

Accurate Band Gaps of Semiconductors and Insulators with a Semilocal Exchange-Correlation Potential

Fabien Tran and Peter Blaha

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

(Received 5 December 2008; published 3 June 2009)

A modified version of the exchange potential proposed by Becke and Johnson [J. Chem. Phys. **124**, 221101 (2006)] is tested on solids for the calculation of band gaps. The agreement with experiment is very good for all types of solids we considered (e.g., wide band gap insulators, *sp* semiconductors, and strongly correlated *3d* transition-metal oxides) and is of the same order as the agreement obtained with the hybrid functionals or the *GW* methods. This semilocal exchange potential, which recovers the local-density approximation (LDA) for a constant electron density, mimics very well the behavior of orbital-dependent potentials and leads to calculations which are barely more expensive than LDA calculations. Therefore, it can be applied to very large systems in an efficient way.

DOI: 10.1103/PhysRevLett.102.226401

PACS numbers: 71.15.Mb, 71.20.Mq, 71.20.Nr, 71.27.+a

In the solid-state community, the vast majority of electronic structure calculations are done using the Kohn-Sham equations [1] with the local-density approximation (LDA) [1,2] or generalized gradient approximation (GGA) [3] for the exchange-correlation energy and potential. The main reason is that these (semi)local approximations yield, in many circumstances, results which are accurate enough to help interpret experimental data or to have some predictive power. Another important advantage of semilocal functionals is that they lead to calculations which are computationally cheap in comparison to more sophisticated methods. However, the application of LDA and GGA to solids can also lead, depending on the studied solid and property, to results which are in very bad agreement with experiment, e.g., for the band gap of semiconductors and insulators which is severely underestimated or even absent (see, e.g., Ref. [4]). More generally, with an orbital-independent potential (i.e., a potential which is the same for all orbitals) the band gap calculated from the eigenvalue spectrum differs from the true band gap (the ionization potential I minus the electron affinity A) by the derivative discontinuity of the exchange-correlation potential Δ_{xc} [5,6]. Better band gaps can be calculated using other functionals or methods. The optimized effective potential (OEP) method applied to the exact exchange functional usually leads to better band gaps, but strong underestimations and overestimations can also be obtained [7,8]. Furthermore, the results of several studies could indicate that these good results are fortuitous, since applying the OEP method to the exact exchange plus some accurate *ab initio* correlation leads to band gaps similar to LDA band gaps. The reason is that Δ_{xc} can be important (see Ref. [9] and references therein). The hybrid functionals [4,10] can be applied to many classes of solids, but lead to calculations which are very expensive (albeit less if a screened version is used). The LDA + U [11] scheme is computationally as cheap as the semilocal functionals, but can only be applied to localized electrons (e.g.,

3d or *4f*) and contains atom-dependent parameters, while the LDA plus dynamic mean-field theory (LDA + DMFT) method is very successful (see, e.g., Ref. [12]), but leads to significantly more expensive calculations. The *GW* method can yield very accurate band gaps (see, e.g., Refs. [13–18]), but leads to very expensive calculations, particularly if done self-consistently [14,18].

Recently [19], we tested an exchange potential proposed by Becke and Johnson (BJ) [20], which was designed to reproduce the shape of the exact exchange OEP potential of atoms. We showed that this potential improves over the LDA and PBE [3] (the standard GGA for solids) potentials for the description of band gaps [19]. Nevertheless, the BJ potential (used in combination with LDA correlation [2]) still underestimates the band gaps significantly, which means that there is room for improvement.

In this Letter, we present a simple modification of the original BJ exchange potential which yields band gaps with an accuracy comparable to approaches which are orders of magnitude more expensive. The modified BJ potential (MBJ) we propose is

$$v_{x,\sigma}^{\text{MBJ}}(\mathbf{r}) = c v_{x,\sigma}^{\text{BR}}(\mathbf{r}) + (3c - 2) \frac{1}{\pi} \sqrt{\frac{5}{12}} \sqrt{\frac{2t_{\sigma}(\mathbf{r})}{\rho_{\sigma}(\mathbf{r})}}, \quad (1)$$

