PHYSICAL REVIEW B **95**, 125309 (2017)

Theory of ionization potentials of nonmetallic solids

Yu Kumagai,[1,2,*] Keith T. Butler,[3] Aron Walsh,[3,4] and Fumiyasu Oba[1,5,6]
[1]*Materials Research Center for Element Strategy, Tokyo Institute of Technology, Yokohama 226-8503, Japan*
[2]*PRESTO, Japan Science and Technology Agency, Tokyo 113-8656, Japan*
[3]*Centre for Sustainable Chemical Technologies and Department of Chemistry, University of Bath, Bath BA2 7AY, United Kingdom*
[4]*Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom*
[5]*Laboratory for Materials and Structures, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan*
[6]*Center for Materials Research by Information Integration, National Institute for Materials Science, Tsukuba 305-0047, Japan*

(Received 3 November 2016; revised manuscript received 20 February 2017; published 14 March 2017)

Since the ionization potential (IP) is one of the fundamental quantities in a solid, ruling the physical and chemical properties and electronic device performances, many researchers have quantified the IPs using first-principles calculations of slab models recently. However, the breakdown into bulk and surface contributions has remained a contentious issue. In this study, we discuss how to decompose the IP into the bulk and surface contributions by using the macroscopic average technique. Although this procedure quantifies well-defined macroscopic dipoles and corroborates with the continuous model, it is not consistent with the physical intuition. This is because the strong charge fluctuation inside solids significantly contributes to the macroscopic dipole potential. We also discuss the possibility of an alternative splitting procedure that can be consistent with the physical intuition, and conclude that it is possible only when both bulk and surface charge density is well decomposed into a superposition of spherical charges. In the latter part, we evaluate the IPs of typical semiconductors and insulators such as Si, diamond, GaAs, GaN, ZnO, and MgO, using atomic-charge and molecular-charge approximations, in which the charge density of a solid is described as a superposition of charge density of the constituent atoms and molecules, respectively. We find that the atomic-charge approximation also known as the model-solid theory can successfully reproduce the IPs of covalent materials, but works poorly for ionic materials. On the other hand, the molecular-charge approximation, which partly takes into account the charge transfer from cations to anions, shows better predictive performance overall.

DOI: 10.1103/PhysRevB.95.125309

I. INTRODUCTION

Band-edge positions with respect to the vacuum level are fundamental physical quantities having great relevance to materials properties and applications such as catalysis and photocatalysis, electronic transport, and electron emission [1]. With x-ray photoelectron spectroscopy or Kelvin probe microscopy, one can measure an ionization potential (IP) that is the energy required for removing an electron from a surface and defined as the position of the valence band maximum (VBM) with respect to the vacuum level [1]. IPs are physically well-defined quantities depending on the surface orientations, and can be computed using first-principles calculations of slab models within the given approximation to the exchange and correlation interactions between electrons [2–5]. The electron affinity (EA) is defined as the conduction band minimum with respect to the vacuum level. IPs and EAs are also utilized to approximately discuss the Schottkey barrier heights of metal-semiconductor interfaces [6], band offsets between semiconductors [5,7], and carrier doping limits [8,9]. It has also been conjectured that the IP is composed of surface dipole potential and a surface independent bulk electronic eigenvalue [10,11]. However, researchers have thought that the latter seems impossible to obtain for a three-dimensionally periodic solid since the average electrostatic potential is arbitrary due to the conditionally convergence nature of the Coulomb summation [12]. Note that such arbitrariness does not affect the total energies of neutral systems, but special care is required for those of charged systems [13–16].

In this paper, our main purposes are twofold. First, we explore a definite way to decompose an IP into surface dependent and independent contributions. We begin with a macroscopic continuous model, and apply its theory to realistic solids in conjunction with first-principles calculations. For metallic systems, the surface dipole potential has been analyzed using the jellium model for discussing the work functions [17,18]. The potential is determined from the electron distributions in the slab model with the positive jellium background charge and vacuum. Such model calculations can reproduce the tendency of work functions experimentally observed. If the dipole potential is a well-defined quantity, it should be plausible to estimate such a dipole potential for real solids quantitatively once the spatial charge distribution is available. However, such a dipole is ill defined in a solid because of the continuous spatial fluctuation of the electron charge density. In this study, we bridge a gap between such realistic solids and the continuous model by adopting the macroscopic average technique. This introduces a well-defined macroscopic dipole and corroborates with the continuous model. However, the calculated bulk contributions of IPs are very much different from those inferred from ionization energies of atoms, molecules, and/or clusters. This is because the macroscopic dipole also includes the influence of the spatial fluctuation of bulk charge density. We also discuss the possibility of other splitting procedures that are consistent with the physical intuition, but is possible only when the bulk charge density is described as a superposition of spherical charges and they can also represent the surface charge.

---

[*] yuuukuma@gmail.com

Second, we evaluate the IPs using simple approximations. The first is the atomic-charge approximation, also known as the model-solid theory. Model-solid theory has been used for evaluating band offsets between conventional semiconductors with similar chemical bonding [19–21], but not for predicting IPs so far. As will be shown, this approximation works well for covalent materials including simple substances, but does not for ionic compounds. Therefore we also considered molecules composed of the constituent ions, in which the charge transfer between atoms is taken into account. We will see that the molecular-charge approximation improves over the atomic-charge approximation drastically. These investigations shed light on understanding of the IPs, and estimating IPs quantitatively, without using slab model calculations.

II. RESULTS AND DISCUSSION

A. Macroscopic model

For discussing the work functions of metals, many researchers have employed a continuous model [17,18]. Following this, let us start with a macroscopic but finite nonmetallic system in which the bulk region is macroscopically charge neutral and surfaces consist of infinitestimal charge dipole layers. To avoid confusion, we consider the potential energies for electrons that are consistent with electron eigenvalues, so the sign is the inverse of the conventional electrostatic potential. The sign of potential and charge follow the conventional rule and thus the electron charge is negative. The electrostatic potential at point $\boldsymbol{x}$ posed by the surface dipole layers is determined by solid angles subtended by the surfaces [22]. When the solid angles are distributed as illustrated in Figs. 1(a) and 1(b), the electrostatic potential arising from the dipoles is given as

$$\phi(\boldsymbol{x}) = \sum_{i=1}^{N} \frac{\Omega_i(\boldsymbol{x}) d_i}{4\pi\epsilon_0}, \tag{1}$$

where N is the number of surfaces, Ω_i is the solid angle in steradian unit for surface i, d_i is the dipole density, and ϵ_0 is the vacuum permittivity. We set the sign of Ω_i to be positive when it is subtended by the front of the crystal surface, in which $\sum_{i=1}^{N} \Omega_i = 0$ and $\sum_{i=1}^{N} \Omega_i = -4\pi$ at a point outside and inside the system, respectively [see Figs. 1(a) and 1(b)]. d_i is defined from inside the system, and so is usually negative because electrons are considered to move toward vacuum. As seen in Eq. (1), the electrostatic potential inside the system depends not on the system size but on the *crystal morphology*. Thus the electrostatic potential inside a solid caused by the surface dipole does not diminish but reaches the value in Eq. (1) as the system size is increasing even if the surface dipole layer is finite. In case the system is surrounded by the same surface with dipole density d, the potential energy reduces to a constant of $(-e)\phi = d/\epsilon_0$, where $e(>0)$ is the elementary charge.

