Optical Absorption of Insulators and the Electron-Hole Interaction: An *Ab Initio* Calculation

Lorin X. Benedict and Eric L. Shirley
Optical Technology Division, Physics Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Robert B. Bohn
High Performance Systems and Services Division, Information Technology Laboratory,
National Institute of Standards and Technology, Gaithersburg, Maryland 20899

(Received 23 December 1997)

We present a computationally efficient scheme to calculate the optical absorption of insulators from first principles, including the electron-hole interaction. Excited states of the solid are chosen to consist of single electron-hole pairs. The electron-hole interaction is statically screened using a model dielectric function. Only two pieces of input are required, the crystal structure of the material and the macroscopic dielectric constant. We apply this scheme to two wide-gap insulators, LiF and MgO, and obtain excellent agreement with experimental measurements of their UV reflectance spectra. [S0031-9007(98)06129-8]

PACS numbers: 78.20.Bh, 71.10.–w, 71.35.–y, 78.66.Nk

The absorption of light by an insulating solid has been important in the development of solid state theory. In the 1960s, practitioners of the empirical pseudopotential method [1] used reflectivity measurements to fit pseudopotential Fourier components, which were then used to calculate the one-electron band structure. With *ab initio* electronic structure methods, it is natural to attack the problem in reverse by using one-electron energies and wave functions to predict optical properties. Two levels of theory have contributed to this. The first is pseudopotential local-density approximation (LDA) calculations [2], where electronic band structures and wave functions are predicted. In the LDA, the band gap is underestimated and the optical absorption edge is too low in energy. This is largely corrected at the next level of theory, quasiparticle calculations [3], in which a better treatment of electron correlations results in quasiparticle band structures that agree well with experiments.

Unfortunately, these theories deal only with *single-particle* states, and do not describe a key aspect of the optical absorption process: the electron-hole interaction. When a photon is absorbed, an electron in the valence band is excited into the conduction band, leaving a hole behind. Single-particle theories regard the excited electron and hole as independent, noninteracting particles. In insulating solids, the electron-hole interaction leads to bound states (excitons) and concentrates oscillator strength near and below the band gap. This is especially important in wide-gap insulators, where the electron-hole interaction is very strong, due to insufficient screening. In these systems, excitonic effects can dominate the spectrum and wash out the features found in one-electron spectra.

Two decades ago, Hanke and Sham introduced a scheme in which the electron-hole interaction can be included [4]. It relies on a calculation of the particle-hole propagator, which describes the dynamics of a particle-hole pair in a quantum many-body system. It can be calculated from single-particle energies and wave functions, and the screened interaction. Hanke and Sham considered tetrahedrally coordinated covalent materials, and used tight-binding band energies and wave functions to construct the particle-hole propagator and resulting absorption spectrum. There is now a need for a first-principles scheme for computing optical absorption, one which may readily be applied to a *wide range* of materials with minimal empirical fitting. Only recently have calculations of the Hanke-Sham type been performed with single-particle energies and wave functions from LDA and/or quasiparticle calculations. One of us [5] has calculated x-ray absorption, when the hole is in a core state and can be considered fixed in position. Albrecht *et al.* [6] considered the more complex problem of the hole in a valence state, but no absorption spectrum was presented.

In this Letter, we present a computationally efficient first-principles scheme for calculating $\epsilon_2(\omega)$, including the electron-hole interaction in the Hanke-Sham framework. The input is energies and wave functions from LDA/quasiparticle calculations. The screened Coulomb interaction is calculated with a model dielectric function [7]. We consider two wide-gap insulators, where the electron-hole interaction is very strong: LiF and MgO. The calculated spectra are dramatically different from the one-electron results, even well above the absorption edge. The results are in excellent agreement with reflectivity experiments performed on these materials. In what follows, we describe the theoretical/computational model, and present the results.

