

# Exciting prospects for solids: Exact-exchange based functionals meet quasiparticle energy calculations

Editor's Choice

Patrick Rinke<sup>\*,1,2</sup>, Abdallah Qteish<sup>1,3</sup>, Jörg Neugebauer<sup>4</sup>, and Matthias Scheffler<sup>1,2,5,6</sup>

<sup>1</sup> Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, 14195 Berlin, Germany

<sup>2</sup> European Theoretical Spectroscopy Facility (ETSF), Spain

<sup>3</sup> Department of Physics, Yarmouk University, 21163 Irbid, Jordan

<sup>4</sup> Max-Planck-Institut für Eisenforschung, Max-Planck-Str. 1, 40237 Düsseldorf, Germany

<sup>5</sup> Department of Chemistry & Biochemistry, University of California Santa Barbara, Santa Barbara, CA, USA

<sup>6</sup> Materials Department, University of California Santa Barbara, Santa Barbara, CA, USA

Received 10 September 2007, revised 21 November 2007, accepted 26 November 2007

Published online 1 February 2008

PACS 71.15.Mb, 71.15.Qe, 71.45.Gm, 71.55.Eq, 71.55.Gs

\* Corresponding author: e-mail rinke@fhi-berlin.mpg.de

Focussing on spectroscopic aspects of semiconductors and insulators we will illustrate how quasiparticle energy calculations in the  $G_0W_0$  approximation can be successfully combined with density-functional theory calculations in the exact-

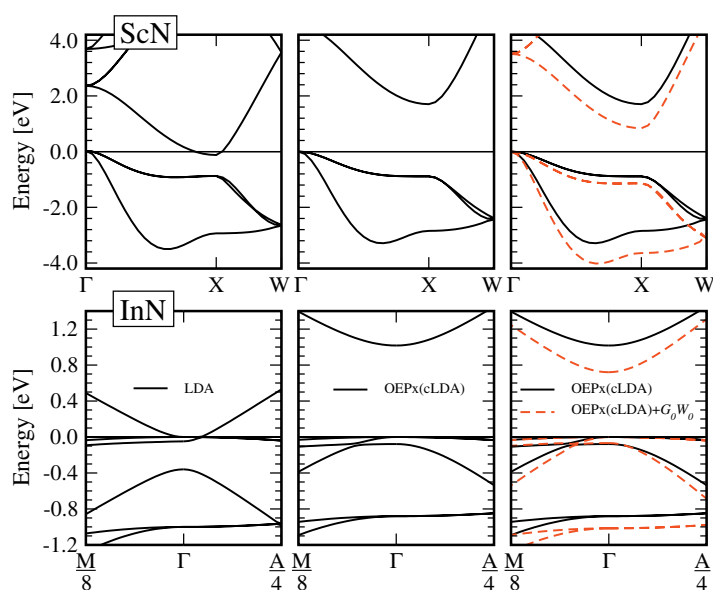
exchange optimised effective potential approach (OEPx) to achieve a first principles description of the electronic structure that overcomes the limitations of local- or gradient-corrected DFT functionals (LDA and GGA).

© 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

**1 Introduction** Density functional theory (DFT) has contributed significantly to our present understanding of a wide range of materials and their properties. As quantum-mechanical theory of the density and the total energy, it provides an atomistic description from first principles and is, in the local-density or generalized gradient approximation (LDA and GGA), applicable to polyatomic systems containing up to several thousand atoms. However, a combination of three factors limits the applicability of LDA and GGA to a range of important materials and interesting phenomena. They are approximate (jellium-based) exchange-correlation functionals, which suffer from incomplete cancellation of artificial self-interaction and lack the discontinuity of the exchange-correlation potential with respect to the number of electrons. As a consequence the Kohn–Sham (KS) single-particle eigenvalue band gap for semiconductors and insulators underestimates the quasiparticle band gap as measured by the difference of ionisation energy (via photoemission spectroscopy (PES)) and electron affinity (via inverse PES (IPES)). This reduces the predictive power for materials whose band gap is not known from experiment and poses a problem for calculations

where the value of the band gap is of importance, e.g. the energies of surface and defect states. Moreover, the incomplete cancellation of the self-interaction questions the applicability to systems with localized defect states, strong magnetic interactions or materials with localized d or f states, such as II–VI compounds, group-III-nitrides, transition metal oxides or lanthanides, actinides and their compounds. It is therefore desirable to have a more advanced electronic structure approach that overcomes these shortcomings of LDA/GGA and provides a reliable description of ground state total energies as well as electronic excitation spectra.

Many-body perturbation theory (MBPT) in the  $GW$  approach [1] has developed into the method of choice for describing the energies of electronic excitation spectra of solids (in the following referred to as quasiparticle energies) [2, 3]. For conceptual as well as computational reasons, however, it is currently not feasible to compute total energies for systems of interest. Exact-exchange based DFT functionals on the other hand give access to total energies and thus atomic structures. In addition they largely or in certain cases completely remove the self-interaction error.



**Figure 1** (online colour at: [www.pss-b.com](http://www.pss-b.com)) LDA KS calculations incorrectly predict wurtzite InN (bottom) to be a zero-gap semiconductor with the wrong band ordering at the  $\Gamma$ -point. ScN (top) is erroneously described as semimetal. In OEPx(cLDA) the band gap opens and InN and ScN correctly become semiconductors, thus providing a more suitable starting point for subsequent quasiparticle energy calculations in the  $G_0W_0$  approximation. All calculations are performed at the experimental lattice constants.

Exact-exchange in the optimized effective potential (OEP) approach (OEPx or EXX or OEPx(cLDA) when LDA correlation is added (see also Section 3)) has become the most prominent approach of this kind. It is self-interaction free and greatly improves the comparison of KS eigenvalue differences with quasiparticle excitations for solids, with respect to Hartree–Fock [4–16].

Alternatively, the removal of the self-interaction can be approached by applying the exchange-operator in a non-local fashion within the framework of the generalized Kohn–Sham (GKS) scheme [17] or by using other non-

local schemes such as the self-interaction corrected LDA (SIC-LDA) method [18]. In the former case the bare exchange interaction is empirically screened and combined with a local potential (commonly referred to as hybrid functional). The most widely used hybrid functionals are PBE0 [19–21], B3LYP [22–24], HSE [25–27] and screened exchange (sX-LDA) [17, 28]. With the exception of SIC and sX-LDA hybrid functionals are only slowly being applied to solids. Almost no experience exists in their combination with  $GW$  calculations [29–31] and we will thus defer a discussion of their properties to the Outlook (Section 7).

Combining exact-exchange based DFT calculations with quasiparticle energy calculations [10, 11, 32–34] offers several advantages and will be an important step towards reaching a thorough understanding of the aforementioned systems. Up to now most  $GW$  calculations are performed non-self-consistently as a single perturbation to an LDA ground state. Unlike the Kohn–Sham scheme in DFT (see Section 3.1) the  $GW$  approach in MBPT is not governed by a closed set of equations. Iterating beyond the first order ( $G_0W_0$ ) in the exact theory would introduce higher order interactions at every step that leads beyond the  $GW$  approximation, while iterating the set of  $GW$  equations ignoring these contributions generally worsens the good agreement achieved after the first iteration (a more detailed account of self-consistency in  $GW$  will be given in Section 2.2). From the view point of perturbation theory it would thus be advantageous to use OEPx instead of the LDA<sup>1</sup> as a starting point for  $G_0W_0$  calculations because the corresponding Kohn–Sham spectrum is closer to the quasiparticle spectrum, thus requiring a smaller correction. This becomes particularly important for materials where LDA severely underestimates the band gap like in



Patrick Rinke graduated from the University of York in England in 2003 with a Ph.D. in theoretical physics. Subsequently, he accepted a researcher position at the Fritz-Haber-Institut (FHI) of the Max-Planck-Society in Berlin, where he worked on the electronic structure and excited states of semiconductors, insulators and their surfaces, combining exact-exchange density-functional theory with quasiparticle energy calculations. While at the FHI he became a node leader of the European Theoretical Spectroscopy Facility (ETSF), an initiative of the Nanoquanta Network of Excellence. He is now continuing his research at the University of California at Santa Barbara.

<sup>1</sup> All  $G_0W_0$  calculations in this work are performed at the experimental lattice constants.

ZnO (by  $\approx 80\%$  for the wurtzite phase, LDA (at exp. lattice const.): 0.7 eV, exp: 3.4 eV) or incorrectly predicts a (semi)metallic state as e.g. in germanium (Ge), indium nitride (InN) and scandium nitride (ScN) (see Figs. 1 and 3).

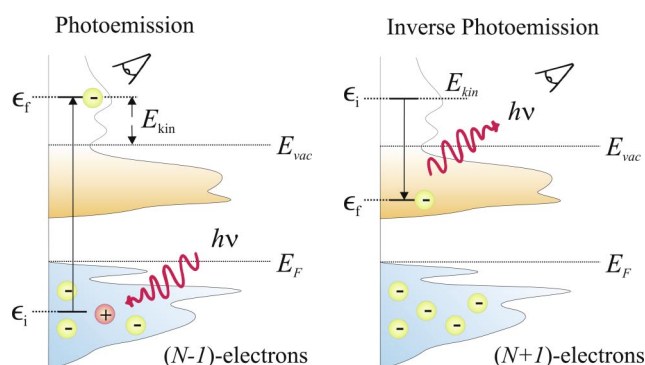
In this article we will focus only on spectroscopic aspects of semiconductors and insulators and illustrate how quasiparticle energy calculations can be successfully combined with OEPx(cLDA) calculations to achieve a first principles description of the electronic structure that overcomes the limitations of LDA and GGA. Figure 1 summarises the structure and the message of this paper. For the example of InN and ScN we demonstrate how OEPx(cLDA) opens the band gap and correctly predicts them to be semiconductors thus providing a better starting point for subsequent  $G_0W_0$  calculations. The quasiparticle corrections to the OEPx(cLDA) band structure are considerable – also for Gallium nitride (GaN) and selected II–VI compounds – bringing the quasiparticle band structure into good agreement with direct and inverse photoemission data. Our OEPx based  $G_0W_0$  calculations have helped to identify the band gap of InN and ScN and to clarify the source for the puzzling, wide interval of experimentally observed band gaps in InN [33, 34].

The article is structured as follows. In Section 2.2 we will briefly review the connection between the single particle Green's function and photo-electron spectroscopy and then introduce the  $GW$  method. The exact-exchange approach will be described in Section 3. In Section 4 we will demonstrate why and how the combination of OEPx(cLDA) and  $GW$  provides an improved description of the electronic structure. Finally we will conclude in Section 6 and give an outlook in Section 7.

## 2 Quasiparticle energy calculations

**2.1 Photo-electron spectroscopy and the quasiparticle concept** In photo-electron spectroscopy (PES) [35–37] electrons are ejected from a sample upon irradiation with visible or ultraviolet light (UPS) or with X-rays (XPS), as sketched in the left panel of Fig. 2. Energies  $\varepsilon_i = h\nu - E_{\text{kin}}$  and lifetimes of holes can be reconstructed from the photon energy  $h\nu$  and the kinetic energy  $E_{\text{kin}}$  of the photo-electrons that reach the detector<sup>2</sup>. By inverting the photoemission process, as schematically shown in the right panel of Fig. 2, the unoccupied states can be probed. An incident electron with energy  $E_{\text{kin}}$  is scattered in the sample emitting *bremstrahlung*. Eventually it will undergo a radiative transition into a lower-lying unoccupied state, emitting a photon that carries the transition energy  $h\nu$ . The energy of the final, unoccupied state can be deduced from the measured photon energy according to  $\varepsilon_f = E_{\text{kin}} - h\nu$ . This technique is commonly referred to as inverse photoemission spectroscopy (IPES) or *bremstrahlung* isochromat spectroscopy (BIS) [38–40].

<sup>2</sup> Throughout this article the top of the valence bands is chosen as energy zero.



