

## All-Electron Exact Exchange Treatment of Semiconductors: Effect of Core-Valence Interaction on Band-Gap and $d$ -Band Position

S. Sharma,\* J. K. Dewhurst, and C. Ambrosch-Draxl

*Institut für Physik, Karl-Franzens-Universität Graz, Universitätsplatz 5, A-8010 Graz, Austria*

(Received 17 January 2005; published 21 September 2005)

We present the first all-electron full-potential exact exchange (EXX) Kohn-Sham density functional calculations on a range of semiconductors and insulators (Ge, GaAs, CdS, Si, ZnS, C, BN, Ne, Ar, Kr, and Xe). We remove one of the main computational obstacles of such calculations by the use of a highly efficient basis for inversion of the response function. We find that the band gaps are not as close to experiment as those obtained from previous pseudopotential EXX calculations. The locations of  $d$  bands, determined using the full-potential EXX method, are in excellent agreement with experiment, irrespective of whether these are core, semicore, or valence states. We conclude that the inclusion of the core-valence interaction is necessary for accurate determination of EXX Kohn-Sham band structures and that EXX alone is not a complete answer to the band-gap problem in semiconductors.

DOI: [10.1103/PhysRevLett.95.136402](https://doi.org/10.1103/PhysRevLett.95.136402)

PACS numbers: 71.15.Mb, 71.10.-w, 71.22.+i

The exact treatment of exchange within the Kohn-Sham (KS) formulation of density functional theory (DFT) has been one of the most interesting developments of *ab initio* theory in recent years [1–7]. It has, however, led to a number of outstanding unresolved issues. One of the most striking examples concerns the band gaps and positions in semiconductors and insulators. As is well known, these properties constitute one of the classic failures of the local density approximation (LDA) to KS-DFT. The self-interaction of the LDA potential and the lack of a discontinuity in the exchange potential lead to both wrong band positions and too small (or absent) gaps. The fundamental gap is defined as the difference of ionization energy and electron affinity. This can be written in terms of ground-state energies of systems with different numbers of electrons, and so in principle is a ground-state property and may be calculated within KS-DFT.

It is expected that exact exchange (EXX) should improve the KS band gaps as it is identically self-interaction-free. Indeed, for the  $sp$  semiconductors the KS band gaps, and many other properties, were found to be in excellent agreement with experiment [3,8]. In KS-DFT the fundamental gap may be expressed as  $E_g = E_g^{\text{KS}} + \Delta_{\text{xc}}$ , where  $\Delta_{\text{xc}}$  is the discontinuity in the exchange-correlation potential. In fact, Städele *et al.* [3] also determined the discontinuities in the exchange potential for Si, Ge, and GaAs, and found them to be about 3–5 times the band gap, leading to speculations that there may be a vast cancellation between  $\Delta_c$  and  $\Delta_x$ . However, this was then followed by calculations of the KS gaps in wide band noble gas solids, where the agreement with experiment was found to be rather poor [5]. Thus it seems that EXX works extremely well for one class of materials but is much less effective for others.

A further problem is the  $d$ -band positions. Although the misplacement of the  $d$  states due to self-interaction is one of the major reasons for the LDA gap underestimation, it was found that the  $d$ -band positions in EXX are wrong

even when there is an improvement in the gap [9]. This contradicts with self-interaction corrected LDA calculations [10–12], where the  $d$ -band positions are improved by the removal of self-interaction in the LDA.

All these EXX calculations have been performed within the pseudopotential approach. There do exist all-electron EXX studies on the band gaps in C, Si, and Ge [4] which, interestingly, give worse results than the corresponding pseudopotentials (PP) ones. However, this work has been done within the atomic sphere approximation (ASA); hence it is not clear what effect this simplification of the potential has.

Thus the natural question that arises is how the basis for the KS wave function affects the performance of EXX. It is interesting to note that an analogous discussion is going on regarding the  $GW$  method: it was found that the quasiparticle spectra gave better gaps in the PP method if one performed one iteration only, even though full self-consistency is essential for charge conservation [13–15]. On the other hand, the  $GW$  method within an all-electron full-potential approach required full self-consistency to produce gaps in agreement with experiment. The success of the “one shot”  $GW$  within the PP approach was shown to be due to a cancellation of errors [16]: These errors arise, on the one hand, from the approximate treatment of the core-valence interaction and, on the other hand, from a non-self-consistent treatment.

