

**Limitations of the DFT-1/2 method for covalent semiconductors and transition-metal oxides**

Jan Doumont,\* Fabien Tran, and Peter Blaha

*Institute of Materials Chemistry, Vienna University of Technology, Getreidemarkt 9/165-TC, A-1060 Vienna, Austria*

(Received 15 January 2019; published 1 March 2019)

The DFT-1/2 method in density functional theory [L. G. Ferreira *et al.*, *Phys. Rev. B* **78**, 125116 (2008)] aims to provide accurate band gaps at the computational cost of semilocal calculations. The method has shown promise in a large number of cases, however some of its limitations or ambiguities on how to apply it to covalent semiconductors have been pointed out recently [K.-H. Xue *et al.*, *Comput. Mater. Science* **153**, 493 (2018)]. In this work, we investigate in detail some of the problems of the DFT-1/2 method with a focus on two classes of materials: covalently bonded semiconductors and transition-metal oxides. We argue for caution in the application of DFT-1/2 to these materials, and the condition to get an improved band gap is a spatial separation of the orbitals at the valence band maximum and conduction band minimum.

DOI: [10.1103/PhysRevB.99.115101](https://doi.org/10.1103/PhysRevB.99.115101)**I. INTRODUCTION**

The calculation of the fundamental band gap of solids in Kohn-Sham (KS) density functional theory [1,2] (DFT) is a long standing problem [3]. The reason is that the exchange-correlation functional of the local density approximation [2] (LDA) severely underestimates band gaps by typically 50–100% [3], and the standard functionals of the generalized gradient approximation (GGA) [4] do not perform much better [5]. The current state-of-the-art in band gap calculations is Hedin's *GW* method [6,7], but it goes beyond DFT and is computationally very demanding especially if applied self-consistently [8]. Within the generalized Kohn-Sham (gKS) scheme [9] (i.e., with nonmultiplicative potentials), hybrid functionals, which mix LDA/GGA functionals with exact exchange [10], do offer greatly improved band gaps [5], but at a computational cost that is also much higher (by one or two orders of magnitude) than LDA/GGA functionals. The meta-GGA (MGGA) approximation [11], which is also of the semilocal type and therefore computationally fast, is a very promising route for improving band gaps within the gKS framework at a modest cost. The MGGA functionals that have been developed so far are however not as accurate as the hybrid or *GW* methods [12–14].

Nevertheless, within the true KS-DFT scheme, i.e., with a multiplicative potential, computationally fast DFT methods have been developed for band gap calculations, like the functional of Armiento and Kümmel [15,16], the potential of Gritsenko *et al.* [17,18] (GLLB), or the modified Becke-Johnson potential [19] (mBJ), the latter being as accurate as the very expensive hybrid or *GW* methods.

Another fast method designed for band gaps is DFT-1/2 [20], which is an application of Slater's half-occupation (transition state) technique [21,22] to periodic solids. It only requires the addition of a self-energy correction potential, calculated from a half-ionized free atom, to the usual KS-DFT

potential (see Sec. II for details). The method has been shown to perform quite well for a number of test sets [23–25] and has been evaluated as a good starting point for  $G_0W_0$  calculations [26]. For instance, an application to metal halide perovskites has found comparable accuracy to *GW* [27]. Thanks to its low computational cost, DFT-1/2 has been regularly applied to systems that require larger unit cells. A study of the negatively charged nitrogen-vacancy ( $NV^-$ ) center in diamond has been performed with a generalized version of DFT-1/2, which is suited not only for band gap but also optical transitions and defect levels [28]. Other applications include studies of doped materials [29,30], heterostructures [31,32], surfaces [33,34], or interfaces [35,36]. Also, in a study of semiconducting indium alloys comparing the DFT-1/2 method with hybrid functionals, it was found that, although the hybrid functionals were slightly more accurate, DFT-1/2 allows for larger supercells and consequently better convergence of the bowing parameter [37]. Another comparative study of DFT-1/2 to the pseudo-self-interaction-corrected approach to DFT was performed on fluorides [38]. Furthermore, a few magnetic systems have been studied, namely GaMnAs [39] and InN doped with Cr [30]. We also mention that the method has recently been applied successfully for the calculation of the ionization potential of atoms and molecules [40].

However, the limitations of the method have not been given much consideration until recently [41]. These limitations stem from the fact that the correction applied in DFT-1/2 has an atomic origin. One of them is the application of the method to covalently bonded semiconductors. Originally, it was argued that group IV semiconductors (diamond, Si, and Ge) need a modified correction that is calculated from a 1/4-ionized atom instead of a 1/2-ionized atom, the argument being that valence band holes of neighboring atoms overlap [20]. In III-V compounds (GaAs, AlP, ...) it is claimed that the valence band hole resembles more closely the photoionization hole in the atom, such that the standard 1/2 ionization is justified [20].

As shown and discussed in detail in this work, another limitation of the DFT-1/2 method is that it performs very poorly for transition-metal (TM) oxides. Many of these materials

\*Corresponding author: [jan.doumont@tuwien.ac.at](mailto:jan.doumont@tuwien.ac.at)

are Mott insulators, where both the highest occupied band and the lowest conduction band have strong TM  $d$ -orbital characters which differ only by their angular shape, such that the spherical atomic DFT-1/2 correction cannot work efficiently for the band gap. The focus of the present work will be on the problems of the DFT-1/2 method mentioned above, namely the ambiguity about the ionization of the free atom to calculate the correction potential in semiconductors and the limited applicability of DFT-1/2 for TM oxides.

The paper is organized as follows. Section II provides a description of the methods and the computational details, while the results are presented and discussed in Sec. III. Finally, Sec. IV gives the summary of this work.

## II. THEORY

In KS-DFT, the so-called KS band gap  $E_g^{\text{KS}}$  is defined as the difference between the KS eigenvalues of the highest occupied  $[\epsilon_N(N)]$  and lowest unoccupied  $[\epsilon_{N+1}(N)]$  orbitals of the  $N$ -electron system. On the other hand, the fundamental band gap  $E_g$ , the physical many-body property one is interested in, is defined as the ionization potential  $I(N)$  minus the electron affinity  $A(N)$  and can be expressed in terms of the (exact) KS eigenvalues of the highest occupied orbitals of the  $N$ - and  $N + 1$ -electron systems [3,42]:

$$\begin{aligned} E_g &= I(N) - A(N) \\ &= -\epsilon_N(N) - (-\epsilon_{N+1}(N+1)) \\ &= \underbrace{\epsilon_{N+1}(N) - \epsilon_N(N)}_{E_g^{\text{KS}}} + \underbrace{\epsilon_{N+1}(N+1) - \epsilon_{N+1}(N)}_{\Delta_{\text{xc}}} \\ &= E_g^{\text{KS}} + \Delta_{\text{xc}}, \end{aligned} \quad (1)$$

where  $\Delta_{\text{xc}}$  is the discontinuity of the exchange-correlation potential at integer values of the number of electrons  $N$ . In KS-DFT calculations employing LDA or GGA functionals, this discontinuity is not captured [43] (but can be calculated by some means in finite systems [44–46]). From Eq. (1), it is clear that a good estimation of the true band gap can, in principle, not be obtained by considering  $E_g^{\text{KS}}$  alone in particular since  $\Delta_{\text{xc}}$  can be of the same order of magnitude as the band gap itself [47,48].

The DFT-1/2 technique aims to correct the band gap problem by adapting Slater's atomic transition state technique to periodic solids. Starting from Janak's theorem [49],

$$\frac{\partial E(f_\alpha)}{\partial f_\alpha} = \epsilon_\alpha(f_\alpha), \quad (2)$$

where  $E(f_\alpha)$  is the total energy of the system and  $f_\alpha$  is the occupation number of orbital  $\phi_\alpha$  relative to the neutral atom ( $f_\alpha = 0$ ), and using the midpoint rule for integrating the right-hand side of Eq. (2), it is trivial to show that the KS eigenvalue for the 1/2-ionized (hence transition) state can be used to calculate the ionization potential of the atom:

$$E(0) - E(-1) \simeq \epsilon_\alpha(-1/2). \quad (3)$$

In order to benefit from Eq. (3) for self-consistent DFT calculations in solids [20,24], a self-energy correction potential  $v_S$  is defined by rewriting the ionization potential the

following way:

$$E(0) - E(-1) = \epsilon_\alpha(0) - \int d^3\mathbf{r} \rho_\alpha(\mathbf{r}) v_S(\mathbf{r}), \quad (4)$$

with  $\rho_\alpha = |\phi_\alpha|^2$  and where  $v_S$  is chosen such that

$$\int d^3\mathbf{r} \rho_\alpha(\mathbf{r}) v_S(\mathbf{r}) \simeq \epsilon_\alpha(0) - \epsilon_\alpha(-1/2) \quad (5)$$

and therefore Eq. (3) is satisfied. Equation (5) shows that adding  $-v_S$  to the effective KS potential  $v_{\text{KS}}$  in a calculation should shift the eigenvalue of orbital  $\alpha$  by  $\epsilon_\alpha(-1/2) - \epsilon_\alpha(0)$  and therefore bring it close to  $\epsilon_\alpha(-1/2)$ , i.e., the ionization potential according to Eq. (3). In practice, the potential  $v_S$  is not obtained from calculations on the solid but on an isolated atom (the one where the orbital  $\alpha$  is mostly located) [20]:

$$v_S = v_{\text{KS}}^{\text{atom}}(f_\alpha = 0) - v_{\text{KS}}^{\text{atom}}(f_\alpha = -1/2), \quad (6)$$

where  $v_{\text{KS}}^{\text{atom}}$  are the KS effective potentials obtained at the end of self-consistent calculations in the neutral and 1/2-ionized states.

Concretely, the DFT-1/2 method consists, first, of two self-consistent calculations on the free atom to calculate  $v_S$  with Eq. (6), and the orbital  $\phi_\alpha$  that is chosen to be ionized is the one that is supposed to contribute the most to the valence band maximum (VBM) in the solid. Then, this atomic potential  $v_S$  is added to the usual LDA or GGA effective KS potential  $v_{\text{KS}}$  for the self-consistent calculation on the solid. However, before  $v_S$  is added to  $v_{\text{KS}}$ , it must be multiplied by a spherical step function

$$\Theta(r) = \begin{cases} \left(1 - \left(\frac{r}{r_c}\right)^8\right)^3 & r \leq r_c \\ 0 & r > r_c \end{cases}. \quad (7)$$

because  $v_S$  falls off only like  $1/r$  at long range which causes divergence when summed over the lattice. The cutoff radius  $r_c$  is the only parameter introduced in the method and is determined variationally by maximizing the band gap [20].

As argued in Ref. [24], the correction to the KS band gap due to  $v_S$  can be somehow identified to the discontinuity  $\Delta_{\text{xc}}$  in Eq. (1) (although it is questionable since the potential is still multiplicative [43,50]). However, we mention that no correction was applied to the conduction band minimum (CBM). As reported in Ref. [24], such correction should affect only little the unoccupied states due to their more delocalized nature.

A few extensions or refinements to the method have been proposed. The shell correction from Ref. [41] uses a step function with an additional (inner) radius to improve the accuracy and will be discussed in detail in Sec. III D. In other works [51,52], an empirical amplification factor (which multiplies  $v_S$  by a constant) to fit experiment was used. In Ref. [51], nonstandard ionization levels for the correction potential (other than 1/4 or 1/2) have been used. The character of the atomic orbital contributions to the VBM is used to determine the ionization levels (normalized to 1/2 across both atomic species). In Ref. [28] a generalization of DFT-1/2 also suited for optical transition levels (including adding self-energy correction to the excited band and nonstandard ionization levels) has been applied to the  $\text{NV}^-$  center of diamond.