**Figure 2** (online colour at: www.pss-b.com) Schematic of the photoemission (PES) and inverse photoemission (IPES) process. In PES (left) an electron is excited by an incoming photon from a previously occupied valence state (lower shaded region) into the continuum (white region, starting above the vacuum level  $E_{\text{vac}}$ ). In IPES (right) an injected electron with kinetic energy  $\varepsilon_i = E_{\text{kin}}$  undergoes a radiative transition into an unoccupied state (upper shaded region) thus emitting a photon in the process.

The experimental observable in photoemission spectroscopy is the photocurrent. Since the energy dependence of the transition matrix elements is usually weak and smooth, structures in the photoemission spectrum can be associated with features in the density of states (DOS), i.e. the imaginary part of the one-particle Green's function<sup>3</sup> [3, 41]

$$A(\mathbf{r}, \mathbf{r}'; \varepsilon) = \frac{1}{\pi} \text{Im} G(\mathbf{r}, \mathbf{r}'; \varepsilon). \quad (1)$$

Peaks due to shake ups and shake downs found in XPS are not described by the  $GW$  approximation and will therefore not be addressed here.

The Green's function is the solution to the many-body Hamiltonian

$$\hat{H}(\mathbf{r}, \mathbf{r}'; \varepsilon) = \hat{h}_0(\mathbf{r}) + \Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon), \quad (2)$$

written here in single-particle form where all electron-electron interaction terms are rolled up in the non-local, energy dependent self-energy  $\Sigma$ . The remaining contributions are given by<sup>4</sup>  $\hat{h}_0(\mathbf{r}) = -1/2\nabla^2 + v_{\text{ext}}(\mathbf{r})$ . The external potential  $v_{\text{ext}}$  is due to the nuclei, after the Born–Oppenheimer approximation of stationary nuclei is taken.

<sup>3</sup> The  $\mathbf{r}$  and  $\mathbf{r}'$  dependence can easily be transformed into a  $\mathbf{k}$  dependence. Furthermore only the spin unpolarised situation is discussed here. For the present discussion a summation over the spin indices in the Green's function is therefore assumed.

<sup>4</sup> Atomic units  $4\pi\varepsilon_0 = \hbar = e = m_e = 1$ , where  $e$  and  $m_e$  are the charge and mass of an electron, respectively, will be used in the remainder of this article. If not otherwise stated energies are measured in hartrees and length in bohr.

The photocurrent is then the surface weighted integral over the diagonal part of the spectral function  $A(\mathbf{r}, \mathbf{r}'; \varepsilon)$ . We note, however, that with respect to the measured intensities a photoemission spectrum is a noticeably distorted spectral function that in addition is weighted over the momentum components normal to the surface ( $\mathbf{k}_\perp$ ). In particular when selection rules and the energy dependence of the transition matrix elements become important certain peaks in the spectral function may be significantly reduced or may even disappear completely.

The excitation of a non-interacting or a bare particle would give rise to a delta peak in the spectral function. When the electron–electron interaction is turned on, the electrons can no longer be regarded as independent particles. As a consequence the matrix elements of the spectral function  $A_{nk}(\varepsilon)$  will contain contributions from many non-vanishing transition amplitudes. If these contributions merge into a clearly identifiable peak that appears to be derived from a single delta-peak broadened by the electron–electron interaction

$$A_{nk}(\varepsilon) \approx \frac{Z_{nk}}{\varepsilon - (\varepsilon_{nk} + i\Gamma_{nk})}, \quad (3)$$

this structure can be interpreted as single-particle-like excitation – the *quasiparticle*<sup>5</sup>. The broadening of the quasiparticle peak in the spectral function is associated with the lifetime  $\tau_{nk} = 2/\Gamma_{nk}$  of the excitation due to electron–electron scattering, whereas the area underneath the peak is interpreted as the renormalisation  $Z_{nk}$  of the quasiparticle. This renormalisation factor quantifies the reduction in spectral weight due to electron–electron exchange and correlation effects compared to an independent electron.

A computational description of the quasiparticle band structure thus requires the calculation of the Green’s function and the self-energy.

**2.2 The GW formalism** An exact solution to Eq. (2) is given by Hedin’s set of coupled integro-differential equations [1]

$$P(1, 2) = -i \int G(1, 3) G(4, 1) \Gamma(3, 4, 2) d(3, 4), \quad (4)$$

$$W(1, 2) = v(1, 2) + \int v(1, 3) P(3, 4) W(4, 2) d(3, 4), \quad (5)$$

$$\Sigma(1, 2) = i \int G(1, 3) \Gamma(3, 2, 4) W(4, 1) d(3, 4), \quad (6)$$

$$\Gamma(1, 2, 3) = \delta(1, 2) \delta(1, 3) + \int \frac{\delta\Sigma(1, 2)}{\delta G(4, 5)} G(4, 6) G(7, 5) \Gamma(6, 7, 3) d(4, 5, 6, 7), \quad (7)$$

<sup>5</sup> A *quasielectron* is a new quantum mechanical entity that combines an electron with its surrounding polarisation cloud. The same concept applies to holes.

where the notation  $1 \equiv (\mathbf{r}_1, t_1, \sigma_1)$  is used to denote a triple of space, time and spin variables<sup>6</sup>. Accordingly  $\int d(1)$  is a shorthand notation for the integration in all three variables of the triple. In Eqs. (4)–(7)  $P$  is the polarisability,  $W$  the screened and  $v$  the bare Coulomb interaction and  $\Gamma$  the so-called vertex function. By means of Dyson’s equation

$$G^{-1}(1, 2) = G_0^{-1}(1, 2) - \Sigma(1, 2), \quad (8)$$

which links the non-interacting system with Green’s function  $G_0$  to the fully interacting one ( $G$ ) via the self-energy  $\Sigma$ , Hedin’s equations could in principle be solved self-consistently starting from a given  $G_0$ . However, the functional derivative in Eq. (7) does not permit a numeric solution, requiring approximations for the vertex function in practise.

Starting from a non-interacting system the initial self-energy is zero and the vertex function becomes  $\Gamma(1, 2, 3) = \delta(1, 2) \delta(1, 3)$ . The first iteration yields Hedin’s *GW* approximation [1] for the self-energy as a function of energy

$$\Sigma_{\text{GW}}(\mathbf{r}, \mathbf{r}'; \varepsilon) = \frac{i}{2\pi} \int_{-\infty}^{\infty} d\varepsilon' e^{i\varepsilon'\delta} G(\mathbf{r}, \mathbf{r}'; \varepsilon + \varepsilon') W(\mathbf{r}, \mathbf{r}'; \varepsilon'), \quad (9)$$

where  $\delta$  is an infinitesimal, positive time. The physical interpretation of the self-energy corresponds to the energy contribution of a particle that arises from the response of the system due to particle’s own presence. The interaction of the electrons in the system is mediated by the screened Coulomb interaction rather than the bare one. In the *GW* formalism  $W$  (Eq. (5)) is calculated in the Random-Phase approximation (i.e. quasiholes and quasielectrons do not interact with each other) from the irreducible polarizability<sup>7</sup> (Eq. (4)). For the zeroth order Green’s function  $G_0$  a Kohn–Sham Green’s function  $G_{\text{KS}}$  (built from Kohn–Sham energies ( $\varepsilon_{nk}$ ) and wavefunctions ( $\phi_{nk}(\mathbf{r})$ ))

$$G_0(\mathbf{r}, \mathbf{r}', \varepsilon) = G_{\text{KS}}(\mathbf{r}, \mathbf{r}', \varepsilon) = \lim_{\delta \rightarrow 0^+} \sum_{nk} \frac{\phi_{nk}(\mathbf{r}) \phi_{nk}^*(\mathbf{r}')}{\varepsilon - (\varepsilon_{nk} + i\delta \operatorname{sgn}(E_F - \varepsilon_{nk}))} \quad (10)$$

is typically taken where  $E_F$  defines the Fermi level.

Keeping  $\Gamma$  fixed (to delta functions) removes the functional derivative from Eqs. (4)–(7). These *GW* equations could then in principle be solved self-consistently via Dyson’s equation (8). However this issue is still a matter of debate [42–46]. Unlike in DFT, a self-consistent solution of the full set of Hedin’s equations would go beyond the

<sup>6</sup> Throughout this article we will only consider systems without any explicit spin dependence and will therefore omit the spin index in the following.

<sup>7</sup> With  $\Gamma(1, 2, 3) = \delta(1, 2) \delta(1, 3)$  Eq. (4) reduces to the much simpler form  $\chi_0 = -iGG$ .



*GW* approximation and successively introduce higher order electron–electron interactions through the vertex function with every iteration step. Solving the *GW* equations self-consistently is therefore inconsistent if no higher order electron–electron interactions are included. It was first observed for the homogeneous electron gas [47] that the spectral features broaden with increasing number of iterations in the self-consistency cycle. Similarly, for closed shell atoms the good agreement with experiment for the ionization energy after the first iteration is lost upon iterating the equations to self-consistency [44]. Imposing self-consistency in an approximate fashion [46, 48–50] is not unique and different methods yield different results. Since the controversies regarding self-consistency within *GW* have not been resolved conclusively yet, and research into vertex corrections is still in its infancy [51–53] the majority of all *GW* calculations is performed employing the zeroth order in the self-energy ( $G_0W_0$ ). This, however, introduces a dependence of the starting Green's function on the self-energy and thus the quasiparticle energies, which can have a noticeable influence, as we will demonstrate in this article.

To obtain the quasiparticle band structure we solve the quasiparticle equation<sup>8</sup>

$$\left[ -\frac{\nabla^2}{2} + v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) \right] \psi_{nq}(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon_{nq}^{qp}) \psi_{nq}(\mathbf{r}') = \varepsilon_{nq}^{qp} \psi_{nq}(\mathbf{r}) \quad (11)$$

with the *GW* self-energy ( $\Sigma = \Sigma_{GW}$ ) by approximating the quasiparticle wavefunctions ( $\psi_{nq}$ ) with the Kohn–Sham ones ( $\phi_{nq}$ ) [54]. The corrections to the Kohn–Sham eigenvalues ( $\varepsilon_{nq}$ ) are then given by

$$\varepsilon_{nq}^{qp} = \varepsilon_{nq} + \langle \phi_{nq} | \Sigma_{GW}(\varepsilon_{nq}^{qp}) - v_{xc} | \phi_{nq} \rangle, \quad (12)$$

where  $\langle \phi_{nq} | \phi_{nq} \rangle$  denotes matrix elements with respect to the wavefunctions  $\phi_{nq}$  of the preceding Kohn–Sham DFT calculation with exchange–correlation potential  $v_{xc}$ .

### 3 Exact-exchange based DFT

#### 3.1 DFT and the Kohn–Sham band structure

With regard to quasiparticle band structures Kohn–Sham DFT calculations not only serve as a starting point for  $G_0W_0$  calculations, they are frequently used to interpret quasiparticle spectra due to the similarity of the Kohn–Sham

$$\left[ -\frac{\nabla^2}{2} + v_{\text{eff}}[n(\mathbf{r})](\mathbf{r}) \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}), \quad (13)$$

with the quasiparticle equation (11) and because the correction can often be expressed in terms of first-order perturbation theory (Eq. (12)). To illustrate the difference between the Kohn–Sham and the quasiparticle picture and to lay the foundations for the discussion of the exact-exchange OEP approach we will briefly review the Kohn–Sham DFT scheme.

The central quantities in DFT are the electron density  $n(\mathbf{r})$  and the total energy  $E_{\text{tot}}$ . The latter is a functional of the former and attains its minimum at the exact ground state density, as proven by Hohenberg and Kohn [55]. This formalism was turned into a tractable computational scheme by Kohn and Sham [56], by assuming that the system of interacting particles can be mapped onto a fictitious system of non-interacting particles moving in an effective local potential  $v_{\text{eff}}(\mathbf{r})$  that reproduce the same density as the many-body problem of interest. The electron density

$$n(\mathbf{r}) = \sum_i^{\text{occ}} |\phi_i(\mathbf{r})|^2, \quad (14)$$

is composed of the occupied Kohn–Sham orbitals  $\phi_i(\mathbf{r})$  that are solutions of the Kohn–Sham equation (13).

