

Local-field and excitonic effects in the calculated optical properties of semiconductors from first-principles

B. Arnaud and M. Alouani

Institut de Physique et Chimie des Matériaux (IPCMS), UMR 7504 du CNRS, Université Louis Pasteur, 23 rue du Loess, 67037 Strasbourg, Cedex, France

(Received 11 September 2000; published 7 February 2001)

The recently developed GW approximation (GWA) based on the all-electron full-potential projector augmented wave method is used to study the local-field (LF) and electron-hole excitation effects in the optical properties of small-, medium-, and large-band-gap semiconductors: Si, InP, AlAs, GaAs, and diamond. It is found that while the use of the GWA energies instead of local-density approximation (LDA) eigenvalues has a tendency to align the calculated structures in the optical spectra with their experimental counterparts, the LF effects do not change these peak positions but systematically reduce the intensities of the so-called E_1 and E_2 structures in all the optical spectra. Taking into account the electron-hole interaction, shifts the theoretical oscillator strength towards lower photon energies and thereby improves considerably the comparison with experiment. It is also shown that the LDA static dielectric constant, a ground-state property, is considerably improved when the LF effects are included. On the other hand, as expected, the static dielectric function obtained using the GW quasiparticle energies, and including the LF effects, is underestimated for all the semiconductors. Including the excitonic effects in the theory is expected to correct this discrepancy with experiment.

DOI: 10.1103/PhysRevB.63.085208

PACS number(s): 71.10.-w, 71.15.Mb, 71.20.Nr

I. INTRODUCTION

The electronic structures of semiconductors and insulators are now well described by means of *ab initio* methods based on the density-functional theory within the local-density approximation¹ in conjunction with the so-called GW approximation of Hedin.² In this approximation the self-energy operator is given as a product of the Green's function G times the screened Coulomb interaction W . The excited states obtained with this approach are in good agreement with angle-resolved photoemission experiments.³⁻⁵ However, the one-electron description of the optical properties of materials based on the knowledge of the GW electronic structure is not satisfactory. In particular, (1) when the GW energies are used the peak positions are much higher in energy than the experimental ones, (2) the relative intensity of the so-called E_1 and E_2 peaks is not reproduced with the one-electron theory, i.e., the E_1 peak is underestimated by as much as 50% of the observed value and the E_2 peak is somewhat larger, and (3) calculations ignoring excitonic effects, but including local-field (LF) effects, reduce the intensities of both E_1 and E_2 peaks.⁶

On the other hand, the local-density approximation (LDA) description of the optical spectra is not satisfactory either, since (1) the peak positions are much lower in energy than the experimental ones due to the underestimation of the energy band gap and (2) the static dielectric constant, which can be obtained from a functional derivative of the electron density with respect to the total Kohn-Sham potential evaluated at the ground state, hence a ground-state property, is overestimated by the LDA calculation.⁷⁻¹³ However, some of this overestimation is primarily due to the neglect of the local-field effects, and as it can be seen in this paper, the inclusion of the LF effects improves somewhat the agreement with experiment.

It was always assumed that the inclusion of the electron-hole excitations in the interaction of light and matter is the missing ingredient for an adequate comparison of the theoretical and experimental optical spectra. Model calculations have somehow qualitatively confirmed this assumption.^{14,15} However, it is only recently that *ab initio* pseudopotential (PP) calculations,¹⁶⁻¹⁸ incorporating the electron-hole interaction into the dielectric function, have been able to determine the importance of this interaction and make a realistic comparison with experiment. To achieve this goal, the so-called Bethe-Salpeter equation has been solved for a range of semiconductors and insulators using the same approach as for the model calculations.¹⁴ The outcome of this hard work was quite an achievement, and led to a good agreement of the optical spectra of Si, Ge, GaAs, diamond, and LiF with experiment.¹⁶⁻¹⁸ Those calculations clearly show that the inclusion of the two-particle effects in the dielectric function, i.e., the interaction of the electron, promoted from a valence band to a conduction band, with the hole left behind, is indeed an important ingredient for the description of the optical spectra.

In this paper, we are also motivated by the same old problem, i.e., computing correctly the optical spectra without any adjustable parameter. But, instead of using the most popular *ab initio* pseudopotential method, we use an all-electron method. The calculation becomes, of course, much more complicated due to the complexity of the basis set; nevertheless, the advantages are well worth the effort. We do *not* have to pseudoize the valence electron in the atomic region. This is a plus over the PP method, since for localized d electrons the optical matrix elements can be computed without any approximation. Even, for semiconductors, it is not clear whether the matrix elements calculated using a PP approach are not necessarily as accurate as those obtained using an all-electron theory. As it can be seen later, the dielectric functions of Si and GaAs are reproduced by two different

all-electron methods and are in excellent agreement. Therefore it is of interest to know whether a PP calculation could reproduce these results with the same accuracy.

To study the local-field and excitonic effects in various types of semiconductors we used an all-electron *ab initio* method, based on the projector-augmented-wave (PAW) method,²⁰ that was previously extended to compute the quasiparticle energies within the GWA.¹⁹ We first used the tetrahedron method²¹ to produce accurate spectra at the random-phase approximation (RPA) level and investigated the effects of local fields with the formulation of Adler-Wiser.²² Then, as it can be seen later, the tetrahedron method can no longer be used to compute the optical spectra including the excitonic effects. The standard technique for performing this calculation is to solve the Bethe-Salpeter equation for the two-particle Green's function of electron-hole pairs with an appropriately screened Coulomb interaction. Such an approach neglects multiply excited final states, i.e., we are dealing with two-particle excited states. These excited states are then used to construct the frequency dependent imaginary part of the dielectric function by a summation over these states [see Eq. (17)].

The remainder of our paper is organized as follows: In Sec. II we describe our method of calculation. In Sec. III we apply it to determine the optical properties of two distinct semiconductor groups: some small- and medium-band-gap semiconductors: Si, InP, AlAs, GaAs, and a large-band-gap semiconductor (diamond). For the diamond the optical spectrum including electron-hole interaction is widely different from the noninteracting one due to the strong electron-hole interaction. We then compare our results with available calculations and experiments. We have also determined the effects of the local fields on the electron-energy-loss function. At the end of this section we explore the static dielectric constant using both the LDA and GWA with and without local-field effects. This leads us to discuss the importance of the excitonic effects for the calculation of the quasiparticle static dielectric constants.

II. METHOD OF CALCULATION

A. Quasiparticles within the GW approximation

As described elsewhere¹⁹ the quasiparticle (QP) energies ϵ_{nk}^{qp} for a state of band index n , and crystal momentum \mathbf{k} are computed by solving a Schrödinger-like equation given by

$$(T + V_{ext} + V_h) \psi_{nk}(\mathbf{r}) + \int d^3 r' \Sigma(\mathbf{r}, \mathbf{r}', \epsilon_{nk}^{qp}) \psi_{nk}(\mathbf{r}') = \epsilon_{nk}^{qp} \psi_{nk}(\mathbf{r}), \quad (1)$$

where T is the kinetic-energy operator, V_{ext} is the external (ionic) potential, V_h is the Hartree potential due to the average Coulomb repulsion of the electrons, Σ is the self-energy operator, and ψ_{nk} is the quasiparticle wave function. In the GWA,² the self-energy Σ is approximated by

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int e^{i\delta\omega'} G(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega') d\omega', \quad (2)$$

where δ is a positive infinitesimal, W is the dynamically screened interaction calculated at the RPA level, and G is the one-electron Green's function given by

$$G(\mathbf{r}, \mathbf{r}', \omega) = \sum_{kn} \frac{\Psi_{nk}(\mathbf{r}) \Psi_{nk}^*(\mathbf{r}')}{\omega - \epsilon_{nk}^{qp} - i\delta \operatorname{sgn}(\mu - \epsilon_{nk}^{qp})}, \quad (3)$$

where μ is the chemical potential and Ψ_{nk} are the LDA wave functions. It is common to use the LDA eigenvalues in the Green's function given by Eq. (3), nonetheless, we have found that when updating the energies in the Green's function, the GWA provides quasiparticle energies that are in better agreement with experiment. This was also noticed by Hybertsen and Louie.²³ Thus, the calculated quasiparticle band gaps, using the updated Green's function, are slightly larger from those reported earlier.¹⁹ To compute the optical spectra we have used two different approaches: (1) the projector augmented wave (PAW) method²⁰ eigenvalues together with the so-called scissors operator shift,¹⁰⁻¹² (2) the quasiparticle energies obtained by using the plasmon-pole model of Engel and Farid²⁴ to describe the frequency dependence of the screened interaction W .