We then consider the *vacuum level*. The solid angles decay with increasing distance from the system and the electrostatic potential finally reaches a plateau at the infinite distance, at which the potential is set to zero by convention. The electronic level corresponds to the *intrinsic* vacuum level, which is known not to be experimentally accessible [1]. We can instead access the vacuum level at a point just outside the surface of the system, which is called the *local* vacuum level. The distance between such a point and a surface is assumed to be large enough compared to the surface dipole dimension, but small to the surface area dimension. The IP measured by experimental techniques such as x-ray photoelectron spectroscopy or Kelvin probe microscopy is an energy required for extracting an electron from a surface to the local vacuum level near the surface [1]. We should emphasize that, only when the system is surrounded by the same surface, the electrostatic potential is zero at any point outside the system as $\sum_{i=1}^{N} \Omega_i = 0$, and therefore the local and intrinsic vacuum levels are exactly the same.

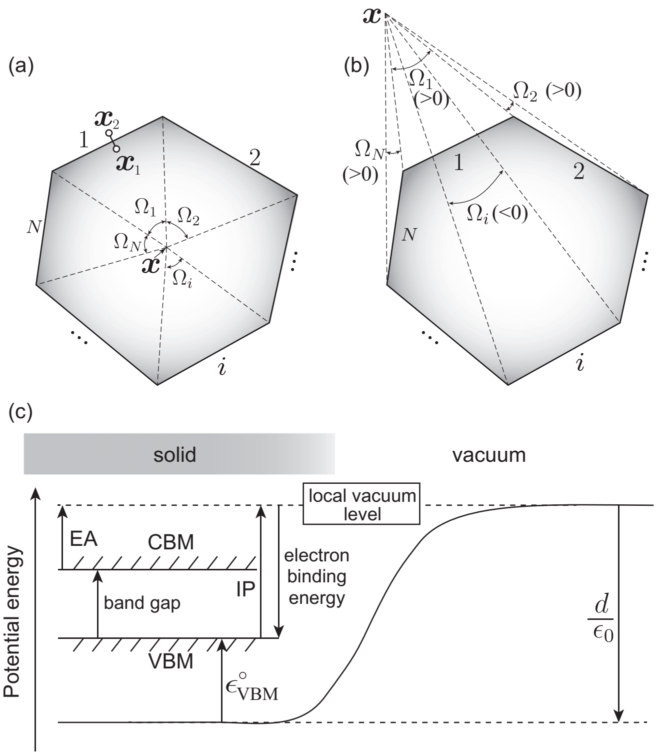

FIG. 1. (a) and (b) Schematics of solid-angle distributions subtended by the surface dipole layers at points (a) inside and (b) outside a finite solid. $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in (a) are located immediate vicinity of surface 1. The surface dipole layers are assumed to have infinitesimal thickness. The signs of the solid angles are the definition in this study. (c) Schematic of the relationship between the ionization potential (IP), electron binding energy, electron affinity (EA), band gap, bulk electronic eigenvalue ($\epsilon^{\circ}_{\mathrm{VBM}}$), and dipole potential energy (d/ϵ_0).

Next, we discuss the IPs. The energy level of the VBM in the system with respect to the intrinsic vacuum level is written as $\epsilon^{\circ}_{\mathrm{VBM}} - e\phi(\boldsymbol{x})$, where $\epsilon^{\circ}_{\mathrm{VBM}}$ is the bulk electronic eigenvalue at the VBM under the situation where the surface dipole is absent. Then, the IP at surface 1, that is, the potential energy difference from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ in Fig. 1(a), can be written as

$$\left(0 - \sum_{i=1}^{N} \frac{\Omega_i(\boldsymbol{x}_2) d_i}{4\pi\epsilon_0}\right) - \left(\epsilon^{\circ}_{\mathrm{VBM}} - \sum_{i=1}^{N} \frac{\Omega_i(\boldsymbol{x}_1) d_i}{4\pi\epsilon_0}\right)$$
$$= -\frac{d_1}{\epsilon_0} - \epsilon^{\circ}_{\mathrm{VBM}}, \tag{2}$$

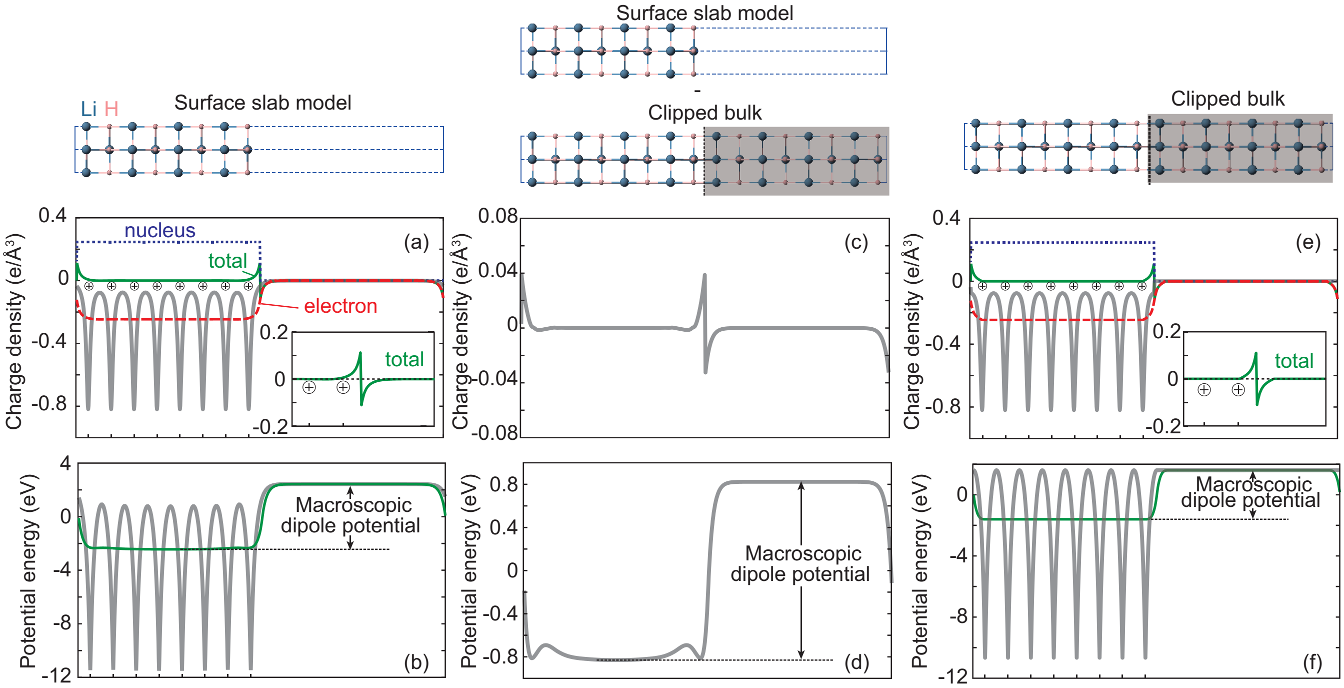

FIG. 2. Results of an all-electron calculation of the LiH (100) surface slab model. (a) Planar average of electron charge density (thick solid gray line) and macroscopic average of electron (dashed red line), nuclear (dotted blue line), and total (solid green line) charge density. (b) Planar average (thick solid gray line) and macroscopic average (solid green line) of electrostatic potential. Note that the average electrostatic potential in the entire supercell is set to zero by convention. (c) The planar average of electron leakage charge, defined as the charge difference between the surface slab model and the clipped bulk model, in which the charge density is halved on the bisecting plane and zero outside (see text for details). (d) Planar average of the potential energy calculated from the charge density in (c) via Poisson equation. (e) and (f) Same as those in (a) and (b) but for the clipped bulk model (see text for details). In (a) and (e), the insets show the enlarged total charge density.

