Improved Electronic Structure and Optical Properties
of sp-Hybridized Semiconductors Using LDA+U^{SIC}

Clas Persson[*] and Susanne Mirbt[†]

[*]*Department of Materials Science and Engineering, Royal Institute of Technology, SE–100 44 Stockholm, Sweden*
[†] *Department of Physics, Uppsala University, SE–751 21 Uppsala, Sweden*

We propose the local density approximation (LDA) plus an on-site Coulomb self-interaction-like correction (SIC) potential for describing *sp*-hybridized bonds in semiconductors and insulators. We motivate the present LDA+U^{SIC} scheme by comparing the exact exchange (EXX) hole with the LDA exchange hole. The LDA+U^{SIC} method yields good band-gap energies E_g and dielectric constants $\varepsilon(\omega \approx 0)$ of Si, Ge, GaAs, and ZnSe. We also show that LDA consistently underestimates the Γ-point effective electron m_c and light-hole m_{lh} masses, and the underlying reason for this is a too strong light-hole–electron coupling within LDA. The advantages of the LDA+U^{SIC} approach are a computational time of the same order as the ordinary LDA, the orbital dependent LDA+U^{SIC} exchange-correlation interaction is asymmetric analogously to the EXX potential, and the method can be used for materials and compounds involving localized *d*- and *f*-orbitals.

I. INTRODUCTION

It is a well known fact that the LDA within the density functional theory (DFT) underestimates the band gap of semiconductors by about 50%. This so band-gap problem has frequently been studied in the past. Its origin has been attributed to the missing discontinuity in the exchange-correlation potential [1] and the self-interaction error within the LDA [2]. In other words, LDA fails to correctly describe the atomic limit [3]. Many physical properties are more or less determined by the band gap. Optical properties, transport properties, properties that depend on band matching, all strongly depend on the band gap and are often not trustworthy when calculated within LDA.

We propose in this work to use the LDA+U scheme for semiconductors in order to cure the band-gap problem. By construction the Coulomb U correction leads to a discontinuity in the exchange-correlation potential and approximately solves the self-interaction error of LDA [4]. The advantage is that one obtains for an appropriate chosen value of U not only the band gap but the entire band structure almost in agreement with experiment and moreover, the charge density and total energy are calculated for the corrected band gap. The tendency of LDA to average the interaction potential (and thus also the exchange hole) is in the present LDA+U corrected by spatially delocalizing (U>0) the *d*-states, and localizing (U<0) the *p*-states.

II. PHYSICAL MOTIVATION

The problem of LDA is its description of the exchange hole. LDA works as good as it does due to it fulfilling the sum rules connected to the exchange hole. But the self-interaction error manifests itself in an insufficient description of the spatial dependence of the exchange hole. Close to the atomic core the LDA exchange hole is usually too small compared to an exact exchange calculation, but further away the LDA exchange hole becomes too large.

For strongly localized states the LDA-correction should thus increase the exchange hole. This corresponds to an increase in the electron-electron repulsion (positive U), because the hole becomes larger if the electrons are kept further away from each other. On the other hand this effectively corresponds to an increased exchange interaction. An increased exchange interaction leads to a lowering of the state, because of the following: The core electrons shield the attractive field of the nuclei, thereby raising the energy of the valence and conduction states. But the exchange interaction counteracts this in lowering the valence and conduction states.

For more delocalized states, the situation is vice versa. The LDA correction should decrease the exchange hole. This corresponds to a decrease of the electron-electron repulsion (negative U), which in turn corresponds to a decreased exchange potential. The energy level in question is thus shifted higher in energy. This now also explains the dependence of the band gap on the self-interaction error. With a too strong electron-electron repulsion within LDA, the electron (wavefunction) overlap is too small and thus also the semiconductor band gap becomes too small. Adding a negative Coulomb U correction increases the overlap, i.e., hybridization, and thus opens up the band gap.

The more delocalized a state becomes, the smaller becomes the self-interaction error. In a semiconductor the conduction-band edge has mainly *s*-character and is more delocalized than the *p*-like valence-band edge. In this work we accordingly correct the *p*-states. We tested also an additional correction of the *s*-states, but find its effect to be relatively small for Si, Ge, GaAs, and ZnSe. In addition, correction to *d*-states modifies the screened core potential and thereby influences the *p*-states.

