Possible Approach to Overcome the Doping Asymmetry in Wideband Gap Semiconductors

Yanfa Yan,* Jingbo Li, Su-Huai Wei, and M. M. Al-Jassim

National Renewable Energy Laboratory, Golden, Colorado 80401, USA

(Received 13 November 2006; published 29 March 2007)

The asymmetry doping problem has severely hindered the potential applications of many wideband gap (WBG) materials. Here, we propose a possible approach to overcome this long-standing doping asymmetry problem for WBG semiconductors. Our approach is based on the reduction of the ionization energies of dopants through introduction and effective doping of mutually passivated impurity bands, which can be realized by doping the host with passive donor-acceptor complexes or isovalent impurities. Our density-functional theory calculations demonstrate that this approach provides excellent explanations for the *n*-type doping of diamond and *p*-type doping of ZnO, which could not be understood by previous theories. In principle, this approach can be applied to any WBG semiconductors and therefore will open a broad vista for the application of WBG materials.

DOI: 10.1103/PhysRevLett.98.135506 PACS numbers: 61.72.Bb, 61.72.Ji, 71.55.Cn, 71.55.Gs

Wideband gap (WBG) semiconductors, such as diamond and ZnO, are essential materials for making short-wavelength lasers and light-emitting diodes. Unfortunately, because WBG semiconductors either have a low valence-band maximum (VBM) or a high conduction-band minimum (CBM), they experience the doping asymmetry problem; i.e., they can be easily doped either *p*-type or *n*-type, but not both [1,2]. For example, diamond can be doped relatively easily *p*-type, but not *n*-type [3–7]. On the other hand, ZnO can easily be made high-quality *n*-type, but not *p*-type [8,9]. These doping asymmetry problems have been the major obstacles for potential applications of many WBG materials, even though they hold excellent properties [3,4,10–12].

In general, the doping asymmetry may be caused by three main reasons: (a) the desirable dopants have limited solubility; (b) the desirable dopants have sufficient solubility, but they produce deep levels, which are not ionized at working temperature; and (c) spontaneous formation of compensating defects. Great efforts have been devoted to overcome the doping asymmetric problem in WBG semiconductors [13–17]. The codoping concept has especially received extensive attention. The original codoping concept suggests that the solubility of desirable dopants can be enhanced through the Coulomb coupling of the donor-acceptor pairs and that the defect levels can be reduced through the donors and acceptors level repulsion [13,14]. However, further study found that although codoping may increase the dopant concentration, it fails to reduce the defect transition energy levels because the donor and acceptor levels have different wave function characters and symmetry; thus, the level repulsion between them is very weak [18].

However, experimental studies have show that in some cases, codoping can indeed drastically reduce the defect levels in some WBG materials. For example, *n*-type doping in diamond is extremely difficult because the donor levels are usually 0.6 eV or deeper below the CBM. However, through codoping of B with deuterium, *n*-type diamond has been realized with an activation energy of about 0.2–0.3 eV [5–7]. Another example is *p*-type doping in ZnO—*p*-type doping has been realized by codoping Ga with N [8,9], and the measured activation energy is less than 150 meV [19,20]. Subsequent theoretical study on these systems has failed to explain the experimental results through the conventional codoping concept. For example, theoretical calculations find that (B,H) complexes are deep donors or even acceptors [21,22]. The calculated acceptor level of isolated $(2N_O + Ga_{Zn})$ in ZnO is also deep, about 0.4 eV above the VBM [23,24]. Therefore, it has been quite puzzling about how to reconcile the discrepancies between the experimental observations and current theoretical understandings, and what is the real mechanism of codoping in overcoming the doping polarity problem in WBG semiconductors.

