

This is the peer reviewed version of the following article:

Implementation and Validation of Fully Relativistic GW Calculations: Spin-Orbit Coupling in Molecules, Nanocrystals, and Solids / Scherpelz, Peter; Govoni, Marco; Hamada, Ikutaro; Galli, Giulia. - In: JOURNAL OF CHEMICAL THEORY AND COMPUTATION. - ISSN 1549-9626. - 12:8(2016), pp. 3523-3544. [10.1021/acs.jctc.6b00114]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

18/12/2025 18:58

Article

Implementation and Validation of Fully-Relativistic GW Calculations: Spin-Orbit Coupling in Molecules, Nanocrystals and Solids

Peter Scherpelz, Marco Govoni, Ikutaro Hamada, and Giulia Galli

J. Chem. Theory Comput., **Just Accepted Manuscript** • DOI: 10.1021/acs.jctc.6b00114 • Publication Date (Web): 22 Jun 2016

Downloaded from <http://pubs.acs.org> on July 4, 2016

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

Implementation and Validation of Fully-Relativistic GW Calculations: Spin-Orbit Coupling in Molecules, Nanocrystals and Solids

Peter Scherpelz,^{*,†,‡} Marco Govoni,^{†,¶} Ikutaro Hamada,[§] and Giulia Galli^{†,¶}

Institute for Molecular Engineering, The University of Chicago, Chicago, IL, USA, Intelligence Community Postdoctoral Research Fellowship Program, Materials Science Division, Argonne National Laboratory, Argonne, IL, USA, and International Center for Materials Nanoarchitectonics, Global Research Center for Environment and Energy based on Nanomaterials Science, and Center for Materials Research by Information Integration, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan

E-mail: pscherpelz@uchicago.edu

Abstract

We present an implementation of G_0W_0 calculations including spin-orbit coupling (SOC) enabling investigations of large systems, with thousands of electrons, and we discuss results for molecules, solids and nanocrystals. Using a newly-developed set

^{*}To whom correspondence should be addressed

[†]Institute for Molecular Engineering, The University of Chicago, Chicago, IL, USA

[‡]Intelligence Community Postdoctoral Research Fellowship Program

[¶]Materials Science Division, Argonne National Laboratory, Argonne, IL, USA

[§]International Center for Materials Nanoarchitectonics, Global Research Center for Environment and Energy based on Nanomaterials Science, and Center for Materials Research by Information Integration, National Institute for Materials Science, 1-1 Namiki, Tsukuba 305-0044, Japan

of molecules with heavy elements (called GW-SOC81), we find that when based upon hybrid density functional calculations, fully-relativistic (FR) and scalar-relativistic (SR) G_0W_0 calculations of vertical ionization potentials both yield excellent performance compared to experiment, with errors below 1.9%. We demonstrate that while SR calculations have higher random errors, FR calculations systematically underestimate the VIP by 0.1 to 0.2 eV. We further verify that SOC effects may be well approximated at the FR density functional level and then added to SR G_0W_0 results for a broad class of systems. We also address the use of different root-finding algorithms for the G_0W_0 quasiparticle equation, and the significant influence of including d electrons in the valence partition of the pseudopotential for G_0W_0 calculations. Finally, we present statistical analyses of our data, highlighting the importance of separating definitive improvements from those that may occur by chance due to a limited number of samples. We suggest the statistical analyses used here will be useful in the assessment of the accuracy of a large variety of electronic structure methods.

1 Introduction

First-principles calculations of the electronic properties of molecules, solids and nanostructures are of great importance, both in interpreting experimental findings and in seeking to predict new materials with targeted properties. Density functional theory (DFT)^{1,2} has been successful in describing the properties of broad classes of systems,³ and is widely used, but it still suffers from a number of shortcomings, especially in the calculation of absolute energies and excited state properties.⁴ While hybrid functionals⁵ have led to improved results compared to semilocal functionals in many applications,^{6–8} and recent work has removed the need for empirical parameters in some cases,^{7,8} the dependence of most density functionals on material-specific parameters still limits their applicability and accuracy, especially in inhomogeneous systems such as interfaces.

The presence of heavy elements further complicates material studies, due to the spin-orbit

effect.⁹ For example, the band gap and ionization potentials of systems containing elements such as lead, gold, and iodine varies by as much as 1 eV when spin-orbit coupling (SOC) is introduced. Further, topological insulators and similar systems are of great current interest, and many topological effects are a direct result of SOC.¹⁰ The latter can be described by fully-relativistic (FR) many-body perturbation theory (MBPT), and specifically FR *GW* calculations.^{11–13}

In this paper we report the inclusion of SOC in a recently developed *GW* method and code, WEST, which can perform calculations at the G_0W_0 level of unprecedented size, including nanoparticles, liquids, and interfaces, in addition to crystals and molecules.^{14,15} This represents the first implementation of a G_0W_0 calculation in which both G and W are computed at the FR level, and the use of empty states is avoided. It is further notable in going beyond plasmon-pole^{12,13,16–21} and analytic continuation^{21–24} methods.

The description of SOC in DFT and beyond-DFT methods requires the use of two-component spinors, instead of scalar wave functions;^{25–29} related 2-component and 4-component methods have also been developed in quantum chemistry.^{30–35} Due to the need for multiple components, its inclusion into existing electronic structure methods is nontrivial. Only in the past few years has SOC been included at the *GW* level of theory in electronic structure codes.³⁶ A small set of systems have now been investigated by SOC *GW* approaches, including mercury chalcogenides,³⁶ Pu, Am,³⁷ and actinide³⁸ metals, atomically thin MoS₂,³⁹ topological insulators and related materials,^{40–46} Pb₄S₄ nanoparticles,⁴⁷ perovskites,^{48–53} and small molecules.³⁴

We go beyond previous work in performing a large survey of the applicability of FR G_0W_0 to a variety of systems. Included in this survey is the new GW-SOC81 set of 81 molecules, which can further serve as a tool for the validation of future SOC implementations. Also included are statistical analyses of the performance of different methods, e.g. SR and FR calculations with multiple functionals at the DFT and G_0W_0 levels of theory. We report a detailed analysis of the actual statistical significance of our results, and not only a comparison

of the mean absolute errors which may differ simply due to the specific test sets that were chosen. These analyses should be broadly applicable to method performance studies in the electronic structure community.

We also use the large set of calculations and systems presented here to investigate two important issues: First, whether using the SOC corrections found at the DFT level of theory can be a good strategy for estimating SOC effects at the G_0W_0 level of theory. This issue is especially relevant for large systems, as the computational penalty for FR GW vs. SR GW is severe, with at least 8 to 16 times larger computational resources required.

Second, we discuss the solutions of the G_0W_0 quasiparticle equation as a function of frequency.^{12,13} This equation must be solved through a root-finding procedure in order to obtain the quasiparticle energies, and its solutions may not be unique.^{54,55} The GW method used here allows for efficient computations of energy as a function of frequency, which in several cases is critical to finding the correct, physical solution.

This paper presents the theoretical methods in Section 2. In Section 3 we use a set of molecular vertical ionization potentials to assess the improvement of FR GW over SR GW calculations for molecules. Section 4 presents an analysis of the method’s performance for bulk semiconductors. Next complex systems, e.g. organic-inorganic hybrid perovskites and lead chalcogenide nanoparticles, are explored in Sections 5 and 6 respectively. Section 7 investigates the use of SOC corrections obtained at a DFT level to approximate FR GW results, and Section 8 describes the results of graphical solutions of the quasiparticle equation. We conclude in Section 9.

2 Method

Throughout this paper we consider two approaches to incorporate relativistic effects in our computations: Scalar-relativistic (SR) and fully-relativistic (FR) calculations. In the former, relativistic effects are included in the construction of the pseudopotential used to model core-

valence electron interactions.^{3,56–58} Scalar-relativistic calculations lead to an error in the total energy, due to the neglect of spin-orbit coupling, proportional to $Z^2\alpha^2$,^{58–60} where Z is the ion's atomic number and α the fine-structure constant.

In contrast, FR calculations explicitly incorporate spin-orbit coupling; we also refer to these as SOC calculations throughout this paper. These calculations use pseudopotentials generated using the full four-component Dirac equation, i.e. all core electron effects are treated using the Dirac equation. Two components are used in DFT and *GW* calculations to include spin-orbit effects on valence electrons. These calculations lead to an error due to relativistic effects of order α^2 ,^{58,59} which is notably smaller than that of SR for heavy elements with e.g. $Z \gtrsim 30$.

2.1 Relativistic effects in the Kohn-Sham equations

Spin-orbit coupling is included in the solution of the Kohn-Sham (KS) equations whose eigenvalues are then used as input for perturbative *GW* calculations. The description of spin-orbit coupling interactions requires electronic states to be noncollinear; they are represented as spinors

$$\psi_i = \begin{pmatrix} \psi_i^\uparrow \\ \psi_i^\downarrow \end{pmatrix}. \quad (1)$$

In our work we used the spin-orbit coupling implementation^{27,29} of QUANTUM ESPRESSO (QE),^{61,62} where spin-orbit (SO) interactions are introduced in the nonlocal part of the pseudopotential, which becomes a 2×2 matrix operating on the spinor wave functions. Separate pseudopotential projectors are used for each $j = l \pm \frac{1}{2}$ where l is the orbital angular momentum component of pseudopotentials, as labeled in ordinary SR calculations.^{59,63}

The reason why it is sufficiently accurate to include SO interactions only in the pseudopotential is that the SO contribution to the Hamiltonian is given by:^{3,64}

$$\hat{H}_{\text{SO}} \propto \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S}, \quad (2)$$

where V is the system potential, r the radial coordinate referenced to the atomic nucleus, \mathbf{L} the orbital angular momentum, and \mathbf{S} the electron spin. The term $\frac{1}{r} \frac{dV}{dr}$ is non-negligible only near the nucleus of a heavy atom, where r is small and $\frac{dV}{dr}$ is large. As a result, SO interactions only need to be included explicitly in the four-component all-electron Dirac equations⁵⁹ solved for the atoms and used to generate an FR pseudopotential.

2.2 G_0W_0 calculations

In perturbative GW (termed G_0W_0) calculations,^{11–13,65} quasiparticle energies E^{QP} are obtained from KS eigenvalues ϵ^{KS} and orbitals $\psi_{\mathbf{k}\mu}$:

$$E_{\mathbf{k}\mu}^{\text{QP}} = \epsilon_{\mathbf{k}\mu}^{\text{KS}} + \sum_{\alpha\beta} \left\langle \psi_{\mathbf{k}\mu}^{\alpha} \left| \hat{\Sigma}_{\alpha\beta} \left(E_{\mathbf{k}\mu}^{\text{QP}} \right) - \hat{V}_{\alpha\beta}^{xc} \right| \psi_{\mathbf{k}\mu}^{\beta} \right\rangle. \quad (3)$$

Here μ denotes the eigenvalue index, α and β denote the spin components of the ψ spinor, and \mathbf{k} labels the momentum in the first Brillouin zone. $\Sigma_{\alpha\beta} = iG_{\alpha\beta}W$ is the GW self-energy with off-diagonal spin elements,^{36,66,67} and $V_{\alpha\beta}^{xc}$ is the exchange-correlation operator in DFT. Note atomic units are used throughout the text.

Our approach is based on the formalism given in Refs. 66,67. We consider non-magnetic systems, for which $V_{\alpha\beta}^{xc} = V^{xc}\delta_{\alpha\beta}$, and the GW equations also simplify significantly. Our approach thus mirrors that implemented by Sakuma *et al.* in the SPEX code.³⁶

In this paper we present calculations using solely the Γ point in the Brillouin zone. These supercell calculations are ideal for large, inhomogeneous systems. Due to our focus on large systems, in this section we use the $\mathbf{k} \rightarrow 0$ limit, hence W does not couple states at different momenta. For Γ -point-only calculations, this limit does not entail a further approximation.

The expectation value of Σ is split, as usual, into two components, a frequency-independent exact exchange Σ^X and a correlation component $\Sigma^C(\omega)$. The former is given by

$$\hat{\Sigma}_{\alpha\beta}^X(\mathbf{r}, \mathbf{r}') = i \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_{\alpha\beta}^{\text{KS}}(\mathbf{r}, \mathbf{r}'; \omega) \nu(\mathbf{r}, \mathbf{r}'), \quad (4)$$

and

$$G_{\alpha\beta}^{\text{KS}}(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{\mathbf{k}\mu} \frac{\psi_{\mathbf{k}\mu}^{\alpha}(\mathbf{r}) \psi_{\mathbf{k}\mu}^{\beta*}(\mathbf{r}')}{\omega - \epsilon_{\mathbf{k}\mu}^{\text{KS}} + i\eta \operatorname{sgn}(\epsilon_{\mathbf{k}\mu}^{\text{KS}} - \epsilon_F)}. \quad (5)$$

Here ν is the bare Coulomb interaction,

$$\nu(\mathbf{r}, \mathbf{r}') = \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \quad (6)$$

where e is the electronic charge. The expectation value of the exchange part of the self-energy is then

$$\begin{aligned} E_{\mathbf{k}\mu}^X &= \sum_{\alpha\beta} \left\langle \psi_{\mathbf{k}\mu}^{\alpha} \left| \hat{\Sigma}_{\alpha\beta}^X \right| \psi_{\mathbf{k}\mu}^{\beta} \right\rangle \\ &= - \sum_{\substack{\mathbf{k}', \kappa \text{ in} \\ \text{occupied states}}} \iint d\mathbf{r} d\mathbf{r}' \tau_{\mathbf{k}\mu, \mathbf{k}'\kappa}(\mathbf{r}) \nu(\mathbf{r}, \mathbf{r}') \tau_{\mathbf{k}\mu, \mathbf{k}'\kappa}^*(\mathbf{r}') \end{aligned} \quad (7)$$

where we define τ as

$$\tau_{\mathbf{k}\mu, \mathbf{k}'\kappa}(\mathbf{r}) = \sum_{\alpha} \psi_{\mathbf{k}\mu}^{\alpha*}(\mathbf{r}) \psi_{\mathbf{k}'\kappa}^{\alpha}(\mathbf{r}). \quad (8)$$

Equation 8 is used to avoid applying operations to each spin component separately.

Writing the screened Coulomb interaction within RPA as

$$W_{\text{RPA}}(\mathbf{r}, \mathbf{r}'; \omega) = \nu(\mathbf{r}, \mathbf{r}') + W_p(\mathbf{r}, \mathbf{r}'; \omega), \quad (9)$$

the correlation self-energy is

$$\hat{\Sigma}_{\alpha\beta}^C(\mathbf{r}, \mathbf{r}'; \omega) = i \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} G_{\alpha\beta}^{\text{KS}}(\mathbf{r}, \mathbf{r}'; \omega + \omega') W_p(\mathbf{r}, \mathbf{r}'; \omega'). \quad (10)$$

Note W has no spin dependence.

G_0W_0 calculations were performed using the formalism proposed in Refs. 14,68,69. This method avoids both the explicit calculation of empty electronic states and the storage and

inversion of large dielectric matrices, unlike the Adler-Wiser formulation.^{70,71} The method relies on the spectral decomposition of the static dielectric matrix (ϵ) through the “projective dielectric eigenpotential” (PDEP) algorithm,^{72,73} which uses iterative diagonalization⁷⁴ and density functional perturbation theory⁷⁵ to solve the Sternheimer equation.⁷⁶ A symmetrization operation is used to transform the dielectric matrix into a Hermitian one.⁷⁷

The eigenvalues of the dielectric matrix decay rapidly to 1,⁷³ and hence only a relatively small number of eigenvectors and eigenvalues is required for an accurate spectral decomposition of ϵ . These eigenvectors are also used as a basis for the representation of ϵ and ϵ^{-1} at finite frequency in *GW* calculations.^{68,69} Dielectric eigenvectors are represented by the same plane-wave energy cutoffs as the DFT calculation, and both the PDEP algorithms and the *GW* calculations are implemented in the WEST massively parallel code.¹⁴

The spectral decomposition of ϵ allows both G and W to be treated as shifted-inverted problems, which are solved recursively using Lanczos chain techniques.⁷⁸ Frequency integration is carried out by contour deformation:^{14,20,23}

$$\begin{aligned}\hat{\Sigma}_{\alpha\beta}^C(\mathbf{r}, \mathbf{r}'; \omega) = & i \int_{-i\infty}^{i\infty} \frac{d\omega'}{2\pi} G_{\alpha\beta}^{\text{KS}}(\mathbf{r}, \mathbf{r}'; \omega + \omega') W_p(\mathbf{r}, \mathbf{r}'; \omega') \\ & - \sum_{z_{\mathbf{k}\mu}^G \in \Gamma^+} \psi_{\mathbf{k}\mu}^\alpha(\mathbf{r}) \psi_{\mathbf{k}\mu}^{\beta*}(\mathbf{r}') W_p(\mathbf{r}, \mathbf{r}'; z_{\mathbf{k}\mu}^G) \\ & + \sum_{z_{\mathbf{k}\mu}^G \in \Gamma^-} \psi_{\mathbf{k}\mu}^\alpha(\mathbf{r}) \psi_{\mathbf{k}\mu}^{\beta*}(\mathbf{r}') W_p(\mathbf{r}, \mathbf{r}'; z_{\mathbf{k}\mu}^G).\end{aligned}\quad (11)$$

The contours Γ^+ and Γ^- are given in Ref. 14. Note that here the poles z^G are energies of KS states, as they come from G^{KS} .

Finally, the contribution of Σ^C to the quasiparticle equation is given by

$$\begin{aligned}E_{\mathbf{k}\mu}^C(E_{\mathbf{k}\mu}^{\text{QP}}) = & \sum_{\alpha\beta} \left[- \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \left\langle \psi_{\mathbf{k}\mu}^\alpha \left| G_{\alpha\beta}^{\text{KS}}(\mathbf{r}, \mathbf{r}'; E_{\mathbf{k}\mu}^{\text{QP}} + i\omega') W_p(\mathbf{r}, \mathbf{r}'; i\omega') \right| \psi_{\mathbf{k}\mu}^\beta \right\rangle \right] \\ & + \sum_{\mathbf{k}'\kappa} \iint d\mathbf{r} d\mathbf{r}' f_{\mathbf{k}'\kappa}^{\mathbf{k}\mu} \tau_{\mathbf{k}\mu, \mathbf{k}'\kappa}(\mathbf{r}) W_p(\mathbf{r}, \mathbf{r}'; \epsilon_{\mathbf{k}'\kappa} - E_{\mathbf{k}\mu}^{\text{QP}}) \tau_{\mathbf{k}\mu, \mathbf{k}'\kappa}^*(\mathbf{r}')\end{aligned}\quad (12)$$

where

$$f_{\mathbf{k}'\kappa}^{\mathbf{k}\mu} = \begin{cases} +1 & \text{if } \epsilon_F < \epsilon_{\mathbf{k}'\kappa} < E_{\mathbf{k}\mu}^{\text{QP}} \\ -1 & \text{if } E_{\mathbf{k}\mu}^{\text{QP}} < \epsilon_{\mathbf{k}'\kappa} < \epsilon_F \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

Due to the use of Lanczos chains, it is computationally straightforward to calculate $\Sigma(\omega)$ at multiple ω . Usually one solves eq 3 by a secant method. In Section 8 we examine this procedure in detail.