because the solid angles subtended by surface 1 are -2π and 2π at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ as shown in Fig. 1(a), respectively, and the other angles are common. In this paper, the electron binding energy, which corresponds to the sign-reversed IP [see Fig. 1(c)], is also used, because it is intuitively understood rather than the IP, and thus $\epsilon^{\circ}_{\mathrm{VBM}} + d/\epsilon_0$ corresponds to the electron binding energy. Note that the calculated IP using the typical slab model for surface 1 corresponds to the value of Eq. (2), since the modeled surface is infinitely repeated two-dimensionally. As shown in Eq. (2), the IP can be divided into the contributions of $\epsilon^{\circ}_{\mathrm{VBM}}$ and dipole potential energy d/ϵ_0 [see Fig. 1(c)]. Therefore the change of the IP by the difference in the surface orientations reflects the change of the surface dipole potential energy as $\epsilon^{\circ}_{\mathrm{VBM}}$ is a surface independent term.

B. Realistic solids

We then discuss the contributions of $\epsilon^{\circ}_{\mathrm{VBM}}$ and d/ϵ_0 using a prototypical solid system, LiH. Its great advantage is that we can perform all-electron (AE) calculations with a plane-wave basis set at a practical computational cost owing to small nuclear charges of Li and H; the extensively used projector augmented-wave (PAW) method [23] requires additional care for the approximation to the treatment of core electrons as discussed in Appendix [24]. The calculations were performed using the Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [25] as implemented in VASP [26]. The lattice constant was fixed at the bulk value of 4.02 Å, and the atomic positions were unrelaxed from their bulk positions so as to simplify the discussion here. The plane-wave cutoff energy was set at 10 keV after checking the convergence of the VBM eigenvalue under zero average electrostatic potential condition. The slab models were constructed from sufficiently large supercells elongated along one direction perpendicular to the associated surface orientation, with the same thickness for the slab and vacuum layers. The FFT grids were set to keep the same densities as those used for bulk calculations so that numerical errors are minimized.

Figures 2(a) and 2(b) show the planar-averaged electron charge density and electrostatic potential of the LiH (100) surface, respectively. As seen, the charge density and electrostatic potential are not constant but drastically oscillate owing to the spatial distribution of electrons and nuclei. In such cases, the surface dipole is ill defined inside the bulk as opposed to the continuous model in Sec. II A. To corroborate with the continuous model, such oscillations should be removed. Since vacuum is viewed as one type of medium, surfaces can be regarded as special cases of interfaces, and therefore, the procedure used for computing the interface dipole can also be applied. Note that such macroscopic average technique has already been applied for determining the reference potential offset between solids and vacuum [27].

Baldereschi *et al.* have proposed the *macroscopic average* technique for analyzing the potential offset between different materials and the dipole charge around interfaces [28]. The macroscopic average function is calculated by $\overline{f}(z) = \frac{1}{L}\int_{z-\frac{L}{2}}^{z+\frac{L}{2}} f(z)dz$, where $f(z)$ is a target function and L the

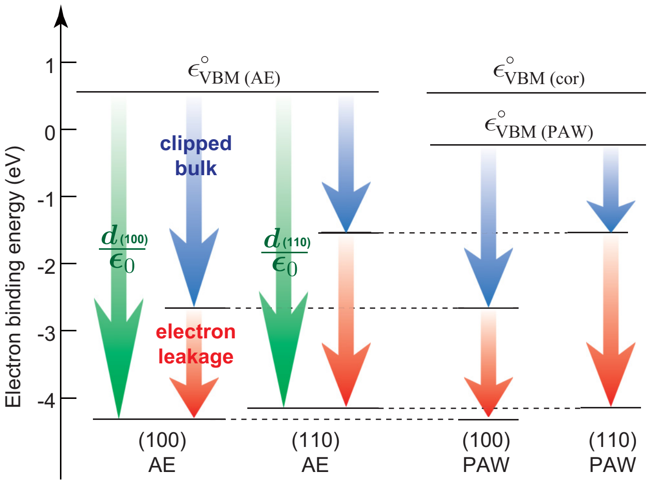

FIG. 3. Electron binding energies of LiH (100) and (110) surfaces, and their individual components [bulk electronic eigenvalue ($\epsilon^{\circ}_{\mathrm{VBM}}$), and electron leakage and clipped bulk potential] calculated using AE and PAW methods. $\epsilon^{\circ}_{\mathrm{VBM}}$ shown are obtained using the AE method ($\epsilon^{\circ}_{\mathrm{VBM(AE)}}$), PAW method ($\epsilon^{\circ}_{\mathrm{VBM(PAW)}}$), and PAW method with a correction derived by Bruneval *et al.* [24] ($\epsilon^{\circ}_{\mathrm{VBM(cor)}}$) (see Appendix for details).

period length of the unit-cell along the direction considered. This technique can filter out rapid oscillations caused by the core electrons and the nuclei. When adopting this approach for the surface model, the dipole potential is defined as the difference in the Coulomb potentials averaged over a unit cell going from inside the slab to vacuum. We can now clearly see the surface dipole density [Fig. 2(a)] and the dipole potential energy [Fig. 2(b)], which we refer to as the *macroscopic* dipole (d) and the *macroscopic* dipole potential energy (d/ϵ_0), respectively, in this study. Note that, when adopting a larger averaging window with an integral multiple of the minimum period distance instead, the dipole charge distribution is changed, but both the dipole moment and potential offset are unaltered. The calculated d/ϵ_0 from the potential offset is -4.87 eV, which can also be estimated from the macroscopic dipole density [solid green line in Fig. 2(a)] via Poisson equation.

$\epsilon^{\circ}_{\mathrm{VBM}}$ can be obtained from first-principles calculations of *bulk* systems under zero average electrostatic potential condition, as often adopted in the plane-wave codes. However, in order to know the relative position of the VBM with respect to the vacuum level, we need the calculation of the slab model to quantify d/ϵ_0. In rocksalt LiH, $\epsilon^{\circ}_{\mathrm{VBM}}$ is 0.55 eV and d/ϵ_0 is -4.87 eV as shown above. Therefore the IP of the (100) surface is $4.87 - 0.55 = 4.32$ eV within the PBE-GGA. This means that the bulk electronic eigenvalue $\epsilon^{\circ}_{\mathrm{VBM}}$ of LiH is positive but the large macroscopic dipole potential pushes the electron's energy level downward significantly, resulting in the positive IP (see Fig. 3). The macroscopic dipole potential energy through the (110) surface was also calculated and found to be -4.71 eV as shown in Fig. 3. Interestingly, the difference of the macroscopic dipole potential between the (100) and (110) surfaces (0.16 eV) is very small compared to their absolute values. This indicates that d/ϵ_0 partly includes the bulk contribution, which will be discussed in Sec. II C.