In this Letter, we propose that the real mechanism for the above experiments is the effective doping of the impurity bands induced by passive donor-acceptor complexes, rather than the traditional codoping concept. Based on first-principles calculations, we found that for (B,H)-doped diamonds, the passive stoichiometric (B + H) complexes create fully unoccupied bands, about 1.0 eV below the CBM of the pure diamond. Further doping of excess H atoms will bind to the (B + H) complexes and effectively dope the unoccupied impurity bands, giving two transition energies, 0.2 eV and 0.3 eV. For (Ga,N)-doped ZnO, the passive (Ga + N) complexes create fully occupied impurity bands at about 0.2 eV above the VBM of the pure ZnO. Further doping of excess N atoms will bind to the (Ga + N) complexes and effectively dope the fully occupied impurity bands, which give ionization energies of 0.1 and 0.2 eV. Thus, for the first time, our concept explains well the enhanced solubility, lowered ionization energies, and reduced intrinsic defect-compensation for the puzzling results of the doping experiments above. Based on these results, we therefore propose a possible approach to over-

 © 2007 The American Physical Society

come the doping polarity problem for WBG semiconductors, i.e., effective doping of the impurity bands created by passive donor-acceptor complexes or isovalent dopants. Our concept agrees well with the observation by Kalish *et al.* [6] and theoretical study by Nishimatsu *et al.* [14] who suggested that impurity bands could be playing a role in codoped diamond.

We performed the calculations using the density-functional theory, as implemented in the VASP codes [25], with general gradient approximation [26], and ultrasoft pseudopotentials [27]. The energy cutoff for the plane wave expansion was 400 eV. For B and H-doped diamond, a 64-atom host supercell is used. For Ga and N-doped ZnO, a 128-atom host supercell is used. Zn $3d$ and Ga $3d$ were treated as valence states. In all calculations, all the atoms are allowed to relax until the Hellmann-Feynman forces acting on them become less than 0.01 eV/Å.

Following the conventional procedure, the ionization energy of a donor with respect to the CBM is calculated by [28]

$$\varepsilon(0/q) = [E(\alpha, q) - (E(\alpha, 0) - q\varepsilon_D^k(0))] - q[\varepsilon_D^\Gamma(0) - \varepsilon_{\mathrm{CBM}}^\Gamma(\mathrm{host})]. \quad (1)$$

The ionization energy of an acceptor with respect to the VBM is calculated by

$$\varepsilon(0/q) = [E(\alpha, q) - (E(\alpha, 0) - q\varepsilon_D^k(0))] - q[\varepsilon_D^\Gamma(0) - \varepsilon_{\mathrm{VBM}}^\Gamma(\mathrm{host})]. \quad (2)$$

where $E(\alpha, q)$ and $E(\alpha, 0)$ are the total energies of the supercell at charge state q or neutral for defect α; $\varepsilon_D^k(0)$ and $\varepsilon_D^\Gamma(0)$ are the defect levels at the special k-points (averaged) and at Γ-point, respectively; q is the defect charge state; and $\varepsilon_{\mathrm{VBM}}^\Gamma$ (host) and $\varepsilon_{\mathrm{CBM}}^\Gamma$ (host) are the VBM and CBM energy, respectively, of the host supercell at Γ-point. The first term on the righthand side of Eqs. (1) and (2) determines the U energy parameter (including both the Coulomb contribution and the atomic relaxation contribution) of the charged defects calculated at the special k-points, which is the extra cost of energy after moving $(-q)$ charge from the VBM or CBM of the host to the neutral defect level. The second term gives the single-electron defect level at the Γ-point.

The formation energy of a charged defect is given by

$$\Delta H_f(\alpha, q) = \Delta H_f(\alpha, 0) + \varepsilon(0/q) + qE_F, \quad (3)$$

where $\Delta H_f(\alpha, 0)$ is the formation energy of the charge-neutral defect and E_F is the Fermi level with respect to the VBM for acceptors ($q < 0$) or CBM for donors ($q > 0$). The chemical potentials for H (μ_H) and N (μ_N) are derived from gaseous H_2 and N_2. We will see later that all the terms regarding the host should be modified accordingly in our new approach.

