APS physics

CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Self-consistent solution of Hedin's equations: Semiconductors and insulators

Andrey L. Kutepov

Phys. Rev. B **95**, 195120 — Published 10 May 2017

DOI: 10.1103/PhysRevB.95.195120

Self-consistent solution of Hedin's equations: semiconductors and insulators

Andrey L. Kutepov*

Department of Physics and Astronomy,

Rutgers University, Piscataway, NJ 08856

Abstract

The band gaps of a few selected semiconductors/insulators are obtained from the self-consistent solution of the Hedin's equations. Two different schemes to include the vertex corrections are studied: (i) the vertex function of the first-order (in the screened interaction W) is applied in both polarizability P and self-energy Σ, and (ii) the vertex function obtained from the Bethe-Salpeter equation is used in P whereas the vertex of the first-order is used in Σ. Both schemes show considerable improvement in the accuracy of the calculated band gaps as compared to the self-consistent GW approach (scGW) and to the self-consistent quasi-particle GW approach (QSGW). To further distinguish between the performances of two vertex-corrected schemes one has to properly take into account the effect of the electron-phonon interaction on the calculated band gaps which appears to be of the same magnitude as the difference between schemes i) and ii).

PACS numbers: 71.20.Mq, 71.20.Nr, 71.20.Ps, 71.45.Gm

* e-mail: kutepov@physics.rutgers.edu

Introduction

The ability to accurately predict band gaps in semiconductors/insulators is a long-standing goal for computational physicists. Density functional theory[1] (DFT) in its Local Density Approximation (LDA) or in Generalized Gradient Approximation (GGA) has been very successful in the prediction of the ground state properties, but it is incapable of describing the excited-state properties such as the quasiparticle band structure in a quantitative way. Specifically, the average deviation in the calculated band gaps of semiconductors/insulators is about 40-50% as can be estimated from Fig.1. The deviation becomes even larger if one takes into account the effect of the electron-phonon (e-ph) interaction. The common approach nowadays to address the band gap problem, therefore, is to use Hedin's GW approximation[2] or its numerous extensions[3–13] which often are more accurate than DFT in this respect. A great variety of different GW-based schemes, which are available today, allows one to select the most accurate one for a given material. On the other hand, the same variety of approaches tells us that the search for an optimal method to calculate the band gaps has not been finished yet. As it was discussed in the recent paper (Ref.[14], which hereafter will be abbreviated as I), the success of many existing extensions of the GW approach is based on the cancellation of errors which renders their systematic improvement complicated. Besides, GW schemes do not provide information on the effect of the diagrams beyond GW. For example, one cannot say whether the perturbation expansion is convergent or not. Common route to go beyond GW approximation nowadays is to combine the GW approach with the Time Dependent Density Functional Theory (TDDFT) through the introduction of the exchange-correlation kernel f_{xc}.[12,15,16] However, to the best of my knowledge, until recently, there were no attempts to apply self-consistent purely many-body extensions of the GW method to crystalline materials.

In I, a few schemes to solve the Hedin's equations[2,25] self-consistently (sc) and with higher order diagrams (vertex corrections) included have been introduced. The schemes are not based on the cancellation of errors. The main advantages of the schemes as compared to the existing GW-based methods are the following: they are diagrammatic and self-consistent, they do not apply the quasi-particle approximation for the Green's function, they treat full frequency dependence of interaction W in higher order diagrams, they apply vertex corrections for both polarizability and self-energy. The schemes were successfully used in I

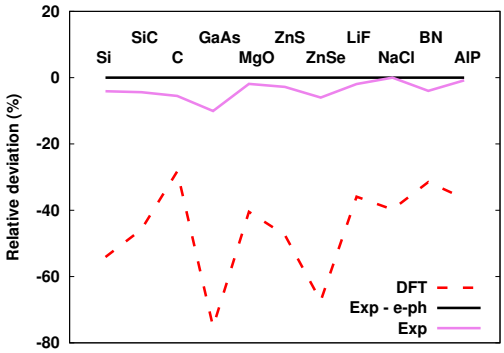

FIG. 1: (Color online) Relative deviation of the calculated (in GGA) band gaps of selected semiconductors/insulators from the experimental band gaps (excluding the electron-phonon interaction). Experimental data has been taken from the Refs.[17–19]. The horizontal line represents the experimental data corrected by the effects of electron-phonon and spin-orbit interactions (zero deviation). The corrections were taken from the Refs.[20–23]. The GGA results have been cited from the Ref.[24].

to calculate the band widths in alkali metals (Na and K) and the band gaps in Si and LiF. The accuracy of the obtained results was superior to the accuracy of the QSGW method and it was generally within uncertainty of the experimental data, but the calculations presented in I were very time consuming.

