ground-state calculations and provides a means for performing excited-state QM/MM molecular dynamics simulations that include the state-specific response of a solvent, i.e., the response due to changes in the electron density of the solute.

Variationally optimized excited states correspond to single Slater determinants of optimal orbitals with a non-aufbau occupation and are typically saddle points on the electronic energy surface. Hence, variational calculations of excited states are prone to collapsing to lower-energy solutions, which preserve the symmetry of the initial guess. A simple maximum overlap method (MOM)[266,267] is available in GPAW to address this problem. At each SCF step, the MOM occupies those orbitals that overlap the most with the orbitals of a non-aufbau initial guess, usually obtained from a ground-state calculation. The MOM, however, does not guarantee that variational collapse is avoided, and convergence issues are common when using SCF eigensolvers with density-mixing schemes developed for ground-state calculations.

1. Direct orbital optimization

To alleviate the issues leading to variational collapse and achieve more robust convergence to excited-state solutions, GPAW contains two alternative strategies that are more reliable than conventional SCF eigensolvers with the MOM. They are based on direct optimization of the orbitals and use saddle point search algorithms akin to those for transition-state searches on potential energy surfaces of atomic rearrangements. These approaches also facilitate variational excited-state calculations of nonunitary invariant functionals, such as self-interaction corrected functionals (see Sec. III C 5).

The first of these methods is a direct orbital optimization approach supplemented with the MOM (DO-MOM). This method is an extension of the direct minimization approach using the exponential transformation illustrated in Sec. III B 5, where the search is for a generic stationary point of $E[\mathbf{\Psi}]$ instead of a minimum,

$$\underset{\mathbf{\Psi}}{\text{stat}}\ E[\mathbf{\Psi}] = \min_{\mathbf{\Psi}_0}\ \underset{\mathbf{A}}{\text{stat}}\ \mathcal{F}[\mathbf{A}, \mathbf{\Psi}_0]. \tag{95}$$

The DO-MOM is available in GPAW for both LCAO,[51,268] real-space grid, and plane wave basis sets.[50] For the LCAO basis, the excited-state optimization only necessitates making the energy stationary with respect to the elements of the anti-Hermitian matrix $\mathbf{A}$ (the orbital rotation angles), while calculations using the real-space grid and plane-wave basis include an outer-loop minimization with respect to the reference orbitals $\mathbf{\Psi}_0$. The optimization in the linear space of anti-Hermitian matrices uses efficient quasi-Newton algorithms that can handle negative Hessian eigenvalues and therefore converge on saddle points. GPAW implements a novel limited-memory SR1 (L-SR1) algorithm, which has proven to be robust for calculations of excitations in molecules.[50,51]

The DO-MOM relies on estimating the degrees of freedom along which the energy needs to be maximized, starting from an initial guess. For valence and Rydberg excitations, the initial guess consisting of ground-state canonical orbitals with non-aufbau occupation numbers and preconditioning with a diagonal approximation of the electronic Hessian using the orbital energies[81] can be sufficient. However, if the excitation involves significant charge transfer, large rearrangements of the energy ordering of the orbitals can occur, and DO-MOM can struggle to converge. A second direct optimization method with generalized mode following (DO-GMF)[49] alleviates these problems and is also implemented. In DO-GMF, the components of the energy gradient $\mathbf{g}$ along the modes $\mathbf{v}_i$ corresponding to the n lowest eigenvalues of the electronic Hessian are inverted, yielding

$$\mathbf{g}^{\text{mod}} = \mathbf{g} - 2\sum_{i=1}^{n} \mathbf{v}_i \mathbf{v}_i^{\text{T}} \mathbf{g}, \tag{96}$$

and a minimization using the modified gradient $\mathbf{g}^{\text{mod}}$ is performed by following the n modes simultaneously. This procedure guarantees convergence to an nth-order saddle point, eliminating the risk of variational collapse altogether. While it is computationally more expensive than DO-MOM due to the need for a partial Hessian diagonalization, DO-GMF is more robust. Hence, it is particularly useful in the exploration of potential energy surfaces because it is able to follow an excited state through bond-breaking configurations, where broken-symmetry solutions appear, by targeting the solution that preserves the saddle point order. This important advantage is exemplified by the challenging double-bond twisting in ethylene,[49,135] where DO-GMF calculations of the lowest doubly excited state provide an avoided crossing with the ground state, whereas other methods fail to do so.

2. Example applications of direct optimization

The efficiency and robustness of the direct-optimization approaches, combined with the possibility of choosing different basis-set types, make variational calculations of excited states in GPAW applicable to a great variety of systems ranging from molecules in the gas phase or solution to solids.

State-specific orbital relaxation enables the description of challenging excitations characterized by large density rearrangements. Figure 13 shows the error on the vertical excitation energy of a charge-transfer excitation in the twisted N-phenylpyrrole molecule[49] obtained with direct optimization in GPAW using the LDA, PBE, and BLYP functionals and an sz+aug-cc-pVDZ[53] basis

12 March 2024 13:19:04

© Author(s) 2024

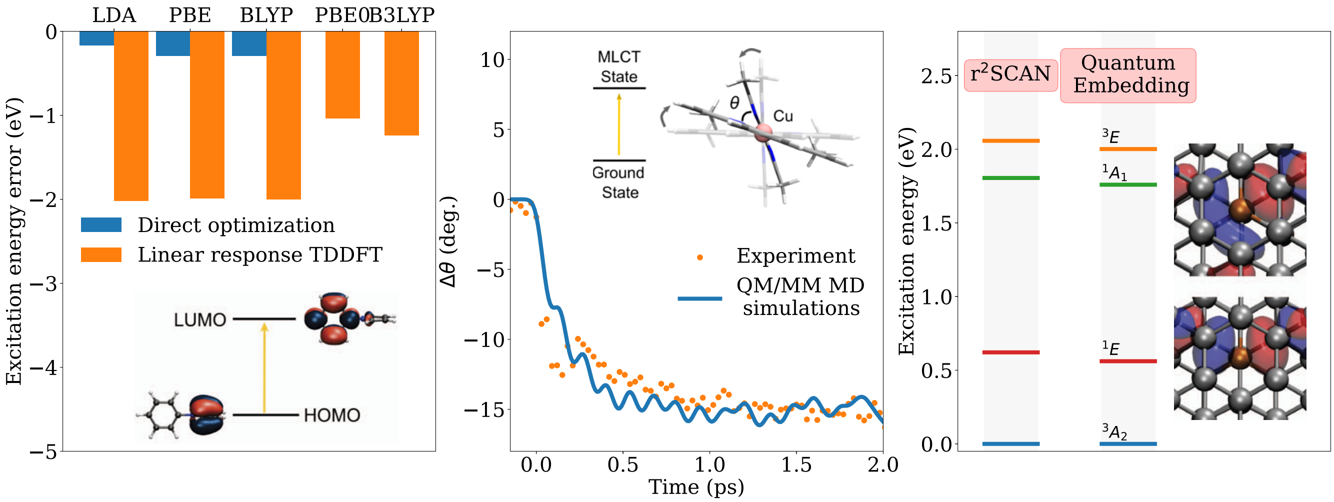


FIG. 13. Applications of time-independent, variational calculations of excited states to a molecule in vacuum (left), a molecule in solution (middle), and a solid-state system (right). Left: Deviation of the calculated excitation energy of a charge-transfer excited state in the N-phenylpyrrole molecule from the theoretical best estimate of 5.58 eV.[269] The results of linear-response TDDFT calculations with hybrid functionals are from Ref. 269. Middle: Time-evolution of the interligand angles of the $[\mathrm{Cu(dmphen)_2}]^+$ complex upon photoexcitation to the lowest metal-to-ligand charge-transfer (MLCT) state in acetonitrile. The experimental results from femtosecond x-ray scattering measurements[270] are compared to the average over excited-state molecular dynamics trajectories obtained using the QM/MM electrostatic embedding scheme in GPAW[167] (see Sec. IV) and convoluted with the experimental instrument-response function.[270] Right: Vertical excitation energy for excitations in the negatively charged nitrogen-vacancy center in diamond obtained with the r^{2}SCAN functional as compared to the results of previous calculations using an advanced quantum embedding approach.[271] The orbitals involved in the electronic transitions are visualized in the inset (the C atoms are gray and the N atom is orange).

set as compared to the results of linear-response TDDFT calculations with the same basis set and functionals, as well as the hybrid functionals PBE0 and B3LYP (results from Ref. 269). For the variational calculations, the energy of the singlet excited state is computed using the spin-purification formula $E_\mathrm{s} = 2E_\mathrm{m} - E_\mathrm{t}$, where E_m is the energy of a spin-mixed state obtained by promoting an electron in one spin channel and E_t is the energy of the triplet state with the same character. The variational calculations underestimate the theoretical best-estimate value of the excitation energy (5.58 eV) in Ref. 269 by 0.15–0.3 eV, an error that is significantly smaller than that of linear-response TDDFT calculations with the same functionals (−2.0 eV) or with the more computationally intensive PBE0 hybrid functional (−0.85 eV).[269]

The method has also been used to simulate the photoinduced structural changes of photocatalytic metal complexes and concomitant solvation dynamics.[167,183,184,270,272] Figure 13 shows an application to the prototypical copper complex photosensitizer $[\mathrm{Cu(dmphen)_2}]^+$ (dmphen = 2,9-dimethyl-1,10-phenanthroline) in acetonitrile, where QM/MM molecular dynamics simulations elucidated an intricate interplay between deformation of the ligands and rearrangement of the surrounding solvent molecules following a photoexcitation to a metal-to-ligand charge-transfer state.[167,270]

The last example of an application shown in Fig. 13 is a calculation of the excited states of a solid state system,[273] the negatively charged nitrogen-vacancy center in diamond, which comprises a prototypical defect for quantum applications. The system is described with a large supercell of 511 atoms, and the calculations use a plane-wave basis set. In contrast to previous reports, a range of different density functionals are found to give the correct energy ordering of the excited states, with the r^{2}SCAN functional providing the best agreement with high-level many-body quantum embedding calculations with an error of less than 0.06 eV.[271,273] This example shows that the direct optimization methods in GPAW are promising tools for simulating excited states in extended systems, where alternative approaches are either computationally expensive or lack accuracy.

IX. OTHER FEATURES

A. Electric polarization

The formal polarization of bulk materials may be calculated from the modern theory of polarization[274,275] as

$$\mathbf{P}_\mathrm{F} = \mathbf{P}_\mathrm{F}^\mathrm{el} + \mathbf{P}_\mathrm{F}^\mathrm{n}, \tag{97}$$

where

$$\mathbf{P}_\mathrm{F}^\mathrm{el} = \frac{e}{(2\pi)^3}\mathrm{Im}\int_\mathrm{BZ} d\mathbf{k}\sum_n f_{\mathbf{k}n}\langle u_{\mathbf{k}n}|\nabla_\mathbf{k}|u_{\mathbf{k}n}\rangle, \tag{98}$$

is the electronic contribution and

$$\mathbf{P}_\mathrm{F}^\mathrm{n} = \frac{1}{V_\mathrm{cell}}\sum_a Z^a\mathbf{r}^a, \tag{99}$$

is the contribution from the nuclei. Here, the sums run over atoms in the unit cell and Z^a is the charge of nucleus a (including core electrons), situated at position $\mathbf{r}^a$. The electronic contribution can be viewed as a Brillouin-zone integral of $\mathbf{k}$-space Berry phases and may be evaluated from a finite-difference version of Eq. (98).[276]

12 March 2024 13:19:04

© Author(s) 2024

This involves the overlaps between Bloch orbitals at neighboring $\mathbf{k}$-points, which are straightforward to evaluate in the PAW formalism.[46]

Equation (97) is only defined modulo $e\mathbf{R}_\mathbf{i}/V_{\text{cell}}$, which follows from the arbitrary choice of unit cell for the atomic positions as well as the choice of phases for $u_{\mathbf{k}n}$, which can shift the Berry phase by 2π. The change in polarization under any adiabatic deformation is, however, well defined and may be calculated as

$$\Delta\mathbf{P} = \int_0^1 \mathrm{d}\lambda \frac{\mathrm{d}\mathbf{P}_\mathrm{F}}{\mathrm{d}\lambda}, \tag{100}$$

where λ is some dimensionless variable parameterizing the adiabatic path. In particular, for ferroelectrics, the spontaneous polarization $\mathbf{P}_\mathrm{S}$ can be evaluated by choosing a path that deforms the structure from the polar ground state at $\lambda = 1$ to a non-polar structure at $\lambda = 0$.

In Fig. 14, we show an example of this for tetragonal $\mathrm{KNbO_3}$, which is a well-known ferroelectric.[277] The polar structure was relaxed under the restraint of tetragonal symmetry using PBE ($\lambda = 1$) and then linearly interpolated to the inverted structure ($\lambda = -1$) passing through a centrosymmetric point at $\lambda = 0$. There are infinitely many polarization branches differing by the polarization quantum ec/V_{cell} (c being the lattice constant in the z-direction), and the spontaneous polarization is obtained by choosing a single branch and evaluating the difference in formal polarization at $\lambda = 1$ and $\lambda = 0$. Interestingly, the centrosymmetric point has a non-vanishing polarization given by half the polarization quantum. This is allowed due to the multi-valued nature of formal polarization, and such "topological polarization" in non-polar materials has been shown to yield gapless surface states.[278,279] Here, however, we merely use the topological polarization to emphasize the importance of evaluating the spontaneous polarization as the change in $\mathbf{P}_\mathrm{F}$ along an adiabatic path.

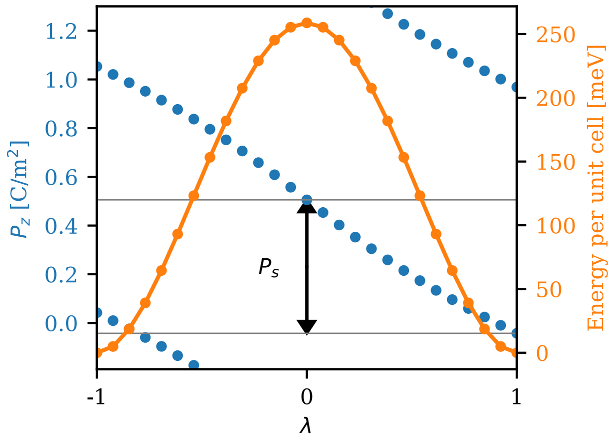


FIG. 14. Formal polarization along an adiabatic path connecting two states of polarization and the energy along the path in tetragonal $\mathrm{KNbO_3}$. The spontaneous polarization is obtained as the difference in polarization between a polar ground state ($\lambda = 1$) and a non-polar reference structure ($\lambda = 0$). The energy along the path is also shown.

The expressions (97)–(100) can be applied to extract various properties of non-polar materials as well, for example, the Born effective charge tensors

$$Z^a_{\alpha\beta} = V_{\text{cell}} \frac{\partial P_\beta}{\partial r^a_\alpha}, \tag{101}$$

which yield the change in polarization resulting from small shifts in atomic positions. In GPAW, these are obtained by a simple call to a module that introduces a (user-defined) shift of all atoms in the unit cell and calculates the resulting change in polarization from Eqs. (97)–(99). The Born charges may be combined with the atomic force matrix to calculate equilibrium positions under an applied static electric field, and the lattice contribution to the dynamic polarizability can be calculated from the eigenvalues and eigenvectors of the force matrix.[280] In addition, the piezoelectric response can be obtained by calculating the change in polarization in response to external strain.[45] The lattice contribution to the polarizability is typically orders of magnitude smaller than the electronic part, but for ferroelectrics, the soft phonon modes associated with spontaneous polarization can give rise to significant lattice polarizabilities.

