PHYSICAL REVIEW B **94**, 035113 (2016)

Accurate ionization potential of semiconductors from efficient density functional calculations

Lin-Hui Ye
Key Laboratory for the Physics and Chemistry of Nanodevices, Department of Electronics, Peking University, Beijing 100871, People's Republic of China

(Received 13 January 2016; revised manuscript received 22 June 2016; published 6 July 2016)

Despite its huge successes in total-energy-related applications, the Kohn-Sham scheme of density functional theory cannot get reliable single-particle excitation energies for solids. In particular, it has not been able to calculate the ionization potential (IP), one of the most important material parameters, for semiconductors. We illustrate that an approximate exact-exchange optimized effective potential (EXX-OEP), the Becke-Johnson exchange, can be used to largely solve this long-standing problem. For a group of 17 semiconductors, we have obtained the IPs to an accuracy similar to that of the much more sophisticated GW approximation (GWA), with the computational cost of only local-density approximation/generalized gradient approximation. The EXX-OEP, therefore, is likely as useful for solids as for finite systems. For solid surfaces, the asymptotic behavior of the v_{xc} has effects similar to those of finite systems which, when neglected, typically cause the semiconductor IPs to be underestimated. This may partially explain why standard GWA systematically underestimates the IPs and why using the same GWA procedures has not been able to get an accurate IP and band gap at the same time.

DOI: 10.1103/PhysRevB.94.035113

I. INTRODUCTION

The ionization potential (IP) of a semiconductor is one of the most important material parameters which govern its optics, transport, electrochemistry, and interface properties in heterostructures. Although, in principle, the IP can be calculated by the many-body Green's function method in the GW approximation (GWA) [1], due to the heavy computational cost, in practice, further approximations have to be made, which has remained unjustified to this date. For finite systems, it is usually more convenient to use density functional theory (DFT) to calculate the neutral and ionized systems separately and then obtain an accurate IP from their total energy difference. This approach, however, cannot be used for infinite solids for which DFT can only treat neutral systems.

Luckily enough, in DFT there exists an "IP theorem" [2,3] which states that the negative eigenvalue of the highest occupied Kohn-Sham orbital, $-\varepsilon_N$, happens to be equal to the IP. The reason for the existence of this theorem is because the electron density at $r = \infty$ is contributed by the most extended orbital in space, which happens to be the highest occupied one. In both the quasiparticle theory and DFT, the corresponding orbital density decays exponentially, with the exponent being related in the same way to the IP and $-\varepsilon_N$, respectively. Therefore, for both theories to give the same density at $r = \infty$, IP and $-\varepsilon_N$ must be equal.

However, the IP theorem relies on the exact exchange-correlation potential v_{xc}, which is unknown and must be approximated. Popular approximations such as the local-density approximation (LDA) or generalized gradient approximations (GGA) suffer from the self-interaction (SI) error [4], which pushes up occupied states and causes LDA/GGA to systematically underestimate the IP. For atoms and molecules [5], the IP error (hereafter, "IP error" specifically refers to the deviation of $-\varepsilon_N$ of LDA/GGA from the true IP) can be as large as 5–10 eV, while for semiconductors [6,7] it is typically 1 eV. Consequently, the powerful DFT-Kohn-Sham scheme has not been able to get reliable IPs for semiconductors even as simple as silicon.

This paper takes one major step forward: We will show that the IP theorem when used with an approximate optimized effective potential [8–10] can predict the IPs of semiconductors to an accuracy similar to that of the much more sophisticated GWA. As v_{xc} is a simple, local potential, our whole procedure is restricted to the DFT-Kohn-Sham realm, and therefore, the computational cost is as low as LDA/GGA. This has made it possible to calculate the IP of a complex system composed of hundreds of atoms.

Strictly speaking, for infinite solids there is no exact IP theorem anymore. That is to say, in general, IP is not equal to $-\varepsilon_{\text{vbm}}$, the negative eigenvalue of the valence-band maximum (VBM). This is because out of a solid surface the most extended orbital is generally not the VBM but the highest occupied orbital at the Γ point of the planar Brillouin zone. This fact was first noted by Engel [11]. We hereby use it to show that the exactness of the DFT density far away from the surface has nothing to do with the VBM energy, unless the VBM happens to locate at Γ. Nevertheless, the possible deviation of $-\varepsilon_{\text{vbm}}$ from the true IP is likely small and is not considered in this work.

