Intrinsic defects in ZnO calculated by screened exchange and hybrid density functionals

S. J. Clark,[1] J. Robertson,[2,*] S. Lany,[3] and A. Zunger[3]
[1]*Physics Department, Durham University, Durham, United Kingdom*
[2]*Engineering Department, Cambridge University, Cambridge CB2 1PZ, United Kingdom*
[3]*National Renewable Energy Laboratory, Golden, Colorado 80401, USA*
(Received 24 December 2009; revised manuscript received 5 February 2010; published 10 March 2010)

The formation energies of intrinsic defects in ZnO are calculated by a family of screened exchange and hybrid density functionals, which include different fractions of Fock exchange and range separation in the hybrids. All functionals improve on local-density methods and agree remarkably well for formation energies of neutral vacancies but show significant variations for the energy of charge transition levels in the gap. This result highlights that a correct prediction of the band gap by a functional does not guarantee a high accuracy for the defect levels. Hybrid functionals obtain the correct localization of trapped hole states at the Zn vacancy.

DOI: 10.1103/PhysRevB.81.115311 PACS number(s): 71.55.Gs, 71.15.Mb, 71.20.Nr

I. INTRODUCTION

The importance of semiconductor point defects in controlling the creation and annihilation of free carriers has long motivated a strong interest in predicting their properties. The three main quantities of interest are (i) the defect formation energy $H(\mu;E_f)$ as a function of the Fermi Energy (E_f) and chemical potential μ of its components. This formation energy sets the defect concentration attainable in the solid at a given temperature and E_F in equilibrium with a dopant source. (ii) The energy $E(q,q')$ required to ionize the center from charge state q to q'. This (donor or acceptor) transition energy sets the ability of the defect to produce carriers at a given temperature. (iii) The spatial localization of deep defect states that show polaronic localization of carriers. For example, acceptors in oxides often introduce hole states localized on single oxygen but which are wrongly described in the local-density approximation (LDA).[1,2]

These defect-related issues are of particular interest for ZnO, which is an important semiconductor widely used as a phosphor, piezoelectric, and transparent electrode for ultraviolet light emission and spintronics.[3–7] It can be easily doped *n* type, but it is difficult to dope *p* type,[5] (attributed to self-compensation of holes by intrinsic donors[8]). It is therefore important to understand the energetics of its intrinsic defects.

Accurate calculations of the above quantities $H(\mu,E_f)$, $E(q,q')$ and localization have often used density-functional-based "supercell" methods. This approach has two classes of difficulties:[9] first, unlike Green's-function methods[10] that treat a single defect in an infinite solid, supercell methods treat a defect in a periodic finite cell. While this allows us to use standard band-structure codes, when the supercell size is restricted typically to ~100 atoms, the results can be clouded by the supercell periodicity. Fortunately, these periodicity errors consisting of spurious electrostatic interactions between periodically repeated charged defects, and the filling of host bands by artificially large carrier densities can be corrected *postfacto* by well-tested procedures.[9] Omitting or approximating such corrections causes some of the scatter in existing ZnO defect calculations.

The second difficulty, the subject of this paper, involves shortcomings of the density-functional depiction of the electron-electron interactions of the electron gas, causing the well-known underestimation of the band gap. This is a large error for ZnO, with LDA giving a gap of ~0.9 eV compared to 3.44 eV, experimentally. Postdensity-functional methods can correct this band-gap error in bulk solids but it is unclear if they suffice to fix the defect levels. Indeed, defect orbitals are constructed from host states spread throughout the Brillouin zone and extend to highly excited bands.[10] Thus, a functional that correctly describes band-edge states may not suffice to correctly describe deep defects or impurity levels.

In this work, we examine the performance for defect calculations of a *class of functionals* with screened exchange (sX) interactions applied to the important test case of ZnO. The local exchange and correlation functionals underlying the LDA and generalized gradient approximation (GGA) lead to a spurious self-interaction which places occupied (unoccupied) states too high (too low) an energy. The Hartree-Fock (HF) method uses a nonlocal exchange, so it can be self-interaction free, but HF lacks interelectronic correlation, and its exchange is unrealistically long ranged due to an absence of screening. The latter shortcoming has been variously addressed by hybrid and screened exchange corrections. The screened exchange functional replaces all the LDA exchange by a Thomas-Fermi screened Coulombic exchange potential,[11,12]

$$V_{SX}(r,r') = -\sum_i \frac{\psi_i(r)e^{-k_{TF}|r-r'|}\psi_j^*(r)}{|r-r'|} \quad (1)$$

leaving the LDA correlation potential, where *i* and *j* label the electronic bands, and k_s is the inverse Thomas-Fermi screening length.