Our calculations were performed using a development version of the WEST code, which is highly scalable on high-performance architectures.^{14,15} As mentioned in the introduction, our implementation is the first to both avoid the use of empty states, and calculate G and W at the FR level of theory. It also uses contour deformation^{14,20,23} rather than plasmon-pole^{12,13,16–21} or analytic continuation^{21–24} methods. The calculations performed here used computational resources ranging from 320 Ivy Bridge processors for small molecules to 262,144 cores on the BlueGene/Q Mira supercomputer at Argonne National Laboratory for the largest calculation of a 1152-electron CdTe supercell. We note that the inclusion of SOC considerably adds to the computational complexity of calculations; when using only the Γ point, the number of bands to be optimized is doubled and it is necessary to use two-component complex wave functions rather than single-component real wave functions. This implies an eight times increase in memory, and due to the overall $N_{\text{occ}}^2 \times N_{\text{PDEP}} \times N_{\text{pw}}$ scaling,¹⁴ we expect an asymptotic sixteen times increase in running time. Here N_{occ} is the number of occupied bands, N_{PDEP} the number of eigenvalues and eigenpotentials included in the spectral decomposition of the dielectric matrix, and N_{pw} the number of plane waves used for the expansion of single-particle wavefunctions. In the calculations reported below we observed an eight to ten times increase in running time.

3 Validation of implementation: vertical ionization potential of small molecules

We validated our implementation for a set of 81 molecules. The NIST Computational Chemistry Comparison and Benchmark Database⁷⁹ and the recent GW27⁸⁰ and GW100⁵⁵ test sets primarily contain systems with light elements, and hence were considered inappropriate for validation purposes. Instead, we used the NIST WebBook database:⁸¹ we selected molecules for which experimental vertical ionization potential (VIP) data and experimental geometries were available, and which were composed of elements with atomic number 29 (Cu) or higher (see Supporting Information (SI)). As detailed in Appendix A and the SI, we also excluded elements which would have had semicore d states belonging to an incomplete shell. This selection process resulted in 81 molecules, which we term here the GW-SOC81 set.

3.1 Computational details

Calculations at the SR and FR levels were performed both with the PBE GGA⁸² and the hybrid PBE0 functional.^{83–85} For all molecules we used experimental geometries. After a self-consistent calculation was performed in QE^{61,62} at the DFT level, a calculation of E^{QP} was carried out in WEST.^{14,15}

For SR calculations, version 1.1 of the SG15 Optimized Norm-Conserving Vanderbilt (ONCV) pseudopotentials were used,^{86,87} generated with the ONCVPSP code.^{88,89} The SG15 set⁸⁷ includes pseudopotentials and input files for all elements up to atomic number 83 (Bi), with the exception of the Lanthanides. Bromine and iodine pseudopotentials were created separately.⁹⁰ For FR pseudopotentials, the same input parameters as for the SR case were used (ONCVPSP version 2.1.1)⁸⁸ and are also available online.⁸⁷ Comparisons of the performance of these pseudopotentials with others reported in literature can be found in Secs. 4.5, 5 and 6.1, and in the SI.

We first determined the VIP of all molecules with a 120 Ry plane-wave kinetic energy

cutoff at the DFT level. We then used for subsequent calculations the cutoff at which the VIP was determined to be within 5 meV of the one obtained at the 120 Ry cutoff (see Table 1). Further comparisons of results obtained at different cutoffs are presented in Ref. 86. All energy levels were referred to vacuum using a Makov-Payne-based correction.⁹¹

We carried out converged SR *GW*@PBE calculations by choosing the number of eigenvectors in the spectral decomposition of the dielectric matrix¹⁴ as $N_{\text{PDEP}} = \max(512, 30n_{\text{elec}})$, and the cell size as $a = \max(40a_0, \bar{a})$, where a_0 is the Bohr radius and \bar{a} is the value that at the DFT level yields VIPs within 5 meV of those obtained with $a = 60a_0$.

The results of these converged calculations served as reference to estimate the errors made when using smaller but more tractable N_{PDEP} and a values. In particular we used $N_{\text{PDEP}} = \max(512, 7.9n_{\text{elec}} + 158)$, with the latter term used to interpolate between $N_{\text{PDEP}} = 20n_{\text{elec}}$ for 13 electrons and $N_{\text{PDEP}} = 10n_{\text{elec}}$ for 74 electrons. For the cell size we used $a = \max(30a_0, \bar{a}')$ where \bar{a}' yield VIPs within 10 meV of those obtained with $a = 60a_0$ at the DFT level of theory. At the SR *GW*@PBE level of theory, VIP results had a mean absolute error (MAE) of 11 meV, and a maximum deviation of 40 meV in $\text{Mo}(\text{CO})_6$, compared to the stricter parameters above. Table 1 provides all parameters used for the G_0W_0 results reported in Table 2.

Table 1: Molecules used in the GW-SOC81 set. The table includes a local GW-SOC81 set number, the linear size of the cubic cell used, the plane-wave energy cutoff for the wavefunctions ($E_{\text{cut}}^{\text{wfc}}$), and the number of eigenvalues used in the spectral decomposition of the dielectric matrix (N_{PDEP}).

Set number	Formula	Name	CAS Number	Cell size (a_0)	$E_{\text{cut}}^{\text{wfc}}$ (Ry)	N_{PDEP}
1	AgBr	Silver bromide	7785-23-1	40	40	364
2	AgCl	Silver chloride	7783-90-6	40	40	364
3	AgI	Silver iodide	7783-96-2	45	40	364
4	Al_2Br_6	Dialuminum hexabromide	18898-34-5	30	40	662
5	AlBr_3	Aluminum bromide	7727-15-3	30	40	410
6	AlI_3	Aluminum iodide	7784-23-8	30	40	410
7	AsBr_3	Arsenic tribromide	7784-33-0	30	40	364
8	AsCl_3	Arsenic trichloride	7784-34-1	30	40	364
9	AsF_3	Trifluoroarsine	7784-35-2	30	70	364
10	AsF_5	Pentafluoroarsorane	7784-36-3	30	70	474
11	AsH_3	Arsine	7784-42-1	30	50	256
12	AsI_3	Arsenic triiodide	7784-45-4	30	40	364
13	Br_2	Bromine	7726-95-6	30	40	268
14	BrCl	Bromine chloride	13863-41-7	30	40	268
15	$\text{C}_2\text{H}_2\text{Se}$	Selenoketene	61134-37-0	30	50	284
16	C_2HBrO	Bromoketene	78957-22-9	30	60	332
17	$\text{C}_4\text{H}_4\text{Se}$	Selenophene	288-05-1	30	50	364
18	$(\text{C}_5\text{H}_5)_2\text{Ru}$	Bis(5-cyclopentadienyl) Ruthenium	1287-13-4	30	60	678

19	CF ₃ I	Trifluoroiodomethane	2314-97-8	30	50	410
20	(CH ₃) ₂ Cd	Dimethylcadmium	506-82-1	30	50	426
21	(CH ₃) ₂ Hg	Dimethylmercury	593-74-8	30	50	426
22	(CH ₃) ₂ Se	Dimethylselenide	593-79-3	30	40	316
23	(CH ₃) ₂ Zn	Dimethylzinc	544-97-8	30	50	426
24	CH ₃ HgBr	Bromomethylmercury	506-83-2	40	40	426
25	CH ₃ HgCl	Chloromethylmercury	115-09-3	40	40	426
26	CH ₃ HgI	Iodomethylmercury	143-36-2	40	40	426
27	CH ₃ I	Iodomethane	74-88-4	30	40	268
28	Cl ₄	Tetraiodomethane	507-25-5	30	40	410
29	CaBr ₂	Calcium bromide	7789-41-5	35	40	348
30	CaI ₂	Calcium iodide	10102-68-8	35	40	348
31	CdBr ₂	Cadmium bromide	7789-42-6	30	40	426
32	CdCl ₂	Cadmium chloride	10108-64-2	30	40	426
33	CdI ₂	Cadmium iodide	7790-80-9	30	40	426
34	CsCl	Cesium chloride	7647-17-8	45	40	284
35	CsF	Cesium fluoride	13400-13-0	35	70	284
36	CsI	Cesium iodide	7789-17-5	50	40	284
37	CuF	Copper monofluoride	13478-41-6	30	80	364
38	HgCl ₂	Mercury(II) chloride	7487-94-7	30	40	426
39	I ₂	Iodine dimer	7553-56-2	30	40	268
40	IBr	Iodine bromide	7789-33-5	30	40	268
41	ICl	Iodine monochloride	7790-99-0	30	40	268
42	IF	Iodine monofluoride	13873-84-2	30	50	268
43	KBr	Potassium bromide	7758-02-3	50	40	284
44	KI	Potassium iodide	7681-11-0	50	40	284
45	Kr ₂	Krypton dimer	12596-40-6	30	40	284
46	KrF ₂	Krypton difluoride	13773-81-4	30	70	332
47	LaBr ₃	Lanthanum tribromide	13536-79-3	30	40	410
48	LaCl ₃	Lanthanum chloride	10099-58-8	30	40	410
49	LiBr	Lithium bromide	7550-35-8	35	40	256
50	LiI	Lithium iodide	10377-51-2	35	40	256
51	MgBr ₂	Magnesium bromide	7789-48-2	30	40	348
52	MgI ₂	Magnesium iodide	10377-58-9	30	40	348
53	Mo(CO) ₆	Molybdenum hexacarbonyl	13939-06-5	30	60	740
54	NaBr	Sodium bromide	7647-15-6	45	40	284
55	NaI	Sodium iodide	7681-82-5	45	40	284
56	OsO ₄	Osmium tetroxide	20816-12-0	30	70	474
57	PBr ₃	Phosphorus tribromide	7789-60-8	30	40	364
58	POBr ₃	Phosphoryl tribromide	7789-59-5	30	50	410
59	RbBr	Rubidium bromide	7789-39-1	50	40	284
60	RbCl	Rubidium chloride	7791-11-9	50	40	284
61	RbI	Rubidium iodide	7790-29-6	50	40	284
62	RuO ₄	Ruthenium tetroxide	20427-56-9	30	70	474
63	SOBr ₂	Thionyl dibromide	507-16-4	30	50	364
64	SPBr ₃	Thiophosphoryl tribromide	3931-89-3	30	40	410
65	SeCl ₂	Selenium dichloride	14457-70-6	30	40	316
66	SeO ₂	Selenium(IV) oxide	7446-08-4	30	50	300
67	SiBrF ₃	Bromotrifluorosilane	14049-39-9	30	50	410
68	SiH ₃ I	Iodosilane	13598-42-0	30	40	268
69	SrBr ₂	Strontium bromide	10476-81-0	40	40	348
70	SrCl ₂	Strontium chloride	10476-85-4	35	40	348
71	SrI ₂	Strontium iodide	10476-86-5	40	40	348
72	SrO	Strontium oxide	1314-11-0	45	60	284
73	TiBr ₄	Titanium tetrabromide	7789-68-6	30	40	474
74	TiI ₄	Titanium tetraiodide	7720-83-4	30	40	474
75	ZnBr ₂	Zinc bromide	7699-45-8	30	50	426
76	ZnCl ₂	Zinc chloride	7646-85-7	30	60	426
77	ZnF ₂	Zinc fluoride	7783-49-5	30	70	426
78	ZnI ₂	Zinc iodide	10139-47-6	30	50	426
79	ZrBr ₄	Zirconium tetrabromide	13777-25-8	30	40	474
80	ZrCl ₄	Zirconium tetrachloride	10026-11-6	30	40	474
81	ZrI ₄	Zirconium tetraiodide	13986-26-0	30	40	474

3.2 The GW-SOC81 set

Results of the VIP for the molecules of the GW-SOC81 set are displayed in Table 2. All experimental values were taken from the NIST WebBook.⁸¹ We found that starting from results obtained with the PBE0 functional provides a significant improvement over the use of the PBE functional, which regularly underestimates the experimental VIP. These results are consistent with those reported in Ref. 14 using SR methods.

At the $GW@PBE$ level, FR calculations yield lower VIPs than SR calculations, and in worse agreement with experiment. However, at the $GW@PBE0$ level, the performance of FR and SR methods is essentially identical (see Fig. 1).

In order to quantify the errors more thoroughly, we considered FR and SR methods with a constant offset applied to the VIP of all molecules, chosen to minimize the MAE of each method, respectively. Such results are shown in Table 3. The shifted FR $GW@PBE0$ results show a noticeable decrease in MAE, while the improvement in SR $GW@PBE0$ results is much smaller.

Table 2: Vertical ionization potential (VIP, in eV) of 81 molecules containing heavy elements ($Z \geq 29$) compared to experiment. We report results obtained with scalar-relativistic (SR) and fully-relativistic (FR) calculations, and using Kohn-Sham wave functions computed with the PBE and hybrid PBE0 functionals. In all calculations we used experimental geometries. Summarized at the end of the table are the mean error (ME), standard deviation of the errors (SE), mean absolute error (MAE), and mean absolute relative error (MARE) with respect to experiment for each method.

Number	Molecule	SR $GW@PBE$	FR $GW@PBE$	SR $GW@PBE0$	FR $GW@PBE0$	Experiment
1	AgBr	9.03	8.90	9.53	9.38	9.59
2	AgCl	9.35	9.31	10.03	9.98	10.08
3	AgI	8.60	8.44	8.96	8.66	8.80
4	Al ₂ Br ₆	10.38	10.34	10.78	10.74	10.97
5	AlBr ₃	10.53	10.48	10.91	10.86	10.91
6	AlI ₃	9.44	9.23	9.78	9.57	9.66
7	AsBr ₃	9.83	9.77	10.17	10.10	10.21
8	AsCl ₃	10.67	10.66	11.00	10.99	10.90
9	AsF ₃	12.49	12.49	12.89	12.89	13.00
10	AsF ₅	14.51	14.49	15.28	15.26	15.53
11	AsH ₃	10.33	10.33	10.55	10.54	10.58
12	AsI ₃	8.99	8.72	9.39	9.11	9.00
13	Br ₂	10.33	10.16	10.59	10.42	10.54
14	BrCl	10.79	10.67	11.06	10.93	11.01
15	C ₂ H ₂ Se	8.48	8.48	8.72	8.72	8.71
16	C ₂ HBrO	8.98	8.97	9.28	9.27	9.10
17	C ₄ H ₄ Se	8.60	8.60	8.86	8.86	8.86
18	(C ₅ H ₅) ₂ Ru	7.00	6.90	7.31		7.45
19	CF ₃ I	10.52	10.20	10.81	10.48	10.45
20	(CH ₃) ₂ Cd	8.82	8.83	9.09	9.10	8.80

21	(CH ₃) ₂ Hg	8.99	9.07	9.20	9.28	9.32
22	(CH ₃) ₂ Se	8.17	8.12	8.41	8.41	8.40
23	(CH ₃) ₂ Zn	9.38	9.38	9.65	9.65	9.40
24	CH ₃ HgBr	9.72	9.59	10.11	9.97	10.16
25	CH ₃ HgCl	10.30	10.26	10.76	10.72	10.84
26	CH ₃ HgI	9.08	8.79	9.38	9.09	9.25
27	CH ₃ I	9.57	9.26	9.78	9.46	9.52
28	Cl ₄	8.91	8.76	9.28	9.04	9.10
29	CaBr ₂	9.80	9.67	10.24	10.10	10.35
30	CaI ₂	9.09	8.79	9.48	9.18	9.39
31	CdBr ₂	10.21	10.06	10.62	10.46	10.58
32	CdCl ₂	10.91	10.87	11.39	11.34	11.44
33	CdI ₂	9.41	9.09	9.76	9.43	9.57
34	CsCl	7.87	7.84	8.51	8.47	8.69
35	CsF	8.24	8.24	9.08	9.08	9.68
36	CsI	7.01	6.64	7.49	7.20	7.46
37	CuF	10.23	10.17	10.03	9.99	10.90
38	HgCl ₂	10.93	10.85	11.35	11.28	11.50
39	I ₂	9.41	9.01	9.64	9.26	9.35
40	IBr	9.81	9.51	10.04	9.73	9.85
41	ICl	10.14	9.85	10.37	10.07	10.10
42	IF	10.55	10.23	10.80	10.47	10.62
43	KBr	7.56	7.37	8.21	8.07	8.82
44	KI	7.12	7.34	7.67	7.38	7.40
45	Kr ₂	13.42	13.27	13.68	13.53	13.77
46	KrF ₂	12.58	12.51	13.30	13.22	13.34
47	LaBr ₃	9.90	9.82	10.41	10.31	10.68
48	LaCl ₃	10.73	10.72	11.26	11.24	11.29
49	LiBr	8.95	8.81	9.35	9.21	9.44
50	LiI	8.35	8.04	8.65	8.36	8.44
51	MgBr ₂	10.49	10.35	10.91	10.76	10.85
52	MgI ₂	9.62	9.31	9.97	9.65	10.50
53	Mo(CO) ₆	8.55	8.50	8.78		8.50
54	NaBr	8.06	7.89	8.67	8.54	8.70
55	NaI	7.65	7.75	8.12	7.82	8.00
56	OsO ₄	11.74	11.74	12.41	12.41	12.35
57	PBr ₃	9.60	9.57	9.92	9.88	10.00
58	POBr ₃	10.55	10.49	11.01	10.93	11.03
59	RbBr	7.41	7.26	8.09	7.95	8.62
60	RbCl	7.75	7.72	8.54	8.50	8.70
61	RbI	6.97	7.19	7.55	7.26	7.30
62	RuO ₄	11.45	11.44	12.19	12.18	12.15
63	SOBr ₂	10.17	10.13	10.57	10.52	10.59
64	SPBr ₃	9.45	9.43	9.82	9.79	9.89
65	SeCl ₂	9.24	9.24	9.53	9.53	9.50
66	SeO ₂	11.03	11.03	11.61	11.60	11.76
67	SiBrF ₃	11.78	11.64	12.10	11.95	12.46
68	SiH ₃ I	9.93	9.64	10.17	9.86	9.82
69	SrBr ₂	9.49	9.35	9.92	9.77	9.82
70	SrCl ₂	10.00	9.97	10.49	10.46	10.20
71	SrI ₂	8.82	8.51	9.21	8.90	9.01
72	SrO	6.26	6.26	6.01	6.01	6.91
73	TiBr ₄	9.98	9.89	10.57	10.47	10.57
74	TiI ₄	8.90	8.65	9.42	9.17	9.27
75	ZnBr ₂	10.52	10.37	10.90	10.75	10.89
76	ZnCl ₂	11.36	11.32	11.79	11.75	11.80
77	ZnF ₂	12.69	12.66	13.42	13.39	13.91
78	ZnI ₂	9.63	9.30	9.96	9.63	9.73
79	ZrBr ₄	10.26	10.17	10.78	10.68	10.86
80	ZrCl ₄	11.35	11.32	11.93	11.91	11.94
81	ZrI ₄	9.17	8.93	9.62	9.36	9.55
<hr/>						
	ME	-0.421	-0.531	-0.014	-0.149	
	SE	0.321	0.306	0.253	0.227	
	MAE	0.431	0.532	0.176	0.184	
	MARE	4.25%	5.27%	1.84%	1.87%	

