PHYSICAL REVIEW B **82**, 205127 (2010)

Electronic excitations from a perturbative LDA+*GdW* approach

Michael Rohlfing*
Fachbereich Physik, Universität Osnabrück, Barbarastraße 7, 49069 Osnabrück, Germany
(Received 17 September 2010; published 29 November 2010)

We discuss an efficient approach to excited electronic states within *ab initio* many-body perturbation theory (MBPT). Quasiparticle corrections to density-functional theory result from the difference between metallic and nonmetallic dielectric screening. They are evaluated as a small perturbation to the density-functional theory (within the local-density approximation) band structure rather than fully calculating the self-energy and evaluating its difference from the exchange-correlation potential. The dielectric screening is described by a model, which applies to bulk crystals, as well as, to systems of reduced dimension, like molecules, surfaces, interfaces, and more. The approach also describes electron-hole interaction. The resulting electronic and optical spectra are slightly less accurate but much faster to calculate than a full MBPT calculation. We discuss results for bulk silicon and argon, for the Si(111)-(2×1) surface, the SiH_4 molecule, an argon-aluminum interface, and liquid argon.

DOI: 10.1103/PhysRevB.82.205127 PACS number(s): 71.15.Qe, 71.20.−b, 71.35.−y, 73.20.−r

I. INTRODUCTION

Many-body perturbation theory (MBPT) has become the state of the art for excited states in electronic-structure theory.[1,2] Starting from a density-functional theory (DFT) calculation, the *GW* method[3] and its combination with the Bethe-Salpeter equation (BSE) (Refs. 1 and 2) allow to investigate the spectra of electrons, holes, and correlated electron-hole pairs. The great success of MBPT is based on the systematic incorporation of Coulomb interaction and polarization effects on all length scales, which is not considered in most other electronic-structure approaches. The significant computational cost of MBPT, however, still constitutes a major obstacle for the widespread use of the method. This holds in particular for larger-scale systems, such as defects, hybrid systems, adsorbates, nanostructures, and others. In this paper we propose a dramatic reduction in the computational requirements of MBPT. As a result, the excellent precision of standard *GW* and *GW*+BSE calculations is slightly reduced, but instead the treatment of much larger systems becomes possible, thus allowing the investigation of spectroscopic features that might be inaccessible otherwise.

As key ingredient we exploit the observation that for many systems MBPT, when carried out by (wrongly) assuming *metallic dielectric screening*, approximately reproduces the band structure of the underlying DFT calculation [when employing the local-density approximation (LDA)]. This had already been observed by Wang and Pickett,[4] as well as, by Gygi, Baldereschi, and Fiorentini[5,6] and was subsequently exploited for model quasiparticle (QP) calculations for various materials.[7,8] As illustration, Fig. 1 shows QP corrections for silicon (Si) and solid argon (Ar). The open circles (○) result from a conventional *GW* calculation (with standard RPA dielectric screening, "*GW*/RPA"), yielding the well-known opening of the band gap (by 0.7 eV for Si and 6.1 eV for Ar). The squares (□), on the other hand, come from *GW* calculations which employ *metallic* screening; these QP shifts are close to zero (at least for states near the Fermi level). The "metallic" dielectric screening is simulated by a dielectric model function.[9–13] Here we use a model based on that of Bechstedt, Enderlein, and Wischnewski,[9] slightly modified (see Sec. II B) for broader applicability. Such models are controlled by a few parameters, most importantly by the macroscopic dielectric constant, ϵ_∞. Setting $\epsilon_\infty \equiv \infty$ turns the screening into that of a metal.

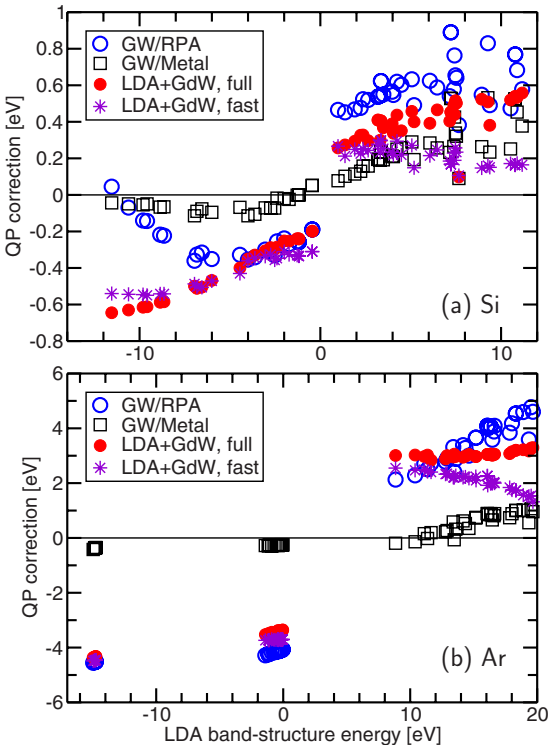

FIG. 1. (Color online) Quasiparticle corrections of (a) bulk Si and (b) bulk Ar. The open circles (○) denote standard *GW* data within RPA screening. The squares (□) result from a standard *GW* calculations, but based on a *metallic* dielectric model function (see text). The filled circles (●) result from the present LDA+*GdW* approach, employing the same basis and band-summation details as in the standard *GW* data. The asterisks (∗, fast) result from LDA+*GdW* with nine plane waves (15 plane waves for Ar) and eight bands, only (see text).

