

Optimized effective potential using the Hylleraas variational method

T. W. Hollins* and S. J. Clark†

Department of Physics, Science Laboratories, University of Durham, South Road, Durham DH1 3LE, England, United Kingdom

K. Refson‡ and N. I. Gidopoulos§

Science and Technology Facilities Council, Rutherford Appleton Laboratory, Didcot, Oxfordshire OX11 0QX, England, United Kingdom

(Received 29 March 2012; published 15 June 2012)

In electronic structure calculations the optimized effective potential (OEP) is a method that treats exchange interactions exactly using a local potential within density-functional theory. In the exchange-only case, this is commonly called exact exchange. We present a method using density-functional perturbation theory combined with the Hylleraas variational method for finding the OEP by direct minimization, which avoids any sum over unoccupied states. The method has been implemented within the plane-wave, pseudopotential formalism. Band structures for zinc-blende semiconductors Si, Ge, C, GaAs, CdTe, and ZnSe; wurtzite semiconductors InN, GaN, and ZnO; and the rocksalt insulators CaO and NaCl have been calculated using the OEP and compared to calculations using the local-density approximation (LDA), a selection of generalized gradient approximations (GGAs) and Hartree-Fock (HF) functionals. The band gaps found with the OEP improve on the calculated results for the LDA, GGAs, or HF, with calculated values of 1.16 eV for Si, 3.32 eV for GaN, and 3.48 eV for ZnO. The OEP energies of semicore d states are also greatly improved compared to LDA.

DOI: [10.1103/PhysRevB.85.235126](https://doi.org/10.1103/PhysRevB.85.235126)

PACS number(s): 71.15.-m, 71.15.Mb, 71.20.-b

I. INTRODUCTION

Density-functional theory (DFT) is one of the great successes in tackling the many-electron problem.¹⁻³ Density-dependent exchange-correlation potentials, such as the local-density approximation (LDA) or generalized gradient approximations (GGAs), are useful and surprisingly accurate for many properties of materials despite the known deficiencies of these functionals. Generally, the LDA slightly underestimates lattice constants, while GGAs slightly overestimate them, and both underestimate single-particle band gaps in solids.⁴⁻⁷ The exchange energy can be calculated exactly using the nonlocal Hartree-Fock (HF) formalism, rather than being approximated as part of the LDA or GGA. However, pure HF grossly overestimates band gaps in solids and is considerably more expensive computationally because of the fully nonlocal nature of the exchange operator.

The central theorem of density-functional theory¹ demonstrates that a local potential $V(\mathbf{r})$ is sufficient to completely define a many-electron system in the ground state (GS). A local potential that includes the effect of the exact exchange interaction was first proposed by Sharp and Horton⁸ and solved nearly a quarter of a century later by Talman and Shadwick.⁹ This potential is known as the optimized effective potential (OEP), because a local potential is created to optimally represent the nonlocal HF exchange potential. The total energy of the OEP system is variational with the potential in the same way the LDA or GGA total energy is variational in the electronic density. This allows the construction of equations to describe the potential within a Kohn-Sham (KS) DFT formalism. The resultant single-particle excitation energies agree much better with experiment than those calculated using the LDA, GGAs, and HF.³

The improved OEP description of excited states is a consequence of the freedom from the self-interaction error of both occupied and unoccupied Kohn-Sham states.¹⁰ Within the LDA and GGAs the potential experienced by all states

includes some degree of self-interaction.¹¹ In HF the exchange term cancels the self-interaction error in the occupied states, but no such cancellation is present for the unoccupied states. This results in too small an electronic band gap using the LDA and GGAs and too large a one in HF.¹²⁻¹⁴

There is a distinction between the fundamental band gap of a material and the corresponding Kohn-Sham gap. The Kohn-Sham gap (E_g^{KS}) is defined as the difference between the eigenvalues of the highest occupied orbital and the lowest unoccupied orbital:

$$E_g^{\text{KS}} = \epsilon_{N+1} - \epsilon_N, \quad (1)$$

where ϵ_i is the eigenvalue of the i th orbital and N is the total number of electrons in the system. Whereas the fundamental band gap (E_g) is defined as the difference in the ionization potential and the electron affinity,

$$E_g = I - A = E(N+1) - 2E(N) + E(N-1), \quad (2)$$

where I is the ionization potential, A is the electron affinity and $E(N)$ is the total energy of an N -electron system. The two are related by

$$E_g = E_g^{\text{KS}} + \Delta_{\text{xc}}, \quad (3)$$

where Δ_{xc} is the derivative discontinuity in XC (exchange-correlation) energy with respect to particle number.^{6,7,15} This constant is the energy associated with a change in the exchange-correlation potential when an infinitesimal charge is added to an N -electron system. While a comparison between the Kohn-Sham and fundamental band gaps would be instructive, the value of the derivative discontinuity is unknown for most real materials (see, however, Godby *et al.*¹⁶ and Stadele *et al.*^{17,18}).

In finite systems, the exchange potential should decay as $-1/r$ in the long-range limit for all states. In the OEP this decay occurs for all states irrespective of occupancy. In HF the nonlocal potential for occupied states behaves correctly but

for the empty states decays exponentially, and in the LDA/GGA the potential for all states decays exponentially.¹⁹ Furthermore, the orbitals in the OEP each correctly decay with an individual exponent, whereas in HF all occupied orbitals decay with the same exponent.^{3,10,12-14}

The OEP is potentially as versatile as LDA/GGA and HF methods. It could be used instead of HF as a foundation for orbital-dependent potentials (so-called hybrid functionals).³ Furthermore once the local potential has been obtained, the calculation of system properties is faster than HF based methods, as the computationally expensive application of the exchange operator is avoided.

While the OEP correctly describes many system properties (see Grabo *et al.*¹²), it is not as widely used as the LDA, GGA, and HF-hybrid based methods. This can be attributed in part to the lack of comparably efficient and robust computational schemes to evaluate it. In most formulations of the OEP method a sum over all excited states is required and truncation to a finite sum yields a slowly convergent series. Furthermore an accurate description of these high-energy states requires large basis sets at high computational expense if a local basis description is used. Despite these deficiencies, some calculations have been done in solids, within the linear muffin-tin orbital (LMTO) basis set²⁰⁻²² and the full potential linearized augmented plane-wave (FPLAPW) method,²³ as well as in plane-wave pseudopotential implementations.^{17,18,24-27} The agreement of the calculated results in the above-mentioned previous work with experiment can be very good (see Ref. 10). However, most applications have been to semiconductors. For wide-gap insulating systems the performance of the OEP can be poor, notably for noble-gas solids.²⁷ See Ref. 10 for full details and further discussion of the OEP method and its results.

The extremely demanding task of calculating the full OEP can be reduced by using the mean-field approximation of Krieger, Li, and Iafrate²⁸ (well known as the KLI approximation), and even in this less precise formalism impressively accurate results have been obtained.^{24-27,29}

The principal aim of this work is to derive and demonstrate a variational method of calculating the full exchange-only OEP without the need for a sum over all unoccupied states of a system. To this end, our method for the calculation of the OEP is formulated using density-functional perturbation theory and the Hylleraas variational principle. As in the case of phonon or electric-field perturbations, only occupied Kohn-Sham orbitals are explicitly included. This method is then applied to a range of semiconductors and insulators.

