Improving the modified Becke-Johnson exchange potential

David Koller, Fabien Tran, and Peter Blaha

Institute of Materials Chemistry, Vienna University of Technology, Getreidemarkt 9/165-TC, A-1060 Vienna, Austria
(Received 30 January 2012; published 5 April 2012)

The modified Becke-Johnson exchange potential [F. Tran and P. Blaha, Phys. Rev. Lett. **102**, 226401 (2009)] (TB-mBJ) yields very accurate electronic band structures and gaps for various types of semiconductors and insulators (e.g., *sp* semiconductors, noble-gas solids, and transition-metal oxides). However, the TB-mBJ potential has, for a few groups of solids, the tendency to underestimate the band gap. This has led us to examine the possibility to further improve over the original TB-mBJ potential by either reparametrizing its coefficients using a larger test set of solids or defining a parametrization for small-/medium-size band-gap semiconductors only. We also checked alternatives to the average of $|\nabla\rho|/\rho$ in the unit cell for the determination of parameter c, which determines the amount of the screening contribution. Among these different possibilities, the best one seems to be a reparametrization of the coefficients, which leads to a much more balanced description of the band gaps.

I. INTRODUCTION

A nontrivial problem, when studying the electronic structure of molecules or solids, is to find the most appropriate method, i.e., the one leading to results, which are accurate enough at a reasonable computational cost. Very often, density functional theory,[1] in its Kohn-Sham (KS) formulation,[2] represents a good compromise between accuracy and computational cost. In the KS method, the term in the total-energy functional accounting for the exchange-correlation (xc) effects must be approximated, and the reliability of the results depends mainly on the chosen approximation. The most commonly used approximate functionals for the xc energy E_{xc} belong to the local density approximation (LDA),[2] the generalized gradient approximation (GGA),[3] or the hybrid approximation.[4]

For solids, the LDA and GGA usually yield ground-state properties, which are in reasonable agreement with the experiment. However, this often is not the case for excited-state properties. For instance, it is well known that, very often, the band gap is strongly underestimated and sometimes even a metallic instead of an insulating state is obtained. There are several reasons for this. One problem is that the LDA and GGA functionals contain the self-interaction error[5] and do not show a derivative discontinuity, which is important when one wants to compare the KS band gap (the difference in the eigenvalues of the conduction-band minimum and the valence-band maximum) with the experimental band gap (the difference between the ionization potential and the electron affinity).[6,7] In this respect, the (screened) hybrid functionals lead to band gaps, which are usually in much better agreement with the experiment[8,9] but lead to calculations which are 1 or 2 orders of magnitude more expensive than with LDA or GGA. Other theoretical methods leading to more accurate excited states are LDA + U,[10] LDA + DMFT (dynamical mean-field theory),[11] and GW.[12] LDA + DMFT and GW are expensive methods, whereas, LDA + U (which is as cheap as LDA) can only be applied to localized states (typically 3*d* and 4*f* electrons).

There are two ways to construct an approximation for the xc functional. The first one is to use only known mathematical conditions satisfied by the exact functional to determine the parameters contained in the functional.[13] The second one is to tune the parameters in order to reproduce the experimental results (of one or several properties) in a chosen test set of systems (see, e.g., Ref. 14).

Recently, Becke and Johnson[15] (BJ) proposed an exchange potential, which was designed to reproduce the exact exchange potential in atoms. The BJ potential, which does not contain any empirical parameter, reads

$$v_x^{BJ}(\mathbf{r}) = v_x^{BR}(\mathbf{r}) + \frac{1}{\pi}\sqrt{\frac{5}{12}}\sqrt{\frac{2t(\mathbf{r})}{\rho(\mathbf{r})}}, \qquad (1)$$

where $\rho = \sum_{i=1}^{N}|\psi_i|^2$ is the electron density,

$$t(\mathbf{r}) = \frac{1}{2}\sum_{i=1}^{N}\nabla\psi_i^*(\mathbf{r})\cdot\nabla\psi_i(\mathbf{r}) \qquad (2)$$

is the KS kinetic-energy density, and

$$v_x^{BR}(\mathbf{r}) = -\frac{1}{b(\mathbf{r})}\left(1 - e^{-x(\mathbf{r})} - \frac{1}{2}x(\mathbf{r})e^{-x(\mathbf{r})}\right) \qquad (3)$$

is the Becke-Roussel[16] (BR) exchange potential. In Eq. (3), x is determined from a nonlinear equation involving ρ, $\nabla\rho$, $\nabla^2\rho$, and t, and then b is calculated with $b = [x^3 e^{-x}/(8\pi\rho)]^{1/3}$. The first contribution to the BJ potential, the BR potential, has been designed using the exchange hole in the hydrogen atom as a model. It has been shown[15] to reproduce Slater's averaged exchange potential[17] and is, in general, a negative (attractive) potential. The second contribution, which is positive and proportional to $\sqrt{t/\rho}$, was introduced to correct the difference between the averaged exchange potential and the exact exchange potential, obtained by applying the optimized effective potential (OEP) method[18,19] to the Hartree-Fock method. The BJ potential as the combination of these two parts is in excellent agreement with exact exchange OEP in atoms.[15]

