

Exact exchange Kohn-Sham formalism applied to semiconductors

M. Städele, M. Moukara, J. A. Majewski, and P. Vogl

Physik-Department und Walter Schottky Institut, Technische Universität München, Am Coulombwall, D-85748 Garching, Germany

A. Görling

Lehrstuhl für Theoretische Chemie, Technische Universität München, D-85748 Garching, Germany

(Received 28 August 1998)

We present a Kohn-Sham method that allows one to treat exchange interactions exactly within density-functional theory. The method is used to calculate lattice constants, cohesive energies, Kohn-Sham eigenvalues, dielectric functions, and effective masses of various zinc-blende semiconductors (Si, Ge, C, SiC, GaAs, AlAs, GaN, and AlN). The results are compared with values obtained within the local-density approximation, generalized gradient approximations, the Krieger-Li-Iafrate approximation for the Kohn-Sham exchange potential, and the Hartree-Fock method. We find that the exact exchange formalism, augmented by local density or generalized gradient correlations, yields both structural and optical properties in excellent agreement with experiment. Exact exchange-only calculations are found to lead to densities and energies that are close to Hartree-Fock values but to eigenvalue gaps that agree with experiment within 0.2 eV. The generalized gradient approximations for exchange yield energies that are much improved compared to local-density values. The exact exchange contribution to the discontinuity of the exchange-correlation potential is computed and discussed in the context of the band-gap problem. [S0163-1829(99)05515-0]

I. INTRODUCTION

Density-functional theory has proven to be an exceedingly powerful tool for predicting material properties quantitatively. It is centered around a variational principle that states that the ground-state energy $E[\rho]$ of a many-body system is a functional of the density ρ and becomes minimal for the ground-state density.¹ The best-established realization of density-functional theory is the Kohn-Sham (KS) method, that maps the many-body problem onto a set of Schrödinger-type KS single-particle equations that yield, in principle, the exact ground-state energy and density of the many-body system. The functional $E[\rho]$ contains a nonclassical contribution, namely the exchange-correlation energy $E_{xc}[\rho]$. Its functional derivative with respect to the electron density $\rho(\mathbf{r})$ is the KS exchange-correlation potential. It enters the KS equations as a multiplicative potential $V_{xc}([\rho];\mathbf{r})$.¹

The central challenge for any practical implementation of KS theory is to approximate $E_{xc}[\rho]$ and $V_{xc}[\rho]$ in a way that yields accurate results for atoms, molecules, and solids. The exchange-correlation energy is conventionally split into its exchange and correlation contributions $E_x[\rho]$ and $E_c[\rho]$, and consequently, $V_{xc}[\rho]$ may also be divided into the exchange and correlation parts $V_x[\rho]$ and $V_c[\rho]$, respectively. Up to now, neither of these potentials has been exactly determined for an extended system, in spite of the fact that the exchange energy is explicitly known in terms of the KS single-particle wave functions. However, it is yet unknown how to evaluate its functional derivative with respect to the density. In this paper, we present an exact expression for the local KS exchange potential $V_x([\rho];\mathbf{r}) = \delta E_x[\rho] / \delta \rho(\mathbf{r})$, and calculate it for several semiconductors that range from wide-gap to small-gap materials.

The most commonly employed approximations for $E_{xc}[\rho]$ and $V_{xc}[\rho]$ are the local-density approximation¹

(LDA) and several variants of the generalized gradient approximation¹⁻¹⁰ (GGA). Both types of schemes exhibit characteristic deficiencies when one applies these methods to real systems. Indeed, the LDA is known to overestimate the binding energy of molecules and solids considerably,^{1,11} whereas the GGA overestimates lattice constants and underestimates the bulk moduli of semiconductors.¹²⁻¹⁴ Both the GGA and LDA grossly fail to predict experimental band gaps of solids,¹⁵⁻²⁰ as well as the correct asymptotic behavior of $V_{xc}(\mathbf{r})$.¹

Several density-functional schemes beyond the LDA have been developed, such as self-interaction corrected methods,^{21,22} the weighted-density approximation,^{23,24} or the average Fock approximation.^{25,26} They attempt to improve both $V_x[\rho]$ and $V_c[\rho]$, albeit approximately. Alternatively, since $E_x[\rho]$ can be evaluated exactly and is typically an order of magnitude larger than $E_c[\rho]$, one might attempt to find a way to calculate $V_x[\rho]$ exactly. This is the concept that we pursue in the present paper. For free atoms, this goal was already achieved a long time ago by what is known as the optimized effective potential (OEP) method.²⁷ The OEP formalism was originally meant to approximate the Hartree-Fock method by constructing a local potential whose eigenfunctions minimize the expectation value of the Hartree-Fock Hamiltonian. Only later was it recognized to represent an exact exchange-only KS method within density-functional theory.²⁸ Unfortunately, the original OEP scheme is only applicable to spherical finite systems where the potential can be chosen to approach zero asymptotically.

Therefore, a straightforward generalization of the OEP scheme for solids is not possible. Nevertheless, several approximate approaches have been developed to extend this method to solids. One approach used a spherical shape approximation for the potential, and treated the ion cores nonrelativistically,²⁹⁻³³ and the other one approximated the

one-particle Green function by replacing eigenvalue differences by a constant.^{34,35} The latter approximation—known as the Krieger-Li-Iafrate (KLI) method—was shown to affect OEP results for atoms only negligibly.^{34,36,37} Importantly, both schemes yielded encouraging results for band structures in semiconductors and insulators.^{33,35,38–40}

We have developed a scheme—termed the exact exchange (EXX) method³¹—that yields an expression for the KS exchange potential that can be evaluated exactly, together with the exact exchange energy. This method requires approximations solely in the treatment of correlations and thus represents a systematic step beyond the LDA.

Effectively, we solve the OEP equations in a way that is free of shape approximations for the potential, eliminates all singularities, and is adapted to periodic boundary conditions. Furthermore, this scheme allows one to calculate the discontinuity Δ_x of $V_x(\mathbf{r})$ explicitly upon the addition of an electron.⁴¹ The present implementation of the EXX method relies on the use of consistently generated *ab initio* relativistic and norm-conserving EXX pseudopotentials.⁴² As we shall show below, such a treatment of the atomic cores is crucial for elements with intermediate and large cores.

This paper is organized as follows. In Sec. II, we outline the EXX method for solids. Section III addresses the EXX pseudopotentials, as well as computational issues and convergence checks. Section IV presents EXX results for cohesive properties, KS eigenvalues, effective masses, and optical response functions in zinc-blende semiconductors, and compares them with results obtained with approximate exchange methods such as LDA, GGA, KLI, and Hartree-Fock. Finally, the paper is summarized in Sec. V.

II. THEORY OF EXACT EXCHANGE

The EXX formalism allows one to calculate the local KS exchange potential $V_x(\mathbf{r})$ exactly. Before we present this scheme, we briefly summarize the relevant key points of the KS framework. Throughout, we focus on perfect bulk semiconductors and insulators. Since the present implementation of the EXX formalism is based on *ab initio* relativistic pseudopotentials, only the valence electrons are treated explicitly. We note, however, that this scheme is applicable to any type of electronic system,^{43,44} and can be easily generalized to the all-electron or spin-polarized case.

KS density-functional theory¹ maps an interacting electron system onto a noninteracting one (termed the KS system) with the same ground-state density $\rho(\mathbf{r})$. For systems with a nondegenerate ground state, the corresponding wave function is a Slater determinant built of KS orbitals which are obtained from the single-particle KS equations

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + V_{KS}(\mathbf{r}) \right] \varphi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} \varphi_{n\mathbf{k}}(\mathbf{r}), \quad (1)$$

with a local effective potential $V_{KS}(\mathbf{r})$. In a periodic solid, the KS orbitals $\varphi_{n\mathbf{k}}(\mathbf{r})$ are Bloch states with band index n , wave vector \mathbf{k} , and eigenenergy $\varepsilon_{n\mathbf{k}}$. The ground-state (valence) density $\rho(\mathbf{r})$ is given by

$$\rho(\mathbf{r}) = 2 \sum_{v\mathbf{k}} |\varphi_{v\mathbf{k}}(\mathbf{r})|^2. \quad (2)$$

Here the \mathbf{k} sum extends over all vectors in the first Brillouin zone, and the index v labels the occupied valence bands. Spin-orbit coupling is neglected in this paper, which gives a factor of 2 for spin degeneracy. The ground-state energy E may be partitioned in the following way:⁴⁵

$$E = T_s[\rho] + E'_{ei}[\rho] + E'_H[\rho] + E_x[\rho] + E_c[\rho] + E'_{ii}, \quad (3)$$

where $T_s[\rho]$, $E'_{ei}[\rho]$, $E'_H[\rho]$, $E_x[\rho]$, $E_c[\rho]$, and E'_{ii} denote the kinetic energy, the electron-ion interaction energy, the classical Hartree energy, the exchange energy, the correlation energy, and the ion-ion interaction energy of the KS system, respectively. The primes indicate that the divergent contributions—that exactly cancel in an infinite neutral system—have been removed. With this notation, the KS potential in Eq. (1) reads

$$\begin{aligned} V_{KS}(\mathbf{r}) &= \frac{\delta}{\delta \rho(\mathbf{r})} (E'_{ei}[\rho] + E'_H[\rho] + E_x[\rho] + E_c[\rho]) \\ &= V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_x(\mathbf{r}) + V_c(\mathbf{r}), \end{aligned} \quad (4)$$

where each potential term represents the functional derivative of the corresponding total energy contribution with respect to the density. The functional derivative of the electron-ion energy $\int d\mathbf{r} \rho(\mathbf{r}) V_{ext}(\mathbf{r})$ is taken for a fixed external potential V_{ext} . The exchange potential is given by¹

$$\begin{aligned} V_x(\mathbf{r}) &= \frac{\delta E_x[\rho]}{\delta \rho(\mathbf{r})}, \\ E_x[\rho] &= -\frac{e^2}{2} \sum_{vv'\mathbf{k}\mathbf{k}'} \iint d\mathbf{r} d\mathbf{r}' \\ &\quad \times \frac{\varphi_{v\mathbf{k}}^*(\mathbf{r}) \varphi_{v\mathbf{k}}(\mathbf{r}') \varphi_{v'\mathbf{k}'}(\mathbf{r}') \varphi_{v'\mathbf{k}'}^*(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}. \end{aligned} \quad (5)$$

Note that the only difference between the KS and the Hartree-Fock exchange energy is that the latter quantity is calculated using Hartree-Fock rather than KS orbitals.

