Agoston, Péter; Körber, Christoph; Klein, Andreas; Puska, Martti; Nieminen, Risto; Albe, Karsten

Limits for n-type doping in In2O3 and SnO2: A theoretical approach by first-principles calculations using hybrid-functional methodology

Published in:
Journal of Applied Physics

DOI:
10.1063/1.3467780

Published: 01/01/2010

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Limits for $n$-type doping in $\text{In}_2\text{O}_3$ and $\text{SnO}_2$: A theoretical approach by first-principles calculations using hybrid-functional methodology

Péter Ágoston, Christoph Körber, Andreas Klein, Martti J. Puska, Risto M. Nieminen, and Karsten Albe

Citation: Journal of Applied Physics 108, 053511 (2010); doi: 10.1063/1.3467780
View online: http://dx.doi.org/10.1063/1.3467780
View Table of Contents: http://aip.scitation.org/toc/jap/108/5
Published by the American Institute of Physics

Articles you may be interested in

Evaporated Sn-doped $\text{In}_2\text{O}_3$ films: Basic optical properties and applications to energy-efficient windows
Journal of Applied Physics 60, R123 (1998); 10.1063/1.337534

First-principles calculations for defects and impurities: Applications to III-nitrides
Journal of Applied Physics 95, 3851 (2004); 10.1063/1.1682673
I. INTRODUCTION

Transparent conducting oxides (TCO) exhibit a unique combination of electrical conductivity and optical transparency in the visual range.\(^1\) \(\text{In}_2\text{O}_3\) and \(\text{SnO}_2\) are widely used in optoelectronic devices\(^2\) and in gas-sensing applications.\(^3\) Under reducing conditions these materials exhibit \(n\)-type conductivity and oxygen deficiency due to the occurrence of oxygen vacancies.\(^4\)–\(^6\) and possibly hydrogen.\(^7\)–\(^9\)

The conductivity in these TCO materials is usually obtained by adding large amounts (2\%–10\%) of \(n\)-type dopants to the host.\(^1\) The conductivity is then mainly determined by the free electron concentration in the conduction band. On the other hand, the conductivity is given by the doping efficiency, i.e., the degree of ionization of the dopant, which is generally high for frequently-used material combinations such as tin in \(\text{In}_2\text{O}_3\) (ITO) or fluorine or antimony in \(\text{SnO}_2\) (FTO and ATO, respectively).

While ITO shows the highest conductivities of all the TCO materials its high price has stimulated research for finding cheaper alternatives.\(^10\),\(^11\) ATO, on the other hand, is cheap but lower conductivities obtained\(^12\) hinder its commercial application. Doping on the anion site with fluorine in FTO increases the conductivities but FTO is still significantly outperformed by ITO.\(^10\) Additionally, it is difficult to use fluorine in conjunction with sputtering techniques. For these materials degenerate carrier densities can easily be obtained, reflected in a Fermi level position within the conduction band. High Fermi energies, however, can only be achieved, if no compensating intrinsic defects such as cation vacancies and oxygen interstitials are limiting the range of accessible Fermi level positions. Therefore a detailed understanding of the properties of intrinsic acceptor-type compensating defects is necessary for estimating \(n\)-type doping limits in these materials.

Frank and Köstlin\(^13\) proposed the well-established view that the doping limit in ITO should be exclusively ruled by the occurrence of interstitial oxygen, which at the same time binds to the tin dopants. To the best of our knowledge a similar model for ATO or FTO has not yet been established. Experimentally, no significant influence of the oxygen partial pressure on the conductivity was found for ATO samples,\(^12\) which suggests that the presence of intrinsic acceptors is of minor importance in this material. In the case of FTO, a decrease in conductivity, however, has been reported for high fluorine contents, which was assumed to be due to the occurrence of fluorine interstitial acceptors.\(^14\),\(^15\) On the other hand, anomalies in electrical conductivity measurements of pure and slightly acceptor-doped \(\text{SnO}_2\) samples as a function of oxygen partial pressure and dilatation measurements\(^16\) hint to the presence of cation vacancies at elevated temperatures.