In Ref. 20, it has been shown that, in various types of solids, the BJ potential leads to only moderately improved band gaps compared to standard LDA and GGA. However, in Ref. 21, a simple modification of the BJ potential (hereafter called

TB-mBJ) was proposed,

$$v_x^{TB\text{-}mBJ}(\mathbf{r}) = c v_x^{BR}(\mathbf{r}) + (3c - 2)\frac{1}{\pi}\sqrt{\frac{5}{12}}\sqrt{\frac{2t(\mathbf{r})}{\rho(\mathbf{r})}}, \quad (4)$$

where c is given by

$$c = A + B\sqrt{\overline{g}}, \quad (5)$$

and

$$\overline{g} = \frac{1}{V_{cell}}\int_{cell}\frac{1}{2}\left(\frac{|\nabla\rho^{\uparrow}(\mathbf{r})|}{\rho^{\uparrow}(\mathbf{r})} + \frac{|\nabla\rho^{\downarrow}(\mathbf{r})|}{\rho^{\downarrow}(\mathbf{r})}\right)d^3r \quad (6)$$

is the average of $g = |\nabla\rho|/\rho$ in the unit cell of volume V_{cell}. In Eq. (5), A and B are two free parameters whose values are $A = -0.012$ and $B = 1.023$ bohr$^{1/2}$ according to a fit to the experimental band gaps. Larger c values than 1 lead to a less negative (less attractive) potential, in particular, in low-density regions. It was shown (see Refs. 21–23) that, for band-gap calculations, the TB-mBJ potential is as accurate as the much more expensive hybrid and GW methods.

The strength of the TB-mBJ exchange potential is that, as a multiplicative potential, it can predict band gaps of semiconductors and insulators with better accuracy than any other multiplicative potential and this at a computational cost of the order of a regular GGA calculation. It already has been used many times,[24–38] and its performance has been analyzed in deep detail.[22] This analysis showed that, although in many cases its performance was very good, there also are cases where it performs less satisfyingly (e.g., for itinerant metals or Cu$_2$O). For this reason, in the present paper, we report three attempts of improving over TB-mBJ. First, we used a larger testing set of solids to optimize the parameters in Eq. (5) and also a second set whose solids have been restricted to small gap semiconductors. The second attempt is to determine c not by $\overline{g}$ but by another quantity. Third, we have checked whether the results could be improved by using a local (i.e., position dependent) $c_{loc}(\mathbf{r})$ instead of the position-independent averaged c as given by Eq. (5).

II. RESULTS AND DISCUSSION

All calculations were performed with WIEN2K,[39] a full-potential (linearized) augmented plane wave and local orbitals software for quantum calculations on periodic systems. The LDA was used for the correlation potential.[40] The **k** meshes and basis sets were chosen to be good enough such that the presented results were well converged. Spin-orbit interaction was included in those cases where it had a significant effect.

A. Different parametrizations

In this section, new parametrizations for c in Eq. (4) are searched for by using a larger testing set of solids (see Table I) than the one used in Ref. 21. They can be grouped into five categories: ionic compounds, noble-gas solids, sp semiconductors, transition-metal oxides, and other transition-metal compounds. First, in Fig. 1, we show the value of $c = c_{opt}$ in Eq. (4), which leads to the experimental band gap together with the corresponding value of $\overline{g}$ [Eq. (6)]. We recall that smaller c values than c_{opt} lead to smaller gaps, and larger c values lead to larger gaps. Several observations can be made.
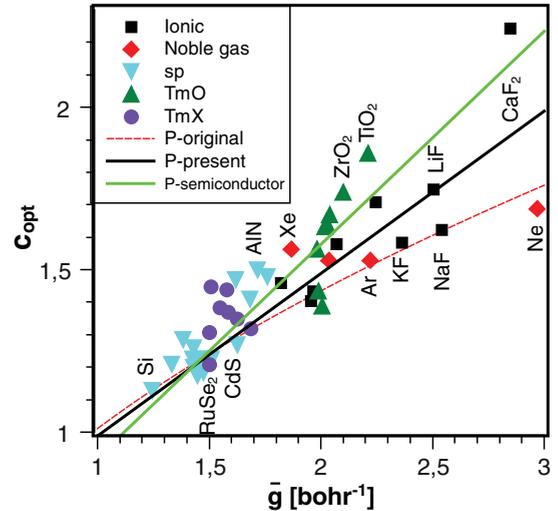
FIG. 1. (Color online) Plot of c_{opt} versus $\overline{g}$ (symbols) and several parametrizations of c as a function of $\overline{g}$ (lines).

