Resolution of the Band Gap Prediction Problem for Materials Design

Thesis by

Jason M. Crowley

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2016

Defended 26 May 2016

ACKNOWLEDGEMENTS

ABSTRACT

An important property with any new material is the band gap. In order to design new materials in silico, it is critical to have an accurate and computationally inexpensive tool for predicting band gaps. Standard density functional theory (DFT) methods are computationally efficient, but grossly underestimate band gaps. Hybrid density functionals are known to improve band gap predictions, but the computational cost in the overwhelmingly popular plane-wave basis set codes used for solids is a serious drawback. Exact exchange can be evaluated much more efficiently using localized Gaussian basis functions; however, the most readily available Gaussian basis periodic quantum chemistry code lacked spin-orbit coupling. This seriously limited the range of compounds that can be studies. In this thesis, spin-orbit coupling was implemented in the periodic, Gaussian basis set code CRYSTAL. Using the modified code, band gaps were computed using the B3PW91 hybrid density functional for 70 compounds spanning the entire periodic table and a factor of 500 in band gap (0.014 - 15 eV). To benchmark the quality of the hybrid method, we compared to the rigorous GW many-body perturbation theory method. Surprisingly, the MAD for B3PW91 is about 1.5 times smaller than the MAD for GW. Furthermore, B3PW91 is three to four orders of magnitude faster computationally. We also show that increasing (decreasing) the amount of exact exchange compared to B3PW91 leads to systematic overestimates (underestimates) of band gaps. Finally, we show that the pathological vanishing of the density of states at the Fermi level of a metal cannot be observed in practical calculations of real metals. Thus, we believe that B3PW91 is a practical tool for predicting the band gaps of materials before they are synthesized while being computationally efficient enough for high-throughput applications and represents a solution to the band gap prediction problem for materials design.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] J. M. Crowley, J. Tahir-Kheli, and W. A. Goddard, "Accurate ab initio quantum mechanics simulations of Bi_2Se_3 and Bi_2Te_3 topological insulator surfaces", J. Phys. Chem. Lett. **6**, 3792–3796 (2015).

[2] J. M. Crowley, J. Tahir-Kheli, and I. William A. Goddard, "Resolution of the band gap prediction problem for materials design", J. Phys. Chem. Lett. **7**, 1198–1203 (2016).

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

C h a p t e r   1

INTRODUCTION

Increases in computing power, improved algorithms, and the development of user-friendly software packages have made computational methods in chemistry and materials science nearly as ubiquitous as standard experimental techniques like NMR. The ultimate goal of computation in chemistry is the reliable prediction of materials before they are synthesized to enable more efficient use of experimental resources. Projects like the Materials Genome Initiative seek to leverage "big data" techniques to realize this goal.

An especially important property for designing new materials is the band gap. A fundamental question about a material is whether it is a metal, an insulator, or a semiconductor, which is the question of the magnitude of the band gap. In order to design new materials for applications like solar energy, it is crucial to have a computational method that can accurately and quickly predict band gaps for materials across the entire periodic table. This is the band gap prediction problem for materials design.

The rigorous approach to computing band gaps is to calculate the Green's function using many-body perturbation theory (referred to as the *GW* method [1]). The rigorous quasiparticle excitation spectrum (giving the exact band gap) is calculated in this method. Unfortunately, it is far too computationally expensive for most systems, and requires non-self-consistent approximations.

It is very well-known [2] that hybrid density functionals, which include a fraction of exact Hartree-Fock exchange, are more accurate for band gaps than local density functionals. However, in plane-wave basis set codes (which are by far the most commonly used), the computational cost of evaluating the exact exchange is a serious drawback. Localized Gaussian basis sets enable much more efficient evaluation of Hartree-Fock exchange.

The most commonly used Gaussian basis set code for solids, CRYSTAL [3], does not include spin-orbit coupling. This limits its applicability for band gap prediction to compounds containing only light elements. In this work, CRYSTAL was extended to include spin-orbit coupling, enabling band gap calculations using hybrid functionals across the entire periodic for a fraction of the computational cost

of plane wave basis set codes. Band gap calculations for 70 compounds with experimental band gaps ranging from 0–15 eV were performed using the B3PW91 hybrid density functional in the modified code. Comparison of these calculations to literature *GW* results shows that the B3PW91 hybrid density functional is more accurate than the theoretically rigorous *GW* (many-body perturbation) method. When used with a Gaussian basis set, hybrid density functional calculations are about 3–4 orders of magnitude faster than *GW* calculations. Across this wide array of compounds, B3PW91 is accurate enough to enable the prediction of band gaps for materials that have not been synthesized. This accuracy and computational speed make Gaussian basis set B3PW91 an ideal choice for high-throughput materials design applications, and constitutes a solution to the band gap prediction problem for materials design. This constitutes the main result of the thesis.

The remainder of the thesis is organized as follows. Chapter 2 describes the implementation of spin-orbit coupling in CRYSTAL. Chapter 3 presents the comparison of B3PW91 to *GW*. In Chapter 4, B3PW91 is systematically compared to five popular density functionals. It is shown that functionals with more (less) exact exchange than B3PW91 systematically predict higher (lower) band gaps. Comparison of B3LYP to B3PW91 shows that the correlation functional has a smaller and not systematic effect. Chapter 5 shows that the oft-criticised "pathology" of long-range exchange (zero density of states at the Fermi level for metals) cannot be observed in B3PW calculations on real metals. Finally, Chapter 6 shows the application of B3PW91 to modelling two-dimensional slabs of the topological insulators Bi_2Te_3 and Bi_2Se_3.

Chapter 2

IMPLEMENTATION OF SPIN-ORBIT COUPLING IN CRYSTAL98

In the nonrelativistic Schrödinger equation, the Hamiltonian does not operate on the electron spin. Because the spin functions α and β are orthonormal, only matrix elements between functions with the same spin can be nonzero. Consequently, separate Fock matrices for α and β spins can be constructed in unrestricted Hartree-Fock (UHF) calculations without spin-orbit coupling. The α and β Fock matrices each have $N \times N$ elements, where N is the number of basis functions.

Spin-orbit coupling introduces explicit spin dependence into the Hamiltonian, and electron spin is not a good quantum number. An orbital cannot necessarily be assigned an α or β spin; instead, the spin part of an orbital is some linear combination of α and β. The wavefunction then takes the form

$$\Psi = \phi_\alpha \alpha + \phi_\beta \beta, \qquad (2.1)$$

and "off-diagonal" terms like $\langle \phi_\alpha \alpha | H | \phi_\beta \beta \rangle$ must be considered. As in UHF calculations, ϕ_α and ϕ_β are each represented as a linear combination of N basis functions. Because off-diagonal matrix elements can be nonzero, we must diagonalize a single $2N \times 2N$ matrix instead of two $N \times N$ matrices as in UHF.

In this chapter, the implementation of spin-orbit coupling in CRYSTAL98 [4], a Gaussian basis set periodic quantum chemistry program, is discussed. CRYSTAL98 was used because the source code was distributed for this version, but for no later versions. This modified code was used for all spin-orbit calculations presented in this thesis. As discussed in Section 2.1, spin-orbit coupling can be fully included implicitly by pseudopotentials when the implications of equation 2.1 are taken into account. Equation 2.1 also necessitates modifications in the treatment of symmetry (Section 2.2), the density matrix (Section 2.3), and the exchange operator (Section 2.4). The all-electron spin-orbit operator is discussed in Section 2.5. However, because the fully relativistic pseudopotentials used in all calculations presented here fully account for spin-orbit coupling, this operator is not used in the code.

2.1 Spin-Orbit Coupling in Pseudopotentials

It is well-established that the valence electrons are most relevant to chemical bonding. In fact, this intuitive picture is the conceptual basis of the periodic table and is routinely taught to high school chemistry students. From the earliest days of quantum chemistry, this intuition has been leveraged to simplify calculations. The basic idea is to replace the Schrödinger equation for the full many-electron system with an equation modified to explicitly treat only a subset of the electrons by adding an effective potential, also known as a pseudopotential.

This approach began with the work of Hellmann [5] and Gombás [6] in 1935. In 1940, Fock, Vesselow, and Petraschen [7] considered the theoretical treatment of two electrons outside a closed shell. The main challenge is in maintaining orthogonality of the valence orbitals to the core orbitals. Explicitly orthogonalizing the valence orbitals would require knowledge of the exact core states, which is the same as solving the Schrödinger equation for the full system. The pseudopotential concept was given sound theoretical basis by Phillips and Kleinman [8], whose work was generalized to many-electron systems by Weeks and Rice [9]. The key result was that the variational optimization of the ground state energy, subject to the constraint that the valence eigenfunctions must be orthogonal to the core states, can be replaced by an *unconstrained* optimization of the ground state energy of a modified Hamiltonian that includes a pseudopotential. In other words, the orthogonality condition can be implicitly included via an effective potential.

In the ensuing years, a number of empirical "model potentials" were developed [10–13]. These model potentials were developed by choosing some plausible functional form for the effective potential and fitting adjustable parameters to experimental data. A more systematic, ab initio, pseudopotential was developed by Goddard and coworkers [14–18] and summarized by Kahn, Baybutt, and Truhlar [19].

Pitzer and coworkers [20–24] showed that not only do pseudopotentials substantially decrease the computational cost of calculations, but they can also be used to implicitly include relativistic effects in non-relativistic calculations. Relativistic effects can be very large (on the order of an electron-volt) for the core orbitals of heavy elements. For the valence orbitals, the matrix elements of the relativistic operators are typically negligible. However, relativistic effects manifest themselves in valence states via orthogonality to the core states, where relativity has a large impact. For this reason, pseudopotentials can be used to include relativistic effects in non-relativistic calculations.

The pseudopotentials of Lee et al. [20], as well as those of Hay and Wadt [25] and Stevens et al. [26], follow the approach pioneered by Kahn, Melius, and Goddard. All of these pseudopotentials use the eigenfunctions from atomic Hartree-Fock (or Dirac-Hartree-Fock for heavy elements) calculations to derive the pseudopotential. An alternative approach is to fit the pseudopotential parameters so that the total energies of atomic pseudopotential calculations match those of reference all-electron calculations. These so-called "energy-adjusted pseudopotentials" [27–34] have been recently reviewed by Dolg and Cao [35]. In this method, a set of atomic configurations of neutral and ions with small charges that is expected to include the most relevant configurations for molecules and solids is chosen. These energy-adjusted pseudopotentials are used in all pseudopotential calculations in this thesis.

The pseudopotential operator is written as a sum of angular-momentum-dependent potentials multiplied by angular momentum projections. This ensures orthogonality of valence and core states. The angular-momentum-dependent potentials are fit to a sum of Gaussian functions where the kth Gaussian for angular momentum l has coefficient c_{lk} and exponent α_{lk}. The pseudopotential then takes the form

$$V_{PP}(r) = -\frac{Z_{eff}}{r} + \sum_{l=0}^{L} \sum_{k} c_{lk} \exp(-\alpha_{lk}\vec{r}^2) \sum_{m=-l}^{l} |lm\rangle\langle lm|, \qquad (2.2)$$

where Z_{eff} is the effective charge of a nucleus and L is the largest angular momentum in the core. The angular momentum eigenfunction $|lm\rangle$ is the spherical harmonic Y_{lm}, and the projections $|lm\rangle\langle lm|$ ensure orthogonality of the valence and core states.

Pitzer and coworkers [20–24] showed that pseudopotentials can also be used to implicitly include relativistic effects such as spin-orbit coupling in non-relativistic calculations. When spin-orbit coupling is included, we must consider the total angular momentum $j = l + s$ instead of separate orbital and spin angular momenta l and s. Thus, in place of the projections onto orbital angular momentum eigenfunctions in equation 2.2, we must use projections onto total angular momentum eigenfunctions $|lsjm_j\rangle$, with $s = \frac{1}{2}$. Thus, the pseudopotential becomes

$$V_{PP}(r) = -\frac{Z_{eff}}{r} + \sum_{l=0}^{L} \sum_{j=|l-\frac{1}{2}|}^{l+\frac{1}{2}} \sum_{k} c_{jk} \exp(-\alpha_{jk}\vec{r}^2) \sum_{m_j=-j}^{j} |lsjm_j\rangle\langle lsjm_j|.$$

This can be rewritten in terms of projections onto $|lm\rangle$ via the Clebsch-Gordan

coefficients [36]:

$$V_{PP}(r) = -\frac{Z_{eff}}{r} + \sum_{l=0}^{L} \sum_{j,k} c_{jk} \exp(-\alpha_{jk} \vec{r}^2)$$
$$\times \sum_{\substack{m_j m_l m_s \\ m'_j m'_l m'_s}} \langle lsjm_j|lsm'_l m'_s\rangle\langle lsm_l m_s|lsjm_j\rangle|lm_l\rangle|sm_s\rangle\langle lm'_l|\langle sm'_s|. \quad (2.3)$$

Carrying out the sums over m_j, m_l, and m_s in equation 2.3 reduces matrix elements of the fully relativistic pseudopotential operator to linear combinations of nonrelativistic pseudopotential matrix elements. The evaluation of nonrelativistic pseudopotential matrix elements was discussed by McMurchie and Davidson [37]. It is important to note that McMurchie and Davidson considered real spherical harmonics. The Clebsch-Gordan coefficients determine a total angular momentum eigenfunction as a linear combination of complex spherical harmonics, and the orbital angular momentum eigenfunctions in equation 2.3 are complex spherical harmonics. Using the formulae in reference [37] requires writing the complex spherical harmonics in terms of real spherical harmonics, which is straightforward.

The highly optimized nature of the pseudopotential subroutines in CRYSTAL made it impractical to "hack" them to compute j-dependent potentials. Therefore, the pseudopotential equations were implemented by the author in a from-scratch subroutine. My subroutine occupies a fraction of the number of lines of the original code, but runs approximately five times slower! Where the original CRYSTAL code uses every trick in the book to evaluate the pseudopotential integrals, my code uses a simple trapezoid rule numerical integration for the radial integral (the angular integrals are much cheaper).

Since $s = \frac{1}{2}$ and $j = l + s$, the possible values of j are $l + \frac{1}{2}$ and $l - \frac{1}{2}$. The fully relativistic energy-consistent pseudopotentials of the Stuttgart group [35] used in this thesis have different parameters for the two values of j corresponding to each l. These parameters are determined from all-electron calculations based on four-component Dirac-Hartree-Fock calculations, and therefore contain all the effects of spin-orbit coupling. Thus, the projections of the valence states onto the potentials for $j = l \pm \frac{1}{2}$ account for the spin-orbit coupling. For example, a p orbital can have $j = \frac{3}{2}$ or $j = \frac{1}{2}$. In a nonrelativistic calculation, these are degenerate. By contrast, when spin-orbit coupling is included, this degeneracy is lifted. Instead of a six-fold degenerate $l = 1$ state, the result is a four-fold degenerate $j = \frac{3}{2}$ state and a two-

fold degenerate $j = \frac{1}{2}$ state. The splitting between these states is determined by the j-dependent potential.

The magnitude of spin-orbit coupling scales as $1/r^3$, where r is the distance from the nucleus, and Z^2, where Z is the nuclear charge. Clearly, spin-orbit coupling is much larger for core electrons than valence electrons, since valence electrons are further from the nucleus than core electrons. Furthermore, valence electrons are shielded from the full nuclear charge by core electrons. Therefore, the effect of spin-orbit coupling on valence electrons is indirect, stemming from orthogonality to core states rather than the direct effect of the spin-orbit operator.

2.2 Symmetry

For computational efficiency, three types of symmetry are used: the crystal symmetry, the Hermitian property of operators, and time reversal symmetry.

Crystal Symmetry

A space group operation rotates a basis function to a linear combination of other basis functions of the same angular momentum on another atom in the crystal (possibly in another unit cell). For instance, a rotation around the z-axis may take a p_x orbital on atom A to a linear combination of p_x and p_y orbitals on atom B. With spin-orbit coupling, the Hamiltonian depends on the electron spin. Because the Hamiltonian must have the full space group symmetry of the crystal, this means that space group operators must apply to electron spins. Thus, a symmetry operation takes an α spin on atom A to a linear combination of α and β spins on atom B. For symmetry operation O,

$$
\begin{aligned}
O\alpha &= a_1\alpha + b_1\beta \\
O\beta &= a_2\alpha + b_2\beta.
\end{aligned}
$$

A general spin $\sigma = u\alpha + v\beta$, which can be written as $\begin{pmatrix} u \\ v \end{pmatrix}$. Operation O takes σ to $\sigma' = \begin{pmatrix} u' \\ v' \end{pmatrix}$. Thus,

$$
O\sigma = O\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u' \\ v' \end{pmatrix},
$$

which can be rewritten as

$$
\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u' \\ v' \end{pmatrix}
$$

The effect of O on the electron spin is thus given by the unitary 2×2 matrix corresponding to the 3×3 spatial rotation by the homomorphism $SO(3) \to SU(2)$. The rotation corresponding to O can be represented in terms of Euler angles α, β, γ. Following the convention used by Tinkham [38], the rotation is given by $R_z(\alpha)R_y(\beta)R_z(\gamma)$, where $R_z(\alpha)$ represents rotation by α around the z-axis. Here, rotation by α will be taken to be counterclockwise (note that Tinkham uses clockwise rotations). In this convention,

$$R_z(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$R_y(\beta) = \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 0 & 1 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix}.$$

Thus, a general rotation is given by

$$R_z(\alpha)R_y(\beta)R_z(\gamma) =$$
$$\begin{pmatrix} \cos(\alpha)\cos(\gamma)\cos(\beta) - \sin(\alpha)\sin(\gamma) & -\cos(\alpha)\cos(\beta)\sin(\gamma) - \sin(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\beta) \\ \sin(\alpha)\cos(\gamma)\cos(\beta) + \cos(\alpha)\sin(\gamma) & -\sin(\alpha)\cos(\beta)\sin(\gamma) + \cos(\alpha)\cos(\gamma) & \sin(\alpha)\sin(\beta) \\ -\sin(\beta)\cos(\gamma) & \sin(\beta)\sin(\gamma) & \cos(\beta) \end{pmatrix}.$$

The corresponding $SU(2)$ matrix is

$$U = \begin{pmatrix} e^{-i(\gamma+\alpha)/2}\cos\left(\frac{\beta}{2}\right) & -e^{i(\gamma-\alpha)/2}\sin\left(\frac{\beta}{2}\right) \\ e^{-i(\gamma-\alpha)/2}\sin\left(\frac{\beta}{2}\right) & e^{i(\gamma+\alpha)/2}\cos\left(\frac{\beta}{2}\right) \end{pmatrix}.$$

For a Fock matrix element,

$$O\langle\phi\sigma|F|\phi'\sigma'\rangle = \langle\tilde{\phi}(a\alpha + b\beta)|F|\tilde{\phi}'(a'\alpha + b'\beta)\rangle$$
$$= \bar{a}a'\langle\tilde{\phi}\alpha|F|\tilde{\phi}'\alpha\rangle + \bar{a}b'\langle\tilde{\phi}\alpha|F|\tilde{\phi}'\beta\rangle + \bar{b}a'\langle\tilde{\phi}\beta|F|\tilde{\phi}'\alpha\rangle + \bar{b}b'\langle\tilde{\phi}\beta|F|\tilde{\phi}'\beta\rangle.$$

Thus, the tensor product $U \otimes U$ generates $\langle\phi\alpha|F|\phi'\alpha\rangle$, $\langle\phi\alpha|F|\phi'\beta\rangle$, $\langle\phi\beta|F|\phi'\alpha\rangle$, and $\langle\phi\beta|F|\phi'\beta\rangle$ from $\langle\tilde{\phi}\alpha|F|\tilde{\phi}'\alpha\rangle$, $\langle\tilde{\phi}\alpha|F|\tilde{\phi}'\beta\rangle$, $\langle\tilde{\phi}\beta|F|\tilde{\phi}'\alpha\rangle$, and $\langle\tilde{\phi}\beta|F|\tilde{\phi}'\beta\rangle$.

2.3 Density Matrix

Using (2.1), the density operator becomes

$$
\begin{aligned}
P & = |\Psi\rangle\langle\Psi| \\
& = (|\phi_\alpha\rangle|\alpha\rangle + |\phi_\beta\rangle|\beta\rangle)(\langle\phi_\alpha|\langle\alpha| + \langle\phi_\beta|\langle\beta|) \\
& = |\phi_\alpha\rangle\langle\phi_\alpha| \cdot |\alpha\rangle\langle\alpha| + |\phi_\beta\rangle\langle\phi_\beta| \cdot |\beta\rangle\langle\beta| \\
& + |\phi_\alpha\rangle\langle\phi_\beta| \cdot |\alpha\rangle\langle\beta| + |\phi_\beta\rangle\langle\phi_\alpha| \cdot |\beta\rangle\langle\alpha| \\
& \equiv P^{\alpha\alpha} + P^{\beta\beta} + P^{\alpha\beta} + P^{\beta\alpha}.
\end{aligned}
$$

As with the pseudopotential and exchange operators, we must add contributions from alpha-beta and beta-alpha. In general, we need density matrix elements of the form

$$
P_{\mu\nu}^{\sigma\sigma'}(k) = \sum_{\mu\nu} c_\mu^\sigma(k)\overline{c_\nu^{\sigma'}(k)},
$$

where σ and σ' can be either α or β. Diagonalization of the Fock matrix at a given k-point returns eigenvectors composed of $c_\mu^\sigma(k)$, the complex weight of basis function μ with spin σ to the wavefunction at k.

The first N components of an eigenvector are the $c_\mu^\alpha(k)$; the second N components are the $c_\mu^\beta(k)$. In a UHF code, these can be treated independently. The code simply reads the eigenvector for α, makes the density matrix P^α, and then repeats the process for β. However, with spin-orbit coupling, the α and β contributions to an eigenvector must be mixed to compute $P^{\alpha\beta}$ and $P^{\beta\alpha}$. Thus, the code must read both the α and β pieces, then compute all four blocks of the density matrix.

The code computes the Fock and density matrices for an irreducible wedge of the Brillouin zone and uses the space group symmetry to sum over the Brillouin zone and form the real space density matrix. In addition to the space group symmetry, the original code also used a combination of time reversal symmetry and the commutation of the Hamiltonian with S_z to further reduce the number of k-points it evaluated. Time reversal symmetry says

$$
\langle\Psi|O|\Phi\rangle = \langle T\Phi|O|T\Psi\rangle,
$$

where T is the anti-unitary time reversal operator. For a periodic system, we have

$$
\langle\Psi\mathbf{0}|O|\Phi\mathbf{R}\rangle = \langle T\Phi\mathbf{R}|O|T\Psi\mathbf{0}\rangle = \langle T\Psi\mathbf{0}|O|T\Phi{-}\mathbf{R}\rangle. \tag{2.4}
$$

The time reversal operator has no effect on the real space vector $\mathbf{R}$. However, time reversal negates momentum, so in k-space we have

$$
\langle\Psi k|O|\Phi k\rangle = \langle T\Phi - k|O|T\Psi - k\rangle. \tag{2.5}
$$

The effect of T on the spin functions α and β is,

$$T\alpha = \beta$$
$$T\beta = -\alpha.$$

Without spin-orbit, these symmetries allow the code to assume $P_{\mu\nu}(-k) = \overline{P_{\mu\nu}(k)}$. Since the density matrix is transformed to real space, the code can simply add the contributions from k and $-k$ with the appropriate phase factors. These are complex conjugates, and the code simply keeps $2Re(P_{\mu\nu}(k))$.

With spin-orbit coupling, the Hamiltonian does not commute with S_z, and we can only use time reversal symmetry and the Hermitian property of the Fock operator. For a Hermitian operator O and states Ψ, Φ,

$$\langle \Psi|O|\Phi \rangle = \overline{\langle \Phi|O|\Psi \rangle}.$$

In a system with translational symmetry, a basis function is associated with a unit cell $\mathbf{R}$. Because of the translational symmetry, we choose a unit cell $\mathbf{0}$ all other cells are referred to. Then,

$$\langle \Psi\mathbf{0}|O|\Phi\mathbf{R} \rangle = \overline{\langle \Phi\mathbf{R}|O|\Psi\mathbf{0} \rangle},$$

or, translating the entire system by -$\mathbf{R}$,

$$\langle \Psi\mathbf{0}|O|\Phi\mathbf{R} \rangle = \overline{\langle \Phi\mathbf{0}|O|\Psi-\mathbf{R} \rangle}. \tag{2.6}$$

In k-space,

$$\langle \Psi k|O|\Phi k \rangle = \overline{\langle \Phi k|O|\Psi k \rangle}. \tag{2.7}$$

Together, time reversal symmetry and Hermiticity tell us $P_{\mu\nu}^{\alpha\beta}(k) = -\overline{P_{\mu\nu}^{\beta\alpha}(-k)}$. Use of these symmetries allows the generation of the density matrix at $-k$ from that at $+k$. However, because of the assumption of time reversal symmetry, the modified code is only applicable to time-reversal-invariant systems. Consequently, the modified code cannot treat open-shell systems such as magnets.

2.4 Exchange Integrals

Let χ represent a general spin function, and $\mu_i(R_i)$ basis function i in unit cell R_i. The Fock matrix contains the Coulomb term (using chemist's notation [39])

$$J_{\mu_1\mu_2}(R_2) = \sum_{\mu_3\mu_4} \sum_{R_3R_4} P_{\mu_3\mu_4}^{\chi_3\chi_4}(R_3, R_4) \left(\mu_1(\mathbf{0})\mu_2(R_2)|\mu_3(R_3)\mu_4(R_4)\right)$$
$$\cdot (\chi_1|\chi_2)(\chi_3|\chi_4).$$

Defining $R' = R_4 - R_3$ and $R'' = R_3$, we can rewrite this as

$$J_{\mu_1\mu_2}(R_2) = \sum_{\mu_3\mu_4} \sum_{R'} P_{\mu_3\mu_4}^{\chi_3\chi_4}(R') \sum_{R''} (\mu_1(\mathbf{0})\mu_2(R_2)|\mu_3(R'')\mu_4(R'+R''))$$
$$\cdot (\chi_1|\chi_2)(\chi_3|\chi_4).$$

Now, since the spin functions are orthogonal, we know $\chi_1 = \chi_2$ and $\chi_3 = \chi_4$. Just like in standard UHF, we get terms from $\chi_3\chi_4 = \alpha\alpha$ and $\beta\beta$ for $\chi_1\chi_2 = \alpha\alpha$ and $\chi_1\chi_2 = \beta\beta$. Thus,

$$J_{\mu_1\mu_2}(R_2) = 2 \sum_{\mu_3\mu_4} \sum_{R'} P_{\mu_3\mu_4}^{\alpha\alpha}(R') \sum_{R''} (\mu_1(\mathbf{0})\mu_2(R_2)|\mu_3(R'')\mu_4(R'+R''))$$
$$+ 2 \sum_{\mu_3\mu_4} \sum_{R'} P_{\mu_3\mu_4}^{\beta\beta}(R') \sum_{R''} (\mu_1(\mathbf{0})\mu_2(R_2)|\mu_3(R'')\mu_4(R'+R''))$$

$$= 2 \sum_{\mu_3\mu_4} \sum_{R'} [P_{\mu_3\mu_4}^{\alpha\alpha}(R') + P_{\mu_3\mu_4}^{\beta\beta}(R')]I_{1234}(R'),$$

where $I_{1234}(R') \equiv \sum_{R''} (\mu_1(\mathbf{0})\mu_2(R_2)|\mu_3(R'')\mu_4(R'+R''))$.

This is the rigorous, full Coulomb term. Now, for speed we should exploit the available symmetries in the density matrix. Combining time reversal symmetry and Hermiticity, $P_{\mu_3\mu_4}^{\alpha\alpha}(R') = P_{\mu_3\mu_4}^{\beta\beta}{}^{\bar{}}(R')$, and $P_{\mu_3\mu_4}^{\alpha\alpha}(R') + P_{\mu_3\mu_4}^{\beta\beta}(R') \in \mathbb{R}$. Using these symmetries, we can write the Coulomb contribution to the Fock matrix as

$$J_{\mu_1\mu_2}(R_2) = 2 \sum_{\mu_3\geq\mu_4} \sum_{R'} \left(P_{\mu_3\mu_4}^{\alpha\alpha}(R') + P_{\mu_3\mu_4}^{\beta\beta}(R')\right) \left[I_{1234}(R') + I_{1243}(-R')\right]. \quad (2.8)$$

We can do this because R' can always be written as $-R'$ in a sum over all R', since for every R' there is a $-R'$ (taking care, of course, to handle the unit cell $R' = 0$.

The general exchange piece of the Fock matrix is

$$K_{\mu_1\mu_2}^{\chi_1\chi_2}(R_2) = - \sum_{\mu_3\mu_4} \sum_{R'} P_{\mu_3\mu_4}^{\chi_3\chi_4}(R')I_{1324}(R')(\chi_1|\chi_3)(\chi_2|\chi_4).$$

So, in contrast to the Coulomb contribution, we get contributions from $K^{\alpha\alpha}$, $K^{\beta\beta}$, $K^{\alpha\beta}$, and $K^{\beta\alpha}$.

$$K_{\mu_1\mu_2}^{\alpha\alpha}(R_2) = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\alpha\alpha}(R')I_{1324}(R')$$
$$-\sum_{\mu_3 < \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\alpha\alpha}(R')I_{1324}(R')$$

$$= -\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\alpha\alpha}(R')I_{1324}(R')$$
$$-\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_4\mu_3}^{\alpha\alpha}(-R')I_{1423}(-R')$$

We can easily show $P_{\mu_3\mu_4}^{\alpha\alpha}(R') = P_{\mu_4\mu_3}^{\alpha\alpha}(-R')^*$. It then follows that

$$Re\left\{K_{\mu_1\mu_2}^{\alpha\alpha}(R_2)\right\} = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} Re\left\{P_{\mu_3\mu_4}^{\alpha\alpha}(R')\right\}\mathbf{K}^+ \tag{2.9}$$

$$Im\left\{K_{\mu_1\mu_2}^{\alpha\alpha}(R_2)\right\} = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} Im\left\{P_{\mu_3\mu_4}^{\alpha\alpha}(R')\right\}\mathbf{K}^-, \tag{2.10}$$

where $\mathbf{K}^{\pm} = I_{1324}(R') \pm I_{1423}(-R')$.