Density-functional theory (DFT) calculations within the generalized gradient approximation (GGA) have confirmed that oxygen interstitials in \(\text{In}_2\text{O}_3\) are more stable than indium vacancies,\(^5\) which is in line with the defect model of Frank and Köstlin. Of intrinsic acceptor in \(\text{SnO}_2\), tin vacancies have lower formation energies than oxygen interstitials.\(^6\) In both studies, however, the formation energies are severely underestimated leading to pinning values of the Fermi energy well below those experimentally accessible.\(^17\) This would imply, that in contrast to experimental findings, \(n\)-type doping is hardly possible especially under more oxygen-rich conditions and that the Fermi level cannot enter into conduction band significantly. If one considers the large formation entropies of acceptor-type defects\(^18\) the situation becomes even worse. The main reason for this deficiency is the shortcomings of local approximations to DFT which are especially significant in the case \(n\)-type TCOs.\(^4\)

In this study, we revisit the problem of the thermodynamic stability of acceptor-type point defects in \(\text{In}_2\text{O}_3\) and \(\text{SnO}_2\) using hybrid-functional DFT in conjunction with finite-size corrections based on the local density approximation (LDA) and large supercells.
II. METHODOLOGY

A. Computational approach

Total energy calculations were carried out using the VENNA AB INITIO SIMULATION PACKAGE.\textsuperscript{19,20} For the representation of exchange and correlation we use the hybrid-functionals HSE06 and PBE0 for In$_2$O$_3$ and SnO$_2$, respectively.\textsuperscript{21–24} The two functionals are closely related and only differ by the value of the range-separation parameter which we have adjusted in order to reproduce the band gaps. Within this approach the band-gap problem is resolved and no correction schemes need to be applied. The potentials due to the nuclei and the core electrons were represented by the projector augmented wave scheme.\textsuperscript{25,26} The plane wave cut-off energy was set to 500 eV to assure well converged results. Within the Brillouin-zones were sampled with $2 \times 2 \times 2$ Monkhorst–Pack $k$-point grids.\textsuperscript{27} Ionic relaxations were carried out until the forces on the unclamped ions decayed to less than 0.01 eV/Å.

Within the hybrid-functionals scheme, structural as well as thermodynamic parameters are reproduced well and the band gaps agree closely with the experimental values.\textsuperscript{28–30} In the case of hybrid-functional calculations we used supercells containing 72 and 80 atoms for SnO$_2$ and In$_2$O$_3$, respectively. Due to the high computational cost a systematic finite-size scaling using hybrid functionals is not yet feasible.

B. Finite-size effects

Because of the existence of highly charged defect states (charge $q = 3$ for V$_{\text{In}}$ and $q = 4$ for V$_{\text{Sn}}$), the effects of the finite cell-size have to be considered and corrected. In this case the electrostatic image charge interactions arising from electrostatic monopoles embedded in a jellium countercharge is the dominating energy contribution. Based on the LDA calculations using cell sizes up to 640 and 576 atoms for In$_2$O$_3$ and SnO$_2$, respectively, we have confirmed the $E \sim V^{1/3}$ long-range scaling behavior in our calculations including full structural relaxation of the supercells at constant volume. Since we observed the expected long-range interaction regime in our calculations we obtained the correction term $\Delta E = E_{\text{LDA}}^{30 \times 72} - E_{\text{LDA}}^\infty$ by fitting the two first terms of the Makov–Payne series.\textsuperscript{31} We have taken the defect formation energies as obtained from the hybrid-functional calculations and added the LDA finite-size correction term to them. For highly charged defects this LDA correction can be seen as a lower boundary of the positive correction values, since the static dielectric constant is overestimated by the LDA. Additional hybrid-functional calculations with cell sizes of 40 and 162 for In$_2$O$_3$ and SnO$_2$, respectively, confirmed this trend. Therefore, also our calculated formation energies mark a lower boundary with respect to the finite-size correction. Using this setup we predicted approximate doping limits arising due to the occurrence of intrinsic acceptor defects in In$_2$O$_3$ and SnO$_2$.