First, the classification of our solids into five categories makes sense since the categories somehow are located in different regions in the $\overline{g} - c_{opt}$ plane. Another observation is that choosing the quantity $\overline{g}$ to parametrize c was a good choice since, from Fig. 1, we can see a rather clear linear relation between c_{opt} and $\overline{g}$ (solids with a high $\overline{g}$ also require a high c_{opt}).

In Fig. 1, the various fits of c as a function of $\overline{g}$ are also shown. We can see that the parametrization P-original given by Eq. (5) [original TB-mBJ (Ref. 21)] yields values for c, which are too small for many solids, resulting in a theoretically predicted band gap, which are too small. This particularly is the case for the transition-metal compounds, which would need a different relation between $\overline{g}$ and c. The reason for the low slope in the original TB-mBJ is the fact that very few TM compounds were included in the test set but the extreme case of Ne was. Ne has a comparably high $\overline{g}$ value of about 3, while its c_{opt} value is about 1.7, which is in the normal range. On the other hand, CaF$_2$ (whose $\overline{g}$ value is also high) was not considered in the original TB-mBJ paper, Ne was the only case in the high $\overline{g}$ region, thus, causing the low slope in the original TB-mBJ function.

Therefore, we have searched for a new parametrization for c in Eq. (4). We used the ansatz,

$$c = A + B\overline{g}^e, \quad (7)$$

which has a more general form than Eq. (5). However, we have observed that changing the value of the exponent e (from 0.2 to 1.5) only affects the quality of the fits a little, and thus, for simplicity, the exponent is fixed at $e = 1$. The parameters A and B in Eq. (7) were then determined by minimizing the mean-absolute relative error (MARE) of a selection of the solids listed in Table I. First, all solids were used to optimize the parameters. This leads to parameters which are shown in Table II and, hereafter, are referred to as P-present. The calculation of the band gaps with the P-present parameters gives the results shown in Table I, which are compared with the band gaps obtained with the original TB-mBJ parameters and with the experimental results.

TABLE I. Fundamental band gaps (in eV) calculated with the mBJ potential using different parametrizations for c (see Table II). The second column indicates the type of solid (TmO = transition-metal oxide, TmX = other transition-metal compounds).

Solid	Type	P-original	P-present	P-semiconductor	Expt.
NaF	Ionic	11.62	14.03	26	11.7[a]
KF	Ionic	10.64	12.13	14.13	10.9[a]
NaCl	Ionic	8.44	9.01	10.61	8.6[a]
KCl	Ionic	8.64	9.29	11.05	8.5[a]
CaF_2	Ionic	10.43	10.86	11.34	11.8[b]
LiF	Ionic	12.94	14.16	17.65	14.2[c]
LiCl	Ionic	8.64	8.83	9.35	9.4[c]
MgO	Ionic	7.17	7.47	8.15	7.83[c]
SiO_2	Ionic	8.89	9.41	10.64	10.3[d]
Ar	Noble gas	13.91	15.51	20.81	14.2[c]
Kr	Noble gas	10.62	11.23	12.93	11.6[c]
Xe	Noble gas	8.10	8.26	8.74	9.8[c]
C	sp	4.93	4.94	5.00	5.48[c]
Si	sp	1.17	1.10	1.00	1.17[c]
Ge	sp	0.76	0.74	0.74	0.74[c]
SiC	sp	2.28	2.26	2.25	2.4[c]
BN	sp	5.85	5.89	6.04	6.25[c]
GaN	sp	2.81	2.87	3.00	3.2[c]
GaAs	sp	1.54	1.52	1.56	1.52[c]
CdS	sp	2.66	2.71	2.90	2.55[e]
AlN	sp	5.55	5.58	5.70	6.13[e]
AlP	sp	2.32	2.27	2.21	2.45[c]
InP	sp	1.60	1.57	1.59	1.42[f]
InAs	sp	0.59	0.57	0.60	0.42[f]
InSb	sp	0.27	0.24	0.24	0.24[f]
GaSb	sp	0.74	0.70	0.70	0.81[f]
AlAs	sp	2.05	2.01	1.98	2.32[g]
CdTe	sp	1.55	1.54	1.60	1.475[h]
ZnO	TmO	2.71	2.89	3.26	3.44[c]
HfO_2	TmO	5.83	6.13	6.65	5.7[i]
$SrTiO_3$	TmO	2.70	2.81	3.04	3.25[j]
TiO_2	TmO	2.57	2.69	2.91	3.3[k]
ZrO_2	TmO	4.73	4.86	5.14	5.5[i]
NiO	TmO	4.25	4.66	5.06	4.3[c]
MnO	TmO	2.98	3.18	3.62	3.9[c]
FeO	TmO	1.83	2.02	2.40	2.4[c]
ZnS	TmX	3.66	3.71	3.86	3.91[c]
RuS_2	TmX	1.28	1.27	1.28	1.38[l]
$RuSe_2$	TmX	0.80	0.79	0.81	0.76[m]
ScN	TmX	0.90	0.94	1.07	0.9[c]
ZrS_2	TmX	1.26	1.25	1.28	1.69[n]
HfS_2	TmX	1.64	1.65	1.71	1.94[n]
$HfSe_2$	TmX	0.82	0.82	0.85	1.14[n]
MoS_2	TmX	1.07	1.07	1.07	1.29[d]

[a]Reference 41.
[b]Reference 42.
[c]Reference 21.
[d]Reference 23.
[e]Reference 9.
[f]Reference 31.
[g]Reference 43.
[h]Reference 25.
[i]Reference 44.
[j]Reference 45.
[k]Reference 46.
[l]Reference 47.
[m]Reference 48.
[n]Reference 49.