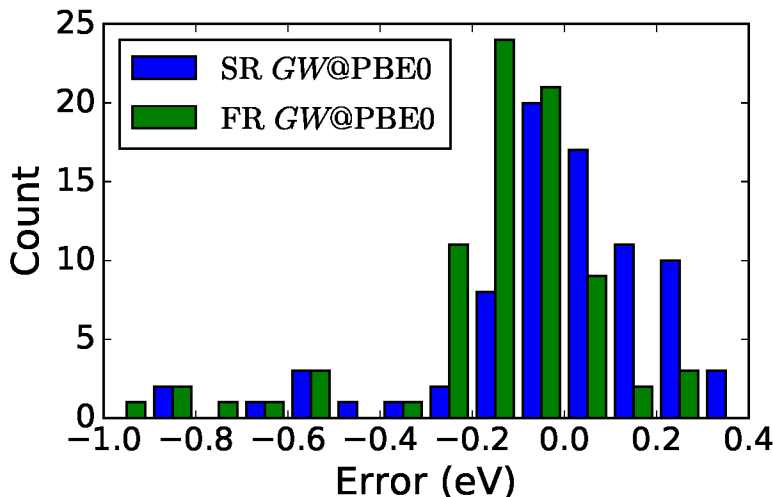


Figure 1: Histogram of the errors in calculated VIPs vs. experiment for the $GW@PBE0$ calculations reported in Table 2. Bins go from $n/10$ to $(n + 1)/10$ for $n \in \mathbb{Z}$.

Table 3: Summary of performance for SR and FR methods when a constant shift, chosen to minimize the calculated MAE, is applied to all VIP values computed using each method. The acronyms for statistics and methods match those in Table 2. Aside from MARE, values are in eV.

Quantity	SR $GW@PBE$	FR $GW@PBE$	SR $GW@PBE0$	FR $GW@PBE0$
VIP Shift	-0.376	-0.499	0.005	-0.106
ME	-0.045	-0.032	-0.019	-0.042
SE	0.321	0.306	0.253	0.227
MAE	0.243	0.226	0.175	0.139
MARE	2.43%	2.33%	1.83%	1.44%

In order to ensure our results are statistically significant, rather than obtained by chance due to the finite sample of molecules used, we performed statistical significance tests for the MAE. We used nonparametric tests, because absolute errors (AEs) are bounded below by zero and thus cannot be normally distributed. For the results described below, we considered the traditional $p \leq 0.05$ as the maximum threshold of significance,⁹² noting that much smaller p values are preferred.⁹³

First, we performed the Friedman test, which takes as a null hypothesis that the median AEs are equal across all computational methods used.^{94,95} It then uses rank ordering of the AEs of the different methods for each molecule (or block in statistical terms). As expected

given the large data set and large differences, in this case and in other multiple-set comparisons shown below (unless otherwise noted), the Friedman test gives $p < 10^{-5}$, confirming there are distinctions in the MAEs of the sets.

To compare individual methods, we performed pairwise Wilcoxon Signed-Rank Tests to determine which median AEs in the set show differences.^{95,96} This test uses rank ordering, but in contrast to the Friedman test it also incorporates for each molecule the magnitude of the difference in AE between the methods. The resulting p values, and associated confidence intervals for differences in the median AEs, are shown in Table 4. Due to the multiple hypotheses tested, we adjusted the p -values produced by the Wilcoxon test using the Holm (often termed Holm-Bonferroni) method.^{97,98} Based on this analysis, for the absolute errors we found all pairwise comparisons show clearly significant differences, except for FR $GW@PBE0$ vs. SR $GW@PBE0$, for which there is insufficient evidence to say either is better.

Table 4: Statistical analysis of the results obtained with different methods for the GW-SOC81 set of Table 2. The acronyms used for the methods are the same as in Table 2. For each pair of methods, the first entry of the table gives the adjusted p -value for the significance of the difference in median absolute errors (AEs). The original p -values were arranged from smallest to largest, denoted as $p_{(1)}, \dots, p_{(n)}$, and then multiplied by factors $C_{(i)} = (n, n-1, \dots, 1)$ to give $p'_{(1)}, \dots, p'_{(n)}$. If $p'_{(j)} > p'_{(k)}$ for any $j < k$, we set $p'_{(k)} = p'_{(j)}$ to maintain the ordering, and p -values are set to a maximum of 1.0 (Holm-Bonferroni method). For each pair of methods, the second entry in parentheses gives the estimated median values for the row AE minus column AE (eV) and corresponding 95% confidence intervals based on the Wilcoxon test.^{95,99} These are not adjusted based on the number of tests, in contrast to the first entry.

	FR $GW@PBE$	SR $GW@PBE0$	FR $GW@PBE0$
SR $GW@PBE$	2×10^{-10} (-0.10 ± 0.03)	1×10^{-8} (0.27 ± 0.08)	2×10^{-10} (0.25 ± 0.06)
FR $GW@PBE$		2×10^{-11} (0.38 ± 0.07)	1×10^{-12} (0.36 ± 0.05)
SR $GW@PBE0$			0.06 (-0.02 ± 0.02)

The statistical significance results for the shifted values in Table 3 are presented in Table 5, and interesting changes are present compared to Table 4. All $GW@PBE0$ methods continue to show improvement over $GW@PBE$ methods, but FR $GW@PBE$ is as good,

and possibly better (but with no statistical significance) compared to SR $GW@PBE$. Most notably, however, FR $GW@PBE0$ shows a statistically significant improvement over SR $GW@PBE0$.

Based on these analyses, it is clear that $GW@PBE0$ offers significant improvements over $GW@PBE$. For FR vs. SR methods, conclusions are less straightforward: The two are statistically similar in terms of accuracy, but FR $GW@PBE0$ does show a statistically significant improvement in the precision of its calculations—in other words, it captures the variations of VIP between molecules more accurately, but systematically underestimates them by about 0.15 eV. Finally, given the generally very good agreements of these results with experiments, and the availability of coordinates in the SI, the GW-SOC81 set can now serve to validate future implementations of SOC in electronic structure methods.

Table 5: Statistical analysis of the results obtained with different methods for the GW-SOC81 set of Table 3, for which VIPs were shifted by a constant for each method in order to minimize MAE. This helps establish the precision of each method. The presentation is identical to that described in Table 4.

	FR $GW@PBE$	SR $GW@PBE0$	FR $GW@PBE0$
SR $GW@PBE$	0.12 (0.03 \pm 0.03)	6×10^{-5} (0.07 \pm 0.03)	2×10^{-5} (0.10 \pm 0.04)
FR $GW@PBE$		0.005 (0.06 \pm 0.04)	2×10^{-7} (0.09 \pm 0.03)
SR $GW@PBE0$			0.01 (0.03 \pm 0.02)

4 Bulk semiconductor calculations

In order to further quantify the performance of FR G_0W_0 calculations, we next considered semiconducting solids incorporating heavy elements, and we computed the splitting of the valence band maximum (VBM) and the electronic gap (as the difference between VBM and conduction band minimum (CBM)). We found that in the case of valence band splittings FR GW at least matches, and may exceed, the performance of FR DFT. For band gaps, FR GW significantly improves upon FR DFT, but any definitive conclusions are complicated by

systematic errors discussed below.

4.1 Computational details

Computations were performed using the same sequence of calculations as in Section 3, and with the same pseudopotentials and energy cutoff; here we only used the PBE functional. The lattice constant was optimized within $\pm 0.005 \text{ \AA}$, using a conventional 8-atom cell and a $4 \times 4 \times 4$ unshifted Monkhorst-Pack¹⁰⁰ (MP) grid of k -points with an FR calculation.

For the G_0W_0 calculations, supercells were used with 64 atoms and sampling only the Γ point. See the SI for notes on convergence of this supercell size. Aside from silicon (256 valence electrons), the number of valence electrons ranged from 576 (e.g. GaAs) to 1152 (CdTe and ZnTe). We chose $N_{\text{PDEP}} = 2048$ for all computations; an SR $GW@PBE$ test with GaAs for $N_{\text{PDEP}} = 4096$ gave a 47 meV change in the band gap.

Semiconductors were selected based on available results in Table 3.3 of Ref. 101 for the valence band splitting. Of those, C was excluded due to its very small splitting. For simplicity we restricted our investigation to semiconductors with diamond or zincblende structures, hence GaN and ZnO were excluded. Finally, GaSb, Ge, InAs, and InSb were not considered due to the lack of a semiconducting gap in the DFT calculation at the FR PBE level.

4.2 Splitting of the highest-energy valence band

In the materials studied here, the valence band maximum is a p -bonding state. When spin-orbit coupling is included, the state splits into $j = 1/2$ and $j = 3/2$ states, whose energies differ by a gap that we call Δ_{so} (see Figure 2).¹⁰¹ By definition, in SR calculations $\Delta_{\text{so}} = 0$.

In our GW calculations, Δ_{so} is reproduced rather accurately, with results shown in Figure 3 and Table 6. Compared to DFT calculations; the splittings for AlSb, CdS, and ZnSe improve substantially. Overall, the MAE of our GW calculations decreases by 14 meV compared to DFT. This may suggest an improvement over DFT, however, the statistical evidence

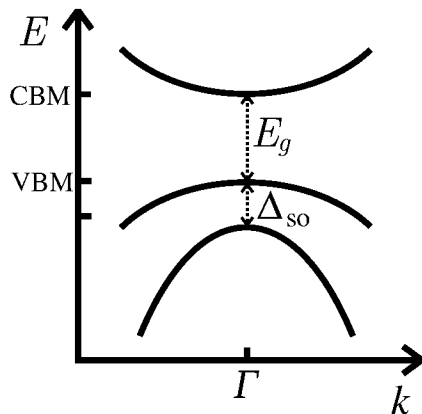


Figure 2: Schematic representation of the bands in the bulk semiconductors studied here, and the associated energies which are measured or calculated. The band gap E_g gives the energy difference between the conduction band minimum (CBM) and valence band maximum (VBM). The splitting Δ_{so} measures the energy difference at the Γ point between the $j = 3/2$ and $j = 1/2$ valence states.

for such a difference is weak: the Wilcoxon signed-rank test leads to an estimate of the difference between median AEs with 95% confidence interval of $0.009^{+0.028}_{-0.009}$ eV, corresponding to $p = 0.06$ for the null hypothesis. In any case, we can say with confidence that GW calculations at least match the accuracy of DFT. Furthermore, the lack of improvement in the two outliers at the GW level—CdTe and ZnTe—could be due to the presence of semicore d electrons in the Te pseudopotential valence partition, as explored in Appendix A.

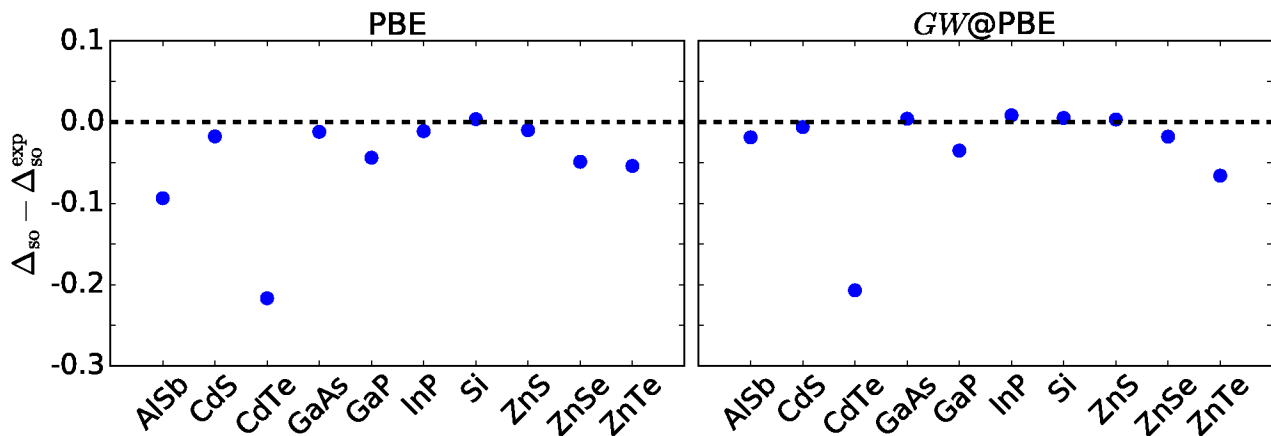


Figure 3: Deviation of the valence band splitting Δ_{so} from the corresponding experimental value Δ_{so}^{exp} . Results at the PBE (left panel) and $GW@PBE$ (right panel) levels of theory are shown.

4.3 Calculation of the electronic band gap

Spin-orbit coupling also affects the magnitude of the band gap, E_g . In all of the semiconductors investigated in this section, SOC led to a decrease in the band gap of 0 to 0.4 eV. As Table 6 shows, these changes are quite similar between DFT and GW calculations. On average the magnitude of the change is larger at the GW level by 0.019 eV, but even this difference has little significance with $p = 0.08$ for the Wilcoxon signed rank test.

As expected, the absolute value of GW -level band gaps are significantly larger than DFT band gaps. However, further conclusions on method performance are hampered by a significant dependence of results on the choice of pseudopotentials, especially the presence of d electrons in the valence partition. These issues are explored further in Appendix A.2, and they are not specific to relativistic effects analyzed in this paper.

Table 6: Calculated and experimental values, and mean errors determined with different methods for the valence band splitting Δ_{so} , as well as differences caused by SOC in the electronic band gap E_g . $\Delta_{\text{so}} = 0$ for all SR calculations.

	Δ_{so}			$E_g^{\text{FR}} - E_g^{\text{SR}}$	
	FR DFT	FR GW	Exp. ¹⁰¹	DFT	GW
AlSb	0.657	0.731	0.75	-0.202	-0.242
CdS	0.049	0.060	0.07	-0.017	-0.022
CdTe	0.703	0.713	0.92	-0.278	-0.314
GaAs	0.328	0.344	0.34	-0.136	-0.123
GaP	0.083	0.092	0.13	-0.028	-0.031
InP	0.099	0.119	0.11	-0.031	-0.037
Si	0.048	0.049	0.04	-0.016	-0.017
ZnS	0.060	0.073	0.07	-0.021	-0.022
ZnSe	0.381	0.412	0.43	-0.128	-0.148
ZnTe	0.876	0.864	0.93	-0.289	-0.380
ME (eV)	-0.050	-0.033			
MAE (eV)	0.051	0.037			
MARE (%)	15.0%	9.8%			

4.4 Choice of lattice constant

In our calculations, the optimized lattice constants at the FR PBE level of theory were used. The absolute value of the electronic band gap can depend heavily on the choice of the lattice constant. This effect was checked at the SR *GW* level of theory. We found a band gap increase up to 0.36 eV when using the experimental lattice constant (average increase of 0.21 eV). The SR DFT increases are smaller but similar, with an average of 0.13 eV. Further analysis of absolute band gap values is presented in Appendix A.2.

Differences due to SOC effects reported here have been checked for lattice constants computed at the DFT level, and appear to be quite small—for Δ_{so} , values shift by -0.005 ± 0.039 eV; for $E_g^{\text{FR}} - E_g^{\text{SR}}$ values shift by 0.007 ± 0.011 where errors are one standard deviation of the distribution of changes. Shifts of this magnitude at the *GW* level would not affect the conclusions we draw above regarding comparisons of Δ_{so} results to experiment.

4.5 Band structure of HgS

Some semiconducting compounds have large SOC shifts that lead to inverted band structures, which is of relevance to topological insulators.¹⁰ Here we study one of these materials, zincblende β -HgS, in order to compare with literature FR DFT and *GW* results³⁶ obtained for an inverted band structure solid.¹⁰² While the identity of HgS as a semimetal or semiconductor is still debated,^{36,103–107} at the DFT level SOC effects consistently open a gap between inverted bands at Γ .^{36,104} In our calculations we used a 60 Ry plane-wave energy cutoff, a 64-atom supercell with Γ -only sampling, and $N_{\text{PDEP}} = 2048$. For PBE calculations we used the SG15 pseudopotential set;^{86,87} for LDA calculations we used the SG15 inputs to the ONCV PSP code,^{88,89} but with the LDA functional used for the actual pseudopotential generation. We checked that when using a 64-atom supercell with a $4 \times 4 \times 4$ unshifted MP grid¹⁰⁰ of k -points, the DFT band separations reported here changed by only 4 meV. The experimental lattice constant was used.³⁶

Reference 36 finds that at the DFT level the states are ordered $\Gamma_6 - \Gamma_8 - \Gamma_7$ from lowest to

highest energy; we find the same ordering in our calculations along with the same symmetry character of the bands (Γ_8 is a fourfold degenerate $p_{3/2}$ band, Γ_7 a twofold degenerate $p_{1/2}$ band, Γ_6 an s band). Results for separations between the gaps, and comparisons with Ref. 36, are shown in Table 7. Both LDA and $GW@LDA$ results are in excellent agreement with those of Ref. 36 except for the GW correction to the $(\Gamma_8 - \Gamma_7)$ energy difference. This 0.09 eV difference could be due e.g. to the use of pseudopotentials, or to differences in how SOC is treated at the DFT level. $GW@PBE$ results are fairly similar to $GW@LDA$, but exhibit a larger magnitude of the inverse gap $(\Gamma_6 - \Gamma_8)$.