In this work, a few improvements of the algorithms presented in I are used, which makes the approaches less computationally demanding. The improvements continue to explore the fact that the diagrams beyond GW (vertex part) can generally be accurately evaluated with less intensive numerical parameters as compared to the parameters one uses in the GW part. This idea was partially used already in I. Namely, the number of band states (N_{bnd}) (they are obtained from the effective Hartree-Fock problem[10] on every iteration) used in the vertex part was less than in the GW part. Also, the number of orbitals (ϕ_{nl}) inside the MT (muffin-tin) spheres and the number of plane waves in the interstitial region ($N_{\mathbf{G}}$) to represent the band states were smaller in the vertex part. In this work, the coarser k-mesh in the Brillouin zone (with $N_{\mathbf{k}}$ points) and a smaller number of the Matsubara's time/frequency points (n_τ/n_ω) for the vertex part have been used.

The plan of the paper is the following. Section I explains the selection of vertex corrected schemes for this study and presents convergence checks. Section II provides the results obtained and the discussion. The conclusions are given afterwords.

I. METHODS AND CONVERGENCE CHECKS

In this work, for a systematic study of the band gap problem in semiconductors, two of the six approaches introduced in I are used: B and D schemes. Scheme B is conserving and represents the first step beyond the GW approximation where all diagrams up to the first order (in the screened interaction W) are included in the vertex function (Fig.2). The same vertex function ($\Gamma = \Gamma_1$) is then applied for both the polarizability

$$P(12) = \sum_{\alpha} G^{\alpha}(13)\Gamma^{\alpha}(342)G^{\alpha}(41), \tag{1}$$

and the self-energy

$$\Sigma^{\alpha}(12) = -G^{\alpha}(14)\Gamma^{\alpha}(425)W(51), \tag{2}$$

where digits inside the brackets represent the space-time arguments and α is the spin index. All relevant quantities (P, W, G, Σ) are iterated till the full self-consistency in scheme B. The satisfaction of the macroscopic conservation laws is appealing, but the limited number of the diagrams included in scheme B breaks down other relationships which should be fulfilled in an exact theory. For instance, the polarizability in scheme B doesn't match its natural definition as a functional derivative of the electron density with respect to the electric field (external plus induced). Thus, the polarizability in scheme B is not physical.

On the other hand, the polarizability in scheme D is physical by construction, as the corresponding vertex function ($\Gamma = \Gamma_{GW}$) is obtained from the Bethe-Salpeter equation (Fig.3) with the kernel ($\Theta = \frac{\delta\Sigma}{\delta G}$) in "GW" approximation (Fig.4). Thus, in scheme D, the polarizability (and the corresponding W) are not iterated till the full self-consistency but are

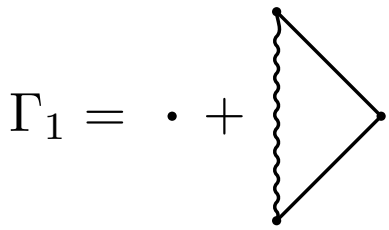

FIG. 2: First order approximation for the 3-point vertex function. Direct lines represent Green's function G and the wavy line represents the screened interaction W. The dot is for the trivial part of the vertex function.

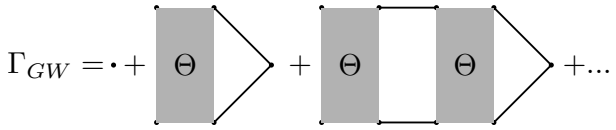

FIG. 3: Ladder sequence for the 3-point vertex function with the kernel Θ (see Fig.4) as the rung of the ladder.

evaluated only once, immediately after scGW calculation. The choice of the diagrams for self energy in scheme D (equation (2) with $\Gamma = \Gamma_1$) is a trade between the consistency with the Bethe-Salpeter equation for the polarizability on one hand (the kernel of the Bethe-Salpeter equation is obtained assuming $\Sigma = GW$) and the explicit "improvement" of self-energy itself by including higher order skeleton diagrams on the other hand. Self energy and Green's function are iterated till full self-consistency in scheme D, similar to scheme B. Scheme D represents an example of a diagrammatic approach where certain infinite series of diagrams are included based on a specific physical principle (in this particular case - the requirement of the microscopic charge preservation).