We use the Tamm-Dancoff approximation (TDA) [8], wherein the ground state, $|0\rangle$, is a filled valence band, and all excited states are *singly excited,* with a hole below the Fermi level and an electron above. Because there are two particles singled out in these states (electron and hole), an excited state wave function depends on two particles' coordinates. We use two different bases for periodic systems: the *real space* basis, and the *one-electron eigenstate* basis. In the first, an excited state is of the form

$$\sum_{\mathbf{x},\mathbf{y},\mathbf{R}} \psi(\mathbf{x},\mathbf{y};\mathbf{R}) a^{\dagger}_{\mathbf{x}+\mathbf{R}} a_{\mathbf{y}} |0\rangle = \sum_{\mathbf{x},\mathbf{y},\mathbf{R}} \psi(\mathbf{x},\mathbf{y};\mathbf{R})\, |\mathbf{x},\mathbf{y};\mathbf{R}\rangle\,, \tag{1}$$

where $\mathbf{R}$ is the lattice vector separating the unit cells containing the electron and hole, and $\mathbf{x}$ and $\mathbf{y}$ are positions of the electron and hole in their unit cells. In the second, an excited state is

$$\sum_{i,j,\mathbf{k}} \psi(i,j;\mathbf{k}) a^{\dagger}_{i,\mathbf{k}+\mathbf{q}} a_{j,\mathbf{k}} |0\rangle = \sum_{i,j,\mathbf{k}} \psi(i,j;\mathbf{k})\, |i,j;\mathbf{k}\rangle\,, \tag{2}$$

where i and j are electron and hole band indices, $\mathbf{k}$ is the hole wave vector, and $\mathbf{k} + \mathbf{q}$ is the electron wave vector ($\hbar\mathbf{q}$ is the photon momentum). The transformation from the one-electron eigenstate basis to the real space basis is

$$\psi(\mathbf{x},\mathbf{y};\mathbf{R}) = \sum_{i,j} \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y}+\mathbf{R})} u_{i,\mathbf{k}+\mathbf{q}}(\mathbf{x}) \times u^{*}_{j,\mathbf{k}}(\mathbf{y}) \psi(i,j;\mathbf{k})\,, \tag{3}$$

where $u_{i,\mathbf{k}+\mathbf{q}}$ and $u_{j,\mathbf{k}}$ are the periodic parts of Bloch functions.

The equation of motion for the electron-hole pair creation operator is

$$\langle\Psi|\,[H, a^{\dagger}_{e} a_{h}]\,|\Psi_0\rangle = E\langle\Psi|a^{\dagger}_{e} a_{h}|\Psi_0\rangle\,, \tag{4}$$

where H is the exact many-body Hamiltonian, $|\Psi\rangle$ is an exact excited state with energy E, and $|\Psi_0\rangle$ is the exact ground state (with energy set to zero). The secular equation for the TDA effective Hamiltonian is obtained by replacing $|\Psi_0\rangle$ with $|0\rangle$, and $|\Psi\rangle$ with the singly excited states above [8]. The resulting equation has two parts: a one-electron part, and the electron-hole interaction, which has both exchange and direct parts. Following Ref. [4], the exchange terms involve the bare Coulomb interaction V, while the direct terms involve the statically screened interaction W [9]. In the one-electron eigenstate basis, we have

$$[E_i(\mathbf{k} + \mathbf{q}) - E_j(\mathbf{k}) - E]\psi(i,j;\mathbf{k}) = \sum_{i',j'} \int \frac{d^3\mathbf{k}'}{(2\pi)^3} [2\delta_{\sigma}\langle i,j;\mathbf{k}|V|i',j';\mathbf{k}'\rangle - \langle i,j;\mathbf{k}|W|i',j';\mathbf{k}'\rangle]\psi(i',j';\mathbf{k}')\,. \tag{5}$$

E_i and E_j are one-electron band energies; $\delta_{\sigma} = 0$ for triplet states, and 1 for singlet states. We consider singlets from now on. The exchange matrix element, $\langle i,j;\mathbf{k}|V|i',j';\mathbf{k}'\rangle$, is

$$\sum_{\mathbf{R}} \int d^3\mathbf{x} \int d^3\mathbf{y}\, e^{-i\mathbf{q}\cdot(\mathbf{x}-\mathbf{y}+\mathbf{R})} u^{*}_{i,\mathbf{k}+\mathbf{q}}(\mathbf{x}) u_{j,\mathbf{k}}(\mathbf{x}) \times V(\mathbf{x},\mathbf{y};\mathbf{R}) u_{i',\mathbf{k}'+\mathbf{q}}(\mathbf{y}) u^{*}_{j',\mathbf{k}'}(\mathbf{y})\,, \tag{6}$$