In most practical calculations, both the exchange and correlation energy have been approximated so far. The most common approach is the local-density approximation,

$$E_\alpha^{LDA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_\alpha^{LDA}[\rho(\mathbf{r})] d\mathbf{r}, \quad (6)$$

where $\alpha = x$ or c and $\varepsilon_\alpha^{LDA}(\rho)$ is the exchange or correlation energy per particle in a homogeneous electron gas with constant density ρ . Alternatively, one employs the GGA scheme,

$$E_\alpha^{GGA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_\alpha^{GGA}[\rho(\mathbf{r}), \nabla \rho(\mathbf{r}), \dots] d\mathbf{r}, \quad (7)$$

where ε_α^{GGA} is constructed to obey various sum rules and other known exact properties of $E_{xc}[\rho]$.^{1–9,46}

The difficulty in calculating the exact $V_x(\mathbf{r})$ lies in the fact that $E_x[\rho]$ in Eq. (5) is explicitly known only in terms of

the KS orbitals but not in terms of the density $\rho(\mathbf{r})$. Unfortunately, the explicit dependence of the KS orbitals $\varphi_{n\mathbf{k}}[\rho]$ on the density is totally unknown. Therefore, $V_x(\mathbf{r}) = \delta E_x[\rho]/\delta\rho(\mathbf{r})$ cannot be calculated straightforwardly.

Nevertheless, this functional derivative can be determined explicitly by noting that it is sufficient to determine the in-

finitesimal (first-order) change in the exchange energy, $\delta E_x[\rho]$, with respect to an infinitesimal (charge-conserving) change $\delta\rho(\mathbf{r})$ in the density. The central idea of the EXX method⁴⁴ is to split the functional derivative in Eq. (5) into separate terms, each of which can be calculated exactly. By employing the chain rule, we can write

$$V_x(\mathbf{r}) = \frac{\delta E_x[\rho]}{\delta\rho(\mathbf{r})} = \sum_{v\mathbf{k}} \int d\mathbf{r}' \int d\mathbf{r}'' \left[\frac{\delta E_x[\rho]}{\delta\varphi_{v\mathbf{k}}(\mathbf{r}')} \frac{\delta\varphi_{v\mathbf{k}}(\mathbf{r}')}{\delta V_{KS}(\mathbf{r}'')} + \text{c.c.} \right] \frac{\delta V_{KS}(\mathbf{r}'')}{\delta\rho(\mathbf{r})}. \quad (8)$$

All three functional derivatives on the right-hand side of Eq. (8) are directly accessible, as we will outline below. The first one is obtained by straightforward differentiation of $E_x[\rho]$ in Eq. (5), and reads

$$\frac{\delta E_x[\rho]}{\delta\varphi_{v\mathbf{k}}(\mathbf{r}')} = -e^2 \sum_{v'\mathbf{k}'} \int d\mathbf{r}_1 \frac{\varphi_{v'\mathbf{k}'}^*(\mathbf{r}') \varphi_{v\mathbf{k}}^*(\mathbf{r}_1) \varphi_{v'\mathbf{k}'}(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}'|}. \quad (9)$$

The second one is the first-order change in the KS orbitals induced by an infinitesimal change in the KS potential. In analogy to first-order perturbation theory, one obtains

$$\frac{\delta\varphi_{v\mathbf{k}}(\mathbf{r}')}{\delta V_{KS}(\mathbf{r}'')} = \sum_{n'\mathbf{k}' \neq v\mathbf{k}} \varphi_{n'\mathbf{k}'}(\mathbf{r}') \frac{\varphi_{n'\mathbf{k}'}^*(\mathbf{r}'') \varphi_{v\mathbf{k}}(\mathbf{r}'')}{\varepsilon_{v\mathbf{k}} - \varepsilon_{n'\mathbf{k}'}}. \quad (10)$$

Insertion of Eqs. (9) and (10) into Eq. (8) leads to

$$V_x(\mathbf{r}) = \int d\mathbf{r}' \sum_{v\mathbf{k}} \left[\langle v\mathbf{k} | \hat{V}_x^{NL} | c\mathbf{k} \rangle \frac{\varphi_{c\mathbf{k}}^*(\mathbf{r}') \varphi_{v\mathbf{k}}(\mathbf{r}')}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} + \text{c.c.} \right] \frac{\delta V_{KS}(\mathbf{r}')}{\delta\rho(\mathbf{r})}, \quad (11)$$

where $\langle \mathbf{r} | c\mathbf{k} \rangle = \varphi_{c\mathbf{k}}(\mathbf{r})$ are the unoccupied conduction-band KS orbitals. Note that the nonlocal operator \hat{V}_x^{NL} has the same form as the Hartree-Fock exchange operator but is built of KS rather than Hartree-Fock orbitals. In real space it is an integral operator with the kernel

$$V_x^{NL}(\mathbf{r}, \mathbf{r}') = -e^2 \sum_{v\mathbf{q}} \frac{\varphi_{v\mathbf{q}}(\mathbf{r}) \varphi_{v\mathbf{q}}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (12)$$

and in reciprocal space it reads

$$V_x^{NL}(\mathbf{k}, \mathbf{G}, \mathbf{G}') = -\frac{4\pi e^2}{\Omega} \sum_{v\mathbf{q}\mathbf{G}_1} \frac{C_{v\mathbf{q}}(\mathbf{G} + \mathbf{G}_1) C_{v\mathbf{q}}^*(\mathbf{G}' + \mathbf{G}_1)}{|\mathbf{q} - \mathbf{k} + \mathbf{G}_1|^2}, \quad (13)$$

where \mathbf{G}, \mathbf{G}' , and \mathbf{G}_1 denote reciprocal-lattice vectors, $C_{v\mathbf{q}}(\mathbf{G}) = \langle \mathbf{q} + \mathbf{G} | v\mathbf{q} \rangle$ is the plane-wave expansion coefficient of the Bloch function $|v\mathbf{q}\rangle$ associated with \mathbf{G} , and Ω is the crystal volume. The third remaining functional derivative in Eq. (8) is the first-order change in the KS potential induced by a small change in the density. In order to calculate it, it is easier to consider the inverse functional derivative, since it can be obtained from linear response theory of independent particles and we determine this quantity first. One has

$$\frac{\delta\rho(\mathbf{r})}{\delta V_{KS}(\mathbf{r}')} = \chi_0(\mathbf{r}, \mathbf{r}'),$$

$$\chi_0(\mathbf{r}, \mathbf{r}') = 2 \sum_{v\mathbf{k}} \frac{\varphi_{v\mathbf{k}}^*(\mathbf{r}) \varphi_{c\mathbf{k}}(\mathbf{r}) \varphi_{c\mathbf{k}}^*(\mathbf{r}') \varphi_{v\mathbf{k}}(\mathbf{r}') + \text{c.c.}}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}}. \quad (14)$$

We require that a first-order change in the KS potential always induces a change in the density that is also of first order. In this case, the Hohenberg-Kohn theorem guarantees a one-to-one mapping between $\delta\rho(\mathbf{r})$ and $\delta V_{KS}(\mathbf{r})$.⁴⁴ However, this mapping excludes constant potential changes, since a rigid potential shift has no effect on the wave functions and the density. With Eq. (14), this implies that one has

$$\int d\mathbf{r}' \chi_0(\mathbf{r}, \mathbf{r}') = 0. \quad (15)$$

Thus the mapping χ_0 cannot be inverted, and we need to consider a restricted function space that excludes constant changes of the KS potential. This is most easily achieved through a Fourier expansion, which leads to⁴⁷

$$\begin{aligned} \delta\rho(\mathbf{G}) &= \sum_{\mathbf{G}'} \chi_0(\mathbf{G}, \mathbf{G}') \delta V_{KS}(\mathbf{G}'), \\ \chi_0(\mathbf{G}, \mathbf{G}') &= \frac{4}{\Omega} \sum_{v\mathbf{c}\mathbf{k}} \frac{\langle v\mathbf{k} | e^{-i\mathbf{G}\mathbf{r}} | c\mathbf{k} \rangle \langle c\mathbf{k} | e^{i\mathbf{G}'\mathbf{r}} | v\mathbf{k} \rangle}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}}. \end{aligned} \quad (16)$$

