Excitonic optical spectrum of semiconductors obtained by time-dependent density-functional theory with the exact-exchange kernel

Yong-Hoon Kim and Andreas Görling
Lehrstuhl für Theoretische Chemie, Technische Universität München, D-85748 Garching, Germany
(Dated: October 30, 2018)

Applying the novel exact-exchange (EXX) Kohn-Sham method within time-dependent density-functional theory, we obtained the optical absorption spectrum of bulk silicon in good agreement with experiments including excitonic features. Analysis of the EXX kernel shows that inclusion of the Coulomb coupling of electron-hole pairs and the correct long-wavelength behavior in the kernel is crucial for the proper description of excitonic effects in semiconductors.

PACS numbers: 71.15.-m, 71.15.Mb, 71.35.-y, 78.20.-e

The calculation of optical spectra of solids has been traditionally one of the most important and challenging areas of first-principles material investigations. The interaction of excited electrons and holes plays a crucial role in optical excitations, however properly incorporating such effects \textit{ab initio} is theoretically and computationally very difficult. In the past several years there has been a major advance in the field \cite{1} based on the solution of the Bethe-Salpeter equation (BSE) \cite{2} for the two-particle Green’s function starting from quasiparticles obtained within the GW approximation (GWA) \cite{3}, but unfortunately the calculation scheme is by nature complicated and demanding.

One attractive alternative is extending density-functional theory (DFT) to the time-dependent (TD) case \cite{4}. TDDFT is a well-founded theory for the treatment of electronic excitations in general and, unlike the GWA-BSE route which is presently implemented as a post-DFT method, provides a computational scheme based directly on Kohn-Sham (KS) one-particle states. Moreover, once the KS one-particle states are given, the electronic linear response properties are completely determined by the orbital-independent Hartree kernel \[ F_{H}(\mathbf{r}, \mathbf{r}'; t - t') \equiv \frac{1}{|\mathbf{r} - \mathbf{r}'|} \] and the \textit{dynamic} exchange-correlation kernel

\[ F_{xc}(\mathbf{r}, \mathbf{r}'; t - t') \equiv \frac{\delta v_{xc}(\mathbf{r}; t)}{\delta n(\mathbf{r}'; t')}. \] (1)

Thus, in principle, one only needs to find a good approximation to the exchange-correlation kernel for practical TDDFT applications, which represents a significant conceptual and computational simplification. In practice, however, the conventional adiabatic local-density approximation (LDA) and generalized-gradient approximation (GGA) kernels are known to be inadequate for the study of optical excitations in solids. A primary example is their incorrect long-wavelength behavior for insulators which has been recently emphasized in several contexts \cite{5, 6, 12}.

Deficiencies of the LDA and GGA in fact appear already on the level of the \textit{ground-state} exchange-correlation energy \(E_{xc} \) and potential \(v_{xc}(\mathbf{r}) \equiv \delta E_{xc}/\delta n(\mathbf{r})\). With respect to energetics, for example, the LDA and GGA inherently fail to describe the quasi-two-dimensional electron gas \cite{13} and the van der Waals interactions \cite{14}. For \(v_{xc} \) and the corresponding KS eigenvalue spectrum, LDA and GGA \(v_{xc} \) incorrectly decay exponentially rather than as \(1/r \) in finite systems and accordingly their highest occupied orbital energies are too high and unoccupied orbitals do not exhibit Rydberg series, and for solids their band gaps are too small. The deficiencies of the LDA and GGA \(v_{xc} \) in particular pose significant practical difficulties for the study of electronic excitations within the DFT formalism.