Finally,

$$K_{\mu_1\mu_2}^{\beta\alpha}(R_2) = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\beta\alpha}(R')I_{1324}(R')$$
$$-\sum_{\mu_3 < \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\beta\alpha}(R')I_{1324}(R')$$

$$= -\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_3\mu_4}^{\beta\alpha}(R')I_{1324}(R')$$
$$-\sum_{\mu_3 \geq \mu_4}\sum_{R'} P_{\mu_4\mu_3}^{\beta\alpha}(-R')I_{1423}(-R').$$

Applying time reversal symmetry gives $P_{\mu_4\mu_3}^{\beta\alpha}(-R') = -P_{\mu_3\mu_4}^{\beta\alpha}(R')$, and it follows that

$$Re\left\{K_{\mu_1\mu_2}^{\beta\alpha}(R_2)\right\} = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} Re\left\{P_{\mu_3\mu_4}^{\beta\alpha}(R')\right\}\mathbf{K}^- \tag{2.11}$$

$$Im\left\{K_{\mu_1\mu_2}^{\beta\alpha}(R_2)\right\} = -\sum_{\mu_3 \geq \mu_4}\sum_{R'} Im\left\{P_{\mu_3\mu_4}^{\beta\alpha}(R')\right\}\mathbf{K}^-. \tag{2.12}$$

As with the density matrix, these invocations of time reversal symmetry limit the applicability of the modified code to time-reversal-invariant systems. It is not physically necessary to use time reversal in this way. However, the unmodified CRYSTAL code does not need to worry about $P^{\alpha\beta}$ or $P^{\beta\alpha}$. Furthermore, its density matrices are always real in real space. Therefore, the code leveraged the Hermitian property of the density matrix here as well. In order to avoid entirely rewriting the code for evaluating the two electron integrals, it was necessary to invoke time reversal symmetry here as well.

The key advantage of CRYSTAL over its competitors is its computational speed, and this speed is the result of very careful optimization of every subroutine. Nowhere is that more evident than in the routines for the two-electron integrals, and a home-built version that could have been written in a reasonable amount of time would have been so slow as to make doing real science impractical.

2.5 Spin-Orbit Operator in All-Electron Calculations

For completeness, the evaluation of the spin-orbit operator is discussed in this section. As detailed in Section 2.1, this is not implemented in the modified code used for calculations in this thesis, as the spin-orbit coupling is accounted for implicitly by the pseudopotentials. The spin-orbit coupling (SOC) operator can be written in atomic units as

$$-\frac{1}{4c^2}(\nabla V \times \vec{p}) \cdot \vec{\sigma},$$

where V is the nuclear potential, $\vec{p}$ is the momentum operator, and $\vec{\sigma}$ is the vector of Pauli matrices. The potential from the nuclei is $\sum_c -Z_c/r_c$, where r_c is the distance between a given electron and nucleus c. The momentum operator $\vec{p} = -i\nabla$.

Let $\phi(l, m, n, \alpha, \vec{A})$ be a Gaussian orbital centered on atom $\vec{A}$: $x^l y^m z^n e^{-\alpha(\vec{r}-\vec{A})^2}$. The matrix elements needed for the spin-orbit operator are thus

$$\frac{i}{4c^2} \sum_c \left\langle \phi_1(l_1, m_1, n_1, \alpha_1, \vec{A}) \left| \frac{Z_c \vec{r_c}}{r_c^3} \times \nabla \right| \phi_2(l_2, m_2, n_2, \alpha_2, \vec{B}) \right\rangle \cdot \langle \chi_1 | \vec{\sigma} | \chi_2 \rangle,$$

where χ_1 is a spin function, that is, α or β. The evaluation of these integrals was discussed by McMurchie and Davidson [40].

Expanding the cross product yields

$$\frac{\vec{r_c}}{r_c^3} \times \nabla = \left(\frac{y_c}{r_c^3} \frac{\partial}{\partial z} - \frac{z_c}{r_c^3} \frac{\partial}{\partial y} \right) \hat{x} + \left(\frac{x_c}{r_c^3} \frac{\partial}{\partial z} - \frac{z_c}{r_c^3} \frac{\partial}{\partial x} \right) \hat{y} + \left(\frac{x_c}{r_c^3} \frac{\partial}{\partial y} - \frac{y_c}{r_c^3} \frac{\partial}{\partial x} \right) \hat{z}.$$

Noticing that

$$\frac{x_c}{r_c^3} = -\frac{\partial \frac{1}{r_c}}{\partial x}$$

and integrating by parts, the spin-orbit matrix element can be written as

$$\left[\int \left(\nabla \phi_1(l_1, m_1, n_1, \alpha_1, \vec{A}) \times \nabla \phi_2(l_2, m_2, n_2, \alpha_2, \vec{B}) \right) \frac{1}{r_c} \mathrm{d}^3 r \right] \cdot \langle \chi_1 | \vec{\sigma} | \chi_2 \rangle.$$

The integrals over spin functions are easily evaluated.

$$< \alpha | \sigma_x | \beta > \quad = \quad < \beta | \sigma_x | \alpha >= 1$$

$$< \alpha | \sigma_x | \alpha > \quad = \quad < \beta | \sigma_x | \beta >= 0$$

$$- < \alpha | \sigma_y | \beta > \quad = \quad < \beta | \sigma_y | \alpha >= i$$

$$< \alpha | \sigma_y | \alpha > \quad = \quad < \beta | \sigma_y | \beta >= 0$$

$$< \alpha | \sigma_z | \beta > \quad = \quad < \beta | \sigma_z | \alpha >= 0$$

$$< \alpha | \sigma_z | \alpha > \quad = \quad - < \beta | \sigma_z | \beta >= 1$$

Chapter 3

RESOLUTION OF THE BAND GAP PREDICTION PROBLEM FOR MATERIALS DESIGN

As we enter the Materials Genome Initiative (MGI) era of designing optimal materials *in silico*, it is essential to quickly and accurately predict the band gaps of proposed new materials. Similarly, experimentalists would benefit from an efficient computational tool for predicting band gaps prior to synthesis. We consider the problem of predicting physical band gaps to be solved by a computational method that delivers an accurate band gap for compounds spanning the whole periodic table and is simultaneously practical to compute.

Standard DFT methods, such as the Perdew-Burke-Ernzerhof (PBE) [41] functional, have been the workhorse of computational materials science for decades. PBE is used by both theorists and experimentalists because of its computational speed and widespread availability in commercial codes. Unfortunately, it contains unphysical self-Coulomb repulsion [42], leading to a systematic underestimate of band gaps [43–45].

It is well known that including exact Hartree-Fock exchange substantially reduces the Coulomb self-repulsion error [2]. The Heyd Scuseria Ernzerhof functional (HSE06) includes short-range exact Hartree Fock exchange and has become a standard in the community [46].

In 2012, Moussa et al. [47] and Lucero et al. [48] benchmarked HSE06 against a set of 33 compounds consisting solely of simple binary materials of the form AB. This set neglects many classes of materials such as: alkali halides, transition metal oxides (Mott insulators), transition metal halides, systems with strong spin-orbit coupling (particularly bismuth- and lead-containing compounds), transition metal chalcogenides, magnetic systems, and two-dimensional systems. Moussa et al. found a mean absolute deviation (MAD) of 0.32 eV with a mean error (ME) of -0.24 eV, while Lucero et al. found a MAD of 0.33 eV and a ME of -0.26 eV for this limited set of compounds.

In Moussa et al. [47], the HSE calculations were done without spin-orbit coupling for experimental crystal structures and were compared to experimental band gaps with the spin-orbit coupling removed. For these compounds, the spin-orbit splitting

is seen at the valence band maximum at the Γ point (wave vector k = 0). In these cases, the spin-orbit coupling originates from the anions (Sb, Se, Te, As). At the Γ point, the valence band consists of anion p orbitals. Spin-orbit coupling splits these triply degenerate p states at Γ into a four-fold degenerate ($j = 3/2$) band and a two-fold degenerate ($j = 1/2$) band. The $j = 3/2$ state is split upwards in energy by $(1/3)\Delta$ and the $j = 1/2$ state is split downwards by $(2/3)\Delta$, where Δ is the energy difference between the $j = 3/2$ and $j = 1/2$ state. For comparison to non-spin-orbit coupling HSE calculations, the known experimental band gaps of the compounds with the above anions must be increased by $(1/3)\Delta$. For the As, Se, Sb, and Te compounds, Δ is approximately 0.3, 0.4, 0.7, and 0.9 eV, respectively.

Inexplicably, Moussa et al. include the spin orbit correction for CdTe (1.61 eV experiment corrected to 1.92 eV), but fail to include the spin-orbit correction for BAs, AlAs, AlSb, GaAs, GaSb, InAs, InSb, ZnSe, ZnTe, CdSe, MgSe, MgTe, BaSe, and BaTe. In addition, Moussa et al. use an experimental band gap of 7.22 eV for MgO instead of the low-temperature value of 7.9 eV [48, 49]. These numbers appear to have been transcribed from Heyd et al. [50]. When we correct these values, the MAD increases to 0.39 eV, and the mean error becomes -0.36 eV.

In Lucero et al. [48], the MgO experimental value was corrected. These authors optimized the geometry and quoted band gaps at the optimized structures. Thus, we cannot directly compare these results to those of Moussa et al [47]. While no spin-orbit corrections were made to the experimental band gaps, it is unclear from the paper whether these calculations include spin-orbit coupling or not. However, in Figure 4, the authors show their band structure for AlSb, and there is no splitting of the valence band at the Γ point. This figure suggests spin-orbit coupling may not have been included in these calculations.

The two authoritative HSE06 studies discussed above did not include the Mott insulators (FeO, NiO, MnO, CoO) or any compounds with band gaps above 8 eV. Thus, we expanded the HSE06 test set to include these Mott insulators and NaCl, β-cristobalite SiO_2, LiCl, α-quartz SiO_2, and LiF because these additional compounds are included in our present work. Using literature HSE06 values for these compounds [51–54], the MAD and ME for Moussa et al. rise to 0.57 eV and -0.51 eV, respectively. The MAD and ME for Lucero et al. rise to 0.52 eV and -0.43 eV, respectively. These numbers are slight underestimates because we were unable to locate a literature HSE06 band gap for β-cristobalite. Thus, we assumed HSE06 obtained perfect agreement to experiment for this compound. As we show below,

the MAD and ME for many-body perturbation theory (*GW* method) are better than these HSE06 numbers. Thus, we do not consider HSE06 any further.

The rigorous approach to solving the band gap problem is to calculate the Green's function using many-body perturbation theory (referred to as the *GW* method [1]). In this method, the quasiparticle excitation spectrum is calculated rigorously, whereas DFT calculates Kohn-Sham eigenvalues that are not rigorously physical. However, full *GW* calculations are far too computationally expensive for routine use. A non-self-consistent *GW* approximation, G_0W_0, has been used to improve standard DFT results for decades [51, 55–58]. Recently, iterative "post-G_0W_0" methods have been used to improve G_0W_0 results [59–63]. As the next term in the expansion of the true Green's function, post-G_0W_0 should systematically improve quasiparticle energies and band gaps, albeit at a considerably higher computational cost. Despite intense effort in recent years, all *GW* methods are too computationally expensive for applications such as MGI.

Here, we show that the B3PW91 hybrid density functional is more accurate than *GW* by approximately a factor of 1.5 and is also 3-4 orders of magnitude faster computationally. Because speed and accuracy are competing properties, it is highly unusual to discover a method that dramatically improves both. This observation constitutes the main result of the thesis, and suggests that B3PW91 should replace PBE as the default computational approach.

Hybrid density functionals include a fraction of exact Hartree-Fock exchange, which dramatically reduces the self-Coulomb error in DFT. Previously, we have shown that hybrid DFT matches the best G_0W_0 bulk band structure for the useful thermoelectrics and topological insulators Bi_2Te_3 and Bi_2Se_3, whereas standard DFT does not [64]. Moreover, with hybrid DFT, we performed calculations of large slabs that are inaccessible to *GW*. Other studies [2, 47, 48, 65–68] of small sets of compounds have also hinted at the usefulness of hybrid functionals for solids.

In order to determine the quality of the B3PW91 functional (referred to as B3PW hereafter), we computed band gaps of 70 insulating compounds with band gaps ranging from 0.014 to 14.2 eV. These compounds span the entire periodic table (except for lanthanides, actinides, and solid noble gases) and include thermoelectrics, topological insulators, transition metal oxides, photovoltaics, elemental and binary semiconductors, and transition metal halides. In order to have the most accurate comparison to experiment, spin-orbit coupling was included.

No lanthanides or actinides were studied because they have strong spin-orbit coupling and partially filled 4f and 5f shells. The CRYSTAL98 source code we modified to include spin-orbit coupling does not include f orbitals. Solid noble gases were excluded because the anion is unbound, and, in the solid phases, the conduction band is above the vacuum energy [69]. We know of no other study that spans as many classes of materials. Of these 70 compounds, we found *GW* published literature results for only 53, and there was no single study of all 53 of these compounds. The largest study [62] we found had 16 compounds. We did not perform any *GW* calculations. Tables A.7 – A.7 contain references to every single *GW* data point used in this chapter.

All of our band gap results are shown in Figure 3.1. Panels 3.1a, b, and c plot the B3PW, PBE, and GW band gaps up to 7 eV, respectively. Panel 3.1d plots the band gaps for all three methods from 7-15 eV. The experimental band gap range is split into $0 - 7$ eV and $7 - 15$ eV for clarity only. Figures A.9 – A.11 show all band gap results for B3PW, PBE, and GW, respectively over the full range of $0 - 15$ eV. All of the compounds are listed in Figures 3.2 and 3.3. The results in Figures 3.1a and 3.1b are for 64 out of the total 70 compounds. The results in Figure 3.1c for *GW* are for 47 out of the 53 total compounds. The mean absolute deviation (MAD) for B3PW is 0.22 eV. The MAD over all *GW* methods (G_0W_0 and post-G_0W_0) is 0.36 eV. The MADs for G_0W_0 and post-G_0W_0 are equal to 0.36 eV and 0.35 eV, respectively. The MAD for PBE is 1.10 eV. The B3PW MAD is 1.6 times smaller than the *GW* MAD, and 5 times smaller than the PBE MAD. As is well known [43], PBE systematically underestimates band gaps. In fact, for ten compounds, PBE predicts a metal (zero band gap) rather than an insulator (nonzero band gap). We found two exceptions (PbSe and Bi_2Se_3), shown in the inset to Figure 3.1c. These compounds have strong spin-orbit coupling, and will be discussed later in the paper.

Figure 3.1d shows B3PW (blue circles), PBE (purple circles), G_0W_0 (red circles), and post-G_0W_0 (green circles) results for six compounds with experimental band gaps above 7 eV. The MAD for each method over these compounds is shown in the upper left corner. The bottom right of the figure shows the MADs for the full band gap range ($0 - 15$ eV). The MADs are worse in the $7 - 15$ eV band gap range for all methods compared to the $0 - 7$ eV range. For the entire $0 - 15$ eV band gap range, post-G_0W_0 is slightly better than G_0W_0. The MAD for B3PW is worse than all *GW* methods in the $7 - 15$ eV range. Over the full $0 - 15$ eV band gap range, B3PW has
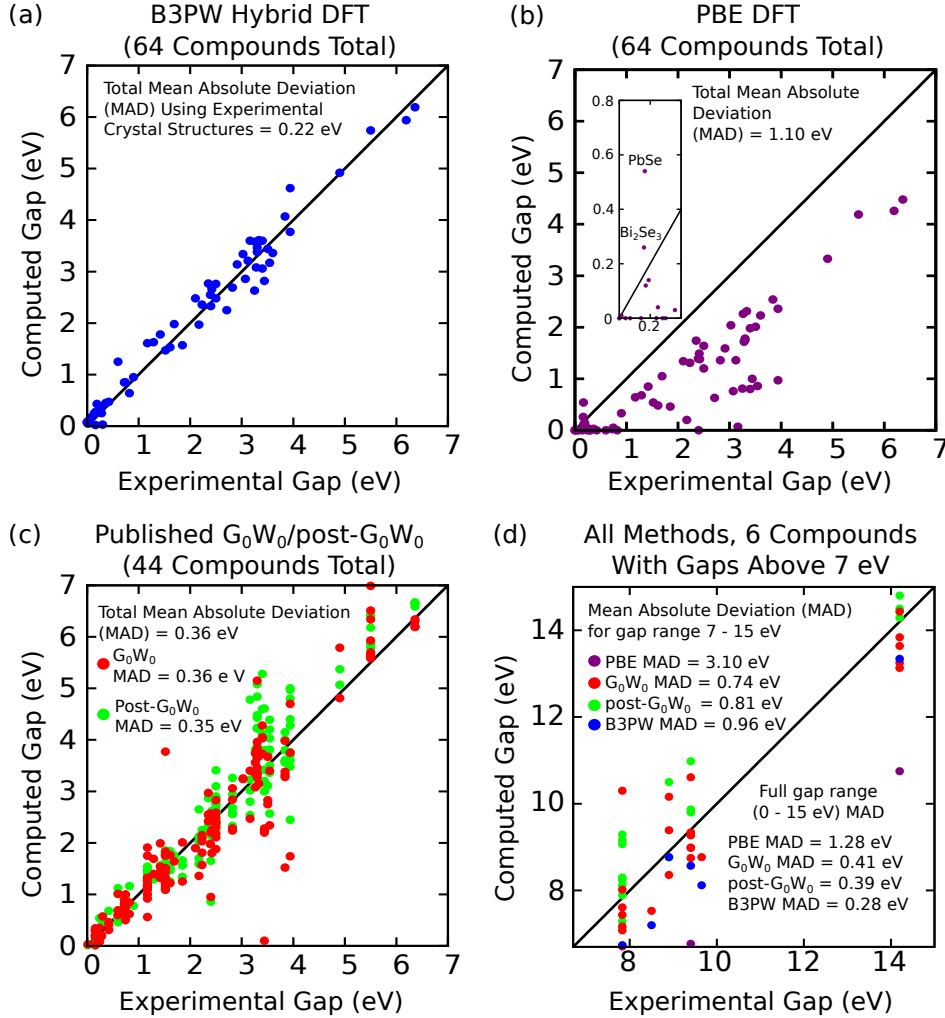
Figure 3.1: Calculated B3PW (hybrid DFT), GW (many-body perturbation), and PBE (standard DFT) band gaps versus low-temperature experiment. A list of the compounds studied may be found in Figures 3.2 and 3.3 and Tables A.7–A.7. (a) B3PW, (b) PBE, (c) GW results (G_0W_0 and post-G_0W_0) versus experiment for compounds with experimental band gaps from 0 to 7 eV. (d) B3PW, PBE, and GW data for eight compounds with experimental band gaps larger than 7 eV. For the full band gap range, the MADs are 1.28, 0.41, 0.39, and 0.28 eV for PBE, G_0W_0, post-G_0W_0, and B3PW, respectively. Overall, the error for B3PW is 1.4 times smaller than that for post-G_0W_0. The computational cost of B3PW is 3.2 ± 2.4 times that of PBE.

the lowest MAD.

In order to further test B3PW against the GW method, we broke down the GW data into four subsets: G_0W_0 with any starting wavefunction, G_0W_0 specifically using LDA as the starting wavefunction, G_0W_0 using PBE as the starting wavefunction, and post-G_0W_0 (Figures A.1 – A.8). In all four cases, B3PW had the lower MAD.

Since the only difference between our B3PW and PBE calculations is the functional, we can make a machine-independent comparison of computational cost. We find that the computational cost of B3PW is 3.2 ± 2.4 times that of PBE. It is important to note that hybrid functionals are typically $\sim 10^3$ times slower than PBE in plane-wave basis set codes. Because we used localized Gaussian basis sets, B3PW became competitive with PBE. In practical terms, this additional cost is negligible compared to the nearly five-fold improvement in accuracy.

Plane-wave basis sets are simpler than Gaussian basis sets because they are characterized by one number: the wave vector cutoff. Gaussian basis sets require Gaussian exponents and contraction coefficients for each basis function, and different Gaussian basis sets have different numbers of basis functions. Thus, the choice of the correct Gaussian basis set is more complicated. Fortunately, computational chemists have developed highly optimized Gaussian basis sets for every element. However, these highly optimized basis sets developed for computational chemistry usually cannot be used without some modification. This is because the extremely diffuse basis function commonly used in high-quality Gaussian basis sets can lead to basis set linear dependence when used in a periodic system [3].

Linear independence can usually be ensured by modifying the basis set so that the most diffuse Gaussian has an exponent of 0.1, as discussed in reference [68]. Although minimizing the energy with respect to other Gaussian exponents can improve band gap predictions, we intentionally chose not to perform this step to keep our evaluation of B3PW as straightforward as possible. Basis sets were not tuned to a particular system, and hence can be transferred to other compounds.

The basis sets used in this paper were chosen according to a systematic recipe which is described in complete detail in section 1 of Appendix A. Additionally, the basis sets used in our input decks (included in the Supplement to ref [70]) are linearly independent. As a result, we believe these basis sets can be used as a template for rapid calculation of systems not considered here.

An issue for both Gaussian and plane-wave basis set calculations is whether or not to optimize the crystal structure or use the experimental crystal structure. A perfect theory would predict zero force for experimental structures, rendering this question irrelevant. However, DFT (regardless of basis set) tends to overestimate lattice constants by a few percent. We did our calculations at the experimental crystal structure because it is a harsher test to demand that a theory predict the correct electronic structure given only the physical structure of a material.

We now return to the inset of Figure 3.1b showing that PBE overestimated the band gaps for PbSe and Bi_2Se_3. These are the only two compounds where PBE overestimated the band gap. In general, spin-orbit coupling decreases the band gap by breaking the spin degeneracy. However, if the PBE band gap without spin-orbit coupling is smaller than the magnitude of the spin-orbit splitting, unphysical band repulsions may occur. Thus, for strong spin-orbit compounds with small gaps, PBE may overestimate the band gap.

Figures 3.2 and 3.3 show the band gap error broken down by compound for all 70 compounds with experimental band gaps in the range 0–14.2 eV. The compounds are ordered on the x-axis by increasing experimental band gap, and compounds with magenta labels are those where we did not find *GW* data. The largest B3PW band gap error occurred for α-quartz. As discussed above, these errors could be reduced by basis set optimizations. The mean signed error of B3PW is -0.03 eV.

In some cases (e.g., CoO), there is more than one measured experimental band gap. Our methodology for choosing a number for comparison is discussed in Table A.8. We did not choose the experimental results closest to the B3PW prediction. Doing so would lower the overall MAD from 0.28 eV to 0.25 eV. On the other hand, choosing the experimental band gaps furthest from the B3PW prediction raises the MAD by 0.03 eV to 0.31 eV, which is less than the post-G_0W_0 MAD of 0.39 eV. Thus, we believe that B3PW solves the band gap problem.
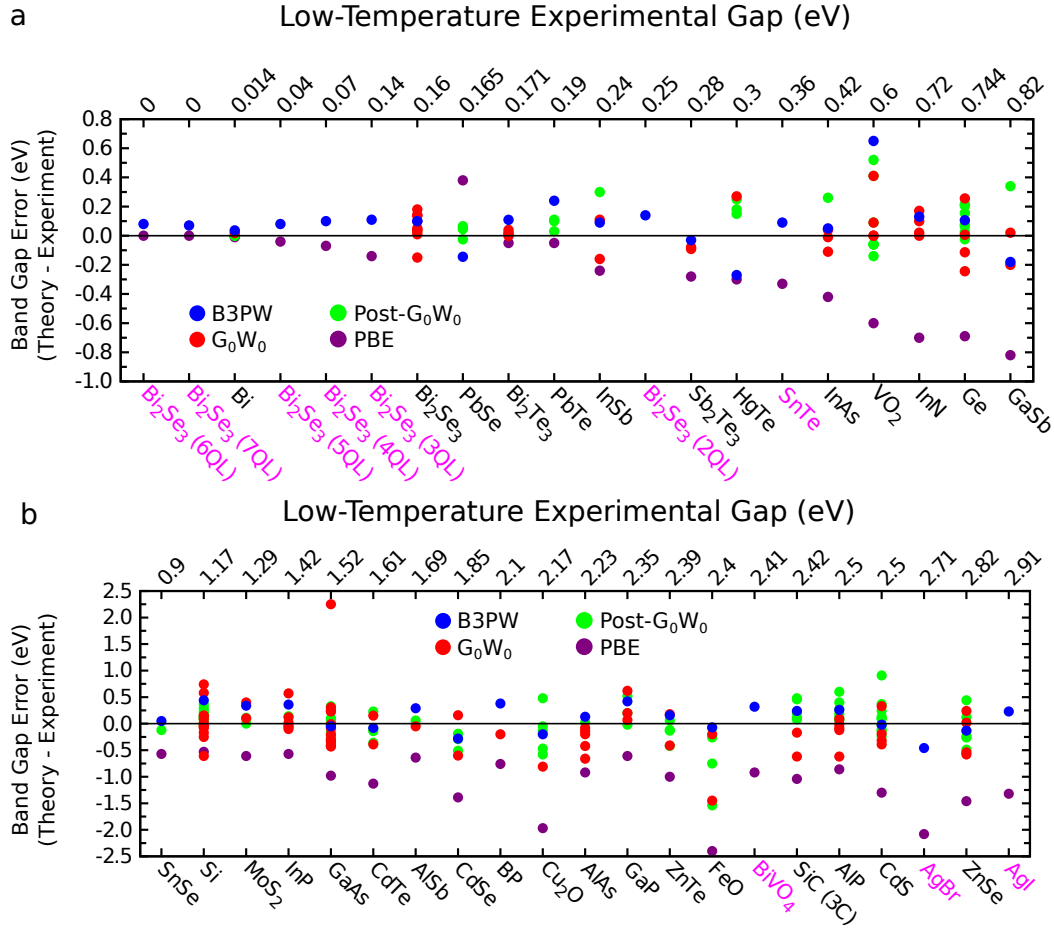
Figure 3.2: The difference between computed and experimental band gaps for 41 of the 70 compounds studied in Figure 1. The compounds are ordered by increasing experimental band gaps. Results for compounds with experimental gaps from (a) 0.0 – 0.82 eV, (b) 0.9 – 2.91 eV. The horizontal lines at zero represent perfect agreement between theory and experiment. Points above these lines indicate theory has overestimated the gap, and points below these lines indicate theory has underestimated the gap. Compound names are listed on the bottom x-axes; experimental band gaps (low temperature or $T \approx 0K$ when available; see Supplemental Table S4 for a discussion of experimental gaps) are listed on the top x-axes. Compound labels in magenta are those for which our literature search found no GW results. Our B3PW results are shown by blue circles, and our PBE results are shown by purple circles. Red and green circles represent literature G_0W_0 and post-G_0W_0 results, respectively. Bi_2Se_3(NQL) refers to a two-dimensional slab of Bi_2Se_3 composed of N quintuple layers (QLs) (see reference [64] for further details). Tables A.7–A.7 contain all of our B3PW and PBE results, the lowest and highest GW band gap for each system, all references from which GW results were taken.
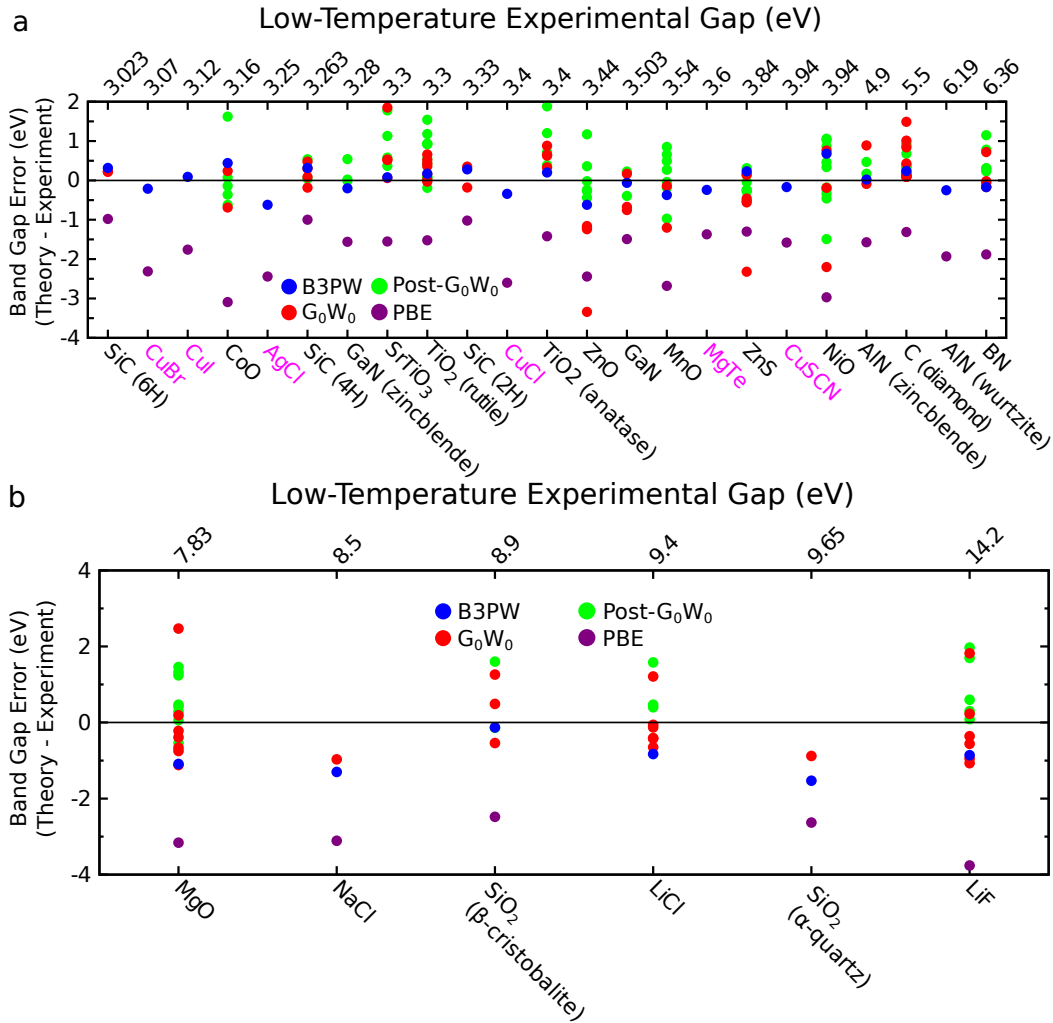
Figure 3.3: Same as Figure 3.2 for the remaining 29 compounds from Figure 3.1. Results for compounds with experimental gaps (a) 3.023 – 6.36 eV, (b) 7.83 – 14.2 eV.