It is thus urgently needed to perform all-electron full-potential EXX calculations to clarify the physics of EXX. Unfortunately, the EXX has not, until now, been implemented in an all-electron full-potential treatment due to the formidable numerical difficulties involved. These are associated with the need to invert the response, which for the complex basis needed in an all-electron full-potential method, is a large matrix. In this Letter, we present a formulation of the EXX method that solves these numerical problems and, by deployment of this method on a wide

range of semiconductors and insulators, resolve the outstanding issues.

We have implemented the EXX potential using the all-electron full-potential linearized augmented-plane wave (FP-LAPW) method [17] within the EXCITING code [18]. The starting point for these calculations is the exchange energy

$$E_x[n] = -\frac{1}{2} \sum_{i\mathbf{k}, j\mathbf{k}'}^{\text{occ}} w_{\mathbf{k}} w_{\mathbf{k}'} \int d\mathbf{r} d\mathbf{r}' \frac{\phi_{i\mathbf{k}}^*(\mathbf{r}) \phi_{j\mathbf{k}'}^*(\mathbf{r}') \phi_{j\mathbf{k}'}(\mathbf{r}) \phi_{i\mathbf{k}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|},$$

where  $\phi_{i\mathbf{k}}$  is the  $i$ th Kohn-Sham orbital of  $k$ -point  $\mathbf{k}$ ,  $w_{\mathbf{k}}$  is the  $k$ -point weight,  $n(\mathbf{r}) \equiv \sum_{i\mathbf{k}}^{\text{occ}} w_{\mathbf{k}} \phi_{i\mathbf{k}}^*(\mathbf{r}) \phi_{i\mathbf{k}}(\mathbf{r})$  is the density, and  $i$  and  $j$  run over both core and valence states, in contrast to PPs where the indices run over only the valence states. Proceeding in the same manner as Görling *et al.* [1,2,7], the functional derivative chain rule is employed to obtain the exchange potential

$$\begin{aligned} v_x[n](\mathbf{r}) &\equiv \frac{\delta E_x[n]}{\delta n(\mathbf{r})} = \sum_{i\mathbf{k}}^{\text{occ}} w_{\mathbf{k}} \int d\mathbf{r}' d\mathbf{r}'' \left[ \frac{\delta E_x}{\delta \phi_{i\mathbf{k}}(\mathbf{r}'')} \frac{\delta \phi_{i\mathbf{k}}(\mathbf{r}'')}{\delta v_s(\mathbf{r}')} + \frac{\delta E_x}{\delta \phi_{i\mathbf{k}}^*(\mathbf{r}'')} \frac{\delta \phi_{i\mathbf{k}}^*(\mathbf{r}'')}{\delta v_s(\mathbf{r}')} \right] \frac{\delta v_s(\mathbf{r}')}{\delta n(\mathbf{r})} \\ &= \int d\mathbf{r}' \left[ \sum_{i\mathbf{k}}^{\text{occ}} \sum_j^{\text{unocc}} w_{\mathbf{k}} \langle \phi_{i\mathbf{k}} | \hat{v}_x^{\text{NL}} | \phi_{j\mathbf{k}} \rangle \frac{\phi_{j\mathbf{k}}^*(\mathbf{r}') \phi_{i\mathbf{k}}(\mathbf{r}')}{\varepsilon_{i\mathbf{k}} - \varepsilon_{j\mathbf{k}}} + \text{c.c.} \right] \frac{\delta v_s(\mathbf{r}')}{\delta n(\mathbf{r})}, \end{aligned} \quad (1)$$

where  $\varepsilon_{i\mathbf{k}}$  are the Kohn-Sham eigenvalues and

$$\langle \phi_{i\mathbf{k}} | \hat{v}_x^{\text{NL}} | \phi_{j\mathbf{k}} \rangle = \sum_{i\mathbf{k}'}^{\text{occ}} w_{\mathbf{k}'} \int d\mathbf{r} d\mathbf{r}' \frac{\phi_{i\mathbf{k}}^*(\mathbf{r}) \phi_{i\mathbf{k}'}^*(\mathbf{r}') \phi_{i\mathbf{k}'}(\mathbf{r}) \phi_{j\mathbf{k}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2)$$

