Efficient calculation of excitonic effects in solids including approximated quasiparticle energies

Filipe Matusalem, Marcelo Marques, Ivan Guilhon and Lara K Teles

Group of Semiconductor Materials and Nanotechnology (GMSN), Technological Institute of Aeronautics (ITA), 12228-900 São José dos Campos/SP, Brazil

E-mail: filipematus@gmail.com and gmsn@ita.br

Received 29 March 2020, revised 19 May 2020
Accepted for publication 3 June 2020
Published 6 July 2020

Abstract

In this work we present a new procedure to compute optical spectra including excitonic effects and approximated quasiparticle corrections with reduced computational effort. The excitonic effects on optical spectra are included by solving the Bethe–Salpeter equation, considering quasiparticle eigenenergies and respective wavefunctions obtained within DFT-1/2 method. The electron–hole ladder diagrams are approximated by the screened exchange. To prove the capability of the procedure, we compare the calculated imaginary part of the dielectric functions of Si, Ge, GaAs, GaP, GaSb, InAs, InP, and InSb with experimental data. The energy position of the absorption peaks are correctly described. The good agreement with experimental results together with the very significant reduction of computational effort makes the procedure suitable on the investigation of optical spectra of more complex systems.

Keywords: excitonic effects, density functional theory, dielectric function, optical properties, semiconductors

(Some figures may appear in colour only in the online journal)

1. Introduction

The density functional theory (DFT) is one of the most powerful and popular theoretical techniques to obtain groundstate and electronic properties of materials. However, it is well known that the excited spectra obtained from standard DFT suffer with the underestimation of energy transitions between Kohn–Sham (KS) eigenvalues. On the other hand, a good description of optical properties requires the knowledge of the excited states and a proper description of many-body effects. In this sense, solving the Bethe–Salpeter equation (BSE) for pair excitations and local-field contributions to the optical response can be taken into account [1–3] to describe spectroscopic properties. Moreover, in some cases, such as the calculation of absorption spectra of low-dimensional materials, the inclusion of excitonic effects may dramatically affect the calculated results due to very large excitonic binding energy [4, 5].

In order to treat excited states, realistic quasiparticle (QP) energies are obtained by applying self-energy corrections to the KS energies, usually evaluated in the very demanding GW approximation [6]. The calculation of absorption spectra is often made in two steps. First, the QP energies and the RPA screened Coulomb interaction $W$ need to be obtained. In a second step the calculated eigenfunctions, together with the QP energies and $W$, are then used in the exciton calculation [1]. The computation of the QP energies within the expensive GW method and the approximation of the electron–hole ladder diagrams by static screened Coulomb interaction, $W(\omega \to 0)$, dramatically increase the time computational requirements and time. Therefore, large numerical effort is required to solve the BSE, restricting such calculations to the interaction of relatively few electron–hole pairs. Therefore, the study of low computational cost methods for obtaining an accurate optical spectrum considering excitonic effects is of great importance, and could open an avenue for studying more...
complex systems. For example, good results are found using an approach including LDA + U and scissor operator with parameter U and scissor operator Δ determined such that the corresponding binding energies and resulting band gaps resemble the HSE + G0W0 values. This is a low cost method however a previous HSE + G0W0 is necessary to set the U and scissor parameters [7].

In order to reduce the computational cost of the first step of the described procedure, in this work, we propose as an alternative the substitution of the GW QP energies by those obtained within DFT-1/2 method [8, 9]. The DFT-1/2 method was developed to improve computational efficiency while keeping a good compromise with accuracy. This method is able to predict the energy gap results with precision similar to GW and HSE, but with the same computational effort of standard DFT calculations. It is worth mentioning, that this formalism is also free of adjusted parameters. In the sequence, the electron–hole ladder diagrams is then approximated by the screened exchange, provided by the hybrid HSE06 [10–12] scheme. The reduction in the computational effort obtained by this methodology resides basically on time and memory requirements during the calculation of quasiparticle eigenvalues. This time and memory are reduced by around 2 or 3 orders of magnitude in comparison with a hybrid HSE06 or single shotG0W0 calculation, respectively [13].

