Bandgaps and band bowing in semiconductor alloys

Titus Sandu\textsuperscript{a,b}, Radu I. Iftimie\textsuperscript{b}

\textsuperscript{a}International Centre of Biodynamics, Intrarea Portocalelor Street, Nr:1B, Postal Code 060101, District 6, Bucharest, Romania
\textsuperscript{b}Département de Chimie, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Quèbec H3C 3J7, Canada

Abstract

The bandgap and band bowing parameter of semiconductor alloys are calculated with a fast and realistic approach. The method is a dielectric scaling approximation that is based on a scissor approximation. It adds an energy shift to the bandgap provided by the local density approximation (LDA) of the density functional theory (DFT). The energy shift consists of a material-independent constant weighted by the inverse of the high-frequency dielectric constant. The salient feature of the approach is the fast calculation of the dielectric constant of alloys via the Green function (GF) of the TB-LMTO (tight-binding linear muffin-tin orbitals) in the atomic sphere approximation (ASA). When it is applied to highly mismatched semiconductor alloys (HMAs) like Zn\textsubscript{1-x}Se\textsubscript{x}, this method provides a band bowing parameter that is different from the band bowing parameter calculated with the LDA due to the bowing exhibited also by the high-frequency dielectric constant.

Keywords: A. disordered systems, A. semiconductors, D. electronic band structure, D. dielectric response
PACS: 71.10.-w, 71.15.Ap, 71.55.Gs

1. Introduction

Semiconductor alloying is a very important way of bandgap engineering with applications in tailoring the materials properties like the electronic structure or optical behaviour. The bandgap of semiconductor alloys can be usually described by a quadratic polynomial in concentration $x$ with the quadratic coefficient being called the bowing coefficient. Accordingly, the bandgap $E_g$ of an alloy $\text{AB}_x\text{C}_{1-x}$ is described by

$$E_g(x) = xE_g(\text{AB}) + (1-x)E_g(\text{AC}) - bx(1-x),$$

where $E_g(\text{AB})$ and $E_g(\text{AC})$ are the bandgaps of the binary constituents $\text{AB}$ and $\text{AC}$. The bandgap bowing coefficient $b$ describes the deviation from linearity and is close to a constant for most alloys.

Highly mismatched semiconductor alloys (HMAs) exhibit a large bowing parameter. HMAs are present in III-V systems like Ga\textsubscript{1-x}As\textsubscript{x}, Ga\textsubscript{1-x}Bi\textsubscript{x}, and Ga\textsubscript{1-x}Sb\textsubscript{x}. II-VI compounds like Zn\textsubscript{1-x}S\textsubscript{x}, Zn\textsubscript{1-x}Te\textsubscript{x}, and Zn\textsubscript{1-x}O\textsubscript{x} are also present. IV-IV alloys like Sn\textsubscript{x}Ge\textsubscript{1-x} are also present. Lattice relaxation, localization, charge transfer, and huge bandgap bowing are encountered in HMA systems due to large differences in atomic sizes and orbital energies, and large lattice mismatch.

The bandgap bowing of HMAs cannot be described well by simple models like virtual crystal approximation, at least in the mid-alloy regime of concentrations. Other models like the band anticrossing (BAC) model\textsuperscript{1}, the valence-band anticrossing (VBAC) model\textsuperscript{2} or combinations of them\textsuperscript{3,4} are able to explain the strong bandgap change with concentration $x$. These models are applied in the dilute limit of concentrations, in which one constituent is considered as an impurity with composition-dependent coupling to the conduction band minimum or valence band maximum. However, the BAC model and its extensions completely neglect cluster states induced by impurities\textsuperscript{5}. Super-cell methods overcome this shortcoming but the price is large unit cells\textsuperscript{10}. More efficient approaches are based on special quasirandom structures (SQSs) having smaller supercells in which the most relevant atom-atom correlation functions are similar to those of random alloys\textsuperscript{11}.

