

CHCRUS

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

Comparative study of defect transition energy calculation methods: The case of oxygen vacancy in $In_{2}O_{3}$ and ZnO

Wan-Jian Yin, Jie Ma, Su-Huai Wei, Mowafak M. Al-Jassim, and Yanfa Yan Phys. Rev. B **86**, 045211 — Published 23 July 2012 DOI: 10.1103/PhysRevB.86.045211

Comparative study of defect transition energy calculation methods: The case of oxygen vacancy in In₂O₃ and ZnO

Wan-Jian Yin,^{1,*} Jie Ma,² Su-Huai Wei,² Mowafak M. Al-Jassim,² and Yanfa Yan¹ ¹Deptment of Physics & Astronomy, University of Toledo, 2801 Bancroft St., Toledo, OH 43606 ²National Renewable Energy Laboratory, Golden, CO 80401

ABSTRACT

Theoretical calculation of defect properties, especially transition energy levels, is typically done by first-principles density-functional theory calculation using supercells with finite size. So far, three approaches—band-filling corrections (BFC), band-edge corrections (BEC), and no corrections (NC)—have been applied to deal with the potential inaccuracy caused by the finite size. In this paper, we compare these three approaches by calculating the (0/2+) ionization energies of the oxygen vacancy (V₀) in In₂O₃ and ZnO. We find that a correction must be included whether or not the defect level is deep or shallow, especially when the defect band has a large dispersion. The BFC approach gives the best correction. The BEC approach works well in GGA calculations only for certain systems in which the bandgap underestimation is partially corrected by choosing effective band edges.

PACS number: 71.55.-I; 71.15.Mb; 88.40.H-

Corresponding author: wanjian.yin@utoledo.edu

The functionality of semiconductors and wide-bandgap oxides depends critically on the properties of point defects. First-principles methods based on density-functional theory (DFT) have been widely used to predict the behaviors of defects and to help experimental design of materials with desirable properties. In most modern first-principles DFT defect calculations, the supercell approach with periodic boundary conditions is used [1,2]. The supercell calculation can describe precisely the impurity states if the cell size is infinitely large. However, due to the current limit on the computational capacity and cost, the supercell size used in a defect calculation is unavoidably finite. The finite cell size inevitably introduces uncertainties [3,4] that should be corrected. Currently, three approaches have been proposed to deal with the uncertainties: band-filling correction (BFC), band-edge correction (BEC), and no correction (NC). The following three methods are essentially the same: mixed k-points scheme described by Wei [5], which corrects the uncertainties using the energy difference between the total energies calculated at special k points and single electron level at Γ point; the shallow-level correction method described by Van de Walle and Neugebauer [6]; and the so-called BFC method described by Lany and Zunger [7]. The BEC approach was proposed by Zhang [8], in which the correction is obtained by averaging the band-edge energy over special k points. In many cases, results with no correction are also reported [9]. Both the BFC and BEC corrections have mostly been applied for shallow defects. So far, there are still issues under debate about these correction methods: Are these corrections necessary for qualitative description of the defects? Are they also needed for deep levels? What are the differences between the different correction methods?

In this paper, we compare these three approaches by calculating the (0/2+) ionization energies of the oxygen vacancy (V₀) in In₂O₃ and ZnO. We also compare the results obtained by

generalized gradient approximation (GGA) and hybrid functional theory (HSE06). We find that in both GGA and HSE06 calculations, without correction, the V_0 could be a shallow donor state in In₂O₃. However, the V_0 level becomes deep when the correction is added. In ZnO, V_0 is found to be deep both in GGA and HSE06 calculations with and without correction. However, the transition energies have different values. Therefore, the corrections are necessary whether or not the defect level is deep or shallow, especially when the defect band has a large dispersion. The BFC approach gives the best correction. The BEC approach works well in the GGA calculation only for certain systems in which the bandgap underestimation is partially corrected. We explain how the corrections work based on the fundamental electronic structure.