The hybrid functionals employed here include (a) Perdew, Burke, Ernzerhof (PBEh),[13] which simply replaces a fraction α of the local DFT exchange by the unscreened and nonlocal HF exchange. A value $\alpha=0.25$ yields reasonable band gaps and energies of many molecules and solids.[13,14] (b) The Heyd, Scuseria, Ernzerhof (HSE) functional[15] uses the error function to separate the exchange into long-range and short-range parts, and replaces a fraction α of the short-range GGA exchange by the respective fraction of a nonlocal Fock exchange potential,

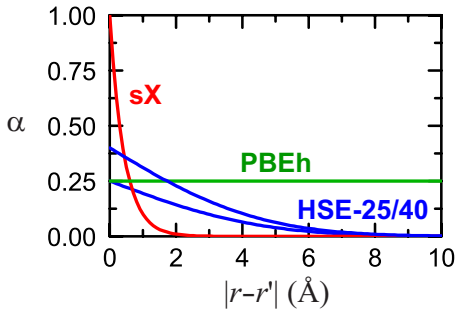

FIG. 1. (Color online) Illustrating the distance-dependent fraction α of HF exchange in different hybrid DFT functionals considered here: sX [see Eq. (1), k_{TF}=2.3 Å^{-1}]. PBEh [see Eq. (2), α =0.25, μ=0] and HSE-25/40 (α=0.25 or 0.40, μ=0.2 Å^{-1}),

$$V_{\text{hyb}}(r,r') = -\alpha \sum_i \frac{\psi_i(r)\text{erfc}(\mu|r-r'|)\psi_j^*(r)}{|r-r'|}. \qquad (2)$$

A range-separation parameter μ=0.2 Å^{-1} is chosen as a reasonable compromise.[15] The PBEh functional corresponds to limiting case of μ=0 in Eq. (2). Since the attenuation of HF exchange in the HSE form usually reduces band gaps, the fraction α is sometimes increased beyond the PBEh value of 0.25, and a value of ~0.40 has been found to reproduce the band gap of ZnO.[16] We use both values (HSE-25/HSE-40) to compare below the results for defects. Figure 1 shows the family of sX and hybrid DFT functionals considered here, in terms of the amplitude and decay length of HF exchange with interelectronic separation ($r-r'$). These functionals have been previously applied to study molecules, semiconductors, and their defects, particularly ZnO.[12,16–21]

Here, we apply the family of mixed HF-DFT and screened HF functionals (Fig. 1) to structural defects in ZnO, performing in all cases systematic corrections to the artificial periodicity errors so that our results adequately reflect the underlying functionals. The functionals are efficient enough so that full geometry relaxations can be carried out on realistic-sized defect supercells, and not just postprocessing of geometries found by LDA. The results are briefly compared with previous LDA-based corrections to defects in ZnO,[9,22] and recent GW supercell calculations.[23]

II. METHOD

The total energy (E_q) is calculated for a defect cell of charge q, for a perfect cell (E_H) of charge q, and a perfect cell of charge 0. This allows us to calculate the defect formation energy, H_q, as a function of the Fermi energy (ΔE_F) from the valence-band edge E_V and the relative chemical potential ($\Delta\mu$) of element α,[9]

$$H_q(E_F,\mu) = [E_q - E_H] + q(E_V + \Delta E_F) + \sum_\alpha n_\alpha(\mu_\alpha^0 + \Delta\mu_\alpha),$$

where qE_V is the change in energy of the Fermi level when charge q is added and n_α is the number of atoms of species α. This is the shift in the average electrostatic potential due to the change in charge of the system with respect to the uncharged system. The corrections for background charge, band filling are included as described in Ref. 9. The oxygen chemical potential (μ^0) is referred to that of the O_2 molecule, taken as zero, which is the O-rich limit. The O-poor limit corresponds to the Zn:ZnO equilibrium.

TABLE I. Bulk properties of wurzite ZnO in sX, compared to experiment.

	GGA	sX	Expt.
a (Å)	3.286	3.267	3.2495
c (Å)	5.299	5.245	5.2069
Direct gap (eV)	0.9	3.41	3.44
Zn 3d (eV)	−4.8	−7.0	−7.3[a]

[a]Reference 26.

