ຽ

Band convergence and linearization error correction of all-electron *GW* calculations: The extreme case of zinc oxide

Christoph Friedrich, Mathias C. Müller, and Stefan Blügel

Peter Grünberg Institut and Institute for Advanced Simulation, Forschungszentrum Jülich and JARA, 52425 Jülich, Germany
(Received 18 December 2010; revised manuscript received 16 January 2011; published 15 February 2011)

Recently, Shih *et al.* [Phys. Rev. Lett. **105**, 146401 (2010)] published a theoretical band gap for wurtzite ZnO, calculated with the non-self-consistent *GW* approximation, that agreed surprisingly well with experiment while deviating strongly from previous studies. They showed that a very large number of empty bands is necessary to converge the gap. We reexamine the *GW* calculation with the full-potential linearized augmented-plane-wave method and find that even with 3000 bands the band gap is not completely converged. A hyperbolical fit is used to extrapolate to infinite bands. Furthermore, we eliminate the linearization error for high-lying states with local orbitals. In fact, our calculated band gap is considerably larger than in previous studies, but somewhat smaller than that of Shih *et al.*.

In a recent Letter Shih *et al.*[1] presented a new calculation for the band gap of ZnO in the wurtzite structure employing the *GW* approximation[2] for the electronic self-energy. They used a conventional non-self-consistent, *one-shot* approach[3,4] in which neither the quasiparticle energies nor the quasiparticle wave functions are updated. Instead, the *GW* self-energy $\Sigma^\sigma(\mathbf{r},\mathbf{r}';\omega)$ [Eq. (4)] is constructed from the Kohn-Sham Green function taken from a density-functional theory calculation with the local-density approximation (LDA) for the exchange-correlation energy functional. The quasiparticle energy $E_{n\mathbf{q}}^\sigma$ with band index n, Bloch vector $\mathbf{q}$, and spin σ is then obtained from the nonlinear equation

$$E_{n\mathbf{q}}^\sigma = \epsilon_{n\mathbf{q}}^\sigma + \left\langle \varphi_{n\mathbf{q}}^\sigma \middle| \Sigma^\sigma\left(E_{n\mathbf{q}}^\sigma\right) - v_{\mathrm{xc}}^\sigma \middle| \varphi_{n\mathbf{q}}^\sigma \right\rangle \qquad (1)$$

with the Kohn-Sham energy $\epsilon_{n\mathbf{q}}^\sigma$, the wave function $\varphi_{n\mathbf{q}}^\sigma(\mathbf{r})$, and the exchange-correlation potential $v_{\mathrm{xc}}^\sigma(\mathbf{r})$. Off-diagonal elements of $\Sigma^\sigma - v_{\mathrm{xc}}^\sigma$ are neglected.

All previous calculations[5–8] invoking the one-shot LDA + *GW* approach showed that the band gap of wurtzite ZnO is underestimated with respect to the experimental value by more than 1 eV. They fall in the range 2.12–2.6 eV while the experimental gap amounts to 3.6 eV,[9] after correction for vibrational effects. This large underestimation is untypical for *GW* calculations of *sp*-bound systems.

The Letter of Shih *et al.* addressed two issues: first, the erroneous hybridization effects between Zn 3*d* and O 2*p* states that results from the self-interaction error within the LDA,[10] and second, the band convergence in the correlation part of the self-energy. The first problem was tackled with the LDA + *U* approach,[11] in which an orbital-dependent potential corrects the position of the 3*d* bands and, thus, reduces hybridization effects with the O 2*p* states. However, the combination LDA + *U* and *GW* yields a band gap that is still well below the experimental value. Therefore, the authors investigated the second issue by carefully converging the correlation self-energy and the dielectric matrix with respect to the number of bands. They performed calculations with up to 3000 bands corresponding to a maximal band energy of 67 Ry as well as a cutoff for the dielectric matrix of up to 80 Ry and showed

that the resulting *GW* band gaps, 3.4 eV for LDA + *GW* and 3.6 eV for LDA + *U* + *GW*, turned out to be in very good agreement with experiment. They also demonstrated that a too small energy cutoff for the dielectric matrix can lead to a false convergence behavior: the band gap seems to converge with many fewer bands, but toward a value that is too small.