We use here the so called SIC LDA+U (or atomic limit LDA+U) [5], where we use the following effective potential and total energy
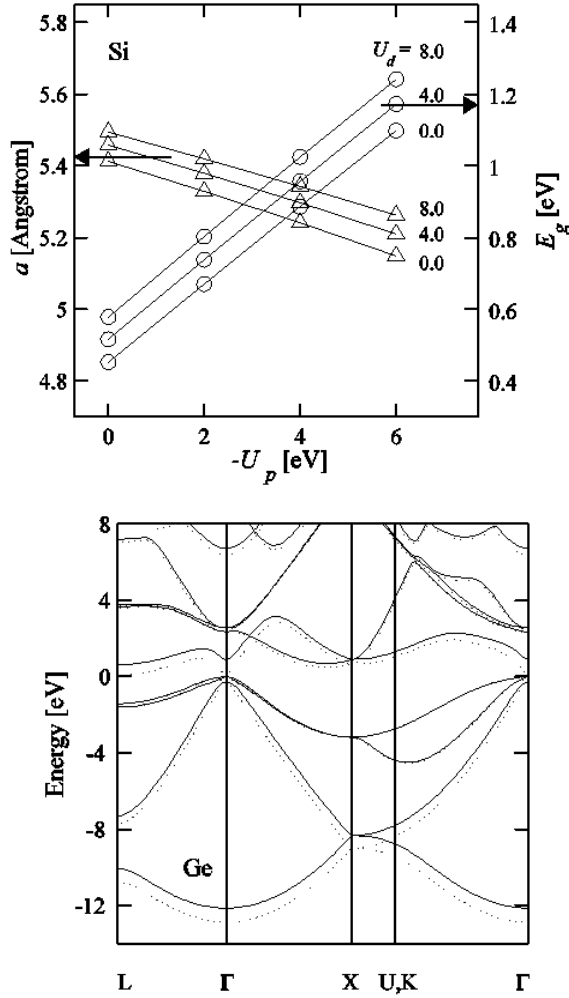
FIG. 1: (a) The LDA+U^{SIC} lattice constant (triangles) and fundamental band gap (circles) of Si as functions of $-U_p$ for U_d = 0.0, 4.0, and 8.0 eV. Arrows indicate the zero-temperature experimental values of Ref. 6. (b) Ge band structure from LDA (dotted) and LDA+U^{SIC} (solid lines).

$$V(\text{LDA} + \text{U}^{SIC}) = V(\text{LDA}) + \frac{U}{2}\left(1 - 2\sum_{m,\sigma} n_{m,\sigma}\right)$$

$$E(\text{LDA} + \text{U}^{SIC}) = E(\text{LDA}) + \frac{U}{2}\sum_{m,\sigma}\left(n_{m,\sigma} - n_{m,\sigma}^2\right)$$

where $n_{m,\sigma}$ is the orbital occupancy of the l-orbital in question (i.e., s, p, or d) and σ is the spin.

The U$_p$ value we find to best reproduce the experimental band structure is rather large (-8 eV for ZnSe). The current LDA+U^{SIC} implementation only allows us to correct the fraction of the p-valence states that are within some core-radius. Thus, the interstitial part of the eigenfunctions is not affected.

This restriction most likely leads to a higher U value than otherwise necessary. The disadvantage of this Coulomb U correction is the worsened ground-state volume. Our correction increases the LDA overbinding error, because we neglect the self-interaction error of almost all localized states. The effect of the self-interaction error of the d-states on the ground-state volume is seen in Fig. 1(a). For a given U correction on the p-level, the lattice constant increases with increased localization of the d-states. To improve the ground-state volume, the self-interaction error of all core states need to be corrected. This agrees with the EXX calculations [2] where it was found, that the exact exchange treatment of the core electrons improved the lattice constant. For our purposes the correct band gap is sufficient, but the optimization of the ground-state volume and cohesive energy is a future task.

III.   RESULTS

The present LDA+U^{SIC} model has been applied to Si, Ge, GaAs, and ZnSe which represent four different types of semiconductors. A relativistic, full-potential linearized augmented plane wave (FPLAPW) method has been employed with Perdew and Wang exchange-correlation potential [7]. The on-site Coulomb U correction was included by a rotational invariant scheme with a SIC-like double counting correction [8]. For comparisons, all LDA+U^{SIC}, LDA, and GGA calculations have been performed with the experimental equilibrium volume [6].