Standard first-principles band theory is based on the Kohn-Sham (KS) scheme\textsuperscript{12} of the density functional theory (DFT), which is, in principle, exact. The DFT is a total-energy method for calculating the elec-
tronic ground state. The local density approximation (LDA) \cite{13} and the generalized gradient approximation (GGA) \cite{14} provide reasonably accurate results for ground-state properties like crystal structure, pressure-induced transformations, phonon spectra, magnetism, etc. Usually the LDA leads to some overbindings. For example, in materials, where narrow bands determined by semi-core states (for instance, Zn 3d states in ZnSe and ZnTe) contribute to the bonding, the LDA predicts these localized semi-core states too high in energy. In addition, the LDA and GGA do not describe well the long-range part of correlations and have the so-called “bandgap problem”. Even though the LDA often yields good band dispersions for the valence bands, the LDA bandgap energy in semiconductors and insulators is much smaller than the experimental bandgap. The LDA bandgap overestimates the bandgap energy in semiconductor materials by a factor of 2. Method

There is still a simpler method to calculate the semiconductor bandgaps. It is based on the scissor operator \cite{16} and is basically a dielectric scaling approximation (DSA) for correcting the LDA bandgaps \cite{26}. Thus for a wide variety of materials the main correction to the LDA bandgap is \cite{26}

\[ \Delta \approx \frac{\alpha}{\varepsilon_{\infty}} \]  

(2)

where \( \alpha = 9.1 \) eV is a universal constant and \( \varepsilon_{\infty} \) is the high-frequency dielectric constant of the material. The method is applied not only to usual zinc blende or wurtzite types but also to other structures like technologically interesting high-dielectric materials \cite{27}. Besides the technical arguments presented in the original paper \cite{26} there are other arguments that ensure the success of this method. First, there is the perturbative argument of the GWA with respect the LDA \cite{18}: the GW corrections are basically the expectation value of the GW self-energy over the LDA eigenvectors. Second, the difference between quasiparticle energies and the LDA eigenvalues is essentially due to the non-local nature of the effective potential \cite{28}, which can be expressed with a scissor-like operator \cite{26}. There is also a more physical picture provided by Harrison \cite{29}. The
exact non-local exchange operator of the Hartree-Fock theory has a long-range part which is unscreened, and hence overestimated. In the LDA the exchange operator is replaced by a local exchange-correlation operator which neglects the long-range part of the exchange altogether. In semiconductors the long-range part of the exchange is screened by the high-frequency dielectric constant. The LDA uses a fixed one-electron potential for all bands including the unoccupied bands. Thus the LDA neglects the electrostatic energy in equilibrium. We wish to calculate the response charge in the random phase approximation \[16\]. CONTB-LMTO-ASA. The procedure is laid out below.

The frequency dielectric constant but not that well for total energy calculations, especially for closely packed systems, in particular for energy bands, but not that well for total energy calculations. Finally the static polarization \( \chi_0 \) and the high-frequency dielectric constant \( \varepsilon_{\infty} \) are calculated with the TB-LMTO-ASA. The procedure is laid out below.

Let us then consider the static screening of a test charge in the random phase approximation \[16\]. Consider a lattice of points (spheres), with the electron density in equilibrium. We wish to calculate the response (screening) in the electron charge \( \delta q_j \) at site \( j \), induced by the addition of a small external potential \( \delta V_i \) at site \( i \). The electron charge \( \delta q_j \) is related to \( \delta V_i \) by the non-interacting response function \( \chi_{ij}^0 \):

\[
\delta q_j^0 \equiv \chi_{ij}^0 \delta V_i.
\]