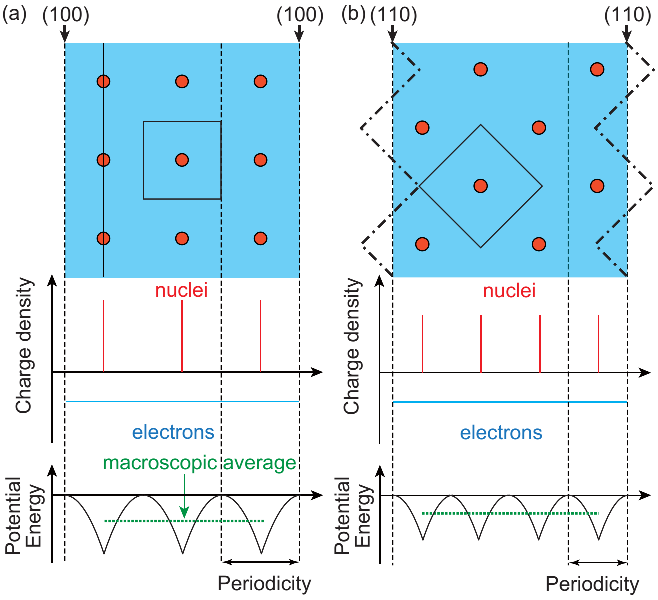

FIG. 4. (a) Schematic of a model composed of uniform electron and nuclear point charges. The electron charge density is zero outside of the model. Planar-averaged charge density, electric field, and electrostatic potential energy are also shown. (b) Same as (a) but 45° rotated. Note that the macroscopic dipole potential energy is higher than that in (a). As with the stepwise surface shown in (b) with dashed-dotted lines, other definitions for the surface are possible.

We also calculated the surface dipole charge using the procedure proposed for estimating the interface dipole by Bylander and Kleinman [29]. They defined the dipole charge as the difference between the charge density of an interface model and that of its constituent solids [29]. The interface was then set at the bisected plane. Figure 2(c) shows the planar-averaged charge density of the surface model minus that of the clipped pristine bulk at the bisected plane. The bulk charge density on the boundary is then reduced to half. It is clearly seen that the dipole layer is caused by the electron leakage at the surface. It generates the potential shift [Fig. 2(d)], but corresponds to just 34% of the macroscopic dipole potential.

What contributes to the remaining macroscopic dipole potential? Figures 2(e) and 2(f) show the macroscopic averages of the charge and electrostatic potential energy of the clipped bulk, respectively. As seen, despite the centrosymmetricity of the unit cell, the clipped bulk charge introduces a shift of the macroscopic average potential. This can be understood from the schematic of the uniform electron model shown in Fig. 4(a), which is comprised of nuclear point charges and uniform electron charge. One can see the electrostatic potential energy is zero only on the cell boundary and negative at any other area inside the system, owing to electrons that inevitably locate outside nuclei. As a consequence, the macroscopic average potential is negative inside the slab, even without spill of electrons into vacuum. This is also related to the seemingly counterintuitive positive value of $\epsilon^{\circ}_{\mathrm{VBM}}$ obtained under the condition of zero average electrostatic potential. We deal with this in more detail in Sec II C.

The partition of the surface dipole into the electron leakage and clipped bulk charge is, however, not unique

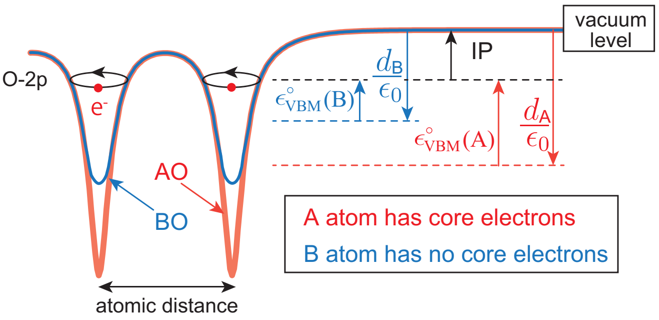

FIG. 5. Schematic of the relationship between electrostatic potential (solid lines), macroscopic dipole potential energy d/ϵ_0, and bulk electronic eigenvalues $\epsilon^\circ_{\mathrm{VBM}}$ of hypothetical oxides AO and BO. We assume that these two oxides have the same crystal structures and same IPs (see text for details). Although the IPs are the same, d/ϵ_0 and $\epsilon^\circ_{\mathrm{VBM}}$ are very different, which is attributed to the core electrons in A atoms.

since it depends on where the clipping surface is chosen [6]. Figure 4(b) shows the schematic of a surface rotated by 45° from the surface in Fig. 4(a). The surface in Fig. 4(a) has larger negative dipole potential than that in Fig. 4(b) since the potential caused by the clipped bulk charge is proportional to the square of the periodic distance in the uniform electron model as easily inferred from Poisson equation. The IPs are almost independent of the surface orientation, and such difference of the clipped bulk potential is compensated by the electron leakage potential. Indeed, as shown in Fig. 3, the LiH (100) surface has a larger clipped bulk potential than the (110) surface although the IPs are nearly the same as the former has the longer periodicity along the surface orientation. Furthermore, we can also apply other definitions for the clipped bulk surfaces such as a stepwise surface as shown in Fig. 4(b). The amount of the dipole of the electron leakage is then decreased whereas that of the clipped bulk is increased with keeping the same total dipole.

C. Physical meaning of the split of IPs from macroscopic viewpoint

The split of IPs based on the macroscopic average technique is mathematically strict and unique; $\epsilon^\circ_{\mathrm{VBM}}$ is obtained from bulk calculation and d/ϵ_0 fully contains surface dependency. However, one might wonder about the physical meaning of this split because the bulk contribution of the electron binding energies should be physically negative and similar to the ionization energies of atoms, molecules, and/or clusters. The discrepancy between our definition and physical intuition is attributed to the significant fluctuation of the bulk charge density. To explain this, Fig. 5 shows the energy diagram of hypothetical oxides AO and BO. Here are two assumptions: first, these two solids have the same IP and the same structure. Second, A ion has core electrons, whereas B ion does not, and they do not affect the valence electrons (O-2p band in our example). Therefore the valence electrons feel the same electrostatic potential and their wave functions are the same. Deduced from physical intuition, these two bulk components of IPs should have the same values. However, based on the macroscopic average technique, $\epsilon^\circ_{\mathrm{VBM}}$ values are very different from each other, and d/ϵ_0 counteracts this difference as shown in Fig. 5. This is due to the core electrons introducing a strong fluctuation of potential and so a shift of d/ϵ_0. Such core effect should be included into the bulk contribution from the aspect of physics. Thus, although the macroscopic average technique is strict for splitting IPs into bulk and surface independent components, it is not useful for discussing physical and chemical tendencies. This is also discussed via systematic calculations in Appendix.

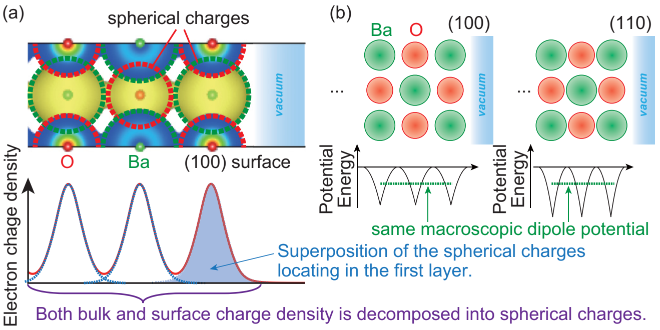

FIG. 6. (a) An illustration of the prerequisites for obtaining the bulk contribution of macroscopic dipole potential energy $d_{\mathrm{bulk}}/\epsilon_0$ (see text for details). Rocksalt BaO (100) surface is taken as an example. (b) Schematics showing that the macroscopic dipole potential calculated using the decomposed spherical charges is the same between different neutral surfaces.