while the direct matrix element, $\langle i,j;\mathbf{k}|W|i',j';\mathbf{k}'\rangle$, is

$$\sum_{\mathbf{R}} \int d^3\mathbf{x} \int d^3\mathbf{y}\, e^{-i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{x}-\mathbf{y}+\mathbf{R})} u^{*}_{i,\mathbf{k}+\mathbf{q}}(\mathbf{x}) u_{j,\mathbf{k}}(\mathbf{y}) \times W(\mathbf{x},\mathbf{y};\mathbf{R}) u_{i',\mathbf{k}'+\mathbf{q}}(\mathbf{x}) u^{*}_{j',\mathbf{k}'}(\mathbf{y})\,. \tag{7}$$

Both terms describe the scattering of pair $(i,j;\mathbf{k})$ into pair $(i'j';\mathbf{k}')$. In the exchange integral, the electron and hole of each pair are at the same spatial point. Thus, it can be rewritten in terms of an effective one-electron potential, V_{eff}. In the direct integral, the electrons in the first and second pair are at the same point (as are the holes). The Levine-Louie-Hybertsen model [7] is used to calculate $W(\mathbf{x},\mathbf{y};\mathbf{R})$ from the ground-state charge density, $\rho(\mathbf{x})$, and the macroscopic dielectric constant, ϵ_{∞}. This model reproduces the correct short- and long-range behavior: $W(\mathbf{x},\mathbf{y};0) \to e^2/|\mathbf{x} - \mathbf{y}|$ as $\mathbf{x} \to \mathbf{y}$, and $W(\mathbf{x},\mathbf{y};\mathbf{R}) \to e^2/[\epsilon_{\infty}|\mathbf{x} - \mathbf{y} + \mathbf{R}|]$ as $\mathbf{R} \to \infty$.

Instead of solving Eq. (5) to obtain the energies, E, and states, $\psi(i,j;\mathbf{k})$, we perform all computations by repeatedly acting with the TDA Hamiltonian on singly excited states. The Hamiltonian has three terms, $H_{1\text{-}e}$, H_{ex}, and H_{dir}. Our strategy is to act with each term separately in the basis in which it is easily handled (the exchange term is treated by standard reciprocal space techniques [10]),

$$\begin{aligned} H_{\mathrm{TDA}}\psi &= H_{1\text{-}e} \sum_{i,j;\mathbf{k}} \psi(i,j;\mathbf{k})\, |i,j;\mathbf{k}\rangle + H_{\mathrm{ex}} \sum_{\mathbf{G}} \psi(\mathbf{G})\, |\mathbf{G}\rangle + H_{\mathrm{dir}} \sum_{\mathbf{x},\mathbf{y};\mathbf{R}} \psi(\mathbf{x},\mathbf{y};\mathbf{R})\, |\mathbf{x},\mathbf{y};\mathbf{R}\rangle \\ &= \sum_{i,j;\mathbf{k}} [E_i(\mathbf{k} + \mathbf{q}) - E_j(\mathbf{k})]\psi(i,j;\mathbf{k})\, |i,j;\mathbf{k}\rangle + 2\sum_{\mathbf{G}} V_{\mathrm{eff}}(\mathbf{G};\mathbf{q})\psi(\mathbf{G})\, |\mathbf{G}\rangle \\ &\quad - \sum_{\mathbf{x},\mathbf{y};\mathbf{R}} W(\mathbf{x},\mathbf{y};\mathbf{R})\psi(\mathbf{x},\mathbf{y};\mathbf{R})\, |\mathbf{x},\mathbf{y};\mathbf{R}\rangle\,. \end{aligned} \tag{8}$$

This is accomplished by first expressing ψ in the various bases using transformation laws, e.g., Eq. (3). Then, the Hamiltonian is applied with the above equation. Finally, the inverse transformations are applied to transform all terms back to one basis. Because $\mathbf{q}$ is small compared to the Brillouin-zone size, we set $\mathbf{k} + \mathbf{q} = \mathbf{k}$ in what follows.