All charge-conserving changes in the density lie within the subspace $\mathbf{G} \neq \mathbf{0}$. In addition, the subspace $\mathbf{G}' \neq \mathbf{0}$ excludes all constant potential changes. Therefore, the submatrix of χ_0 with both \mathbf{G} and $\mathbf{G}' \neq \mathbf{0}$ —that we denote by $\tilde{\chi}_0$ —is regular, and can be inverted unambiguously. After a straightforward transformation of Eq. (11) to reciprocal-lattice space, we arrive at the following exact expression for the exchange potential in Eq. (5):

$$\begin{aligned} V_x(\mathbf{G}) &= \sum_{\mathbf{G}' \neq \mathbf{0}} [E(\mathbf{G}') + E^*(-\mathbf{G}')] \tilde{\chi}_0^{-1}(\mathbf{G}, \mathbf{G}'), \\ E(\mathbf{G}) &= \frac{2}{\Omega} \sum_{v\mathbf{c}\mathbf{k}} \frac{\langle v\mathbf{k} | \hat{V}_x^{NL} | c\mathbf{k} \rangle \langle c\mathbf{k} | e^{-i\mathbf{G}\mathbf{r}} | v\mathbf{k} \rangle}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}}. \end{aligned} \quad (17)$$

In a self-consistent Kohn-Sham calculation, both quantities $\tilde{\chi}_0(\mathbf{G}, \mathbf{G}')$ and $E(\mathbf{G})$ in Eq. (17) have to be updated in each step of the iteration cycle. For a system with inversion symmetry, one has $E^*(-\mathbf{G}) = E(\mathbf{G})$. The present formalism does not yield the spatial average of the exchange potential, i.e., its $\mathbf{G} = \mathbf{0}$ component. Indeed, exchange-correlation potentials are defined only up to an additive constant in the KS formalism with a fixed particle number. $V_x(\mathbf{G} = \mathbf{0})$ does not enter the total energy or eigenvalue differences, and will be set to zero in the remainder of the paper without loss of generality.

By construction, the exact exchange energy cancels the self-interaction contributions in the Hartree energy. Therefore, a calculation of the exact exchange potential Eq. (17) yields a realization of KS theory that is rigorously self-interaction free, in contrast to the LDA and GGA. In addition, expression Eq. (17) satisfies all scaling relations and asymptotic convergence laws that are known for the KS exchange potential.^{48,49} Even though Eqs. (11) and (17) show that the value of $V_x(\mathbf{r})$ at any point \mathbf{r} depends on the KS wave functions and eigenvalues in the whole unit cell, the locality of the exchange potential significantly simplifies band-structure calculations, since it is independent of the Bloch wave vector.

In atoms, the analog of the present EXX formalism is known as the optimized effective potential scheme, as we pointed out in Sec. I.²⁷ It is based on an integral equation that can be derived by acting with the operator $\delta\rho/\delta V_{KS}$ on both sides of Eq. (11). So far, this integral equation has only been solved for finite spherical systems, where $V_x(\mathbf{r})$ can be chosen to vanish asymptotically. This is related to the divergence of the full inverse χ_0^{-1} , as expressed by Eq. (15). By contrast, the present reciprocal-space formalism eliminates this divergence exactly.

We would like to emphasize that there is an important conceptual difference between the EXX exchange scheme and the Hartree-Fock scheme. In the Hartree-Fock formalism, the single-particle equations contain a *nonlocal* exchange potential, and they provide an approximate solution of the many-electron problem. By contrast, the EXX formalism leads to a *local* exchange potential and provides, in principle, an exact solution if it is augmented by the exact correlation potential. Finally, we point out that the procedure outlined in Eqs. (8)–(17) can be applied not only to the KS realization of density functional theory but also to generalized KS schemes,⁵⁰ the present procedure allows one to take the functional derivative of any orbital dependent expression with respect to the electron density.

III. COMPUTATIONAL DETAILS

A. Pseudopotentials

We have implemented the EXX scheme within the scalar relativistic pseudopotential framework. From a rigorous density-functional point of view, the external (pseudo)potential $V_{ext}(\mathbf{r})$ should be local. We have, therefore, employed local⁵¹ as well as semilocal pseudopotentials. For the latter, we have used the Troullier-Martins procedure⁵² to determine consistent EXX pseudopotentials⁴² based on atomic EXX calculations. For comparison, we have also employed Bachelet-Hamann-Schlüter LDA as well as KLI pseudopotentials.^{52,38,40}

All nonlocal pseudopotentials include angular momentum components V_l^{ion} up to $l=2$, and we have chosen the cutoff radii r_c to be independent of l . In the present calculations, the values $r_c^{Si, Ga, As} = 2.2$, $r_c^{Al, Ge} = 1.9$, $r_c^N = 1.7$, and $r_c^C = 1.5$ in units of Bohr radii were used. They are either equal to or smaller than previously applied values.⁵³

It has been noted before^{40,42} that EXX and KLI ionic pseudopotentials for atoms of valency Z converge rather slowly to $-Ze^2/r$ outside the core radius, as opposed to LDA pseudopotentials. A measure of this convergence is given by

$$\alpha = \int_b^\infty 4\pi r'^2 \left| V_l^{ion}(r') + \frac{Ze^2}{r'} \right| dr', \quad (18)$$

where the parameter b has to be chosen larger than r_c . We have chosen the pseudopotential cutoff radii r_c such that the values of α are kept as small as possible. The optimum values for r_c do not depend significantly on the parameter b , which therefore has been set to 2.8 Bohr radii in all cases. In addition, we have followed the procedure of Ref. 39 and cut off the tails beyond a certain ionic radius r_0 that was set to approximately half the bond length in the corresponding el-

emental materials. In units of Bohr radii, we have chosen $r_0=2.5$ for Si, Ge, Ga, and As, $r_0=2.2$ for Al, and $r_0=2.0$ for C and N, respectively. Fortunately, the calculated solid-state results have been found to be insensitive to this cutoff: a variation of r_0 by ± 0.3 Bohr radii changes lattice constants typically by $\leq 0.01 \text{ \AA}$, cohesive energies by $\leq 0.2 \text{ eV/atom}$, and band gaps by $\leq 0.05 \text{ eV}$.

B. Convergence of solid-state properties

The nonlocal exchange integrals in $E_x[\rho]$ and $V_x(\mathbf{r})$ in Eqs. (5) and (17) contain an integrable Coulomb singularity that we have taken into account analytically, following Ref. 54. This greatly reduces the number of wave vectors \mathbf{k} required for Brillouin-zone integrations. We have performed calculations with up to 28 special \mathbf{k} points for the eight semiconductors studied in this paper. For these materials, however, ten \mathbf{k} points already guarantee convergence of exchange energies to $\leq 0.05 \text{ eV/atom}$ (i.e., within 0.1%) and band gaps to 0.02 eV. Therefore, we have used ten \mathbf{k} points throughout.

The kinetic-energy cutoffs used for the presently studied materials were chosen to be 25 Ry (Si, Ge, GaAs, AlAs), 45 Ry (GaN, AlN), 50 Ry (SiC), and 65 Ry (C), respectively. We have carefully checked that these values suffice to yield a plane-wave convergence error smaller than $\leq 0.01 \text{ eV}$ for band gaps, and of $\leq 0.01 \text{ eV/atom}$ for exchange and total energies. In Eqs. (16) and (17), all unoccupied states (between 570 and 800) included in the Hamiltonian have been taken into account. The dimension of the matrix $\tilde{\chi}_0$ has been set to typically half the value for the Hamiltonian matrix. This gives plane-wave cutoffs for $\tilde{\chi}_0$ of 14 Ry (Si, Ge), 16 Ry (GaAs, AlAs), and 34 Ry (C, SiC, GaN, AlN), respectively, and guarantees convergence of *all* matrix elements $\tilde{\chi}_0(\mathbf{G}, \mathbf{G}')$.⁴⁷ We note that the corresponding cutoffs used in standard *GW* calculations^{55,56} are usually considerably smaller.

We have estimated the scaling of the method with respect to the number of basis functions and atoms in the unit cell. With fast Fourier transform or real-space techniques, an EXX calculation roughly scales as $N_v^2 N_c N_k^2 N$ where N_v, N_c, N_k , and N are the numbers of valence bands, conduction bands, special \mathbf{k} points, and basis functions, respectively. An EXX calculation for systems with large unit cells is comparable to a *GW* (Ref. 56) calculation. As noted above, the locality of the exchange potential in the EXX method renders the calculation of a band structure very efficient.