The condition for an accurate IP is much more stringent than for accurate total energy since v_{xc} depends on every detail of the exchange-correlation hole, while total energy only requires the spherical average of the hole to be correct. This explains why the same LDA/GGA behave drastically different for the two quantities. For finite systems, it has long been established [5] that the exact v_{xc} must decay as slowly as $-1/r$ away from the physical system. However, due to the SI error, LDA/GGA decay exponentially fast. Consequently, a large negative portion of v_{xc} is missing which is responsible for most of the eigenvalue overestimation. Often, the accuracy of IP can be improved by correcting the wrong asymptotic behavior of the v_{xc}. However, for solid surfaces asymptotically long-range potentials are very hard to use with the supercell method [12], so that the role of the long-range part of v_{xc} has remained unexplored. As revealed by this work, the effects are not as strong as for finite systems but still too significant to be simply ignored.

©2016 American Physical Society

The exact-exchange optimized effective potential (EXX-OEP, or simply OEP, as in this work) is the local multiplicative potential minimizing the Hartree-Fock expression of total energy. Using OEP to calculate a semiconductor IP is inspired by the fact that most SI resides in the exchange part of v_{xc}, while OEP is self-interaction free [13] and satisfies many exact conditions, especially the desired $-1/r$ asymptotic behavior [14]. For neutral and ionized atoms [15] as well as small molecules [6], OEP removes the IP errors by more than 80%.

The capability of OEP to calculate the IP for semiconductors is not trivially seen because its success for atoms and molecules seems to rely on the overwhelmingly dominant role of exchange, while in solids correlation is much stronger. In fact, so far there has been only one work [11] using OEP to calculate the IPs of two semiconductor, Si (111) and graphene. For Si, the obtained value is 5.45 eV. However, this case is questionable because the slab used in the calculation is too thin and the surface is not relaxed. Let us explain both reasons in more detail: First, to get a reliable IP, the slab used in the calculation must be thick enough. Convergence testing has shown that the minimum thickness is nine atomic layers, while in Ref. [11] the slab only contains three bilayers. Second, the IP sensitively depends on the formation of the surface dipole, which means that the surface density must be determined to high accuracy. For this reason the surface of the slab has to be relaxed since the density of the unrelaxed surface is highly unrealistic. In fact, a comparison between relaxed and unrelaxed surfaces has shown that their IPs can differ by as much as 0.5 eV [7]. The other case, graphene, is simpler [16] because there is no issue of slab thickness or surface relaxation. Nevertheless, the obtained IP is 8.13 eV, which severely overestimates the experimental value of 4.6 eV. To compensate for the huge 3.53 eV difference, the author had to conclude that the correlation potential for graphene must be strongly repulsive. In general, according to this only existing work [11], OEP does not seem to be useful for semiconductor IPs. This paper, however, will arrive at a different conclusion.

The application of OEP for solids has not become popular mainly because the construction of the exact OEP is technically very challenging: OEP is computationally costly and is extremely sensitive to the subtle balance of the basis sets [17–19]. For semiconductor IP the problem is more agonized due to its extra surface dependence, which therefore requires the OEP of the complex surface structure. Moreover, OEP decays as slowly as $-1/z$ from the surface. Like other long-range potentials, it is very hard to use with the supercell method [11,12].

This work is made possible by implementing an approximate OEP, the Becke-Johnson 2006 (BJ06) exchange potential [20], to the authentic slab geometry (ASG) [21,22], which is a surface technique suitable for both short- and long-range potentials [12]. Even for simple systems, using approximate OEP to avoid the intricacy of the exact OEP is common. For example, the Krieger-Li-Iafrate (KLI) approximation [23] has been much more popular than the exact OEP itself. Compared to KLI, BJ06 is less accurate. But for the following reasons BJ06 will well represent the performance of OEP: (i) For atoms BJ06 approaches the exact OEP very closely. (ii) For solids it satisfies the uniform electron-gas limit. (iii) For finite systems BJ06 decays as $-1/r + C$ (C is positive and system dependent), and it is found [24] to reduce the IP errors by as much as 60%. (iv) For solid surfaces BJ06 decays as [12] $-1/z + C$, which will be shown to be equivalent to the asymptotic behavior of the exact OEP. To account for the stronger correlation in solids, in this work BJ06 is amended by the LDA correlation [25]. The total v_{xc} is called BJ06c throughout this paper.