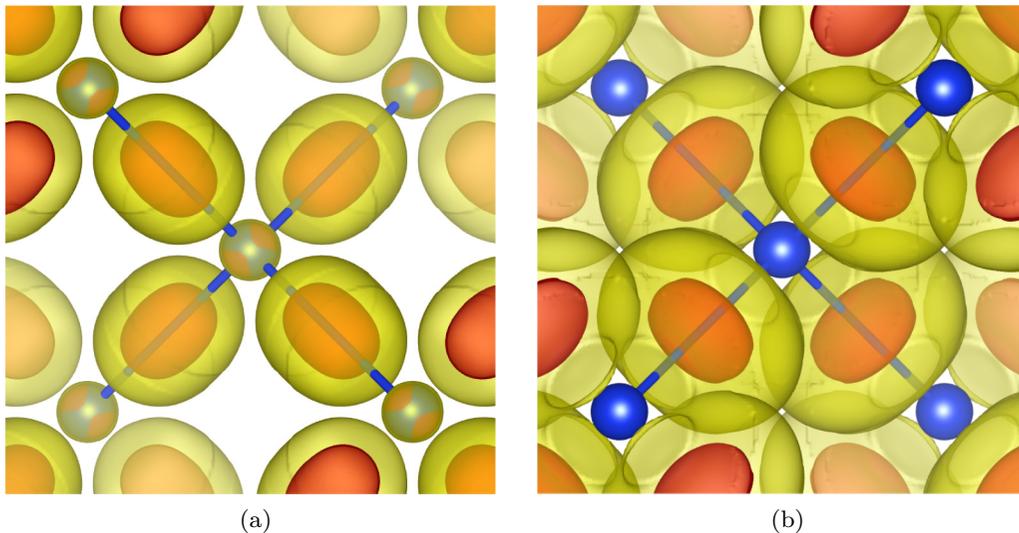


FIG. 1. Illustration of the “overlapping holes” argument in silicon. Isosurfaces are chosen for best visibility, in both cases with a lower value in yellow and a higher one in red. The electron density close to the VBM (a) sits mostly around the bond centers. The correction potential  $v_S$  (using PBE-1/2) (b) is the largest in the middle of the Si-Si bonds, since the sum of the cutoff radii  $r_c$  is larger than the nearest-neighbor distance, even though the  $v_S$  of the individual atoms is spherically distributed.

For the present work, the DFT-1/2 method has been implemented into the all-electron WIEN2K [53] code which is based on the linearized-augmented plane-wave (LAPW) method [54,55]. The implementation is very similar to the one reported recently [25] in EXCITING which is also an LAPW-based code. The calculations were done at the experimental lattice parameters (specified in Table S1 of Ref. [56]) for all compounds. A dense  $24 \times 24 \times 24$   $\mathbf{k}$  mesh was used for all cubic solids, while for other structures a proportional mesh with 24  $\mathbf{k}$  points along the direction corresponding to the shortest lattice constant was used. For some of the TM oxides [notably those with antiferromagnetic (AFM) ordering, which have larger unit cells], a less dense  $\mathbf{k}$  mesh was used, but care was taken that convergence is reached. The same applies to the basis set size. For all compounds containing Ga or heavier atoms, the calculations were done with spin-orbit coupling included. The cutoff radius  $r_c$  in Eq. (7) is optimized using a multidimensional search with a precision of 0.01 eV, which corresponds to a precision in  $r_c$  of about  $0.05 a_0$  (the band gap is not very sensitive to  $r_c$  close to the extremum). Furthermore, the optimal cutoff radii of different atoms in binary compounds are to a large extent independent [20]. LDA and GGA [using the functional of Perdew *et al.* [4] (PBE)] calculations were done with and without the 1/2 correction. For comparison purposes, calculations with the mBJ potential [19,57,58], which has been shown to be the most accurate semilocal potential for band gap calculations and is even superior to hybrid functionals [57–60], will also be reported.

### III. RESULTS

#### A. Group IV and III-V semiconductors

We start by mentioning that how to apply the DFT-1/2 method to the group IV semiconductors C, Si, and Ge is unclear. In contrast to binary compounds, the self-energy correction potential  $v_S$  has not always been calculated from 1/2-ionized free atoms but from 1/4-ionized ones [i.e., with

$f_\alpha = -1/4$  in the second term of Eq. (6)]. Actually, for diamond  $v_S$  was calculated in Refs. [20,24] by ionizing both the  $p$  and  $s$  bands by a 1/4-electron charge (in total, removing half an electron), whereas for Si and Ge only the  $p$  band receives a 1/4-ionization correction. The argument behind this is that the orbital at the VBM overlaps with the correction potential of both atoms in the unit cell, such that only a 1/4 electron should be removed on each atom to avoid a correction that is too large. This is illustrated for Si in Fig. 1, where we can see that  $v_S$  is the largest at the Si-Si bond center.

Turning to our DFT-1/2 calculations, Table I shows the results for a set of covalent semiconductors that were obtained with a 1/2- or 1/4-ionization correction. Furthermore, both LDA and PBE were considered for the underlying semilocal functional. All atoms were corrected and the ionized orbital is the one with the largest contribution to the VBM. For SiC and AIP, an additional calculation was done where the correction is applied only to the anion.

Indeed, we can see that the band gaps obtained using a 1/4-ionization correction (i.e., LDA-1/4 and PBE-1/4) are very accurate for Si and SiC, since the values differ by at most 0.2 eV compared to experiment, while using a 1/2-ionization correction (i.e., LDA-1/2 and PBE-1/2) leads to overestimations of at least 0.8 eV. For diamond, the results show that using a 1/2-ionized (1/4-ionized) correction leads to an overestimation (underestimation) of about 0.5 eV. For Ge, the experimental gap of 0.74 eV lies above the LDA-1/2 and PBE-1/2 values by about 0.4 and 0.2 eV, respectively, while using a 1/4-ionized correction leads to strongly underestimated values. Note the contrast between Si and Ge which require different ionization, despite having relatively similar valence band density and optimized cutoff radius  $r_c$  in Eq. (7).

Another issue that may arise is the ambiguity in choosing the atom(s) and/or orbital(s) on which the correction should be applied. For instance in the case of binary semiconductors, it has been claimed [20] that in most cases (but not always) only the correction on the anion has an impact on the results.

TABLE I. Band gaps (in eV) of groups IV and III-V semiconductors calculated using the DFT-1/2 method with different underlying functionals (LDA or PBE) and ionization degrees (1/2 or 1/4). The orbitals that were ionized and the cutoff radii  $r_c$  (in  $a_0$ ) in Eq. (7) are indicated in the second and third columns, respectively. Only the cutoff radii from LDA-1/2 calculations are shown, however we checked that those for PBE-1/2, LDA-1/4, and PBE-1/4 calculations are practically identical (the difference is below  $0.05 a_0$ , which does not impact the band gap). A dash “-” in the orbitals column indicates a noncorrected atom. Literature results (all based on LDA) are also given and are marked according to the reference/code: (S) for SIESTA and (V) for VASP calculations [20], (E) for EXCITING calculations [25], and (X) for VASP calculations [41]. When the correction details of the literature results do not correspond to ours (which atoms, orbitals are corrected), they are specified in parenthesis. Note that in Ref. [41], the calculations were done at the LDA lattice constant. For comparison purposes, LDA, PBE, and mBJ results are also shown. The experimental results are from Refs. [61] and [62]. The MAE (in eV) and MARE (in %) are calculated when the correction is applied on all atoms and for an ionized  $p$  orbital in Ga (as deduced from a partial charge analysis at VBM). The most accurate values among the DFT-1/2 methods are underlined.

Solid	Orbitals	$r_c$	LDA	LDA-1/4	LDA-1/2	PBE	PBE-1/4	PBE-1/2	mBJ	Other works	Expt.
C	$p$	2.41	4.10	4.95	<u>5.82</u>	4.14	5.04	5.95	4.92	5.25 (S,1/4- $sp$ )	5.50
Si	$p$	3.78	0.47	<u>1.20</u>	1.96	0.57	1.35	2.16	1.15	1.21 (S,1/4)	1.17
SiC	$p, p$	2.80, 3.00	1.31	<u>2.31</u>	3.40	1.35	<u>2.43</u>	3.59	2.25	2.32 (E) <sup>a</sup>	2.42
SiC	$-, p$	3.00	1.31	2.19	3.16	1.35	<u>2.28</u>	3.31	2.25		2.42
Ge	$p$	4.11	metal	0.04	0.37	metal	0.27	<u>0.59</u>	0.76	0.27 (X,1/4)	0.74
BN	$p, p$	2.41, 2.47	4.35	5.24	<u>6.78</u>	4.47	5.79	7.06	5.80		6.36
BP	$p, p$	3.17, 3.09	1.18	<u>1.97</u>	2.79	1.24	2.48	2.95	1.85		2.10
BA <sub>s</sub>	$p, p$	3.22, 3.25	1.04	<u>1.82</u>	2.62	1.09	1.93	2.77	1.58		1.46
AlN	$p, p$	2.91, 2.92	3.25	4.48	5.80	3.34	<u>4.66</u>	6.08	4.88		4.90
AIP	$p, p$	3.74, 3.69	1.45	2.29	3.15	1.59	<u>2.50</u>	3.43	2.31	2.23 (X,1/4)	2.5
AIP	$-, p$	3.69	1.45	2.18	2.95	1.59	<u>2.38</u>	3.21	2.31	2.96 (E) <sup>a</sup>	2.5
AlAs	$p, p$	4.39, 3.88	1.25	2.08	2.92	1.35	<u>2.26</u>	3.17	2.05	2.73 (V,1/2-As)	2.23
AlSb	$p, p$	6.28, 4.11	0.89	1.51	2.17	0.97	<u>1.62</u>	2.30	1.51	1.97 (X,1/2-Sb)	1.69
GaN	$p, p$	1.30, 3.00	1.66	2.50	<u>3.42</u>	1.66	<u>2.55</u>	3.53	2.86		3.28
GaN	$d, p$	1.30, 3.00	1.66	2.56	<u>3.54</u>	1.66	2.61	3.66	2.86	3.56 (E), <sup>a</sup> 3.52 (V,1/2)	3.28
GaP	$p, p$	1.15, 3.50	1.41	2.00	<u>2.51</u>	1.57	<u>2.21</u>	2.80	2.22	2.57 (X,1/2-P)	2.35
GaAs	$p, p$	1.14, 3.82	0.19	0.71	1.27	0.43	0.97	<u>1.57</u>	1.54	1.41 (V,1/2-As)	1.52
GaAs <sup>b</sup>	$d, p$	1.15, 3.83	0.30	0.86	<u>1.47</u>	0.54	1.13	1.77	1.64	1.46 (E) <sup>a</sup>	1.52
GaSb	$p, p$	1.14, 4.19	metal	0.05	0.48	metal	0.30	<u>0.75</u>	0.74	0.67 (X,1/2-Sb)	0.82
MAE			1.10	0.44	0.56	1.02	0.32	0.67	0.20		
MARE			52	25	30	48	19	32	7		

<sup>a</sup>Reference [25] does not provide details about the ionization correction, but the very close agreement with one of our results indicates which ionization correction was applied.

<sup>b</sup>Calculation without spin-orbit coupling for comparison with the result from other work. The effect of the spin-orbit coupling is to reduce the band gap by about 0.1 eV.