The next question is whether it is possible to exclude such bulk fluctuation effect from d/ϵ_0 in a consistent manner. By splitting d/ϵ_0 into bulk contribution without surface dependence $d_{\mathrm{bulk}}/\epsilon_0$ and surface contribution $d_{\mathrm{surface}}/\epsilon_0$ from a physics viewpoint, the electron binding energy can be rewritten as

$$\left(\epsilon^\circ_{\mathrm{VBM}} + \frac{d_{\mathrm{bulk}}}{\epsilon_0}\right) + \frac{d_{\mathrm{surface}}}{\epsilon_0}. \tag{3}$$

The first parenthesis and the third term correspond to the redefined bulk and surface contributions, respectively. The IPs of unrelaxed Tasker type 1 surfaces denoted as neutral surfaces in this study are almost independent of the surface orientations, especially when surface chemical bonds are absent. Such IPs would be close to bulk components of IPs, i.e., $\epsilon^\circ_{\mathrm{VBM}} + d_{\mathrm{bulk}}/\epsilon_0$. When the charge density of both bulk and neutral surfaces is well described with a superposition of spherical charges as illustrated in Fig. 6(a), the macroscopic dipole potentials d/ϵ_0 estimated from these spherical charges are accurately predicted without surface dependencies [see Fig. 6(b)].

One possibility of realizing such a situation is the model-solid theory, in which electron charges in solids are supposed to be a superposition of atomic charges. However, when atoms condensate, chemical bonds and charge transfer emerge, and consequently, such models do not satisfy the aforementioned prerequisites. Indeed, model-solid theory cannot well predict the macroscopic dipole potential of ionic compounds as shown later. Therefore it is a future work to construct a unique and nonbiased procedure fulfilling the conditions.

D. Model-solid theory

In the previous section, we considered how to split the IPs of the realistic materials into the bulk electronic eigenvalue

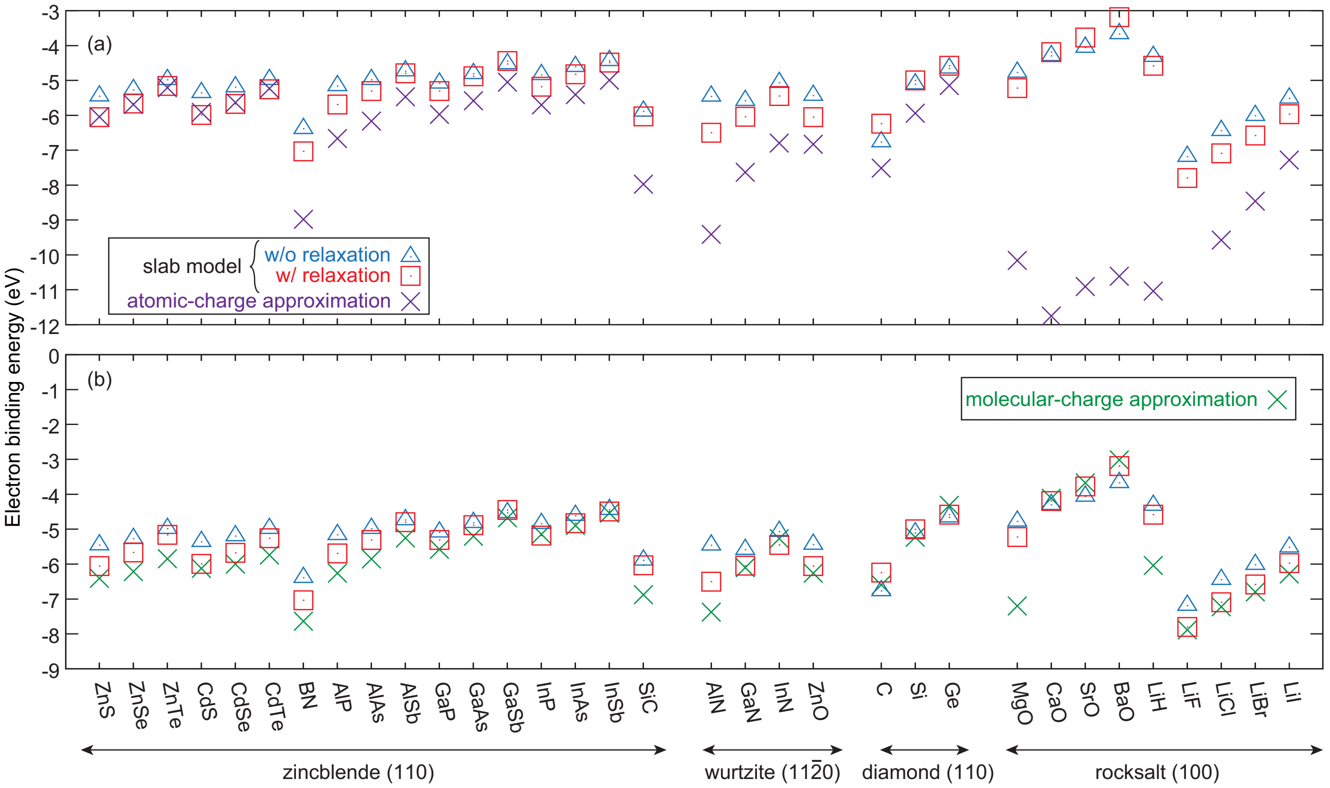

FIG. 7. Estimated electron binding energies from (a) atomic-charge and (b) molecular-charge approximations compared to those calculated using slab models with and without atomic relaxation.

$\epsilon^{\circ}_{\text{VBM}}$ and macroscopic dipole potential energy d/ϵ_0. Here, we discuss procedures to predict IPs without slab model calculations, which requires a way to approximately estimate d/ϵ_0.

A few simple methods that predict the *band offset* between semiconductors solely from bulk information have been proposed so far. For example, Baldereschi *et al.* used the *bulk* charge density of the Wigner-Seiz cell, which corresponds to the clipped charge density in this study, for estimating the band offset [28]. This works well for interfaces composed of isostructural materials with similar chemical bonding, in which the electron leakage would be mostly canceled out between two materials. However, this does not work for predicting the IPs because the electron leakage contribution to the dipole potential is significant as shown in Fig. 2(d). Furthermore, the surface boundary is ill defined in general as discussed above.

Van de Walle and Martin have proposed an alternative way to predict the band offset using atomic-charge density also known as the model-solid theory [19–21]. It is successful for predicting the band offset between the conventional semiconductors such as II-VI and III-V zincblende materials as the Baldereschi's procedure. In this study, we applied the model-solid theory to predict the IPs. Although the PAW radii affect both $\epsilon^{\circ}_{\text{VBM}}$ and d/ϵ_0, their deviations are canceled out when calculating the IPs. Therefore calculations in this subsection were performed with the PAW method (see Appendix for details).

First, we have checked how atomic relaxation in the slab models affects the IPs. As seen in Fig. 7(a) showing a comparison between the IPs with and without atomic relaxation, the IPs are mostly increased by the surface reconstruction, but the variations are not significantly large (the mean absolute variation is 0.36 eV).