In analogy to Hartree theory Kohn and Sham divided the total energy into known contributions such as the kinetic energy of the non-interacting particles  $T_s$ , the Hartree energy

$$E_H[n] = \frac{1}{2} \int d\mathbf{r} n(\mathbf{r}) v_H(\mathbf{r}) = \frac{1}{2} \sum_{ij}^{\text{occ}} \int d\mathbf{r} d\mathbf{r}' \frac{\phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (15)$$

the external energy

$$E_{\text{ext}}[n] = \int d\mathbf{r} n(\mathbf{r}) v_{\text{ext}}(\mathbf{r}), \quad (16)$$

and an unknown remainder. This last term includes all electron–electron interactions beyond the Hartree mean-field and is defined as the exchange–correlation energy

$$E_{xc}[n] = E_{\text{tot}}[n] - T_s[n] - E_{\text{ext}}[n] - E_H[n]. \quad (17)$$

Performing the variation with respect to the density then yields the effective potential

$$v_{\text{eff}}[n](\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_H[n](\mathbf{r}) + v_{xc}[n](\mathbf{r}), \quad (18)$$

where each term in the sum is obtained as functional derivative of the corresponding energy expression. Since the exact form of the exchange–correlation functional is un-

<sup>8</sup> Note that the Hartree potential  $v_H(\mathbf{r})$  defined in the next Section has been separated from the self-energy.

known<sup>9</sup> suitable approximations have to be found in practice.

**3.2 LDA and self-interaction** Probably the most popular and efficient approximation for the exchange-correlation energy to date is the local-density approximation (LDA) [56]

$$E_{xc}^{LDA}[n] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_{xc}^{HEG}(n)|_{n=n(\mathbf{r})} \quad (19)$$

where  $\varepsilon_{xc}^{HEG}(n(\mathbf{r}))$  is the exchange-correlation energy density of the homogeneous electron gas (HEG). Whilst the LDA<sup>10</sup> describes even inhomogeneous systems with startling success in many cases, it does so at the expense of a non-physical electron self-interaction. This is introduced by the Hartree term, which contains a spurious interaction of an electron with itself since the sum in Eq. (15) includes all occupied states. Although the LDA fulfills the sum rule and thus correctly removes the self-interaction with respect to the particle number, the shape of the exchange-correlation hole is not correct<sup>11</sup>. Applying the definition of Perdew and Zunger [18] the ensuing self-interaction error can be quantified for every state as

$$\delta_i = \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{|\phi_i(\mathbf{r})|^2 |\phi_i(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[|\phi_i(\mathbf{r})|^2], \quad (20)$$

which follows directly from Eq. (14) and (15) for  $i = j$ . The self-interaction error is largest for localised states and has a tendency to delocalise the electronic wavefunction [32, 59], a point that we will return to later.

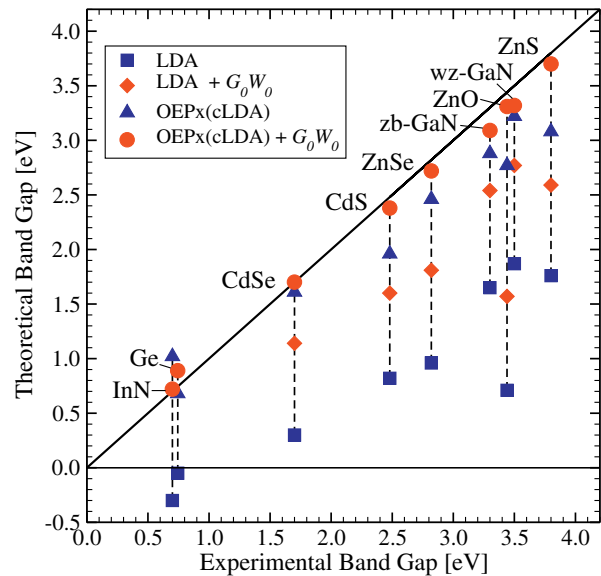
Perdew and Zunger proposed to subtract the sum over all self-interaction contributions  $\delta_i$  for all occupied states from the expression for the total energy in the LDA (or alternatively the spin-dependent LDA) [18]

$$E_{xc}^{SIC-LDA}[n] = E_{xc}^{LDA}[n] - \sum_i^{occ} \delta_i. \quad (21)$$

<sup>9</sup> To be more precise: according to the Hohenberg-Kohn theorem  $E_{xc}$  is a unique functional of the density. However, this does not necessarily imply that  $E_{xc}$  can be written in a closed mathematical form as functional of the density. Analogous examples are  $T_e$  (the kinetic energy of interacting electrons) and  $T_s$ . For the latter one may write down a series expansion, but this does not converge properly. Its evaluation therefore requires the detour via the Kohn–Sham formalism. In fact this exact evaluation of  $T_s$  is the strongest reason for using the Kohn–Sham scheme. As far as the exact  $v_{xc}$  is concerned it can be expressed in terms of the Green's function and the self-energy via the Sham–Schlüter equation in the context of many-body perturbation theory. Alternatively an exact representation of  $v_{xc}$  can be obtained in Görling–Levy perturbation theory [57].

<sup>10</sup> The statements made in this subsection apply also to the generalised gradient approximation.

<sup>11</sup> Only the spherical average of the exchange-correlation hole enters in the expression for the total energy [58], which, at least partly, explains the remarkable success of the LDA.



**Figure 3** (online colour at: [www.pss-b.com](http://www.pss-b.com)) Theoretical versus experimental band gaps: the OEPx(cLDA) based schemes systematically open the band gap compared to the LDA based calculations. Our OEPx(cLDA) +  $G_0W_0$  calculations with the cation d-electrons included as valence electrons agree very well with the experimental values. (For ZnO an estimate of 0.2 eV was added to the zinc-blende values in order to compare to the experimental results for wurtzite.)

The expression for the self-interaction corrected LDA (SIC-LDA) total energy can be minimised according to the variational principle. However, since the energy functional is now orbital dependent, the computational simplicity of the LDA is lost, making SIC-LDA calculations computationally much more demanding than LDA or GGA calculations. Moreover, this form of self-interaction correction vanishes for completely delocalised states, which makes a direct application of this formalism to Bloch states in solids difficult.

It is this self-interaction error in the LDA that is responsible for the fact that InN, ScN as well as Ge are incorrectly predicted to be (semi)metals (cf. Figs. 1 and 3) and that the band gap of GaN and II–VI compounds is underestimated severely as shown in Fig. 3. Removing the self-interaction error, as done for instance in the exact-exchange approach, alleviates this problem, despite the fact that the valence and conduction band edge are given by Bloch states, as we will demonstrate in the following.

**3.3 The OEP method and exact-exchange** Following Kohn and Sham's concept of dividing the total energy into known and unknown contributions the exact-exchange energy  $E_x$

$$E_x = -\frac{1}{2} \sum_{ij}^{occ} \iint d\mathbf{r} d\mathbf{r}' \frac{\phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (22)$$

can be isolated from  $E_{xc}$  leaving only the correlation part  $E_c$  to be approximated. In the exact-exchange only approach this correlation term is omitted<sup>12</sup> so that the total energy becomes

$$E_{\text{tot}}^{x\text{-only}} = T_s + E_{\text{ext}} + E_H + E_x. \quad (23)$$

For occupied states the exact-exchange term cancels exactly with the corresponding term in the Hartree potential for  $i = j$  and thus removes the self-interaction in the Perdew–Zunger sense ( $\delta_i = 0$ ).

Applying the variational principle to Eqs. (22) and (23) with respect to the orbitals  $\phi_i(\mathbf{r})$  yields the Hartree–Fock approach. Like in the  $GW$  approach the eigenenergies and wavefunctions are solutions to a quasiparticle equation – albeit with a hermitian operator

$$\left[ -\frac{\nabla^2}{2} + v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) \right] \phi_i(\mathbf{r}) + \int d\mathbf{r}' \Sigma_x(\mathbf{r}, \mathbf{r}') \phi_i(\mathbf{r}') = \varepsilon_i \phi_i(\mathbf{r}). \quad (24)$$

$\Sigma_x$  is the non-local Fock-operator or exchange self-energy

$$\Sigma_x(\mathbf{r}, \mathbf{r}') = - \sum_j^{\text{occ}} \frac{\phi_j(\mathbf{r}) \phi_j^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (25)$$

To stay within the framework of density-functional theory the variation of Eq. (22) with respect to the density has to be performed instead. This can be done analytically for the exact-exchange energy expression Eq. (22) in the optimised effective potential approach [60–63] and yields the local exchange potential

$$v_x^{\text{OEP}}(\mathbf{r}) = \int d\mathbf{r}' \sum_i^{\text{occ}} \sum_j^{\text{unocc}} \left[ \langle \phi_i | \Sigma_x | \phi_j \rangle \frac{\phi_j(\mathbf{r}') \phi_i(\mathbf{r}')}{\varepsilon_i - \varepsilon_j} + \text{c.c.} \right] \times \chi_0^{-1}(\mathbf{r}', \mathbf{r}). \quad (26)$$

$\langle \phi_i | \Sigma_x | \phi_j \rangle$  are the matrix elements of the exchange self-energy of Eq. (25) and  $\chi_0^{-1}(\mathbf{r}, \mathbf{r}')$  is the inverse of the static independent particle polarisability

$$\chi_0(\mathbf{r}, \mathbf{r}') = \sum_k^{\text{occ}} \sum_{n \neq k}^{\infty} \frac{\phi_k^*(\mathbf{r}) \phi_n(\mathbf{r}) \phi_n^*(\mathbf{r}') \phi_k(\mathbf{r}')}{\varepsilon_k - \varepsilon_n} + \text{c.c.} \quad (27)$$

The exchange-potential  $v_x$  can be thought of as the best local potential approximating the non-local Fock operator [60]. It is important to emphasise, however, that by construction the total energy in Hartree–Fock is always lower (or at most equal), and thus better than in the OEPx formalism [64], because the energy minimisation in the optimised effective potential method is subject to the constraint of the wavefunctions being solutions to the Kohn–Sham Eq. (13). The eigenvalues of the OEPx

formalism for the unoccupied states, on the other hand, are closer to the photo-electron excitation energies for the materials presented in this article than the Hartree–Fock single-particle energies. The difference is due to the derivative discontinuity of the exchange potential (see Section 4.2 and 7). In the exchange-only case the discontinuity is particularly large [8, 9, 65, 66], because the conduction band states are poorly accounted for in Hartree–Fock. They are subject to a different potential than the valence states, since the Fock-operator contains the self-interaction correction only for the valence electrons. In OEPx, in contrast, the Kohn–Sham valence and conduction band states are governed by the same effective potential, which exhibits the correct asymptotic behaviour ( $1/r$  decay for large distances in finite systems). The contribution arising from the discontinuity can be calculated separately [17] and when added to the OEPx band gaps the Hartree–Fock band gaps are recovered [66].

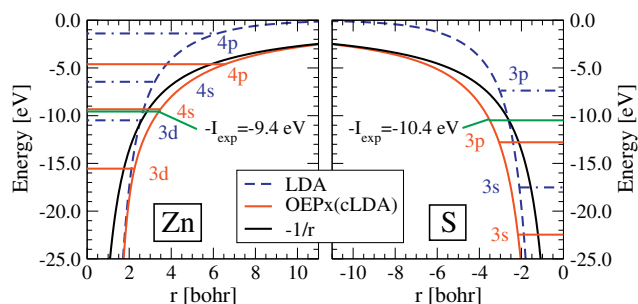
In OEPx calculations *local* correlation is frequently added by including the LDA correlation energy in the expression of the total energy (Eq. (23)). Here we follow the parametrization of Perdew and Zunger [18] for the correlation energy density  $\varepsilon_c^{\text{HEG}}[n]$  of the homogeneous electron gas based on the data of Ceperley and Alder [67]. This combination will in the following be denoted OEPx(cLDA).