Optical absorption is characterized by the imaginary part of the dielectric function, $\epsilon_2(\omega)$. Within the TDA, one has

$$\epsilon_2(\omega) = \frac{4\pi^2}{\omega^2} \sum_f |\langle 0|\hat{\lambda} \cdot \mathbf{J}|f\rangle|^2 \delta(\omega - E_f) = 4\pi \,\mathrm{Im}\left[\sum_{b,b'} \langle 0| \frac{\hat{\lambda} \cdot \mathbf{J}}{\omega_b} |b\rangle\langle b| \left(\frac{1}{H_{\mathrm{TDA}} - \omega - i\eta}\right) |b'\rangle\langle b'| \frac{\hat{\lambda} \cdot \mathbf{J}}{\omega_{b'}} |0\rangle\right], \tag{9}$$

where $\hat{\lambda} \cdot \mathbf{J}$ is the component of the macroscopic current operator along the polarization direction ($\hbar = 1$). The $|f\rangle$'s are eigenstates of H_{TDA}, while $|b\rangle$ and $|b'\rangle$ are members of the one-electron eigenstate basis (with one-electron energies ω_b and $\omega_{b'}$). The radiation field acts on $|0\rangle$ to create an electron-hole pair. The second line of Eq. (9) has the form of a resolvent element and is readily calculated with the Haydock recursion method [11], in which the partial density of states of H_{TDA} weighted by $\hat{\lambda} \cdot \mathbf{J}$ is calculated by repeated actions of H_{TDA}. Thus, calculation of $\epsilon_2(\omega)$ is reduced to the action of the Hamiltonian on state vectors. The advantage of this is that the number of actions needed depends chiefly on the energy broadening parameter, η, and is roughly independent of the dimension of the problem.

We obtain u functions from pseudopotential LDA calculations performed on a dense mesh of $\mathbf{k}$ points in the first Brillouin zone. An optimized basis set for large-$\mathbf{k}$-point calculations is used [12]. Band energies are adjusted to mimic the results of our quasiparticle calculations [13] when used in the first term of Eq. (8). We calculate the values of u functions on a mesh of $\mathbf{x}$ points in the unit cell, used in the transformation of Eq. (3). The mesh of lattice vectors, $\mathbf{R}$, is chosen to be compatible with the $\mathbf{k}$-point mesh so fast Fourier transform techniques can be employed in $\mathbf{k} \leftrightarrow \mathbf{R}$ transformations. The $u_{j,\mathbf{k}}(\mathbf{x})$ are also used to compute $\rho(\mathbf{x})$, needed in the determination of $W(\mathbf{x}, \mathbf{y}; \mathbf{R})$.

The time it takes to compute $H_{\mathrm{TDA}}\psi$ is dominated by the evaluation of the direct interaction term. There are three steps involved: (i) transformation from $\psi(i, j; \mathbf{k}) \rightarrow \psi(\mathbf{x}, \mathbf{y}; \mathbf{R})$ via Eq. (3), (ii) local multiplication $\phi(\mathbf{x}, \mathbf{y}; \mathbf{R}) = W(\mathbf{x}, \mathbf{y}; \mathbf{R})\psi(\mathbf{x}, \mathbf{y}; \mathbf{R})$, and (iii) transformation from $\phi(\mathbf{x}, \mathbf{y}; \mathbf{R}) \rightarrow \phi(i, j; \mathbf{k})$ via the inverse of Eq. (3). Each action of H_{TDA} scales as $N_c N_v N_{\mathbf{k}} \log(N_{\mathbf{k}}) N_{\mathbf{x}}^2$, where N_c and N_v are the numbers of conduction and valence bands, and $N_{\mathbf{k}}$ and $N_{\mathbf{x}}$ are the numbers of $\mathbf{k}$ and $\mathbf{x}$ points. Since the number of actions of H_{TDA} necessary to determine $\epsilon_2(\omega)$ is roughly independent of the dimension of H_{TDA}, the computation as a whole scales roughly as $N_c N_v N_{\mathbf{k}} \log(N_{\mathbf{k}}) N_{\mathbf{x}}^2$.

