

Excitonic effects in time-dependent density-functional theory: An analytically solvable model

R. Stubner,* I. V. Tokatly, and O. Pankratov

Lehrstuhl für Theoretische Festkörperphysik, Universität Erlangen-Nürnberg, Staudtstrasse 7, 91058 Erlangen, Germany

(Received 15 June 2004; published 23 December 2004)

We investigate the description of excitonic effects within time-dependent density-functional theory (TD-DFT). The exchange-correlation kernel f_{xc} introduced in TDDFT allows a clear separation of quasiparticle and excitonic effects. Using a diagrammatic representation for f_{xc} , we express its excitonic part f_{xc}^{Ex} in terms of the effective vertex function Λ . The latter fulfills an integral equation that thereby establishes the exact correspondence between TDDFT and the standard many-body approach based on the Bethe-Salpeter equation (BSE). The diagrammatic structure of the kernel in the equation for Λ suggests the possibility of strong cancellation effects. Should the cancellation take place, already the first-order approximation to f_{xc}^{Ex} is sufficient. A potential advantage of TDDFT over the many-body BSE method is thus dependent on the efficiency of the above-quoted cancellation. We explicitly verify this for an analytically solvable two-dimensional two-band model. The calculations confirm that the low-order f_{xc}^{Ex} perfectly describes the bound exciton as well as the excitonic effects in the continuous spectrum in a wide range of the electron-hole coupling strength.

DOI: 10.1103/PhysRevB.70.245119

PACS number(s): 71.10.-w, 71.15.Qe, 71.35.-y, 78.20.Bh

I. INTRODUCTION

Calculation of electronic excitation spectra remains one of the central problems of the quantum theory of solids. Of special interest are two-particle electron-hole excitations, which determine the material's optical properties. In semiconductors and insulators the electronic screening is suppressed by the energy gap and the interaction of the excited quasiparticles may substantially modify the excitation spectrum. The excitonic effects stemming from this interaction comprise the formation of bound electron-hole states as well as the alteration of the absorption in the continuum spectrum above the band edge. The latter is commonly referred to as unbound exciton effects or the Sommerfeld absorption enhancement.

In many-body perturbation theory two-particle excitations are characterized by the two-particle Green function that satisfies the Bethe-Salpeter equation (BSE).¹ Already in 1980 Hanke and Sham,² using an approximate tight-binding representation, showed that the BSE correctly describes the strong excitonic features above the optical absorption edge in Si. In the current "state of the art" procedure (see Ref. 3 for a recent review) the calculation of excitonic effects involves three steps. First, a density-functional theory (DFT) calculation in the local density approximation (LDA) is performed. In the second stage, the LDA Kohn-Sham (KS) energies and wave functions are used as a starting point for the *GW* calculation of the quasiparticle spectrum. Finally, the BSE is solved numerically, using the *GW* eigenvalues and the LDA wave functions as input characteristics of the noninteracting quasiparticles. The outlined procedure leads to highly accurate results, as has been shown for a number of relatively simple systems, mostly bulk semiconductors (see Ref. 3 and references therein). However, this method is extremely laborious, and for more complex systems the calculations become prohibitively expensive.

A promising alternative that has been intensively developed over recent years relies on time-dependent density-functional theory (TDDFT).⁴ This theory allows us to calcu-

late (formally exactly) the linear density-density response function and thereby the excitation energies.⁵ Since in the framework of DFT the exchange-correlation (*xc*) effects are lumped in a local *xc* potential v_{xc} , the TDDFT equation for the response function contains the variational derivative of v_{xc} with respect to density $f_{xc}(\mathbf{r}, t; \mathbf{r}', t') = \delta v_{xc}(\mathbf{r}, t) / \delta n(\mathbf{r}', t')$. This *xc* kernel f_{xc} is the central unknown quantity of TDDFT in the linear response regime. In their pioneering work Zangwill and Soven⁶ calculated the photoabsorption in rare gases in a self-consistent field manner. They used what later became known as the adiabatic local density approximation (ALDA), simply substituting the time-dependent density in the LDA *xc* potential $v_{xc}^{\text{ALDA}} = v_{xc}^{\text{LDA}}(n(\mathbf{r}, t))$. The resulting *xc* kernel is local in space and time: $f_{xc}^{\text{ALDA}} = \delta(\mathbf{r} - \mathbf{r}') \delta(t - t') dv_{xc}^{\text{LDA}} / dn(\mathbf{r})$. ALDA has been successfully applied to various finite systems like atoms or molecules.^{5,7-9} Typically in these systems already the random phase approximation (RPA) response function calculated with KS eigenvalues and eigenfunctions gives good results. The correction due to f_{xc}^{ALDA} is quite small, which signifies that Hartree effects dominate in the response function. Unfortunately, f_{xc}^{ALDA} remains insignificant also in extended systems such as semiconductors or insulators, where KS-RPA gives a very poor description of the absorption spectra.^{10,11} Thus, whereas a correct accounting for *xc* effects becomes crucial in extended systems, the ALDA kernel f_{xc}^{ALDA} fails to provide even a reasonable starting approximation.

In the late 1990s it has been realized that ALDA cannot serve as the basis approximation for the dynamic *xc* response of an inhomogeneous electron gas, because of the intrinsically nonlocal nature of f_{xc} .¹² For extended systems with an energy gap this remains valid even in the static case.^{13,14}

The importance of the nonlocality of f_{xc} was highlighted by the work of Reining and co-workers,^{15,16} who were able to describe the contributions of unbound excitons in several diamond or zinc-blende type semiconductors with a static *xc* kernel proportional to $1/|\mathbf{r} - \mathbf{r}'|$. Other examples for this are

the exact exchange kernel^{17,18} and the results of de Boeji *et al.*,¹⁹ where in the context of time-dependent current-density functional theory^{20,21} the nonlocal effects were crucial for accurately describing the effects of unbound excitons.²²

It can be easily understood that a nonlocal f_{xc} is crucial for describing excitonic effects. Within TDDFT the proper polarization operator is defined via the RPA-like equation

$$\tilde{\chi}(\omega) = \chi_S(\omega) + \chi_S(\omega) \cdot f_{xc}(\omega) \cdot \tilde{\chi}(\omega),$$

where in a crystalline solid $\tilde{\chi}$, the Kohn-Sham response function χ_S and f_{xc} are matrices in reciprocal space. The matrix structure of $\tilde{\chi}$ is responsible for local-field effects. However, these are relatively small in typical semiconductors and can be neglected for a qualitative analysis. Keeping only diagonal elements with zero reciprocal lattice vectors, we can easily solve for $\tilde{\chi}$, obtaining

$$\tilde{\chi}(\omega, \mathbf{q}) = \frac{\chi_S(\omega, \mathbf{q})}{1 - f_{xc}(\omega, \mathbf{q})\chi_S(\omega, \mathbf{q})}.$$

The macroscopic dielectric function ε_M is given by

$$\varepsilon_M(\omega) = 1 - \lim_{q \rightarrow 0} V_C(\mathbf{q})\tilde{\chi}(\omega, \mathbf{q}) = 1 - \lim_{q \rightarrow 0} \frac{V_C(\mathbf{q})\chi_S(\omega, \mathbf{q})}{1 - f_{xc}(\omega, \mathbf{q})\chi_S(\omega, \mathbf{q})},$$

with the Coulomb interaction $V_C(\mathbf{q}) = 4\pi e^2/q^2$. An additional excitonic peak in $\varepsilon_M(\omega)$ appears when the denominator vanishes. However, it is well known that χ_S is proportional to q^2 in the limit $q \rightarrow 0$ for systems with an energy gap.^{23,24} Hence f_{xc} must behave as $1/q^2$ in this limit to counterbalance χ_S . Otherwise the xc kernel would have no effect on $\varepsilon_M(\omega)$ at all. For the static long-ranged xc kernel of Reining and co-workers we have $f_{xc}(\omega, \mathbf{q}) = 4\pi e^2 \beta/q^2$ with some constant β . The macroscopic dielectric function thus reads

$$\varepsilon_M = 1 - \frac{4\pi e^2 \alpha_S(\omega)}{1 - 4\pi e^2 \beta \alpha_S(\omega)},$$

where $\alpha_S(\omega)$ is the macroscopic polarizability of the Kohn-Sham system $\alpha_S(\omega) = \lim_{q \rightarrow 0} \chi_S(\omega, \mathbf{q})/q^2$. For a typical $\alpha_S(\omega)$ close to the band edge this formula suggests the existence of only *one* excitonic peak. However, one expects several peaks from unbound excitons above the band gap and bound excitons within the gap. Phenomenologically one could overcome this problem by introducing a frequency-dependent β . One though would need to introduce very rapid oscillations in the region of the Rydberg series of bound excitonic states.

Probably the most promising path in the quest for a good approximation to f_{xc} is a direct comparison of the TDDFT formalism with the BSE.^{15,25–27} Simply comparing the calculated spectra, it was found that it is often sufficient to use an approximation to f_{xc} that is of the first-order in the screened particle-hole interaction. Although these results are very encouraging, it is unclear why this approximation is so efficient and what its range of validity is.