We first explain how and why our new concept can explain the experimental results of n-type doping by deuteration of B-doped diamonds. It is reported that the deuteration of B-doped diamond undergoes two clear steps [6,7]. The first step is the passivation of B acceptors by deuterium. The second step is the excess deuterium doping that leads to the formation of shallow donors. The experiments suggest strongly that (B,D) complexes are responsible for the shallow donors; here, D indicates deuterium. In our calculation, we use H for deuterium. Our calculation shows that the ionization energy level for an isolated H in diamond is about 2.8 eV below the CBM, which is consistent with the calculated results reported by others [29]. Isolated B + 2H complexes in diamond have also been found theoretically to be deep donors [22]. Our calculations reveal that the passive (B + H) complexes generate fully unoccupied impurity bands, which lie about 1.0 eV below the host CBM. An isolated H atom in diamond has two low-energy sites: bond-center (C-H-C) or antibond (C-C-H) sites. When B atoms are available in diamond, H atoms preferentially bond to B atoms because in their mutual presence, B atoms are negatively charged and H atoms are positively charged. The energy of the bond-center configuration is lower than the antibond configuration because an H^+ ion prefers to sit at a high electron-density site. Figure 1 shows the calculated total density of states (DOS) for pure diamond host (curve marked by D) and a supercell containing a (B + H) complex [curve marked by D(B,H)], with the B-H-C configuration. It reveals clearly that the formation of a passive (B + H) complex does not change the basic electronic structure, but only generates an unoccupied impurity band below the CBM. Our results, therefore, suggest that the first step of the deuteration of B-doped diamonds is to passivate the B acceptors and create the fully unoccupied impurity bands below the CBM.

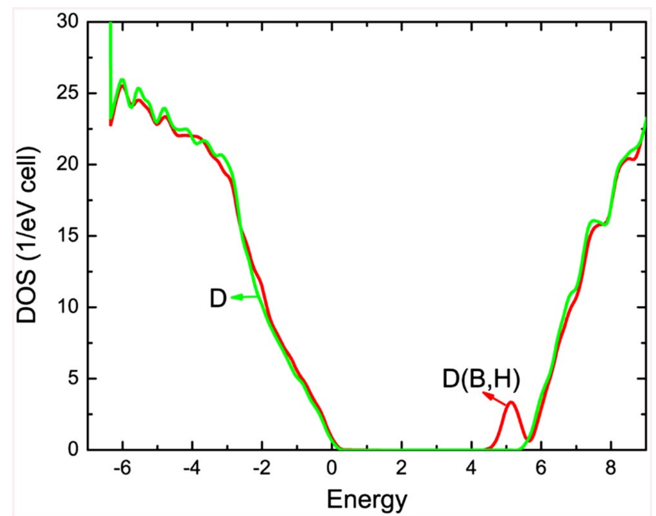

FIG. 1 (color online). Calculated total DOS of supercells with pure diamond (curve marked by D) and diamond containing one passive (B + H) complex [curve marked by D(B,H)].

When excess deuterium or H atoms are available after the first step, they will start to dope the passivated system; i.e., they effectively dope the new host with the unoccupied impurity band rather than the original conduction band. Thus, in calculating the ionization energy, the term $\varepsilon^{\Gamma}_{\mathrm{CBM}}$ (host) in Eq. (1) should now be replaced by the impurity-band minimum (IBM), $\varepsilon^{\Gamma}_{\mathrm{IBM}}$. In other words, the transition now occurs between the H defect levels and the unoccupied impurity bands, rather than the original conduction bands. As a result, the transition energy can be reduced dramatically.