The rest of this paper is organized as follows: Sec. II explains the major technological issues, first the method for calculating semiconductor IPs and then the details of our calculations. Section III presents the main IP results, which are also compared to GWA and experiment. Discussion of the performance of our method is offered, which is then followed by analysis of the effects of the long-range part of the v_{xc}. After that, we comment on the existing GWA calculations of semiconductor IPs. Finally, Sec. IV summarizes the major achievements of this work.

II. TECHNOLOGICAL ISSUES

A. Method for semiconductor IP calculations

The IP calculation follows the usual two-step procedure: first a bulk calculation (with the primitive cell) to get $\varepsilon_{\mathrm{vbm}}$, and then a surface calculation using a slab like Fig. 1 for the alignment of the Coulomb potential. The surface calculation is needed since in the bulk calculation the Coulomb potential is determined only up to a global constant due to the use of periodic boundary conditions. The expression of IP used in this work is

$$\mathrm{IP} = -\varepsilon^{\mathrm{p.cell}}_{\mathrm{vbm}} + \overline{v}^{\mathrm{p.cell}}_{\mathrm{Coul}} - \overline{v}^{\mathrm{s.cell}}_{\mathrm{Coul}}, \tag{1}$$

in which $\overline{v}^{\mathrm{p.cell}}_{\mathrm{Coul}}$ is the average Coulomb potential of the bulk and $\overline{v}^{\mathrm{s.cell}}_{\mathrm{Coul}}$ is the drop of the average Coulomb potential from the vacuum edge to the center of the slab. We emphasize that $\overline{v}^{\mathrm{s.cell}}_{\mathrm{Coul}}$ is caused by the formation of surface dipoles; therefore, it is very sensitive to the surface charge distribution. This is the first hint that the long-range part of v_{xc} will play an important role in the final IP.

B. Details of the surface calculations

The 17 semiconductors belong to the diamond, wurtzite, and zinc-blende structures; for surface calculations the slab thickness is set to 15, 23, and 24 atomic layers, respectively. To account for the long-range character of BJ06, the vacuum thickness is set to the huge value of 440 bohrs, facilitated by the authentic slab geometry [21,22]. The two-dimensional Brillouin zone integrations are performed by the special **k**-point method using a Monkhorst-Pack mesh of 6 × 8 for

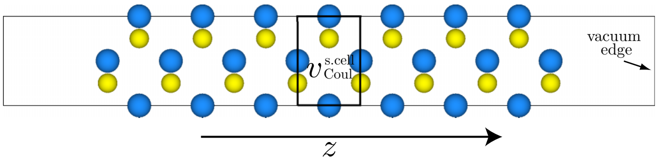

FIG. 1. A ZnS slab of 15 atom layers. The highlighted region is used for macroscopic averaging [26].

diamond and zinc blende and 8 × 8 for wurtzite. All above parameters are either the same as or more stringent than those used in the earlier works [7,27,28]. Geometry relaxation is performed with GGA, although atoms within the slab cell are kept fixed to avoid too much distortion of the density in the cell. Note that even so, the density near the cell boundary is still unavoidably affected. For this reason, one extra atom layer on both sides of the slab cell is also kept fixed to minimize the effect. Both the bulk and the slab calculations are iterated to self-consistency until the final distances of the density and the potential are both smaller than 0.01 atomic units. The full-potential linearized augmented plane wave (FLAPW) method [29] is used for all calculations with spin-orbital effects included.