We exemplify this in the well-known case of 2D ferroelectric GeS,[281–283] where we obtain a spontaneous polarization of 490 pC/m[44,284] in excellent agreement with previous calculations.[283] The lattice and electronic 2D in-plane polarizabilities are $\alpha^{\text{2D}}_{\text{lat}} = 4.32$ Å and $\alpha^{\text{2D}}_{\text{el}} = 3.75$ Å, respectively.[45] For comparison, the non-polar case of 2D $\mathrm{MoS_2}$ yields $\alpha^{\text{2D}}_{\text{lat}} = 0.09$ Å and $\alpha^{\text{2D}}_{\text{el}} = 6.19$ Å.[45]

B. Berry phases and band topology

Topological phases such as the quantum spin Hall state and Chern insulator depend crucially on the presence of spin–orbit coupling, and the band topology may be obtained from the evolution of $\mathbf{k}$-space Berry phases across the Brillouin zone. In particular, for insulators, the eigenvalues of the Berry phase matrix of occupied states

$$\gamma^i_{mn}(\mathbf{k}_\perp) = i\int_0^{2\pi} \mathrm{d}k_i \langle u_m(\mathbf{k})|\partial_{k_i}|u_n(\mathbf{k})\rangle, \tag{102}$$

must change by an integer multiple of 2π when a component of $\mathbf{k}_\perp$ (components of $\mathbf{k}$ orthogonal to k_i) is cycled through the Brillouin zone.[285] Here, $u_n(\mathbf{k})$ are spinor Bloch states, which are typically not smooth functions of $\mathbf{k}$. The evaluation of Eq. (102) thus requires the construction of a smooth gauge, and in GPAW, this is handled by the parallel-transport algorithm of Ref. 286.

The method has been applied in a high-throughput search for new topological two-dimensional materials,[46] and in Fig. 15, we show the calculated Berry phases of 1T′-$\mathrm{MoS_2}$.[287] Due to time-reversal symmetry, the Berry phases at Γ and Y are twofold degenerate, and therefore no perturbation that conserves time-reversal symmetry can open a gap in the Berry-phase spectrum. This property is diagnostic for the quantum spin Hall insulating state and closely related to the presence of gapless edge states.[285]

The eigenvalues of Eq. (102) may also be used to calculate the electronic contribution to the formal polarization since the sum of all individual Berry phases yields the same value as one may obtain from Eq. (98). The present approach is, however, more involved since it requires the construction of a smooth gauge, which is not needed in Eq. (98).

12 March 2024 13:19:04

© Author(s) 2024

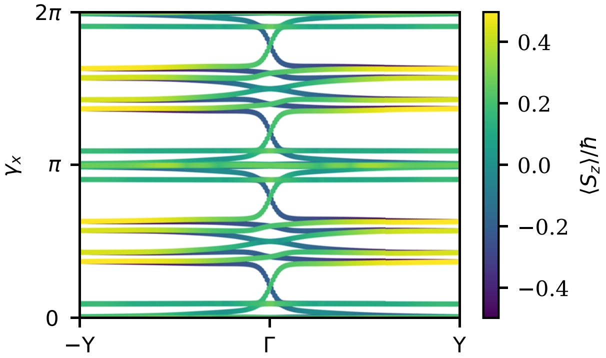


FIG. 15. Berry phases of the quantum spin Hall insulator 1T′-MoS_2 obtained from PBE with non-selfconsistent spin–orbit coupling. The colors indicate the expectation value of S_z for each state as defined in Ref. 46.

C. Wannier functions

Wannier functions (WFs) provide a localized representation of the electronic states of a solid. The WFs are defined by a unitary transformation of the Bloch eigenstates that minimizes the spatial extent of the resultant orbitals. Specifically, the nth Wannier function in unit cell i is written as

$$w_{in}(\mathbf{r}) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} \tilde{\psi}_{n\mathbf{k}}(\mathbf{r}), \tag{103}$$

where $\tilde{\psi}_{n\mathbf{k}}$ is a generalized Bloch function (a superposition of Bloch eigenstates at $\mathbf{k}$).

The minimization of the spatial extent of a set of WFs $\{w_n(\mathbf{r})\}_{n=1}^{N_w}$ is equivalent to the maximization of the spread functional[288]

$$\Omega = \sum_{n=1}^{N_w} \sum_{\alpha=1}^{N_q} W_\alpha |Z_{\alpha,nn}|^2, \tag{104}$$

where

$$Z_{\alpha,nn} = \langle w_n | e^{-i\mathbf{q}_\alpha\cdot\mathbf{r}} | w_n \rangle. \tag{105}$$

The $\{\mathbf{q}_\alpha\}$ is a set of at most six reciprocal vectors connecting a $\mathbf{k}$-point to its neighbors, and W_α are corresponding weights accounting for the shape of the unit cell.[289]

The generalized Bloch functions of Eq. (103) are determined by minimizing Ω using, e.g., a conjugate gradient scheme as implemented in the ASE Wannier module. The inputs to this Wannierization algorithm are the matrices

$$Z^{(0),\mathbf{k}}_{\alpha,ij} = \langle \tilde{\psi}_{\mathbf{k}i} | e^{-i\mathbf{q}_\alpha\cdot\mathbf{r}} | \tilde{\psi}_{\mathbf{k}+\mathbf{q}_\alpha,j} \rangle + \sum_{aii'} e^{-i\mathbf{q}_\alpha\cdot\mathbf{R}^a} \Delta S^a_{ii'} \times \langle \tilde{\psi}_{\mathbf{k}i} | \tilde{p}^a_i \rangle \langle \tilde{p}^a_{i'} | \tilde{\psi}_{\mathbf{k}+\mathbf{q}_\alpha,j} \rangle, \tag{106}$$

where $\Delta S^a_{ii'}$ are the PAW corrections from Eq. (8). From these matrices, the ASE Wannier module can be used to construct partially occupied Wannier functions,[290,291] which are a generalization of maximally localized Wannier functions[286] for entangled bands and non-periodic systems. Recently, a further improvement in terms of robustness of the Wannierisation procedure was achieved using a modified spread functional containing a penalty term proportional to the variance of the spread distribution of the WFs, which leads to a more uniform spread distribution.[292]

D. Point defect calculations with hybrid functionals

Point defects play a crucial role in many applications of semiconductors.[293,294] First-principles calculations can be used to determine the atomic structure, formation energy, and charge-transition levels of point defects. It is well established that the best description of point defects in semiconductors/insulators is obtained using range separated hybrids, such as the HSE06 xc-functional.[5,295] To illustrate the use of GPAW for point defect calculations, we determine the formation energy diagrams of the C_N and C_B defects in the hexagonal boron nitride (hBN) crystal with the HSE06 functional. These defects have been proposed to be responsible for the deep-level luminescence signal with a zero-phonon line (ZPL) around 4.1 eV.[295,296] The results are compared to similar results obtained with the Vienna *Ab initio* Simulation Package (VASP) software package.

For a point-defect D in charge state q, the formation energy E^f is calculated from the formula,

$$E^f[D^q] = E_{tot}[D^q] - E^{bulk}_{tot} - \sum_i n_i \mu_i + qE_F + E_{corr}. \tag{107}$$

Here, $E_{tot}[D^q]$ and E^{bulk}_{tot} are the total energies of the crystal with the point defect in charge state q and of the neutral pristine crystal, respectively. μ_i is the chemical potential of the element i, while n_i is the number of atoms added to ($n_i > 0$) or removed from ($n_i < 0$) the crystal to create the defect. E_F is the chemical potential of the electrons, i.e., the Fermi level, which is written as $E_F = E_{VBM} + \Delta E_F$, where VBM is the valence band maximum. Finally, E_{corr} is a correction term that accounts for (i) the spurious electrostatic interaction between the periodic images of the defect and their interaction with the compensating homogeneous background charge and (ii) the potential shift between the pristine and defect systems.

For more details on the methodology of point defect calculations, we refer the reader to the excellent review papers on the topic.[297–300]

All calculations have been performed using the HSE06 functional with the default mixing parameter $\alpha = 0.25$, a plane wave cutoff of 800 eV, and forces converged to 0.01 eV/Å. The lattice of the hBN crystal was fixed at the experimental parameters ($a = 2.50$ Å and $c = 6.64$ Å).[301] The bandgap of the pristine crystal was determined to be 5.58 eV (using $8 \times 8 \times 4k$-points), in good agreement with the experimental bandgap of 6.08 eV.[302] The structure of the point defects was relaxed in a 4 × 4 × 2 (128 atom) supercell using Γ-point k-point sampling. For each defect, three different charge states ($q = 1, 0, -1$) were considered. The corrections (E_{corr}) due to image charges and potential alignment are evaluated following Freysoldt, Neugebauer and Van de Walle[303] as implemented in GPAW.

Figure 16 shows the defect formation energies as a function of Fermi level for C_N and C_B at N-rich and N-poor conditions, respectively. We can see that under N-rich conditions, C_B is energetically

12 March 2024 13:19:04

© Author(s) 2024

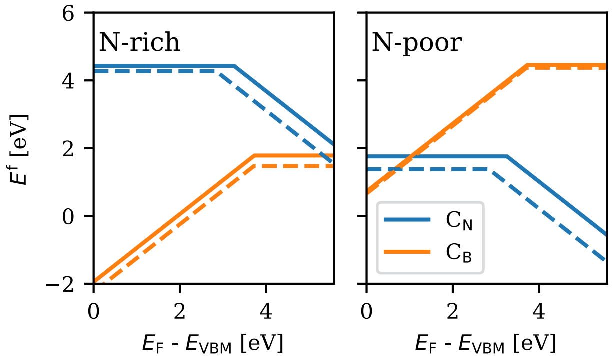


FIG. 16. Defect formation energies for C_N and C_B under N-rich and N-poor conditions, respectively. The dashed lines are reproduced from Ref. 295.

lower, whereas C_N is favorable under N-poor conditions. C_B shows a 1/0 charge transition at 3.73 eV above the VBM, whereas C_N has a 0/−1 charge transition at 3.26 eV deep inside the bandgap. We find good agreement with a similar study[295] also employing plane waves and the HSE06 functional (VASP calculations). Minor discrepancies can be attributed to the use of a different supercell size and a slightly higher fraction of the non-local mixing parameter ($\alpha = 0.31$).

E. Point-group symmetry representations

GPAW allows for the automatized assignment of point-group symmetry representations for the pre-computed Kohn–Sham wavefunctions. This can be used for determining the wave-function symmetry representations for both molecules[304] and extended structures[305] to analyze, for example, the symmetry-involved degeneracy of the bands and selection rules for dipole transitions.

The analysis follows directly from group theory,[306] stating that the solutions to the Schrödinger equation inherit the symmetry group of the respective Hamiltonian, or essentially the external potential invoked by the atomic configuration. The representation matrices Γ are computed as

$$\Gamma_n(T) = \int \phi_n^\dagger(\mathbf{r}) P(T) \phi_n(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{108}$$

where ϕ is a normalized wavefunction, n is the eigenstate label, and $P(T)$ is an operation that corresponds to the transformation T of the symmetry group of the Hamiltonian. The operations $P(T)$ include rotations that are non-trivial for the rectangular grid that the computed wavefunctions are projected onto. The wavefunction rotations are performed on the grid by cubic interpolation. The output of the analysis contains the irreducible representation weights $c_{\alpha,n}$ for each eigenstate n, as solved from

$$\Gamma_n(T) = \sum_\alpha c_{\alpha,n} \chi_\alpha(T), \tag{109}$$

where χ_α are the character vectors of the group (i.e., rows of the character table).

When doing the analysis, the user needs to input the coordinates of the center of symmetry (typically the coordinates of a single atom) and the point group for which the analysis is run. It is possible to analyze only a part of the wave function by selecting a cutoff radius from the center of symmetry beyond which the parts of the wave function are neglected. This enables the investigation of the purity of the local symmetry even if the symmetry of the Hamiltonian is broken far from the center of symmetry.[304,305] To date, point groups of C_2, C_{2v}, C_{3v}, D_{2d}, D_{3h}, D_5, D_{5h}, I, I_h, O_h, T_d, and T_h are implemented.

F. Band-structure unfolding

When studying defect formation, charge-ordered phases, or structural phase transitions, it is often necessary to perform DFT calculations on a supercell. A super-cell (SC) calculation comes with the cost of having to account for many more electrons in the unit cell when compared to the primitive cell (PC). This implies that, besides the increased computational effort, the band structure of a SC contains more bands in a smaller Brillouin zone as compared to the PC. In order to compare electronic band structures between SC and PC, unfolding the band structure of the SC into that of the primitive cell (PC) becomes convenient.

GPAW features the possibility of performing band-structure unfolding in the real-space grid, plane wave, and LCAO modes. The implementation allows to unfold the SC band structure without the explicit calculation of the overlap between SC and PC wavefunctions, following the procedure described in Ref. 307. The unfolded band structure is given in terms of the spectral function

$$A(\mathbf{k} = \mathbf{K} + \mathbf{G}, \epsilon) = \sum_m P_{\mathbf{K}m}(\mathbf{k}) \delta(\epsilon_{\mathbf{K}m} - \epsilon), \tag{110}$$

with $\mathbf{k}$ and $\mathbf{K}$ the momenta in the PC and SC Brillouin zones, respectively ($\mathbf{k}$ is uniquely related to $\mathbf{K}$ by a $\mathbf{G}$ vector of the SC). $\epsilon_{\mathbf{K}m}$ are the SC eigenvalues obtained for momentum $\mathbf{K}$ and band index m, and $P_{\mathbf{K}m}(\mathbf{k})$ is calculated as

$$P_{\mathbf{K}m}(\mathbf{k}) = \sum_{\{\mathbf{G}\}} |C_{\mathbf{K}m}(\mathbf{G} + \mathbf{k} - \mathbf{K})|^2, \tag{111}$$

where $C_{\mathbf{K}m}$ are the Fourier coefficients of the eigenstate $|\mathbf{K}m\rangle$ and $\{\mathbf{G}\}$ is the subset of reciprocal space vectors of the SC that match the reciprocal space vectors of the PC. A more detailed explanation and technical details on how to perform a band-structure unfolding can be found on the GPAW webpage.[62]

G. The QEH model

The quantum-electrostatic heterostructure (QEH)[308] model is an add-on GPAW feature for calculating the dielectric response and excitations in vertical stacks of 2D materials, also known as van der Waals (vdW) heterostructures. The QEH model can be used independently from the GPAW code, but it relies on the GPAW implementation for the calculation of the fundamental building blocks used by the model, as elaborated below.

The dielectric screening in 2D materials is particularly sensitive to changes in the environment and depends on the stacking order and thickness of the 2D heterostructure, providing a means to tune the electronic excitations, including quasi-particle bandgaps and excitons. While the dielectric response of freestanding layers can be explicitly represented *ab initio* in GPAW at the linear-response TDDFT, GW, and BSE levels of theory, the lattice mismatch between different 2D layers often results in large supercells that make these many-body approaches infeasible. Since the interaction between

12 March 2024 13:19:04

© Author(s) 2024

stacked layers is generally governed by van der Waals interactions, the main modification to the non-interacting layers' dielectric response arises from the long-range electrostatic coupling between layers.