In order to obtain the functional derivative  $\delta v_s(\mathbf{r}')/\delta n(\mathbf{r})$ , the linear-response operator  $\chi$ , defined by the relation

$$\delta n(\mathbf{r}) = \int d\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}') \delta v_s(\mathbf{r}'),$$

must be inverted. This is not possible directly as  $\chi$  has a zero eigenvalue corresponding to a constant eigenfunction. However, in a basis that excludes such functions, the inversion of  $\chi$  can be performed. Application of elementary perturbation theory to  $n$  leads to an explicit expression for the response

$$\chi(\mathbf{r}, \mathbf{r}') = \sum_{i\mathbf{k}}^{\text{occ}} \sum_j^{\text{unocc}} w_{\mathbf{k}} \frac{\phi_{i\mathbf{k}}^*(\mathbf{r}) \phi_{j\mathbf{k}}(\mathbf{r}) \phi_{j\mathbf{k}}^*(\mathbf{r}') \phi_{i\mathbf{k}}(\mathbf{r}')}{\varepsilon_{i\mathbf{k}} - \varepsilon_{j\mathbf{k}}} + \text{c.c.}$$

Implementations of EXX for pseudopotentials have used the plane wave basis for expanding  $\chi$ . This is not feasible in a FP-LAPW approach because of the strongly varying wave functions inside the muffin tin. Instead, we propose a new basis described below. The overlap densities  $\rho_{\alpha}(\mathbf{r}) \equiv \phi_{i\mathbf{k}}^*(\mathbf{r}) \phi_{j\mathbf{k}}(\mathbf{r})$  and their complex conjugates are used as a spanning set, where  $\alpha \equiv (i\mathbf{k}, j\mathbf{k})$  with  $i$  and  $j$  labeling occupied and unoccupied states, respectively. These are the states of some representative  $k$  point that is chosen in advance. This spanning set is then reduced to a basis set as follows. First, the overlap matrix of the spanning set elements

$$O_{\alpha\beta} \equiv \int d\mathbf{r} \rho_{\alpha}^*(\mathbf{r}) \rho_{\beta}(\mathbf{r})$$

is obtained and diagonalized. Note that  $O$ , which is positive semidefinite, has non-negative eigenvalues. Next, all the eigenvectors of  $O$  corresponding to eigenvalues smaller than a certain tolerance,  $\epsilon$ , are discarded. Finally, denoting the remaining eigenvectors as  $v_{\gamma}^{\beta}$ , where  $\beta$  labels the

vector and  $\gamma$  its coordinate, we seek a transformation matrix  $C$  such that if

$$\tilde{\rho}_{\alpha}(\mathbf{r}) = \sum_{\beta} C_{\beta}^{\alpha} \sum_{\gamma} v_{\gamma}^{\beta} \rho_{\gamma}(\mathbf{r})$$

then

$$\int d\mathbf{r} \tilde{\rho}_{\alpha}^*(\mathbf{r}) \tilde{\rho}_{\beta}(\mathbf{r}) = \delta_{\alpha\beta}.$$

Straightforward algebra shows that  $C$  should satisfy

$$CC^{\dagger} = (v^{\dagger} O v)^{-1}, \quad (3)$$

where  $v$  is the matrix of eigenvectors  $v_{\gamma}^{\beta}$  in columnwise form. The matrix  $C$  is obtained by performing a Cholesky decomposition on the right-hand side of Eq. (3). By virtue of their construction and the fact that  $\int d\mathbf{r} \rho_{\alpha}(\mathbf{r}) = 0$ , the set of functions  $\{\tilde{\rho}_{\alpha}\}$  form an optimal basis for the expansion and inversion of  $\chi$ , as well as for the term in square brackets in Eq. (1). We should also point out that this basis may be useful for FP-LAPW time-dependent DFT response and  $GW$  methods.

