Dopability, Intrinsic Conductivity, and Nonstoichiometry of Transparent Conducting Oxides

Stephan Lany and Alex Zunger

National Renewable Energy Laboratory, Golden, Colorado 80401, USA
(Received 7 September 2006; published 23 January 2007)

Existing defect models for In$_2$O$_3$ and ZnO are inconclusive about the origin of conductivity, nonstoichiometry, and coloration. We apply systematic corrections to first-principles calculated formation energies $\Delta H$, and validate our theoretical defect model against measured defect and carrier densities. We find that (i) intrinsic acceptors (“electron killers”) have a high $\Delta H$ explaining high $n$-dopability, (ii) intrinsic donors (“electron producers”) have either a high $\Delta H$ or deep levels, and do not cause equilibrium-stable conductivity, (iii) the O vacancy $V_O$ has a low $\Delta H$ leading to O deficiency, and (iv) $V_O$ has a metastable shallow state, explaining the paradoxical coexistence of coloration and conductivity.

DOI: 10.1103/PhysRevLett.98.045501

PACS numbers: 61.72.Bb, 61.50.Nw, 65.40.Gr, 71.15.Nc

Most metal oxides are insulators, even when band theory predicts them to be metallic (i.e., “Mott Insulators” [1]). An interesting opposite case is when oxides are electrically conductive, even though band theory predicts them to have wide band gaps. This behavior defines the materials class of “transparent conductive oxides” (TCO’s) [2] such as ZnO, In$_2$O$_3$, and SnO$_2$, which have large band gaps in the ultraviolet, yet the Fermi level is in the conduction band due to free-carrier producing centers. These materials exhibit a remarkable phenomenology: (i) Doping by extrinsic electron producers (e.g., Sn donors in In$_2$O$_3$ and Al donors in ZnO) leads to enormous electron densities approaching 10$^{21}$ cm$^{-3}$. Thus, in contrast to most other gap materials, “electron-killers” (i.e., intrinsic acceptor-defects such as the cation vacancy) do not form spontaneously even at extremely high Fermi levels [3] well above the conduction band minimum (CBM). (ii) In contrast to common oxides like MgO, even the pure (undoped) In$_2$O$_3$ and ZnO show considerable free-carrier densities up to the 10$^{17}$–10$^{19}$ cm$^{-3}$ range [4,5]. The source of this conductivity remains, however, a mystery: While unintentional $H$ donor impurities were implicated [6], it is now clear that the conductivity exists even when $H$ has been annealed out or is absent [5,7], (iii) In$_2$O$_3$, ZnO, and SnO$_2$ show a pronounced nonstoichiometry, with an O deficiency up to 1%, even under equilibrium growth conditions, at high-temperature [4,8]. (iv) After annealing (in reducing, metal-rich conditions), In$_2$O$_3$ and ZnO develop an apparently paradoxical coexistence of coloration, which is indicative of deep centers ($F^-$ centers [9]), and conductivity, which is indicative of shallow centers [4,5].

To date, no defect model exists that would account for the above phenomenology in TCO’s. Early models [4,8] had difficulties to estimate the formation energy of defects and remained speculative as to the source of intrinsic electron-producers. In recent years, first-principles calculations in In$_2$O$_3$ [10–12], ZnO [7,13–18], or SnO$_2$ [19] helped to clarify many of the factors involved in TCO behavior. However, the lack of systematic corrections for deficiencies in the local density or generalized gradient approximations (LDA or GGA) to density functional theory (DFT), as well as for spurious energy contributions due to the finite-supercell formalism, lead to a large spread in the predictions for ZnO. For example, published results for the formation energy $\Delta H$ of the neutral zinc vacancy $V_{Zn}^0$ under Zn-poor conditions ranges from 1.5 [13] to 7.5 eV [14], while $\Delta H$ of the neutral oxygen vacancy $V_O^0$ under O-poor conditions ranges from (unphysically negative) $-0.8$ [15] to 3.9 eV [17].

