Polariton propagation in high quality semiconductors: Microscopic theory and experiment versus additional boundary conditions

H. C. Schneider, F. Jahnke, and S. W. Koch
Department of Physics and Materials Sciences Center, Philipps-University, 35032 Marburg, Germany

J. Tignon, T. Hasche, and D. S. Chemla
Department of Physics, University of California at Berkeley and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 24 May 2000; published 8 January 2001)

Linear exciton-polariton propagation in semiconductors is analyzed using a microscopic theory. Numerical results are compared with various approximation schemes based on additional boundary conditions, and with phase-amplitude linear spectroscopy experiments in high-quality GaAs. A simultaneous description of the measured amplitude and phase of the transmitted electric field is only possible with the full model.

DOI: 10.1103/PhysRevB.63.045202 PACS number(s): 71.35.Cc, 71.36.+c, 68.35.Ja

I. INTRODUCTION

Since the introduction of the polariton concept[1,2] for the interplay of a propagating light field with the polarization it induces in a semiconductor, this topic has received a great deal of experimental and theoretical interest. Even in the case of the linear optical properties considered in the present paper, the problem of light propagation through samples with surfaces presents considerable theoretical and experimental difficulties if the polarization in the sample exchanges momentum with the electromagnetic field, leading to spatial dispersion. For realistic conditions, the presence of surfaces raises serious complications because the Coulomb-bound electron-hole states (excitons) that show up as resonances in the optical response have a non-negligible spatial extension (typically, from 3 to 30 nm in the most common materials). Therefore, near the surfaces in regions much larger than the crystal primitive cell, the semiconductor response function deviates strongly from that of an infinitely extended medium. In fact, these deviations have a profound impact on the observed optical properties.[3]

Recently, there has been a renewed debate on how an approximate description of polariton propagation in a sample with surfaces can be obtained from the expression for the macroscopic polarization of an infinitely extended medium only.[4–6] Because of the microscopic structure of excitons, such an approach cannot determine how the macroscopic polarization falls off at the surfaces and, therefore, various assumptions have been proposed to describe the spatial behavior of the polarization field near these surfaces. Pekar[2] was the first to propose an approximate treatment by imposing an ''additional boundary condition'' (since then called the ABC in the literature): the macroscopic polarization should vanish on all surfaces. Pekar's ABC was subsequently augmented by other phenomenological restrictions, e.g., exciton-free dead layers at the surfaces and surface potentials.[7] Their influence was studied in Refs. 8 and 9, respectively. Several attempts were made to justify macroscopic approaches, leading to contradicting results: Zeyher and co-workers concluded that the excitonic polarization should vanish at the surfaces according to Pekar's ABC,[10,11] whereas Ting *et al.*,[12] using other arguments, suggested another ABC: the polarization flux should vanish at the surfaces. Finally, yet another method was presented recently.[4]

In contrast to using approximations for the polarization space dependence from the outset, microscopic treatments are free from the conceptual prolems of these ABC-like theories. However, the complexity of the polariton problem has so far limited microscopic treatments to specific geometries, or has necessitated the use of further simplifications: Direct calculations of the microscopic polarization coupled to the propagating field, were applied to half-space[13,14] and slab geometries[15] using a contact interaction instead of the Coulomb interaction between electron-hole pairs. A similar direct calculation was carried out for thick quantum wells using an expansion of the polarization in terms of quantum-well envelope functions.[16] Other approaches used an expansion into a finite number of excitonic wave functions in slab and half-space geometries to obtain the nonlocal semiconductor response function,[17–19] or analyzed the half-space problem in wave-vector space.[20]

From an experimental point of view, a large variety of techniques was devoted to the study of polariton effects in GaAs as well as other materials; see e.g., Refs. 7, and 21–27. But since the longitudinal-transverse splitting is so small in bulk GaAs, direct and unambiguous experimental evidence (e.g., experiments on a high-quality sample) of genuine polariton propagation effects in this material was still unsatisfactory, because features on the order of the longitudinal-transverse splitting could not be spectrally resolved. Even though the ''transmission configuration'' used here is the most common one in practice, and transmission experiments are trivial in principle, a considerable amount of care in processing the samples and conducting the experiments is required to obtain reliable spectra.

The purpose of this paper is a critical assessment of approximate treatments of the polariton problem. We first compare transmission spectra obtained with approximations to the full microscopic solution, and point out how the shortcomings of the approximations are related to the space dependence of the semiconductor polarization in the sample.

 ©2001 The American Physical Society

We then present a detailed comparison between the application of our microscopic and of the approximate treatments to recent experimental results.[28] We also show how the application of an approximation can lead to an ''explanation'' for experimental spectra in terms of wrong material parameters.

The paper is organized as follows. In Sec. II we formulate the theory of light propagation in semiconductor bulk materials in the linear regime, but with the presence of surfaces without referring to approximations using ABC's. In Sec. III the assumptions for various ABC schemes are discussed. A comparison of the full solution with these approximations and with experimental results is given in Secs. IV and V, respectively. The treatment of band-structure effects on the material polarization and details of the numerical method are presented in the Appendixes.

II. LINEAR LIGHT PROPAGATION IN SEMICONDUCTOR HETEROSTRUCTURES

A. Material polarization

When a light field propagates through a bulk semiconductor, it induces a macroscopic polarization $\vec{P}$ which is expressed in terms of the electron-hole transition amplitude Ψ and the dipole matrix element $\vec{d}$ by

$$\vec{P}(\vec{r})=\int d^3\rho \vec{d}^*(\vec{\rho})\Psi(\vec{r}_e,\vec{r}_h), \tag{1}$$

where $\vec{r}=\vec{r}_e+\vec{r}_h$ and $\vec{\rho}=\vec{r}_e-\vec{r}_h$ are the sum and difference of the electron and hole coordinates, respectively, and the integration runs over ρ. In the linear regime, the electron-hole transition amplitude obeys a two-particle Schrödinger equation,[3,29] which in the effective-mass approximation reads

$$\left[i\hbar\frac{\partial}{\partial t}-E_{\mathrm{Gap}}+\frac{\hbar^2}{2m_e}\nabla_e^2+\frac{\hbar^2}{2m_h}\nabla_h^2+V(|\vec{r}_e-\vec{r}_h|)\right]\Psi(\vec{r}_e,\vec{r}_h,t)=-\vec{d}(\rho)\cdot\vec{E}(\vec{r},t). \tag{2}$$

Here V is the Coulomb interaction between electrons and holes, and the driving term is due to the external electromagnetic field $\vec{E}$.

In this paper we consider a slab geometry which well describes most experimental conditions. The electromagnetic field is a plane wave propagating in the z direction, and the semiconductor sample is bounded by two surfaces perpendicular to the z axis, but extends homogeneously in the x-y plane. Under these conditions, it is useful to Fourier transform the in-plane relative coordinates $\vec{\rho}_\perp=\vec{r}_{e\perp}-\vec{r}_{h\perp}$ to introduce the in-plane momentum $\vec{k}_\perp$ according to

$$\Psi(\vec{k}_\perp;z_e,z_h)=\int d^2\rho_\perp e^{i\vec{k}_\perp\cdot\vec{\rho}_\perp}\Psi(\vec{r}_{e\perp},z_e,\vec{r}_{h\perp},z_h). \tag{3}$$

For rotational invariance around the direction of propagation Ψ depends only on the modulus of the wave vector in the x-y plane denoted by $k_\perp$. The essential nonlocality of the electron-hole transition amplitude $\Psi(k_\perp;z_e,z_h)$ is expressed by the dependence on electron and hole z coordinates, it obeys the two-particle Schrödinger equation

$$\begin{aligned}\hbar\frac{\partial}{\partial t}\Psi(k_\perp;z_e,z_h)=&\left[E_{\mathrm{Gap}}+\Delta E+i\gamma+\frac{\hbar^2}{2\mu_\perp}k_\perp^2+\frac{\hbar^2}{2m_{ez}}\frac{\partial^2}{\partial z_e^2}+\frac{\hbar^2}{2m_{hz}}\frac{\partial^2}{\partial z_h^2}\right]\Psi(k_\perp;z_e,z_h)\\&-\frac{1}{(2\pi)^2}\int_0^\infty dk_\perp' k_\perp' V(k_\perp,k_\perp',|z_e-z_h|)\\&\times\Psi(k_\perp',z_e,z_h)-d_{\mathrm{cv}}E(z,t)\delta(z_e-z_h).\end{aligned} \tag{4}$$

In Eq. (4) we account for the different effective masses in the x-y plane in the z direction, and for the energy shift ΔE from the band-gap energy E_{Gap}. We use the standard approximation of a momentum-independent dipole matrix element which corresponds to a delta-like space dependence of the dipole matrix element in the real-space formulation, $d(z_e-z_h)=d_{\mathrm{cv}}\delta(z_e-z_h)$. In this approximation, dipole transitions involve only electrons and holes at the same space point. The only phenomenological parameter introduced in this equation is the damping constant γ. The mass anisotropy and the energy shift have their origin in band-structure effects due to strain and valence-band mixing, which are discussed in Appendix A within the framework of the Kohn-Luttinger theory. The Coulomb matrix element is

$$V(k_\perp,k_\perp';|z_e-z_h|)=\frac{e^2}{2\varepsilon_0 n_{\mathrm{bg}}^2}\int_0^{2\pi}d\phi_\perp\frac{e^{-|\vec{k}_\perp-\vec{k}_\perp'||z_e-z_h|}}{|\vec{k}_\perp-\vec{k}_\perp'|}, \tag{5}$$

where $n_{\mathrm{bg}}=\sqrt{\varepsilon}_{\mathrm{bg}}$ is the background refractive index, $\phi_\perp$ is the angle between $\vec{k}_\perp$ and $\vec{k}_\perp'$, and $|\vec{k}_\perp-\vec{k}_\perp'|=\sqrt{k_\perp^2+k_\perp'^2-2k_\perp k_\perp'\cos\phi_\perp}$. Note that in Eq. (4) a z-dependent band-gap energy can be introduced which allows the modeling of heterostructures.

B. Electromagnetic field

The evolution of the electromagnetic field is determined by Maxwell's equations which, for circularly polarized transverse fields propagating in the z direction, $\vec{E}(\vec{r},t)=E(z,t)\vec{\sigma}_+$ and $\vec{B}(\vec{r},t)=B(z,t)\vec{\sigma}_+$, are

$$n^2(z)\frac{\partial}{\partial t}E(z,t)=-c_0^2\frac{\partial}{\partial z}B(z,t)-\frac{1}{\varepsilon_0}\frac{\partial}{\partial t}P(z,t), \tag{6}$$

$$\frac{\partial}{\partial t}B(z,t)=-\frac{\partial}{\partial z}E(z,t) \tag{7}$$

where $n(z)$ is the nonresonant (background) refractive index profile along the propagation direction. Within a homogeneous slab, $n(z)=n_{\mathrm{bg}}$. The electromagnetic field is coupled through the source term in Eq. (6) to the macroscopic polar-

ization $\vec{P}(\vec{r},t)=P(z,t)\vec{\sigma}_+$. The latter can be expressed in terms of the transition amplitude as

$$P(z,t)=2\pi\int_0^\infty dk_\perp k_\perp d_{cv}^*\Psi(k_\perp;z,z,t) \tag{8}$$

with the help of Eqs. (1) and (3). In Eq. (8) we have again used a local dipole matrix element.

III. COMPARISON TO MACROSCOPIC APPROACHES

In this section we discuss various approximate schemes that used the propagating solutions for *homogeneous* media to describe nonlocal (spatially inhomogeneous) situations. For reference, in Appendix B we briefly recall the dielectric theory of polariton propagation in *homogeneous* media with spatial dispersion.

For a finite-size sample with surfaces in the propagation direction, the system becomes spatially inhomogeneous, and specific boundary conditions for the interband transition amplitude apply. In the case of the slab geometry these are

$$\Psi(\vec{r}_e,\vec{r}_h)=0 \text{ for } \vec{r}_e \text{ or } \vec{r}_h \text{ at a surface,} \tag{9}$$

i.e., the transition amplitude vanishes whenever the electron or the hole coordinate is on one of the surfaces (or outside of the sample). This boundary condition introduces an entanglement between the electron-hole relative motion and the motion of their center of mass (COM), which is not present for an infinitely extended medium where the only requirement is that Ψ vanishes if the electron-hole distance approaches infinity.