It is important to mention that the frequency dependence of W is taken into account without approximations in both schemes B and D. This allows one to consider them as advanced approaches as compared to the commonly used static approximation for W (taken at zero frequency) or the plasmon-pole approximation.[4] One more technical detail is of importance: in the diagrams beyond "GG" approximation for the polarizability and "GW" approximation for the self-energy all building blocks (G and W) are treated with the same (reduced) basis set. Thus, no asymmetry is introduced in self-energy from the terms like $G * W * \Gamma[G, W]$ because all G and W (not only those which are the arguments of the Γ) are expressed in the same reduced basis set.

As the above two vertex-corrected schemes represent different approaches for the selection of diagrams, it seems to be interesting to apply them systematically for semicon-

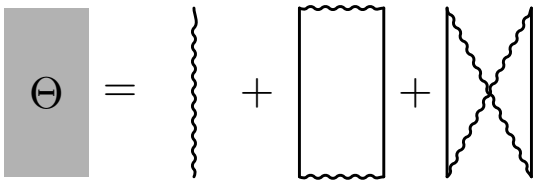

FIG. 4: The GW approximation for the irreducible 4-point kernel Θ. Direct lines represent the Green's function and wavy lines represent the screened interaction W.

TABLE I: Dependence of the calculated band gap of MgO on the calculation setup for the diagrams beyond *GW*. Scheme B (see the text for the details) has been used. ϕ_{nl} enumerates the orbitals (shown as sequences of principal quantum numbers and orbital characters) inside MT spheres which were used to represent the band states in the diagrams beyond GW approximation. The temperature was 1000K.

Parameter	Setup	Band gap
N_{bnd}	0	9.31
	5	8.81
	10	8.40
	20	8.28
	30	8.29
ϕ_{nl}	3s(Mg)/2p(O)	8.56
	3s3p(Mg)/3s2p(O)	8.40
	3s3p3d(Mg)/3s2p3d(O)	8.28
	3s3p3d4f(Mg)/3s2p3d4f(O)	8.27
$N_{\mathbf{G}}$	26	8.19
	59	8.25
	92	8.28
$N_{\mathbf{k}}$	2^3	8.28
	3^3	8.24
	4^3	8.27
n_τ, n_ω, n_ν	46	8.28
	62	8.29
	94	8.29

ducting/insulating materials. For the comparative purposes, the calculations with scGW approach (scheme A) and with QSGW approach have also been performed in this study.

An important step is checking that the band gaps are converged with respect to the calculation setup for the vertex part, which was less intensive as compared to the setup for the GW part. Table I shows how the evaluated band gap of MgO (all studied in this work materials show similar convergence) depends on the calculation setup for the vertex part. The setup for the *GW* part was fixed at the level presented below in the Table II. When a certain setup parameter (in the vertex part) was varied, the rest of the vertex part parameters were kept at their values also indicated in Table II. As it follows, the most sensitive parameter is the number of band states included in the basis set. However, with about 20-30 bands being sufficient for the vertex part, this number is approximately

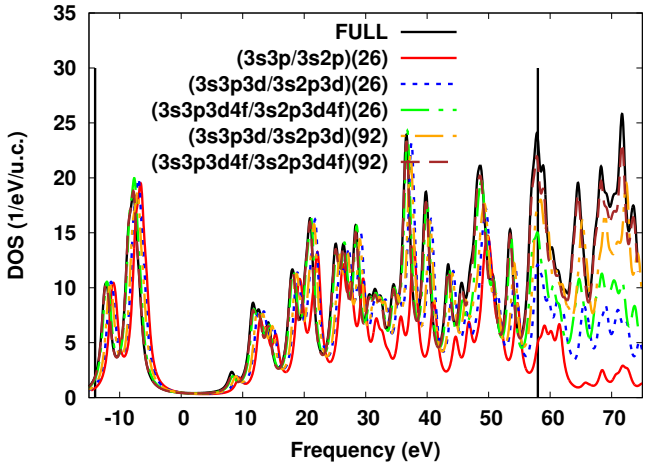

FIG. 5: (Color online) Dependence of the density of states of MgO (from the effective Hartree-Fock problem) on the quality of the representation of the band states. Line 'FULL' corresponds to the properly normalized band states (in this case they are represented with the sets of plane waves and ϕ's as in GW part, see Table II). The rest of the lines have been obtained with the sets of plane waves and ϕ's of different completeness. Each line is marked with the corresponding set of ϕ's (first entry, Mg/O), and the number of plane waves $N_{\mathbf{G}}$ in the interstitial region (second entry). Two vertical lines show the range of the band states ($\sim$20 bands) used in the vertex part as a basis set. Formula (D4) from Ref.[10] has been used to evaluate the spectral function. Smearing parameter was set to 0.005 Ry to avoid extremely sharp peaks. The k-mesh $3 \times 3 \times 3$ corresponding to the vertex part was used.