Therefore, in the QEH model, the dielectric function of the vdW heterostructure is obtained through an electrostatic coupling of the quantum dielectric building blocks of individual 2D layers.[309] The dielectric building blocks consist of monopole and dipole components of the density-response function of the freestanding layers, $\tilde{\chi}_i(\mathbf{q}_\parallel, \omega)$, calculated from *ab initio* on the RPA level. Subsequently, the full density-response function $\chi_{i,j}(\mathbf{q}_\parallel, \omega)$ (density perturbation on layer i due to a monopole or dipole component of the perturbing field acting on layer j) is calculated by solving the Dyson equation

$$\chi_{i,j}(\mathbf{q}_\parallel, \omega) = \tilde{\chi}_i(\mathbf{q}_\parallel, \omega)\delta_{i,j} + \tilde{\chi}_i(\mathbf{q}_\parallel, \omega)\sum_{k\neq i} V_{i,k}(\mathbf{q}_\parallel)\chi_{k,j}(\mathbf{q}_\parallel, \omega), \quad (112)$$

where the Coulomb matrix element is obtained as the real-space overlap over the density, ρ, and potential, ϕ, basis functions on the different layers,

$$V_{i,k}(\mathbf{q}_\parallel) = \int \rho_i(z, \mathbf{q}_\parallel)\phi_k(z, \mathbf{q}_\parallel)\mathrm{d}z. \quad (113)$$

From the density-response function, the dielectric function is obtained on the basis of monopole/dipole perturbations on each layer in the heterostructure. While building blocks pre-computed with GPAW for a large variety of 2D materials are provided with the QEH package, GPAW offers the possibility of calculating custom building blocks for any 2D material, as explained in Ref. 310.

As an illustrative example of the QEH model, we show how the static dielectric function of a heterostructure can be engineered by multi-layer stacking. Figure 17 shows the static dielectric function of a vdW heterostructure made up of N MoS_2 layers and N WSe_2 layers. We see that the dielectric function increases significantly as a function of the number of layers, eventually approaching a bulk limit. The knowledge of the layer-dependence of the dielectric response for such a heterostructure could be further exploited to investigate inter- and intra-layer excitonic properties and band-edge renormalization effects.

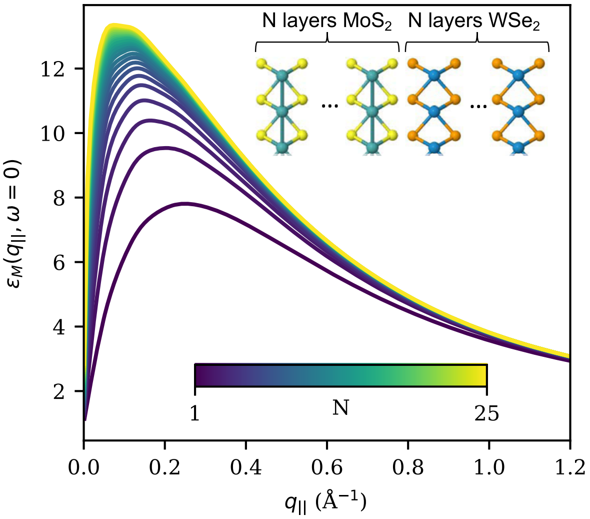


FIG. 17. The static macroscopic dielectric function of a vdW heterostructure interface as a function of the number of layers. The heterostructure is made up of N MoS_2 layers on one half and N WSe_2 layers on the other half (see inset). Increasing the number of layers eventually leads to a bulk-like limit for the chosen stacking configuration.

H. Solvent models

The presence of a solvent has a large effect on the energetics and the electronic structure of molecules or extended surfaces. In particular, the arguably most important solvent, water, is able to stabilize ions, or zwitterions, that would not form in the gas phase. The main effect relates to the large permittivity of water ($\varepsilon_r = 78$) that effectively screens Coulomb interactions.

A convenient and computationally lean method to describe this effect is the inclusion of a position-dependent solvent permittivity $\varepsilon(\mathbf{r})$ in the electrostatics via the Poisson solver.[311] The solvent is represented solely as a polarizable continuum that averages out all movements and re-arrangements of the solvent molecules and their electrons. The computational cost is, therefore, practically the same as a calculation in a vacuum.

This implementation allows the calculation of the solvation free energies of neutral and ionic species in solution.[311,312] The nature of the solvent is described by its permittivity, an effective repulsive potential characterizing the smooth solvent distribution around the solute, and an effective surface tension that summarizes dispersive and repulsive interactions between the solute and solvent.[311] The description is, therefore, well suited for any solvent that can be effectively represented by a continuum, where usually high-permittivity solvents show the largest effects. Local solute–solvent interactions like hydrogen bonds can be considered by the explicit inclusion of single solvent molecules as part of the solute.[313] The solvent model works with periodic as well as finite boundary conditions but is so far available for LCAO and FD modes only. The model can be further applied to periodic surfaces interfaced with an electrolyte to reproduce reasonable potential drops within the simulation of electrochemical reaction processes, as we will elaborate in the following.

I. Charged electrochemical interfaces

Simulating atomistic processes at a solid–liquid interface held at a controlled electrode potential is most appropriately performed in the electronically grand-canonical ensemble.[316,318,319] Here, electrons can be exchanged dynamically with an external electron reservoir at a well-defined electrochemical potential. In a periodic system, a non-zero net charge would lead to divergence of the energy; therefore, any fractional electrons that are added to or removed from the system must be compensated by an equal amount of counter charge. Several approaches able to account for this change in boundary conditions have recently been brought forward.[317–321] In GPAW, this is conveniently accomplished with the introduction of a jellium slab of equal and opposite charge to the needed electronic charge; the jellium is embedded in an implicit solvent localized in the vacuum region above the simulated surface (cf. Sec. IX H). As a particular highlight of this *Solvated Jellium Method* (SJM),[322] as it is known in GPAW, we are able to localize the excess charge on only the top side of the simulated atomistic surface, which occurs naturally by

12 March 2024 13:19:04

© Author(s) 2024

introducing the jellium region solely in the top-side vacuum and electrostatically decoupling the two sides of the cell via a dipole correction. Both a purely implicit and a hybrid explicit-implicit solvent can be applied.

In the SJM, the simulated electrode potential is a monotonic function of the number of electrons in the simulation; calculations can be run in either a constant-charge or a constant-potential ensemble. The electrode potential (ϕ_e) within SJM is defined as the Fermi level (μ) referenced to an electrostatic potential deep in the solvent (the solution inner potential Φ_w), where the whole charge on the electrode has been screened and no electric field is present,

$$\phi_\mathrm{e} = \Phi_\mathrm{w} - \mu. \tag{114}$$

We can relate ϕ_e to the commonly used reference potentials, for example, the standard hydrogen electrode, by subtracting its absolute potential such as the experimental value of 4.44 V as reported by Trasatti.[323] In practice, the reference potentials depend on the used solvent model,[324] and the reference can be calibrated using computed and measured potentials of zero charge.[325]

The energy used in the analysis of electrode reactions is the grand-potential energy Ω,

$$\Omega \equiv E_\mathrm{tot} + \phi_\mathrm{e} N_\mathrm{e}, \tag{115}$$

where N_e are the excess electrons in the simulation. This allows for energetic comparisons between calculations with different numbers of electrons and is the default energy returned to atomistic methods by SJM. While E_tot is consistent with the forces in traditional electronic-structure calculations, the grand-potential energy Ω is consistent with the forces in constant-potential simulations.[316,326] This means that relaxations that follow forces will correctly find local minima in Ω, and any kind of structure optimization or dynamical routine can be performed on the grand-canonical potential energy surface, such as the search for saddle points[327–329] or molecular-dynamics simulations.[330]

In constant-potential mode, the potential is controlled by a damped iterative technique that varies N_e to find the target ϕ_e; in practice, a trajectory (such as a relaxation or nudged elastic band) is run, where in a first series of SCF cycles the potential equilibrates. Upon achieving a target potential within the given threshold, the code will conduct the chosen geometry-optimization routine under constant potential control.

J. Constrained DFT

Constrained DFT (cDFT)[331–333] is a computationally efficient method for constructing diabatic or charge/spin-localized states. GPAW includes a real-space implementation of cDFT,[334] which can be used in both the FD and LCAO modes. Compared to most cDFT implementations, in GPAW, the periodicity can be chosen flexibly between isolated molecules and periodic systems in one-, two-, or three-dimensions with **k**-point sampling.

The key difference between cDFT and normal DFT is the introduction of an auxiliary potential to force a certain region (in real space around a molecule, molecular fragment, or atom) to carry a predefined charge or spin. This leads to a modified energy functional

$$F[n(\mathbf{r}),\{V_i\}] = E^{\mathrm{KS}}[n(\mathbf{r})] + \sum_i V_i \sum_\sigma \left(\int \mathrm{d}\mathbf{r}\, w_i^\sigma(\mathbf{r}) n^\sigma(\mathbf{r}) - N_i\right), \tag{116}$$

where E^{KS} is the Kohn–Sham energy functional, σ denotes the spin, $n^\sigma(\mathbf{r})$ is the spin-dependent electron density, and N_i is the predefined charge or spin restraint. V_i acts as a Lagrange multiplier, which determines the strength of the auxiliary potential and needs to be determined self-consistently, as discussed below. $w_i^\sigma(\mathbf{r})$ is the weight function that defines how the charge or spin are to be partitioned, i.e., the regions where charge/spin are to be localized. This will be discussed in some detail below.

Introducing the constraining term in Eq. (116) leads to a new localized, spin-dependent external potential

$$v_{\mathrm{eff}}^\sigma(\mathbf{r}) = \frac{\delta F[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \frac{\delta E^{\mathrm{KS}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} + \sum_i V_i \sum_\sigma w_i^\sigma(\mathbf{r}). \tag{117}$$

The restraint is further enforced by demanding that the V_i satisfy the chosen restraints

$$C \geq \left| \sum_\sigma \int \mathrm{d}\mathbf{r}\, w_i^\sigma(\mathbf{r}) n^\sigma(\mathbf{r}) - N_i \right|. \tag{118}$$

In practice, the strength of the constraining potential V_i is found through a self-consistent two-stage optimization of both $\{V_i\}$ and $n(\mathbf{r})$. As the derivatives of $F[n(\mathbf{r}),\{V_i\}]$ with respect to V_i are readily available,[334] gradient-based optimization algorithms in SciPy[69] are used for optimizing $\{V_i\}$. The weight functions are defined by a Hirshfeld-type partitioning scheme with Gaussian atomic densities, and $w_i^\sigma(\mathbf{r})$ and the resulting external potential are presented on the grid. With these definitions, the forces resulting from the cDFT external potential can be analytically computed and used in, e.g., geometry optimization or molecular-dynamics simulations.

cDFT has been widely used for computing transfer rates within Marcus theory, which depends on the reorganization and reaction (free) energies and the diabatic coupling matrix element; the GPAW-cDFT implementation includes all the needed tools for obtaining these parameters for bulk, surface, and molecular systems.[334,335] Recently, the cDFT approach has been combined with molecular-dynamics methods to compute the reorganization energy at electrochemical interfaces[336] as well as with grand-canonical DFT methods (see Sec. IX A) to construct fixed electron-potential diabatic states.[337]

K. Orbital-free DFT

Orbital-free DFT (OFDFT) approximates the DFT energy functional by modeling the kinetic energy as a direct functional of the density

$$E_{\mathrm{OF}}[n] = \underbrace{\int \mathrm{d}\mathbf{r}\, n^{1/2}(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) n^{1/2}(\mathbf{r})}_{T_{\mathrm{W}}[n]} + J[n] + V[n] + E_{\mathrm{xc}}[n] + T_\mathrm{s}[n] - T_{\mathrm{W}}[n]. \tag{119}$$

12 March 2024 13:19:04

© Author(s) 2024

Levy *et al.* showed that a Kohn–Sham-like equation derived variationally from the equation above holds for the square-root of the density[338]

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right) n^{1/2}(\mathbf{r}) = \mu\, n^{1/2}(\mathbf{r}). \tag{120}$$

OFDFT approximately enforces the Pauli principle, partially accounting for quantum effects in an averaged way.

The OFDFT scheme implemented in GPAW offers the advantage of accessing all-electron values while maintaining computational linear-scaling time with respect to system size. To achieve this, we employ the PAW method in conjunction with real-space methods, obtaining a mean absolute error of 10 meV per atom when compared to reference all-electron values.[339]

While OFDFT functionals perform better using local pseudopotentials in bulk materials, the OFDFT PAW implementation can be interesting for assessing density functionals. For example, in studying large-Z limits or semiclassical limits of density functionals, the all-electron values allow us to find highly performing OFDFT functionals.[340]

L. Zero-field splitting

Zero-field splitting (ZFS) refers to the energetic splitting of the magnetic sub-levels of a localized triplet state in the absence of a magnetic field.[341] The origin of the ZFS is the magnetic dipole–dipole interactions between the two electrons of the triplet. This interaction is described by a spin Hamiltonian of the form $(\alpha, \beta = x, y, z)$[342,343]

$$\hat{H}_{\text{ZFS}} = \sum_{\alpha\beta} \hat{S}_a D_{\alpha\beta} \hat{S}_\beta, \tag{121}$$

where $\hat{\mathbf{S}}$ is the total spin operator and $\mathbf{D}$ is the ZFS tensor given by

$$D_{\alpha\beta} = \frac{1}{2}\frac{\mu_0}{4\pi} g_e^2 \mu_{\text{B}}^2 \int \rho_2(\mathbf{r}_1, \mathbf{r}_2) \frac{\delta_{\alpha\beta} r^2 - 3 r_\alpha r_\beta}{r^5} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2, \tag{122}$$

where r_α and r_β denote the Cartesian components of $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, ρ_2 is the two-particle density matrix of the Kohn–Sham ground-state Slater determinant, μ_{B} is the Bohr magneton, and g_e is the Landé splitting factor. GPAW computes the $\mathbf{D}$-tensor by evaluating the double integral in reciprocal space using the pseudo density including compensation charges.[305]

M. Hyperfine coupling

The hyperfine coupling describes the interaction between the magnetic dipole of a nuclear spin, $\hat{\mathbf{I}}^N$, and the magnetic dipole of the electron-spin distribution, $\hat{\mathbf{S}}(\mathbf{r})$. The interaction is described by the spin Hamiltonian $(\alpha, \beta = x, y, z)$

$$\hat{H}^N_{\text{HF}} = \sum_{\alpha\beta} \hat{S}_\alpha A^N_{\alpha\beta} \hat{I}^N_\beta, \tag{123}$$

where the hyperfine tensor of nucleus N at $\mathbf{R} = \mathbf{0}$ is given by[344]

$$\begin{aligned} A^N_{\alpha\beta} = {} & \frac{2\mu_0}{3} g_e \mu_B g_N \mu_N \int \delta_T(\mathbf{r}) \rho_s(\mathbf{r}) \mathrm{d}\mathbf{r} \\ & + \frac{\mu_0}{4\pi} g_e \mu_B g_N \mu_N \int \frac{3 r_\alpha r_\beta - \delta_{\alpha\beta} r^2}{r^5} \rho_s(\mathbf{r}) \mathrm{d}\mathbf{r}. \end{aligned} \tag{124}$$

The first term is the isotropic Fermi contact term, which is proportional to the spin density, $\rho_s(\mathbf{r})$, at the center of the nucleus. $\delta_T(\mathbf{r})$ is a smeared-out δ-function. g_e and g_N are the gyromagnetic ratios for the electron and nucleus, and μ_N is the nuclear magneton. The second term represents the anisotropic part of the hyperfine coupling tensor and results from dipole–dipole interactions between nuclear and electronic magnetic moments. GPAW evaluates A^N using the pseudo-spin density with compensation charges.[305]

X. OUTLOOK

As described in this review, GPAW is a highly versatile code that is both maintenance-, user-, and developer-friendly at the same time. The continued expansion of the code requires substantial effort and can be lifted only because of the dedicated team of developers contributing at all levels. There are currently a number of ongoing as well as planned developments for GPAW, which will further improve the performance and applicability of the code. We are currently finishing a major refactoring of the code, which will make it even more developer-friendly and facilitate easier implementation of new functionality.