III. RESULTS AND DISCUSSIONS

Using BJ06c within ASG, we have calculated the IPs of 17 semiconductors. The results are presented in Table I and Fig. 2 and are compared to GWA [7,27] and experiment. Through the two-step procedure, the accuracy of the IP is determined by the surface part of the calculation, which may be checked by comparing Eq. (1) with $-\varepsilon_{\mathrm{vbm}}^{\mathrm{slab}}$, the negative eigenvalue of the VBM of the slab. This is because, providing surface states do not cover up the bulk VBM, Eq. (1) and $-\varepsilon_{\mathrm{vbm}}^{\mathrm{slab}}$ give the same IP value unless the slab is not sufficiently thick. Indeed, for 14 semiconductors Eq. (1) and $-\varepsilon_{\mathrm{vbm}}^{\mathrm{slab}}$ differ by less than 0.1 eV. The exceptions are BN, SiC, and CdS, for which the VBM is from a surface state. Therefore, the numerical accuracy of our IPs is better than 0.1 eV.

TABLE I. The GGA and BJ06c IPs of the 17 semiconductors calculated from Eq. (1) and the negative of the VBM eigenvalue of the slab. The *GW* results are calculated by GW@GGA (Ref. [7]) and GWΓ@HSE (Ref. [27]). Experimental results are collected from these two works. For the original reports see references therein. Units are eV.

	GGA		BJ06c		*GW*		
	Eq. (1)	$\varepsilon_{\mathrm{vbm}}^{\mathrm{slab}}$	Eq. (1)	$\varepsilon_{\mathrm{vbm}}^{\mathrm{slab}}$	Ref. [7]	Ref. [27]	Experiments
Si	4.66	4.68	5.36	5.38	5.45	5.47	5.13, 5.33, 5.1, 5.25, 5.35
Ge	4.08	4.07	5.09	5.03	4.55	5.07	4.75, 4.8, 4.74
GaN	5.73	5.68	6.60	6.53	6.97	7.12	6.6, 6.8, 7.75
GaP	5.06	5.04	5.87	5.83	5.82	6.10	5.95, 6.01
ZnO	5.95	6.00	6.87	6.88	7.46	8.19	7.82
ZnS	5.93	5.90	6.86	6.81	7.01	7.40	7.5
ZnSe	5.53	5.50	6.41	6.37	6.40	6.92	6.82
ZnTe	4.71	4.69	5.79	5.76	5.67	5.89	5.76, 5.75
CdS	5.89	5.79	6.79	6.62	6.83	7.14	6.1, 7.26
CdSe	5.43	5.38	6.42	6.33	6.29	6.79	6.62
CdTe	4.79	4.75	5.90	5.82	5.90	5.93	5.78, 5.8
BN	6.87	6.51	7.93	7.57		8.52	
AlP	5.57	5.51	6.37	6.32		6.62	
AlAs	5.10	5.03	6.06	5.99		6.26	
AlSb	4.57	4.52	5.57	5.51		5.36	5.22
C	5.45	5.53	6.41	6.49		6.74	5.85, 6.0, 6.5
SiC	5.80	5.60	6.66	6.45		7.00	5.9, 6.0

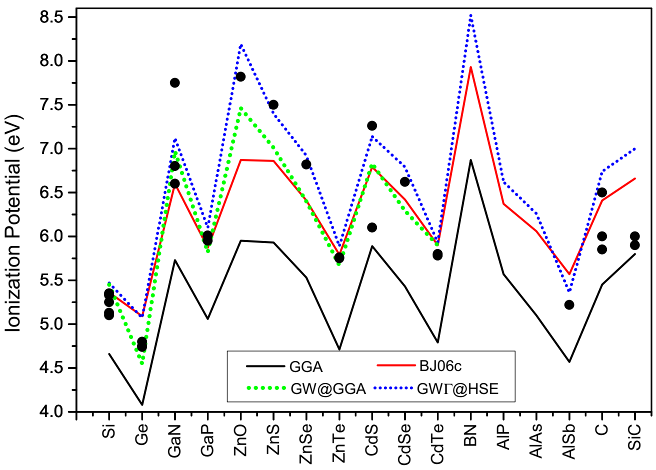

FIG. 2. IPs of the 17 semiconductors from GGA (black solid line) and BJ06c (red solid line) calculated by Eq. (1). The experimental data (black dots) and the GW results (blue and green dotted lines) are the same as in Table I.