While this may be true for ionic solids, where the states at the VBM come only from the anion, such a choice cannot always be justified in the case of binary semiconductors where both atoms may contribute to the VBM. Thus, in addition to the degree of ionization correction (e.g., 1/4 or 1/2), it may not always be clear on which atoms the potential  $v_S$  should be applied. Since in SiC the VBM has a dominant  $p$ -orbital character from the C atom, we did an alternative calculation where the correction is applied only to the C atom. Compared to the usual procedure where the orbitals on all atoms are corrected, a reduction of the band gap by 0.1 eV to 0.3 eV is observed. Good agreement with experiment is obtained with 1/4-ionized correction (even though there is very little correction potential overlap at the VBM in this case), while a 1/2-ionized correction leads to large overestimations of  $\sim 1$  eV similar to Si.

Considering the III-V compounds, we see that LDA-1/2 and PBE-1/2 clearly overestimate the band gaps for the BX and AlX compounds, while a moderate overestimation is observed for GaN and GaP. On the other hand, PBE-1/2

performs very well for GaAs and GaSb since the error is below 0.1 eV.

On average, PBE-1/4 is the most accurate of the DFT-1/2 methods for this test set. It provides in eight cases the best agreement with experiment and leads to a MAE of only 0.32 eV; this is half of the one for PBE-1/2 (0.67 eV) which is the worst of the DFT-1/2 methods. However, note that the mBJ potential which has MAE of 0.20 eV and MARE of 7% is clearly more accurate. In comparison, LDA and PBE lead to MAE that are around 1 eV. The general observation is that a 1/4 ionization is more appropriate for the light systems but not sufficient for the heavier ones, i.e., those with Ga or Ge atoms, for which a 1/2-ionization correction, either with LDA or PBE, is usually more suitable. Nevertheless, a few borderline cases are C, BN, and GaP, where the best correction also depends on the underlying semilocal functional. We also mention that for only one system (GaN) there is no overlap (loosely defined as whether the sum of the cutoff radii of two nearest-neighboring atoms is larger than their distance) between the correction potentials  $v_S$ , while

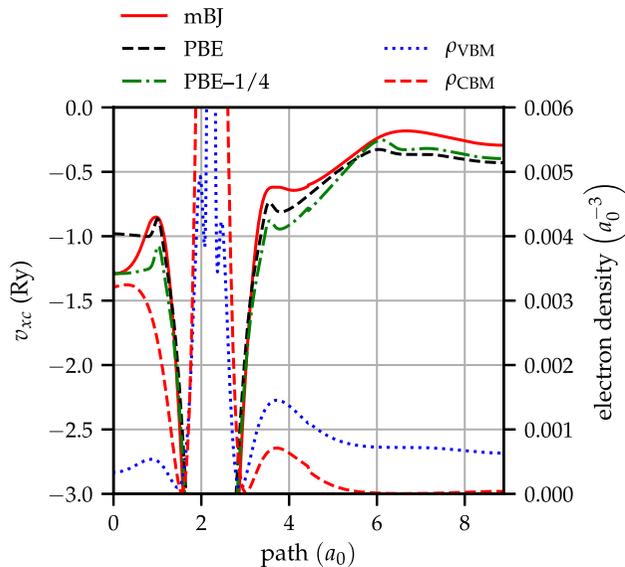


FIG. 2. Plots of mBJ, PBE, and PBE-1/4 exchange-correlation potentials  $v_{xc}$  in Si and densities of the VBM and CBM. The path is from (0, 0, 0) to (1/2, 1/2, 1/2) in the unit cell fractional coordinates, thus starting at the midpoint of the Si-Si bond, passing through the atom at (1/8, 1/8, 1/8) and terminating at the middle of the unit cell. The densities are taken from the PBE calculation.

for the other Ga compounds the overlap is small (tenths of one  $a_0$ , compared to an overlap of  $3a_0$  to  $4a_0$  in Si and Ge).

In Fig. 2, the exchange-correlation potentials  $v_{xc}$  mBJ, PBE, and PBE-1/4 in Si are compared. The band gaps from mBJ (1.15 eV) and PBE-1/4 (1.35 eV) are relatively close to each other, but the corresponding potentials show noticeable differences. Compared to PBE, PBE-1/4 corrects the band gap by lowering the energy in the region where the VBM density  $\rho_{\text{VBM}}$  is very large (in the region within  $2a_0$  from the atom), whereas mBJ has a smaller correction. On the other hand, at the CBM mBJ leads to a larger upshift than PBE-1/4.

Concerning the orbital to which the ionization should be applied, the Ga compounds are interesting since they are not always treated the same way. For some reported calculations [20,39], the  $d$  orbital was ionized for all Ga compounds, while in Ref. [35], the Ga  $p$  orbital in GaAs was ionized as deduced from a partial charge analysis at the VBM.

In order to find which orbital should be corrected, we used the new PES (photo-electron spectrum) module in WIEN2K [63]. Using this module, we can decompose the interstitial charge into their atomic orbital contributions and get atomic partial charges uniquely and independently on the choice of the atomic sphere radii and the localization of different orbitals. For instance in GaN only 16.2% of the Ga- $4p$  charge, but 97.5% of Ga- $3d$  charge are enclosed inside the atomic sphere, and thus the Ga- $3d$  charge dominates over Ga- $4p$  when considering the charges within the atomic sphere. However, the rescaled orbital character contributions at the VBM are 12.1% and 8.2% of Ga- $4p$  and Ga- $3d$ , respectively, and 79.8% of N- $2p$ . For the heavier GaX compounds, we find progressively larger Ga- $p$  and smaller Ga- $d$  character contributions at the VBM. Thus, that means that a proper

ionization correction for the Ga compounds should be applied to the Ga- $4p$  orbital.

The comparison of our calculations to those found in literature needs to be done carefully, because the correction potential  $v_S$  is not always calculated the same way (e.g., 1/2- or 1/4-ionization and on which atoms) and, furthermore, the details are not always specified. For instance, our LDA-1/4 result for Si agrees perfectly with the one from Ferreira *et al.* [20] while in this same work C was corrected with a 1/4-ionization for both  $p$  and  $s$  orbitals, leading to a value of 5.25 eV that differs substantially from our result even when we use the same ionization scheme (5.87 eV, which is very close to 5.82 eV with LDA-1/2). This discrepancy for C is unclear.

The comparison with the results from the recent implementation of the DFT-1/2 method in the LAPW EXCITING code [25] shows perfect agreement, but it also shows the importance of knowing the exact correction procedure, since for AlP the agreement is obtained if the correction is applied only to the P atom, although we cannot be sure that this scheme was used in Ref. [25]. However, for GaN and GaAs our results in Table I (without spin-orbit coupling for GaAs) show that agreement with those of Pela *et al.* [25] is only obtained if the Ga  $d$  orbitals (and not the  $p$  orbitals) and N/As  $p$  orbitals are corrected, although in GaAs the Ga  $p$  orbital contributes non-negligibly to the VBM. Thus, these examples show that for a meaningful comparison of results of two sets of DFT-1/2 calculations, one needs to know the details of the calculations, since depending on the ionization correction (1/4 or 1/2) and on which atoms/orbitals it is applied, a sizable variation in the results can be obtained.

In general, a more valid explanation for some of the overestimations found in covalent materials is that these are not necessarily due to overlapping holes but simply due to the fact that the assumptions used in deriving the method (see Sec. II) may be too crude. The larger the difference between the VBM density and the corresponding atomic density (from which the self-correction potential is calculated) is, the worse the DFT-1/2 method should perform. This is illustrated with the case of BAs, where even the 1/4-ionization correction clearly overestimates the experimental band gap. The VBM in BAs has very little pure atomic character but is strongly  $sp^3$  hybridized and thus very aspherical, as seen in Fig. 3. The asphericity in the valence distribution causes an overestimation of the band gap, because the matrix element of  $v_S$  [Eq. (5)] will be too large when the charge distribution is spread out compared to the nonhybridized atomic case. In many cases, this will also cause an overlap, but not always (see for example BeTe below).

## B. Be compounds

An interesting case study for the DFT-1/2 method that has not been considered previously consists of the Be compounds BeO, BeS, BeSe, and BeTe, where the first one has the wurtzite structure, while the others have the zincblende structure. We chose these compounds to investigate the behavior of DFT-1/2 because of the descending order of ionicity along the series. The results for the band gap can be found in Table II where we can see that the standard 1/2-ionization

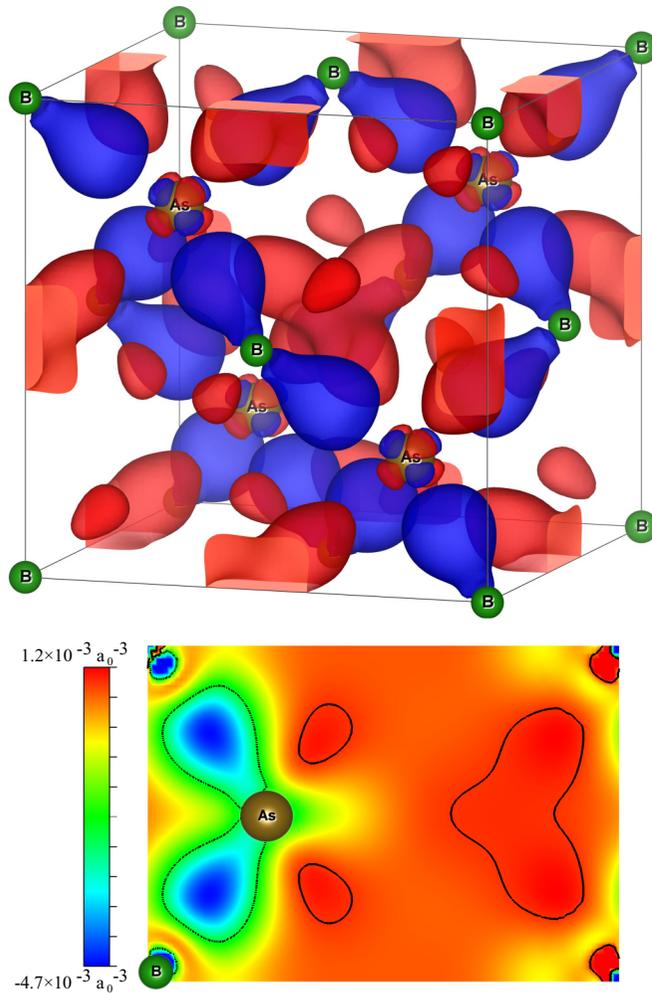


FIG. 3. Three-dimensional (upper panel) and two-dimensional (lower panel) plots of the electron density difference CBM – VBM in BAs. On the three-dimensional plot, the positive (CBM, in red) and negative (VBM, in blue) isosurfaces are defined at  $0.9 \times 10^{-3} a_0^{-3}$  and  $-2 \times 10^{-3} a_0^{-3}$ , respectively. On the two-dimensional plot, the slice corresponds to a  $1\bar{1}0$  plane with B atoms at the corners and an As atom at the left center, and the black contour lines correspond to the isosurfaces on the three-dimensional plot.

correction is the most effective for the first three compounds. PBE–1/2, for instance, yields a band gap of 10.32 eV for BeO, which is within a few percent of the experimental

TABLE II. Band gaps (in eV) of Be compounds calculated using the DFT–1/2 method with different underlying functionals (LDA or PBE) and ionization degrees (1/2 or 1/4). In all cases the ionization was applied to the Be  $s$  and anion  $p$  orbitals. The cutoff radii  $r_c$  (in  $a_0$ ) in Eq. (7) are indicated in the second column. Only the cutoff radii from LDA–1/2 calculations are shown, however we checked that those for PBE–1/2, LDA–1/4, and PBE–1/4 calculations are practically identical (the difference is below  $0.05 a_0$ , which does not impact the band gap). For comparison, LDA, PBE, and mBJ results are also shown. The experimental and  $G_0W_0$  results are from Refs. [59,64,65]. The most accurate values among the DFT–1/2 methods are underlined.