In summary, we believe hybrid DFT can successfully resolve the band gap prediction problem for practical materials design. Standard DFT methods such as PBE have sufficient computational speed, but suffer from very poor accuracy. In particular, we found the PBE functional to have a mean absolute deviation from experiment (MAD) of 1.28 eV over a class of 70 insulating compounds. While many-body perturbation methods (GW) are significantly more accurate than PBE (MAD = 0.40 eV), they are 3–4 orders of magnitude slower. Here, we have shown that the B3PW91 hybrid density functional has the best of both worlds. The MAD of B3PW91 was 0.28 eV. When used with localized Gaussian basis sets, the computational cost of B3PW91 is only 3.2 ± 2.4 times that of PBE. This remarkable

result was achieved without optimization of any computational parameter, including crystal structure. Thus, B3PW91 can be employed "off the shelf" by non-expert users to deliver excellent predictions of band gaps for materials that have not yet been synthesized.

3.1 Computational Methods

The GW data we used for comparison consisted of 387 calculations from 42 publications dating from 1987 to 2015. Of these 42 publications, 23 were from 2010 or later. Tables A.7 – A.7 contain all references to the *GW* data, and all references to experimental band gaps. Hyperlinks to the input decks for all of our hybrid DFT calculations are included in the corresponding tables in the supplement to ref [70]. In this chapter, we only study band gaps because *GW* is primarily benchmarked against band gaps. We chose low-temperature experimental band gaps from reference [49] for comparison whenever possible. Supplemental Table S4 details our choice of experimental band gaps.

In this work, we use the first proposed hybrid functional, B3PW91 [71], based on its prior success with topological isulators [64], binary semiconductors [68], and band offsets [72]. We refer to the B3PW91 functional as B3PW for simplicity. This functional was developed by Becke using the adiabatic connection [71]. B3PW makes three corrections to the local density approximation (LDA) exchange correlation energy. First, a fraction a_0 of the LDA exchange is replaced by Hartree-Fock exact exchange. Second, Becke's 1988 gradient correction [73] for exchange is included, weighted by the parameter a_x. Finally, the Perdew-Wang 1991 (PW91) gradient correction [74] for correlation is included, weighted by the parameter a_c. These parameters were determined by a least squares fit to 56 atomization energies, 42 ionization potentials, 8 proton affinities, and 10 total atomic energies [71]. This fitting gave $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.81$. Becke's parameterization made no use of band gaps or any property of a solid, and we made no adjustment to these parameters.

Computing exact Hartree-Fock exchange is about three orders of magnitude faster with localized Gaussian basis sets [64] compared to plane-wave basis sets. Therefore, we use the CRYSTAL [3] code for all of our calculations. CRYSTAL does not include spin-orbit coupling. For the cases with strong spin-orbit coupling, we used our modified version of CRYSTAL98 [64]. When used for solids, the Gaussian basis sets commonly used and optimized for molecules can lead to linear dependen-

cies. Hence, modifications for diffuse orbitals are required to remove these linear dependencies, as described in the CRYSTAL manual [3]. We also used Stuttgart pseudopotentials [75, 76]. The basis set and pseudopotential (if used) for all of our calculations can be seen in the input decks included in the supplement to ref [70]. Our systematic basis set recipe is described in full detail in section 1 of Appendix A. It is known that optimization of basis sets and geometries improves band gaps [68]. However, we chose not to optimize basis sets and to use experimental crystal structures because these optimizations are impractical for high-throughput applications such as MGI.

In addition to the choice of basis set, it is important to have enough irreducible k-points to obtain a fully converged result. We chose on average 103 irreducible k-points in our calculations. This number is far larger than necessary, but our calculations are so computationally inexpensive that this overkill is irrelevant. High-throughput screening becomes more practical when one can ensure convergence with a default k-point sampling. In fact, we believe our basis sets and input decks (included in the Supplement to ref [70]) can be used as a template for studies on new materials.

Chapter 4

ANALYSIS OF THE ROLE OF EXACT EXCHANGE IN HYBRID DENSITY FUNCTIONALS

4.1 Introduction

In Chapter 3, we found the surprising result that the B3PW hybrid density functional is about 1.5 times more accurate for band gaps and 3-4 orders of magnitude faster than *GW*. However, B3PW has not been systematically compared to any of the many flavors of hybrid density functional that have been developed over the last two decades. Naturally, different hybrid functionals use different exchange and correlation functionals, and contain different fractions of exact Hartree-Fock exchange. It is not clear how these differences affect band gap predictions from one functional relative to another. The purpose of the present chapter is to compare B3PW to several other hybrid functionals in order to gain some insight into the reasons behind the success of B3PW for band gap prediction.

Soon after the introduction of B3PW came the B3LYP functional, which quickly became the default method in molecular quantum chemistry. Both B3PW and B3LYP use the B88 gradient correction to the LDA exchange and 20% Hartree-Fock exchange. The difference between the two functionals is the correlation functional: B3LYP uses the Lee-Yang-Parr [77] correlation functional, and B3PW uses the Perdew-Wang 1991 [74] correlation functional.

The fraction of exact exchange in Becke's three-parameter functionals B3LYP and B3PW was determined by fitting to experimental data. Perdew and coworkers [78] argued from perturbation theory that 25% exact exchange should be used. This idea was implemented by Adamo and Barone in the PBE0 functional [79]. In this functional, PBE is used for both the exchange and correlation. The only modification from PBE is the replacement of 25% of the PBE exchange with Hartree-Fock exchange. Thus, PBE0 has no empirically adjusted parameters, while B3PW and B3LYP have three.

The HSE functional (briefly discussed in Chapter 3) differs from PBE0 in the treatment of exact exchange. While the Hartree-Fock exchange in PBE0 is evaluated over all space, it is only evaluated over a short range in HSE. There is one parameter determining the distance over which to evaluate the exact exchange. It is worth

noting that in the formulation of HSE [46], this parameter was not fully optimized. Rather, the authors chose the value they considered to have the best balance of accuracy and speed. These authors used the Gaussian software for their studies. We can make no independent study of the performance of this program, since we are forbidden from buying a license (see http://www.bannedbygaussian.org). Thus, we can only state that the computational cost of evaluating the full exact exchange in CRYSTAL is only 3 times the cost of a local density functional.

All of the hybrid functionals discussed above use GGA functionals like PBE or PW91. A more recent development that has become very widely used in molecular quantum chemistry is the meta-GGA, which includes dependence on the spin kinetic energy density [80]. One such meta-GGA, M06L, was found to outperform B3LYP for main-group thermochemistry [80]. When applied to a small set of nine semiconductors, M06L was found to predict band gaps more accurately than PBE, but less accurately than HSE [81]. While M06L is a purely local density functional, M06 is a global hybrid version of the same functional with 27% exact exchange [82]. These functionals contain many more parameters than those described above, but are highly regarded by computational chemists.

In this chapter, we systematically compare B3LYP, M06L, HSE06, PBE0, and M06 to B3PW. We find that M06L and HSE06, which have less Hartree-Fock exchange than B3PW, always return lower band gaps than B3PW. By contrast, PBE0 and M06, which have more Hartree-Fock exchange than B3PW, systematically predict higher band gaps than B3PW. The B3LYP functional, which differs only in the correlation functional, does not systematically predict higher or lower band gaps than B3PW. The error of the B3LYP band gaps compared to experiment is slightly higher than the error for B3PW. Based on these trends, we suggest two possible avenues for improving hybrid functionals even further for band gap prediction.

4.2   Computational Methods

Band gap calculations using these functionals were carried out for 64 of the 70 compounds from Chapter 3. The Bi_2Se_3 quintuple layer systems were excluded due to their high computational cost. Exactly the same geometries and basis sets were used in these calculations as those in Chapter 3. Because only the B3PW and B3LYP hybrid functionals were implemented in CRYSTAL98, these were the only functionals that could be used for spin-orbit coupling (SOC) calculations. For all 6 functionals, the 2014 version of CRYSTAL was used to perform calculations

without SOC. Band gaps from non-SOC calculations should not be compared to experiment for systems with non-negligible SOC, since these band gaps are not physical. Despite this, some authors have used non-SOC band gaps for assessing the quality of density functionals [47, 48].

One possibility to make consistent comparisons to experiment using non-SOC calculations is to "correct" the experimental band gap values to remove the effect of SOC. This is possible in simple zincblende systems where the spin-orbit coupling comes from the anion atomic-like spin-orbit splitting. For more complex materials, there is no clear way to make this correction, so we choose not to take this approach. Another alternative is to use the difference between SOC and non-SOC B3PW band gaps to correct the band gaps calculated using other functionals. In many cases, SOC is a first-order perturbation and this approach is reasonable. However, in strongly covalent systems, there can be significant projection of orbitals from an atom onto a neighboring atom. In these cases, the effect of SOC depends strongly on the orbital character of the valence and conduction bands, which is arrived at self-consistently and varies significantly from one functional to another.

In order to have a completely consistent comparison, we chose to compare non-SOC calculations of all the functionals considered here to non-SOC B3PW calculations. The figures in the next section plot non-SOC band gaps from various functionals versus non-SOC B3PW band gaps. The line $y = x$ represents the B3PW non-SOC band gap. In these plots, points above the diagonal are systems where the functional being considered predicted a higher band gap than B3PW. Conversely, points below the diagonal are systems where the given functional predicted a lower band gap than B3PW. The purpose of these plots is to show how computed band gaps change when using different exchange and/or correlation functionals than those used in B3PW. These plots do not give direct information about the accuracy of a functional for predicting band gaps.

When SOC is included, B3PW has a mean signed error of near zero compared to experiment. Therefore, if a method systematically returns higher non-SOC band gaps compared to non-SOC B3PW band gaps, to first order that method will likely have a positive mean signed error when SOC calculations are compared to experiment. Similarly, a method that systematically returns lower non-SOC band gaps compared to non-SOC B3PW band gaps will have a negative mean signed error when SOC calculations are compared to experiment.

4.3 Results and Discussion

The non-SOC band gaps computed with every functional considered here are tabulated in Table 4.4.

B3LYP

Table 4.3 shows band gaps computed with B3PW and B3LYP including spin-orbit coupling for the 64 compounds considered in this chapter. Because spin-orbit coupling is included in these calculations, the computed band gaps can be compared to experiment. Figure 4.3 shows the results. The MAD and ME for B3PW are 0.30 and -0.03 eV, respectively. The MAD and ME for B3LYP are 0.35 and +0.03 eV, respectively. The MAD is about 17% larger than the MAD for B3PW, but the ME has the same magnitude (near zero). Thus, B3LYP is as likely to overestimate a band gap as it is to underestimate one.

Figure 4.3 shows non-spin-orbit B3LYP gaps versus non-spin-orbit B3PW band gaps. The mean signed difference between B3LYP and B3PW, defined as

$$\frac{1}{N} \sum_{i=1}^{N} E_g^{B3LYP} - E_g^{B3PW},$$

is 0.05 eV, indicating a very slight tendency of B3LYP to return higher gaps than B3PW. The mean unsigned difference, defined as

$$\frac{1}{N} \sum_{i=1}^{N} |E_g^{B3LYP} - E_g^{B3PW}|,$$

is 0.14 eV.

The differences between band gaps computed with B3LYP and B3PW are caused by the different correlation functionals, since this is the only difference between B3LYP and B3PW. Thus, it is clear that LYP is not quite as well-suited for band gap prediction as B3PW. As indicated by the small mean signed difference, B3LYP does not systematically predict higher or lower gaps than B3PW.

M06L

Figure 4.3 shows non-spin-orbit M06L gaps versus non-spin-orbit B3PW band gaps. The mean signed difference between M06L and B3PW is -0.94 eV, and the mean unsigned difference is 0.94 eV. As is evident in Figure 4.3, M06L always predicts lower band gaps than B3PW. This is consistent with the results of ref [81], which considered band gaps for only 9 compounds.

Table 4.1: Band Gaps (eV) from experiment, B3PW, and B3LYP calculations including spin-orbit coupling (SOC). The abbreviations zb and w stand for zincblende and wurtzite, respectively. For TiO_2, (r) and (a) refer to rutile and anatase. For SiO_2, (a) and (b) refer to α-quartz and β-cristobalite. All calculations were performed using exactly the same input decks used in Chapter 3.

System	Exp	B3PW	B3LYP	System	Exp	B3PW	B3LYP
Bi	0.014	0.00	0.02	AgBr	2.71	2.24	2.23
Bi_2Se_3	0.16	0.12	0.16	ZnSe	2.82	2.69	2.65
PbSe	0.165	0.25	0.38	AgI	2.91	3.13	3.09
Bi_2Te_3	0.171	0.23	0.22	SiC (6H)	3.023	3.34	3.59
PbTe	0.190	0.59	0.66	CuBr	3.07	2.87	2.84
InSb	0.24	0.34	0.26	CuI	3.12	3.21	3.20
Sb_2Te_3	0.28	0.22	0.18	CoO	3.16	3.60	3.84
HgTe	0.3	0.01	0.00	AgCl	3.25	2.63	2.66
SnTe	0.36	0.44	0.56	SiC (4H)	3.263	3.58	3.83
InAs	0.42	0.47	0.42	GaN (zb)	3.28	3.08	3.24
VO_2	0.6	1.25	1.31	$SrTiO_3$	3.3	3.38	3.40
InN	0.72	0.87	0.93	TiO_2 (r)	3.3	3.47	3.50
Ge	0.744	0.86	0.72	SiC (2H)	3.33	3.61	3.88
GaSb	0.82	0.66	0.54	CuCl	3.4	3.06	3.06
SnSe	0.90	0.95	1.04	TiO_2 (a)	3.4	3.60	3.63
Si	1.17	1.61	1.79	ZnO	3.44	2.82	2.90
MoS_2	1.29	1.63	1.71	GaN	3.503	3.43	3.55
InP	1.42	1.78	1.74	MnO	3.54	3.17	3.33
GaAs	1.52	1.48	1.40	MgTe	3.6	3.31	3.42
CdTe	1.61	1.53	1.50	ZnS	3.84	4.08	4.07
AlSb	1.69	1.83	2.01	NiO	3.94	4.62	4.56
CdSe	1.85	1.57	1.56	CuSCN	3.94	3.97	3.94
BP	2.1	2.48	2.71	AlN (zb)	4.9	4.92	5.20
Cu_2O	2.17	2.20	2.29	C	5.5	5.74	5.96
AlAs	2.23	2.37	2.63	AlN (w)	6.19	5.94	6.14
GaP	2.35	2.71	2.88	BN	6.36	6.19	6.49
ZnTe	2.39	2.55	2.50	MgO	7.83	6.74	6.90
FeO	2.4	2.33	2.58	NaCl	8.5	7.20	7.05
$BiVO_4$	2.41	2.98	3.02	SiO_2 (b)	8.9	8.77	8.46
SiC (3C)	2.42	2.66	2.93	LiCl	9.4	8.57	8.57
AlP	2.5	2.76	2.98	SiO_2 (a)	9.65	8.12	8.05
CdS	2.5	2.49	2.50	LiF	14.2	13.33	13.35

Figure 4.1: Band gaps computed using the (a), (b) B3PW and (c), (d) B3LYP hybrid density functionals including spin-orbit coupling versus experiment.

HSE06

Figure 4.3 shows non-spin-orbit HSE06 gaps versus non-spin-orbit B3PW band gaps. The mean signed difference between HSE06 and B3PW is -0.26 eV, and the mean unsigned difference is 0.26 eV. For each of the 64 compounds tested, HSE06 returned a lower band gap than B3PW. In cases where B3PW overestimates band gaps, such as transition metal oxides, HSE06 will likely give better agreement with experiment. By contrast, in cases where B3PW underestimates band gaps, such as alkali halides, B3PW will likely give better agreement with experiment.

PBE0

Figure 4.3 shows non-spin-orbit PBE0 gaps versus non-spin-orbit B3PW band gaps. The mean signed difference between PBE0 and B3PW is +0.32 eV, and the mean
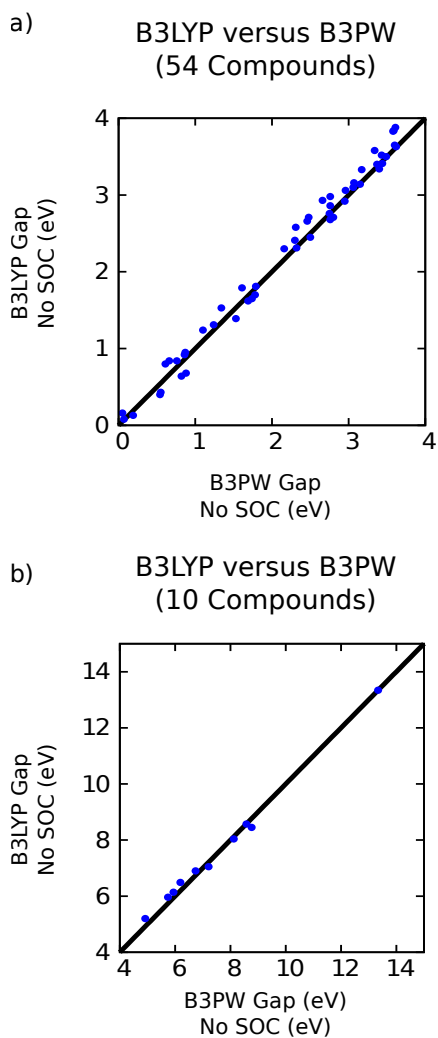
Figure 4.2: B3LYP band gaps excluding SOC versus B3PW band gaps excluding SOC for a) 54 compounds with experimental band gaps from 0 to 4 eV and b) 10 compounds with experimental band gaps from 4 to 15 eV. Exactly the same basis sets and crystal structures were used for the B3PW and B3LYP calculations. The mean signed difference between B3LYP and B3PW is 0.05 eV. The mean unsigned difference is 0.14 eV.

Figure 4.3: M06L band gaps excluding SOC versus B3PW band gaps excluding SOC for a) 54 compounds with experimental band gaps from 0 to 4 eV and b) 10 compounds with experimental band gaps from 4 to 15 eV. Exactly the same basis sets and crystal structures were used for the B3PW and M06L calculations. The mean signed difference between M06L and B3PW is -0.94 eV. The mean unsigned difference is 0.94 eV.
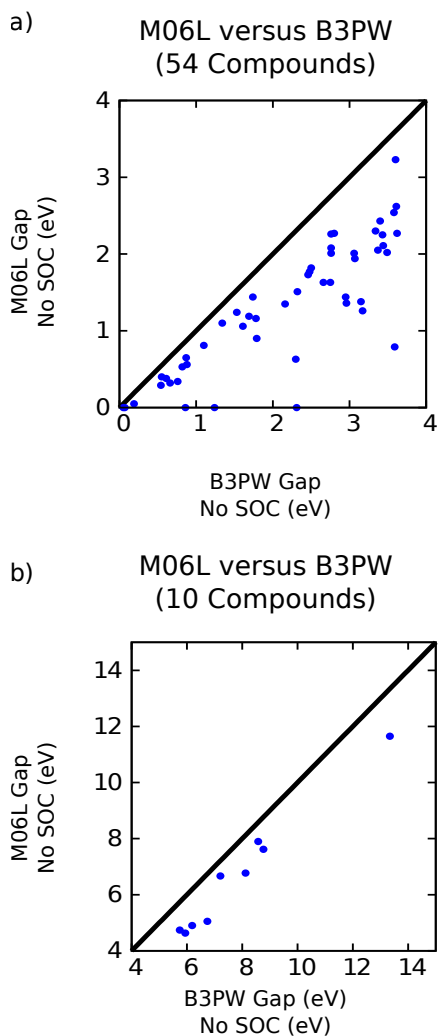
Figure 4.4: HSE06 band gaps excluding SOC versus B3PW band gaps excluding SOC for a) 54 compounds with experimental band gaps from 0 to 4 eV and b) 10 compounds with experimental band gaps from 4 to 15 eV. Exactly the same basis sets and crystal structures were used for the B3PW and HSE06 calculations. The mean signed difference between HSE06 and B3PW is -0.26 eV. The mean unsigned difference is 0.26 eV.
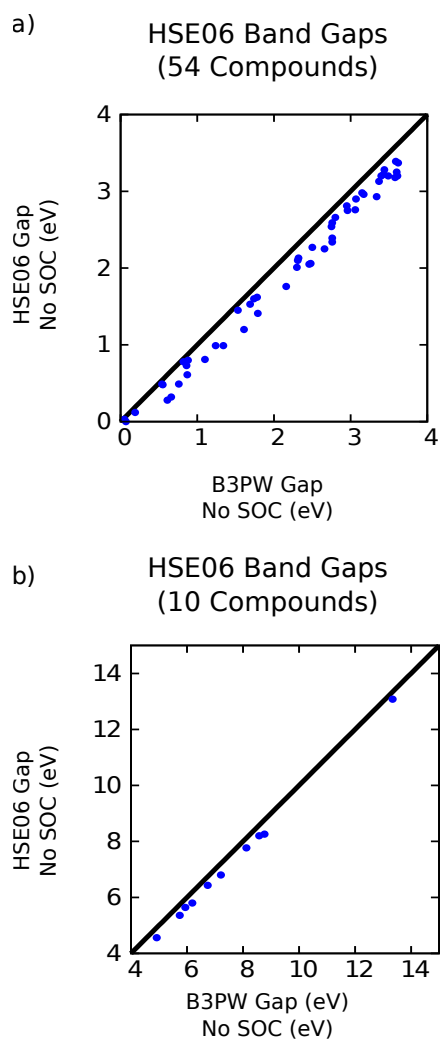
Figure 4.5: PBE0 band gaps excluding SOC versus B3PW band gaps excluding SOC for a) 54 compounds with experimental band gaps from 0 to 4 eV and b) 10 compounds with experimental band gaps from 4 to 15 eV. Exactly the same basis sets and crystal structures were used for the B3PW and PBE0 calculations. The mean signed difference between PBE0 and B3PW is +0.32 eV. The mean unsigned difference is 0.32 eV.
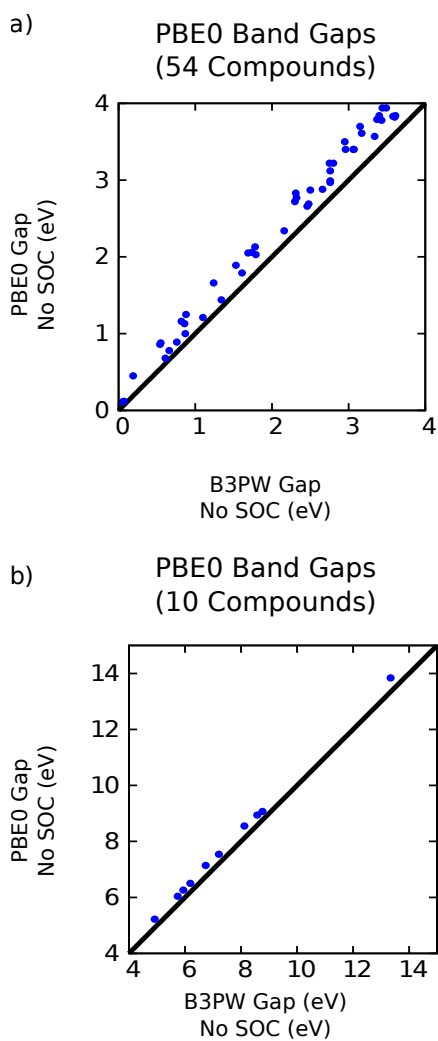
Table 4.2: Band Gaps (eV) Computed with B3PW and B3LYP with and without spin-orbit coupling (SOC). Abbreviations are defined in Table 4.3.

System	B3PW w/ SOC	B3PW No SOC	B3LYP w/ SOC	B3LYP No SOC	System	B3PW w/ SOC	B3PW No SOC	B3LYP w/ SOC	B3LYP No SOC
Bi	0.00	0.07	0.02	0.08	AgBr	2.24	2.32	2.23	2.31
Bi_2Se_3	0.12	0.76	0.16	0.84	ZnSe	2.69	2.80	2.65	2.71
PbSe	0.25	0.66	0.38	0.84	AgI	3.13	3.40	3.09	3.34
Bi_2Te_3	0.23	0.87	0.22	0.95	SiC (6H)	3.34	3.34	3.59	3.58
PbTe	0.59	1.34	0.66	1.53	CuBr	2.87	2.95	2.84	2.92
InSb	0.34	0.54	0.26	0.40	CuI	3.21	3.44	3.20	3.41
Sb_2Te_3	0.22	0.05	0.18	0.16	CoO	3.60	3.59	3.84	3.84
HgTe	0.01	0.19	0.00	0.13	AgCl	2.63	2.75	2.66	2.76
SnTe	0.44	0.61	0.56	0.80	SiC (4H)	3.58	3.58	3.83	3.83
InAs	0.47	0.55	0.42	0.43	GaN (zb)	3.08	3.07	3.24	3.16
VO_2	1.25	1.24	1.31	1.31	$SrTiO_3$	3.38	3.37	3.40	3.40
InN	0.87	0.86	0.93	0.92	TiO_2 (r)	3.47	3.49	3.50	3.50
Ge	0.86	0.88	0.72	0.68	SiC (2H)	3.61	3.61	3.88	3.88
GaSb	0.66	0.82	0.54	0.64	CuCl	3.06	3.15	3.06	3.14
SnSe	0.95	1.10	1.04	1.24	TiO_2 (a)	3.60	3.62	3.63	3.63
Si	1.61	1.61	1.79	1.79	ZnO	2.82	2.96	2.90	3.06
MoS_2	1.63	1.79	1.71	1.81	GaN	3.43	3.43	3.55	3.52
InP	1.78	1.74	1.74	1.65	MnO	3.17	3.17	3.33	3.33
GaAs	1.48	1.53	1.40	1.39	MgTe	3.31	3.60	3.42	3.65
CdTe	1.53	1.78	1.50	1.70	ZnS	4.08	4.07	4.07	4.00
AlSb	1.83	2.16	2.01	2.30	NiO	4.62	4.62	4.56	4.07
CdSe	1.57	1.69	1.56	1.62	CuSCN	3.97	4.08	3.94	4.56
BP	2.48	2.48	2.71	2.71	AlN (zb)	4.92	4.92	5.20	5.20
Cu_2O	2.20	2.30	2.29	2.41	C	5.74	5.74	5.96	5.96
AlAs	2.37	2.46	2.63	2.66	AlN (w)	5.94	5.94	6.14	6.14
GaP	2.71	2.76	2.88	2.86	BN	6.19	6.19	6.49	6.49
ZnTe	2.55	2.76	2.50	2.68	MgO	6.74	6.74	6.90	6.90
FeO	2.33	2.31	2.58	2.58	NaCl	7.20	7.21	7.05	7.05
$BiVO_4$	2.98	3.06	3.02	3.10	SiO_2 (b)	8.77	8.77	8.46	8.45
SiC (3C)	2.66	2.66	2.93	2.93	LiCl	8.57	8.58	8.57	8.57
AlP	2.76	2.76	2.98	2.98	SiO_2 (a)	8.12	8.12	8.05	8.04
CdS	2.49	2.50	2.50	2.45	LiF	13.33	13.34	13.35	13.34

unsigned difference is 0.32 eV. Contrary to HSE06, for each of the 64 compounds tested, PBE0 returned a higher band gap than B3PW. In cases where B3PW overestimates band gaps, such as transition metal oxides, B3PW will likely give better agreement with experiment. By contrast, in cases where B3PW underestimates band gaps, such as alkali halides, PBE0 will likely give better agreement with experiment. This is consistent with the results of "dielectric-dependent" hybrids [83], where the amount of exact exchange varies for different systems. In general, it is found that a higher fraction of exact exchange is necessary for materials with high band gaps.