We then show the calculated IPs using the atomic charge. In this approximation, the macroscopic dipole potential energy can be calculated as

$$\frac{d}{\epsilon_0} = \frac{1}{\Omega_{\text{cell}}} \sum_{i=1}^{N} \int V_i d\boldsymbol{r}, \tag{4}$$

where Ω_{cell} is the volume of the unit cell, N the number of atoms in the unit cell, and V_i the electrostatic potential caused by the ith atom, and the integral interval is infinite space [30]. Under the zero average electrostatic potential condition, $\int V_i d\boldsymbol{r}$ can be simply obtained from $-V_{\text{cell}} \cdot \overline{\phi}$, where V_{cell} is the volume of the supercell used for a calculation of ith atom and $\overline{\phi}$ the potential at the outermost point from the atom in the supercell. Note that, $\int V_i d\boldsymbol{r}$ only depends on the atomic species, so we do not have to recalculate these once obtained. Therefore this approximation allows us to estimate the IPs at great speed as long as the bulk calculations are finished. Note also that this approximation and the following molecular-charge approximation do not depend on the surface orientation but on the volume of the system involved. Therefore, if this works well for predicting IPs of neutral surfaces, we can adopt the IPs as a definition of bulk contributions of IPs as discussed in Sec. II C. As shown in Fig. 7(a), this can predict the IPs of zinc-blende compounds well. Especially, the predicted IPs for II-VI compounds are almost identical to those of the slab models. However, it works poorly for other solids, especially ionic compounds such as rocksalt oxides and LiH, because the representation of the charge density as a superposition of neutral atoms is very far from the physical situation, owing to significant charge transfer between these ionic species.

To improve the predictive capability, we consider molecules comprised of the constituent atoms and estimated the surface dipole potential from the molecular-charge density in the same manner as the model-solid theory. The surface dipole potential energy can be then calculated as

$$\frac{d}{\epsilon_0} = \frac{Z}{\Omega_{\text{cell}}} \int V_{\text{mol}} d\boldsymbol{r}, \tag{5}$$

where Z is the number of formula unit in the unit cell, V_{mol} the electrostatic potential caused by the diatomic molecule. The interatomic distances were optimized within PBE-GGA. Note that we can calculate the molecules much faster than the slab models. In Fig. 7(b), the IPs estimated from the molecular-charge approximation are shown. It is clearly seen that this approximation drastically improves the prediction of the IPs over the atomic-charge approximation. Especially, it can reproduce the IPs calculated from slab models with atomic relaxation. This is because atoms near surfaces relax in a similar manner to the molecules. The mean absolute error estimated in our test set is 0.4 eV. Since the IPs generally range between 4 and 12 eV [31], its prediction error roughly corresponds to 5%. Combining this method with an accurate calculation of the bulk eigenvalue, for example using the GW method (which provides a calculate error of approximately 0.2 eV compared to experiment [4,5]) would result in IPs with an expected error of ~0.6 eV, compared to experiment; this is still an acceptable accuracy for application in a high-throughput screening approach. A drawback of this approximation is that the calculations of molecules is straightforward only for such simple binaries, and we need a further consideration for complex compounds because their constituent molecules can have various configurations.

The MgO (100) surface shows the largest error in the molecular-charge approximation. Therefore we checked the convergence of the dipole potential as a function of the size of the MgO clusters surrounded by the (100) surface. Figure 8 shows the calculated dipole potential from the Mg-O molecule and MgO clusters with various sizes. For 64- and 216-atom clusters, the dipole potential is estimated from the average electrostatic potential in the central region of each cluster with respect to the vacuum potential at the outermost point in the supercell (see Fig. 8). One can see that the dipole potential is well calculated even with the eight-atom cluster and almost converged in the 64-atom cluster. This indicates that the IP of the MgO (100) surface, having the largest error in our test set, can be very accurately calculated only with half unit-cell width for the surface layer.

III. CONCLUSIONS

In this paper, we theoretically investigated the IPs of nonmetallic solids. First, we discussed how the IPs are decomposed into the bulk electronic eigenvalue $\epsilon^{\circ}_{\text{VBM}}$ and surface dipole potential energy d/ϵ_0 from a macroscopic viewpoint (Sec. II A).

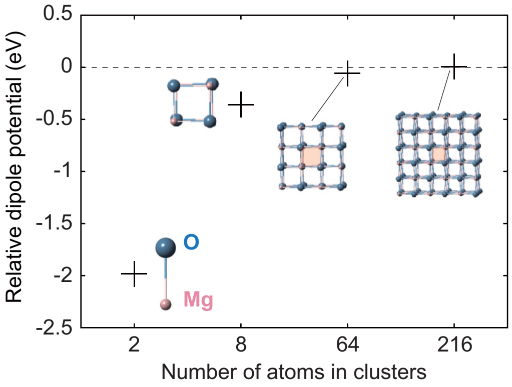

FIG. 8. Relative dipole potential of the MgO (100) surface calculated using the Mg-O molecule and eight-, 64-, and 216-atom clusters. The reference is set to the value of the relaxed MgO (100) slab model. The atomic positions in the molecule and clusters are relaxed except for the eight atoms at the center of 64- and 216-atom clusters, for which the macroscopic dipole potential is estimated from the average electrostatic potential at the central region colored by light red. On the other hand, for the eight-atom cluster, it is estimated in the same manner as the molecule.

Next, we considered the IPs of realistic materials in Sec. II B. A detailed argument was developed with the (100) and (110) surfaces of the prototypical solid, LiH. To comply with the macroscopic viewpoint, we adopted the macroscopic average technique, by which we define the macroscopic dipole density d and macroscopic dipole potential energy d/ϵ_0. These are composed of the electron leakage and clipped bulk contributions. However, since the surface boundary is ill defined, this partition is not unique. It is found that the bulk electronic eigenvalues are mostly positive, and a large macroscopic dipole potential significantly pushes the electron's energy levels downward, resulting in the positive IPs (see also Appendix). Although the partition of IPs into $\epsilon^{\circ}_{\text{VBM}}$ and d/ϵ_0 is well defined from macroscopic viewpoint, it is difficult to find the clear physical meaning of this split as neither quantity correlates strongly to the IPs, and they largely cancel out each other when constructing the IPs. We also discussed the possibility of an alternative splitting procedure that is consistent with the physical intuition, and conclude that it is possible only when the charge density of both bulk and neutral surfaces is well decomposed into a superposition of spherical charges.

Finally, we estimated the IPs using two approximations, *i.e.*, atomic-charge and molecular-charge approximations in Sec. II D. The former can estimate the IPs of covalent systems quite successfully, but works poorly for ionic systems. On the contrary, the latter can predict the IPs of both covalent and ionic materials overall. This is because the calculations of molecules can take into account the charge transfer from cations to anions. Furthermore, we found that the MgO (100) surface has the largest discrepancy between the IPs of the slab model and the molecular-charge approximation in our calculation set, but its IP can be accurately calculated with the 64-atom cluster which has just half unit-cell width for the surface layer, indicating a rather localized nature of the surface region at this surface.