IV. EXX RESULTS AND DISCUSSION

As shown in Sec. III, it is possible to determine the exact KS exchange potential explicitly. This is not the case for the correlation potential. In order to obtain realistic predictions, however, correlations need to be included in a self-consistent total-energy calculation. We have, therefore, performed calculations both with and without several model correlation functionals and will adopt the following notation. We will term a calculation ‘‘EXX’’ if it is carried out self-consistently with the exact exchange potential [Eq. (17)] and exchange energy [Eq. (5)] and with the standard LDA corre-

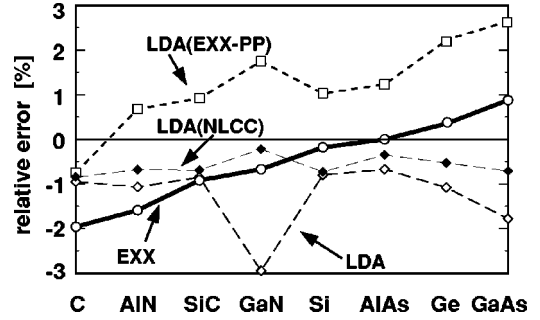


FIG. 1. Relative deviations (given in percent) of self-consistently calculated LDA and EXX lattice constants of various semiconductors from the experimental values that are taken from Refs. 73, 87, and 93. The results denoted by LDA(NLCC) and LDA(EXX-PP) were obtained with LDA pseudopotentials that include nonlinear core corrections (see Ref. 58) and EXX pseudopotentials, respectively.

lation potential.²² If correlations are omitted in the self-consistent ground-state calculation, on the other hand, we term the method ‘‘EXX(X).’’ In all other cases, the correlation functional is specified explicitly such as ‘‘EXX(GGA)’’ if a GGA correlation functional⁵⁷ is employed. Often, we compare to self-consistent LDA calculations that use both the standard LDA exchange and correlation functional; those calculations are denoted by ‘‘LDA.’’ Except when noted otherwise, we have always used *consistent* pseudopotentials in the sense that they have been constructed with the same Hamiltonian in the atom as is employed in the solid.

A. Lattice constants and bulk moduli

In Fig. 1, we compare the lattice constants as obtained by the LDA and EXX methods with the experimental values. As is well known, self-consistent LDA pseudopotential calculations yield lattice constants in very good agreement with experiment, provided the so-called nonlinear core corrections⁵⁸ in the exchange and correlation potentials are included. Otherwise, LDA calculations underestimate lattice constants by 1–3%, as shown in the figure.

The lattice constants predicted by the EXX method generally agree as well with experiment as the LDA ones, the average deviation for the eight semiconductors in Fig. 1 being 0.8% and 0.6%, respectively. This good agreement is obtained in spite of the fact that the EXX calculations do not include a nonlinear core correction.⁵⁸ Since this correction vanishes for the Hartree-Fock method, where the potential can be rigorously divided into a core and valence contribution, we expect this correction to be small for EXX as well.

In order to study separately the effect of the exact exchange functional on the core and valence electrons, respectively, we have performed the LDA solid-state calculations using either LDA (labeled LDA in Fig. 1) or EXX pseudopotentials [labeled LDA(EXX-PP)]. In the latter case, we find an increase and overall overestimate of the lattice constants compared to the former approach. This can be physically understood as follows. Since the EXX scheme is self-interaction free, an atomic EXX calculation causes the core electrons to screen the nucleus more efficiently than in the LDA. Accordingly, EXX pseudopotentials are less attractive than LDA pseudopotentials and therefore yield larger lattice

TABLE I. Theoretical LDA and EXX lattice constants (in Å) and bulk moduli (in Mbar) of Si, Ge, and GaAs, compared with experimental data (taken from Ref. 73 except where noted otherwise). The LDA values include nonlinear core corrections according to Ref. 58.

	LDA		EXX		Expt.	
	a_0	B_0	a_0	B_0	a_0	B_0
Si	5.39	0.96	5.42	1.15	5.43	0.99
GaAs	5.61	0.74	5.70	0.85	5.65	0.77
GaN	4.49	1.93	4.47	2.39	4.50 ^a	

^aReference 87.

constants. Clearly, this effect is more pronounced for atoms with a larger number of core electrons (which increases in Fig. 1 from left to right).

If one uses the EXX method for the valence electrons as well, on the other hand, the absence of the unphysical self-repulsion causes the valence charge density to shrink relative to the LDA valence density, and to counteract the effect from the core electrons. For small core ions, the valence effect dominates, whereas, for the large core ions, the core effect is more pronounced. This explains the trends in the EXX versus LDA lattice constants in Fig. 1. We note that similar trends have been obtained in LDA calculations with self-interaction-corrected pseudopotentials.^{59,60}

The calculated EXX bulk moduli are summarized in Table I for a few selected materials, and compared to the LDA as well as to experiment. The EXX method is seen to overestimate experimental bulk moduli by about 20%. This originates mostly in the valence electrons, since we find LDA calculations with EXX pseudopotentials to underestimate the experimental bulk moduli by 5–10%. It has been noticed before⁶¹ that valence-electron correlation effects tend to reduce Hartree-Fock bulk moduli by typically 20%. Thus a treatment of correlations beyond the LDA might be needed for obtaining more accurate bulk moduli in an exact exchange calculation.

B. Cohesive energies

Figure 2 compares the EXX and EXX(GGA) cohesive energies with results from LDA and Hartree-Fock calculations and with experimental data. Whereas the LDA results exhibit the well-known *overbinding* effect, and overestimate the experimental data by 0.6–2.1 eV/atom, the EXX(GGA) cohesive energies are seen to agree excellently with experiment.

To assess the origin of this good agreement, we have computed cohesive energies with related methods for comparison. We find that EXX(X), i.e., the EXX method without correlations, leads to cohesive energies that are virtually identical to Hartree-Fock values (cf. Fig. 2). Once one adds either LDA correlations or GGA correlations, the cohesive energies systematically increase and become closer to the experimental values. With LDA correlations, cohesive energies are still underestimated by 0.5–0.8 eV/atom, whereas GGA correlations produce those excellent EXX(GGA) results shown in Fig. 2. Similar findings have been reported by adding a GGA correlation functional to the Fock operator.⁶²

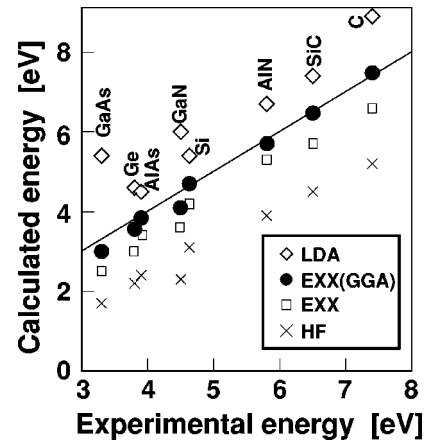


FIG. 2. Comparison of self-consistently calculated LDA, EXX, EXX(GGA) (i.e., the EXX method including GGA correlations from Ref. 2), and Hartree-Fock cohesive energies (given in eV/atom) of various semiconductors with experimental data. The latter are taken from Refs. 73, 94, and 95. The solid line corresponds to perfect agreement between theory and experiment, and is only drawn to guide the eye.

C. KS eigenvalues and related properties

Recently, significant efforts have been made to develop self-consistent *ab initio* density-functional methods that simultaneously predict accurately cohesive properties and energy gaps.^{26,30,35,39,40} We are now going to show below that the EXX method indeed predicts band gaps very accurately, in addition to cohesive properties.

1. EXX band structures

Figure 3 compares the calculated fundamental LDA and EXX band gaps for eight semiconductors with experimental data. Several additional energy gaps and total valence-bandwidths are listed in Table II. As is evident from these results, the EXX formalism leads to band gaps that are in excellent agreement with experiment, the deviation being less than 0.1–0.3 eV. By contrast, the corresponding LDA values are too small by 1–1.5 eV. Noticeably, EXX gaps are much smaller than the corresponding Hartree-Fock values, which are known to overestimate semiconductor band gaps dramatically. For example, we have computed the minimal Hartree-Fock energy gaps of Si, Ge, and GaAs to be 7.4, 6.4, and 7.7 eV, respectively, which is in accord with previously calculated results.⁶³

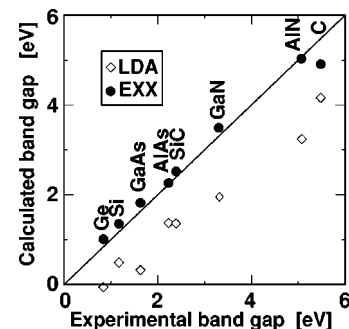


FIG. 3. Comparison of self-consistently calculated LDA and EXX band gaps (in eV) of various semiconductors with experimental data from Refs. 73 and 89–91.

TABLE II. Comparison of theoretical LDA and EXX Kohn-Sham eigenvalue gaps (in eV) and total valence band widths (VBW, in eV) with experimental data, taken from Ref. 73 except where otherwise noted. The gaps have been calculated at the experimental lattice constants and are given relative to the top of the valence band. Si* refers to a calculation with a local pseudopotential. The experimental data have been corrected for spin-orbit splitting effects since the latter are not included in the calculations.