A direct computation of the transition amplitude from the two-particle Schrödinger equation, as discussed in Sec. II, together with boundary condition (9), is numerically very demanding. Thus for more than four decades many approaches have been proposed based on an approximate separation of the electron-hole relative and COM motion. Usually one introduces expansion wave functions for the excitonic relative and COM wave functions, denoted by φ_i and Φ_q, respectively, as outlined in Appendix B for a homogeneous medium. *However, in the inhomogeneous case, one cannot deduce boundary conditions for the COM and relative wave functions from* Eq. (9). An approximate boundary condition is imposed by specifying the COM wave functions for the COM coordinate $\vec{r}$ at the surface:

$$H_{com}\Phi_q(\vec{r})=-\frac{\hbar^2}{2M}\nabla_r^2\Phi_q(\vec{r})=\hbar\omega_q\Phi_q(\vec{r}), \tag{10}$$

$$\Phi_q(\vec{r})|_{\vec{r}\text{ at surface}}=\Phi_{surf}. \tag{11}$$

The quantum number q attached to the COM wave function can stand for the continuous COM momentum q in semi-infinite slab geometries, or for the discrete quantum number n, which labels quantized COM states in bounded geometries. A constant Φ_{surf} must now be specified for each surface of the geometry under consideration. The electron-hole relative motion, on the other hand, is approximated by eigenfunctions φ_i of the Wannier equation (B2), which is strictly valid only for homogeneous systems. Clearly, this scheme cannot account for the distortion of the relative motion at the boundaries. Furthermore, different choices for Φ_{surf}, i.e., different macroscopic boundary conditions for the COM problem, lead to qualitatively different macroscopic approximations because the spatial behavior of the macroscopic polarization near the surfaces critically influences the propagating solutions even in thick samples.

For comparison with our full calculation, we now outline briefly approximate treatments based on Eqs. (10) and (11). Taking into account only the interaction with the $1s$-exciton resonance, one uses the two complex polariton wave vectors $q_{1,2}(\omega)$ from Eq. (B12), which describe the propagation of an optical field in the infinite system, as an ansatz for the propagating electromagnetic field inside a finite semiconductor sample:

$$E_{mat}(z,\omega)=\sum_{p=1,2}E_p^+(\omega)e^{iq_p(\omega)z}+E_p^-(\omega)e^{-iq_p(\omega)z}. \tag{12}$$

Here, propagating (E^+) and counter-propagating waves (E^-) are considered for the slab geometry. Then Maxwell's boundary conditions are applied to connect the polariton waves inside the slab (12) to free solutions of Maxwell's equations outside the sample. The simplest possible geometry involves an incident wave from the left and the reflected and transmitted components, E_r and E_t, with wave vectors $q_{left}=n_{left}\omega/c$ and $q_{right}=n_{right}\omega/c$,

$$E_{left}(\omega,z)=e^{iq_{left}z}+E_r(\omega)e^{-iq_{left}z}, \tag{13}$$

$$E_{right}(\omega,z)=E_t(\omega)e^{iq_{right}z}. \tag{14}$$

The continuity of E and $\partial E/\partial z$ on both boundaries determines only four of the six unknowns $E_r(\omega)$, $E_t(\omega)$, $E_1^\pm(\omega)$, $E_2^\pm(\omega)$. The remaining two conditions are obtained by using macroscopic boundary conditions for the excitonic COM wave functions, [Eq. (11)]. This procedure defines an ABC for the macroscopic polarization. For instance, the choice $\Phi_{surf}=0$ at the boundary leads to the ABC originally introduced by Pekar[2] (also see Ref. 10),

$$P(z,\omega)|_{z\text{ at boundary}}=0, \tag{15}$$

whereas the choice $\Phi_{surf}=1$ results in the Ting-Frankel-Birman (TFB) ABC[12]

$$\left[\frac{\partial}{\partial z}P(z,\omega)\right]_{z\text{ at boundary}}=0, \tag{16}$$

as can be seen with Eq. (B7). In both cases the macroscopic polarization is given by

$$P(z,\omega)=\sum_{p=1,2}\chi(q,\omega)|_{q=q_p(\omega)}[E_p^+(\omega)e^{iq_p(\omega)z}+E_p^-(\omega)e^{-iq_p(\omega)z}], \tag{17}$$

which can be used to determine the remaining two coefficients.

In a recent paper,[4] a macroscopic treatment of the polariton problem that is different from the Pekar and TFB analysis was suggested. The macroscopic dielectric function of an infinitely extended medium is combined with an analysis of surface effects to describe light propagation with spatial dispersion in finite samples. The underlying assumption is that the deviations from the bulk properties are important only in a thin layer near the surfaces. Then the influence of surfaces is described simply by source layers with strength s_i in the wave equation for the infinitely extended medium, written in momentum space as

$$\left[\frac{\omega^2}{c_0^2}\varepsilon(q,\omega)-q^2\right]E(q,\omega)=\sum_i s_i(q,\omega). \tag{18}$$

The solutions of Eq. (18) for sources located at the position of the sample boundaries are taken as the electromagnetic field inside of the finite sample. By matching the material solution to the outside solution, the source terms can be determined. The validity of this matching procedure has to be checked carefully, because it connects solutions for infinitely extended media at the surface though these solutions are valid only away from the surface. Therefore, matching the two solutions at a boundary is a good approximation only if the difference between the outside solution E_0 and material solution E_{mat} is small. Following the original analysis,[4] both fields are expanded up to second order in the distance from the surface. For a wave propagating in z direction through a boundary at $z=0$, one obtains

$$E_0(z)-E_{\mathrm{mat}}(z)=\left[\frac{\partial^2}{\partial z^2}E_0\bigg|_{z=0}-\frac{\partial^2}{\partial z^2}E_{\mathrm{mat}}\bigg|_{z=0}\right]\frac{z^2}{2}+O(z^3)$$

$$\approx[(q_1^2-q_0^2)E_1^+ +(q_2^2-q_0^2)E_2^+]\frac{z^2}{2}, \tag{19}$$

where the continuity of the fields and their first derivatives have been used. Here q_0 is the wave number of the external field, and $q_{1,2}$ are the complex polariton solutions for the $1s$-exciton resonance; cf. Eq. (B12). At resonances of the optical susceptibility the modulus of $q_{1,2}^2$ becomes large and it is not clear what thickness of the transition layer should be used in Eq. (19). Hence it is difficult to show that the last line in Eq. (19) is indeed small enough.

IV. NUMERICAL RESULTS: FULL SOLUTION VS MACROSCOPIC APPROXIMATIONS

In this section we compare the approximate treatments mentioned in Sec. III with the full solution of the propagation problem for samples with surfaces. We focus on samples of intermediate thickness where the geometrical confinement of the semiconductor polarization as well as the genuine polariton propagation effects over distances of the order of a wavelength are present. The combination of these mechanisms leads to several interesting features in the optical transmission spectra.

To simplify the comparison of various theoretical results in this section we use a two-band semiconductor model

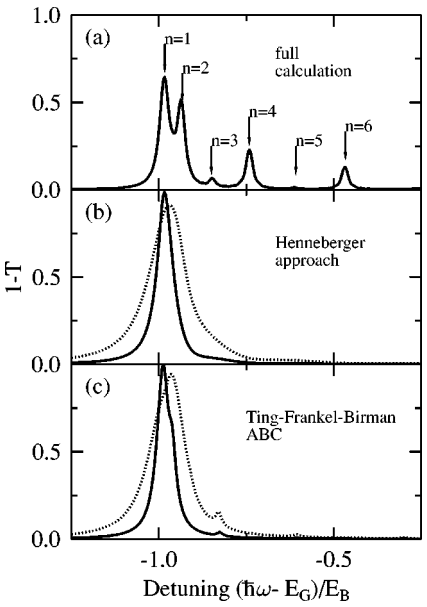

FIG. 1. Calculated transmission T for a sample length $L=10a_B$, using the full calculation (a), the Henneberger approach (b), and the ABC of Ting, Frankel, and Birman (c). The detuning is relative to the bulk semiconductor band-gap energy E_G in units of the three-dimensional exciton Rydberg energy E_B. Solid lines correspond to $r_{\mathrm{cv}}=5$ Å, and dotted lines to $r_{\mathrm{cv}}=10$ Å.

whereas for the experiment-theory comparison in Sec. V a more complicated band structure will be considered. We apply standard GaAs parameters: electron and hole masses $m_e=0.067m_0$ and $m_h=0.457m_0$, exciton Bohr radius $a_B=125$ Å, exciton binding energy $E_B=4.2$ meV, and background refractive index $n_{\mathrm{bg}}=3.71$. The microscopic dipole strength is taken to be $r_{\mathrm{cv}}=0.5$ nm, corresponding to the LT splitting $\Delta_{\mathrm{LT}}=|d_{\mathrm{cv}}|^2/(\pi n_{\mathrm{bg}}^2\varepsilon_0 a_B^3)=0.06$ meV. The only phenomenological parameter in our theory is the polarization dephasing rate γ defined in Eq. (4). In the following calculations we use decay time $T=1/\gamma=15.7$ ps which corresponds to a broadening $\gamma=0.04$ meV. To concentrate on the effects due to the excitonic polarization we use the same background refractive index inside and outside of the sample. Thus there is no semiconductor air interface that could lead to Fabry-Pérot resonances in the optical transmission. This model describes the experimental situation realized by a semiconductor heterostructure with an applied antireflection coating. In such a configuration the optical field resonantly interacts with the active semiconductor material but not with the buffer layers, which have a similar refractive index but a larger band gap.

Approximate solutions based on ABC models and Henneberger's approach are calculated only for the contribution of the $1s$-exciton resonance as usually done in the literature, although the extension to other bound exciton states is possible. The full solution always includes both exciton bound and continuum states.

The three panels of Fig. 1 show transmission spectra for a sample thickness $L=10a_B$ calculated using the microscopic theory [Fig. 1(a)], using the Henneberger approach[4] [Fig. 1(b)], and the Ting-Frankel-Birman (TFB) ABC [Fig. 1(c)].

In the spectrum of Fig. 1(a) the $1s$-excitonic resonance exhibits a clear splitting, labeled $n=1$ and 2, as well as additional peaks at higher energies ($n \geqslant 3$). To give a rough interpretation of this multiple-peak structure, we consider the quantization of the exciton COM motion.[23,19] Neglecting the coupling between the relative and COM exciton motion, the eigenvalue equation (10) together with boundary condition (11) and $\Phi_{\mathrm{surf}}=0$ for a finite sample length L in the propagation direction leads to discrete energies for the COM motion:

$$E_n=\frac{\hbar^2}{2M}\left(\frac{\pi}{L}\right)^2 n^2. \tag{20}$$

Here $M=m_e+m_h$ is the excitonic mass, and the eigenvalues are labeled by the discrete index n. Though the approximate condition (20) cannot consistently explain all the features we find in Fig. 1(a), a fit for the COM resonances with quantum numbers $n=4$, 5, and 6 using Eq. (20) gives a length $L \simeq 9a_B$. This effective length is smaller than the actual sample thickness as a result of the ''dead layer'' effect:[7] Due to the finite extension of the exciton relative motion, the macroscopic polarization remains small in regions near the surface and in the ''dead layer'' picture the exciton COM motion is confined in a reduced slab length. Note that the odd numbered peaks have a much weaker oscillator strength, for instance, $n=5$ at $\hbar\omega=-0.61E_B$ is almost indiscernible. (Whether even or odd peaks dominate in the optical spectra generally depends on the sample length and the exciton momentum; see Ref. 23.) The structure of the double peak with the largest oscillator strength, on the other hand, is dominated by polariton effects: Using Eq. (20) to fit the double peak leads to $L \simeq 25a_B$ which clearly illustrates that the concept of excitons with mass M confined to an effective length is not applicable near the main resonance.

Henneberger's approach [Fig. 1(b)], does not account for the additional maxima above the main exciton resonance. Instead of a multiple peak structure a single asymmetric line is predicted. The ABC of Ting *et al.* [Fig. 1(c)], leads to only one very small replica and no satisfactory agreement with the full calculation concerning the line shape of the main resonance. Furthermore, it is not possible to fit the results of Henneberger or Ting *et al.* to the full calculation by artificially increasing the dipole coupling. This is shown by the dashed lines in Fig. 1, where $r_{\mathrm{cv}}=1$ nm has been used.