The sX calculations use the CASTEP plane-wave code,[24] norm-conserving pseudopotentials, and a cutoff energy of 800 eV. The Zn pseudopotential includes the shallow Zn 3d core states. k_{TF} is determined from the valence-electron density, and for elements such as Zn with shallow core states, from s,p electrons only. The sX defect calculations use a 120-atom supercell. The internal geometry is relaxed within sX, using one special k point (1/4,1/4,1/3). The HSE and PBEh calculations use VASP,[25] a 64-atom supercell of zinc-blende (ZB) ZnO, a $2\times2\times2$ k-point mesh (excluding Γ), and projector-augmented wave pseudopotentials with 283 eV cutoff.

III. RESULTS

The results for the basic quantities discussed in Sec. I are as follows.

A. Ground-state formation enthalpies and bulk bands

Hybrid DFT and sX produce a similar, robust result. For bulk ZnO, they give a heat of formation of −3.1 to −3.3 eV (experiment: −3.63 eV) whereas GGA (PBE) gives −2.8 eV. The converged lattice parameters of ZnO are within 0.5% of experiment whereas GGA is 1% too large (Table I). Figure 2 shows the calculated sX band structure of wurzite ZnO. The calculated gap is 3.41 eV, very close to experiment: 3.44 eV.[4] Our sX value is much closer to experiment than HSE with the standard α=0.25 [2.49 eV in wurtzite (WZ) (Ref. 19)], or even the expensive GW (2.7 eV).[27] Part of the LDA band-gap error in ZnO arises from the Zn 3$d(t_{2g})$ levels lying too high, and repelling the Γ_{15} O 2p states which form the valence-band maximum. In sX, Zn 3d states now lie at −7.0 eV below the VB top (Table I), close to their experimental energy in angle-resolved photoemission.[26]

B. Defect enthalpies

For the charge-neutral oxygen vacancy V_O^0 (in O-poor conditions), the formation energy is ~0.85–1.0 eV, Fig. 3, suggesting a reasonable vacancy concentration of 10^{19} cm^{-3} at 700 °C, consistent with experiment.[28] It is interesting that the formation energy for V_O^0 is within ±0.1 eV for all hybrid

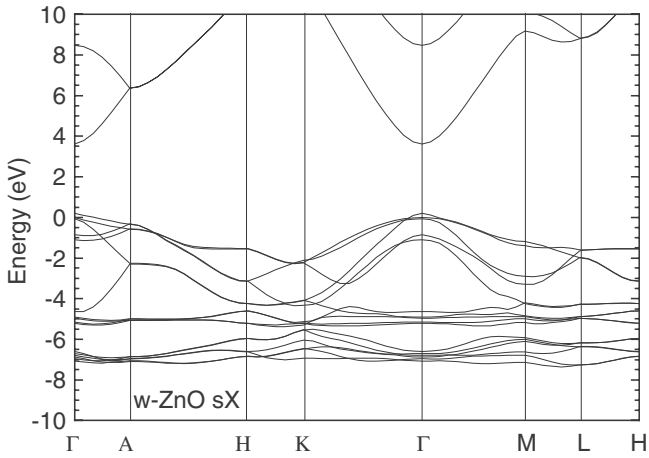

FIG. 2. The band structure of ZnO in the wurtzite structure evaluated using sX functional.

functionals (sX, PBEh, HSE-25, and HSE-40) in Table II, and also for GGA (see below). On the other hand, the formation energy for V_O^{2+} varies because of the variation in the (0/++) transition energy.

For the charge-neutral Zn vacancy (under Zn-poor conditions), HSE and sX also give consistently a formation energy of 3.9–4.1 eV, a very close range. This suggests an extremely low equilibrium concentration of (electron-compensating) metal vacancies; consistent with the propensity of ZnO to be dopable by electrons without opposition.

C. Electrical transition levels

In sX, V_O is a negative U center, as found in most previous calculations. The negative U arises from the large lattice relaxations with varying charge state around the vacancy as seen in Figs. 4(a)–4(c). The metastable (0/+) level lies at E_V+0.9 eV agreeing perfectly with one interpretation of optical measurements.[29] The Zn interstitial is shallow in sX, also agreeing with experiment.[30] Comparing the results of different functionals (Table II), we see that unlike the uniformity of the results on formation energies of charge-neutral