Solid	$r_c$	LDA	LDA–1/4	LDA–1/2	PBE	PBE–1/4	PBE–1/2	mBJ	$G_0W_0$	Expt.
BeO	0.00, 2.52	7.49	8.71	10.05	7.57	8.88	<u>10.32</u>	9.58		10.60
BeS	0.44, 3.28	2.92	3.74	4.60	3.13	4.01	<u>4.94</u>	4.13	4.92	> 5.5
BeSe	0.48, 3.44	2.34	3.14	4.00	2.51	3.36	<u>4.26</u>	3.39	4.19	4.0–4.5
BeTe	0.33, 3.79	1.57	2.25	<u>2.97</u>	1.69	<u>2.41</u>	3.17	2.33		2.7

value of 10.60 eV, and band gaps for BeS (4.94 eV) and BeSe (4.26 eV) that coincide with  $G_0W_0$  results. For BeTe, the PBE–1/2 value is too large by 0.5 eV, while the errors of  $\sim 0.3$  eV with PBE–1/4 and LDA–1/2 are somewhat smaller.

In order to investigate the difference between, e.g., BeSe and BeTe, we now consider the PBE and PBE–1/2 band structures as well as the electron density close to the Fermi energy. The band structures for both compounds (see Fig. 4) show a very similar change when the 1/2-ionization correction is applied. Compared to PBE, the gap separating the valence and conduction bands is larger and the bands are more flat. The shift of the bands is not uniform, but no dramatic change in the shape of the bands is induced.

Figure 5 shows plots of electron density difference between the VBM and the CBM that are calculated in a small energy range above the CBM and below the VBM, respectively, while ensuring that the total charge in both cases is equal. Two isosurfaces are shown: one in red with a positive sign (corresponding to the CBM) and one in blue (corresponding to the VBM). In the case of the Be compounds this is almost equivalent to simply superposing both densities in different colors, because both are well separated spatially (which is not always true, see Sec. III C below). The main observation is that there are no distinctive features that could be used to clearly judge *a priori* which correction (1/2 or 1/4) would be most suitable. Moreover, in both cases the valence density is mostly distributed around the anion. This is reflected in the values of the cutoff radii  $r_c$  of Be, which in both cases is optimized to very small values (see Table II), such that the correction potential on the cation is therefore negligible. Thus, in the case of BeTe, overlapping holes cannot explain the overestimation of the band gap in PBE–1/2. Also, a partial charge analysis (again using the PES module) of the VBM reveals a nearly identical atomic  $p$ -orbital character of the anion in BeSe and BeTe of 96.5% and 94.3%, respectively.

### C. Transition-metal oxides

Another class of materials that provides a challenge for the DFT–1/2 method are TM oxides (see Refs. [51,68] for results for a few nonmagnetic TM oxides). Results for some representative nonmagnetic and AFM cases are shown in Table III. For all TM oxides the correction is based on a 1/2-ionization of the TM atom  $d$  orbital and O  $p$  orbital. For the AFM systems the self-energy correction potential  $v_S$  is spin

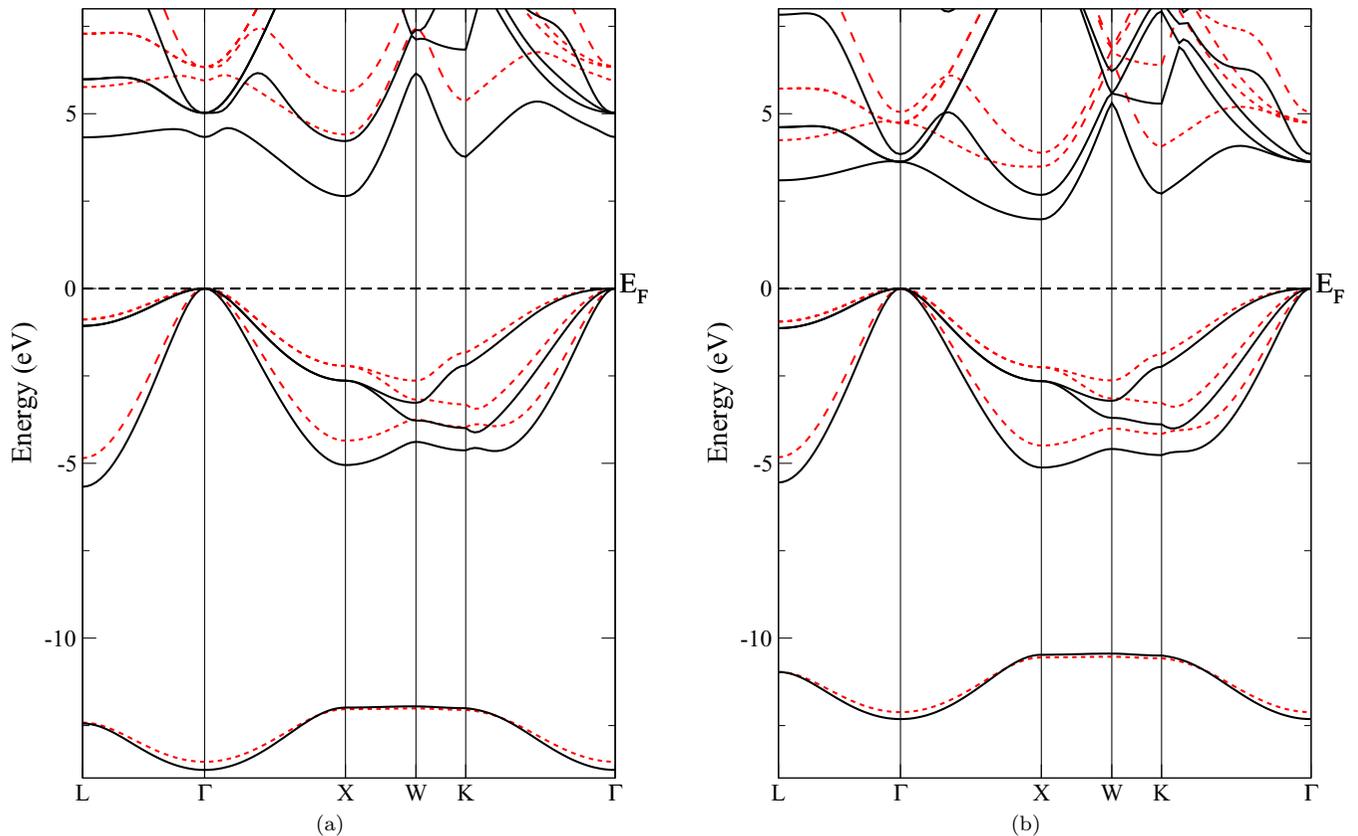


FIG. 4. PBE (black solid line) and PBE-1/2 (red dashed line) band structures for BeSe (a) and BeTe (b). For better visibility, we show band structures calculated without spin-orbit effects.

dependent (and respects the AFM ordering) and calculated from a 1/2-ionization of the spin with the largest contribution to the VBM.

From the results we can see that the DFT-1/2 band gaps for the nonmagnetic  $\text{TiO}_2$  and  $\text{ZnO}$  are much larger than LDA/PBE and very close to experiment (errors are below 0.2 eV). However, for all other systems the band gaps calculated using DFT-1/2 are still much smaller than experiment. Actually, the DFT-1/2 method works very well for  $\text{TiO}_2$  and  $\text{ZnO}$  since these two systems have a charge-transfer band gap and thus a clear spatial separation between the VBM and CBM. This is not the case for the other systems ( $\text{Cu}_2\text{O}$ ,  $\text{VO}_2$ , and the AFM strongly correlated systems), where a significant  $d$  character is present in both the VBM and CBM, such that the spherical atomic self-energy correction  $v_S$  can not really distinguish these split bands; any correction that is applied to the VBM will also influence the energy level of the CBM in a similar way and thus fails to increase the band gap as much as one would like. It is well known that the standard (semi)local functionals like PBE are not able to describe strongly correlated systems properly even at the qualitative level [69], and only more advanced methods like DFT +  $U$ , hybrid functionals, or mBJ can lead to reasonable results [58,70–73]. Note, however, that although mBJ performs better than DFT-1/2 overall, it fails for some cases, notably  $\text{TiO}_2$ ,  $\text{Cu}_2\text{O}$ , and  $\text{ZnO}$  (see Table III and Ref. [74]).

In more detail, for  $\text{MnO}$ ,  $\text{Fe}_2\text{O}_3$ , and  $\text{CuO}$ , DFT-1/2 leads to a clear (but not sufficient) improvement of at least 1 eV in

the band gap, which is, however, not as impressive as in  $\text{TiO}_2$  and  $\text{ZnO}$ . This has to be compared to the small improvement of a few tenths of an eV obtained for  $\text{NiO}$  and  $\text{Cr}_2\text{O}_3$  and a metallic character that persists for  $\text{FeO}$ ,  $\text{CoO}$ , and  $\text{VO}_2$ . For the latter three solids, we checked that the results are the same whatever is the electron density (e.g., PBE or mBJ) that is used to initialize the self-consistent field procedure. Actually, the common feature of  $\text{Fe}_2\text{O}_3$ ,  $\text{MnO}$ , and  $\text{CuO}$  is to have CBM and VBM that are made of states of opposite spins, as a consequence of the large exchange splitting [69]. Since the correction potential  $v_S$  is spin dependent (the ionization is done for the spin with the largest contribution to the VBM), a sizable increase in the band gap is possible. In other TM oxides like  $\text{NiO}$ ,  $\text{FeO}$ , or  $\text{CoO}$  the crystal field splitting is dominant, such that the VBM has a more mixed spin population.