PBE0 is the limiting case of HSE06 where exact exchange is considered over all of space, rather than just a short range. Since HSE06 gives lower gaps than B3PW and PBE0 gives higher gaps, there is some parameterization of an HSE-type functional

Figure 4.6: M06 band gaps excluding SOC versus B3PW band gaps excluding SOC for a) 54 compounds with experimental band gaps from 0 to 4 eV and b) 10 compounds with experimental band gaps from 4 to 15 eV. Exactly the same basis sets and crystal structures were used for the B3PW and M06 calculations. The mean signed difference between PBE0 and B3PW is +0.33 eV. The mean unsigned difference is 0.34 eV.

a)

M06 Band Gaps
(54 Compounds)

b)

M06 Band Gaps
(10 Compounds)
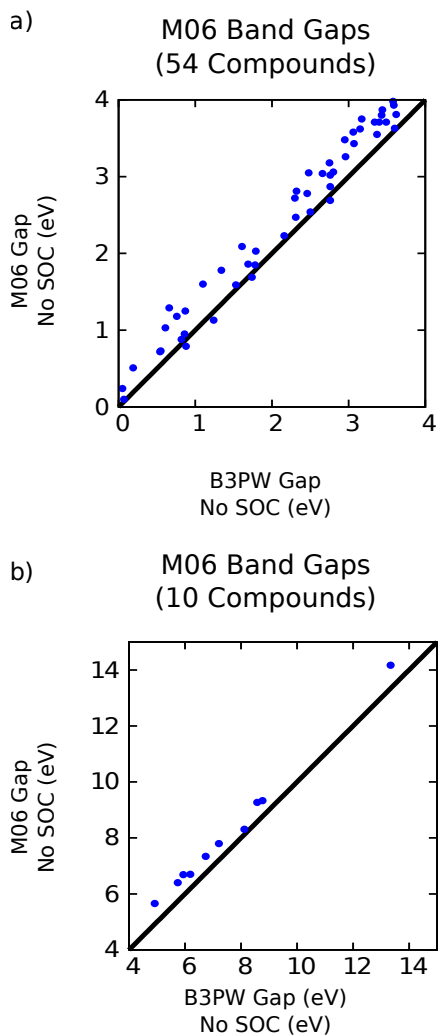
that will return band gaps very close to the B3PW gaps. This functional would have the advantage of having near zero mean error as B3PW does, but may be more computationally affordable than global hybrids in a plane-wave code. We note that this was done by Moussa et al. in 2012 [47], who indeed found a set of parameters for an "HSE12" functional that gave zero mean error.
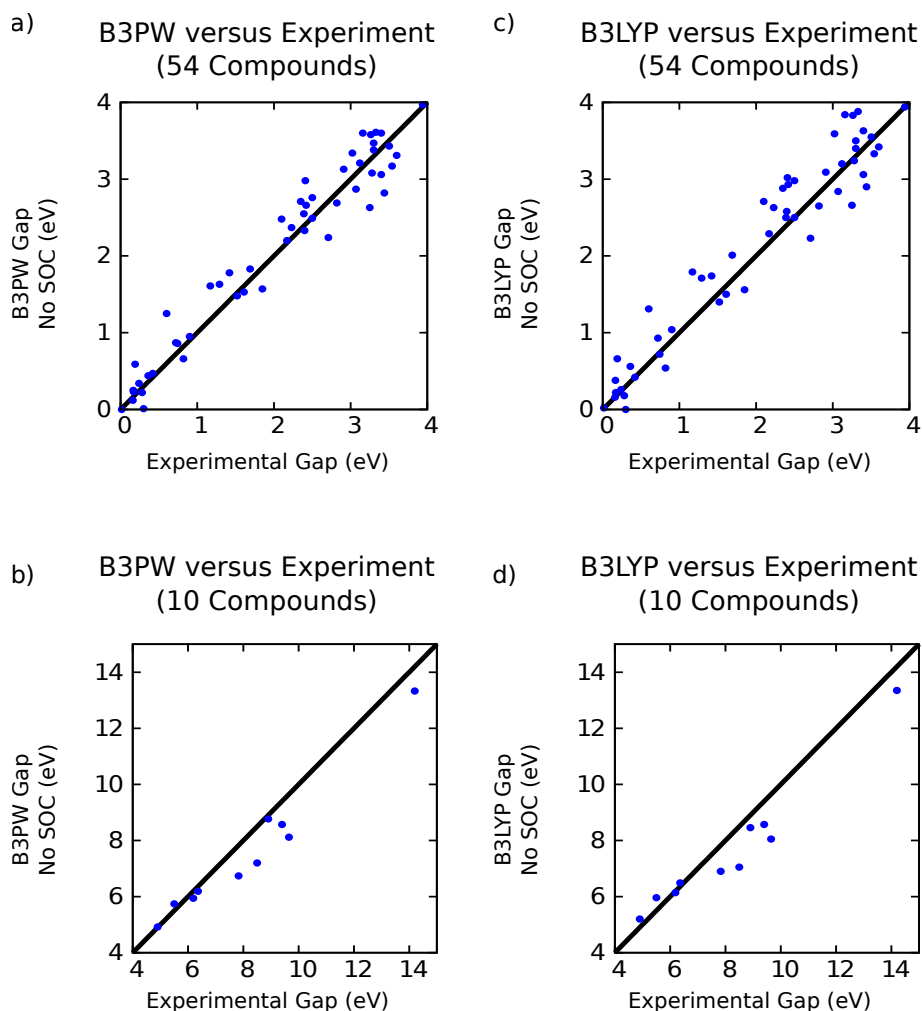
M06

Figure 4.3 shows non-spin-orbit M06 gaps versus non-spin-orbit B3PW band gaps. The mean signed difference between M06 and B3PW is +0.33 eV, and the mean unsigned difference is 0.34 eV. For 4 of the 64 cases tested, M06 returned a lower band gap than B3PW. In the remaining 60 systems, M06 predicted a larger band gap. These results suggest that the emphasis on molecular thermochemistry in the parametrization of M06 limits its applicability for band gap prediction.

4.4   Conclusions

For exact exchange, Gaussian basis sets have a significant advantage in computational speed compared to plane waves [84], plane wave basis sets remain by far the more popular choice in solid state calculations. In plane-wave codes, range-separated functionals like HSE are less computationally demanding than global hybrids. The results presented here suggest that band gap predictions in plane wave codes could be improved by increasing the fraction of exact exchange or the distance over which the exact exchange is evaluated.

The band gaps computed with all of the hybrid functionals considered here are surprisingly similar. Hybrid functionals with less exact exchange than B3PW return band gaps that are lower than the corresponding B3PW value by a nearly constant amount. Similarly, functionals with more exact exchange return band gaps that are higher by a nearly constant amount. While the ideal amount of exact exchange varies from one system to another [83], the results presented here suggest that the treatment of exact exchange is a well-balanced choice, neither systematically overestimating nor underestimating band gaps. Given that the mean signed error comparing B3PW with SOC to experiment is very close to zero (-0.03) and the mean signed error for B3LYP with SOC is +0.03, it does not seem likely that much improvement in band gap predictions can be achieved by simply altering the (fixed) amount of exact exchange in the functional. Dielectric-dependent hybrids that use a system-dependent amount of exact exchange may perform better than B3PW, but involve additional complexity. Therefore, developing a hybrid density functional that improves B3PW's band gap predictions would involve choosing or developing a correlation functional to replace PW91.

Table 4.3: Band gaps (eV) computed without spin-orbit coupling (SOC) using the B3PW, B3LYP, M06L, HSE06, PBE0, and M06 density functionals. Abbreviations are defined in Table 4.3.

System	B3PW	B3LYP	M06L	HSE	PBE0	M06	System	B3PW	B3LYP	M06L	HSE	PBE0	M06
Bi	0.07	0.08	0.00	0.00	0.12	0.10	AgBr	2.32	2.31	1.51	2.13	2.77	2.81
Bi_2Se_3	0.76	0.84	0.34	0.49	0.89	1.18	ZnSe	2.80	2.71	2.27	2.66	3.22	3.06
PbSe	0.66	0.84	0.32	0.32	0.78	1.29	AgI	3.40	3.34	2.43	3.20	3.84	3.71
Bi_2Te_3	0.87	0.95	0.65	0.61	1.00	1.25	SiC (6H)	3.34	3.58	2.30	2.93	3.57	3.71
PbTe	1.34	1.53	1.10	0.99	1.44	1.78	CuBr	2.95	2.92	1.44	2.81	3.50	3.48
InSb	0.54	0.40	0.29	0.49	0.86	0.72	CuI	3.44	3.41	2.11	3.28	3.94	3.87
Sb_2Te_3	0.05	0.16	0.00	0.03	0.11	0.24	CoO	3.59	3.84	0.79	3.39	4.10	3.93
HgTe	0.19	0.13	0.05	0.12	0.45	0.51	AgCl	2.75	2.76	1.63	2.54	3.22	3.18
SnTe	0.61	0.80	0.38	0.28	0.68	1.03	SiC (4H)	3.58	3.83	2.54	3.18	3.83	3.98
InAs	0.55	0.43	0.40	0.48	0.88	0.73	GaN (zb)	3.07	3.16	1.94	2.90	3.40	3.43
VO_2	1.24	1.31	0.00	0.99	1.66	1.13	$SrTiO_3$	3.37	3.40	2.05	3.13	3.79	3.55
InN	0.86	0.92	0.00	0.73	1.13	0.95	TiO_2 (r)	3.49	3.50	2.02	3.20	3.94	3.71
Ge	0.88	0.68	0.56	0.80	1.25	0.79	SiC (2H)	3.61	3.88	2.62	3.20	3.84	4.08
GaSb	0.82	0.64	0.53	0.78	1.16	0.88	CuCl	3.15	3.14	1.38	2.98	3.70	3.62
SnSe	1.10	1.24	0.81	0.81	1.21	1.60	TiO_2 (a)	3.62	3.63	2.27	3.37	4.07	3.81
Si	1.61	1.79	1.06	1.20	1.79	2.09	ZnO	2.96	3.06	1.36	2.75	3.40	3.26
MoS_2	1.79	1.81	0.90	1.41	2.03	2.03	GaN	3.43	3.52	2.25	3.22	3.78	3.80
InP	1.74	1.65	1.44	1.60	2.06	1.69	MnO	3.17	3.33	1.26	2.96	3.61	3.75
GaAs	1.53	1.39	1.24	1.45	1.89	1.59	MgTe	3.60	3.65	3.23	3.25	3.82	3.63
CdTe	1.78	1.70	1.16	1.62	2.13	1.85	ZnS	4.07	4.00	3.35	3.86	4.50	4.17
AlSb	2.16	2.30	1.35	1.76	2.34	2.23	NiO	4.62	4.07	2.71	4.59	5.14	4.95
CdSe	1.69	1.62	1.19	1.53	2.05	1.86	CuSCN	4.08	4.56	2.14	3.76	4.46	4.41
BP	2.48	2.71	1.77	2.06	2.69	3.05	AlN (zb)	4.92	5.20	3.61	4.56	5.22	5.66
Cu_2O	2.30	2.41	0.63	2.01	2.72	2.72	C	5.74	5.96	4.74	5.36	6.04	6.40
AlAs	2.46	2.66	1.73	2.05	2.66	2.78	AlN (w)	5.94	6.14	4.63	5.64	6.26	6.69
GaP	2.76	2.86	2.08	2.39	2.99	2.87	BN	6.19	6.49	4.90	5.80	6.50	6.70
ZnTe	2.76	2.68	2.26	2.59	3.12	2.69	MgO	6.74	6.90	5.05	6.43	7.14	7.34
FeO	2.31	2.58	0.00	2.10	2.83	2.47	NaCl	7.21	7.05	6.67	6.80	7.54	7.80
$BiVO_4$	3.06	3.10	2.01	2.76	3.40	3.58	SiO_2 (b)	8.77	8.45	7.62	8.26	9.07	9.33
SiC (3C)	2.66	2.93	1.63	2.25	2.88	3.04	LiCl	8.58	8.57	7.90	8.20	8.94	9.27
AlP	2.76	2.98	2.01	2.34	2.97	3.02	SiO_2 (a)	8.12	8.04	6.77	7.77	8.55	8.31
CdS	2.50	2.45	1.82	2.27	2.87	2.54	LiF	13.34	13.34	11.65	13.08	13.84	14.17

C h a p t e r 5

DO HYBRID FUNCTIONALS REALLY FAIL FOR METALS?

All of the systems considered thus far have had bulk band gaps. For metallic systems, long-range Hartree-Fock exchange rigorously causes an unphysical logarithmic singularity in the Fermi velocity [85]. This is standard enough to be included in basic solid-state physics textbooks [86]. For the Hartree-Fock Hamiltonian,

$$\epsilon(k) = \frac{\hbar^2 k^2}{2m} - \frac{2e^2}{\pi} k_{fermi} F\left(\frac{k}{k_{fermi}}\right),$$

where

$$F(x) = \frac{1}{2} + \frac{1 - x^2}{4x} ln \left|\frac{1 + x}{1 - x}\right|.$$

The logarithmic singularity arises from the long-range part of the exchange operator. It is often used as a reason to avoid global hybrid functionals like B3PW, and is central to the rationale behind the HSE functional [87]. However, calculations of the Hartree-Fock density of states by Fry, Brener, and Bruyere [88] did not display the theoretically required zero density of states at the Fermi level for the nearly-free electron gas. They computed the density of states using the tetrahedron method on a grid of up to 1505 k points.

The prevailing attitude in the literature echoes the opinion of Monkhorst [85]:

"The fact that numerical inaccuracies tend to wash out these HF-caused singularities (or near sinularities) is irrelevant, although expedient in practice. The point is that the theory allows for them, and therefore they should be either shown faithfully or corrected for."

Moussa et al. [47], state that "...exact exchange interactions in three-dimensional metals are doomed to fail".

In this brief chapter, we show that any practical Hartree-Fock calculation of a metal will not display this theoretically required pathology. We extended the study of Fry et al. by computing the rigorous Hartree-Fock density of states using up to 8 billion k points. The same tetrahedral interpolation method was used. The results are shown in Figure 5.
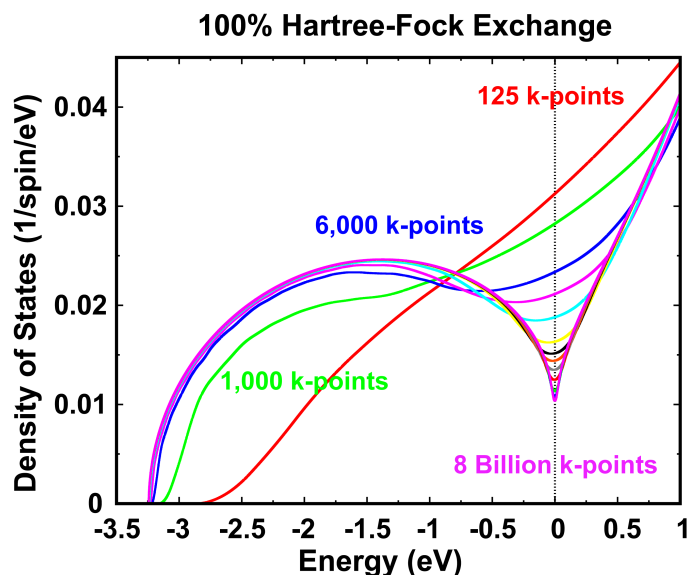
Figure 5.1: Hartree-Fock density of states evaluated by tetrahedral interpolation.

For an extremely large number of k points, the effect of the logarithmic singularity on the density of states can be observed. However, even for 8 billion k points, many orders of magnitude more than can be used in any practical calculation, the theoretically mandated zero of the density of states at the Fermi level has not yet appeared. This is the worst case scenario for exact exchange because the logarithmic singularity is manifest in the equation being evaluated. In a calculation of a real system, this is not the case.

We now wish to study the Hartree-Fock density of states of some real metals in order to determine the effect of the "pathology" on real calculations of metals. We performed B3PW CRYSTAL14 (unmodified, no spin-orbit coupling) calculations on sodium metal using the 6-311G* basis set. Calculations were performed using 413, 8797, and 13728 irreducible k points. No smearing was employed, which would artificially remove the singularity. The latter requires approximately 8 gigabytes of memory, and represents a far larger calculation than would performed under normal practical circumstances. The density of states for each of these calculations was calculated using CRYSTAL and an 81x81x81 k-point grid for interpolation. This corresponds to approximately 12000 irreducible k points, and is again many more k points than would be used normally. As shown in Figure 5, the three densities of states are indistinguishable. No evidence of a singularity is seen.

As a second example, B3PW calculations of the density of states of copper were performed using 72, 413, and 2769 irreducible k points for the SCF calculation. For
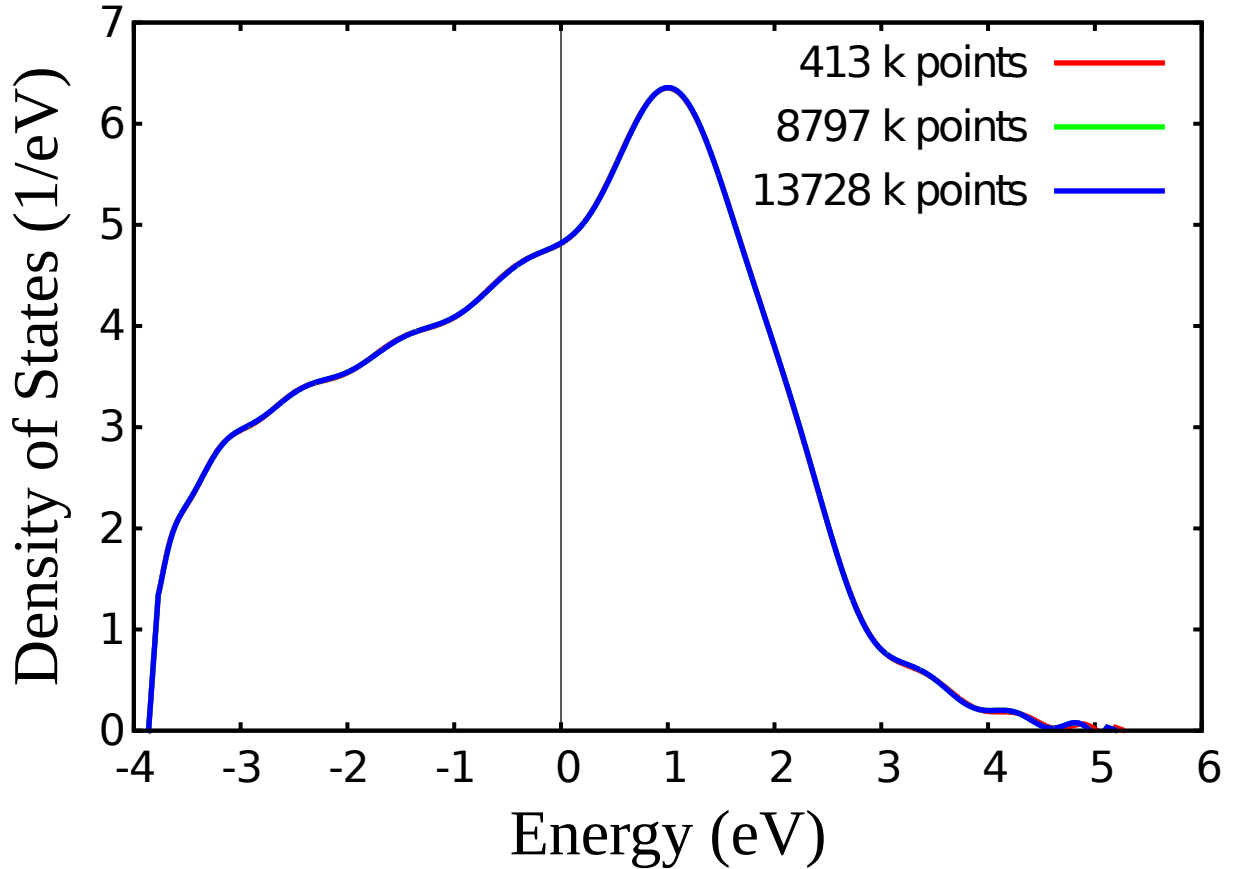
Figure 5.2: Density of states for sodium metal computed with B3PW with 413, 8797, and 13728 irreducible k points. The Fermi energy is set to 0, shown by the vertical line. For all three density of states calculations, an 81x81x81 k-point grid was used for the interpolation. The differences among the three densities of states are negligible, and no evidence of the theoretically mandated singularity is observed.

all three calculations, an 81x81x81 k-point grid was used in the density of states interpolation. The SBKJC [26] pseudopotential and basis set were used for copper. The same 81x81x81 k-point grid was used for the interpolation in the density of states calculation. The results are shown in Figure 5. Again, no evidence of a singularity is observed as the number of k points in the SCF calculation is increased.

Finally, to investigate the effect of the number of k points used for the interpolation in a density of states calculation, the density of states for the same copper calculation (using 2769 irreducible k points) was calculated using 72, 413, 2769, 8797, 13728 irreducible k points. The results are shown in Figure 5. CRYSTAL uses Legendre polynomials for the k space interpolation. The "wiggles" in the densities of states seen in Figure 5 even for the densest grid used are the result of this poly-
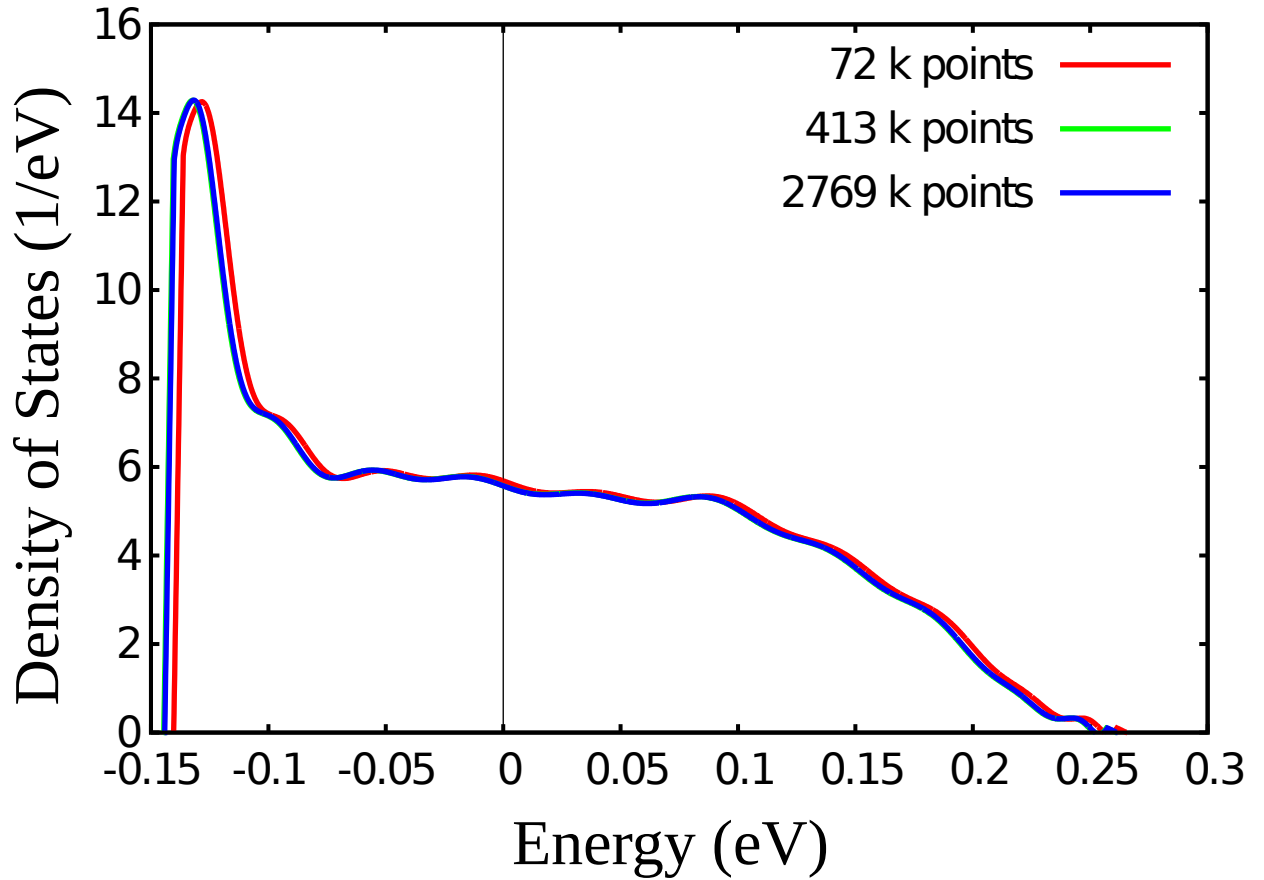
Figure 5.3: Density of states for copper metal computed with B3PW using 72, 413, and 2769 irreducible k points. The Fermi energy is set to 0, shown by the vertical line. For all three density of states calculations, an 81x81x81 k-point grid was used for the interpolation.

nomial interpolation. However, in this case as with Figure 5, no singularity can be observed.

The long-range part of the exact exchange rigorously requires the density of states of a metal to vanish at the Fermi energy. This fact is not disputed here. However, we find that even for the largest calculations we could afford, this singularity could not be observed in hybrid density functional calculations of sodium or copper. Our results do not remove the theoretical problem. Instead, they show that for any practical calculation of a metal, the logarithmic singularity manifests itself too slowly to be observed.

We philosophically disagree with Monkhorst that theoretical rigor must always be prioritized over practical efficacy. As famously stated by R.W. Hamming [89], "The purpose of computing is insight, not numbers". (The more cynical reader may pre-
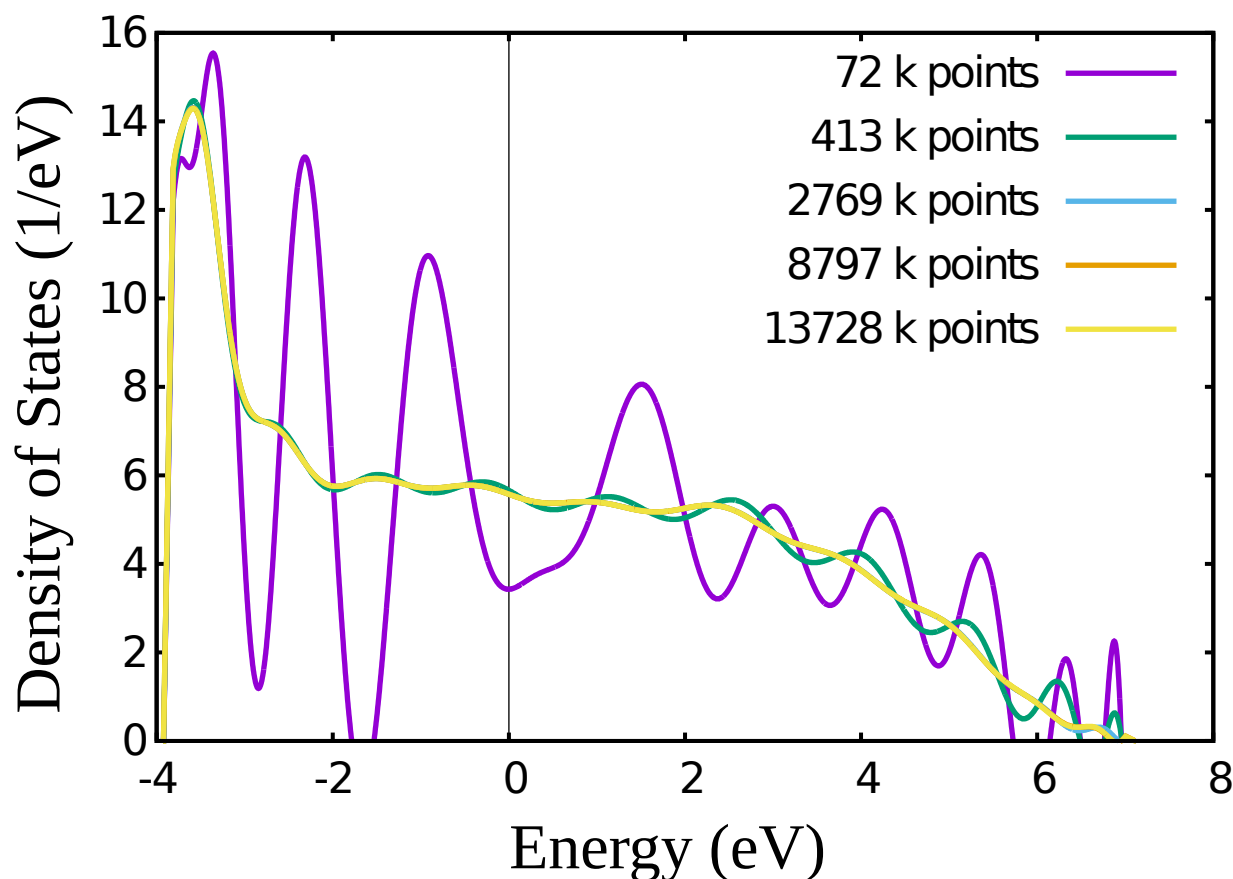
Figure 5.4: Density of states for copper metal computed with B3PW using 72, 413, 2769, 8797, and 13728 irreducible k points for the interpolation. For the SCF calculation, 2769 irreducible k points were used.

fer the 1984 Macintosh's admonition that one should "never trust a computer that you can't lift".) There is absolutely no theoretical justification for the excellent performance of hybrid density functionals (or any density functional) for band gaps. As is well-known, Kohn-Sham eigenvalues are not physical, and their differences have no physical meaning. Even worse, global hybrid density functionals like B3PW have parameters that were fit using experimental data (although we emphasize here that no properties of solids were used in these fits). Consider the stark contrast to *GW* methods, with their exact quasiparticle excitation spectra. Despite the unforgivable lack of theoretical rigor in B3PW and the obvious superiority of *GW*, the engineered density functional is, at present, more accurate for band gap prediction. Of course, this is not an argument to eschew theory entirely in favor of blind empiricism, but it would be foolish to ignore the excellent performance of B3PW for

band gaps prediction.

Chapter 6

ACCURATE AB INITIO QUANTUM MECHANICS SIMULATIONS OF BI$_2$SE$_3$ AND BI$_2$TE$_3$ TOPOLOGICAL INSULATOR SURFACES

Topological insulators are bulk insulators for which a strong spin-orbit interaction inverts the orbital character of the conduction and valence bands. In a topological insulator, there exist surface states at all energies within the bulk band gap. These surface states have a linear dispersion with respect to the surface momentum **k**, and the spin polarization varies with **k** [90, 91]. Angle-resolved photoemission experiments on finite slabs observing the linear dispersion [92, 93] show that bulk Bi$_2$Te$_3$ and Bi$_2$Se$_3$ are topological insulators.

In order to better understand the nature of these topological insulators, it is essential to determine the electronic structure. This requires an accurate prediction of band gaps (including relativistic effects) with computational efficiency sufficient for realistic surfaces and interfaces. It is generally accepted that the ideal *ab initio* calculation would be fully self-consistent *GW* [1] because it rigorously approximates the true quasiparticle energies. However, *GW* is computationally impractical even for bulk Bi$_2$Te$_3$ and Bi$_2$Se$_3$. A non-self-consistent *GW* approximation, G_0W_0, can be performed for bulk materials, but it is also impractical for slab calculations to examine the surface states. Consequently, comparisons of G_0W_0 and density functional theory (DFT) calculations on bulk solids are used to empirically correct the electronic states of the surface [94]. We refer to this approach as EC-LDA ("empirically-corrected LDA").

Density functional (DFT) methods based on the local density approximation (LDA) or the generalized gradient approximation of Perdew, Burke, and Ernzerhof (PBE) [41] are practical for computation of realistic surfaces. Unfortunately, these methods underestimate band gaps in solids because of a discontinuity of the derivative of the energy with respect to the number of electrons [43, 44], the electron self Coulomb repulsion, and the nonlinear dependence of energy on number of electrons [45]. This underestimate is especially problematic in small-gap semiconductors with strong spin-orbit coupling. Viewed as a first-order perturbation, spin-orbit coupling tends to decrease band gaps by lifting spin degeneracies in the valence and

conduction bands. If the *ab initio* method already underestimates band gaps without spin-orbit coupling, then including it may cause an unphysical band inversion, or a "false positive" prediction that a material is a topological insulator [95].