In this regard, recent realizations of multi-dimensional KS exact-exchange (EXX) \cite{15, 16, 17} and effective EXX \cite{18, 19, 20, 21} methods provide an interesting opportunity. The orbital-based self-interaction-free EXX methods not only provide realistic local multiplicative KS exchange potentials and KS eigenvalue spectra for molecules \cite{20, 21, 22, 23, 24} but also give band gaps of semiconductors in good agreement with experiments \cite{25, 26}. We have also recently shown \cite{27} that the EXX spectrum at the one-particle level without any previously applied post-DFT modification such as a quasi-particle shift \cite{28} gives a very good description of the absorption spectrum of semiconductors excluding excitonic features resulting from two-particle interactions, which supports the notion that KS eigenvalue differences represent well-defined approximations for excitation energies \cite{29, 30}. However, the remaining excitonic character in the spectrum is not properly treated by the LDA or GGA kernel, so a complete set of DFT methods for the investigation of electronic optical excitations in solids is still missing.

Given the importance of proper inclusion of excitonic effects in the study of electronic optical excitations and the encouraging performance of the EXX method for band structures of semiconductors, we consider in this work a TD-DFT scheme based on the EXX kernel. At the formal level, in contrast to LDA and GGA kernels, the EXX kernel is nonlocal in real space, depends explicitly on the frequency, and has the correct long-wavelength limit behavior \cite{1, 22}. Taking bulk silicon as the repre-
sentative semiconductor system, we show that the EXX kernel indeed provides optical spectra of semiconductors in good agreement with experiments including excitonic effects and further analyze the origin of the excitonic peak and the validity of locality approximations in space and time. We additionally make comparisons between EXX-TDEXX and GWA-BSE, which renders new insight into the TDDFT approach in general \(^{23}\).

To obtain the optical spectrum of a solid, we computed the frequency \(\omega\)-dependent macroscopic dielectric function \(\varepsilon_M(\omega)\) which can be written in terms of the modified full linear response matrix \(\tilde{\chi}\) as

\[
\varepsilon_M(\omega) = 1 - \lim_{q \to 0} \frac{F_H(q)\tilde{\chi}(q; \omega)|_{q = \omega = 0}}{F_H(q)} ,
\]

where \(q\) is the photon momentum, and \(F_H(q) \equiv F_H(G, G', q) = \delta_{G, G'} 4\pi q/G^2\). Within TDDFT, \(\tilde{\chi}\) is completely determined by the KS linear response matrix \(\chi_0\) and the kernel matrix \(F_{xc}\) according to

\[
\tilde{\chi}(q; \omega) = \left[1 - \chi_0(q; \omega) \left(\tilde{F}_H + F_{xc}(q; \omega)\right)\right]^{-1} \chi_0(q; \omega),
\]

where \(\tilde{F}_H = 0\) for \(G = G' = 0\) and \(\tilde{F}_H = F_H\) otherwise. At the independent-particle level, \(\tilde{\chi} = \chi_0\), the two-particle interaction effects due to \(\tilde{F}_H\) and \(F_{xc}\) are both ignored, while only the \(F_{xc}\) part is ignored at the time-dependent Hartree level. As is apparent from Eqs. \(^{3}\) and \(^{4}\), the long-wavelength \(q \to 0\) behavior of the "head" \((G = G' = 0)\) and "wing" \((G = 0\) and \(G' \neq 0\) or vice versa) elements of \(F_{xc}\) (in insulators, \(O(1/q^2)\) for the head and \(O(1/q)\) for the wings \(^{3}\)\(^{,}^{23}\)) plays an important role.

The computation of \(\chi_0\) is straightforwardly done with KS one-particle states and energies, and consequently \(F_{xc}\) remains as the only important quantity to be determined. So far, for \(F_{xc}\), the adiabatic LDA kernel \(F_{xc}^{LDA} = \delta^2 E_{xc}^{LDA}/(\delta n \delta n')\) has been almost exclusively employed. It is local in real space and frequency-independent, which results in a reciprocal-space representation independent of \(q\) and \(\omega\), \(F_{xc}^{LDA}(G, G', q; \omega) = F_{xc}^{LDA}(G - G')\). While numerical advantageous, this drastic simplification is known to be problematic for the study of optical responses in solids. For example, the head and wings of the LDA kernel are incorrectly nondifferential in the \(q \to 0\) limit and this defect cannot be corrected by a semilocal GGA kernel.