	LDA				EXX				Expt.			
	L	Γ	X	VBW	L	Γ	X	VBW	L	Γ	X	VBW
Si*	1.54	2.79	0.52	12.77	2.36	3.46	1.43	12.34	2.4	3.34	1.25 ^a	12.5
Si	1.43	2.56	0.64	11.95	2.35	3.26	1.50	11.58				
Ge	0.13	-0.09	0.75	12.86	1.01	1.28	1.34	12.48	0.84	1.00	1.3	12.6
C	8.43	5.56	4.78	21.29	9.18	6.28	5.43	21.52		7.3 ^b		24.2 ^c
GaAs	0.89	0.32	1.41	12.78	1.93	1.82	2.15	12.33	1.85	1.63	2.18	13.1
AlAs	2.03	1.84	1.37	11.89	2.99	3.20	2.26	11.53	2.54	3.11	2.24	
SiC	5.36	6.33	1.36	15.32	6.30	7.37	2.52	15.23		7.59 ^c	2.39	
GaN	4.58	1.90	3.38	15.75	6.23	3.49	4.95	15.64		3.30 ^d		
AlN	7.16	4.20	3.24	14.74	8.58	5.66	5.03	14.86				

^aReference 88.

^bReference 89.

^cReference 90.

^dReference 91.

Correlations seem to have a weak influence on the energy gaps, as far as we can judge from the available correlation functionals. Within an exchange-only EXX(X) calculation, we find the energy gaps to decrease by typically $\lesssim 0.1$ eV relative to the EXX results. For example, the minimal gaps within EXX(X) are 1.23, 0.94, and 1.78 eV for Si, Ge, and GaAs, respectively. By using GGA rather than LDA correlations in the solid-state calculation, the energy gaps decrease slightly by typically 0.2 eV and amount to 0.97, 0.72, and 1.55 eV for Si, Ge, and GaAs, respectively.

The physical origin of the excellent agreement between EXX energy gaps and experiment lies in the absence of self-interaction and the locality of the exact exchange potential. Since the occupied valence states do not “feel” a self-repulsion in the EXX method, they become more localized and are energetically lowered relative to the LDA. The Hartree-Fock approach is self-interaction free for the occupied states but not for the unoccupied states. The nonlocality of the Fock exchange potential causes the unoccupied states to effectively “see” a different potential (N electrons) than the valence states ($N-1$ electrons) which gives rise to the huge energy gap in Hartree-Fock. By contrast, the EXX exchange potential is local and state independent, and therefore the *same* for occupied as well as for unoccupied states.

The EXX approach not only shifts all band gaps upwards in energy, but additionally yields the correct ordering of the conduction-band minima for all studied semiconductors. Thus it faithfully reproduces the known \mathbf{k} dependence of the band structure across the whole Brillouin zone. For Si and Ge, the dispersion relations are depicted in Fig. 4. In the case of Ge, the differences between LDA and EXX gaps are particularly strongly \mathbf{k} dependent. Whereas the LDA incorrectly predicts a direct and negative energy gap at Γ , the EXX method correctly yields the conduction-band minimum at L . The indirect nature of the energy gap of Ge originates predominantly in the Ge core. Indeed, by performing a LDA calculation for Ge bulk but employing EXX pseudopotentials, we already find an indirect band gap at L that is lower than the direct one at Γ by 0.18 eV (the experimental value is

0.16 eV). Indeed, the importance of the core for reproducing the indirect band gap in Ge has been pointed out before.^{59,64}

These results lead to the question of whether good band gaps may be obtained by carrying out standard LDA calculations for the valence electrons and by employing pseudopotentials that are constructed with the EXX scheme. We have performed LDA calculations for several semiconductors using both LDA and EXX pseudopotentials. The resulting band gaps are summarized in Table III. As one can deduce from these results, especially the direct band gaps at Γ of GaAs and Ge are increased significantly relative to a strict LDA calculation.

Finally, we compare the theoretical LDA and EXX valence-band-widths with the experimental values in Table II. We find the EXX method to predict generally smaller bandwidths than the LDA, irrespective of the type of pseudopotential used. This may be explained by the enhanced electron localization induced by the absence of self-repulsion in the EXX method. The only exceptions are the small core

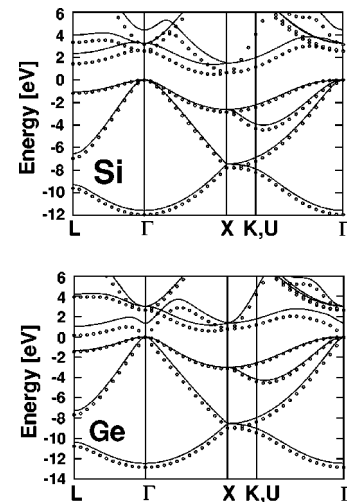


FIG. 4. Band structures for Si and Ge, as calculated within the LDA (open circles) and EXX (solid lines) methods, respectively.

TABLE III. Calculated LDA band gaps (in eV) of Si, Ge, and GaAs at L , Γ , and X , calculated with EXX pseudopotentials. The numbers in parentheses are the LDA band gaps obtained with LDA pseudopotentials.

	L	Γ	X
Si	1.49 (1.43)	2.53 (2.56)	0.49 (0.64)
Ge	0.24 (0.13)	0.42 (-0.09)	0.50 (0.75)
GaAs	1.11 (0.89)	0.86 (0.32)	1.23 (1.41)

compounds C and AlN. The results in Table II indicate that the EXX method slightly overestimates the localization of valence states, since EXX bandwidths are generally a few percent smaller than the experimental values.

2. EXX and the band-gap problem

The exact (but unknown) KS eigenvalue gap ε_{gap}^{KS} of a semiconductor differs from the true band gap E_{gap} . The latter may be defined as the ground-state energy difference between the N and $N \pm 1$ particle system,^{18,19,65}

$$\begin{aligned} E_{gap} &= E(N+1) + E(N-1) - 2E(N) \\ &= \varepsilon_{gap}^{KS} + \Delta_{xc} = \varepsilon_{gap}^{EXX(X)} + \varepsilon_{gap,c}^{KS} + \Delta_{xc}. \end{aligned} \quad (19)$$

The difference between the true gap and the eigenvalue gap originates in the fact that the exchange-correlation potential changes discontinuously by an amount Δ_{xc} upon adding one electron to the N -electron ground state of the semiconductor. In spite of several intriguing model calculations, the question whether this discontinuity Δ_{xc} is small or large compared to E_{gap} is still open.^{56,66-68,20} In Eq. (19), we have split up ε_{gap}^{KS} into two contributions, namely, the eigenvalue gap resulting from a self-consistent exchange-only scheme [EXX(X)] and a term that takes into account the remaining correlation-related contributions.

The present results shed some new light on the band-gap problem. First of all, the good agreement between the exchange-only EXX(X) eigenvalue gaps and experiment, $E_{gap} \approx \varepsilon_{gap}^{EXX(X)}$, yields

$$\Delta_{xc} \approx -\varepsilon_{gap,c}^{KS}. \quad (20)$$

With LDA and GGA correlations, one can evaluate the right-hand side of this equation and finds rather small values of $\Delta_{xc} \sim -0.1$ (LDA) and $\sim +0.2$ eV (GGA), respectively.

Recently, it has been shown that the exchange part Δ_x of the discontinuity $\Delta_{xc} = \Delta_x + \Delta_c$ can be evaluated exactly in terms of the exact exchange potential.^{41,69} It is given by

$$\Delta_x = \langle c\mathbf{k} | \hat{V}_x^{NL} - V_x | c\mathbf{k} \rangle - \langle v\mathbf{k}' | \hat{V}_x^{NL} - V_x | v\mathbf{k}' \rangle, \quad (21)$$

where $|c\mathbf{k}\rangle$ and $|v\mathbf{k}'\rangle$ denote the lowest conduction-band state and highest valence-band state, respectively, and V_x is the exact local exchange potential from Eq. (11). For Si, Ge, GaAs, we find Δ_x to be very large, namely, 5.48, 4.81, and 5.28 eV, respectively.⁴¹ This finding indicates a significant cancellation between the correlation part Δ_c and the exchange contribution Δ_x of the discontinuity. Thus, both Δ_x and the correlation contribution Δ_c are individually large compared to the true band gap.

TABLE IV. Effective electron masses of various zinc-blende semiconductors (in units of the free-electron mass) in the lowest conduction-band minima at L , Γ , or X , or along Δ , respectively. For the indirect semiconductors, the longitudinal (m_l) and transverse (m_t) masses are given. Experimental data are taken from Ref. 73 except for GaN; the latter value is from Ref. 92.

		LDA	EXX	Expt.
Si	m_l^Δ	0.95	0.97	0.92
	m_t^Δ	0.19	0.22	0.19
Ge	m_l^L	1.71	1.70	1.57
	m_t^L	0.07	0.10	0.08
C	m_l^Δ	1.68	1.59	1.4
	m_t^Δ	0.29	0.29	0.36
GaAs	m_Γ	0.02	0.10	0.07
AlAs	m_l^X	0.84	0.95	1.0
	m_t^X	0.24	0.27	0.25
GaN	m_Γ	0.17	0.26	0.20
AlN	m_Γ	0.30	0.36	
SiC	m_l^X	0.68	0.67	0.68
	m_t^X	0.23	0.26	0.25

3. Effective masses

Since the KS potential in the EXX approach is local, all momentum states see the same potential. This greatly simplifies band-structure calculations relative to Hartree-Fock type or GW -type methods.^{56,63} Therefore, EXX allows one to compute effective band masses as easily as the LDA does. It is well known that LDA effective masses agree rather poorly with experimental data in some materials such as GaAs.^{70,71} Table IV summarizes the effective conduction-band masses for several semiconductors. Generally, EXX electron masses are found to be almost equal to or larger than the corresponding LDA masses and show systematically good agreement with experiment in cases where the LDA fails markedly. This is in qualitative accord with standard $\mathbf{k}\cdot\mathbf{p}$ theory, which predicts band masses to change roughly in proportion to energy gaps.