These new results for ZnO are in striking contrast to previous studies. If they are correct, they cast doubt on all *GW* calculations published so far, not only for ZnO but also for other materials, especially for systems with localized states. In fact, Shih *et al.* point out at the end of their paper that "many of the previous quasiparticle calculations . . . involving localized states may need to be reexamined." This will likely provoke a controversial debate that reaches different fields of solid state physics and requests a rapid clarification.

Since its publication the paper has aroused criticism, mostly from the all-electron community who pointed out that calculations with the pseudopotential approximation cannot give a definite answer for the real *GW* gap because of the neglect of the core-valence exchange effects, the pseudized form of the valence wave functions, and the inaccurate description of high-lying states. Therefore, they attributed the large difference between the new result and the previous all-electron calculations[5–7] (2.12–2.44 eV) to the approximations inherent to the pseudopotential approach or to numerical errors in the calculation. It is, thus, vitally important that the *GW* band gap of ZnO is reinvestigated and thoroughly converged with a genuine all-electron method to provide a standard for the discussion that will follow.

In this paper, we present an all-electron LDA + *GW* calculation for wurtzite ZnO that is based on the full-potential linearized augmented-plane-wave (FLAPW) method.[12] For simplicity, we restrict ourselves to the standard LDA approach for the noninteracting starting point without an additional Hubbard *U* parameter. Our calculation yields a band gap that is, in fact, much larger than that of the previous calculations, but somewhat smaller than the result of Shih *et al.* We go beyond their approach in two respects: we employ neither the

pseudopotential approximation nor a plasmon-pole model for the dielectric matrix. Instead, the screened interaction

$$W(\mathbf{r},\mathbf{r}';\omega) = v(\mathbf{r},\mathbf{r}') + \int\int v(\mathbf{r},\mathbf{r}'')P(\mathbf{r}'',\mathbf{r}''';\omega)$$
$$\times W(\mathbf{r}''',\mathbf{r}';\omega)\, d^3r''\, d^3r''' \qquad (2)$$

is calculated explicitly within the random-phase approximation for the polarization function

$$P(\mathbf{r},\mathbf{r}';\omega) = -\frac{i}{2\pi}\sum_\sigma \int_{-\infty}^{\infty} G^\sigma(\mathbf{r},\mathbf{r}',\omega')$$
$$\times G^\sigma(\mathbf{r}',\mathbf{r},\omega'-\omega)\, d\omega', \qquad (3)$$

where the Green function $G^\sigma(\mathbf{r},\mathbf{r}',\omega)$ is constructed from Kohn-Sham energies and wave functions. The frequency convolution of the self-energy

$$\Sigma^\sigma(\mathbf{r},\mathbf{r}';\omega) = \frac{i}{2\pi}\int_{-\infty}^{\infty} G^\sigma(\mathbf{r},\mathbf{r}';\omega+\omega')W(\mathbf{r},\mathbf{r}';\omega')e^{i\eta\omega'}\, d\omega' \qquad (4)$$

(η is a positive infinitesimal) is evaluated analytically for $v(\mathbf{r},\mathbf{r}')$ [see Eq. (2)], and with a contour integration[4,13] on the complex frequency plane for the remainder $W(\mathbf{r},\mathbf{r}';\omega) - v(\mathbf{r},\mathbf{r}')$. The nonlinear equation (1) is solved on an energy mesh around $\epsilon_{n\mathbf{q}}^\sigma$ with a cubic spline interpolation between the mesh points. Thus, no additional Taylor expansion of the self-energy is needed. Details of the implementation can be found in Ref. 14.

We carefully converged the number of empty bands for the calculation of both the polarization function and the correlation self-energy as well as the parameters for the all-electron mixed product basis,[14,15] in which we represent the dielectric matrix. While the ground-state electron density was converged with a standard LAPW basis with moderate cutoff parameters ($l_{\max} = 8$, $G_{\max} = 4.3a_0^{-1}$, where a_0 is the Bohr radius), we had to employ much larger cutoffs to generate enough wave functions for the GW calculation: a linear momentum cutoff of $G_{\max} = 8.0a_0^{-1}$ and an angular momentum cutoff in the muffin-tin (MT) spheres of $l_{\max} = 12$. Furthermore, in order to avoid linearization errors in the MT part of the LAPW basis,[16,17] we added local orbitals[18,19] (LOs) with different angular momentum quantum numbers and energy parameters distributed over the relevant energy range: 292 LOs for Zn (five LOs for each lm channel with $l = 0$–3, three for $l = 4$, two for $l = 5$, and one for $l = 6$) and 186 for O (four LOs for $l = 0$–3, two for $l = 4$, and one for $l = 5$). We also treat the $3s$ and $3p$ semicore states of Zn explicitly with LOs.