Figure 2 shows the corresponding transmission spectra calculated with Pekar's ABC. The sample length is varied to account for an exciton-free dead layer. Transmission spectrum (a) is obtained without assuming a dead layer, i.e., using Pekar's ABC in their original form with $L=10a_B$. This gives a qualitatively different picture compared to the full calculation because neither the double-peak structure of the main resonance nor the energetic position of the higher replicas are reproduced. For Fig. 2(b) the effective length $L=9a_B$ extracted from the full calculation by applying Eq. (20) has been used. Then the high-energy replicas are at the correct energetic position, but the agreement for the line shape of the main peak remains unsatisfactory. Figure 2(c) is a fit aimed at reproducing the typical double-peak structure

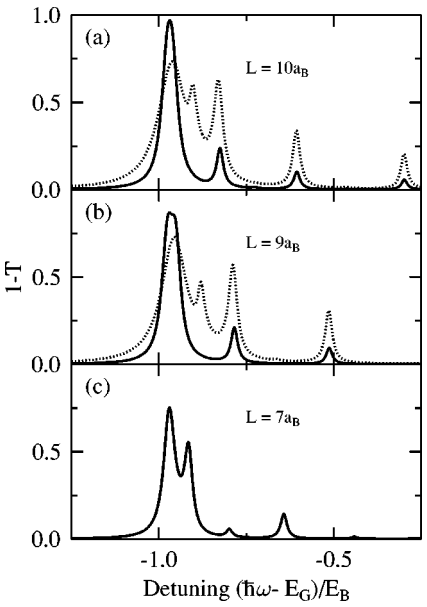

FIG. 2. Calculated transmission T using Pekar's ABC with effective sample lengths $10a_B$ (a), $9a_B$ (b), and $7a_B$ (c) for the same parameters as in Fig. 1. Solid lines correspond to $r_{\mathrm{cv}}=5$ Å, and dotted lines to $r_{\mathrm{cv}}=10$ Å.

of the main resonance as well as possible. This fitting leads to an effective length $L=7a_B$, and gives a satisfactory agreement for the main resonance lineshape, but then the high-energy part of the spectrum is not correctly described because the COM replicas are now shifted. In summary, when the effective sample length is regarded as a fitting parameter, Pekar's ABC allows qualitatively better agreement with the full calculation than the Henneberger or TFB approach. It is impossible, however, to obtain a good overall agreement by fixing the effective sample length, since this has several opposing consequences: optimizing the line shape of the main peak, which is dominated by interferences of the propagating polariton, leads to less satisfactory results for the COM quantization energies of the replicas. Note that it is also impossible to improve the agreement by artificially increasing the dipole coupling, as shown by the dotted lines in Fig. 2.

The full calculation of the propagation problem combined with the solution of two-particle Schrödinger equation for the electron-hole motion in a finite sample geometry can also be used to study the transition from bulk material to quantum wells where eventually a series of subbands emerges. Figure 3(a) shows results of the full calculation for the same parameters as above and decreasing sample lengths of $L=10a_B$, $3a_B$, and $2a_B$. We obtain an increasing energetic shift of the whole spectrum due to the confinement of the electron and hole motion. Furthermore the spacing of the exciton replica at higher energies increases. For $L=3a_B$ and $2a_B$ the light-matter coupling results in a broadening of the exciton line. Figure 3(b) shows the corresponding spectra obtained with Pekar's ABC. In this approximation the energy shifts are caused only by the COM quantization of ideal three-dimensional excitons formed from electron and hole bands unaffected by a geometrical confinement. Thus for smaller

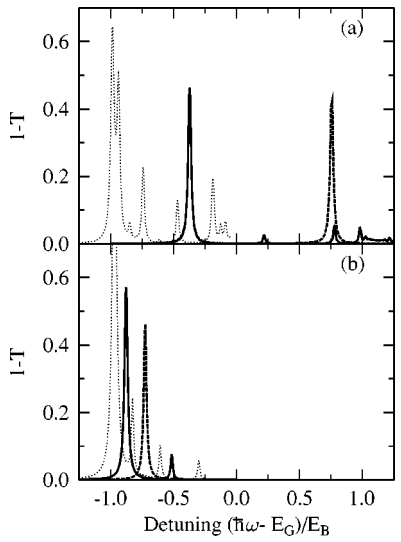

FIG. 3. Calculated transmission T for various sample lengths using the full calculation (a) and Pekar's ABC (b): $L=10a_B$ (thin line), $L=3a_B$ (solid line), and $L=2a_B$ (dash-dotted line). The dipole coupling is $r_{cv}=5$ Å.

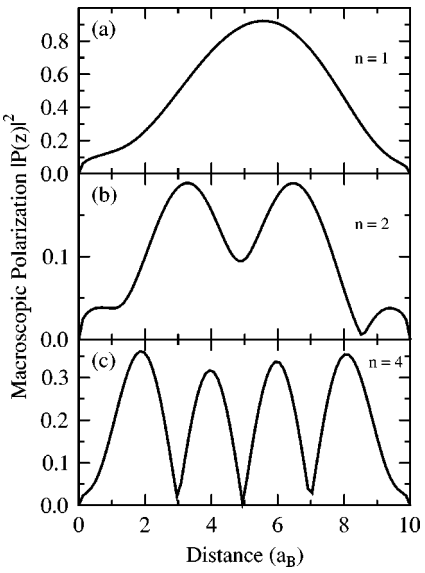

FIG. 4. Spatial distribution of the macroscopic polarization for stationary, monochromatic excitation. In (a) the excitation frequency is tuned to the absorption peak $n=1$, in (b) to $n=2$, and in (c) to $n=4$, where the peaks are marked in Fig. 1(a).

samples Pekar's ABC gives approximately right trends for the line shapes, but fail to reproduce the energy shifts due to carrier confinement.

The series of spectra shown in Fig. 3 also reveals the nature of the splitting of the exciton resonance for $L=10a_B$, and especially that of the $n=2$ peak of Fig. 1(a). The position of this peak cannot be explained as a replica due to COM quantization if one treats the COM motion independently of the exciton relative motion. Moreover, the $n=2$ peak shifts away from the exciton resonance to higher energies when the length is decreased to $L=3a_B$, $\hbar\omega-E_G=0.23E_B$. (The corresponding peak for $L=2a_B$, found at $\hbar\omega-E_G=1.9E_B$, is not shown in Fig. 3.) Hence, for $L=10a_B$, the $n=2$ resonance is due to the *combined influence* of the COM quantization and the energy dressing of the coupled exciton-photon states, which is described for the spatially homogeneous system in Appendix B in terms of polariton states. Clearly, for $L=3a_B$ and $L=2a_B$ the COM quantization energies are already so large that the $n=2$ replica does not interfere with the remaining radiatively broadened exciton resonance.

So far we have shown that calculated transmission spectra are strongly influenced by approximations affecting the *spatial distribution of the macroscopic polarization* which determines the source in Maxwell's equations. On the other hand, the full solution of the propagation problem, based on Eq. (4), can be used to directly analyze this space dependence of the macroscopic polarization, given by Eq. (8). For stationary monochromatic driving fields, the macroscopic polarization exhibits a very distinct spatial distribution depending on the excitation frequency. Figure 4 shows the distribution of $|P(z)|$ inside a sample of length $L=10a_B$ for stationary monochromatic excitation at the (a) $n=1$, (b) $n=2$, and (c) $n=4$ peaks of Fig. 1(a). This spatial distribution further supports the aforementioned relation between the transmission peaks and the quantization of the COM motion. For example, if one only considers the COM motion, i.e., Eq. (10), with the approximative boundary condition [Eq. (11)] and $\Phi_{surf}=0$, the eigenfunctions $\Phi_n(z)$ are sinusoidal standing waves with n maxima. Since the approximative boundary condition neglects effects due to electron-hole relative motion near the surface, the deviations of the full solution (Fig. 4), from standing waves show the influence of the surface on the macroscopic polarization. Furthermore, we can now investigate the validity of the concept of a uniform polarization-free region near the surface: We take as the extension of this polarization-free dead layer the distance over which $|P(z)|$ is strongly reduced close to the surface. Clearly, this dead-layer thickness depends on the excitation frequency. For the $n=1$ and 2 peaks in Fig. 1(a), these regions at both surfaces are approximately $l\approx 1.5a_B$ thick, which determines an effective sample length $L_{eff}=L-2l=7a_B$, in agreement with the results of the fitting procedure using approximation (20) to reproduce the transmission spectrum in Fig. 2(c). In Fig. 4(c), which corresponds to the $n=4$ peak, the ''dead layer'' is reduced to $l=0.5a_B$ also in good agreement with length $L_{eff}=L-2l=9a_B$ used to fit the corresponding peak position in Fig. 2(b).

Some approaches, in which a dead-layer thickness is computed, predict that the quantity $l/(ia_B)$, where i denotes the principal exciton quantum number, is *constant* for a given material. (For an overview, see Ref. 3.) Using these approaches, one finds numerical values for the constant dead-layer depths in the range 0.5–2,[30] comparable to our results, as well as in the range of 2–3.[3] However, using the full $|P(z)|$ distribution excited in the sample by stationary excitation is a direct proof that even COM replicas belonging to the same 1*s*-resonance exhibit different dead layers.[15]

For completeness, in Fig. 5 we show the spatial distribution of the macroscopic polarization for the same configuration as in Fig. 4, but for excitation energies away from the

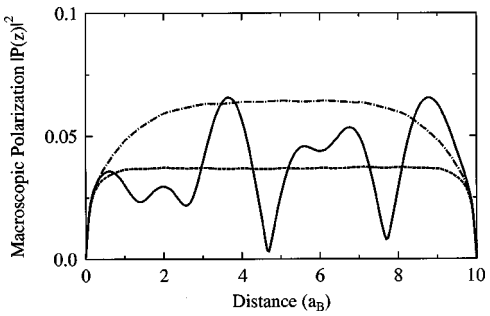

FIG. 5. Spatial distribution of the macroscopic polarization for stationary optical excitation below ($\hbar\omega - E_G = -2.5E_B$; dashed line) and near ($\hbar\omega - E_G = -1.25E_B$; dash-dotted line) the main absorption peak, as well as above the $n=4$ peak ($\hbar\omega - E_G = -0.6E_B$; solid line), corresponding to Fig. 1(a).

resonances. In all cases we find a rapid growth of the polarization directly from the surface into the sample without dead layer. Finally, Fig. 6 gives the $|P(z)|$ in the $L=3a_B$ sample, whose linear spectrum is shown as solid line in Fig. 3. Only small regions of reduced polarization at the surfaces are found for the two resonances at $\hbar\omega - E_G = -0.38E_B$ and $+0.22E_B$, whereas below the resonances ($-0.9E_B$) these regions are completely absent.

V. EXPERIMENTAL RESULTS

A. Sample

The measurements were performed on a very high quality 0.25-μm-thick GaAs layer, clad between a $Al_{0.3}Ga_{0.7}As$ cap layer and a $Al_{0.7}Ga_{0.3}As$ stop-etch layer. In order to allow transmission experiments, the GaAs substrate has been etched and the sample antireflection (AR) coated on both sides before being mounted on a sapphire window. The sample was immersed in superfluid He and the transmission was obtained by exciting the system with a 100-fs Ti:sapphire laser. The absorption spectrum, at $T=2$ K and very low carrier density, $n_{eh} \leqslant 10^{11}$ cm^{-3} is shown in Fig. 7(a), along with the laser profile (dotted line). This absorption spectrum is similar to that reported recently in the literature[31]: the light-hole–heavy-hole (LH-HH) exciton degeneracy is lifted due to the strain induced by the difference in thermal expansion between the sapphire substrate and the sample. Additionally, the transition linewidths are very small (below 0.25 meV), showing the high quality of the sample. Even with this careful sample preparation we found that the sapphire slab was responsible for the small Fabry-Pérot fringes, which appear close to the base-line of the spectrum in Fig. 7(a). Therefore, in a second preparation step, the sapphire substrate was also AR coated. In order to obtain a satisfactory suppression of the residual Fabry-Pérot fringes an excellent matching between the refractive indices of the AR layer and sapphire is necessary, this condition is only met by the hydrophobic NaF_2 coating material. Thus, a new coating had to be reprocessed for each new set of measurements. For comparison, Fig. 7(b) shows the absorption spectrum obtained using this second coating in the same excitation conditions as in Fig. 7(a). Now the small but important features marked by arrows are resolved unambiguously.

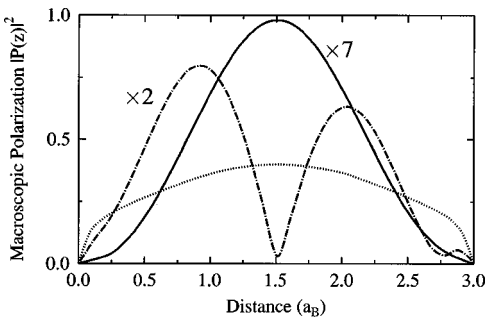

FIG. 6. Spatial distribution of the macroscopic polarization for a sample length $L=3a_B$. The stationary optical driving field is tuned below the main resonance ($\hbar\omega - E_G = -0.9E_B$; dotted line), and to the first ($\hbar\omega - E_G = -0.38E_B$; solid line), and the second ($\hbar\omega - E_G = 0.22E_B$; dash-dotted line) peak in the transmission spectrum is shown by the solid line in Fig. 3(a).