Consistent with earlier works, GGA [30] consistently underestimates the experimental IPs by about 1 eV. With BJ06c, all IPs are upshifted, and the corrections upon GGA are from 0.70 eV (for Si) to 1.11 eV (for CdTe). Experimentally, the IP measurement is extremely sensitive to the sample's surface condition, which has led to large scattering of the IP data. Consequently, a strict quantitative comparison between theories and experiment is impossible. Nevertheless, the extent of the corrections by BJ06c is in the desired range and is similar to that of GWA. The largest discrepancy is found for ZnO [31], which is also a difficult case for GWA. In general, the performance of BJ06c is quite impressive since it is a local potential, requiring essentially negligible computational cost compared to GWA.

From Table I and Fig. 2 the performance of BJ06c is about equally good for solids as for atoms [15] and small molecules [6]. This seems to be inconsistent with the only earlier work [11] which implies that OEP may not be useful for semiconductor IPs. In fact, we have also calculated the IP of graphene by BJ06c and obtained 4.6 eV, which is in excellent agreement with experiment. Therefore, a strongly repulsive correlation potential is not needed. We note in Ref. [11] the OEP was generated by supercell. Limited by the low efficiency of the supercell method, the sheet separation is only 14 bohrs, which is definitely not sufficient to achieve the asymptotic value of the OEP at the vacuum edge.

The long-range part of v_{xc} has been well understood for finite systems but is widely ignored in surface calculations for which the erroneously short range LDA/GGA are almost exclusively used. To check its effects for a solid surface, we have decomposed ΔIP, the total IP correction upon GGA, into the three individual contributions of $\Delta\varepsilon_{\mathrm{vbm}}^{\mathrm{p.cell}}$, $\Delta\overline{v}_{\mathrm{Coul}}^{\mathrm{p.cell}}$, and $\Delta\overline{v}_{\mathrm{Coul}}^{\mathrm{s.cell}}$, following Eq. (1). With this decomposition the effects of the long-range part of v_{xc} are only reflected in the surface term since in the bulk calculation there is no asymptotic region. Note that the electron self-energy also shares a similar asymptotic behavior [32]; the existing GWA studies [7,27,28] can therefore be used as convenient examples

TABLE II. Total IP correction ΔIP and the individual contributions of $\Delta\varepsilon_{\text{vbm}}^{\text{p.cell}}$, $\Delta\overline{v}_{\text{Coul}}^{\text{p.cell}}$, and $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$ from Eq. (1). PCT gives the percentages of the contributions. Units are eV.

		$\Delta\varepsilon_{\text{vbm}}^{\text{p.cell}}$		$\Delta\overline{v}_{\text{Coul}}^{\text{p.cell}}$		$\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$	
	ΔIP	Value	PCT	Value	PCT	Value	PCT
Si	0.70	−0.62	89	−0.03	−4	−0.11	16
Ge	1.00	−0.88	88	−0.08	−8	−0.20	20
GaN	0.87	−0.78	90	−0.02	−2	−0.11	13
GaP	0.80	−0.56	70	−0.02	−3	−0.27	34
ZnO	0.93	−1.07	115	−0.03	−3	0.11	−12
ZnS	0.93	−0.73	78	−0.02	−2	−0.22	24
ZnSe	0.88	−0.74	84	−0.04	−5	−0.18	20
ZnTe	1.07	−0.75	70	−0.13	−12	−0.45	42
CdS	0.90	−0.75	83	−0.06	−7	−0.21	23
CdSe	1.00	−0.76	76	−0.07	−7	−0.30	30
CdTe	1.11	−0.78	70	−0.15	−14	−0.47	42
BN	1.06	−0.83	78	0.01	1	−0.22	21
AlP	0.80	−0.64	80	−0.02	−3	−0.18	23
AlAs	0.96	−0.72	75	−0.04	−4	−0.28	29
AlSb	1.00	−0.80	80	−0.12	−12	−0.33	33
C	0.96	−0.73	76	0.01	1	−0.22	23
SiC	0.87	−0.72	83	−0.01	−1	−0.16	18

to assist our analysis. For this purpose, let us recall that all GWA calculations are only applied to the bulk, while the slabs are always treated by GGA so that

$$\text{IP}^{\text{GWA}} = \text{IP}^{\text{GGA}} - \Delta\varepsilon_{\text{vbm}}^{\text{p.cell}} + \Delta\overline{v}_{\text{Coul}}^{\text{p.cell}}. \tag{2}$$