seven times less than the number of band states needed for the GW part (Table II). One can estimate from Table I that the uncertainty of the calculated band gap related to the changing of the vertex part setup is about $0.03eV$, which comprises less than 5% of the vertex correction to the band gap (see Table III below). It is also a few times smaller

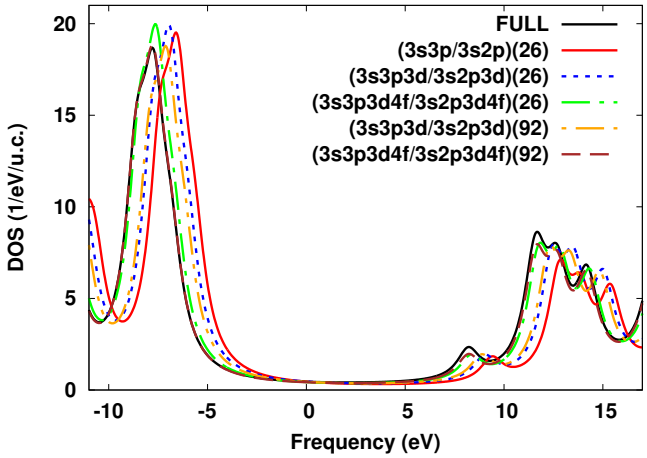

FIG. 6: (Color online) The same as Figure 5 but for the low energy part of the spectrum.

TABLE II: Setups of the calculations. $N_{\mathbf{k}}$ stands for the number of **k**-points in the Brillouin zone. N_{lapw} is the number of linearized augmented plane waves (LAPW's) in the basis set (GW part). ϕ_{nl} enumerates all orbitals (shown as sequences of principal quantum numbers and orbital characters) inside MT spheres which were used to represent the band states. Those of them shown in bold were used to augment the plane waves. The rest are the local orbitals (LO's). The total number of bands used as a basis for the GW part of the calculations is the sum of all LO's and LAPW's. The columns $(L_{max}/N)^{PB}_{MT}$ show maximal orbital character (L_{max}) and the total number of product basis (PB) functions in the MT sphere for a specific atom. $N^{PB}_{\mathbf{G}}$ stands for the number of PB functions in the interstitial region (GW part). $N_{\mathbf{G}}$ shows (for the Vertex part) the number of plane waves used to represent the band states and the PB functions in the interstitial region. The number of points on the Matsubara's time/frequency meshes was 62 in the GW part and 46 in the vertex part. The basis set of the band states in the vertex part included 20-30 bands closest to the chemical potential.