B. Dielectric function

1. Inclusion of local-field effects at the RPA level

In a crystal, that possesses the lattice translation symmetry, a small electric perturbation $E_0(\mathbf{q} + \mathbf{G}, \omega)$ having wave vector $\mathbf{q} + \mathbf{G}$ and frequency ω produces responses $E(\mathbf{q} + \mathbf{G}', \omega)$ of wave vectors $\mathbf{q} + \mathbf{G}'$, the \mathbf{G} and \mathbf{G}' being reciprocal-lattice vectors. The dielectric matrix describing these responses, is of the form $\epsilon_{\mathbf{G}', \mathbf{G}}(\mathbf{q}, \omega)$ and it can be written as

$$E(\mathbf{q} + \mathbf{G}', \omega) = \sum_{\mathbf{G}} \epsilon_{\mathbf{G}', \mathbf{G}}^{-1}(\mathbf{q}, \omega) E_0(\mathbf{q} + \mathbf{G}, \omega). \quad (4)$$

An external macroscopic electric field can be viewed as a perturbation of vanishingly small wave vector \mathbf{q} and, therefore, the screening of the external macroscopic field is given by the matrix element $\epsilon_{0,0}^{-1}(\mathbf{q}, \omega)$ of the inverse dielectric matrix. In insulating crystals, this results in a formula for the macroscopic dielectric function:

$$\epsilon(\omega) = \lim_{\mathbf{q} \rightarrow 0} \frac{1}{\epsilon_{\mathbf{G}, \mathbf{G}'}^{-1}(\mathbf{q}, \omega)} \Big|_{\mathbf{G} = \mathbf{G}' = 0}, \quad (5)$$

which can be rewritten as

$$\begin{aligned} \epsilon(\omega) = & \lim_{\mathbf{q} \rightarrow 0} \epsilon_{0,0}(\mathbf{q}, \omega) \\ & - \lim_{\mathbf{q} \rightarrow 0} \sum_{\mathbf{G}, \mathbf{G}' \neq 0} \epsilon_{0, \mathbf{G}}(\mathbf{q}, \omega) \epsilon_{\mathbf{G}, \mathbf{G}'}^{-1}(\mathbf{q}, \omega) \epsilon_{\mathbf{G}', 0}(\mathbf{q}, \omega). \end{aligned} \quad (6)$$

The first term of this equation is the interband contribution to the macroscopic dielectric function and the second term represents the local-field contribution to ϵ . The determination of the macroscopic dielectric constant amounts to the computation of the inverse of $\epsilon_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}, \omega)$. Adler and Wiser²² have

derived, essentially by an extension of the RPA, an approximation to $\epsilon_{\mathbf{G},\mathbf{G}'}$ for longitudinal fields,

$$\begin{aligned} \epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) &= \delta_{\mathbf{G},\mathbf{G}'} - \frac{8\pi}{\Omega|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|} \\ &\times \sum_{\mathbf{k},n,m} \frac{f_{n,\mathbf{k}-\mathbf{q}} - f_{m,\mathbf{k}}}{\epsilon_{n\mathbf{k}-\mathbf{q}}^{qp} - \epsilon_{m\mathbf{k}}^{qp} + \omega + i\delta} \\ &\times \langle n\mathbf{k}-\mathbf{q} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | m\mathbf{k} \rangle \\ &\times \langle m\mathbf{k} | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}} | n\mathbf{k}-\mathbf{q} \rangle, \end{aligned} \quad (7)$$

where n and m are the band indices, $f_{n,\mathbf{k}}$ is the zero temperature Fermi distribution, and Ω is the crystal volume. The matrix elements $\langle n\mathbf{k}-\mathbf{q} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | m\mathbf{k} \rangle$ are calculated as described in Ref. 19 in the context of the GW approximation. In this expression, the time dependence of the field was assumed to be $e^{-i\omega t}$ and the small positively defined constant δ guarantees that the matrix elements of $\epsilon(\omega)$ are analytic functions in the upper-half-plane. Such a matrix could be separated into an Hermitian part $\epsilon_{\mathbf{G},\mathbf{G}'}^{(1)}$ and an anti-Hermitian part $i\epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}$ according to

$$\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) = \epsilon_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{q},\omega) + i\epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{q},\omega) \quad (8)$$

with $\epsilon^{(2)}$ for positive ω given by

$$\begin{aligned} \epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{q},\omega) &= \frac{8\pi^2}{\Omega|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|} \\ &\times \sum_{\mathbf{k},v,c} \langle v\mathbf{k}-\mathbf{q} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | c\mathbf{k} \rangle \\ &\times \langle c\mathbf{k} | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}} | v\mathbf{k}-\mathbf{q} \rangle \\ &\times \delta(\omega - \epsilon_{c\mathbf{k}}^{qp} + \epsilon_{v\mathbf{k}-\mathbf{q}}^{qp}), \end{aligned} \quad (9)$$

and $\epsilon^{(1)}$ defined by a Kramers-Kronig (KK) transform as

$$\epsilon_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{q},\omega) = \delta_{\mathbf{G},\mathbf{G}'} + \frac{2}{\pi} P \int_0^\infty d\omega' \frac{\omega' \epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{q},\omega')}{\omega'^2 - \omega^2}. \quad (10)$$

It should be noted here that the matrix elements of $\epsilon^{(2)}$ and $\epsilon^{(1)}$ could be chosen to be real if the inversion is contained in the point group of the crystal. The calculation of the head element $\lim_{\mathbf{q}\rightarrow 0} \epsilon_{\mathbf{0},\mathbf{0}}^{(2)}(\mathbf{q},\omega)$ and of the wing elements $\lim_{\mathbf{q}\rightarrow 0} \epsilon_{\mathbf{0},\mathbf{G}}^{(2)}(\mathbf{q},\omega)$ necessitate special care because the quasiparticle energies calculated within the GWA are used to determine the optical properties of semiconductors. Instead of handling numerically $\lim_{\mathbf{q}\rightarrow 0} \langle v\mathbf{k}-\mathbf{q} | e^{-i\mathbf{q}\cdot\mathbf{r}} | c\mathbf{k} \rangle / q$, when the quasiparticle wave functions are to be used, it is reasonable to approximate the quasiparticle wave function with the LDA wave function, and take the limit analytically:²⁵

$$\lim_{|\mathbf{q}|\rightarrow 0} \langle v\mathbf{k}-\mathbf{q} | e^{-i\mathbf{q}\cdot\mathbf{r}} | c\mathbf{k} \rangle / q = \hat{\mathbf{q}} \cdot \langle v\mathbf{k} | \mathbf{p} | c\mathbf{k} \rangle / (\epsilon_{v\mathbf{k}} - \epsilon_{c\mathbf{k}}). \quad (11)$$

Here $\epsilon_{v\mathbf{k}}$ and $\epsilon_{c\mathbf{k}}$ are the LDA valence and conduction energies for wave vector \mathbf{k} , \mathbf{p} is the momentum operator, and $\hat{\mathbf{q}} = \mathbf{q}/q$. Indeed, it was shown by inspection that for Si the LDA and the GW wave functions have more than 99% overlap.²³

On the other hand, the size of the dielectric matrix is critical for the convergence of the optical spectrum. We have found that a size of 65×65 for all systems studied here is good for the convergence of the optical spectra. The imaginary part of each matrix element $\epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{q}\rightarrow 0, \omega)$ is evaluated in energy intervals of 0.1 eV up to 200 eV. Then the ‘‘real part’’ $\epsilon^{(1)}$ is deduced by means of a KK transformation defined previously, see Eq. (10). The linear tetrahedron method²¹ is employed to perform the summation over the Brillouin zone, which appears in Eq. (9). We use 8000 \mathbf{k} points in the full Brillouin zone to calculate the head element and 1000 \mathbf{k} points to calculate the wing elements and the body elements. The Hermiticity of $\epsilon_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{q}\rightarrow 0, \omega)$, the time-reversal symmetry, and the symmetry properties are used to reduce the number of independent matrix elements to be computed.

2. Inclusion of excitonic effects

We follow here another approach for the computation of the optical properties of semiconductors, that allows us to include both the local-field effects and the electron-hole interaction in the theoretical optical-absorption spectrum. Because of the electron-hole interaction, the excited states $|\lambda\rangle$ of the system are described by a linear combination of free quasielectron-quasihole pairs

$$|\lambda\rangle = \sum_{v\mathbf{c}\mathbf{k}} A_{v\mathbf{c}\mathbf{k}}^\lambda |v\mathbf{c}\mathbf{k}\rangle. \quad (12)$$

Here, $|v\mathbf{c}\mathbf{k}\rangle$ is defined by $|v\mathbf{c}\mathbf{k}\rangle = a_{v\mathbf{k}}^\dagger a_{c\mathbf{k}} |0\rangle$ where $|0\rangle$ is the ground state and $a_{c\mathbf{k}}^\dagger$ creates a quasielectron in the GW state $|c\mathbf{k}\rangle$. We limit ourselves to zero-momentum excited states and neglect multiply excited states. The electron-hole amplitudes $A_{v\mathbf{c}\mathbf{k}}^\lambda$ and the excitation energies E^λ are obtained by solving an effective two-particle Schrödinger equation, which originates from the Bethe-Salpeter equation²⁶

$$(\epsilon_{c\mathbf{k}}^{qp} - \epsilon_{v\mathbf{k}}^{qp}) A_{v\mathbf{c}\mathbf{k}}^\lambda + \sum_{v'\mathbf{c}'\mathbf{k}'} \langle v\mathbf{c}\mathbf{k} | \Xi | v'\mathbf{c}'\mathbf{k}' \rangle A_{v'\mathbf{c}'\mathbf{k}'}^\lambda = E^\lambda A_{v\mathbf{c}\mathbf{k}}^\lambda, \quad (13)$$

where the kernel Ξ represents the electron-hole interaction, and ϵ^{qp} are the quasiparticle energies obtained in the GW-PAW approximation. The relevant parameters of our calculations are the number of valence bands N_v , the number of conduction bands N_c , and the number of \mathbf{k} points $N_{\mathbf{k}}$. The set of \mathbf{k} points belong to a regular grid ($2N_{\text{gr}} \times 2N_{\text{gr}} \times 2N_{\text{gr}}$) defined by

$$\mathbf{k} = \mathbf{s} + (n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3) / 2N_{\text{gr}}, \quad (14)$$

where N_{gr} is the number of divisions along the primitive vectors \mathbf{b}_i of the reciprocal lattice, the n_i are integers ranging from $-N_{\text{gr}} + 1$ to N_{gr} , and \mathbf{s} is a small arbitrary vector that shifts the grid slightly from the origin. The symmetry breaking vector \mathbf{s} produces converged spectra with a limited num-

ber of \mathbf{k} points by avoiding degenerate eigenstates. It should be noted that no symmetry reduction holds for these \mathbf{k} points, making the calculation of the quasiparticle energies with the GW approximation more time consuming.