II. THEORY

A. Optimized effective potential

We first summarize the OEP before deriving a variational implementation within the density-functional perturbation formalism. For the remainder of this article we consider only the exchange-only OEP (exact exchange) and will simply use the term OEP from here on. A starting point for finding the local exchange-only OEP is the nonlocal Hartree-Fock equation

$$[\hat{T}(\mathbf{r}) + \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}_{\text{H}}(\mathbf{r}) + \hat{V}_{\text{X}}^{\sigma}(\mathbf{r})]\phi_i^{\sigma}(\mathbf{r}) = \epsilon_i^{\sigma} \phi_i^{\sigma}(\mathbf{r}), \quad (4)$$

where $\hat{T}(\mathbf{r})$ is the kinetic energy, $\hat{V}_{\text{ext}}(\mathbf{r})$ is the external potential, $\hat{V}_{\text{H}}(\mathbf{r})$ is the Hartree potential, $\hat{V}_{\text{X}}^{\sigma}(\mathbf{r})$ is the HF exchange potential, i indexes the electronic states, σ is the spin index, and $\phi_i^{\sigma}(\mathbf{r})$ is the i th orbital with spin σ . The Kohn-Sham (KS) equation^{1,2} that incorporates the OEP is given by

$$[\hat{T}(\mathbf{r}) + \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}^{\sigma}(\mathbf{r})]\phi_i^{\sigma}(\mathbf{r}) = \epsilon_i^{\sigma} \phi_i^{\sigma}(\mathbf{r}), \quad (5)$$

where \hat{V}^{σ} is the effective potential fulfilling the role of both the Hartree and exchange potentials. The usual derivation of the OEP uses a chain rule expansion of the derivative of the exchange-correlation energy with respect to the density,¹² however here a treatment inspired by perturbation theory is used.

For a potential which differs from the ground-state (GS) potential by an amount $\delta V^{\sigma}(\mathbf{r})$ such that $V^{\sigma}(\mathbf{r}) \rightarrow V^{\sigma}(\mathbf{r}) + \delta V^{\sigma}(\mathbf{r})$ the GS KS orbitals change by

$$\phi_i^{\sigma}(\mathbf{r}) \rightarrow \phi_i^{\sigma}(\mathbf{r}) + \int d\mathbf{x} \delta V^{\sigma}(\mathbf{x}) \frac{\delta \phi_i^{\sigma}(\mathbf{r})}{\delta V^{\sigma}(\mathbf{x})}, \quad (6)$$

where from first-order perturbation theory³⁰ the last term of the above equation becomes

$$\begin{aligned} \frac{\delta \phi_i^{\sigma}(\mathbf{r})}{\delta V^{\sigma}(\mathbf{x})} = & - \sum_{j=1}^{N^{\sigma}} \frac{\phi_j^{\sigma}(\mathbf{r}) \phi_j^{\sigma*}(\mathbf{x})}{\epsilon_j^{\sigma} - \epsilon_i^{\sigma}} \phi_i^{\sigma}(\mathbf{x}) \\ & - \sum_{a=N^{\sigma}+1}^{\infty} \frac{\phi_a^{\sigma}(\mathbf{r}) \phi_a^{\sigma*}(\mathbf{x})}{\epsilon_a^{\sigma} - \epsilon_i^{\sigma}} \phi_i^{\sigma}(\mathbf{x}). \end{aligned} \quad (7)$$

The change in the effective potential also gives a first-order change in the total energy, thus

$$\Delta E[V^{\sigma}] = \int d\mathbf{x} \delta V^{\sigma}(\mathbf{x}) \frac{\delta E[V^{\sigma}]}{\delta V^{\sigma}(\mathbf{x})}. \quad (8)$$

The OEP energy $E[V^{\sigma}]$ is given by the Hartree-Fock functional evaluated using the Kohn-Sham OEP orbitals. Using the chain rule and the perturbation of the orbitals, the derivative of the energy with respect to the potential is given by

$$\frac{\delta E[V^{\sigma}]}{\delta V^{\sigma}(\mathbf{x})} = \int d\mathbf{r} \sum_{i=1}^{N^{\sigma}} \frac{\delta E[V^{\sigma}]}{\delta \phi_i^{\sigma}(\mathbf{r})} \frac{\delta \phi_i^{\sigma}(\mathbf{r})}{\delta V^{\sigma}(\mathbf{x})} + \text{H.c.} \quad (9)$$

By applying the Hellmann-Feynman theorem (for brevity the explicit \mathbf{r} dependence of the potentials has been omitted),

$$\frac{\delta E[V^{\sigma}]}{\delta \phi_i^{\sigma}(\mathbf{r})} = [\hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{H}} + \hat{V}_{\text{X}}^{\sigma}]\phi_i^{\sigma*}(\mathbf{r}), \quad (10)$$

and substituting Eqs. (7) and (10) into Eq. (9), we obtain

$$\begin{aligned} \frac{\delta E[V^{\sigma}]}{\delta V^{\sigma}(\mathbf{x})} = & - \int d\mathbf{r} \sum_{i=1}^{N^{\sigma}} \sum_{a=N^{\sigma}+1}^{\infty} \frac{\phi_a^{\sigma}(\mathbf{r}) \phi_a^{\sigma*}(\mathbf{x}) \phi_i^{\sigma}(\mathbf{r})}{\epsilon_a^{\sigma} - \epsilon_i^{\sigma}} \\ & \times [\hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{H}} + \hat{V}_{\text{X}}^{\sigma}]\phi_i^{\sigma*}(\mathbf{r}) + \text{H.c.} \end{aligned} \quad (11)$$

By substitution of the Kohn-Sham equation [Eq. (5)] the above can be written as³¹

$$\frac{\delta E[V^\sigma]}{\delta V^\sigma(\mathbf{x})} = - \sum_{i=1}^{N^\sigma} \sum_{a=N^\sigma+1}^{\infty} \left[\frac{\langle \phi_i^\sigma | \hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma | \phi_a^\sigma \rangle}{\epsilon_a^\sigma - \epsilon_i^\sigma} \phi_a^{*\sigma}(\mathbf{x}) \phi_i^\sigma(\mathbf{x}) \right] + \text{H.c.}, \quad (12)$$

which is known as the OEP equation and was first derived by Sharp and Horton.⁸ Here it is important to note that the requirement to sum over all unoccupied states presents a challenge for convergence of the system properties. A very large number of unoccupied orbitals must be included to ensure adequate convergence, and consequently the calculation is extremely demanding.³² The truncation of the infinite sum has also been the subject of concerns about the analytic properties of the solutions.³³

B. Applying the Hylleraas variational principle

We present the following formalism to evaluate the OEP, which avoids an explicit sum over states by borrowing ideas from density-functional perturbation theory^{30,34} using the Hylleraas variational method. The effective potential is obtained by variational minimization of $E[V^\sigma(\mathbf{r})]$. The minimization direction for the potential is defined by the functional derivative of the energy with respect to the effective potential:

$$V^\sigma(\mathbf{r}) \rightarrow V^\sigma(\mathbf{r}) - \lambda \frac{\delta E[V^\sigma]}{\delta V^\sigma(\mathbf{r})}. \quad (13)$$