	GW part						Vertex part		
	$N_{\mathbf{k}}$	N_{lapw}	ϕ_{nl}	$(L_{max}/N)^{PB}_{MT}$	$N^{PB}_{\mathbf{G}}$	$N_{\mathbf{k}}$	ϕ_{nl}	$(L_{max}/N)^{PB}_{MT}$	$N_{\mathbf{G}}$
Si	8^3	190	**3s3p3d4f5g6h7i**2s4s5s2p4p5p4d5d5f6g7h(Si)	6/332(Si)	652	4^3	**3s3p3d**(Si)	4/44(Si)	59
SiC	6^3	144	**3s3p3d4f5g6h7i**2s4s2p4p4d5f6g(Si)	6/294(Si)	650	3^3	**3s3p3d**(Si)	4/45(Si)	92
			2s2p3d4f5g6h7i3s4s3p4p4d5f6g(C)	6/291(C)			**2s2p3d** (C)	4/44(C)	
C	8^3	144	**2s2p3d4f5g6h7i**1s3s4s3p4p4d5f(C)	6/275(C)	652	4^3	**2s2p3d**(C)	4/47(C)	59
GaAs	6^3	144	**4s4p4d4f5g6h7i**3s5s3p5p3d5d5f6g(Ga)	6/344(Ga)	650	3^3	**4s4p4d**(Ga)	4/90(Ga)	92
			4s4p4d4f5g6h7i3s5s3p5p3d5d5f6g(As)	6/344(As)			**4s4p4d**(As)	4/90(As)	
MgO	6^3	113	**3s3p3d4f5g6h7i**2s4s5s2p4p5p4d5d5f6g7h(Mg)	6/332(Mg)	380	3^3	**3s3p3d**(Mg)	4/97(Mg)	92
			3s2p3d4f5g6h7i2s4s3p4p4d5d5f6g7h(O)	6/314(O)			**3s2p3d**(O)	4/90(O)	
ZnS	6^3	190	**4s4p3d4f5g6h7i**3s5s6s3p5p6p4d5d6d5f6f6g7h(Zn)	6/381(Zn)	650	3^3	**4s4p**(Zn)	2/19(Zn)	92
			3s3p3d4f5g6h7i2s4s5s2p4p5p4d5d5f6g7h(S)	6/292(S)			**3s3p**(S)	2/19(S)	
ZnSe	6^3	190	**4s4p3d4f5g6h7i**3s5s6s3p5p6p4d5d6d5f6f6g7h(Zn)	6/381(Zn)	648	3^3	**4s4p**(Zn)	2/19(Zn)	92
			4s4p4d4f5g6h7i3s5s6s3p5p6p3d5d6d5f6g7h(Se)	6/226(Se)			**4s4p** (Se)	2/19(Se)	
LiF	6^3	113	**2s2p3d4f5g6h7i**1s3s4s3p4p4d5f6g7h(Li)	6/286(Li)	376	3^3	**2s2p3d**(Li)	4/44(Li)	92
			2s2p3d4f5g6h7i3s4s3p4p4d5f6g7h(F)	6/286(F)			**2s2p3d**(F)	4/44(F)	
NaCl	6^3	113	**3s3p3d4f5g6h7i**2s4s5s2p4p5p4d5d5f6g7h(Na)	6/332(Na)	376	3^3	**3s3p**(Na)	2/37(Na)	92
			3s3p3d4f5g6h7i4s5s4p5p4d5d5f6g7h(Cl)	6/329(Cl)			**3s3p**(Cl)	2/33(Cl)	
BN	6^3	113	**2s2p3d4f5g6h7i**1s3s3p4d(B)	6/255(B)	380	3^3	**2s2p3d**(B)	4/36(B)	92
			2s2p3d4f5g6h7i3s3p4d(N)	6/255(N)			**2s2p3d**(N)	4/36(N)	
AlP	6^3	190	**3s3p3d4f5g6h7i**2s4s2p4p4d5f(Al)	6/290(Al)	650	3^3	**3s3p3d**(Al)	4/45(Al)	92
			3s3p3d4f5g6h7i2s4s2p4p4d5f(P)	6/290(P)			**3s3p3d**(P)	4/44(P)	

than the difference between band gaps obtained in schemes B and D. Thus, the accuracy is sufficiently good to allow us to compare the different schemes (at least at the level when one neglects by the e-ph interaction).

As it was mentioned above, the basis set for the vertex part consists from the smaller number of band states and the representation of these "lower energy" band states inside the MT spheres and in the interstitial region is provided with a reduced number of ϕ_{nl}-orbitals and the plane waves correspondingly. So, it is interesting to check how this approximate

TABLE III: Theoretical and experimental band gaps (eV). Abbreviations A, B, and D have been introduced in I. Experimental data have been cited from Refs.[9,17–19]. Corrected experimental results (electron-phonon/spin-orbit interaction) are based on Refs.[20–23].

	GGA	G_0W_0	Earlier QSGW			Present work				exp	exp+corr
	Ref.[24]	Ref.[24]	Ref.[9]	Ref.[7]	Ref.[26]	QSGW	A	B	D		
Si	0.56	1.12	1.23	1.41	1.47	1.41	1.55	1.32	1.26	1.17	1.22-1.24
SiC	1.36	2.38	2.52	2.88	2.90	2.79	2.89	2.52	2.42	2.40	2.51
C	4.16	5.69	5.94	6.18	6.40	6.18	6.15	5.80	5.73	5.48	5.80-5.88
GaAs	0.42	1.15	1.93	1.85	1.75	1.96	2.27	1.80	1.72	1.52	1.69
MgO	4.75	7.52		9.16	9.29	9.42	9.31	8.24	7.96	7.83	7.98
ZnS	2.07	3.35	4.04	4.15		4.19	4.28	3.90	3.79	3.83	3.94
ZnSe	1.15	2.34	3.08			3.17	3.32	2.96	2.80	2.82	3.00
LiF	9.28	14.27		15.9		16.63	16.30	15.02	14.39	14.2	14.48
NaCl	5.12	7.92				9.81	9.25	8.55	8.49	8.5	
BN	4.46	6.36		7.14	7.51	7.06	7.06	6.37	6.30	6.25	6.51,6.6
AlP	1.57	2.37		2.90	3.10	2.80	2.84	2.53	2.44	2.45	2.47