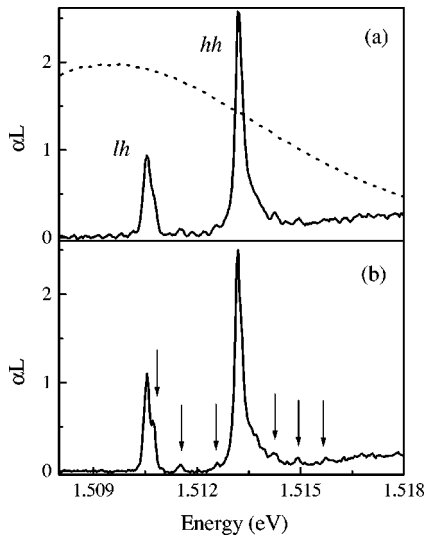

FIG. 7. Absorption spectra before (a) and after (b) AR coating, together with the excitation pulse spectrum (dotted line). The features that emerge in the spectrum after AR coating are marked by arrows.

B. Experimental setup

The amplitude-phase transmission measurements are performed by Fourier-transform spectral interferometry in a Mach-Zehnder interferometer[32] (Fig. 8). The Ti:sapphire laser used in the experiments delivers transform limited 100-fs pulses. The phase of these pulses was carefully characterized by second-harmonic generation FROG,[33] and the laser was adjusted until that phase was essentially flat across the whole spectrum, varying by less than $4.5\times10^{-2}\pi$ between the 2% points of the power spectrum. The output of the laser is split into two replicas using a first beam splitter. One beam is sent through the sample, and the other is used as a reference beam. The beam transmitted through the sample and the reference beam are colinearly recombined using a second beam splitter and sent to a 75-cm spectrometer (1200 gr/mm) and a

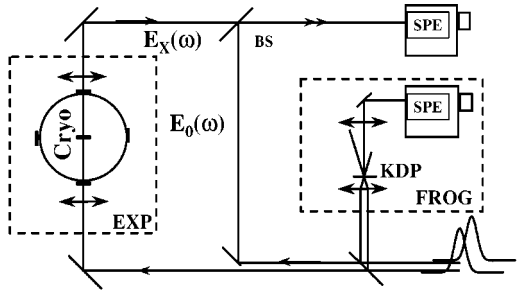

FIG. 8. Experimental setup.

CCD camera. The overall system resolution is 0.06 meV. The intensity of the measured interferogram spectrum is

$$I(\omega) = I_0(\omega) + I_X(\omega) + 2\sqrt{I_0(\omega) I_X(\omega)} \cos[\phi_0(\omega) - \phi_X(\omega) + \omega\Delta\tau], \quad (21)$$

where $I_0(\omega)$, $I_X(\omega)$, $\phi_0(\omega)$, and $\phi_X(\omega)$ are the intensity and phase of the reference beam and the probe beam, respectively. The time delay between the probe beam and the reference beam, $\Delta\tau$, is adjusted using a delay line. The interferometric measurements are performed differentially by measuring the amplitude and the phase of the pulses going through the optical system, first with, then without, the sample. This procedure allows one to eliminate any phase contribution due to the small difference between the optical dispersions of the two arms in the experimental setup. A typical interferogram $I(\omega)$ is shown in Fig. 9(b). The power spectrum of the reference beam pulse [Fig. 9(a), dotted line] and the transmitted power spectrum [Fig. 9(a), solid line], are then subtracted from the interferogram. The result $S(\omega)$ is plotted in panel (c) of Fig. 9. The amplitude and phase (relative to the reference pulse) of the transmitted pulse are then extracted from $S(\omega)$ using Fourier analysis.[32]

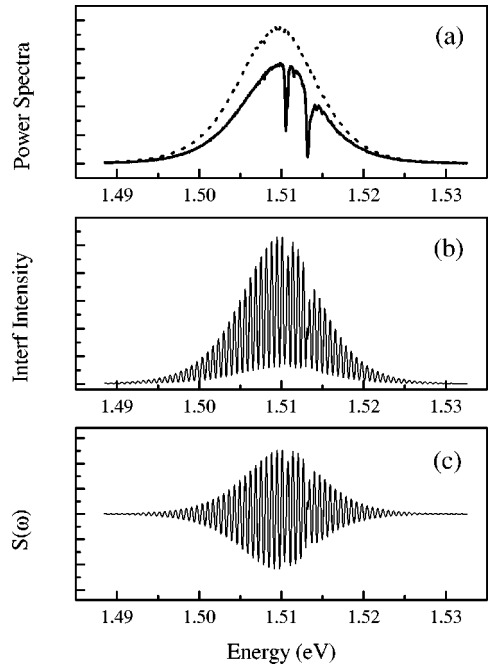

FIG. 9. (a) Power spectra of the incident (dotted line) and transmitted (solid line) field. (b) Intensity pattern of the interferogram. (c) Difference between the transmitted power spectrum and interferogram.

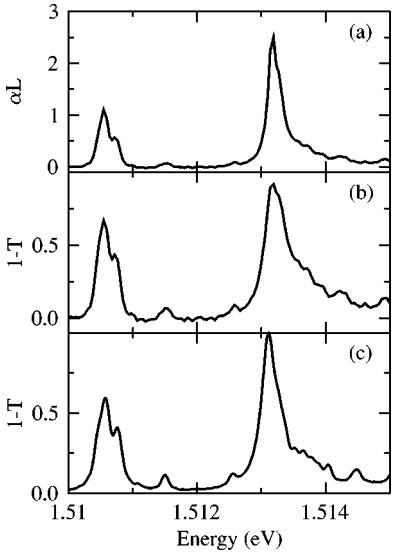

FIG. 10. Results for a high-quality 0.25-μm GaAs sample: (a) Experimental absorption and (b) experimental transmission spectrum. (c) Theoretical transmission spectrum.

C. Experimental results

Figures 10(a) and 10(b) and Fig. 11(a) and 11(b) show the retrieved absorption and phase spectra, respectively. In order to allow an easy comparison with previous works, in Fig. 10(a) we display the conventional absorption $\alpha(\omega)L = -\log T(\omega)$ calculated from the intensity transmission $T(\omega) = |t(\omega)|^2$, and given in terms of the incident and transmitted electromagnetic fields by

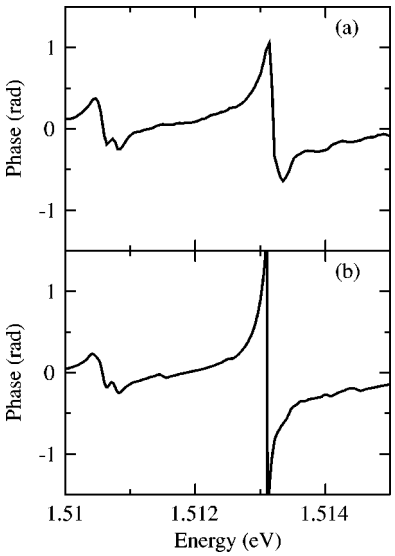

FIG. 11. Experimental (a) and theoretical (b) phase spectra corresponding to Fig. 10.

$$t(\omega)=\frac{E_{\text{trans}}(\omega)}{E_{\text{inc}}(\omega)}. \tag{22}$$

It is important to note that in the context of the analysis presented in this paper, the use of a simple exponential absorption law for the transmitted and incident intensities,

$$I_{\text{trans}}(\omega)=I_{\text{inc}}(\omega)e^{-\alpha(\omega)L}, \tag{23}$$

is no longer valid, and the quantity $\alpha(\omega)L$ not physically meaningful. Our analysis is based on the transmission $T(\omega)$, shown in Fig. 10(b), where small features are more pronounced than in the $\alpha(\omega)L$ spectrum.

The high resolution obtained using our experimental procedure shows that the excitonic lines are not Lorentzian: the HH transition at 1.5132 eV is asymmetric with a shoulder on the high-energy side, and the LH transition at 1.5105 eV reveals a clear fine structure. It is split in two components, and followed by several reproducible secondary peaks of much smaller magnitude at higher energy. All the features seen in the transmission spectrum have a counterpart in the phase spectrum $\phi(\omega)$ [Fig. 11(a)].

D. Discussion of the phase

The information contained in the phase $\phi(\omega)$ of the electric field, which also characterizes the optical response of the system is usually neglected. In a linear absorption experiment and if there are no zeros in the real part of the transmission $t(\omega)$, the phase is related to the absorption via the usual Kramers-Kronig (KK) relations. However, the KK relations have to be modified if there are zeros in the real part of $t(\omega)$. In this case, $\phi(\omega)$ is given by

$$\phi(\omega)=\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}d\omega'\frac{\log|t(\omega')|}{\omega-\omega'}+\sum_i \arg\left(\frac{\omega-\omega_i}{\omega-\omega_i^*}\right), \tag{24}$$

where the first term on the right-hand side is the usual KK integral ($\mathcal{P}$ denotes the principal value). The second term, known as Blaschke factors,[34] contains the complex zeros of $|t(\omega)|$ at ω_i. Such contributions beyond the KK relation are expected to occur for stronger absorption or thicker samples than in the case discussed presently.

Generally, a KK integral computation of the phase spectrum using the absorption data suffers from conceptual as well as practical difficulties. First, it is not possible, in the most general case, to infer the zeros of the real part of $t(\omega)$ from the knowledge of $|t(\omega)|$ alone. Second, the transmission spectrum is always experimentally measured with a finite resolution and within a finite window of frequencies, which results in a loss of resolution and divergence at the boundaries in the computed phase spectrum. This last point is illustrated in Fig. 12 which shows both the measured $\phi(\omega)$ (solid line) and the phase computed via KK relations from the absorption data (dotted line). For clarity, the latter was slightly shifted toward the bottom of the figure. It appears that the fine polariton features are washed out and that the magnitude of the HH phase shift is too small. Thus the true absorption at the HH must be slightly larger than the measured $\alpha L=2.5$ due to the finite experimental resolution. This is shown in Fig. 13(a), where we artificially increased the absorption data around the HH transition (dotted line), which then allows us to obtain a good agreement between the experimental (solid line) and computed (dotted line) phase in Fig. 13(b).

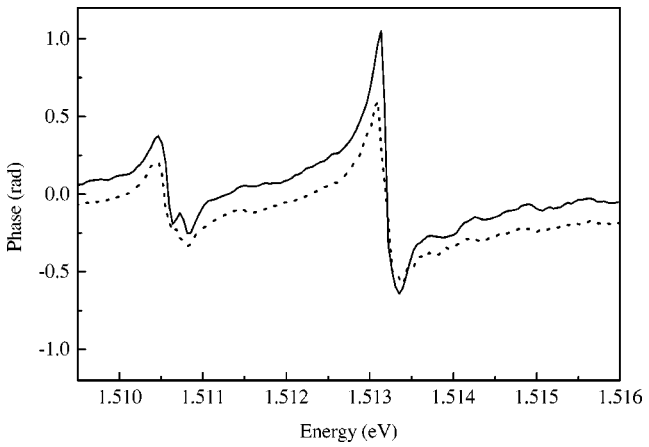

FIG. 12. Comparison of the measured phase spectrum (solid line) with a computation via the Kramers-Kronig relation from measured absorption spectrum (dotted line).

In summary, amplitude and phase measurements are complementary, and must be recorded at the same time in order to obtain a better overall resolution and, more fundamentally, to perform a complete comparison with theoretical models. As we will see below, requiring that the theory describes simultaneously the amplitude and phase of the transmitted electric field imposes very severe constraints.