The calculations were performed using DFT as implemented in the VASP code [10] using the standard frozen-core projector augmented-wave (PAW) method [11]. The cut-off energy for basis functions is 400 eV. Both GGA [12] and HSE06 [13] are used for comparative studies. The $(2\times2\times2)$ special *k* point [14] and Γ point are used in the total-energy calculations for comparison. For HSE06, the portion of the exact exchange potential has been chosen as 0.27 for In₂O₃ and 0.36 for ZnO to correct their bandgaps to match experimental values.

The formation energy for an V_0 in the charge state q could be written as

$$\Delta H_{f}(V_{O},q) = E(V_{O},q) - E(bulk) + \mu_{O} + q(E_{VBM,\Gamma} + E_{F}), \qquad (1)$$

where $E(V_0,q)$ and E(bulk) are total energies of supercells with a *q*-charged V₀ and without V₀. μ_0 , E_{VBM} , and E_F are chemical potential of oxygen, valence-band maximum (VBM) of the host, and the Fermi energy referred to VBM, respectively. For V₀, which is usually a negative U system due to large *q*-dependent structural relaxation [15], the most interesting thermodynamic transition energy level $\varepsilon_t(0/2+)$ is determined by the Fermi energy at which

$$\Delta H_{f}(V_{0},0) = \Delta H_{f}(V_{0},2+).$$
⁽²⁾

Using Eqs. (1) and (2), the transition energy level could be calculated by

$$\varepsilon_{t}(0/2+) = \left[E(V_{0},0) - E(V_{0},2+) \right] / 2 - E_{VBM},$$
(3)

For a donor state, the behavior is usually described by the ionization energy ε_i (with $\varepsilon_i(0/2+)$) and $\varepsilon_t(0/2+)$ written in abbreviated form as ε_i and ε_t in the following), which is calculated by

$$\varepsilon_{i} = E_{g} - \varepsilon_{t} = E_{CBM} - [E(V_{O}, 0) - E(V_{O}, 2+)] / 2.$$
(4)

To solve Eq. (3) or Eq. (4), one often does the supercell calculation with special k points to get good k-point convergence on E(V₀,0) and E(V₀,2+) and choose the values of VBM and conduction-band minimum (CBM), which is usually at Γ point [2], and the transition energy and ionization energy are then derived by

$$\varepsilon_{t,NC} = [E(V_0, 0) - E(V_0, 2^+)] / 2 - E_{VBM,\Gamma}$$
(5a)

$$\varepsilon_{i,NC} = E_{CBM,\Gamma} - [E(V_0, 0) - E(V_0, 2+)] / 2, \qquad (5b)$$

where index NC in $\varepsilon_{L,NC}$ and $\varepsilon_{i,NC}$ indicates no corrections. However, the special *k* point used may overestimate total energy of neutral V₀ within the finite cell. The schematic band structure diagrams of the shallow and deep donor states are shown in Fig. 1. A defect band should be flat and coverge to a single point in the limit of an infinitely large supercell, whereas it is more dispersive in a finite supercell. For the neutral V₀, the extra two electrons will occupy the defect band. In the approximations adopted in Refs. 5 and 6, the shallow level is supposed to be the perturbation to the host band and with the same dispersion as in Fig. 1(a), and the deep level is supposed to be flat within the bandgap as in Fig. 1(b). In the limit of shallow level [Fig. 1(a)], the extra two electrons will occupy the dispersive band with an average energy ($E_{D,k-av} = \sum_i \omega_i \varepsilon_{D,i}$, with $\sum_i \omega_{D,i} = 2$). Here, ω_i is the weight of the defect state at the *i*th *k*-point, and $\varepsilon_{D,i}$ is the eigenvalue of the defect band at the *i*th *k*-point. In the infinite cell approximation, these two electrons will only occupy state at the Γ point. However, in a calculation using a finite cell size, the electrons will occupy the defect levels at many *k* points, resulting in errors. The so-called BFC term of $2(E_{D,k-av} - E_{D,\Gamma})$, which is described as E_{corr} [6] or ΔE_{bf} [7], should be subtracted from the E(V₀,0), and its impact on transition energy is explicitly given as Eq. (3) in Ref. [5]. Here, we rewrite the formula for calculating the transition and ionization energy with BFC as:

$$\varepsilon_{t,BFC} = [E(V_0, 0) - E(V_0, 2+)] / 2 - E_{VBM,\Gamma} - (E_{D,k-av} - E_{D,\Gamma}), \qquad (6a)$$

$$\varepsilon_{i,BFC} = E_{CBM,\Gamma} - [E(V_0, 0) - E(V_0, 2+)] / 2 + (E_{D,k-av} - E_{D,\Gamma}).$$
(6b)



Fig. 1: (Color Online) The schematic band structure of the deep and shallow V_0^0 states. The dashed lines represent the special *k* points used in the supercell calculations. The energy position of $E_{CBM,k-av}$, $E_{CBM,\Gamma}$, $E_{D,k-av}$, and $E_{D,\Gamma}$ are also indicated.

For the BEC method, the band edge (VBM/CBM) is also determined by the special *k*-point eigenvalues ($E_{CBM,k-av} = \sum_{i} \omega_{i} \varepsilon_{CBM,i}$ and $E_{VBM,k-av} = \sum_{i} \omega_{i} \varepsilon_{VBM,i}$ with $\sum_{i} \omega_{i} = 2$), rather than the band-edge states at the Γ point ($E_{CBM,\Gamma}$). In this way, all the energy terms are kept, and the band

edge in Eq. (3) and (4) is chosen as $E_{VBM,k-av}$ and $E_{CBM,k-av}$, respectively, and the transition and ionization energies in the BEC approach are calculated by

$$\varepsilon_{t,BEC} = [E(V_0, 0) - E(V_0, 2+)] / 2 - E_{VBM,k-av}$$
(7a)

$$\varepsilon_{i,BEC} = E_{CBM,k-av} - [E(V_0,0) - E(V_0,2+)] / 2$$
(7b)



Figure 2: (Color Online) The ionization energy in (a) In_2O_3 and (b) ZnO as functions of the supercell size in different methods, namely, Γ -point only calculation (red lines) and (2×2×2) *k*-points calculation with NC by Eq. (5b) (black lines), with BFC by Eq. (6b) (blue lines), with BEC by Eq. (7b) (green lines). The results of In_2O_3 are calculated by GGA and those of ZnO are by HSE06.



Figure 3: (Color Online) The V_O transition energy levels of (a) In_2O_3 and (b) ZnO as functions of the supercell size in different methods. The black and red lines are calculated transition levels directly from Eq. (5a) with NC. The black lines are calculated on (2×2×2) *k* points and the red lines are on Γ -point only. The blue lines are calculated on (2×2×2) *k* points by Eq. (6a) with BFC. The averaged CBM (E_{CBM,k-av}) and CBM at Γ point (E_{CBM,\Gamma}) are also indicated. The results of In₂O₃ are calculated by GGA and those of ZnO are by HSE06. All the energies are referred to E_{VBM,Γ}(E_{VBM,Γ}=0).



Fig. 4: (Color Online) The band structures of (a) 80-atom In_2O_3 and (b) 96-atom ZnO supercells. Partial band structures with one V_0 inside the supercell are highlighted as red (gray) lines. Their eigenvalues are aligned by comparing the O 1*s* core level away from V_0 site with the bulk value. To clearly show the V_0 defect bands derived from conduction band, all the valence bands are labeled as black. The results of In_2O_3 are calculated by GGA. The results of ZnO are calculated by HSE06 because the V_0 band of ZnO is resonant inside the valence band in GGA calculation.