The functional derivative of the energy can alternatively be written as

$$\frac{\delta E[V^\sigma]}{\delta V^\sigma(\mathbf{x})} = \sum_{i=1}^{N^\sigma} \phi_i^\sigma(\mathbf{x}) [\tilde{\phi}_i^\sigma(\mathbf{x})]^* + \text{H.c.}, \quad (14)$$

where

$$\tilde{\phi}_i^\sigma(\mathbf{x}) = - \sum_{a=N^\sigma+1}^{\infty} \frac{\phi_a^\sigma(\mathbf{x}) \langle \phi_a^\sigma | \hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma | \phi_i^\sigma \rangle}{\epsilon_a^\sigma - \epsilon_i^\sigma} \quad (15)$$

and $\tilde{\phi}_i^\sigma(\mathbf{x})$ is the first-order correction to an unperturbed orbital $\phi_i^\sigma(\mathbf{x})$. These first-order corrections to the orbitals were named ‘‘orbital shifts’’ by Kümmel and Perdew.^{13,35}

An alternative method of calculating $\tilde{\phi}_i^\sigma(\mathbf{x})$ without explicitly including any unoccupied states is to use the Hylleraas variational principle.³⁶ We define a second-order variational functional

$$G_i^\sigma[\tilde{\phi}_i^\sigma] = \langle \tilde{\phi}_i^\sigma | (\hat{T} + \hat{V}_{\text{ext}} + \hat{V}^\sigma - \epsilon_i^\sigma) | \tilde{\phi}_i^\sigma \rangle + \langle \tilde{\phi}_i^\sigma | (\hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma - \tilde{\epsilon}_i^\sigma) | \phi_i^\sigma \rangle + \langle \phi_i^\sigma | (\hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma - \tilde{\epsilon}_i^\sigma) | \tilde{\phi}_i^\sigma \rangle, \quad (16)$$

where $\tilde{\epsilon}_i^\sigma$ is the first-order correction to the eigenvalues. $G_i^\sigma[\tilde{\phi}_i^\sigma]$ is also variational with respect to $\tilde{\phi}_i^\sigma$ under the constraint

$$\langle \phi_i^\sigma | \tilde{\phi}_i^\sigma \rangle + \langle \tilde{\phi}_i^\sigma | \phi_i^\sigma \rangle = 0, \quad (17)$$

and using the Hylleraas variational principle it can be shown that the orbital shifts which minimize the second-order functional also satisfy Eq. (15). By substitution of the projection operator

$$P_c = \hat{I} - \sum_{j=1}^{N^\sigma} |\phi_j^\sigma\rangle\langle\phi_j^\sigma| = \sum_{a=N^\sigma+1}^{\infty} |\phi_a^\sigma\rangle\langle\phi_a^\sigma|, \quad (18)$$

Eq. (16) may be rewritten

$$G_i^\sigma[\tilde{\phi}_i^\sigma] = \langle \tilde{\phi}_i^\sigma | (\hat{T} + \hat{V}_{\text{ext}} + \hat{V}^\sigma - \epsilon_i^\sigma) | \tilde{\phi}_i^\sigma \rangle + \langle \tilde{\phi}_i^\sigma | \left(\hat{I} - \sum_{j \neq i}^{N^\sigma} |\phi_j^\sigma\rangle\langle\phi_j^\sigma| \right) (\hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma) | \phi_i^\sigma \rangle + \langle \phi_i^\sigma | (\hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma) \left(\hat{I} - \sum_{j \neq i}^{N^\sigma} |\phi_j^\sigma\rangle\langle\phi_j^\sigma| \right) | \tilde{\phi}_i^\sigma \rangle. \quad (19)$$

The exact first-order correction to the orbitals $\tilde{\phi}_i^\sigma$ that minimizes $G_i^\sigma[\tilde{\phi}_i^\sigma]$ can also be found using the Sternheimer-like equation³⁷

$$(\hat{T} + \hat{V}_{\text{ext}} + \hat{V}^\sigma - \epsilon_i^\sigma) | \tilde{\phi}_i^\sigma \rangle + \left(\hat{I} - \sum_{j=1}^{N^\sigma} |\phi_j^\sigma\rangle\langle\phi_j^\sigma| \right) (\hat{V}_H + \hat{V}_X^\sigma - \hat{V}^\sigma) | \phi_i^\sigma \rangle = 0. \quad (20)$$

Equation (20) can be solved using iterative methods.³⁴ Replacing the infinite sum over states in Eq. (15) with an iterative procedure and considering only the occupied Kohn-Sham subspace avoids the severe convergence difficulties of the former method. Instead of determining how many additional states to include in Eq. (15), at a computational cost increasing with their number, the convergence problem is transformed to one of determining how many iterative cycles are required. Consequently a reliable solution may be achieved at considerably lower expense than summing Eq. (15) directly.

By combining Eqs. (14) and (20), the OEP can be found variationally while the first-order correction to the orbitals is obtained by solving the Sternheimer equation directly.

Equations (14) and (20) are similar to those given by Kümmel and Perdew.^{13,35} Our method differs from theirs by exploiting the variational character of $E[V^\sigma]$ using the explicit gradient of Eq. (14) to perform a direct minimization. Their method involves a self-consistency cycle with an update procedure for v_{XC} using a KLI-like expression and involving a division by the density. Our direct minimization is conceptually simpler and, because of the variational principle, is likely to be more numerically robust.

III. IMPLEMENTATION AND CONVERGENCE

Our procedure for finding the OEP is to solve a double-nested loop of minimizations. A flow diagram illustrating the full procedure is shown in Fig. 1. The inner of the two loops represents the solution of the Sternheimer equation [Eq. (20)] to find the first-order correction to the orbitals. The method is derived from the Baroni Green’s-function solver technique, employing a conjugate gradient minimization scheme to find $\tilde{\phi}_i^\sigma(\mathbf{r})$ in a fixed potential $V^\sigma(\mathbf{r})$.^{34,38,39}

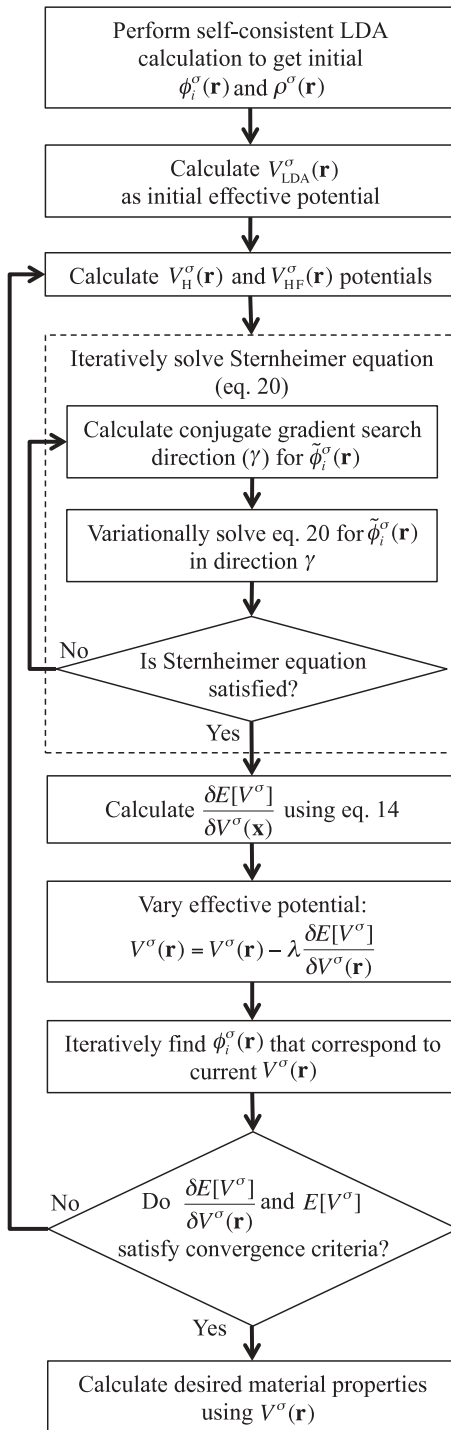


FIG. 1. Flow diagram showing procedure for finding the OEP.