where $\rho_{\sigma} = \sum_{i=1}^{N_{\sigma}} |\psi_{i,\sigma}|^2$ is the electron density, $t_{\sigma} = (1/2) \sum_{i=1}^{N_{\sigma}} \nabla \psi_{i,\sigma}^* \cdot \nabla \psi_{i,\sigma}$ is the kinetic-energy density, and

$$v_{x,\sigma}^{\text{BR}}(\mathbf{r}) = -\frac{1}{b_{\sigma}(\mathbf{r})} \left(1 - e^{-x_{\sigma}(\mathbf{r})} - \frac{1}{2} x_{\sigma}(\mathbf{r}) e^{-x_{\sigma}(\mathbf{r})} \right) \quad (2)$$

is the Becke-Roussel (BR) [21] potential which was proposed to model the Coulomb potential created by the exchange hole. x_{σ} [in Eq. (2)] is determined from an equation involving ρ_{σ} , $\nabla \rho_{\sigma}$, $\nabla^2 \rho_{\sigma}$, and t_{σ} , and then b_{σ} is calculated with $b_{\sigma} = [x_{\sigma}^3 e^{-x_{\sigma}} / (8\pi \rho_{\sigma})]^{1/3}$. Originally, BJ used the Slater potential $v_{x,\sigma}^{\text{Slater}}$ [22] instead of $v_{x,\sigma}^{\text{BR}}$, but

they showed that these two potentials are quasi-identical for atoms [20]. In Eq. (1), c was chosen to depend linearly on the square root of the average of $|\nabla\rho|/\rho$:

$$c = \alpha + \beta \left(\frac{1}{V_{\text{cell}}} \int_{\text{cell}} \frac{|\nabla\rho(\mathbf{r}')|}{\rho(\mathbf{r}')} d^3r' \right)^{1/2}, \quad (3)$$

where α and β are two free parameters and V_{cell} is the unit cell volume. Minimization of the mean absolute relative error for the band gap of the solids listed in Table I leads to $\alpha = -0.012$ (dimensionless) and $\beta = 1.023 \text{ bohr}^{1/2}$. Equation (1) was chosen such that the LDA exchange potential $v_{x,\sigma}^{\text{LDA}} = -(3/\pi)^{1/3}(2\rho_\sigma)^{1/3}$ is approximately recovered (for any value of c) for a constant electron density. Indeed, $v_{x,\sigma}^{\text{BR}} \approx v_{x,\sigma}^{\text{Slater}}$, which reduces to $(3/2)v_{x,\sigma}^{\text{LDA}}$ for a constant ρ , while the second term of Eq. (1) [without $(3c-2)$] reduces to $-(1/2)v_{x,\sigma}^{\text{LDA}}$ since $t_\sigma = (3/20)(3\pi^2)^{2/3} \times (2\rho_\sigma)^{5/3}$ for a constant ρ . For $c = 1$ the original BJ potential is recovered [20]. We mention that it was shown in Ref. [28] that due to the second term in Eq. (1), the BJ potential reproduces very well the step structure and derivative discontinuity of the exact exchange potential,

TABLE I. Theoretical and experimental fundamental band gaps (in eV). The structure is indicated in parenthesis. For comparison, results from the literature which were obtained by other methods are also shown (HSE03, HSE06, G_0W_0 , and GW). The experimental values were taken from Refs. [4,7,10,14,18,23–26].