III. RESULTS

A. Defect energetics

The formation energies were obtained as a function of the chemical potentials of the constituents $\mu_i$ and the Fermi energy $E_F$ in all relevant charge states $q$ as\textsuperscript{32}

$$\Delta G^0_i = G^0_{\text{def}} - G_{\text{host}} - \sum_i n_i \mu_i + q(E_{\text{VB}} + E_F), \quad (1)$$

where the Gibbs free energies of the supercells with $(G^0_{\text{def}})$ and without $(G_{\text{host}})$ the defect are taken at the zero-temperature and zero-pressure limits. The reference for the Fermi energy is the valence band maximum (VBM) of the host material. The allowed stability region is given by the heat of formation of the host compound which we have obtained using the hybrid functionals (see Table I).

Figure 1 shows the finite-size corrected formation energies of the acceptor defects in In$_2$O$_3$ and SnO$_2$ for maximally reducing conditions. At this limit the formation energies of all acceptors attain their highest values and are therefore suitable for the discussion of maximum doping limits. For the two materials the doping limits are approximately given by the lowest-energy intersections of the formation energies with the zero energy line in the left part of the figure. For both materials the intersections are significantly ($>2$ eV) beyond the conduction band minimum (CBM) rendering the materials truly $n$-dopable. Moreover, it is very clear that SnO$_2$ is intrinsically compensated only at Fermi energies.

<table>
<thead>
<tr>
<th>Material</th>
<th>Experiment</th>
<th>LDA</th>
<th>HSE06</th>
<th>PBE0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indium oxide</td>
<td>$a_0$</td>
<td></td>
<td></td>
<td>$\cdots$</td>
</tr>
<tr>
<td></td>
<td>$E_G$</td>
<td>10.117\textsuperscript{a}</td>
<td>10.15</td>
<td>10.23</td>
</tr>
<tr>
<td></td>
<td>$\Delta H'$</td>
<td>2.6–2.9\textsuperscript{b}</td>
<td>1.2</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−9.47\textsuperscript{a}</td>
<td>−10.6</td>
<td>−10.1</td>
</tr>
<tr>
<td>Tin oxide</td>
<td>$a$</td>
<td>4.738\textsuperscript{a}</td>
<td>4.73</td>
<td>$\cdots$</td>
</tr>
<tr>
<td></td>
<td>$c$</td>
<td>3.188\textsuperscript{a}</td>
<td>3.20</td>
<td>$\cdots$</td>
</tr>
<tr>
<td></td>
<td>$E_G$</td>
<td>3.6</td>
<td>1.2</td>
<td>$\cdots$</td>
</tr>
<tr>
<td></td>
<td>$\Delta H'$</td>
<td>−6.01\textsuperscript{d}</td>
<td>−6.8</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>Indium (tetragonal)</td>
<td>$a$</td>
<td>3.332\textsuperscript{c}</td>
<td>3.37</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>$c$</td>
<td>4.471\textsuperscript{c}</td>
<td>4.35</td>
<td>4.64</td>
</tr>
<tr>
<td>$\beta$-tin</td>
<td>$a$</td>
<td>5.831\textsuperscript{c}</td>
<td>5.72</td>
<td>$\cdots$</td>
</tr>
<tr>
<td></td>
<td>$c$</td>
<td>3.181\textsuperscript{c}</td>
<td>3.20</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>Oxygen (dimer)</td>
<td>$r_0$</td>
<td>1.208\textsuperscript{c}</td>
<td>1.22</td>
<td>1.21</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Reference 33.\textsuperscript{b}References 28 and 34.\textsuperscript{c}Reference 35.\textsuperscript{d}Reference 36.
considerably higher than In$_2$O$_3$. Namely, in this limit the compensation occurs above ~4.7 eV and ~5.7 eV for In$_2$O$_3$ and SnO$_2$, respectively.