In this paper we derive a diagrammatic expression that exactly relates the excitonic part of f_{xc} to the BSE. We start with splitting f_{xc} into two parts, separately accounting for quasiparticle and excitonic effects. We then apply the diagrammatic rules we previously derived²⁸ to these two parts

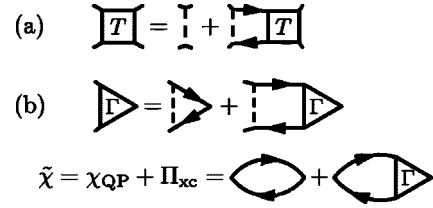


FIG. 1. Diagrammatic representation of the Bethe-Salpeter equation.

of f_{xc} . This leads us to an expression for the excitonic part of f_{xc} in terms of the three-point function Λ . The latter satisfies an integral equation similar to the BSE that establishes the exact correspondence between TDDFT and common many-body theory. The main advantage of this approach is that the possibility of cancellation effects, which have been conjectured in Ref. 27, is directly seen in the kernel of the equation for Λ .

In order to investigate the properties of our integral equation and the applicability of low-order approximations we study a model two-band system. In this model both the BSE and TDDFT equation can be solved analytically, which offers an ideal test bed for approximations to the exact f_{xc} . We find that indeed there are strong cancellation effects in the integral equation for Λ in the energy range close to the band gap. For this reason both the shallow excitons and the unbound excitonic effects are well described with a first-order approximation to the excitonic part of f_{xc} .

The paper is organized as follows. In Sec. II we investigate diagrammatic properties of the excitonic part of f_{xc} and derive an exact correspondence between TDDFT and BSE. In Sec. III we introduce the model system that is used in further calculations. Sections IV and V are devoted to analytic calculations of excitonic effects with a general short-ranged and the Coulomb interaction, respectively. Finally, we present our conclusions in Sec. VI.

II. DIAGRAMMATIC MEANING OF f_{xc}

The Bethe-Salpeter equation in the particle-hole channel is commonly formulated as an integral equation for the particle-hole propagator or the scattering matrix T in the ladder approximation. This integral equation is equivalent to a summation of all ladder diagrams. The diagrammatic representation of the T -matrix formulation is depicted in Fig. 1(a), where the full lines are the quasiparticle Green functions and the dashed lines represent the screened interaction. In the context of a connection to TDDFT we are not interested in the four-point particle-hole propagator but rather in the response functions. Therefore we consider a modified BSE where two of the external lines of the scattering matrix have been contracted to form the three-point function Γ . The diagrammatic representation of the integral equation for Γ and its relation to the proper polarization operator $\tilde{\chi}$ are depicted in Fig. 1(b). For our purposes this equation is equivalent to the BSE. In the following we refer to the equation in Fig. 1(b) as the Bethe-Salpeter equation.

In terms of TDDFT the proper polarization operator $\tilde{\chi}$ is given as

$$\tilde{\chi}(\omega) = \chi_S(\omega) + \chi_S(\omega) \cdot f_{xc}(\omega) \cdot \tilde{\chi}(\omega), \quad (1)$$

where χ_S represents the density-density response function of the noninteracting KS particles, i.e., a bare loop of two KS Green functions. Equation (1) looks like the RPA equation, although it relates χ_S with the full $\tilde{\chi}$, which includes all self-energy and ladder diagrams. We can therefore interpret f_{xc} as an effective interaction that describes self-energy and ladder diagrams in the annihilation channel. Thereby f_{xc} contains both quasiparticle and excitonic effects. As in this paper we are only interested in the excitonic effects, it is tempting to separate these two contributions to f_{xc} , as suggested in previous works.^{15,25} This separation is indeed possible without approximations, because we have

$$f_{xc} = \chi_S^{-1} - \tilde{\chi}^{-1} = \underbrace{\chi_S^{-1} - \chi_{QP}^{-1}}_{=: f_{xc}^{QP}} + \underbrace{\chi_{QP}^{-1} - \tilde{\chi}^{-1}}_{=: f_{xc}^{Ex}}, \quad (2)$$

where χ_{QP} is the density-density response function for the noninteracting quasiparticles (QP). By definition, f_{xc}^{QP} and f_{xc}^{Ex} are the kernels of the following RPA-type equations:

$$\chi_{QP}(\omega) = \chi_S(\omega) + \chi_S(\omega) \cdot f_{xc}^{QP}(\omega) \cdot \chi_{QP}(\omega), \quad (3a)$$

$$\tilde{\chi}(\omega) = \chi_{QP}(\omega) + \chi_{QP}(\omega) \cdot f_{xc}^{Ex}(\omega) \cdot \tilde{\chi}(\omega). \quad (3b)$$

The newly introduced quantities f_{xc}^{QP} and f_{xc}^{Ex} describe quasiparticle and excitonic effects, respectively. This can be visualized by applying the diagrammatic rules for f_{xc} as derived in Ref. 28. The structure of the diagrammatic representation of f_{xc}^{QP} and f_{xc}^{Ex} is similar to the one for f_{xc} , except that for f_{xc}^{QP} one has to use KS Green functions and should only account for all possible self-energy insertions in every order of the perturbation theory. This clearly describes the quasiparticle corrections. For f_{xc}^{Ex} one should use the quasiparticle Green functions with all possible particle-hole interactions. Obviously, this reflects excitonic contributions. It is important to note that the general properties of the perturbative expansion of f_{xc} obtained in Ref. 28 remain valid separately for f_{xc}^{QP} and f_{xc}^{Ex} . In particular, f_{xc}^{QP} remains finite at KS excitation energies in every order of the perturbation theory. The same holds for f_{xc}^{Ex} at excitation energies of the noninteracting quasiparticle system. Note that our f_{xc}^{Ex} is the same as the f_{xc}^{FQP} of Refs. 26 and 29.

Let us briefly outline the diagrammatic rules for the excitonic part of f_{xc} . According to Ref. 28 with the above-mentioned modifications we have to draw all loops with n particle-hole interactions to construct the n th order correction. These diagrams serve as parent graphs for the construction of the n th order f_{xc}^{Ex} . To comply with the BSE we must use the ladder approximation here as well.³² Therefore, only one diagram with n interactions is left. To the two ends of the diagram we have to attach wiggled lines representing χ_{QP}^{-1} . Next, we work out all possibilities to separate this parent graph into two by cutting two fermionic lines. Then we join the external fermionic lines of these parts, connect them by a wiggled line, and change the sign of the resulting diagram. Obviously, the only way to separate the parent graph is to cut between adjacent interaction lines. The cutting does not change the ladder structure of the diagrams as seen in Fig.

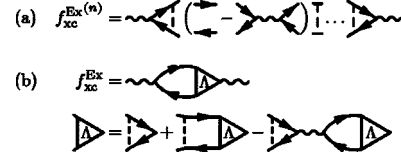


FIG. 2. (a) The n th order of f_{xc}^{Ex} and (b) the integral equation for the three-point function Λ , which is an essential part of f_{xc}^{Ex}

2(a). The summation of all ladder diagrams can be cast in an integral equation as displayed in Fig. 2(b). If one inserts f_{xc}^{Ex} obtained by solving the equation of Fig. 2(b) into Eq. (3b) and calculates the response function $\tilde{\chi}$, the result will be the same as the $\tilde{\chi}$ obtained from the BSE. In this sense the equation of Fig. 2(b) gives the exact “translation” of the BSE into the TDDFT language.

At this point we would like to highlight the differences between the integral equation of Fig. 2(b) and the equation for the excitonic part of f_{xc} derived in Ref. 27 [Eqs. (4) and (5) therein]. The iterative equation of Marini *et al.*²⁷ is for the two-point xc kernel and is based on (finite order approximations to) the xc part of the response function and is logically analogous to our diagrammatic expansion of f_{xc} obtained in Ref. 28. However, the equation of Fig. 2(b) is an integral equation for a three-point function analogous to the BSE (Ref. 33) and can be solved *instead* of the BSE to obtain the same results.

Unfortunately, the exact calculation of f_{xc}^{Ex} from the equation of Fig. 2(b) is at least as difficult as obtaining an exact solution of the BSE. However, one can hope that the two-point kernel f_{xc}^{Ex} is more suitable for approximations. An indication in this direction can be seen directly from the diagrammatic equation of Fig. 2(b). Comparing this equation to Fig. 1(b) we see that it can be obtained from the BSE by substitution of the particle-hole propagator as shown in Fig. 3(a). This is the same replacement that was used in Ref. 28 to prove the cancellation of divergencies in f_{xc} at KS excitation energies. Similarly, it facilitates the cancellation of divergencies in f_{xc}^{Ex} at QP excitation energies.

Alternatively we can interpret the difference between the BSE and our equation for Λ as a modification of the interaction process in the second diagram of the right-hand side of the equation in Fig. 1(b). This change of the interaction is displayed in Fig. 3(b). As f_{xc}^{Ex} is responsible for translating the ladder diagrams into the annihilation channel, it is non-trivial only if the ladder channel and the annihilation channel are distinguishable. This is easily seen from the replacement of Fig. 3(b). Indeed the ladder and the annihilation channel coincide for nonrelativistic systems with a static point interaction. In this case the quasiparticle propagators in Fig. 3(b)

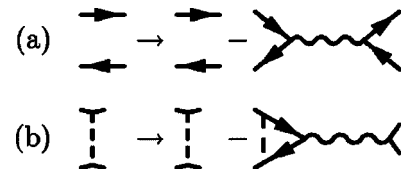


FIG. 3. Diagrammatic representation of the two possible interpretations of the difference between BSE and our equation for Λ .

form a polarization loop which cancels the wiggled line. As a result the two diagrams of Fig. 3(b) cancel exactly. In the equation for Λ in Fig. 2(b) this means that the last two terms cancel and Λ reduces to the first term. The excitonic part of the xc kernel then reduces to the interaction itself. One can therefore expect that in systems with a short-ranged and almost static effective interaction, the two terms of Fig. 3(b) will cancel to a large extent, and a low-order approximation to f_{xc}^{Ex} will be sufficient. Conditions like this can, e.g., be found in simple metals. On the contrary, in semiconductors screening is less efficient and the effective interaction is long-ranged. Further research is needed to verify to what extent the cancellation is efficient.