Using the first-order correction to the orbitals and the orbitals themselves, the derivative of the total energy with respect to the potential can be found from Eq. (14). This gradient is used in the direction-set methods to variationally optimize the potential, and its magnitude (residual norm) provides an indication of the level of convergence of the effective potential. The variation of the potential is accomplished via either one of two methods; the first is a conjugate gradient scheme;⁴⁰ the second method is a modified steepest-descent method, known as the Barzilai-Borwein (BB) method, which requires

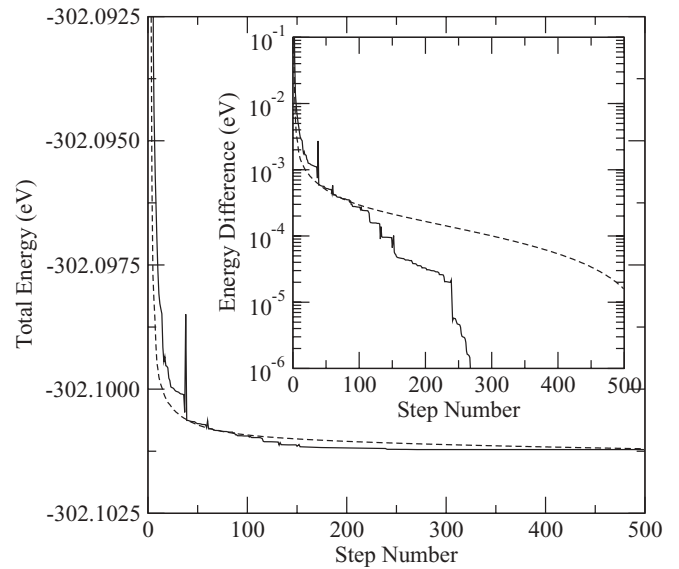


FIG. 2. Convergence of total Kohn-Sham energy for the OEP against the number of steps. The solid line is the Barzilai-Borwein minimizer and the dashed line is for the conjugate gradient minimizer. Inset: energy residual on a log-linear scale.

knowledge of the gradient at the current point and previous point only.⁴¹

After each step of varying the effective potential, Kohn-Sham orbitals in this new, fixed potential are found non-self-consistently. When the gradient and difference in total OEP Kohn-Sham energy per step are smaller than a predetermined threshold, the calculation is considered to be converged.

The procedure is initialized by first solving the KS equation using a local-density-dependent exchange-correlation potential, in this case the LDA. From this self-consistent calculation the initial orbitals and density are obtained. Using

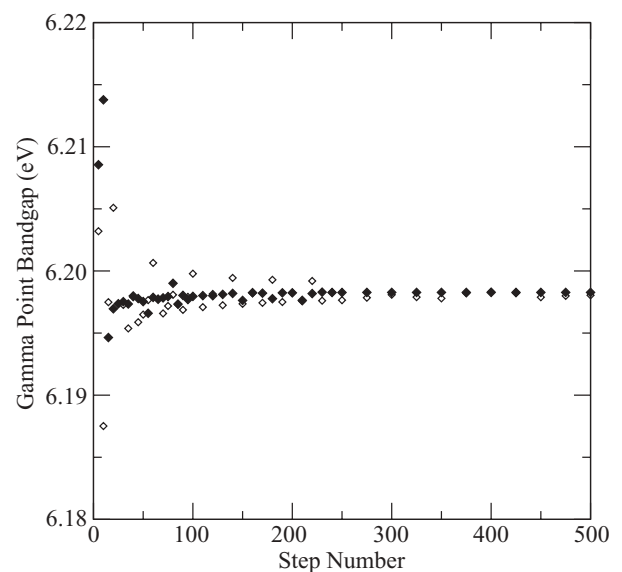


FIG. 3. Convergence of the Γ -point band gap of diamond using the OEP against the number of steps. Solid dots are for the Barzilai-Borwein minimizer and open dots are for the conjugate gradient minimizer.

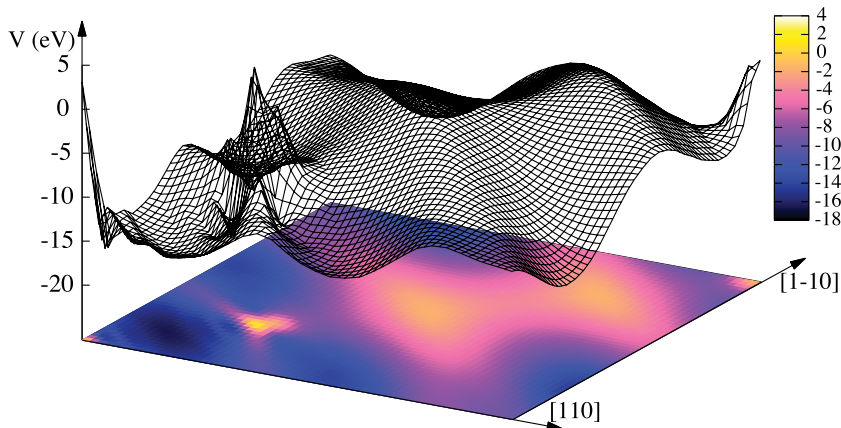


FIG. 4. (Color online) The exchange potential ($V_{\text{OEP}} - V_{\text{H}}$) as a slice perpendicular to the $[111]$ direction (bond axis) in diamond structured Si for the primitive unit cell. Atomic positions are at the origin and quarter of the long diagonal. The color key is in eV.

this density the effective potential is then constructed explicitly using the LDA functional for the exchange and correlation. Solving the KS equation to find the ground-state orbitals is accomplished variationally using a conjugate gradient method.⁴⁰ The nonlocal HF exchange energy and the Hartree energy are also initially calculated using the LDA orbitals.

The OEP has been implemented within the pseudopotential plane-wave code CASTEP.^{42,43} The orbitals, density, and potentials are represented on rectilinear grids in the usual manner of a plane-wave DFT implementation.⁴⁴ Kohn-Sham orbitals are described on reciprocal space grid points \mathbf{G} within a sphere bounded by the cutoff wave vector, G_{max} , and the density and potentials are nonzero within a sphere of radius $2G_{\text{max}}$. Therefore, the grids in both real and reciprocal space used to represent the density and potentials have twice the dimensions of the grid used for the orbitals. The Hylleraas minimization scheme is performed explicitly on these real space grids by direct variation, so that the effective basis used to represent $V^{\sigma}(\mathbf{r})$ is the set of grid points $\{\mathbf{G}\} : |\mathbf{G}| \leq 2G_{\text{max}}$.