In  $\text{CuO}$ , the choice for the spin for calculating the potential  $v_S$  is crucial. The calculation done with the ionization correction applied to the correct spin (i.e., the one which has the largest population at the VBM for the given atom) gives a PBE-1/2 result of 1.17 eV, whereas calculating  $v_S$  using the other spin results in a band gap of only 0.49 eV. Another particular feature of  $\text{CuO}$  is the cutoff radius  $r_c$  of the Cu atom, which is extremely large ( $5.42 a_0$ ). As shown in Ref. [41], the band gap as a function of  $r_c$  may consist of several maxima when  $r_c$  reaches the next coordination shell. Usually, one would expect the global maximum of the band gap with respect to  $r_c$  to be at the first maximum (see

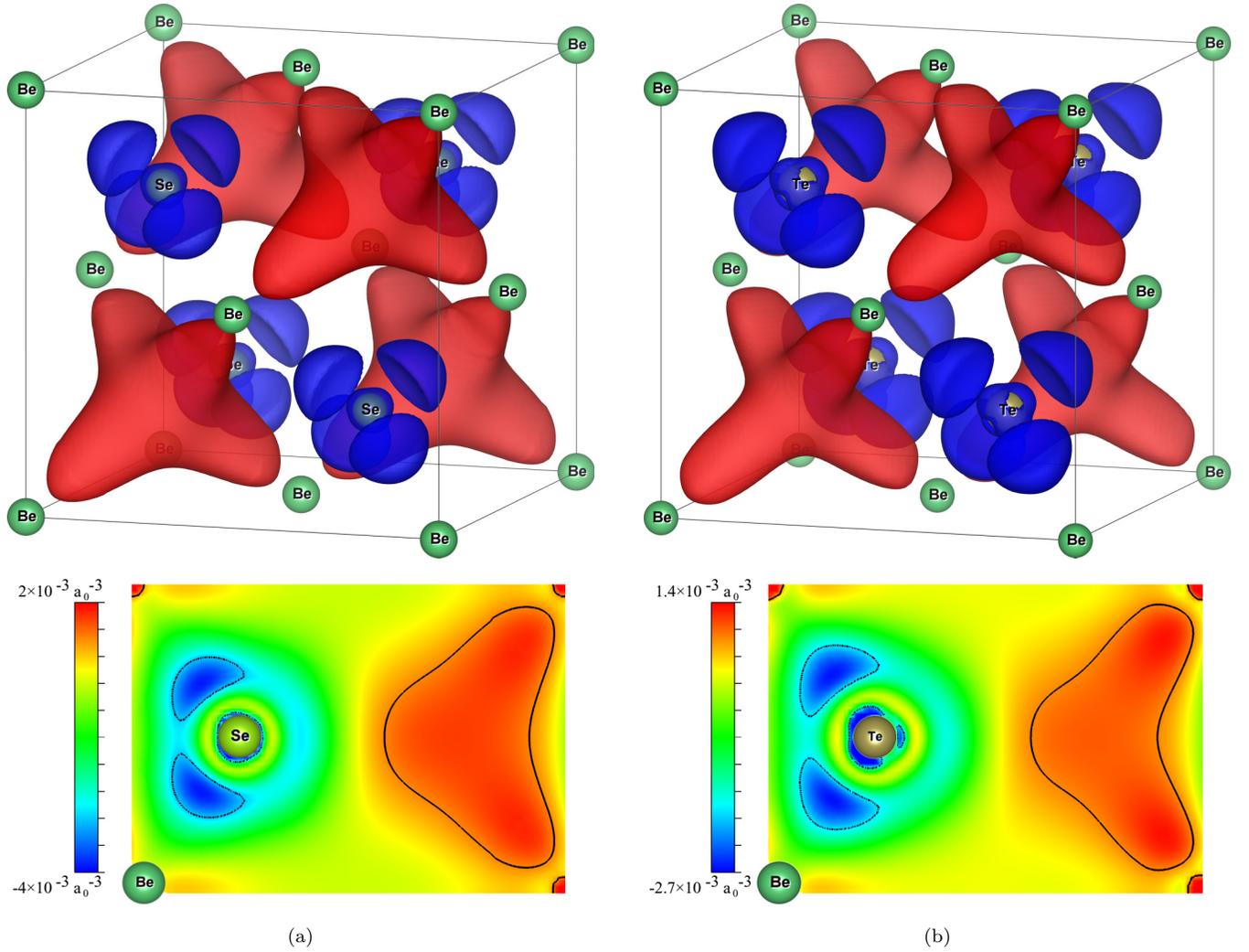


FIG. 5. Three-dimensional (upper panel) and two-dimensional (lower panel) plots of the electron density difference CBM – VBM in BeSe (a) and BeTe (b). On the three-dimensional plots, the positive (CBM, in red) and one negative (VBM, in blue) isosurfaces are defined at  $1.25 \times 10^{-3} a_0^{-3}$  and  $-2.75 \times 10^{-3} a_0^{-3}$  for BeSe and at  $0.9 \times 10^{-3} a_0^{-3}$  and  $-1.75 \times 10^{-3} a_0^{-3}$  for BeTe. On the two-dimensional plots, the slice corresponds to a  $1\bar{1}0$  plane with Be atoms at the corners and the anion atom at the left center, and the black contour lines correspond to the isosurfaces on the three-dimensional plots.

TABLE III. Band gaps (in eV) of TM oxides calculated using the DFT-1/2 method with different underlying functionals (LDA or PBE). In all cases the ionization was applied to the TM  $d$  and O  $p$ . The cutoff radii  $r_c$  (in  $a_0$ ) in Eq. (7) are indicated in the second column. Only the cutoff radii from LDA-1/2 calculations are shown, however we checked that those for PBE-1/2 calculations are practically identical (the difference is below  $0.05 a_0$ , which does not impact the band gap). For comparison purposes, LDA, PBE, and mBJ results are also shown. The experimental results are from Refs. [59,62,66,67]. The most accurate values among the DFT-1/2 methods are underlined.

Solid	$r_c$	LDA	LDA-1/2	PBE	PBE-1/2	mBJ	Expt.
TiO <sub>2</sub>	0.29, 2.76	1.80	3.16	1.89	<u>3.38</u>	2.56	3.30
VO <sub>2</sub>		metal	metal	metal	metal	0.51	0.6
Cu <sub>2</sub> O	2.73, 2.21	0.53	1.09	0.53	<u>1.14</u>	0.82	2.17
ZnO	1.68, 2.80	0.74	3.26	0.81	<u>3.50</u>	2.65	3.44
Cr <sub>2</sub> O <sub>3</sub> (AFM)	0.24, 2.0	1.20	1.35	1.64	<u>1.76</u>	3.68	3.4
MnO (AFM)	1.44, 2.90	0.74	1.89	0.89	<u>2.33</u>	2.94	3.9
FeO (AFM)		metal	metal	metal	metal	1.84	2.4
Fe <sub>2</sub> O <sub>3</sub> (AFM)	0.35, 2.87	0.33	1.33	0.56	<u>1.66</u>	2.35	2.2
CoO (AFM)	1.72, 2.57	metal	metal	metal	<u>0.17</u>	3.13	2.5
NiO (AFM)	1.35, 2.17	0.43	0.66	0.95	<u>1.33</u>	4.14	4.3
CuO (AFM)	5.42, 2.10	metal	0.84	0.06	<u>1.17</u>	2.27	1.44

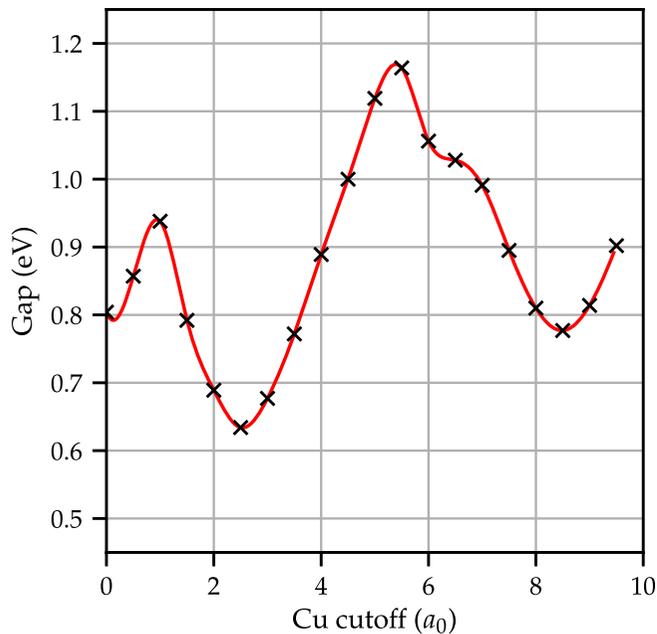


FIG. 6. CuO band gap using PBE-1/2 as a function of the cutoff radius of Cu. Splines (red line) are fitted through the calculated results (black crosses).

Ref. [41]), however, as shown in Fig. 6 the second maximum at  $5.42 a_0$  is higher than the first one at about  $1 a_0$ . Note that  $r_c = 5.42 a_0$  is very close to the distance to the nearest Cu atom of  $5.48 a_0$ , which means an overlap with a large portion of the neighboring Cu  $d$  orbitals. Such overlap introduces a

small anisotropy in the superimposed correction potential  $v_S$  around each Cu atom. We mention that because of numerical problems in the calculations when cutoff radii larger than  $10 a_0$  are used, we could not verify whether the third local maximum would be even higher or not.

Interestingly, in  $\text{Fe}_2\text{O}_3$  the reverse is observed. A slightly larger band gap (1.87 eV with PBE-1/2) is obtained when the wrong spin is ionized for calculating the correction potential. We hypothesize that this behavior is caused by a larger bonding-antibonding splitting of the Fe- $d_{eg}$  - O- $p$  interaction when the wrong spin is chosen.

Now, a comparison between  $\text{TiO}_2$  (accurately described by DFT-1/2) and  $\text{Cu}_2\text{O}$  (inaccurately described by DFT-1/2) is made. Figure 7 shows difference density plots, where the density around the VBM is subtracted from the one around the CBM. In  $\text{TiO}_2$  the VBM and CBM are, as expected, spatially well separated with the conduction band consisting primarily of the Ti  $d$  orbitals and the valence band of the O  $p$  orbitals. In  $\text{Cu}_2\text{O}$ , however, both bands are predominantly composed of Cu  $d$  orbitals, i.e., the  $d$  orbitals are split across the Fermi level due to the crystal field. The conduction band (in red) has a strong  $d_{z^2}$  character with lobes pointing towards the O atom, while the lobes of the valence band (in blue) point in other directions. Thus, as clearly visible, in  $\text{Cu}_2\text{O}$  the VBM and CBM are located on the same atom such that the spherical correction potential  $v_S$  can barely increase the energy difference between the VBM and CBM.

The PBE and PBE-1/2 band structures of  $\text{TiO}_2$  and  $\text{Cu}_2\text{O}$  are shown in Figs. 8(a) and 8(b), respectively. For  $\text{TiO}_2$ , the band gap is approximately twice as large. Changes in the shape of the bands are rather minor for the conduction

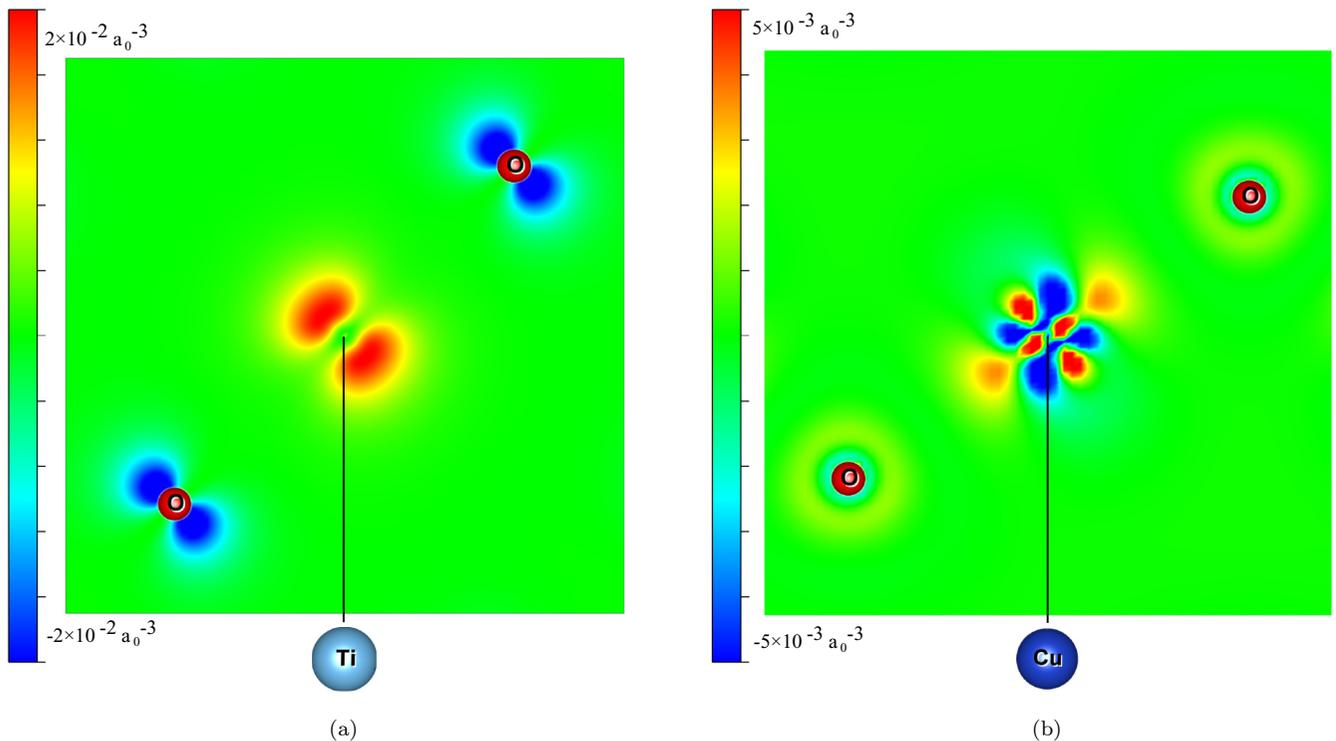


FIG. 7. Density difference (CBM - VBM) plots for  $\text{TiO}_2$  (a) and  $\text{Cu}_2\text{O}$  (b). For  $\text{TiO}_2$  the 001 plane is shown, with a Ti atom at the center and two nearest-neighbor O atoms at the corners. For  $\text{Cu}_2\text{O}$ , the  $1\bar{1}0$  plane is shown, with a Cu atom at the center and two nearest-neighbor O atoms at the corners.