## 4 OEPx(cLDA) + $G_0W_0$

**4.1 From the atom to the solid** Having established that the removal of the self-interaction in the OEPx(cLDA)-KS approach is the distinguishing feature compared to KS-LDA or KS-GGA calculations; we will now illustrate how this leads to an opening of the band gap in solids. For this it is illuminating to start from the eigenvalues of the isolated atoms, depicted in Figs. 4 and 5. The Kohn–Sham potential in OEPx(cLDA) (red curves) is essentially self-interaction free and follows the correct asymptotic  $-e^2/r$  potential outside the atom (black curves), whereas the LDA potential (blue, dashed curves) decays exponentially fast. The strong underbinding of the electrons inherent to the LDA is thus greatly reduced in the OEPx(cLDA) approach resulting in a downward shift of all atomic states. The valence electrons of the anion (N, O, S, Se) that take part in the chemical bonding are localised stronger than the respective valence electrons in the cation (Al, Ga, In, Sc, Zn, Cd). The stronger localisation leads to a larger self-interaction correction resulting in a net relative downwards shift of the anion levels with respect to the relevant cation levels (see Fig. 5).

The same behaviour is also observed in SIC-LDA calculations [68] as Table 1 demonstrates. In fact for the anions O, S and Se and the 4(5)s state in Zn(Cd), Kohn–Sham eigenvalues obtained in the SIC-LDA calculations agree closely with the OEPx(cLDA) ones. The most striking differences are found for the semicore d-electrons, which are significantly lower in SIC-LDA than in OEPx(cLDA).

<sup>12</sup> Later in this section we will reintroduce the correlation energy in an approximate form that is commonly used in connection with exact-exchange DFT calculations.



**Figure 4** (online colour at: [www.pss-b.com](http://www.pss-b.com)) Effective Kohn–Sham potential for the neutral Zn (left panel) and S (right panel) atom: the OEPx(cLDA) potential (red curves) reproduces the correct asymptotic decay  $-e^2/r$  (black curves), whereas the LDA (blue, dashed curves) decays exponentially and thus underbinds the electrons. The atomic Kohn–Sham eigenvalues (shown as horizontal lines) are lowered in the OEPx(cLDA) approach compared to the LDA resulting in better agreement with the experimentally measured ionisation potentials (green horizontal lines).

This is due to the fact that the potential in SIC-LDA is non-local. Localising it by means of the optimised effective potential approach denoted here by OEP(SIC-LDA) [69] leads to a significantly narrower spectrum, i.e. the d-electrons (as well as all lower lying states) are moved up in energy. The same behaviour has also been observed for the local/non-local pair of OEPx(cLDA)/Hartree–Fock [71]. The Kohn–Sham energies for both the valence as well as the semicore d-electrons then agree quite well in OEPx(cLDA) and OEP(SIC-LDA).

Taking the example of ScN we demonstrate that the large relative shift of the atomic anion state (N 2p) also translates to the solid<sup>13</sup>. In ScN the scandium atom donates

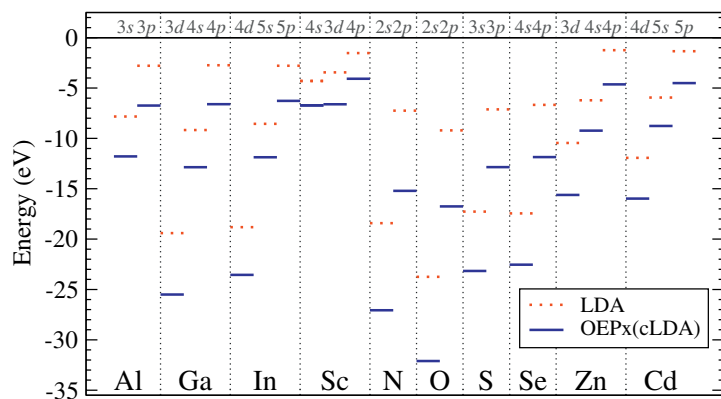
<sup>13</sup> This was for example realised by Vogl et al. who therefore suggested to use SIC-LDA pseudopotentials in regular LDA calculations for the solid [68, 72, 73]. While this idea is appealing, using a different exchange–correlation functionals for the pseudopotential than for the solid has to be taken with care. In fact LDA calculations performed with OEPx(cLDA) pseudopotentials, that include the d-electrons but not the remaining semicore states in the valence, yield greatly improved d-electron binding energies in GaN and InN, whereas OEPx(cLDA) calculations using the same OEPx(cLDA) pseudopotentials worsens the results again [15].

its two 4s and single 3d-electron to the nitrogen atom (cf. Fig. 6). The five d-states of Sc hybridize with the three valence p-states of the neighboring N atoms in the rock-salt structure of ScN, forming three p-like bonding, three d-like anti-bonding  $t_{2g}$  and two d-like non-bonding  $e_g$  bands at the  $\Gamma$ -point<sup>14</sup>. The upper three valence bands correspond to the bonding states and originate mainly from the N 2p-states with some admixture of the Sc 3d-states, while the lowest conduction bands are the anti-bonding  $t_{2g}$  states with Sc 3d-character. The two bands derived from the non-bonding  $e_g$  states are more than 1 eV higher in energy [34].

Inspection of the difference between the exchange potential in OEPx(cLDA) ( $v_x^{\text{OEPx(cLDA)}}$ ) and LDA ( $v_x^{\text{LDA}}$ ) shown in Fig. 6(a) reveals that the large relative shift of the atomic anion state (N 2p) indeed translates to the solid. Figure 6(a) illustrates that  $v_x^{\text{OEPx(cLDA)}}$  is significantly less attractive than  $v_x^{\text{LDA}}$  in the Sc regions and more attractive around the N atoms. This difference in  $v_x$  leads to a significant electron density redistribution (shown in Fig. 6(b)). The electron transfer from the Sc to the N regions gives rise to an increase in the electron negativity difference between the cation and the anion and thus increases the bond polarisation. This in turn, leads to an opening of the band gap – consistent with our OEPx(cLDA) band structure calculations [34].

In the II–VI compounds and group-III-nitrides, this mechanism is also responsible for an opening of the band gap in OEPx(cLDA) compared to LDA, but it is complemented by a contribution arising from the coupling between the anion semicore d-electrons and the 2p-electrons of the anion. Contrary to ScN the cation d-shell of these compounds is fully filled and the bonding and anti-bonding bands are formed between the 2p-state of the anion and the highest occupied s-states of the respective cation. The bands derived from the d-electrons fall energetically between the anion 2p- and 2s-bands. The difference in the partial electron densities plotted along the [111] direction

<sup>14</sup> At other points in the Brillouin zone the 4s states hybridise, too, which does not change the argument, since the relative shift of the 4s and the 3d to the N 2p level are almost identical.



**Figure 5** (online colour at: [www.pss-b.com](http://www.pss-b.com)) Kohn–Sham eigenenergies of the isolated atoms: The removal of the self-interaction in the OEPx(cLDA) (solid lines) leads to a systematic lowering of all atomic levels compared to the LDA (dotted lines). The downward shift is larger for more localised atomic states, such as the outer valence s- and p-states in N, O, S and Se.



**Table 1** Kohn–Sham eigenvalue spectrum of selected isolated atoms in eV. For the valence states SIC-LDA [68], OEPx(cLDA) and OEP calculations for the SIC-LDA functional (OEP(SIC-LDA)) [69] yield very similar energies and improve significantly on the LDA compared to experiment [70] (remaining difference in the two SIC formulations are due to different parametrisation of the LDA functional). For the semicore d-electrons (as well as the lower states not shown here) calculations with a local potential (OEP) give higher eigenvalues than their non-local counterpart.

	Zn		Cd		O		S		Se	
	4s	3d	5s	4d	2p	2s	3p	3s	4p	4s
exp.	−9.4	−17.2	−9.0	−17.6	−13.6	−28.5	−10.4	−20.3	−9.8	−20.2
LDA	−6.2	−10.4	−6.0	−11.9	−9.2	−23.8	−7.1	−17.3	−6.7	−17.5
SIC-LDA	−9.3	−20.0	−8.9	−18.9	−16.5	−31.0	−11.4	−22.4	−10.5	−22.3
OEPx(cLDA)	−9.2	−15.6	−8.8	−16.0	−16.8	−32.1	−12.9	−23.2	−11.9	−22.5
OEP(SIC-LDA)	−9.1	−15.0	−8.4	−15.8	−17.9	−32.8	−12.1	−22.2	−11.1	−21.3

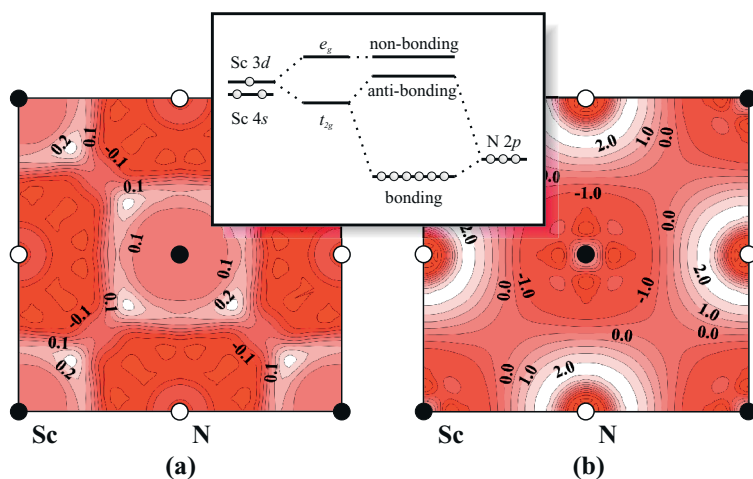
for the example of ZnS and GaN in Fig. 7 reveals that the removal of the self-interaction leads to a localisation of the d-electrons in OEPx(cLDA) compared to LDA. In ZnS, where the d-bands are energetically much closer to the p-derived valence bands than in GaN, the stronger localisation of the d-electrons is accompanied by a delocalisation of the p-electrons into the bonding region [32]. The localisation of the d-electrons reduces the strength of the pd repulsion and the valence bands are lowered in energy leading to a further opening of the band gap in OEPx(cLDA) [15].