\( \chi_{ij}^0 \) can be calculated in the first-order perturbation theory from the eigenvectors of a non-interacting one-electron Schrödinger equation with the Adler-Wiser formalism \[37, 38\] or directly via \( G^0 \), the Green function (GF) calculated from the one-electron Hamiltonian of the LDA. By linearizing the Dyson equation for the perturbed GF, i.e. \( \delta G = G^0 \delta VG \approx G^0 \delta V G^0 \), one obtains an explicit representation of \( \chi_{ij}^0 \) in terms of \( G^0 \):

\[
\delta q_j^0 = \frac{1}{\pi} \text{Im} \sum_{k} \delta G_{ij}(k, z) = \frac{1}{\pi n} \sum_{k} \delta G_{ij}^0(k, z) G_{ji}^0(k, z) \delta V_j^0 = \chi_{ij}^0 \delta V_j^0.
\]

In Eq. 3 the \( k \) summation is the summation in the Brillouin zone and the energy-integration contour encloses the occupied states. A large part of the computer resources is allocated to the calculation of \( \chi^0 \). The calculation of \( \chi^0 \) by the Adler-Wiser formula would require the utilization of large computer resources because it needs all the occupied orbitals at once. In contrast, Eq. 4 shows clearly that \( \chi^0 \) can be calculated piecwisely, and hence efficiently once the GF is obtained. Equation 3 is the TB-LMTO-ASA counterpart of the static polarization function defined in real space \[16\]. Similarly, the dielectric function is:

\[
\varepsilon = 1 - V \chi^0,
\]

where \( V \) is the unscreened Coulomb matrix. The macroscopic dielectric constant \( \varepsilon_{\infty} \) is calculated as \[39\]

\[
\varepsilon_{\infty} = \frac{1}{N_o} \sum_{\beta} \lim_{k \to 0} 1/[\epsilon_{\beta \beta}(k)]^{-1}.
\]

In equation 6 \( N_o \) is the total number of orbitals. The indices \( \beta = (R, l, m) \) denote the lattice \( R \) and the angular-momentum quantum numbers \( l \) and \( m \). Nevertheless equation 6 neglects "local field" effects and it represents the upper bound of \( \varepsilon_{\infty} \) when "local field" effects are considered \[39\].

The GF can be directly calculated within the LDA of the TB-LMTO-ASA by using the potential parameters, \( C, \Delta, \) and \( \gamma \) \[33, 34, 35\]:

\[
G(z) = -\frac{1}{2} \rho_c(z) + \sqrt{\rho_c(z)} \left[ P^c(z) - S^c(z)^{-1} \right]^{-1} \sqrt{\rho_c(z)}.
\]
where \( \alpha \) is the representation, determined by the screening constants \( \alpha_{RI} \) and

\[
P_R^\alpha(\gamma) = \frac{z - C_R}{\Delta_R + (\gamma - \alpha_R) (z - C_R)}
\]

is the potential function. \( P^\alpha \) and \( P^{\bar{\alpha}} \) are, respectively, the first and the second derivatives of the potential function with respect to energy. \( S^\alpha = S_{Rm}^\alpha R_{m'}^{\bar{\alpha}} \) is the matrix of structure constants in the \( \alpha \) representation. In the LMTO-ASA approach, the effective potential is entirely described by the potential parameters: the band center \( C \), bandwidth \( \Delta \), and the distortion parameter \( \gamma \). The potential function provides the boundary conditions for the radial Schrödinger equation inside each WS sphere [33, 34, 40]. The fact that both the charge density and the effective potential are described by a few atomic parameters makes the calculation of the response functions a very simple task.