Instead, we adopted in this work the EXX kernel,

\[
F_x^{EXX}(G, G', q; \omega) = \sum_{G_1, G_2} \chi_0^{-1}(G, G_1, q; \omega) \times H_x(G_1, G_2, q; \omega) \chi_0^{-1}(G_2, G', q; \omega),
\]

which is nonlocal in real space and explicitly depends on the frequency. The expression of the EXX kernel "core" \(H_x\), which was interpreted in the many-body diagrammatic picture as the first-order self-energy and vertex corrections to the irreducible polarization \(\tilde{\chi} = \chi_0 + \chi_0 F_{xc} \chi_0\), has been presented in Ref. \(^{22}\) for the case of periodic insulators. We have further shown that \(F_x^{EXX}\) has the \(q \to 0\) behavior of the exact \(F_{xc}\) and thus rectifies one serious deficiency of the LDA and GGA kernels. The full expression for \(H_x\) is rather lengthy and we list here only the three resonant terms \(^{25}\):

\[
H_x^{A - res}(q; \omega) = - \frac{2}{\Omega} \sum_{a \neq b \neq k' k} \sum_{\Sigma \chi} \left[ \frac{\langle ak | e^{-i(q+G_1^a) \cdot r} | sk + q \rangle \langle sk + q | bk' \rangle \Delta C(tk' + q; ak)}{(\epsilon_{ak} - \epsilon_{sk+q} + \omega + i\delta)(\epsilon_{bk'} - \epsilon_{sk+q} + \omega + i\delta)} \right]
\]

\[
H_x^{B - res}(q; \omega) = - \frac{2}{\Omega} \sum_{a \neq b} \left[ \frac{\langle ak | e^{-i(q+G_1^a) \cdot r} | sk + q \rangle \langle sk + q | \Sigma x \rangle \Delta C(tk + q; ak)}{(\epsilon_{ak} - \epsilon_{sk+q} + \omega + i\delta)(\epsilon_{bk} - \epsilon_{sk+q} + \omega + i\delta)} \right]
\]

\[
+ \frac{2}{\Omega} \sum_{a \neq s \neq k} \left[ \frac{\langle ak | e^{-i(q+G_1^a) \cdot r} | sk + q \rangle \langle sk + q | \Sigma x \rangle \Delta C(tk + q; ak)}{(\epsilon_{ak} - \epsilon_{sk+q} + \omega + i\delta)(\epsilon_{bk} - \epsilon_{sk+q} + \omega + i\delta)} \right],
\]

where two is the spin factor, \(\Omega\) is the crystal volume, \((a, b)\) are valence bands, \(\{s, t\}\) are conduction bands, the matrix elements \(\langle i; j | \Delta C | l; m \rangle\) are four-index integrals defined as

\[
\langle ik + q; jk' | \Delta C | lk' + q; mk \rangle \equiv \int dr \int dr' \phi_{ik+q}(r) \phi_{jk'}(r') \times \left. \phi_{lk'+q}(r) \phi_{mk}(r) \right|_{r = r'},
\]

\(\Sigma x\) denotes an orbital-dependent exchange operator of the form of the Hartree-Fock-like exchange operator but
constructed from the KS orbitals, i.e.,
\[\langle \mathbf{k} + \mathbf{q} | \hat{\Sigma}_x | \mathbf{j} \rangle = -\sum_{\mathbf{a} \mathbf{k}'} \langle \mathbf{k} + \mathbf{q} : \mathbf{a} \mathbf{k}' | \tilde{\omega}_C | \mathbf{a} \mathbf{k}' : \mathbf{j} \rangle, \tag{7}\]
and \(\tilde{\omega}_x\) is generated by the orbital-independent EXX KS potential \(v_x(r)\). Here \(\tilde{\omega}_C\) yielding \(\tilde{\omega}_C\) denotes the \textit{generalized} Coulomb interaction, e.g., \(1/|r - r'|\) for the bare Coulomb interaction.