TABLE II. Summary of the different parametrizations for c [Eq. (7)].

Parametrization	A	B	e
P-original	-0.012	1.023	0.5
P-present	0.488	0.500	1
P-semiconductor	0.267	0.656	1

In Fig. 2, the errors in the calculated band gaps with respect to the experiment are shown, and Table III lists the mean error (ME), mean-relative error (MRE), mean-absolute error (MAE), and MARE. From Fig. 2, we can see that, for the case of the other transition-metal compounds (TmX) and the sp semiconductors, there is not much change. The reason for that is that their values for $\overline{g}$ lie in a region where the P-original and P-present parameters lead to similar c. They are already quite well described by the original TB-mBJ so that there is not much room for improvement. The values of $\overline{g}$ for the solids in the other three categories lie in a region where the two parametrizations lead to a larger difference in c. In this region, the P-present parameters lead to larger c and, therefore, to larger band gaps. Since the band-gap values of most solids in these categories are too low, the new parameters lead to better values. However, there are solids, such as NaF for which the original TB-mBJ potential is very good such that the new parameters overestimate the band gap.

Actually, the statistical values in Table III reflect this trend. The MARE is only marginally better when the new parameters are used, while the MAE is even slightly worse. The signed errors (ME and MRE) for P-present, however, show a clear improvement over the original TB-mBJ. This is not surprising
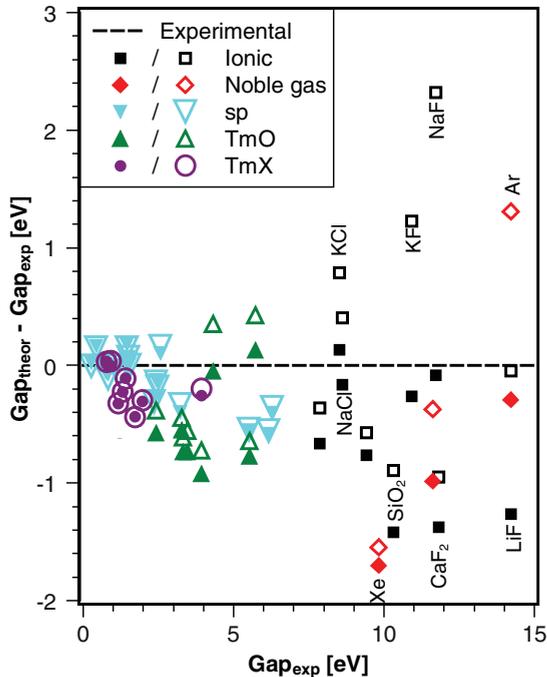
since the original TB-mBJ leads, in almost all cases, to values below the experiment, whereas, P-present underestimates some gaps and overestimates others, making it more balanced overall than the original TB-mBJ.

While mBJ with the P-present parameters works well for all five material classes, from the point of view of technical relevance, good prediction of semiconductors, such as Si or CdTe is much more important than the gap of Ar, and thus, it makes sense to search for parameters which are optimized for small-gap semiconductors. If we arbitrarily consider 7 eV as the border between small and large gaps, the small-gap solids in our list are the sp semiconductors, the transition-metal oxides, and the other transition-metal compounds, whereas, the ionic compounds and the noble-gas solids are large-gap solids. Already, Fig. 1 suggests that the P-present parameters are not optimal for the small-gap semiconductors, although they are better than original TB-mBJ. Instead, we searched for new parameters which are optimal for the set of sp semiconductors, transition-metal oxides, and other transition-metal compounds. These parameters (P-semiconductors) are also listed in Table II. Figure 1 shows that, in the high $\overline{g}$ region, where the ionic compounds and the noble-gas solids need a lower c than the solids in the P-semiconductor group, the P-semiconductor parameters lead to a higher c than the P-present parameters so that they are better suited for the small-gap semiconductors. However, this leads to severe overestimations for some of the ionic compounds and noble-gas solids. Furthermore, in the case of NaF, it leads to convergence problems in the self-consistency cycle.

As can be seen in Fig. 3, the use of the P-semiconductor parameters leads in most cases to better agreement with the experiment than both original TB-mBJ and P-present. The statistical data in Table III also show the superiority of the P-semiconductor parameters. While changing from P-original to P-present already reduces the MARE of the semiconductor test set, going to P-semiconductor leads to a further reduction in the MARE close to 10%, which is quite good for such small band gaps. Comparing the ME and the MRE of the three parametrizations applied to the semiconductor test set is a very strong indication that the P-semiconductor parameters are the best in this case. Especially the ME, whose absolute value is below 0.1 eV, is of a precision hardly reachable by any other method.