Solid	LDA	MBJLDA	HSE	G_0W_0	GW	Expt.
Ne (A1)	11.42	22.72		19.59 ^e	22.1 ^g	21.70
Ar (A1)	8.16	13.91	10.34 ^d	13.28 ^e	14.9 ^g	14.20
Kr (A1)	6.76	10.83				11.6
Xe (A1)	5.78	8.52				9.8
C (A4)	4.11	4.93	5.49 ^b	5.50 ^e	6.18 ^g	5.48
Si (A4)	0.47	1.17	1.28 ^b	1.12 ^e	1.41 ^g	1.17
Ge (A4)	0.00	0.85	0.83 ^b	0.66 ^f	0.95 ^g	0.74
LiF (B1)	8.94	12.94		13.27 ^e	15.9 ^g	14.20
LiCl (B1)	6.06	8.64				9.4
MgO (B1)	4.70	7.17	6.67 ^b	7.25 ^e	9.16 ^g	7.83
ScN (B1)	-0.14	0.90		0.95 ^f	1.4 ^h	~0.9
MnO (B1)	0.76	2.95	2.8 ^c		3.5 ⁱ	3.9 ± 0.4
FeO (B1)	-0.35	1.82	2.2 ^c			2.4
NiO (B1)	0.42	4.16	4.2 ^c	1.1 ^f	4.8 ⁱ	4.0, 4.3
SiC (B3)	1.35	2.28	2.40 ^b	2.27 ^e	2.88 ^g	2.40
BN (B3)	4.39	5.85	5.99 ^b	6.10 ^e	7.14 ^g	~6.25
GaN (B3)	1.63	2.81	3.14 ^b	2.80 ^e	3.82 ^g	3.20
GaAs (B3)	0.30	1.64	1.12 ^b	1.30 ^e	1.85 ^g	1.52
AIP (B3)	1.46	2.32	2.51 ^b	2.44 ^e	2.90 ^g	2.45
ZnS (B3)	1.84	3.66	3.49 ^b	3.29 ^e	4.15 ^g	3.91
CdS (B3)	0.86	2.66	2.25 ^b	2.06 ^e	2.87 ^g	2.42
AlN (B4)	4.17	5.55	5.81 ^b	5.83 ^f		6.28
ZnO (B4)	0.75	2.68	2.49 ^d	2.51 ^f	3.8 ^g	3.44

^aHSE06, erratum of Ref. [10].

^bHSE03, supplementary material of Ref. [4].

^cHSE03 [26]. ^dHSE06 [27]. ^eReference [17]. ^fReference [15].

^gReference [18]. ^hReference [16]. ⁱReference [14].

which is an important result since only semilocal quantities are used. The BJ potential was also studied in Ref. [29]. Correlation effects were taken into account by adding a LDA correlation potential [2] to $v_{x,\sigma}^{\text{MBJ}}$ (MBJLDA); however, in Ref. [19] we showed that adding LDA correlation to the BJ potential has a relatively small effect.

Varying c in Eq. (1), we observed that for all studied solids the band gap increases monotonically with respect to c , and since the values obtained with $c = 1$ are for all cases too small with respect to experiment [19], a larger value for c leads to better agreement with experiment. More specifically, for solids with small band gaps, c_{opt} (the value of c which leads to a perfect agreement with experiment) lies within the range 1.1–1.3, while for large band gaps, c_{opt} is larger (1.4–1.7). Therefore, our goal was to find a property of the systems which could uniquely define a value for c close to c_{opt} . Such a property could, e.g., be the dielectric constant, but finally we got inspired by Ref. [30] where it is proposed to use the expression $\omega = \eta|\nabla\rho|/\rho$ for the parameter ω which defines the separation between short- and long-range exchange in the screened hybrid functional HSE [4]. Among the different possibilities we have tried, Eq. (3) is the one which leads to the most satisfying results. Note that since $v_{x,\sigma}^{\text{BR}} \approx v_{x,\sigma}^{\text{Slater}}$ and $v_{x,\sigma}^{\text{Slater}}$ is an average of the Hartree-Fock potential [22], Eq. (1) can be seen as a kind of “hybrid” potential whose amount of “exact exchange” is given by c .

Table I and Fig. 1 show the results obtained with the LDA and MBJLDA potentials for the fundamental band gap of 23 solids. We used the WIEN2K package [31] which is based on the full-potential (linearized) augmented plane-wave and local orbitals [FP-(L)APW + lo] method (see Ref. [32] and references therein). For comparison pur-