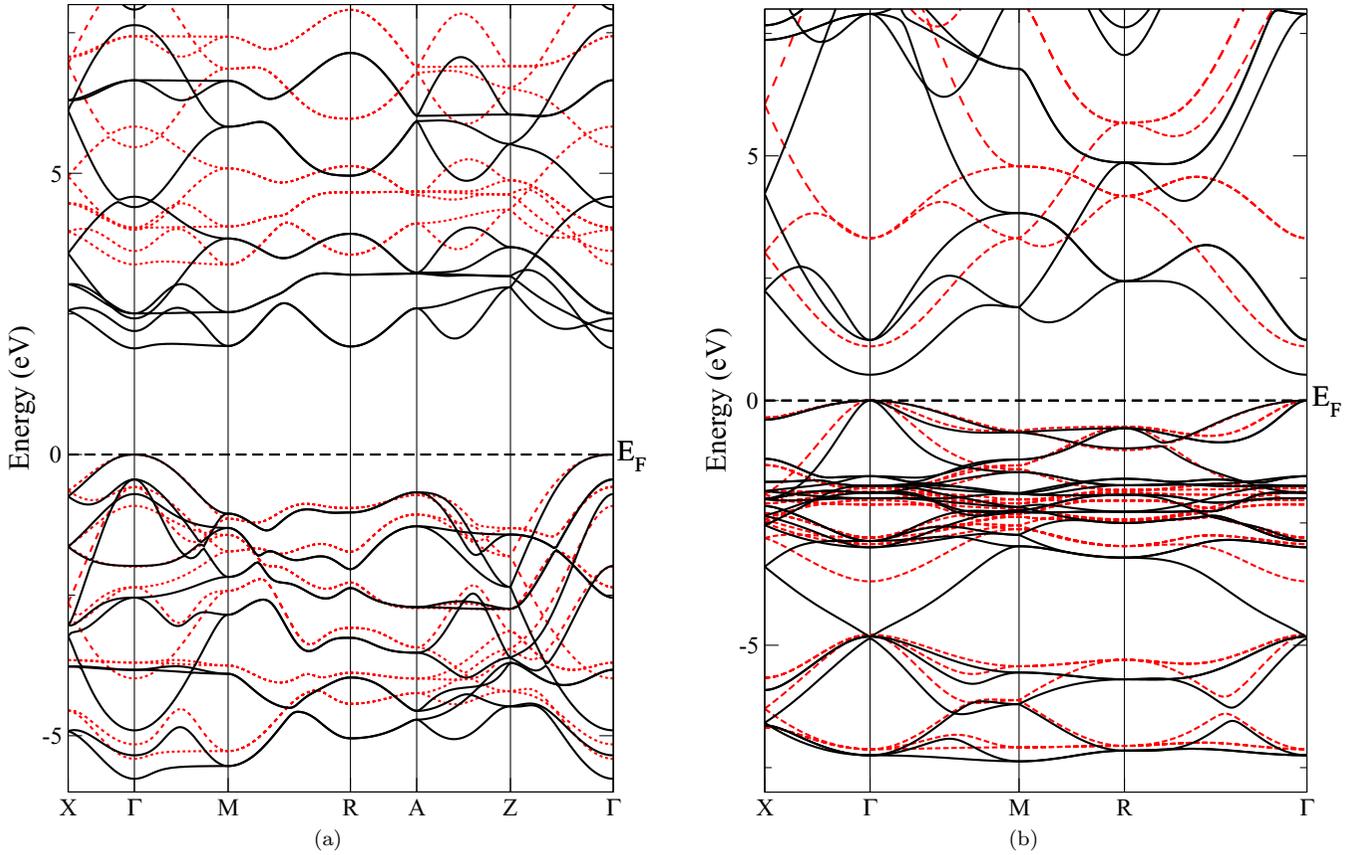


FIG. 8. PBE (black solid line) and PBE-1/2 (red dashed line) band structures for  $\text{TiO}_2$  (a) and  $\text{Cu}_2\text{O}$  (b).

bands, but more pronounced differences can be observed in the occupied bands, e.g., at the  $\Gamma$  and  $R$  points in the range 4–5 eV below the Fermi energy, where the changes do not consist of a simple shift. Among the differences in the shape of the bands in  $\text{Cu}_2\text{O}$ , there is for instance the crossing of bands at  $\Gamma$  at  $-5$  eV with PBE, while they are clearly separated with PBE-1/2. The band that is significantly raised in energy has a strong Cu  $d$  character, whereas the bands that are not shifted relative to the Fermi energy have strong O  $p$  character.

Figure 9 shows the mBJ, PBE, and PBE-1/2 exchange-correlation potentials in  $\text{TiO}_2$ , which lead to band gaps of 2.56 eV, 1.89 eV, and 3.38 eV, respectively. As mentioned above, mBJ performs badly. We can see that compared to PBE, mBJ raises the energy in the interstitial and has peaks at the outer atomic orbitals for both Ti and O (where, respectively, the CBM and VBM have a large density). With PBE-1/2 a much more accurate band gap is achieved thanks to a significantly more negative potential at the VBM region around the O atom.

To finish, we mention that Xue *et al.* [41] reported a similar issue in  $\text{Li}_2\text{O}_2$  as in  $\text{Cu}_2\text{O}$ . In this case the O  $p$  bands are split across the Fermi level, with the VBM formed by the (degenerate)  $p_x$  and  $p_y$  orbitals and the CBM by the  $p_z$  band, while the correction potential  $v_S$  is calculated from an atomic calculation, which is spherically symmetric. Thus, as in Ref. [41], a severe underestimation of the band gap for  $\text{Li}_2\text{O}_2$  is obtained and our LDA-1/2 and PBE-1/2 values are 2.52 and 2.71 eV, respectively (only  $\sim 0.5$  eV larger than

LDA and PBE), while experiment is 4.91 eV [75]. With a value of 4.81 eV, the mBJ potential succeeds in describing the band gap very accurately.

#### D. Shell correction for DFT-1/2

Xue *et al.* [41] proposed a more general version of DFT-1/2, called shDFT-1/2 (sh is a shorthand for shell), which employs a modified, shell-like cutoff function

$$\Theta(r) = \begin{cases} \left(1 - \left[\frac{2(r-r_{\text{in}})}{r_{\text{out}}-r_{\text{in}}}\right]^{20}\right)^3 & r_{\text{in}} < r < r_{\text{out}} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

with two variationally determined parameters  $r_{\text{in}}$  and  $r_{\text{out}}$  and a sharper cutoff compared to Eq. (7). Note that the shell-like cutoff function reduces to the spherical one of Eq. (7) when the inner radius is chosen as  $r_{\text{in}} = -r_{\text{out}}$  but with an exponent of 20 instead of 8. The inner radius  $r_{\text{in}}$  is also used to maximize the band gap, which implies that the band gap calculated with shDFT-1/2 should be larger compared to optimizing only  $r_{\text{out}}$  as done with DFT-1/2. The aim of introducing an inner radius is to avoid unwanted interaction of (semi)core electrons with the correction potential  $v_S$ . However, optimizing the radii in the shDFT-1/2 method is more tedious, since  $r_{\text{in}}$  and  $r_{\text{out}}$  need to be optimized simultaneously and may be interdependent. For example in GaAs, DFT-1/2 requires a Ga cutoff of  $1.23 a_0$  [20], whereas shLDA-1/4 requires  $r_{\text{in}} = 2.1 a_0$ , and  $r_{\text{out}} = 3.9 a_0$  [41] for the same atom. In this case (as in some others) the correction potential of both atoms

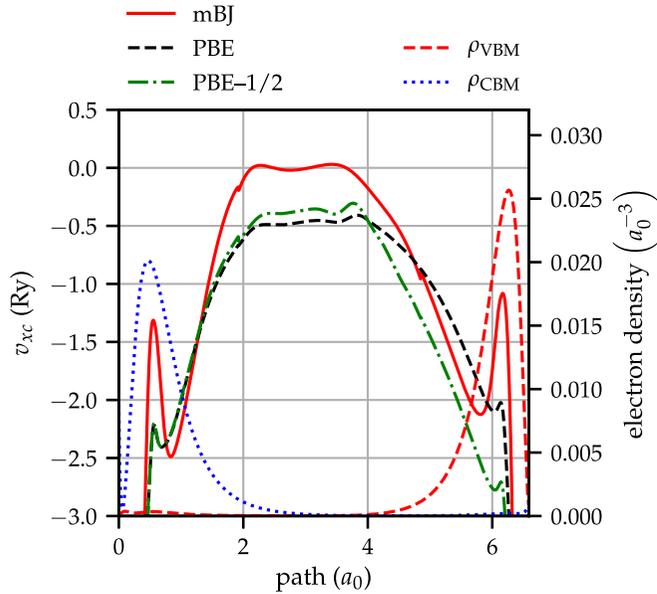


FIG. 9. Plots of mBJ, PBE, and PBE-1/2 exchange-correlation potentials  $v_{xc}$  in  $\text{TiO}_2$  and densities of the VBM and CBM. The path is from  $(0, 1, 0)$  to  $(0.305, 0.305, 0)$  in the unit cell fractional coordinates, thus from the Ti atom through the interstitial region and terminating at an O atom. The densities are taken from the PBE calculation.

overlap with the valence density that is distributed around one of the atoms as illustrated in Fig. 10 for GaAs. However, a well-founded explanation why this approach should yield more accurate band gaps is not provided in Ref. [41].