We have also investigated the Luttinger parameters⁷² L , M , and N that characterize the warped structure of the three topmost valence bands near Γ . In this paper, these parameters were determined numerically by fitting the three-band $\mathbf{k}\cdot\mathbf{p}$ band structure⁷² to the LDA and EXX bands. In units of $\hbar^2/2m$, the triplet $(-L, -M, -N)$ for GaAs is given by (56.56, 4.01, 57.71) within the LDA and by (11.20, 3.48, 11.97) within the EXX method. In this case, the EXX method agrees substantially better with the experimental values⁷³ of (15.49, 3.94, 16.09) than the LDA does. In some other materials, however, the EXX method tends to predict hole masses that are slightly too heavy which is consistent with the trends in the valence band widths in Table II and Fig. 4. In Si, for example, we find $(-L, -M, -N)$ to be given by (7.15, 4.75, 9.17), (5.64, 3.89, 6.86), and (6.64, 4.60, 8.68), within the LDA, the EXX method, and experiment, respectively.

4. Dielectric functions

In this section we will show that the EXX approach not only yields excellent band gaps but also accurately predicts

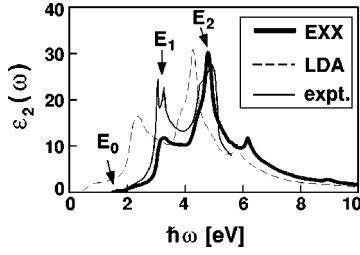


FIG. 5. Comparison of the calculated EXX and LDA imaginary part of the GaAs dielectric function as a function of energy (in eV) with experimental data from Ref. 96. Note that the present calculations do not include spin-orbit-splitting effects that lead to the splitting in the experimental E_1 peak.

optical properties such as dielectric functions and reflectivities. Figure 5 depicts a representative example, namely, the imaginary part of the dielectric function of GaAs, which is given by

$$\varepsilon_2(\omega) = \frac{4\pi^2 e^2 \hbar^2}{3m^2 \Omega \omega^2} \sum_{vc\mathbf{k}} |\langle v\mathbf{k} | \mathbf{p} | c\mathbf{k} \rangle|^2 \delta(\hbar\omega - \varepsilon_{c\mathbf{k}} - \varepsilon_{v\mathbf{k}}), \quad (22)$$

where \mathbf{p} is the momentum operator. We have neglected exchange-correlation, local-field and nonlocal pseudopotential effects in this expression since they are known to have only a minor effect on the peak positions.^{74,75} For the \mathbf{k} summation in Eq. (22), 328 \mathbf{k} points in the irreducible wedge of the Brillouin zone were taken into account. It is apparent from Fig. 5 that the position of the EXX calculated absorption edge E_0 as well as of the two prominent peaks labeled E_1 and E_2 are much closer to the measured data than the LDA. As is well known, a \mathbf{k} -independent scissors operator correction to the LDA band structure cannot achieve such an improvement.⁷⁶ In accord with other calculations,^{74,75,77} the height of the E_1 peak still does not fit the experimental one very well, which may be caused by the neglect of excitonic, many-body, and surface effects.

D. Comparison with approximate exchange functionals

In this section, we will assess the accuracy of various approximate exchange-related KS functionals by comparing their values for different densities to the values of the exact EXX functionals. To be consistent, all calculations in this section have been carried out by setting all correlation functionals equal to zero. This also applies to the LDA and GGA calculations. We have employed exact exchange-only [EXX(X)] semilocal, norm-conserving pseudopotentials throughout,⁴² except for Si, where a local pseudopotential⁵¹ has been used. As noted above, the latter serves as a test of the EXX scheme with a strictly local external potential which is a formal requirement of KS theory. All GGA results are based on the functional proposed by Becke;⁴ other published functionals^{2,3} yield very similar results.

In contrast to the LDA, the exact exchange potential Eq. (11) is not a simple function of the local density $\rho(\mathbf{r})$. Consequently, there is no simple local relation between the values of the density and the exchange potential $V_x(\mathbf{r})$ at any given point in the unit cell. Figure 6 shows that, for given values of the density [i.e., the self-consistent EXX(X) den-

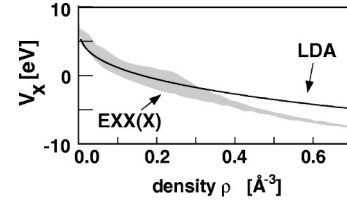


FIG. 6. Values of the local EXX(X) exchange potential $V_x(\mathbf{r})$ in Si, plotted vs the corresponding values of the density $\rho(\mathbf{r})$ for a set of positions \mathbf{r} that sample the unit cell. Since there is a simple local relation between V_x and ρ in the LDA, this relation leads to a simple curve for the LDA which is depicted for comparison. All potentials are given in eV. The spatial averages of the EXX(X) and LDA exchange potential have been set to zero.

sity], the corresponding values of the exchange potential in the unit cell may differ by as much as 4 eV in Si. Interestingly, the LDA mimics the average density dependence of $V_x(\mathbf{r})$ quite well. Both the exact and LDA potentials become more attractive when the density increases, but this trend is less pronounced for the LDA potential. Qualitatively similar findings hold for all eight semiconductors studied in this paper.

The spatial variation of the exchange potentials along the [111] direction in Si and GaAs is shown in Fig. 7. As one can see, LDA or GGA potentials show characteristic deviations from the exact one. All depicted valence exchange potentials are attractive in the bond regions and repulsive in the interstitial and core regions. The LDA exchange potential is much smoother than the exact one, which is a consequence of the self-repulsion that plagues this approximation. The GGA potential is significantly more accurate in the bond region but exhibits artificial peaks in regions of low density (see also Refs. 78 and 79).

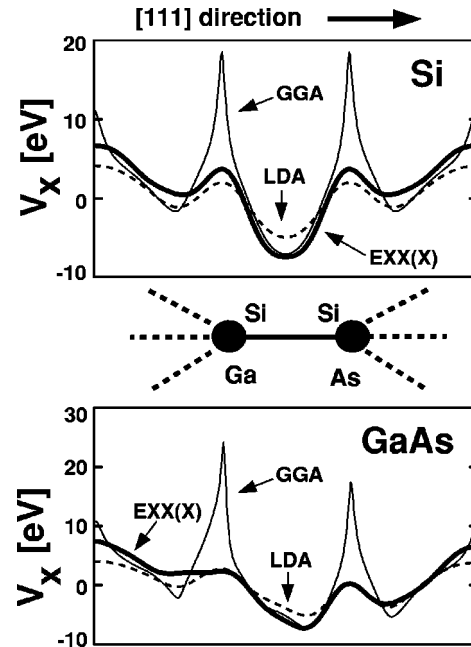


FIG. 7. Comparison of the exact EXX(X), LDA, and GGA exchange potentials in Si and GaAs along the bond axis ([111] direction). The positions of the nearest-neighbor atoms are indicated by black circles. The LDA and GGA potentials have been evaluated with the EXX(X) density. The spatial average values of all exchange potentials have been set to zero.

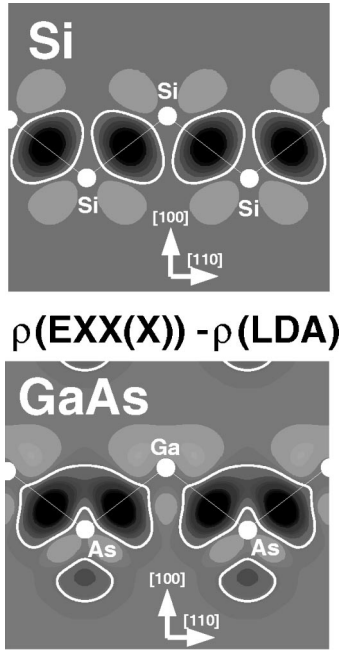


FIG. 8. Contour plots of the calculated density differences $\rho^{EXX(X)}(\mathbf{r}) - \rho^{LDA}(\mathbf{r})$ for Si and GaAs (in units of electrons per unit cell), respectively. The grey scales run from +3 (black areas) to -3 (white areas). White lines separate regions with positive and negative density differences, and mark the zero contour line. All results have been obtained with exchange only calculations.

Figure 8 shows contour plots of the differences between the self-consistently computed valence charge densities $\rho^{EXX(X)}(\mathbf{r})$ and $\rho^{LDA}(\mathbf{r})$. Analogously to the exchange potentials, the LDA densities are more homogeneous than the EXX(X) densities, while the GGA densities resemble the latter more closely. This is most pronounced in the bond regions. The same trend has been found before in atoms.^{36,37,48,80} As has been found before in atoms,^{36,81} we obtain Hartree-Fock densities that are nearly identical to the EXX(X) densities.