Special attention should now be drawn to the calculation of the nonlocal matrix elements [Eq. (2)], which may be determined by a well-established method for solving Poisson's equation in a FP-LAPW environment [19]. The differences here are that the densities are now complex and there is a long-range term arising when  $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}'$  is close to zero. This is treated by considering the so-called pseudocharge density [17], which is chosen to be sufficiently smooth within the muffin tins so that it may be expanded in terms of plane waves, and yet has the same multipole expansion as the real density. If we restrict  $\mathbf{k}$  and  $\mathbf{k}'$  such that  $\mathbf{q}$  is in the first Brillouin zone (BZ), then the long-range (LR) contribution to the matrix elements from

the pseudocharge is

$$\langle \phi_{ik} | \hat{v}_x^{\text{NL}} | \phi_{jk} \rangle_{\text{LR}} = \sum_{\mathbf{q}}^{\text{occ}} w_{\mathbf{q}} \frac{4\pi\Omega}{q^2} \rho_{i\ell}^*(\mathbf{q}) \rho_{l_j}(\mathbf{q}),$$

where  $\rho_{i\ell}(\mathbf{q})$  and  $\rho_{l_j}(\mathbf{q})$  are the pseudocharge densities in reciprocal space. This sum suffers from poor convergence with respect to the number of  $q$  points [20]. We therefore approximate it by an integral over a sphere of volume equivalent to that of the BZ. The final expression is

$$\langle \phi_{ik} | \hat{v}_x^{\text{NL}} | \phi_{jk} \rangle_{\text{LR}} \approx 2 \left( \frac{6\Omega^5}{\pi} \right)^{1/3} \sum_{\mathbf{q}}^{\text{occ}} w_{\mathbf{q}} \rho_{i\ell}^*(\mathbf{q}) \rho_{l_j}(\mathbf{q}).$$

All the calculations in the present work are performed on a mesh of 27 special  $k$  points in the irreducible BZ. As the response and its inverse are very sensitive to the number of empty states, we have taken special care to use a sufficient number to converge the band gaps to within 0.01 eV. We find that almost all the materials studied here have converged  $d$ -band positions and band gaps with around 25 empty states. All the results presented are generated using 30 empty states. One final practical note is that the eigenvalue cutoff,  $\epsilon$ , discussed above, is taken to be  $10^{-5}$  times the largest eigenvalue of  $O$ . For example, for Si this results in 73 elements in the full basis set. One should note that expressed in terms of the LAPW basis the corresponding number of elements in the basis would be 700.

We now turn to the question of the band gap in EXX. In Fig. 1 we present the difference between the experimental fundamental band gap and the KS band gap for various semiconductors and insulators. (The absolute experimental band gaps can be found in tables in Refs. [5,8].) One can

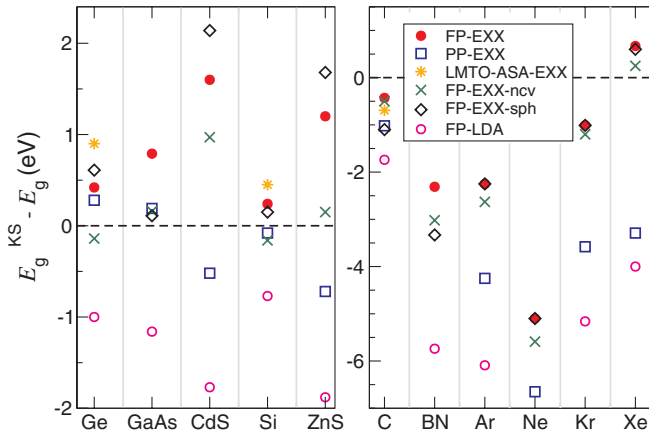


FIG. 1 (color online). Difference between theoretical KS gaps and experimental fundamental gaps in eV. Full-potential exact exchange (FP-EXX), FP-EXX without core-valence interaction (FP-EXX-ncv), FP-EXX with only spherical exchange (FP-EXX-sph), and LDA (FP-LDA) results are from this work. The PP-EXX results are from Refs. [5,8,9]. The LMTO-ASA-EXX results are from Ref. [4] and the experimental data are taken from Refs. [22–24]. Note that the LDA gap difference for Ne is  $-10.25$  eV and is off the scale of the plot.

immediately see that, in contrast to the PP approach, the agreement of KS gaps for the  $sp$  semiconductors is not good. In fact, for semiconductors the FP-EXX band gaps are overestimated with respect to experiments by up to 60%, whereas for all insulators except Xe we find an underestimate of up to 24%. We further find that for Xe the KS band gap is overestimated with respect to the experimental fundamental and optical gap by 3.8% and 24.6%, respectively, while for other insulators it is underestimated up to 6%. Thus the contrasting behavior of the  $sp$  semiconductors and insulators is not found in our all-electron full-potential calculations. Finally, the worst agreement, for both LDA and EXX, is seen in the  $sp$  semiconductors ZnS and CdS, and one notes that both these materials have very shallow  $d$  band positions.