B. Other quantities to determine c

Although the results from Sec. II A are excellent in most cases, there is still room for improvement. One possibility, which could lead to even better results, is to determine c not from $\overline{g}$ but instead from another system-specific quantity. An example would be the average value $\overline{s}$ of the reduced density gradient $s = |\nabla \rho|/[2(3\pi^2)^{1/3}\rho^{4/3}]$, which differs from the definition of g in the power of the density in the denominator (1 for g versus 4/3 for s). However, when comparing Fig. 1 with Fig. 4, this idea does not turn out to be very useful since there is a large collection of points in the region $0.5 < \overline{s} < 1$ with very different corresponding values of c_{opt}, while in the range $1 < \overline{s} < 1.6$, there are two separate branches. It means that a nice fit, such as in Fig. 1, for the case of $\overline{g}$, is clearly

FIG. 2. (Color online) Difference between calculated and experimental band gaps for the different categories of solids and different parametrizations. Full symbols use P-original, open symbols use the new P-present parameters.

TABLE III. Statistics on the fundamental band gaps for the different parametrizations of c and the testing set of the solids. Here, "semiconductors" means those compounds with an experimental band gap of less than 7 eV. ME and MAE are in eV, and MRE and MARE are in percentages.

Parametrization	Testing set	ME	MRE	MAE	MARE
P-original	All	−0.37	−6.51	0.42	10.51
P-present	All	−0.11	−4.20	0.45	10.43
P-original	Semiconductor	−0.24	−6.40	0.29	11.80
P-present	Semiconductor	−0.19	−6.03	0.28	11.18
P-semiconductor	Semiconductor	−0.07	−2.10	0.24	10.82

not possible. Actually, this result might lead one to think that, instead of increasing the exponent in the denominator, decreasing it might help. However, in the end, it turns out that the best choice for the exponent of ρ in the denominator is 1 as used in Sec. II A.

Two other possibilities that we have tested involve electron-density-based approximations to the kinetic-energy density, namely, the Weizsäcker kinetic-energy density,

$$t_W(\mathbf{r}) = \frac{1}{8} \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})}, \qquad (8)$$

and the Thomas-Fermi kinetic-energy density,

$$t_{TF}(\mathbf{r}) = \frac{3}{10} (3\pi)^{2/3} \rho^{5/3}(\mathbf{r}). \qquad (9)$$

The first quantity is $\overline{\alpha}$, the average of

$$\alpha(\mathbf{r}) = \frac{t(\mathbf{r}) - t_W(\mathbf{r})}{t_{TF}(\mathbf{r})}, \qquad (10)$$

which is used in the meta-GGA Tao-Perdew-Staroverov-Scuseria (Ref. 50), and the second one is the average of $t_W(\mathbf{r})/t(\mathbf{r})$, which has been used in local hybrid functionals

where the relative weights of exact and semilocal exchange are position dependent.[51] We have observed that both these approaches have similar problems as $\overline{s}$ and, therefore, do not lead to improvements compared to $\overline{g}$.