Following the usual approximation,²⁷ the kernel Ξ is an effective interaction, which could be expressed as the sum of two terms. The first term Ξ^{exch} is obtained as the functional derivative of the Hartree self-energy with respect to the single-particle Green's function, while the second term Ξ^{dir} is expressed as the functional derivative of the self-energy in the GW approximation, neglecting a term $G(\delta W/\delta G)$, which is expected to be small. In the basis of electron-hole pairs, the matrix elements of the exchange term Ξ^{exch} are given by

$$\langle v c \mathbf{k} | \Xi^{\text{exch}} | v' c' \mathbf{k}' \rangle = 2 \times \frac{4\pi}{\Omega} \sum_{\mathbf{G} \neq 0} \frac{1}{|\mathbf{G}|^2} \langle c \mathbf{k} | e^{i\mathbf{G}\mathbf{r}} | v \mathbf{k} \rangle \times \langle v' \mathbf{k}' | e^{-i\mathbf{G}\mathbf{r}} | c' \mathbf{k}' \rangle. \quad (15)$$

The volume of the crystal is $\Omega = N_{\mathbf{k}} \times V$, with $N_{\mathbf{k}}$ being the number of \mathbf{k} points in the sampling over the Brillouin zone and V the volume of the unit cell. The factor of 2 stems from the fact that we are dealing with singlet excited states. The summation over the reciprocal-lattice vectors is restricted to nonzero \mathbf{G} vectors because the Coulomb interaction without the long-range term of vanishing wave vector should be used to obtain the macroscopic dielectric constant.²⁸ It should be emphasized that the Bethe-Salpeter equation could be easily solved in the plane-wave basis when retaining only the exchange term in the kernel Ξ . Such a procedure leads to the expression given by Eq. (7) for the dielectric matrix, making the connection with the previous part, where the local-field effects in the macroscopic dielectric constant were obtained by inverting the dielectric-function matrix.

The second term Ξ^{dir} is related to the direct screened electron-hole interaction. We have neglected the energy dependence of Ξ^{dir} to facilitate the resolution of the Bethe-Salpeter equation. This approximation results from taking into account only the static screening of the electron-hole interaction and neglecting any dynamical screening of the electron-hole interaction. It was argued by Bechstedt *et al.*²⁹ that the dynamical effects of the electron-hole interaction are compensated by the dynamical effects of the renormalized Green's function. Within this approximation, the matrix elements of the screened electron-hole interaction are given by

$$\langle v c \mathbf{k} | \Xi^{\text{dir}} | v' c' \mathbf{k}' \rangle = - \frac{4\pi}{\Omega} \sum_{\mathbf{G}, \mathbf{G}'} \frac{\tilde{\epsilon}_{\mathbf{G}, \mathbf{G}'}^{-1}(\mathbf{q}, \omega = 0)}{|\mathbf{q} + \mathbf{G}| |\mathbf{q} + \mathbf{G}'|} \times \langle v' \mathbf{k}' | e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}} | v \mathbf{k} \rangle \times \langle c \mathbf{k} | e^{-i(\mathbf{q} + \mathbf{G}') \cdot \mathbf{r}} | c' \mathbf{k}' \rangle \delta_{\mathbf{q}, \mathbf{k}' - \mathbf{k}}, \quad (16)$$

where the static symmetrized dielectric matrices $\tilde{\epsilon}(\mathbf{q}, \omega = 0)$ are obtained within the RPA. The sizes of the dielectric matrices are 137×137 for Si and 169×169 for the other semiconductors. For a grid specified by $N_{gr} = 4$ ($8 \times 8 \times 8$ mesh),

the use of symmetry reduces the number of dielectric matrices to be evaluated from 512 to 29. This is because, although the \mathbf{k} points given by Eq. (14) have no special symmetry, the difference between \mathbf{k} and \mathbf{k}' belongs to a regular grid centered at the Γ point. When $\mathbf{q} = 0$, special attention has to be paid to the calculation of the matrix elements defined by Eq. (16). The divergence of $1/q^2$ type ($\mathbf{G} = 0, \mathbf{G}' = 0$) is integrated out over a small sphere of volume $V_{\text{BZ}}/N_{\mathbf{k}}$, where V_{BZ} is the volume of the Brillouin zone. This divergence contributes only notably when $c = c'$ and $v = v'$. The divergence of $1/|\mathbf{q}|$ type ($\mathbf{G} = 0, \mathbf{G}' \neq 0$ or $\mathbf{G} \neq 0, \mathbf{G}' = 0$) is neglected, because its contribution either averages to zero or vanishes quickly in the limit of a large number of \mathbf{k} points ($N_{\mathbf{k}} \rightarrow \infty$).

Finally, the effective Hamiltonian is diagonalized to obtain the eigenvectors A^λ and eigenvalues E_λ , which are necessary ingredients for the computation of the imaginary part of the dielectric function:¹⁶

$$\epsilon^{(2)}(\omega) = \lim_{\eta \rightarrow 0} \frac{8\pi^2}{\Omega q^2} \sum_{\lambda} \left| \sum_{v c \mathbf{k}} \langle v \mathbf{k} | e^{-i\mathbf{q}\mathbf{r}} | c \mathbf{k} \rangle A_{v c \mathbf{k}}^\lambda \right|^2 \delta(E_\lambda - \omega). \quad (17)$$

Since the tetrahedron method can no longer be used to evaluate $\epsilon^{(2)}(\omega)$, the delta function appearing in Eq. (17) is replaced by a Lorentzian function with a finite width η . On the other hand, this method, as compared to the previous one, suffers from the limited number of bands and \mathbf{k} points included in the calculation. In computing Eq. (17), we have used $N_v = 3$, $N_c = 4$, and $N_{\mathbf{k}} = 512$ ($8 \times 8 \times 8$ mesh) for all systems considered here. Such a set of parameters defines the maximum size of the Hamiltonian which can be computed on a sequential computer of 1 Gbyte of memory; it generally produced a converged optical spectra below 6 eV. Unfortunately, with this limited set, the real part of the dielectric function is found to be inaccurate, and its convergence will require a large number of bands and \mathbf{k} points. Moreover, the solution of the Bethe-Salpeter equation requires an important numerical effort because the basis set for the electron-hole wave function contains a large number, $N_v \times N_c \times N_{\mathbf{k}}$, of functions. For the set of parameters specified previously, the number of matrix elements $\langle v c \mathbf{k} | \Xi | v' c' \mathbf{k}' \rangle$ to be computed is about 3.8×10^7 . This number can be halved by using the Hermiticity of the effective Hamiltonian.

III. RESULTS AND DISCUSSION

A. Calculated optical spectra with and without local-field effects

The dynamical dielectric function of all the semiconductors studied here are calculated using the tetrahedron method²¹ to evaluate the Brillouin-zone summation in Eq. (9). To test the accuracy of the all-electron PAW method we have used the LDA energies to compute the imaginary part of the dielectric function of Si and GaAs without local field and we have compared the spectra to the full-potential linear muffin-tin orbital (FPLMTO) results.¹³ Figure 1 shows that the agreement with the FPLMTO spectra is excellent, and sets the standard for an accurate LDA calculation of the di-

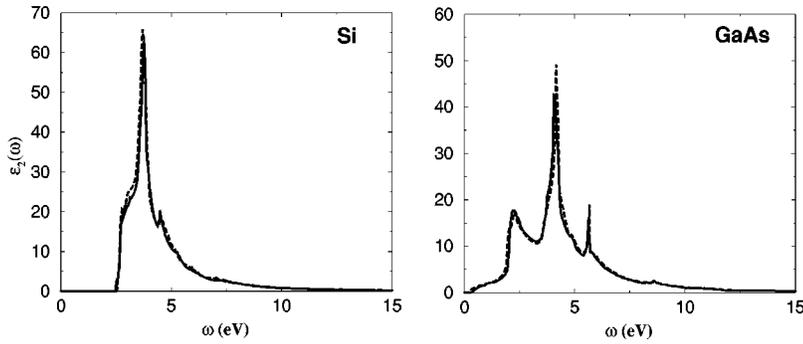


FIG. 1. Calculated LDA imaginary part of the dielectric function of Si and GaAs vs photon energy using the PAW method (solid line) and the FPLMTO method (Ref. 13) (dashed line).

electric function of Si and GaAs. This is encouraging since the FPLMTO method is a state-of-the-art first-principles method for calculating the electronic structure of materials, and in comparison, the PAW formalism is much simpler, but nevertheless the method does not lose any accuracy.