In order to determine the reason for this behavior we performed ground-state calculations for these materials by ignoring the core-valence interaction term in the EXX potential. We emphasize here that this term was removed only from the exchange potential which, however, has nevertheless a substantial effect on the KS band gaps. The results are marked as FP-EXX-ncv in Fig. 1. One notes that the deviation from the experimental gap is now much less for all the semiconductors, while, on the other hand, it increases for the insulators.

This trend is concomitant with the findings of the PP-EXX calculations [5,8] where the performance of the EXX potential in determining the KS gaps is excellent for semiconductors but not so for insulators. Previous all-electron EXX calculations performed within the ASA show a similar overestimation of the gaps for semiconductors [4]. All this suggests that the lack of core-valence interaction improves the agreement of the KS band gaps with experiments for  $sp$  semiconductors, whereas it worsens in the noble gas solids and insulators in general.

In our work on magnetic metals [21] we found that inclusion of the nonspherical contributions to the exchange potential were crucial for obtaining the correct ground state. It is thus interesting to determine the effect here of retaining only the spherical part of the EXX potential. These are marked as FP-EXX-sph in Fig. 1. Clearly the effect of the shape approximation to the potential in all the materials other than noble gas solids is substantial, and the latter are, of course, expected to be well treated by the spherical approximation being composed of almost independent atoms.

Another property that is known to lead to errors in the band gaps [10] and for which the LDA performs badly is the position of the semicore  $d$  states. The LDA underbinds, leading to displacement of  $d$  bands much above the experimental value. In the past, quasiparticle band structures have been determined using the self-energy calculated with the  $GW$  approximation and shown to correct the  $d$ -band position [11]. Also the SIC to the potential was found to lead to good agreement between the theoretical and experimental position of these  $d$  bands. Since the EXX is self-interaction-free in nature, it should improve the

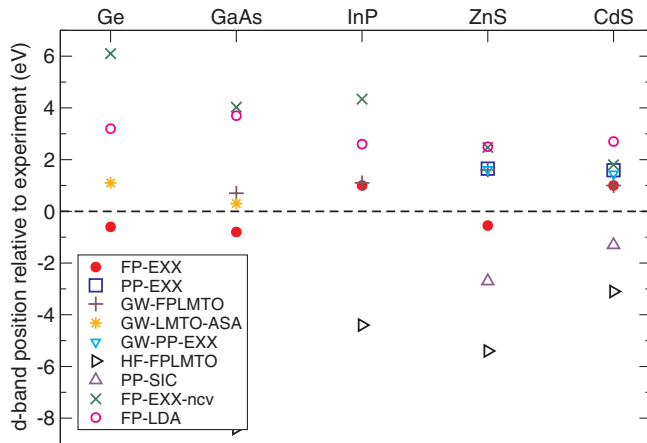


FIG. 2 (color online). Difference between theoretical  $d$ -band eigenvalues and experimental data in eV. Full-potential exact exchange (FP-EXX), FP-EXX without core-valence interaction (FP-EXX-ncv), and LDA (FP-LDA) results are from this work. The results obtained using the  $GW$  calculation on top of the PP-EXX ground state ( $GW$ -PP-EXX) are from Ref. [9], and the self-interaction corrected potential within the PP approach (PP-EXX) are from Ref. [12].  $GW$  results calculated using a FP-LMTO method ( $GW$ -FP-LMTO) and Hartree-Fock FP-LMTO (HF-FP-LMTO) are from Ref. [10]. The  $GW$  calculation on top of the LMTO-ASA-LDA ground state ( $GW$ -LMTO-ASA) are from Ref. [11].

$d$ -band positions. Counterintuitively Rinke *et al.* [9] found that even though the PP-EXX calculations result in a downward motion of the  $d$  bands with respect to LDA, still  $GW$  calculations are needed to fully correct the positions. It would be enlightening to see how the FP-EXX potential, in which the core and valence states are treated on the same footing, and which is truly self-interaction-free, effects these semicorelike or corelike  $d$  bands.