Compared with the standard BSE approach, although we also have to perform the four-index Coulomb integrals of Eqs. (3) and (4), we can calculate optical spectra without diagonalizations in the space of occupied and unoccupied single-particle states. Due to the huge number of \(k\)-points involved in the optical spectrum calculations, the size of the matrix to be diagonalized can be very large, and the fact that we can avoid the diagonalization process altogether in principle represents a significant numerical advantage. In addition, note that \(F_{\chi}^\text{EXX}\) is free of Coulomb singularities \(24\) and thus the formal validity of the EXX TDDFT approach does not depend on the consideration of the thermodynamic limit. In the BSE equation, on the other hand, integrable singularities appear which require special numerical care \(1\).

We have implemented the \textit{full dynamic} EXX kernel in reciprocal space employing the plane wave basis. The accuracy of the code has been carefully checked by numerically testing if the calculated exchange kernel acts as a functional derivative of the exchange potential with respect to the electron density at \(\omega = 0\), \(\delta v_x^\text{EXX} = F_{\chi}^\text{EXX} \delta n\). To obtain \(F_{\chi}^\text{EXX}\), we first carried out self-consistent EXX ground-state calculations at the experimental lattice constant of Si, 5.43\(\AA\), and generated the KS potential \(11, 18\). Ten special-\(k\) points and an orbital kinetic energy cutoff of 12.5 Ha have been used. In the response calculation step, we solved the KS equations once more at a larger number of \(k\)-points and obtained KS orbitals and eigenvalues as the input for the construction of \(F_{\chi}^\text{EXX}\). We adopted a shifted uniform \(k\)-mesh, for which we employed up to 22 \(\times\) 22 \(k\)-points for \(\chi_0\) and up to 9 \(\times\) 9 \(k\)-points for \(F_{\chi}^\text{EXX}\). The kinetic energy cutoff of 10 Ha and \(\delta = 0.15\)\(eV\) have been used, and ten conduction bands were included.

The first question we addressed with small \(k\)-points was whether the adiabatic approximation, defined as the \(\omega = 0\) limit of Eqs. (4) and (5), is justified. Although the nature of the approximation within TDEXX is different from that of the BSE approach in which the static screened interaction is taken, we found that the adiabatic EXX kernel generates an overall similar spectrum as the dynamic one. Computing \(F_{\chi}^\text{EXX}\) only at a single \(\omega\) results in a significant reduction of computational workload, and thus the converged calculation was performed in the adiabatic approximation.

Figure 1 shows the absorption spectrum of Si obtained at the single-particle EXX level and by taking into account two-particle interaction effects via \(\hat{F}_H\) plus \(F_{\chi}^{LDA}\) (EXX+TDLDA) and via \(\hat{F}_H\) plus \(F_{\chi}^{EXX}\) (EXX+TDEXX). Compared with the much discussed LDA (and LDA+TDLDA) spectrum, which is incorrectly shifted to the lower frequency region by about 1 eV due to its well-known band gap underestimation, the single-particle EXX absorption edge and second \((E_2)\) peak are in excellent agreement with the experimental curve \(26\) due to the realistic EXX band structure \(18\). However, the first \((E_1)\) peak originating from electron-hole attractions is much underestimated at the single-particle EXX level, and it is not recovered by taking into account electron-hole interactions via TDLDA, demonstrating the failure of \(F_{\chi}^{LDA}\) to describe excitonic effects.

The EXX+TDEXX spectrum, on the other hand, is in excellent agreement with the measured data: The absorption strength at the \(E_1\) peak region is correctly enhanced while that of the higher frequency region is reduced, which shows that \(F_{\chi}^{EXX}\) indeed provides the correct description of the important excitonic effects.

To obtain the EXX+TDEXX spectrum in Fig. 1, we used a slightly modified bare Coulomb interaction as \(w_{C}\). Employing the bare Coulomb interaction resulted in a collapse of the spectrum due to too strong long-range Coulomb interaction of electron-hole pairs at different \(k\)-points. We therefore cut off the long-range Coulomb interaction for these pairs, or set to zero the contributions to the integrals of Eq. (1) for which \((G + k - k')\) lies within the first Brillouin zone \(27\). Devising a more systematic strategy might be desirable in the future.