Table 7: Values for splittings (eV) between given bands of β -HgS. All calculations here include SOC.

Calculation	$\Gamma_6 - \Gamma_8$	$\Gamma_8 - \Gamma_7$
PBE	-0.42	-0.11
$GW@PBE$	-0.10	-0.06
LDA	-0.64	-0.11
$GW@LDA$	-0.02	-0.10
Ref. 36 LDA	-0.66	-0.12
Ref. 36 $GW@LDA$	-0.02	-0.19

5 Band gap of $MAPbI_3$ perovskite

As another example of an important application of GW methods incorporating spin-orbit coupling, we computed the quasiparticle energy levels of the perovskite $CH_3NH_3PbI_3$, illustrated in Figure 4. This perovskite has been found to be promising for solar energy applications,^{108,109} and has drawn significant attention from the computational community, where several investigations of alternative compositions, the roles of the organic cations, and the interplay of spin-orbit coupling, crystal structures, and hydrogen bonding have been reported.^{48–53,110–113} Given the complexity of this system, it was surprising to find that the band gap was well-determined by SR DFT calculations.⁴⁸ Further investigation, however, revealed that this is primarily due to the cancellation of two very large (order 1 eV) unrelated

errors in the semilocal DFT calculation: the neglect of the spin-orbit interaction, as well as the lack of quasiparticle or other corrections to the band gap.^{48,51,110} The fully-relativistic G_0W_0 method we present here accounts for both of these effects.

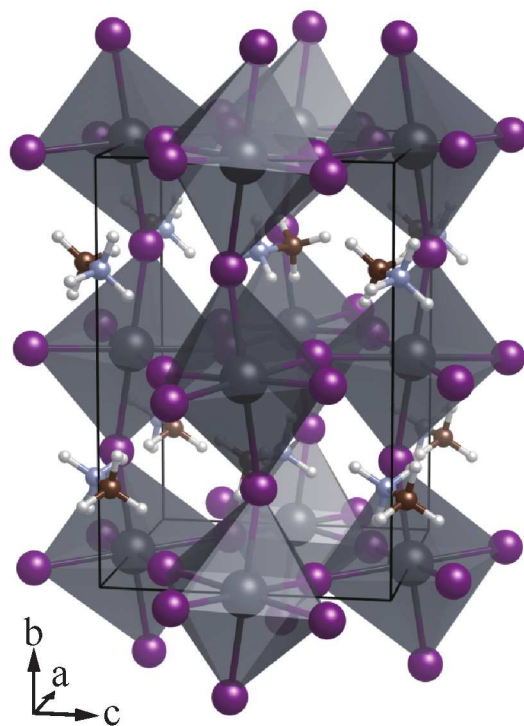


Figure 4: Perspective image of the perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$, with elements Pb (dark gray), I (purple), C (brown), N (light blue) and H (white). The unit cell and lattice vectors a , b , and c are shown, as well as the PbI_6 octahedra.

We again used the SG15 ONCV pseudopotential sets described in Section 3.1, including the original I^{17+} SG15 pseudopotential for iodine, with a 60 Ry plane-wave energy cutoff. We employed the structure given in Ref. 114, with added hydrogens in the positions they occupy in isolated methylammonium molecules and with arbitrary rotations. We used a 48-atom orthorhombic unit cell, and sampled only the Γ point. We employed $N_{\text{PDEP}} = 2048$. Convergence with respect to N_{PDEP} was checked by performing a computation with $N_{\text{PDEP}} = 1024$; the gap changed by only 0.01 eV, highlighting the straightforward and robust convergence of the GW method used here.¹⁴

The results reported in Table 8 show the incidental agreement between the SR DFT and FR *GW* calculations, with only a 0.03 eV difference. The agreement with experiment is good, with 1.51 eV for the computed gap vs. the experimental optical band gaps ranging from 1.51 to 1.64 eV;^{48,51,114–118} note that an exciton binding energy of approximately 0.05 eV^{117,119} should be added to the measured optical gap in order to compare with the quasiparticle gaps in Table 8.

Our results fall in the range of those of previous calculations^{48,51,110} (also shown in Table 8), with the SR *GW* gap slightly higher and the FR DFT gap slightly lower than the results reported in the literature. Some differences are expected due to the fact that we used explicit frequency integration rather than a plasmon-pole model,¹⁴ included FR effects in both *G* and *W*, and avoided sums over empty states. Remaining differences may also be due to the unoptimized crystal structure used here, and the lack of *k*-points at the DFT level (SR DFT with a $4 \times 4 \times 4$ unshifted MP grid increased the gap by 0.10 eV, to 1.58 eV, with respect to Γ -only calculations).

Table 8: Computed band gap (eV) of solid $\text{CH}_3\text{NH}_3\text{PbI}_3$, compared to the results of previous work. Values come from (a) Ref. 51, (b) Ref. 48, (c) Ref. 110 [note that the LDA functional is used in (a) and (c) and PBE in (b)]. See also Table 9 for a crucial dependence of these numbers on pseudopotential choices.

Method	This work	Previous work
SR DFT	1.49	1.50 ^a , 1.68 ^b , 1.46 ^c
FR DFT	0.44	0.58 ^a , 0.60 ^b , 0.53 ^c
SR G_0W_0	2.84	2.55 ^a , 2.68 ^b , 2.73 ^c
FR G_0W_0	1.51	1.20 ^a , 1.32 ^a , 1.67 ^b , 1.27 ^c

Due to significant effects from semicore d electrons found in other calculations carried out in the present work (see Appendix A), we also investigated the role of pseudopotentials and that of valence electron configurations in the case of perovskites. Table 9 shows our results, and in Appendix A we give a description of iodine pseudopotentials. Goedecker-Teter-Hutter (GTH) pseudopotentials^{120,121} were taken from the library generated by Krack.^{122,123} SG15-

based LDA pseudopotentials used the SG15 inputs into the ONCVSP code,⁸⁸ but with the LDA functional used for the pseudopotential generation as in Section 4.5. We used plane-wave energy cutoffs of 60 Ry, except for calculations with I^{25+} , for which we increased this cutoff to 100 Ry yielding converged band gaps within 6 meV at the DFT level.

Our results are consistent with those of Ref. 51, with at most a few tenths of an eV difference in absolute values and all of the same trends present. In agreement with Ref. 51, at the FR *GW* level using the I^{17+} pseudopotential configuration yields the best agreement with experiment for the MAPbI_3 gap. Unfortunately, our work indicates that this result is due to the incomplete semicore electron shell included in the valence; using an I^{25+} configuration gives results very close to that of the I^{7+} configuration. Thus, the I^{17+} pseudopotential appears to yield good agreement with experiments for unphysical reasons. This calls for further investigations, beyond the scope of this work, into the band gaps of solar perovskites: both in understanding how to accurately reproduce the measured experimental gap while carefully controlling the role of the pseudopotential in the calculation, and in understanding why different calculations using the iodine 5s5p pseudopotential give significantly different results in Ref. 48 compared to those in Ref. 51 and here. In addition, hybrid functionals may be necessary to provide wavefunctions that are a sufficiently accurate starting point for G_0W_0 calculations.

6 Pb_4S_4 and $\text{Pb}_{14}\text{Se}_{13}$ nanocrystals

We now turn to semiconducting nanocrystals, intriguing hybrids between bulk semiconductors and molecular systems, with properties, especially the band gap, that are heavily size-dependent. This tunability has made them promising candidates for e.g. photovoltaic applications.^{124–126} While many calculations have been done on nanocrystals, most are at the DFT level,¹²⁵ and higher-accuracy *GW* calculations are desirable. A separate technique, ΔSCF , can give relatively accurate results for the electronic gap of finite systems,⁴⁷ but it

Table 9: Computed band gap (eV) of solid $\text{CH}_3\text{NH}_3\text{PbI}_3$ as determined with SR and FR calculations, for different pseudopotential configurations as described in the text.

Pseudo. Set	Functional	I config.	Pb config.	DFT E_g	GW E_g
SR					
ONCV (I^{25+})	PBE	4s4p4d5s5p	5d6s6p	1.50	2.27
SG15 (I^{17+})	PBE	4d5s5p	5d6s6p	1.49	2.84
SG15 (I^{7+})	PBE	5s5p	5d6s6p	1.48	2.26
GTH	PBE	5s5p	5d6s6p	1.46	2.17
GTH	PBE	5s5p	6s6p	1.44	2.22
SG15-based (I^{17+})	LDA	4d5s5p	5d6s6p	1.36	2.86
SG15-based (I^{7+})	LDA	5s5p	5d6s6p	1.37	2.39
GTH	LDA	5s5p	6s6p	1.32	2.18
Ref. 51	LDA	4d5s5p	5d6s6p	1.50	2.55
Ref. 51	LDA	5s5p	5d6s6p	1.42	2.16
FR					
ONCV (I^{25+})	PBE	4s4p4d5s5p	5d6s6p	0.46	0.85
SG15 (I^{17+})	PBE	4d5s5p	5d6s6p	0.44	1.51
SG15 (I^{7+})	PBE	5s5p	5d6s6p	0.43	0.87
SG15-based (I^{17+})	LDA	4d5s5p	5d6s6p	0.32	1.41
SG15-based (I^{7+})	LDA	5s5p	5d6s6p	0.32	0.98
Ref. 51	LDA	4d5s5p	5d6s6p	0.58	1.32
Ref. 51	LDA	5s5p	5d6s6p	0.52	0.98

cannot easily provide full quasiparticle spectra in contrast to GW , and it involves calculations for both neutral and charged systems.

Here we performed FR G_0W_0 calculations for two sample systems, both lead chalcogenides: Pb_4S_4 , for which our results can be compared to existing ones,⁴⁷ and $\text{Pb}_{14}\text{Se}_{13}$, which to our knowledge has not been investigated at the GW level. The non-stoichiometric structure of the latter system has been found to lead to “trap” states that change the character of the HOMO and/or LUMO levels compared to stoichiometric nanocrystals;^{125–128} non-stoichiometric ones have also been observed experimentally.¹²⁹ The geometries of the two lead chalcogenide clusters considered here are nearly cubic with (100) surfaces, which have been found to be the most stable surfaces for these materials.^{130–132}

6.1 Pb_4S_4

Pb_4S_4 is the smallest nanocrystal that displays the bulk PbS rock salt structure, and it has been shown to be exceptionally stable (a “magic-number” behavior).¹²⁵ Recent work reported GW and ΔSCF calculations for this nanocrystal,⁴⁷ which are compared to our results in Table 10. We used the same ONCV pseudopotentials described in Section 3.1 with an 85 Ry cutoff, and a $(20 \text{ \AA})^3$ unit cell. Our DFT results were compared with those obtained using PAW pseudopotentials from pslibrary 1.0.0,¹³³ (and the recommended kinetic energy cutoffs contained with that library) showing deviations of at most 0.04 eV in the gap. The structure, available in the SI, was generated by taking a Pb_4S_4 cubic unit from bulk PbS, randomly perturbing atoms to break symmetries, and then performing an FR DFT relaxation. We used $N_{\text{PDEP}} = 256$.

Table 10 shows good agreement between our results and those of Ref. 47, especially given the different pseudopotentials and details of the GW computations. We note that both the spin-orbit (~ 0.9 eV) and the GW correction (~ 3.5 eV) to SR DFT are quite large, but in contrast to the accidental cancellation of the two terms in the MAPbI_3 perovskite, these corrections are different in magnitude for PbS clusters. In addition, we found that SOC

mostly affects the energy of the LUMO level, with minor corrections for the HOMO, which is to be expected, given that the LUMO state has a predominantly Pb character.^{126,134}

Table 10: Highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of Pb_4S_4 referred to the vacuum level (see Section 3.1),⁹¹ as obtained with different methods. We also report the gap calculated as the energy difference of the LUMO and HOMO levels. The results of Ref. 47 were obtained using a ΔSCF procedure for the HOMO and LUMO levels. All entries are in eV.

Method	HOMO	LUMO	Gap
SR DFT	-5.36	-2.45	2.90
FR DFT	-5.37	-3.31	2.06
SR <i>GW</i>	-7.60	-1.17	6.43
FR <i>GW</i>	-7.56	-2.00	5.55
Ref. 47 FR DFT			1.98
Ref. 47 FR ΔSCF	-7.66	-1.74	5.92
Ref. 47 FR <i>GW</i>			5.5

6.2 $\text{Pb}_{14}\text{Se}_{13}$

We chose to investigate this non-stoichiometric nanoparticle, displayed in Figure 5, due to its interesting electronic structure which is not closed shell, given the excess Pb atom. The HOMO of this cluster is an intragap state (IGS) primarily localized on Pb surface atoms.^{127,128} The LUMO also has Pb character, and hence both HOMO and LUMO are expected to be affected by the inclusion of SOC.

6.2.1 Calculation details

We used the same computational procedure as in previous sections. A cubic cell with sides of 20 Å was adopted; using the Makov-Payne correction, the difference in the HOMO and LUMO energies obtained with a 30 Å cell was less than 1 meV. We employed a cutoff of 60 Ry, which led to less than 1 meV difference in the absolute HOMO/LUMO levels compared to calculations with a 120 Ry cutoff at the DFT level. For the dielectric matrix $N_{\text{PDEP}} = 1024$ was used.

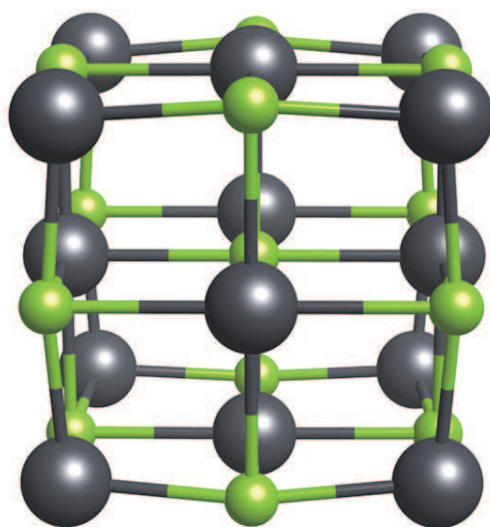


Figure 5: Perspective view of the $\text{Pb}_{14}\text{Se}_{13}$ nanoparticle with dark gray Pb atoms and green Se atoms.

To determine the geometry of the nanocrystal, we cut out a 3×3 cubic section of PbSe bulk. Atomic positions were then randomly perturbed by up to 0.5 \AA , and the structure relaxed. Up to eight relaxations with different random perturbations were performed to search for a reasonable state.

We used the geometry from an FR DFT relaxation ("FR geometry") for the FR *GW* calculation, and the geometry from an SR DFT relaxation ("SR geometry") for the SR *GW* calculation (both are available in the SI). This choice was made due to the fact that when using the FR geometry, the nanocrystal behaved as a metallic system in SR calculations (no intragap state formed). The two geometries had a root mean square (RMS) change in atomic positions of 0.08 \AA , and a maximum change in atomic positions of 0.10 \AA , including possible rotations and translations. At the FR DFT level of theory, the total energy of the SR geometry was found to be 0.88 eV larger than that of the FR geometry. Table 11 shows further details about the differences between the two nanocrystal configurations.

Table 11: Structural and electronic properties of the relaxed geometries of $\text{Pb}_{14}\text{Se}_{13}$, obtained using SR and FR calculations (acronyms are the same as in Table 2). Face separation measures the minimum distance between two atoms centered on opposing faces of the cluster. Corner distances are measured from the central atom to each of the 8 atoms in corners; the root mean square (RMS) average is given. Maximum diameter is the maximum distance between two corner atoms. All distances are in Å. E_{g1} (eV) and E_{g2} (eV) were calculated as described in Figure 6. The entry marked with a * only converged using fractional occupation numbers with Fermi-Dirac smearing of 0.002 Ry.

Quantity	SR geometry	FR geometry
Face separation	6.10	6.12
RMS corner distance	4.84	4.93
Max. diameter	9.68	9.86
SR PBE E_{g1}	1.97	1.86
FR PBE E_{g1}	1.24	1.61
SR PBE E_{g2}	0.87	< 0.03*
FR PBE E_{g2}	0.30	0.59

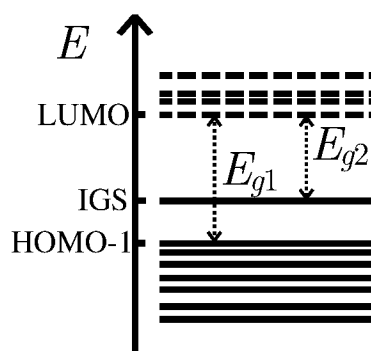


Figure 6: Schematic representation of energy levels in the $\text{Pb}_{14}\text{Se}_{13}$ nanocluster. Solid horizontal lines indicate occupied energy levels, including the intragap state (IGS), coinciding with the HOMO, and the state below the IGS, "HOMO-1." Dashed horizontal lines indicated unoccupied energy levels, including the LUMO. The energy differences E_{g1} and E_{g2} measure the indicated energy gaps in the nanocluster.

6.2.2 Results

As might be expected, the *GW* corrections are quite large for $\text{Pb}_{14}\text{Se}_{13}$, and SOC corrections, while comparatively smaller, are also substantial. We report both the location of the highest state delocalized over Se (“HOMO-1”), and the IGS and LUMO locations (see Figure 6). Results are summarized in Table 12. In general, we find that the SR DFT results compare reasonably well with the results of recent calculations,¹²⁸ while both FR and *GW* corrections are necessary to approach an accuracy of ± 0.2 eV compared to experiments. As expected, the effect of SOC on the highest Se state is rather small. States with Pb character are affected to a greater extent: the intragap state by 0.07 eV to 0.11 eV, and the LUMO by 0.37 eV to 0.39 eV.