Hybrid density functionals include a fraction of exact Hartree-Fock exchange, which allows much of the self Coulomb repulsion error to be eliminated [68]. This leads to much more accurate reaction barriers so that the B3LYP hybrid method has been the *de facto* standard DFT approach in molecular computational chemistry for decades. Computing exact exchange is impractical for codes (such as VASP or Quantum Espresso) that use plane wave basis sets. This has lead to increased popularity of the hybrid HSE functional [96], which only computes the exact exchange operator over a short range. While still costly, this approach has been applied to some topological insulators using plane-wave codes [97, 98].

In contrast, with Gaussian basis sets, hybrid functionals such as B3PW91 [71] (referred to as B3PW in this paper for brevity) involve computational costs comparable to PBE. We have shown elsewhere [68] that the B3PW hybrid functional used in this paper gives excellent agreement (mean absolute error = 0.09 eV) with experimental band gaps across a wide range of semiconductors.

In this chapter, we show that the B3PW hybrid DFT method leads to excellent agreement with the best available G_0W_0 calculations of the bulk band structure for the topological insulators Bi_2Te_3 and Bi_2Se_3. We also report slab calculations on Bi_2Te_3 and Bi_2Se_3 using B3PW. We find that B3PW, with no empirical corrections, is in excellent agreement with the EC-LDA slab calculations. This result is the main point of this chapter.

Figure 6.1 shows the crystal structure and Brillouin zone of bulk Bi_2Te_3 and a two quintuple layer slab. Both Bi_2Te_3 and Bi_2Se_3 crystallize in a layered structure composed of repeating quintuple layers (QLs). Each quintuple layer is composed of three bismuth and two tellurium layers, with the Bi and Te layers alternating. Within a QL, atoms in adjacent layers are covalently bound. In contrast, there is only van der Waals interaction between QLs.

In order to assess the quality of our hybrid approach, we compared the results of bulk band structure calculations of Bi_2Te_3 and Bi_2Se_3 with PBE and G_0W_0 [99]. Our PBE calculations are in good agreement with previously published calculations [100–102]. The G_0W_0 calculation we chose for comparison was performed by Aguilera *et al.*[99]. They used an explicitly spin-dependent G_0W_0 approxima-
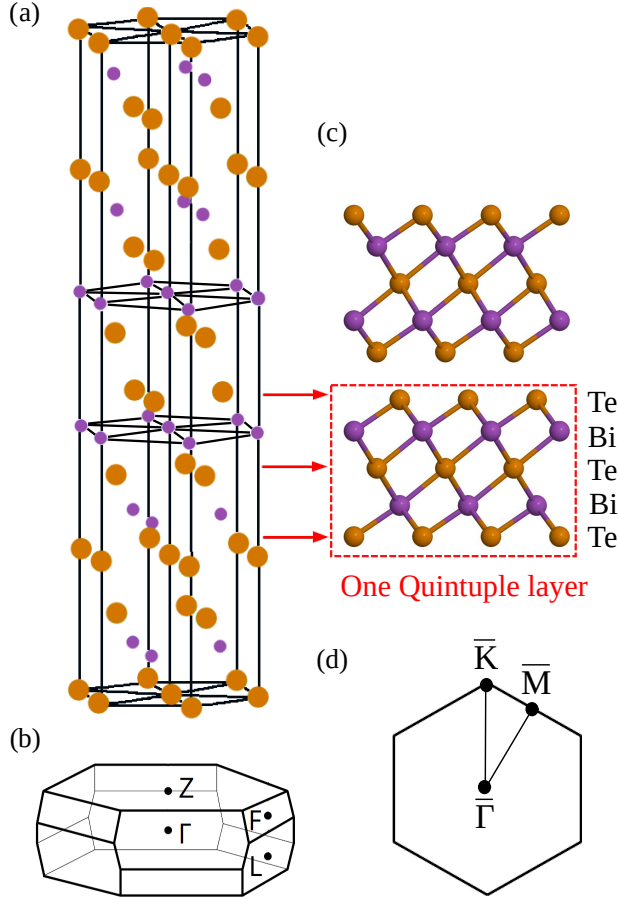
Figure 6.1: (a) Bulk crystal structure of Bi_2Te_3. Te and Bi atoms are colored orange and purple, respectively. (b) Bulk Brillouin zone of Bi_2Te_3 showing high-symmetry k-points. (c) Side view of a two-quintuple-layer slab of Bi_2Te_3. The red box indicates a single quintuple layer, and the arrows show the quintuple layer within the bulk structure. (d) Brillouin zone of a two-dimensional slab of Bi_2Te_3. The $\bar{\Gamma}$ point is the zone center.

tion, thereby fully accounting for the spin-orbit coupling rather than adding it as a perturbation to a scalar-relativistic calculation. The comparison is shown in Figure 6.2.

Physically, the most important energy regime is near the conduction and valence band edges (in the vicinity of the Γ point). Here, B3PW is in excellent agreement with G_0W_0. In contrast, PBE fails to obtain the correct band structure in the valence and conduction bands at the Γ point. In the valence band for both materials, PBE always obtains a pronounced "m" shape that is not seen for Bi_2Se_3 and appears weakly for Bi_2Te_3. Additionally, B3PW and G_0W_0 predict a direct band gap for Bi_2Se_3 and an indirect gap for Bi_2Te_3 (between the Z and F points in the Brillouin

(a)

(b)

(c)

(d)
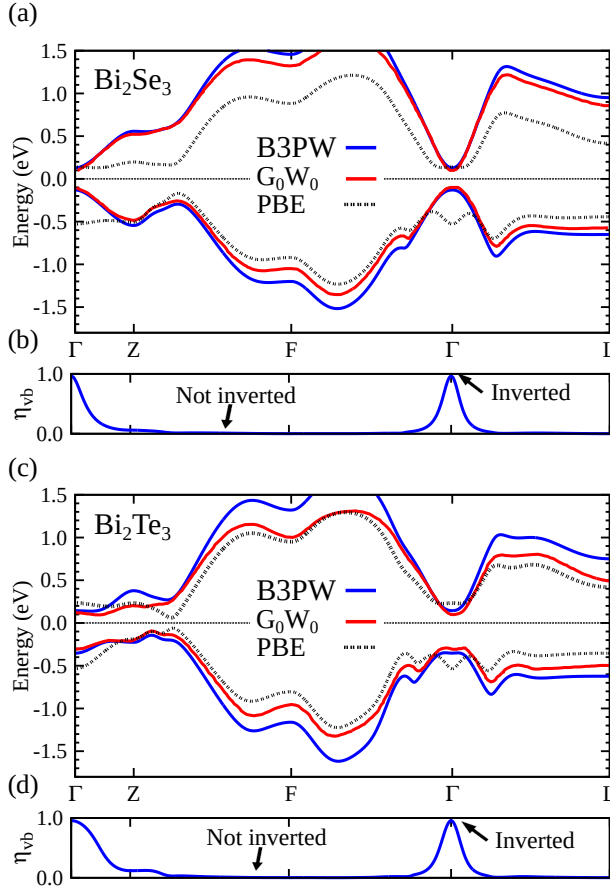
Figure 6.2: Comparison of the B3PW hybrid density functional (blue line) to high-quality G_0W_0 (red line) calculations of the band structures of (a) Bi_2Se_3, and (c) Bi_2Te_3. The Fermi level is set to zero. The black dashed line is the PBE result.

zone), in agreement with recent photoemission experiments [103, 104]. On the other hand, PBE finds an indirect band gap for both materials.

The excellent match we find between B3PW and the G_0W_0 valence band structure does not prove the viability of hybrid functionals. We must also ascertain if a band inversion has occurred. Thus, we computed the degree of valence band inversion, using a previously defined expression [94]. The valence band inversion parameter, η_{vb}, is equal to one when the valence band is completely inverted. We find exactly the same inversion at the Γ point as G_0W_0, as shown in Figure 6.2.

Having established the accuracy of the B3PW functional for bulk band structures, we now turn to calculations for realistic slabs. A direct calculation of these slabs requires up to 35 atoms per periodic cell (whereas only 5 are required for the bulk system), making G_0W_0 impractical for the slabs. Yazyev *et al.* [94] empirically cor-

rected LDA calculations to estimate G_0W_0 results for slabs (EC-LDA). In contrast, using the B3PW functional with a Gaussian basis set is quite practical for slabs.

We computed the $\bar{\Gamma}$-point band gap for slabs with 1 to 6 QLs (Figure 6.3) in order to determine the fewest number of quintuple layers that leads to a topological insulator. Both EC-LDA and B3PW hybrid calculations agree well with each other and the experimental results [105] for Bi_2Se_3. While the experiment found closure of the $\bar{\Gamma}$ energy gap at a minimum thickness of six QLs, both EC-LDA and B3PW did not. For B3PW, the mean absolute error in band gaps was found to be 0.09 eV in ref [68]. We find band gaps below this mean absolute error for 6QL and 7QL slabs (0.08 and 0.07 eV, respectively). We believe a Dirac point exists at 6 QLs within our computational error. Our PBE calculations, as well as those of reference [106], predict zero gap by three QLs.

Relative to Bi_2Se_3, all three methods predict a faster decrease of the energy gap with film thickness in Bi_2Te_3. While the band gap is smaller than our estimated computational error by 2QLs, we see a true gap closing in Bi_2Te_3 slabs composed of 5 or more QLs, in contrast to EC-LDA. There is currently no experimental data for Bi_2Te_3 for QLs.

The B3PW band structures of all QL slabs for both materials are shown in Appendix B. There is a dramatic difference between the B3PW and EC-LDA band structures for 5QL Bi_2Te_3. The B3PW band structure has a clear Dirac cone, whereas the empirically corrected band structure of Yazyev *et al.* [94] does not. A true G_0W_0 calculation of a 5QL Bi_2Te_3 would most likely give a similar band structure to B3PW. Therefore, we believe this is the first time an *ab initio* method has simultaneously provided accurate band gaps and demonstrated the onset of topological insulating behavior in a Bi_2Te_3 slab. At $\bar{\Gamma}$, we find a Fermi velocity of 4×10^5 m/s, in agreement with experiment [107].

For 5QL Bi_2Se_3, the B3PW and EC-LDA band structures are similar. Neither method predicts a Dirac point at 5QL, so in this case the Fermi velocity is zero at $\bar{\Gamma}$. Thus, we estimated the slope of the Dirac cone by fitting to the linear part of our band structure near $\bar{\Gamma}$ (0.006 to 0.05 inverse Bohr along the $\bar{M}$ direction). We find a velocity of 4×10^5 m/s, slightly below the experimental value of 5×10^5 m/s. [108].

We have shown that the hybrid B3PW functional returns *GW*-quality results for band gaps. We emphasize here that hybrid functionals are only practically use-
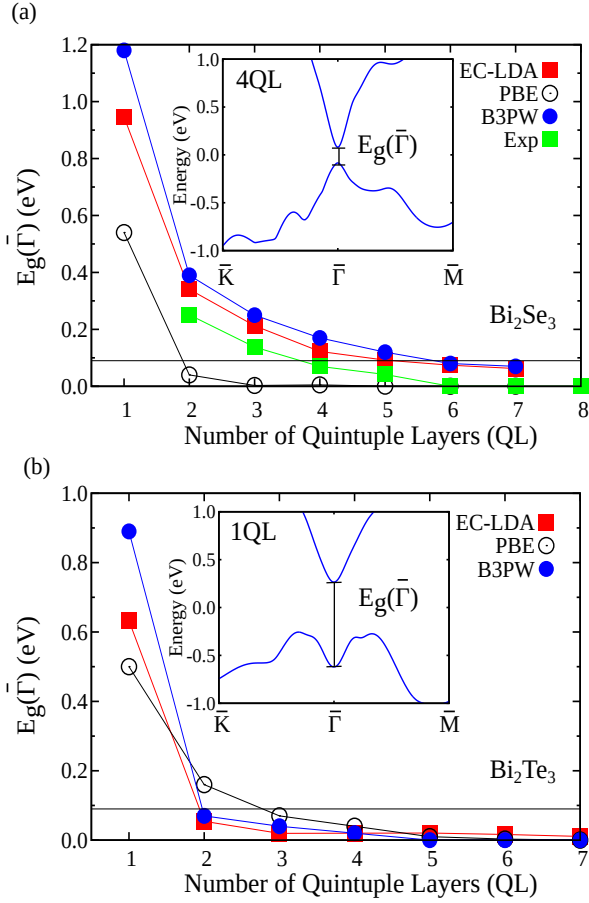
Figure 6.3: Comparison of our B3PW calculations (blue circles), empirically corrected LDA (EC-LDA, red squares) [94], and our PBE (open circles) calculations of the energy gap at $\bar{\Gamma}$, $E_g(\bar{\Gamma})$, for (a) Bi_2Se_3 and (b) Bi_2Te_3. The lines between the points are guides to the eye. The insets show the valence and conduction bands of (a) a four quintuple layer (4QL) slab and (b) a 1QL slab computed with B3PW. B3PW band structures for 1-7QL slabs of Bi_2Se_3 and Bi_2Se_3 are shown in the Supporting Information. The horizontal lines at 0.09 eV indicate the mean absolute error in band gaps computed with B3PW [68]. The green squares are the experimental results [105] for Bi_2Se_3. Currently, there are no experimental results for Bi_2Te_3.

ful for most material systems, and surfaces in particular, with the use of localized Gaussian basis sets. In order to compare the speed of calculations using Gaussians to equivalent calculations using plane waves, we performed B3LYP calculations of a 40-atom single-walled carbon nanotube system using VASP [109] version 5.2.11 and CRYSTAL09. We used 36 k-points for both calculations. The kinetic energy cutoff chosen for the VASP calculation was 400eV; the standard 6-21G* basis set for carbon was used in CRYSTAL. Using 4 CPUs on exactly the same machine, VASP required 900 times more CPU time than CRYSTAL, and 9 times more memory.

With a Gaussian basis set, B3PW delivers comparable results to the highest-level G_0W_0 calculations available for significantly less cost. Indeed, this off-the-shelf functional allows direct computation of systems that are presently beyond the reach of G_0W_0. The results presented here pave the way to calculations of topological insulator surfaces and interfaces with realistic treatments of defects, doping, and surface reconstruction. All of these effects are important to accurate simulations of topological phase transitions and new spintronic devices, and our approach is expected to be effective for other semiconductors with strong spin-orbit coupling.

6.1 Computational Methods

Using the B3PW functional, we calculated the electronic structure of bulk Bi_2Te_3 and Bi_2Se_3 and of slabs with 1 to 6 QLs. All calculations were performed using the CRYSTAL program [3, 110], which we modified to include spin-orbit coupling. We used the fully relativistic large-core pseudopotentials and valence basis sets of Stoll *et al.* for Bi and Te [76], and the small-core pseudopotential and valence basis set of Peterson *et al.* for Se [75]. In the valence basis sets, all exponents smaller than 0.1 were removed to ensure linear independence. A $10 \times 10 \times 10$ k-point grid was used for bulk calculations, and a $10 \times 10 \times 1$ grid was used for the slab calculations. These grids include more k-points than necessary, but the calculations are fast enough to render this overkill irrelevant. All calculations were converged to a 10^{-6} Hartree root mean square difference in Fock matrix elements. Convergence was accelerated using a modified Broyden approach [111]. PBE calculations were also performed using exactly the same basis sets, crystal structures, and computational parameters as our B3PW calculations in order to have a direct comparison of the results of the two functionals. For all the calculations in this paper (bulk and slabs), we find B3PW to be 2-3 times slower than an identical PBE calculation. Experimental crystal structures for Bi_2Se_3 [112] and Bi_2Te_3 [113] were used for both bulk and

(111) slab calculations.

A p p e n d i x   A

SUPPLEMENTAL INFORMATION FOR CHAPTER 3

A.1 Basis Set Recipe

Our basis sets were chosen according to the following systematic recipe:

- For elements up to chlorine, we chose 6-311++G** by default. Because very diffuse basis functions can cause linear dependence in basis sets, basis functions were removed so that there was only one basis function with an exponent less than 0.1. We only retain the basis function with the largest exponent less than 0.1. All basis functions with smaller exponents are removed.

- If this modified basis could be used without numerical (linear dependence) problems, we used it. Otherwise, we replaced the most diffuse exponent with 0.1.

- If this modified basis could be used without numerical (linear dependence) problems, we used it. Otherwise, we replaced the most diffuse exponent with 0.12.

- If this modified basis could be used without numerical (linear dependence) problems, we used it. Otherwise, this basis function was removed.

In several cases, the above procedure did not yield a linearly independent basis set. Thus, we made the following modifications.

- For oxygen, we used the 6-31d1 basis set of Gatti et al [114].

- For LiF, we removed all basis functions with exponents below 0.1 from the aug-cc-PCVTZ basis set.

- For Co, Ni, and Mn, we used 6-31G* and followed the above procedure for treating diffuse exponents.

- We used 6-31G for Fe and followed the above procedure for treating diffuse exponents.

- For all SiC polymorphs, we used 6-31+G** for Si and followed the above procedure for treating diffuse exponents.

- For AlN (zincblende and wurtzite), AlP, AlAs, and AlSb, we used 6-31+G** for Al and followed the above procedure for treating diffuse exponents.

For Sr and all elements from Cu to Bi, we used Stuttgart fully relativistic pseudopotentials and the accompanying cc-PVDZ-level basis sets with the same treatment of diffuse exponents outlined above.

A.2   Comparison of B3PW and G_0W_0
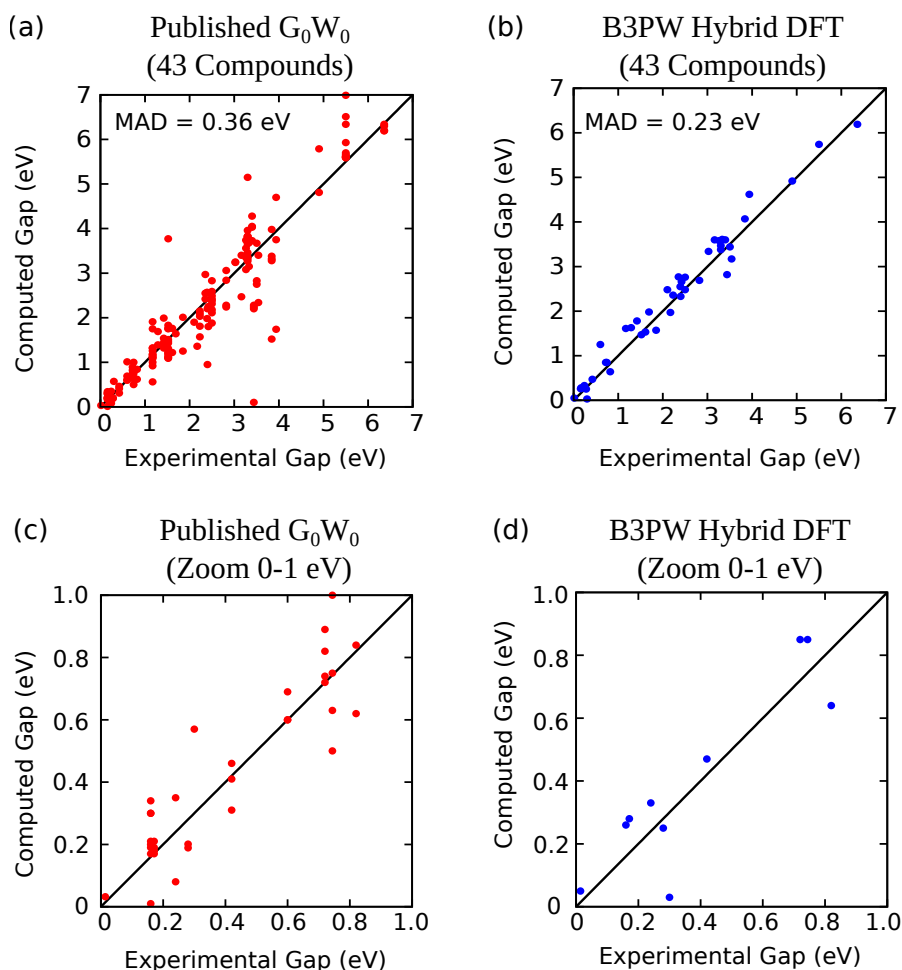
Figure A.1: G_0W_0 and B3PW calculated band gaps versus experiment (low-temperature) for 43 compounds with band gaps below 7 eV. These 43 compounds are those for which our literature search found G_0W_0 results, and are a subset of the 70 compounds listed in Figures 3.2 and 3.3. These 43 compounds are listed in Figure A.2. (a) Published G_0W_0 results (164 data points, 43 compounds, 32 publications). The mean absolute deviation (MAD) is 0.36 eV. (b) B3PW hybrid DFT calculations on the same 43 compounds. The MAD is 0.23 eV. (c) Zoom of (a) from 0–1 eV. (d) Zoom of (b) from 0–1 eV.
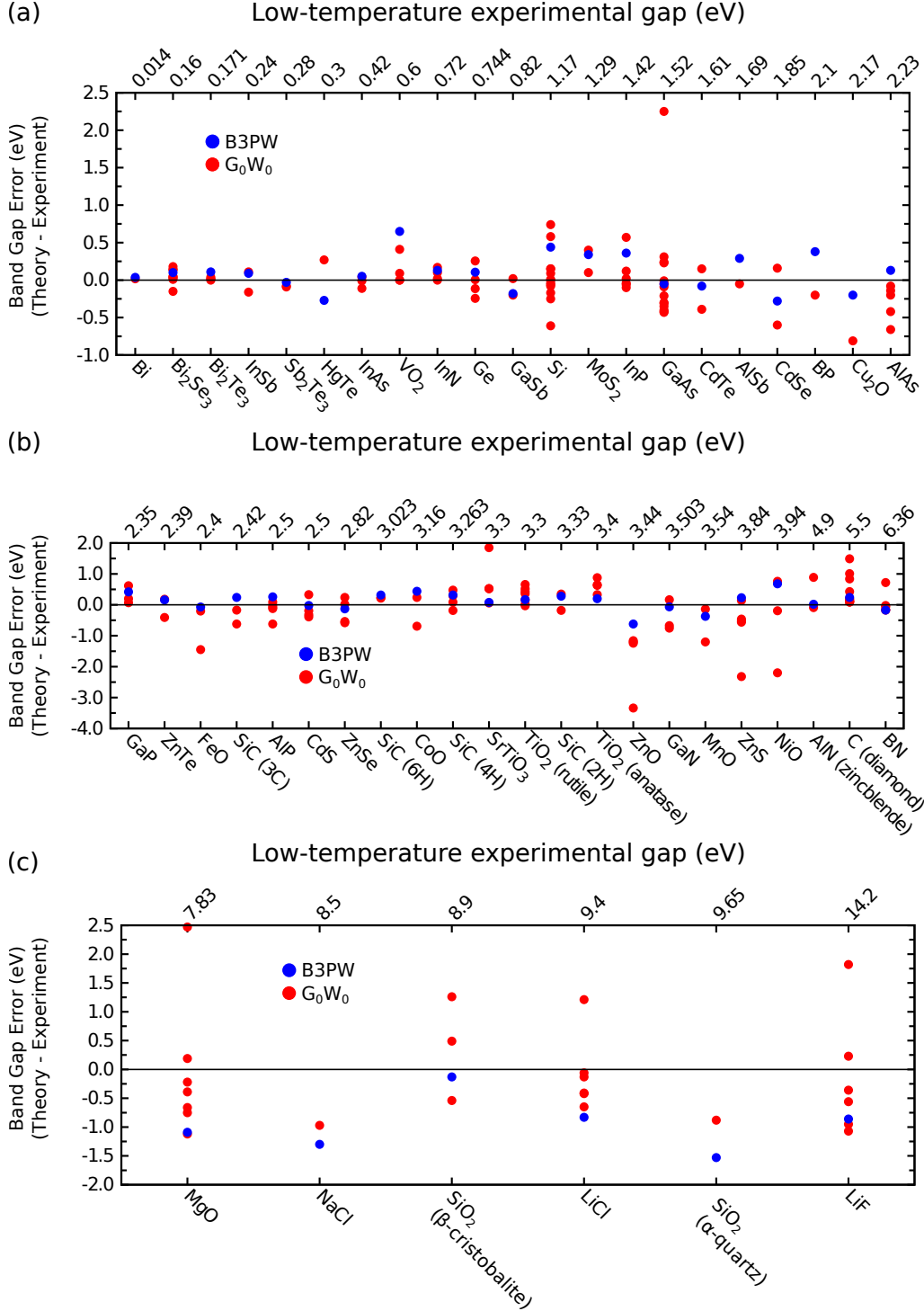
Figure A.2: Difference between computed and low-temperature experimental band gaps for the 49 compounds for which our literature search found G_0W_0 results. Our B3PW results are shown by blue circles, and red circles represent literature G_0W_0 results. Compound names are listed on the bottom x-axes; experimental band gaps (low temperature or $T \approx 0K$ when available; see Table A.8 for a discussion of experimental gaps) are listed on the top x-axes. (a) Results for Bi to AlAs (experimental band gaps 0.014 – 2.23 eV). (b) Results for GaP to BN (experimental gaps 2.35 – 6.36 eV). (c) Results for MgO to LiF (experimental gaps 7.83 – 14.2 eV)

A.3 Comparison of B3PW and G_0W_0 @ LDA
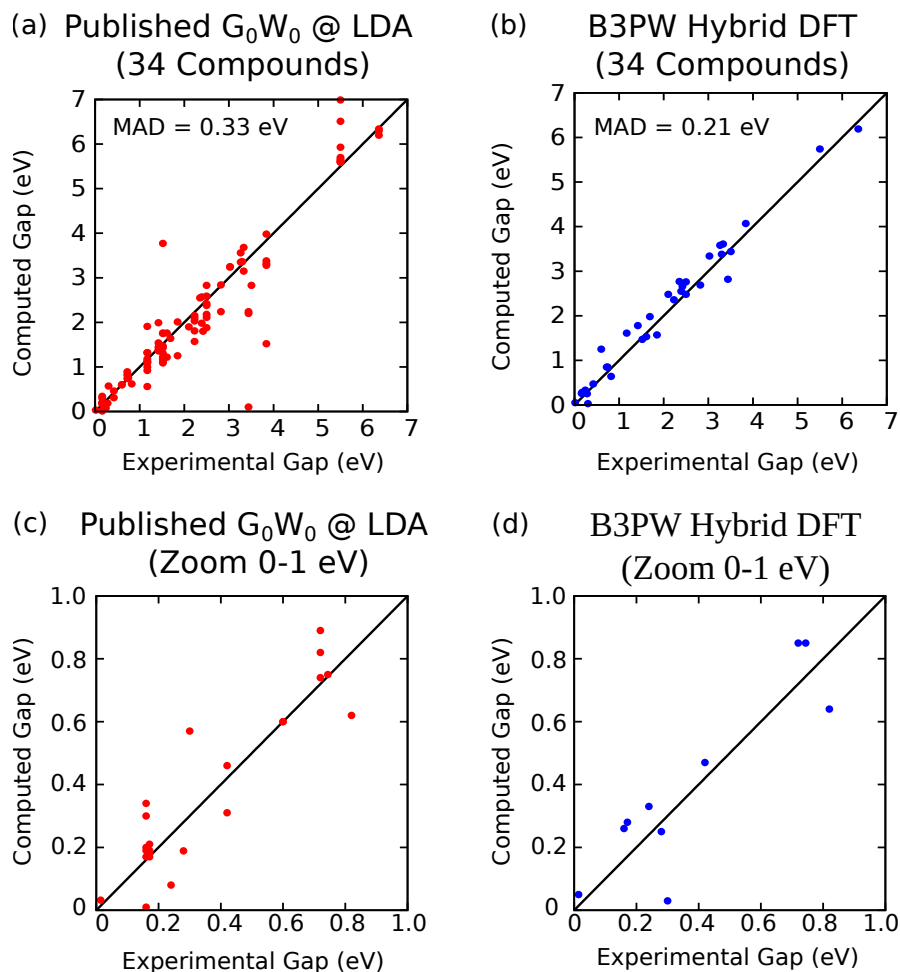
Figure A.3: G_0W_0 @ LDA (G_0W_0 using LDA DFT as the starting point) and B3PW calculated band gaps versus experiment (low-temperature) for 34 compounds with band gaps below 7 eV. These 34 compounds are those for which our literature search found G_0W_0 @ LDA results, and are a subset of the 70 compounds listed in Figures 3.2 and 3.3. These 34 compounds are listed in Figure A.4 (a) Published G_0W_0 @ LDA results (100 data points, 34 compounds, 22 publications). The mean absolute deviation (MAD) is 0.33 eV. (b) B3PW hybrid DFT calculations on the same 34 compounds. The MAD is 0.21 eV. (c) Zoom of (a) from 0–1 eV. (d) Zoom of (b) from 0–1 eV.
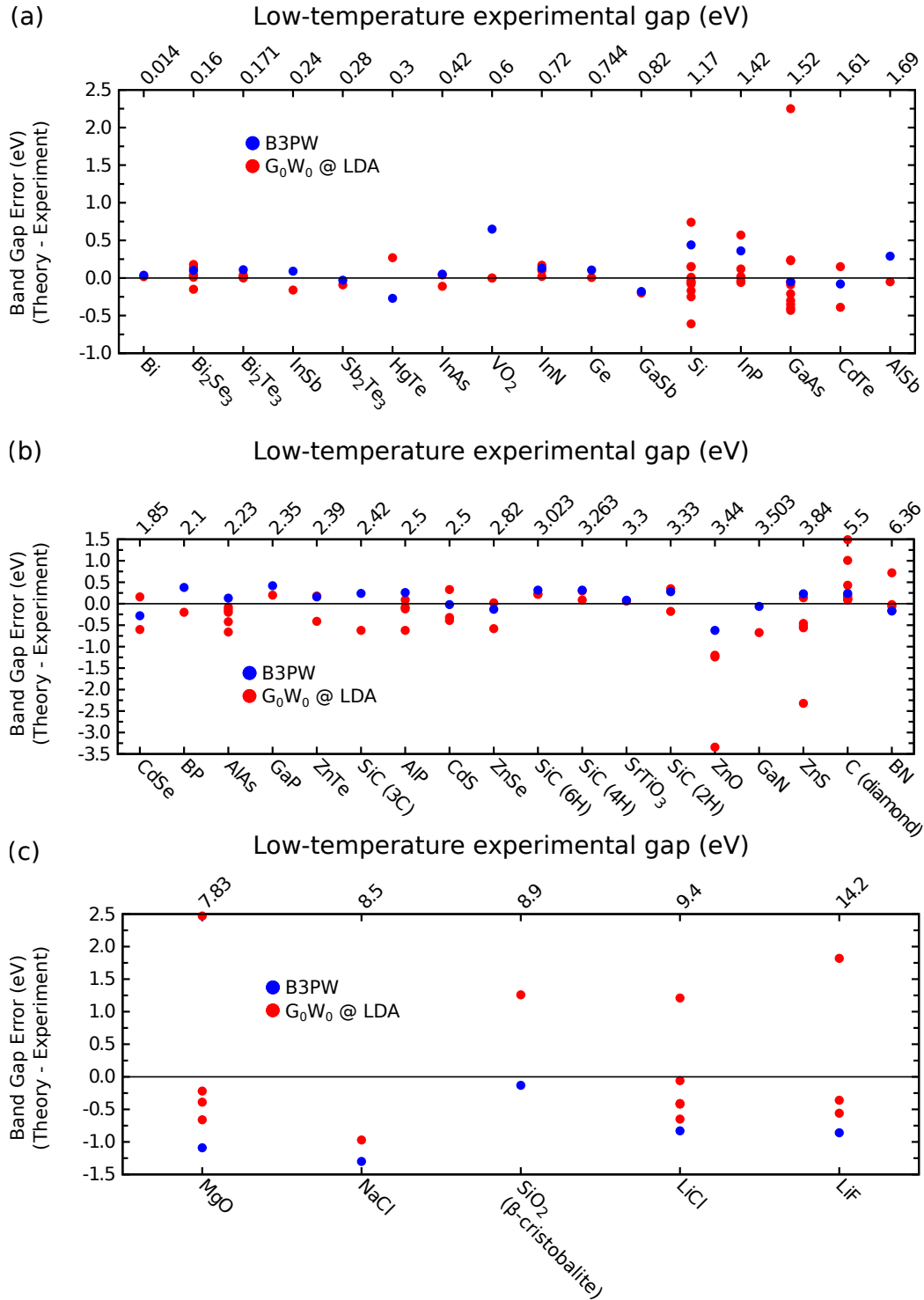
Figure A.4: Difference between computed and low-temperature experimental band gaps for the 39 compounds for which our literature search found G_0W_0 @ LDA results. Our B3PW results are shown by blue circles, and red circles represent literature G_0W_0 @ LDA results. Compound names are listed on the bottom x-axes; experimental band gaps (low temperature or $T \approx 0K$ when available; see Table A.8 for a discussion of experimental gaps) are listed on the top x-axes. (a) Results for Bi to AlSb (experimental band gaps $0.014 - 1.69$ eV). (b) Results for CdSe to BN (experimental gaps $1.86 - 6.36$ eV). (c) Results for MgO to LiF (experimental gaps $7.83 - 14.2$ eV)