We compare FP-EXX  $d$ -band eigenvalues with those of previous PP-EXX, experiment, and various  $GW$  and SIC calculations in Fig. 2. The compounds are chosen so that  $d$  bands are in various energy regimes to get an overall picture: corelike (Ge: experimentally  $-29.6$  eV from the top of the valence band), semicore (GaAs:  $-18.8$  eV, InP:  $-16.8$  eV), and valence (ZnS:  $-8.7$  eV, CdS:  $-9.2$  eV). As can be seen, the eigenvalues of  $d$  states obtained using the FP-EXX method are in excellent agreement with experiments. In fact, in several cases it is better than that of the FP- $GW$  results. Here again the agreement between the FP-EXX-ncv and PP-EXX values of Rinke *et al.* [9] is quite good. This indicates that the inclusion of core-valence interaction is crucial for determination of the correct  $d$ -band eigenvalues.

To summarize, we find that the lack of core-valence interactions leads to an anomalously good agreement of the KS gap with the experimental fundamental gap in  $sp$  semiconductors. Inclusion of these interactions worsens the band gaps compared to the experiments but leads to a

consistent treatment of semiconductors and insulators. We conclude that EXX alone certainly does not solve the band-gap problem. However, we find  $d$ -band eigenvalues for a wide variety of cases to be in agreement with experiments, which is as good or better than that achieved by full-potential  $GW$  calculations.

We acknowledge the Austrian Science Fund (Project No. P16227) and the EXCITING network funded by the EU (Contract No. HPRN-CT-2002-00317) for financial support.

\*Electronic address: sangeeta.sharma@uni-graz.at

- [1] A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994).
- [2] A. Görling, Phys. Rev. B **53**, 7024 (1996).
- [3] M. Städele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. **79**, 2089 (1997).
- [4] T. Kotani, Phys. Rev. Lett. **74**, 2989 (1995).
- [5] R. J. Magyar, A. Fleszar, and E. K. U. Gross, Phys. Rev. B **69**, 045111 (2004).
- [6] Y. M. Niquet and X. Gonze, Phys. Rev. B **70**, 245115 (2004).
- [7] A. Görling, Phys. Rev. B **59**, 10370 (1999).
- [8] M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).
- [9] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).
- [10] T. Kotani and M. V. Schilfgarde, Solid State Commun. **121**, 461 (2002).
- [11] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. B **54**, 17564 (1996).
- [12] D. Vogel, P. Krüger, and J. Pollmann, Phys. Rev. B **54**, 5495 (1996).
- [13] B. Holm and U. von Barth, Phys. Rev. B **57**, 2108 (1998).
- [14] W.-D. Schöne and A. G. Eguiluz, Phys. Rev. Lett. **81**, 1662 (1998).
- [15] B. Holm, Phys. Rev. Lett. **83**, 788 (1999).
- [16] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).
- [17] D. J. Singh, *Planewaves Pseudopotentials and the LAPW Method* (Kluwer Academic Publishers, Boston, 1994).
- [18] J. K. Dewhurst, S. Sharma, and C. Ambrosch-Draxl, Graz, 2004, <http://exciting.sourceforge.net/>.
- [19] M. Weinert, J. Math. Phys. (N.Y.) **22**, 2433 (1981).
- [20] S. Massidda, M. Posternak, and A. Baldereschi, Phys. Rev. B **48**, 5058 (1993).
- [21] S. Sharma, J. K. Dewhurst, C. Ambrosch-Draxl, and C. Persson, <http://arxiv.org/abs/cond-mat/0501258>.
- [22] *Semiconductors, Physics of Group IV Elements and III-IV Compounds*, edited by K.-H. Hellwege, O. Madelung, M. Schulz, and H. Weiss (Springer Berlin, New York, 1982), Vol. 17.
- [23] A. Mainwood, in *Properties and Growth of Diamond* (INSPEC, London, 1993).
- [24] D. R. T. Zahn, G. Kudlek, U. Rossow, A. Hoffman, I. Broser, and W. Richter, Adv. Mater. Opt. Electron. **3**, 11 (1994).