**4.2 Discontinuity and the band gap** In the Kohn–Sham formalism the system of interacting electrons is mapped onto a fictitious system of non-interacting electrons as alluded to in Section 3. Even in exact DFT only the eigenvalue corresponding to the highest occupied Kohn–Sham state of a finite system can be rigorously assigned to the ionisation potential. For an extended system with well defined chemical potential (Fermi level) this is equivalent to stating that the electron chemical potential in DFT is the same as the true one [56, 74–76]. In compari-

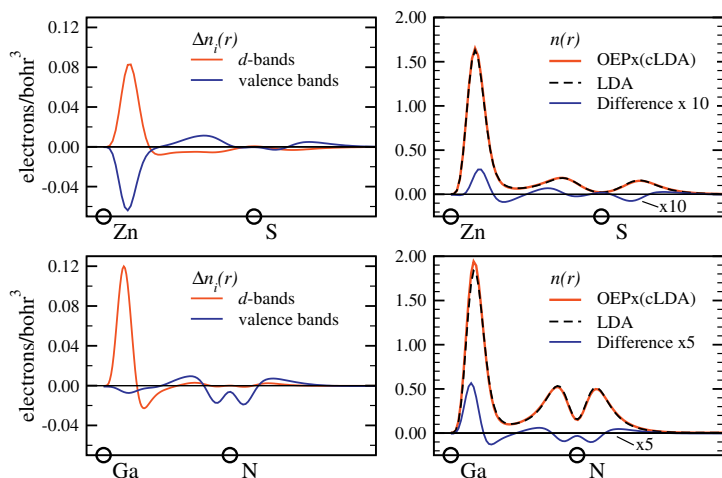
son to quasiparticle energies this leads to an important difference between the Kohn–Sham and the quasiparticle band structure. This is best illustrated by the example of the band gap of semiconductors and insulators. It can be expressed in terms of total energy differences of the  $N$ - and  $(N \pm 1)$ -particle system ( $E_{\text{gap}} = E(N+1) - 2E(N) + E(N-1)$ ). Alternatively, the band gap can be entirely formulated in terms of KS eigenvalues as the difference between the electron affinity and the ionisation potential:

$$E_{\text{gap}} = \varepsilon_{N+1}^{\text{KS}}(N+1) - \varepsilon_N^{\text{KS}}(N) \\ = \underbrace{\varepsilon_{N+1}^{\text{KS}}(N+1) - \varepsilon_{N+1}^{\text{KS}}(N)}_{\Delta_{xc}} + \underbrace{\varepsilon_{N+1}^{\text{KS}}(N) - \varepsilon_N^{\text{KS}}(N)}_{E_{\text{gap}}^{\text{KS}}}. \quad (28)$$

Here  $\varepsilon_i^{\text{KS}}(N)$  denotes the  $i$ th Kohn–Sham state of an  $N$ -particle system.  $E_{\text{gap}}^{\text{KS}}$  is the eigenvalue gap of a Kohn–Sham calculation for the  $N$ -particle system, given by the difference between the highest occupied and lowest unoccupied state. In a solid, in which  $N \gg 1$ , the addition of an extra electron only induces an infinitesimal change of the density. Therefore, the two Kohn–Sham potentials must



**Figure 6** (online colour at: www.pss-b.com) The difference between (a) the OEPx(cLDA) and LDA exchange potentials (in Hartree atomic units) and (b) the valence electron densities (in electrons/unit cell) of ScN (the 1s of N and the 1s, 2s and 2p of Sc form the core of the pseudopotentials) for one of the square faces of the conventional rock-salt unit cell. Black circles denote Sc and white circles N atoms. The inset shows the bonding scheme together with the electron filling.



**Figure 7** (online colour at: [www.pss-b.com](http://www.pss-b.com)) Comparison between OEPx(cLDA) and LDA results for the electron density and the partial densities difference for ZnS (top) and GaN (bottom). Positive density differences indicate an accumulation in OEPx(cLDA). *Left hand side*: Partial density differences ( $\Delta n_i(r)$ ) along the [111] direction through the unit cell. The removal of the self-interaction in OEPx(cLDA) leads to a stronger localisation of the d-electrons (red) on the Zn and Ga atoms. In GaN this localisation is not accompanied by a delocalisation of the valence electrons as in ZnS leading to a visible change in the electron density. *Right hand side*: OEPx(cLDA) (red), LDA (black dashed) and their difference (blue line – magnified by a factor of 10 for ZnS and 5 for GaN.)

be practically the same inside the solid up to a constant shift and, consequently, the Kohn–Sham wavefunctions do not change. The energy difference  $\Delta_{xc}$  can then only arise from a spatially constant discontinuity of the exchange–correlation energy upon changing the particle number

$$\Delta_{xc} = \left( \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \Big|_{N+1} - \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \Big|_N \right) + \mathcal{O}\left(\frac{1}{N}\right), \quad (29)$$

since changes in the Hartree potential will be negligible for  $\Delta n(\mathbf{r}) \rightarrow 0$  [77–79]. This also implies that the dispersion of bands will not be affected by the discontinuity. The conduction bands will merely be shifted relative to the valence bands.

Whether the (considerable) underestimation of the LDA Kohn–Sham band gaps reported here and elsewhere is a deficiency of the LDA itself or a fundamental property of the Kohn–Sham approach has been a longstanding debate. Similarly it has been argued that the band gaps in the OEPx approach should be larger than the true Kohn–Sham gap, since correlation is omitted [9]. The exchange–correlation energy in the LDA is a continuous function of the density with respect to changes in the particle number and will thus not exhibit a discontinuity even if the band gap were calculated by means of total energy differences. The OEPx formalism, on the other hand, exhibits a derivative discontinuity [76, 80], which would be taken into account if the excitation energies were calculated by computing total energy differences between the  $N$  and the  $N \pm 1$  electron system [80]. In KS-DFT, however, this derivative discontinuity does not enter the calculation. The opening of the band gap discussed in Section 4.1 is therefore due to the removal of the self-interaction and not the derivative discontinuity. Adding the corresponding derivative discontinuity would make the band gaps comparable to the Hartree–Fock ones [9, 65, 81], which are significantly too large.

Further substance to the notion that even the exact Kohn–Sham potential would give rise to a band gap underestimation was first given by Gunnarsson and Schön-

hammer [82, 83] and Godby et al. [79, 84] and recently by Grüning et al. [81]. Gunnarsson and Schönhammer derive their conclusions from an exactly solvable, Hubbard-like model, whereas Godby et al. and Grüning et al. use the Sham–Schlüter formalism to generate the local exchange–correlation potential that corresponds to the  $G_0W_0$  self-energy. They show that the resulting potential closely resembles that of the LDA. Also the Kohn–Sham eigenvalue differences are very similar to the LDA ones, whereas  $G_0W_0$  calculations, that incorporate information of the  $(N \pm 1)$ -particle system in a natural way (see Section 2), generally give quasiparticle band gaps to within 0.1–0.2 eV of the experimental values [2].

The good agreement of the OEPx(cLDA) band gaps with experiment reported previously [7–13] can therefore be regarded as fortuitous. In fact inspection of Fig. 3 shows that the OEPx(cLDA) band gaps are lower than the experimental ones [32] and hence lower than those of previous studies. This is due to the fact, that unlike in the earlier work we have explicitly included the d-electrons as valence electrons in our pseudopotential calculations. The pd repulsion pushes the valence bands up in energy and shrinks the gap [15]. All-electron OEPx calculations for CdS and ZnS, on the other hand, report band gaps considerably higher than the experimental ones [65]. The origin of this discrepancy between pseudopotential and all-electron calculations is still under debate. A band gap underestimation in OEPx has also been observed for noble gas solids and insulators [14, 65, 66].

Table 2 illustrates for four representative compounds that adding LDA correlation to the OEPx approach has a marginal effect on the calculated band gaps (and the band structure in general – not shown). ZnO is the only material of those studied here where the difference exceeds 0.1 eV. The remaining difference between OEPx(cLDA) and experiment is recovered to a large degree in the  $G_0W_0$  quasiparticle energy calculations as Table 2 and Fig. 3 demonstrate.

The advantage of the OEPx(cLDA) +  $G_0W_0$  approach is that it proves to be sufficient to include the d-electrons of

**Table 2** DFT and quasiparticle band gaps in eV for ZnO, ZnS, CdS, and GaN in the zinc-blende structure sorted in increasing energy from top to the experimental values. The first column lists the DFT scheme and the second column denoted PP the pseudopotential used. For all-electron calculations this column denotes if the atomic sphere approximation (ASA) or the full potential (FP) was employed. "Conf." refers to the configurations of the (pseudo)atoms: d-electrons included (d), as described in the previous section, valence only (no d's), d-electrons and their respective shell included (d-shell) and all-electron (all e<sup>-</sup>). Experimental results are taken from: ZnO [85], ZnS [86], CdS [87], GaN [88] and the OEPx(cLDA) and *GW* data from: <sup>a</sup> Ref. [89], <sup>b</sup> Ref. [90], <sup>c</sup> Ref. [13], <sup>d</sup> Ref. [8], <sup>e</sup> Ref. [10], <sup>f</sup> Ref. [91], <sup>g</sup> Ref. [48], <sup>h</sup> Ref. [92], <sup>i</sup> Ref. [49], <sup>j</sup> Ref. [93], <sup>k</sup> Ref. [94], <sup>l</sup> Ref. [95], <sup>m</sup> Ref. [96]. Numbers in round brackets refer to wurtzite structures. In references with superscript *a* and *b* a model dielectric function was employed and in *e, f, g, h* a plasmon pole model was used.

DFT	PP	conf.	<i>GW</i>	ZnO	ZnS	CdS	GaN
OEPx	OEPx	d		2.34	2.94	1.84	2.76
OEPx(cLDA)	OEPx(cLDA)	d		2.57	3.08	1.96	2.88
OEPx	OEPx	d	<i>GW</i>	3.07	3.62	2.36	3.09
OEPx(cLDA)	OEPx(cLDA)	d	<i>GW</i>	3.11	3.70	2.39	3.09
experiment				(3.44)	3.80	2.48	3.30
LDA	LDA	no d's	<i>GW</i>		3.98 <sup>a</sup>	2.83 <sup>a</sup>	3.10 <sup>b</sup>
OEPx(cLDA)	OEPx(cLDA)	no d's			3.74 <sup>c</sup>		3.46 <sup>d</sup>
OEPx(cLDA)	OEPx(cLDA)	no d's	<i>GW</i>				3.49 <sup>e</sup>
LDA	LDA	d	<i>GW</i>			1.50 <sup>f</sup>	
LDA	LDA	d-shell	<i>GW</i>		3.64 <sup>g</sup>		
LDA	LDA	d-shell	<i>GW</i>		3.50 <sup>h</sup>	2.45 <sup>h</sup>	2.88 <sup>h</sup>
LDA	LDA	d-shell	<i>GW</i>		3.38 <sup>i</sup>	2.11 <sup>i</sup>	
LDA	FP	all e <sup>-</sup>	<i>GW</i>	(244) <sup>j</sup>	3.24 <sup>k</sup>		(3.03) <sup>k</sup>
LDA	ASA	all e <sup>-</sup>	<i>GW</i>	(4.06) <sup>m</sup>	3.97 <sup>l</sup>		(3.25) <sup>m</sup>

the cation explicitly as valence electrons in the pseudopotential without having to include the remaining electrons of the semicore shell, provided OEPx(cLDA) pseudopotentials [97] are used. Taking GaN as an example, the 3s-electrons in the Gallium atom are approximately 130 eV and the 3p-electrons approximately 80 eV lower than the 3d states. Resolving these more strongly localized 3s and 3p-electrons in GaN with a plane-wave basis set will hence require significantly higher plane-wave cutoffs than the 70 Ry required for the d-electrons [32]. In a pseudopotential framework it would thus make sense to explicitly include the d-electrons of the cations in the II–VI compounds and group-III-nitrides as valence electrons, but to freeze the chemically inert semicore s- and p-electrons in the core of the pseudopotential. However, due to the large spatial overlap of the atomic semicore s and p with the d wavefunctions, core-valence exchange-correlation is large in these compounds. As a consequence core-valence exchange-correlation is treated inconsistently when going from LDA to LDA +  $G_0W_0$ , if pseudopotentials are used in this fashion, because the exchange-correlation self-energy in the *GW* approach acts on the d-electrons in the solid, but cannot act on the s- and p-electrons in the semicore shell, too. The result is a severe underestimation of the LDA +  $G_0W_0$  band gaps (cf. red diamonds in Fig. 3) and d-bands that are pushed energetically into the p-derived valence bands in the II–VI compounds [32, 91, 92] (cf. Table 3). The only way to remedy this problem within LDA +  $G_0W_0$  is to free the electrons in question by performing all-electron  $G_0W_0$  calculations [94] or by using pseudopotentials that include the entire shell as valence

electrons [48, 91, 92], which in the latter case introduces formidably high plane-wave cutoffs. If, on the other hand, OEPx or OEPx(cLDA) is used for the ground state calculation, then the exchange self-energy already acts on the semicore s- and p-states in the generation of the pseudopotential. Since the exchange self-energy can be linearly decomposed into a core and a valence contribution no non-linear core corrections [98] arise in the Hartree–Fock case and they are expected to be small for OEPx(cLDA) pseudopotentials [9]. When going from OEPx(cLDA) to OEPx(cLDA) +  $G_0W_0$  core-valence exchange is therefore treated consistently, even when the semicore s- and p-electrons are frozen in the core, as long as OEPx(cLDA) pseudopotentials are used [32].