The strength of the U correction was determined by optimizing the band-gap energies to the experimental values [6]. In order to correct the localized cation d-like states in Ge, Ga, and Zn (which are known to have too high energies within LDA [9]), we use the Coulomb parameter U_d = 8 eV, which lowers the energy of the d-bands by about 2.6, 2.9, and 3.8 eV in Ge, GaAs, and ZnSe, respectively. The Coulomb U interaction parameter for the p-like states was optimized to U_p =-5.0, -5.7, -7.0, -8.0 eV for Si, Ge, GaAs, and ZnSe, respectively.

The band-gap energies are strongly improved using the LDA+U^{SIC} approach (Table I), especially for Ge for which LDA produces metallic ground state with an incorrect band-edge degeneration (i.e., inverted valence-band maximum: $E_g(\Gamma_{8v}^+ - \Gamma_{7c}^-)$= - 0.30 eV). In Fig. 1(b), the electronic structure of Ge is shown, obtained with LDA and LDA+U^{SIC}. With the LDA+U^{SIC} potential (solid lines), the degeneration at the Γ-point comes out correctly. One also obtains the correct position of the conduction-band minimum (at the L-point) with a band-gap energy of 0.73 eV, which can be compared to the experimental value of 0.74 eV [6]. The direct band-gap energy of 0.93 eV at the Γ-point agrees very well with measured 0.90 eV [6].

GaAs is known to be almost metallic in the fully relativistic LDA calculation since the valence-band maximum is lifted up by $\sim$0.1 eV by the spin-orbit interaction [10]. We obtain $E_g(\Gamma_{8v} - \Gamma_{6c})$ = 0.20 and 0.45 eV with the LDA and GGA potentials, respectively. LDA+U^{SIC} approach yields $E_g(\Gamma_{8v} - \Gamma_{6c})$= 1.37 eV.

Also in Si the LDA+U^{SIC} model can reproduce the exper-

TABLE I: The band-gap energy E_g (in eV), the valence-band width ΔW (in eV), and the dielectric constant ε.

	LDA	GGA	LDA+U	Expt.[6]	GW [11,12]
<u>Si</u>					
$E_g(\Gamma_{8v}^+ - \Delta_{6c})$	0.45	0.55	1.17	1.17	1.18
$E_g(\Gamma_{8v}^+ - \Gamma_{6c}^-)$	2.48	2.52	2.74	3.35	3.22
ΔW	12.0	12.0	11.5	12.5	11.6
ε	14.7	14.1	10.0	11.9-12.1	
<u>Ge</u>					
$E_g(\Gamma_{8v}^+ - L_{4c}^-)$	−0.07	0.10	0.73	0.74	0.62
$E_g(\Gamma_{8v}^+ - \Gamma_{7c}^-)$	−0.30	−0.29	0.93	0.90	0.57
ΔW	13.0	12.9	11.8	12.7	12.8
ε		21.3	16.3	16.0-16.5	
<u>GaAs</u>					
$E_g(\Gamma_{8v} - \Gamma_{6c})$	0.20	0.45	1.37	1.52	1.22
$E_g(\Gamma_{8v}-L_{4c})$	0.74	0.91	1.47	1.82	1.61
ΔW	12.9	12.9	11.8	13.1	12.5
ε_∞	16.0	14.5	12.0	10.6	
<u>ZnSe</u>					
$E_g(\Gamma_{8v} - \Gamma_{6c})$	1.06	1.31	1.80	2.82	2.96
$E_g(\Gamma_{8v}-L_{4c})$	2.36	2.55	2.70		
ΔW	13.2	13.1	12.7	15.2	13.4
ε_∞	8.9	8.3	8.8	7.1-8.8	

TABLE II: Effective electron and hole masses (in m_0).

	LDA	GGA	LDA+U	Expt. [6]	EXX [2]
<u>Si</u>					
$m_c^{\perp\Delta}$	0.19	0.20	0.22	0.19	0.22
$m_c^{\parallel\Delta}$	0.96	0.97	1.02	0.92-0.98	0.97
m_{hh}	0.46	0.47	0.46	0.54	0.60
m_{lh}	0.14	0.14	0.16	0.15	0.19
m_{so}	0.22	0.22	0.23	0.23	0.29
<u>Ge</u>					
$m_c^{\perp L}$	0.07	0.07	0.10	0.08	0.10
$m_c^{\parallel L}$	1.68	1.70	1.83	1.57-1.74	1.59
m_c^{Γ}	0.02	0.005	0.05	0.04	
m_{hh}		0.38	0.39	0.30-0.33	
m_{lh}		0.006	0.06	0.04	
m_{so}		0.03	0.11	0.08	
<u>GaAs</u>					
m_c^{Γ}	0.01	0.03	0.09	0.07	0.10
m_{hh}	0.49	0.50	0.54	0.53-0.59	0.64
m_{lh}	0.02	0.03	0.10	0.08	0.12
m_{so}	0.07	0.09	0.16	0.13	0.20
<u>ZnSe</u>					
m_c^{Γ}	0.08	0.10	0.14	0.13-0.17	
m_{hh}	0.94	0.99	0.90	1.1	
m_{lh}	0.10	0.12	0.16	0.12	
m_{so}	0.18	0.22	0.28		