In this way, $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$, together with all the effects of the long-range part of v_{xc}, is completely ignored. In fact, if the GWA of the bulk is performed in the one-shot fashion, i.e., without self-consistency, then $\Delta\overline{v}_{\text{Coul}}^{\text{p.cell}}$ is also ignored. Since Coulomb potential is uniquely determined by density, the underlying assumption is that the GGA densities are already accurate enough to be directly used for the IP calculations.

From Table II, the average contributions of the three terms to the total ΔIP are 82%, 23%, and −5%, respectively. Therefore, the bulk part of the corrections indeed accounts for most of the ΔIP, which lends fundamental support to the GWA treatment. Especially, the smallness of $\Delta\overline{v}_{\text{Coul}}^{\text{p.cell}}$ implies that the GGA bulk density may be indeed of high quality. A problem, however, exists for $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$ since its 23% contribution to the total ΔIP is certainly non-negligible. Especially, for the two Te compounds the contributions are larger than 40%. This means that the GGA slab density is not sufficiently accurate.

The reason why the two GGA densities are of different quality is because SI behaves differently in the bulk and the slab: Within the bulk, SI usually causes the density to overly delocalize. However, this problem mostly affects strongly correlated materials, while for weak to intermediate correlations it is not very serious. In fact, within the bulk the SI error is largely screened out by the response of the Coulomb potential, as illustrated by Li *et al.* [14]. On the other hand, in the slab case the Coulomb potential cannot compensate for the missing tail of GGA since it decays even faster than v_{xc}. Consequently, the SI error is more prevailing in the slab than in the bulk.

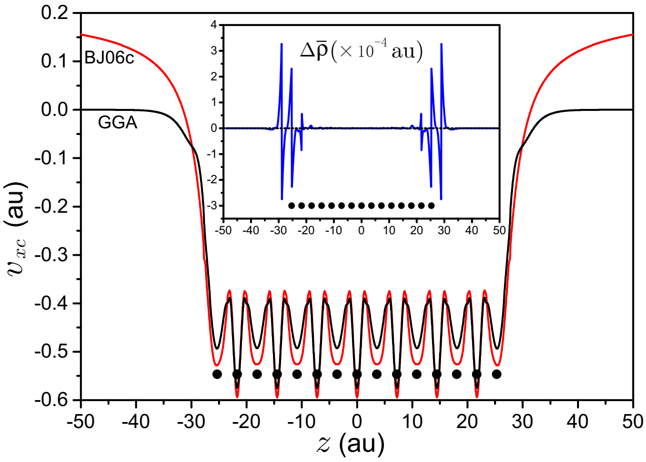

FIG. 3. v_{xc} of the GGA (black) and BJ06c (red) for the ZnS slab in Fig. 1 along $(0,0,z)$. The plot path is between atomic sites to avoid the divergence of the GGA at the nuclei. The inset shows the difference of the macroscopically averaged [26] slab density $\Delta\bar{\rho}(z) = \bar{\rho}^{\text{BJ06c}}(z) - \bar{\rho}^{\text{GGA}}(z)$. Black dots indicate the location of the atomic planes.

To make sure that $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$ is indeed related to the long-range part of BJ06c, we have plotted the two v_{xc} for Fig. 1 together with the difference of their macroscopically averaged [26] densities. Figure 3 shows that within the bulk BJ06c is deeper than GGA, which explains the downshift of ε_{vbm}. The most obvious distinction, however, is around the slab surface: While GGA already decays to zero at about $z = 40$, BJ06c climbs up very slowly to its 0.20 value at the vacuum edge of $z = 247$ (not shown). Correspondingly, within the bulk the electron density is essentially unchanged. Only near the surface does the slow variation of BJ06c strongly perturb the surface density. For the relaxed slab, this enhances the surface dipole and changes $\overline{v}_{\text{Coul}}^{\text{s.cell}}$ by −0.22 eV, which contributes positively to the total ΔIP of 0.93 eV.