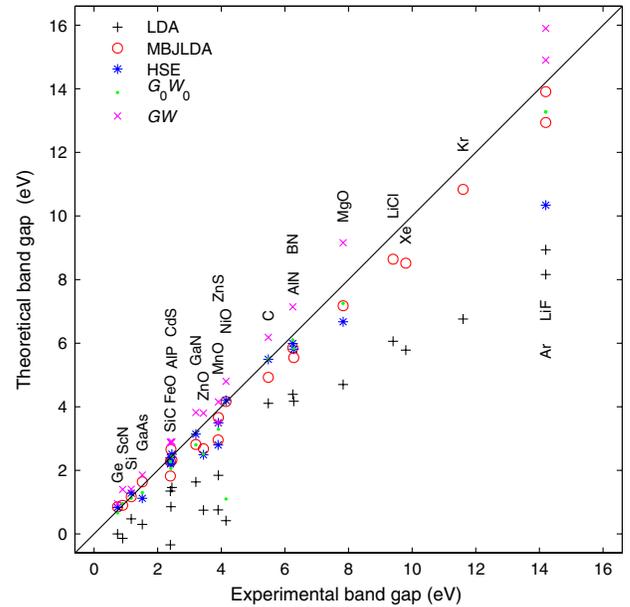


FIG. 1 (color online). Theoretical versus experimental band gaps. The values are given in Table I (Ne is omitted).

poses, experimental [4,7,10,14,18,23–26], HSE (HSE03 or HSE06) [4,10,26,27], and *GW* results are also given. For *GW*, we list all-electron non-self-consistent (G_0W_0) [15,17] and self-consistent (*GW*) [14,16,18] results. Our results as well as most of the literature theoretical results were obtained at the experimental geometry. Note that among the considered solids there are not only simple *sp* semiconductors, but also rare-gas solids (from Ne to Xe) with band gaps up to 22 eV and the antiferromagnetic MnO, FeO, and NiO whose 3*d* electrons are strongly correlated. We can see that for most cases the MBJLDA potential yields band gaps which are in (very) good agreement with experiment leading to typical errors of less than 10%. For instance, for Ge and ScN the band gaps of 0.85 and 0.90 eV are obtained, respectively, in very good agreement with experiment (0.74 and 0.9 eV). For medium-size band gaps, very good agreement with experiment is obtained for SiC, ZnS, CdS, and AlP with errors smaller than 0.25 eV. We also note the good results obtained for the rare-gas solids, LiF, and LiCl which have very large band gaps. Among the 23 considered solids only MnO, FeO, and ZnO have relative errors larger than 15%. Note, however, that FeO is metallic with LDA.

A comparison with other theoretical methods shows that for solids with small- or medium-size band gaps, the HSE hybrid method leads to rather similar results. For larger band gap systems HSE is more accurate for C and BN, while the MBJLDA results are much closer to experiment for Ar and MgO. The trend of the G_0W_0 method is to underestimate band gaps (in particular for NiO, GaN, ZnS, CdS, and ZnO), while the self-consistent *GW* band gaps are slightly too large compared to experiment, with errors of similar magnitude than with MBJLDA. Overall, the semilocal MBJLDA potential leads to results that are of similar accuracy as other much more sophisticated theories, which are computationally orders of magnitude more demanding. We have not reviewed the many non-self-consistent G_0W_0 calculations based on the pseudopotential approach, although they often lead to very accurate results, since this scheme may benefit from some cancellation of errors. It seems that only when electron-hole interactions in self-consistent *GW* calculations are taken into account, further systematical improvement can be achieved [18].

We also compared the full band structures of Si and GaAs. With respect to the LDA eigenvalues, the MBJLDA eigenvalues for Si are shifted upwards by 0.1–0.3 eV for the occupied states and 0.5–1 eV for the unoccupied states, and for the states where the experimental data are reliable (see Ref. [13] for a compilation of experimental values), the MBJLDA potential brings the eigenvalues closer to the experimental values than LDA does. For instance, for the state Γ_{2c} , LDA gives 3.17 eV, while MBJLDA gives 4.34 eV (4.1–4.2 eV for experiment). While for Si the shift is rather uniform in *k* space, for GaAs it is not. For instance, MBJLDA shifts up the Γ_{1c} , X_{1c} , and L_{1c} states by 1.34, 0.71, and 0.95 eV with respect to LDA. This yields MBJLDA eigenvalues of 1.64, 2.06, and 1.80 eV, which

compare very well with the experimental values of 1.52, 1.90, and 1.74 eV [13].