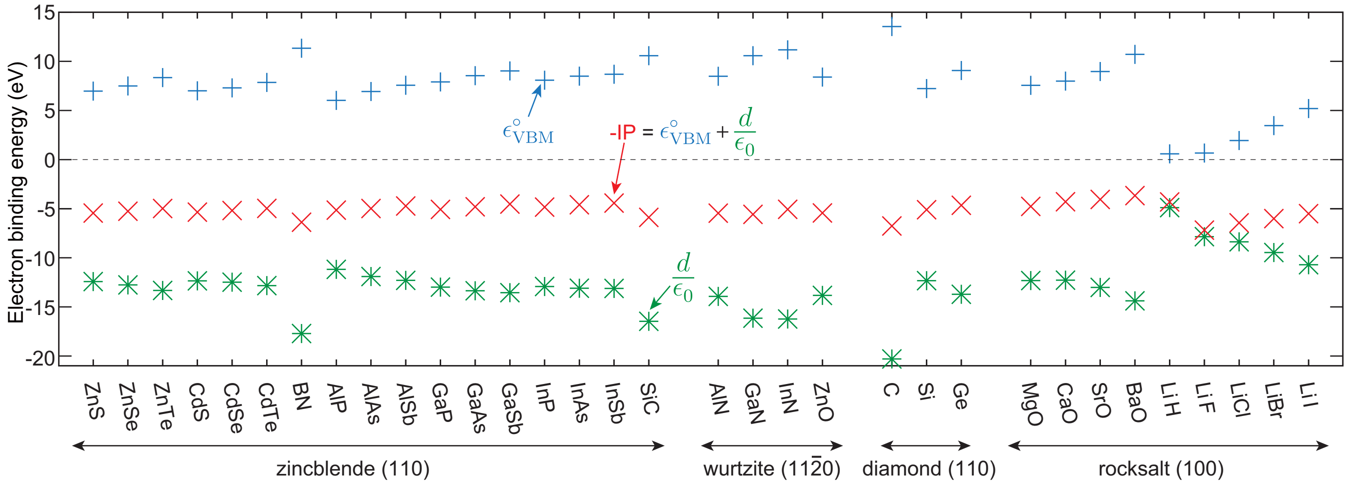

FIG. 9. Calculated electron binding energies without atomic relaxation for materials taking zinc-blende, wurtzite, diamond, and rocksalt structures, and bulk electronic eigenvalues ($\epsilon^{\circ}_{\mathrm{VBM}}$) and macroscopic surface dipole potential energy (d/ϵ_0) (see text for details). The calculated surface orientations shown are classified to the Tasker type 1 surface. Note that $\epsilon^{\circ}_{\mathrm{VBM}}$ and d/ϵ_0 are corrected to the all-electron ones (see text for details).

ACKNOWLEDGMENTS

We sincerely acknowledge discussions with J. Buckeridge and A. A. Sokol. This work was supported by the MEXT Elements Strategy Initiative to Form Core Research Center, Grants-in-Aid for Young Scientists A (Grant No. 15H05541), Scientific Research B (Grant No. 15H04125), and JSPS Fellows (Grant No. 26-04792) from JSPS, and PRESTO and Support Program for Starting Up Innovation Hub MI^{2}I from JST, Japan. K.T.B. and A.W. acknowledge support from EP/M009580/1 and the Royal Society for a University Research Fellowship, respectively. Computing resources of ACCMS at Kyoto University were partly used in this work.

APPENDIX

To see the meaning of the split of the IP into bulk electronic eigenvalues and a macroscopic surface dipole potential, we also show the results of systematic calculations for a wide variety of materials taking the diamond, zinc-blende, and rocksalt structures in this Appendix. An IP is almost independent of the choice of PAW data set as long as electron reconstruction at the surface is accurately described by the PAW method. However, since the PAW method does not explicitly consider charge density near nuclei, its calculated $\epsilon^{\circ}_{\mathrm{VBM}}$ and d/ϵ_0 are decreased and increased from the AE values, respectively. The change of $\epsilon^{\circ}_{\mathrm{VBM}}$ in LiH as a function of PAW radii has also been discussed in Ref. [24]. It is, however, prohibitively computationally demanding to calculate other materials including heavier elements with the AE method. Therefore we here use a correction scheme developed by Bruneval *et al.* to estimate $\epsilon^{\circ}_{\mathrm{VBM}}$ using the PAW method [23,24]. We can indirectly estimate d/ϵ_0 from $\epsilon^{\circ}_{\mathrm{VBM}}$ and IPs via Eq. (2). Figure 3 shows an example of LiH, where $\epsilon^{\circ}_{\mathrm{VBM}}$ was calculated using the PAW potential with a Bruneval correction ($\epsilon^{\circ}_{\mathrm{VBM(cor)}}$) by the ABINIT code [32]. We have found that the discrepancy between $\epsilon^{\circ}_{\mathrm{VBM(cor)}}$ and AE $\epsilon^{\circ}_{\mathrm{VBM}}$ ($\epsilon^{\circ}_{\mathrm{VBM(AE)}}$) is less than 10 meV, and the obtained IPs at the LiH (100) surface are almost identical within 3 meV. Therefore d/ϵ_0 can be estimated using the PAW method via Eq. (2) to be -4.86 eV, which is almost identical to the AE value of -4.87 eV.

Figure 9 shows the calculated electron binding energies $d/\epsilon_0 + \epsilon^{\circ}_{\mathrm{VBM}}$ with the PBE-GGA functional, and their decomposition into d/ϵ_0 and $\epsilon^{\circ}_{\mathrm{VBM}}$. The surfaces considered are all classified as Tasker type 1 surfaces at which each plane holds neutral charge [33]. As represented by the LiH (100) and (110) surfaces shown in Fig. 3, the IPs of the Tasker type 1 surfaces are very similar to each other in the same system [34,35]. The atomic positions are fixed at those in the PBE-GGA structures for simplicity. Note that PBE-GGA typically underestimates band gaps and GaSb, InAs, and InSb show metallic behaviors in our test set although they are semiconductors in experiments. Their IPs are set to the Fermi levels in this study. Such underestimated IPs

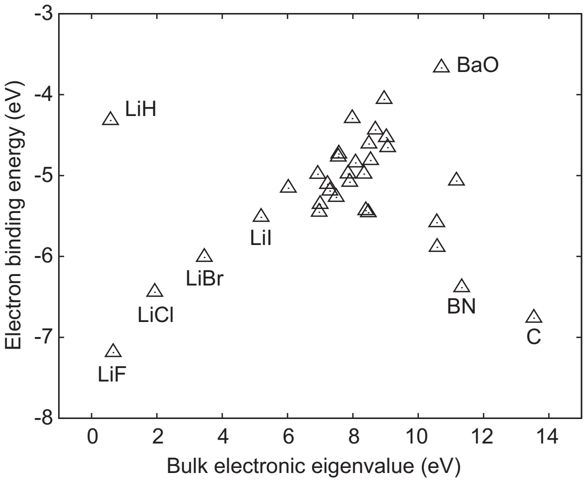

FIG. 10. Electron binding energies against the bulk electronic eigenvalues ($\epsilon^{\circ}_{\mathrm{VBM}}$). Here, $\epsilon^{\circ}_{\mathrm{VBM}}$ are corrected to the all-electron values (see text for details).

can be improved by using the hybrid functionals or *GW* approximations compared to the experimental values, which can be significant for practical applications [4,5,36]. It is generally seen that, when calculating IPs, $\epsilon^{\circ}_{\mathrm{VBM}}$ and d/ϵ_0 largely counteract each other, and when $\epsilon^{\circ}_{\mathrm{VBM}}$ is high, d/ϵ_0 tends to be low. As a result, the variations of IPs are not so large compared to both $\epsilon^{\circ}_{\mathrm{VBM}}$ and d/ϵ_0. It is also seen that $\epsilon^{\circ}_{\mathrm{VBM}}$ is increased as the lattice constant is decreased whereas d/ϵ_0 is decreased, which is also discussed with the jellium model in Sec. II B.