Cancellation effects as we are expecting them from our diagrammatic equation have been seen in Ref. 27. The success of the lowest-order f_{xc} found in Ref. 27 implies that *de facto* the cancellation can be efficient in materials with a band gap as well.

III. MODEL SYSTEM

In this section we consider a model system that reveals both bound and unbound excitonic effects and where both the BSE and the equation for Λ can be solved analytically. With the exact f_{xc}^{Ex} at hand we can verify under what circumstances the first-order approximation to f_{xc}^{Ex} may be sufficient. The approximate kernel must describe bound as well as unbound excitons. A simple system with a bound exciton is given by the two-band Dirac model with a static density-density interaction. Moreover, we consider the two-dimensional (2D) case in order to avoid technical difficulties with diverging integrals. The model Hamiltonian is given by

$$\mathcal{H} = \int d^2r \psi^\dagger(\mathbf{r}) \hat{H} \psi(\mathbf{r}) + \frac{1}{2} \int d^2r \int d^2r' \hat{n}(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \hat{n}(\mathbf{r}'), \quad (4)$$

where $\hat{n}(\mathbf{r}) = \psi^\dagger(\mathbf{r}) \psi(\mathbf{r})$ is the density operator, $\psi(\mathbf{r})$ is a two-component field operator, and $V(\mathbf{r} - \mathbf{r}')$ describes the interaction between the particles. The Hamiltonian for the noninteracting particles reads

$$\hat{H} = \hat{k}_x \sigma_x + \hat{k}_y \sigma_y + \Delta \sigma_z = \begin{pmatrix} \Delta & \hat{k}_- \\ \hat{k}_+ & -\Delta \end{pmatrix}, \quad (5)$$

where $\sigma_{x,y,z}$ are the Pauli matrices, $\hat{k}_\pm = \hat{k}_x \pm i\hat{k}_y$, the 2D momentum operator $\hat{\mathbf{k}} = (\hat{k}_x, \hat{k}_y)$, and the band gap equals 2Δ . The energy dispersion for the noninteracting particles has two branches that we label c and v for the (unoccupied) conduction band and the (occupied) valence band:

$$E_{c/v}(\mathbf{k}) = \pm E_k = \pm \sqrt{\Delta^2 + k^2}. \quad (6)$$

The eigenvectors of \hat{H} are

$$\Psi_{c\mathbf{k}} = \begin{pmatrix} u_k \\ \frac{k_+}{k} v_k \end{pmatrix}, \quad \Psi_{v\mathbf{k}} = \begin{pmatrix} -\frac{k_-}{k} v_k \\ u_k \end{pmatrix} \quad (7)$$

with

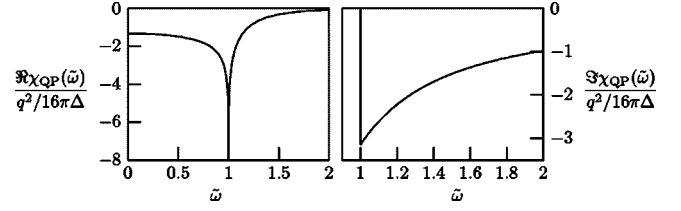


FIG. 4. The real and the imaginary part of the exact χ_{QP} .

$$u_k = \sqrt{\frac{1}{2} \left(1 + \frac{\Delta}{E_k} \right)}, \quad v_k = \sqrt{\frac{1}{2} \left(1 - \frac{\Delta}{E_k} \right)}. \quad (8)$$

Note that this model can be understood as a two-component (“relativistic”) system. The interactions in the ladder channel and in the annihilation channel in such a system are always distinct regardless whether the interaction is long- or short-ranged. Therefore in this model f_{xc}^{Ex} is non-trivial even in the case of a point interaction.

We are going to solve this model in the ladder approximation, i.e., ignoring self-energy terms and higher-order corrections to the irreducible scattering matrix. Comparing to the BSE this implies that \hat{H} refers to the independent quasiparticles and V is the screened interaction between them. Therefore the one-particle Green functions of \hat{H} are the quasiparticle Green functions:

$$G_{\text{QP}}(\omega, \mathbf{k}) = \frac{\Psi_{c\mathbf{k}} \Psi_{c\mathbf{k}}^\dagger}{\omega - E_k + i\delta} + \frac{\Psi_{v\mathbf{k}} \Psi_{v\mathbf{k}}^\dagger}{\omega + E_k - i\delta}. \quad (9)$$

Note that $\Psi_{v\mathbf{k}}$ and $\Psi_{c\mathbf{k}}$ are two-component vectors and G_{QP} is a 2×2 matrix. To lowest order in the wave vector \mathbf{q} the quasiparticle response function $\chi_{\text{QP}}(\omega, \mathbf{q})$ is given by

$$\begin{aligned} \chi_{\text{QP}}(\tilde{\omega}, \mathbf{q}) &= -i \int \frac{d\epsilon}{2\pi} \int \frac{d^2k}{(2\pi)^2} \text{tr} G_{\text{QP}}(\epsilon + \tilde{\omega}, \mathbf{k} + \mathbf{q}) G_{\text{QP}}(\epsilon, \mathbf{k}) \\ &= -\frac{q^2}{16\pi\Delta} \left(\frac{\tilde{\omega}^2 + 1}{2\tilde{\omega}^3} \ln \frac{1 + \tilde{\omega}}{1 - \tilde{\omega}} - \frac{1}{\tilde{\omega}^2} \right), \end{aligned} \quad (10)$$

where $\tilde{\omega} = \omega / (2\Delta)$. The real and the imaginary part of this function are displayed in Fig. 4. A nonvanishing imaginary part occurs only at frequencies above the quasiparticle gap $\tilde{\omega} > 1$, when the argument of the logarithm becomes negative. Note that one sees here explicitly the q^2 dependence of the response functions mentioned in the Introduction.

In principle, we are now in the position to solve the BSE in this model analytically. However, to proceed further we need to introduce some technical issues. Due to the matrix structure of G_{QP} one has to compute traces when calculating χ_{QP} . Similarly, the three-point functions Γ and Λ are 2×2 matrices and traces have to be calculated over internal indices in the diagrams of Figs. 1(a) and 2(b). Since all these matrices have nonvanishing off-diagonal elements, evaluation of these traces becomes quite tedious. Therefore it is convenient to choose the eigenstates of \hat{H} as the basis. The Green function G_{QP} becomes then a diagonal matrix with the elements

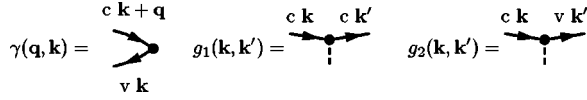


FIG. 5. The diagrammatic representation of the “bare” vertex and the “charges” of Eqs. (12) and (13).

$$G_{c/v}(\omega, \mathbf{k}) = \frac{1}{\omega \mp E_{\mathbf{k}} \pm i\delta}. \quad (11)$$

We thus can abandon the matrix notation altogether and use the two scalar Green functions G_c and G_v instead of the now diagonal 2×2 matrix G_{QP} . Now, every full line is either a conduction- or a valence-state propagator (11). This, of course, increases the number of diagrams we need to draw, because we have to consider all possible combinations of conduction- and valence-band states. However, in all diagrams the “upper” and “lower” Green functions that constitute a particle-hole propagator must always be of a different type (c or v). The diagrams with c - c or v - v two-particle propagators vanish due to the integration over frequency since both “upper” and “lower” Green function have their pole in the same half of the complex plane. This means that these diagrams do not contribute to the polarization. The same holds for the three-point functions Γ and Λ . Instead of one equation for the 2×2 matrix Γ as depicted in Fig. 1(b) we obtain two coupled equations for the scalar functions Γ_{cv} and Γ_{vc} . For Γ_{cv} the upper line is a conduction-band state and the lower line a valence-band state, whereas for Γ_{vc} it is vice versa.

The transfer to the diagonal representation of G_{QP} is equivalent to a corresponding transformation of the field operators in Eq. (4). This transformation generates the “bare” vertex, describing interaction with an external field in the polarization diagrams, as well as the “interaction vertices.” In the original representation (5) all these vertices are simply unit matrices. They become, however, nontrivial matrices in the diagonal representation. In fact, only three matrix elements of these vertices are essential. These are depicted in Fig. 5. To the lowest order in the transferred wave vector \mathbf{q} the bare vertex computes to

$$\gamma(\mathbf{q}, \mathbf{k}) = \Psi_{c\mathbf{k}+\mathbf{q}}^\dagger \Psi_{v\mathbf{k}} = \frac{1}{2E_{\mathbf{k}}} \left[u_{\mathbf{k}}^2 q_- - \left(v_{\mathbf{k}} \frac{k_-}{k} \right)^2 q_+ \right]. \quad (12)$$

Note that $\gamma(\mathbf{q}, \mathbf{k})$ is linear in q because of the opposite parity of c and v eigenfunctions at $\mathbf{k}=\mathbf{0}$. Thus the linear q dependence stems from the off-diagonal momentum operator in the Hamiltonian (5). This is actually the cause of the q^2 dependence of the response function we referred to above in conjunction with Eq. (10) and in the Introduction. The “charges,” i.e., the vertices associated with the interaction, are given by

$$g_1(\mathbf{k}, \mathbf{k}') = \Psi_{c\mathbf{k}}^\dagger \Psi_{c\mathbf{k}'} = u_{\mathbf{k}} u_{\mathbf{k}'} + \frac{k_-}{k} v_{\mathbf{k}} v_{\mathbf{k}'} \frac{k'_+}{k'}, \quad (13a)$$

$$g_2(\mathbf{k}, \mathbf{k}') = \Psi_{c\mathbf{k}}^\dagger \Psi_{v\mathbf{k}'} = \frac{k_-}{k} v_{\mathbf{k}} u_{\mathbf{k}'} - u_{\mathbf{k}} v_{\mathbf{k}'} \frac{k'_-}{k'}. \quad (13b)$$

All other possible combinations of valence- and conduction-

band states differ from Eqs. (12) and (13) only by sign changes or complex conjugation.