For InN the LDA starting point is so problematic (cf. Fig. 1), that even subsequent all-electron  $G_0W_0$  calculations only open the gap to 0.02–0.05 eV [94, 99]. Here the importance of removing the self-interaction from the ground state calculation has been demonstrated before by combining SIC-LDA calculations with  $G_0W_0$  calculations (albeit in a rather approximate way by adjusting the pd repulsion and combining calculations with and without the 4d-electrons in the core of the In pseudopotential) [100, 101]. The OEPx(cLDA) approach, on the other hand, predicts InN to be a semiconductor, as Fig. 1 illustrates, and  $G_0W_0$  calculations can be applied without further approximations. The size of the band gap of InN and the origins for the considerable spread in the experimental data has been a matter of intense debate over the last years. For wurtzite InN our value of 0.7 eV [33] strongly supports recent experimental findings [102–104]. The OEPx(cLDA)

**Table 3** d-electron binding energies referenced to the top of the valence band: The layout is the same as in Table 2. Experimental values taken from: <sup>a</sup> Ref. [106], <sup>b</sup> Ref. [107] <sup>c</sup> Ref. [108], <sup>d</sup> Ref. [109], and the *GW* data from: <sup>e</sup> Ref. [91], <sup>f</sup> Ref. [48], <sup>g</sup> Ref. [92], <sup>h</sup> Ref. [49], <sup>i</sup> Ref. [110], <sup>j</sup> Ref. [111] ( $U = 8$  eV,  $J = 1$  eV), <sup>k</sup> Ref. [93], <sup>l</sup> Ref. [94], <sup>m</sup> Ref. [96], <sup>n</sup> Ref. [95], <sup>o</sup> Ref. [65]. SAT denotes *GW* calculations including plasmon satellites,  $G'W'T$  calculations performed with eigenvalue self-consistency and addition of a vertex contribution. Numbers in round brackets refer to wurtzite structures. In references denoted by *d, e, f* a plasmon pole model was used.

DFT	PP	conf.	<i>GW</i>	ZnO	ZnS	CdS	GaN
1 LDA	LDA	d	<i>GW</i>	4.29	4.30	6.17	13.05
2 OEPx(cLDA)	LDA	d	<i>GW</i>	4.98	5.02	6.40	13.58
3 OEPx(cLDA)	LDA	d		4.36	5.33	6.54	12.75
4 LDA	LDA	d		5.20	6.32	7.72	14.25
5 OEPx	OEPx	d		5.12	6.91	7.57	14.85
6 OEPx(cLDA)	OEPx(cLDA)	d		5.20	7.05	7.61	15.02
7 OEPx	OEPx	d	<i>GW</i>	6.68	6.97	7.66	16.12
8 OEPx(cLDA)	OEPx(cLDA)	d	<i>GW</i>	6.87	7.08	7.75	16.15
9 experiment				(9.00) <sup>a</sup>	8.97 <sup>a</sup> 9.03 <sup>d</sup>	9.20 <sup>b</sup> 9.50 <sup>d</sup>	17.70 <sup>c</sup>
10 LDA	LDA	d	<i>GW</i>			5.20 <sup>e</sup>	
11 LDA	LDA	d-shell	<i>GW</i>		7.40 <sup>f</sup>		
12 LDA	LDA	d-shell	<i>GW</i>		6.40 <sup>g</sup>	8.10 <sup>g</sup>	15.70 <sup>g</sup>
13 LDA	LDA	d-shell	<i>GW</i>		6.87 <sup>h</sup>	8.15 <sup>h</sup>	
14 LDA	LDA	d-shell	SAT		7.90 <sup>g</sup>	9.10 <sup>i</sup>	17.30 <sup>g</sup>
15 LDA	LDA	d-shell	$G'W'T$		8.02 <sup>h</sup>	8.99 <sup>h</sup>	
16 LDA + <i>U</i>	LDA	d-shell			8.78 <sup>j</sup>		
17 LDA + <i>U</i>	LDA	d-shell	<i>GW</i>		7.10 <sup>j</sup>		
18 LDA	FP	all e <sup>-</sup>	<i>GW</i>	6.16 <sup>k</sup>	7.10 <sup>l</sup>	8.20 <sup>l</sup>	(16.40) <sup>l</sup>
19 LDA	ASA	all e <sup>-</sup>	<i>GW</i>	(5.94) <sup>m</sup>	8.33 <sup>n</sup>		(17.60) <sup>m</sup>
20 OEPx	FP	all e <sup>-</sup>			9.1 <sup>o</sup>	8.2 <sup>o</sup>	

+  $G_0W_0$  calculations have further helped to clarify the source for wide interval of experimentally observed band gaps [33].

For ScN, an emerging versatile material for promising technological applications, the electronic band structure has also been difficult to access experimentally, due to growth related problems. Similar to the metallic state in InN the strong self-interaction effects in the LDA predict ScN to be semimetallic (cf. Fig. 1) preventing a direct application of the LDA +  $G_0W_0$  approach. By using OEPx(cLDA) the ground state becomes essentially self-interaction free. This leads to an opening of the band gap and a suitable semiconducting starting point for  $G_0W_0$  calculations [34], as described in Section 4.1. The OEPx(cLDA) +  $G_0W_0$  calculations lower the OEPx(cLDA) band gap of 1.7 eV down to 0.8 eV [34], clearly supporting recent experimental findings of an indirect gap of  $0.9 \pm 0.1$  eV [105].

While LDA based  $G_0W_0$  calculations generally open the band gap from the underestimated LDA value, we observe here that the  $G_0W_0$  corrections lower the respective gaps of InN and ScN obtained in the OEPx(cLDA) approach. Since  $G_0W_0$  falls into the realm of perturbation theory (cf. Eq. (12)) this is not unusual, because it is not *a priori* clear if the quasiparticle corrections are positive or negative.

**4.3 d-electron binding energies** A known problem for the group-III-nitrides and the II–VI compounds is, that both KS-LDA and LDA based  $G_0W_0$  calculations underes-

timate the binding of the d-bands – regardless of whether applied in a pseudopotential or an all-electron fashion. For four selected compounds the center of the d-bands<sup>15</sup> is listed in Tab. 3 for the different computational schemes. For ZnS and CdS OEPx(cLDA) and OEPx(cLDA) +  $G_0W_0$  produce essentially the same d-electron binding energies. Only in ZnO quasiparticle corrections are found to lower the d-states by 1.5 eV compared to OEPx(cLDA), further reducing the pd coupling. Again, adding LDA correlation to the OEPx has essentially no effect on the energies of the bands (energy differences <0.2 eV).

At first sight it appears surprising that the removal of the self-interaction does not automatically lower the position of the d-bands in OEPx(cLDA). Closer inspection of the atomic eigenvalues in Fig. 5 reveals, however, that the cation d-states are shifted down by almost the same amount as the anion p-states. The relative position to the p-states, from which the top of the valence band is derived, therefore remains similar in OEPx(cLDA) and LDA [15]. Applying exact-exchange in a non-local fashion (Hartree–Fock) instead lowers the d-states considerably [30]. The narrower spectrum of the local exchange potential therefore results in d-electron binding energies that are lower in OEPx(cLDA) than in HF. This difference between the local and the non-local incarnation of an exchange-correlation potential pertaining to the same total energy functional is iden-

<sup>15</sup> This has been obtained by averaging over the d-bands at the  $\Gamma$ -point.



tical to that alluded to previously for the case of isolated atoms in Section 4.1. Core-valence exchange-correlation effects beyond those included via the pseudopotential are thus not expected to change this conclusion.

A recent all-electron OEPx study (line 20) has reported values for ZnS and CdS in better agreement with experiment [65]. However, it is not clear at the present stage, why for CdS the difference to our pseudopotential calculations (line 5) is merely 0.6 eV, whereas for ZnS it amounts to 2.2 eV. It is also not clear if the discrepancy is caused by core-valence linearisation in the pseudopotentials, the use of pseudo- rather than all-electron wavefunctions in the Fock-operator or if it has an entirely different origin.

Overall the binding energies obtained with our OEPx(cLDA) +  $G_0W_0$  approach agree well with other available  $GW$  calculations (lines 11–13 and 18 in Table 3), but are still about 2 eV at variance with experiment. The reason for this could be twofold. Either interactions beyond the  $GW$  approximation are required for describing the excitation of these semicore d-electrons and/or one needs to go beyond even OEPx(cLDA) as a starting point. Evidence that both self-consistency in the  $GW$  calculations as well as the inclusion of vertex effects (see Section 2.2) introduce changes in the right direction was recently given by Fleszar and Hanke. Starting from an LDA ground state only eigenvalues (and not also the wavefunctions) were iterated to self-consistency in the  $GW$  calculations. This increases the binding energy by 0.4–0.5 eV in ZnS and CdS [49]. Adding a vertex function that makes the calculation “consistent” with the LDA starting point<sup>16</sup> increases the binding energy by additional 0.4–0.5 eV (line 15) [49]. Similar conclusions were drawn from a recent study, in which  $G_0W_0$  calculations were based on LDA+ $U$  ground states [111]. With increasing  $U$  the binding energy of the d-state increases linearly in the LDA+ $U$  calculations (line 16) as the pd hybridisation reduces. The  $G_0W_0$  calculations, however, prove to be insensitive to this change and shift the d-bands up again, close to their LDA+ $G_0W_0$  positions (line 17) [111]. Previously Rohlfing et al. had devised a way to go beyond the  $GW$  approximation by including plasmon satellites in the Green’s function, denoted here by SAT (line 14 in Table 3). Although the SAT improves on the d-electron binding energies it considerably overcorrects the band gap and the valence part of the band structure [110]. Work towards a consistent description of excitation energies of the semicore d-electrons in these materials is clearly required in the future.

**5 Comments on self-sistency** As alluded to in Section 2.2 it is still a matter of debate how to perform self-consistent  $GW$  calculations for quasiparticle band structures. For the non self-consistent  $G_0W_0$  scheme our

calculations for the selected II–VI compounds and GaN have revealed a significant dependence on the starting point, if the semicore d-electrons, but not the remaining electrons of the semicore shell, are explicitly taken into account in the pseudopotential calculation. Since for the reasons given in Section 4.2 no meaningful comparison can be made between LDA +  $G_0W_0$  and OEPx(cLDA) +  $G_0W_0$  calculations unless the entire semicore shell is included in the calculation these systems are not suitable to assess the influence of the starting point on the quasiparticle band structure. In ScN, on the other hand, the semicore d-shell of Sc is not fully filled and the remaining s- and p-electrons in the semicore shell are much higher in energy. Taking GaN as an example the 3s-electrons in the Scandium atom are approximately 100 eV and the 3p-electrons approximately 60 eV higher than in Gallium. Resolving the 3s- and 3p-derived bands in ScN with plane-waves thus only requires a cutoff of 80 Ry [34] and makes ScN an ideal candidate for constructing a comparison between LDA +  $G_0W_0$  and OEPx(cLDA) +  $G_0W_0$  calculations.

However, the *negative* LDA band gap (see Table 4) impedes the direct application of the LDA +  $G_0W_0$  formalism with our  $GW$  code, since in its current implementation [112–114] a clear separation between conduction and valence bands is required. Therefore, an indirect approach is adopted. First, LDA +  $G_0W_0$  calculations are performed at a lattice constant ( $a_0 = 4.75$  Å) larger than the experimental one ( $a_0 = 4.50$  Å), where the fundamental band gap in the LDA is small but positive. We then use the LDA *volume* deformation potentials (which agree well with the OEPx(cLDA) +  $G_0W_0$  ones) to determine the corresponding LDA +  $G_0W_0$  band gaps at the equilibrium lattice constant<sup>17</sup>.