Optimized norm-conserving pseudopotentials used in this work were generated using the OPIUM code⁴⁵ developed by Rappe *et al.* The Hartree-Fock approximation was used and the nonanalytic behavior of HF pseudopotentials was treated using the localization and optimization scheme of Al-Saidi *et al.*⁴⁶ Including exchange exactly in the pseudopotentials is required to ensure accurate core-valence interaction within the current formalism. Incorrect core-valence interaction has

been shown to have noticeable effects on calculated electronic structures.^{47,48} We found that calculations using pseudopotentials which include d states in the core typically predict band gaps 20–70% larger than if the d states are treated as valence. Therefore, semicore d states were treated as valence for As, Cd, Ga, Ge, In, Se, Te, and Zn.

The convergence characteristics of the conjugate gradient and Barzilai-Borwein variational minimizers are compared in Fig. 2, which plots the total OEP energy for diamond as a function of iteration number. An initial rapid decrease in the total energy is followed by a long tail of decreasing change in the total energy in both cases. The conjugate gradient energy decreases smoothly and monotonically towards the converged result, but the Barzilai-Borwein energy does not, exhibiting sharp drops and an occasional spiked increase (behavior which has been noted previously⁴⁹). After 100 or so steps the BB method still converges rapidly while the conjugate gradient method exhibits a very slow energy decrease. The difference in total energy per step is below $2.5 \mu\text{eV}$ per atom after 250 iterations of the BB minimizer and the solution is stationary after 300 iterations. For the conjugate gradients method, however, the energy difference per step does not drop below $5 \mu\text{eV}$ per atom even after 500 iterations. After 500 iterations the energy difference between the two methods is within $8 \mu\text{eV}$ per atom. In both cases we find that the calculated gradient tends to zero as the total energy converges, as expected in a variational method. The convergence rates of the Γ -point band

TABLE I. Energy gaps (in eV) for some semiconductors and insulators. Calculated values are shown for LDA, PBE, PBESOL, PW91, WC, Hartree-Fock, and the OEP method, using HF pseudopotentials.

	LDA	PBE	PBESOL	PW91	WC	HF	OEP	Experimental
Ge	0.09	0.13	0.04	0.14	0.04	6.01	0.86	0.79 ⁵⁰
InN	0.21	0.38	0.27	0.39	0.27	7.46	1.39	0.93 ⁵¹
Si	0.43	0.60	0.45	0.62	0.47	6.45	1.16	1.16 ⁵⁰
GaAs	0.99	1.08	0.97	1.09	0.96	7.32	1.86	1.52 ⁵⁰
CdTe	1.46	1.79	1.64	1.74	1.61	8.16	2.20	1.61 ⁵²
ZnSe	1.88	2.15	2.02	2.13	1.99	9.35	2.86	2.80 ⁵³
GaN	2.13	2.35	2.23	2.36	2.23	10.38	3.32	3.39 ⁵⁴
ZnO	1.93	2.28	2.09	2.29	2.10	11.51	3.48	3.43 ⁵⁵
C	3.98	4.21	4.03	4.24	4.08	12.76	4.87	5.47 ⁵¹
CaO	3.93	4.11	4.01	4.12	4.02	14.63	6.09	7.09 ⁵⁶
NaCl	4.84	5.34	5.13	5.35	5.19	13.74	6.27	8.97 ⁵⁷

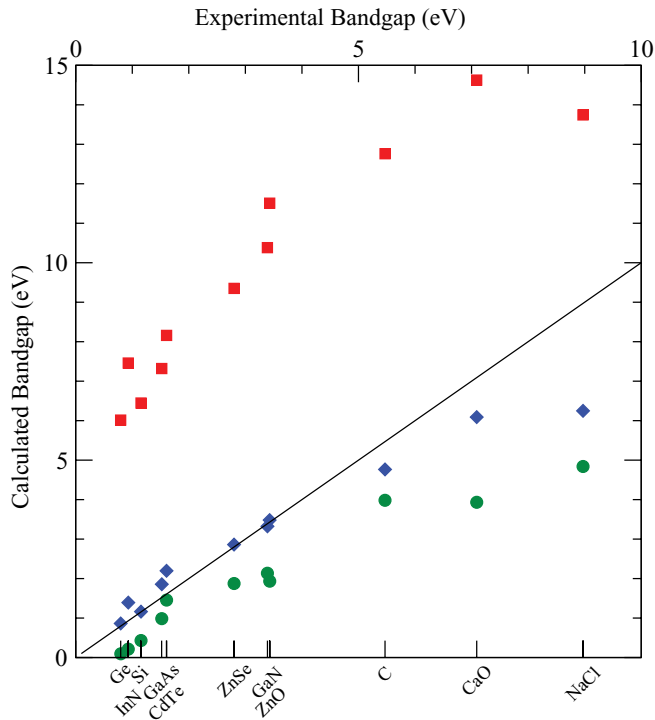


FIG. 5. (Color online) Comparison of calculated and experimental band gaps: red squares are HF, blue diamonds are OEP, and green circles are LDA.

gap of diamond behave in a similar fashion as shown in Fig. 3. After 250 iterations the gap is converged to within $25 \mu\text{eV}$ for the BB minimizer. By comparison, with the conjugate gradient minimizer the band gap is converged to $200 \mu\text{eV}$ in the same number of iterations. All calculations presented below using the OEP method were run for at least 250 iterations to ensure sufficient accuracy. These calculations take about the same CPU time as the commonly used hybrid XC functionals.

In the calculations that follow, the basis set size (plane-wave cutoff energy) and Brillouin-zone sampling were chosen so that total energy differences, evaluated using the LDA, were less than 1 meV/atom . The same settings were used for the OEP calculations, which resulted in a very similar convergence error with cutoff as the LDA case. However, the OEP calculations were more sensitive to Brillouin-zone sampling error; for example, in the case of diamond, a

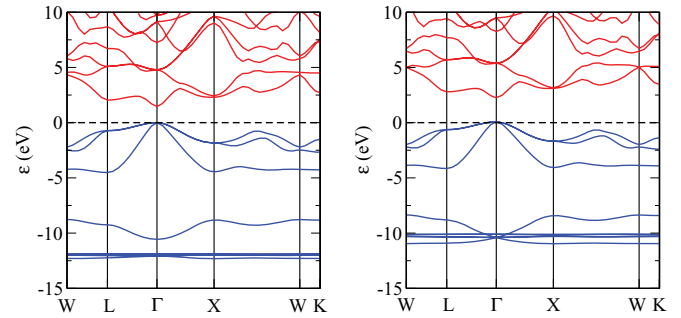


FIG. 7. (Color online) Band structure of CdTe (zinc-blende structure), calculated using the LDA (left) and the OEP (right).

$8 \times 8 \times 8$ Monkhorst-Pack grid⁵⁸ resulted in a 1 meV/atom error using the LDA but 3 meV for the OEP, rising to 9 meV for ZnO. Convergence testing was also performed on the fast Fourier-transform grid used to represent the potential, and the energy error was determined to be $\approx 10 \mu\text{eV}$.

IV. CALCULATED ELECTRONIC PROPERTIES

An example OEP is displayed in Fig. 4, in this case a slice through the primitive unit cell of Si in the diamond structure. The potential is smoothly varying outside the pseudopotential regions, which are the only regions where the potential is positive. The potential is at its deepest in the bonding region between the two ions and (outside the pseudopotential core region) is always negative but tending to zero in the low-density regions away from the ions.