From now on we use a diagram technique with two different types of full lines representing conduction- and valence-band states. With these lines we associate the scalar Green functions of Eq. (11). Vertices are associated with the scalar functions of Eqs. (12) and (13).

There is an alternative interpretation of this basis transformation. Consider one of the traces that has to be calculated for χ_{QP} ,

$$\text{tr}(\Psi_{c\mathbf{k}+\mathbf{q}} \Psi_{c\mathbf{k}+\mathbf{q}}^\dagger)(\Psi_{v\mathbf{k}} \Psi_{v\mathbf{k}}^\dagger), \quad (14)$$

where parentheses show the grouping of the matrix multiplication. One thus has to calculate the outer products of two vectors, multiply the resulting matrices, and in the end take the trace. However, this grouping can be changed as follows:

$$\text{tr} \Psi_{c\mathbf{k}+\mathbf{q}} (\Psi_{c\mathbf{k}+\mathbf{q}}^\dagger \Psi_{v\mathbf{k}}) \Psi_{v\mathbf{k}}^\dagger = (\Psi_{c\mathbf{k}+\mathbf{q}}^\dagger \Psi_{v\mathbf{k}}) (\Psi_{v\mathbf{k}}^\dagger \Psi_{c\mathbf{k}+\mathbf{q}}). \quad (15)$$

Now one computes inner products of two vectors and multiplies the resulting *scalar* functions. Taking the trace is accounted for automatically in Eq. (15). The change of the diagram technique outlined above is in effect a way to incorporate this regrouping into the formalism.

As noted above, working in the basis of the conduction- and valence-band states we have to split the BSE of Fig. 1(b) into two scalar equations for $\Gamma_{cv}(\omega, \mathbf{q}, \mathbf{k})$ and $\Gamma_{vc}(\omega, \mathbf{q}, \mathbf{k})$. Note that while Γ_{cv} and Γ_{vc} depend on two momenta, they depend only on one frequency for a frequency-independent interaction. These two three-point functions are in fact not independent but related by the replacement $\mathbf{q} \rightarrow -\mathbf{q}$, $\omega \rightarrow -\omega$ and complex conjugation. We can therefore derive one equation for Γ_{cv} :

$$\begin{aligned} \Gamma_{cv}(\omega, \mathbf{q}, \mathbf{k}) &= \Gamma_1(\omega, \mathbf{q}, \mathbf{k}) \\ &+ \sum_{\mathbf{k}'} V_{\mathbf{k}, \mathbf{k}'} \frac{g_1(\mathbf{k}, \mathbf{k}') \Gamma_{cv}(\omega, \mathbf{q}, \mathbf{k}') g_1^*(\mathbf{k}', \mathbf{k})}{2E_{\mathbf{k}'} - \omega} \\ &+ \sum_{\mathbf{k}'} V_{\mathbf{k}, \mathbf{k}'} \frac{g_2(\mathbf{k}, \mathbf{k}') \Gamma_{cv}^*(-\omega, -\mathbf{q}, \mathbf{k}') g_2(\mathbf{k}', \mathbf{k})}{2E_{\mathbf{k}'} + \omega}, \end{aligned} \quad (16a)$$

$$\begin{aligned} \Gamma_1(\omega, \mathbf{q}, \mathbf{k}) &= \sum_{\mathbf{k}'} V_{\mathbf{k}, \mathbf{k}'} \frac{g_1(\mathbf{k}, \mathbf{k}') \gamma(\mathbf{q}, \mathbf{k}') g_1^*(\mathbf{k}', \mathbf{k})}{2E_{\mathbf{k}'} - \omega} \\ &+ \sum_{\mathbf{k}'} V_{\mathbf{k}, \mathbf{k}'} \frac{g_2(\mathbf{k}, \mathbf{k}') \gamma^*(-\mathbf{q}, \mathbf{k}') g_2(\mathbf{k}', \mathbf{k})}{2E_{\mathbf{k}'} + \omega}. \end{aligned} \quad (16b)$$

Note that we omit the \mathbf{q} dependence in the “charges” g_1 and g_2 as well as in the energy denominators, as we are only interested in the lowest-order expansion in \mathbf{q} , which stems from the dipole matrix elements in the external vertices. Here $V_{\mathbf{k}, \mathbf{k}'}$ are the matrix elements of the interaction between the particles. Now we investigate this equation for two dif-

ferent types of interaction: a general short-range interaction and the long-ranged Coulomb interaction.

IV. A SHORT-RANGE INTERACTION

A. The solution of the BSE

In this section we solve Eq. (16) for a short-ranged interaction, i.e., an interaction with a characteristic length scale shorter than Δ^{-1} . The final results are expressed in terms of the physical (renormalized) scattering length, which includes the high-energy contribution to the integrals in Eq. (16). Having in mind this renormalization we can formally use a momentum independent bare interaction $V_{\mathbf{k},\mathbf{k}'} = V$ in Eq. (16). Note that this does not make f_{xc}^{Ex} trivially equal to the interaction itself as discussed in the last paragraph of Sec. II, because for our two-band model the annihilation channel and the ladder channel remain different even for a contact interaction. From Eq. (16) we see, that the \mathbf{k} dependence of Γ_{cv} is given by the ‘‘charges’’ from Eq. (13) and has the same general form as for the bare vertex (12). The same is true for the \mathbf{q} dependence. Therefore we can write the following ansatz for Γ_{cv} :

$$\Gamma_{cv}(\omega, \mathbf{q}, \mathbf{k}) = u_k^2 q_- \Gamma_{cv}^{(s)}(\omega) + \left(v_k \frac{k_-}{k} \right)^2 q_+ \Gamma_{cv}^{(d)}(\omega). \quad (17)$$

Inserting this ansatz into Eq. (16) we obtain two coupled equations:

$$\begin{aligned} \Gamma_{cv}^{(s)}(\omega) &= V \sum_{\mathbf{k}} \frac{1}{2E_k} \left(\frac{u_k^4}{(2E_k - \omega)} - \frac{v_k^4}{(2E_k + \omega)} \right) \\ &+ V \sum_{\mathbf{k}} \frac{u_k^4}{2E_k - \omega} \Gamma_{cv}^{(s)}(\omega) + V \sum_{\mathbf{k}} \frac{v_k^4}{2E_k + \omega} \Gamma_{cv}^{(d)*}(-\omega), \end{aligned} \quad (18a)$$

$$\begin{aligned} \Gamma_{cv}^{(d)}(\omega) &= V \sum_{\mathbf{k}} \frac{1}{2E_k} \left(\frac{u_k^4}{(2E_k + \omega)} - \frac{v_k^4}{(2E_k - \omega)} \right) \\ &+ V \sum_{\mathbf{k}} \frac{v_k^4}{2E_k - \omega} \Gamma_{cv}^{(d)}(\omega) + V \sum_{\mathbf{k}} \frac{u_k^4}{2E_k + \omega} \Gamma_{cv}^{(s)*}(-\omega). \end{aligned} \quad (18b)$$

From these equations it immediately follows that $\Gamma_{cv}^{(s)}(\omega) = \Gamma_{cv}^{(d)*}(-\omega)$ and the equation to solve reduces to

$$\begin{aligned} \Gamma_{cv}^{(s)}(\omega) &= V \sum_{\mathbf{k}} \frac{1}{2E_k} \left(\frac{u_k^4}{(2E_k - \omega)} - \frac{v_k^4}{(2E_k + \omega)} \right) \\ &+ V \sum_{\mathbf{k}} \left(\frac{u_k^4}{2E_k - \omega} + \frac{v_k^4}{2E_k + \omega} \right) \Gamma_{cv}^{(s)}(\omega) \\ &=: V\gamma_1(\omega) + VK_0(\omega)\Gamma_{cv}^{(s)}(\omega). \end{aligned} \quad (19)$$

At this point it is convenient to perform the above-mentioned renormalization of the interaction by splitting K_0 into a low- and a high-energy part, the latter being equal to $\Sigma_{\mathbf{k}} 1/(4E_k)$. The high-energy part logarithmically diverges at

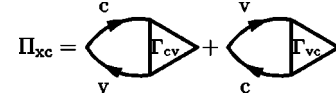


FIG. 6. The diagrammatic expression for Π_{xc} .