3. Results and Discussions

For illustration we show the calculations of the bandgap corrections in the ZnTe\(_S\)Se\(_{1-x}\) alloys. The random zinc blende alloys were modeled as supercells of 54 atoms (27 cations) with special quasirandom structures (SQSs) [11]. The supercells have the lattice vectors \((3a_1, 3a_2, 3a_3)\) with \((a_1, a_2, a_3)\) the unit cell vectors of the binary constituents. The SQS structures are constructed such that the most relevant atom-atom correlation functions of the SQS structure are similar to those of random alloys [11]. We assumed that the alloy lattice constants are determined by the weighted average of the lattice constants of the constituents (Vegard’s law) [41]. In the ASA, the spherical average of the charge in each sphere makes the TB-LMTO approach less reliable and more cumbersome in calculating the equilibrium positions of the atoms in the SQS cells. For that reason the atomic positions inside the SQS cells were relaxed using VASP with the LDA exchange-correlation potential; at relaxed positions, the quantum-mechanical forces on each atom were less than 0.03 eV/Å. As we have already mentioned the atomic sphere approximation assumes that the whole space is filled with (overlapping) spheres and that the volume of the interstitial region vanishes. This is achieved by using empty spheres [36]. The empty spheres must satisfy several criteria: the sum of the volume of all spheres should equal the volume of all space, the average overlap between the spheres should be minimal, and the radii of the spheres should be chosen such that the spheres overlap in the regions close to the local maxima of the electrostatic potential. The positions of the empty spheres are well known in the high-symmetry systems like zinc blende structures. In relaxed SQS cells the positions and the radii of the empty spheres were chosen to satisfy the above criteria starting from the initial positions of the non-relaxed (higher-symmetry) SQS cells. The Bril-louin zone integration was performed on a \(4\times4\times4\) mesh by a Monkhorst-Pack \(k\)-point sampling scheme [42].

In Fig. 1 we plot the calculated dielectric constant \( \varepsilon_\infty \) estimated with the TB-LMTO-ASA. There is an estimated bowing of about \(-0.89\).

![Figure 1: Macroscopic dielectric constant \( \varepsilon_\infty \) estimated with the TB-LMTO-ASA.](image)

The calculated values of \( \varepsilon_\infty \) for the binary constituents are 10 – 20% larger than the experimental values and 5 – 15% larger than the time-dependent density functional theory (TD-DFT) calculations [43]. There is a bowing shown also by the macroscopic dielectric constant \( \varepsilon_\infty \), which is \( b \approx -0.89 \). Equation (2) and the bowing exhibited by \( \varepsilon_\infty \) show us that the DFT-LDA alone cannot reproduce the bowing of the bandgap in ZnTe\(_S\)Se\(_{1-x}\).

Fig. 2 shows the bandgaps calculated with GGA-VASP and LDA-VASP, with the TB-LMTO-ASA in the LDA, and with the DSA defined by (2). The TB-LMTO-ASA results were considered with combined corrections [40]. The bandgaps calculated with GGA-VASP are larger than the results of LDA-VASP over entire range of concentrations. The bandgaps of the binary constituents calculated with LDA-VASP are both 1.06 eV. The LDA ASA gives a bandgap value of 1.07 eV for ZnSe and 1.05 eV for ZnTe. The ASA values are pretty close to the LMTO full potential values which are 1.05 eV and 1.03 eV, respectively [17]. Over the whole range of Te concentration \( x \), the ASA results are in quite good agreement with the results obtained with PAW-VASP which, in principle, is a full potential method with no shape approximation. Fig. 2 also shows the bandgaps calculated with the DSA, which ex-
hbit much larger bandgaps than both the LDA and the GGA. The bandgaps of ZnSe and ZnTe calculated with the DSA are 2.38 eV and 2.07 eV, respectively, being comparable with the results of the GW A, which, in general, underestimates the bandgaps because it generates too much screening due to overestimation of \( \chi^{\mu} \) [17]. As we expected, the bandgaps show a bowing, i.e., the bandgap deviates from linear interpolation with respect to composition \( x \). The bowing parameter \( b \) is 1.64 eV for the VASP-GGA and 1.57 eV for the VASP-LDA. The LDA of the TB-LMTO-ASA produces a bowing of 1.50 eV, while the DSA estimates the bowing parameter to 1.71 eV. The bowing parameters predicted with methods that go beyond the LDA are larger. The fact that the LDA does not reproduce the entire bowing of the bandgap is rooted not only in the LDA's inability to predict the bandgap but also in the overbinding produced by the LDA. For ZnTe, ZnSe, and ZnTe\(_{1-x}\)Se\(_x\), the LDA overbinding pushes the cation \( d \) levels closer to the top of the valence band. Thus the \( d \) levels will mix too much with the \( p \) states at the maximum of the valence band; therefore additional errors in estimating the band bowing occur. On the other hand it is known that the GWA moves these cation \( d \) states in the right direction downwards [17].