The bulk electronic eigenvalue $\epsilon^{\circ}_{\mathrm{VBM}}$ can be obtained from a simple bulk calculation. Therefore, if the bulk eigenvalue were significantly correlated to the IP, one could predict the IP without requiring an expensive slab calculation. In order to assess this correlation, Fig. 10 shows the plot of the electron binding energies against $\epsilon^{\circ}_{\mathrm{VBM}}$. The electron binding energy seems positively correlated with $\epsilon^{\circ}_{\mathrm{VBM}}$, but the correlation is unfortunately not strong, as demonstrated by the discrepancy between LiH and Li halides. Therefore it would be difficult to deduce the IPs only from $\epsilon^{\circ}_{\mathrm{VBM}}$ obtained through a bulk calculation.

[1] D. Cahen and A. Kahn, Electron energetics at surfaces and interfaces: Concepts and experiments, Adv. Mater. **15**, 271 (2003).

[2] V. Stevanović, S. Lany, D. S. Ginley, W. Tumas, and A. Zunger, Assessing capability of semiconductors to split water using ionization potentials and electron affinities only, Phys. Chem. Chem. Phys. **16**, 3706 (2014).

[3] W. Chen and A. Pasquarello, Band-edge positions in *gw*: Effects of starting point and self-consistency, Phys. Rev. B **90**, 165133 (2014).

[4] A. Grüneis, G. Kresse, Y. Hinuma, and F. Oba, Ionization Potentials of Solids: The Importance of Vertex Corrections, Phys. Rev. Lett. **112**, 096401 (2014).

[5] Y. Hinuma, A. Grüneis, G. Kresse, and F. Oba, Band alignment of semiconductors from density-functional theory and many-body perturbation theory, Phys. Rev. B **90**, 155405 (2014).

[6] A. Franciosi and C. G. Van de Walle, Heterojunction band offset engineering, Surf. Sci. Rep. **25**, 1 (1996).

[7] Y. Hinuma, F. Oba, Y. Kumagai, and I. Tanaka, Band offsets of $CuInSe_2$/CdS and $CuInSe_2$/ZnS (110) interfaces: A hybrid density functional theory study, Phys. Rev. B **88**, 035305 (2013).

[8] J. Robertson and S. J. Clark, Limits to doping in oxides, Phys. Rev. B **83**, 075205 (2011).

[9] S. B. Zhang, S.-H. Wei, and A. Zunger, A phenomenological model for systematization and prediction of doping limits in ii-vi and i-iii-vi_2 compounds, J. Appl. Phys. **83**, 3192 (1998).

[10] K. T. Butler, J. Buckeridge, C. R. A. Catlow, and A. Walsh, Crystal electron binding energy and surface work function control of tin dioxide, Phys. Rev. B **89**, 115320 (2014).

[11] A. J. Logsdail, D. O. Scanlon, C. R. A. Catlow, and A. A. Sokol, Bulk ionization potentials and band alignments from three-dimensional periodic calculations as demonstrated on rocksalt oxides, Phys. Rev. B **90**, 155106 (2014).

[12] L. Kleinman, Comment on the average potential of a Wigner solid, Phys. Rev. B **24**, 7412 (1981).

[13] G. Makov and M. C. Payne, Periodic boundary conditions in ab initio calculations, Phys. Rev. B **51**, 4014 (1995).

[14] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, Finite-size supercell correction schemes for charged defect calculations, Phys. Rev. B **86**, 045112 (2012).

[15] Y. Kumagai and F. Oba, Electrostatics-based finite-size corrections for first-principles point defect calculations, Phys. Rev. B **89**, 195205 (2014).

[16] Y. Kumagai, M. Choi, Y. Nose, and F. Oba, First-principles study of point defects in chalcopyrite $ZnSnP_2$, Phys. Rev. B **90**, 125202 (2014).

[17] J. Bardeen, Theory of the work function. ii. the surface double layer, Phys. Rev. **49**, 653 (1936).

[18] N. D. Lang and W. Kohn, Theory of metal surfaces: Work function, Phys. Rev. B **3**, 1215 (1971).

[19] C. G. Van de Walle and R. M. Martin, Theoretical calculations of heterojunction discontinuities in the Si/Ge system, Phys. Rev. B **34**, 5621 (1986).

[20] C. G. Van de Walle and R. M. Martin, Theoretical study of band offsets at semiconductor interfaces, Phys. Rev. B **35**, 8154 (1987).

[21] C. G. Van de Walle, Band lineups and deformation potentials in the model-solid theory, Phys. Rev. B **39**, 1871 (1989).

[22] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1999), pp. 31–34.

[23] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[24] F. Bruneval, J.-P. Crocombette, X. Gonze, B. Dorado, M. Torrent, and F. Jollet, Consistent treatment of charged systems within periodic boundary conditions: The projector augmented-wave and pseudopotential methods revisited, Phys. Rev. B **89**, 045116 (2014).

[25] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

[26] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[27] P. G. Moses, M. Miao, Q. Yan, and C. G. Van de Walle, Hybrid functional investigations of band gaps and band alignments for AlN, GaN, InN, and InGaN, J. Chem. Phys. **134**, 084703 (2011).

[28] A. Baldereschi, S. Baroni, and R. Resta, Band Offsets in Lattice-Matched Heterojunctions: A Model and First-Principles Calculations for Gaas/Alas, Phys. Rev. Lett. **61**, 734 (1988).

[29] D. M. Bylander and L. Kleinman, Comparison of Dipole Layers, Band Offsets, and Formation Enthalpies of Gaas-Alas(110) and (001) Interfaces, Phys. Rev. Lett. **59**, 2091 (1987).

[30] C. G. Van de Walle, Theoretical calculations of semiconductor heterojunction discontinuities, J. Vac. Sci. Technol. B **4**, 1055 (1986).

[31] K. T. Butler, Y. Kumagai, F. Oba, and A. Walsh, Screening procedure for structurally and electronically matched contact layers for high-performance solar cells: Hybrid perovskites, J. Mater. Chem. C **4**, 1149 (2016).

[32] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf,

M. Torrent, M. Verstraete, G. Zerah, and J. Zwanziger, Abinit: First-principles approach to material and nanosystem properties, Comput. Phys. Commun. **180**, 2582 (2009).

[33] P. W. Tasker, The stability of ionic crystal surfaces, J. Phys. C **12**, 4977 (1979).

[34] Y. Hinuma, Y. Kumagai, F. Oba, and I. Tanaka, Categorization of surface polarity from a crystallographic approach, Comput. Mater. Sci. **113**, 221 (2016).

[35] Y. Kumagai, L. A. Burton, A. Walsh, and F. Oba, Electronic Structure and Defect Physics of Tin Sulfides: SnS, Sn_2S_3, and SnS_2, Phys. Rev. Appl. **6**, 014009 (2016).

[36] Y. Hinuma, Y. Kumagai, I. Tanaka, and F. Oba, Band alignment of semiconductors and insulators using dielectric-dependent hybrid functionals: Toward high-throughput evaluation, Phys. Rev. B **95**, 075302 (2017).