Before presenting our calculated optical properties of semiconductors, we give some details of how we performed the calculations of the optical spectra. We have used the GWA quasiparticle energies as well as the so-called scissors operator energy shift to the LDA eigenvalues to compute the dielectric function of Si and diamond. As it will be shown later, we have found that the dielectric function of Si is almost unchanged in the two calculations, and that of diamond changed only slightly near the main peak position. Because of this small change of the dielectric function due to the use of the quasiparticle energies, and because of the high CPU cost in obtaining the quasiparticle energies across the whole Brillouin zone, all the other small- and medium-band-gap semiconductor optical properties are computed using only the scissors-operator shift corresponding to the difference between the quasiparticle and LDA energies at the X point, except for GaAs where we have used a shift between the experimental and LDA energies. This is because for GaAs, where the $3d$ semicore states hybridize significantly with the valence states, the GW approximation without core polarization underestimated the experimental band gap.¹⁹

The accuracy of the macroscopic function depends on the convergence of all the elements of the microscopic dielectric matrix. As stated in the previous section we have found that a matrix of 65 by 65 \mathbf{G} vectors and the use of 200 bands in the interband transitions produce a well-converged $\epsilon(\omega)$. Figure 2 shows different elements of the microscopic dielectric function of silicon versus photon energy up to 70 eV. The highest intensity of these elements is at least one order of magnitude smaller than the $\epsilon_{(000),(000)}$ element. We compared our results to the only available results of Gavrilenko and Bechstedt,³⁰ and found that our results are in reasonable agreement with theirs. As for the reason why their imaginary parts of the off-diagonal matrix elements $\epsilon_{\mathbf{G},\mathbf{G}'}$ have nonzero contributions inside the band gap, this is because Gavrilenko and Bechstedt have chosen the origin of the reference system on one Si atom.³¹ This choice makes their optical matrix elements complex, and hence the imaginary part of the off-diagonal $\epsilon_{\mathbf{G},\mathbf{G}'}$ is a mixture of the Hermitian and anti-Hermitian parts.³¹ The same thing is true for the real part. The Hermitian part of the dielectric function, the so-called ϵ_1 , is nonzero inside the band gap, and hence their imaginary

parts of the off-diagonal elements $\epsilon_{\mathbf{G},\mathbf{G}'}$ have nonzero contributions inside the band gap. The real and imaginary parts of their diagonal element $\epsilon_{(111),(111)}$ are, respectively, Hermitian and anti-Hermitian, and are very close to our results. The Hermitian and anti-Hermitian off-diagonal parts can be obtained by multiplying the optical matrix elements by the phase factor $e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{d}}$, where \mathbf{d} is the vector representing the distance between the two atoms, or by choosing the origin of the reference frame of the atoms in the middle of the distance between the two silicon atoms.

Figure 3 shows the calculated imaginary part of the dielectric function of Si, InP, AlAs, GaAs, and diamond versus photon energies up to 8 eV with and without LF. These calculations are compared to the experimental results.^{32–35} For Si and diamond, the gray-dashed curve represents the difference between the calculated optical spectra using the GWA quasiparticle energies and the scissors-operator energy shift to the LDA eigenvalues. As stated above, these spectral differences are small and justify the use of the scissors-operator shift for the calculation of the optical spectra of the other small- and medium-band-gap semiconductors.

For Si the agreement with the empirical pseudopotential (EPP) calculation of Louie *et al.*, is good.⁶ We have found that the LF effects do not change the peak positions, but systematically reduce the intensities of the so-called E_1 and E_2 peaks.⁶ Thus, the local-field effects seem to improve the agreement with experiments concerning the intensity of the E_2 peak and the structures in the higher-energy part of the spectrum, and worsen the agreement with experiment regarding the low-energy part, where the E_1 peak is located. Nevertheless, we find it surprising that our calculations do not agree well with the *ab initio* PP calculations of Gavrilenko and Bechstedt³⁰ that are supposed to be similar to the empirical pseudopotential method calculation. The latter calculation found that while the LF underestimates the E_1 peak intensity in agreement with our calculation and EPP, it overestimates the E_2 peak intensity in disagreement with our calculation and with EPP. Gavrilenko pointed out that their off-diagonal elements of the dielectric function were highly oscillating with the increasing size of the \mathbf{G} vectors and the conduction-band index.³¹ This could be the reason for the discrepancy, because in our case we did not have any convergence problem.

The calculated spectrum of Albrecht *et al.*,¹⁶ obtained by solving the Bethe-Salpeter equation, using the PP method, in a special limit where only local-field effects are included (see

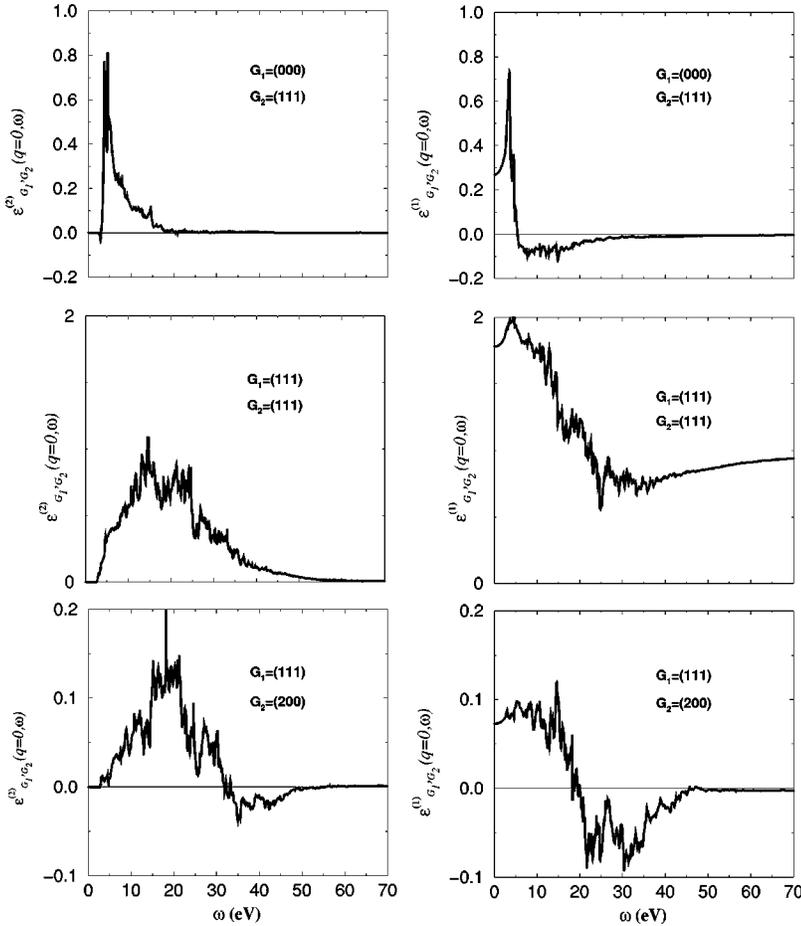


FIG. 2. Calculated elements of the real (right column) and imaginary part (left column) of the symmetrized microscopic dielectric matrix $\epsilon(\mathbf{q}, \omega)_{\mathbf{G}, \mathbf{G}'}$ of silicon for the limit $\mathbf{q} \rightarrow \mathbf{0}$ and for $(\mathbf{G}_1, \mathbf{G}_2) = (000, 111), (111, 111),$ and $(111, 200)$.

Sec. III B), agrees only qualitatively with our calculation and with the EPP results.⁶ Their E'_1 structure, which is located in energy above the E_2 structure, has a large intensity and disagrees with our calculation and other *ab initio* or EPP calculations.^{6,13} It is then not clear what makes the extra reduction of their E'_1 peak when the excitonic effects are included.

For all the other small-, and medium-band-gap semiconductors, the LF effects do not change the peak positions, but systematically reduce the intensities of the so-called E_1 and E_2 peaks. Hence, as for Si, the LF effects do not improve the agreement between theory and experiment concerning the peak positions and the intensity of the E_1 peak. Notice that for diamond the difference (shown by a gray-dashed curve) between the optical spectrum calculated using the GW quasiparticle energies and that using the scissors-operator shift is much larger than that of Si. This reflects the deviation of the GWA conduction bands from those obtained using a simple energy shift of the LDA eigenvalues toward higher energies¹⁹ to match the GWA band gap. This difference shows that the intensity of the E_2 peak of the spectrum calculated using the LDA shifted energies is slightly increased. This was also observed by Adolph *et al.*³⁶ Regarding the LF effects, our results are not in agreement with Gavrilenko and Bechsted³⁰ for the same reasons invoked for Si,³¹ and they are also in disagreement with the early empirical work of Van Vechten and Martin.³⁷ Indeed, we find that LF effects

decrease the intensity of the E_2 peak without changing its position and in addition transfer spectral weight from the low-energy side to the high-energy side, beyond the position of the E_2 peak. This trend is the same as the one observed for Si, InP, AlAs, and GaAs. Moreover, the comparison of the optical spectrum of diamond with experiment³⁵ is much worse than the other smaller-band-gap semiconductors. This is because the small dielectric constant of diamond indicates that excitonic effects are much more important. For example, most of the large discrepancy between our calculated spectrum and experiment can be attributed to these effects (see Sec. III B).

B. Calculated optical spectra including the electron-hole interaction

As stated earlier, we have used $N_v=3$, $N_c=4$, $N_{\mathbf{k}}=512$ ($8 \times 8 \times 8$ mesh), and $\eta = 0.30$ eV for all the semiconductors, except for diamond, where we used $\eta = 0.60$ eV. The choice of η is such that the agreement of the RPA dielectric functions calculated using this latter method (see Figs. 4 and 5) with those calculated using the tetrahedron method (see Fig. 3) is optimal.