A.4 Comparison of B3PW and G_0W_0 @ PBE
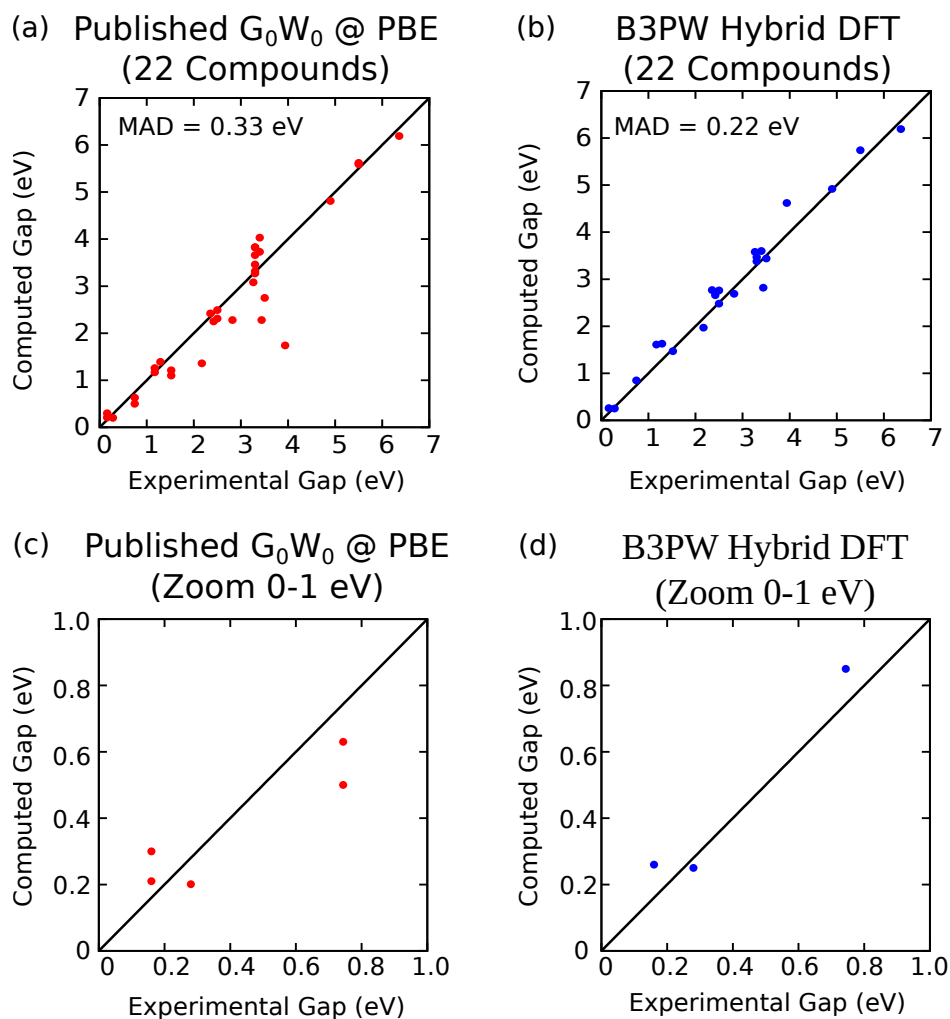
Figure A.5: G_0W_0 @ PBE (G_0W_0 using PBE DFT as the starting point) and B3PW calculated band gaps versus experiment (low-temperature) for 22 compounds with gaps below 7 eV. These 22 compounds are those for which our literature search found G_0W_0 @ PBE results, and are a subset of the 70 compounds listed in Figures 3.2 and 3.3. These 22 compounds are listed in Figure A.6. (a) Published G_0W_0 @ PBE results (33 data points, 22 compounds, 8 publications). The mean absolute deviation (MAD) is 0.33 eV. (b) B3PW hybrid DFT calculations on the same 22 compounds. The MAD is 0.22 eV. (c) Zoom of (a) from 0–1 eV. (d) Zoom of (b) from 0–1 eV.
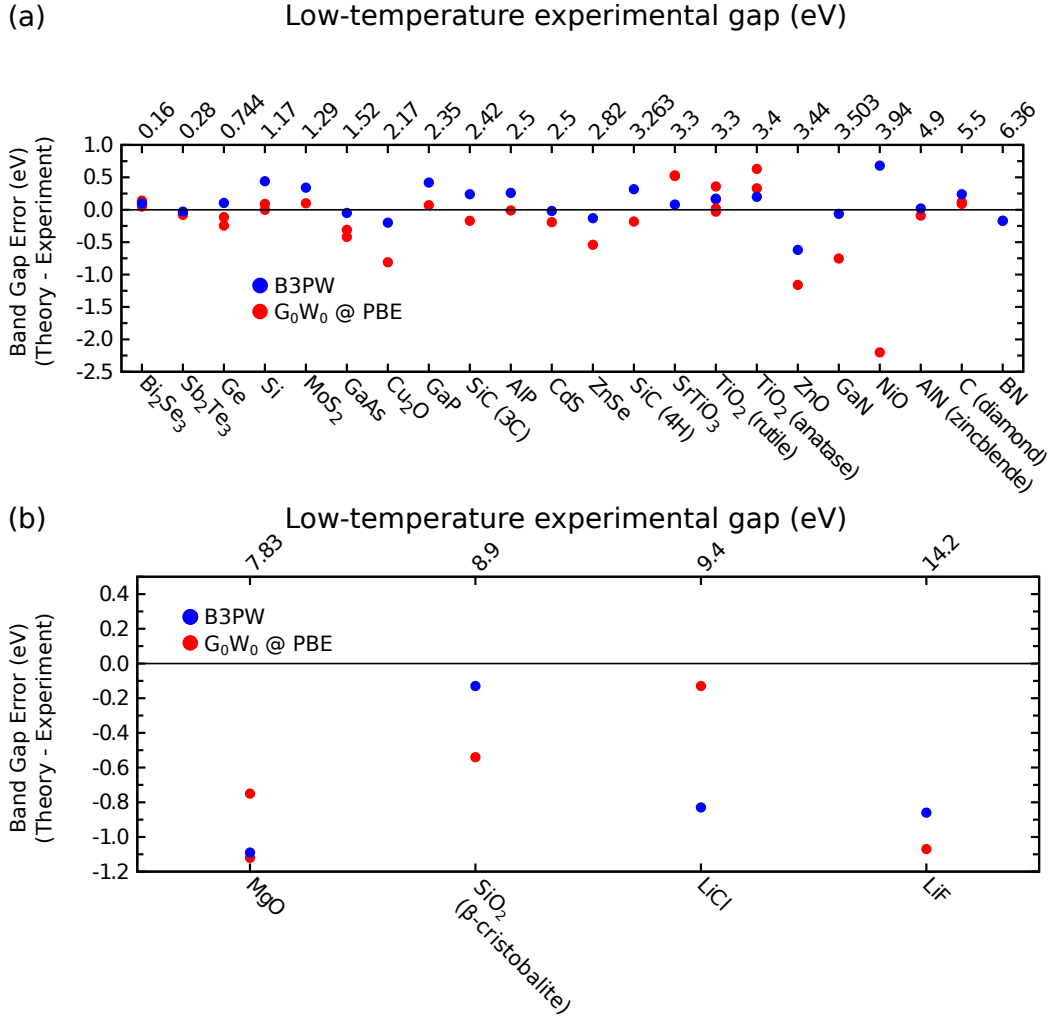
Figure A.6: Difference between computed and low-temperature experimental band gaps for the 26 compounds for which our literature search found G_0W_0 @ PBE results. Our B3PW results are shown by blue circles, and red circles represent literature G_0W_0 @ PBE results. Compound names are listed on the bottom x-axes; experimental band gaps (low temperature or $T \approx 0K$ when available; see Table A.8 for a discussion of experimental gaps) are listed on the top x-axes. (a) Results for Bi_2Se_3 to BN (experimental band gaps $0.16 - 6.36$ eV). (b) Results for MgO to LiF (experimental gaps $7.83 - 14.2$ eV)

A.5 Comparison of B3PW and post-G_0W_0
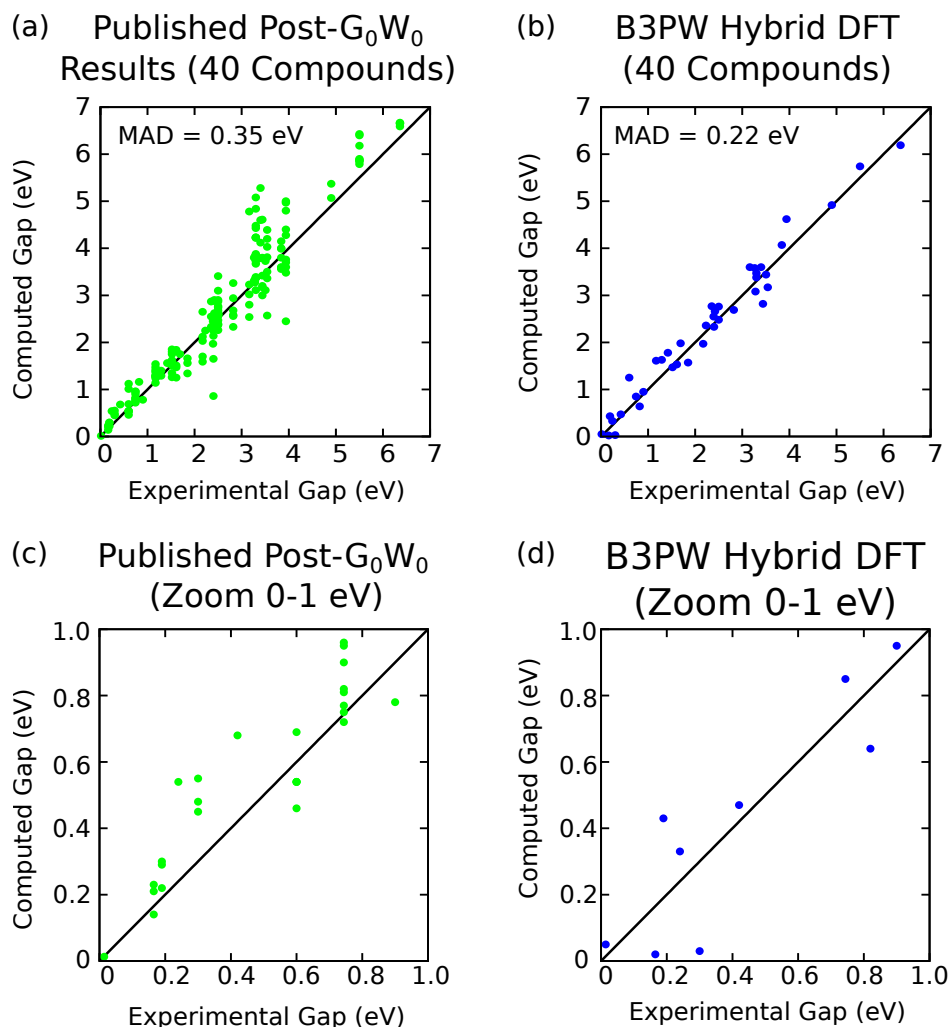
Figure A.7: Post-G_0W_0 and B3PW calculated band gaps versus experiment (low-temperature) for 40 compounds with band gaps below 7 eV. These 40 compounds are those for which our literature search found post-G_0W_0 results, and are a subset of the 70 compounds listed in Figures 3.2 and 3.3. These 40 compounds are listed in Figure A.8. (a) Published post-G_0W_0 results (180 data points, 40 compounds, 17 publications). The mean absolute deviation (MAD) is 0.35 eV. (b) B3PW hybrid DFT calculations on the same 40 compounds. The MAD is 0.22 eV. (c) Zoom of (a) from 0–1 eV. (d) Zoom of (b) from 0–1 eV.
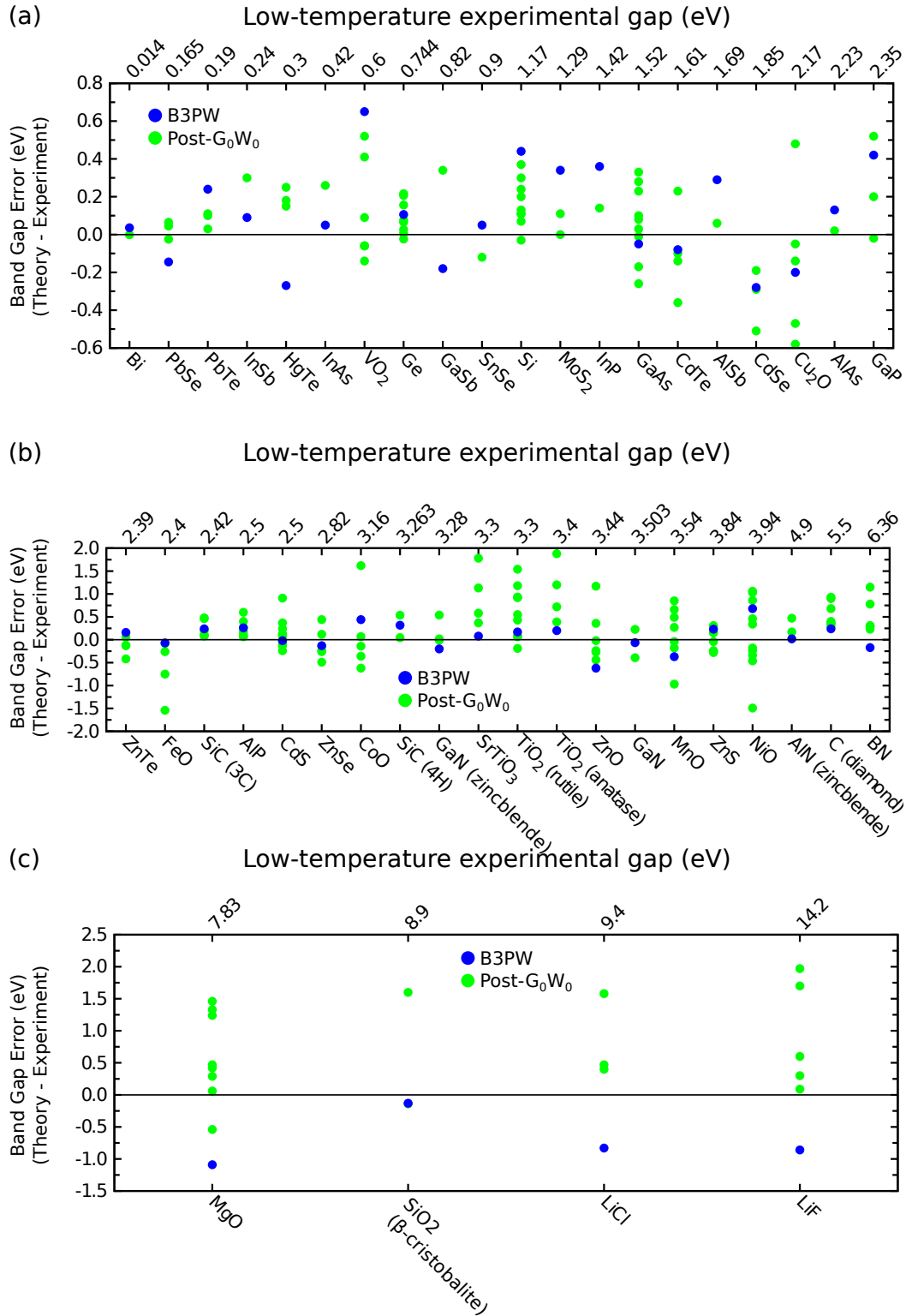
Figure A.8: Difference between computed and low-temperature experimental band gaps for the 46 compounds for which our literature search found post- G_0W_0. Our B3PW results are shown by blue circles, and red circles represent literature post-G_0W_0 results. Compound names are listed on the bottom x-axes; experimental band gaps (low temperature or $T \approx 0K$ when available; see Table A.8 for a discussion of experimental gaps) are listed on the top x-axes. (a) Results for Bi to GaP (experimental band gaps 0.014 – 2.35 eV). (b) Results for ZnTe to BN (experimental band gaps 2.39 – 6.36 eV). (c) Results for MgO to Ne (experimental gaps 7.83 – 21.7 eV)

A.6 Calculated Band Gaps Versus Low-Temperature Experiments for all Compounds
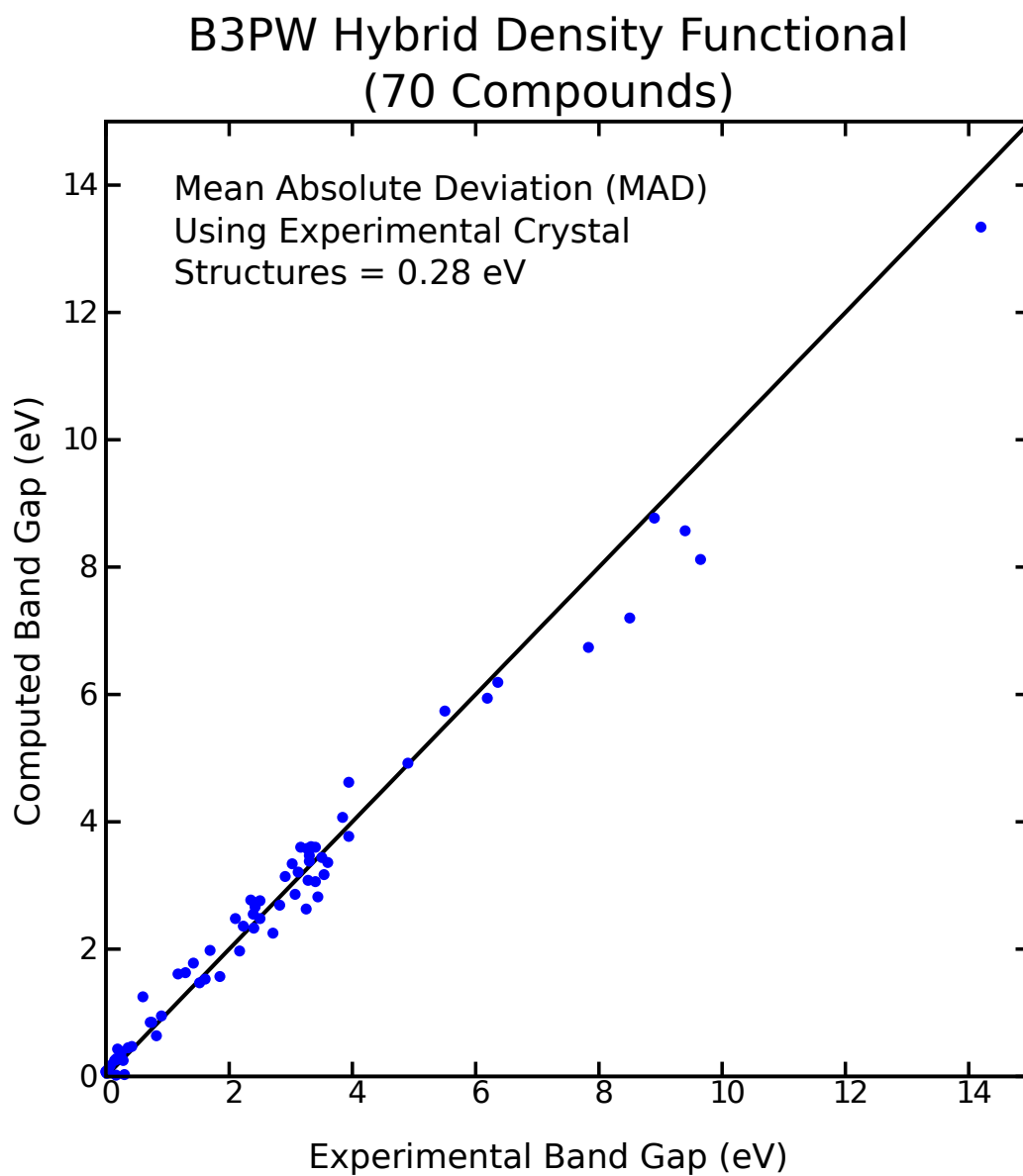
Figure A.9: B3PW calculated band gaps versus experiment (low-temperature) for all 70 compounds listed in Figures 3.2 and 3.3. The mean absolute deviation is 0.28 eV.
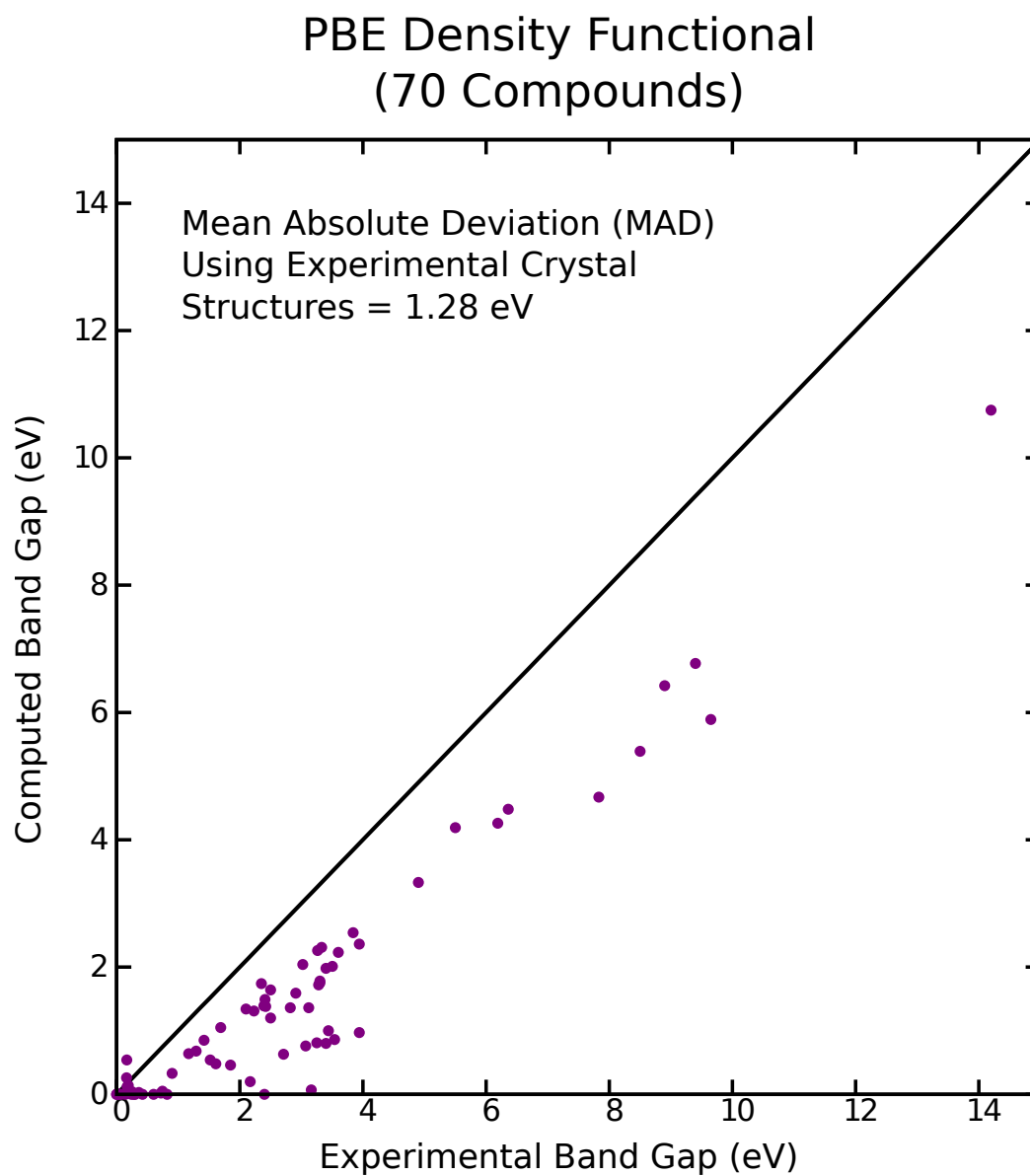
Figure A.10: PBE calculated band gaps versus experiment (low-temperature) for PBE for all 70 compounds listed in Figures 3.2 and 3.3. The mean absolute deviation is 1.28 eV.
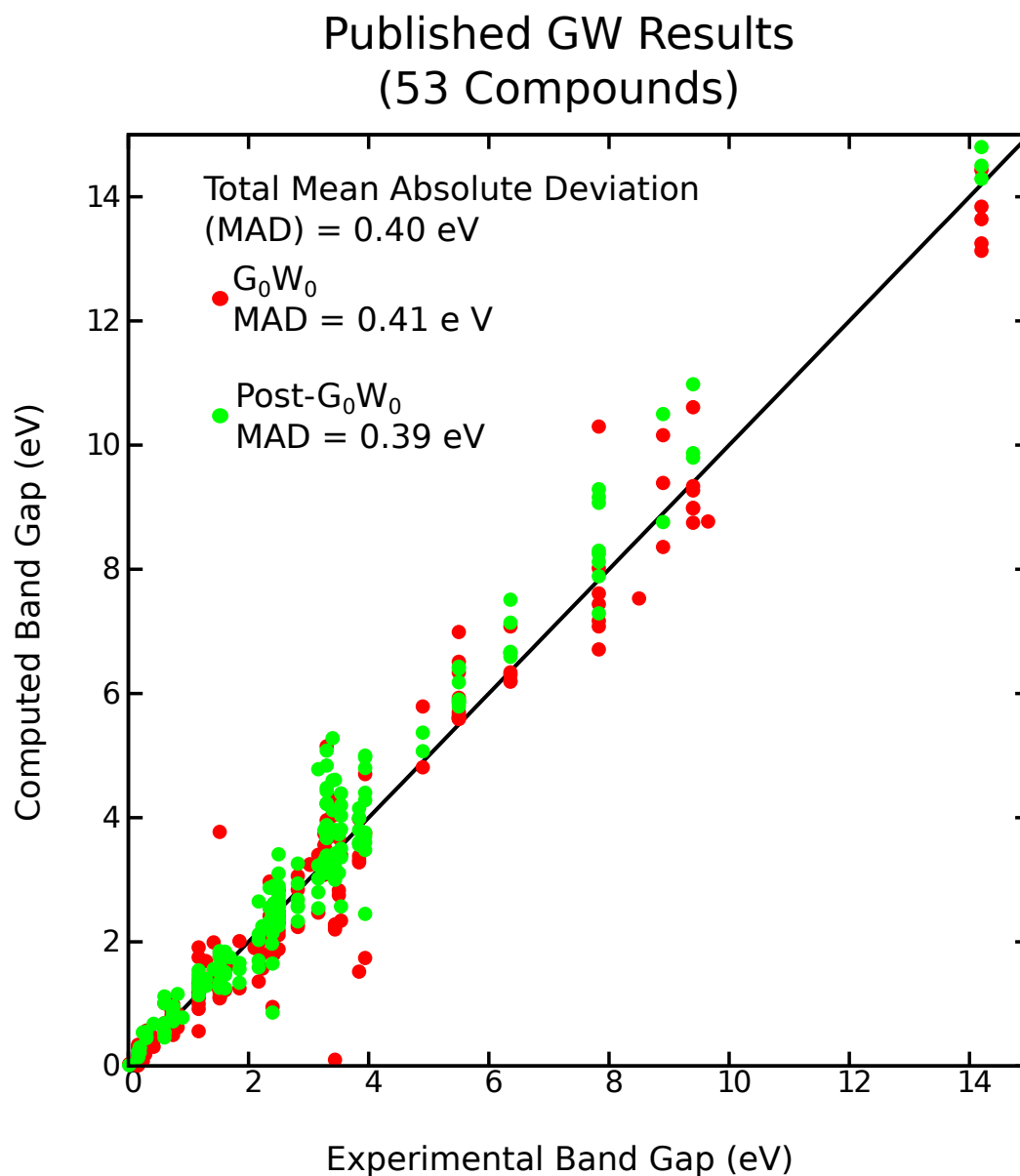
Figure A.11: G_0W_0 (red circles) and post-G_0W_0 (green circles) calculated band gaps versus experiment (low-temperature) for the 53 compounds where our literature search found G_0W_0 or post-G_0W_0 results. These 53 compounds form a subset of the 70 compounds listed in Figures 3.2 and 3.3. The mean absolute deviation is 0.40 eV overall, 0.41 eV for G_0W_0, and 0.39 eV for post-G_0W_0.