Now, to understand the underlying mechanism of the encouraging EXX+TDEXX result, we analyze \(F_{\chi}^{EXX}\) by investigating the role of different contributions to \(H_\chi\) as shown in Fig. 3. First, by taking only \(H_\chi^B\), we obtain a spectrum resembling that of the one-particle EXX but shifted upward by about 0.3 eV. No excitonic feature arises at this level. This spectrum is similar to that of the GWA, which is in agreement with previous predictions \(23, 24\) that GWA results should be recovered with the \(H_\chi^B\) part of the screened EXX kernel. However, note that we have used an only slightly modified bare Coulomb interaction, and, while the GWA spectrum results from independent quasiparticles that involve \((N \pm 1)\)-electron excitations, the DFT “quasiparticles” spectrum already represents electron-number conserving \((N\)-electron\) one-particle excitations at the time-independent level \(18, 20, 21\). Second, if we take only \(H_\chi^A\), the spectrum is strongly enhanced at the lower frequency region \((\lesssim 4\ eV)\) while it is much reduced at the higher frequency region \((\gtrsim 4\ eV)\). The excitonic \(E_1\) peak appears predominantly while the \(E_2\) peak is slightly redshifted by about 0.2 eV. This shows that the Coulomb coupling of electron-hole pairs is precisely the origin of the excitonic \(E_1\) peak. Finally, we set the head and wings of \(F_{\chi}^{EXX}\) to zero and thus make them \(q\)-independent as in the case of the LDA or GGA. By doing so, we obtain
a spectrum similar to that of the EXX+TDLDA without excitonic peaks or whatsoever. This demonstrates that taking into account the correct $q \to 0$ behavior of $F_{xc}$ is crucial.

We finally comment on other recent TDDFT works that have also obtained optical spectra of semiconductors in good agreement with experiments. First, de Boeij et al. have left the domain of conventional DFT and resorted to TD current-DFT [28]. Albeit they had to introduce an empirical energy shift and a prefactor for the polarization functional, their approach may represent a different route to include the space nonlocality discussed above via a macroscopic functional. However, our result shows that employing the current response derived from a macroscopic functional is not necessary if one properly incorporates the space nonlocality in the microscopic density response. In this context, recent work of Reining et al. which was based on the idea of mapping the BSE to the TDDFT is in agreement with ours [3]. But in their scheme they first invoked the GWA quasiparticle shift and second employed an empirical static scaled Coulomb kernel as $F_{xc}$. Although their approach is not rigorous, we note that their simple static kernel nevertheless exhibits a correct $q \to 0$ divergence.

In summary, we have reported calculations of the optical absorption spectrum of Si with the novel TDEXX approach. Of particular interest was the nature of the electronic absorption spectrum of Si with the novel TDEXX Coulomb kernel as $F_{xc}$. Although their approach is not rigorous, we note that their simple static kernel nevertheless exhibits a correct $q \to 0$ divergence.

We thank L. Reining, L. X. Benedict, and E. L. Shirley for helpful discussions. This work has been supported by the Humboldt Foundation (Y.-H.K.), the Deutsche Forschungsgemeinschaft, and the Fonds der Chemischen Industrie (A.G.).

* Present address: Materials and Process Simulation Center (139-74), California Institute of Technology, Pasadena, CA 91125-7400