Another priority is to improve the parallelization of hybrid functional calculations in plane wave mode by enabling parallelization over bands and **k**-points. In the same vein, there is ongoing work to support LCAO-based hybrid functional calculations using a resolution-of-identity approach. A natural next step would then be LCAO-based GW calculations. Such a method could potentially be very efficient compared to plane wave calculations, but it is currently unclear if the accuracy can be maintained with the limited LCAO basis. In relation to quasiparticle calculations, there are plans to implement (quasiparticle) self-consistent GW and vertex corrected GW using nonlocal xc-kernels from TDDFT. Constrained RPA calculations that provide a partially screened Coulomb interaction useful for *ab initio* calculation of interaction parameters in low-energy model Hamiltonians are currently being implemented.

Underlying any GPAW calculation are the PAW potentials. The current potentials date back to 2009. A new set of potentials, including both soft and norm-conserving potentials (for response function calculations), is being worked on.

As described herein, GPAW already has an efficient implementation of real-time TDDFT in the LCAO basis, while Ehrenfest dynamics is supported only in the comparatively slower grid mode. Work to enable Ehrenfest dynamics in LCAO mode is ongoing.

The current version of GPAW supports GPU acceleration only for standard ground-state calculations. The CuPy library greatly simplifies the task of porting GPAW to GPU, and we foresee that large parts of the code, including more advanced features such as linear response and GW calculations, will become GPU compatible.

ACKNOWLEDGMENTS

K.S.T. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program Grant No. 773122 (LIMA) and Grant Agreement No. 951786 (NOMAD CoE). K.S.T. is a Villum Investigator supported by VILLUM FONDEN (Grant No. 37789). Funding for A.O.D., G.L., and Y.L.A.S. was provided by the Icelandic

12 March 2024 13:19:04

Research Fund (Grant Nos. 196279, 217734, and 217751, respectively). F.N. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 899987. M.M.M. was supported by the Academy of Finland (Grant No. 338228). T.O. acknowledges support from the Villum Foundation Grant No. 00029378. S.K. and K.W.J. acknowledge support from the VILLUM Center for Science of Sustainable Fuels and Chemicals, which is funded by the VILLUM Fonden research (Grant No. 9455). T.B. was funded by the Danish National Research Foundation (DNRF Grant No. 146). J.S. acknowledges funding from the Independent Research Fund Denmark (DFF-FTP) through Grant No. 9041-00161B. C.S. acknowledges support from the Swedish Research Council (VR) through Grant No. 2016-06059 and funding from the Horizon Europe research and innovation program of the European Union under the Marie Skłodowska-Curie Grant Agreement No. 101065117. Partially funded by the European Union. The views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them. T.S. received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 756277-ATMEN). O.L.-A. has been supported by Minciencias and the University of Antioquia (Colombia). K.T.W. was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program, and the SUNCAT Center for Interface Science and Catalysis. G.K. acknowledges funding from V-Sustain: The VILLUM Centre for the Science of Sustainable Fuels and Chemicals (Grant No. 9455). Additional funding: Knut and Alice Wallenberg Foundation (Grant No. 2019.0140; J.F. and P.E.), the Swedish Research Council (Grant No. 2020-04935; J.F. and P.E.), the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101065117 (C.S.). Computations and code development work have been enabled by the resources provided by the Niflheim Linux cluster supercomputer installed at the Department of Physics at the Technical University of Denmark, CSC–IT Center for Science, Finland, through national supercomputers and through access to the LUMI supercomputer owned by the EuroHPC Joint Undertaking, and the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at NSC, PDC, and C3SE, partially funded by the Swedish Research Council through Grant Agreement No. 2022-06725. We gratefully acknowledge these organizations for providing computational resources and facilities.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jens Jørgen Mortensen: Writing – original draft (equal). **Ask Hjorth Larsen**: Writing – original draft (equal). **Mikael Kuisma**: Writing – original draft (equal). **Aleksei V. Ivanov**: Writing – original draft (equal). **Alireza Taghizadeh**: Writing – original draft (equal). **Andrew Peterson**: Writing – original draft (equal). **Anubhab Haldar**: Writing – original draft (equal). **Asmus Ougaard Dohn**: Writing – original draft (equal). **Christian Schäfer**: Writing – original draft (equal). **Elvar Örn Jónsson**: Writing – original draft (equal). **Eric D. Hermes**: Writing – original draft (equal). **Fredrik Andreas Nilsson**: Writing – original draft (equal). **Georg Kastlunger**: Writing – original draft (equal). **Gianluca Levi**: Writing – original draft (equal). **Hannes Jónsson**: Writing – original draft (equal). **Hannu Häkkinen**: Writing – original draft (equal). **Jakub Fojt**: Writing – original draft (equal). **Jiban Kangsabanik**: Writing – original draft (equal). **Joachim Sødequist**: Writing – original draft (equal). **Jouko Lehtomäki**: Writing – original draft (equal). **Julian Heske**: Writing – original draft (equal). **Jussi Enkovaara**: Writing – original draft (equal). **Kirsten Trøstrup Winther**: Writing – original draft (equal). **Marcin Dulak**: Writing – original draft (equal). **Marko M. Melander**: Writing – original draft (equal). **Martin Ovesen**: Writing – original draft (equal). **Martti Louhivuori**: Writing – original draft (equal). **Michael Walter**: Writing – original draft (equal). **Morten Gjerding**: Writing – original draft (equal). **Olga Lopez-Acevedo**: Writing – original draft (equal). **Paul Erhart**: Writing – original draft (equal). **Robert Warmbier**: Writing – original draft (equal). **Rolf Würdemann**: Writing – original draft (equal). **Sami Kaappa**: Writing – original draft (equal). **Simone Latini**: Writing – original draft (equal). **Tara Maria Boland**: Writing – original draft (equal). **Thomas Bligaard**: Writing – original draft (equal). **Thorbjørn Skovhus**: Writing – original draft (equal). **Toma Susi**: Writing – original draft (equal). **Tristan Maxson**: Writing – original draft (equal). **Tuomas Rossi**: Writing – original draft (equal). **Xi Chen**: Writing – original draft (equal). **Yorick Leonard A. Schmerwitz**: Writing – original draft (equal). **Jakob Schiøtz**: Writing – original draft (equal). **Thomas Olsen**: Writing – original draft (equal). **Karsten Wedel Jacobsen**: Writing – original draft (equal). **Kristian Sommer Thygesen**: Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in the git-repository at https://gitlab.com/jensj/gpaw-paper-2023.

REFERENCES

[1] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. **136**(3B), B864 (1964).

[2] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. **140**(4A), A1133 (1965).

[3] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**(18), 3865 (1996).

[4] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," J. Chem. Phys. **98**, 5648–5652 (1993).

[5] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential," J. Chem. Phys. **118**(18), 8207–8215 (2003).

[6] G. M. Torrie and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," J. Comput. Phys. **23**(2), 187–199 (1977).

[7] P. Souvatzis, O. Eriksson, M. I. Katsnelson, and S. P. Rudin, "Entropy driven stabilization of energetically unstable crystal structures explained from first principles theory," Phys. Rev. Lett. **100**(9), 095901 (2008).

12 March 2024 13:19:04

© Author(s) 2024

[8] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies," Phys. Rev. B **34**, 5390–5413 (1986).

[9] D. Golze, M. Dvorak, and P. Rinke, "The *GW* compendium: A practical guide to theoretical photoemission spectroscopy," Front. Chem. **7**, 377 (2019).

[10] M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, "Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold," J. Chem. Phys. **108**(11), 4439–4449 (1998).

[11] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: Density-functional versus many-body Green's-function approaches," Rev. Mod. Phys. **74**(2), 601 (2002).

[12] J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, "Towards the computational design of solid catalysts," Nat. Chem. **1**(1), 37–46 (2009).

[13] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The materials project: A materials genome approach to accelerating materials innovation," APL Mater. **1**(1), 011002 (2013).

[14] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," Nat. Mater. **12**(3), 191–201 (2013).

[15] S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen *et al.*, "The computational 2D materials database: High-throughput modeling and discovery of atomically thin crystals," 2D Mater. **5**(4), 042002 (2018).

[16] I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, and K. W. Jacobsen, "Computational screening of perovskite metal oxides for optimal solar light capture," Energy Environ. Sci. **5**(2), 5814–5819 (2012).

[17] A. Marek, V. Blum, R. Johanni, V. Havu, B. Lang, T. Auckenthaler, A. Heinecke, H.-J. Bungartz, and H. Lederer, "The ELPA library: Scalable parallel eigenvalue solutions for electronic structure theory and computational science," J. Phys.: Condens. Matter **26**(21), 213201 (2014).

[18] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, "Cubic scaling *GW*: Towards fast quasiparticle calculations," Phys. Rev. B **94**(16), 165109 (2016).

[19] A. Jain, S. P. Ong, W. Chen, B. Medasani, X. Qu, M. Kocher, M. Brafman, G. Petretto, G.-M. Rignanese, G. Hautier *et al.*, "Fireworks: A dynamic workflow system designed for high-throughput applications," Concurrency Comput.: Pract. Exper. **27**(17), 5037–5059 (2015).

[20] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, "AiiDA: Automated interactive infrastructure and database for computational science," Comput. Mater. Sci. **111**, 218–230 (2016).

[21] J. J. Mortensen, M. Gjerding, and K. S. Thygesen, "MyQueue: Task and workflow scheduling system," J. Open Source Softw. **5**(45), 1844 (2020).

[22] M. Gjerding, T. Skovhus, A. Rasmussen, F. Bertoldo, A. H. Larsen, J. J. Mortensen, and K. S. Thygesen, "Atomic simulation recipes: A Python framework and library for automated workflows," Comput. Mater. Sci. **199**, 110731 (2021).

[23] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," npj Comput. Mater. **5**(1), 83 (2019).

[24] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," Phys. Rev. Lett. **104**(13), 136403 (2010).

[25] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. Von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," Phys. Rev. Lett. **108**(5), 058301 (2012).

[26] A. M. Teale, T. Helgaker, A. Savin, C. Adamo, B. Aradi, A. V. Arbuznikov, P. W. Ayers, E. J. Baerends, V. Barone, P. Calaminici *et al.*, "DFT exchange: Sharing perspectives on the workhorse of quantum chemistry and materials science," Phys. Chem. Chem. Phys. **24**(47), 28700–28781 (2022).

[27] X. Ren, P. Rinke, C. Joas, and M. Scheffler, "Random-phase approximation and its applications in computational chemistry and materials science," J. Mater. Sci. **47**, 7447–7471 (2012).

[28] J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional," Phys. Rev. Lett. **115**(3), 036402 (2015).

[29] T. Olsen and K. S. Thygesen, "Random phase approximation applied to solids, molecules, and graphene-metal interfaces: From van der Waals to covalent bonding," Phys. Rev. B **87**(7), 075111 (2013).

[30] A. Tkatchenko and M. Scheffler, "Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data," Phys. Rev. Lett. **102**(7), 073005 (2009).

[31] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der Waals density functional for general geometries," Phys. Rev. Lett. **92**(24), 246401 (2004).

[32] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, "Real-space grid implementation of the projector augmented wave method," Phys. Rev. B **71**, 035109 (2005).

[33] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen, and K. W. Jacobsen, "Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method," J. Phys.: Condens. Matter **22**(25), 253202 (2010).

[34] M. Walter, H. Häkkinen, L. Lehtovaara, M. Puska, J. Enkovaara, C. Rostgaard, and J. J. Mortensen, "Time-dependent density-functional theory in the projector augmented-wave method," J. Chem. Phys. **128**(24), 244101 (2008).

[35] A. H. Larsen, M. Vanin, J. J. Mortensen, K. S. Thygesen, and K. W. Jacobsen, "Localized atomic basis set in the projector augmented wave method," Phys. Rev. B **80**, 195112 (2009).

[36] J. Yan, J. J. Mortensen, K. W. Jacobsen, and K. S. Thygesen, "Linear density response function in the projector augmented wave method: Applications to solids, surfaces, and interfaces," Phys. Rev. B **83**(24), 245122 (2011).

[37] T. Olsen and K. S. Thygesen, "Extending the random-phase approximation for electronic correlation energies: The renormalized adiabatic local density approximation," Phys. Rev. B **86**(8), 081103 (2012).

[38] F. Hüser, T. Olsen, and K. S. Thygesen, "Quasiparticle GW calculations for solids, molecules, and two-dimensional materials," Phys. Rev. B **87**(23), 235132 (2013).

[39] J. Yan, K. W. Jacobsen, and K. S. Thygesen, "Optical properties of bulk semiconductors and graphene/boron nitride: The Bethe-Salpeter equation with derivative discontinuity-corrected density functional energies," Phys. Rev. B **86**(4), 045208 (2012).

[40] J. Sødequist and T. Olsen, "Type II multiferroic order in two-dimensional transition metal halides from first principles spin-spiral calculations," 2D Mater. **10**(3), 035016 (2023).

[41] D. Torelli and T. Olsen, "Calculating critical temperatures for ferromagnetic order in two-dimensional materials," 2D Mater. **6**, 015028 (2018).

[42] F. L. Durhuus, T. Skovhus, and T. Olsen, "Plane wave implementation of the magnetic force theorem for magnetic exchange constants: Application to bulk Fe, Co and Ni," J. Phys.: Condens. Matter **35**(10), 105802 (2023).

[43] T. Skovhus and T. Olsen, "Dynamic transverse magnetic susceptibility in the projector augmented-wave method: Application to Fe, Ni, and Co," Phys. Rev. B **103**, 245110 (2021).

[44] M. Kruse, U. Petralanda, M. N. Gjerding, K. W. Jacobsen, K. S. Thygesen, and T. Olsen, "Two-dimensional ferroelectrics from high throughput computational screening," npj Comput. Mater. **9**, 45 (2023).

[45] M. N. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N. R. Knøsgaard, M. Kruse, A. H. Larsen, S. Manti, T. G. Pedersen, U. Petralanda, T. Skovhus, M. K. Svendsen, J. J. Mortensen, T. Olsen, and K. S. Thygesen, "Recent progress of the computational 2D materials database (C2DB)," 2D Mater. **8**, 044002 (2021).