Table 12: Computed energies (eV) of the second-highest occupied molecular orbital (HOMO-1), intragap state (IGS), lowest unoccupied molecular orbital (LUMO), and energy gaps (E_{g1} and E_{g2} , see Figure 6) of the $\text{Pb}_{14}\text{Se}_{13}$ nanoparticle, as obtained with different methods (acronyms are the same as in Table 2). SR and FR calculations were performed with slightly different geometries (see text).

Method	HOMO-1	IGS	LUMO	E_{g1}	E_{g2}
SR DFT	-5.02	-3.92	-3.05	1.97	0.87
FR DFT	-5.05	-4.03	-3.44	1.61	0.59
SR <i>GW</i>	-6.58	-5.45	-2.37	4.21	3.08
FR <i>GW</i>	-6.56	-5.52	-2.74	3.81	2.77

7 Validity of separate *GW*/SOC corrections

A common approximation to FR *GW* results adopted in the literature is to calculate the SOC effect at the DFT level and add it to SR *GW* results.^{135–139} Here we assess the validity of this approximation by considering the difference between FR DFT and SR DFT energy levels, which we add to SR *GW* results. We term this approximation an “SR-*GW*+ ΔFR ” calculation.

SR *GW* and FR DFT calculations can both be viewed as yielding perturbative corrections

to SR DFT results, hence one may expect that combining the two would only omit second-order perturbative corrections with respect to full FR *GW* results. This view is bolstered by a recent analysis of the SR-*GW*+ Δ FR method on a small set of atomic and molecular VIPs;³⁴ formal analyses also established that low-order perturbative terms are sufficient to accurately describe SOC contributions.^{33,140,141} However, for systems whose electronic structure is dramatically affected by spin-orbit coupling, e.g. topological insulators,¹⁰ the use of the SR-*GW*+ Δ FR method would seem to be an uncontrolled, and potentially poor, approximation to the correct many-body electronic structure.⁴¹

Our results allowed us to assess the validity of this approximation for all of the systems presented above. In Table 13, we show the VIP results of the GW-SOC81 data set when using the SR-*GW*+ Δ FR method. Comparing to the results in Table 2 (see summary of deviations from the FR *GW* calculation at the bottom of Table 13), we concluded that in general differences are quite small and the SR-*GW*+ Δ FR method is a good approximation to the FR *GW* results. Statistically, we found unadjusted *p*-values for a difference in medians between FR *GW* and SR-*GW*+ Δ FR of *p* = 0.06 at the PBE level and *p* = 0.15 at the PBE0 level. Thus, while in both cases FR *GW* shows lower MAE values than SR-*GW*+ Δ FR, more data is needed to conclude that actual improvements are present.

In Table 13 there are three outliers at the *GW*@PBE level that show corrections larger than 0.2 eV: KI, NaI, and RbI. In these cases, the result is due to the FR *GW* and SR *GW* calculations differing just enough for the secant solver of the quasiparticle equation to lead to different roots which are both valid (see Section 8 below).

Table 13: Vertical ionization potential (VIP, in eV) of the GW-SOC81 set of molecules. Here we report results obtained with the “SR-*GW*+ Δ FR” method (see text), using Kohn-Sham wave functions computed with the PBE and hybrid PBE0 functionals. Summarized at the end of the table are the mean error (ME), standard deviation of the errors (SE), mean absolute error (MAE), and mean absolute relative error (MARE) for each method, compared both to experiment, and to FR *GW* results from Table 2.

Number	Molecule	<i>GW</i> @PBE	<i>GW</i> @PBE0
1	AgBr	8.89	9.38
2	AgCl	9.29	9.97
3	AgI	8.33	8.67
4	Al ₂ Br ₆	10.34	10.74

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

5	AlBr ₃	10.49	10.86
6	AlI ₃	9.27	9.59
7	AsBr ₃	9.77	10.10
8	AsCl ₃	10.67	10.99
9	AsF ₃	12.49	12.89
10	AsF ₅	14.49	15.26
11	AsH ₃	10.33	10.54
12	AsI ₃	8.81	9.12
13	Br ₂	10.17	10.42
14	BrCl	10.67	10.93
15	C ₂ H ₂ Se	8.48	8.72
16	C ₂ HBrO	8.97	9.27
17	C ₄ H ₄ Se	8.60	8.86
18	(C ₅ H ₅) ₂ Ru	7.00	
19	CF ₃ I	10.21	10.48
20	(CH ₃) ₂ Cd	8.83	9.10
21	(CH ₃) ₂ Hg	9.06	9.27
22	(CH ₃) ₂ Se	8.17	8.41
23	(CH ₃) ₂ Zn	9.38	9.65
24	CH ₃ HgBr	9.58	9.96
25	CH ₃ HgCl	10.25	10.71
26	CH ₃ HgI	8.79	9.08
27	CH ₃ I	9.28	9.46
28	Cl ₄	8.70	9.06
29	CaBr ₂	9.67	10.10
30	CaI ₂	8.81	9.18
31	CdBr ₂	10.06	10.46
32	CdCl ₂	10.86	11.34
33	CdI ₂	9.10	9.43
34	CsCl	7.84	8.47
35	CsF	8.24	9.08
36	CsI	6.75	7.20
37	CuF	10.17	10.00
38	HgCl ₂	10.85	11.27
39	I ₂	9.07	9.27
40	IBr	9.53	9.74
41	ICl	9.88	10.08
42	IF	10.25	10.48
43	KBr	7.44	8.07
44	KI	6.86	7.39
45	Kr ₂	13.27	13.52
46	KrF ₂	12.50	13.20
47	LaBr ₃	9.80	10.30
48	LaCl ₃	10.72	11.24
49	LiBr	8.82	9.21
50	LiI	8.08	8.36
51	MgBr ₂	10.36	10.76
52	MgI ₂	9.33	9.66
53	Mo(CO) ₆	8.51	
54	NaBr	7.93	8.54
55	NaI	7.39	7.84
56	OsO ₄	11.74	12.38
57	PBr ₃	9.57	9.88
58	POBr ₃	10.45	10.89
59	RbBr	7.28	7.95
60	RbCl	7.72	8.50
61	RbI	6.71	7.27
62	RuO ₄	11.44	12.19
63	SOBr ₂	10.07	10.50
64	SPBr ₃	9.43	9.79
65	SeCl ₂	9.24	9.54
66	SeO ₂	11.03	11.60
67	SiBrF ₃	11.65	11.95
68	SiH ₃ I	9.65	9.87
69	SrBr ₂	9.35	9.77
70	SrCl ₂	9.98	10.47
71	SrI ₂	8.54	8.91
72	SrO	6.26	6.01

73	TiBr ₄	9.89	10.47
74	TiI ₄	8.67	9.17
75	ZnBr ₂	10.37	10.75
76	ZnCl ₂	11.31	11.75
77	ZnF ₂	12.66	13.39
78	ZnI ₂	9.32	9.63
79	ZrBr ₄	10.17	10.67
80	ZrCl ₄	11.32	11.90
81	ZrI ₄	8.94	9.36
Versus experiment			
	ME	-0.540	-0.149
	SE	0.296	0.228
	MAE	0.541	0.185
	MARE	5.41%	1.88%
Versus FR <i>GW</i>			
	ME	-0.009	-0.000
	SE	0.090	0.008
	MAE	0.033	0.005
	MARE	0.42%	0.05%

In the case of valence band splitting presented in Section 4.2, FR *GW* may give better results than an FR DFT-based correction, but without strong statistical support and with an average improvement of 0.012 eV in the absolute error. The largest improvement was found for AlSb (0.075 eV); we note that Ref. 36 shows a similar improvement of about 0.1 eV in the same quantity for HgSe and HgTe.

Finally, in Section 5 a larger correction between FR *GW* and SR-*GW*+ Δ FR was found for MAPbI₃, with a shift of -0.29 eV to -0.41 eV from SR-*GW*+ Δ FR to FR *GW*, depending on the pseudopotential and functional. We note that Ref. 51 found similar differences between these approaches, of -0.23 eV to -0.31 eV depending on the level of theory, while Ref. 48 showed a smaller shift in the opposite direction of 0.07 eV. In contrast to the perovskite case, we found maximum differences of only 0.04 eV for lead chalcogenide nanoparticles (see Tables 10 and 12 in Section 6).

In summary, while in a few cases the difference between the results of FR *GW* and SR-*GW*+ Δ FR approaches can be 0.3 eV or more, in most cases we found that these methods performed in a very similar fashion. More work is necessary to understand if FR *GW* performs better than SR-*GW*+ Δ FR.

8 Graphical solutions to the quasiparticle equation

In G_0W_0 calculations, the function

$$F(\omega) = \epsilon_{\mathbf{k}\mu}^{\text{KS}} + \sum_{\alpha\beta} \left\langle \psi_{\mathbf{k}\mu}^{\alpha} \left| \hat{\Sigma}_{\alpha\beta}(\omega) - \hat{V}_{\alpha\beta}^{xc} \right| \psi_{\mathbf{k}\mu}^{\beta} \right\rangle - \omega, \quad (14)$$

is used to find the quasiparticle energies by solving $F(\omega) = 0$. While various approximations or root-finding algorithms can be used to obtain these solutions, the most thorough method is a graphical solution, in which $F(\omega)$ is plotted for all ω and its roots are determined from the graph. An example is presented in Figure 7. The WEST code can perform this calculation over all ω at minimal computational cost, due to its use of a Lanczos chain approach,^{14,78} which also allows for the use of contour deformation techniques,^{14,20,23} rather than plasmon-pole^{12,13,16–21} or analytic continuation^{21–24} methods. Note that all figures in this section were obtained at the FR $GW@PBE$ level of theory.

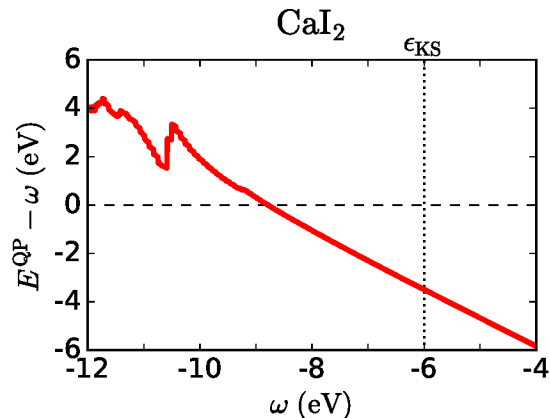


Figure 7: Computed quasiparticle energy $E^{\text{QP}}(\omega)$ as a function of frequency; when the curve $E^{\text{QP}}(\omega) - \omega = 0$, a solution of eq 14 is found. The vertical dotted line labels the Kohn-Sham energy.

We compare the graphical solutions of eq 14 to the results obtained with the secant method used in the literature and in previous calculations with the WEST code,¹⁴ and to those of the so-called “one-shot” approximation which estimates the slope of $F(\omega)$ near ϵ_{KS} in order to determine the root’s location.¹² Equation 14 may have multiple roots,^{54,55,142}

which in general cannot be detected by the secant or one-shot methods. Three different systems exhibiting multiple roots are shown in Figure 8. Specifically, in CuF a pole causes the one-shot approximation to deviate significantly from the solution of eq 14, which is found by both the secant and graphical methods. In RbCl, the graphical solution reveals a second candidate quasiparticle energy at $\omega = -8.71$ eV. Finally, for SrO, both the one-shot and secant methods give inaccurate results due to the presence of a pole.

In order to treat cases of multiple roots, additional information is needed to establish the physical quasiparticle energy, such as the slope of the self-energy.¹² Here, we make use of the spectral function $A_{\mathbf{k}\mu}(\omega)$ defined below. The area of the peak in the spectral function corresponding to each root represents the “weight” with which each root contributes to the full quantum state. The general retarded Green’s function is

$$G_{\mathbf{k}\mu}^{\text{ret}}(\omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}\mu} - \Sigma_{\mathbf{k}\mu}^{\text{ret}}(\omega)}, \quad (15)$$

and the spectral function is given by

$$A_{\mathbf{k}\mu}(\omega) = -2\text{Im } G^{\text{ret}} = \frac{2\Gamma}{\sigma^2 + \Gamma^2} \quad (16)$$

where $\sigma = \omega - \epsilon_{\mathbf{k}\mu} - \text{Re } \Sigma_{\mathbf{k}\mu}^{\text{ret}}(\omega)$ and $\Gamma = -\text{Im } \Sigma_{\mathbf{k}\mu}^{\text{ret}}(\omega)$.¹⁴³ The real and imaginary parts of Σ are obtained by the contour deformation technique. For integer occupation numbers, $\Sigma^{\text{ret}} = \Sigma^*$ for occupied states and $\Sigma^{\text{ret}} = \Sigma$ for unoccupied states.¹⁴³ In both cases

$$A_{\mathbf{k}\mu}(\omega) = \frac{2|\text{Im } \Sigma_{\mathbf{k}\mu}(\omega)|}{\sigma^2 + |\text{Im } \Sigma_{\mathbf{k}\mu}(\omega)|^2}. \quad (17)$$

The spectral function computed from eq 17 is reported in Figure 9 for CuF.

The results presented in Table 14 were obtained by using the value of $A_{\mathbf{k}\mu}(\omega)$ at each intercept ω_i to sort the roots. We also considered an averaged $A_{\mathbf{k}\mu}(\omega)$ in order to capture the full quasiparticle peak weight. To do so, for each ω_i , A was averaged over an energy of ± 0.1

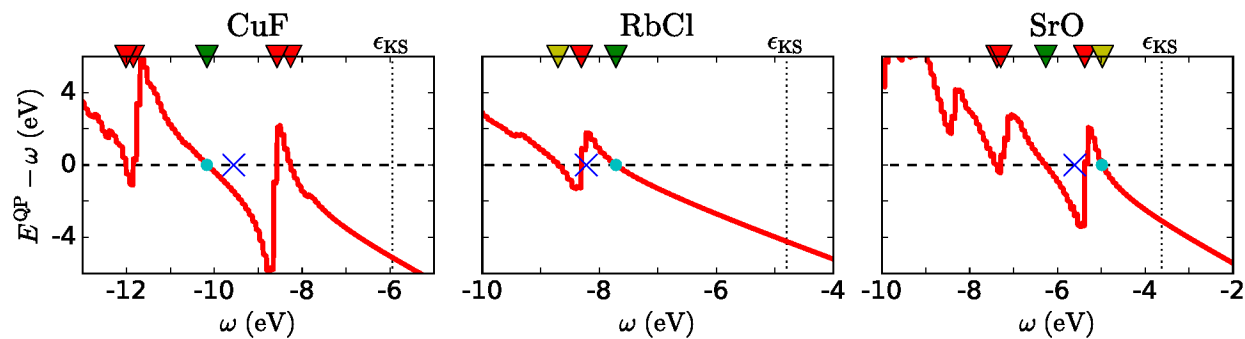


Figure 8: Computed quasiparticle energies (E^{QP}) as a function of frequency. Blue crosses represent the result obtained by the one-shot approximation; the cyan circles represent the secant method result. Arrows represent the locations at which $E^{\text{QP}}(\omega) = \omega$. Green, yellow, and red arrows correspond to the ratio $A(\omega)/A_{\text{max}} \geq 2/3$, $\geq 1/3$, and $< 1/3$ for green, yellow, and red respectively.

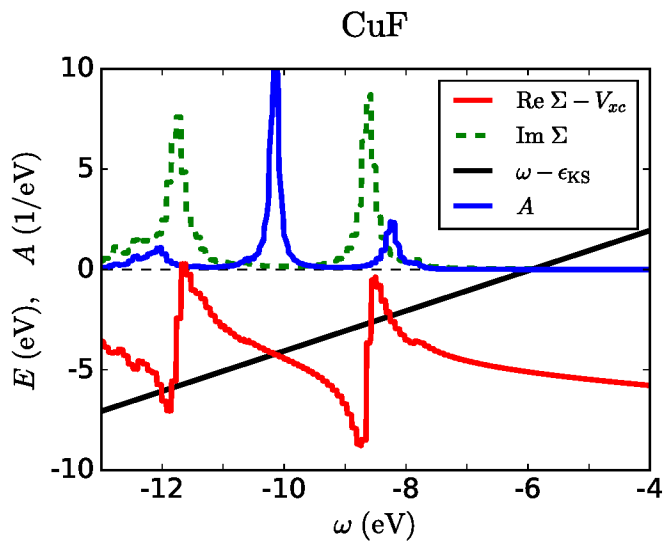


Figure 9: The real and imaginary part of Σ , together with the spectral function A (see eq 17) are shown as a function of frequency for the molecule CuF. The solution corresponding to the largest weight of the spectral function is -10.17 eV, also given by the secant method, while the one-shot approximation gives -9.44 eV.

eV over ω_i . For quasiparticle solutions with $A(\omega)/A_{\max} > 0.1$, the only ordering change this method caused was for the $\omega = 8.77$ eV and 8.70 eV solutions in Cl_4 , which is a consequence of having three solutions within the interval of 0.1 eV over which A is averaged.

We note that secant values were used in Table 2, *except* for cases in which the secant solver never converged and the root of the graphical solution was found to differ from the secant one, in which case the final 30th iteration is listed for the secant method in Table 14, whereas the graphical solution value with largest A is listed in Table 2. For SR GW@PBE , this leads to small changes in VIPs for KBr and RbBr; for FR GW@PBE it causes large changes for CsI and SrO. For secant calculations at the GW@PBE0 level, all VIP values converged, and in no cases was there a second quasiparticle solution with larger A than the one found using the secant solution.

There are some systems in which the graphical solution for FR GW@PBE yields a significantly better result if the optimal root is chosen, such as CsCl, CsI, KI, RbI, and SrO. However, in most of these cases there are multiple roots, and while the spectral function is generally a good guide of which root to choose, it does not always yield the root closest to the experimental value. In all multi-root cases, though, the graphical method does show that there are intrinsic uncertainties in the GW solution that one-shot or secant methods are oblivious to. The errors of the various approximations used to solve eq 14 are summarized in Table 15. We see that one-shot biases all VIP values to higher energies; this gives superior MAE values for SR GW@PBE , FR GW@PBE , and FR GW@PBE0 , all with $p < 10^{-7}$. In comparing secant and graphical (A) solutions, the number of differences are not enough to form any statistically significant conclusions.