To prove the accuracy of this methodology, we compute the imaginary part of dielectric function of a set of semiconductors, namely, Si, Ge, GaAs, GaP, GaSb, InAs, InP and InSb, and compare the obtained results with available experimental data. The article is organized as follows. In section 2, we describe the methods adopted and computational details of our calculations. In section 3, we present and discuss the results. Finally, in section 4 we briefly summarize the paper and present our conclusions.

2. Computational details

We perform DFT calculations employing a semi-local approach including a second-order gradient coefficient for the exchange energy named PBEsol [14, 15]. The PBEsol is employed to determine the theoretical lattice parameters and relaxed positions of ions. The Kohn–Sham equations are solved self-consistently within the projector augmented wave (PAW) for generation of wavefunctions and pseudopotentials [16], as implemented in VASP. For the plane wave expansion of the electronic wave functions we use an energy cutoff of 245 eV, 174 eV, 209 eV, 255 eV, 283 eV, 209 eV, 255 eV and 172 eV for Si, Ge, GaAs, GaP, GaSb, InAs, InP and InSb, respectively. Spin–orbit coupling was taken into account for all structures including Sb atoms, in calculation of band structures. The Brillouin zone (BZ) integrations are carried out with a $15 \times 15 \times 15 \Gamma$-centered k-point mesh.

The general rule for the inclusion of the self-energy correction in the DFT-1/2 approach is to distribute the half-ionization between the atomic orbitals forming the valence band maximum (VBM) according to the percentage of each orbital contribution [8]. For GaAs, GaSb and InSb around 90% of the VBM is formed by the p orbital of the correspondent anion atom in each case. Therefore an accurate result is achieve with all the half-ionization included for the corresponding orbital. In Si and Ge, a p character of VBM states and covalent bonds between atoms of same species are observed. In these cases the half-ionization is shared for each pair of atoms, so the atomic potential is generated with only 1/4 ionization to avoid the overestimation of the self-energy correction on the QP eigenenergies. As described in the papers introducing the DFT-1/2 [8, 9], the true band gap is a result of a sum of the standard DFT Kohn–Sham band gap plus the self-energy contribution relative to the VBM, $\Delta E_V$, minus the self-energy contribution relative to the conduction band minimum (CBM), $\Delta E_C$.

For three-dimensional compounds in general, the contribution relative to the CBM can be ignored and a good description of the band gap already achieved by considering only corrections on the valence states. However, in the case of cubic germanium compound and GaP, the self-energy contribution relative to the CBM could not be ignored, mainly due the known existence of an indirect band gap. The self-energy $\Delta E_C$ correction is then made in this case applying the DFT-1/2 procedure for the Ge(s), Ga(s) and P(s) orbitals forming the CBM. To adjust the sign of the contribution, in order to consider the addition of half electron to the unoccupied band, a negative amplitude is used [17]. For InAs and InP, the half electron ionization is shared between both species according to the VBM composition. This is controlled by the amplitude parameter included in the generation of the DFT-1/2 Paw potential [18]. The details for the generation of the Paw corrected potentials used in this work are listed in table 2.

Then, the DFT-1/2 QP eigenenergies and respective wavefunctions are used to build a two-particle Hamiltonian of singlet excitations to describe excitonic effects. For given sets of valence bands, conduction bands, and k-points we solve the eigenvalue problem of the two-particle Hamiltonian, the homogeneous Bethe–Salpeter equation [19]. The dielectric functions are calculated considering both the independent quasiparticle and excitonic effects for comparison.

3. Results and discussion

Lattice constants calculated using PBEsol exchange–correlation functional are listed in table 1. A very good agreement between the calculated results and experimental findings is observed, only exhibiting overestimations not much larger

|                | $a_{\text{PBEsol}}$ (Å) | $a_{\text{EXP}}$ (Å) |
|----------------|------------------------|----------------------|
| Si             | 5.439                  | 5.431                |
| Ge             | 5.704                  | 5.658                |
| GaAs           | 5.685                  | 5.653                |
| GaP            | 5.477                  | 5.450                |
| GaSb           | 6.115                  | 6.096                |
| InAs           | 6.123                  | 6.058                |
| InP            | 5.932                  | 5.869                |
| InSb           | 6.543                  | 6.479                |
than 1%. The resulting relaxed structures were used for the calculation of electronic and optical properties.