We now turn to an assessment of exchange and total energies. In principle, the accuracy of any approximate energy functional $E^{app}[\rho]$ can be checked in two ways. One may determine the deviation from the exact functional $E^{EXX}[\rho]$ at a given, fixed, reference density ($\rho^{EXX(X)}$, for example). Alternatively, one may determine the self-consistent density ρ^{app} (which carries implicitly information on the potential obtained in the approximate method) and compare the self-consistently calculated $E^{app}[\rho^{app}]$ with the exact value $E^{EXX}[\rho^{EXX(X)}]$. We have carried out both types of comparisons for the LDA and GGA exchange energy functionals for eight semiconductors, and the results are given in Fig. 9. The approximate (LDA and GGA) functionals were evaluated both at the reference densities $\rho^{EXX(X)}$ and at the corresponding self-consistent densities (ρ^{LDA} and ρ^{GGA}), respectively. We find the LDA exchange energies to be too small by 4–7 %, whereas the GGA exchange energies deviate from the exact values by typically 1%, no matter which density is used. One reason is that GGA exchange functionals benefit from error cancellations between different spatial regions.⁸² The corresponding errors in atomic GGA exchange energies are of the same order of magnitude.^{36,48,65,80,82,83}

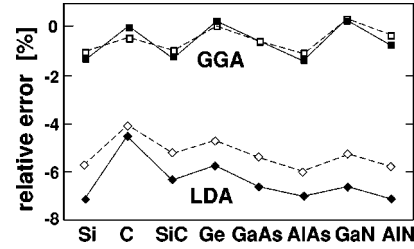


FIG. 9. Relative deviations (in percent) of LDA and GGA (Ref. 4) exchange energies from the exact EXX(X) values for various semiconductors. Solid symbols denote self-consistent exchange-only results, and open symbols refer to energies evaluated with the density $\rho^{EXX(X)}$.

Importantly, LDA and GGA exchange energies change only slightly when evaluated at the EXX(X) density $\rho^{EXX(X)}$. This shows that the main part of the error both in the LDA and GGA originates directly in the energy functional and not in the density that affects the energy indirectly.

In Table V, we list the individual contributions to the total energy Eq. (3) of GaAs within the EXX(X), LDA, GGA, and Hartree-Fock methods, respectively. The results are also typical of the other semiconductors that have been studied here. The LDA is seen to underestimate the absolute values of the kinetic energy, Hartree energy, exchange energy, and electron-ion interaction energy by about 1–2 eV/atom, which predominantly results from the LDA densities being too homogeneous. The GGA results, on the other hand, are much closer to the EXX(X) values.

Since the self-consistent EXX(X) and Hartree-Fock electron densities are almost identical, the Hartree and the electron-ion interaction energies exhibit very little difference in these two methods. The kinetic energy and the exchange energy, on the other hand, differ by a larger amount from each other. The reason for this lies in that the KS and Hartree-Fock determinants differ from one another even when they yield identical densities. For a given density, the KS determinant minimizes the kinetic energy,^{1,50} whereas the Hartree-Fock determinant minimizes the expectation value $\langle \hat{T} + \hat{V}_{ee} + V_{ext} \rangle$. As a consequence, the kinetic energy T_s of the KS determinant is lower than that of the Hartree-Fock determinant. The exchange energy, on the other hand, turns out to be lower in the Hartree-Fock approach than in the EXX(X) approach, which is consistent with $E^{HF} < E^{EXX(X)}$.

TABLE V. Calculated contributions to the total energy E in GaAs in eV/atom, as obtained by different self-consistent methods. Correlations are consistently set to zero in all cases. The symbol T^{HF} denotes the Hartree-Fock kinetic energy, whereas T_s is the kinetic energy in the EXX(X), LDA, and GGA method, respectively.

	EXX(X)	LDA	GGA	HF
T_s, T^{HF}	40.63	39.60	40.94	40.72
E_H^l	10.89	9.84	10.80	10.88
E_x, E_x^{HF}	-28.99	-27.07	-28.81	-29.17
E_{ei}^l	-19.13	-17.53	-19.41	-19.12
E_{ii}^l	-114.59	-114.59	-114.59	-114.59
E	-111.19	-109.74	-111.07	-111.28

TABLE VI. Predicted electronic properties of Ge, as obtained within the EXX method (first column), the KLI method (third column), and the EXX method with KLI rather than consistent (EXX) pseudopotentials (second column), respectively.

valence pseudopot.	EXX EXX-PP	EXX KLI-PP ^a	KLI ^a KLI-PP ^a
$-E^{ps}$ (eV/atom)		107.3	107.1
E_{coh} (eV/atom)	3.0	2.8	2.6
a_0 (Å)	5.63	5.63	5.67
B_0 (Mbar)	1.00	1.00	0.99
E_{gap}^L (eV)	1.01	1.16	0.77
E_{gap}^I (eV)	1.28	1.56	1.26

^aReference 39.

E. Comparison with KLI method

Krieger, Li, and Iafate³⁶ developed an approximation to the exact exchange potential that has been applied to atoms,³⁶ molecules,⁸⁴ jellium clusters,⁸⁵ and semiconductors.^{38–40} It is based on a simplification of the OEP integral equation,²⁷ where the energy denominators in the Green function are replaced by a constant. While this approach agrees with the EXX method in atoms very well, it is still an open question how well it compares to the EXX method in solids. Here, we compare the results of three different approaches, namely, EXX, KLI, and EXX with KLI pseudopotentials. In all of these cases, LDA correlations have been used. The latter two approaches allow one to directly compare the KLI versus EXX treatment of the valence electrons, since the same KLI pseudopotentials are employed.⁸⁶ For the example of Ge, the results are summarized in Table VI.

With KLI pseudopotentials, we find the EXX total energies to be lower than the KLI total energies. In the exchange-only case, this follows from the variational nature of the EXX potential, and correlations apparently do not alter this ordering. This energy difference is fairly small for Ge, but reaches 0.6 eV per atom for Si. The dominant difference

between the two approaches lies in the energy gaps, however. In Si, Ge, and GaAs, we obtain EXX gaps that are larger than the KLI gaps by 0.4–0.5 eV.

These conclusions do not change much when the valence and core electrons are treated within the same approach (KLI or EXX). For Si and Ge, respectively, the consistent KLI cohesive energies lie 0.8 and 0.4 eV per atom below the consistent EXX ones. Also, the EXX band gaps substantially exceed the KLI values.

V. SUMMARY AND CONCLUSIONS

A self-consistent first-principles method has been presented that goes beyond the standard local-density approximation in a rigorous and systematic fashion. It treats the local exchange potential $V_x(\mathbf{r})$ and the exchange energy E_x of Kohn-Sham density functional theory exactly (the EXX scheme). The ion core potentials have been calculated in a consistent EXX pseudopotential framework, and correlations have been treated within the LDA or GGA. The scheme yields a local multiplicative total potential which allows one to carry out extensive band-structure calculations with a moderate computational cost.

The major success of this method lies in the finding that it predicts both the eigenvalue gaps throughout the Brillouin zone, effective masses, and cohesive properties in very good agreement with experiment. In addition, the EXX scheme allows one to assess approximate exchange functionals such as the LDA and GGA quantitatively. Generally, GGA exchange and total energies are found to be significantly closer to the exact values than the corresponding LDA results. Finally, we find the exchange part Δ_x of the discontinuity Δ_{xc} to be 3–5 times the experimental band gap.

ACKNOWLEDGMENTS

We are indebted to L. Kleinman and M. Bylander for sending us their KLI pseudopotentials, and to E. Engel for making his OEP results for jellium spheres available to us prior to publication. Financial support by the Deutsche Forschungsgemeinschaft (DFG) and Bayerische Forschungsvorhaben FOROPTO II is gratefully acknowledged. A.G. personally thanks the DFG for financial support.

- ¹R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
- ²J. P. Perdew and Y. Wang, *Phys. Rev. B* **33**, 8800 (1986).
- ³J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).
- ⁴A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- ⁵J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ⁶D. C. Langreth and M. J. Mehl, *Phys. Rev. B* **28**, 1809 (1983).
- ⁷L. C. Wilson and M. Levy, *Phys. Rev. B* **41**, 12 930 (1990).
- ⁸C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- ⁹D. J. Lacks and R. G. Gordon, *Phys. Rev. A* **47**, 4681 (1993).
- ¹⁰R. Colle and D. Salvetti, *Theor. Chim. Acta.* **37**, 329 (1975); **53**, 55 (1979).
- ¹¹R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
- ¹²M. Fuchs, M. Bockstedte, E. Pehlke, and M. Scheffler, *Phys. Rev. B* **57**, 2134 (1998).
- ¹³A. Dal Corso, A. Pasquarello, A. Baldereschi, and R. Car, *Phys. Rev. B* **53**, 1180 (1996).
- ¹⁴Y.-M. Juan, E. Kaxiras, and R. G. Gordon, *Phys. Rev. B* **51**, 9521 (1995).
- ¹⁵C. Filippi, D. J. Singh, and C. J. Umrigar, *Phys. Rev. B* **50**, 14 947 (1994).
- ¹⁶G. Ortiz, *Phys. Rev. B* **45**, 11 328 (1992).
- ¹⁷J. P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).
- ¹⁸L. J. Sham and M. Schlüter, *Phys. Rev. Lett.* **51**, 1888 (1983).
- ¹⁹L. J. Sham and M. Schlüter, *Phys. Rev. B* **32**, 3883 (1985).
- ²⁰J. P. Perdew, *Int. J. Quantum Chem., Symp.* **19**, 497 (1986).
- ²¹M. M. Rieger and P. Vogl, *Phys. Rev. B* **52**, 16 567 (1995).
- ²²J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- ²³J. A. Alonso and L. A. Girifalco, *Solid State Commun.* **24**, 135 (1977); *Phys. Rev. B* **17**, 3735 (1978).
- ²⁴O. Gunnarsson, M. Jonson, and B.I. Lundqvist, *Solid State Commun.* **24**, 765 (1977); *Phys. Rev. B* **20**, 3136 (1979).