E. Comparison with the theory

For a detailed comparison of the discussed experiments, we include in the theory transitions from both LH and HH-

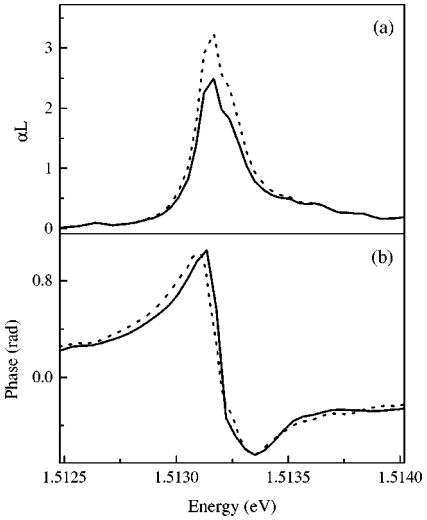

FIG. 13. (a) Measured αL spectrum (solid line) and artificially increased αL spectrum (dashed line). (b) Phase spectra calculated via Kramers-Kronig relations from the corresponding spectra in (a) and experimental phase (solid line).

bands to the electron bands. As discussed briefly in Appendix A, the symmetry of the one-electron states at the zone center leads to a different behavior of electron and hole bands. The electrons can be described by a single isotropic effective mass, but one has to include the mass anisotropy for the hole bands, i.e., the hole masses for the in-plane and z directions in Eq. (4) are different. The hole bands are conventionally labeled ''light'' and ''heavy'' according to their z masses, but the hole subband curvature is reversed in the x-y plane giving a HH transverse mass smaller than that of the LH. In addition, the thin GaAs sample on a sapphire substrate experiences a strain due to the lattice mismatch of the sample and substrate material. This strain field leads affects the whole band structure, but its most significant result is a energy difference between HH and LH that is much larger than the their small splitting to due exchange interaction effects. We use the standard values for GaAs,[35,36] $m_{\mathrm{LHz}}=0.08m_0$, $m_{\mathrm{LH}\perp}=0.23m_0$, $m_{\mathrm{HHz}}=0.53m_0$, and $m_{\mathrm{HH}\perp}=0.11m_0$, obtained from the Luttinger parameters $\gamma_1=6.9$ and $\gamma_{23}=2.5$. This leads to very similar exciton binding energies for heavy- and light-hole bands, so that the strain-induced energy difference between the LH and HH excitons can be taken to be the splitting energy $2\Delta E=2.4$ meV, observed in the experimental spectra. In addition, we use the dipole coupling strength $r_{\mathrm{cv}}=0.5$ nm corresponding to $\Delta_{\mathrm{LT}}=0.17$ meV and a phenomenological dephasing time $T=1/\gamma=10$ ps which corresponds to a broadening of 0.06 meV.

Figure 10(c) shows the result of the full calculation in direct comparison with the above discussed experimental spectrum. The characteristic splitting of the LH exciton resonance and the asymmetric line shape of the HH exciton line are well described. These two features are explained unambiguously by the interplay of polaritonic effects and quantization of the COM motion, as discussed in Sec. IV. Also, the calculated spectrum reproduces very well the high-energy secondary peaks seen experimentally, and explains their origin as COM replicas. In Fig. 11 the measured and calculated phase of the transmitted field are compared. We obtain an excellent overall agreement between the experimental and theoretical results for transmission and phase spectra.

To complete the discussion of theoretical results, in Fig. 14 we show transmission spectra obtained with various macroscopic approaches using the same parameters as for the full calculation. The 1s contribution of the numerically calculated susceptibility, including band-structure effects, is used. The Henneberger approach neither reproduces the characteristic features of HH and LH excitonic line shapes nor the COM replicas. The Ting-Frankel-Birman ABC exhibits the characteristic double-peak structure of the LH exciton line, but the HH line shape remains symmetrical and there are no pronounced COM replicas as in the experimental spectrum. Pekar's original ABC qualitatively reproduces most of the experimental signatures. Using a dead-layer thickness as fitting parameter adds the same ambiguity as in the $L=10a_B$ case discussed in Sec. IV. In the context of these approximate solutions, adjusting the material parameters can lead to unphysical results, especially if one chooses models, such as

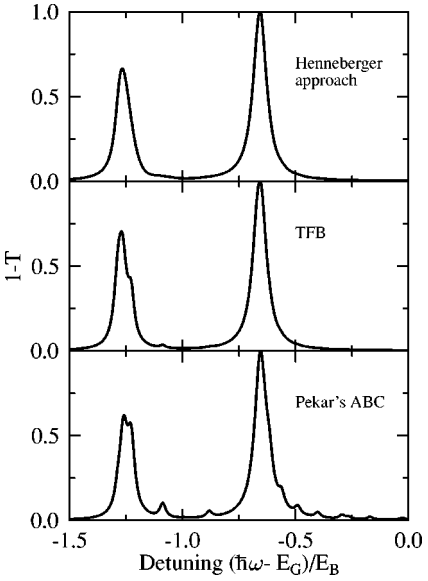

FIG. 14. Transmission spectra obtained from the Henneberger approach (a), Ting, Frankel, Birman's ABC, (b) and Pekar's ABC (c) using the same parameters as for the full calculation.

that of Henneberger, that have an indirect dependence on the physical quantities entering the calculation.

We would like to point out some very serious shortcomings of the macroscopic models for a comparison with experimental results. Indeed, very often one is confronted with situations where the material or growth parameters are not exactly known. This is especially true when new materials or heterostructures are fabricated for specific applications. Then hole masses, the sign of the strain energy, and excitonic resonance oscillator strengths have to be deduced from a fit based on a model calculation. To illustrate this point in Fig.

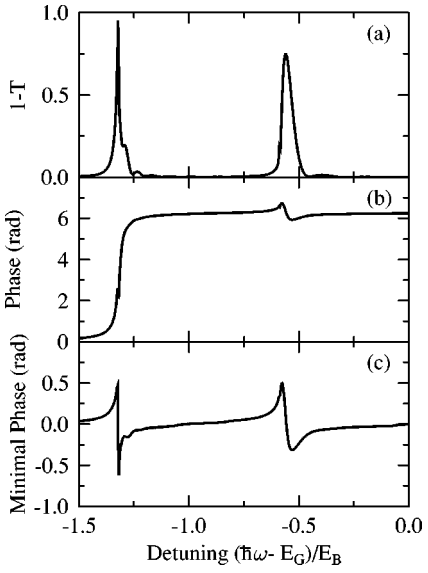

FIG. 15. (a) Transmission spectrum obtained with the Henneberger approach using artificial mass parameters and an enlarged phenomenological dephasing time $T=150$ ps. (b) Corresponding phase spectrum. (c) Minimal phase spectrum calculated from transmission spectrum via Kramers-Kronig relations.

15(a) we show the calculated transmission spectrum obtained from Henneberger's approach where, in contrast to Fig. 14, the mass and dipole coupling of the *heavy* and *light* holes have been exchanged, and the dephasing time T has been increased by a factor of 10 to $T=1/\gamma=150$ ps. One sees that with this unphysical choice of parameters Henneberger's model can roughly reproduce both the double-peak structure of the lower resonance and the asymmetry of the higher resonance. Note that the main difference between experimental and calculated spectra appears as a narrow peak at the energetically lowest resonance. Therefore, due to the finite experimental resolution a qualitative agreement between experiment and calculation can be obtained, and lead to interpretation in terms of blatantly wrong parameters. It is important to note, however, that although a reasonable fit can be obtained for the transmission spectrum, the situation is completely different for the phase as seen in Fig. 15(b). There we show the phase calculated with the parameters corresponding to the best fit for the transmission. The dominant feature is that the phase increases by about 2π, and shows a very weak dip at lower resonance and signature of the usual anomalous dispersion at higher resonance. Since the overall phase behavior deviates drastically from the experimental phase spectrum over the whole frequency range, it is not possible to explain the experiment in terms of Henneberger's approach. This again demonstrates the importance of simultaneous measurements of the phase and transmission. To further develop this point we have calculated the KK transform of $\log|t(\omega)|$ [see Eq. (24)], which yields the ''minimal phase''[34] but not the additional contributions of the Blaschke factors. The resulting phase spectrum is shown in Fig. 15(c). Its overall resemblance to the experimental spectrum is much better than that of the full phase.

VI. CONCLUSIONS

We have presented a detailed investigation of the transmission and reflection of light through layered semiconductor materials including amplitude-phase measurements on ultra high quality sample and full microscopic theory. Our method for calculating the transmission and reflection is in excellent agreement with the experimental results. The theory accounts for spatial dispersion, center-of-mass quantization, carrier confinement and band-structure effects and it is based on a direct solution of the coupled evolution equations of the microscopic electron-hole transition amplitude and of the electromagnetic fields. It explains very well the interplay of polariton effects and COM quantization for intermediate sample lengths (several times the Bohr radius) and in particular for increasing carrier confinement that occurs when the sample length becomes of the order or smaller than the Bohr radius.

To clarify the situation that has resulted from numerous phenomenological macroscopic models that have been proposed over more than three decades, we have compared calculations of the most popular ones with the full theory. Considering samples of intermediate thickness, we find that the macroscopic models cannot reproduce important features of the transmission spectra because both COM quantization and polariton effects contribute equally. Among the macroscopic models Pekar's ABC always yields better agreement with the full theory and with experiment than the Ting-Frankel-Birman or Henneberger models. If Pekar's ABC is augmented by a dead layer, the agreement with the full result can be improved. However, the thickness of that layer is an additional parameter that must be adjusted for each frequency if one wishes to account for the spectral features seen for intermediate length samples. As the sample thickness increases, however, the dead layer effects become less pronounced.

ACKNOWLEDGMENTS

We thank K. Henneberger, M. Kira, A. Knorr, and R. Zimmermann for helpful discussions. The Marburg Group acknowledges support from the Deutsche Forschungsgemeinschaft through the SFB 383, the Heisenberg (F.J.) and Leibniz (S.W.K.) programs, as well as a grant for CPU time from the John von Neumann Institute for Computing at the Forschungszentrum Jülich. The work of the Berkeley Group was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Division of Material Sciences of the U.S. Department of Energy, under Contract No. DE-AC03-76SF00098. D.S.C. and J.T. thank the France-Berkeley foundation for its support.

APPENDIX A: MULTIBAND POLARIZATION

In this appendix we derive the evolution equation of the interband transition amplitudes when optical transitions from both heavy- and light-hole bands are excited. A Kohn-Luttinger Hamiltonian is used to describe hole-band-mixing effects in the presence of an elastic strain.

1. Momentum-space formulation

The equations of motion for the nonlocal electron-hole polarizations are derived from the nonrelativistic many-body Hamiltonian

$$\hat{H}_{\mathrm{tot}}=\hat{H}_{\mathrm{kin}}+\hat{H}_{\mathrm{Coul}}+\hat{H}_{\mathrm{int}}. \tag{A1}$$

The operators are expressed in terms of electron and hole spinor wave functions at the zone center ($\vec{k}=0$). Measuring energies from the bottom of the hole band, for the kinetic part one has

$$\hat{H}_{\mathrm{kin}}=\sum_{\vec{k},s}[E_G+\epsilon_e(k)]c_s^{\dagger}(\vec{k})c_s(\vec{k})+\sum_{\vec{k},jj'}\mathcal{H}_{jj'}c_j^{\dagger}(\vec{k})c_{j'}(\vec{k}). \tag{A2}$$

Here s refers to the z component of the electron spin, and is used to label the electron bands assumed to be degenerate with an effective mass m_e^*:

$$\epsilon_e(k_{\perp},k_z)=\frac{\hbar^2}{2m_e^*}(k_{\perp}^2+k_z^2). \tag{A3}$$

The hole bands are described by a 4×4 Kohn-Luttinger Hamiltonian $\mathcal{H}$ discussed below. The hole states are labeled

by the z component of the total angular momentum j for all $\vec{k}$ values, though they are eigenstates with well-defined angular momentum projection numbers only at the zone center $\vec{k}=0$. The Coulomb Hamiltonian in this basis reads

$$\hat{H}_{\mathrm{Coul}}=\frac{1}{2\mathcal{V}}\sum_{\vec{k}\vec{k}',\vec{q}\neq 0}V(q)\Bigg[\sum_{ss'}c_s^{\dagger}(\vec{k}+\vec{q})c_{s'}^{\dagger}(\vec{k}'-\vec{q})\times c_{s'}(\vec{k}')c_s(\vec{k})+\sum_{jj'}c_j^{\dagger}(\vec{k}+\vec{q})c_{j'}^{\dagger}(\vec{k}'-\vec{q})\times c_{j'}(\vec{k}')c_j(\vec{k})-2\sum_{sj}c_s^{\dagger}(\vec{k}+\vec{q})c_j^{\dagger}(\vec{k}'-\vec{q})\times c_j(\vec{k}')c_s(\vec{k})\Bigg], \tag{A4}$$

with the 3D normalization volume $\mathcal{V}$ and the Coulomb potential in momentum space

$$V(q)=\frac{e^2}{4\pi\varepsilon_0 n_{\mathrm{bg}}^2}\frac{1}{q^2}. \tag{A5}$$

The interaction Hamiltonian in dipole approximation is

$$\hat{H}_{\mathrm{int}}=\sum_{\vec{k}\vec{k}'}\sum_{sj}\vec{d}_{sj}\cdot\vec{E}(\vec{k}-\vec{k}',t)c_s^{\dagger}(\vec{k})c_j^{\dagger}(-\vec{k}')+\mathrm{H.c.}, \tag{A6}$$

with the dipole matrix elements $\vec{d}_{sj}$ for the transition from hole band j to electron band s.