large k . This divergence can be removed by the standard renormalization of the interaction

$$\tilde{V} = \frac{V}{1 - V \sum_{\mathbf{k}} 1/(4E_k)} =: \frac{4\pi}{\Delta} a \quad (20)$$

where we introduce the dimensionless scattering length a . With this renormalized interaction, $\Gamma_{cv}^{(s)}(\omega)$ fulfills the following equation:

$$\Gamma_{cv}^{(s)}(\omega) = \tilde{V}\gamma_1(\omega) + \tilde{V}\tilde{K}_0(\omega)\Gamma_{cv}^{(s)}(\omega) \quad (21)$$

with

$$\tilde{K}_0(\omega) = K_0(\omega) - \sum_{\mathbf{k}} \frac{1}{4E_k} = \sum_{\mathbf{k}} \frac{(2\Delta + \omega)^2}{4E_k(4E_k^2 - \omega^2)}. \quad (22)$$

Equation (21) has the obvious solution

$$\Gamma_{cv}^{(s)}(\omega) = \frac{\tilde{V}\gamma_1(\omega)}{1 - \tilde{V}\tilde{K}_0(\omega)}. \quad (23)$$

Calculating the 2D integrals in Eq. (22) and $\gamma_1(\omega)$ in Eq. (19), we obtain

$$\tilde{K}_0(\tilde{\omega}) = \frac{\Delta}{4\pi} \frac{(1 + \tilde{\omega})^2}{4\tilde{\omega}} \ln \frac{1 + \tilde{\omega}}{1 - \tilde{\omega}} =: \frac{\Delta}{4\pi} F(\tilde{\omega}) \quad (24)$$

and

$$\gamma_1(\tilde{\omega}) = \frac{1}{16\pi} \left(\frac{(1 + \tilde{\omega})^2}{2\tilde{\omega}^2} \ln \frac{1 + \tilde{\omega}}{1 - \tilde{\omega}} - \frac{1}{\tilde{\omega}} \right) = \frac{1}{16\pi} \frac{2}{\tilde{\omega}} \left(F(\tilde{\omega}) - \frac{1}{2} \right). \quad (25)$$

Inserting these integrals into the solution (21) we can compute the response function's xc part $\Pi_{xc} = \tilde{\chi} - \chi_{\text{QP}}$ and its first-order approximation $\Pi_{xc}^{(1)}$. In diagrammatic form Π_{xc} is displayed in Fig. 6. The results are

$$\Pi_{xc}(\tilde{\omega}, \mathbf{q}) = - \frac{q^2}{16\pi\Delta} \frac{a}{\tilde{\omega}^2} \left(\frac{[F(\tilde{\omega}) - \frac{1}{2}]^2}{1 - aF(\tilde{\omega})} + \frac{[F^*(-\tilde{\omega}) - \frac{1}{2}]^2}{1 - aF^*(-\tilde{\omega})} \right) \quad (26)$$

and

$$\Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q}) = - \frac{q^2}{16\pi\Delta} \frac{a}{\tilde{\omega}^2} \left\{ \left[F(\tilde{\omega}) - \frac{1}{2} \right]^2 + \left[F^*(-\tilde{\omega}) - \frac{1}{2} \right]^2 \right\}, \quad (27)$$

where the function $F(\omega)$ is defined in Eq. (24).

The equation for Λ given in Fig. 2(b) can be solved in a similar fashion, which allows us then to calculate the excitonic part of the exact xc kernel. However, it is easier to obtain f_{xc}^{Ex} directly from Eq. (2) and the exact response function $\tilde{\chi} = \chi_{\text{QP}} + \Pi_{xc}$ with Π_{xc} from Eq. (26):

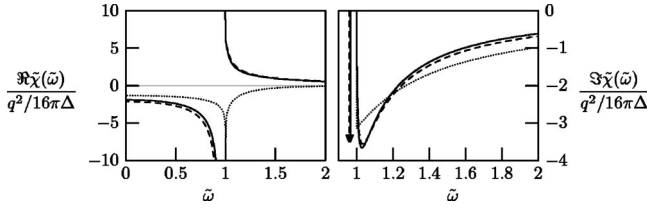


FIG. 7. The real and the imaginary parts of the exact $\tilde{\chi}$ (full) and the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ (dashes) for the scattering length $a=0.25$. The real and the imaginary part of χ_{QP} are shown by the dotted line.

$$f_{xc}^{\text{Ex}}(\tilde{\omega}, \mathbf{q}) = \frac{\chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q}) \Pi_{xc}(\tilde{\omega}, \mathbf{q}) \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q})}{1 + \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q}) \Pi_{xc}(\tilde{\omega}, \mathbf{q})}. \quad (28)$$

From here the first-order approximation to f_{xc}^{Ex} immediately follows:

$$f_{xc}^{\text{Ex}(1)}(\tilde{\omega}, \mathbf{q}) = \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q}) \Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q}) \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q}). \quad (29)$$

Inserting this $f_{xc}^{\text{Ex}(1)}$ in Eq. (3b) we arrive at an approximate solution for the response function's xc part

$$\Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q}) = \frac{\Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q})}{1 - \Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q}) \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q})}. \quad (30)$$

Note that although $f_{xc}^{\text{Ex}(1)}$ is based on $\Pi_{xc}^{(1)}$, this formula does *not* coincide with $\Pi_{xc}^{(1)}$. In accordance with Eq. (3b) it accounts for an infinite series of diagrams instead. This way, the excitonic pole in Π_{xc} , Eq. (26), which has been lost in $\Pi_{xc}^{(1)}$, Eq. (27), reappears in Eq. (30).

B. Results

Having calculated the exact and the approximate expressions for the excitonic part of the xc kernel and for the response function we are now in the position to compare these results. Let us start with the excitonic peak in the absorption spectrum. In both the exact and the approximate response functions the excitonic peak originates from the divergence of the xc part. The exact Π_{xc} in Eq. (26) has a pole at

$$1 - aF(\tilde{\omega}) = 0 \quad (31)$$

for positive a and $0 \leq \tilde{\omega} \leq 1$. When the dimensionless exciton binding energy $\tilde{\varepsilon} = 1 - \tilde{\omega}$ is small, Eq. (31) has an approximate solution

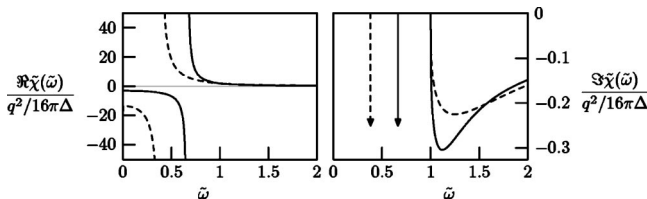


FIG. 8. The real and the imaginary parts of the exact $\tilde{\chi}$ (full) and the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ (dashes) for the scattering length $a=0.6$.

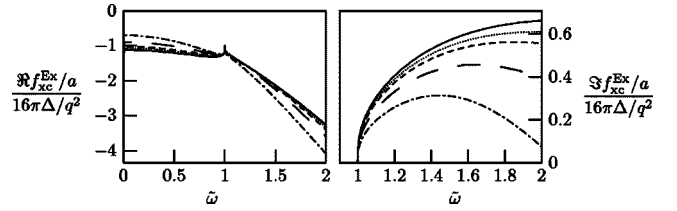


FIG. 9. The real and the imaginary part of the exact f_{xc}^{Ex}/a for $a=0.05$ (dots), $a=0.1$ (dashes), $a=0.25$ (long dashes), and $a=0.6$ (dot-dashes) compared with $f_{xc}^{\text{Ex}(1)}/a$ (full).

$$\tilde{\varepsilon} = 2 \exp\left(-\frac{1}{a}\right). \quad (32)$$

The approximate response function of Eq. (30) which is based on $f_{xc}^{\text{Ex}(1)}$ has a pole

$$1 - \chi_{\text{QP}}^{-1}(\tilde{\omega}, \mathbf{q}) \Pi_{xc}^{(1)}(\tilde{\omega}, \mathbf{q}) = 0. \quad (33)$$

Similarly to the solution (32) of the exact equation (31), for a small binding energy $\tilde{\varepsilon}$ this equation can be approximately solved by

$$\tilde{\varepsilon} = 2 \exp\left(-\frac{1}{a} + \frac{1-a-\sqrt{1-2a-a^2}}{2a}\right), \quad (34)$$

which differs from the exact result (32). However, e.g., at $a=0.2$ this is only about 14% larger than the exact solution (32), which gives $\tilde{\varepsilon} \approx 0.013$. For a smaller scattering length and therefore a smaller binding energy the agreement between the exact result and the one based on $f_{xc}^{\text{Ex}(1)}$ is even better.

In Fig. 7 the real and the imaginary part of the exact $\tilde{\chi}$ and the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ are shown for the scattering length $a=0.25$. For comparison, χ_{QP} is also displayed. One clearly observes a very good agreement between the exact and the approximate response functions. Both main features, the enhancement of the imaginary part (the Sommerfeld factor) and the excitonic peak are correctly reproduced. The latter is indicated by the arrows on the plots of the imaginary part. The apparent difference between the exact and the approximate response function is the exact position of the excitonic peak as described above. The oscillator strengths should also be different, which is not, however, reflected in the figure.

From the above discussion one could conjecture that the approximation for the response function based on $f_{xc}^{\text{Ex}(1)}$ is always sufficient. This is, of course, not true, as we can see from Fig. 8, where the approximate and the exact response function are compared for the larger value of the scattering length $a=0.6$. With this large interaction³⁴ the position of the excitonic peak in the approximate response function is clearly wrong. In addition, the imaginary part has the wrong magnitude above the band gap.