This result is not obtained using the LDA, GGA, or GGA+U functionals. Very generally, the (semi-)local functionals are not able to provide a description of these TCO materials in accordance with experiments, i.e., the formation energies are significantly overestimated for donors, whereas they are underestimated for acceptors as we showed in this study. In practice, under experimental deposition/annealing conditions, the oxygen chemical potential is not fully at the reducing limit but may have typically values between $\mu_{O_2} = (-1.0) - (-2.0)$ eV (this is true, e.g., for the conditions of T=600 °C and $p_{O_2} < p_{O_2}^{ambient}$). For this reason we have plotted in the right part of Fig. 1 the formation energies of the acceptor defects in their predominant charge states for the Fermi level at the CBM as a function of the oxygen chemical potential. Thus, the n-dopability of both materials can be compared irrespective of the different band gaps. The figure illustrates the fact that unless degenerate doping is achieved, in both materials no acceptor defect can contribute significantly to the defect equilibria. For example, the formation energy of the double negative oxygen interstitial in In$_2$O$_3$ is ~1 eV in the oxidizing limit and for the Fermi energy at the CBM. Since the oxidizing limit is unrealistic at elevated temperatures and the Fermi energy is generally below the CBM for undoped and oxidized In$_2$O$_3$ samples, we can exclude the occurrence of intrinsic acceptor defects under any experimentally accessible conditions. As can be seen in the left part of Fig. 1 this effect is even more pronounced for SnO$_2$. For the Fermi energy at the CBM and at most oxidizing conditions the formation energies of both acceptors are larger than 1.8 eV. Note also that the use of the calculated heats of formation for the determination of the stability range underestimates the formation energies in the oxygen-rich limit whereas the values are more reliable in the metal-rich limit.  

In In$_2$O$_3$ oxygen interstitials are the predominant acceptor defects with the charge state $q = -2$. However, the indium vacancy is energetically very close to the interstitial at the doping limit, which is easily reached in ITO. In fact, based on the estimated extrapolation error of the finite-size scaling, the remaining uncertainties of the exchange-correlation functionals, and the zero-temperature approximation, it is not possible to surely predict the energetic order of the acceptor defects at the doping limit. Based on experiment the oxygen interstitial was suggested to compensate the donors. According to our calculations it is safe to assume that indium vacancies will be present in considerable numbers at high Fermi energy values. This finding is also consistent with the remarkable mobility observed for Sn cation dopants in ITO presuming a vacancy-mediated migration mechanism. For a further clarification of this point, defect-defect interactions among the donor dopants and acceptors should be still investigated.

In SnO$_2$ the energetic order of the defects is unambiguous with the vacancies more stable than the oxygen interstitials. This is consistent with the previous LDA calculations and can be explained by the close packing of the rutile structure which makes the incorporation of large anion interstitials energetically expensive. The close packing is also reflected in a strong structural relaxation of the neighboring atoms around the oxygen interstitial in SnO$_2$ (see Fig. 2).

FIG. 1. (Color online) On the left: formation energies of acceptor defects in SnO$_2$ and In$_2$O$_3$ in the metal-rich limit (left). On the right: formation energies of acceptor defects at the conduction band minima as a function of the oxygen chemical potential. The stability limits for the two materials are similar (see Table I).

FIG. 2. (Color online) Comparison of the structural relaxations of oxygen interstitials in SnO$_2$ and In$_2$O$_3$. The outward relaxation of neighboring oxygen atoms is especially large in SnO$_2$. Oxygen anions and the metal cations are marked with O and M, respectively.
the same reason also the charge-state transition \((0/-2)\) is found at high Fermi level positions for SnO\(_2\) (left part of Fig. 1).

The formation volume of the interstitial oxygen, i.e., the elastic strain can be decreased by releasing two electrons and forming a neutral oxygen dumbbell configuration on a regular oxygen lattice site. This covalently bonded configuration is, therefore, more stable than the negative charge state for Fermi level positions throughout the whole band-gap. In the negative state the surplus atom is necessarily located in the interstitial region due to the destabilization of the covalent bond upon electron addition. In comparison, the structural relaxation around the interstitial in In\(_2\)O\(_3\) is much weaker due to the presence of the large interstitial sites within the bixbyite structure (Fig. 2). Therefore the energetic cost for the accommodation of the negative interstitial is low and the charge transition level \((0/-2)\) is below the CBM.