imental value of the fundamental band gap (1.17 eV). The conduction-band energy difference $\Delta E_g(\Gamma_{6c}^- - \Delta_{6c}) = 1.57$ eV, is about 0.61 eV too small compared to the experimental value [6]. The balance between the two on-site Coulomb interaction parameters, U_p and U_d, has an affect on the **k**-dependence of the energy distribution of the lowest conduction band in Si. Therefore one has to carefully consider the electronic struc-

ture of the whole Brillouin zone (and not only the fundamental band-gap energy) when one chooses the U correction parameters. Since the LDA+U^{SIC} can be regarded as an energy independent Hartree–Fock like GW theory, at least for the localized states [13], one might expect that an energy dependence of the Coulomb interaction parameter U will result in better **k**-dependence of the lowest conduction band of Si. The
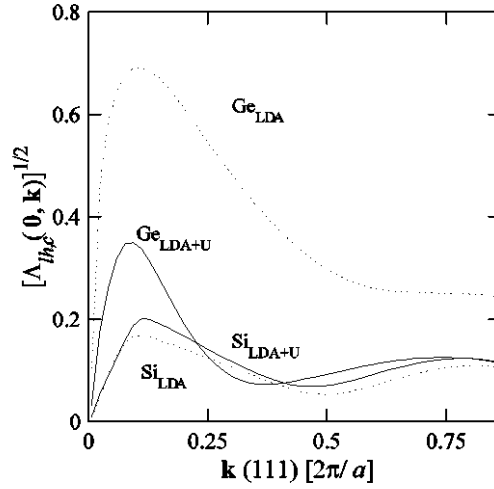
FIG. 2: $[\Lambda_{lh,c}(\mathbf{0},\mathbf{k})]^{1/2}$ between light-hole states at the Γ- point and the lowest conduction band along the Γ L $\equiv \Lambda$ line.

width of the valence bands ΔW in both LDA+U^{SIC} and EXX is somewhat decreased compared to LDA, whereas GW better reproduces the experimental valence-band width (Table I). The narrowing of the band width will give somewhat too large values of the effective masses, seen also for the EXX potential [2].

For ZnSe, the present LDA+U^{SIC} requires large U to correct the LDA band gap. There are three main reasons for that: First, the anion-*s* band is below the cation-*d* band, and consequently the *s*-states are very localized. Thus, one needs a SIC for these states (i.e., U_s). Second, we correct the potential only in the atomic muffin-tin regions. By using empty spheres in the interstitial one can use smaller U_l value. Third, one may also correct the core states.

The optical properties in terms of the dielectric constant are improved with the LDA+U^{SIC} approach (Table I). Primarily, this is explained by the band-gap correction. However, the **k**-dependence of the interaction energy across the band gap has an effect on the optical absorption.

The curvature of the energy dispersion near the band edges can be represented by the effective electron and hole masses. Normally, LDA predict the effective electron and hole masses fairly accurately despite incorrect band gap [14], but it is also well known that LDA fails to describe the curvature of the conduction band in GaAs properly [10]. It is however not known that this is a generic LDA failure for Γ-point effective masses of semiconductors. In Table II we present the the effective electron and hole masses. The LDA fails to predict the Γ-point electron and light-hole masses accurately. We have seen this effect also for other semiconductors with small or moderate Γ-point band-gap energy. This shows that a constant energy shift (i.e., the "sissor operator") is not sufficient to correct the band-edge properties of semiconductors.

We explain the LDA problem of describing the Γ-point effective masses by showing the strength of the coupling between the conduction- and valence-band edges. Since LDA

couples the bonding and antibonding states this also explains the zero band gap of Ge. This bonding-antibondig interaction can be visualized by the overlap integral of the Bloch functions of the valence-band maximum and the lowest conduction band:

$$\Lambda_{j,c}(\mathbf{0},\mathbf{k}) = \frac{1}{2} \sum_{\sigma,\sigma'=\uparrow\downarrow} \left| \int u_{j\sigma}(\mathbf{0},\mathbf{r})u_{c\sigma'}(\mathbf{k},\mathbf{r})d\mathbf{r} \right|.$$

where $j = hh$, lh, and *so* denote the heavy-hole, light-hole, and spin-orbit split-off bands, respectively.