In the present Letter, we develop a predictive defect model for In$_2$O$_3$ and ZnO from first-principles calculations, applying systematic corrections to LDA/GGA and supercell errors. In order to validate our theoretical results against available experimental data, we compute experimentally accessible quantities, i.e., defect and carrier densities as shown in Figs. 1 and 2 (whereas the commonly computed defect formation energies are rarely measured). We find the following: (i) “electron killers,” such as the O interstitial $O_i$ in In$_2$O$_3$ and the $V_{Zn}$ in ZnO, are remarkably unstable in In$_2$O$_3$ and ZnO, so high concentrations ($\sim$10$^{23}$ cm$^{-3}$) of electron-producing Sn or Al donors remain uncompensated, in contrast to the case in main group oxides MgO and CaO [20]. (ii) Intrinsic electron producers, i.e., $V_O$ and the cation interstitials In$_i$/Zn$_i$, do not lead to more than 10$^{14}$ cm$^{-3}$ free electrons at room temperature in equilibrium, and, thus, fail to explain the observed electron densities of $\sim$10$^{18}$ cm$^{-3}$. (iii) The nonstoichiometry up to 1% is due to abundant O vacancies, and not due to cation interstitials. (iv) The paradoxical coexistence of coloration and conductivity can be explained by a metastable conductive state of $V_O$ in both materials. While the O vacancy causes optical $F$-center absorption similar as in MgO, the $V_O$ defect state becomes—unlike in MgO—resonant inside the conduction band after photoexcitation, leading to persistent photocconductivity (PPC), which may account for the observed free-carrier densities in pure In$_2$O$_3$ and ZnO.

Methods.—We use the pseudopotential-momentum-space-formalism [21], with the GGA exchange correlation of Ref. [22], as implemented in the VASP code [23]. Defects are calculated in 72-atom supercells of wurtzite ZnO, and 80-atom supercells of bixbyte In$_2$O$_3$, at their respective...
GGA equilibrium lattice constant. The GGA band gaps are 0.94 eV in In$_2$O$_3$ and 0.73 eV in ZnO, much smaller than the experimental gaps 3.50 and 3.45 eV. We correct both GGA errors and finite-supercell errors according to the scheme described in Ref. [24] (see also below). We calculate the defect formation energy as

$$\Delta H_D(E_F, \mu) = [E_D - E_H] + qE_F + \sum \pm (\mu^0 + \Delta \mu),$$

where $E_D$ and $E_H$ are host + defect and host-only supercell energies, respectively, and $q$ is the defect charge state. The chemical potentials $\Delta \mu_{In^+}$, $\Delta \mu_{Zn^2+}$, $\Delta \mu_{Sn^{2+}}$, $\Delta \mu_{Al^3+}$, and $\Delta \mu_O$ for atoms added to (−) or removed from (+) the lattice are taken with respect to the GGA energy $\mu^0$ of the elementary metals and the O$_2$ molecule. We apply the thermodynamic host stability condition, i.e., $\Delta \mu_{In^+} + 3\Delta \mu_{O} = \Delta H_f(\text{In}_2\text{O}_3) = -8.12$ eV and $\Delta \mu_{Zn^2+} + \Delta \mu_{O} = \Delta H_f(\text{ZnO}) = -2.93$ eV, where $\Delta H_f$ is the calculated oxide formation enthalpy. For $\Delta \mu_{Sn^{2+}}$ and $\Delta \mu_{Al^3+}$, we further consider the bounds imposed by formation of SnO$_2$ and Al$_2$O$_3$. The dependence of $\Delta \mu_O$ on the temperature and oxygen partial pressure $p(O_2)$ is calculated according to Ref. [20]. We also take into account the temperature dependence [4,25] of the band gap $E_F$ by which the energy $E_C$ of the CBM is lowered. Having computed the corrected $\Delta H_D$'s, we perform thermodynamic simulations to obtain the defect and carrier densities, as well as the equilibrium Fermi energy $E_F$ [26].