Table 2 shows the electronic band gaps calculated considering PBEsol functional and DFT-1/2 method, together with experimental data. As expected, the PBEsol band gaps are underestimated in comparison with experiment, including the wrong prediction of metallic behavior for Ge, GaSb, InAs, and InSb. After the inclusion of the self-energy correction following the DFT-1/2 method, no prediction of metallic behavior is observed and more accurate band gaps are obtained for the semiconductor materials. Good agreement between calculated and experimental results is achieved.

In figure 1 are show the band structures calculated using PBEsol and DFT-1/2 approaches. We clear see the opening of the band gaps due the inclusion of the DFT-1/2 correction, corroborating the data of table 1. Also, a known characteristic of DFT-1/2 can be observed, the trend to flattening the bands. The effect occurs mainly on the valence bands since are the ones formed by the atomic orbitals subject to the applied half- ionization (see last column of table 2).

The imaginary part of the dielectric function, calculated using PBEsol functional (PBEsol), PBEsol functional including excitonic effects (PBEsol + BSE), DFT-1/2 functional (DFT-1/2), DFT-1/2 functional including excitonic effects (DFT-1/2 + BSE), and HSE functional including excitonic effects (HSE + BSE), are shown in figure 2 in comparison with experimental results (EXP). At least four valence and four conducting bands are included in the calculation of optical spectra, which guarantee that transitions up to 10 eV are included, as can be inferred from band structures in figure 1. The obtained optical spectra are all depicted in 2. A common characteristic observed in all spectra is that the curves obtained without inclusion of QP corrections (PBEsol) are severe redshifted with respect to the experimental results. This behavior is justified by the underestimation of the fundamental band gap, since the imaginary part of the dielectric function is strictly related with the interband transitions from the valence to conduction bands. Considering the green curves in figure 1, one observes that the inclusion of the DFT-1/2 QP correction shifts the previous PBEsol results to greater energies in the same rate that the ones observed for the increasing in the band gaps shown in table 2. As for the band gaps, the inclusion of the QP correction resulted in imaginary dielectric functions in very good agreement with the experimental results when looking only for the position in energy of the main peaks present in the spectra. The exception, although improvement is achieved, occurs for the silicon case. As can be observed in table 2 the direct gap of silicon is still 0.4 eV underestimated in comparison with the experiment. This difference is approximately the same observed in the absorption spectra of silicon (figure 1(a)). Better agreement of silicon dielectric function in comparison with experiment is obtained only with BSE plus GW approach [1, 21] or with time-dependent density-functional theory (TDDFT), however, still with a small blueshift for low frequencies [22]. The inclusion of the excitonic effects on the other hand yields in a small blueshift of the spectra. This is expected and can be associated with the decrease in band gap due the influences of the bound energy of the electron–hole pair (exciton) generated by the excitation of an electron from the valence to conduction band. Peaks in spectra related to bound excitons are not visible but could appear below the onset of the absorption if a much more refined k-point meshes is used [3, 21, 23]. Another consequence of the inclusion of the excitonic effects is related with the intensities. Peaks located in the lowest energy of the spectra (to the left side) have their intensities decreased while peaks located in the highest energy of spectra (to the right side) have their intensities increased. The spectra of GaAs, GaSb, InAs, InP and InSb presents a double peak centered at about 3.1, 2.5, 2.8, 3.4, and 2.5 eV, respectively. The inclusion of the excitonic effects increase the intensity of both peaks, however, the left one that before had lower intensity becomes more pronounced than the right one, in better agreement with experimental results, except for InP that presents only one peak at the refereed energy position. In this case we argue that possibly the second peak in InP spectra do not appear in experiment due the small number of observed points.