We now present results for LiF and MgO, two rocksalt-structure crystals. We use experimental lattice constants for both materials. For LiF, we include the F $2s$ and $2p$ bands in the valence, and the first four conduction bands. Valence bands are stretched by 15% and conduction bands by 2%, and the direct band gap is moved to 14.2 eV, bringing LDA bands into accord with quasiparticle calculations [14]. The dielectric constant (used in the determination of the screened interaction) is $\epsilon_\infty = 1.92$. A $\mathbf{k}$-point mesh of $6 \times 6 \times 6$ and an $\mathbf{x}$-point mesh of $3 \times 3 \times 3$ are sufficient to converge the calculation to compare with experiment. A broadening parameter of $\eta = 0.2$ eV is used, and 100 iterations (actions of H_{TDA}) are performed. The calculated $\epsilon_2(\omega)$ is presented in Fig. 1(a), together with results deduced from a reflectance experiment [15]. Note that the calculation shows significant features below the 14.2 eV band gap, the most significant of which is the huge peak at 12.4 eV. This is the Γ exciton, which is at 12.6 eV in the reflectance data. Analysis reveals that the peak is comprised of three states, each consisting primarily of a hole in the three nearly degenerate F $2p$ bands, and an electron in the first conduction band. The general agreement between theory and experiment is quite good, although there is a large discrepancy in the height of the peak at 22 eV. Note that neglect of the electron-hole interaction [where we use only the first term in Eq. (8)] yields an $\epsilon_2(\omega)$ which is drastically different at *all* energies. These one-electron results are shown in Fig. 1(b). Thus, excitonic effects are important even far above the absorption edge [4].

For MgO, we use the O $2s$ and $2p$ valence bands, and the first four conduction bands. Valence and conduction

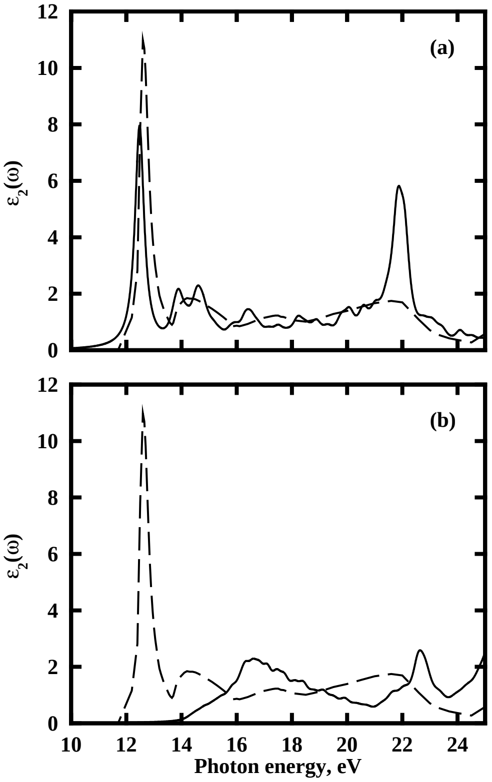

FIG. 1. (a) $\epsilon_2(\omega)$ vs $\hbar\omega$ for LiF. Solid line: calculation; dashed line: experimental data of Ref. [12]. (b) Solid line: $\epsilon_2(\omega)$ for LiF calculated with no electron-hole interaction; dashed line: experiment.

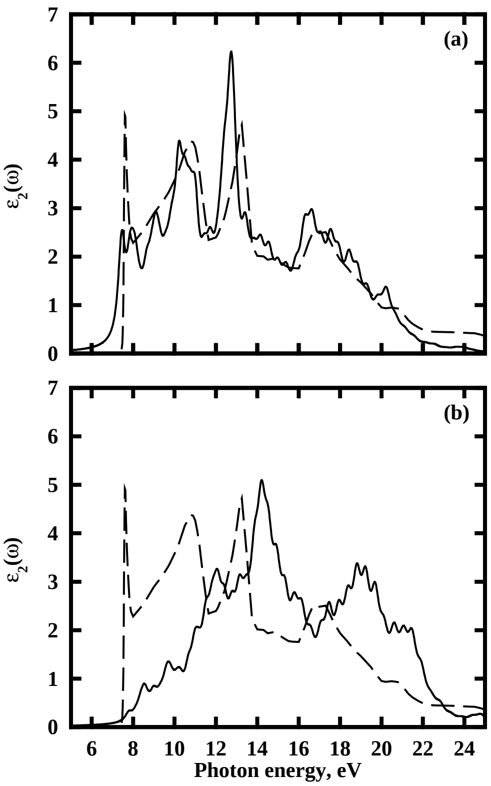

FIG. 2. (a) $\epsilon_2(\omega)$ vs $\hbar\omega$ for MgO. Solid line: calculation; dashed line: experimental data of Ref. [12]. (b) Solid line: $\epsilon_2(\omega)$ for MgO calculated with no electron-hole interaction; dashed line: experiment.