A.7 Tables of B3PW, PBE, GW, and Experimental Band Gaps

Table A.1: Band Gaps (eV) from B3PW, GW, PBE and Experiment. Table A.8 lists the experimental band gap from every reference.

System	B3PW	Exp	Refs	PBE	GW	Refs
Bi_2Se_3 (6 QL)	0.08	0.0	[105]	0.00	N/A	N/A
Bi_2Se_3 (7 QL)	0.07	0.0	[105]	0.00	N/A	N/A
Bi	0.00	0.011 – 0.015	[49, 115–119]	0.03	0.013 – 0.032	[63]
Bi_2Se_3 (5 QL)	0.12	0.04	[105]	0.00	N/A	N/A
Bi_2Se_3 (4 QL)	0.17	0.07	[105]	0.00	N/A	N/A
Bi_2Se_3 (3 QL)	0.25	0.14	[105]	0.00	N/A	N/A
Bi_2Se_3	0.18	0.16-0.35	[49, 120]	0.24	0.01–0.34	[94, 121, 122]
PbSe	0.25	0.145–0.165	[49, 123]	0.43	0.14–0.23	[124, 125]
Bi_2Te_3	0.22	0.13–0.171	[49, 126]	0.09	0.17–0.21	[94, 121]
PbTe	0.59	0.190	[123]	0.05	0.22–0.3	[124, 125]
InSb	0.34	0.24	[49]	0.00	0.08–0.54	[56, 60, 127]
Bi_2Se_3 (2 QL)	0.39	0.25	[105]	0.04	N/A	N/A
Sb_2Te_3	0.22	0.28	[49]	0.00	0.189–0.201	[128]
HgTe	0.01	0.3	[49]	0.00	0.45–0.57	[129]
SnTe	0.44	0.36	[49]	0.05	N/A	N/A
InAs	0.47	0.42	[49]	0.00	0.31–0.68	[56, 60, 127, 130]
VO_2	1.25	0.6	[131]	0.00	0.46–1.12	[132, 133]
InN	0.86	0.7 – 1.0	[134–136]	0.01	0.72 – 0.89	[137–139]
Ge	0.86	0.744	[49]	0.07	0.5–1.0	[58, 61, 62, 140]
GaSb	0.66	0.82	[49]	0.00	0.62–1.16	[56, 60, 127]
SnSe	0.95	0.86–0.95	[141–143]	0.33	0.78	[144]
Si	1.61	1.17	[49]	0.64	0.56–1.91	[56, 58, 61, 62, 130] [59, 140, 145–147]
MoS_2	1.63	1.29	[148]	0.68	1.29–1.69	[62]
InP	1.78	1.42	[49]	0.86	1.32–1.99	[56, 60, 127, 146, 147]
GaAs	1.48	1.52	[49]	0.55	1.09–3.77	[56, 58, 60–62, 127, 140] [55, 145–147]
CdTe	1.53	1.61	[149]	0.48	1.22–1.84	[60, 129, 150]
AlSb	1.83	1.69	[49]	0.92	1.64–1.75	[56, 60]
CdSe	1.57	1.85	[151]	0.46	1.25–2.01	[129, 150]
BP	2.48	2.1	[49]	1.34	1.9	[152]
Cu_2O	1.97	2.17	[49, 153]	0.20	1.36–2.65	[133, 140]
AlAs	2.37	2.23	[49]	1.31	1.57–2.25	[56, 130, 146] [55, 57, 60]

Table A.2: Band Gaps (eV) from B3PW, GW, PBE and Experiment. Table A.8 lists the experimental band gap from every reference.

System	B3PW	Exp	Refs	PBE	GW	Refs
GaP	2.71	2.35	[49]	1.69	2.33–2.97	[56, 60, 62]
ZnTe	2.55	2.39	[49]	1.39	1.97–2.67	[60, 129, 150]
FeO	2.33	2.4	[154]	0.00	0.86–2.32	[51, 133, 155, 156]
BiVO$_4$	2.98	2.41	[157]	1.72	N/A	N/A
SiC (3C)	2.66	2.42	[49]	1.38	1.8 – 2.88	[61, 130, 140]
AlP	2.59	2.5	[49]	1.51	1.88–3.1	[56, 61, 140] [60, 147]
CdS	2.49	2.5	[151]	1.20	2.11–3.41	[58, 61, 129] [140, 150, 158]
AgBr	2.24	2.71	[49]	0.62	N/A	N/A
ZnSe	2.69	2.82	[49]	1.37	2.24–3.26	[129, 150] [60, 62]
AgI	3.13	2.91	[49]	1.57	N/A	N/A
SiC (6H)	3.34	3.023	[159]	2.04	3.24–3.25	[160, 161]
CuBr	2.87	3.07	[49]	0.76	N/A	N/A
CuI	3.21	3.12	[49]	1.36	N/A	N/A
CoO	3.54	2.1–5.43	[162–166]	0.13	2.4–4.78	[51, 133, 155] [156, 167, 168]
AgCl	2.63	3.25	[49]	0.82	N/A	N/A
SiC (4H)	3.58	3.263	[159]	2.26	3.08–3.8	[62, 160, 161]
GaN (zincblende)	3.08	3.28	[169]	1.73	3.27 – 3.82	[61]
SrTiO$_3$	3.38	3.25–3.3	[170–172]	1.75	3.36–5.15	[58, 173, 174]
TiO$_2$ (Rutile)	3.47	3.3	[172]	1.78	3.11–4.84	[133, 174, 175] [62, 140]
SiC (2H)	3.60	3.33	[176]	2.30	3.15–3.68	[160, 161]
CuCl	3.06	3.4	[49]	0.81	N/A	N/A
TiO$_2$ (Anatase)	3.60	3.4	[177]	1.98	3.73–5.28	[174, 175]
ZnO	2.82	3.44	[49]	1.00	0.1–4.61	[61, 140, 147]
GaN	3.43	3.503	[49]	2.01	2.75–3.82	[58, 61, 62]
MnO	3.16	2.0–4.2	[163, 178–181]	0.86	2.34–4.39	[51, 133, 155] [158, 167] [59, 156]
MgTe	3.31	3.6	[182]	2.18	N/A	N/A
ZnS	4.07	3.84	[151]	2.55	1.52–4.15	[61, 129, 150] [60, 147]
CuSCN	3.97	3.94	[183]	2.36	N/A	N/A
NiO	4.57	3.7–4.3	[163, 165, 184] [166, 185]	0.97	1.74–5.0	[51, 133, 155] [140, 158, 167] [59, 156]

Table A.3: Band Gaps (eV) from B3PW, GW, PBE and Experiment. Table A.8 lists the experimental band gap from every reference.

System	B3PW	Exp	Refs	PBE	GW	Refs
AlN (zincblende)	4.94	4.9	[186]	3.33	N/A	N/A
C (diamond)	5.74	5.5	[49]	4.19	5.59–6.99	[58, 145, 146] [62, 140] [61, 147]
AlN (wurtzite)	6.05	6.19	[49]	4.31	4.81–5.79	[62]
BN	6.19	6.36	[187]	4.48	6.19–7.51	[58, 61, 140, 147]
MgO	6.74	7.83	[188]	4.67	6.71–10.3	[58, 61, 62, 140, 147]
NaCl	9.08	8.5–8.69	[189, 190]	7.27	7.53	[58]
SiO_2 (β-cristobalite)	8.90	8.9	[191]	6.52	8.36–10.5	[62, 192]
LiCl	8.76	9.4	[193]	7.00	8.75–10.98	[56, 140, 145, 146]
SiO_2 (α-quartz)	10.18	9.65	[194]	7.79	8.77	[53]
LiF	13.33	14.2	[195]	10.75	13.13–16.17	[53, 61, 62, 147]

A.8 Table of Published Experimental Band Gaps

Table A.4: Experimental Band Gaps (caption below). When there is more than one experimental result, our choice is in bold.

System	Band Gap (eV)	Temperature (K)	Reference
Bi_2Se_3 (6QL)	0.00	?	[105]
Bi_2Se_3 (7QL)	0.00	?	[105]
Bi	0.015	1.2	[118]
Bi	0.015	4	[115]
Bi	0.011	4	[116]
Bi	0.0136	0 (extrapolated)	[117]
Bi	0.0153	1.4	[119]
Bi	**0.0136**	?	**[49]**
Bi_2Se_3 (5QL)	0.04	?	[105]
Bi_2Se_3 (4QL)	0.07	?	[105]
Bi_2Se_3 (3QL)	0.14	?	[105]
Bi_2Se_3	0.35	RT	[120]
Bi_2Se_3	**0.16**	**77**	**[49]**
PbSe	0.145	4	[49]
PbSe	**0.165**	**4.2**	**[123]**
Bi_2Te_3	0.13	293	[49]
Bi_2Te_3	**0.171**	**0 (extrapolated)**	**[126]**
PbTe	0.190	4.2	[123]
InSb	0.24	1.7	[49]
Bi_2Se_3 (2QL)	0.25	?	[105]
Sb_2Te_3	0.28	299	[49]
HgTe	0.304	0 (extrapolated)	[49]
SnTe	0.36	12	[49]
InAs	0.42	4.2	[49]
VO_2	0.6	300	[131]
InN	0.7 – 0.8	12 – 300	[134]
InN	0.7 – 1.0	RT	[135]
InN	**0.72**	**4–6**	**[136]**
Ge	0.744	1.5	[49]
GaSb	0.82	0 (extrapolated)	[49]
SnSe	0.86	RT	[143]
SnSe	0.898	RT	[141]
SnSe	0.95	RT	[142]
SnSe (average)	**0.90**		
Si	1.17	0 (extrapolated)	[49]
MoS_2	1.29	?	[148]
InP	1.42	1.6	[49]

System	Band Gap (eV)	Temperature (K)	Reference
GaAs	1.52	0 (extrapolated)	[49]
CdTe	1.48	300	[49]
CdTe	**1.61**	**4.2**	**[149]**
AlSb	1.69	27	[49]
CdSe	1.73	300	[49]
CdSe	**1.85**	**0**	**[151]**
BP	2.1	RT	[49]
Cu_2O	2.17	4.2	[153]
Cu_2O	2.17	4.2	[49]
AlAs	2.23	4	[49]
GaP	2.35	0 (extrapolated)	[49]
ZnTe	2.39	< 2	[49]
FeO	2.4	77	[154]
$BiVO_4$	2.41	?	[157]
SiC (3C)	2.42	2K	[49]
AlP	2.5	2	[49]
CdS	2.5	300	[151]
AgBr	2.71	1.8	[49]
ZnSe	2.82	6	[49]
AgI	2.91	4	[49]
SiC (6H)	3.023	4.2	[159]
CuBr	3.07	1.6	[49]
CuI	3.12	80	[49]
CoO	3.6	?	[162]
CoO	2.6	?	[163]
CoO	2.5	?	[164]
CoO	2.1	78	[165]
CoO	2.7	?	[166]
CoO	5.43	?	[168]
CoO (average)	**3.16**		
AgCl	3.25	1.8	[49]
SiC (4H)	3.263	4.2	[159]
GaN (zincblende)	3.28	0	[169]
$SrTiO_3$	3.3	?	[172]
$SrTiO_3$	3.25	?	[171]
$SrTiO_3$	**3.3**	**20**	**[170]**
TiO_2 (rutile)	3.3	?	[172]
SiC (2H)	3.33	4.2	[176]
CuCl	3.4	2	[49]
TiO_2 (anatase)	3.4	0 (extrapolated)	[177]
ZnO	3.44	6	[49]
GaN	3.503	1.6	[49]

System	Band Gap (eV)	Temperature (K)	Reference
MnO	3.6–3.8	?	[178]
MnO	3.8–4.2	?	[179]
MnO	4.1	?	[163]
MnO	3.9	?	[180]
MnO	2.0	77	[181]
MnO (average)	**3.54**		
MgTe	3.49	RT	[49]
MgTe	**3.6**	**0 (extrapolated)**	**[182]**
ZnS	3.72	300	[49]
ZnS	**3.84**	**0**	**[151]**
CuSCN	3.94	RT	[183]
NiO	4.3	?	[184]
NiO	4.0	?	[163]
NiO	4.0	78	[165]
NiO	3.7	?	[185]
NiO	3.7	?	[166]
NiO (average)	**3.94**		
AlN (zincblende)	4.9	0	[186]
C (diamond)	5.5	RT	[49]
AlN (wurtzite)	6.19	7	[49]
BN	6.2	?	[49]
BN	**6.36**	**8**	**[187]**
MgO	7.22	?	[196]
MgO	7.9	0 (extrapolated)	[49]
MgO	**7.83**	**83**	**[188]**
NaCl	8.69	77	[190]
NaCl	8.75	10	[197]
NaCl	**8.5**	**?**	**[189]**
SiO_2 (β-cristobalite)	8.9	?	[191]
LiCl	9.4	55	[193]
SiO_2 (α-quartz)	9.65	?	[194]
LiF	14.2	77	[195]

Table A.8 lists the experimental band gap in every reference we found. We also list the temperature for each result or a question mark when we were unsure. In many cases there are several experimental results to choose from. When this is the case, our choice is shown in bold in Table A.8. In general, when a low-temperature result from reference [49] (a standard and well-known reference work) was available, we used it. We followed this approach in all cases except the following.

PbSe. We chose reference [123], as this result is more commonly used in the literature.

Bi$_2$Te$_3$. We chose reference [126] because it is a 0 K result, whereas the band gap in [49] is at room temperature.

SnSe. We found three recent room-temperature measurements. In order to avoid biasing our comparison, we chose to average these results.

CoO, MnO, NiO. There appears to be no clear consensus as to the band gaps of these materials. We gathered the results most commonly cited in the *GW* literature and averaged. As a result, the MADs for *GW* decreased by $\sim$ 0.01 eV and the MADs for B3PW increased by $\sim$ 0.01 eV.

MgO and NaCl. We chose the references commonly used in the *GW* literature. This choice biases the comparison in favor of *GW*.

BP. Lucero et al. [48] quote a low-temperature experimental band gap of 2.4 eV. This number seems plausible given the room temperature value of 2.1 eV. However, we were unable to locate any experimental paper containing this value. Thus, we used the room temperature band gap from reference [49].

A p p e n d i x   B

B3PW BAND STRUCTURES OF SLABS

Figures B.1 and B.2 show the 1-7QL band structures for Bi_2Te_3 and Bi_2Se_3, respectively.
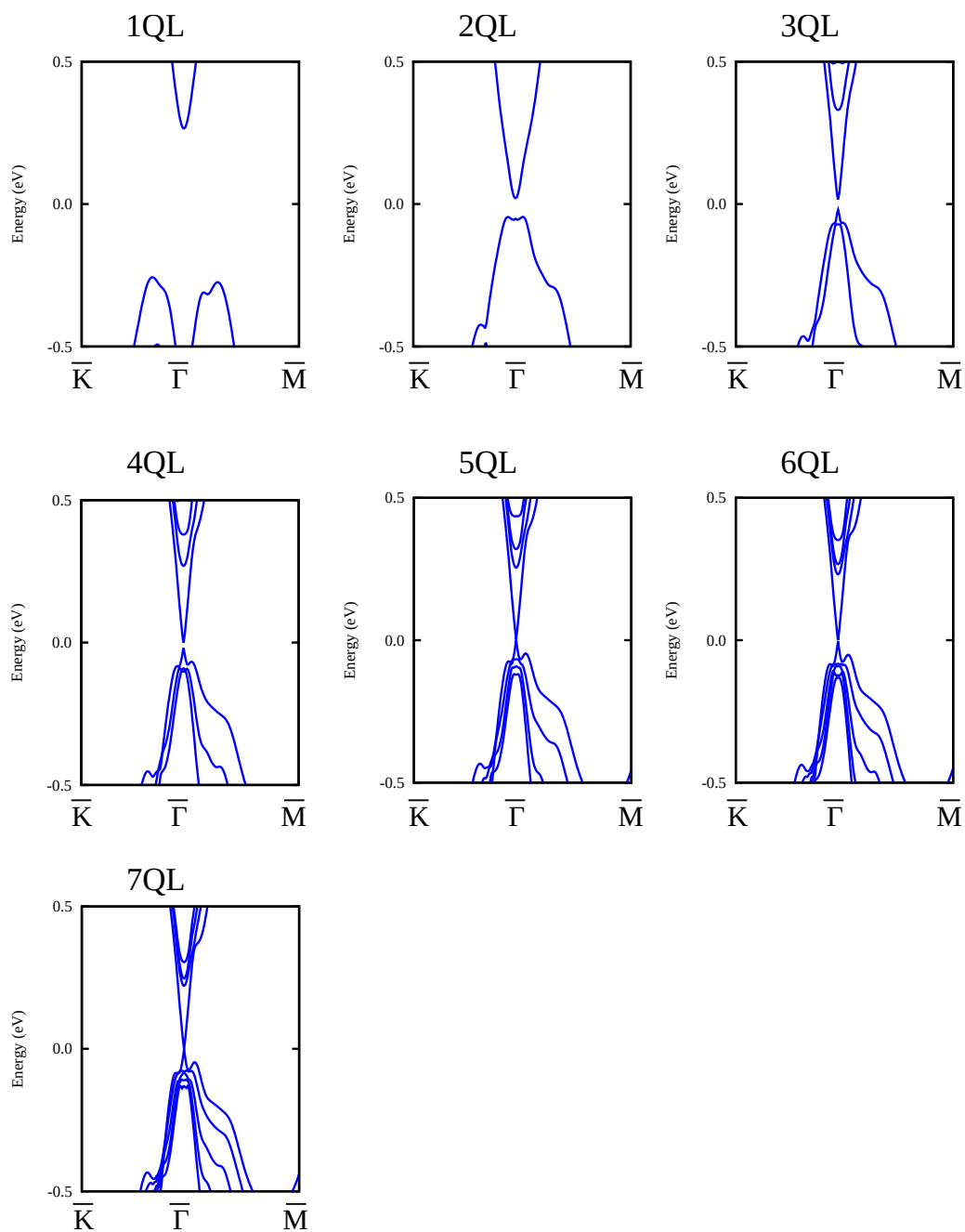
Figure B.1: B3PW band structures of Bi$_2$Te$_3$ 1-7QL slabs
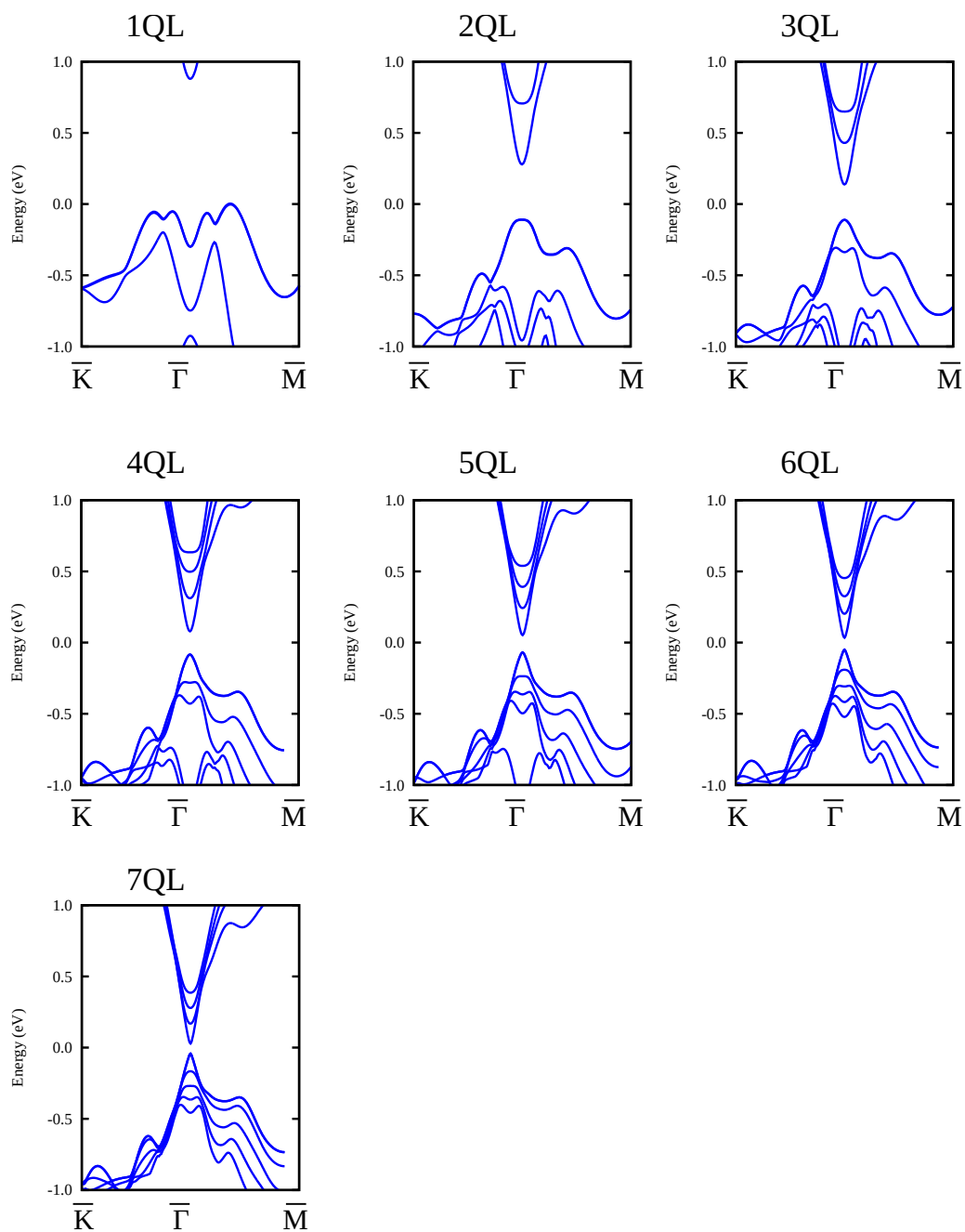
Figure B.2: B3PW band structures of Bi_2Se_3 1-7QL slabs

BIBLIOGRAPHY

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

[2] J. K. Perry, J. Tahir-Kheli, and W. A. Goddard, "Antiferromagnetic band structure of La_2CuO_4 : Becke-3-Lee-Yang-Parr calculations", Phys. Rev. B **63**, 144510 (2001).

[3] R. Dovesi, V. Saunders, C. Roetti, R. Orlando, C. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. Harrison, I. Bush, P. D'Arco, M. Llunell, M. Causà, and Y. Noël, *Crystal14 user's manual* (University of Torino: Torino, 2014).

[4] V. Saunders, R. Dovesi, C. Roetti, M. Causà, N. Harrison, R. Orlando, and C. Zicovich-Wilson, *Crystal98 user's manual* (University of Torino: Torino, 1998).

[5] H. Hellmann, "A new approximation method in the problem of many electrons", J. Chem. Phys. **3**, 61–61 (1935).

[6] P. Gombás, "Über die metallische bindung", Zeitschrift für Physik **94**, 473–488 (1935).

[7] V. Fock, M. Veselov, and M. Petrashen, Zh. Eksp. Theor. Fiz. **10**, 723 (1940).

[8] J. C. Phillips and L. Kleinman, "New method for calculating wave functions in crystals and molecules", Phys. Rev. **116**, 287–294 (1959).

[9] J. D. Weeks and S. A. Rice, "Use of pseudopotentials in atomic-structure calculations", J. Chem. Phys. **49**, 2741–2755 (1968).

[10] I. V. Abarenkov and V. Heine, "The model potential for positive ions", Philos. Mag. **12**, 529–537 (1965).

[11] A. Dalgarno, C. Bottcher, and G. Victor, "Pseudo-potential calculation of atomic interactions", Chem. Phys. Lett. **7**, 265–267 (1970).

[12] G. Simons, "New model potential for pseudopotential calculations", J. Chem. Phys. **55**, 756–761 (1971).

[13] W. L. McMillan, "Theory of one-electron molecules. i. Li_2^+", Phys. Rev. A **4**, 69–81 (1971).

[14] W. A. Goddard, "New foundation for the use of pseudopotentials in metals", Phys. Rev. **174**, 659–662 (1968).

[15] L. Kahn and W. Goddard, "A direct test of the validity of the use of pseudopotentials in molecules", Chem. Phys. Lett. **2**, 667–670 (1968).

[16] L. R. Kahn and W. A. Goddard, "Ab initio effective potentials for use in molecular calculations", J. Chem. Phys. **56**, 2685–2701 (1972).

[17]C. F. Melius, W. A. Goddard, and L. R. Kahn, "Use of ab initio g1 effective potentials for calculations of molecular excited states", J. Chem. Phys. **56**, 3342–3348 (1972).

[18]C. F. Melius and W. A. Goddard, "*Ab initio* effective potentials for use in molecular quantum mechanics", Phys. Rev. A **10**, 1528–1540 (1974).

[19]L. R. Kahn, P. Baybutt, and D. G. Truhlar, "Ab initio effective core potentials: reduction of all-electron molecular structure calculations to calculations involving only valence electrons", J. Chem. Phys. **65**, 3826–3853 (1976).

[20]Y. S. Lee, W. C. Ermler, and K. S. Pitzer, "Abinitio effective core potentials including relativistic effects. I. formalism and applications to the Xe and Au atoms", J. Chem. Phys. **67**, 5861–5876 (1977).

[21]W. C. Ermler, Y. S. Lee, K. S. Pitzer, and N. W. Winter, "Abinitio effective core potentials including relativistic effects. II. potential energy curves for Xe_2, Xe_2^+, and Xe_2^*", J. Chem. Phys. **69**, 976–983 (1978).

[22]Y. S. Lee, W. C. Ermler, K. S. Pitzer, and A. D. McLean, "Abinitio effective core potentials including relativistic effects. III. Ground state Au_2 calculations", J. Chem. Phys. **70**, 288–292 (1979).

[23]W. C. Ermler, Y. S. Lee, P. A. Christiansen, and K. S. Pitzer, "Ab initio effective core potentials including relativistic effects. A procedure for the inclusion of spin-orbit coupling in molecular wavefunctions", Chemical Physics Letters **81**, 70–74 (1981).

[24]R. M. Pitzer and N. W. Winter, "Electronic-structure methods for heavy-atom molecules", J. Phys. Chem. **92**, 3061–3063 (1988).

[25]W. R. Wadt and P. J. Hay, "Ab initio effective core potentials for molecular calculations. potentials for main group elements na to bi", J. Chem. Phys. **82**, 284–298 (1985).

[26]W. J. Stevens, M. Krauss, H. Basch, and P. G. Jasien, "Relativistic compact effective potentials and efficient, shared-exponent basis sets for the third-, fourth-, and fifth-row atoms", Can. J. Chem. **70**, 612–630 (1992).

[27]M. Dolg, U. Wedig, H. Stoll, and H. Preuss, "Energy-adjusted abinitio pseudopotentials for the first row transition elements", The Journal of Chemical Physics **86**, 866–872 (1987).

[28]W. Küchle, M. Dolg, H. Stoll, and H. Preuss, "Ab initio pseudopotentials for hg through rn", Molecular Physics **74**, 1245–1263 (1991).

[29]M. Dolg, H. Stoll, H. Preuss, and R. M. Pitzer, "Relativistic and correlation effects for element 105 (hahnium, ha): a comparative study of m and mo (m = nb, ta, ha) using energy-adjusted ab initio pseudopotentials", J. Phys. Chem. **97**, 5852–5859 (1993).

[30] P. Schwerdtfeger, T. Fischer, M. Dolg, G. Igel-Mann, A. Nicklass, H. Stoll, and A. Haaland, "The accuracy of the pseudopotential approximation. i. an analysis of the spectroscopic constants for the electronic ground states of incl and incl3 using various three valence electron pseudopotentials for indium", J. Chem. Phys. **102**, 2050–2062 (1995).

[31] T. Leininger, A. Nicklass, H. Stoll, M. Dolg, and P. Schwerdtfeger, "The accuracy of the pseudopotential approximation. ii. a comparison of various core sizes for indium pseudopotentials in calculations for spectroscopic constants of inh, inf, and incl", J. Chem. Phys. **105**, 1052–1059 (1996).

[32] B. Metz, M. Schweizer, H. Stoll, M. Dolg, and W. Liu, "A small-core multiconfiguration dirac–hartree–fock-adjusted pseudopotential for tl – application to tlx (x = f, cl, br, i)", Theor. Chem. Acc. **104**, 22–28.

[33] H. Stoll, B. Metz, and M. Dolg, "Relativistic energy-consistent pseudopotentials–recent developments", J. Comput. Chem. **23**, 767–778 (2002).

[34] D. Figgen, K. A. Peterson, M. Dolg, and H. Stoll, "Energy-consistent pseudopotentials and correlation consistent basis sets for the 5d elements hf–pt", J. Chem. Phys. **130**, 164108 (2009) `http://dx.doi.org/10.1063/1.3119665`.

[35] M. Dolg and X. Cao, "Relativistic pseudopotentials: their development and scope of applications", Chem. Rev. **112**, 403–480 (2012).

[36] M. Weissbluth, *Atoms and molecules* (Academic Press, New York, 1978).

[37] L. E. McMurchie and E. R. Davidson, "Calculation of integrals over ab initio pseudopotentials", J. Comp. Phys. **44**, 289–301 (1981).

[38] M. Tinkham, *Group theory and quantum mechanics* (Dover, Mineola, New York, 2003).

[39] A. Szabo and N. S. Ostlund, *Modern quantum chemistry* (Dover, Mineola, New York, 1996).