| Atom (orbital) fraction | CUT amplitude |
|-------------------------|---------------|
| Si(p) 1/4 3.78 1.00     |               |
| S(p) 1/4 3.78 1.00      |               |
| Ge(p) 1/4 4.23 1.00     |               |
| Ge(s) 1/4 2.27 –1.00    |               |
| As(p) 1/2 3.84 1.00     |               |
| Sh(p) 1/2 4.27 1.00     |               |
| In(p) 1/2 4.14 0.86     |               |
| In(p) 1/2 1.36 0.14     |               |
| Sb(p) 1/2 4.53 1.00     |               |

Table 2. PBEsol (E\textsubscript{PBEsol}), DFT-1/2 (E\textsubscript{DFT-1/2}) and experimental [20] (E\textsubscript{EXP}) electronic band gaps. The last column show detailed information about the DFT-1/2 correction. The plus(minus) sign means correction applied to the valence(conduction) level. 0.5 and 0.25 are the amount of electronic ionization in the atomic calculation applied to p or s orbitals. The superscript d and i stands for direct and indirect gap, respectively.
A common feature observed in all DFT-1/2 spectra depicted in figure 1 is the decrease of the intensities in relation with both PBEsol and experimental results in the medium energy region of the transition spectra. As already pointed out the dielectric function is directly related with electronic transitions between valence to conduction bands. Peaks or increase of intensity occurs when the valence and conduction bands exhibit parallel segments in the energy–momentum band dispersion, i.e., the same amount of energy produce excitation of electrons with different momenta in the $k$-space. In this sense, slight modification of the bands participating of the transitions, in such a way there is a decrease in the parallelism, will cause a reduction of the intensity of absorbance for the corresponding excitation energy. The DFT-1/2 correction in general tends to slightly change the shape of the bands derived from the orbitals for which the self-energy is included. In the most cases only the valence bands are affected, therefore decreasing the parallelism of the bands participating of the transitions which by consequence decrease the intensity of the spectra. Looking the last column of table 2 we can see that except for Ge and GaP the self-energy correction is applied only for orbitals forming the valence band (positive amplitude), resulting in a decrease in intensities. For Ge and GaP we also include self-energy correction for orbitals forming the conduction bands (negative amplitude). For these two cases there is some compensation, i.e., also the bands in the conduction part are modified in such a way that the reduction in intensities of the spectra is less pronounced. Moreover, one verifies in figure 1 a similar performance of DFT-1/2 + BSE and HSE + BSE approaches. The later methodology is a cheaper alternative to GW + BSE state-of-art approach, but still orders of magnitude more expensive than former one.

To explain why the intensities are in better agreement with experiment for high energy regions we have to regard to an intrinsic characteristic of the DFT-1/2 method, the ability to well describe the eigenvalues near the band edges. Far from this region the bands are in general not affected, except by an energy shift, which do not change the parallelism of the bands and consequently not affect the intensities.
the other hand, for low energy region, the transitions occur for $k$-points around the valence band maximum (VBM) and conduction band minimum (CBM). As the DFT-1/2 method is calibrated by the orbital character of VBM and/or CBM, the intensities of the beginning of the spectra is also well described.
4. Conclusions

In summary, we propose a new procedure to calculate the absorption spectra of semiconductors including quasiparticle energies and excitonic effects with significant reduction in computational effort and in good agreement with experimental results. It consist in use the DFT-1/2 method to generate an approximated quasiparticle spectra which serves as input to solve the Bethe–Salpeter equation responsible to take into account the excitonic effects. The methodology was tested for a series of semiconductors, namely, Si, Ge, GaAs, GaP, GaSb, InAs, InP and InSb, and result in very good agreement with experimental results.

The calculated intensities of absorption spectra in the medium energy region are slight lower than experimental and standard DFT functional results, which can be explained by the stronger localization of valence bands in the DFT-1/2 picture. Although the intensities of the spectra do not exactly fit the experimental results, their qualitative behavior and the energy positions of the peaks of absorption present a very good agreement. Being DFT-1/2 a low-cost quasiparticle correction method with respect to computational resources, this combination (DFT-1/2 + BSE) can be used to obtain optical spectra including excitonic effects for systems in which GW + BSE approach is unfeasible.