To uncover the background of the good agreement between the exact and the approximate response functions for a “weak” interaction, let us compare $f_{xc}^{\text{Ex}(1)}$ with the exact f_{xc}^{Ex} for different interaction strengths. We note first that $f_{xc}^{\text{Ex}(1)}$ is proportional to the scattering length a . Therefore, more adequate is to compare $f_{xc}^{\text{Ex}(1)}/a$ with f_{xc}^{Ex}/a , as done in Fig. 9.

It is clearly visible that $f_{xc}^{\text{Ex}^{(1)}}$ is a very good approximation to the exact f_{xc}^{Ex} in a frequency range close to the band gap. Actually, for $\tilde{\omega}=1$ we obtain from our exact expressions

$$f_{xc}^{\text{Ex}^{(1)}}(\tilde{\omega}=1) = -\frac{16\pi\Delta}{q^2}a \quad (35)$$

and

$$f_{xc}^{\text{Ex}}(\tilde{\omega}=1) = -\frac{16\pi\Delta}{q^2} \frac{a}{1-a^2/4}. \quad (36)$$

Thus even for a very strong interaction of $a=0.6$ where $f_{xc}^{\text{Ex}^{(1)}}$ leads to the rather poor response function of Fig. 8, there is only a 10% error in $f_{xc}^{\text{Ex}^{(1)}}$ at $\tilde{\omega}=1$. In the static case the errors are larger, as we get for $\tilde{\omega}=0$:

$$f_{xc}^{\text{Ex}^{(1)}}(\tilde{\omega}=0) = -\frac{16\pi\Delta}{q^2} \frac{9}{8}a \quad (37)$$

and

$$f_{xc}^{\text{Ex}}(\tilde{\omega}=0) = -\frac{16\pi\Delta}{q^2} \frac{9}{8} \frac{a}{1+a}. \quad (38)$$

Here a 10% error is already reached for $a=0.1$, as the ratio of f_{xc}^{Ex} and $f_{xc}^{\text{Ex}^{(1)}}$ is of order a , whereas it is of order a^2 at $\tilde{\omega}=1$.

It is also interesting to look at the difference $\delta f_{xc}^{\text{Ex}}$ between the exact f_{xc}^{Ex} and $f_{xc}^{\text{Ex}^{(1)}}$ for different interaction strengths. Since the leading order term in $\delta f_{xc}^{\text{Ex}}$ is of the order a^2 , we normalize these differences by a^2 , when plotting them in Fig. 10. The general behavior of the different curves is quite similar, which indicates that it is mostly the second-order approximation $f_{xc}^{\text{Ex}^{(2)}}$ that contributes to $\delta f_{xc}^{\text{Ex}}$. One may be surprised that the curves for stronger interaction are closer to 0 for $\tilde{\omega}<1$, which might indicate that $f_{xc}^{\text{Ex}^{(1)}}$ is a better approximation for *stronger* interactions. However, this is not the case. This only tells us that $f_{xc}^{\text{Ex}^{(3)}}$ is positive for $\tilde{\omega}<1$.

C. The validity of the first-order approximation

From the preceding section we can conclude that $f_{xc}^{\text{Ex}^{(1)}}$ is a good approximation to f_{xc}^{Ex} close to the band gap $\tilde{\omega}=1$ practically for any interaction strength. Even for a strong

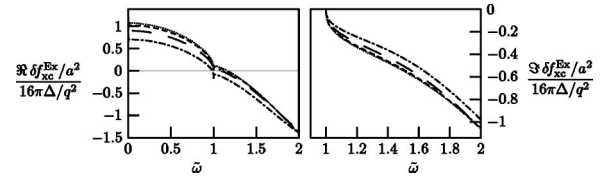


FIG. 10. The real and the imaginary parts of $\delta f_{xc}^{\text{Ex}}/a^2 = (f_{xc}^{\text{Ex}} - f_{xc}^{\text{Ex}^{(1)}})/a^2$ for $a=0.05$ (dots), $a=0.1$ (dashes), $a=0.25$ (long dashes), and $a=0.6$ (dot-dashes).

interaction, where the response function stemming from $f_{xc}^{\text{Ex}^{(1)}}$ is quite wrong, the first-order kernel $f_{xc}^{\text{Ex}^{(1)}}$ is still very good close to $\tilde{\omega}=1$. This is the reason for the success of $f_{xc}^{\text{Ex}^{(1)}}$ in describing the bound excitonic states as shown in the preceding section. Since around $\tilde{\omega}=1$ the first-order kernel $f_{xc}^{\text{Ex}^{(1)}}$ is a good approximation, if the (exact) bound excitonic state lies within this region, $f_{xc}^{\text{Ex}^{(1)}}$ will describe it correctly. If, however, the binding energy is outside this region, $f_{xc}^{\text{Ex}^{(1)}}$ will fail. Note that as seen from Fig. 9 this region gets smaller as the interaction increases and at the same time the binding energy of the exciton increases. Hence the error in the exciton binding energy increases with the increase of the interaction strength.

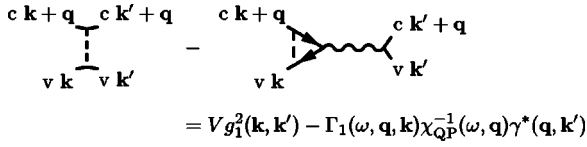
Can we understand the good agreement between $f_{xc}^{\text{Ex}^{(1)}}$ and f_{xc}^{Ex} for $\tilde{\omega} \approx 1$ in terms of the integral equation for Λ ? For this we explicitly calculate the diagrams involved in the replacement shown in Fig. 3(b). Working with separate Green functions for the valence- and conduction-band states we have four possible combinations of the external lines and can therefore split this replacement into four parts according to these combinations. At energies close to the band gap the two-particle propagator with a conduction-band state in the upper line should dominate. Hence, the most important part of this replacement is the diagram with a conduction-band state in the upper line on both sides of the graphs. This diagram together with its translation into quantities introduced in the preceding section is displayed in Fig. 11. Note that we use the bare interaction V here, as the interaction renormalization in these diagrams means simply a replacement $V \rightarrow \tilde{V}$.

Inserting $g_1, \Gamma_1, \chi_{\text{QP}}$, and γ in the expression in Fig. 11 and taking the limit $\tilde{\omega} \rightarrow 1$ we obtain:

$$\begin{aligned} & V g_1^2(\mathbf{k}, \mathbf{k}') - \Gamma_1(\tilde{\omega}=1, \mathbf{q}, \mathbf{k}) \chi_{\text{QP}}^{-1}(\tilde{\omega}=1, \mathbf{q}) \gamma^*(\mathbf{q}, \mathbf{k}') \\ &= V \left[u_k^2 u_{k'}^2 \left(1 - \frac{\Delta}{E_{k'}} \right) + \left(\frac{k_-}{k} v_k \right)^2 \left(\frac{k'_+}{k'} v_{k'} \right)^2 + 2 \frac{k_-}{k} v_k u_k u_{k'} v_{k'} \frac{k'_+}{k'} + u_k^2 \left(\frac{k'_+}{k'} v_{k'} \right)^2 \left(\frac{q_-}{q} \right)^2 \frac{\Delta}{E_{k'}} \right]. \quad (39) \end{aligned}$$

Not surprisingly the whole expression (39) is proportional to the interaction V (or \tilde{V} after renormalization). In the proportionality factor in square brackets all summands contain (1

$-\Delta/E_{k'}$) (or powers thereof). For the first summand this is directly visible, in the others it is “hidden” in $v_{k'}$. Let us look at this factor more closely. When the diagram of Fig. 11 is



$$= V g_1^2(\mathbf{k}, \mathbf{k}') - \Gamma_1(\omega, \mathbf{q}, \mathbf{k}) \chi_{QP}^{-1}(\omega, \mathbf{q}) \gamma^*(\mathbf{q}, \mathbf{k}')$$

FIG. 11. Explicit calculation of one replacement. Here ω and \mathbf{q} are the transferred energy and momentum.

inserted into the equation for Λ of Fig. 2(b) the integration over \mathbf{k}' has to be performed. The main contribution to this integral comes from small momenta, due to the small energy denominators in the particle-hole propagator. However, for these small momenta the expression (39) is small as $\Delta \approx E_{k'}$.

We can therefore conclude that close to the band gap the kernel in the equation for Λ is indeed small for those states that mainly contribute to the integral. Thus we explicitly observe the cancellation we qualitatively discussed in Sec. II. This cancellation explains the excellent agreement between f_{xc}^{Ex} and $f_{xc}^{\text{Ex}(1)}$ for $\tilde{\omega} \approx 1$.

A complimentary interpretation can be obtained from looking at the diagrammatic expansion of the response functions. In Fig. 12 the third-order term in the expansion of the exact $\tilde{\chi}$ and the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ is displayed. The difference between the two expressions is similar to the replacements discussed in Sec. II. This remains true in all orders of the perturbation theory. From the previous calculation of the effect of this replacement follows, that the exact $\tilde{\chi}$ and the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ are almost identical in the vicinity of the band gap, independently of the position of the excitonic peak.