The nonlocal electron-hole transition amplitude is expressed in terms of electron and hole creation operators

$$\Psi_{js}(\vec{k}_h,\vec{k}_e)=\langle c_j(-\vec{k}_h)c_s(\vec{k}_e)\rangle, \tag{A7}$$

and the time evolution of the transition amplitude is determined by the equations of motion for c_j and c_s in a time-dependent Hartree-Fock decoupling scheme,[35,37,38]

$$i\hbar\frac{\partial}{\partial t}\Psi_{js}(\vec{k}_h,\vec{k}_e)=\sum_{j'}\mathcal{H}_{jj'}(\vec{k}_h)\Psi_{j's}(\vec{k}_h,\vec{k}_e)+\epsilon(\vec{k}_e)\Psi_{js}(\vec{k}_h,\vec{k}_e)-\frac{1}{\mathcal{V}}\sum_{q\neq 0}V(q)\times\Psi_{js}(\vec{k}_h-\vec{q},\vec{k}_e+\vec{q})-\vec{d}_{js}\cdot\vec{E}(\vec{k}_h-\vec{k}_e), \tag{A8}$$

where we have consistently suppressed the time dependence.

Since the propagation effects studied here are due to surfaces in the direction of propagation, i.e., in the z direction, it is assumed in Eq. (A8) that the exciting laser field is spatially homogenous in transverse direction perpendicular to z. Then $\vec{E}(\vec{k}_h-\vec{k}_e)=\vec{E}(k_{hz}-k_{ez})$, and the homogeneous excitation in the plane perpendicular to the z axis leads to conservation of the COM momentum

$$\vec{K}_{\perp}\equiv\vec{k}_{h\perp}+\vec{k}_{e\perp}=0 \tag{A9}$$

in that plane. Consequently,

$$\vec{k}_{\perp}\equiv\frac{m_e\vec{k}_{h\perp}-m_h\vec{k}_{e\perp}}{m_h+m_e} \tag{A10}$$

can be set

$$\vec{k}_{\perp}=\vec{k}_{h\perp}=-\vec{k}_{e\perp}. \tag{A11}$$

Also, $\vec{K}_{\perp}$ will be dropped and only the dependence on the relative momentum $\vec{k}_{\perp}$ will be kept in the transition amplitude $\Psi(\vec{k}_{\perp};k_{\mathrm{h}},k_{\mathrm{e}})$.

The 4×4 Kohn-Luttinger Hamiltonian in the axial approximation[35] is given by

$$\mathcal{H}=\begin{pmatrix}P+Q+\Delta E & R & -S & 0\\ R^* & P-Q-\Delta E & 0 & R\\ -S^* & 0 & P-Q-\Delta E & S\\ 0 & S^* & R^* & P+Q+\Delta E\end{pmatrix}. \tag{A12}$$

Its matrix elements include the band-gap energy shift ΔE due to elastic strain that can be determined from the sample's material parameters,[36] and

$$P(k_{\perp},k_z)=\frac{\hbar^2}{2m_0}\gamma_1({k_{\perp}}^2+k_z^2), \tag{A13}$$

$$Q(k_{\perp},k_z)=\frac{\hbar^2}{2m_0}\gamma_2({k_{\perp}}^2-2k_z^2), \tag{A14}$$

$$R(\vec{k}_{\perp})=-\frac{\hbar^2}{2m_0}\sqrt{3}\gamma_2(k_x^2-k_y^2-2ik_xk_y)^2=\frac{\hbar^2}{2m_0}\sqrt{3}\gamma_2{k_{\perp}}^2e^{-2i\phi_{\perp}}, \tag{A15}$$

$$S(\vec{k}_{\perp},k_z)=\frac{\hbar^2}{2m_0}2\sqrt{3}\gamma_3k_z(k_x-ik_y)=\frac{\hbar^2}{2m_0}2\sqrt{3}\gamma_3k_zk_{\perp}e^{-i\phi_{\perp}}. \tag{A16}$$

Note that the diagonal matrix elements only depend on the modulus of k_z and $k_\perp \equiv |\vec{k}_\perp|$, whereas the off-diagonal elements also depend on the angle $\phi_\perp \equiv \tan^{-1}(k_y/k_x)$ of the in-plane vector $\vec{k}_\perp$. Defining

$$\epsilon_{\mathrm{HH}}(k_\perp, k_z) \equiv P + Q = \frac{\hbar^2}{2m_{\mathrm{HH}\perp}} k_\perp{}^2 + \frac{\hbar^2}{2m_{\mathrm{HHz}}} k_z^2, \tag{A17}$$

$$\epsilon_{\mathrm{LH}}(k_\perp, k_z) \equiv P - Q = \frac{\hbar^2}{2m_{\mathrm{LH}\perp}} k_\perp{}^2 + \frac{\hbar^2}{2m_{\mathrm{LHz}}} k_z^2, \tag{A18}$$

we obtain the energy dispersion in terms of heavy- and light-hole masses.

The hole-band coupling described by the $\phi_\perp$ dependence is analyzed using the expansion

$$\Psi^{(m)}(k_\perp; k_h, k_e) = \int_0^{2\pi} d\phi_\perp e^{-im\phi_\perp} \Psi(\vec{k}_\perp; k_h, k_e). \tag{A19}$$

The in-plane integration variable is changed from $\vec{q}_\perp$ to $\vec{k}_\perp{}' = \vec{k}_\perp - \vec{q}_\perp$ and the vector sum is replaced by an integral according to

$$\frac{1}{\mathcal{V}} \sum_{\vec{k}_\perp} \sum_{q_z} \rightarrow \frac{1}{(2\pi)^3} \int_0^\infty k_\perp dk_\perp \int_0^{2\pi} d\phi_\perp \int_{-\infty}^{+\infty} dq_z. \tag{A20}$$

Then the expansion of Eq. (A8), using Eq. (A19), becomes

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \Psi_{js}^{(m)}(k_\perp; k_{hz}, k_{ez}) = & -\vec{d}_{sj} \cdot \vec{E}(k_{hz} - k_{ez}) \delta_{m,m'} + [E_{\mathrm{Gap}} + \epsilon_e(k_\perp, k_z)] \Psi_{js}^{(m)}(k_\perp; k_{hz}, k_{ez}) \\ & + \sum_{j',m'} \mathcal{H}_{jj'}^{(m-m')}(k_\perp, k_z) \Psi_{j's}^{(m')}(k_\perp; k_{hz}, k_{ez}) - \frac{1}{(2\pi)^3} \int_0^\infty dk_\perp' k_\perp' \int_{-\infty}^{+\infty} dq_z \\ & \times \sum_{m'} V_{m,m'}(k_\perp, k_\perp', q_z) \Psi_{js}^{(m')}(k_\perp; k_{hz} + q_z, k_{ez} + q_z), \end{aligned} \tag{A21}$$

where the Coulomb matrix element is given by

$$\begin{aligned} V_{m,m'}(k_\perp, k_\perp', q_z) = & \frac{e^2}{\varepsilon_0 n_{\mathrm{bg}}^2} \int_0^{2\pi} d\phi_\perp \\ & \times \int_0^{2\pi} d\phi_\perp' \frac{e^{-im\phi_\perp} e^{im'\phi_\perp'}}{(\vec{k}_\perp - \vec{k}_\perp')^2 + q_z^2}. \end{aligned} \tag{A22}$$

Here the integrand depends on the angular variables through

$$(\vec{k}_\perp - \vec{k}_\perp')^2 = k_\perp^2 + k_\perp'^2 - 2k_\perp k_\perp' \cos\phi_\perp. \tag{A23}$$

By transforming to $\tilde{\phi} = \phi - \phi'$ it can be seen that the Coulomb potential [Eq. (A22)] is proportional to $\delta_{m,m'}$. The $m - m'$ coefficients of $\mathcal{H}_{jj'}$ are determined from Eqs. (A13)–(A16). Because of the phase factors only the diagonal terms, Eqs. (A17) and (A18) are proportional to $\delta_{m,m'}$. The off-diagonal terms Eqs. (A15) and (A16) couple m to $m' = m-2$ and $m' = m-1$, respectively. Because neither the driving term nor the Coulomb potential couple different ms, only the $m=0$ components of the transition amplitudes are significant when the system is driven by a weak electric field.[39] Therefore the $m \neq 0$ components can be neglected, and we consider only the $m=0$ component

$$\Psi(k_\perp; k_{hz}, k_{ez}) \equiv \Psi^{(0)}(k_\perp, k_{hz}, k_{ez}). \tag{A24}$$

The hole bands are now effectively decoupled because the restriction of the Hamiltonian to its $m = m' = 0$ components is diagonal in the hole subspace.

We assume in the following polarized excitation $\vec{E}(\vec{r}, t) = E(z,t)\vec{\sigma}_+$. The dipole matrix elements

$$d_{\uparrow,+3/2} = \sqrt{3} d_{\downarrow,+1/2} = -d_{\mathrm{cv}} \vec{\sigma}_+ \tag{A25}$$

$$d_{\downarrow,-3/2} = \sqrt{3} d_{\uparrow,-1/2} = d_{\mathrm{cv}} \vec{\sigma}_- \tag{A26}$$

then yield that only the $j = -\frac{3}{2} \rightarrow s = \downarrow$ and the $j = -\frac{1}{2} \rightarrow s = \uparrow$ transitions are driven. We have used here the definition for the complex polarization vectors

$$\vec{\sigma}_\pm = \frac{1}{\sqrt{2}} (\vec{e}_x \pm i \vec{e}_y) \tag{A27}$$

using the real Cartesian basis vectors $\vec{e}_x$ and $\vec{e}_y$. Therefore, only two interband transition amplitudes remain which will be designated by $\nu = \mathrm{HH}$ and LH, respectively. In particular, we have $\Psi_{\mathrm{LH}} \equiv \Psi_{-1/2,\uparrow}$ and $\Psi_{\mathrm{HH}} \equiv \Psi_{-3/2,\downarrow}$. The corresponding dipole matrix elements are defined by

$$d_{\mathrm{HH}} = \sqrt{3} d_{\mathrm{LH}} \equiv d_{\mathrm{cv}}. \tag{A28}$$

Therefore, we can combine the ''free'' contributions to Eq. (A21) into interband energies for the electron-HH and electron-LH transition energies

$$\tilde{\epsilon}_{e,\nu}(k_\perp,k_{ez},k_{hz}) \equiv E_G \pm \Delta E + \epsilon_\nu(k_\perp,k_z) + \epsilon_e(k_\perp,k_z), \tag{A29}$$

where ν=HH and LH, and the ''+'' refers to the *e*-HH transition energy. Equation (A29) contains anisotropic HH and LH energies and energy shifts ΔE due to strain, respectively.