Xue *et al.* [41] also prescribed a procedure to choose the correction. In monoatomic compounds, there is a choice to apply either a 1/2- or a 1/4-ionization correction. The former should be used when the VBM density is distributed around the atom (like in diamond), while the latter should be used when the VBM density is distributed around the bond center (like in Si, see Sec. III A). In binary compounds, either a 1/2-ionization correction is applied to the anion or a 1/4 ionization to both the anion and cation. Which one of these two corrections is applied depends on the CBM density distribution. When the CBM density is distributed close to the cation-cation bonds (AIP is the example given by Xue *et al.* [41], but this would apply also to BeSe and BeTe, see Fig. 5), only the anion should be corrected by a 1/2-ionization. However, when the CBM density is distributed around the atoms, like in GaAs (Fig. 10), a 1/4-ionization correction should be applied to both atoms, and in such a case a large  $r_{in}$  should minimize the interaction of the correction potential  $v_S$  with the CBM. However, how to deal with a case like ZnO where a 1/2-ionization correction on both atoms is needed to obtain a reasonable band gap [24] is not discussed. It is also clear that other situations exist, like BAs (see Fig. 3) where the CBM density is distributed along the cation bonds but also around the anion (note the CBM lobes around the As atom, which are absent in BeSe and BeTe, see Fig. 5).

Before discussing the results obtained with shPBE-1/2, the influence of the steepness of the outer part of Eq. (8) is now discussed. As noted above, the outer cutoff is sharper in the shell function [Eq. (8)] than in the original spherical function

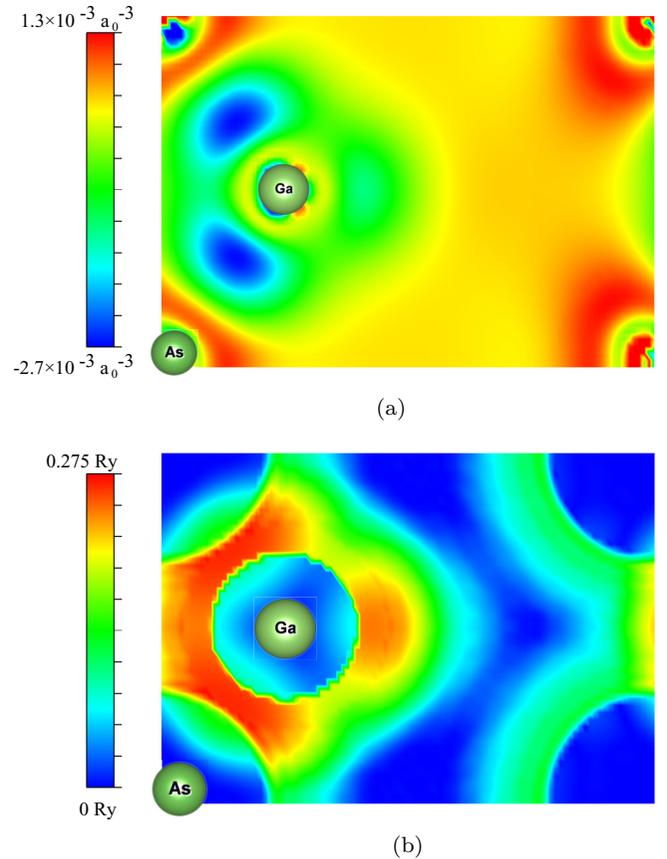


FIG. 10. Density difference (CBM - VBM) (a) and shPBE-1/4 potential  $v_S$  (b) for GaAs. The correction potentials of both atoms spatially overlap with the valence density maximum around the As atom. The procedure optimizes the four radial parameters ( $r_{in}$  and  $r_{out}$  of both atoms) such that the gap is maximized. This clearly means that the overlap of the total correction potential with the VBM density will be large while it will be minimized with the CBM density.

[Eq. (7)]. In order to test the influence of the outer steepness on the results, calculations with Eq. (8) were done using no inner cutoff (i.e., with  $r_{in} = -r_{out}$ , see discussion above) and the results are compared to those obtained with Eq. (7). The band gaps obtained with the two cutoff functions are shown in Table IV, where we can see that Eq. (8) leads to values

TABLE IV. Band gap (in eV) calculated with different cutoff functions: spherical [Eq. (7)] and sharp spherical [Eq. (8) with  $r_{in} = -r_{out}$  and  $r_{out}$  determined variationally]. The reason for choosing this value for  $r_{in}$  is that the shell-like cutoff function becomes very similar to the spherical one [Eq. (7)], with the only difference being the exponent of the reduced radius term  $r/r_{out}$  (see discussion in text).

Solid	Eq. (7)	Eq. (8)
Ge-1/4	0.27	0.41
AIP-0-1/2	3.21	3.33
BN-0-1/2	6.79	6.97
BeSe-0-1/2	4.26	4.40
GaAs-1/4	0.97	1.06

TABLE V. Band gaps (in eV) of various compounds using the shPBE method. The cutoff radii  $r_{\text{in}}$  and  $r_{\text{out}}$  are given for the specific correction that is required for the compound according to the rules of Xue *et al.* [41]. That means that when two cutoff radii are given, a 1/4-ionization correction is applied to both atoms, and when one is given a 1/2-ionization correction is applied only to the anion. For each of the shPBE calculations the cutoff radii were optimized for the specific correction, as the cutoff radii between shPBE-1/4-1/4 and shPBE-0-1/2 are not always transferable. For NiO, a normal 1/2-ionization correction is applied to all atoms in the unit cell. The most accurate values among the DFT-1/2 methods are underlined.

Solid	$r_{\text{in}}$	$r_{\text{out}}$	PBE	PBE-1/4	PBE-0-1/2	shPBE-1/4	shPBE-0-1/2	Expt.
Ge	1.71	3.30	metal	0.27	<u>0.59</u>	<u>0.93<sup>a</sup></u>	1.78	0.74
SiC	0.12	2.61	1.35	<u>2.43</u>	3.31	2.55	3.52 <sup>a</sup>	2.42
BN	0.15	2.14	4.47	5.79	<u>6.79</u>	5.87	7.03 <sup>a</sup>	6.36
BAs	0.18, 0.89	2.81, 2.70	1.09	<u>1.93</u>	2.00	2.03	2.14	1.46
AlN	0.13	2.51	3.34	4.66	5.96	<u>4.78</u>	6.19	4.90
AlP	0.66	3.15	1.59	<u>2.50</u>	3.21	<u>2.59</u>	3.36 <sup>a</sup>	2.5
GaN	0.16	2.53	1.66	2.55	<u>3.41</u>	2.71	3.55 <sup>a</sup>	3.28
GaAs	2.17, 1.48	4.21, 3.25	0.43	0.97	<u>1.49</u>	<u>1.54<sup>a</sup></u>	2.04	1.52
BeSe	0.83	2.89	2.51	3.36	<u>4.26</u>	3.55	<u>4.44<sup>a</sup></u>	4.0-4.5
BeTe	1.15	3.22	1.69	2.41	3.17	<u>2.69</u>	<u>3.42<sup>a</sup></u>	2.7
ZnO	0.13, 0.10	1.43, 2.38	0.81	2.08	<u>2.32</u>	2.21	<u>2.36<sup>a</sup></u>	3.44
NiO	-0.25, 0.12	1.29, 1.25	0.95	metal	<u>1.33</u>	metal	<u>1.36<sup>a</sup></u>	4.3

<sup>a</sup>Obtained with the preferred correction according to Xue *et al.* [41].

that are moderately larger by 0.1 eV to 0.2 eV. This is easily explained by noting that a steeper cutoff can more effectively maximize (minimize) the overlap of  $v_S$  with the VBM (CBM).

Representative compounds were considered for calculations with the shDFT-1/2 method. We chose border cases (BeTe and BeSe), some of the group IV and III-V semiconductors whose band gaps are significantly underestimated in a 1/4-ionization correction and BAs to check if the overestimation found even in LDA-1/4 is worsened or not. We also included one nonmagnetic (ZnO) and one AFM TM oxide (NiO) to see the influence of the inner cutoff on this class of materials. We limited our calculations to the PBE functional.

The results obtained with the shPBE-1/2 methods are shown in Table V. Compared to the corresponding PBE-1/2 methods with the same ionization correction, the improvement is in most cases rather small or nonexistent. Actually, comparing the results to those from Table IV discussed above, in most cases (e.g., AlP, BN, or BeSe) the increase in the band gap is mostly due to the sharper cutoff and not to the inner radius  $r_{\text{in}}$ . It is only for Ge and GaAs that the inner radius has a large influence on the results. For these two latter cases, excellent agreement with experiment is obtained with shPBE-1/4.

The case of Ge is particularly noteworthy. All our DFT-1/2 results predict a direct band gap, whereas shPBE-1/4 correctly predicts an indirect  $\Gamma \rightarrow L$  band gap of 0.93 eV (0.74 eV for experiment) and a slightly larger  $\Gamma \rightarrow \Gamma$  gap of 0.95 eV (0.89 eV for experiment) [76]. PBE-1/2 predicts a much smaller direct band gap of 0.59 eV and a  $\Gamma \rightarrow L$  gap of 1.18 eV. Except shDFT, the only other DFT-1/2 method that leads to an indirect band gap uses an additional correction to the conduction band by adding a 1/4 charge to an  $s$  orbital [77]. Note that also the mBJ potential correctly predicts an indirect  $\Gamma \rightarrow L$  band gap (0.76 eV) that is smaller than the direct one  $\Gamma \rightarrow \Gamma$  (0.91 eV) and furthermore in excellent agreement with experiment. The shLDA-1/4 results of Ref. [68], 0.80 eV for  $\Gamma \rightarrow L$  and 0.86 eV for  $\Gamma \rightarrow \Gamma$ , are also very close to experiment.

The other main observation is that shPBE-0-1/2 (1/2-ionization correction applied only to the anion) strongly overestimates the band gap in all cases except ZnO and NiO. In the case of NiO, it is expected since also the shell correction cannot capture the  $d$ - $d$  transition that makes up the fundamental gap. The case of ZnO shows that sometimes a 1/2-ionization correction on both atoms is required to obtain a good band gap (see Table III).

Actually, with the larger set of solids used by Xue *et al.* [41] to test shLDA-1/2, the overall improvement is rather modest, in particular when taking into account the fact that an extra parameter ( $r_{\text{in}}$ ) is introduced. This leads to a two-dimensional search for the cutoff radii, since the optimal  $r_{\text{out}}$  and  $r_{\text{in}}$  are in general interdependent as observed in GaAs for instance.

#### IV. SUMMARY

Since the DFT-1/2 method has been proposed, a large number of works reporting accurate results for the band gap have been published. However, as discussed in Xue *et al.* [41] and in the present work, the method has flaws which prevent its straightforward application. Firstly, for the cases where the states around the band gap, i.e., both at the VBM and CBM, come from orbitals centered at the same atom, the method will most likely fail. Such examples discussed in this work are many TM oxides but also  $\text{Li}_2\text{O}_2$ .

Secondly, the method cannot be blindly applied to covalent semiconductors and it is only recently [41] that this discussion has been extended beyond the group IV semiconductors. It is rather clear that there is no *unique* way (1/2- or 1/4-ionization correction, which atoms, and which orbital) to calculate reliably the band gap for these materials using (sh)DFT-1/2, without prior knowledge of the experimental band gap.

The comparison with the mBJ potential shows that mBJ is superior to DFT-1/2 for the test set considered in this work. The most visible differences in the performance of DFT-1/2 and mBJ are for the TM oxides. While DFT-1/2 is very

accurate for TiO<sub>2</sub> and ZnO, but very inaccurate for the AFM oxides, the reverse is observed with mBJ.

We also considered the shell correction (shDFT-1/2). It requires the introduction of an extra parameter, which leads to a more tedious application of the method. Furthermore, it is only for a few cases that shDFT-1/2 clearly improves the results.