C. Using local values of c

$g(\mathbf{r}) = |\nabla \rho(\mathbf{r})|/\rho(\mathbf{r})$ can vary strongly throughout the unit cell as shown in Fig. 5 for the case of LiF. It has a large value in the anion core region and a lower one in the anion valence region, whereas, for the cation, $g(\mathbf{r})$ is rather large in the $1s$ region, but outside the core, it is much smaller. In essence, the average in both atomic spheres is very similar ($\overline{g}_{Li} = 2.79$ bohr^{-1}, $\overline{g}_F = 2.88$ bohr^{-1}). However, between the atoms, where the density has minima and saddle points, $g(\mathbf{r})$ even goes to zero. This means that, while in traditional TB-mBJ, the constant c in Eq. (4) is determined by the average of $g(\mathbf{r})$, it is also possible to replace this by a position dependent $c(\mathbf{r})$ based on $g(\mathbf{r})$. There are two reasons to consider this idea. First, it may lead to improved results compared to a constant c. Second, for systems, such as molecules or surfaces, the region over which $g(\mathbf{r})$ should be averaged is not clearly defined. Since the idea of a constant c, determined by the average of $g(\mathbf{r})$, leads to good results, we tested two ways which use $g(\mathbf{r})$ as a basis for the construction of $c(\mathbf{r})$. In the first strategy, the unit cell is divided into atomic spheres and an interstitial region in the same way as it is
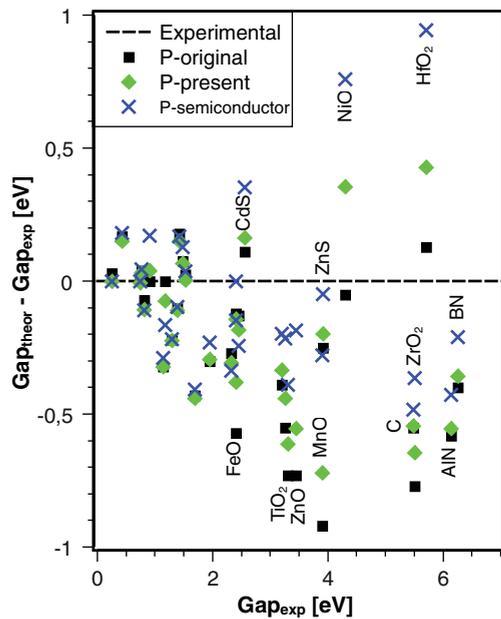
FIG. 3. (Color online) Difference between calculated and experimental band gaps for small-gap semiconductors and different parametrizations.
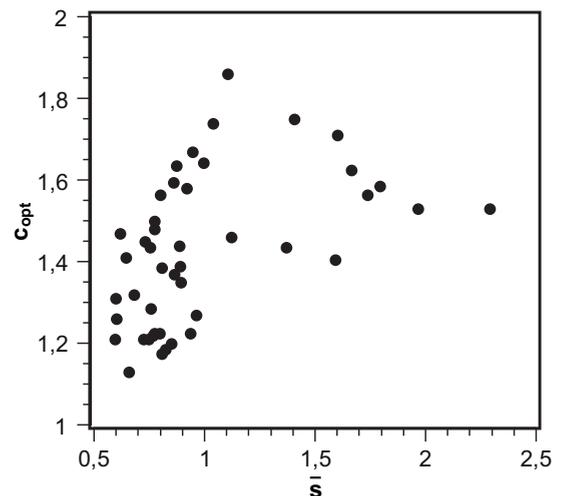
FIG. 4. Plot of c_{opt} versus $\overline{s}$.
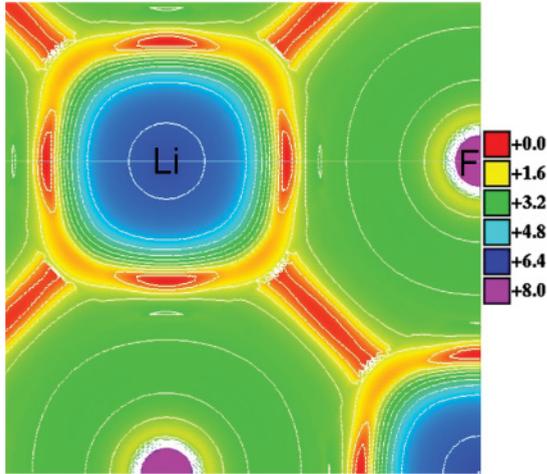
FIG. 5. (Color online) $g(\mathbf{r})$ in the LiF-(001) plane. The numbers are given in bohr^{-1}.

FIG. 6. (Color online) Comparison of band gaps predicted by the original TB-mBJ method and an approach, which includes a semilocal $c_{SL}(\mathbf{r})$ as described in Sec. II C with experimental results (see Table I).

performed in the augmented-plane-wave method.[52] $g(\mathbf{r})$ then is averaged separately for each sphere and for the interstitial. From these averages, a semilocal c_{SL} is obtained with Eq. (5) and is applied to its corresponding region. It is important to keep in mind that the results obtained with this strategy depend on the choice of the radius of the atomic spheres. However, we just wanted a comparison with the original TB-mBJ and, in a first step, ignored this dependence. The second scheme uses a fully local $c_{loc}(\mathbf{r})$, which is determined in the following way. Already, since a constant c (calculated from $\overline{g}$) leads to (very) good results, the idea is to have a $c_{loc}(\mathbf{r})$, which does not deviate too much from the constant c. We chose to keep $c_{loc}(\mathbf{r})$ in a ±10% range around the constant c: The minimum of $g(\mathbf{r})$ is 0, and there we set our local $c_{loc}(\mathbf{r}) = 0.9c$. Where $g(\mathbf{r})$ has its average value $\overline{g}$, we set $c_{loc}(\mathbf{r}) = c$. This is achieved by the relation $c_{loc}(\mathbf{r}) = c + \lambda[g(\mathbf{r}) - \overline{g}]$ with $\lambda = 0.1c/\overline{g}$. Furthermore, a cutoff is introduced which does not allow $c_{loc}(\mathbf{r})$ to become larger than $1.1c$.