To make the connection with the previous method of computing the local-field effects, we have neglected the direct screened interaction term Ξ^{dir} in the effective two-particle kernel interaction Ξ and have solved the Bethe-

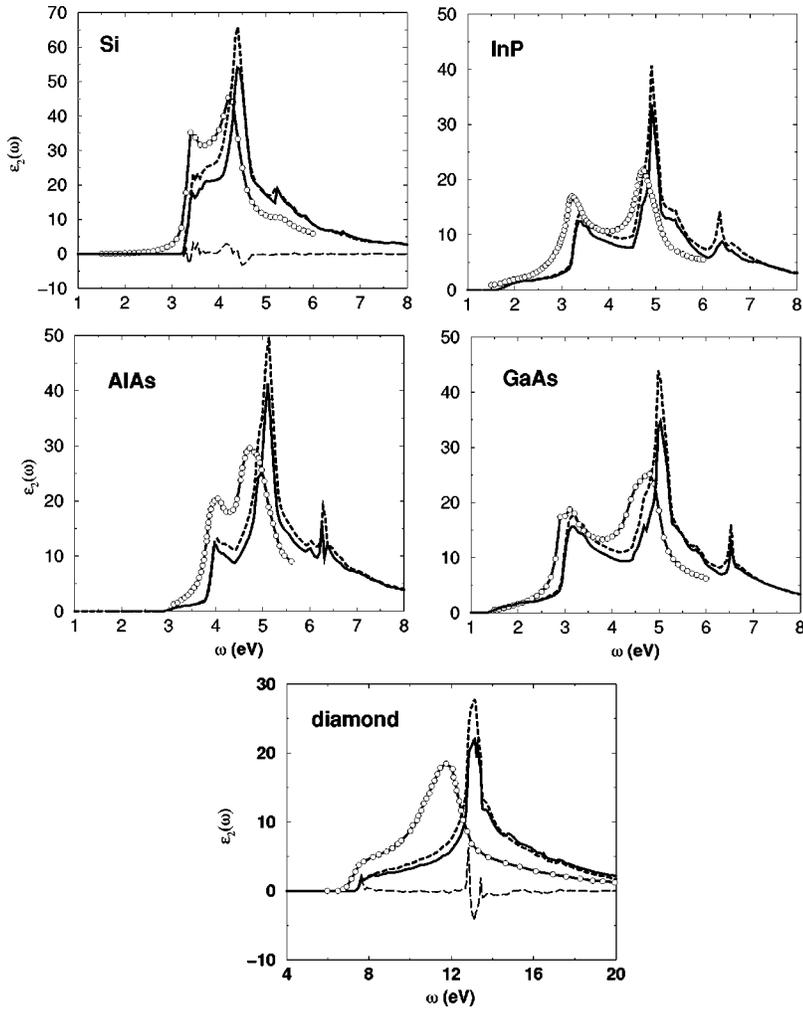


FIG. 3. Calculated noninteracting imaginary part of the dielectric function of Si, InP, AlAs, GaAs, and diamond with (full curve) and without (dashed curve) local-field effects vs photon energy compared to the experimental data (curve with empty circles) of Ref. 32 for Si, GaAs, and diamond, Ref. 33 for AlAs, Ref. 34 for InP, and Ref. 35 for diamond. The calculations have been done using the so called scissors-operator energy shift to the LDA eigenvalues, except for Si and diamond, where the quasiparticle energies across the Brillouin zone have been used. For Si and diamond, the long-dashed curve represents the difference between the calculation without local-field effects using the quasiparticle energy across the Brillouin zone and that using the rigid energy shift. This small difference justifies the use of the scissors-energy shift for the calculation of the optical properties.

Salpeter equation in this limit. This procedure is, in principle, equivalent to the inclusion of LF effects in the matrix inversion of the standard RPA calculation, i.e., to the previously followed approach. Figure 4 shows the results of the calculation for Si, which can be compared to the results

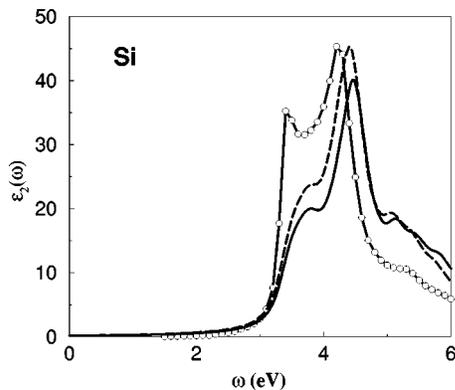


FIG. 4. Calculated noninteracting imaginary part of the dielectric function of Si with (full curve) and without (dashed curve) local-field effects versus photon energy compared to the experimental data (Ref. 32) (curve with empty circles). The local-field effects are included by solving the Bethe-Salpeter equation in an approach where the direct screened electron-hole interaction is removed.

shown in Fig. 3. It is very interesting to notice that these two calculations, although very different in nature, led to the same results at a semiquantitative level. The LF effects decrease the intensity of the E_1 and E_2 peaks, and in addition transfer spectral weight from the low-energy side to the high-energy side. Moreover, LF effects shift the E_2 peak by about 0.06 eV towards higher energies, while the results of the standard RPA calculation, see Fig. 3, indicate that the position of this peak remains unchanged. This small difference is not really significant and could be traced back to the limited number of \mathbf{k} points used in the Bethe-Salpeter equation approach.

Figure 5 shows the calculated imaginary part of the dielectric function of Si, InP, AlAs, GaAs, and diamond versus photon energy up to 8 eV with and without electron-hole interaction (excitonic effects). The calculations, including the excitonic effects, are in better agreement with the experimental results,³²⁻³⁵ regarding the absolute and relative intensities as well as the peak positions. It should be noted that the dielectric functions of InP and GaAs, with or without excitonic effects, are not fully converged with respect to the number of \mathbf{k} points. Indeed, we found that a $14 \times 14 \times 14$ mesh is necessary to get converged spectra at the RPA level. Such a large number of \mathbf{k} points cannot be handled when including the excitonic effects in the calculation of the di-

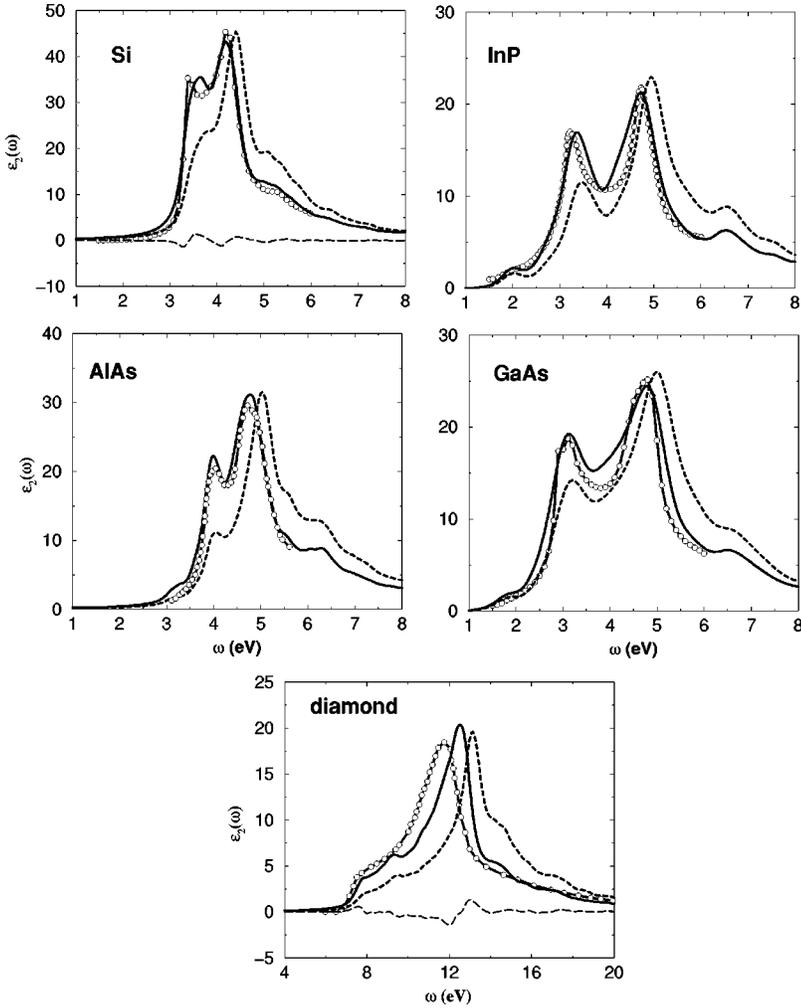


FIG. 5. Calculated imaginary part of the dielectric function of Si, InP, AlAs, GaAs, and diamond with (full curve) and without (dashed curve) excitonic effects vs photon energy compared to the experimental data (curves with empty circles) of Ref. 32 for Si and GaAs, Ref. 33 for AlAs, Ref. 34 for InP, and Ref. 35 for diamond. The calculations have been done using the so-called scissors-operator energy shift to the LDA eigenvalues, except for Si and diamond, where the quasiparticle energies across the Brillouin zone have been used. For Si and diamond, the long-dashed curve is the difference between the calculation including excitonic effects using the quasiparticle energy across the Brillouin zone and that using the rigid energy shift.

electric function of those compounds. Thus, we believe that the remaining discrepancy between theory and experiment for InP and GaAs could be attributed to the limited number of \mathbf{k} points used in the calculation. Figure 5 shows also that the electron-hole interaction seems to (1) redistribute the oscillator strength, i.e., add oscillator strength in the low-energy part of the optical spectrum and reduce it from the high-energy part and (2) shift the energy peaks towards lower energies. Notice that the electron-hole interaction does not significantly change the joint density-of-states compared to the noninteracting one. It only produces new peak positions and intensities due to a constructive and destructive interference phenomena caused by the mixing of electron-hole pairs in the excited wave function. It is worth noticing that the dielectric function of Si, obtained using the GWA quasiparticle energies or the scissors-operator shift, is almost the same. This justifies the use of the scissors-operator approximation for the calculation of the optical spectra of small- and medium-band-gap semiconductors.