For MnO, FeO, and NiO, LDA + *U* calculations were also done (with *U* and *J* given in Ref. [11]). From Fig. 2, which shows the density of states (DOS) of NiO, we can see that the effect of the LDA + *U* and MBJLDA potentials is to shift up (with respect to LDA) the unoccupied Ni 3*d* states. For the MBJLDA potential this shift leads to very good agreement with photoelectron spectra [25]. However, three major differences between the on-site LDA + *U* and MBJLDA methods can be noticed. (a) Since the MBJLDA potential is applied to all electrons, the tail in the unoccupied LDA + *U* DOS which is just below the Ni 3*d* peak (mainly of Ni 4*s* character) is absent in the MBJLDA DOS. (b) The “lower Hubbard band” of dominant Ni 3*d* character at the bottom of the valence band is very weak in MBJLDA calculations (note that in experiment [33] and LDA + DMFT calculations [12] this band is also rather weak). (c) The Mott-Hubbard character of NiO is much more pronounced with the MBJLDA than with LDA + *U*, but MBJLDA still leads to an appreciable amount of O 2*p* states at the top of the valence bands. Recently, NiO has been experimentally described as a mixed Mott-Hubbard–charge-transfer insulator [33]. From Fig. 2, we can also see that the Ni 3*d* and O 2*p* DOSs obtained with the MBJLDA potential agree very well with the Ni and O spectra obtained from x-ray-emission spectroscopy (XES) measurements [33], while the Ni 3*d* DOS obtained with LDA + *U* shows a completely different structure. For MnO and FeO, the structure of the DOS is in better agreement with experiment with LDA + *U* and MBJLDA, respectively, and we note that for these two

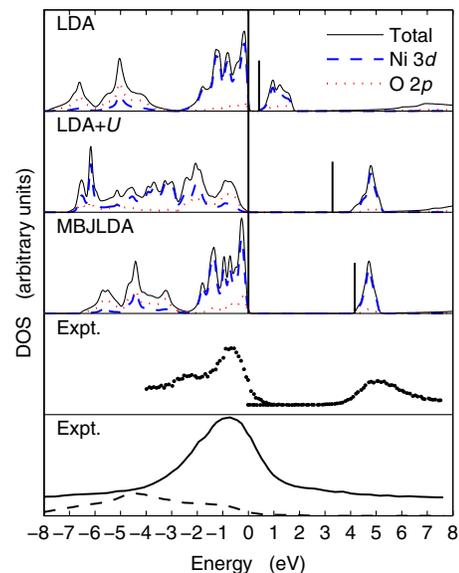


FIG. 2 (color online). DOS of NiO. The vertical bars indicate the end of the fundamental band gap which starts at $E = 0$ eV. The panels labeled “Expt.” show photoelectron [25] (upper panel) and XES [33] [lower panel, Ni (solid line) and O (dashed line) spectra] measurements.

compounds the two methods lead to similar lower Hubbard bands (unlike in NiO). We also mention that LDA + U and MBJLDA potentials open a band gap (2.52 and 2.69 eV) in ferromagnetic NiO (with broken symmetry), while LDA does not; thus, we expect the MBJLDA potential to open also a band gap in paramagnetic NiO. The atomic spin magnetic moments obtained with the LDA + U and MBJLDA methods are very similar. These two methods lead to moments of 4.54 and $4.51\mu_B$ for MnO, 3.64 and $3.63\mu_B$ for FeO, and 1.74 and $1.75\mu_B$ for NiO. These values are much larger than the LDA values (4.23, 3.38, and $1.21\mu_B$ for MnO, FeO, and NiO), and hence in much better agreement with experiment (see Ref. [34] for a compilation of experimental values).

We mention that since there is no exchange functional E_x such that $v_{x,\sigma}^{\text{MBJ}} = \delta E_x / \delta \rho_\sigma$, we recommend to use first a modern GGA functional for the structural properties, and then to use the MBJLDA potential for the calculation of the band structure.