Old measurements of the bandgap bowing parameters \( b \) with and without spin-orbit interaction are presented in [44]. The authors estimated that the bandgap bowing is \( b = 1.266 \) eV. They also estimated the spin-orbit splitting, which shows a concave bowing. Thus their estimated bandgap bowing without spin-orbit coupling is \( b = 1.07 \) eV. Other measurements estimate a bowing parameter of 1.62 eV (at 5 K) and 1.51 eV (at 300 K) [45]. More recent experiments set the bowing of the bandgap between 1.25 eV and 1.5 eV at 300 K [4, 46] and the bowing of the spin-orbit splitting at -0.33 eV [4]. Other measurements at low temperature, namely at 13 K, indicate a bowing parameter around 2.21 eV [47].

Several general features can be outlined from all experimental data and our calculations. More recent experiments reveal larger band bowing coefficients than the older ones. The bowing is larger at low temperatures. There is also a bowing of the spin-orbit splitting but it is negative. The bandgap bowing calculated with spin-orbit coupling is larger than the bandgap bowing without spin-orbit coupling by one third of the bowing parameter modulus of the spin-orbit splitting. Overall, our results for bandgap bowing are in rather good agreement with recent experimental results.

### 4. Summary and Conclusions

To summarize, we have presented a computational scheme that realistically estimates the bandgaps and band bowing parameters with applicability to semiconductors and their alloys. The scheme uses the static polarization function \( \chi^{\mu} \) to calculate the high-frequency dielectric constant \( \varepsilon_{\text{HF}} \) in the GF-TB-LMTO formalism. The corrections to the LDA bandgaps are made with a dielectric scaling approximation (DSA), which estimates the correction to the LDA bandgap as a universal constant divided by the high-frequency dielectric constant.

The approach inherits its speed from the fact that it works piecewisely to calculate \( \chi^{\mu} \) via GF formalism and its efficiency from the TB-LMTO (the use of the muffin-tin orbitals and a minimal basis).

The limitations of the method were already discussed in [26]. The narrow-bandgap semiconductors are particularly difficult to describe with the DSA because the other corrections of the GWA to the LDA bandgap become as important as the DSA. Also, often in these semiconductors there might appear a wrong ordering of the LDA levels around the bandgap with poor results in GWA too [17].

We applied the scheme to the ternary alloy ZnTe\(_{1-x}\)Se\(_x\) by calculating the bandgap and band bowing in a supercell with 54 atoms that mimics the alloying of ZnTe\(_{1-x}\)Se\(_x\). The ternary alloy ZnTe\(_{1-x}\)Se\(_x\) exhibits large bandgap bowing. The bandgap bowing computed with the DSA is in good agreement with the experimental data. It was also shown that the bowing of the bandgap in ZnTe\(_{1-x}\)Se\(_x\), cannot be explained by the LDA alone due to inherent bowing shown also by the dielectric constant.
5. Acknowledgments

This work has been supported in part by the Romanian Ministry of Education and Research under the project “Ideas” No.120/2007. The work was also partially supported by the Natural Science and Engineering Research Council of Canada (NSERC) grant and was made possible by the computational resources of the Reseau Qu´e´ebecois de Calcul de Haute Performance (RQCHP).

[1] W. Shan, W. Walukiewicz, J. W. A. III, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, S. R. Kurz, Phys. Rev. Lett. 82 (1999) 1221.