For the mixed product basis we found an angular momentum cutoff of 4 in the spheres and a suprisingly small linear momentum cutoff of $2.4a_0^{-1}$ in the interstitial region to be sufficient. However, we had to take into account many product functions in the MT spheres, which after optimization led to 177 MT functions for Zn (ten, eight, eight, seven, and six radial functions per lm channel for $l = 0$–4, respectively) and 127 for O (eight, six, six, five, and four radial functions for $l = 0$–4, respectively). Obviously, the rapid variations close to the atomic nuclei must be accurately described. Within the all-electron mixed product basis this is possible

with a relatively modest number of MT functions, while in a pure plane-wave approach a very large number of plane waves is necessary to resolve the variations adequately. This explains the finding of Shih *et al.* that the dielectric matrix must be converged to very large energy cutoffs. The total number of mixed product basis functions in the calculations is less than 700 per $\mathbf{k}$ point. This number is further reduced to around 490 by constructing linear combinations that are continuous in value and radial derivative at the MT sphere boundaries.[20]

Figure 1 shows the quasiparticle band gap of ZnO as a function of the number of states included in the calculation of the polarization function and the correlation self-energy. The calculations were performed with a $4 \times 4 \times 4$ $\mathbf{k}$-point sampling of the Brillouin zone. There is a large difference between calculations with (pluses) and without the LOs for unoccupied states (crosses), which shows the importance of eliminating the linearization error of the conventional LAPW basis. As the linearization error becomes larger for higher and higher bands, it is not surprising that the difference between the convergence curves grows toward increasing numbers of bands. We find an asymptotic difference of 0.5 eV. The calculations without LOs for unoccupied states already converge with a few hundred bands, which could have led the authors of the previous all-electron studies to believe that their calculations are sufficiently converged. In fact, when the converged value of 2.27 eV is corrected with respect to finer $\mathbf{k}$-point samplings, we arrive at 2.44 eV, which lies at the upper edge of the range of the all-electron GW band gaps published so far.

The linear momentum cutoff allows the interstitial part of the LAPW basis to be converged in a systematic way. In the MT spheres, however, the basis is linearized around predefined energy parameters, which gives rise to the linearization error
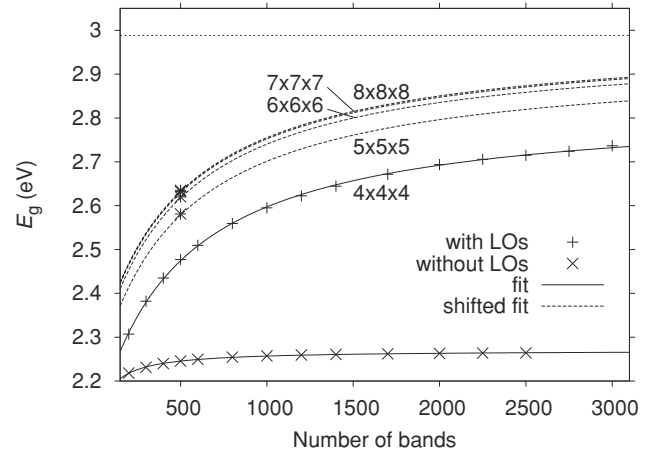