**B. Electron concentration**

According to our hybrid-functional calculations the formation energies of acceptor defects generally increase in comparison with the LDA/GGA results. Within our choice of hybrid functionals, which reproduce the experimental band gaps, the increase is clearly larger for SnO\(_2\) than in In\(_2\)O\(_3\). This contradicts with the experimental trend indicating doping difficulties for SnO\(_2\) but not for In\(_2\)O\(_3\).\(^{10}\) In order to illustrate this finding more clearly, we have calculated the electron concentrations self-consistently on the basis of the finite-size corrected formation energies of acceptor defects obtained from hybrid-DFT total energy calculations.

The defect concentrations were obtained by the usual Boltzmann type expression,

\[
c = c_0 \exp \left( -\frac{\Delta G_f}{k_B T} \right),
\]

where \(c_0\) is the concentration of available sites for the defect and \(\Delta G_f\) the free energy of defect formation. Since we are presently only interested in the defect properties of the intrinsic acceptors, the \(n\)-dopant is assumed to be ideal in this calculation, i.e., it has the ionization probability of unity and it is ideally soluble. For both materials we use temperature-independent parabolic band edges with effective electron and hole masses of 0.3 \(m_e\) and 0.6 \(m_h\), respectively.\(^1\) Because the Fermi level can enter into the conduction band, we use the Fermi function instead of the Boltzmann approximation in order to integrate the charge carriers densities. Note that by relaxing the above approximations our arguments for the theory-experiment discrepancy in \(n\)-type doping are further strengthened. Namely, nonparabolic bands, band-gap renormalization,\(^{44}\) and a temperature dependent band gap\(^{45}\) would result in a slower increase in the Fermi level in the conduction band as a function of the free electron concentration. Therefore our calculated carrier concentrations represent lower bounds.

Figure 3 shows the free electron concentration due to heavy \(n\)-type doping as a function of the oxygen partial pressure in the range from \(10^{-15}\) to \(10^5\) Pa, which is accessible within experiments. The oxygen partial pressure is obtained from the oxygen chemical potential via the ideal gas law and using electrochemical tables for the temperature of 600 °C.\(^{39,43}\) The conductivities are given for three different doping concentrations of \(2 \times 10^{20}\), \(6 \times 10^{20}\), and \(2 \times 10^{21}\) \(\text{cm}^{-3}\), corresponding to the range of \(\sim 1\% - 10\%\) of substituted cations.

The free electron concentrations are highest at low oxygen pressures and mainly determined by the doping concentrations. For every dopant concentration there is a characteristic pressure at which the electron concentration begins to decay with increasing oxygen pressure. This characteristic transition is also found in experiments for ITO (Ref. 46) and has a temperature dependence such that it shifts to higher oxygen pressures at higher temperatures. Our selected temperature of 600 °C corresponds to an average deposition/annealing temperature\(^{12}\) and results for other temperatures can be obtained by shifting the oxygen pressure scale in the figure.

For the dopant concentration of \(6 \times 10^{20}\) \(\text{cm}^{-3}\) the decay of free electron concentration for SnO\(_2\) is found at oxygen pressures which are more than eight orders of magnitude higher than those for In\(_2\)O\(_3\). Further, for the lower dopant concentration of \(2 \times 10^{20}\) \(\text{cm}^{-3}\) neither of the materials suffers from compensation effects in the whole range of oxygen pressures. In comparison, as shown in Fig. 3 the acceptor defects lead to a strong compensation when the smaller LDA formation energies are used instead of the hybrid-functional ones. Additionally, according to the LDA results In\(_2\)O\(_3\) appears to be more dopable than SnO\(_2\), a trend which is inverted with respect to the hybrid-functional results. This is due to the larger LDA band-gap error for SnO\(_2\) than for In\(_2\)O\(_3\). The failure of the LDA to describe the strong \(n\)-type behavior of these TCOs is reflected in the fact that using LDA formation energies for the acceptor defects a partial-pressure independent region is not reached for the selected realistic dopant concentrations and environmental conditions (Fig. 3).

As we have pointed out above, our calculated free electron concentrations are likely to represent lower boundaries. The uncertainties, which are still connected to the actual experimental band-gap of In\(_2\)O\(_3\) cannot alter our conclusions. Namely, by using the largest presently-suggested band-gap
value for In$_2$O$_3$ (Refs. 34 and 47) for tuning the range-separation parameter of the exchange correlation functional does not lead to increased formation energies for the Fermi level at the CBM. Band-gap-related ambiguities do not arise in the case of SnO$_2$.