We find (Table 4) that the LDA +  $G_0W_0$  and OEPx(cLDA) +  $G_0W_0$  calculations, starting from the two extremes (negative band gap in LDA, 0.8 eV overestimation in OEPx(cLDA)), yield quasiparticle band gaps that agree to within 0.3 eV. Since the LDA-based calculations are close to the limit of metallic screening, whereas the OEPx(cLDA)-based calculations form the opposite ex-

**Table 4** Calculated and experimental band gaps ( $E_g$ ) of ScN (in eV). [ $G_0W_0$ ]<sub>average</sub> denotes the arithmetic average between the OEPx(cLDA) +  $G_0W_0$  and LDA +  $G_0W_0$  results (see text).

approach	$E_g^{r-r}$	$E_g^{r-x}$	$E_g^{x-x}$
OEPx(cLDA) + $G_0W_0$	3.51	0.84	1.98
LDA + $G_0W_0$	3.71	1.14	2.06
[ $G_0W_0$ ] <sub>average</sub>	3.62	0.99	2.02
OEPx(cLDA)	4.53	1.70	2.59
GGA	2.43	−0.03	0.87
LDA	2.34	−0.15	0.75
experiment			
Ref. [12]	~3.8	−1.30	2.40
Ref. [105]		0.9 ± 0.1	2.15

<sup>16</sup> Strictly spoken  $\Sigma$  is only zero on the first iteration when starting from Hartree theory (see Section 2.2). Starting from a Kohn–Sham calculations implies that  $\Sigma = v_{xc}$  in which case the vertex function is not given by delta functions, but can still be analytically derived [51].

<sup>17</sup> This approach is in principle not limited to band gaps and can equally well be applied to the full band structure for example.

treme of starting from a completely self-interaction free exchange-correlation functional, we expect the results of a self-consistent  $GW$  calculation to fall in the range between the LDA +  $G_0W_0$  and OEPx(cLDA) +  $G_0W_0$  calculations. From these results we estimate the maximum error bar associated with omitting self-consistency in  $GW$  to be the difference from the arithmetic averages for  $E_g^{I-I}$ ,  $E_g^{I-X}$  and  $E_g^{X-X}$  (reported in Table 4) to the largest deviation between LDA +  $G_0W_0$  and OEPx(cLDA) +  $G_0W_0$ , i.e. of the order of 0.15 eV for ScN [34].

**6 Conclusions** We have presented the combination of quasiparticle energy calculations in the  $G_0W_0$  approximation with DFT calculations in the OEPx(cLDA) approach. Using OEPx(cLDA) instead of LDA or GGA removes the inherent self-interaction of the latter from the ground state calculation. Starting from the individual atoms we have illustrated how this leads to an opening of the Kohn–Sham band gap in the solid. In the spirit of perturbation theory OEPx(cLDA) thus provides a more suitable starting point for  $G_0W_0$  calculations than LDA or GGA. For materials for which LDA erroneously predicts a (semi)metallic state, like e.g. InN or ScN, OEPx(cLDA) yields a semiconducting ground state, which unlike in the LDA case, permits a direct application of the  $G_0W_0$  approximation. For the II–VI compounds and group-III-nitrides presented here as an example, the band gaps in the OEPx(cLDA) +  $G_0W_0$  approach are in excellent agreement with experiment and the position of the semicore d-electron bands no worse than in previous LDA +  $G_0W_0$  calculations.

**7 Outlook** A disadvantage of the OEPx approach is that it yields total energies and thus structural properties comparable to Hartree–Fock. For the materials considered here all lattice parameters are known from experiment, but this might not be the case for surfaces, defects, nanostructures or new materials. Then, in cases where LDA and GGA are known to fail, exchange-correlation functionals that appropriately include exact-exchange are needed.

In the most pragmatic approach a portion of Hartree–Fock exchange is mixed with a different portion of local exchange and correlation<sup>18</sup>

$$E_{xc}^{\text{hyb}} = E_{xc}^{\text{DFT}} + \alpha(E_x^{\text{HF}} - E_x^{\text{DFT}}), \quad (30)$$

as already alluded to in the introduction. Prominent examples of these so called hybrid functionals are PBE0 [19–21] and B3LYP [22–24]. The factor  $\alpha$  in front of the Hartree–Fock exact-exchange term is equivalent to a constant static screening function. Alternatively a more complex screening function can be chosen as in the screened-exchange (sX-LDA) approach [17, 28, 115, 116], the  $\Sigma$ -GKS scheme [117] or the Heyd–Scuseria–Ernzerhof (HSE) functional [25–27]. An appealing feature of the hy-

brid functional approach is that it incorporates part of the derivative discontinuity of the exchange-correlation energy [17], because a quasiparticle equation (Eq. (11)) with a non-local potential is solved. For semiconductors and insulators this leads to a much improved description of band gaps [27, 66, 115, 116, 118–120]. From the point of perturbation theory this is again beneficial since the perturbation required by the  $GW$  (or eventually the true) self-energy is smaller. However, the dependence on universal (i.e. material independent) parameters questions the *ab initio* character of today’s hybrid functionals.

While ample experience with hybrid functionals exists in the quantum chemistry community, applications to solids are only slowly emerging and many issues remain open or controversial i.e. computational efficiency, basis sets, pseudopotentials or the choice of functional. In particular with regard to structural properties systematic studies are required. Conversely,  $GW$  quasiparticle energy calculations have traditionally fallen into the realm of condensed matter physics, but also here similar open questions remain (most of which have not been explicitly discussed in this article) including that of self-consistency. Again in a more pragmatic fashion, the latter has recently been approached by mapping the  $G_0W_0$  self-energy back onto a static, non-local potential. The eigenenergies and eigenfunctions of the new non-local, hermitian Hamiltonian then serve as input for the next  $G_0W_0$  cycle until self-consistency is reached [46, 121]. This approach shows promising success for different types of materials ranging from semiconductors, insulators and metals to transition metal oxides and f-electron systems [29, 46, 121, 122].

Although the  $G_0W_0$  approximation is currently the state of the art approach to calculate defect levels for solids from first principles it has so far only been applied in a few cases [123–128]. Ideally the  $GW$  calculations would encompass the determination of the defect structure, too, but so far  $GW$  total energy calculations, that would be required for this task, are still in their infancy. Meanwhile a combination of exact-exchange-based DFT functionals (for the structural properties) with quasiparticle energy calculations in the  $GW$  approach and the  $G_0W_0$  approximation (for the spectral properties) offer the possibility to develop a better understanding of defect properties untainted by structural artefacts caused by the self-interaction or the band gap underestimation of Kohn–Sham. First  $G_0W_0$  calculations for bulk semiconductors based on hybrid functionals appear to be promising [29–31], but more work along these lines is needed in the future.

**Acknowledgements** We acknowledge stimulating discussions with Sixten Boeck, Martin Fuchs, Matthias Wahn, Hong Jiang, Wolfram Zieger, Claudia DeRubeis, Heidi Henrickson, and Christoph Freysoldt. This work was in part supported by the Volkswagen Stiftung/Germany, the DFG research group “nitride based nanostructures” and the EU’s 6th Framework Programme through the NANOQUANTA (Contract No. NMP4-CT-2004-500198) Network of Excellence.

<sup>18</sup> Some hybrid schemes introduce additional parameters for mixing different portions of the local-spin density (LSD) functional with GGA exchange and correlation in the  $E_{xc}^{\text{DFT}}$  term [21].

## References

- [1] L. Hedin, *Phys. Rev.* **139**, A 796 (1965).
- [2] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, *Solid State Phys.*, *Adv. Res. Appl.* **54**, 1 (2000).
- [3] G. Onida, L. Reining, and A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).
- [4] T. Kotani, *Phys. Rev. B* **50**, 14816 (1994); *Phys. Rev. B* **51**, 13903(E) (1995).
- [5] T. Kotani, *Phys. Rev. Lett.* **74**, 2989 (1995).
- [6] T. Kotani and H. Akai, *Phys. Rev. B* **54**, 16502 (1996).
- [7] A. Görling, *Phys. Rev. B* **53**, 7024 (1996).
- [8] M. Städele, J. A. Majewski, P. Vogl, and A. Görling, *Phys. Rev. Lett.* **79**, 2089 (1997).
- [9] M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, *Phys. Rev. B* **59**, 10031 (1999).
- [10] W. G. Aulbur, M. Städele, and A. Görling, *Phys. Rev. B* **62**, 7121 (2000).
- [11] A. Fleszar, *Phys. Rev. B* **64**, 245204 (2001).
- [12] D. Gall, M. Städele, K. Järrendahl, I. Petrov, P. Desjardins, R. T. Haasch, T. Y. Lee, and J. E. Greene, *Phys. Rev. B* **63**, 125119 (2001).
- [13] N. Fitzner, A. Kuligk, R. Redmer, M. Städele, S. M. Goodnick, and W. Schattke, *Phys. Rev. B* **67**, R201201 (2003).
- [14] R. J. Magyar, A. Fleszar, and E. K. U. Gross, *Phys. Rev. B* **69**, 045111 (2004).
- [15] A. Qteish, A. I. Al-Sharif, M. Fuchs, M. Scheffler, S. Boeck, and J. Neugebauer, *Phys. Rev. B* **72**, 155317 (2005).
- [16] S. Kümmel and L. Kronik, *Rev. Mod. Phys.* **80**, 3 (2007).
- [17] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, *Phys. Rev. B* **53**, 3764 (1996).
- [18] J. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [19] J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
- [20] M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).
- [21] C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- [22] A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- [23] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- [24] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- [25] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
- [26] J. Heyd and G. E. Scuseria, *J. Chem. Phys.* **120**, 7274 (2004).
- [27] J. Heyd and G. E. Scuseria, *J. Chem. Phys.* **121**, 1187 (2004).
- [28] D. M. Bylander and L. Kleinman, *Phys. Rev. B* **41**, 7868 (1990).
- [29] F. Bruneval, N. Vast, and L. Reining, *Phys. Rev. B* **74**, 045102 (2006).
- [30] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, *Phys. Rev. B* **76**, 115109 (2007).
- [31] Z. Zanolli, F. Fuchs, J. Furthmüller, U. von Barth, and F. Bechstedt, *Phys. Rev. B* **75**, 245121 (2007).
- [32] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, *New J. Phys.* **7**, 126 (2005).
- [33] P. Rinke, A. Qteish, M. Winkelkemper, D. Bimberg, J. Neugebauer, and M. Scheffler, *Appl. Phys. Lett.* **89**, 161919 (2006).
- [34] A. Qteish, P. Rinke, J. Neugebauer, and M. Scheffler, *Phys. Rev. B* **74**, 245208 (2006).
- [35] F. J. Himpsel, *Adv. Phys.* **32**, 1 (1983).
- [36] E. W. Plummer and W. Eberhardt, *Adv. Chem. Phys.* **49**, 533 (1982).
- [37] L. Kevan (ed.), *Angle-Resolved Photoemission* (Elsevier, Amsterdam, 1992).
- [38] V. Dose, *Surf. Sci. Rep.* **5**, 337 (1985).
- [39] N. V. Smith, *Rep. Prog. Phys.* **51**, 1227 (1988).
- [40] J. C. Fuggle and J. E. Inglesfield (eds.), *Unoccupied Electronic States* (Springer-Verlag, 1992).
- [41] C. O. Almbladh and L. Hedin, in: *Handbook on Synchrotron Radiation*, edited by E. E. Koch (North Holland, 1983), pp. 607–904.
- [42] W. Ku and A. G. Eguiluz, *Phys. Rev. Lett.* **89**, 126401 (2002).
- [43] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, *Phys. Rev. B* **69**, 125212 (2003).
- [44] K. Delaney, P. García-González, A. Rubio, P. Rinke, and R. W. Godby, *Phys. Rev. Lett.* **93**, 249701 (2004).
- [45] N. E. Zein, S. Y. Savrasov, and G. Kotliar, *Phys. Rev. Lett.* **96**, 226403 (2006).
- [46] M. van Schilfgaarde, T. Kotani, and S. Faleev, *Phys. Rev. Lett.* **96**, 226402 (2006).
- [47] B. Holm and U. von Barth, *Phys. Rev. B* **57**, 2108 (1998).
- [48] W. Luo, S. Ismail-Beigi, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **66**, 195215 (2002).
- [49] A. Fleszar and W. Hanke, *Phys. Rev. B* **71**, 045207 (2005).
- [50] M. Marsili, O. Pulci, F. Bechstedt, and R. Del Sole, *Phys. Rev. B* **72**, 115415 (2005).
- [51] R. Del Sole, L. Reining, and R. W. Godby, *Phys. Rev. B* **49**, 8024 (1994).
- [52] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, *Phys. Rev. Lett.* **94**, 186402 (2005).
- [53] A. J. Morris, M. Stankovski, K. T. Delaney, P. Rinke, P. García-González, and R. W. Godby, *Phys. Rev. B* **76**, 155106 (2007).
- [54] For the upper valence and conduction bands of standard semiconductors numerical investigations indicate that this approximation is well justified [2, 129], but it breaks down for certain surface [130–133] and cluster states [134, 135].
- [55] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [56] W. Kohn and K. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [57] A. Görling and M. Levy, *Phys. Rev. A* **50**, 196 (1994); *Phys. Rev. B* **47**, 13105 (1993).
- [58] O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).
- [59] W. Nelson, P. Bokes, P. Rinke, and R. Godby, *Phys. Rev. A* **75**, 032505 (2007).
- [60] M. E. Casida, *Phys. Rev. A* **51**, 2005 (1995).
- [61] R. Grabo, T. Kreibich, S. Kurth, and E. K. U. Gross, in: *Strong Coulomb Correlations in Electronic Structure Calculations: Beyond the Local Density Approximation*, edited by V. I. Anisimov (Gordon and Breach, New York, 2000), pp. 203–311.
- [62] Y. M. Niquet, M. Fuchs, and X. Gonze, *J. Chem. Phys.* **118**, 9504 (2003).
- [63] E. Engel, in: *A Primer in Density-Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. Marques (Springer, Berlin, 2003), pp. 56–122.