For H doping in the (B + H)-passivated diamonds, the excess H atoms bind to the (B + H) complexes, forming (H-B-H) triplets. For charge-neutral H atoms, the lowest energy configuration is shown in Fig. 2(a), where the excess H is at the B antibonding site. We call this configuration (H-B-H)-AB. When the excess H atom is positively charged ($q = +1$), the fully relaxed structure is shown in Fig. 2(b). We see in Fig. 2(b) that the H^+ ion at the antibonding site becomes energetically unstable, and it moves to a bond-center site with high electron density to lower the Coulomb energy. This atomic displacement results in significant bond rearrangements and a large energy lowering of the charged defect (-1.8 eV), which leads to significant reduction of the ionization energy [see Eq. (1)]. The calculated $\varepsilon(0/+)$ transition energy level is 0.3 eV below the unoccupied impurity-band edge. We have also studied a metastable (H-B-H)-BC triplet defect, where both H atoms are at the puckered B-C bond-center sites. The atomic configurations for neutral and charged defect complexes are shown in Figs. 2(c) and 2(d), respectively. This configuration is about 0.6 eV higher in energy than the (H-B-H)-AB complex due to strong H^+-H^+ Coulomb repulsion; but the calculated transition energy level is 0.2 eV, which is 0.1 eV lower than that for the (H-B-H)-AB complex due to less crystal-field splitting.

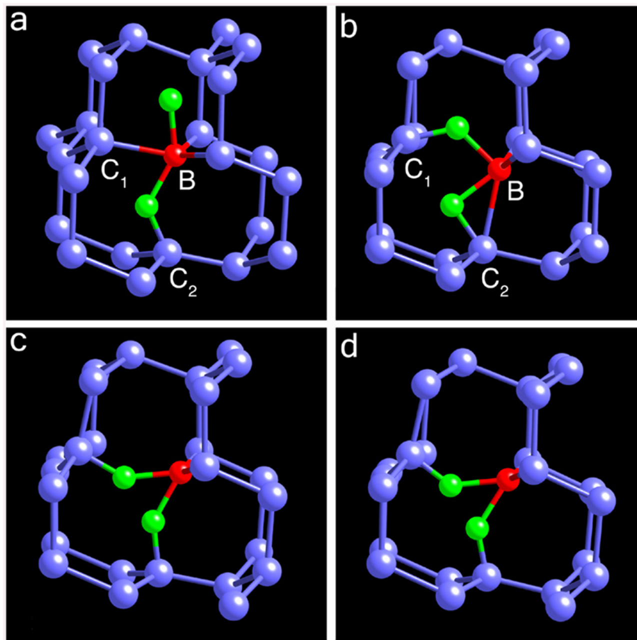

FIG. 2 (color online). (a through d) Relaxed structures for B + 2H complexes in diamond with charge-neutral and +1 charged states. The balls marked by C are C atoms. The balls marked by B are B atoms. The small balls bonding to B atoms are H atoms. (a) Neutral state for complex (H-B-H)-AB; (b) +1 charged state for complex (H-B-H)-AB; (c) Neutral state for complex (H-B-H)-BC; (d) +1 charged state for complex (H-B-H)-BC.

The calculated transition energies agree very well with the experimentally measured ionization energies, suggesting that the second step of deuteration of B-doped diamond is to effectively dope the (B + H) impurity bands. This new concept, therefore, explains for the first time why (B,H) codoping can create shallow donors in diamonds. It should be noted that to form the impurity bands and have reasonable transport properties, a critical concentration threshold is needed. Furthermore, the edge of the impurity band depends on the concentration of B atoms. The higher B concentration results in a more-broadened (B+H) impurity band. Consequently, the ionization energy will be reduced. This explains another experimental observation; i.e., diamonds with a higher B concentration exhibit shallower donor levels.

We now apply our concept to p-type doping in ZnO. So far, Ga and N codoping has produced the best p-type ZnO among other codoped ZnO [8,9]. However, the doping mechanism is not well understood. Most reliable theoretical calculations predicted that the ionization energy for N acceptors in ZnO is about 0.4 eV above the VBM [13,17,23,24], but the experimentally measured N acceptor ionization energy in p-type ZnO is much shallower, only 0.1–0.2 eV above the VBM [19,20]. The conventional codoping concept cannot explain the discrepancy because the calculated ionization level of an isolated Ga + 2N impurity is still deep, at about 0.4 eV [24].