With the definition

$$\frac{1}{\mu_{\nu\perp}} = \frac{1}{m_{\nu\perp}} + \frac{1}{m_e^*} \tag{A30}$$

for ν=LH and HH, we obtain

$$\tilde{\epsilon}_{e,\nu}(k_\perp,k_{ez},k_{hz}) \equiv E_G \pm \Delta E + \frac{\hbar^2}{2\mu_{\nu\perp}}k_\perp^2 + \frac{\hbar^2}{2m_{\nu z}}k_{hz}^2 + \frac{\hbar^2}{2m_e^*}k_{ez} \tag{A31}$$

2. Real-space representation

In the *z* direction the real-space dependence is obtained by the Fourier transform

$$\Psi_\nu(k_\perp,z_e,z_h) \equiv \int_{-\infty}^{+\infty} dz_h \int_{-\infty}^{+\infty} dz_e e^{ik_{hz}z_h} e^{ik_{ez}z_e} \times \Psi_\nu(k_\perp,k_{hz},k_{ez}), \tag{A32}$$

together with the appropriate boundary conditions discussed below. The equation of motion for the ''wave functions'' $\Psi_\nu(k_\perp,z_e,z_h)$ associated with the interband transition from the LH and HH, respectively, reads

$$i\hbar\frac{\partial}{\partial t}\Psi_\nu(k_\perp,z_h,z_e) = \tilde{\epsilon}_\nu(k_\perp,\hat{k}_{hz},\hat{k}_{ez})\Psi_\nu(k_\perp,z_e,z_h) - \frac{1}{(2\pi)^2}\int dk_\perp' k_\perp' V(k_\perp,k_\perp',|z_h-z_e|) \times \Psi_\nu(k_\perp',z_h,z_e) - d_\nu E(z,t) \times \delta(z_h-z_e), \tag{A33}$$

where $z=(z_e+z_h)/2$. The Fourier-transformed Coulomb potential is

$$V_{k_\perp,k_\perp'}(|z_h-z_e|) = \frac{e^2}{2\varepsilon_0 n_{\rm bg}^2}\int_0^{2\pi} d\phi_\perp \frac{e^{-|\vec{k}_\perp-\vec{k}_\perp'||z_h-z_e|}}{|\vec{k}_\perp-\vec{k}_\perp'|}, \tag{A34}$$

together with Eq. (A23). Equation (A32) also transforms the k_{hz} and k_{ez} dependences of the kinetic energy in differential operators according to $k_{ez}\rightarrow\hat{k}_{ez}=-i\partial/\partial z_e$ and $k_{hz}\rightarrow\hat{k}_{hz}=-i\partial/\partial z_h$, which gives Eq. (A33) the form of a Schrödinger equation. A boundary condition must now be specified for the differential equation for the electron-hole transition amplitude Ψ that we have obtained. This is the essential difference between the microscopic theory and the formulations that use a macroscopic polarization calculated from the transition amplitude: the macroscopic polarization does not obey a differential equation. For the slab geometry we impose $\Psi(z_e,z_h)=0$ for z_e or z_h on one of the surfaces. As emphasized previously, more complex heterostructures can easily be handled by also including a *z*- dependent band-gap energy in Eq. (A29).

The macroscopic polarization which acts a source for the electromagnetic field is computed from the microscopic polarization according to Eq. (1) using Fourier transformation (A32),

$$\vec{P}(z,t) = \sum_{\nu=\rm HH,LH}\sum_{\vec{k}_\perp}\int_{-\infty}^{+\infty} dz'\vec{d}_\nu(z')^*\Psi_\nu(\vec{k}_\perp,z_e,z_h,t), \tag{A35}$$

where $z=(z_h+z_e)/2$, and the integration is over $z'=z_h-z_e$. Assuming that $d_\nu(z')\propto\delta(z')$, and using the vector form of the dipole moment, the above equation results in a summation of two transition amplitudes:

$$P(z,t) = 2\pi\sum_{\nu=\rm LH,HH}\int_0^\infty dk_\perp k_\perp d_{\rm cv}^*\Psi_\nu^{(0)}(k_\perp,z,z,t). \tag{A36}$$

For clarity, we have explicitly indicated that only the $m=0$ component of expansion (B4) enters into Eq. (A36). If only one type of hole band is considered, there is no summation over transition amplitudes, and we recover Eq. (8).

APPENDIX B: POLARITONS IN HOMOGENOUS MEDIA

For a homogeneous, infinitely extended sample, the solution of the propagation problem can be simplified because it is possible to separate relative and COM motion of the excited electron-hole pairs. For a better comparison with the general case where this separation cannot be made, we recall some results of the polariton propagation treatment in homogeneous media.

We start from the equation of motion for the interband-transition amplitude [Eq. (2)], and transform from electron and hole coordinates to COM and relative variables:

$$\left[i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2M}\nabla_r^2 + \frac{\hbar^2}{2\mu}\nabla_\rho^2 + V(\rho)\right]\Psi(\vec{r},\vec{\rho},t) = -\vec{d}(\rho)\cdot\vec{E}(\vec{r},t). \tag{B1}$$

The left-hand side of Eq. (B1) contains Hamiltonians for the exciton relative and COM motion which define the eigenvalue equations

$$H_{\rm rel}\varphi_i(\vec{\rho}) = -\left[\frac{\hbar^2}{2\mu}\nabla_\rho^2 + V(\rho)\right]\varphi_i(\vec{\rho}) = \epsilon_i\varphi_i(\vec{\rho}), \tag{B2}$$

$$H_{\rm com}\Phi_q(\vec{r}) = -\frac{\hbar^2}{2M}\nabla_r^2\Phi_q(\vec{r}) = \hbar\omega_q\Phi_q(\vec{r}). \tag{B3}$$

The Wannier equation [Eq. (B2)], which includes the attractive Coulomb interaction *V*, describes bound and scattering states with eigenvalues ϵ_i and eigenfunctions $\varphi_i(\vec{\rho})$, whereas

Eq. (B3) leads to plane wave solutions $\Phi_{\vec{q}}(\vec{r})=e^{i\vec{q}\cdot\vec{r}}$ for the COM motion with the energy $\hbar\omega_q=\hbar^2q^2/(2M)$. The complete set of eigenfunctions can be used to expand the solution of the Schrödinger equation for the combined electron-hole motion [Eq. (B1)],

$$\Psi(\vec{r},\vec{\rho},t)=\sum_i \int d^3q\Psi_{\vec{q},i}(t)\varphi_i(\vec{\rho})\Phi_{\vec{q}}(\vec{r}), \tag{B4}$$

which leads to

$$\left(i\hbar\frac{\partial}{\partial t}-\epsilon_i-\hbar\omega_q\right)\Psi_{\vec{q},i}(t)=-d_iE_{\vec{q}}(t). \tag{B5}$$

Here we have introduced the components of the dipole matrix element, $d_i=\int d^3\rho\varphi_i^*(\vec{\rho})d(\rho)$, and of the optical field $E_{\vec{q}}(\omega)=\int d^3r\Phi_{\vec{q}}(\vec{r})E(\vec{r},\omega)$, interacting with the corresponding excitonic states. The solution of Eq. (B5) can be found directly by a Fourier transform to the frequency domain:

$$\Psi_{\vec{q},i}(\omega)=-\frac{d_i}{\hbar\omega-\epsilon_i-\hbar\omega_q+i\gamma}E_{\vec{q}}(\omega). \tag{B6}$$

This solution allows us to determine the macroscopic polarization defined by Eq. (1). With Eq. (B4) we obtain

$$\begin{aligned}P(\vec{r})&=\sum_i \int d^3q d_i^*\Psi_{\vec{q},i}\Phi_{\vec{q}}(\vec{r})\\&=-\sum_i \int d^3q|d_i|^2\frac{\int d^3r'\Phi_{\vec{q}}^*(\vec{r}')E(\vec{r}')}{\hbar\omega-\epsilon_i-\hbar\omega_q+i\gamma}\Phi_{\vec{q}}(\vec{r}),\end{aligned} \tag{B7}$$

where the definitions of d_i and $E_{\vec{q}}$ have been used. Clearly Eq. (B7) has the form $P(\vec{r})=\int d^3r'\chi(\vec{r},\vec{r}')E(\vec{r}')$, which defines a macroscopic susceptibility:

$$\chi(\vec{r},\vec{r}')=-\sum_i |d_i|^2\int d^3q\frac{e^{i\vec{q}\cdot(\vec{r}-\vec{r}')}}{\hbar\omega-\epsilon_i-\hbar\omega_q+i\gamma}. \tag{B8}$$

As expected for the homogeneous case, the translation invariance of the COM coordinate allows the introduction of the COM momentum $\vec{q}$ independent of the relative motion. In momentum space, we directly obtain $P_{\vec{q}}(\omega)=\chi(\vec{q},\omega)E_{\vec{q}}(\omega)$. The corresponding susceptibility

$$\chi(\vec{q},\omega)=-\sum_i\frac{|d_i|^2}{\hbar\omega-\epsilon_i-\hbar\omega_q+i\gamma} \tag{B9}$$

describes excitonic resonances at ϵ_i as well as the energetic contribution $\hbar\omega_q$ of the COM motion. Since the boundary condition for the homogeneous case,

$$\Psi(\vec{r},\vec{\rho})\to 0 \quad \text{for} \quad \rho\to\infty, \tag{B10}$$

involves only the relative variable ρ, the excitonic relative and COM motion are decoupled.

A propagating electromagnetic wave in an infinitely extended medium whose response is described by Eq. (B9) must be a solution of the electromagnetic field equations (6) and (7). Defining the dimensionless dielectric function $\varepsilon(q,\omega)=n_{\rm bg}^2+\chi(q,\omega)/\varepsilon_0$, the Fourier-transformed wave equation for E becomes[35]

$$\left[\frac{\omega^2}{c_0^2}\varepsilon(q,\omega)-q^2\right]E_q(\omega)=0. \tag{B11}$$

For a propagating solution the difference in the brackets must vanish. If only a $1s$-exciton resonance is considered ($i=1$), from Eqs. (B11) and (B9) one obtains two *complex* solutions

$$q_{1,2}^2(\omega)=\frac{1}{2}(q_{\rm bg}^2+q_X^2)\pm\sqrt{\frac{1}{4}(q_{\rm bg}^2-q_X^2)^2+\kappa^4}. \tag{B12}$$

Here we have defined $q_{\rm bg}^2=n_{\rm bg}^2\omega^2/c_0^2$, $q_X^2=(2M/\hbar^2)(\hbar\omega-\epsilon_1+i\gamma)$, and $\kappa^4=(\omega^2/c_0^2)(2M/\hbar^2)\varepsilon_0|d_1|^2$. Again assuming $d(\rho)\equiv d_{\rm cv}\delta(\rho)$, and using the normalization of the excitonic wave functions, the dipole matrix element takes the form $|d_1|^2=|d_{\rm cv}|^2/(\pi a_B^3)$.

The complex wave vectors [Eq. (B12)] form the two polariton dispersion branches. The real part of these wave-vector frequency dispersions describes the propagation of polaritons in the infinitely extended medium, and the imaginary part is associated with a damping of the polariton waves. The propagating solutions from the infinitely extended medium were used in various forms in the construction of macroscopic approximations for finite samples, some of which are discussed in Sec. IV.

APPENDIX C: DISCRETIZATION AND NUMERICAL TECHNIQUES

In this appendix we outline the numerical treatment of the coupled equations (4), (6), (7), and (8). We first transform the dynamical quantities in order to eliminate the explicit appearance of the gap energy in Eq. (4), since the numerically large value necessitates very small time steps. Therefore we define, e.g.,

$$\tilde{\Psi}(t)=e^{-i\omega_0t}\Psi(t), \tag{C1}$$

$$\tilde{E}(z,t)=e^{-i\omega_0t}E(z,t). \tag{C2}$$

Here we have specialized ω_0 to be the gap frequency, $\hbar\omega_0\equiv E_{\rm Gap}$. Since this transformation affects all dynamical variables, i.e., polarizations, electromagnetic fields, and currents, we will omit the tilde in the following.

To focus on the discretization in the time domain we suppress space and momentum variables and rewrite the Schrödinger-type equation (4) symbolically as

$$i\hbar\frac{\partial}{\partial t}\Psi(t)=\hat{H}\Psi(t)+\Omega(t), \tag{C3}$$

where $\Omega(t)$ is a generalized external driving field, and $\hat{H}$ includes both the kinetic Hamiltonian and the Coulomb interaction. For the numerical integration of Eq. (C3) we have used the explicit differencing scheme[40,41]

$$\Psi(t+\Delta t)\approx\Psi(t-\Delta t)-i\frac{2\Delta t}{\hbar}\hat{H}\Psi(t)+\Omega(t). \quad \text{(C4)}$$

This scheme has two advantages: First, it is explicit, i.e., the ''new'' $\Psi(t+\Delta t)$ can be computed directly from values at earlier times without solving a system of linear equations. Second, it respects the time-reversal symmetry of the Schrödinger equation, and therefore conserves the energy and the norm of the complex wave function.[40] Also, if the chosen time step exceeds the stability limit for Eq. (C4), exponential solutions become dominant, and lead to divergent results which allows the empirical determination of the stability limit.

The amplitude Ψ in Eq. (C4) depends on two space variables and one momentum variable. The space dependence of the microscopic polarization is discretized using a two-dimensional equidistant grid (z_α, z_β) with spacing Δz_{micro}. For the momentum space variable $k_\perp$ Gaussian quadrature points $k_{\perp\kappa}$ indexed by κ are used. This leads to the discretized quantity

$$\Psi^n_\kappa(\alpha,\beta)=\Psi(k_{\perp\kappa},z_\alpha,z_\beta,t_i). \quad \text{(C5)}$$

The singularity of the Coulomb potential is removed numerically,[35] and we typically use 90 $k_\perp$ quadrature points which are accumulated below $k_\perp=4a_B^{-1}$. Choosing the spacing of $\Delta z_{\text{micro}}=0.1a_B$, we reach the stability limit of the scheme [Eq. (C4)] at $\Delta t\simeq 2.0$ fs. For our computations we used $\Delta t=1.5$ and 1.0 fs, which lead to the same results.