[40] L. E. McMurchie and E. R. Davidson, "One- and two-electron integrals over cartesian gaussian functions", Journal of Computational Physics **26**, 218–231 (1978).

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

[42] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", Phys. Rev. B **23**, 5048–5079 (1981).

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

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

[45] P. Mori-Sánchez, A. J. Cohen, and W. Yang, "Localization and delocalization errors in density functional theory and implications for band-gap prediction", Phys. Rev. Lett. **100**, 146401 (2008).

[46] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the exchange screening parameter on the performance of screened hybrid functionals", J. Chem. Phys. **125**, 224106 (2006) `http://dx.doi.org/10.1063/1.2404663`.

[47] J. E. Moussa, P. A. Schultz, and J. R. Chelikowsky, "Analysis of the heyd-scuseria-ernzerhof density functional parameter space", J. Chem. Phys. **136**, 204117 (2012).

[48] M. J. Lucero, T. M. Henderson, and G. E. Scuseria, "Improved semiconductor lattice parameters and band gaps from a middle-range screened hybrid exchange functional", Journal of Physics: Condensed Matter **24**, 145504 (2012).

[49] O. Madelung, *Semiconductors: data handbook*, 3rd ed. (Springer, New York, 2004).

[50] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, "Energy band gaps and lattice parameters evaluated with the heyd-scuseria-ernzerhof screened hybrid functional", J. Chem. Phys. **123**, 174101 (2005) `http://dx.doi.org/10.1063/1.2085170`.

[51] C. Rödl, F. Fuchs, J. Furthmüller, and F. Bechstedt, "Quasiparticle band structures of the antiferromagnetic transition-metal oxides mno, feo, coo, and nio", Phys. Rev. B **79**, 235114 (2009).

[52] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining, and S. Botti, "Density-based mixing parameter for hybrid functionals", Phys. Rev. B **83**, 035119 (2011).

[53] W. Chen and A. Pasquarello, "Correspondence of defect energy levels in hybrid density functional theory and many-body perturbation theory", Phys. Rev. B **88**, 115104 (2013).

[54] W. Chen and A. Pasquarello, "Band-edge levels in semiconductors and insulators: hybrid density functional theory versus many-body perturbation theory", Phys. Rev. B **86**, 035134 (2012).

[55] R. W. Godby, M. Schlüter, and L. J. Sham, "Quasiparticle energies in gaas and alas", Phys. Rev. B **35**, 4170–4171 (1987).

[56] X. Zhu and S. G. Louie, "Quasiparticle band structure of thirteen semiconductors and insulators", Phys. Rev. B **43**, 14142–14156 (1991).

[57] E. L. Shirley, X. Zhu, and S. G. Louie, "Core polarization in solids: formulation and application to semiconductors", Phys. Rev. B **56**, 6648–6661 (1997).

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

[59] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, "All-electron self-consistent *GW* approximation: application to si, mno, and nio", Phys. Rev. Lett. **93**, 126406 (2004).

[60] A. N. Chantis, M. van Schilfgaarde, and T. Kotani, "*Ab Initio* prediction of conduction band spin splitting in zinc blende semiconductors", Phys. Rev. Lett. **96**, 086405 (2006).

[61] M. Shishkin, M. Marsman, and G. Kresse, "Accurate quasiparticle spectra from self-consistent *GW* calculations with vertex corrections", Phys. Rev. Lett. **99**, 246403 (2007).

[62] W. Chen and A. Pasquarello, "Band-edge positions in *GW*: effects of starting point and self-consistency", Phys. Rev. B **90**, 165133 (2014).

[63] I. Aguilera, C. Friedrich, and S. Blügel, "Electronic phase transitions of bismuth under strain from relativistic self-consistent *GW* calculations", Phys. Rev. B **91**, 125129 (2015).

[64] J. M. Crowley, J. Tahir-Kheli, and W. A. Goddard, "Accurate ab initio quantum mechanics simulations of Bi_2Se_3 and Bi_2Te_3 topological insulator surfaces", J. Phys. Chem. Lett. **6**, 3792–3796 (2015).

[65] J. Muscat, A. Wander, and N. Harrison, "On the prediction of band gaps from hybrid functional theory", Chem. Phys. Lett. **342**, 397–401 (2001).

[66] J. E. Peralta, J. Heyd, G. E. Scuseria, and R. L. Martin, "Spin-orbit splittings and energy band gaps calculated with the heyd-scuseria-ernzerhof screened hybrid functional", Phys. Rev. B **74**, 073101 (2006).

[67] Y. Matsuda, J. Tahir-Kheli, and W. A. Goddard, "Definitive band gaps for single-wall carbon nanotubes", J. Phys. Chem. Lett. **1**, 2946–2950 (2010).

[68] H. Xiao, J. Tahir-Kheli, and W. A. Goddard, "Accurate band gaps for semiconductors from density functional theory", J. Phys. Chem. Lett. **2**, 212–217 (2011).

[69] N. Schwentner, F. .-.-J. Himpsel, V. Saile, M. Skibowski, W. Steinmann, and E. E. Koch, "Photoemission from rare-gas solids: electron energy distributions from the valence bands", Phys. Rev. Lett. **34**, 528–531 (1975).

[70] J. M. Crowley, J. Tahir-Kheli, and I. William A. Goddard, "Resolution of the band gap prediction problem for materials design", J. Phys. Chem. Lett. **7**, 1198–1203 (2016).

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

[72] H. Xiao and W. A. Goddard, "Predicted roles of defects on band offsets and energetics at cigs (cu(in,ga)se2/cds) solar cell interfaces and implications for improving performance", J. Chem. Phys. **141**, 094701 (2014) `http://dx.doi.org/10.1063/1.4893985`.

[73] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior", Phys. Rev. A **38**, 3098–3100 (1988).

[74] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation", Phys. Rev. B **46**, 6671–6687 (1992).

[75] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg, "Systematically convergent basis sets with relativistic pseudopotentials. ii. small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements", J. Chem. Phys. **119**, 11113–11123 (2003).

[76] H. Stoll, B. Metz, and M. Dolg, "Relativistic energy-consistent pseudopotentials–recent developments", J. Comput. Chem. **23**, 767–778 (2002).

[77] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density", Phys. Rev. B **37**, 785–789 (1988).

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

[79] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: the pbe0 model", The Journal of Chemical Physics **110**, 6158–6170 (1999).

[80] Y. Zhao and D. G. Truhlar, "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions", The Journal of Chemical Physics **125**, 194101 (2006) `http://dx.doi.org/10.1063/1.2370993`.

[81] Y. Zhao and D. G. Truhlar, "Calculation of semiconductor band gaps with the m06-l density functional", The Journal of Chemical Physics **130**, 074103, 074103 (2009).

[82] Y. Zhao and D. G. Truhlar, "The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals", Theoretical Chemistry Accounts **120**, 215–241 (2008).

[83] J. H. Skone, M. Govoni, and G. Galli, "Self-consistent hybrid functional for condensed systems", Phys. Rev. B **89**, 195112 (2014).

[84] M. Gerosa, C. E. Bottani, L. Caramella, G. Onida, C. Di Valentin, and G. Pacchioni, "Electronic structure and phase stability of oxide semiconductors: performance of dielectric-dependent hybrid functional dft, benchmarked against *GW* band structure calculations and experiments", Phys. Rev. B **91**, 155201 (2015).

[85] H. J. Monkhorst, "Hartree-fock density of states for extended systems", Phys. Rev. B **20**, 1504–1513 (1979).

[86] N. W. Ashcroft and N. D. Mermin, *Solid state physics* (Brooks/Cole, Belmont, CA, 1976).

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

[88] J. L. Fry, N. E. Brener, and R. K. Bruyere, "Hartree-fock formalism for solids. ii. application to the nearly-free-electron gas", Phys. Rev. B **16**, 5225–5232 (1977).

[89] R. Hamming, *Numerical methods for scientists and engineers* (Dover, Mineola, New York, 1986).

[90] M. Z. Hasan and C. L. Kane, "*Colloquium*: topological insulators", Rev. Mod. Phys. **82**, 3045–3067 (2010).

[91] X.-L. Qi and S.-C. Zhang, "Topological insulators and superconductors", Rev. Mod. Phys. **83**, 1057–1110 (2011).

[92] Y. L. Chen, J. G. Analytis, J.-H. Chu, Z. K. Liu, S.-K. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z.-X. Shen, "Experimental realization of a three-dimensional topological insulator, Bi_2Te_3", Science **325**, 178–181 (2009).

[93] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. Hor, R. Cava, and M. Hasan, "Observation of a large-gap topological-insulator class with a single dirac cone on the surface", Nature Physics **5** (2009).

[94] O. V. Yazyev, E. Kioupakis, J. E. Moore, and S. G. Louie, "Quasiparticle effects in the bulk and surface-state bands of bi_2se_3 and bi_2te_3 topological insulators", Phys. Rev. B **85**, 161101 (2012).

[95] J. Vidal, X. Zhang, L. Yu, J.-W. Luo, and A. Zunger, "False-positive and false-negative assignments of topological insulators in density functional theory and hybrids", Phys. Rev. B **84**, 041109 (2011).

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

[97] B. Sa, J. Zhou, Z. Sun, J. Tominaga, and R. Ahuja, "Topological insulating in $GeTe/Sb_2Te_3$ phase-change superlattice", Phys. Rev. Lett. **109**, 096802 (2012).

[98] B. Yan, M. Jansen, and C. Felser, "A large-energy-gap oxide topological insulator based on the superconductor $BaBiO_3$", Nat. Phys. **9**, 709–711 (2013).

[99] I. Aguilera, C. Friedrich, and S. Blügel, "Spin-orbit coupling in quasiparticle studies of topological insulators", Phys. Rev. B **88**, 165136 (2013).

[100] G. Wang and T. Cagin, "Electronic structure of the thermoelectric materials Bi_2Te_3 and Sb_2Te_3 from first-principles calculations", Phys. Rev. B **76**, 075201 (2007).

[101] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, "Topological insulators in bi$_2$se$_3$, bi$_2$te$_3$, and sb$_2$te$_3$ with a single dirac cone on the surface", Nature Physics **5**, 438 (2009).

[102] X. Luo, M. B. Sullivan, and S. Y. Quek, "First-principles investigations of the atomic, electronic, and thermoelectric properties of equilibrium and strained bi$_2$se$_3$ and bi$_2$te$_3$ including van der waals interactions", Phys. Rev. B **86**, 184111 (2012).

[103] M. Michiardi, I. Aguilera, M. Bianchi, V. E. de Carvalho, L. O. Ladeira, N. G. Teixeira, E. A. Soares, C. Friedrich, S. Blügel, and P. Hofmann, "Bulk band structure of Bi$_2$Te$_3$", Phys. Rev. B **90**, 075105 (2014).

[104] I. A. Nechaev, R. C. Hatch, M. Bianchi, D. Guan, C. Friedrich, I. Aguilera, J. L. Mi, B. B. Iversen, S. Blügel, P. Hofmann, and E. V. Chulkov, "Evidence for a direct band gap in the topological insulator Bi$_2$Se$_3$ from theory and experiment", Phys. Rev. B **87**, 121111 (2013).

[105] Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, "Crossover of the three-dimensional topological insulator bi$_2$se$_3$ to the two-dimensional limit", Nature Physics **6**, 584 (2010).

[106] O. V. Yazyev, J. E. Moore, and S. G. Louie, "Spin polarization and transport of surface states in the topological insulators Bi$_2$Se$_3$ and Bi$_2$Te$_3$ from first principles", Phys. Rev. Lett. **105**, 266806 (2010).

[107] D.-X. Qu, Y. Hor, J. Xiong, R. Cava, and N. Ong, "Quantum oscillations and hall anomaly of surface states in the topological insulator Bi$_2$Te$_3$", Science **329**, 821–824 (2010).

[108] Y. Xia, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. Hor, R. Cava, and M. Hasan, "Observation of a large-gap topological-insulator class with a single dirac cone on the surface", Nat. Phys. **5**, 398–402 (2009).

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

[110] R. Dovesi, R. Orlando, A. Erba, C. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, P. D'Arco, Y. Noel, M. Causa, M. Rerat, and B. Kirtman, Int. J. Quantum Chem. **114**, 1287 (2014).

[111] D. D. Johnson, "Modified Broyden's method for accelerating convergence in self-consistent calculations", Phys. Rev. B **38**, 12807–12813 (1988).

[112] C. Pérez Vicente, J. L. Tirado, K. Adouby, J. C. Jumas, A. A. Touré, and G. Kra, "X-ray diffraction and [119]Sn Mössbauer spectroscopy study of a new phase in the Bi$_2$Se$_3$-SnSe system: SnBi$_4$Se$_7$", Inorg. Chem. **38**, PMID: 11670996, 2131–2135 (1999).

[113] Y. Feutelais, B. Legendre, N. Rodier, and V. Agafonov, "A study of the phases in the bismuth - tellurium system", Mater. Res. Bull. **28**, 591–596 (1993).

[114] C. Gatti, V. R. Saunders, and C. Roetti, "Crystal field effects on the topological properties of the electron density in molecular crystals: the case of urea", The Journal of Chemical Physics **101**, 10686–10696 (1994).

[115] R. N. Brown, J. G. Mavroides, and B. Lax, "Magnetoreflection in bismuth", Phys. Rev. **129**, 2055–2061 (1963).

[116] M. Maltz and M. S. Dresselhaus, "Magnetoreflection studies in bismuth", Phys. Rev. B **2**, 2877–2887 (1970).

[117] M. P. Vecchi and M. S. Dresselhaus, "Temperature dependence of the band parameters of bismuth", Phys. Rev. B **10**, 771–774 (1974).

[118] R. T. Isaacson and G. A. Williams, "Alfvén-wave propagation in solid-stae plasmas. iii. quantum oscillations of the fermi surface of bismuth", Phys. Rev. **185**, 682–688 (1969).

[119] G. E. Smith, G. A. Baraff, and J. M. Rowell, "Effective g factor of electrons and holes in bismuth", Phys. Rev. **135**, A1118–A1124 (1964).

[120] J. Black, E. Conwell, L. Seigle, and C. Spencer, "Electrical and optical properties of some m2v-bn3vi-b semiconductors", J. Phys. Chem. Solids **2**, 240–251 (1957).

[121] I. Aguilera, C. Friedrich, and S. Blügel, "Spin-orbit coupling in quasiparticle studies of topological insulators", Phys. Rev. B **88**, 165136 (2013).

[122] I. A. Nechaev and E. V. Chulkov, "Quasiparticle band gap in the topological insulator bi_2te_3", Phys. Rev. B **88**, 165135 (2013).

[123] R. Dalven, "Electronic structure of pbs, pbse, and pbte", in, Vol. 28, edited by F. S. Henry Ehrenreich and D. Turnbull, Solid State Physics (Academic Press, 1974), pp. 179–224.

[124] K. Hummer, A. Grüneis, and G. Kresse, "Structural and electronic properties of lead chalcogenides from first principles", Phys. Rev. B **75**, 195211 (2007).

[125] A. Svane, N. E. Christensen, M. Cardona, A. N. Chantis, M. van Schilfgaarde, and T. Kotani, "Quasiparticle self-consistent GW calculations for pbs, pbse, and pbte: band structure and pressure coefficients", Phys. Rev. B **81**, 245120 (2010).

[126] C.-Y. Li, A. L. Ruoff, and C. W. Spencer, "Effect of pressure on the energy gap of bi2te3", Journal of Applied Physics **32**, 1733–1735 (1961).

[127] Y.-S. Kim, M. Marsman, G. Kresse, F. Tran, and P. Blaha, "Towards efficient band structure and effective mass calculations for iii-v direct band-gap semiconductors", Phys. Rev. B **82**, 205212 (2010).

[128] I. A. Nechaev, I. Aguilera, V. De Renzi, A. di Bona, A. Lodi Rizzini, A. M. Mio, G. Nicotra, A. Politano, S. Scalese, Z. S. Aliev, M. B. Babanly, C. Friedrich, S. Blügel, and E. V. Chulkov, "Quasiparticle spectrum and plasmonic excitations in the topological insulator Sb_2Te_3", Phys. Rev. B **91**, 245123 (2015).

[129] A. Fleszar and W. Hanke, "Electronic structure of II^B-VI semiconductors in the GW approximation", Phys. Rev. B **71**, 045207 (2005).

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

[131] H. W. Verleur, A. S. Barker, and C. N. Berglund, "Optical properties of vO_2 between 0.25 and 5 ev", Phys. Rev. **172**, 788–798 (1968).

[132] J. E. Coulter, E. Manousakis, and A. Gali, "Limitations of the hybrid functional approach to electronic structure of transition metal oxides", Phys. Rev. B **88**, 041107 (2013).

[133] S. Lany, "Band-structure calculations for the $3d$ transition metal oxides in GW", Phys. Rev. B **87**, 085112 (2013).

[134] J. Wu, W. Walukiewicz, K. M. Yu, J. W. Ager, E. E. Haller, H. Lu, W. J. Schaff, Y. Saito, and Y. Nanishi, "Unusual properties of the fundamental band gap of inn", Applied Physics Letters **80**, 3967–3969 (2002).

[135] T. Matsuoka, H. Okamoto, M. Nakao, H. Harima, and E. Kurimoto, "Optical bandgap energy of wurtzite inn", Applied Physics Letters **81**, 1246–1248 (2002).

[136] P. A. Anderson, C. E. Kendrick, R. J. Kinsey, V. J. Kennedy, A. Markwitz, R. J. Reeves, and S. M. Durbin, "Optical and compositional properties of indium nitride grown by plasma assisted molecular beam epitaxy", Smart Materials and Structures **15**, S87 (2006).

[137] F. Bechstedt and J. Furthmüller, "Do we know the fundamental energy gap of inn?", Journal of Crystal Growth **246**, Proceedings of the International Workshop on Bulk Nitride Semicon ductors, 315–319 (2002).

[138] J. Furthmüller, P. H. Hahn, F. Fuchs, and F. Bechstedt, "Band structures and optical spectra of inn polymorphs: influence of quasiparticle and excitonic effects", Phys. Rev. B **72**, 205106 (2005).

[139] P. Rinke, M. Scheffler, A. Qteish, M. Winkelnkemper, D. Bimberg, and J. Neugebauer, "Band gap and band parameters of inn and gan from quasiparticle energy calculations based on exact-exchange density-functional theory", Applied Physics Letters **89**, 161919 (2006) http://dx.doi.org/10.1063/1.2364469.

[140] W. Chen and A. Pasquarello, "Accurate band gaps of extended systems via efficient vertex corrections in GW", Phys. Rev. B **92**, 041115 (2015).

[141] M. Parenteau and C. Carlone, "Influence of temperature and pressure on the electronic transitions in sns and snse semiconductors", Phys. Rev. B **41**, 5227–5234 (1990).

[142] P. A. Fernandes, M. G. Sousa, P. M. P. Salome, J. P. Leitao, and A. F. da Cunha, "Thermodynamic pathway for the formation of snse and snse2 polycrystalline thin films by selenization of metal precursors", CrystEngComm **15**, 10278–10286 (2013).

[143] L.-D. Z. and Shih-Han Lo, Y. Zhang, H. S. and Gangjian Tan and Ctirad Uher and C. Wolverton and Vinayak P. Dravid, and M. G. Kanatzidis, "Ultralow thermal conductivity and high thermoelectric figure of merit in snse crystals", Nature **508**, 373 (2014).

[144] R. Guo, X. Wang, Y. Kuang, and B. Huang, "First-principles study of anisotropic thermoelectric transport properties of iv-vi semiconductor compounds snse and sns", Phys. Rev. B **92**, 115202 (2015).

[145] W. Kang and M. S. Hybertsen, "Enhanced static approximation to the electron self-energy operator for efficient calculation of quasiparticle energies", Phys. Rev. B **82**, 195108 (2010).

[146] B. Arnaud and M. Alouani, "All-electron projector-augmented-wave GW approximation: application to the electronic properties of semiconductors", Phys. Rev. B **62**, 4464–4476 (2000).

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

[148] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically thin MoS_2: a new direct-gap semiconductor", Phys. Rev. Lett. **105**, 136805 (2010).

[149] G. Fonthal, L. Tirado-Mejía, J. Marín-Hurtado, H. Ariza-Calderón, and J. Mendoza-Alvarez, "Temperature dependence of the band gap energy of crystalline cdte", J. Phys. Chem. Solids **61**, 579–583 (2000).

[150] O. Zakharov, A. Rubio, X. Blase, M. L. Cohen, and S. G. Louie, "Quasiparticle band structures of six ii-vi compounds: zns, znse, znte, cds, cdse, and cdte", Phys. Rev. B **50**, 10780–10787 (1994).

[151] S. Sze and K. K. Ng, *Physics of semiconductor devices*, 3rd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).

[152] M. P. Surh, S. G. Louie, and M. L. Cohen, "Quasiparticle energies for cubic bn, bp, and bas", Phys. Rev. B **43**, 9126–9132 (1991).

[153] P. W. Baumeister, "Optical absorption of cuprous oxide", Phys. Rev. **121**, 359–362 (1961).

[154] H. Bowen, D. Adler, and B. Auker, "Electrical and optical properties of feo", Journal of Solid State Chemistry **12**, 355–359 (1975).

[155]L.-H. Ye, R. Asahi, L.-M. Peng, and A. J. Freeman, "Model gw study of the late transition metal monoxides", J. Chem. Phys. **137**, 154110, 154110 (2012).

[156]H. Jiang, R. I. Gomez-Abal, P. Rinke, and M. Scheffler, "First-principles modeling of localized d states with the GW@LDA + U approach", Phys. Rev. B **82**, 045108 (2010).

[157]S. Tokunaga, H. Kato, and A. Kudo, "Selective preparation of monoclinic and tetragonal bivo4 with scheelite structure and their photocatalytic properties", Chemistry of Materials **13**, 4624–4628 (2001).

[158]S. Massidda, A. Continenza, M. Posternak, and A. Baldereschi, "Quasiparticle energy bands of transition-metal oxides within a model gw scheme", Phys. Rev. B **55**, 13494–13502 (1997).

[159]W. J. Choyke, D. R. Hamilton, and L. Patrick, "Optical properties of cubic sic: luminescence of nitrogen-exciton complexes, and interband absorption", Phys. Rev. **133**, A1163–A1166 (1964).

[160]B. Wenzien, P. Käckell, F. Bechstedt, and G. Cappellini, "Quasiparticle band structure of silicon carbide polytypes", Phys. Rev. B **52**, 10897–10905 (1995).

[161]R. T. M. Ummels, P. A. Bobbert, and W. van Haeringen, "*Ab initio* quasiparticle energies in $2H$, $4H$, and $6H$ sic", Phys. Rev. B **58**, 6795–6799 (1998).

[162]M. Gvishi and D. Tannhauser, "Hall mobility and defect structure in undoped and cr or ti-doped coo at high temperature", J. Phys. Chem. Solids **33**, 893–911 (1972).

[163]E. Z. Kurmaev, R. G. Wilks, A. Moewes, L. D. Finkelstein, S. N. Shamin, and J. Kune š, "Oxygen x-ray emission and absorption spectra as a probe of the electronic structure of strongly correlated oxides", Phys. Rev. B **77**, 165127 (2008).

[164]J. van Elp, J. L. Wieland, H. Eskes, P. Kuiper, G. A. Sawatzky, F. M. F. de Groot, and T. S. Turner, "Electronic structure of coo, li-doped coo, and $LiCoO_2$", Phys. Rev. B **44**, 6090–6103 (1991).

[165]G. W. Pratt and R. Coelho, "Optical absorption of coo and mno above and below the néel temperature", Phys. Rev. **116**, 281–286 (1959).

[166]R. J. Powell and W. E. Spicer, "Optical properties of nio and coo", Phys. Rev. B **2**, 2182–2193 (1970).

[167]S. Das, J. E. Coulter, and E. Manousakis, "Convergence of quasiparticle self-consistent GW calculations of transition-metal monoxides", Phys. Rev. B **91**, 115105 (2015).

[168]T. D. Kang, H. S. Lee, and H. Lee, "Optical properties of black nio and coo single crystals studied with spectroscopic ellipsometry", J. Korean Phys. Soc. **50**, 632.

[169] V. Bougrov, M. Levinshtein, S. Rumyantsev, and A. Zubrilov, "Gallium nitride (gan)", in *Properties of advanced semiconductor materials*, edited by M. Levinshtein, S. Rumyantsev, and M. Shur (John Wiley & Sons, Inc., New York, 2001), pp. 1–30.

[170] M. Takizawa, K. Maekawa, H. Wadati, T. Yoshida, A. Fujimori, H. Kumigashira, and M. Oshima, "Angle-resolved photoemission study of nb-doped $SrTiO_3$", Phys. Rev. B **79**, 113103 (2009).

[171] K. van Benthem, C. Elsässer, and R. H. French, "Bulk electronic structure of srtio3: experiment and theory", J. Appl. Phys. **90**, 6156–6164 (2001).

[172] Y. Tezuka, S. Shin, T. Ishii, T. Ejima, S. Suzuki, and S. Sato, "Photoemission and bremsstrahlung isochromat spectroscopy studies of tio2 (rutile) and srtio3", J. Phys. Soc. Jpn. **63**, 347–357 (1994).

[173] A. Benrekia, N. Benkhettou, A. Nassour, M. Driz, M. Sahnoun, and S. Lebègue, "Structural, electronic and optical properties of cubic srtio3 and ktao3: ab initio and {gw} calculations", Physica B: Condensed Matter **407**, 2632–2636 (2012).

[174] G. Kang, Y. Kang, and S. Han, "Influence of wave-function updates in *GW* calculations on titanates", Phys. Rev. B **91**, 155141 (2015).

[175] M. Landmann, E. Rauls, and W. G. Schmidt, "The electronic structure and optical response of rutile, anatase and brookite tio 2", Journal of Physics: Condensed Matter **24**, 195503 (2012).

[176] L. Patrick, D. R. Hamilton, and W. J. Choyke, "Growth, luminescence, selection rules, and lattice sums of sic with wurtzite structure", Phys. Rev. **143**, 526–536 (1966).

[177] H. Tang, F. Lévy, H. Berger, and P. E. Schmid, "Urbach tail of anatase TiO_2", Phys. Rev. B **52**, 7771–7774 (1995).

[178] R. Iskenderov, I. Drabkin, L. Emel'yanova, and Y. Ksendzov, Fiz. Tverd. Tela **10**, 2573 (1968).

[179] I. A. Drabkin, L. T. Emel'yanova, R. N. Iskenderov, and Y. M. Ksendzov, Fiz. Tverd. Tela **10**, 3082 (1968).

[180] J. van Elp, R. H. Potze, H. Eskes, R. Berger, and G. A. Sawatzky, "Electronic structure of mno", Phys. Rev. B **44**, 1530–1537 (1991).

[181] D. R. Huffman, R. L. Wild, and M. Shinmei, "Optical absorption spectra of crystal-field transitions in mno", J. Chem. Phys. **50**, 4092–4094 (1969).

[182] J. M. Hartmann, J. Cibert, F. Kany, H. Mariette, M. Charleux, P. Alleysson, R. Langer, and G. Feuillet, "Cdte/mgte heterostructures: growth by atomic layer epitaxy and determination of mgte parameters", Journal of Applied Physics **80**, 6257–6265 (1996).

[183] J. E. Jaffe, T. C. Kaspar, T. C. Droubay, T. Varga, M. E. Bowden, and G. J. Exarhos, "Electronic and defect structures of cuscn", J. Phys. Chem. C **114**, 9111–9117 (2010).

[184] G. A. Sawatzky and J. W. Allen, "Magnitude and origin of the band gap in nio", Phys. Rev. Lett. **53**, 2339–2342 (1984).

[185] Y. M. Ksendzov and I. A. Drabkin, Fiz. Tverd. Tela, 1884 (1965).

[186] I. Vurgaftman and J. R. Meyer, "Band parameters for nitrogen-containing semi-conductors", J. Appl. Phys. **94**, 3675–3696 (2003).

[187] D. A. Evans, A. G. McGlynn, B. M. Towlson, M. Gunn, D. Jones, T. E. Jenkins, R. Winter, and N. R. J. Poolton, "Determination of the optical band-gap energy of cubic and hexagonal boron nitride using luminescence excitation spectroscopy", Journal of Physics: Condensed Matter **20**, 075233 (2008).

[188] R. Whited, C. J. Flaten, and W. Walker, "Exciton thermoreflectance of mgo and cao", Solid State Communications **13**, 1903–1905 (1973).

[189] R. T. Poole, J. G. Jenkin, J. Liesegang, and R. C. G. Leckey, "Electronic band structure of the alkali halides. i. experimental parameters", Phys. Rev. B **11**, 5179–5189 (1975).

[190] D. M. Roessler and W. C. Walker, "Electronic spectra of crystalline nacl and kcl", Phys. Rev. **166**, 599–606 (1968).

[191] T. DiStefano and D. Eastman, "The band edge of amorphous sio2 by photo-injection and photoconductivity measurements", Solid State Communications **9**, 2259–2261 (1971).

[192] L. E. Ramos, J. Furthmüller, and F. Bechstedt, "Quasiparticle band structures and optical spectra of β-cristobalite SiO_2", Phys. Rev. B **69**, 085102 (2004).

[193] G. Baldini and B. Bosacchi, "Optical properties of na and li halide crystals at 55 °k", physica status solidi (b) **38**, 325–334 (1970).

[194] L. A. Garvie, P. Rez, J. R. Alvarez, P. R. Buseck, A. J. Craven, and R. Brydson, "Bonding in alpha-quartz (SiO_2): a view of the unoccupied states", American Mineralogist **85**, 732 (2000).

[195] M. Piacentini, D. W. Lynch, and C. G. Olson, "Thermoreflectance of lif between 12 and 30 ev", Phys. Rev. B **13**, 5530–5543 (1976).

[196] B. Rafferty and L. M. Brown, "Direct and indirect transitions in the region of the band gap using electron-energy-loss spectroscopy", Phys. Rev. B **58**, 10326–10337 (1998).

[197] T. Miyata and T. Tomiki, "Optical studies of nacl single crystals in 10 ev region ii. the spectra of conductivity at low temperatures, absorption constant and energy loss", J. Phys. Soc. Jpn. **24**, 1286–1302 (1968).