representation for the "lower energy" band states affects their normalization and how it is related to the convergence of the calculated band gap presented in the Table I. A convenient way to compare the normalization of the band states is to plot the density of states (DOS) (Figures 5 and 6). As one can see, the accuracy of the representation is very good at the energies close to the chemical potential. The accuracy is under control in the interval of the band states used as a basis in the vertex part of the calculations. The most important ϕ_{nl}-orbitals for the vertex part are the $3s$-orbitals of Mg and the $2p$-orbitals of O. They ensure proper normalization of the band states at low energies and they provide most of the band gap correction (see Table I). The *spd*-basis set in the MT spheres and 59 plane waves in the interstitial region allow to obtain well converged band gaps for this material. Inclusion of the f-orbitals inside the MT spheres and further increasing the number of plane waves in the interstitial region improve the normalization at higher energies (Fig.5), but do not affect the calculated band gap much, providing a further support that the selected basis set (Table II below) for the vertex part is sufficiently good for our purposes.

For the materials with a band gap, studied in this work, the quality of the **k**-mesh in the Brillouin zone is not very critical and, correspondingly, it was safe to use the coarser meshes like $3 \times 3 \times 3$ or even $2 \times 2 \times 2$ for the vertex part. In the case of metals, the using of the

coarser meshes for the vertex part should still be appropriate with the exception of when the vertex function is evaluated for zero (or very small) external bosonic frequency. In this case, one has to make sure that the details of the Fermi surface are handled with sufficient accuracy.

Table II shows the setups of the calculations for all materials studied in this work, distinguishing the basis sets used in the GW part and in the vertex part. The information presented in Table II might be useful as a reference in the future. Experimental crystal structures have been used: the diamond structure for Si and C, the zinc-blende structure for SiC, GaAs, ZnS, ZnSe, BN, AlP, and the rocksalt (B1) structure for LiF, NaCl. Temperature was fixed at 1000K.

II. RESULTS

The main results of this work are collected in Table III, where the band gaps obtained with two vertex corrected schemes (B and D in the notations of I) are compared with other theoretical results and with the experimental data. GGA and G_0W_0 results are cited from the work [24] where the completeness of the basis set of the LAPW+LO method with respect to the increase in the number of LO's was specifically checked. The principal conclusion from Table III is the following: vertex corrected schemes (B and D) allow one to considerably improve the results from scGW/QSGW calculations. As one can see, the biggest correction comes from the first order vertex (scheme B), and scheme D reduces the gaps from scheme B a little more. Essentially, the difference between band gaps obtained with schemes B and D (and the deviation of the corresponding gaps from the experimental data) is of about the same magnitude or less than the correction of the band gaps originating from the electron-phonon/spin-orbit interaction (MgO and LiF slightly deviate from this rule). The fact that the effects of the electron-phonon interaction were evaluated only approximately (often with use of the LDA wafe functions to evaluate the correction) makes it difficult to conclude decisively wchich scheme (B or D) is better. Thus, further improvement in the theoretical approach should, obviously, include electron-phonon interaction in a certain way.

As it was mentioned in the Introduction and as one can judge from the Table III, the GGA band gaps deviate strongly from the experimental ones (often more than 50%). G_0W_0 calculations also underestimate the gaps with an average deviation of about 10-15% (in the case

of GaAs the underestimation is somewhat bigger - 40%). QSGW calculations in this work have shown only marginal (if any) improvement as compared to the scGW approach and the corresponding band gaps differ noticeably from the experimental data. Among the QSGW results published earlier, the band gaps reported in Ref.[26] are the closest ones to the QSGW gaps obtained in this study. There are considerable differences in the gaps reported in the earlier QSGW works. Particularly, the gaps from Ref.[9] are systematically smaller (GaAs is an exception) than the gaps reported in Refs.[7,26] and obtained in this study. Differences in numerical methods (linear muffin-tin orbitals (LMTO) in [9], projector-augmented-wave (PAW) in [7], and norm-conserving pseudo-potentials in [26]) can contribute to the discrepancies among calculated band gaps. Besides, the construction of the quasiparticle spectrum in [7] is somewhat different. The disagreements may also arise from the different degree of convergence with respect to the basis set. This speculation appears when one thinks about a rather strong effect of the high energy local orbitals (LO) in the LAPW+LO basis set upon the calculated band gaps reported recently in G_0W_0 calculations.[24,27–30]

Analyzing the calculations performed with scheme D, one can learn about the convergence of the diagrammatic series. Namely, every subsequent iteration of the Bethe-Salpeter equation represents an addition of the more complicated diagram to the vertex function. Figure 7 shows the convergence of this process. All studied materials show very good convergence. After six iterations (9 in case of GaAs), the correction is reduced by two orders,