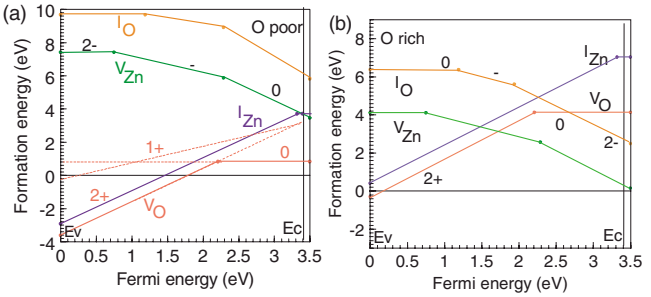

FIG. 3. (Color online) The formation energies of native defects in ZnO evaluated using the sX functional under (a) oxygen-poor and (b) oxygen-rich conditions.

defects, the electrical transition levels show a larger spread among the different functionals: in sX and HSE-40, which both fully correct the ZnO gap, the (0/++) level of V_O lies at E_V+2.3 eV. In PBEh and HSE-25, which only partially correct the band gap, the levels lie lower at E_V+2.0 and E_V +1.7 eV, respectively. Nevertheless, in all cases, the V_O donor is too deep to lead to any measurable free electrons in equilibrium.[30]

A similar variation in transition energies is also observed for the Zn vacancy V_{Zn}, where the first (second) acceptor ionization energy range from 0.7 to 1.5 eV (1.2 to 2.3 eV). Here, there are small differences even between the gap-corrected sX and HSE-40 functionals (Table II). Our comparison illustrates the variation in the predicted defect levels that can result from different choices of functional (sX, PBEh, and HSE), or hybrid parameters (HSE-25 vs HSE-40). A recent GW study[23] on V_O found that the level at E_V +1.7 eV in HSE-25 remains constant (relative to E_V) when the remaining gap error is corrected by GW, i.e., the defect level does not follow the upward shift found for HSE-40. Thus, a good reproduction of a perfect host property (e.g., band gap) by a given functional does not guarantee a good reproduction of all its defect properties.

D. Carrier localization

The LDA is known to underestimate the localization of hole states, such as the trapped hole of Al_{Si} in quartz.[1] Figure

TABLE II. Host and defect properties using the exchange functionals of Fig. 1, PBEh, HSE-25, HSE-40, and sX: heat of formation ΔH_f of bulk ZnO, formation energies of neutral O and Zn vacancies under conditions favoring their formation (O-poor and Zn-poor, respectively), band gap E_g, and charge transition levels of the vacancies relative to VBM. Results in the last rows use GGA with "postprocessor" corrections applied (GGA-C). All numbers in electron volts.

	ΔH_f	$\Delta H(V_O^0)$ (O poor)	$\Delta H(V_{Zn}^0)$ (Zn poor)	E_g	V_O $\varepsilon(2+/0)$	V_{Zn} $\varepsilon(0/1-)$	V_{Zn} $\varepsilon(1-/2-)$
sX (WZ)	−3.31	0.85	4.1	3.41	2.20	0.7	2.3
PBEh (ZB)	−3.08	0.94	3.91	3.02	2.02	1.15	1.55
HSE25(ZB)	−3.07	0.96	3.87	2.34	1.67	0.79	1.20
HSE40(ZB)	−3.20	1.01	3.92	3.46	2.34	1.53	1.94
GGA-C[a]		4			2.2		
GGA-C[b]	−2.93	0.83	2.37	3.44	1.30	0.91	1.48

[a]Reference 22.
[b]Reference 9.

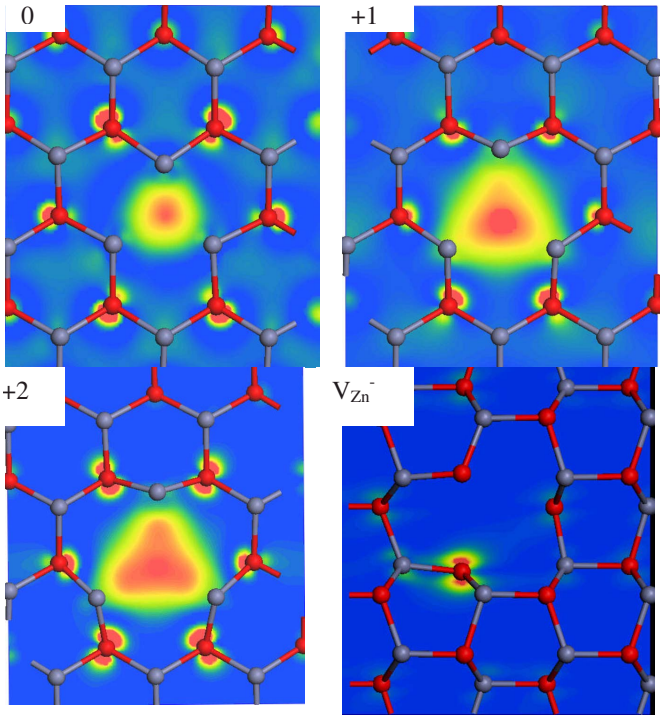

FIG. 4. (Color online) (a), (b), and (c) charge density of the oxygen vacancy for its neutral, +1 and +2 charge states, respectively. Note relaxation of Zn ions. (d) Defect state of the Zn vacancy is a *p* state localized on only *one* of the adjacent oxygen atoms.