In the following, $\Psi^n(\alpha,\beta)$ always refers to the whole ''vector'' formed by the $\Psi^n_\kappa(\alpha,\beta)$. Derivatives are approximated on the space grid by using a second order scheme,[42] i.e., the terms in Eq. (4) containing derivatives become

$$\left[\frac{1}{m_{ez}}\frac{\partial^2}{\partial z_e^2}+\frac{1}{m_{hz}}\frac{\partial^2}{\partial z_h^2}\right]\Psi(z_e,z_h,t)$$
$$\approx\frac{1}{m_{ez}}\frac{\Psi^n(\alpha+1,\beta)-2\Psi^n(\alpha,\beta)+\Psi^n(\alpha-1,\beta)}{(\Delta z)^2}$$
$$+\frac{1}{m_{hz}}\frac{\Psi^n(\alpha,\beta+1)-2\Psi^n(\alpha,\beta)+\Psi^n(\alpha,\beta-1)}{(\Delta z)^2}. \quad \text{(C6)}$$

Consequently, for the computation of $\Psi^{n+1}(\alpha,\beta)$ At time $t_{n+1}=t_i+\Delta t$ we only need $\Psi^n(\alpha,\beta)$, and the four neighboring values $\Psi^n(\alpha+1,\beta)$, $\Psi^n(\alpha-1,\beta)$, $\Psi^n(\alpha,\beta+1)$, and $\Psi^n(\alpha,\beta-1)$. Note also that the Coulomb matrix element in Eq. (4) only couples $\Psi^n_\kappa(\alpha,\beta)$ values with the same spatial (α,β) indices, but a different momentum (κ) index. For numerical purposes it is effective to regard $\Psi^n(\alpha,\beta)$ as defined on a two-dimensional space grid where (α,β) numbers the grid points, each of which is now a vector with the discrete momentum index κ.

Grouping the grid points into square cells in a checkerboard pattern leads to a two-dimensional mesh of cells such that every grid point belongs to only one cell. The numerical computation is very well suited for parallel computers if each of these cells is assigned to one processor. Since a space grid point needs only its four neighboring points, the calculation of the right-hand side of Eq. (C4) using Eq. (C6) for a time step can be carried out directly for all *inner* points, i.e., for points whose four neighbors belong to the same cell. The points that require the values of $\Psi^n(\alpha,\beta)$ from *another cell* for the computation of $\Psi^{n+1}(\alpha,\beta)$ are lined up along the square-shaped boundaries of each cell because discretization (C6) makes only use of α and β values incremented by 1. Therefore each processor must receive the values of the adjacent boundary points from the four neighboring cells before Eq. (C4) can be computed for the boundary points of itsown cell. This data exchange between all processors must occur in each time step.

The grid structure outlined above can be implemented using the message passing interface.[43,44] Technically, nonblocking send and receive commands handle the data exchange necessary for the boundary points, *while* the $\Psi^{n+1}(\alpha,\beta)$ values are computed for all inner points. When the data exchange is complete, $\Psi^{n+1}(\alpha,\beta)$ is computed for the boundary points of each cell.

When treating the time evolution of the electromagnetic field one faces the problem that the space grid encompasses the sample and a finite space outside the sample where the electromagnetic field is produced and detected. Now the time development must be computed for several picoseconds to monitor the radiation from the slowly decaying polarization in the slab. Therefore, the initial pulse has long reached the end of the computational domain before the computation has finished and the spectrum can be computed. Thus multiple reflections occur at the ends of the space grid if one discretizes Maxwell's equation straightforwardly, and chooses a boundary condition for the computational domain because any boundary condition will lead to at least partial reflection. This difficulty is avoided by discretizing the equations for the electomagnetic field [Eqs. (6) and (7)] on a one-dimensional grid using Hartree's method. Since this method is described comprehensively in Ref. 42, we omit a detailed discussion and only quote the results here: The characteristic curves for Maxwell's equations are straight lines, $\Delta z=c\Delta t$, and the partial differential equations (6) and (7) reduce to ordinary differential equations along these lines. These are used to derive the discretized versions of the transformed Eqs. (6) and (7):

$$E_j^{n+1}=\frac{1}{2n_{\text{bg}}^2}\eta(B_{j-1}^n-B_{j+1}^n)+\frac{1}{2}\eta(E_{j-1}^n+E_{j+1}^n)$$
$$+\frac{\Delta t}{4n_{\text{bg}}^2\xi}(2J_j^{n+1}+J_{j-1}^n+J_{j+1}^n), \quad \text{(C7)}$$

$$B_j^{n+1}=\frac{1}{2n_{\text{bg}}^2}\eta(B_{j-1}^n+B_{j+1}^n)+\frac{1}{2}\eta(E_{j-1}^n-E_{j+1}^n)$$
$$+\frac{\Delta t}{4n_{\text{bg}}^2\xi}(J_{j-1}^n-J_{j+1}^n). \quad \text{(C8)}$$

Here we have discretized $E_j^n \equiv E(z_j,t_n)$, $B_j^n \equiv c_0 B(z_j,t_n)$, and $J_j^n \equiv c_0[J(z_j,t_n)+(\partial/\partial t)P(z_j,t_n)/\varepsilon_0]$. Also, we have defined $\xi = 1 - i\omega_0 \Delta t/2$ and $\eta \equiv \xi^*/\xi$. These additional factors are due to transformation (C2); the discretized versions of the original Eqs. (6) and (7) are obtained by letting $\omega_0 = 0$. We also have introduced a current $J(z,t)$ in Eq. (6), which acts as an additional source term localized outside of the sample. It is introduced in our numerical scheme as a convenient device to generate the external electromagnetic driving field.

The space grid for the macroscopic fields [Eqs. (C7) and (C8)], Δz_{macro}, is chosen such that $\Delta z_E = c\Delta t$ holds after Δt is determined according to the stability limit of Eq. (C4). Typically, we obtain Δz_E on the order of one Bohr radius. This method is capable of handling semiconductor heterostructures of arbitrary composition via the z-dependent background refractive index $n(z)$. Since we are only interested in polaritonic effects due to the excitonic polarization, we use the constant GaAs value for the refractive index $n(z) \equiv n_{\text{bg}} = 3.71$ in Eqs. (C7) and (C8).

[1] J.J. Hopfield, Phys. Rev. **112**, 1555 (1958).

[2] S.I. Pekar, Zh. Éksp. Teor. Fiz. **33**, 1022 1958 [Sov. Phys. JETP **6**, 785 (1958)].

[3] A. Stahl and I. Balslev, *Electrodynamics of the Semiconductor Band Edge*, Springer Tracts in Modern Physics Vol. 10 (Springer, New York, 1987).

[4] K. Henneberger, Phys. Rev. Lett. **80**, 2889 (1998).

[5] D.F. Nelson and B. Chen, Phys. Rev. Lett. **83**, 1263 (1999).

[6] R. Zeyher, Phys. Rev. Lett. **83**, 1264 (1999).

[7] J.J. Hopfield and D.G. Thomas, Phys. Rev. **132**, 563 (1963).

[8] V.M. Agranovich, V.E. Kravtsov, T.A. Leskova, A.G. Mal'shukov, G. Hernández-Cocoletzi, and A.A. Maradudin, Phys. Rev. B **29**, 976 (1984).

[9] R. Ruppin and R. Englman, Phys. Rev. Lett. **53**, 1688 (1984).

[10] R. Zeyher, J.L. Birman, and W. Brenig, Phys. Rev. B **6**, 4613 (1972).

[11] R. Zeyher, in *Recent Developments in Condensed Matter Physics*, edited by J. T. Devreese (Plenum, New York, 1981), Vol. I, p. 807.

[12] C.S. Ting, M.J. Frankel, and J.L. Birman, Solid State Commun. **17**, 1285 (1975).

[13] V.M. Axt and A. Stahl, Solid State Commun. **77**, 189 (1991).

[14] K. Victor, V.M. Axt, and A. Stahl, Z. Phys. B: Condens. Matter **92**, 35 (1993).

[15] F. Bassani, Y. Chen, G. Czajkowski, and A. Tredicucci, Phys. Status Solidi B **180**, 115 (1993).

[16] S. Bischoff, A. Knorr, and S.W. Koch, Phys. Rev. B **55**, 7715 (1997).

[17] K. Cho, J. Phys. Soc. Jpn. **55**, 4113 (1986).

[18] H. Ishihara and K. Cho, Phys. Rev. B **41**, 1424 (1990).

[19] A. D'Andrea and R. Del Sole, Phys. Rev. B **41**, 1413 (1990).

[20] B. Chen and D.F. Nelson, Phys. Rev. B **48**, 15 372 (1993).

[21] K.-H. Pantke *et al.*, Phys. Rev. B **70**, 327 (1993); U. Neukirch *et al.*, Phys. Rev. B **57**, 9208 (1998).

[22] D. Fröhlich *et al.*, Phys. Rev. Lett. **67**, 2343 (1991).

[23] A. Tredicucci, Y. Chen, F. Bassani, J. Massies, C. Deparis, and G. Neu, Phys. Rev. B **47**, 10 348 (1993).

[24] R.G. Ulbrich and C. Weisbuch, Phys. Rev. Lett. **38**, 865 (1977).

[25] H.J. Bakker *et al.*, Phys. Rev. B **50**, 7805 (1994); T. Rappen *et al.*, *ibid.* **47**, 9658 (1993); G.W. Fehrenbach *et al.*, Phys. Rev. Lett. **49**, 1281 (1982).

[26] L. Schultheis *et al.*, in *Excitons in Confined Systems*, edited by R. Del Sole, A. D'Andrea, and A. Lapiccirella (Springer, Berlin, 1987), p. 110.

[27] H. Schneider and K. Köhler, Phys. Rev. B **52**, R14 364 (1995); Y. Zhilyaev *et al.*, Fiz. Tverd. Tela (Leningrad **32**, 1801 (1990) [Sov. Phys. Solid State **32**, 1048 (1990)]; J. Kusano *et al.*, Solid State Commun. **72**, 215 (1989); L. Schultheis and K. Ploog, Phys. Rev. B **29**, 7058 (1984); D.D. Sell *et al.*, Phys. Rev. B **7**, 4568 (1973).

[28] J. Tignon, T. Hasche, D.S. Chemla, H.C. Schneider, F. Jahnke, and S.W. Koch, Phys. Rev. Lett. **84**, 3382 (2000).

[29] I. Balslev, R. Zimmermann, and A. Stahl, Phys. Rev. B **40**(6), 4095 (1989).

[30] D. Viri, R. Del Sole, and A. D'Andrea, Phys. Rev. B **48**, 9110 (1993).

[31] A.C. Schaefer and D.G. Steel, Phys. Rev. Lett. **79**, 4870 (1997).

[32] L. Lepetit, G. Chériaux, and M. Joffre, J. Opt. Soc. Am. B **12**, 2467 (1995).

[33] K.W. DeLong, D.N. Fittinghoff, and R. Trebino, IEEE J. Quantum Electron. **32**, 1253 (1996).

[34] J. Perina, *Coherence of Light*, 2nd ed. (Reidel, Dordrecht, 1985).

[35] H. Haug and S. W. Koch, 3rd ed. *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (World Scientific, Singapore, 1995).

[36] S.L. Chuang, *Physics of Optoelectronic Devices* (Wiley, New York, 1995).

[37] R. Binder and S.W. Koch, Prog. Quantum Electron. **19**(4/5), 307 (1995).

[38] F. Steininger, A. Knorr, T. Stroucken, P. Thomas, and S.W. Koch, Phys. Rev. Lett. **77**(3), 550 (1996).

[39] Y.Z. Hu, R. Binder, and S.W. Koch, Phys. Rev. B **47**, 15 679 (1993).

[40] C. Leforestier *et al.*, J. Comput. Phys. **94**, 59 (1991).

[41] M. Kira, I. Tittonen, and S. Stenholm, Phys. Rev. B **52**, 10 972 (1995).

[42] W.F. Ames, *Numerical Methods for Partial Differential Equations*, 3rd ed. Computer Science and Scientific Computing (Academic, San Diego, 1992).

[43] W. Gropp, E. Lusk, and A. Skjellum, *Using MPI* (MIT Press, Cambridge, MA, 1994).

[44] E.F. Van de Velde, *Concurrent Scientific Computing*, Texts in Applied Mathematics Vol. 16 (Springer, New York, 1994).