Compared to atoms and molecules, the effects of the long-range part of the v_{xc} are weaker for the solid surface. This is because the wave functions are mainly bound to the slab but less exposed to the vacuum or the asymptotic region. At this point, it is necessary to point out that the exact asymptotic behavior of v_{xc} for the solid surface is still unresolved [11,12]. In BJ06c, the asymptotic behavior $-1/z + C$ of the exchange part is exact according to the form found in Ref. [3] since an overall shift of the potential by the C constant has no effect on the slab density or $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$. On the other hand, the LDA correlation is still erroneously short ranged. Consequently, the $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$ data in Table II contain systematic errors which are hard to evaluate. Nevertheless, $\Delta\overline{v}_{\text{Coul}}^{\text{s.cell}}$ is at least system dependent, and there is no apparent reason why it can be universally neglected. For example, it is found that GGA can successfully line up the Coulomb potential at the interface of heterostructures. This is frequently cited to support the similar use of GGA for the potential alignment in the IP calculation by GWA [7,27,28]. However, the two interfaces are fundamentally different: The asymptotic behavior of v_{xc} is operative only at the slab-vacuum interface but has little effect at the interface of heterostructures because, again, there is no asymptotic region.

Indirect evidence supporting non-negligible $\Delta\overline{v}^{\text{s.cell}}_{\text{Coul}}$ may be found in existing GWA results. As Table II shows, except for ZnO, all $\Delta\overline{v}^{\text{s.cell}}_{\text{Coul}}$ contribute positively to ΔIP. Therefore, by neglecting this term GWA shall systematically underestimate the semiconductor IPs. Indeed, it has been consistently found [7,27,28] that semiconductor IPs calculated by GWA starting from GGA are systematically too low. Nevertheless, since the same GWA procedure also underestimates the band gaps, it is envisioned that both problems are due to the same origin of insufficient quasiparticle corrections.

If $\Delta\overline{v}^{\text{s.cell}}_{\text{Coul}}$ is indeed universally negligible, then by merely improving GWA it will be possible to achieve accurate band gaps and accurate IPs at the same time. To enhance quasiparticle corrections, hybrid functionals have been used to replace the GGA starting point [27,28]. As they seem to overcorrect the problems, the fraction of the Fock exchange is changed to an adjustable parameter [28]. Alternatively, the vertex function from certain cross diagrams is included to scale down the corrections [27]. Also, quasiparticle self-consistency [33] with the screened interaction of GGA [28] is also attempted. It is found, however, that when the IPs seem to agree with experiment, the band gaps are unevenly overestimated [27]. On the other hand, when the band gaps seem to agree with experiment, half of the IPs are overestimated [28]. Although in both cases it is claimed that further vertex corrections would reconcile the conflicts, it is not clear how the extra corrections would affect only either the IPs or the band gaps while leaving the other set unchanged.

IV. SUMMARY

This work has shown that it is indeed possible to stay within the DFT-Kohn-Sham realm while calculating semiconductor IPs to far better accuracy than LDA/GGA by using simple, local potentials. This achievement is very inspiring since it has greatly reduced the challenge of large-scale IP calculations for very complex systems containing, for example, hundreds of atoms. Moreover, the improvement of the excitation energy is actually not limited to only the VBM but is also transferred to neighboring **k** points in the Brillouin zone. This is because the dispersions of the DFT bands are often quite accurate. Therefore, following the downshift of the VBM, a bunch of valence bands near the Fermi level are also pushed closer to their true excitation energies. Consequently, we expect that the present work shall be very useful for a large range of applications.

In addition to the above, we have also performed a systematic study of the long-range effects of the v_{xc}. Although a similar topic on finite systems has been under extensive discussion for at least two decades, for solid surfaces much less has been done. Theoretically, the asymptotic behavior of the total v_{xc} has not been made clear. Our work has warned that the effects of the long-range part of v_{xc} on the excitation energies of solids cannot simply be ignored, although the strength is not as strong as for atoms and molecules because the wave functions are less exposed to the vacuum region.