Table 14: Values for the VIP (eV) of small molecules at the FR GW@PBE level, obtained using various techniques to find quasiparticle energies, as described in the text. For the graphical method, the main roots listed have a spectral function (A) value within a factor of 10 of the maximum spectral function value found at an intercept, $A(\omega_0)$. While most roots with smaller spectral values were at much higher energies, those roots with small $A(\omega)$ but within 2 eV or less of the experimental VIP are given in parentheses. When multiple roots of eq 14 met the criteria for inclusion, they are listed in descending order of $A(\omega_i)$, with the value closest to experiment highlighted in green (not considering those roots in parentheses).

Number	Molecule	1-shot	Secant	Graphical	Experiment
1	AgBr	9.27	8.90	8.91, 10.01 (9.70)	9.59
2	AgCl	9.82	9.31	9.31, 10.43 (9.95)	10.08
3	AgI	8.59	8.44	8.44, 9.44 (9.36)	8.80
4	Al ₂ Br ₆	10.44	10.34	10.34	10.97
5	AlBr ₃	10.59	10.48	10.48	10.91
6	AlI ₃	9.34	9.23	9.23	9.66
7	AsBr ₃	9.86	9.77	9.77	10.21
8	AsCl ₃	10.75	10.66	10.66	10.90
9	AsF ₃	12.57	12.49	12.49	13.00
10	AsF ₅	14.68	14.49	14.49	15.53
11	AsH ₃	10.40	10.33	10.33	10.58
12	AsI ₃	8.89	8.72	8.72	9.00
13	Br ₂	10.24	10.16	10.16	10.54
14	BrCl	10.74	10.67	10.67	11.01
15	C ₂ H ₂ Se	8.54	8.48	8.48	8.71
16	C ₂ HBrO	9.02	8.97	8.97	9.10
17	C ₄ H ₄ Se	8.65	8.60	8.60	8.86
18	(C ₅ H ₅) ₂ Ru	6.97	6.90	6.90	7.45
19	CF ₃ I	10.27	10.20	10.20	10.45
20	(CH ₃) ₂ Cd	8.90	8.83	8.83	8.80
21	(CH ₃) ₂ Hg	9.12	9.07	9.07	9.32
22	(CH ₃) ₂ Se	7.88	8.12	8.12	8.40
23	(CH ₃) ₂ Zn	9.46	9.38	9.38	9.40
24	CH ₃ HgBr	9.74	9.59	9.59	10.16
25	CH ₃ HgCl	10.44	10.26	10.26	10.84
26	CH ₃ HgI	8.91	8.79	8.79	9.25
27	CH ₃ I	9.34	9.26	9.26	9.52
28	Cl ₄	8.81	8.76	8.76, 8.70, 8.69	9.10
29	CaBr ₂	9.84	9.67	9.67	10.35
30	CaI ₂	8.95	8.79	8.79	9.39
31	CdBr ₂	10.21	10.06	10.06	10.58
32	CdCl ₂	11.05	10.87	10.87	11.44
33	CdI ₂	9.22	9.09	9.09	9.57
34	CsCl	8.21	7.84	7.84, 8.51 (8.43)	8.69
35	CsF	8.60	8.24	8.24	9.68
36	CsI	7.03	7.23	6.64, 7.15, 6.90	7.46
37	CuF	9.44	10.17	10.17, 8.26 (12.01, 11.85, 8.57)	10.90
38	HgCl ₂	11.00	10.85	10.85	11.50
39	I ₂	9.10	9.01	9.01	9.35
40	IBr	9.57	9.51	9.51	9.85
41	ICl	9.91	9.85	9.85	10.10
42	IF	10.29	10.23	10.23	10.62
43	KBr	7.85	7.37	7.37, 8.40 (8.15)	8.82
44	KI	7.20	7.34	6.70, 7.34, 7.02	7.40
45	Kr ₂	13.37	13.27	13.27	13.77
46	KrF ₂	12.73	12.51	12.51	13.34
47	LaBr ₃	9.98	9.82	9.82	10.68
48	LaCl ₃	10.89	10.72	10.72	11.29
49	LiBr	9.03	8.81	8.81	9.44
50	LiI	8.23	8.04	8.04	8.44
51	MgBr ₂	10.52	10.35	10.35	10.85
52	MgI ₂	9.45	9.31	9.31	10.50
53	Mo(CO) ₆	8.52	8.50	8.50	8.50
54	NaBr	8.31	7.89	7.89, 9.17, 15.10 (8.96)	8.70
55	NaI	7.66	7.75	7.24, 7.75, 7.57, 7.49, 7.50	8.00
56	OsO ₄	11.87	11.74	11.74	12.35
57	PBr ₃	9.64	9.57	9.57	10.00
58	POBr ₃	10.59	10.49	10.49	11.03
59	RbBr	7.72	7.26	7.26, 8.31 (8.00)	8.62
60	RbCl	8.22	7.72	7.72, 8.71, 15.09, 11.66 (8.31)	8.70
61	RbI	7.07	7.19	6.56, 7.19, 6.85	7.30
62	RuO ₄	11.61	11.44	11.44	12.15
63	SOBr ₂	10.26	10.13	10.13	10.59
64	SPBr ₃	9.55	9.43	9.43	9.89
65	SeCl ₂	9.30	9.24	9.24	9.50
66	SeO ₂	11.20	11.03	11.03	11.76
67	SiBrF ₃	11.72	11.64	11.64	12.46

68	SiH ₃ I	9.71	9.64	9.64	9.82
69	SrBr ₂	9.53	9.35	9.35	9.82
70	SrCl ₂	10.18	9.97	9.97	10.20
71	SrI ₂	8.70	8.51	8.51	9.01
72	SrO	5.61	4.99	6.26, 4.98, 7.38 (7.29, 5.38)	6.91
73	TiBr ₄	10.07	9.89	9.89	10.57
74	TiI ₄	8.84	8.65	8.65	9.27
75	ZnBr ₂	10.50	10.37	10.37	10.89
76	ZnCl ₂	11.48	11.32	11.32	11.80
77	ZnF ₂	12.93	12.66	12.66	13.91
78	ZnI ₂	9.42	9.30	9.30	9.73
79	ZrBr ₄	10.31	10.17	10.17	10.86
80	ZrCl ₄	11.49	11.32	11.32	11.94
81	ZrI ₄	9.07	8.93	8.93	9.55

Table 15: Mean error (ME), standard deviation of the errors (SE), mean absolute error (MAE), and mean absolute relative error (MARE) in eV of results presented in Table 14 for the different root-finding methods described in Section 8. Graphical solution method “Graph. *A*” yields the root with the largest spectral function value $A(\omega_i)$. Graphical solution method “Graph. opt” yields the root closest to the experimental value (see green values in Table 14).

Run	Measure	One-shot	Secant	Graph. <i>A</i>	Graph. opt
SR <i>GW</i> @PBE	ME	-0.276	-0.420	-0.421	-0.320
	SE	0.321	0.319	0.321	0.338
	MAE	0.312	0.430	0.431	0.378
	MARE	3.06%	4.24%	4.25%	3.64%
FR <i>GW</i> @PBE	ME	-0.407	-0.540	-0.553	-0.422
	SE	0.277	0.343	0.298	0.329
	MAE	0.411	0.541	0.554	0.465
	MARE	4.10%	5.41%	5.56%	4.50%
SR <i>GW</i> @PBE0	ME	0.024	-0.014	-0.014	-0.014
	SE	0.241	0.253	0.253	0.253
	MAE	0.174	0.176	0.175	0.175
	MARE	1.84%	1.84%	1.84%	1.84%
FR <i>GW</i> @PBE0	ME	-0.109	-0.149	-0.149	-0.149
	SE	0.215	0.227	0.227	0.227
	MAE	0.155	0.184	0.184	0.184
	MARE	1.58%	1.87%	1.87%	1.87%

9 Conclusions

We reported the results of large-scale G_0W_0 calculations including spin-orbit coupling for molecules, solids and nanocrystals, as carried out using fully-relativistic pseudopotentials

and two-component spinor wave functions. Our implementation, which avoids the use of empty states and plasmon-pole models, enabled G_0W_0 -SOC calculations for large systems (with thousands of electrons), including complex semiconductors and nanocrystals. We presented the newly-developed GW-SOC81 set, which contains 81 small molecules with at least one heavy element ($Z \geq 29$) and associated experimental VIP data and geometries. Through the use of the GW-SOC81 set, we reported evidence that aside from a systematic underestimation of VIP, FR GW calculations perform better than SR ones for the prediction of VIPs of small molecules. In effectively all cases studied here, we found that energy shifts from both spin-orbit and GW corrections are essential to obtain results in agreement with experiments.

We also showed that in the majority of cases, the SR- $GW+\Delta$ FR method is a good approximation to fully FR GW calculations. However more work is necessary to establish in what cases such an approximation is most susceptible to failure.

In the case of many small molecules, we showed that the use of “one-shot” or secant methods to solve the quasiparticle equation as a function of frequency may be susceptible to failure, especially at the $GW@PBE$ level of theory. We note that the GW method used here permits efficient computation of the quasiparticle equation over a large range of frequencies, which may lead to much more precise solutions of the quasiparticle equation—or to the observation of multiple, potentially valid solutions.

Finally we discussed the importance in separating out definitive improvements from those that occur by chance, when comparing results obtained at different levels of theory. We suggest the statistical analyses used here should be widely adopted in the assessment of the accuracy of electronic structure methods.

Acknowledgement

We acknowledge Nicholas Brawand, Matthew Goldey, François Gygi, and Márton Vörös for many useful discussions, and Christopher Knight for assistance with the use of computa-

tional resources at the Argonne Leadership Computing Facility. We especially thank Martin Schlipf for creating an SG15 iodine pseudopotential with d electrons in the core for this work. This research was supported by an appointment (P.S.) to the Intelligence Community Postdoctoral Research Fellowship Program at The University of Chicago, administered by Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy (DOE) and the Office of the Director of National Intelligence; by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (I.H.) through the “World Premier International Research Center Initiative” and “Development of Environmental Technology using Nanotechnology” programs, and by Japan Science and Technology Agency through “Materials Research by Information Integration Initiative”; by MICCoM (M.G. and G.G.) as part of the Computational Materials Sciences Program funded by the U.S. DOE, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, under grant #DOE/BES 5J-30161-0010A. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357, and resources provided by the University of Chicago Research Computing Center. The crystal graphics were generated with the VESTA software.¹⁴⁴

Supporting Information Available

A document is included with further tables and descriptions of additional small molecule results, GW-SOC81 selection procedures, DFT-level calculations, and calculations of the absolute band gap in semiconductors. A text file is also provided with coordinates used (in Å) for the GW-SOC81 set, other small molecules noted in the appendix, Pb₄S₄, and Pb₁₄Se₁₃ (FR and SR geometries). Finally, inputs to the ONCVSP code^{88,89} are provided for the I²⁵⁺ pseudopotential. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

A Effect of including d electrons in the valence partition for pseudopotential calculations at the G_0W_0 level

Pseudopotential calculations require the definition of core and valence electron partitions. Sets of pseudopotentials, including SG15, commonly include semicore d electrons in the valence partition, without including the s and p electrons of the same shell in the valence. It has been shown, however, that excluding s and p semicore electrons can lead to substantial errors in GW calculations, as observed for systems including elements such as Cd,^{145,146} Zn,^{147,148} Cu,¹⁴⁹ and Au.¹³⁶ The results of GW calculations turn out to be quite sensitive to the inclusion of incomplete semicore shells in the valence partition (ISSV), even in cases where the corresponding DFT results were unaffected. This sensitivity of GW is due in large part to the dependence of the exchange integral only on wavefunction overlaps, not on relative energies of single-particle states,^{149,150} though the correlation part of the self-energy also showed sensitivity in some systems.^{51,151} Alternative methods or corrections to the use of ISSV have been explored in the literature.^{151–153}

In spite of these concerns, GW calculations with ISSV pseudopotentials are still frequently employed.^{48,51,150,154–156} Avoiding the use of ISSV pseudopotentials can be cumbersome, as including all semicore states in the valence requires high energy cutoffs, and pre-generated pseudopotential sets often contain only ISSV pseudopotentials. Here we present the results for molecules and bulk systems obtained with different core-valence partitions. We find cases, e.g. systems containing iodine, where the use of ISSV pseudopotentials in GW calculations leads to calculated VIP errors of over 1 eV.

A.1 Effect of iodine semicore states on VIP results

We analyzed the case of I in detail. We used the original 4d5s5p SG15 pseudopotential for iodine,^{86,87} a second pseudopotential produced in the same manner but with only a 5s5p valence configuration, and a third iodine pseudopotential with a 4s4p4d5s5p valence

configuration which was also created using the ONCVSP code⁸⁸ but simply choosing input parameters by hand (i.e. without the optimizations used by SG15 pseudopotentials;⁸⁶ inputs available in the SI).

Calculations presented in this section are for a subset of molecules with two or three atoms and one or two elements with $Z \geq 29$, for which experimental VIP data available in the NIST WebBook.⁸¹ Geometries were obtained by relaxing molecules at the FR PBE level until forces were less than 10^{-5} Ry/ a_0 (MAE of 64 meV vs. FR GW@PBE results with experimental geometries). SG15 pseudopotentials, or fully-relativistic versions derived with the same input parameters, were used, except for some iodine calculations as specified above. The kinetic energy cutoff for wavefunctions was chosen to be 60 Ry for I^{7+} and I^{17+} calculations (MAE vs. 90 Ry cutoffs of 3 meV for FR PBE HOMO values), while for I^{25+} calculations the cutoff was chosen for each system to converge the PBE HOMO value, at both the SR and FR levels, within 10 meV of a 160 Ry calculation; such cutoffs ranged from 80 Ry to 120 Ry. Finally, a cubic cell of 30 a_0 was used (MAE vs. a 45 a_0 cell of 3 meV for FR PBE HOMO values), and $N_{\text{PDEP}} = 256$ (MAE of 25 meV compared to $N_{\text{PDEP}} = 512$ for FR GW@PBE for a 16-molecule subset).

Results for all configurations at the GW@PBE level, and for selected configurations at the GW@PBE0 level, are shown in Table 16. We find that in all cases the 4d5s5p valence configuration increases the average VIP significantly (by 0.65 eV to 0.82 eV depending on the method), whereas the 4s4p4d5s5p and 5s5p valence configurations behave very similarly. While systematic errors in the VIP at the PBE level make it hard to determine which results best fit experiment, at the PBE0 level there is a clear improvement in experimental agreement when using the 5s5p valence configuration, as opposed to the 4d5s5p configuration (unadjusted p -value of 0.001 for difference in median AE). Finally, we note that Table 16 includes a number of elements—Ga, Ge, In, Pb, Sn and Tl—for which SG15 pseudopotentials also include semicore d electrons without the corresponding s and p electrons of the same shell, which may cause similar errors. Further calculations for molecules containing ISSV

pseudopotentials can be found in the SI.

Table 16: VIPs calculated for small molecules including iodine, with three different valence configurations of iodine pseudopotentials, and various G_0W_0 -based methods. Summary statistics for each method compared to experiment are also given.

Molecule	SR $GW@PBE$			FR $GW@PBE$			FR $GW@PBE0$		Expt.
	I ⁷⁺	I ¹⁷⁺	I ²⁵⁺	I ⁷⁺	I ¹⁷⁺	I ²⁵⁺	I ⁷⁺	I ¹⁷⁺	
AgI	8.58	10.00	8.58	8.41	9.69	8.42	8.63	9.25	8.80
CaI ₂	9.09	9.77	9.11	8.80	9.51	8.84	9.18	9.84	9.39
CdI ₂	9.34	10.11	9.35	9.01	9.74	9.03	9.37	10.09	9.57
CsI	7.06	8.10	7.06	7.18	7.85	7.20	7.23	7.81	7.46
GaI	9.14	9.79	9.17	8.88	9.48	8.90	9.22	9.77	9.19
GeI ₂	8.96	9.73	8.96	8.76	9.42	8.77	9.12	9.80	9.08
I ₂	9.44	10.38	9.45	9.01	10.12	9.03	9.29	10.13	9.35
ICl	10.18	10.88	10.20	9.90	10.58	9.93	10.12	10.77	10.10
IF	10.63	11.46	10.67	10.33	11.13	10.38	10.54	11.33	10.62
InI	8.71	9.29	8.72	8.44	8.98	8.45	8.87	9.41	8.88
KI	7.18	8.39	7.18	7.34	8.18	7.34	7.41	8.00	7.40
LiI	8.32	8.98	8.34	8.01	8.79	8.04	8.35	8.99	8.44
MgI ₂	9.56	10.31	9.58	9.23	9.94	9.26	9.59	10.31	10.50
NaI	7.66	8.60	7.63	7.73	8.11	7.78	7.82	8.43	8.00
PbI ₂	8.61	9.21	8.61	8.37	9.46	8.37	8.85	9.51	8.90
RbI	7.05	8.29	6.98	7.26	8.04	7.22	7.29	7.88	7.30
SnI ₂	8.84	9.53	8.84	8.63	9.26	8.63	9.00	9.64	9.05
TlI	8.44	8.99	8.43	8.20	8.78	8.21	8.68	9.28	8.84
ZnI ₂	9.58	10.37	9.58	9.26	10.01	9.34	9.58	10.31	9.73
ME	-0.222	0.609	-0.218	-0.413	0.340	-0.394	-0.129	0.524	
SE	0.225	0.330	0.231	0.263	0.343	0.258	0.208	0.206	
MAE	0.241	0.629	0.245	0.413	0.405	0.394	0.140	0.545	
MARE	2.71%	7.18%	2.77%	4.48%	4.62%	4.28%	1.50%	6.11%	

A.2 Absolute values for semiconductor band gaps, and the effect of semicore states

As mentioned in Section 4, we found that numerous factors had significant effects on the calculation of the electronic band gap E_g , which prevented us from drawing conclusions about the importance of changes in the band gap between FR and SR calculations. Table 17 gives the E_g values using the methods described in Section 4, as well as the results of SR calcula-

tions using the experimental lattice parameter rather than the optimized lattice parameter. All DFT calculations, as well as *GW* calculations at the optimized lattice parameter, show gaps that are overall much smaller than experimental gaps. In all cases, errors are still large, and we find that both the valence configuration chosen for the pseudopotentials, and the type of pseudopotential have significant effects on E_g .

Table 17: Calculated and experimental values, and mean errors for the band gap E_g (eV), using the parameters described in Section 4 (acronyms are the same as in Table 2). Results for both the optimized lattice constant, (a_{opt} , Å) and the experimental lattice constant (a_{expt} , Å) are given.