Our findings are surprising in the light of the experiments, since the conductivities reported for ITO are generally higher than in any TCO material related to SnO$_2$. Therefore, according to our hybrid-DFT results the origin of the significantly lower free electron concentrations in SnO$_2$ is not due to intrinsic acceptor defects. The results therefore suggest that the limitations of $n$-type doping in SnO$_2$ mainly arise due to dopant-specific properties rather than properties intrinsically related to SnO$_2$. A further optimization of $n$-type doping with respect to the resulting free electron concentrations is therefore possible.

For FTO the doping limit is conjectured to be caused by interstitial fluorine defects$^{12,15}$ which is now convincingly supported by our calculations. To the best of our knowledge the reason for the low electron concentrations of ATO is presently not known. Beside the occurrence of Sb$_5^+$ instead of Sb$_5^+$ (Refs. 48 and 49) which could act as acceptor, the segregation of the Sb-dopant is a possible, often-considered origin for low electron concentrations.$^{50}$ However, our calculations predict a low abundance of cation vacancies and thereby a good kinetic stability of cation dopants against vacancy-mediated migration in SnO$_2$. The kinetics will therefore be slow even if there is a tendency for segregation. More specifically our calculations explain that segregation cannot be observed for ATO (Ref. 12) but it is possible for ITO (Ref. 40) using a comparable experimental setup. As we have shown above, higher cation vacancy concentrations can be expected for highly-doped In$_2$O$_3$ samples. Since the mobility of indium vacancies is not prohibitively large$^{51}$ a higher mobility of cation dopants and higher segregation kinetics can be expected in In$_2$O$_3$ than in SnO$_2$. In this study we do not present results on specific donor dopants. A comprehensive study of the compensation mechanisms with specific dopant/material combinations is to follow.

**IV. SUMMARY AND CONCLUSION**

We have reinvestigated the electron compensation in two TCO materials In$_2$O$_3$ and SnO$_2$. We have shown that within the hybrid-functional-DFT description In$_2$O$_3$ and SnO$_2$ are highly-$n$-type dopable against the formation of intrinsic acceptors. We have obtained for In$_2$O$_3$ a doping limit which is in good agreement with experiments. This reflects the robustness of the methodology used. Most importantly, we have obtained for SnO$_2$ a doping limit, which is beyond the experimentally-observed one. We conclude that for SnO$_2$ the lower measured electron concentrations are therefore not a consequence of any intrinsic acceptor of the material.

Our general result is that SnO$_2$ is more robust toward high Fermi level values and should allow for a higher maximum doping than has so far been reached in experiment. This conclusion is unlikely to be altered by any approximation used in our calculations. Our findings in turn state that the source for lower conductivities in SnO$_2$ in comparison with In$_2$O$_3$ are related to the dopants presently used (F and Sb). While the doping limit in In$_2$O$_3$ is given by intrinsic acceptors the conductivities are limited by other processes in SnO$_2$. Likely explanations for the presently observed doping limits are therefore either the low ionization rate of the dopants, extrinsic acceptors, low solubility, or defect-defect interactions. Further improvements of cheap SnO$_2$ based TCO materials are therefore possible by using other dopants and dopant combinations.

**ACKNOWLEDGMENTS**

We acknowledge the financial support through the Sonderforschungsbereich 595 “Fatigue of functional materials” of the Deutsche Forschungsgemeinschaft and the Academy of Finland through the center of Excellence Program (2006–2011). Moreover, this work was made possible by grants for computing time at CSC computing facilities in Espoo, Finland, and FZ-Juelich. We also acknowledge financial support through a bilateral travel program funded by the German foreign exchange server (DAAD).

$^{29}$A. Walsh, J. L. F. D. Silva, S.-H. Wei, C. Körber, A. Klein, L. F. J. Piper, A. DeMasi, K. E. Smith, G. Panaccione, P. Torelli, D. J. Payne, A. Bou-


36. Thermodynamic Properties of Compounds SbO$_2$ to Rh$_2$O$_3$ (ed.), Springer Materials - The Landolt-Börnstein Database (http://www.springermaterials.com);