The electronic structures of a selection of insulating and semiconducting materials were calculated using the LDA, some GGAs (PBE,⁵⁹ PBESOL,⁵⁹ PW91,⁶⁰ and WC⁶¹), HF, and the OEP. The resultant band gaps are shown in Table I and for the LDA, OEP, and Hartree-Fock calculations are also plotted in Fig. 5. The mean absolute error is 0.55 eV for the OEP, 1.39 eV for the LDA, and 1.19 eV for PBE, although these values are skewed by the underestimated values for the band gap found for diamond, NaCl, and CaO materials. The corresponding band structures for Ge, CdTe, InN, GaN, and ZnO are plotted in Figs. 6–10.

Figure 6 shows the band structure of Ge in the diamond structure calculated using the LDA and the OEP. For Ge the OEP improves the gap from the LDA value of 0.09 eV to 0.86 eV , which is very close to the experimental gap. Similarly for Si

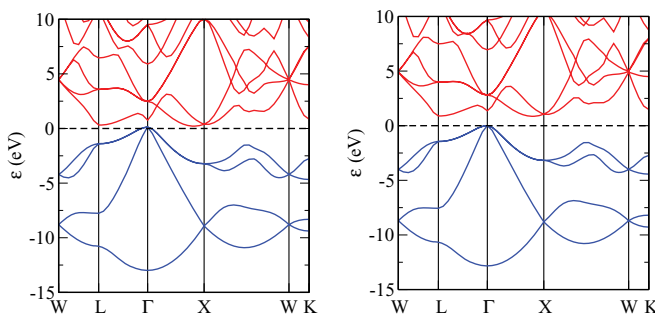


FIG. 6. (Color online) Band structure of Ge (diamond structure), calculated using the LDA (left) and the OEP (right).

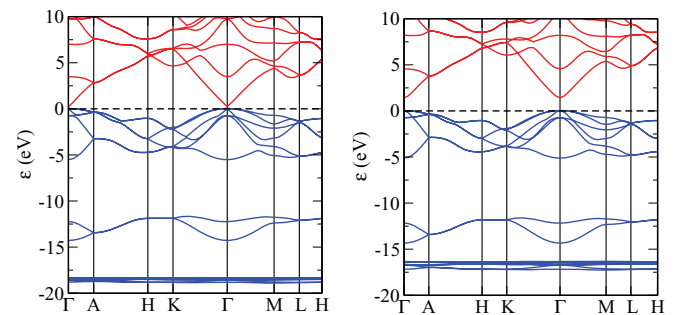


FIG. 8. (Color online) Band structure of InN (wurtzite structure) using the LDA (left) and the OEP (right).

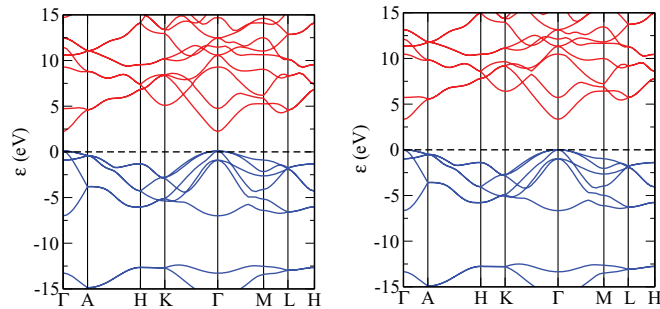


FIG. 9. (Color online) Band structure of GaN (wurtzite structure) using the LDA (left) and the OEP (right).

the gap is improved from 0.43 eV using the LDA to 1.16 eV using the OEP. For GaAs the OEP opens the gap from the LDA value of 0.99 to 1.86 eV. The OEP gap is greater than the experimental gap by 20%; a similar overestimation of the gap by the OEP in GaAs was also observed by Stadele *et al.*¹⁷ Figure 7 shows the band structures for CdTe calculated using the LDA and the OEP. The OEP gap is 35% larger than the experimental gap, opening from 1.46 eV for the LDA to 2.20 eV for the OEP.

Figure 8 shows the band structures for the wurtzite semiconductor InN calculated using the LDA and the OEP. The calculated band gap for InN has a small direct gap of 0.21 eV for the LDA which is opened to a gap of 1.39 eV for the OEP. Similarly the gap in GaAs and CdTe is overestimated when compared to the experimental gap. The band structures calculated for GaN in the wurtzite structure are shown in Fig. 9. The OEP band gap of 3.32 eV is very close to the experimental gap of 3.39 eV and greatly improved on the LDA band gap 2.13 eV. The wurtzite structure ZnO band structures calculated using the LDA and the OEP are shown in Fig. 10. The calculated band gap is greatly improved from 1.93 eV when using the LDA to 3.48 eV for the OEP, which is very close to the experimental gap of 3.43 eV.

For ZnSe in the zinc-blende structure the OEP calculated gap of 2.83 eV is remarkably close to the experimental gap and greatly improved over the LDA value of 1.88 eV. However, for diamond the OEP gap of 4.87 eV underestimates the experimental value by 12% but still improves on the LDA

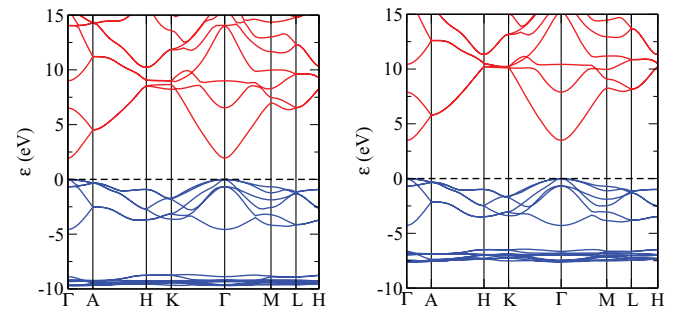


FIG. 10. (Color online) Band structure of ZnO (wurtzite structure) using the LDA (left) and the OEP (right).

value of 3.98 eV; the underestimation by the OEP was again noted by Stadele *et al.*¹⁷

For the wide-gap rocksalt structure insulators CaO and NaCl the OEP band gaps are opened up compared to the LDA band gaps, with the LDA calculated band gap being 3.93 eV for CaO and 4.84 eV for NaCl and the OEP calculated band gap being 6.09 eV for CaO and 6.32 eV for NaCl. For CaO the gap is underestimated by 14% for the OEP, and for NaCl the predicted gap is 30% smaller for the OEP than the experimental value. The OEP value for CaO is very similar to that found by Kotani and Akai⁶² and for NaCl is in line with that found by Li *et al.*⁶³

Table II compares the OEP band gaps obtained here to those obtained by others using a variety of basis sets. The band gaps calculated here are very similar to previous values and confirm the progressive underestimation of the band gap for wide-gap materials.