FIG. 1. Band convergence of the quasiparticle band gap of ZnO employing a $4 \times 4 \times 4$ $\mathbf{k}$-point set and calculated with (pluses) and without local orbitals (LOs) (crosses) for high-lying states. The solid lines show the hyperbolical fits. We also indicate results with finer $\mathbf{k}$-point samplings (stars) calculated with LOs and 500 bands. The dashed lines show the hyperbolical fit shifted to align with these results. The fit asymptote for the $8 \times 8 \times 8$ $\mathbf{k}$-point set at 2.99 eV (dotted line) is considered the best estimate for the all-electron one-shot GW band gap.
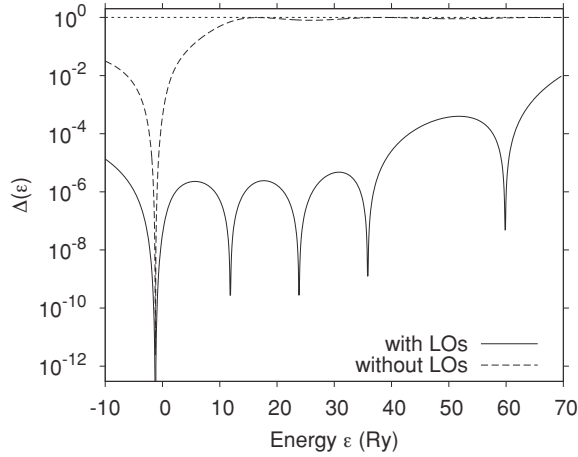
FIG. 2. Representation error $\Delta(\epsilon)$ (see text) of the set of spherical MT functions at an oxygen atom with (solid line) and without local orbitals (LOs) (dashed line) as a function of energy. The Fermi energy is set to $\epsilon = 0$. The "spikes" clearly identify the positions of the energy parameters. In the case without LOs, the representation error soon approaches unity (dotted line) for energies above 10 Ry, which means that the basis becomes practically orthogonal to the exact solution of the radial scalar-relativistic Dirac equation.

for high-lying states. We now analyze this error in more detail and show that it can be eliminated very effectively with the LOs. Figure 2 shows the representation error $\Delta(\epsilon)$ of the MT basis at an oxygen atom for the angular momentum $l = 0$ as a function of energy. We define

$$\Delta(\epsilon) = \int_0^S r^2 [R(\epsilon,r) - R_{\mathrm{r}}(\epsilon,r)]^2 dr, \qquad (5)$$

where S is the oxygen MT sphere radius, $R(\epsilon,r)$ is the normalized solution of the radial scalar-relativistic Dirac equation (cf. Ref. 17) for the energy ϵ and the angular momentum $l = 0$, and $R_{\mathrm{r}}(\epsilon,r)$ is its best representation in terms of linear combination of the MT functions contained in the LAPW basis. In the conventional basis there are only two radial functions available: the solution $R(\epsilon_{\mathrm{par}},r)$ of the radial Dirac equation for $l = 0$ at the energy parameter ϵ_{par}, in this case $\epsilon_{\mathrm{par}} = -1.24$ Ry, and its energy derivative $dR(\epsilon_{\mathrm{par}},r)/d\epsilon_{\mathrm{par}}$. [In the literature these functions are commonly denoted by $u_{l=0}(r)$ and $\dot{u}_{l=0}(r)$, respectively.] The dashed line in Fig. 2 shows that this conventional basis represents the occupied states very accurately, which are located in the small energy range $-1.33 \leqslant \epsilon \leqslant 0$ Ry, but fails to describe unoccupied states at higher energies. In fact, at energies above 10 Ry the exact solution $R(\epsilon,r)$ becomes practically orthogonal to the MT basis, and the representation error approaches unity. Interestingly, at this energy, which roughly corresponds to the 500th band, the convergence curve for the calculations without the LOs levels off, and the band gap seemingly converges (see Fig. 1). The description of wave functions at higher energies is considerably improved, if LOs are added to the LAPW basis. The solid line shows the corresponding representation error with four additional LOs at 12, 24, 36, and 60 Ry. As can be seen, the error remains below 10^{-3} for a very large energy range up to 65 Ry.

As Fig. 1 shows, the calculations with eliminated linearization error (pluses) take far more bands to converge. In fact, even with 3000 bands the band gap is still not completely converged. Therefore, we fitted the values with the hyperbolical function

$$f(N) = \frac{a}{N - N_0} + b, \qquad (6)$$

where a, b, and N_0 are fit parameters. It is surprising how closely the fitted curve (solid line) follows the calculated data points. This makes us confident in taking the fit asymptote b as the band gap extrapolated to infinite bands. Furthermore, we have recalculated the band gap with finer **k**-point meshes (up to $8 \times 8 \times 8$) and 500 bands (crosses). The dashed lines show correspondingly shifted hyperbolical fits. The asymptote of the fit corresponding to an $8 \times 8 \times 8$ **k**-point sampling is found at 2.99 eV, which we take as the final best estimate for the all-electron one-shot GW band gap.