[1] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. 80, 4510 (1998); L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. 80, 4514 (1998); M. Rohlfing and S. G. Louie, Phys. Rev. Lett. 81, 2312 (1998).
[2] W. Hanke and L. J. Sham, Phys. Rev. Lett. 43, 387 (1979).
[3] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. 55, 1418 (1985).
[4] E. Runge and E. K. U. Gross, Phys. Rev. Lett. 52, 997 (1984).
[5] We use Hartree atomic units, $\hbar = e = m_e = 4\pi\epsilon_0 = 1$, throughout.
[6] W. G. Aulbur, L. Jönsson, and J. W. Wilkens, Phys. Rev. B 54, 8540 (1996).
[7] Ph. Ghosez, X. Gonze, and R. W. Godby, Phys. Rev. B 56, 12811 (1997).
[8] L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett. 88, 066404 (2002); G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).
[9] Y.-H. Kim et al., Phys. Rev. B 61, 5202 (2000); L. Pollock and J. P. Perdew, J. Chem. Phys. 12, 1239 (2000); P. Garcia-Gonzalez, Phys. Rev. B 62, 2321 (2000).
[10] W. Kohn, Y. Meir, and D. E. Makarov, Phys. Rev. Lett. 80, 4153 (1998); J. F. Dobson and J. Wang, Phys. Rev. Lett. 82, 2123 (1999); P. Garcia-Gonzalez et al., Phys. Rev. Lett. 88, 056406 (2002).
[11] M. Städele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. 79, 2089 (1997); M. Städele and Richard M. Martin, Phys. Rev. Lett. 84, 6070 (2000).
[12] A. Görling, Phys. Rev. Lett. 83, 5459 (1999).
[13] S. Ivanov, S. Hirata, and R. J. Bartlett, Phys. Rev. Lett. 83, 5455 (1999).
[14] D. M. Bylander and L. Kleinman, Phys. Rev. Lett. 74, 3660 (1995).
[15] T. Grabo and E. K. U. Gross, Int. J. Quant. Chem. 64, 95 (1997).
[16] Y.-H. Kim, M. Städele, and R. M. Martin, Phys. Rev. A 60, 3633 (1999).
[17] F. Della Sala and A. Görling, J. Chem. Phys. 115, 5718 (2001); 116, 5374 (2002); O. V. Gritsenko and E. J. Baerends, Phys. Rev. A 64, 042506 (2001).
[18] Y.-H. Kim, M. Städele, and A. Görling, Int. J. Quant. Chem. (to be published).
[19] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. 63, 1719 (1989).
[20] A. Görling, Phys. Rev. A 54, 3912 (1996).
[21] C. J. Umrigar, A. Savin, and X. Gonze, in Electronic Density Functional Theory: Recent Progress and New Directions, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, New York, 1998); A. I. Al-Sharif, R. Resta, and C. J. Umrigar, Phys. Rev. A 57, 2466 (1998).
[22] Y.-H. Kim and A. Görling, Phys. Rev. B 66, 035114 (2002); see also A. Görling, Int. J. Quant. Chem. 69, 265 (1998); Phys. Rev. A 57, 3433 (1998).
[23] X. Gonze and M. Scheffler, Phys. Rev. Lett. 82, 4416 (1999).
[24] I. V. Tokatly and O. Pankratov, Phys. Rev. Lett. 86, 2078 (2001).
[25] $H^\alpha_k$ and $H^\beta_k$ represent $(H^\alpha_1 + H^\alpha_2)$ and $(H^\beta_1 + H^\beta_2)$ of Ref. [23]. We assume that the external perturbation and other quantities have the time dependence of $e^{-i\omega t}e^\delta$, where $\delta \to 0^+$ is a convergence factor.
[26] C. M. Herzinger, B. Johs, W. A. McGahan, and A. Wool- lam, J. Appl. Phys. 83, 3323 (1998).
[27] This neglect of long-range contribution is necessary only for $k \neq k'$, because $k = k'$ contributions with $|G + k - k'| = |G| = 0$ which would lead to Coulomb singularities do not occur in the EXX kernel expression.
[28] P. L. de Boeij, F. Kootstra, J. A. Berger, R. van Leeuwen, and J. G. Snijders, J. Chem. Phys. 115, 1995 (2001).
FIG. 1: Calculated optical absorption spectrum of Si from EXX (dashed lines), EXX+TDLDA (dot-dashed lines), and EXX+TDEXX (solid lines). Circles denote experimental data of Ref. 26.

FIG. 2: Calculated optical absorption spectrum of Si from EXX+TDEXX with full EXX kernel (solid lines), $H^B$ only (dashed lines), $H^A$ only (dot-dashed lines), and full EXX kernel with head and wings set to zero (long-dashed lines). Circles: experiment.