- ²⁵J. C. Slater, Phys. Rev. **81**, 385 (1951).
- ²⁶Y. M. Gu, D. M. Bylander, and L. Kleinman, Phys. Rev. B **50**, 2227 (1994).
- ²⁷J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).
- ²⁸V. Sahni, J. Gruenebaum, and J. P. Perdew, Phys. Rev. B **26**, 4371 (1982).
- ²⁹T. Kotani, Phys. Rev. Lett. **74**, 2989 (1995).
- ³⁰T. Kotani and H. Akai, Phys. Rev. B **54**, 16 502 (1996).
- ³¹T. Kotani, Phys. Rev. B **50**, 14 816 (1994).
- ³²T. Kotani and H. Akai, Phys. Rev. B **52**, 17 153 (1995).
- ³³Y. Li, J. B. Krieger, M. R. Norman, and G. J. Iafrate, Phys. Rev. B **44**, 10 437 (1991).
- ³⁴J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).
- ³⁵D. M. Bylander and L. Kleinman, Phys. Rev. Lett. **74**, 3660 (1995).
- ³⁶J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **47**, 165 (1993).
- ³⁷J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **46**, 5453 (1992).
- ³⁸D. M. Bylander and L. Kleinman, Phys. Rev. B **52**, 14 566 (1995).
- ³⁹D. M. Bylander and L. Kleinman, Phys. Rev. B **54**, 7891 (1996).
- ⁴⁰D. M. Bylander and L. Kleinman, Phys. Rev. B **55**, 9432 (1997); **56**, 7022(E) (1997).
- ⁴¹M. Städele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. **79**, 2089 (1997).
- ⁴²M. Moukara, M. Städele, J. A. Majewski, A. Görling, and P. Vogl (unpublished).
- ⁴³A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994); Int. J. Quantum Chem., Symp. **29**, 93 (1995).
- ⁴⁴A. Görling, Phys. Rev. B **53**, 7024 (1996).
- ⁴⁵J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
- ⁴⁶J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992), and references therein.
- ⁴⁷M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).
- ⁴⁸C. Filippi, X. Gonze, and C. J. Umrigar, in *Recent Developments and Applications of Density Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam, 1996), and references therein.
- ⁴⁹C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).
- ⁵⁰A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).
- ⁵¹J. Ihm, M. L. Cohen, and D. J. Chadi, Phys. Rev. B **21**, 4592 (1980).
- ⁵²N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991); G. B. Bachelet, D. R. Hamann, and M. Schlüter, *ibid.* **26**, 4199 (1982).
- ⁵³M. Städele, J. A. Majewski, and P. Vogl, Phys. Rev. B **56**, 6911 (1997).
- ⁵⁴F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).
- ⁵⁵M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
- ⁵⁶R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10 159 (1988).
- ⁵⁷J. P. Perdew, Phys. Rev. B **33**, 8822 (1986); **34**, 7406(E) (1986).
- ⁵⁸S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).
- ⁵⁹M. M. Rieger, Ph.D. thesis, Technical University of Munich, München, 1995.
- ⁶⁰D. Vogel, P. Krüger, and J. Pollmann, Phys. Rev. B **55**, 12 836 (1997).
- ⁶¹S. Kalvoda, B. Paulus, P. Fulde, and H. Stoll, Phys. Rev. B **55**, 4027 (1997).
- ⁶²M. Causa, R. Dovesi, and C. Roetti, Phys. Rev. B **43**, 11 937 (1991).
- ⁶³S. Massidda, M. Posternak, and A. Baldereschi, Phys. Rev. B **48**, 5058 (1993), and references therein.
- ⁶⁴E. Shirley, X. Zhu, and S. G. Louie, Phys. Rev. B **56**, 6648 (1997).
- ⁶⁵J. P. Perdew, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985), p. 265.
- ⁶⁶O. Gunnarsson and K. Schönhammer, Phys. Rev. Lett. **56**, 1968 (1986).
- ⁶⁷W. Knorr and R. W. Godby, Phys. Rev. Lett. **68**, 639 (1992).
- ⁶⁸M. Lannoo, M. Schlüter, and L. J. Sham, Phys. Rev. B **32**, 3890 (1985).
- ⁶⁹A. Görling and M. Levy, Phys. Rev. A **52**, 4493 (1995); A. Görling, *ibid.* **54**, 3912 (1996).
- ⁷⁰X. Zhu, M. S. Hybertsen, and S. G. Louie, in *Atomic-Scale Calculations of Structures in Materials*, edited by M. S. Dowe and M. Schlüter, MRS Symposia Proceedings No. 193 (Materials Research Society, Pittsburgh, 1990), p. 113.
- ⁷¹V. Fiorentini, Phys. Rev. B **46**, 2086 (1992).
- ⁷²G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368 (1955).
- ⁷³*Semiconductors. Physics of Group IV Elements and III-IV Compounds*, edited by K.-H. Hellwege, O. Madelung, M. Schulz, and H. Weiss, Landolt-Börnstein, New Series, Group III, Vol. 17, Pt. a (Springer, Berlin, 1982).
- ⁷⁴V. I. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 4343 (1997).
- ⁷⁵B. Adolph, V. I. Gavrilenko, K. Tenelsen, F. Bechstedt, and R. Del Sole, Phys. Rev. B **53**, 9797 (1996).
- ⁷⁶R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11 789 (1993).
- ⁷⁷M. Alouani, L. Brey, and N. E. Christensen, Phys. Rev. B **37**, 1167 (1988).
- ⁷⁸J. A. White and D. M. Bird, Phys. Rev. B **50**, 4954 (1994).
- ⁷⁹E. Engel and S. H. Vosko, Phys. Rev. B **47**, 13 164 (1993).
- ⁸⁰C. J. Umrigar and X. Gonze, Phys. Rev. A **50**, 3827 (1994).
- ⁸¹T. Grabo, T. Kreibich, and E. K. U. Gross, in *Quantum Systems in Chemistry and Physics: Trends in Methods and Applications* (Kluwer Academic, Dordrecht, 1997).
- ⁸²J. P. Perdew, K. Burke, and M. Ernzerhof, in *Density Functional Methods in Chemistry*, edited by B. B. Laird, R. Ross, and T. Ziegler, American Chemical Society Symposium Series, May 1996 (American Chemical Society, Washington, DC, 1996).
- ⁸³T. Zhu, C. Lee, and W. Yang, J. Chem. Phys. **98**, 4814 (1993).
- ⁸⁴T. Grabo and E. K. U. Gross, Int. J. Quantum Chem. **64**, 95 (1997).
- ⁸⁵E. Engel (private communication).
- ⁸⁶We have used the KLI pseudopotentials in semilocal form, in contrast to Refs. 35, 39, and 40 where a separable form is used. This leads to very minor differences in band gaps ≤ 0.05 eV, lattice constants ≤ 0.01 Å, cohesive energies ≤ 0.02 eV/atom, and bulk moduli $\leq 5\%$.
- ⁸⁷S. Strite, J. Ruan, Z. Li, N. Manning, A. Salvador, H. Chen, D. J. Smith, W. J. Choyke, and H. Morkoç, J. Vac. Sci. Technol. B **9**, 1924 (1991).
- ⁸⁸J. E. Ortega and F. J. Himpsel, Phys. Rev. B **47**, 2130 (1993).
- ⁸⁹A. Mainwood, in *Properties and Growth of Diamond*, edited by G. Davies (Electronic Materials Information Service, London, 1994), p. 3.

- ⁹⁰W. R. L. Lambrecht, B. Segall, M. Suttrop, M. Yoganathan, R. P. Devathy, W. J. Choyke, J. A. Edmond, J. A. Powell, and M. Alouani, *Appl. Phys. Lett.* **63**, 2747 (1993); *Phys. Rev. B* **50**, 10 722 (1994).
- ⁹¹G. Ramirez-Flores, H. Navarro-Contreras, A. Lastras-Martinez, R. C. Powell, and J. E. Greene, *Phys. Rev. B* **50**, 8433 (1994).
- ⁹²M. Drechsler, D. M. Hofmann, B. K. Meyer, T. Detchprohm, H. Amano, and I. Akasaki, *Jpn. J. Appl. Phys.* **34**, L1178 (1995).
- ⁹³H. Morkoç, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, *J. Appl. Phys.* **76**, 1363 (1994).
- ⁹⁴M. Palummo, C. M. Bertoni, L. Reining, and F. Finocchi, *Physica B* **185**, 404 (1993).
- ⁹⁵*CRC Handbook of Chemistry and Physics*, 73rd ed., edited by D. R. Lide (Chemical Rubber Company, Boca Raton, FL, 1992).
- ⁹⁶P. Lautenschlager, M. Garriga, S. Logothetidis, and M. Cardona, *Phys. Rev. B* **35**, 9174 (1987), and references therein.