bands are stretched by 14% and 2%. The quasiparticle direct band gap is 7.8 eV [16], and ϵ_∞ is 3.03. A $10 \times 10 \times 10$ **k**-point mesh and a $3 \times 3 \times 3$ **x**-point mesh are used. Again, 100 iterations are performed, and η is chosen to be 0.2 eV. The calculated $\epsilon_2(\omega)$ is presented in Fig. 2(a), together with the results deduced from a reflectance measurement [15]. The agreement is very good, with peak positions well reproduced. Note that there is no significant absorption at energies well below the band gap; unlike for LiF, the exciton binding energy is small. This is due to both the smaller band gap and the larger bandwidths of conduction and valence bands. Nevertheless, there is a sharp onset at ~7.8 eV which is not present in the one-electron result, shown in Fig. 2(b). As for LiF, the electron-hole interaction plays an important role at all energies. Features in the MgO spectrum at high energies (~17 and ~20.5 eV) are reproduced well in the calculation, even though these energies are *greater than twice the band gap*. Thus, we suggest that multielectron excitations are not crucially important in understanding $\epsilon_2(\omega)$ for MgO.

We have presented a computationally efficient scheme for calculating $\epsilon_2(\omega)$ for insulators, including the electron-hole interaction. The only input is the material's crystal structure and dielectric constant. The model was applied to two materials where the electron-hole interaction is very strong. With this first-principles computational scheme, accurate absorption spectra can now be calculated for a wide range of materials, and optical properties of new materials can be predicted.

We are grateful to D. W. Lynch, D. M. Roessler, R. Haydock, and Z. H. Levine for helpful discussions. Finally, we thank S. G. Louie for stimulating us to consider this problem.

[1] See, for example, M. L. Cohen *et al.,* Phys. Rev. **155**, 992 (1966).
[2] M. L. Cohen, Phys. Today **33**, No. 7, 40 (1979).
[3] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[4] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1965); W. Hanke and L. J. Sham, Phys. Rev. Lett. **33**, 582 (1974); W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980).
[5] E. L. Shirley, Phys. Rev. Lett. **80**, 794 (1998).
[6] S. Albrecht, G. Onida, and L. Reining, Phys. Rev. B **55**, 10 278 (1997).
[7] Z. H. Levine and S. G. Louie, Phys. Rev. B **25**, 6310 (1982); M. S. Hybertsen and S. G. Louie, Phys. Rev. B **37**, 2733 (1988).
[8] See the discussion in A. Fetter and J. D. Walecka, *Quantum Theory of Many Particle Systems* (McGraw-Hill Book Company, San Francisco, 1971), pp. 538–539.
[9] In principle, the screening should be time dependent. We choose static screening because it is computationally convenient. This should be sufficient for frequencies much less than the plasma frequency ω_p (~25 eV here). However, we do present results for $\omega \sim \omega_p$. Further work must be done to study the effect of dynamical screening.
[10] We neglect the $\mathbf{G} = 0$ component of the bare Coulomb interaction, as it should have a negligible effect in a cubic system.
[11] R. Haydock, Comput. Phys. Commun. **20**, 11 (1980).
[12] E. L. Shirley, Phys. Rev. B **54**, 16 464 (1996).
[13] Ideally, quasiparticle wave functions should be used. However, it has been shown that LDA single-particle states are very similar to the quasiparticle states (at least for the cases tested). We use LDA wave functions for computational convenience.
[14] E. L. Shirley, L. J. Terminello, J. E. Klepeis, and F. J. Himpsel, Phys. Rev. B **53**, 10 296 (1996).
[15] E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, Orlando, 1985); *Handbook of Optical Constants of Solids II* (Academic Press, Boston, 1991); D. M. Roessler and W. C. Walker, J. Opt. Soc. Am. **57**, 835 (1967).
[16] U. Schonberger and F. Aryasetiawan, Phys. Rev. B **52**, 8788 (1995); E. L. Shirley (unpublished).