- [64] S. Ivanov and M. Levy, *J. Chem. Phys.* **119**, 7087 (2003).
- [65] S. Sharma, J. K. Dewhurst, and C. Ambrosch-Draxl, *Phys. Rev. Lett.* **95**, 136402 (2005).
- [66] M. Grüning, A. Marini, and A. Rubio, *Phys. Rev. B* **74**, R161103 (2006).
- [67] D. Ceperley and B. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- [68] D. Vogel, P. Krüger, and J. Pollmann, *Phys. Rev. B* **54**, 5495 (1996).
- [69] Hong Jiang, private communications.
- [70] C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Stand. USA, Circ. No. 467, Vol. I (U.S. GPO, Washington, DC, 1949); Vol. II (U.S. GPO, Washington, DC, 1952); Vol. III (U.S. GPO, Washington, DC, 1958).
- [71] K. Aashamar, T. M. Luke, and J. D. Talman, *Phys. Rev. A* **19**, 6 (1979).
- [72] D. Vogel, P. Krüger, and J. Pollmann, *Phys. Rev. B* **52**, R14316 (1995).
- [73] D. Vogel, P. Krüger, and J. Pollmann, *Phys. Rev. B* **55**, 12836 (1997).
- [74] C. Almbladh and U. von Barth, *Phys. Rev. B* **31**, 3231 (1985).
- [75] M. Levy, J. P. Perdew, and V. Sahni, *Phys. Rev. A* **30**, 2745 (1984).
- [76] M. E. Casida, *Phys. Rev. B* **59**, 4694 (1999).
- [77] J. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).
- [78] L. Sham and M. Schlüter, *Phys. Rev. Lett.* **51**, 1888 (1983).
- [79] R. W. Godby, M. Schlüter, and L. J. Sham, *Phys. Rev. Lett.* **56**, 2415 (1986).
- [80] J. B. Krieger, Y. Li, and G. J. Iafrate, *Phys. Rev. B* **45**, 101 (1992).
- [81] M. Grüning, A. Marini, and A. Rubio, *J. Chem. Phys.* **124**, 154108 (2006).
- [82] K. Schönhammer and O. Gunnarsson, *Phys. Rev. Lett.* **56**, 1968 (1986).
- [83] K. Schönhammer and O. Gunnarsson, *J. Phys. C, Solid State Phys.* **20**, 3675 (1987).
- [84] R. W. Godby, M. Schlüter, and L. J. Sham, *Phys. Rev. B* **37**, 10159 (1988).
- [85] O. Madelung (ed.), *Numerical Data and Functional Relationships in Science and Technology*, New Series, Group III, Vol. 22, Pt. a (Springer, Berlin, 1987).
- [86] O. Madelung (ed.), *Landolt-Börnstein, Semiconductors: Intrinsic Properties of Group IV Elements and III–V, II–VI and I–VII Compounds*, New Series, Group III, Vol. 22, Pt. a (Springer, Berlin, 1987).
- [87] D. R. T. Zahn, G. Kudlek, U. Rossow, A. Hoffmann, I. Broser, and W. Richter, *Adv. Mater. Opt. Electron.* **3**, 11 (1994).
- [88] G. Ramírez-Flores, H. Navarro-Contreras, A. Lastras-Martínez, R. C. Powell, and J. E. Greene, *Phys. Rev. B* **50**, 8433 (1994).
- [89] O. Zakharov, A. Rubio, X. Blase, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **50**, 10780 (1994).
- [90] A. Rubio, J. L. Corkill, M. L. Cohen, E. L. Shirley, and S. G. Louie, *Phys. Rev. B* **48**, 11810 (1993).
- [91] M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. Lett.* **75**, 3489 (1995).
- [92] M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. B* **57**, 6485 (1998).
- [93] M. Usuda, N. Hamada, T. Kotani, and M. van Schilf-gaarde, *Phys. Rev. B* **66**, 125101 (2002).
- [94] T. Kotani and M. van Schilf-gaarde, *Solid State Commun.* **121**, 461 (2002).
- [95] M. Oshikiri and F. Aryasetiawan, *Phys. Rev. B* **60**, 10754 (1999).
- [96] M. Oshikiri and F. Aryasetiawan, *J. Phys. Soc. Jpn.* **69**, 2113 (2000).
- [97] M. Moukara, M. Städele, J. A. Majewski, P. Vogl, and A. Görling, *J. Phys.: Condens. Matter* **12**, 6783 (2000).
- [98] S. G. Louie, S. Froyen, and M. L. Cohen, *Phys. Rev. B* **26**, 1738 (1982).
- [99] M. Usuda, N. Hamada, K. Shiraiishi, and A. Oshiyama, *Jpn. J. Appl. Phys.* **43**, L407 (2004).
- [100] F. Bechstedt and J. Furthmüller, *J. Cryst. Growth* **246**, 315 (2002).
- [101] J. Furthmüller, P. H. Hahn, F. Fuchs, and F. Bechstedt, *Phys. Rev. B* **72**, 205106 (2005).
- [102] V. Y. Davydov, A. A. Klochikhin, R. P. Seisyan, V. V. Emtsev, S. Ivanov, F. Bechstedt, J. Furthmüller, H. Harima, A. V. Mudryi, J. Aderhold, O. Semchinova, and J. Graul, *phys. stat. sol. (b)* **229**, R1 (2002).
- [103] J. Wu, W. Walukiewicz, K. M. Yu, J. W. Ager III, E. E. Haller, H. Lu, W. Schaff, Y. Saiton, and Y. Nanishi, *Appl. Phys. Lett.* **80**, 3967 (2002).
- [104] Y. Nanishi, Y. Saito, and T. Yamaguchi, *Jpn. J. Appl. Phys.* **42**, 2549 (2003).
- [105] H. A. Al-Britthen, A. R. Smith, and D. Gall, *Phys. Rev. B* **70**, 045303 (2004).
- [106] R. Weidemann, H. E. Gumlich, M. Kupsch, H. U. Mid-delmann, and U. Becker, *Phys. Rev. B* **45**, 1172 (1992).
- [107] O. Madelung (ed.), *Semiconductors: Technology of III–V, II–VI and Non-Tetrahedrally Bonded Compounds*, Landolt-Börnstein, New Series, Group III, Vol. 17, Pt. d (Springer, Berlin, 1982).
- [108] S. A. Ding, G. Neuhold, J. H. Weaver, P. Häberle, K. Horn, O. Brandt, H. Yang, and K. Ploog, *J. Vac. Sci. Technol. A* **14**, 819 (1996).
- [109] L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalczyk, and D. A. Shirley, *Phys. Rev. B* **9**, 600 (1974).
- [110] M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. B* **56**, R7065 (1997).
- [111] T. Miyake, P. Zhang, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **74**, 245213 (2006).
- [112] M. M. Rieger, L. Steinbeck, I. White, H. Rojas, and R. Godby, *Comput. Phys. Commun.* **117**, 211 (1999).
- [113] L. Steinbeck, A. Rubio, L. Reining, M. Torrent, I. White, and R. Godby, *Comput. Phys. Commun.* **125**, 105 (2000).
- [114] C. Freysoldt, P. Eggert, P. Rinke, A. Schindlmayr, R. W. Godby, and M. Scheffler, *Comput. Phys. Commun.* **176**, 1 (2007).
- [115] R. Asahi, W. Mannstadt, and A. J. Freeman, *Phys. Rev. B* **59**, 7486 (1999).
- [116] S. H. Rhim, M. Kim, A. J. Freeman, and R. Asahi, *Phys. Rev. B* **71**, 045202 (2005).
- [117] P. Sánchez-Friera and R. W. Godby, *Phys. Rev. Lett.* **85**, 5611 (2000).
- [118] J. Muscat, A. Wander, and N. M. Harrison, *Chem. Phys. Lett.* **342**, 397 (2001).



- [119] F. Cora, M. Alfredsson, G. Mallia, D. S. Middlemiss, W. C. Mackrodt, R. Dovesi, and R. Orlando, in: *Principles and applications of density functional theory in inorganic chemistry II*, edited by N. Kaltsoyannis and J. E. McGrady (Springer, Berlin, 2004), pp. 171–232.
- [120] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, *J. Chem. Phys.* **124**, 154709 (2006).
- [121] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, *Phys. Rev. Lett.* **93**, 126406 (2004).
- [122] F. Bruneval, N. Vast, L. Reining, M. Izquierdo, F. Sirotti, and N. Barrett, *Phys. Rev. Lett.* **97**, 267601 (2006).
- [123] B. Králík, E. K. Chang, and S. G. Louie, *Phys. Rev. B* **57**, 7027 (1998).
- [124] M. P. Surh, H. Chacham, and S. G. Louie, *Phys. Rev. B* **51**, 7464 (1995).
- [125] M. L. Tiago and J. R. Chelikowsky, *Phys. Rev. B* **73**, 205334 (2006).
- [126] M. Hedström, A. Schindlmayr, and M. Scheffler, *phys. stat. sol. (b)* **234**, 346 (2002).
- [127] M. Hedström, A. Schindlmayr, G. Schwarz, and M. Scheffler, *Phys. Rev. Lett.* **97**, 226401 (2006).
- [128] J. Weber, A. Janotti, P. Rinke, and C. G. Van de Walle, *Appl. Phys. Lett.* **91**, 142101 (2007).
- [129] F. Aryasetiawan and O. Gunnarsson, *Rep. Prog. Phys.* **61**, 237 (1998).
- [130] I. D. White, R. W. Godby, M. M. Rieger, and R. J. Needs, *Phys. Rev. Lett.* **80**, 4265 (1997).
- [131] M. Rohlfing, N. P. Wang, P. Krüger, and J. Pollmann, *Phys. Rev. Lett.* **91**, 256802 (2003).
- [132] G. Fratesi, G. P. Brivio, P. Rinke, and R. Godby, *Phys. Rev. B* **68**, 195404 (2003).
- [133] G. Fratesi, G. P. Brivio, and L. G. Molinari, *Phys. Rev. B* **69**, 245113 (2004).
- [134] O. Pulci, L. Reining, G. Onida, R. D. Sole, and F. Bechstedt, *Comput. Mater. Sci.* **20**, 300 (2001).
- [135] P. Rinke, K. Delaney, P. García-González, and R. W. Godby, *Phys. Rev. A* **70**, 063201 (2004).