Using GGA, we have calculated the $V_O(0/2+)$ transition energy level and ionization energies of In₂O₃ defined by different correction methods above using various supercell sizes. The HSE06 is also used to check the results in the small supercell (\leq 80 atoms/supercell), and the physics discussed below does not change because the defect band structures of V_O in the 80-atom supercell calculated by GGA and HSE06 are similar. *The comparison of ionizations energies:* With an 80-atom supercell, our calculated ionization energy of V₀ in In₂O₃ using BFC ($\varepsilon_{i,BFC}$) is 0.18 eV (see Fig. 2), which is quite consistent with previous reported results using the same correction [16]. Without correction, the transition level of V₀ ($\varepsilon_{t,NC}$) is above E_{CBM,\Gamma} (see Fig. 3) and the ionization energy ($\varepsilon_{i,NC}$) is -0.64 eV, consistent with the result in Ref. [9]. With BEC, the transition energy level is referred to E_{CBM,k-av} and the calculated ionization energy $\varepsilon_{i,BEC}$ is 0.56 eV. The V₀ of In₂O₃ is predicted to be the resonant donor state without correction [9], which is obviously inconsistent with the band structure in Fig. 3(a) and localization character of V₀ state at Γ point (not shown). From above, we see that the correction is indispensible to determine the behaviors of V₀ in In₂O₃. It is interesting to see that the calculated $\varepsilon_{i,\Gamma}$ (ionization energy with total energy calculated at only Γ point) by HSE06 is 0.57 eV, comparable to 0.56 eV of $\varepsilon_{i,BEC}$ calculated by GGA. We will discuss this occurrence later.

The origin of the correction term: The transition levels $\varepsilon_{i,NC}$ calculated using (2×2×2) special *k* points and Γ point have large discrepancies, and their trends are opposite when cell size becomes large. For example, their difference is 0.82 eV for the 80-atom supercell and 0.39 eV for the 640-atom supercell. This difference is actually due to the correction term of $E_{D,k-av}-E_{D,\Gamma}$, which is about half of the V₀ defect band width and could be understood easily from Fig. 1(a). As for the V₀ in In₂O₃, the defect band is dispersive, and the band width could remain a fraction of an eV, even when the supercell contains hundreds of atoms [see Fig. 4(a)]. In this case, the correction term is essential. It was stated that such a correction was only used in shallow defect [6,7]. We observed here that even for some relatively deep defect, such as V₀ in In₂O₃, the defect band is still dispersive. The width of the defect band

should be a criterion for the correction, not the simple estimate of the deepness or shallowness of the defect level.

The difference between BFC and BEC: With an 80-stom supercell, the calculated ionization energies with BFC and BEC are 0.18 and 0.56 eV, respectively. Comparing Eqs. (6b) and (7b), the discrepancy is due to the difference between $(E_{CBM,k-av}-E_{CBM,\Gamma})$ and $(E_{D,k-av}-E_{D,\Gamma})$, which is roughly the difference between the band width of the lowest conduction band (CB) of bulk In₂O₃ and that of the V₀ defect band. As shown in Fig. 3(a), even though the V₀ band is dispersive and has similar curvature near Γ point as the bulk lowest CB, they are not exactly parallel, especially at *k* points far away from the Γ point. As a result, the band width of the lowest CB is wider than that of the V₀ band. However, in the limit of shallow defect approximation, which considers the defect band as the perturbation of the host band, or the infinitely large supercell size in which both of the band widths are close to zero, the BFC and BEC will result in the same ionization energy.

The convergence on the supercell size: When total energy is calculated using Γ -point only (in this case, $E_{CBM,k-av} = E_{CBM,\Gamma}$, $E_{VBM,k-av} = E_{VBM,\Gamma}$, $E_{D,k-av} = E_{D,\Gamma}$), the results by BFC, BEC, and NC are all the same. It is expected that the results by BFC and BEC using (2×2×2) special *k* points should converge to the result obtained in the infinitely large supercell faster than the one with no correction. As shown in Fig. 2, ionization energy for the converged result is 0.08 eV. The result by BFC at (2×2×2) special *k* points converges at the 320-atom supercell, whereas the result by only Γ converges at the 640-atom supercell. However, the result by BEC (2×2×2) special *k* points does not converge well, even at the 640-atom supercell ($\varepsilon_{i,BEC} = 0.11$ eV). The behaviors of V_0 in ZnO are a little different from that in In_2O_3 . Because the V_0 band of ZnO is resonant inside the valence band in the GGA calculation, the HSE06 functional is used in this case. Comparing to In_2O_3 , we observed:

Fast convergence on the supercell size: The convergence of $\varepsilon_{i,BFC}$, $\varepsilon_{i,NC}$, and $\varepsilon_{i,\Gamma}$ on the supercell size are quite good using a 72-atom cell. The difference between $\varepsilon_{i,BFC}$ and $\varepsilon_{i,\Gamma}$ is negligible and that between $\varepsilon_{i,BFC}$ and $\varepsilon_{i,NC}$ is small compared to that in In₂O₃. The above difference between ZnO and In₂O₃ is due to different dispersion or band width of V₀ band in ZnO and In₂O₃. As shown in Fig. 4, the V₀ band in ZnO is much flatter and narrower than that in In₂O₃.

Qualitatively correct value without correction: In In₂O₃, $\varepsilon_{t,NC}$ is above $E_{CBM,\Gamma}$, whereas in ZnO, $\varepsilon_{t,NC}$ is below $E_{CBM,\Gamma}$, because the whole V₀ band in ZnO is much below the $E_{CBM,\Gamma}$. The calculated $\varepsilon_{i,BFC}$, $\varepsilon_{i,BEC}$, $\varepsilon_{i,NC}$, and $\varepsilon_{i,\Gamma}$ using a 72-atom cell are 1.32, 2.42, 1.02, and 1.31 eV, respectively. Because the $\varepsilon_{i,NC}$ is close to $\varepsilon_{i,BFC}$ and $\varepsilon_{i,\Gamma}$, the correction term is not important for qualitatively predicting V₀ in ZnO, which is also due to the flat V₀ band. Our calculated $\varepsilon_{i,NC}$ is consistent with the reported results in Ref. [9].

Overestimation by BEC: The $\varepsilon_{i,BEC}$ is much larger since the band width of the lowest CB is much wider than that of V₀ band, and thus, the BEC is inapplicable when the defect band is flat. Actually, the BEC is often used in the GGA/LDA calculation and sometimes gives reasonable results, such as the consistency of 0.56 eV of $\varepsilon_{i,BEC}$ by GGA and 0.57 eV by HSE06. One of the reasons for the sometimes reasonable results is that in GGA/LDA calculation, the effective band gap ($E_{CBM,k-av}$ – $E_{VBM,k-av}$) is larger than the fundamental gap ($E_{CBM,\Gamma}$ – $E_{VBM,\Gamma}$) and the band gap problem of GGA/LDA is partially corrected in the BEC approach with small supercell [8]. However, in hybrid calculation, since the band gap ($E_{CBM,\Gamma}$ – $E_{VBM,\Gamma}$) has already

been recovered, the BEC results always overestimate the transition level and ionization energy in this case.

The corrections for the finite supercell may include many aspects [3,17-20] and for each aspect, various schemes have been applied [4-8]. The widely scattered data in literature for a particular defect were often attributed to the different corrections without giving further details, which could lead to contradict results in defect calculation. So far, there is no systematic study or comparison on different schemes of each aspect. In this paper, we focused on the particular issue of spurious band dispersion and compared the results obtained using various correction schemes. For a meaningful comparison, we kept all the correction methods for other aspects the same, i.e., the core level was chosen as the potential alignment and no Makov-Payne correction was applied for charged defects, because it could induce additional errors by overestimating the interaction [5, 6]. By considering the spurious defect band dispersion only, our results, which are consistent with many of previous published results, suggest that BFC was the most reliable one.

The correction for the spurious defect band dispersion should be larger when the defect band is more dispersive. For transparent conductive oxides (TCO), which are the interest in this paper, the donor state of V_0 is considered since the lowest conduction band of TCO is mainly the dispersive cation s band so that the amount of correction could be non-negligible. For defects such as V_{In} and V_{Zn} , which are derived from the less dispersive valence band, the correction is relatively small. The calculated amount of corrections for V_{In} and V_{Zn} are within 0.02 eV and 0.14 eV.