In Fig. 2, we show the square root of the overlap integral $\Lambda_{lh,c}(\mathbf{0},\mathbf{k})$ between the light-hole and conduction bands of Ge and Si along the ΓL$\equiv \Lambda$ line. We do not show $\Lambda_{hh,c}$ or $\Lambda_{so,c}$ since these overlap integrals are almost zero. Clearly, LDA has a strong overlap between the light-hole band and the conduction band near the Γ-point. Due to this too strong LDA overlap integral of the $\Lambda_{lh,c}$ near the Γ-point for semiconductors with small of moderate Γ-point band-gap energy, the LDA yields too small Γ-point effective electron and light-hole masses. This is corrected by LDA+U^{SIC}. Thus, the correction of the band-gap energy within the LDA+U^{SIC} leads also to more accurate values of the Γ-point effective electron and light-hole masses. For instance, the measured effective electron masses at the Γ-point of Ge and GaAs are 0.04m_0 and 0.07m_0, respectively. The calculated LDA values are 0.02m_0 and 0.01m_0, whereas the corresponding LDA+U^{SIC} values are 0.05$m_0$and 0.09m_0, respectively. The effective hole masses of Ge are improved considerably by the LDA+U^{SIC}: measured value is $m_{lh} = 0.04m_0$, the GGA value is 0.006m_0, and the LDA+U^{SIC} value is 0.06m_0. On the whole, the LDA+U^{SIC} potential yields accurate values of both the electron and the hole masses. Fig. 2 explains why only Γ-point masses are affected and not the electron masses away from the Γ-point (e.g. Si effective electron masses at the Δ-line): the overlap integral $\Lambda_{lh,c}(\mathbf{0},\mathbf{k})$ is small for **k** at the Brillouin zone edge. We have seen the same effect for several low- and moderate band-gap semiconductors.

IV. CONCLUSION

The LDA+U^{SIC} improves (*i*) the band-gap energies of *sp*-hybridized semiconductors, (*ii*) the effective electron and light-hole masses, as well as (*iii*) the optical properties. Moreover, LDA+U^{SIC} scheme is also appropriate to lower the cation *d* energies in Ge, GaAs, and ZnSe, and this energy shift was found to have an affect on the *sp*-hybridization at the valence-band maximum. LDA+U^{SIC} has a more asymmetric potential than LDA [15]. LDA+U^{SIC} model can therefore be used for a wide range of different semiconductor-metal-oxid systems.

[1] J. P. Perdew *et al*., Phys. Rev. Lett. **51**, 1884 (1983); L. J. Sham *et al*., *ibid*. 1888 (1983).

[2] M. Städele *et al*., Phys. Rev. Lett. **79**, 2089 (1997).

[3] W. Koch *et al.,* A Chemists Guide to Density Functional Theory, (Wiley, Weinheim, 2001).

[4] I. V. Solovyev, *et al*., Phys. Rev. B **50**, 16861 (1994).

[5] M. T. Czyzyk *et al*., Phys. Rev. B **49**, 14211 (1994).

[6] *Semiconductor Basic Data*, ed. by O. Madelung (Springer, Berlin, 1996).

[7] P. Blaha *et al*., WIEN2k, An Augmented Plane Wave +Local Orbitals Program, 2001.

[8] V. I. Anisimov *et al*., Phys. Rev. B **48**, 16929 (1993); A. I. Liechtenstein *et al., ibid*. **52**, 5467 (1995); P. Novák *et al., ibid*

[9] C. Persson *et al*., Phys. R **68**, 073205 (2003).

[10] N. E. Christensen, Phys. Rev. B **30**, 5753 (1984); C. Persson *et al*., Rev. B **64**, 033201 (2001).

[11] A. Fleszar, Phys. Rev. B **64**, 245204 (2001).

[12] M. Oshikiri et al., Phys. Rev. B **60**, 10754 (1999).

[13] V. I. Anisimov *et al*., J. Phys.: Condens. Matter **9**, 767 (1997).

[14] C. Persson *et al*., J. Appl. Phys. **82**, 5496 (1997); *ibid*, B **54**, 10257 (1996).

[15] P. Mohn *et al*., Phys. Rev. Lett. **87** 196401 (2001); S. Sharma *et al*., (unpublished)