The same tendency of the shift of the oscillator strength towards the lower photon energy region is also observed for a wide-band-gap semiconductor like diamond, see Fig. 5. However, the position of the main peak is shifted by about 0.6 eV towards lower photon energy that is about three times larger than for small- and medium-band-gap semiconductors.

Despite this large energy shift the agreement of the main peak with experiment is not fully recovered. Indeed, as compared to the experimental main peak, the theoretical peak is still about 0.8 eV higher in energy and its intensity is slightly overestimated. Notice that the widening of the main peak due to the inclusion of excitonic effects improves the agreement with experiment. The small shoulder around 9.3 eV is absent in the more accurate spectrum obtained using the tetrahedron method (see Fig. 3); this is also due to the limitation of the \mathbf{k} points in the calculation.

It is of interest to compare our spectra to the recent PP results available in the literature.^{16–18,38} For Si and GaAs our results are in good agreement with those of Benedict, Shirley, and Bohn¹⁷ (BSB), whereas the agreement with the results of Rohlfing and Louie¹⁸ concerning GaAs is somewhat less good. This is probably because the intensity of their E_2 peak at the GW level, without electron-hole interaction, is surprisingly much lower than the experimental one. For the diamond, our optical spectrum differ somewhat from that of BSB.¹⁷ The main peak of our spectrum is shifted, by about 0.8 eV, towards higher photon energies compared to experiment whereas theirs is not. However, our onset of absorption is in good agreement with experiment whereas theirs is shifted by about 0.5 eV towards lower photon energies.

The large discrepancy of our spectrum with that of BSB requires a detailed comparison. The model of Hybertsen-Levine-Louie (HLL) used by BSB seems to produce a somewhat less effective screening for small electron-hole separation. This resulted in a slightly larger band gaps for large-band-gap semiconductors.³⁹ The use of the HLL model together with the shifted LDA conduction bands to produce the experimental band gap of diamond could partially explain the shift of the oscillator strength towards lower photon energies in their optical spectrum of diamond. Similar down shift is obtained for their LiF optical spectrum.¹⁷ Nevertheless, we should stress that we have also calculated the optical spectrum of diamond where the excitonic effects are calculated using a model dielectric function, used previously by Gygi and Baldereschi to calculate quasiparticle energies.⁴⁰ We obtained an optical spectrum that is almost identical to that obtained using the RPA dielectric function, except that the main peak is shifted by about 0.1 eV towards lower photon energies. This shift is really small and seems to indicate that the discrepancy between our results and those of BSB could not be fully explained by the use of a model dielectric function. It seems that the largest part of the discrepancy happen already at the RPA level. Notice that our calculated RPA main peak is at 13.13 eV, whereas the corresponding peak of the BSB spectrum is at approximately 12.7 eV (this number is extracted from their published RPA spectrum). Thus, at the RPA level, the BSB spectrum is already shifted by about 0.4 eV towards lower energies with respect to ours. A 0.4 eV shift at the RPA level, plus a 0.1 to 0.2 eV shift due to the use of the model dielectric function, corresponds approximately to discrepancy between the two spectra. The discrepancy at the RPA level explains also why the onset of absorption of their optical spectrum, including the excitonic effects, is shifted by about 0.5 eV towards lower photon energies with respect to ours and experiment. It is worth mentioning that a similar unpublished PP calculation using the GW energies and the RPA screening found the main peak of diamond is about 0.7 eV too high compared to experiment and in agreement with our calculation.³⁸

The large down shift of BSB RPA spectrum with respect to our spectrum is probably due to (1) their use of the experimental band gap to produce shifted conduction bands, (2) the stretch of the valence bands by 7%, but not of the conduction bands, and (3) the type of pseudopotential could also contribute at the 0.1 eV level.

Table I gathers the calculated direct band gaps at Γ , X , and L using the LDA and GWA where the Coulomb interaction is screened using the plasmon pole of Engel and Farid,²⁴ as well as the peak positions of E_1 and E_2 peaks of the noninteracting and interacting spectra of Si, InP, AlAs, GaAs, and diamond. The calculations are compared to the angle-resolved-photoemission experiments for the energy-band gaps,^{41–44} and the calculated peak positions are compared to these obtained from the experimental optical spectra.^{32–34} As it can be seen from this table, the direct band gaps of Si, InP, and AlAs are in good agreement with the experimental results to within 0.1 eV. For GaAs the situation is more complicated due to the presence of the $3d$ semicore states and the core polarization is shown to have a large

effect.⁴⁵ It is for this reason that the experimental energy shift, as stated earlier, was used to produce the optical spectrum for this system instead of the GWA energy shift.

To conclude this section, it is worth mentioning that the calculation of excitonic effects is a difficult task. One cannot use an accurate integration method, like the tetrahedron method,²¹ which has been used to accurately determine the dielectric function at the RPA level (see the previous section). Here, we are obliged to make a sum over the excitonic states as shown in Eq. (17), and replace the delta function by a Lorentzian of a given width. This method of integration cannot resolve the fine structures in the optical spectra obtained at low temperatures because of (1) the limited number of \mathbf{k} points and (2) the smearing of the spectra caused by the Lorentzian width. For this reason the suggestion of Cardona *et al.*,⁴⁶ to make detailed comparisons of theoretical spectra with low-temperature data, remains a challenging task. To be able to resolve fine structures a large number of \mathbf{k} points as well as a small Lorentzian width are necessary. The large number of \mathbf{k} points leads to a large excitonic Hamiltonian making its diagonalization prohibitive on sequential computers.

C. Electron-energy-loss function

Figure 6 shows our calculated electron-energy-loss (EEL) functions $-\text{Im}[\epsilon^{-1}(q=0,\omega)]_{0,0}$ for Si, InP, AlAs, GaAs, and diamond. The calculations are done within the LDA with and without the local-field effects. Whenever possible the calculation is compared to available EEL spectra.^{48,49} The local-field effects seem to improve the agreement with experiment by reducing significantly the intensity of the main peak. The calculated EEL function of diamond is relatively more complicated, it has two maxima at 31.5 and 34.5 eV, and these values are shifted to 31.4 and 35.2 eV, respectively, when the LF effects are included. The experimental curve seems to present only one resonance at 32 eV. This discrepancy could be easily due to a small inaccuracy in the calculated dielectric function at these high photon energies.

We did not calculate $-\text{Im}[\epsilon^{-1}(q=0,\omega)]_{0,0}$ for the quasiparticle energy because we believe that GWA is not valid at high energies, and as pointed out in Ref. 13, the plasma resonance will be pushed towards higher energies in disagreement with experiment. This is because the electronic structure at higher energy is most probably much better described using the LDA than the GWA because (1) at these higher energies the scattering of an electron with the atomic potential is small. In this respect, these high electronic states can be obtained, most likely, from an almost free-electron theory. (2) The plasmon-pole model is not valid at these high energies.²⁴ It is, however, of interest to mention that Soinen and Shirley⁴⁷ have calculated the EEL function of diamond, LiF, and GaN including the excitonic effects but for \mathbf{q} not equal to zero. They used the scissors-operator to correct the LDA results instead of using directly the GW band structure. They found that the effect of the scissors-operator shift is canceled out by the excitonic effects.

Table II shows the values of the maxima of the EEL function compared to the experimental results obtained from the

TABLE I. Calculated direct band gaps at Γ , X , and L within the LDA and the GWA for Si, InP, AlAs, GaAs and diamond compared to the positions of the E_1 and E_2 peaks in the dielectric function with and without excitonic effects (in eV). The calculation of the self-energy is performed using 10 special \mathbf{k} points in the Brillouin zone and 200 bands. The size of the polarizability matrix is 137×137 for Si, and 169×169 for the other semiconductors. The number of reciprocal lattice vectors is 283 for Si, 307 for GaAs and AlAs, and 331 for InP and 387 for C. The QP GWA results are obtained using the plasmon-pole model of Engel-Farid (Ref. 24).

	QP energies			E_1			E_2		
	LDA	GW	Expt.	Noninterac.	Interac.	Expt.	Noninterac.	Interac.	Expt.
Si									
$E_g(\Gamma)$	2.51	3.23	3.40 ^a						
$E_g(L)$	2.57	3.35	3.30 ^a		3.68	3.4 ^e			
$E_g(X)$	3.43	4.16	4.15 ^b				4.4	4.18	4.2 ^e
InP									
$E_g(\Gamma)$	0.77	1.65	1.42 ^c						
$E_g(L)$	2.38	3.29	3.30 ^a	3.49	3.39	3.22 ^f			
$E_g(X)$	3.70	4.51	4.58 ^a				4.94	4.74	4.73 ^f
AlAs									
$E_g(\Gamma)$	1.95	2.97	3.11 ^d						
$E_g(L)$	2.90	3.90	3.92 ^a	4.05	3.99	4.04 ^g			
$E_g(X)$	3.49	4.42	4.54 ^a				5.03	4.78	4.72 ^g
GaAs									
$E_g(\Gamma)$	0.38	1.16	1.52 ^a						
$E_g(L)$	1.97	2.72	3.15 ^a	3.2	3.15	3.1 ^e			
$E_g(X)$	3.88	4.57	4.81 ^a				5.0	4.8	4.8 ^e
Diamond									
$E_g(\Gamma)$	5.53	7.42							
$E_g(L)$	11.17	13.36							
$E_g(X)$	10.92	12.93				13.13	12.53	11.75 ^h	

^aReference 41.

^bReference 42.

^cReference 43.

^dReference 44.

^eReference 32.

^fReference 34

^gReference 33.

^hReference 35.