Figure 6 shows the band gaps of some solids obtained by the c_{SL} approach and compares it to original TB-mBJ and experimental values. The reasons why these samples were chosen are the following: Kr, HfO$_2$, and MnO require very different c_{opt} of 1.51, 1.39, and 1.64, respectively, but have similar $\overline{g}$ of around 2. This means that any relation between $\overline{g}$ and c cannot treat them differently, while a local approach might be helpful. CaF$_2$ and LiF are included because their gaps are underestimated strongly by TB-mBJ, while those of KCl, InAs, and InP are overestimated. This means that these are the most demanding cases where improvement is the most necessary. As Fig. 6 shows, changing from TB-mBJ to the semilocal strategy does not improve the description of the gaps. For solids, such as CaF$_2$ or LiF, whose gaps are already underestimated by TB-mBJ, the underestimation becomes even more pronounced when using the local approach. The reverse is observed for HfO$_2$, InAs, and InP where the band gap increases when changing to the local approach so that it is not possible to get an improvement even by optimizing the parameters again. The trend is very similar for the fully local $c_{loc}(\mathbf{r})$ strategy, which, therefore, is not included in Fig. 6.
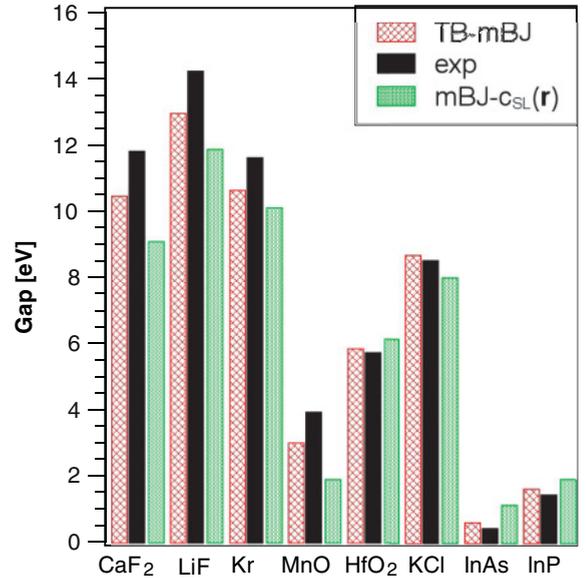
However, the differences for TB-mBJ are smaller because the change in $c_{loc}(\mathbf{r})$ was restricted to 10%.

In order to understand why the approach of a different c is problematic, we take a look at two examples where the effect due to a local c is different: LiF and InAs. Table IV shows some quantities relevant for this discussion. In the case of LiF, the VBM electrons are located mainly in the anion spheres, whereas, a large part of the CBM states is located in the interstitial region where the c value is lower as explained in the discussion related to Fig. 5. In InAs, the VBM states are not localized so much in the anion spheres, but a considerable part is also in the interstitial region. In LiF, both atoms have

TABLE IV. Fractions of the charge densities of the valence-band maximum (VBM) and conduction-band minimum (CBM) in LiF and InAs together with values for c from TB-mBJ [Eq. (5)], c_{opt}, which reproduces the experimental gap, and c_{SL} calculated by Eq. (5) using $\overline{g}$ in the atomic spheres (RMT$_{Li}$ = 1.88, RMT$_F$ = 1.88, RMT$_{In}$ = 2.46, RMT$_{As}$ = 2.46 bohr) and in the interstitial. c_{eff} VBM and c_{eff} CBM are explained in the text.

	LiF	InAs
VBM cation charge [%]	3	14
VBM anion charge [%]	87	52
VBM interstitial charge [%]	10	34
CBM cation charge [%]	22	25
CBM anion charge [%]	22	26
CBM interstitial charge [%]	56	51
c (TB-mBJ)	1.57	1.23
c_{opt}	1.75	1.19
c_{SL} cation	1.70	1.58
c_{SL} anion	1.72	1.35
c_{SL} interstitial	1.36	1.03
c_{eff} VBM	1.68	1.27
c_{eff} CBM	1.51	1.27

a similar c_{SL} value in the atomic spheres, whereas, in InAs, the c_{SL} value is considerably higher in In than in As. We can calculate an effective c value for the VBM and CBM states of both compounds by averaging the atomic and interstitial c_{SL} values weighted with the charge fraction. As shown in Table IV, in the case of LiF, the semilocal strategy increases the effective c for the VBM states and decreases it for CBM compared to TB-mBJ. This results in a reduced gap. In the case of InAs, both c_{eff} are increased slightly, leading to a larger gap.

The fact that a local $c(\mathbf{r})$ does not improve the results is not really surprising. An analysis by Marques *et al.*[23] concluded that screening is a nonlocal effect and cannot be described properly by a local parameter. This does not mean that mBJ cannot be used for finite systems. However, the local parameter should be determined by averaging $g(\mathbf{r})$ over a large enough region which can, for instance, be defined by including an exponential function in Eq. (6), which decays over a few atomic distances.

III. SUMMARY AND CONCLUSION

In this paper, three different strategies to further improve the performance of the TB-mBJ potential have been presented: (a) reoptimizing parameters A and B in Eq. (7), (b) using a different quantity than $\overline{g}$ (e.g., $\overline{s}$) for the determination of c, and (c) considering a position dependent $c_{loc}(\mathbf{r})$ instead of a constant one.