As mentioned above, only orbital-dependent potentials are in principle able to yield orbital band gaps; thus, the good results obtained with the MBJLDA potential could be considered as a surprise. However, we would like to stress that our intention was not to find a good approximation to the exact orbital-independent Kohn-Sham potential, but instead a potential which tries to mimic the behavior of orbital-dependent potentials. This is in principle feasible since the overlap between the occupied and unoccupied orbitals around the band gap is usually small, and thus an orbital-independent potential could catch the essentials of orbital-dependent potentials.

In summary, we have obtained very accurate band gaps of semiconductors and insulators with an orbital-independent exchange-correlation potential which depends solely on semilocal quantities. To our knowledge, this is the first semilocal potential that is able to compete in accuracy with the expensive hybrid and GW methods. The proposed potential leads to calculations of the band structure which are as cheap as LDA or GGA and thus can be applied to very large systems in an efficient way, which is certainly not possible with the hybrid or GW methods.

This work was supported by the project P20271-N17 of the Austrian Science Fund.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
 [2] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
 [3] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997).
 [4] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005).
 [5] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).
 [6] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[7] R. J. Magyar, A. Fleszar, and E. K. U. Gross, Phys. Rev. B **69**, 045111 (2004).
 [8] S. Sharma, J. K. Dewhurst, and C. Ambrosch-Draxl, Phys. Rev. Lett. **95**, 136402 (2005).
 [9] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).
 [10] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, J. Chem. Phys. **124**, 154709 (2006); **125**, 249901 (2006).
 [11] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).
 [12] J. Kuneš, V. I. Anisimov, S. L. Skornyakov, A. V. Lukoyanov, and D. Vollhardt, Phys. Rev. Lett. **99**, 156404 (2007).
 [13] W. G. Aulbur, M. Städele, and A. Görling, Phys. Rev. B **62**, 7121 (2000).
 [14] S. V. Faleev, M. van Schilfhaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).
 [15] M. van Schilfhaarde, T. Kotani, and S. V. Faleev, Phys. Rev. B **74**, 245125 (2006).
 [16] A. N. Chantis, M. van Schilfhaarde, and T. Kotani, Phys. Rev. B **76**, 165126 (2007).
 [17] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).
 [18] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).
 [19] F. Tran, P. Blaha, and K. Schwarz, J. Phys. Condens. Matter **19**, 196208 (2007).
 [20] A. D. Becke and E. R. Johnson, J. Chem. Phys. **124**, 221101 (2006).
 [21] A. D. Becke and M. R. Roussel, Phys. Rev. A **39**, 3761 (1989).
 [22] J. C. Slater, Phys. Rev. **81**, 385 (1951).
 [23] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
 [24] S. Hüfner, J. Osterwalder, T. Riesterer, and F. Hulliger, Solid State Commun. **52**, 793 (1984).
 [25] G. A. Sawatzky and J. W. Allen, Phys. Rev. Lett. **53**, 2339 (1984).
 [26] M. Marsman, J. Paier, A. Stroppa, and G. Kresse, J. Phys. Condens. Matter **20**, 064201 (2008).
 [27] F. Oba, A. Togo, I. Tanaka, J. Paier, and G. Kresse, Phys. Rev. B **77**, 245202 (2008).
 [28] R. Armiento, S. Kümmel, and T. Körzdörfer, Phys. Rev. B **77**, 165106 (2008).
 [29] A. P. Gaiduk and V. N. Staroverov, J. Chem. Phys. **128**, 204101 (2008).
 [30] A. V. Krukau, G. E. Scuseria, J. P. Perdew, and A. Savin, J. Chem. Phys. **129**, 124103 (2008).
 [31] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, WIEN2K: An Augmented Plane Wave and Local Orbitals Program for Calculating Crystal Properties, edited by K. Schwarz, Vienna University of Technology, Austria, 2001.
 [32] G. K. H. Madsen, P. Blaha, K. Schwarz, E. Sjöstedt, and L. Nordström, Phys. Rev. B **64**, 195134 (2001).
 [33] T. M. Schuler, D. L. Ederer, S. Itza-Ortiz, G. T. Woods, T. A. Callcott, and J. C. Woicik, Phys. Rev. B **71**, 115113 (2005).
 [34] F. Tran, P. Blaha, K. Schwarz, and P. Novák, Phys. Rev. B **74**, 155108 (2006).