[46] T. Olsen, E. Andersen, T. Okugawa, D. Torelli, T. Deilmann, and K. S. Thygesen, "Discovering two-dimensional topological insulators from high-throughput computations," Phys. Rev. Mater. **3**, 024005 (2019).

[47] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson,

12 March 2024 13:19:04

© Author(s) 2024

T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, “The atomic simulation environment—A Python library for working with atoms,” J. Phys.: Condens. Matter **29**(27), 273002 (2017).

[48] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, “wannier90: A tool for obtaining maximally-localised Wannier functions,” Comput. Phys. Commun. **178**(9), 685–699 (2008).

[49] Y. L. A. Schmerwitz, G. Levi, and H. Jónsson, “Calculations of excited electronic states by converging on saddle points using generalized mode following,” J. Chem. Theory Comput. **19**(12), 3634–3651 (2023).

[50] A. V. Ivanov, G. Levi, E. Ö. Jónsson, and H. Jónsson, “Method for calculating excited electronic states using density functionals and direct orbital optimization with real space grid or plane-wave basis set,” J. Chem. Theory Comput. **17**, 5034–5049 (2021).

[51] G. Levi, A. V. Ivanov, and H. Jónsson, “Variational density functional calculations of excited states via direct optimization,” J. Chem. Theory Comput. **16**, 6968–6982 (2020).

[52] M. Kuisma, A. Sakko, T. P. Rossi, A. H. Larsen, J. Enkovaara, L. Lehtovaara, and T. T. Rantala, “Localized surface plasmon resonance in silver nanoparticles: Atomistic first-principles time-dependent density-functional theory calculations,” Phys. Rev. B **91**(11), 115431 (2015).

[53] T. P. Rossi, S. Lehtola, A. Sakko, M. J. Puska, and R. M. Nieminen, “Nanoplasmonics simulations at the basis set limit through completeness-optimized, local numerical basis sets,” J. Chem. Phys. **142**(9), 094114 (2015).

[54] T. P. Rossi, M. Kuisma, M. J. Puska, R. M. Nieminen, and P. Erhart, “Kohn–Sham decomposition in real-time time-dependent density-functional theory: An efficient tool for analyzing plasmonic excitations,” J. Chem. Theory Comput. **13**(10), 4779–4790 (2017).

[55] E. Makkonen, T. P. Rossi, A. H. Larsen, O. Lopez-Acevedo, P. Rinke, M. Kuisma, and X. Chen, “Real-time time-dependent density functional theory implementation of electronic circular dichroism applied to nanoscale metal–organic clusters,” J. Chem. Phys. **154**(11), 114102 (2021).

[56] A. Taghizadeh, U. Leffers, T. G. Pedersen, and K. S. Thygesen, “A library of ab initio Raman spectra for automated identification of 2D materials,” Nat. Commun. **11**(1), 3011 (2020).

[57] A. Taghizadeh, K. S. Thygesen, and T. G. Pedersen, “Two-dimensional materials with giant optical nonlinearities near the theoretical upper limit,” ACS Nano **15**(4), 7155–7167 (2021).

[58] M. O. Sauer, A. Taghizadeh, U. Petralanda, M. Ovesen, K. S. Thygesen, T. Olsen, H. Cornean, and T. G. Pedersen, “Shift current photovoltaic efficiency of 2D materials,” npj Comput. Mater. **9**(1), 35 (2023).

[59] See https://wiki.fysik.dtu.dk/gpaw/documentation/basic.html\#features-in-fd-lcao-and-pw-modes for table of avilability of feature in FD, LCAO and PW mode.

[60] “GPAW’s git-repository,” GPAW. https://gitlab.com/gpaw/gpaw/

[61] H. Krekel, B. Oliveira, R. Pfannschmidt, F. Bruynooghe, B. Laugher, and F. Bruhin, pytest, 2004.

[62] See https://wiki.fysik.dtu.dk/gpaw/tutorialsexercises/tutorialsexercises.html for GPAW tutorials and exercises.

[63] M. Frigo and S. G. Johnson, “The design and implementation of FFTW3,” Proc. IEEE **93**(2), 216–231 (2005), part of the Special Issue: Program Generation, Optimization, and Platform Adaptation.

[64] L. S. Blackford, J. Choi, A. Cleary, E. D’Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, *ScaLAPACK Users’ Guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997).

[65] M. A. L. Marques, M. J. T. Oliveira, and T. Burnus, “LIBXC: A library of exchange and correlation functionals for density functional theory,” Comput. Phys. Commun. **183**(10), 2272–2281 (2012).

[66] S. Lehtola, C. Steigemann, M. J. T. Oliveira, and M. A. L. Marques, “Recent developments in LIBXC—A comprehensive library of functionals for density functional theory,” SoftwareX **7**, 1–5 (2018).

[67] A. H. Larsen, M. Kuisma, J. Löfgren, Y. Pouillon, P. Erhart, and P. Hyldgaard, “libvdwxc: A library for exchange–correlation functionals in the vdW-DF family,” Modell. Simul. Mater. Sci. Eng. **25**(6), 065004 (2017).

[68] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, “Array programming with NumPy,” Nature **585**(7825), 357–362 (2020).

[69] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar *et al.*, “SciPy 1.0: Fundamental algorithms for scientific computing in Python,” Nat. Methods **17**, 261–272 (2020).

[70] R. Würdemann, H. H. Kristoffersen, M. Moseler, and M. Walter, “Density functional theory and chromium: Insights from the dimers,” J. Chem. Phys. **142**(12), 124316 (2015).

[71] P. E. Blöchl, “Projector augmented-wave method,” Phys. Rev. B **50**, 17953–17979 (1994).

[72] T. Susi, J. Madsen, U. Ludacka, J. J. Mortensen, T. J. Pennycook, Z. Lee, J. Kotakoski, U. Kaiser, and J. C. Meyer, “Efficient first principles simulation of electron scattering factors for transmission electron microscopy,” Ultramicroscopy **197**, 16–22 (2019).

[73] J. Madsen, T. J. Pennycook, and T. Susi, “*Ab initio* description of bonding for transmission electron microscopy,” Ultramicroscopy **231**, 113253 (2021).

[74] J. Madsen and T. Susi, “The abTEM code: Transmission electron microscopy from first principles,” Open Res. Eur. **1**(24), 13015 (2021).

[75] 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).

[76] N. Marzari, D. Vanderbilt, A. De Vita, and M. C. Payne, “Thermal contraction and disordering of the Al(110) surface,” Phys. Rev. Lett. **82**, 3296–3299 (1999).

[77] P. E. Blöchl, O. Jepsen, and O. K. Andersen, “Improved tetrahedron method for Brillouin-zone integrations,” Phys. Rev. B **49**, 16223–16233 (1994).

[78] S. Lehtola, F. Blockhuys, and C. Van Alsenoy, “An overview of self-consistent field calculations within finite basis sets,” Molecules **25**(5), 1218 (2020).

[79] T. Van Voorhis and M. Head-Gordon, “A geometric approach to direct minimization,” Mol. Phys. **100**, 1713–1721 (2002).

[80] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, “Iterative minimization techniques for *ab initio* total-energy calculations: Molecular dynamics and conjugate gradients,” Rev. Mod. Phys. **64**(4), 1045–1097 (1992).

[81] M. Head-Gordon and J. A. Pople, “Optimization of wave function and geometry in the finite basis Hartree-Fock method,” J. Phys. Chem. **92**(11), 3063–3069 (1988).

[82] A. V. Ivanov, E. Ö. Jónsson, T. Vegge, and H. Jónsson, “Direct energy minimization based on exponential transformation in density functional calculations of finite and extended systems,” Comput. Phys. Commun. **267**, 108047 (2021).

[83] See https://esl.cecam.org/en/data/paw-xml/ for PAW-XML specification.

[84] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger, “The ABINITproject: Impact, environment and recent developments,” Comput. Phys. Commun. **248**, 107042 (2020).

[85] F. Jollet, M. Torrent, and N. Holzwarth, “Generation of projector augmented-wave atomic data: A 71 element validated table in the XML format,” Comput. Phys. Commun. **185**(4), 1246–1254 (2014).

[86] C. Hartwigsen, S. Goedecker, and J. Hutter, “Relativistic separable dual-space Gaussian pseudopotentials from H to Rn,” Phys. Rev. B **58**, 3641–3662 (1998).

12 March 2024 13:19:04

© Author(s) 2024

[87] M. Schlipf and F. Gygi, "Optimization algorithm for the generation of ONCV pseudopotentials," Comput. Phys. Commun. **196**, 36–44 (2015).

[88] See https://cupy.dev/ for CuPy.

[89] "Header only porting," GitHub. https://github.com/cschpc/hop/

[90] S. Hakala, V. Havu, J. Enkovaara, and R. M. Nieminen, "Parallel electronic structure calculations using multiple graphics processing units (GPUs)," Lect. Notes Comput. Sci. **7782**, 63–76 (2013).

[91] S. Hakala, J. Enkovaara, V. Havu, J. Yan, L. Li, C. O'Grady, and R. M. Nieminen, "Grid-based projector-augmented wave method," in *Method in Electronic Structure Calculations on Graphics Processing Units* (Wiley, 2015).

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

[93] 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**(14), 146401 (2003).

[94] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," Phys. Rev. B **45**(23), 13244–13249 (1992).

[95] G. Román-Pérez and J. M. Soler, "Efficient implementation of a van der Waals density functional: Application to double-wall carbon nanotubes," Phys. Rev. Lett. **103**, 096102 (2009).

[96] O. Gritsenko, R. van Leeuwen, E. van Lenthe, and E. J. Baerends, "Self-consistent approximation to the Kohn-Sham exchange potential," Phys. Rev. A **51**(3), 1944 (1995).

[97] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, "Kohn-Sham potential with discontinuity for band gap materials," Phys. Rev. B **82**(11), 115106 (2010).

[98] J. Yan, K. W. Jacobsen, and K. S. Thygesen, "First-principles study of surface plasmons on Ag(111) and H/Ag(111)," Phys. Rev. B **84**(23), 235430 (2011).

[99] J. Yan, K. W. Jacobsen, and K. S. Thygesen, "Conventional and acoustic surface plasmons on noble metal surfaces: A time-dependent density functional theory study," Phys. Rev. B **86**(24), 241404 (2012).

[100] J. M. Rahm, C. Tiburski, T. P. Rossi *et al.*, "A library of late transition metal alloy dielectric functions for nanophotonic applications," Adv. Funct. Mater. **30**(35), 2002122 (2020).

[101] K. Bashyal, C. K. Pyles, S. Afroosheh, A. Lamichhane, and A. T. Zayak, "Empirical optimization of DFT + U and HSE for the band structure of ZnO," J. Phys.: Condens. Matter **30**(6), 065501 (2018).

[102] M. K. Y. Chan and G. Ceder, "Efficient band gap prediction for solids," Phys. Rev. Lett. **105**, 196403 (2010).

[103] J. P. Perdew, "Density functional theory and the band gap problem," Int. J. Quantum Chem. **28**(S19), 497–523 (1985).

[104] K. J. May and A. M. Kolpak, "Improved description of perovskite oxide crystal structure and electronic properties using self-consistent Hubbard *U* corrections from ACBN0," Phys. Rev. B **101**, 165117 (2020).

[105] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study," Phys. Rev. B **57**, 1505–1509 (1998).

[106] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, "Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators," Phys. Rev. B **52**, R5467–R5470 (1995).

[107] F. Aryasetiawan, T. Miyake, and R. Sakuma, "The constrained RPA method for calculating the Hubbard *U* from first-principles," in *The LDA+DMFT Approach to Strongly Correlated Materials* (Forschungszentrum Jülich GmbH Institute for Advanced Simulations, 2011).

[108] M. Cococcioni and S. de Gironcoli, "Linear response approach to the calculation of the effective interaction parameters in the LDA + U method," Phys. Rev. B **71**(3), 035105 (2005).

[109] G. C. Moore, M. K. Horton, A. M. Ganose, M. Siron, E. Linscott, D. D. O'Regan, and K. A. Persson, "High-throughput determination of Hubbard U and Hund J values for transition metal oxides via linear response formalism," Phys. Rev. Materials **8**, 014409 (2024).

[110] L. Wang, T. Maxisch, and G. Ceder, "Oxidation energies of transition metal oxides within the GGA + U framework," Phys. Rev. B **73**, 195107 (2006).

[111] M. F. M. Taib, D. T. Mustaffa, N. H. Hussin, M. H. Samat, A. M. M. Ali, O. H. Hassan, and M. Z. A. Yahya, "First principles study on Zn doped MgO using Hubbard U correction," Mater. Res. Express **6**(9), 094012 (2019).

[112] M. Yu, S. Yang, C. Wu, and N. Marom, "Machine learning the Hubbard *U* parameter in DFT+*U* using Bayesian optimization," npj Comput. Mater. **6**(1), 180 (2020).

[113] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, "A long-range-corrected time-dependent density functional theory," J. Chem. Phys. **120**(18), 8425–8433 (2004).

[114] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Insights into current limitations of density functional theory," Science **321**(5890), 792–794 (2008).

[115] A. Dreuw, J. L. Weisman, and M. Head-Gordon, "Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange," J. Chem. Phys. **119**(6), 2943–2946 (2003).

[116] E. J. Baerends, O. V. Gritsenko, and R. van Meer, "The Kohn–Sham gap, the fundamental gap and the optical gap: The physical meaning of occupied and virtual Kohn–Sham orbital energies," Phys. Chem. Chem. Phys. **15**(39), 16408–16425 (2013).

[117] S. Kümmel, "Charge-transfer excitations: A challenge for time-dependent density functional theory that has been met," Adv. Energy Mater. **7**(16), 1700440 (2017).

[118] R. Baer, E. Livshits, and U. Salzner, "Tuned range-separated hybrids in density functional theory," Annu. Rev. Phys. Chem. **61**(1), 85–109 (2010).

[119] C. Adamo and V. Barone, "Toward chemical accuracy in the computation of NMR shieldings: The PBE0 model," Chem. Phys. Lett. **298**(1–3), 113–119 (1998).

[120] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP)," Chem. Phys. Lett. **393**(1–3), 51–57 (2004).

[121] Y. Akinaga and S. Ten-no, "Range-separation by the Yukawa potential in long-range corrected density functional theory with Gaussian-type basis functions," Chem. Phys. Lett. **462**(4–6), 348–351 (2008).

[122] M. Seth and T. Ziegler, "Range-separated exchange functionals with slater-type functions," J. Chem. Theory Comput. **8**(3), 901–907 (2012).

[123] E. Livshits and R. Baer, "A well-tempered density functional theory of electrons in molecules," Phys. Chem. Chem. Phys. **9**(23), 2932–2941 (2007).

[124] R. Würdemann, "Berechnung optischer spektren und grundzustandseigenschaften neutraler und geladener moleküle mittels dichtefunktionaltheorie," Ph.D. thesis, Universität Freiburg, Freiburg, Germany, 2016.

[125] R. Würdemann and M. Walter, "Charge transfer excitations with range separated functionals using improved virtual orbitals," J. Chem. Theory Comput. **14**(7), 3667–3676 (2018).

[126] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," Phys. Rev. B **23**, 5048 (1981).

[127] S. Klüpfel, P. J. Klüpfel, and H. Jónsson, "Importance of complex orbitals in calculating the self-interaction-corrected ground state of atoms," Phys. Rev. A **84**, 050501(R) (2011).