Out of these three possibilities, the second and third did not turn out to be useful. However, reparametrizing the coefficients using more appropriate test sets of solids leads to more balanced results. In particular, if the test set is restricted to the technologically relevant solids (band gap of less than 7 eV), the accuracy of the results can hardly be reached by any other methods including the very expensive GW.

ACKNOWLEDGMENTS

[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, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997).

[4] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).

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

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

[7] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[8] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).

[9] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005).

[10] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

[11] J. Kuneš, V. I. Anisimov, S. L. Skornyakov, A. V. Lukoyanov, and D. Vollhardt, Phys. Rev. Lett. **99**, 156404 (2007).

[12] F. Bechstedt, F. Fuchs, and G. Kresse, Phys. Status Solidi B **246**, 1877 (2009).

[13] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, J. Chem. Phys. **123**, 062201 (2005).

[14] F. A. Hamprecht, A. J. Cohen, D. J. Tozer, and N. C. Handy, J. Chem. Phys. **109**, 6264 (1998).

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

[16] A. D. Becke and M. R. Roussel, Phys. Rev. A **39**, 3761 (1989).

[17] J. C. Slater, Phys. Rev. **81**, 385 (1951); **82**, 538 (1951).

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

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

[20] F. Tran, P. Blaha, and K. Schwarz, J. Phys.: Condens. Matter **19**, 196208 (2007).

[21] F. Tran and P. Blaha, Phys. Rev. Lett. **102**, 226401 (2009).

[22] D. Koller, F. Tran, and P. Blaha, Phys. Rev. B **83**, 195134 (2011).

[23] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining, and S. Botti, Phys. Rev. B **83**, 035119 (2011).

[24] W. Al-Sawai, H. Lin, R. S. Markiewicz, L. A. Wray, Y. Xia, S.-Y. Xu, M. Z. Hasan, and A. Bansil, Phys. Rev. B **82**, 125208 (2010).

[25] W. Feng, D. Xiao, Y. Zhang, and Y. Yao, Phys. Rev. B **82**, 235121 (2010).

[26] W. Feng, D. Xiao, J. Ding, and Y. Yao, Phys. Rev. Lett. **106**, 016402 (2011).

[27] S.-D. Guo and B.-G. Liu, Europhys. Lett. **93**, 47006 (2011).

[28] D. J. Singh, Phys. Rev. B **82**, 155145 (2010).

[29] D. J. Singh, S. S. A. Seo, and H. N. Lee, Phys. Rev. B **82**, 180103(R) (2010).

[30] D. J. Singh, Phys. Rev. B **82**, 205102 (2010).

[31] Y.-S. Kim, M. Marsman, G. Kresse, F. Tran, and P. Blaha, Phys. Rev. B **82**, 205212 (2010).

[32] A. F. Lima, S. A. S. Farias, and M. V. Lalic, J. Appl. Phys. **110**, 083705 (2011).

[33] S.-D. Guo and B.-G. Liu, J. Appl. Phys. **110**, 073525 (2011).

[34] P. V. Smith, M. Hermanowicz, G. A. Shah, and M. W. Radny, Comput. Mater. Sci. **52**, 37 (2012).

[35] H. Dixit, N. Tandon, S. Cottenier, R. Saniz, D. Lamoen, B. Partoens, V. Van Speybroeck, and M. Waroquier, New J. Phys. **13**, 063002 (2011).

[36] M. César, Y. Ke, W. Ji, H. Guo, and Z. Mi, Appl. Phys. Lett. **98**, 202107 (2011).

[37] N. Yedukondalu, K. R. Babu, C. Bheemalingam, D. J. Singh, G. Vaitheeswaran, and V. Kanchana, Phys. Rev. B **83**, 165117 (2011).

[38] A. Ghafari, A. Boochani, C. Janowitz, and R. Manzke, Phys. Rev. B **84**, 125205 (2011).

[39] 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* (Vienna University of Technology, Austria, 2001).

[40] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[41] S. C. Erwin and C. C. Lin, J. Phys. C **21**, 4285 (1988).

[42] L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999).

[43] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, Phys. Rev. B **82**, 115106 (2010).

[44] E. Bersch, S. Rangan, R. A. Bartynski, E. Garfunkel, and E. Vescovo, Phys. Rev. B **78**, 085114 (2008).

[45] K. van Benthem, C. Elsässer, and R. H. French, J. Appl. Phys. **90**, 6156 (2001).

[46] Y. Tezuka, S. Shin, T. Ishii, T. Ejima, S. Suzuki, and S. Sato, J. Phys. Soc. Jpn. **63**, 347 (1994).

[47] Y.-S. Huang and S.-S. Lin, Mater. Res. Bull. **23**, 277 (1988).

[48] H. P. Vaterlaus, R. Bichsel, F. Levy, and H. Berger, J. Phys. C **18**, 6063 (1985).

[49] H. Jiang, J. Chem. Phys. **134**, 204705 (2011).

[50] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[51] J. Jaramillo, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 1068 (2003).

[52] J. C. Slater, Phys. Rev. **51**, 846 (1937).