System	a_{opt}	SR DFT	FR DFT	SR <i>GW</i>	FR <i>GW</i>	a_{expt}^{157}	SR DFT	SR <i>GW</i>	Expt. ¹⁵⁷
AlSb	6.23	1.22	1.02	1.89	1.65	6.136	1.20	2.23	1.60
CdS	5.93	1.02	1.01	1.61	1.59	5.832	1.14	1.77	2.42
CdTe	6.63	0.56	0.29	1.07	0.76	6.477	0.76	1.29	1.44
GaAs	5.75	0.15	0.01	0.25	0.13	5.653	0.50	0.62	1.35
GaP	5.51	1.54	1.51	1.95	1.91	5.450	1.63	2.17	2.24
InP	5.96	0.42	0.39	0.61	0.57	5.869	0.68	0.90	1.27
Si	5.48	0.72	0.71	1.38	1.36	5.431	0.67	1.29	1.12
ZnS	5.45	1.99	1.97	3.08	3.06	5.409	2.09	3.26	3.54
ZnSe	5.73	1.13	1.01	1.89	1.74	5.668	1.26	2.06	2.58
ZnTe	6.18	1.08	0.79	1.87	1.49	6.101	1.24	2.08	2.26
ME		-1.00	-1.11	-0.42	-0.55		-0.87	-0.22	
SE		0.42	0.43	0.44	0.44		0.38	0.40	
MAE		1.00	1.11	0.53	0.61		0.87	0.38	
MARE		51.7%	58.3%	30.3%	34.6%		43.9%	21.4%	

To understand our sources of error, we first performed computations with a number of different pseudopotentials at the DFT level, followed by further computations at the *GW* level. Details of these computations, and of the other pseudopotentials used,^{88,122,123,133,158,159} are given in the SI. In Table 18, we show G_0W_0 results for the band gap E_g . Comparisons with Ref. 160 are also given. Unfortunately, none of the pseudopotentials used here approach the performance shown in Ref. 160. In general, it seems that when semicore d electrons are included without corresponding s and p electrons of the same shell, band gaps are too low. However, when d electrons are omitted entirely, band gaps are too high. AlSb is the only compound to reverse this trend, while CdTe, GaAs, GaP, InP, and ZnTe all follow it to some degree. In contrast, Ref. 160 uses full semicore electron shells in the valence; our results indicate that such a choice is necessary for accurate gaps. We note that the gaps of CdS,

GaAs, and Zn compounds show variation between different pseudopotentials derived with the same valence partition. Further work is necessary to understand these discrepancies; such an investigation is beyond the scope of this paper.

Table 18: SR $GW@PBE$ results for the band gap E_g (eV) using a number of different pseudopotentials. Pseudopotentials labeled “alt.” differ in valence configurations from SG15 by placing semicore d electrons in the core.

System	SG15	PseudoDojo	PseudoDojo alt.	Ceresoli	GTH	GTH alt.	Ref. 160	Expt.
AlSb	2.23	2.27	1.41			1.60	1.87	1.60
CdS	1.77	2.23					2.15	2.42
CdTe	1.29		1.98				1.57	1.44
GaAs	0.62	1.26				1.27	1.23	1.35
GaP	2.17	2.17		2.18		2.39	2.30	2.24
InP	0.90	0.70			0.95	1.65	1.23	1.27
Si	1.29	1.23		1.31	1.31		1.28	1.12
ZnS	3.26	3.56					3.46	3.54
ZnSe	2.06						2.55	2.58
ZnTe	2.08	2.39	2.97				2.27	2.26
vs. Ref. 160								
ME	-0.22	0.00	0.22	-0.04	-0.13	0.07		
SE	0.27	0.26	0.60	0.11	0.22	0.28		
MAE	0.30	0.18	0.52	0.07	0.15	0.20		
MARE	17.1%	11.0%	27.1%	3.8%	12.5%	13.8%		
vs. Experiment								
ME	-0.22	0.00	0.35	0.07	-0.07	0.11	0.01	
SE	0.40	0.35	0.48	0.18	0.36	0.20	0.15	
MAE	0.38	0.23	0.48	0.12	0.25	0.15	0.12	
MARE	21.4%	15.1%	26.9%	9.8%	21.0%	10.6%	7.0%	

Notes and References

- (1) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (2) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (3) Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*; Cambridge University Press, 2008.
- (4) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659.
- (5) Kümmel, S.; Kronik, L. Orbital-dependent density functionals: Theory and applications. *Rev. Mod. Phys.* **2008**, *80*, 3–60.
- (6) Freysoldt, C.; Grabowski, B.; Hickel, T.; Neugebauer, J.; Kresse, G.; Janotti, A.; Van de Walle, C. G. First-principles calculations for point defects in solids. *Rev. Mod. Phys.* **2014**, *86*, 253–305.
- (7) Skone, J. H.; Govoni, M.; Galli, G. Self-consistent hybrid functional for condensed systems. *Phys. Rev. B* **2014**, *89*, 195112.
- (8) Skone, J. H.; Govoni, M.; Galli, G. Nonempirical range-separated hybrid functionals for solids and molecules. *Phys. Rev. B* **2016**, *93*, 235106.
- (9) Winkler, R. *Spin–Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*; Springer Tracts in Modern Physics; Springer, 2003; Vol. 191.
- (10) Hasan, M. Z.; Kane, C. L. Colloquium: Topological insulators. *Rev. Mod. Phys.* **2010**, *82*, 3045–3067.

- (11) Hedin, L. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.* **1965**, *139*, A796–A823.
- (12) Hybertsen, M. S.; Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B* **1986**, *34*, 5390–5413.
- (13) Aryasetiawan, F.; Gunnarsson, O. The GW method. *Rep. Prog. Phys.* **1998**, *61*, 237.
- (14) Govoni, M.; Galli, G. Large Scale GW Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 2680–2696.
- (15) WEST. <http://www.west-code.org/>, (accessed Jan 26, 2016).
- (16) Godby, R. W.; Needs, R. J. Metal-insulator transition in Kohn-Sham theory and quasiparticle theory. *Phys. Rev. Lett.* **1989**, *62*, 1169–1172.
- (17) Engel, G. E.; Farid, B. Generalized plasmon-pole model and plasmon band structures of crystals. *Phys. Rev. B* **1993**, *47*, 15931–15934.
- (18) Reining, L.; Onida, G.; Godby, R. W. Elimination of unoccupied-state summations in *ab initio* self-energy calculations for large supercells. *Phys. Rev. B* **1997**, *56*, R4301–R4304.
- (19) Shaltaf, R.; Rignanese, G.-M.; Gonze, X.; Giustino, F.; Pasquarello, A. Band Offsets at the Si/SiO₂ Interface from Many-Body Perturbation Theory. *Phys. Rev. Lett.* **2008**, *100*, 186401.
- (20) Stankovski, M.; Antonius, G.; Waroquiers, D.; Miglio, A.; Dixit, H.; Sankaran, K.; Giantomassi, M.; Gonze, X.; Côté, M.; Rignanese, G.-M. G^0W^0 band gap of ZnO: Effects of plasmon-pole models. *Phys. Rev. B* **2011**, *84*, 241201.
- (21) Giustino, F.; Cohen, M. L.; Louie, S. G. GW method with the self-consistent Sternheimer equation. *Phys. Rev. B* **2010**, *81*, 115105.

- (22) Rieger, M. M.; Steinbeck, L.; White, I. D.; Rojas, H. N.; Godby, R. W. The GW space-time method for the self-energy of large systems. *Comput. Phys. Commun.* **1999**, *117*, 211–228.
- (23) Lebègue, S.; Arnaud, B.; Alouani, M.; Bloechl, P. E. Implementation of an all-electron GW approximation based on the projector augmented wave method without plasmon pole approximation: Application to Si, SiC, AlAs, InAs, NaH, and KH. *Phys. Rev. B* **2003**, *67*, 155208.
- (24) Umari, P.; Stenuit, G.; Baroni, S. GW quasiparticle spectra from occupied states only. *Phys. Rev. B* **2010**, *81*, 115104.
- (25) Hybertsen, M. S.; Louie, S. G. Spin-orbit splitting in semiconductors and insulators from the *ab initio* pseudopotential. *Phys. Rev. B* **1986**, *34*, 2920–2922.
- (26) Oda, T.; Pasquarello, A.; Car, R. Fully Unconstrained Approach to Noncollinear Magnetism: Application to Small Fe Clusters. *Phys. Rev. Lett.* **1998**, *80*, 3622–3625.
- (27) Theurich, G.; Hill, N. A. Self-consistent treatment of spin-orbit coupling in solids using relativistic fully separable *ab initio* pseudopotentials. *Phys. Rev. B* **2001**, *64*, 073106.
- (28) Oda, T.; Hosokawa, A. Fully relativistic two-component-spinor approach in the ultrasoft-pseudopotential plane-wave method. *Phys. Rev. B* **2005**, *72*, 224428.
- (29) Dal Corso, A.; Conte, A. M. Spin-orbit coupling with ultrasoft pseudopotentials: Application to Au and Pt. *Phys. Rev. B* **2005**, *71*, 115106.
- (30) Liu, W. Ideas of relativistic quantum chemistry. *Mol. Phys.* **2010**, *108*, 1679–1706.
- (31) Dolg, M.; Cao, X. Relativistic Pseudopotentials: Their Development and Scope of Applications. *Chem. Rev.* **2012**, *112*, 403–480.
- (32) Nakajima, T.; Hirao, K. The Douglas-Kroll-Hess Approach. *Chem. Rev.* **2012**, *112*, 385–402.

- (33) Liu, W. Advances in relativistic molecular quantum mechanics. *Phys. Rep.* **2014**, *537*, 59–89.
- (34) Kühn, M.; Weigend, F. One-Electron Energies from the Two-Component GW Method. *J. Chem. Theory Comput.* **2015**, *11*, 969–979.
- (35) Zhao, R.; Zhang, Y.; Xiao, Y.; Liu, W. Exact two-component relativistic energy band theory and application. *J. Chem. Phys.* **2016**, *144*, 044105.
- (36) Sakuma, R.; Friedrich, C.; Miyake, T.; Blügel, S.; Aryasetiawan, F. GW calculations including spin-orbit coupling: Application to Hg chalcogenides. *Phys. Rev. B* **2011**, *84*, 085144.
- (37) Kutepov, A.; Haule, K.; Savrasov, S. Y.; Kotliar, G. Electronic structure of Pu and Am metals by self-consistent relativistic GW method. *Phys. Rev. B* **2012**, *85*, 155129.
- (38) Ahmed, T.; Albers, R. C.; Balatsky, A. V.; Friedrich, C.; Zhu, J.-X. GW quasiparticle calculations with spin-orbit coupling for the light actinides. *Phys. Rev. B* **2014**, *89*, 035104.
- (39) Molina-Sánchez, A.; Sangalli, D.; Hummer, K.; Marini, A.; Wirtz, L. Effect of spin-orbit interaction on the optical spectra of single-layer, double-layer, and bulk MoS₂. *Phys. Rev. B* **2013**, *88*, 045412.
- (40) Aguilera, I.; Friedrich, C.; Bihlmayer, G.; Blügel, S. GW study of topological insulators Bi₂Se₃, Bi₂Te₃, and Sb₂Te₃: Beyond the perturbative one-shot approach. *Phys. Rev. B* **2013**, *88*, 045206.
- (41) Aguilera, I.; Friedrich, C.; Blügel, S. Spin-orbit coupling in quasiparticle studies of topological insulators. *Phys. Rev. B* **2013**, *88*, 165136.
- (42) Nechaev, I. A.; Hatch, R. C.; Bianchi, M.; Guan, D.; Friedrich, C.; Aguilera, I.; Mi, J. L.; Iversen, B. B.; Blügel, S.; Hofmann, P.; Chulkov, E. V. Evidence for a direct

- band gap in the topological insulator Bi_2Se_3 from theory and experiment. *Phys. Rev. B* **2013**, *87*, 121111.
- (43) Rusinov, I. P.; Nechaev, I. A.; Ereemeev, S. V.; Friedrich, C.; Blügel, S.; Chulkov, E. V. Many-body effects on the Rashba-type spin splitting in bulk bismuth tellurohalides. *Phys. Rev. B* **2013**, *87*, 205103.
- (44) Michiardi, M.; Aguilera, I.; Bianchi, M.; de Carvalho, V. E.; Ladeira, L. O.; Teixeira, N. G.; Soares, E. A.; Friedrich, C.; Blügel, S.; Hofmann, P. Bulk band structure of Bi_2Te_3 . *Phys. Rev. B* **2014**, *90*, 075105.
- (45) Förster, T.; Krüger, P.; Rohlfing, M. Two-dimensional topological phases and electronic spectrum of Bi_2Se_3 thin films from *GW* calculations. *Phys. Rev. B* **2015**, *92*, 201404.
- (46) Aguilera, I.; Friedrich, C.; Blügel, S. Electronic phase transitions of bismuth under strain from relativistic self-consistent *GW* calculations. *Phys. Rev. B* **2015**, *91*, 125129.
- (47) Bisri, S. Z.; Degoli, E.; Spallanzani, N.; Krishnan, G.; Kooi, B. J.; Ghica, C.; Yarema, M.; Heiss, W.; Pulci, O.; Ossicini, S.; Loi, M. A. Determination of the Electronic Energy Levels of Colloidal Nanocrystals using Field-Effect Transistors and Ab-Initio Calculations. *Adv. Mater.* **2014**, *26*, 5639–5645.
- (48) Umari, P.; Mosconi, E.; De Angelis, F. Relativistic *GW* calculations on $\text{CH}_3\text{NH}_3\text{PbI}_3$ and $\text{CH}_3\text{NH}_3\text{SnI}_3$ Perovskites for Solar Cell Applications. *Sci. Rep.* **2014**, *4*.
- (49) Amat, A.; Mosconi, E.; Ronca, E.; Quarti, C.; Umari, P.; Nazeeruddin, M. K.; Grätzel, M.; De Angelis, F. Cation-Induced Band-Gap Tuning in Organohalide Perovskites: Interplay of Spin-Orbit Coupling and Octahedra Tilting. *Nano Lett.* **2014**, *14*, 3608–3616.

- (50) Mosconi, E.; Umari, P.; Angelis, F. D. Electronic and optical properties of mixed Sn-Pb organohalide perovskites: a first principles investigation. *J. Mater. Chem. A* **2015**, *3*, 9208–9215.
- (51) Filip, M. R.; Giustino, F. *GW* quasiparticle band gap of the hybrid organic-inorganic perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$: Effect of spin-orbit interaction, semicore electrons, and self-consistency. *Phys. Rev. B* **2014**, *90*, 245145.
- (52) Filip, M. R.; Verdi, C.; Giustino, F. *GW* Band Structures and Carrier Effective Masses of $\text{CH}_3\text{NH}_3\text{PbI}_3$ and Hypothetical Perovskites of the Type APbI_3 : $\text{A} = \text{NH}_4$, PH_4 , AsH_4 , and SbH_4 . *J. Phys. Chem. C* **2015**, *119*, 25209–25219.
- (53) Filip, M. R.; Giustino, F. Computational Screening of Homovalent Lead Substitution in Organic-Inorganic Halide Perovskites. *J. Phys. Chem. C* **2016**, *120*, 166–173.
- (54) Lischner, J.; Deslippe, J.; Jain, M.; Louie, S. G. First-Principles Calculations of Quasiparticle Excitations of Open-Shell Condensed Matter Systems. *Phys. Rev. Lett.* **2012**, *109*, 036406.
- (55) van Setten, M. J.; Caruso, F.; Sharifzadeh, S.; Ren, X.; Scheffler, M.; Liu, F.; Lischner, J.; Lin, L.; Deslippe, J. R.; Louie, S. G.; Yang, C.; Weigend, F.; Neaton, J. B.; Evers, F.; Rinke, P. *GW100*: Benchmarking G_0W_0 for molecular systems. *J. Chem. Theory Comput.* **2015**,
- (56) Cowan, R. D.; Griffin, D. C. Approximate relativistic corrections to atomic radial wave functions. *J. Opt. Soc. Am.* **1976**, *66*, 1010.
- (57) Koelling, D. D.; Harmon, B. N. A technique for relativistic spin-polarised calculations. *J. Phys. C: Solid State Phys.* **1977**, *10*, 3107.
- (58) Bachelet, G. B.; Schlüter, M. Relativistic norm-conserving pseudopotentials. *Phys. Rev. B* **1982**, *25*, 2103–2108.

- (59) Kleinman, L. Relativistic norm-conserving pseudopotential. *Phys. Rev. B* **1980**, *21*, 2630–2631.
- (60) Pyykko, P. Relativistic effects in structural chemistry. *Chem. Rev.* **1988**, *88*, 563–594.
- (61) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; Gironcoli, S. d.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
- (62) QUANTUM ESPRESSO. <http://www.quantum-espresso.org/>, (accessed Jan 26, 2016).
- (63) Kleinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.
- (64) MacDonald, A. H.; Pickett, W. E.; Koelling, D. D. A linearised relativistic augmented-plane-wave method utilising approximate pure spin basis functions. *J. Phys. C: Solid State Phys.* **1980**, *13*, 2675.
- (65) Hybertsen, M. S.; Louie, S. G. First-Principles Theory of Quasiparticles: Calculation of Band Gaps in Semiconductors and Insulators. *Phys. Rev. Lett.* **1985**, *55*, 1418–1421.
- (66) Aryasetiawan, F.; Biermann, S. Generalized Hedin's Equations for Quantum Many-Body Systems with Spin-Dependent Interactions. *Phys. Rev. Lett.* **2008**, *100*, 116402.
- (67) Aryasetiawan, F.; Biermann, S. Generalized Hedin equations and $\sigma G\sigma W$ approxi-

- mation for quantum many-body systems with spin-dependent interactions. *J. Phys.: Condens. Matter* **2009**, *21*, 064232.
- (68) Nguyen, H.-V.; Pham, T. A.; Rocca, D.; Galli, G. Improving accuracy and efficiency of calculations of photoemission spectra within the many-body perturbation theory. *Phys. Rev. B* **2012**, *85*, 081101.
- (69) Pham, T. A.; Nguyen, H.-V.; Rocca, D.; Galli, G. *GW* calculations using the spectral decomposition of the dielectric matrix: Verification, validation, and comparison of methods. *Phys. Rev. B* **2013**, *87*, 155148.
- (70) Adler, S. L. Quantum Theory of the Dielectric Constant in Real Solids. *Phys. Rev.* **1962**, *126*, 413–420.
- (71) Wiser, N. Dielectric Constant with Local Field Effects Included. *Phys. Rev.* **1963**, *129*, 62–69.
- (72) Wilson, H. F.; Gygi, F.; Galli, G. Efficient iterative method for calculations of dielectric matrices. *Phys. Rev. B* **2008**, *78*, 113303.
- (73) Wilson, H. F.; Lu, D.; Gygi, F.; Galli, G. Iterative calculations of dielectric eigenvalue spectra. *Phys. Rev. B* **2009**, *79*, 245106.
- (74) Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.
- (75) Baroni, S.; de Gironcoli, S.; Dal Corso, A.; Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **2001**, *73*, 515–562.
- (76) Sternheimer, R. M. Electronic Polarizabilities of Ions from the Hartree-Fock Wave Functions. *Phys. Rev.* **1954**, *96*, 951–968.