EEL experiment⁴⁸ and from EEL spectra obtained by inverting the complex dielectric function.^{49,50} The free-electron plasma frequency is also shown for comparison. The plasma resonance of the EEL spectra of Si is in good agreement with experiment and with the free-electron plasma frequency, whereas for GaAs and diamond, the agreement is only at the semiquantitative level. It is, nevertheless, worth mentioning that the experimental plasma frequencies of GaAs and diamond are not obtained from the measured EEL spectra but rather from the EEL spectra obtained by inverting the complex dielectric functions. The experimental plasma frequencies may then be less accurate because the complex dielectric function is obtained from the reflectivity spectra using Kramers-Kronig relations. This may explain part of the discrepancy with our calculations.

D. The static dielectric constant

The electronic static dielectric function ϵ_∞ , with or without local-field effects, is computed using the Kramers-Kronig relations. The calculations were produced using the RPA dielectric function by performing analytically the limit

$\mathbf{q} \rightarrow 0$. Table III presents ϵ_∞ for all semiconductors studied here and compares them to other calculations^{7-13,51,52} and with available experimental results.⁴¹ To illustrate our data and stress the agreement with experiment we show in Fig. 7 our results versus experiment and the PP results including both LF effects and the exchange-correlation kernel.^{7-9,12} A perfect agreement with experiment is achieved when the calculated value is on the dashed line. Because ϵ_∞ is a ground-state property, we expect, *naively*, that the calculation using the LDA and including the LF effects would produce the experimental results. However, we observe only an improvement due to this effect. It is important to mention that our calculation neglects the exchange-correlation contribution to the static dielectric function. To include this contribution one basically has to compute the exchange-correlation kernel $K_{xc}(\mathbf{r}, \mathbf{r}') = \partial^2 E_{xc} / \partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}') = dV_{xc} / d\rho|_{\rho(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}')$, where E_{xc} and V_{xc} are the exchange-correlation energy and potential, respectively, and $\rho(\mathbf{r})$ is the charge density at \mathbf{r} . The calculation of K_{xc} is much more complicated in an all-electron than in a PP method, since one has to determine the matrix of the Kernel

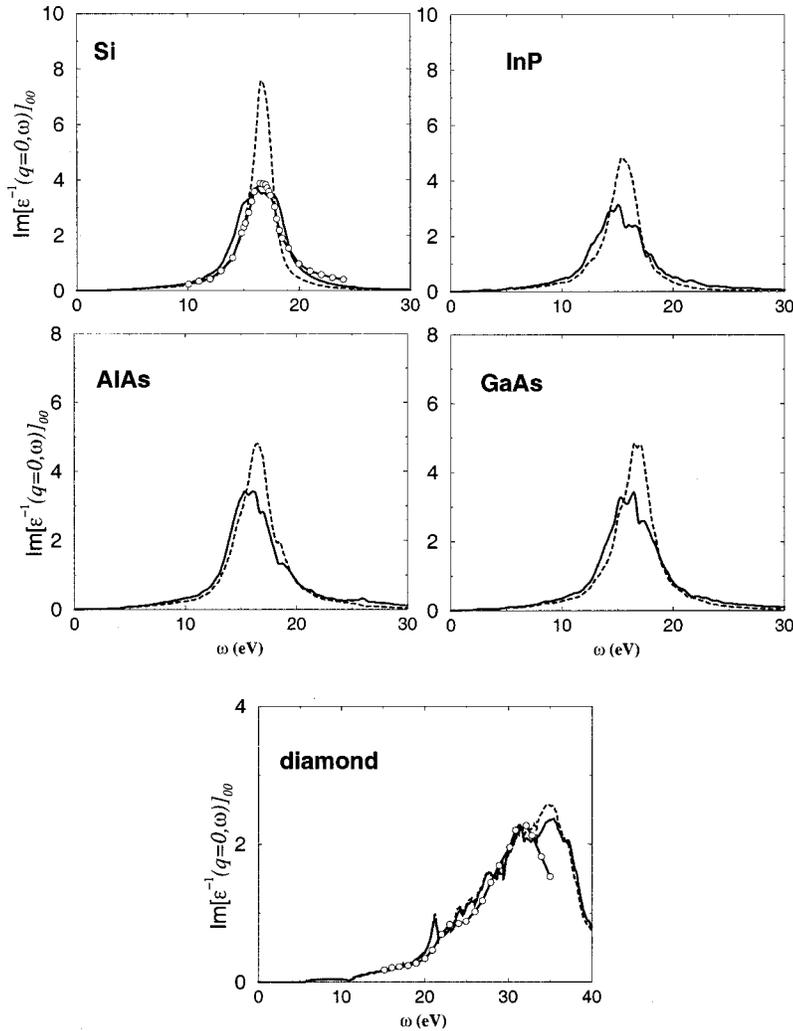


FIG. 6. Calculated electron-energy-loss function with (solid line) and without local-field effects (dashed line) of small-, medium-, and large-band-gap semiconductors: Si, InP, AlAs, GaAs, and diamond compared to available experimental results (solid line with open circle) of Ref. 48 for Si, and Ref. 49 for diamond.

in the Fourier space. The fast Fourier transform convergence of the exchange-correlation kernel with the number of \mathbf{G} vectors is very slow because of the oscillating nature of the full charge density in real space. Fortunately, it can be shown that the effects of the exchange-correlation kernel within the LDA is to increase slightly the static dielectric function with respect to that including the LF effects. This is because the K_{xc} is negative definite.⁷ Without doing any calculation, the inclusion of the K_{xc} is expected to increase the dielectric

function, and thus worsening the agreement with experiment. Such an effect is shown in Fig. 7, where the static dielectric function, as obtained using the PP method,^{7-9,12} including both the exchange-correlation kernel and the LF effects, is represented by the plus signs. The exchange-correlation kernel increases the value of ϵ_∞ including the LF effects by at most 6.4% for diamond.⁸ This is to be compared to the LF effects that reduce the LDA values by as much as 14% for AlAs. It is then tempting to attribute the remainder of the

TABLE II. Influence of the LF on the energy position of the plasmon peak of the electron-energy-loss spectra. Our calculations are compared to available experimental results and to the free-electron plasma frequency (in eV).

Material	LDA	LDA+LF	Free electron	Expt.
Si	16.6	16.5	16.6	16.4 ^a , 16.9 ^b
InP	15.5	15.0	14.8	
AlAs	16.45	15.8	15.8	
GaAs	16.8	16.4	15.6	14.7 ^a
Diamond	31.5 and 34.5	31.4 and 35.2	31.2	32 ^c

^aReference 50.

^bReference 48.

^cReference 49.

TABLE III. Effects of the LF and the QP shifts in the macroscopic dielectric constant ϵ_∞ compared to other calculations and to experiments (the experimental data are from Ref. 41).

Material	LDA	LDA+LF	QP shift	QP shift+LF	Expt.
Si	13.78	(13.41 ^a , 13.6 ^b) (13.8 ^c , 13.75 ^d)	12.39	(12.04 ^a , 12.2 ^b) (12.4 ^c , 12.8 ^e)	11.82
InP	10.71		9.50		8.89
AlAs	10.20		8.98		8.59
GaAs	14.23	(14.44 ^e , 14.17 ^f)	12.73	(13.1 ^e)	11.15
Diamond	5.90	(6.06 ^b)	5.47	(5.62 ^b , 5.5 ^e)	5.15

^aReference 7.

^bReference 8.

^cReference 11.

^dReference 13.

^eReference 51.

^fReference 52.

discrepancy with experiment to the LDA.⁵³ Notice, however, that for a large band-gap semiconductor, like diamond, the LDA results including the LF effects are in good agreement with experiment. This trend was also observed by other researchers.^{13,54}

On the other hand, the quasiparticle computation of ϵ_∞ is of importance, since the results can be directly compared to these obtained within the LDA. A comparison with experiment will show whether the electron-hole interaction is important or not. It is of importance to point out that when the static dielectric function is calculated using GWA energies, including the LF effects, the calculation underestimates the experimental ϵ_∞ for all the semiconductors studied here re-

gardless of the size of the band gap or the type of semiconductor. This suggests the importance of the excitonic effects, which are expected to produce a positive contribution, leading to a better agreement with experiment. This positive contribution arises from the transfer of the oscillator strength towards lower photon energies. In the case of large-band-gap material like the diamond, where the static dielectric function within LDA including LF effects agrees nicely with experiment, the excitonic contribution is expected to cancel out the QP correction. It is however unfortunate that we cannot converge the static dielectric function including the excitonic effects. Our estimation suggests that its convergence requires a large number of both \mathbf{k} points and bands, making the size of the excitonic Hamiltonian exceedingly large. It is however interesting to mention that a recent calculation by Chang *et al.*, found that the excitonic effects increase the calculated GW static dielectric of α quartz leading to an excellent agreement with the experimental result.⁵⁵

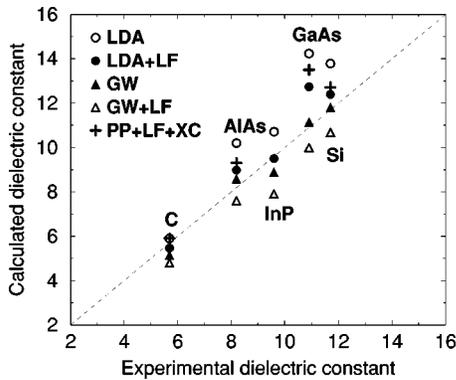


FIG. 7. Calculated static dielectric function compared to experimental results and with the PP results including the local-field effects and the exchange-correlation kernel (Ref. 7 for Si, Ref. 9 for AlAs, Ref. 12 for GaAs, and Ref. 8 for diamond). The PP results are represented by the plus signs. The open circles represent the LDA values without local-field effects (LF), the filled circles the LDA values with LF, and the up triangles the LDA without LF but with an energy shift corresponding to the GW correction of the direct band gap at the X point, the empty up triangles are the LDA values with the GW energy shift and the LF (see the text). A perfect agreement with experiment is achieved when a calculated value is on the dashed line. Notice that when the GW energy shift and the LF are included, the calculation underestimates the static dielectric constant for all these semiconductors regardless of the size of the band gap. This suggests the importance of the excitonic effects, which are expected to produce a positive correction leading to a better agreement with experiment.