4(d) shows the calculated sX wave function of the trapped hole state of V_{Zn}^{-}. It is localized on only one of the four oxygen neighbors. This is consistent with its spin-resonance signature.[31,32] Even LDA+*U* does not localize the hole correctly.[22] On the other hand, the wave functions of V_O localize symmetrically over all four Zn neighbors, see Figs. 4(a)–4(c).

E. Assessment of previous LDA-corrected calculations

The significant computational cost of post-LDA methods such as sX, HSE, or GW makes them useful benchmarks of lower cost LDA-based methods, applying them to postprocessor corrections. We will refer to these as "GGA-C" (C=Corrected). Examples include the LDA+*U* method used by Janotti and Van de Walle[22] to partly correct the band structure, where an empirical energy *U* repels Zn 3*d* states downward, and so partly opens up the gap. Lany[9] used instead a modest *U* (=6 eV) to shift E_V down by 0.7 eV, accounting for the remaining gap error by shifting upward the conduction band minimum (CBM). These two types of LDA-C calculations differ also in how the defect levels track the host band edges. Whereas a general conclusion cannot yet be drawn on current post-LDA calculations, we note the following: (i) Fig. 5 and Table II compare our sX formation energies of V_O to those calculated by LDA-C methods. The

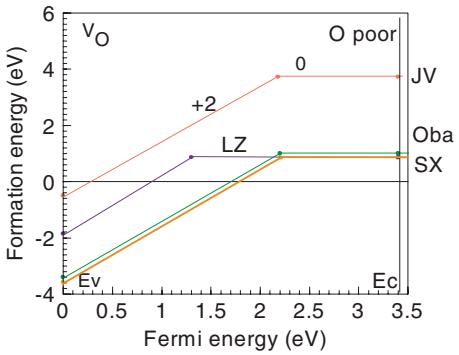

FIG. 5. (Color online) Comparison of the oxygen-vacancy formation energy under O-poor conditions as a function of Fermi energy, for the present sX study and previous studies. JV=Janotti and Van de Walle (Ref. 22), LZ=Lany and Zunger (Ref. 9), and Oba *et al.* (Ref. 16).

formation energy of V^0 of +0.85 eV in sX and +1 eV in HSE is the same as found by Lany[9] and slightly less than the 1.0 eV found by Oba.[16] It is much less than the 4 eV of Janotti.[22] The +0.85 eV formation energy corresponds to a frozen-in vacancy density of 10^{19} cm^{-3} at 700 °C, consistent with experiment.[28]

(ii) The calculated (0/++) levels of V_O divide into two groups: in HSE-40 and sX, they appear at E_V +(2.2–2.3) eV, as in Janotti.[22] In the second camp, we have HSE-25 that gives E_V+1.7 eV, GW calculations based on HSE-25 (Ref. 23) giving E_V+1.7 eV or even only 1.4 eV (GW based on GGA+*U*). This second camp is closer to the LDA+C of Lany[9] which gave a (0/++) level at E_V +1.3 eV.

IV. CONCLUSIONS

In summary, the formation energies and electrical transition levels of intrinsic defects of ZnO were calculated using a family of screened exchange and hybrid density functionals. All functionals provide large improvements over standard DFT calculations, in particular, in regard of the magnitude of the band gap, and the localization of the deep hole states of V_{Zn}. They give remarkably uniform results for the formation energies of the charge-neutral lattice vacancies of ZnO. A more differentiated picture emerges for the transition levels, where our comparison shows there remain significant differences between different choices for the functional and/or hybrid DFT parameters.

ACKNOWLEDGMENTS

The NREL portion of this work is supported by the U. S Department of Energy, Office of Science, and Office of Basic Energy Sciences under Contract No. DE-AC36–08G028308 to NREL. The Center of Inverse Design is a DOE Energy Frontier Research Center. The use of MPP capabilities at the National Energy Research Scientific Computing Center is gratefully acknowledged.