V. THE COULOMB INTERACTION

In this section, we consider the 2D Dirac model of Eq. (4) with the Coulomb interaction between the particles, i.e.,

$$V(\mathbf{r} - \mathbf{r}') = \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}, \quad (40)$$

where e is the particle's charge. Note that Eq. (40) is the 3D interaction although our model system is 2D. We will not attempt to solve exactly the BSE with this interaction, but rather focus on the properties of shallow excitons. The weak binding limit $\tilde{\varepsilon} = 1 - \tilde{\omega} \ll 1$ is, in fact, a “nonrelativistic” limit where the BSE reduces to the two-particle Schrödinger equation.³⁰ The quasiparticle energy eigenvalues are approximately

$$E_k \approx \Delta + \frac{k^2}{2\Delta}, \quad (41)$$

and for the eigenvectors holds



FIG. 12. Third-order term in the expansion of (a) the exact $\tilde{\chi}$ and (b) the $\tilde{\chi}$ based on $f_{xc}^{\text{Ex}(1)}$ in terms of the interaction.

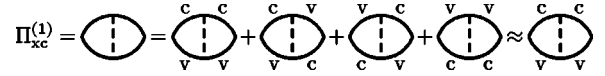


FIG. 13. Approximation for $\Pi_{xc}^{(1)}$ in the limit of small $\tilde{\varepsilon}$.

$$u_k \approx 1 \quad \text{and} \quad v_k \approx 0. \quad (42)$$

Solving the BSE thus reduces to solving the positronium problem in 2D for particles with a mass Δ , i.e., for a reduced mass $\Delta/2$. The response function can be written in the spectral representation as³¹

$$\begin{aligned} \tilde{\chi}(\omega, \mathbf{q}) &= |\gamma(\mathbf{q}, 0)|^2 \sum_n \frac{|\psi_n(r=0)|^2}{\omega - (2\Delta - \varepsilon_n)} \\ &\approx |\gamma(\mathbf{q}, 0)|^2 \frac{|\psi_0(r=0)|^2}{\omega - (2\Delta - \varepsilon_0)}, \end{aligned} \quad (43)$$

where the ψ_n are the eigenfunctions with eigenvalues ε_n (positive for bound states) of the above-mentioned Schrödinger equation. The approximation in Eq. (43) is valid close to the excitonic peak of the $1s$ ground state, in which we are interested in here. The $1s$ wave function needed in this approximation is given by

$$\psi_0(r) = \sqrt{\frac{2\Delta\varepsilon_0}{\pi}} \exp\{-\sqrt{\Delta\varepsilon_0}r\} \quad (44)$$

and the exciton binding energy is

$$\varepsilon_0 = e^4 \Delta. \quad (45)$$

Introducing dimensionless variables $\tilde{\varepsilon}$ and $\tilde{\varepsilon}_0 = \varepsilon_0/(2\Delta)$, we arrive at

$$\tilde{\chi}(\tilde{\varepsilon}, \mathbf{q}) \approx \frac{q^2}{2\pi\Delta} \frac{\tilde{\varepsilon}_0}{\tilde{\varepsilon}_0 - \tilde{\varepsilon}} \quad (46)$$

for $\tilde{\varepsilon}$ close to the excitonic peak of the $1s$ ground state.

We now want to compare the position and the oscillator strength of the excitonic peak of the $1s$ ground state with the results obtained from $f_{xc}^{\text{Ex}(1)}$. For this we first expand χ_{QP} of Eq. (10) for small $\tilde{\varepsilon}$,

$$\chi_{QP}(\tilde{\varepsilon}, \mathbf{q}) \approx \frac{q^2}{16\pi\Delta} \left[\ln\left(\frac{2}{\tilde{\varepsilon}}\right) - 1 \right]. \quad (47)$$

In order to obtain $f_{xc}^{\text{Ex}(1)}$ we also need an expression for $\Pi_{xc}^{(1)}$ in the limit of small $\tilde{\varepsilon}$. In diagrammatic form $\Pi_{xc}^{(1)}$ is given by a single graph as displayed in Fig. 13. There are four possibilities to distribute the conduction- and valence-band states in the two particle-hole propagators. The particle-hole propagators with a conduction-band state in the upper line and a valence-band state in the lower line diverge at $\tilde{\omega}=1$. The particle-hole propagators with a valence-band state in the upper line and a conduction-band state in the lower line diverge at $\tilde{\omega}=-1$. Therefore, in the limit where the energy goes to the band gap (or equivalently $\tilde{\varepsilon}$ is small), the main contribution to $\Pi_{xc}^{(1)}$ comes from the diagram with both particle-hole propagators of the upper line belonging to the conduction-band state and the lower line to the valence-band state. Neglecting the other contributions is our first approxi-

mation in calculating $\Pi_{xc}^{(1)}$. The other approximation is using Eqs. (41) and (42). With these approximations we obtain

$$\Pi_{xc}^{(1)}(\tilde{\epsilon}, \mathbf{q}) \approx \frac{|\gamma(\mathbf{q}, 0)|^2}{(2\Delta)^2} \sum_{\mathbf{k}, \mathbf{k}'} \frac{V(\mathbf{k} - \mathbf{k}')}{\left[\frac{1}{2}(k/\Delta)^2 + \tilde{\epsilon}\right] \left[\frac{1}{2}(k'/\Delta)^2 + \tilde{\epsilon}\right]}, \quad (48)$$

where $V(\mathbf{k} - \mathbf{k}') = 2\pi e^2/|\mathbf{k} - \mathbf{k}'|$ is the Fourier representation of $V(\mathbf{r} - \mathbf{r}')$ from Eq. (40). The easiest way to solve this double integral is to look at it in the real space, instead of the Fourier space, where it becomes a single integral. The 2D Fourier transformation of the particle-hole propagator $[\frac{1}{2}(k/\Delta)^2 + \tilde{\epsilon}]^{-1}$ gives the modified Bessel function of the second kind K_0 , so that we can write

$$\Pi_{xc}^{(1)}(\tilde{\epsilon}, \mathbf{q}) \approx \frac{e^2 q^2}{8\pi\Delta\sqrt{2\tilde{\epsilon}}} \int_0^\infty d\rho K_0^2(\rho). \quad (49)$$

The integral over K_0^2 computes to $(\pi/2)^2$, and we arrive at

$$\Pi_{xc}^{(1)}(\tilde{\epsilon}, \mathbf{q}) \approx \frac{q^2}{16\pi\Delta} \frac{\pi^2}{2} \sqrt{\frac{\tilde{\epsilon}_0}{\tilde{\epsilon}}}. \quad (50)$$

Now we have all ingredients to build $f_{xc}^{\text{Ex}(1)} = \chi_{\text{QP}}^{-1} \Pi_{xc}^{(1)} \chi_{\text{QP}}^{-1}$ and insert it in Eq. (3b). This gives the following approximation for the response function based on $f_{xc}^{\text{Ex}(1)}$:

$$\begin{aligned} \tilde{\chi}^{(1)}(\tilde{\epsilon}, \mathbf{q}) &= \frac{\chi_{\text{QP}}(\tilde{\epsilon}, \mathbf{q})}{1 - \Pi_{xc}^{(1)}(\tilde{\epsilon}, \mathbf{q}) \chi_{\text{QP}}^{-1}(\tilde{\epsilon}, \mathbf{q})} \\ &= \frac{q^2}{16\pi\Delta} \frac{[\ln(2/\tilde{\epsilon}) - 1]}{1 - [\ln(2/\tilde{\epsilon}) - 1]^{-1} (\pi^2/2) \sqrt{\tilde{\epsilon}_0/\tilde{\epsilon}}}. \end{aligned} \quad (51)$$

Note that this equation is valid only for small $\tilde{\epsilon}$, since we derived χ_{QP} and $\Pi_{xc}^{(1)}$ only for small values of $\tilde{\epsilon}$.

From Eq. (51) we clearly see that $\tilde{\chi}^{(1)}$ contains an additional pole where the denominator vanishes. This is the excitonic peak in this approximation and its energy $\tilde{\epsilon}'_0$ is defined by

$$1 - \left[\ln\left(\frac{2}{\tilde{\epsilon}'_0}\right) - 1 \right]^{-1} \frac{\pi^2}{2} \sqrt{\frac{\tilde{\epsilon}_0}{\tilde{\epsilon}'_0}} = 0. \quad (52)$$

To evaluate the approximation used let us reformulate Eq. (52) as

$$\frac{e^4}{2} = \tilde{\epsilon}_0 = 2 \frac{\tilde{\epsilon}_0}{\tilde{\epsilon}'_0} \exp\left[- \left(\frac{\pi^2}{2} \sqrt{\frac{\tilde{\epsilon}_0}{\tilde{\epsilon}'_0}} + 1 \right) \right]. \quad (53)$$

We see that the approximate binding energy $\tilde{\epsilon}'_0$ and the exact binding energy $\tilde{\epsilon}_0$ are identical for

$$\frac{e^4}{2} = \tilde{\epsilon}_0 = 2 \exp\left[- \left(\frac{\pi^2}{2} + 1 \right) \right] \approx \frac{1}{189}. \quad (54)$$

In other words, for the interaction strength that corresponds to this quite realistic exciton binding energy, the approximation using $f_{xc}^{\text{Ex}(1)}$ gives the exact binding energy. For stronger and *weaker* interaction strength there is some error. To be

more precise, the error is below 10% for $\tilde{\epsilon}_0$ between 1/165 and 1/222. It is below 20% for $\tilde{\epsilon}_0$ between 1/148 and 1/271. Note that in the range of energies where the approximation gives good results, $\tilde{\epsilon}$ is indeed small, so that our approximation is consistent. It is interesting to note that for the Coulomb interaction—in contrast to a short-range interaction—one does not get the correct exciton binding energy from $f_{xc}^{\text{Ex}(1)}$ in the limit of the interaction strength going to zero. Note also that while the exact solution of the 2D hydrogen problem gives rise to an infinite series of excitonic peaks in the exact response function, we obtain only *one* excitonic peak in the approximate approach. This problem was already touched upon in the Introduction. Although $f_{xc}^{\text{Ex}(1)}$ is frequency dependent, it does not contain the rapid oscillations needed to describe the whole series of excitonic states.