IV. CONCLUSION

We have used our previously implemented GWA within the all-electron PAW method to study the optical properties of some small-, medium-, and large-band-gap semiconductors: Si, InP, AlAs, GaAs, and diamond. In general, the inclusion of the QP energy shift and the LF effects improves only slightly the agreement with experiment. In particular, the LF effects reduce the intensities of the so-called E_1 and E_2 peaks without changing their energy positions. This reduction of the peak intensity worsens the agreement for the E_1 peak but improves it for the E_2 peak. This trend is observed for all the studied semiconductors and is found to be in agreement with the EPP calculation of Louie *et al.*⁶ The QP energy shift pushes the calculated peaks towards higher energies in agreement with experiment.

The inclusion of the excitonic effects produces an excellent agreement with the experimental data, i.e., the agreement with experiment concerning the positions and intensities of these peaks is recovered, except for diamond where the agreement was at the semiquantitative level. Thus, the shift of the peaks towards higher energy, when the GWA

energies are used, is canceled out by the excitonic effects.

The static dielectric function ϵ_{∞} , with or without local-field effects, is computed by means of the Kramers-Kronig relations. The calculations were performed using the RPA dielectric function by performing analytically the limit $\mathbf{q} \rightarrow 0$. Because the static dielectric function is a ground-state property, one may naively expect that a calculation using the LDA and including the LF effect would describe the experimental results. However, our calculations and other PP calculations^{7-9,12} indicate that for small- and medium-band-gap semiconductors, we observe only an improvement due to the LF effects. Thus, we concluded that the remainder of the discrepancy is due to the use of the LDA.⁵³ However, for large-band-gap semiconductors, the LDA results including the LF effects are found to be in good agreement with experiment.

On the other hand, the QP calculation, when both the

GWA energy shift and LF effects are included, underestimates the static dielectric function for all the semiconductors studied here. This suggests the importance of the excitonic effects, which are expected to increase the static dielectric function.⁵⁵ At the present time, the excitonic effects in the static dielectric function cannot be accurately determined because of the large number of both \mathbf{k} points and bands required for its convergence.

ACKNOWLEDGMENTS

We thank P. Blöchl for providing us with his PAW code and for useful discussions. Part of this work was done during our visit to the Ohio State University, and we thank J. W. Wilkins and W. Aulbur for useful discussions. The supercomputer time was granted by CINES on the IBM SP2 supercomputer (Project No. gem1100).

-
- ¹P. C. Hohenberg and W. L. Kohn, Phys. Rev. **136**, B864 (1964); W. L. Kohn and L. J. Sham, Phys. Rev. A **140**, A1113 (1965).
- ²L. Hedin, Phys. Rev. **139**, A796 (1965); L. Hedin and S. Lundquist, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1969), Vol. 23, p. 1.
- ³F. Araysetianwan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).
- ⁴W. G. Aulbur, L. Jönsson, and J. W. Wilkins, Solid State Physics **54**, 1 (1999).
- ⁵L. Hedin, J. Phys.: Condens. Matter **11**, R489 (1999).
- ⁶S. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. **34**, 155 (1975).
- ⁷S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).
- ⁸M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).
- ⁹S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989).
- ¹⁰Z. H. Levine and D. C. Allan, Phys. Rev. B **43**, 4187 (1991).
- ¹¹Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).
- ¹²Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **66**, 41 (1991).
- ¹³M. Alouani and J. M. Wills, *Electronic Structure and Physical Properties of Solids*, edited by H. Dreyssé, Lecture Notes in Physics, Mont St. Odile (Springer-Verlag, Berlin, 1998); Phys. Rev. B **54**, 2480 (1996).
- ¹⁴W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980).
- ¹⁵M. del Castillo-Mussot and L. J. Sham, Phys. Rev. B **31**, 2092 (1985).
- ¹⁶S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998).
- ¹⁷L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998); Phys. Rev. B **57**, R9385 (1998).
- ¹⁸M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).
- ¹⁹B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).
- ²⁰P. E. Blöchl, Phys. Rev. B **50**, 17 953 (1994).
- ²¹O. Jepsen and O. K. Andersen, Solid State Commun. **9**, 1763 (1971); G. Lehmann and M. Taut, Phys. Status Solidi B **54**, 469 (1972).
- ²²S. L. Adler, Phys. Rev. **126**, 413 (1962); N. Wiser, *ibid.* **129**, 62 (1963).
- ²³M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986); Comments Condens. Matter Phys. **13**, 223 (1987).
- ²⁴G. E. Engel and B. Farid, Phys. Rev. B **47**, 15 931 (1993).
- ²⁵R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11 789 (1993).
- ²⁶G. Strinati, Phys. Rev. B **29**, 5718 (1984).
- ²⁷G. Baym and L. Kadanoff, Phys. Rev. **124**, 287 (1961).
- ²⁸V. Ambegaokar and W. Kohn, Phys. Rev. **117**, 423 (1960).
- ²⁹F. Bechstedt, K. Tenelsen, B. Adolph, and R. Del Sole, Phys. Rev. Lett. **78**, 1528 (1997).
- ³⁰V. I. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 4343 (1997); **54**, 13 416 (1996). Notice that the curve labels of Figs. 4, 5, and 6 of their 1997 PRB paper are accidentally switched (Ref. 31). The legend should read: ‘‘Dashed line: with LF and XC effects; solid line: without LF and XC effects (only $\mathbf{G}=\mathbf{G}'=0$).’’
- ³¹V. I. Gavrilenko (private communication).
- ³²D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).
- ³³M. Garriga, M. Kelly, and M. Cardona, Thin Solid Films **233**, 122 (1993).
- ³⁴S. M. Kelso, D. E. Aspnes, M. A. Pollack, and R. E. Nahory, Phys. Rev. B **26**, 6669 (1982).
- ³⁵H. R. Philipp and E. A. Taft, Phys. Rev. **136**, A1445 (1964); see also A. D. Papadopoulos and E. Anastassakis, Phys. Rev. B **43**, 5090 (1991).
- ³⁶B. Adolph, V. I. Gavrilenko, K. Tenelsen, F. Bechstedt, and R. Del Sole, Phys. Rev. B **53**, 9797 (1996).
- ³⁷J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett. **28**, 446 (1972).
- ³⁸S. Albrecht, Ph.D. thesis, CEA, Saclay, 1999, France, p. 118.
- ³⁹X. Zhu and S. G. Louie, Phys. Rev. **43**, 14 142 (1991).
- ⁴⁰F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).
- ⁴¹*Numerical Data and Functional Relationships in Science and Technology*, edited by K. H. Hellwege and O. Madelung, Landolt-Börnstein, New Series, Group III (Springer, Berlin, 1982), Vols. 17a and 22a.
- ⁴²J. E. Ortega and F. J. Himpsel, Phys. Rev. B **47**, 2130 (1993).
- ⁴³M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. B **38**, 1806 (1988).

- ⁴⁴D. J. Wolford and J. A. Bradley, *Solid State Commun.* **53**, 1069 (1985).
- ⁴⁵E. L. Shirley, X. Zhu, and S. G. Louie, *Phys. Rev. B* **56**, 6648 (1997).
- ⁴⁶M. Cardona, L. F. Lastras-Martinez, and D. E. Aspnes, *Phys. Rev. Lett.* **83**, 3970 (1999); S. Albrecht *et al.*, *ibid.* **83**, 3971 (1999).
- ⁴⁷J. A. Soininen and E. L. Shirley, *Phys. Rev. B* **61**, 16 423 (2000), and references therein.
- ⁴⁸H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons*, Springer Tracts in Modern Physics (Springer, New York, 1980), Vol. 88.
- ⁴⁹H. R. Phillips and A. Taft, *Phys. Rev.* **136**, A1445 (1964).
- ⁵⁰H. R. Philipp and H. Ehrenreich, *Phys. Rev.* **129**, 1550 (1963).
- ⁵¹M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. B* **48**, 17 791 (1993).
- ⁵²J. E. Reynolds, Z. H. Levine, and J. W. Wilkins, *Phys. Rev. B* **51**, 10 477 (1995).
- ⁵³X. Gonze, Ph. Ghosez, and R. W. Godby, *Phys. Rev. Lett.* **74**, 4035 (1995). It was shown in this paper that the exchange and correlation energy of an insulator is also dependent on the macroscopic polarization. Nevertheless, the importance of this dependence has never been shown, i.e., whether it may affect significantly the electronic and optical properties of semiconductors and insulators. Further discussion and references can be found in paper of R. Resta, *Phys. Rev. Lett.* **77**, 2265 (1996).
- ⁵⁴J. Chen, Z. H. Levine, and J. W. Wilkins, *Appl. Phys. Lett.* **66**, 1129 (1995).
- ⁵⁵E. K. Chang, M. Rohlfing, and S. G. Louie, *Phys. Rev. Lett.* **85**, 2613 (2000).