*jr@eng.cam.ac.uk

[1] G. Pacchioni, F. Frigoli, D. Ricci, and J. A. Weil, Phys. Rev. B **63**, 054102 (2000).
[2] S. Lany and A. Zunger, Phys. Rev. B **80**, 085202 (2009).
[3] K. Vanheusden, W. L. Warren, C. H. Seager, D. R. Tallant, J. A. Voigt, and B. E. Gnade, J. Appl. Phys. **79**, 7983 (1996).
[4] Ü. Özgür, Ya. I. Alivov, C. Liu, A. Teke, M. A. Reshchikov, S. Doğan, V. Avrutin, S.-J. Cho, and H. Morkoç, J. Appl. Phys. **98**, 041301 (2005).
[5] D. C. Look, B. Chaflin, Ya. I. Alivov, and S. J. Park, Phys. Status Solidi A **201**, 2203 (2004); D. C. Look and B. Claflin, Phys. Status Solidi B **241**, 624 (2004); J. L. Lyons, A. Janotti, and C. G. Van de Walle, Appl. Phys. Lett. **95**, 252105 (2009).
[6] S. J. Pearton, W. H. Heo, M. Ivill, D. P. Norton, and T. Steiner, Semicond. Sci. Technol. **19**, R59 (2004).
[7] A. Tsukazaki, A. Ohtomo, T. Kita, Y. Ohno, H. Ohno, and M. Kawasaki, Science **315**, 1388 (2007).
[8] S. B. Zhang, S. H. Wei, and A. Zunger, J. Appl. Phys. **83**, 3192 (1998).
[9] S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).
[10] U. Lindefelt and A. Zunger, Phys. Rev. B **26**, 846 (1982).
[11] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).
[12] K. Xiong, J. Robertson, M. C. Gibson, and S. J. Clark, Appl. Phys. Lett. **87**, 183505 (2005); J. Robertson, K. Xiong, and S. J. Clark, Phys. Status Solidi B **243**, 2054 (2006).
[13] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[14] J. Paier, R. Hirschl, M. Marsman, and G. Kresse, J. Chem. Phys. **122**, 234102 (2005).
[15] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, *ibid.* **125**, 224106 (2006).
[16] F. Oba, A. Togo, I. Tanaka, J. Paier, and G. Kresse, Phys. Rev. B **77**, 245202 (2008).
[17] C. H. Patterson, Phys. Rev. B **74**, 144432 (2006).
[18] F. Fuchs, J. Furthmuller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).
[19] J. Uddin and G. E. Scuseria, Phys. Rev. B **74**, 245115 (2006).
[20] P. Ágoston, K. Albe, R. M. Nieminen, and M. J. Puska, Phys. Rev. Lett. **103**, 245501 (2009).
[21] A. Carvalho, A. Alkauskas, A. Pasquarello, A. K. Tagantsev, and N. Setter, Phys. Rev. B **80**, 195205 (2009).
[22] A. Janotti and C. G. Van de Walle, Appl. Phys. Lett. **87**, 122102 (2005); Phys. Rev. B **76**, 165202 (2007).
[23] S. Lany and A. Zunger, arXiv:0910.2962 (unpublished).
[24] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Condens. Matter **14**, 2717 (2002).
[25] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[26] R. T. Girard, O. Tjernberg, G. Chiaia, S. Soderholm, U. O. Jarlsson, C. Wigren, H. Nylen, and I. Lindau, Surf. Sci. **373**, 409 (1997); K. Ozawa, K. Sawada, Y. Shirotori, and K. Edamoto, J Phys.: Condens. Matter **17**, 1271 (2005).
[27] M. Usuda, N. Hamada, T. Kotani, and M. vanSchilfgaarde, Phys. Rev. B **66**, 125101 (2002).
[28] K. I. Hagemark and P. E. Toren, J. Electrochem. Soc. **122**, 992 (1975); F. Tuomisto, V. Ranki, K. Saarinen, and D. C. Look, Phys. Rev. Lett. **91**, 205502 (2003).
[29] L. S. Vlasenko and G. D. Watkins, Phys. Rev. B **71**, 125210 (2005).
[30] F. A. Selim, M. H. Weber, D. Solodovnikov, and K. G. Lynn, Phys. Rev. Lett. **99**, 085502 (2007).
[31] D. Galland and A. Herve, Phys. Lett. A **33**, 1 (1970).
[32] L. S. Vlasenko and G. D. Watkins, Phys. Rev. B **72**, 035203 (2005).