[128] S. Lehtola and H. Jónsson, "Variational, self-consistent implementation of the Perdew–Zunger self-interaction correction with complex optimal orbitals," J. Chem. Theory Comput. **10**, 5324 (2014).

[129] S. Lehtola, M. Head-Gordon, and H. Jónsson, "Complex orbitals, multiple local minima, and symmetry breaking in Perdew–Zunger self-interaction corrected density functional theory calculations," J. Chem. Theory Comput. **12**, 3195 (2016).

[130] H. Jónsson, "Simulation of surface processes," Proc. Natl. Acad. Sci. U. S. A. **108**, 944 (2011).

[131] S. Klüpfel, P. Klüpfel, and H. Jónsson, "The effect of the Perdew-Zunger self-interaction correction to density functionals on the energetics of small molecules," J. Chem. Phys. **137**, 124102 (2012).

[132] H. Gudmundsdóttir, Y. Zhang, P. M. Weber, and H. Jónsson, "Self-interaction corrected density functional calculations of molecular Rydberg states," J. Chem. Phys. **139**, 194102 (2013).

[133] A. V. Ivanov, T. Ghosh, E. Ö. Jónsson, and H. Jónsson, "Mn dimer can be described accurately with density functional calculations when self-interaction correction is applied," J. Phys. Chem. Lett. **12**, 4240 (2021).

12 March 2024 13:19:04

[134] H. Gudmundsdóttir, E. Ö. Jónsson, and H. Jónsson, "Calculations of Al dopant in α-quartz using a variational implementation of the Perdew–Zunger self-interaction correction," New J. Phys. **17**, 083006 (2015).

[135] Y. L. A. Schmerwitz, A. V. Ivanov, E. Ö. Jónsson, H. Jónsson, and G. Levi, "Variational density functional calculations of excited states: Conical intersection and avoided crossing in ethylene bond twisting," J. Phys. Chem. Lett. **13**, 3990–3999 (2022).

[136] J. J. Mortensen, K. Kaasbjerg, S. L. Frederiksen, J. K. Nørskov, J. P. Sethna, and K. W. Jacobsen, "Bayesian error estimation in density-functional theory," Phys. Rev. Lett. **95**(21), 216401 (2005).

[137] J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, "Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation," Phys. Rev. B **85**, 235149 (2012).

[138] J. Wellendorff, K. T. Lundgaard, K. W. Jacobsen, and T. Bligaard, "mBEEF: An accurate semi-local Bayesian error estimation density functional," J. Chem. Phys. **140**(14), 144107 (2014).

[139] K. T. Lundgaard, J. Wellendorff, J. Voss, K. W. Jacobsen, and T. Bligaard, "mBEEF-vdW: Robust fitting of error estimation density functionals," Phys. Rev. B **93**(23), 235162 (2016).

[140] A. J. Medford, J. Wellendorff, A. Vojvodic, F. Studt, F. Abild-Pedersen, K. W. Jacobsen, T. Bligaard, and J. K. Nørskov, "Assessing the reliability of calculated catalytic ammonia synthesis rates," Science **345**(6193), 197–200 (2014).

[141] G. Ciccotti, M. Ferrario, and J. P. Ryckaert, "Molecular dynamics of rigid systems in Cartesian coordinates a general formulation," Mol. Phys. **47**, 1253–1264 (1982).

[142] A. A. Peterson, "Global optimization of adsorbate–surface structures while preserving molecular identity," Top. Catal. **57**, 40–53 (2014).

[143] E. B. Tadmor, G. S. Smith, N. Bernstein, and E. Kaxiras, "Mixed finite element and atomistic formulation for complex crystals," Phys. Rev. B **59**, 235 (1999).

[144] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," IMA J. Appl. Math. **6**, 222–231 (1970).

[145] R. Fletcher, "A new approach to variable metric algorithms," Comput. J. **13**, 317–322 (1970).

[146] D. Goldfarb, "A family of variable-metric methods derived by variational means," Math. Comput. **24**, 23–26 (1970).

[147] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," Math. Comput. **24**, 647–656 (1970).

[148] D. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," Math. Program. **45**, 503–528 (1989).

[149] E. Bitzek, P. Koskinen, F. Gahler, M. Moseler, and P. Gumbsch, "Structural relaxation made simple," Phys. Rev. Lett. **97**, 170201 (2006).

[150] D. Packwood, J. R. Kermode, L. Mones, N. Bernstein, J. Woolley, N. Gould, C. Ortner, and G. Csanyi, "A universal preconditioner for simulating condensed phase materials," J. Chem. Phys. **144**, 164109 (2016).

[151] E. Garijo del Rio, J. J. Mortensen, and K. W. Jacobsen, "Local Bayesian optimizer for atomic structures," Phys. Rev. B **100**, 104103 (2019).

[152] G. Henkelman and H. Jonsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," J. Chem. Phys. **111**, 7010 (1999).

[153] H. Jonsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Systems*, edited by B. J. Berne, G. Cicotti, and D. F. Coker (World Scientific, 1998).

[154] G. Henkelman and H. Jonsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," J. Chem. Phys. **113**, 9978 (2000).

[155] G. Henkelman, B. P. Uberuaga, and H. Jonsson, "A climbing-image nudged elastic band method for finding saddle points and minimum energy paths," J. Chem. Phys. **113**, 9901 (2000).

[156] S. Smidstrup, A. Pedersen, K. Stokbro, and H. Jonsson, "Improved initial guess for minimum energy path calculations," J. Chem. Phys. **140**, 214106 (2014).

[157] E. L. Kolsbjerg, M. N. Groves, and B. Hammer., "An automated nudged elastic band method," J. Chem. Phys. **145**, 094107 (2016).

[158] P. Lindgren, G. Kastlunger, and A. A. Peterson, "Scaled and dynamic optimizations of nudged elastic bands," J. Chem. Theory Comput. **15**, 5787–5793 (2019).

[159] S. Makri, C. Ortner, and J. R. Kermode, "A preconditioning scheme for minimum energy path finding methods," J. Chem. Phys. **150**, 094109 (2019).

[160] A. A. Peterson, "Acceleration of saddle-point searches with machine learning," J. Chem. Phys. **145**, 074106 (2016).

[161] O. Koistinen, F. B. Dagbjartsdottir, V. Asgeirsson, A. Vehtari, and H. Jonsson, "Nudged elastic band calculations accelerated with Gaussian process regression," J. Chem. Phys. **147**, 152720 (2017).

[162] J. A. Garrido Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard, "Low–scaling algorithm for nudged elastic band calculations using a surrogate machine learning model," Phys. Rev. Lett. **122**, 156001 (2019).

[163] M. Šarić, J. Rossmeisl, and P. G. Moses, "Modeling the adsorption of sulfur containing molecules and their hydrodesulfurization intermediates on the Co-promoted MoS_2 catalyst by DFT," J. Catal. **358**, 131–140 (2018).

[164] T. M. Østergaard, L. Giordano, I. E. Castelli, F. Maglia, B. K. Antonopoulos, Y. Shao-Horn, and J. Rossmeisl, "Oxidation of ethylene carbonate on Li metal oxide surfaces," J. Phys. Chem. C **122**(19), 10442–10449 (2018).

[165] L. Arnarson, P. S. Schmidt, M. Pandey, A. Bagger, K. S. Thygesen, I. E. L. Stephens, and J. Rossmeisl, "Fundamental limitation of electrocatalytic methane conversion to methanol," Phys. Chem. Chem. Phys. **20**(16), 11152–11159 (2018).

[166] H. Wan, A. W. Jensen, M. Escudero-Escribano, and J. Rossmeisl, "Insights in the oxygen reduction reaction: From metallic electrocatalysts to diporphyrins," ACS Catal. **10**(11), 5979–5989 (2020).

[167] G. Levi, E. Biasin, A. O. Dohn, and H. Jónsson, "On the interplay of solvent and conformational effects in simulated excited-state dynamics of a copper phenanthroline photosensitizer," Phys. Chem. Chem. Phys. **22**, 748–757 (2020).

[168] D. J. Wales and H. A. Scheraga, "Global optimization of clusters, crystals, and biomolecules," Science **285**, 1368 (1999).

[169] S. Goedecker, "Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems," J. Chem. Phys. **120**, 9911 (2004).

[170] L. B. Vilhelmsen and B. Hammer, "A genetic algorithm for first principles global structure optimization of supported nano structures," J. Chem. Phys. **141**, 044711 (2014).

[171] M.-P. V. Christiansen, N. Rønne, and B. Hammer, "Atomistic global optimization X: A Python package for optimization of atomistic structures," J. Chem. Phys. **157**(5), 054701 (2022).

[172] P. C. Jennings, S. Lysgaard, J. S. Hummelshøj, T. Vegge, and T. Bligaard, "Genetic algorithms for computational materials discovery accelerated by machine learning," npj Comput. Mater. **5**(1), 46 (2019).

[173] M. K. Bisbo and B. Hammer, "Efficient global structure optimization with a machine-learned surrogate model," Phys. Rev. Lett. **124**(8), 086102 (2020).

[174] S. Kaappa, C. Larsen, and K. W. Jacobsen, "Atomic structure optimization with machine-learning enabled interpolation between chemical elements," Phys. Rev. Lett. **127**(16), 166001 (2021).

[175] S. Kaappa, E. G. del Río, and K. W. Jacobsen, "Global optimization of atomic structures with gradient-enhanced Gaussian process regression," Phys. Rev. B **103**(17), 174114 (2021).

[176] C. Larsen, S. Kaappa, A. L. Vishart, T. Bligaard, and K. W. Jacobsen, "Machine-learning enabled optimization of atomic structures using atoms with fractional existence," Phys. Rev. B **107**, 214101 (2022).

[177] R. Wanzenböck, M. Arrigoni, S. Bichelmaier, F. Buchner, J. Carrete, and G. K. H. Madsen, "Neural-network-backed evolutionary search for $SrTiO_3$(110) surface reconstructions," Digital Discovery **1**(5), 703–710 (2022).

[178] A. Møgelhøj, A. K. Kelkkanen, K. T. Wikfeldt, J. Schiøtz, J. J. Mortensen, L. G. M. Pettersson, B. I. Lundqvist, K. W. Jacobsen, A. Nilsson, and J. K. Nørskov, "Ab initio van der Waals interactions in simulations of water alter structure from mainly tetrahedral to high-density-like," J. Phys. Chem. B **115**(48), 14149–14160 (2011).

[179] M. H. Hansen and J. Rossmeisl, "pH in grand canonical statistics of an electrochemical interface," J. Phys. Chem. C **120**(51), 29135–29143 (2016).

[180] A. O. Dohn, E. Ö. Jónsson, G. Levi, J. J. Mortensen, O. Lopez-Acevedo, K. S. Thygesen, K. W. Jacobsen, J. Ulstrup, N. E. Henriksen, K. B. Møller, and

12 March 2024 13:19:04

© Author(s) 2024

H. Jónsson, "Grid-based projector augmented wave (GPAW) implementation of quantum mechanics/molecular mechanics (QM/MM) electrostatic embedding and application to a solvated diplatinum complex," J. Chem. Theory Comput. **13**(12), 6010–6022 (2017).

[181] A. O. Dohn, E. Ö. Jónsson, K. S. Kjær, T. B. van Driel, M. M. Nielsen, K. W. Jacobsen, N. E. Henriksen, and K. B. Møller, "Direct dynamics studies of a binuclear metal complex in solution: The interplay between vibrational relaxation, coherence, and solvent effects," J. Phys. Chem. Lett. **5**(14), 2414–2418 (2014).

[182] T. B. van Driel, K. S. Kjær, R. W. Hartsock, A. O. Dohn, T. Harlang, M. Chollet, M. Christensen, W. Gawelda, N. E. Henriksen, J. G. Kim, K. Haldrup, K. H. Kim, H. Ihee, J. Kim, H. Lemke, Z. Sun, V. Sundström, W. Zhang, D. Zhu, K. B. Møller, M. M. Nielsen, and K. J. Gaffney, "Atomistic characterization of the active-site solvation dynamics of a model photocatalyst," Nat. Commun. **7**(1), 13678 (2016).

[183] G. Levi, M. Pápai, N. E. Henriksen, A. O. Dohn, and K. B. Møller, "Solution structure and ultrafast vibrational relaxation of the PtPOP complex revealed by ΔSCF-QM/MM direct dynamics simulations," J. Phys. Chem. C **122**, 7100–7119 (2018).

[184] K. Haldrup, G. Levi, E. Biasin, P. Vester, M. G. Laursen, F. Beyer, K. S. Kjær, T. Brandt van Driel, T. Harlang, A. O. Dohn, R. J. Hartsock, S. Nelson, J. M. Glownia, H. T. Lemke, M. Christensen, K. J. Gaffney, N. E. Henriksen, K. B. Møller, and M. M. Nielsen, "Ultrafast X-ray scattering measurements of coherent structural dynamics on the ground-state potential energy surface of a diplatinum molecule," Phys. Rev. Lett. **122**(6), 063001 (2019).

[185] A. O. Dohn, K. S. Kjær, T. B. Harlang, S. E. Canton, M. M. Nielsen, and K. B. Møller, "Electron transfer and solvent-mediated electronic localization in molecular photocatalysis," Inorg. Chem. **55**(20), 10637–10644 (2016).

[186] E. Ö. Jónsson, A. O. Dohn, and H. Jónsson, "Polarizable embedding with a transferable H_2O potential function I: Formulation and tests on dimer," J. Chem. Theory Comput. **15**(12), 6562–6577 (2019).

[187] A. O. Dohn, E. Ö. Jónsson, and H. Jónsson, "Polarizable embedding with a transferable H_2O potential function II: Application to $(H_2O)_n$ clusters and liquid water," J. Chem. Theory Comput. **15**(12), 6578–6587 (2019).

[188] T. Olsen, "Designing in-plane heterostructures of quantum spin Hall insulators from first principles: $1T'$–MoS_2 with adsorbates," Phys. Rev. B **94**, 235106 (2016).

[189] A. I. Liechtenstein, M. I. Katsnelson, V. P. Antropov, and V. A. Gubanov, "Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys," J. Magn. Magn. Mater. **67**, 65–74 (1987).

[190] N. D. Mermin and H. Wagner, "Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models," Phys. Rev. Lett. **17**, 1133–1136 (1966).

[191] J. L. Lado and J. Fernández-Rossier, "On the origin of magnetic anisotropy in two dimensional CrI_3," 2D Mater. **4**, 035002 (2017).

[192] D. Torelli, K. S. Thygesen, and T. Olsen, "High throughput computational screening for 2D ferromagnetic materials: The critical role of anisotropy and local correlations," 2D Mater. **6**, 045018 (2019).

[193] D. Torelli, H. Moustafa, K. W. Jacobsen, and T. Olsen, "High-throughput computational screening for two-dimensional magnetic materials based on experimental databases of three-dimensional compounds," npj Comput. Mater. **6**, 158 (2020).

[194] U. v. Barth and L. Hedin, "A local exchange-correlation potential for the spin polarized case. I," J. Phys. C: Solid State Phys. **5**, 1629–1642 (1972).

[195] G. Scalmani and M. J. Frisch, "A new approach to noncollinear spin density functional theory beyond the local density approximation," J. Chem. Theory Comput. **8**, 2193–2196 (2012).