Thus, we conclude that while DFT-1/2 is a computationally fast method and can be accurate for band gap calculations, one should be careful in its application. In particular, the method can be applied efficiently only when the VBM and CBM are spatially well separated, like in ionic solids, such that predominantly the VBM is shifted down by the correction potential and not the CBM. When these conditions are met, DFT-1/2 is certainly useful especially in systems

with large unit cells, like for example for the calculations of defect levels [28,78], surfaces [33], or interfaces [35,36]. An interesting perspective opened by the DFT-1/2 technique is the semiempirical application to larger structures. One can fit or tune the correction to a reference (e.g., bulk) configuration by parametrizing either the ionization level or the correction factor (multiplying the correction potential by a constant factor) [51,52] and consequently applying this semiempirical correction in the structure of interest, e.g., defects, interfaces or surfaces.

#### ACKNOWLEDGMENT

This work was supported by projects F41 (SFB ViCoM), W1243 (Solids4Fun), and P27738-N28 of the Austrian Science Fund (FWF).

- 
- [1] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).  
 [2] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).  
 [3] J. P. Perdew, *Int. J. Quantum Chem.* **28**, 497 (1986).  
 [4] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996); **78**, 1396(E) (1997).  
 [5] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, *J. Chem. Phys.* **123**, 174101 (2005).  
 [6] F. Aryasetiawan and O. Gunnarsson, *Rep. Prog. Phys.* **61**, 237 (1998).  
 [7] L. Hedin, *J. Phys.: Condens. Matter* **11**, R489 (1999).  
 [8] M. Shishkin, M. Marsman, and G. Kresse, *Phys. Rev. Lett.* **99**, 246403 (2007).  
 [9] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, *Phys. Rev. B* **53**, 3764 (1996).  
 [10] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).  
 [11] F. Della Sala, E. Fabiano, and L. A. Constantin, *Int. J. Quantum Chem.* **116**, 1641 (2016).  
 [12] B. Xiao, J. Sun, A. Ruzsinszky, J. Feng, R. Haunschild, G. E. Scuseria, and J. P. Perdew, *Phys. Rev. B* **88**, 184103 (2013).  
 [13] Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, *Phys. Rev. B* **93**, 205205 (2016).  
 [14] S. Jana, A. Patra, and P. Samal, *J. Chem. Phys.* **149**, 044120 (2018).  
 [15] R. Armiento and S. Kümmel, *Phys. Rev. Lett.* **111**, 036402 (2013).  
 [16] V. Vlček, G. Steinle-Neumann, L. Leppert, R. Armiento, and S. Kümmel, *Phys. Rev. B* **91**, 035107 (2015).  
 [17] O. Gritsenko, R. van Leeuwen, E. van Lenthe, and E. J. Baerends, *Phys. Rev. A* **51**, 1944 (1995).  
 [18] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, *Phys. Rev. B* **82**, 115106 (2010).  
 [19] F. Tran and P. Blaha, *Phys. Rev. Lett.* **102**, 226401 (2009).  
 [20] L. G. Ferreira, M. Marques, and L. K. Teles, *Phys. Rev. B* **78**, 125116 (2008).  
 [21] J. C. Slater, in *Advances in Quantum Chemistry*, Vol. 6 (Elsevier, London, 1972), pp. 1–92.  
 [22] J. C. Slater and K. H. Johnson, *Phys. Rev. B* **5**, 844 (1972).  
 [23] L. G. Ferreira, R. R. Pelá, L. K. Teles, M. Marques, M. Ribeiro Jr., and J. Furthmüller, in *Proceedings of the 31st International Conference on the Physics of Semiconductors (ICPS) 2012*, edited by T. Ihn, C. Rössler, and A. Kozikov, AIP Conf. Proc. No. 1566 (AIP, New York, 2013), p. 27.  
 [24] L. G. Ferreira, M. Marques, and L. K. Teles, *AIP Adv.* **1**, 032119 (2011).  
 [25] R. R. Pela, A. Gulans, and C. Draxl, *Comput. Phys. Commun.* **220**, 263 (2017).  
 [26] R. R. Pela, U. Werner, D. Nabok, and C. Draxl, *Phys. Rev. B* **94**, 235141 (2016).  
 [27] S. X. Tao, X. Cao, and P. A. Bobbert, *Sci. Rep.* **7**, 14386 (2017).  
 [28] B. Lucatto, L. V. C. Assali, R. R. Pela, M. Marques, and L. K. Teles, *Phys. Rev. B* **96**, 075145 (2017).  
 [29] M. Ribeiro Jr., *Can. J. Phys.* **93**, 261 (2015).  
 [30] A. Belabbes, A. Zaoui, and M. Ferhat, *Appl. Phys. Lett.* **97**, 242509 (2010).  
 [31] J. P. T. Santos, M. Marques, L. G. Ferreira, R. R. Pelá, and L. K. Teles, *Appl. Phys. Lett.* **101**, 112403 (2012).  
 [32] M. Ribeiro Jr., L. R. C. Fonseca, T. Sadowski, and R. Ramprasad, *J. Appl. Phys.* **111**, 073708 (2012).  
 [33] A. Belabbes, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* **84**, 205304 (2011).  
 [34] S. Küfner, A. Schleife, B. Höffling, and F. Bechstedt, *Phys. Rev. B* **86**, 075320 (2012).  
 [35] M. Ribeiro Jr., L. R. C. Fonseca, and L. G. Ferreira, *Europhys. Lett.* **94**, 27001 (2011).  
 [36] M. Ribeiro Jr., L. R. C. Fonseca, and L. G. Ferreira, *Phys. Rev. B* **79**, 241312(R) (2009).  
 [37] R. R. Pela, M. Marques, and L. K. Teles, *J. Phys.: Condens. Matter* **27**, 505502 (2015).  
 [38] F. Matusalem, M. Marques, L. K. Teles, A. Filippetti, and G. Cappellini, *J. Phys.: Condens. Matter* **30**, 365501 (2018).  
 [39] R. R. Pelá, M. Marques, L. G. Ferreira, J. Furthmüller, and L. K. Teles, *Appl. Phys. Lett.* **100**, 202408 (2012).  
 [40] R. R. Pela, A. Gulans, and C. Draxl, *J. Chem. Theory Comput.* **14**, 4678 (2018).  
 [41] K.-H. Xue, J.-H. Yuan, L. R. C. Fonseca, and X.-S. Miao, *Comput. Mater. Sci.* **153**, 493 (2018).  
 [42] J. P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).  
 [43] W. Yang, A. J. Cohen, and P. Mori-Sánchez, *J. Chem. Phys.* **136**, 204111 (2012).  
 [44] X. Andrade and A. Aspuru-Guzik, *Phys. Rev. Lett.* **107**, 183002 (2011).

- [45] J.-D. Chai and P.-T. Chen, *Phys. Rev. Lett.* **110**, 033002 (2013).
- [46] E. Kraissler and L. Kronik, *J. Chem. Phys.* **140**, 18A540 (2014).
- [47] M. Grüning, A. Marini, and A. Rubio, *J. Chem. Phys.* **124**, 154108 (2006).
- [48] M. Grüning, A. Marini, and A. Rubio, *Phys. Rev. B* **74**, 161103(R) (2006).
- [49] J. F. Janak, *Phys. Rev. B* **18**, 7165 (1978).
- [50] S. Kümmel and L. Kronik, *Rev. Mod. Phys.* **80**, 3 (2008).
- [51] C. A. Ataide, R. R. Pelá, M. Marques, L. K. Teles, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* **95**, 045126 (2017).
- [52] M. Ribeiro Jr., *J. Appl. Phys.* **117**, 234302 (2015).
- [53] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, J. Luitz, R. Laskowski, F. Tran, and L. D. Marks, *WIEN2k: An Augmented Plane Wave plus Local Orbitals Program for Calculating Crystal Properties* (Vienna University of Technology, Austria, 2018).
- [54] O. K. Andersen, *Phys. Rev. B* **12**, 3060 (1975).
- [55] D. J. Singh and L. Nordström, *Planewaves, Pseudopotentials, and the LAPW Method, 2nd ed.* (Springer, New York, 2006).
- [56] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.99.115101> for the experimental lattice constants of the solids used for the calculations.
- [57] F. Tran and P. Blaha, *J. Phys. Chem. A* **121**, 3318 (2017).
- [58] F. Tran, S. Ehsan, and P. Blaha, *Phys. Rev. Materials* **2**, 023802 (2018).
- [59] J. Lee, A. Seko, K. Shitara, K. Nakayama, and I. Tanaka, *Phys. Rev. B* **93**, 115104 (2016).
- [60] K. Nakano and T. Sakai, *J. Appl. Phys.* **123**, 015104 (2018).
- [61] M. J. Lucero, T. M. Henderson, and G. E. Scuseria, *J. Phys.: Condens. Matter* **24**, 145504 (2012).
- [62] J. M. Crowley, J. Tahir-Kheli, and W. A. Goddard III, *J. Phys. Chem. Lett.* **7**, 1198 (2016).
- [63] M. Bagheri and P. Blaha, *J. Electron Spectrosc. Relat. Phenom.* **230**, 1 (2019).
- [64] W. M. Yim, J. P. Dismukes, E. J. Stofko, and R. J. Paff, *J. Phys. Chem. Solids* **33**, 501 (1972).
- [65] M. Nagelstraßer, H. Dröge, H.-P. Steinrück, F. Fischer, T. Litz, A. Waag, G. Landwehr, A. Fleszar, and W. Hanke, *Phys. Rev. B* **58**, 10394 (1998).
- [66] R. Gillen and J. Robertson, *J. Phys.: Condens. Matter* **25**, 165502 (2013).
- [67] Y. Wang, S. Lany, J. Ghanbaja, Y. Fagot-Revurat, Y. P. Chen, F. Soldera, D. Horwat, F. Mücklich, and J. F. Pierson, *Phys. Rev. B* **94**, 245418 (2016).
- [68] J.-H. Yuan, Q. Chen, L. R. C. Fonseca, M. Xu, K.-H. Xue, and X.-S. Miao, *J. Phys. Commun.* **2**, 105005 (2018).
- [69] K. Terakura, T. Oguchi, A. R. Williams, and J. Kübler, *Phys. Rev. B* **30**, 4734 (1984).
- [70] V. I. Anisimov, J. Zaanen, and O. K. Andersen, *Phys. Rev. B* **44**, 943 (1991).
- [71] F. Tran, P. Blaha, K. Schwarz, and P. Novák, *Phys. Rev. B* **74**, 155108 (2006).
- [72] M. Marsman, J. Paier, A. Stroppa, and G. Kresse, *J. Phys.: Condens. Matter* **20**, 064201 (2008).
- [73] M. Gerosa, C. E. Bottani, C. Di Valentin, G. Onida, and G. Pacchioni, *J. Phys.: Condens. Matter* **30**, 044003 (2018).
- [74] D. Koller, F. Tran, and P. Blaha, *Phys. Rev. B* **83**, 195134 (2011).
- [75] J. S. Hummelshøj, J. Blomqvist, S. Datta, T. Vegge, J. Rossmeisl, K. S. Thygesen, A. C. Luntz, K. W. Jacobsen, and J. K. Nørskov, *J. Chem. Phys.* **132**, 071101 (2010).
- [76] R. Pässler, *J. Appl. Phys.* **89**, 6235 (2001).
- [77] F. Matusalem, M. Marques, L. K. Teles, and F. Bechstedt, *Phys. Rev. B* **92**, 045436 (2015).
- [78] F. Matusalem, M. Ribeiro, Jr., M. Marques, R. R. Pelá, L. G. Ferreira, and L. K. Teles, *Phys. Rev. B* **88**, 224102 (2013).