- (77) Car, R.; Tosatti, E.; Baroni, S.; Leelaprute, S. Dielectric band structure of crystals: General properties and calculations for silicon. *Phys. Rev. B* **1981**, *24*, 985–999.
- (78) Cini, M. *Topics and Methods in Condensed Matter Theory: From Basic Quantum Mechanics to the Frontiers of Research*; Springer, 2007.
- (79) Johnson III, R. D., Ed. *NIST Computational Chemistry Comparison and Benchmark Database*; NIST Standard Reference Database 101; National Institute of Standards and Technology, 2015.
- (80) van Setten, M. J.; Weigend, F.; Evers, F. The GW-Method for Quantum Chemistry Applications: Theory and Implementation. *J. Chem. Theory Comput.* **2013**, *9*, 232–246.
- (81) Linstrom, P., Mallard, W., Eds. *NIST Chemistry WebBook*; NIST Standard Reference Database 69; National Institute of Standards and Technology, <http://webbook.nist.gov>.
- (82) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (83) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985.
- (84) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (85) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110*, 5029–5036.
- (86) Schlipf, M.; Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **2015**, *196*, 36–44.

- (87) SG15 ONCV Potentials. http://www.quantum-simulation.org/potentials/sg15_oncv/, (accessed Jan 26, 2016).
- (88) Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B* **2013**, *88*, 085117.
- (89) Mat-Sim Research. <http://www.mat-simresearch.com/>, (accessed Jan 26, 2016).
- (90) The bromine pseudopotential parameters in the SG15 1.1 set led to a ghost state in the FR pseudopotential which in turn caused large discrepancies in results. Instead, for both SR and FR pseudopotentials the PseudoDojo v0.2 inputs¹⁵⁸ for ONCVSP^{88,89} were used. For iodine, as explained in Appendix A, a pseudopotential with 4d electrons in the core was used. The scalar relativistic version was generated by Martin Schlipf using the same techniques as the SG15 set; the FR version was based on those parameters.
- (91) Makov, G.; Payne, M. C. Periodic boundary conditions in *ab initio* calculations. *Phys. Rev. B* **1995**, *51*, 4014–4022.
- (92) Fisher, R. A. In *Statistical Methods, Experimental Design, and Scientific Inference*; Bennett, J. H., Ed.; Oxford University Press, 1990.
- (93) Colquhoun, D. An investigation of the false discovery rate and the misinterpretation of p-values. *R. Soc. Open Sci.* **2014**, *1*, 140216.
- (94) Friedman, M. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Am. Stat. Assoc.* **1937**, *32*, 675–701.
- (95) Hollander, M.; Wolfe, D. A. *Nonparametric statistical methods*, 2nd ed.; Wiley series in probability and statistics; Wiley, 1999.
- (96) Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1*, 80–83.

- (97) Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
- (98) Wright, S. P. Adjusted P-Values for Simultaneous Inference. *Biometrics* **1992**, *48*, 1005–1013.
- (99) Bauer, D. F. Constructing Confidence Sets Using Rank Statistics. *J. Am. Stat. Assoc.* *67*, 687–690.
- (100) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188–5192.
- (101) Winkler, D. R. *Spin–Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*; Springer Tracts in Modern Physics 191; Springer Berlin Heidelberg, 2003; pp 21–34.
- (102) Other comparisons were not made with Ref. 36 as the implementation of fractional occupation states, necessary for the simulation of gapless materials, is still in progress in WEST.
- (103) Svane, A.; Christensen, N. E.; Cardona, M.; Chantis, A. N.; van Schilfgaarde, M.; Kotani, T. Quasiparticle band structures of β -HgS, HgSe, and HgTe. *Phys. Rev. B* **2011**, *84*, 205205.
- (104) Cardona, M.; Kremer, R. K.; Lauck, R.; Siegle, G.; Muñoz, A.; Romero, A. H. Electronic, vibrational, and thermodynamic properties of metacinnabar β -HgS, HgSe, and HgTe. *Phys. Rev. B* **2009**, *80*, 195204.
- (105) Hao, A.; Gao, C.; Li, M.; He, C.; Huang, X.; Zhang, D.; Yu, C.; Hongwu Liu.; Ma, Y.; Tian, Y.; Zou, G. A study of the electrical properties of HgS under high pressure. *J. Phys.: Condens. Matter* **2007**, *19*, 425222.

- (106) Dybko, K.; Szuszkiewicz, W.; Dynowska, E.; Paszkowicz, W.; Witkowska, B. Band structure of β -HgS from Shubnikov-de Haas effect. *Physica B* **1998**, *256-258*, 629–632.
- (107) Zallen, R.; Slade, M. Plasma edge and band structure of cubic HgS. *Solid State Commun.* **1970**, *8*, 1291–1294.
- (108) Kojima, A.; Teshima, K.; Shirai, Y.; Miyasaka, T. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. *J. Am. Chem. Soc.* **2009**, *131*, 6050–6051.
- (109) Zhou, H.; Chen, Q.; Li, G.; Luo, S.; Song, T.-b.; Duan, H.-S.; Hong, Z.; You, J.; Liu, Y.; Yang, Y. Interface engineering of highly efficient perovskite solar cells. *Science* **2014**, *345*, 542–546.
- (110) Brivio, F.; Butler, K. T.; Walsh, A.; van Schilfgaarde, M. Relativistic quasiparticle self-consistent electronic structure of hybrid halide perovskite photovoltaic absorbers. *Phys. Rev. B* **2014**, *89*, 155204.
- (111) Motta, C.; El-Mellouhi, F.; Kais, S.; Tabet, N.; Alharbi, F.; Sanvito, S. Revealing the role of organic cations in hybrid halide perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$. *Nat. Commun.* **2015**, *6*, 7026.
- (112) Lee, J. H.; Lee, J.-H.; Kong, E.-H.; Jang, H. M. The nature of hydrogen-bonding interaction in the prototypic hybrid halide perovskite, tetragonal $\text{CH}_3\text{NH}_3\text{PbI}_3$. *Sci. Rep.* **2016**, *6*, 21687.
- (113) Gao, W.; Gao, X.; Abtew, T. A.; Sun, Y.-Y.; Zhang, S.; Zhang, P. Quasiparticle band gap of organic-inorganic hybrid perovskites: Crystal structure, spin-orbit coupling, and self-energy effects. *Phys. Rev. B* **2016**, *93*, 085202.

- (114) Baikie, T.; Fang, Y.; Kadro, J. M.; Schreyer, M.; Wei, F.; Mhaisalkar, S. G.; Graetzel, M.; White, T. J. Synthesis and crystal chemistry of the hybrid perovskite (CH₃NH₃)PbI₃ for solid-state sensitised solar cell applications. *J. Mater. Chem. A* **2013**, *1*, 5628–5641.
- (115) Kim, H.-S.; Lee, C.-R.; Im, J.-H.; Lee, K.-B.; Moehl, T.; Marchioro, A.; Moon, S.-J.; Humphry-Baker, R.; Yum, J.-H.; Moser, J. E.; Grätzel, M.; Park, N.-G. Lead Iodide Perovskite Sensitized All-Solid-State Submicron Thin Film Mesoscopic Solar Cell with Efficiency Exceeding 9%. *Sci. Rep.* **2012**, *2*, 591.
- (116) Stoumpos, C. C.; Malliakas, C. D.; Kanatzidis, M. G. Semiconducting Tin and Lead Iodide Perovskites with Organic Cations: Phase Transitions, High Mobilities, and Near-Infrared Photoluminescent Properties. *Inorg. Chem.* **2013**, *52*, 9019–9038.
- (117) D’Innocenzo, V.; Grancini, G.; Alcocer, M. J. P.; Kandada, A. R. S.; Stranks, S. D.; Lee, M. M.; Lanzani, G.; Snaith, H. J.; Petrozza, A. Excitons versus free charges in organo-lead tri-halide perovskites. *Nat. Commun.* **2014**, *5*, 3586.
- (118) Yamada, Y.; Nakamura, T.; Endo, M.; Wakamiya, A.; Kanemitsu, Y. Near-band-edge optical responses of solution-processed organic-inorganic hybrid perovskite CH₃NH₃PbI₃ on mesoporous TiO₂ electrodes. *Appl. Phys. Express* **2014**, *7*, 032302.
- (119) Tanaka, K.; Takahashi, T.; Ban, T.; Kondo, T.; Uchida, K.; Miura, N. Comparative study on the excitons in lead-halide-based perovskite-type crystals CH₃NH₃PbBr₃ CH₃NH₃PbI₃. *Solid State Commun.* **2003**, *127*, 619–623.
- (120) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.
- (121) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641–3662.

- (122) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* **2005**, *114*, 145–152.
- (123) Goedecker-Teter-Hutter (GTH) pseudopotentials. <http://cp2k.web.psi.ch/potentials/>, (accessed May 20, 2016).
- (124) Chuang, C.-H. M.; Brown, P. R.; Bulović, V.; Bawendi, M. G. Improved performance and stability in quantum dot solar cells through band alignment engineering. *Nat. Mater.* **2014**, *13*, 796–801.
- (125) Fernando, A.; Weerawardene, K. L. D. M.; Karimova, N. V.; Aikens, C. M. Quantum Mechanical Studies of Large Metal, Metal Oxide, and Metal Chalcogenide Nanoparticles and Clusters. *Chem. Rev.* **2015**, *115*, 6112–6216.
- (126) Kim, D.; Kim, D.-H.; Lee, J.-H.; Grossman, J. C. Impact of Stoichiometry on the Electronic Structure of PbS Quantum Dots. *Phys. Rev. Lett.* **2013**, *110*, 196802.
- (127) Gai, Y.; Peng, H.; Li, J. Electronic Properties of Nonstoichiometric PbSe Quantum Dots from First Principles. *J. Phys. Chem. C* **2009**, *113*, 21506–21511.
- (128) Grassi, F.; Argeri, M.; Marchese, L.; Cossi, M. First Principle Study of Capping Energies and Electronic States in Stoichiometric and Nonstoichiometric PbSe Nanoclusters. *J. Phys. Chem. C* **2013**, *117*, 26396–26404.
- (129) Moreels, I.; Lambert, K.; De Muynck, D.; Vanhaecke, F.; Poelman, D.; Martins, J. C.; Allan, G.; Hens, Z. Composition and Size-Dependent Extinction Coefficient of Colloidal PbSe Quantum Dots. *Chem. Mater.* **2007**, *19*, 6101–6106.
- (130) Fang, C.; van Huis, M. A.; Vanmaekelbergh, D.; Zandbergen, H. W. Energetics of Polar and Nonpolar Facets of PbSe Nanocrystals from Theory and Experiment. *ACS Nano* **2010**, *4*, 211–218.

- (131) Schapotschnikow, P.; van Huis, M. A.; Zandbergen, H. W.; Vanmaekelbergh, D.; Vlugt, T. J. H. Morphological Transformations and Fusion of PbSe Nanocrystals Studied Using Atomistic Simulations. *Nano Lett.* **2010**, *10*, 3966–3971.
- (132) Argeri, M.; Fraccarollo, A.; Grassi, F.; Marchese, L.; Cossi, M. Density Functional Theory Modeling of PbSe Nanoclusters: Effect of Surface Passivation on Shape and Composition. *J. Phys. Chem. C* **2011**, *115*, 11382–11389.
- (133) Dal Corso, A. Pseudopotentials periodic table: From H to Pu. *Comput. Mater. Sci.* **2014**, *95*, 337–350.
- (134) Kaushik, A. P.; Lukose, B.; Clancy, P. The Role of Shape on Electronic Structure and Charge Transport in Faceted PbSe Nanocrystals. *ACS Nano* **2014**, *8*, 2302–2317.
- (135) Fleszar, A.; Hanke, W. Electronic structure of II^B-VI semiconductors in the GW approximation. *Phys. Rev. B* **2005**, *71*, 045207.
- (136) Rangel, T.; Kekik, D.; Trevisanutto, P. E.; Rignanese, G.-M.; Van Swygenhoven, H.; Olevano, V. Band structure of gold from many-body perturbation theory. *Phys. Rev. B* **2012**, *86*, 125125.
- (137) Jiang, H.; Shen, Y.-C. Ionization potentials of semiconductors from first-principles. *J. Chem. Phys.* **2013**, *139*, 164114.
- (138) Grüneis, A.; Kresse, G.; Hinuma, Y.; Oba, F. Ionization Potentials of Solids: The Importance of Vertex Corrections. *Phys. Rev. Lett.* **2014**, *112*, 096401.
- (139) Hinuma, Y.; Grüneis, A.; Kresse, G.; Oba, F. Band alignment of semiconductors from density-functional theory and many-body perturbation theory. *Phys. Rev. B* **2014**, *90*, 155405.
- (140) Li, Z.; Xiao, Y.; Liu, W. On the spin separation of algebraic two-component relativistic Hamiltonians. *J. Chem. Phys.* **2012**, *137*, 154114.

- (141) Cheng, L.; Gauss, J. Perturbative treatment of spin-orbit coupling within spin-free exact two-component theory. *J. Chem. Phys.* **2014**, *141*, 164107.
- (142) Lundqvist, B. I. Single-particle spectrum of the degenerate electron gas. *Phys. Kondens. Mater.* **1967**, *6*, 193–205.
- (143) Mahan, G. D. *Many-particle physics*, 2nd ed.; Physics of solids and liquids; Plenum Press, 1990.
- (144) Momma, K.; Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **2011**, *44*, 1272–1276.
- (145) Rohlfing, M.; Krüger, P.; Pollmann, J. Quasiparticle Band Structure of CdS. *Phys. Rev. Lett.* **1995**, *75*, 3489–3492.
- (146) Shishkin, M.; Kresse, G. Implementation and performance of the frequency-dependent *GW* method within the PAW framework. *Phys. Rev. B* **2006**, *74*, 035101.
- (147) Luo, W.; Ismail-Beigi, S.; Cohen, M. L.; Louie, S. G. Quasiparticle band structure of ZnS and ZnSe. *Phys. Rev. B* **2002**, *66*, 195215.
- (148) Gori, P.; Rakel, M.; Cobet, C.; Richter, W.; Esser, N.; Hoffmann, A.; Del Sole, R.; Cricenti, A.; Pulci, O. Optical spectra of ZnO in the far ultraviolet: First-principles calculations and ellipsometric measurements. *Phys. Rev. B* **2010**, *81*, 125207.
- (149) Marini, A.; Onida, G.; Del Sole, R. Quasiparticle Electronic Structure of Copper in the *GW* Approximation. *Phys. Rev. Lett.* **2001**, *88*, 016403.
- (150) Filip, M. R.; Patrick, C. E.; Giustino, F. *GW* quasiparticle band structures of stibnite, antimonelite, bismuthinite, and guanajuatite. *Phys. Rev. B* **2013**, *87*, 205125.
- (151) Umari, P.; Fabris, S. Importance of semicore states in *GW* calculations for simulating accurately the photoemission spectra of metal phthalocyanine molecules. *J. Chem. Phys.* **2012**, *136*, 174310.

- (152) Qteish, A.; Al-Sharif, A. I.; Fuchs, M.; Scheffler, M.; Boeck, S.; Neugebauer, J. Role of semicore states in the electronic structure of group-III nitrides: An exact-exchange study. *Phys. Rev. B* **2005**, *72*, 155317.
- (153) Rinke, P.; Qteish, A.; Neugebauer, J.; Freysoldt, C.; Scheffler, M. Combining GW calculations with exact-exchange density-functional theory: an analysis of valence-band photoemission for compound semiconductors. *New J. Phys.* **2005**, *7*, 126.
- (154) Zanolli, Z.; Fuchs, F.; Furthmüller, J.; von Barth, U.; Bechstedt, F. Model GW band structure of InAs and GaAs in the wurtzite phase. *Phys. Rev. B* **2007**, *75*, 245121.
- (155) Kim, Y.-S.; Marsman, M.; Kresse, G.; Tran, F.; Blaha, P. Towards efficient band structure and effective mass calculations for III-V direct band-gap semiconductors. *Phys. Rev. B* **2010**, *82*, 205212.
- (156) Govoni, M.; Marri, I.; Ossicini, S. Auger recombination in Si and GaAs semiconductors: *Ab initio* results. *Phys. Rev. B* **2011**, *84*, 075215.
- (157) Haynes, W. M., Ed. *CRC Handbook of Chemistry and Physics, 96th Edition*; CRC Press, 2015.
- (158) van Setten, M. J.; Giantomassi, M. PseudoDojo - Abinit. <http://www.abinit.org/downloads/pseudodojo/pseudodojo>, (accessed Apr 8, 2016).
- (159) Ceresoli, D. Pseudopotentials. <https://sites.google.com/site/dceresoli/pseudopotentials>, (accessed May 20, 2016).
- (160) Klimeš, J.; Kaltak, M.; Kresse, G. Predictive GW calculations using plane waves and pseudopotentials. *Phys. Rev. B* **2014**, *90*, 075125.

Graphical TOC Entry

The diagram shows a standard periodic table with elements color-coded by groups: Group 1 (H, Li, Na, K, Rb, Cs) is light blue; Groups 2-10 (Be, Mg, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr) are light green; Groups 11-18 (B, C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr) are light red. A large, stylized cloud labeled 'W' is positioned above the central part of the table. A thick black arrow labeled 'G' points from left to right, passing through the center of the cloud, indicating the direction of the strong interaction.