The OEP also improves values for semicore energy levels relative to the LDA, as shown in Table III. This is to be expected given that the self-interaction error present in the LDA has the largest effect on the tightly bound *d* states and that there is no self-interaction error within the OEP. In Ge the 3*d* electrons occupy states in the range -31.9 to -32.0 eV below the valence-band maximum with the experimental binding energy being determined as 29.1 to 29.6 eV. The LDA puts them in the range -35.0 to -35.2 eV. The experimental In 4*d* electron binding energy in InN is 17.4 eV; the LDA predicts between

TABLE II. Comparison of calculated OEP band gaps with previous work. The basis sets are plane-wave pseudopotential (PW PP), full potential linearized augmented plane wave (FPLAPW), linear muffin-tin orbitals with the atomic sphere approximation (LMTO-ASA), and the Korringa-Kohn-Rostoker method in the atomic sphere approximation (KKR-ASA).

	PP PW ^{17,18,64}	FPLAPW ^{23,63}	LMTO-ASA ^{21,62}	KKR-ASA ^{21,62}	This work	Experimental
Ge	0.94	0.89	1.57, 1.12	1.03	0.86	0.79
InN	1.40 ^a				1.39	0.93
Si	1.14, 1.23	1.30	1.93, 1.25	1.12	1.16	1.16
GaAs	1.78	1.74			1.86	1.52
GaN	2.76, ^a 3.46, ^a 3.49 ^a				3.32	3.39
ZnO	2.34 ^a				3.48	3.43
C	5.06, 4.90	5.20	5.12, 4.65	4.58	4.87	5.47
CaO			6.15	6.29	6.09	7.09
NaCl		6.3 ^b			6.27	8.97

^aThese results are calculated for the zinc-blende structure.

^bThis result is obtained with the self-interaction corrected functional of Perdew and Zunger¹¹ rather than the Hartree-Fock exchange functional.

TABLE III. Binding energies (eV) relative to the valence-band maxima for semicore electrons calculated using the LDA and OEP.

Material	State	LDA	OEP	Experimental
Ge	Ge $3d$	35.01–35.16	31.89–32.01	29.1–29.6 ⁶⁵
InN	In $4d$	16.34–17.22	18.38–18.94	17.4 ⁶⁶
GaAs	Ga $3d$	22.95–23.01	20.06–20.21	18.60–19.04 ⁶⁷
GaAs	As $3d$	47.45–47.46	44.45–44.49	40.37–41.07 ⁶⁷
CdTe	Cd $4d$	11.89–12.30	10.16–11.03	10.10 ⁶⁸
ZnSe	Zn $3d$	11.65–11.81	9.45–9.74	9.2 ⁶⁹
ZnSe	Se $3d$	61.02–61.03	58.50–58.52	55.5 ⁷⁰
GaN	Ga $3d$	21.23–21.55	18.30–18.85	17.94 ⁷¹
ZnO	Zn $3d$	8.88–9.69	6.65–7.61	7.4 ⁷²

–18.4 and –18.9 eV and the OEP predicts between –16.3 and –17.2 eV.

For GaAs the experimental Ga $3d$ states lie between 18.6 and 19.0 eV below the valence-band maximum; the OEP predicts a range of between –20.0 and –20.2 eV and the LDA predicts between –22.9 and –23.0 eV. Experiment gives As $3d$ states at 40.4 to 41.0 eV below the valence-band maximum; the OEP puts them at –44.5 eV and the LDA predicts –47.5 eV. In CdTe the Cd $4d$ electrons lie between –10.2 and –11.0 eV for the OEP, –11.9 and –12.3 eV for the LDA, compared with an experimental value of 10.10 eV below the valence-band maximum. The experimental results also place the Te $5s$ electrons very close to the Cd $4d$ states with a binding energy of 9.2 eV below the valence-band maximum, which the OEP also appears to predict. For ZnSe the Zn $3d$ electrons lie in the range –9.4 to –9.7 eV below the valence-band maximum for the OEP, –11.6 to –11.8 eV for the LDA, and 9.2 eV from experiment. For the Se $3d$ electrons the LDA gives energies of –61.02 compared with –58.5 eV with the OEP and 55.5 eV from experiment.

For GaN the LDA gives the Ga d electrons lying between –21.2 and –21.6 eV, the OEP between –18.3 and –18.9 eV, with an experimental value of 17.9 eV below the valence-band maximum. Using the OEP for ZnO gives Zn $3d$ states in the range –6.6 to –7.6 eV relative to the valence-band maximum

compared with an experimental value of 7.4 eV and a range of –8.9 to –9.7 eV with the LDA. As can be seen from the band structures for the above materials in the valence, the s and p states are almost identical when using the LDA and OEP. It is the d states which display the greatest difference.

The OEP also improves upon the predicted electronic structure given by the GGA methods of PBE, PBESOL, PW91, and WC, which underestimate the gaps of the materials investigated here, and Hartree-Fock, which greatly overestimates the gaps. Further work on magnetic metal oxides will be reported in a future publication.

V. CONCLUSIONS

A method of obtaining the OEP which treats the local exchange potential exactly without using a sum over all unoccupied states has been derived using the Hylleraas variational method and ideas borrowed from density-functional perturbation theory. This allows for the calculation of the OEP using a variational minimization scheme in real space. The electronic structures of well-known materials with a wide selection of band gaps have been calculated and the band gaps for semiconductors are found to be in good agreement with experimental values, although for the larger band gap materials—diamond, CaO, and NaCl—the calculated band gaps are still underestimated by 10–30%. Hartree-Fock pseudopotentials were found to give more accurate results than LDA pseudopotentials. The absence of self-interaction error within the OEP is manifest in a better description of semicore d states compared to the LDA. Their energies with respect to the top valence band are much closer to experimental spectroscopic measurements than within the LDA.

ACKNOWLEDGMENTS

T. W. H. acknowledges the Engineering and Physical Sciences Research Council for financial support, the UK national supercomputing facility (HECToR), and finally the United Kingdom Car-Parrinello Consortium for support under Grant No. EP/F037481/1.

*t.w.hollins@durham.ac.uk

†s.j.clark@durham.ac.uk; cmt.dur.ac.uk/sjc

‡keith.refson@stfc.ac.uk

§nikitas.gidopoulos@stfc.ac.uk

¹P. Hohnberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).

²W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965)

³J. Kohanoff and N. Gidopoulos, *Handbook Of Molecular Physics and Quantum Chemistry* (Wiley, New York, 2002).

⁴R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).

⁵C. Filippi, D. J. Singh, and C. J. Umrigar, *Phys. Rev. B* **50**, 14947 (1994).

⁶J. P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).

⁷L. J. Sham and M. Schlüter, *Phys. Rev. Lett.* **51**, 1888 (1983).

⁸R. T. Sharp and G. K. Horton, *Phys. Rev.* **90**, 317 (1953).

⁹J. D. Talman and W. F. Shadwick, *Phys. Rev. A* **14**, 36 (1976).

¹⁰S. Kümmel and L. Kronik, *Rev. of Modern Phys.* **80**, 3 (2008).

¹¹J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).

¹²T. Grabo, T. Kreibich, S. Kurth, and E. K. U. Gross, *Strong Coulomb Correlations in Electronic Structure: Beyond the Local Density Approximation* (Gordon and Breach, Tokyo, 1998).

¹³S. Kümmel and J. P. Perdew, *Phys. Rev. Lett.* **90**, 043004 (2003).

¹⁴F. Della Sala and A. Gorling, *Phys. Rev. Lett.* **89**, 033003 (2002).

¹⁵L. J. Sham and M. Schlüter, *Phys. Rev. B* **32**, 3883 (1985).