This band gap is 0.4–0.9 eV larger than the previously reported values. Both the large number of bands that are needed for proper convergence and the elimination of the linearization error, which has not been undertaken in the previous all-electron studies, are responsible for this large difference. Our value is still about 0.4 eV smaller than the band gap of Shih *et al.*, though. In fact, a certain discrepancy should be expected because of the pseudopotential approximation and the plasmon-pole model for the dielectric function used in Ref. 1. The pseudopotential approximation not only neglects the important contribution of core-valence exchange. It also yields accurate wave functions only in the vicinity of the atomic electron energies, but not for high-lying states. This error is very similar in spirit to the linearization error of the LAPW basis and is also of the same magnitude.[17] Thus, it should have an impact on the GW results comparable in size to the linearization error, whose elimination gives rise to a sizable correction of 0.5 eV, as we have shown in the present work.

With the LDA band gap of only 0.84 eV the quasiparticle correction amounts to more than 2 eV. It can be expected that a treatment beyond the one-shot approach, for example, by taking into account off-diagonal elements of the self-energy, by updating the Green function, or by including vertex corrections, will further increase the value and, thus, bring it even closer to the experimental value. As was shown in Ref. 1, already using LDA + U instead of LDA as the mean-field starting point, which corrects the $2p$-$3d$ hybridization, gives an upward correction of 0.2 eV in the resulting GW band gap.

In conclusion, the band convergence is a serious issue in GW calculations and must be thoroughly dealt with. ZnO is an extreme case in this respect. We have reexamined the calculation of the ZnO band gap by Shih *et al.*,[1] and could confirm their main result: the GW band gap of ZnO shows a very slow convergence with respect to the number of states used to construct the polarization function and the correlation self-energy. Furthermore, when high-lying bands are included in the calculation, the linearization error of all-electron approaches becomes another important issue. As we have shown, it can be eliminated systematically within the FLAPW method by augmenting the basis in the MT spheres with LOs. In the case of ZnO this procedure yields a correction of about 0.5 eV and brings the calculated band gap (2.99 eV)

much closer to experiment than in previous studies. We believe that our study helps to clarify the contradiction between the pseudopotential results of Ref. [1] and the previous results based on all-electron approaches.[5–7]

[1] B.-C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **105**, 146401 (2010).

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

[3] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[4] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[5] M. Usuda, N. Hamada, T. Kotani, and M. van Schilfgaarde, Phys. Rev. B **66**, 125101 (2002).

[6] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).

[7] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).

[8] P. Gori, M. Rakel, C. Cobet, W. Richter, N. Esser, A. Hoffmann, R. Del Sole, A. Cricenti, and O. Pulci, Phys. Rev. B **81**, 125207 (2010).

[9] S. Tsoi, X. Lu, A. K. Ramdas, H. Alawadhi, M. Grimsditch, M. Cardona, and R. Lauck, Phys. Rev. B **74**, 165203 (2006); H. Alawadhi, S. Tsoi, X. Lu, A. K. Ramdas, M. Grimsditch, M. Cardona, and R. Lauck, *ibid.* **75**, 205207 (2007).

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

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

[12] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981); M. Weinert, E. Wimmer, and A. J. Freeman, *ibid.* **26**, 4571 (1982); H. J. F. Jansen and A. J. Freeman, *ibid.* **30**, 561 (1984).

[13] F. Aryasetiawan, in *Advances in Condensed Matter Science*, edited by V. I. Anisimov (Gordon and Breach, New York, 2000).

[14] C. Friedrich, S. Blügel, and A. Schindlmayr, Phys. Rev. B **81**, 125102 (2010); [http://www.flapw.de/spex].

[15] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).

[16] E. E. Krasovskii, A. N. Yaresko, and V. N. Antonov, J. Electron Spectrosc. Relat. Phenom. **68**, 157 (1994).

[17] C. Friedrich, A. Schindlmayr, S. Blügel, and T. Kotani, Phys. Rev. B **74**, 045104 (2006).

[18] D. Singh, Phys. Rev. B **43**, 6388 (1991).

[19] E. Sjöstedt, L. Nordström, and D. J. Singh, Solid State Commun. **114**, 15 (2000).

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