[196] T. Y. Kim and C.-H. Park, "Magnetic anisotropy and magnetic ordering of transition-metal phosphorus trisulfides," Nano Lett. **21**, 10114–10121 (2021).

[197] D. Ceresoli, U. Gerstmann, A. P. Seitsonen, and F. Mauri, "First-principles theory of orbital magnetization," Phys. Rev. B **81**(6), 060409 (2010).

[198] A. J. P. Meyer and G. Asch, "Experimental g' and g values of Fe, Co, Ni, and their alloys," J. Appl. Phys. **32**(3), S330–S333 (1961).

[199] S. Foner, "High-field antiferromagnetic resonance in Cr_2O_3," Phys. Rev. **130**, 183–197 (1963).

[200] T. A. Kaplan, "Classical spin-configuration stability in the presence of competing exchange forces," Phys. Rev. **116**(4), 888–889 (1959).

[201] L. M. Sandratskii, "Energy band structure calculations for crystals with spiral magnetic structure," Phys. Status Solidi B **136**(1), 167–180 (1986).

[202] L. M. Sandratskii, "Insight into the Dzyaloshinskii–Moriya interaction through first-principles study of chiral magnetic structures," Phys. Rev. B **96**(2), 024450 (2017).

[203] R. Kubo, "Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems," J. Phys. Soc. Jpn. **12**(6), 570–586 (1957).

[204] E. K. U. Gross and W. Kohn, "Local density-functional theory of frequency-dependent linear response," Phys. Rev. Lett. **55**, 2850–2852 (1985).

[205] E. Runge and E. K. U. Gross, "Density-functional theory for time-dependent systems," Phys. Rev. Lett. **52**, 997–1000 (1984).

[206] S. Sharma, J. K. Dewhurst, A. Sanna, and E. K. U. Gross, "Bootstrap approximation for the exchange-correlation kernel of time-dependent density-functional theory," Phys. Rev. Lett. **107**, 186401 (2011).

[207] A. H. MacDonald, S. H. Vosko, and P. T. Coleridge, "Extensions of the tetrahedron method for evaluating spectral properties of solids," J. Phys. C: Solid State Phys. **12**(15), 2991 (1979).

[208] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, "Exact Coulomb cutoff technique for supercell calculations," Phys. Rev. B **73**, 205119 (2006).

[209] F. Hüser, T. Olsen, and K. S. Thygesen, "How dielectric screening in two-dimensional crystals affects the convergence of excited-state calculations: Monolayer MoS_2," Phys. Rev. B **88**(24), 245309 (2013).

[210] S. Latini, T. Olsen, and K. S. Thygesen, "Excitons in van der Waals heterostructures: The important role of dielectric screening," Phys. Rev. B **92**, 245123 (2015).

[211] C. E. Patrick and K. S. Thygesen, "Hubbard-U corrected Hamiltonians for non-self-consistent random-phase approximation total-energy calculations: A study of ZnS, TiO_2, and NiO," Phys. Rev. B **93**, 035133 (2016).

[212] P. S. Schmidt and K. S. Thygesen, "Benchmark database of transition metal surface and adsorption energies from many-body perturbation theory," J. Phys. Chem. C **122**(8), 4381–4390 (2018).

[213] T. Olsen, J. Yan, J. J. Mortensen, and K. S. Thygesen, "Dispersive and covalent interactions between graphene and metal surfaces from the random phase approximation," Phys. Rev. Lett. **107**, 156401 (2011).

[214] F. Furche and T. Van Voorhis, "Fluctuation-dissipation theorem density-functional theory," J. Chem. Phys. **122**(16), 164106 (2005).

[215] T. Olsen and K. S. Thygesen, "Beyond the random phase approximation: Improved description of short-range correlation by a renormalized adiabatic local density approximation," Phys. Rev. B **88**, 115131 (2013).

[216] T. Olsen and K. S. Thygesen, "Accurate ground-state energies of solids and molecules from time-dependent density-functional theory," Phys. Rev. Lett. **112**, 203001 (2014).

[217] C. E. Patrick and K. S. Thygesen, "Adiabatic-connection fluctuation-dissipation DFT for the structural properties of solids—The renormalized ALDA and electron gas kernels," J. Chem. Phys. **143**(10), 102802 (2015).

[218] T. Olsen, C. E. Patrick, J. E. Bates, A. Ruzsinszky, and K. S. Thygesen, "Beyond the RPA and GW methods with adiabatic xc-kernels for accurate ground state and quasiparticle energies," npj Comput. Mater. **5**(1), 106 (2019).

[219] P. Buczek, A. Ernst, and L. M. Sandratskii, "Different dimensionality trends in the Landau damping of magnons in iron, cobalt, and nickel: Time-dependent density functional study," Phys. Rev. B **84**, 174418 (2011).

[220] L. Van Hove, "Time-dependent correlations between spins and neutron scattering in ferromagnetic crystals," Phys. Rev. **95**(6), 1374–1384 (1954).

[221] T. Skovhus, T. Olsen, and H. M. Rønnow, "Influence of static correlation on the magnon dynamics of an itinerant ferromagnet with competing exchange interactions: First-principles study of mnbi," Phys. Rev. Mater. **6**, 054402 (2022).

[222] T. Skovhus and T. Olsen, "Magnons in antiferromagnetic bcc Cr and Cr_2O_3 from time-dependent density functional theory," Phys. Rev. B **106**, 085131 (2022).

[223] H. N. Rojas, R. W. Godby, and R. J. Needs, "Space-time method for *ab initio* calculations of self-energies and dielectric response functions of solids," Phys. Rev. Lett. **74**(10), 1827 (1995).

[224] C. Friedrich, S. Blügel, and A. Schindlmayr, "Efficient implementation of the *GW* approximation within the all-electron FLAPW method," Phys. Rev. B **81**(12), 125102 (2010).

12 March 2024 13:19:04

© Author(s) 2024

[225] I. Duchemin and X. Blase, "Robust analytic-continuation approach to many-body *GW* calculations," J. Chem. Theory Comput. **16**(3), 1742–1756 (2020).
[226] D. A. Leon, C. Cardoso, T. Chiarotti, D. Varsano, E. Molinari, and A. Ferretti, "Frequency dependence in *GW* made simple using a multipole approximation," Phys. Rev. B **104**(11), 115157 (2021).
[227] F. A. Rasmussen, P. S. Schmidt, K. T. Winther, and K. S. Thygesen, "Efficient many-body calculations for two-dimensional materials using exact limits for the screened potential: Band gaps of MoS_2, H-BN, and phosphorene," Phys. Rev. B **94**(15), 155406 (2016).
[228] A. Marini, C. Hogan, M. Grüning, and D. Varsano, "yambo: An *ab initio* tool for excited state calculations," Comput. Phys. Commun. **180**(8), 1392–1403 (2009).
[229] T. Olsen, S. Latini, F. Rasmussen, and K. S. Thygesen, "Simple screened hydrogen model of excitons in two-dimensional materials," Phys. Rev. Lett. **116**, 056401 (2016).
[230] T. Olsen, "Unified treatment of magnons and excitons in monolayer CrI_3 from many-body perturbation theory," Phys. Rev. Lett. **127**, 166402 (2021).
[231] Z. Li, P. Graziosi, and N. Neophytou, "Deformation potential extraction and computationally efficient mobility calculations in silicon from first principles," Phys. Rev. B **104**(19), 195201 (2021).
[232] A. Y. Liu and A. A. Quong, "Linear-response calculation of electron-phonon coupling parameters," Phys. Rev. B **53**(12), R7575 (1996).
[233] F. Giustino, "Electron-phonon interactions from first principles," Rev. Mod. Phys. **89**, 015003 (2017).
[234] M. Walter and M. Moseler, "Ab initio wavelength-dependent Raman spectra: Placzek approximation and beyond," J. Chem. Theory Comput. **16**(1), 576–586 (2020).
[235] A. Jorio, M. S. Dresselhaus, R. Saito, and G. Dresselhaus, *Raman Spectroscopy in Graphene Related Systems* (John Wiley & Sons, 2011).
[236] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, "Implementation strategies in phonopy and phono3py," J. Phys.: Condens. Matter **35**(35), 353001 (2023).
[237] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Elsevier Science Publishing Co, Inc., Amsterdam, 2008).
[238] A. Taghizadeh, F. Hipolito, and T. G. Pedersen, "Linear and nonlinear optical response of crystals using length and velocity gauges: Effect of basis truncation," Phys. Rev. B **96**, 195413 (2017).
[239] C. Aversa and J. E. Sipe, "Nonlinear optical susceptibilities of semiconductors: Results with a length-gauge analysis," Phys. Rev. B **52**, 14636–14645 (1995).
[240] A. Sakko, T. P. Rossi, and R. M. Nieminen, "Dynamical coupling of plasmons and molecular excitations by hybrid quantum/classical calculations: Time-domain approach," J. Phys.: Condens. Matter **26**(31), 315013 (2014).
[241] T. P. Rossi, K. T. Winther, K. W. Jacobsen, R. M. Nieminen, M. J. Puska, and K. S. Thygesen, "Effect of edge plasmons on the optical properties of MoS_2 monolayer flakes," Phys. Rev. B **96**, 155407 (2017).
[242] P. V. Kumar, T. P. Rossi, M. Kuisma, P. Erhart, and D. J. Norris, "Direct hot-carrier transfer in plasmonic catalysis," Faraday Discuss. **214**, 189–197 (2019).
[243] P. V. Kumar, T. P. Rossi, D. Marti-Dafcik, D. Reichmuth, M. Kuisma, P. Erhart, M. J. Puska, and D. J. Norris, "Plasmon-induced direct hot-carrier transfer at metal–acceptor interfaces," ACS Nano **13**(3), 3188–3195 (2019).
[244] T. P. Rossi, T. Shegai, P. Erhart, and T. J. Antosiewicz, "Strong plasmon-molecule coupling at the nanoscale revealed by first-principles modeling," Nat. Commun. **10**(1), 3336 (2019).
[245] T. P. Rossi, P. Erhart, and M. Kuisma, "Hot-carrier generation in plasmonic nanoparticles: The importance of atomic structure," ACS Nano **14**(8), 9963–9971 (2020).
[246] J. Fojt, T. P. Rossi, M. Kuisma, and P. Erhart, "Hot-carrier transfer across a nanoparticle–molecule junction: The importance of orbital hybridization and level alignment," Nano Lett. **22**(21), 8786–8792 (2022).
[247] D. Sorvisto, P. Rinke, and T. P. Rossi, "Single-atom dopants in plasmonic nanocatalysts," J. Phys. Chem. C **127**(18), 8585–8590 (2023).
[248] K. Yabana and G. F. Bertsch, "Time-dependent local-density approximation in real time," Phys. Rev. B **54**(7), 4484–4487 (1996).
[249] S. Malola, L. Lehtovaara, J. Enkovaara, and H. Häkkinen, "Birth of the localized surface plasmon resonance in monolayer-protected gold nanoclusters," ACS Nano **7**(11), 10263–10270 (2013).
[250] C. Schäfer and G. Johansson, "Shortcut to self-consistent light-matter interaction and realistic spectra from first principles," Phys. Rev. Lett. **128**, 156402 (2022).
[251] C. Schäfer, "Polaritonic chemistry from first principles via embedding radiation reaction," J. Phys. Chem. Lett. **13**(30), 6905–6911 (2022).
[252] A. Ojanperä, V. Havu, L. Lehtovaara, and M. Puska, "Nonadiabatic Ehrenfest molecular dynamics within the projector augmented-wave method," J. Chem. Phys. **136**(14), 144103 (2012).
[253] A. Ojanperä, A. V. Krasheninnikov, and M. Puska, "Electronic stopping power from first-principles calculations with account for core electron excitations and projectile ionization," Phys. Rev. B **89**(3), 035120 (2014).
[254] O. A. Syzgantseva, M. Puska, and K. Laasonen, "Charge transfer at the hybrid interfaces in the presence of water: A theoretical study," J. Phys. Chem. C **119**(51), 28347–28352 (2015).
[255] C. Brand, M. Debiossac, T. Susi, F. Aguillon, J. Kotakoski, P. Roncin, and M. Arndt, "Coherent diffraction of hydrogen through the 246 pm lattice of graphene," New J. Phys. **21**(3), 033004 (2019).
[256] E. A. Buntov and A. F. Zatsepin, "Carbon bond breaking under Ar^+-ion irradiation in dependence on sp hybridization: Car–parrinello, Ehrenfest, and classical dynamics study," J. Phys. Chem. A **124**(44), 9128–9132 (2020).
[257] Y. Aierken, A. Agrawal, M. Sun, M. Melander, E. J. Crumlin, B. A. Helms, and D. Prendergast, "Revealing charge-transfer dynamics at electrified sulfur cathodes using constrained density functional theory," J. Phys. Chem. Lett. **12**(2), 739–744 (2021).
[258] E. Artacho and D. D. O'Regan, "Quantum mechanics in an evolving Hilbert space," Phys. Rev. B **95**(11), 115155 (2017).
[259] A. García *et al.*, "SIESTA: Recent developments and applications," J. Chem. Phys. **152**(20), 204108 (2020).
[260] A. Dreuw and M. Head-Gordon, "Failure of time-dependent density functional theory for long-range charge-transfer excited states: The zincbacteriochlorin–bacteriochlorin and bacteriochlorophyll–spheroidene complexes," J. Am. Chem. Soc. **126**(12), 4007–4016 (2004).
[261] I. Seidu, M. Krykunov, and T. Ziegler, "Applications of time-dependent and time-independent density functional theory to Rydberg transitions," J. Phys. Chem. A **119**(21), 5107–5116 (2015).
[262] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, "Conical intersections and double excitations in time-dependent density functional theory," Mol. Phys. **104**, 1039–1051 (2006).
[263] S. Huzinaga and C. Arnau, "Virtual orbitals in Hartree-Fock theory," Phys. Rev. A **1**(5), 1285–1288 (1970).
[264] S. Huzinaga and C. Arnau, "Virtual orbitals in Hartree–Fock theory. II," J. Chem. Phys. **54**(5), 1948–1951 (1971).
[265] D. Hait and M. Head-Gordon, "Orbital optimized density functional theory for electronic excited states," J. Phys. Chem. Lett. **12**, 4517–4529 (2021).
[266] G. M. J. Barca, A. T. B. Gilbert, and P. M. W. Gill, "Simple models for difficult electronic excitations," J. Chem. Theory Comput. **14**, 1501–1509 (2018).
[267] A. T. B. Gilbert, N. A. Besley, and P. M. W. Gill, "Self-consistent field calculations of excited states using the maximum overlap method (MOM)," J. Phys. Chem. A **112**, 13164–13171 (2008).
[268] G. Levi, A. V. Ivanov, and H. Jónsson, "Variational calculations of excited states *via* direct optimization of the orbitals in DFT," Faraday Discuss. **224**, 448–466 (2020).
[269] P. F. Loos, M. Comin, X. Blase, and D. Jacquemin, "Reference energies for intramolecular charge-transfer excitations," J. Chem. Theory Comput. **17**, 3666–3686 (2021).
[270] T. Katayama, T. K. Choi, D. Khakhulin, A. O. Dohn, C. J. Milne, G. Vankó, Z. Németh, F. A. Lima, J. Szlachetko, T. Sato, S. Nozawa, S. I. Adachi, M. Yabashi, T. J. Penfold, W. Gawelda, and G. Levi, "Atomic-scale observation of solvent reorganization influencing photoinduced structural dynamics in a copper complex photosensitizer," Chem. Sci. **14**, 2572–2584 (2023).
[271] H. Ma, M. Govoni, and G. Galli, "Quantum simulations of materials on near-term quantum computers," npj Comput. Mater. **6**(1), 85 (2020).
[272] M. Abedi, G. Levi, D. B. Zederkof, N. E. Henriksen, M. Pápai, and K. B. Møller, "Excited-state solvation structure of transition metal complexes from molecular dynamics simulations and assessment of partial atomic charge methods," Phys. Chem. Chem. Phys. **21**, 4082–4095 (2019).