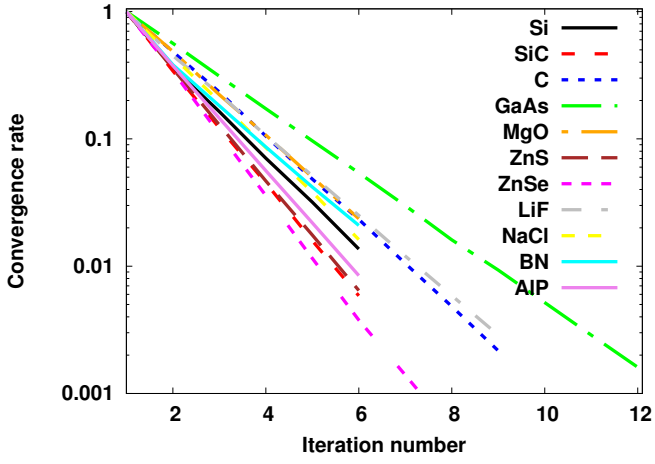

FIG. 7: (Color online) Convergence of the vertex function with respect to the iteration number when one solves the Bethe-Salpeter equation. Shown are the ratios of the largest element of the correction to the vertex at a given iteration and at the first iteration.

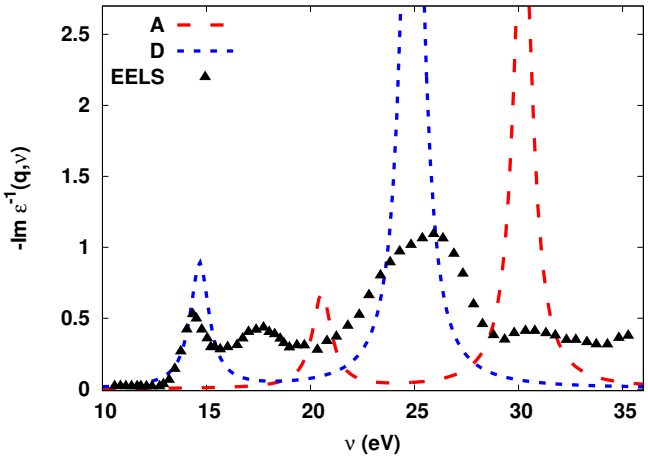

FIG. 8: (Color online) Calculated electron energy loss spectrum of LiF at q=0.5GX. Symbols represent the EELS data reproduced from the Ref.[31]

which is enough for a very good convergence of the band gaps. The reason for a somewhat slower convergence in the case of GaAs is not clear at this point and can be an objective of a separate study.

Whereas the principal goal of this study was to clarify the effect of vertex corrections on the calculated band gaps, it is obvious that other information can be extracted from the calculations. As an example, Fig.8 shows the imaginary part of the inverse dielectric function $\epsilon^{-1}(\mathbf{q},\nu)$ as a function of real frequency ν evaluated for the momentum $\mathbf{q} = 0.5\Gamma X$ (halfway between Γ and X points in the Brillouin zone). The calculated results are compared with the experimentally measured electron energy loss spectrum (EELS). Presented theoretical curves were obtained by analytically continuing[32] the $\epsilon^{-1}(\mathbf{q},i\nu)$ calculated as a function of Matsubara's frequency. As one can see, by taking into account the multiple electron-hole scattering through the solution of the Bethe-Salpeter equation (scheme D) one can considerably improve the agreement with the experimental data. Two principal peaks at approximately 15eV and 25eV are reproduced sufficiently well. The smaller peak at about 18eV noticeable in the experimental data is absent in the calculations, presumably because of insufficient accuracy of the analytical continuation. Another drawback of the need to perform analytical continuation is that the original data (on the imaginary axis) have to be of a very high quality. Particularly, the function $\epsilon^{-1}(\mathbf{q},\nu)$ with ν on the real axis was stabilized only when the number of Matsubara's time points was increased up to 100-120. It is interesting, that the band gaps, the evaluation of which also includes analytical continuation (of the

self-energy), show considerably faster convergence (Table I).

Of course, the calculation of the dielectric function and of the proprties related with it, by itself does not represent anything new in computational solid state physics. In fact, the Bethe-Salpeter equation as a tool for studying the excitons was introduced more than 50 years ago[33] and the applications of it began as early as 1980.[34] What is new in the present work is that the kernel of the Bethe-Salpeter equation is obtained consistently from the scGW calculation. At the same time, it is quite common even nowadays to use the LDA or the G_0W_0 approximation to generate G and W for using in the Bethe-Salpeter equation. Often one supplements the LDA with the scissor operator to correct the one-electron spectrum. These common approximations are either not self-consistent or (like LDA) not diagrammatic and, as a result, their connection with the Bethe-Salpeter equation is not clear. Another step forward is the elimination of the so called static approximation for W in the kernel of the Bethe-Salpeter equation which still is in common use.