The closeness of BJ06 to the exact OEP for atoms has been well established and is likely inherited in solid environment. Therefore, the performance of BJ06 for semiconductor IPs revealed in this work will reflect well that of the exact OEP. So far, construction and applications of the OEP for solids have not become common. The success of the present work will invite more studies in this direction.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (Grants No. 61390504, No. 91221202) and by the National Basic Research Program of China (Grants No. 2012CB932700, No. 2012CB932703).

---

[1] L. Hedin, Phys. Rev. **139**, A796 (1965).

[2] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).

[3] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).

[4] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[5] R. van Leeuwen and E. J. Baerends, Phys. Rev. A **49**, 2421 (1994).

[6] A. Görling, Phys. Rev. Lett. **83**, 5459 (1999).

[7] H. Jiang and Y.-C. Shen, J. Chem. Phys. **139**, 164114 (2013).

[8] R. T. Sharp and G. K. Horton, Phys. Rev. **90**, 317 (1953).

[9] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).

[10] For a detailed review of the OEP method, see S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).

[11] E. Engel, J. Chem. Phys. **140**, 18A505 (2014).

[12] L.-H. Ye, Phys. Rev. B **92**, 115132 (2015), and the references therein.

[13] Strictly speaking, OEP is SI free only at the one-electron level. Higher-order SI may exist at the N-electron level, and its effects on the IPs remain to be seen. For more information about the N-electron SI, see P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **125**, 201102 (2006).

[14] Y. Li, J. B. Krieger, and G. J. Iafrate, Phys. Rev. A **47**, 165 (1993).

[15] T. Grabo and E. K. U. Gross, Int. J. Quantum Chem. **64**, 95 (1997).

[16] An accurate IP of graphene can be straightforwardly obtained from the energy of the Dirac point measured from the reference value of the Coulomb potential. It is not calculated from Eq. (1).

[17] T. Heaton-Burgess, F. A. Bulat, and W. Yang, Phys. Rev. Lett. **98**, 256401 (2007).

[18] S. Hirata, S. Ivanov, I. Grabowski, R. J. Bartlett, K. Burke, and J. D. Talman, J. Chem. Phys. **115**, 1635 (2001).

[19] M. Betzinger, C. Friedrich, S. Blügel, and A. Görling, Phys. Rev. B **83**, 045105 (2011).

[20] A. D. Becke and E. R. Johnson, J. Chem. Phys. **124**, 221101 (2006).

[21] H. Krakauer, M. Posternak, and A. J. Freeman, Bull. Am. Phys. Soc. **23**, 258 (1978); M. Posternak, H. Krakauer, A. J. Freeman, and D. D. Koelling, *ibid.* **23**, 259 (1978); H. Krakauer, M. Posternak, and A. J. Freeman, Phys. Rev. B **19**, 1706 (1979).

[22] O. Jepsen, J. Madsen, and O. K. Andersen, Phys. Rev. B **18**, 605 (1978).

[23] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).

[24] T. F. T. Cerqueira, M. J. T. Oliveira, and M. A. L. Marques, J. Chem. Theory Comput. **10**, 5625 (2014).

[25] The choice of LDA over GGA is purely technical. Since SI typically behaves similarly in LDA and GGA, we expect a similar performance for the IPs if BJ06c is amended by a GGA correlation.

[26] C. J. Fall, N. Binggeli, and A. Baldereschi, J. Phys. Condens. Matter **11**, 2689 (1999).

[27] A. Grüneis, G. Kresse, Y. Hinuma, and F. Oba, Phys. Rev. Lett. **112**, 096401 (2014).

[28] W. Chen and A. Pasquarello, Phys. Rev. B **90**, 165133 (2014).

[29] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).

[30] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[31] Albeit being an insulator, ZnO is well known for its persistent and mysterious n-type conductivity. It is possible that the IP measurement suffers from a strong surface charging effect.

[32] I. D. White, R. W. Godby, M. M. Rieger, and R. J. Needs, Phys. Rev. Lett. **80**, 4265 (1998).

[33] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006).