Acknowledgments

We thank the Brazilian Funding Agency Coordination for Improvement of Higher Level Education (CAPES), Grant No. 23038.005810/2014-34, the National Council for Scientific and Technological Development (CNPq), Grant Nos. 310422/2019-1 and 306322/2017-0, and the São Paulo Research Foundation (FAPESP) for the Grant Nos. 2006/05858-0 and 2012/50738-3 for the financial supports.

ORCID iDs

Filipe Matusalem  https://orcid.org/0000-0002-5305-9693
Marcelo Marques  https://orcid.org/0000-0002-9241-5048
Ivan Guilhon  https://orcid.org/0000-0001-8882-4490
Lara K Teles  https://orcid.org/0000-0002-8713-2094

References

[1] Albrecht S, Reining L, Del Sole R and Onida G 1998 Phys. Rev. Lett. 80 4510–5
[2] Benedict L X, Shirley E L and Bohn R B 1998 Phys. Rev. Lett. 80 4514
[3] Rohlfing M and Louie S 1998 Phys. Rev. Lett. 81 2312–5
[4] Wirtz L, Marini A and Rubio A 2006 Phys. Rev. Lett. 96 126104
[5] Guilhon I, Marques M, Teles L K, Palummo M, Pulci O, Botti S and Bechstedt F 2019 Phys. Rev. B 99 1611204
[6] Hedin L 1965 Phys. Rev. 139 A796–823
[7] de Carvalho L C, Schleife A, Furthmüller J and Bechstedt F 2013 Phys. Rev. B 87 195211
[8] Ferreira L G, Marques M and Teles L K 2008 Phys. Rev. B 78 125116
[9] Ferreira L G, Marques M and Teles L K 2011 AIP Adv. 1 032119
[10] Heyd J, Scuseria G E and Ernzerhof M 2003 J. Chem. Phys. 118 8207
[11] Heyd J and Scuseria G E 2003 J. Chem. Phys. 118 1187–92
[12] Heyd J, Scuseria G E and Ernzerhof M 2006 J. Chem. Phys. 124 219906
[13] Pela R R, Marques M and Teles L K 2015 J. Phys.: Condens. Matter 27 505502
[14] Perdew J P, Ruzsinszky A, Csonka G I, Vydrov O A, Scuseria G E, Constantin L A, Zhou X and Burke K 2008 Phys. Rev. Lett. 100 136406
[15] Perdew J P, Ruzsinszky A, Csonka G I, Vydrov O A, Scuseria G E, Constantin L A, Zhou X and Burke K 2009 Phys. Rev. Lett. 102 039902(E)
[16] Kresse G and Joubert D 1999 Phys. Rev. B 59 1758–75
[17] Matusalem F, Marques, M Teles L K and Bechstedt F 2015 Phys. Rev. B 92 045436
[18] Ataide C A, Pelá R R, Marques M, Teles L K, Furthmüller J and Bechstedt F 2017 Phys. Rev. B 95 0345126
[19] Bechstedt F 2015 Many-Body Approach to Electronic Excitations (Springer Series in Solid-State Sciences vol 181) (Berlin: Springer)
[20] Levinshtein M E, Rumyantsev S L and Shur M 1996 Handbook on Semiconductor Parameters: Si, Ge, C (Diamond), GaAs, GaP, GaSb, InAs, InP, InSb (Singapore: World Scientific)
[21] Hahn P H, Seino K, Schmidt W G, Furthmüller J and Bechstedt F 2005 Phys. Status Solidi Basic Res. 242 2720–8
[22] Reining L, Olevano V, Rubio A and Onida G 2002 Phys. Rev. Lett. 88 066404
[23] Hahn P H, Schmidt W G, Seino K, Preuss M, Bechstedt F and Bernholc J 2005 Phys. Rev. Lett. 94 1–4
[24] Aspnes D E and Studna A A 1983 Phys. Rev. B 27 985–1009