Here, we show that to successfully use Ga and N codoping to obtain p-type ZnO, the first step is to form passive stoichiometric (Ga + N) complexes, and create a fully occupied impurity band above the VBM of ZnO. Ga and N bind together strongly in ZnO because they passivate each other. Figure 3 shows the calculated total DOS for pure ZnO host (curve marked by ZnO) and a supercell containing a (Ga + N) complex [curve marked by ZnO(Ga,N)]. It reveals clearly that the formation of a passive (Ga + N) complex does not change the basic electronic structure, but only generates an additional fully occupied band above the VBM. When excess N atoms are available, they will dope the passivated system. The transition will occur between the N defect levels and the fully occupied impurity bands, rather than the original valence bands. Thus, the term $\varepsilon^{\Gamma}_{\mathrm{VBM}}$ (host) in Eq. (2) should now be replaced by the impurity-band maximum (IBM), $\varepsilon^{\Gamma}_{\mathrm{IBM}}$.

Previous calculations suggested that for the Ga + 2N complexes, the first N occupies the first nearest-

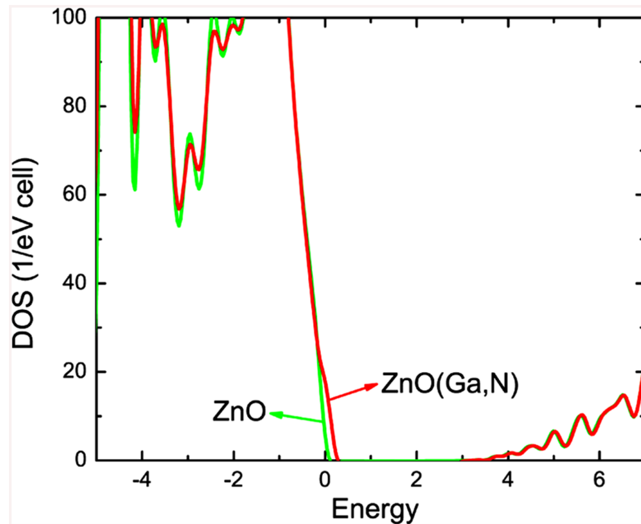

FIG. 3 (color online). Calculated total DOS of supercells with pure ZnO (curve marked by ZnO) and ZnO containing one passive (Ga + N) complex [curve marked by ZnO(Ga,N)].

neighboring O site of the Ga, which occupies at a Zn site. The second N occupies the second nearest-neighboring O site. This N atom does not bind directly to the Ga atom. We call this configuration (N-Ga–N)-A. However, our calculations reveal that the excess N atoms bind to the (Ga + N) sites, forming a (N-Ga-N)-B complex with both N atoms occupying the first nearest-neighboring O sites of the Ga atom. The B configuration is about 0.5 eV lower in energy than the A configuration. We have calculated the acceptor ionization energies for both configurations, considering effective doping of the passive (Ga + N) impurity bands. The calculated ionization energies are 0.2 and 0.1 eV for configurations *A* and *B*, respectively. Our results, therefore, are able to explain, for the first time, the puzzling experimentally measured ionization energies for N acceptors. Again, we want to point out that to form (Ga + N) impurity bands and have reasonable transport properties, heavy doping should be needed. The transition energy is also expected to be reduced in the ZnO with higher Ga concentration.

With our new concept and proposed approach, we are able to explain experimentally observed B and H codoped *n*-type diamonds and Ga and N codoped *p*-type ZnO, which could not be understood by previous theories. The physical principle behind this new concept is clear; that is, we can first create a fully passivated impurity band and then dope the impurity band. This approach can be applied, in principle, to any WBG semiconductors to overcome the doping polarity problems found in these materials. It should be pointed out that to be successful, the concentration of the defects inducing the impurity band must exceed a certain percolation limit, which is material dependent, so that reasonable transport properties can be achieved. This new approach may come at a cost of small band gap reduction caused by forming impurity bands. Fortunately, the reduction can be easily adjusted by alloying with other elements. For example, adding a small amount of Mg or Be in ZnO can easily open the band gap without changing the doping property [30,31].