Intrinsic defects do not lead to shallow and abundant donors.—The calculated equilibrium concentrations of intrinsic defects and free carriers are shown in Fig. 1. While O vacancies form at concentrations up to the $10^{20}$ cm$^{-3}$ range, cation interstitials are scarce, staying below $10^7$ cm$^{-3}$ in In$_2$O$_3$ and below $10^{14}$ cm$^{-3}$ in ZnO. The calculated room-temperature electron density is very low, i.e., $n(\text{RT}) \approx 10^7$ cm$^{-3}$ in In$_2$O$_3$ and $n(\text{RT}) \approx 10^{14}$ cm$^{-3}$ in ZnO. The reason for the moderate carrier density is that $V_O$ has a deep level which does not create free carriers, and that the formation energy of the cation interstitials is high (see Fig. 3). The low formation energy of the O vacancy originates from the formation of stable metal-metal bonds [16] between the In or Zn neighbors of $V_O$, which leads to strong atomic-relaxation and energy-lowering of the doubly occupied $d^2$ gap level. In contrast, the electrons introduced by In$_i$, or Zn$_i$ occupy a shallow level located at high energies close to the CBM, which leads to a large formation energy when $E_F$ is high in the gap (Fig. 3). We conclude that intrinsic defects do not cause n-type conductivity under equilibrium conditions (see, however, below how the metastable state of $V_O$ could lead to conductivity). The earlier suggestion of the Zn interstitial being the origin of conductivity in undoped ZnO [6,7,27] is in conflict with its large formation energy found here as well as in some of the previous calculations [14]. Indeed, Zn$_i$ was experimentally observed only after its artificial generation by high-energy electron-irradiation [27,28]. Also, the suggestion [7] that a (Zn$_i$ – N$_O$)$^+$ complex could cause $n$ conductivity is, in fact, in contradiction with the theoretical results of the same work, which predicts pinning of the Fermi level at $E_F + 1$ eV [7], and, hence, insulating behavior [29]. Thus, a (Zn$_i$ – N$_O$)$^+$ complex may form as a compensating donor in insulating ZnO, but it can not cause conductivity.

Even though the $(2 + /0)$ donor level of $V_O$ is deep at room temperature, we find in In$_2$O$_3$ an interesting deep-shallow transition due to the pronounced reduction of the band gap with temperature, bringing the donor level (cp. Fig. 3) closer to the CBM. This transition explains the observed drastic increase of in situ measured conduc-

![Figure 1](http://example.com/fig1.png)  
**FIG. 1.** Simulation of equilibrium grown pure In$_2$O$_3$ and ZnO: Calculated defect and electron ($n$) densities, as a function of growth temperature $T_g$. Top: Metal-rich conditions; $\Delta \mu_{In^+} = \Delta \mu_{Zn^2+} = 0$. Bottom: O-rich conditions; $p(O_2)$ = 1 atm. The electron density $n_{\text{eff}}(T_g)$ due to thermal excitation across the band gap (without defects) is shown for comparison (dotted line).

![Figure 2](http://example.com/fig2.png)  
**FIG. 2** (color online). Compensation of electron doped In$_{1.98}$Sn$_{0.02}$O$_3$ and Zn$_{0.98}$Al$_{0.02}$O$_3$ during annealing ($T_a = 1073$ K) in O$_2$ atmosphere. The thermodynamic simulations are constrained to a constant doping level of 1%, e.g., $[\text{Sn}] / ([\text{In}] + [\text{Sn}]) = 0.01$. Shown are the calculated defect and carrier ($n$) densities, as a function of the oxygen partial pressure $p(O_2)$. 