[2] K. Alberi, J. Wu, W. Walukiewicz, K. M. Yu, O. D. Dubon, S. P. Watkins, C. X. Wang, X. Liu, Y.-J. Cho, J. Furdyna, Phys. Rev. B 75 (2007) 045203.

[3] W. Walukiewicz, W. Shan, K. M. Yu, J. W. A. III, E. E. Haller, I. Miotkowski, M. J. Seong, H. Alawadhi, A. K. Ramdas, Phys. Rev. Lett. 85 (2000) 246403.

[4] J. Wu, W. Walukiewicz, K. M. Yu, J. W. A. III, E. E. Haller, I. Miotkowski, A. K. Ramdas, C. H. Su, I. K. Sou, R. C. C. Pereira, J. D. Denlinger, Phys. Rev. B 67 (2003) 035207.

[5] K. M. Yu, W. Walukiewicz, J. Wu, W. Shan, J. W. Beeman, M. A. Scarpulla, O. D. Dubon, P. Becla, Phys. Rev. Lett. 91 (2003) 246403.

[6] K. Alberi, J. Blacksherg, L. D. B. S. Nikzad, K. M. Yu, O. D. Dubon, W. Walukiewicz, Phys. Rev. B 77 (2008) 073202.

[7] N. Dit, J. M. Obaidat, H. Alawadhi, J. Phys.: Condens. Matter 21 (2009) 075802.

[8] T. Sandu, W. P. Kirk, Phys. Rev. B 72 (2005) 073204.

[9] P. R. C. Kent, A. Zunger, Phys. Rev. B 64 (2001) 115208.

[10] A. Lindsay, E. P. O’Reilly, Phys. Rev. Lett. 93 (2004) 196402.

[11] A. Zunger, S. W. Wei, L. G. Ferreira, J. E. Bernard, Phys. Rev. Lett. 65 (1990) 353.

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

[13] P. Hohenberg, W. Kohn, Phys. Rev. 136 (1964) B864.

[14] J. P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865.

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

[16] F. Aryasetiawan, O. Gunnarsson, Rep. Prog. Phys. 61 (1998) 237.

[17] M. van Schilfgaarde, T. Kotani, S. V. Faleev, Phys. Rev. B 74 (2006) 245125.

[18] L. Hedin, Int. J. Quantum Chem. 56 (1995) 445.

[19] M. L. del Puerto, M. L. Tiago, J. R. Chelikowsky, Phys. Rev. B 77 (2008) 045404.

[20] F. Gygi, A. Baldereschi, Phys. Rev. Lett. 62 (1989) 2160.

[21] F. Bechstedt, R. D. Sole, G. Cappellini, L. Reining, Solid State Commun. 84 (1992) 765.

[22] D. M. Bylander, L. Kleinman, Phys. Rev. B 41 (1990) 7868.

[23] A. Seidl, A. Gorling, P. Vogl, J. A. Majewski, M. Seidl, Phys. Rev. B 53 (1996) 3764.

[24] J. Furthmuller, G. Cappellini, H. C. Weissker, F. Bechstedt, Phys. Rev. B 66 (2002) 045110.

[25] B. Lee, L. W. Wang, Phys. Rev. B 73 (2006) 153309.

[26] V. Fiorentini, A. Baldereschi, Phys. Rev. B 51 (1995) 17196.

[27] A. Kersch, D. Fischer, J. Appl. Phys. 106 (2009) 041405.

[28] M. Gruning, A. Marini, A. Rubio, Phys. Rev. B 74 (2006) 161103(R).

[29] W. A. Harrison, Phys. Rev. B 31 (1985) 2121.

[30] F. Tran, P. Blaha, Phys. Rev. Lett. 102 (2009) 226401.

[31] G. Kresse, J. Furthmuller, Phys. Rev. B 54 (1996) 11169.

[32] G. Kresse, D. Joubert, Phys. Rev. B 59 (1999) 1758.