To compare the approximate response function of Eq. (51) to the exact one from Eq. (46) and determine the residual it is best to expand $\tilde{\chi}^{(1)}$ around $\tilde{\epsilon}'_0$. This can be done by calculating the first-order Taylor expansion of the denominator in Eq. (51):

$$\tilde{\chi}^{(1)}(\tilde{\epsilon}, \mathbf{q}) \approx \frac{q^2}{16\pi\Delta} \frac{(\pi^2/2) \sqrt{\tilde{\epsilon}_0/\tilde{\epsilon}'_0}}{\frac{1}{2} - (2/\pi^2) \sqrt{\tilde{\epsilon}'_0/\tilde{\epsilon}_0}} \frac{\tilde{\epsilon}'_0}{\tilde{\epsilon} - \tilde{\epsilon}'_0}. \quad (55)$$

In the case where the exciton energy is exactly reproduced, i.e., for $\tilde{\epsilon}'_0 = \tilde{\epsilon}_0$, the second fraction in Eq. (55) is about 16.6. The oscillator strength is thus too large by a factor of 2 in spite of the fact that the binding energy is exact. The value of this second fraction is almost constant for a large range of ratios $\tilde{\epsilon}'_0/\tilde{\epsilon}_0$, so that mainly the $\tilde{\epsilon}'_0$ in the numerator of the third fraction in Eq. (55) produces additional errors in the oscillator strength. The larger oscillator strength is not too surprising, as the excitonic oscillator strength is distributed over more peaks in the exact response function.

Similar to the results for the short-range interaction of the preceding section, we observe that the cancellation effects are also effective for the long-ranged Coulomb interaction. However, the general structure of this cancellation is different and cannot be universally characterized. Hence the accuracy of $f_{xc}^{\text{Ex}(1)}$ has to be checked for every particular system. The integral equation for Λ together with the replacement procedure of Fig. 3 provides an appropriate tool for this task.

VI. SUMMARY AND CONCLUSIONS

We have investigated the excitonic effects within the linear response TDDFT. In general, two different types of correlation effects are important for the response function. These are the excitonic and the quasiparticle effects, which are both absent in the adiabatic local density approximation. Within TDDFT both types of correlations are contained in the exchange-correlation kernel f_{xc} . As we have shown in Sec. II, the xc kernel can be unambiguously split into a quasiparticle part f_{xc}^{QP} and an excitonic part f_{xc}^{Ex} , which separately describe these two effects. We focus on the excitonic part in this paper and did not consider the quasiparticle part. The latter is usually accounted for by using the *GW* approximation.

In the absorption spectrum of semiconductors and insulators the two-particle excitonic correlations produce two types of effects: sharp peaks in the band gap stemming from bound excitons and broad peaks above the band gap due to the Coulomb correlation between the electron and the hole (unbound excitons). These effects are commonly described by the BSE within standard many-body theory. Solving the BSE is very demanding in the general case, though, TDDFT could provide a significant simplification, if a good approximation for f_{xc} is available. To link TDDFT with the BSE we have expressed the excitonic part f_{xc}^{Ex} in terms of an effective vertex function Λ , for which we diagrammatically derived an integral equation. This equation is structurally similar to the BSE and establishes the connection between the common many-body theory and TDDFT. However, solving this integral equation is at least as difficult as solving the BSE. Nevertheless this approach has a certain advantage. Namely, the kernel of the integral equation for Λ shows the possibility of cancellation effects. If the cancellation were complete, f_{xc}^{Ex} would be simply equal to the first-order approximation $f_{xc}^{\text{Ex}^{(1)}}$. This suggests that in some situations $f_{xc}^{\text{Ex}^{(1)}}$ can provide a good substitute for f_{xc}^{Ex} . Generally, the possibility to use $f_{xc}^{\text{Ex}^{(1)}}$ instead of f_{xc}^{Ex} is a tremendous simplification, as there would be *no need to solve an integral equation* to obtain f_{xc}^{Ex} in this case. It is therefore crucial for the application of TDDFT to excitonic effects to understand the nature and to develop a criterion for the cancellation effects.

We approached this question using analytic calculations for a model two-band semiconductor with different interactions. For a short-ranged interaction we computed the exact response function as well as the response function derived from $f_{xc}^{\text{Ex}^{(1)}}$. Comparing these two functions we find a very good agreement for a weak interaction, where both the binding energy of the shallow exciton and the Sommerfeld factor stemming from unbound excitons are correctly described by the first-order kernel. We were able to trace this to the strong cancellation in the kernel of the equation for Λ for energies close to the band gap. Calculations with the Coulomb interaction give similar results, though the deficiencies of $f_{xc}^{\text{Ex}^{(1)}}$ are somewhat worse in this case. The cancellation effects occur in both cases, and we were able to explicitly calculate them. For a given model we gave a criterion of the cancellation, yet it is not possible to make a generalization for an arbitrary system.

These calculations represent an example that the integral equation for Λ can serve as a tool for evaluation of the validity of low-order approximations to f_{xc}^{Ex} . Testing the predictive power of this approach for other systems will be the subject of future work.

ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft under Grant No. PA 516/2-3.

*Email address: ralf.stubner@physik.uni-erlangen.de

- ¹A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N.J., 1963).
- ²W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980).
- ³G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
- ⁴E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
- ⁵M. Petersilka, U. J. Gossmann, and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).
- ⁶A. Zangwill and P. Soven, Phys. Rev. A **21**, 1561 (1980).
- ⁷M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, J. Chem. Phys. **108**, 4439 (1998).
- ⁸S. J. A. van Gisbergen, F. Kootstra, P. R. T. Schipper, O. V. Gritsenko, J. G. Snijders, and E. J. Baerends, Phys. Rev. A **57**, 2556 (1998).
- ⁹T. Grabo, M. Petersilka, and E. K. U. Gross, J. Mol. Struct.: THEOCHEM **501–502**, 353 (2000).
- ¹⁰V. I. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 4343 (1997).
- ¹¹Y.-H. Kim, M. Städele, and A. Görling, Int. J. Quantum Chem. **91**, 257 (2003).
- ¹²G. Vignale, Phys. Lett. A **209**, 206 (1995).
- ¹³X. Gonze, P. Ghosez, and R. W. Godby, Phys. Rev. Lett. **74**, 4035 (1995).
- ¹⁴P. Ghosez, X. Gonze, and R. W. Godby, Phys. Rev. B **56**, 12811 (1997).
- ¹⁵L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett.

- 88**, 066404 (2002).
- ¹⁶S. Botti, F. Sottile, N. Vast, V. Olevano, L. Reining, H.-C. Weisker, A. Rubio, G. Onida, R. Del Sole, and R. W. Godby, Phys. Rev. B **69**, 155112 (2004).
- ¹⁷Y.-H. Kim and A. Görling, Phys. Rev. B **66**, 035114 (2002).
- ¹⁸Y.-H. Kim and A. Görling, Phys. Rev. Lett. **89**, 096402 (2002).
- ¹⁹P. L. de Boeij, F. Kootstra, J. A. Berger, R. van Leeuwen, and J. G. Snijders, J. Chem. Phys. **115**, 1995 (2001).
- ²⁰G. Vignale and W. Kohn, Phys. Rev. Lett. **77**, 2037 (1996).
- ²¹G. Vignale and W. Kohn, in *Electronic Density Functional Theory*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum Press, New York, 1998), pp. 199–216.
- ²²F. Kootstra, P. L. de Boeij, and J. G. Snijders, Phys. Rev. B **62**, 7071 (2000).
- ²³S. L. Adler, Phys. Rev. **126**, 413 (1962).
- ²⁴N. Wiser, Phys. Rev. **129**, 62 (1963).
- ²⁵F. Sottile, V. Olevano, and L. Reining, Phys. Rev. Lett. **91**, 056402 (2003).
- ²⁶G. Adragna, R. Del Sole, and A. Marini, Phys. Rev. B **68**, 165108 (2003).
- ²⁷A. Marini, R. Del Sole, and A. Rubio, Phys. Rev. Lett. **91**, 256402 (2003).
- ²⁸I. V. Tokatly, R. Stubner, and O. Pankratov, Phys. Rev. B **65**, 113107 (2002).
- ²⁹R. Del Sole, G. Adragna, V. Olevano, and L. Reining, Phys. Rev. B **67**, 045207 (2003).
- ³⁰L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966).

³¹G. D. Mahan, *Many-Particle Physics* (Plenum Press, New York, 1990), 2nd ed.

³²In systems where the ladder approximation is not applicable one would have to replace the screened interaction (here and in the BSE) with the scattering amplitude.

³³As noted above the BSE is normally formulated in terms of the four-point function T . Our integral equation for Λ can straight-

forwardly be rewritten in terms of four-point functions as well. However, in the context of the present work it is more convenient to use three-point functions.

³⁴The natural scale for the interaction strength in our system is given by the value $a=2$. At this value for a the exciton binding energy is equal to the band gap, which allows for a spontaneous creation of excitons (excitonic instability) in the system.