045501-2
Extrinsic donors do lead to degenerate doping.—As shown in Fig. 3, Sn\textsubscript{In} in In\textsubscript{2}O\textsubscript{3} and Al\textsubscript{Zn} in ZnO are shallow donors and have low $\Delta H$ even when $E_F$ is near the CBM, and, therefore, can produce free carriers. We find that in In\textsubscript{2}O\textsubscript{3}, even the large donor concentration of 1% remains uncompensated for a wide range of growth conditions (O\textsubscript{2} partial pressures); i.e., the electron density equals the concentration of Sn donors (Fig. 2). In quantitative agreement with annealing experiments ($T_a = 1073$ K) [30], compensation occurs only at rather oxygen-rich conditions above $p(O_2) \approx 10^{-6}$ atm, and leads to the observed [30] $p(O_2)^{-1/8}$ dependence of the electron density (Fig. 2). As seen in Fig. 3, the dominating electron killer $O_1$ is ineffective under metal-rich conditions [low $p(O_2)$], having a formation energy still above 1 eV even when $E_F$ is at the CBM. In ZnO, a donor concentration of 1% remains also uncompensated under extreme metal-rich conditions, as shown in Fig. 2. With increasing $p(O_2)$, however, the electron density falls short of the donor density, as a result of compensation by the $V_{Zn}$ acceptor. The electron density follows a $p(O_2)^{-1/4}$ dependence above around $p(O_2) \approx 10^{-6}$ atm (Fig. 2), where strong compensation leads to nondegenerate electron densities and to a Fermi level inside the band gap ($E_F < E_C$). Note that vanishing formation energies of $V_{Zn}$ when $E_F$ approaches the CBM, as found in Refs. [13,15], are inconsistent with degenerate doping levels achievable in ZnO.

In nominally undoped ZnO with, e.g., $n = 10^{17}$ cm$^{-3}$ (growth at $T_g = 1423$ K) [31], we obtain, in quantitative agreement with the positron annihilation experiments of Ref. [31], only small concentrations of compensating $V_{Zn}$ defects, $c(V_{Zn}) = 2 \times 10^{15}$ cm$^{-3}$ at $p(O_2) = 1$ atm.

Large equilibrium oxygen deficiency.—In\textsubscript{2}O\textsubscript{3} and ZnO are characterized by a large (∼1%) O-deficient nonstoichiometry [4,8], which, historically, has been attributed to either the O vacancy, or the cation interstitial. Our calculations show that $V_O$, not the cation interstitial is the by far most abundant point defect in equilibrium grown In\textsubscript{2}O\textsubscript{3} and ZnO (Fig. 1). In In\textsubscript{2}O\textsubscript{3}, the maximal calculated O vacancy concentration is $2 \times 10^{20}$ cm$^{-3}$ or 0.4% of the O lattice sites at $T = 1673$ K (Fig. 1, metal rich), agreeing well with the maximal O deficiency of 1% found in Ref. [4] from thermogravimetric analysis at this temperature. In ZnO, the calculated concentration of $V_O$ is $4 \times 10^{19}$ cm$^{-3}$ (0.1%) at $T = 1373$ K (Fig. 1) under Zn-rich growth, where we chose conditions comparable with the coloration experiments in Ref. [5]. This maximal $V_O$ concentration is consistent with the range of O deficiency of high-temperature equilibrium grown ZnO [8], and with $c(V_O) = 10^{17}$ cm$^{-3}$ determined by positron annihilation spectroscopy in chemical vapor transport grown ZnO [32].

Excited O vacancies can lead to (persistent) photocurrent.—The possibility that $V_O$ could cause PPC in ZnO was surmised already in Ref. [14], and we recently developed a detailed model for PPC due to anion vacancies [16], finding that the ground state of $V_O^0$ has a deep, non-conductive $a_1^0$ level, but that the excited vacancy has a metastable conductive state. Following photoexcitation, the emptied $a_1^0$ level moves deep into the conduction band, giving up its two electrons to a shallow, conductive state near the CBM [16]. This behavior is found here also for In\textsubscript{2}O\textsubscript{3}, but not for the main group oxides, e.g., CaO [20], where the empty $a_1^0$ stays inside the gap. The back transition into the nonconducting ground state is impeded by an energy barrier, and ambient background illumination could be sufficient to constantly regenerate the conductive state due to the large optical cross section for $F^-$center ($V_O^+$) excitation [9], explaining the residual conductivity of pure In\textsubscript{2}O\textsubscript{3} and ZnO.