12 March 2024 13:19:04

© Author(s) 2024

[273] A. V. Ivanov, Y. L. A. Schmerwitz, G. Levi, and H. Jónsson, "Electronic excitations of the charged nitrogen-vacancy center in diamond obtained using time-independent variational density functional calculations," SciPost Phys. **15**, 009 (2023).

[274] R. D. King-Smith and D. Vanderbilt, "Theory of polarization of crystalline solids," Phys. Rev. B **47**, 1651–1654 (1993).

[275] R. Resta, "Macroscopic polarization in crystalline dielectrics: The geometric phase approach," Rev. Mod. Phys. **66**, 899–915 (1994).

[276] R. Resta and D. Vanderbilt, "Theory of polarization: A modern approach," in *Physics of Ferroelectrics* (Springer, 2007), Vol. 105, pp. 31–68.

[277] Y. Yoneda, K. Ohara, and H. Nagata, "Local structure and phase transitions of KNbO$_3$," Jpn. J. Appl. Phys. **57**, 11UB07 (2018).

[278] M. Gibertini and N. Marzari, "Emergence of one-dimensional wires of free carriers in transition-metal-dichalcogenide nanostructures," Nano Lett. **15**, 6229–6238 (2015).

[279] J. Sødequist, U. Petralanda, and T. Olsen, "Abundance of second order topology in C$_3$ symmetric two-dimensional insulators," 2D Mater. **10**, 015009 (2023).

[280] 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).

[281] R. Fei, W. Kang, and L. Yang, "Ferroelectricity and phase transitions in monolayer group-IV monochalcogenides," Phys. Rev. Lett. **117**, 097601 (2016).

[282] T. Rangel, B. M. Fregoso, B. S. Mendoza, T. Morimoto, J. E. Moore, and J. B. Neaton, "Large bulk photovoltaic effect and spontaneous polarization of single-layer monochalcogenides," Phys. Rev. Lett. **119**, 067402 (2017).

[283] H. Wang and X. Qian, "Two-dimensional multiferroics in monolayer group IV monochalcogenides," 2D Mater. **4**, 015042 (2017).

[284] U. Petralanda and T. Olsen, "Polarization switching induced by domain wall sliding in two-dimensional ferroelectric monochalcogenides," 2D Mater. **10**, 015001 (2023).

[285] M. Taherinejad, K. F. Garrity, and D. Vanderbilt, "Wannier center sheets in topological insulators," Phys. Rev. B **89**, 115102 (2014).

[286] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," Phys. Rev. B **56**(20), 12847 (1997).

[287] X. Qian, J. Liu, L. Fu, and J. Li, "Quantum spin Hall effect in two-dimensional transition metal dichalcogenides," Science **346**, 1344–1347 (2014).

[288] R. Resta and S. Sorella, "Electron localization in the insulating state," Phys. Rev. Lett. **82**(2), 370 (1999).

[289] G. Berghold, C. J. Mundy, A. H. Romero, J. . Hutter, and M. Parrinello, "General and efficient algorithms for obtaining maximally localized Wannier functions," Phys. Rev. B **61**(15), 10040 (2000).

[290] K. S. Thygesen, L. B. Hansen, and K. W. Jacobsen, "Partly occupied Wannier functions," Phys. Rev. Lett. **94**(2), 026405 (2005).

[291] K. S. Thygesen, L. B. Hansen, and K. W. Jacobsen, "Partly occupied Wannier functions: Construction and applications," Phys. Rev. B **72**(12), 125119 (2005).

[292] P. F. Fontana, A. H. Larsen, T. Olsen, and K. S. Thygesen, "Spread-balanced Wannier functions: Robust and automatable orbital localization," Phys. Rev. B **104**(12), 125140 (2021).

[293] C. E. Dreyer, A. Alkauskas, J. L. Lyons, A. Janotti, and C. G. Van de Walle, "First-principles calculations of point defects for quantum technologies," Annu. Rev. Mater. Res. **48**, 1–26 (2018).

[294] J. S. Park, S. Kim, Z. Xie, and A. Walsh, "Point defect engineering in thin-film solar cells," Nat. Rev. Mater. **3**(7), 194–210 (2018).

[295] L. Weston, D. Wickramaratne, M. Mackoit, A. Alkauskas, and C. G. Van de Walle, "Native point defects and impurities in hexagonal boron nitride," Phys. Rev. B **97**(21), 214104 (2018).

[296] T. Q. P. Vuong, G. Cassabois, P. Valvin, A. Ouerghi, Y. Chassagneux, C. Voisin, and B. Gil, "Phonon-photon mapping in a color center in hexagonal boron nitride," Phys. Rev. Lett. **117**(9), 097402 (2016).

[297] C. G. Van de Walle and J. Neugebauer, "First-principles calculations for defects and impurities: Applications to III-nitrides," J. Appl. Phys. **95**(8), 3851–3879 (2004).

[298] S. Lany and A. Zunger, "Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs," Phys. Rev. B **78**(23), 235104 (2008).

[299] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, "First-principles calculations for point defects in solids," Rev. Mod. Phys. **86**(1), 253 (2014).

[300] A. Zunger and O. I. Malyi, "Understanding doping of quantum materials," Chem. Rev. **121**(5), 3031–3060 (2021).

[301] O. O. Kurakevych and V. L. Solozhenko, "Rhombohedral boron subnitride, B$_{13}$N$_2$, by X-ray powder diffraction," Acta Crystallogr., Sect. C: Cryst. Struct. Commun. **63**(Pt9), i80–i82 (2007).

[302] G. Cassabois, P. Valvin, and B. Gil, "Hexagonal boron nitride is an indirect bandgap semiconductor," Nat. Photonics **10**(4), 262–266 (2016).

[303] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, "Fully *ab initio* finite-size corrections for charged-defect supercell calculations," Phys. Rev. Lett. **102**(1), 016402 (2009).

[304] S. Kaappa, S. Malola, and H. Häkkinen, "Point group symmetry analysis of the electronic structure of bare and protected metal nanocrystals," J. Phys. Chem. A **122**(43), 8576–8584 (2018).

[305] F. Bertoldo, S. Ali, S. Manti, and K. S. Thygesen, "Quantum point defects in 2D materials–the QPOD database," npj Comput. Mater. **8**(1), 56 (2022).

[306] J. F. Cornwell, *Group Theory in Physics*, *Techniques of Physics Vol. 1* (Academic Press, San Diego, 1997).

[307] V. Popescu and A. Zunger, "Extracting E versus $\vec{k}$ effective band structure from supercell calculations on alloys and impurities," Phys. Rev. B **85**(8), 085201 (2012).

[308] Quantum electrostatic heterostructure model, https://qeh.readthedocs.io/en/latest/ (accessed 15 May 2023).

[309] K. Andersen, S. Latini, and K. S. Thygesen, "Dielectric genome of van der Waals heterostructures," Nano Lett. **15**(7), 4616–4621 (2015).

[310] Calculating new building blocks, https://qeh.readthedocs.io/en/latest/tutorials.html (accessed 15 May 2023).

[311] A. Held and M. Walter, "Simplified continuum solvent model with a smooth cavity based on volumetric data," J. Chem. Phys. **141**(17), 174108 (2014).

[312] R. Wessling, R. Delgado Andrés, I. Morhenn, P. Acker, W. Maftuhin, M. Walter, U. Würfel, and B. Esser, "Phenothiazine-based donor–acceptor polymers as multifunctional materials for charge storage and solar energy conversion," Macromol. Rapid Commun. **45**, 2200699 (2022).

[313] O. Brügner, T. Reichenbach, M. Sommer, and M. Walter, "Substituent correlations characterized by Hammett constants in the spiropyran–merocyanine transition," J. Phys. Chem. A **121**(13), 2683–2687 (2017).

[314] A. Y. Lozovoi, A. Alavi, J. Kohanoff, and R. M. Lynden-Bell, "*Ab initio* simulation of charged slabs at constant chemical potential," J. Chem. Phys. **115**(4), 1661–1669 (2001).

[315] K. Sakaushi, T. Kumeda, S. Hammes-Schiffer, M. M. Melander, and O. Sugino, "Advances and challenges for experiment and theory for multi-electron multi-proton transfer at electrified solid–liquid interfaces," Phys. Chem. Chem. Phys. **22**(35), 19401–19442 (2020).

[316] P. Lindgren, G. Kastlunger, and A. A. Peterson, "Electrochemistry from the atomic scale, in the electronically grand-canonical ensemble," J. Chem. Phys. **157**(18), 180902 (2022).

[317] M. Otani and O. Sugino, "First-principles calculations of charged surfaces and interfaces: A plane-wave nonrepeated slab approach," Phys. Rev. B **73**(11), 115407 (2006).

[318] K. Mathew, R. Sundararaman, K. Letchworth-Weaver, T. A. Arias, and R. G. Hennig, "Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways," J. Chem. Phys. **140**(8), 084106 (2014).

[319] K. Letchworth-Weaver and T. A. Arias, "Joint density functional theory of the electrode-electrolyte interface: Application to fixed electrode potentials, interfacial capacitances, and potentials of zero charge," Phys. Rev. B **86**(7), 075140 (2012).

[320] N. G. Hörmann, O. Andreussi, and N. Marzari, "Grand canonical simulations of electrochemical interfaces in implicit solvation models," J. Chem. Phys. **150**(4), 041730 (2019).

[321] S. Ringe, N. G. Hörmann, H. Oberhofer, and K. Reuter, "Implicit solvation methods for catalysis at electrified interfaces," Chem. Rev. **122**(12), 10777–10820 (2022).

12 March 2024 13:19:04

© Author(s) 2024

[322] G. Kastlunger, P. Lindgren, and A. A. Peterson, "Controlled-potential simulation of elementary electrochemical reactions: Proton discharge on metal surfaces," J. Phys. Chem. C **122**(24), 12771–12781 (2018).

[323] S. Trasatti, "The absolute electrode potential: An explanatory note (Recommendations 1986)," J. Electroanal. Chem. Interfacial Electrochem. **209**(2), 417–428 (1986).

[324] R. Sundararaman and K. Schwarz, "Evaluating continuum solvation models for the electrode-electrolyte interface: Challenges and strategies for improvement," J. Chem. Phys. **146**(8), 084111 (2017).

[325] M. M. Melander, M. J. Kuisma, T. E. K. Christensen, and K. Honkala, "Grand-canonical approach to density functional theory of electrocatalytic systems: Thermodynamics of solid-liquid interfaces at constant ion and electrode potentials," J. Chem. Phys. **150**(4), 041706 (2019).

[326] M. Weinert and J. W. Davenport, "Fractional occupations and density-functional energies and forces," Phys. Rev. B **45**, 13709–13712 (1992).

[327] P. Lindgren, G. Kastlunger, and A. A. Peterson, "A challenge to the $G \sim 0$ interpretation of hydrogen evolution," ACS Catal. **10**(1), 121–128 (2020).

[328] G. Kastlunger, L. Wang, N. Govindarajan, H. H. Heenen, S. Ringe, T. Jaramillo, C. Hahn, and K. Chan, "Using pH dependence to understand mechanisms in electrochemical CO reduction," ACS Catal. **12**(8), 4344–4357 (2022).

[329] G. Kastlunger, H. H. Heenen, and N. Govindarajan, "Combining first-principles kinetics and experimental data to establish guidelines for product selectivity in electrochemical CO_2 reduction," ACS Catal. **13**(7), 5062–5072 (2023).

[330] M. Melander, T. Wu, and K. Honkala, "Constant inner potential DFT for modelling electrochemical systems under constant potential and bias," npj Comput. Mater. **10**, 5 (2024).

[331] Q. Wu and T. Van Voorhis, "Direct optimization method to study constrained systems within density-functional theory," Phys. Rev. A **72**, 024502 (2005).

[332] Q. Wu and T. Van Voorhis, "Extracting electron transfer coupling elements from constrained density functional theory," J. Chem. Phys. **125**(16), 164105 (2006).

[333] B. Kaduk, T. Kowalczyk, and T. Van Voorhis, "Constrained density functional theory," Chem. Rev. **112**(1), 321–370 (2012).

[334] M. Melander, E. Ö. Jónsson, J. J. Mortensen, T. Vegge, and J. M. García Lastra, "Implementation of constrained DFT for computing charge transfer rates within the projector augmented wave method," J. Chem. Theory Comput. **12**(11), 5367–5378 (2016).

[335] H. Park, N. Kumar, M. Melander, T. Vegge, J. M. Garcia Lastra, and D. J. Siegel, "Adiabatic and nonadiabatic charge transport in Li–S batteries," Chem. Mater. **30**(3), 915–928 (2018).

[336] T. Kumeda, L. Laverdure, K. Honkala, M. M. Melander, and K. Sakaushi, "Cations determine the mechanism and selectivity of alkaline oxygen reduction reaction on Pt(111)," Angew. Chem., Int. Ed. **62**, e202312841 (2022).

[337] M. M. Melander, "Grand canonical rate theory for electrochemical and electrocatalytic systems I: General formulation and proton-coupled electron transfer reactions," J. Electrochem. Soc. **167**(11), 116518 (2020).

[338] M. Levy, J. P. Perdew, and V. Sahni, "Exact differential equation for the density and ionization energy of a many-particle system," Phys. Rev. A **30**, 2745–2748 (1984).

[339] J. Lehtomäki, I. Makkonen, M. A. Caro, A. Harju, and O. Lopez-Acevedo, "Orbital-free density functional theory implementation with the projector augmented-wave method," J. Chem. Phys. **141**(23), 234102 (2014).

[340] J. Lehtomäki and O. Lopez-Acevedo, "Large-Z limit in atoms and solids from first principles," J. Chem. Phys. **151**(24), 244101 (2019).

[341] V. Ivády, T. Simon, J. R. Maze, I. A. Abrikosov, and A. Gali, "Pressure and temperature dependence of the zero-field splitting in the ground state of NV centers in diamond: A first-principles study," Phys. Rev. B **90**(23), 235205 (2014).

[342] M. J. Rayson and P. R. Briddon, "First principles method for the calculation of zero-field splitting tensors in periodic systems," Phys. Rev. B **77**(3), 035119 (2008).

[343] T. Biktagirov, W. G. Schmidt, and U. Gerstmann, "Spin decontamination for magnetic dipolar coupling calculations: Application to high-spin molecules and solid-state spin qubits," Phys. Rev. Res. **2**(2), 022024 (2020).

[344] P. E. Blöchl, "First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen," Phys. Rev. B **62**(10), 6158 (2000).

[345] 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 (1999).

[346] Y. Yang and W. Yang, "Comment on 'Generalized Gradient Approximation Made Simple," Phys. Rev. Lett. **80**, 890 (1998).

12 March 2024 13:19:04

© Author(s) 2024