In conclusion, we have compared different approaches for correcting the errors to calculated defect properties. We find that correction must be added when special k points are used in the calculation. Otherwise, some unphysical properties such as negative ionization energy (shallow

level) may emerge as an artifact. If the cell size is too small, the results obtained by both BEC and BFC may not converge, but their results should be qualitatively correct. In GGA/LDA calculations, the BEC on a finite supercell could give reasonable results compared to the experiment, due to the bandgap correction by effective band edges. However, it overcorrects by predicting much deeper levels when the bandgap has already been corrected by advanced method such as HSE06.

The computation was partially supported by the Ohio Supercomputer Center. Y.Y. acknowledges the support from the Ohio Research Scholar Program (ORSP). Work at NREL was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308.

References

[1] *Calculations of Defects in Materials*, edited by A. Alkauskas, P. Deak, J. Neugebauer, A. Pasquarello, and C.G. Van de Walle (Wiley-VCH, 2011).

[2] S.B. Zhang and J.E. Northrup, *Phys. Rev. Lett.* 67, 2339 (1991).

[3] The finite cell corrections needing to be considered may include: (i) The k-point sampling corrections; (ii) The potential alignment corrections; (iii) Total energy corrections for charged systems; (iv) LDA/GGA bandgap corrections. To emphasize the effects by different approaches for (i), the corrections methods used in (ii), (iii), and (iv) are kept the same and are not discussed in the present paper.

[4] N.D.M. Hine, K. Frensch, W.M.C. Foulkes, and M.W. Finnis, *Phys. Rev. B* **79**, 024112 (2009).

[5] S.-H. Wei, Comp. Mater. Sci. 30, 337 (2004).

- [6] C.G. Van de Walle and J. Neugebauer, J. Appl. Phys. 95, 3851 (2004).
- [7] S. Lany and A. Zunger, Phys. Rev. B 78, 235104 (2008).
- [8] S.B. Zhang, J. Phys.: Condens. Matter 14, R881 (2002).
- [9] P. Agoston, K. Albe, R.M. Nieminen, and M.J. Puska, Phys. Rev. Lett. 103, 245501 (2009).
- [10] G. Kresse and J. Furthmuller, *Phys. Rev. B* 54, 11169 (1996); *Comput. Mater. Sci.* 6, 15 (1996).
- [11] P.E. Blochl, Phys. Rev. B 50, 17953 (1994); G. Kresse and D. Joubert, ibid. 59, 1758 (1999).
- [12] J.P. Perdew and Y. Wang, Phys. Rev. B 45, 13244 (1992).
- [13] J. Heyd, G.E. Scuseria, and M. Ernzerhof, J. Chem. Phys. 118, 8207 (2003); J. Paier, M.
- Marsman, K. Hummer, G. Kresse, I.C. Gerber, and J.G. Ángyán, J. Chem. Phys. 124, 154709 (2006).
- [14] H.J. Monkhorst and J.D. Pack, *Phys. Rev. B* 13, 5188 (1976).
- [15] The V₀ in most of oxides is a negative U system, i.e., the V₀¹⁺ is unstable. We will focus on the (0/2+) transition in this paper.
- [16] S. Limpijumnong, P. Reunchan, A. Janotti, and C.G. Van de Walle, *Phys. Rev. B* 80, 193202 (2009).
- [17] Peter A. Schultz, Phys. Rev. Lett. 84, 1942 (1999).
- [18] Y. Yan and S.-H. Wei, phys. stat. sol. (b) 245, 641 (2008).
- [19] S.-H. Wei and Yanfa Yan, in Electronic Structure Methods for the Study of Defects, edited by Peter Deak and Vera Palmer (Wiley, New York, 2010)
- [20] *Advanced Calculations of Defects in Materials*, edited by A. Alkauskas, P. Deak, J. Neugebauer, A. Pasquarello, and C. G. Van de Walle, (Wiley-VCH, 2011).