Coexistence of coloration and conductivity.—The simultaneous occurrence of coloration and conductivity after metal-rich growth, found both in In\textsubscript{2}O\textsubscript{3} (gray) [4] and in ZnO (red) [5], creates an apparent paradox: On one hand, color centers are indicative of a deep level (i.e., the excitation $V_O^0 \rightarrow V_O^+ + e$ requires an energy in the visible range), but, the existence of conductivity is indicative of a shallow level. For the optical $V_O^0 \rightarrow V_O^+ + e$ and $V_O^0 \rightarrow V_O^{2+} + e$ excitations, we calculate 1.8 and 1.6 eV in In\textsubscript{2}O\textsubscript{3} (present Letter), compared to 2.8 and 2.4 eV in ZnO [16].
These absorption energies, along with the large concentrations of $V_0$ (Fig. 1), explain the gray and red coloration in reduced In$_2$O$_3$ and ZnO, as well as the recent observation of photoemission from a gap state [33] in In$_2$O$_3$. Room-temperature electron-densities in the $10^{17}$–$10^{19}$ cm$^{-3}$ range [4,5,7] would require the simultaneous presence of large quantities of another, donorlike defect with a shallow level, such as hydrogen. However, doping levels up to the $10^{18}$ cm$^{-3}$ range are also found in H-free ZnO [5,7]. Therefore, we propose that our model of $V_0$ causing PPC would resolve the apparent paradox of coexisting coloration and conductivity.

**Total-energy corrections.**—(i) We correct the band gap by a downshift $\Delta E_V$ of the valence-band maximum, determined through GGA + $U$ calculations, and by an upward shift $\Delta E_C$ of the CBM by the remaining band gap error [16,24]:

$\Delta E_V = 0.28$ eV in In$_2$O$_3$ and $\Delta E_V = 0.69$ eV in ZnO. (ii) For shallow, hostlike levels, we apply corrections to $\Delta H$ according to the respective band-edge corrections $\Delta E_C$ and $\Delta E_V$. For deep levels such as the oxygen and the cation vacancies [28], this correction is not applied. (iii) We eliminate band-filling effects due to the finite-supercell size [24], and account for the increased [34] donor formation energies due to the Moss-Burstein shift by considering the actual equilibrium electron density. (iv) The total energy of charged supernclells is corrected by a potential alignment procedure [24]. (v) The effective (i.e., screened) Madelung energy of charged defects in the jellium background is corrected to $O(L^{-3})$.

Our correction scheme yields for the $O$ vacancy in ZnO a deep $\varepsilon(2 + /0)$ transition level at $E_C = 2.2$ eV, contrasting the much shallower level at $E_C = 1.0$ eV obtained by Janotti and van de Walle [17] who used an extrapolation scheme based on LDA + $U$. Recent, more advanced hybrid-DFT results [18], however, which yield the correct band gap energy of bulk ZnO, find the $\varepsilon(2 + /0)$ level of $V_0$ to be deep at $E_C = 3.0$ eV, and confirm our interpretation [16] of the optical magnetic resonance experiments by Vlasenko and Watkins [28]. We further note that the formation energy $\Delta H(V_0) \approx 3.9$ eV in $n$-type ZnO found by Janotti and van de Walle [17] implies $V_0$ concentrations $c(V_0) \leq 10^8$ cm$^{-3}$ far below the experimentally observed level of $c(V_0) \approx 10^{17}$ cm$^{-3}$ [32]. Although self-interaction correction (SIC) of the LDA is expected to yield an accurate electronic structure, its limited implementation, i.e., the use of SIC-corrected atomic pseudopotentials in a non-SIC LDA supercell calculation [14], is not accurate enough to reproduce the deep state of the $O$ vacancy in ZnO. Thus, the rather shallow energy $E_C = 1.0$ eV of the occupied $2t^2$ level of $V_0^+$ in the pseudopotential-SIC scheme [14] is inconsistent with the experimental color-center absorption peak at 3.0 eV [5], and with the hybrid-DFT results [18].

Our present correction scheme also successfully explained the optical absorption energies in the well-studied case of the $F^-$-center (sulfur vacancy) in ZnS [35].

In conclusion, we developed from first-principles calculations and thermodynamic simulations a comprehensive and validated defect model for the TCOs In$_2$O$_3$ and ZnO.

This Letter was funded by the U.S. Department of Energy, DOE-BES and DOE-EERE, under Contract No. DE-AC36-99GO10337.

[29] From our thermodynamic simulations we find that a Fermi level within 0.5 eV from the CBM is needed during growth or annealing to warrant a net $n$ doping of $10^{17}$ cm$^{-3}$.