S. H. W. is grateful to Professor H. Katayama-Yoshida for helpful discussions on the codoping concept. This work is supported by the U.S. DOE under Contract No. DE-AC36-99GO10337 and by NERSC for MPP time.

*Electronic mail: yanfa_yan@nrel.gov

[1] S. B. Zhang, S.-H. Wei, and A. Zunger, J. Appl. Phys. **83**, 3192 (1998).
[2] S. B. Zhang, S.-H. Wei, and A. Zunger, Physica B (Amsterdam) **273–274**, 976 (1999).
[3] S. Koizumi, K. Watanabe, M. Hasegawa, and H. Kanda, Science **292**, 1899 (2001).
[4] J. Isberg *et al.*, Science **297**, 1670 (2002).
[5] Z. Teukam *et al.*, Nat. Mater. **2**, 482 (2003).
[6] R. Kalish *et al.*, J. Appl. Phys. **96**, 7060 (2004).
[7] J. Chevallier *et al.*, Phys. Status Solidi A **201**, 2444 (2004).
[8] M. Joseph, H. Tabata, and T. Kawai, Jpn. J. Appl. Phys. **38**, L1205 (1999).
[9] D. C. Look, B. Claflin, Y. I. Alivov, and S. J. Park, Phys. Status Solidi A **201**, 2203 (2004).
[10] D. M. Bagnall *et al.*, Appl. Phys. Lett. **70**, 2230 (1997).
[11] P. Yu *et al.*, J. Cryst. Growth **184–185**, 601 (1998).
[12] M. H. Huang *et al.*, Science **292**, 1897 (2001).
[13] H. Katayama-Yoshida and T. Yamamoto, Phys. Status Solidi B **202**, 763 (1997).
[14] T. Nishimatsu, H. Katayama-Yoshida, and N. Orita, Jpn. J. Appl. Phys. **41**, 1952 (2002).
[15] Y. Yan, S. B. Zhang, and S. T. Pantelides, Phys. Rev. Lett. **86**, 5723 (2001).
[16] P. Strobel, M. Riedel, J. Ristein, and L. Ley, Nature (London) **430**, 439 (2004).
[17] J. Li, S.-H. Wei, S.-S. Li, and J.-B. Xia, Phys. Rev. B **74**, 081201 (2006).
[18] S.-H. Wei and S. B. Zhang, Phys. Rev. B **66**, 155211 (2002).
[19] A. Zuener *et al.*, Phys. Status Solidi B **234**, R7 (2002).
[20] A. Tsukazaki *et al.*, Natural Materials **4**, 42 (2004).
[21] E. B. Lombardi, A. Mainwood, and K. Osuch, Diam. Relat. Mater. **12**, 490 (2003).
[22] J. P. Goss, P. R. Briddon, S. Sque, and R. Jones, Phys. Rev. B **69**, 165215 (2004).
[23] C. H. Park, S. B. Zhang, and S. H. Wei, Phys. Rev. B **66**, 073202 (2002).
[24] L. G. Wang and A. Zunger, Phys. Rev. Lett. **90**, 256401 (2003).
[25] G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996).
[26] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
[27] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
[28] S.-H. Wei, Comput. Mater. Sci. **30**, 337 (2004).
[29] J. P. Goss *et al.*, Phys. Rev. B **65**, 115207 (2002).
[30] A. Ohtomo *et al.*, Appl. Phys. Lett. **72**, 2466 (1998).
[31] Y. R. Ryu *et al.*, Appl. Phys. Lett. **88**, 052103 (2006).