¹⁶R. W. Godby, M. Schlüter, and L. J. Sham, *Phys. Rev. Lett.* **56**, 2415 (1986).

¹⁷M. Stadele, J. A. Majewski, P. Vogl, and A. Gorling, *Phys. Rev. Lett.* **79**, 2089 (1997).

¹⁸M. Stadele, M. Moukara, J. A. Majewski, P. Vogl, and A. Gorling, *Phys. Rev. B* **59**, 10031 (1999).

¹⁹GGA's decay exponentially except for the B88 GGA functional, which decays as $1/r^2$.

²⁰T. Kotani, *Phys. Rev. B* **50**, 14816 (1994).

- ²¹T. Kotani, *Phys. Rev. Lett.* **74**, 2989 (1995).
- ²²T. Kotani and H. Akai, *Phys. Rev. B* **52**, 17153 (1995).
- ²³M. Betzinger, C. Friedrich, S. Blügel, and A. Gorling, *Phys. Rev. B* **83**, 045105 (2011).
- ²⁴D. M. Bylander and L. Kleinman, *Phys. Rev. Lett.* **74**, 3660 (1995).
- ²⁵D. M. Bylander and L. Kleinman, *Phys. Rev. B* **55**, 9432 (1997).
- ²⁶D. M. Bylander and L. Kleinman, *Phys. Rev. B* **52**, 14566 (1995).
- ²⁷R. J. Magyar, A. Fleszar, and E. K. U. Gross, *Phys. Rev. B* **69**, 045111 (2004).
- ²⁸J. B. Krieger, Y. Li, and G. J. Iafrate, *Phys. Rev. A* **46**, 5453 (1992).
- ²⁹T. Fukazawa and H. Akai, *J. Phys. Condens. Matter* **22**, 405501 (2010).
- ³⁰X. Gonze, *Phys. Rev. A* **52**, 1086 (1995); **52**, 1096 (1995).
- ³¹W. Yang and Q. Wu, *Phys. Rev. Lett.* **89**, 143002 (2002).
- ³²A. Gorling, A. Heßelmann, M. Jones, and M. Levy, *J. of Chem. Phys.* **128**, 104104 (2008).
- ³³N. I. Gidopoulos and N. N. Lathiotakis, *Phys. Rev. A* **85**, 052508 (2012).
- ³⁴S. Baroni, S. de Gironcoli, and A. Dal Corso, *Rev. Mod. Phys.* **73**, 515 (2001).
- ³⁵S. Kümmel and J. P. Perdew, *Phys. Rev. B* **68**, 035103 (2003).
- ³⁶E. A. Hylleraas, *Z. Phys.* **65**, 209 (1930).
- ³⁷R. M. Sternheimer, *Phys. Rev.* **96**, 951 (1954).
- ³⁸C. J. Pickard and F. Mauri, *Phys. Rev. B* **63**, 245101 (2001).
- ³⁹J. R. Yates, C. J. Pickard, and F. Mauri, *Phys. Rev. B* **76**, 024401 (2007).
- ⁴⁰M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1046 (1992).
- ⁴¹J. Barzilai and J. M. Borwein, *J. Num. Analysis* **8**, 141 (1988).
- ⁴²M. D. Segall, L. J. D. Lindan, M. I. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, *J. Phys.: Condens. Matter* **14**, 2717 (2002).
- ⁴³S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, *Z. Kristallogr.* **220**, 567 (2005).
- ⁴⁴R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- ⁴⁵A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, *Phys. Rev. B* **41**, 1227 (1990); For the code, examples, tutorials, and support, see [<http://opium.sourceforge.net/index.html>].
- ⁴⁶W. A. Al-Saidi, E. J. Walter, and A. M. Rappe, *Phys. Rev. B* **77**, 075112 (2008).
- ⁴⁷E. Engel, A. Höck, R. N. Schmid, R. M. Dreizler, and N. Chetty, *Phys. Rev. B* **64**, 125111 (2001).
- ⁴⁸E. Engel, *Phys. Rev. B* **80**, 161205(R) (2009).
- ⁴⁹R. Fletcher, Dundee Numerical Analysis Report NA/207, (2001).
- ⁵⁰*Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer-Verlag, Heidelberg, 1982).
- ⁵¹V. Y. Davydov, A. A. Klochikhin, R. P. Seisyan, V. V. Emtsev, S. V. Ivanov, F. Bechstedt, J. Furthmüller, H. Harima, A. V. Mudryi, J. Aderhold, O. Semchinova, and J. Graul, *Phys. Status Solidi B* **229**, 1R (2002).
- ⁵²G. Fonthal, L. Tirado-Mejia, J. I. Marin-Hurtado, H. Ariza-Calderon, and J. G. Mendoza, *J. Phys. Chem. Solids* **61**, 579 (2000).
- ⁵³D. J. Chadi, *Phys. Rev. Lett.* **72**, 534 (1994).
- ⁵⁴S. Strite and H. Morko, *J. Vac. Sci. Technol. B* **10**, 1237 (1992).
- ⁵⁵D. C. Look, *Mater. Sci. Eng. B* **80**, 383 (2001).
- ⁵⁶R. C. Whited, C. J. Flaten, and W. C. Walker, *Solid State Commun.* **13**, 1903 (1973).
- ⁵⁷S. Nakai and T. Sagawa, *J. Phys. Soc. Jpn.* **26**, 1427 (1969).
- ⁵⁸H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- ⁵⁹J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ⁶⁰J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**, 13244 (1992).
- ⁶¹Z. Wu and R. E. Cohen, *Phys. Rev. B* **73**, 235116-1 (2006).
- ⁶²T. Kotani and H. Akai, *Phys. Rev. B* **54**, 16502 (1996).
- ⁶³Y. Li, J. B. Krieger, M. R. Norman, and G. J. Iafrate, *Phys. Rev. B* **44**, 10437 (1991).
- ⁶⁴P. Rinke, *Phys. Status Solidi B* **245**, 929 (2008).
- ⁶⁵W. D. Grobman, D. E. Eastman, and J. L. Freeouf, *Phys. Rev. B* **12**, 4405 (1975).
- ⁶⁶L. F. J. Piper, T. D. Veal, P. H. Jefferson, C. F. McConville, F. Fuchs, J. Furthmüller, F. Bechstedt, H. Lu, and W. J. Schaff, *Phys. Rev. B* **72**, 245319 (2005).
- ⁶⁷E. A. Kraut, R. W. Grant, J. R. Waldrop, and S. P. Kowalczyk, *Phys. Rev. B* **28**, 1965 (1983).
- ⁶⁸A. Wall, Y. Gao, A. Raisanen, A. Franciosi, and J. R. Chelikowsky, *Phys. Rev. B* **43**, 4988 (1991).
- ⁶⁹L. Ley, R. A. Pollak, F. R. McFecly, S. I. Kowalczyk, and D. A. Shirley, *Phys. Rev. B* **9**, 600 (1974).
- ⁷⁰M. Cardona and L. Ley, *Photoemission in Solids: General Principles* (Springer-Verlag, New York, 1978).
- ⁷¹J. Hedman and N. Martensson, *Phys. Scr.* **22**, 176 (1980).
- ⁷²R. A. Powell, W. E. Spicer, and J. C. McMennamin, *Phys. Rev. B* **6**, 3056 (1972).