III. CONCLUSIONS

In conclusion, two vertex corrected schemes to solve the Hedin's equations self-consistently have been applied to calculate the band gaps for a number of semiconductors/insulators. Undoubtedly, they both improve the results from scGW and QSGW calculations considerably. For this class of materials, the approach based on the physical polarizability (scheme D) results in slightly smaller band gaps than the conserving scheme B. However, it is hard to say decisively which scheme is better because the difference in their results is small and, in fact, is often less than the size of the electron-phonon effects, which are known only approximately.

Comparisons with the experimental data suggest that one has to take the electron-phonon effects into consideration if one wants to enhance the predictive power of the theoretical approach.

This work was supported by the U.S. Department of energy, Office of Science, Basic Energy Sciences as a part of the Computational Materials Science Program. I thank Bartomeu Monserrat for pointing out the importance of the electron-phonon effects when comparing the theoretical and the experimental band gaps. Numerous discussions with Gabriel Kotliar

are highly appreciated.

---

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[2] L. Hedin, Phys. Rev. **139**, A796 (1965).

[3] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).

[4] M. S. Hybertsen and S. G. Louie, Phys. Rev.B **34**, 5390 (1986).

[5] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[6] J. E. Northrup, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987).

[7] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).

[8] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).

[9] T. Kotani and M. van Schilfgaarde, S. V. Faleev, Phys. Rev.B **76**, 165106 (2007).

[10] A. Kutepov, K. Haule, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **85**, 155129 (2012).

[11] R. T. M. Ummels, P. A. Bobbert, and W. van Haeringen, Phys. Rev. B **57**, 11962 (1998).

[12] L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett. **88**, 066404 (2002).

[13] A. Grüneis, G. Kresse, Y. Hinuma, and F. Oba, Phys. Rev. Lett. **112**, 096401 (2014).

[14] A. L. Kutepov, Phys. Rev. B **94**, 155101 (2016).

[15] R. Del Sole, L. Reining, R. W. Godby, Phys. Rev. B **49**, 8024 (1994).

[16] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, Phys. Rev. Lett. **94**, 186402 (2005).

[17] F. J. Himpsel, J. A. Knapp, J. A. VanVechten, and D. E. Eastman, Phys. Rev. B **20**, 624 (1979).

[18] P. Lautenschlager, M. Garriga, S. Logothetidis, and M. Cardona, Phys. Rev. B **35**, 9174 (1987).

[19] A. Mang, K. Reimann, St. Rübenacke, and M. Steube, Phys. Rev. B **53**, 16283 (1996).

[20] B. Monserrat and R. J. Needs, Phys. Rev. B **89**, 214304 (2014).

[21] G. Antonius, S. Ponce, P. Boulanger, M. Cote, and X. Gonze, Phys. Rev. Lett. **112**, 215501 (2014).

[22] M. Cardona, M. L. W. Thewalt, Rev. Mod. Phys. **77**, 1173 (2005).

[23] B. Monserrat, Phys. Rev. B **93**, 100301 (2016).

[24] H. Jiang, and P. Blaha, Phys. Rev. B **93**, 115203 (2016).

[25] G. Strinati, Riv. del Nuovo Cimento **11**, 1 (1988).

[26] W. Chen and A. Pasquarello, Phys. Rev. B **92**, 041115 (2015).

[27] C. Friedrich, A. Schindlmayr, and S. Blügel, T. Kotani, Phys. Rev. B **74**, 045104 (2006).

[28] C. Friedrich, M. C. Müller, and S. Blügel, Phys. Rev. B **83**, 081101 (2011).

[29] C. Friedrich, M. C. Müller, and S. Blügel, Phys. Rev. B **84**, 039906(E) (2011).

[30] D. Nabok, A. Gulans, and C. Draxl, Phys. Rev. B **94**, 035118 (2016).

[31] W. A. Caliebe, J. A. Soininen, E. L. Shirley, C. -C. Kao, K. Hämäläinen, Phys. Rev. Lett. **84**, 3907 (2000).

[32] H. J. Vidberg and J. W. Serene, J. Low Temp. Phys. **29**, 179 (1977).

[33] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966).

[34] W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980).