If the *GW* method with metallic screening, W_{metal}, reproduces the DFT-LDA band structure, one can arrive at the true QP band structure by adding a self-energy

$$\Delta\Sigma := iG_1(W - W_{metal}) = iG_1\Delta W \tag{1}$$

to the DFT-LDA Hamiltonian.[4–6] This procedure, which we label "LDA+*GdW*" throughout the paper (see below for details), yields the filled circles (●) in Fig. 1. These data agree fairly well with the *GW*/RPA data, at least in the important region near the fundamental gap; as a rule of thumb, we find that the gaps from $\Delta\Sigma := iG_1(W-W_{metal})$ agree with experiment to about 10%, which largely corrects the DFT-LDA band-gap error of 30–50 %. Most importantly, long-range polarization effects are included, which is completely missing from the short-sighted DFT-LDA. This systematic improvement due to $\Delta\Sigma$ is much more important than plain agreement of band-structure data with experiment.

While Eq. (1) describes single-particle states of electrons and holes, a straight-forward extension to coupled electron-hole pairs and their optical response is easily possible by calculating the electron-hole interaction (from *W*) and solving the Bethe-Salpeter equation, BSE.[1,2] Since the self-energy considered in this work is in principle still the one resulting from the *GW* approximation (GWA), the corresponding approximations to the electron-hole interaction are meaningful in the context of LDA+*GdW*, as well.

As an important consequence of the above findings, the calculation of *GW*-like band structures via $\Delta\Sigma$ allows for a tremendous gain in numerical efficiency in four respects: (i) the use of model dielectric functions, (ii) small basis-set requirement, (iii) small band-summation requirement, and (iv) weak influence of dynamical effects. The underlying reason for all four issues is that the current approach calculates QP corrections to DFT-LDA as a *perturbation*. Most other *GW* implementations simply replace V_{xc} by Σ^{GW}, both of which are in the order of magnitude of −10 eV or more. In order to get their difference to within 0.1 eV, the *GW* calculation must be carried out with numerical precision better than 1%. Our present approach, on the other hand, starts directly at the QP *correction* (i.e., $\Delta\Sigma$), which is much smaller in magnitude (~1 eV) and much more robust. Here it is fully sufficient to evaluate all quantities to within 10%, only, to achieve the same final numerical accuracy in the band structure.

The approach to be proposed in this paper is similar to the method by Gygi, Baldereschi, and Fiorentini,[5,6] who employed their perturbative *GW* method for the comprehensive analysis of bulk crystals. As a key difference to their approach, here we employ a different, more general model dielectric function which is flexible enough to also describe composite systems containing metals, nonmetals, molecules, surfaces, interfaces, and more. As illustration, we discuss in this paper bulk materials, a semiconductor surface, a molecule, a metal-insulator junction, and a disordered insulator. In all cases our approach yields spectroscopic data of high quality (although slightly less accurate than the corresponding full *GW* or *GW*/BSE calculation), demonstrating an appealing combination of predictive power, broad applicability, and numerical efficiency.

The paper is organized as follows. In Sec. II we discuss the computational approach and the dielectric model function employed in this work. In Sec. III characteristic results for bulk silicon and bulk argon are discussed. Section IV presents results for more complicated systems, indicating the potential of the method. The paper is concluded by a summary in Sec. V.

II. THEORETICAL APPROACH

In this section we discuss the computational method used in this work, including its practical realization and underlying physical principles.

A. Perturbative quasiparticle corrections

Ab initio QP band structures result from the electron self-energy operator $\Sigma(E)$. The state-of-the-art approach to Σ is given by Hedin's *GW* approximation,[3] which is usually evaluated and employed on top of an underlying DFT calculation. The typical procedure employs DFT data to generate the single-particle Green's function G_1 and the screened interaction *W* [usually within the random-phase approximation (RPA)]. Thereafter, the resulting self-energy operator $\Sigma=iG_1W$ replaces the DFT exchange-correlation potential, V_{xc}, arriving at a QP Hamiltonian of

$$\hat{H}^{QP} := \hat{H}^{DFT} + iG_1W - V_{xc}. \tag{2}$$

This procedure is commonly labeled MBPT. However, this does *not* mean that Eq. (2) would be evaluated truly perturbatively in the sense that the smallness of the difference (iG_1W-V_{xc}) would be exploited. Instead, both terms, iG_1W and V_{xc}, are evaluated separately, taking their difference afterwards. The QP corrections are thus obtained as (small) differences between two (rather large) quantities, both of which have to be evaluated independently and with high precision. Simply speaking, in order to get their difference (often ~1 eV) to within 0.1 eV (which is the accuracy expected from MBPT), both iG_1W and V_{xc} (being of the order of ~−10 eV or more) need to be evaluated with a precision of 1%. The underlying reason for this problem is the quite different conceptual origin of the two terms, which makes it difficult to formulate their difference in analytic terms.[14]

Fortunately, there does exist some pragmatic link between Σ and V_{xc}: the self-energy of the homogeneous electron gas (as a function of the energy of a given state) is nearly constant (see Ref. 3) and thus nearly coincides with V_{xc}. (Note that this might not be truly fulfilled by approximations to Σ, like the GWA, which might suffer from offsets.) This behavior is reflected by the observation that in bulk metals, QP corrections (from GWA) to DFT-LDA band structures are very small.[15,16] In other words, DFT (at least within the LDA) does contain correct spectral properties of the quasiparticles, at least for homogeneous systems. Within LDA, however, these spectral properties are by construction still those of the metallic system (jellium) from which the LDA exchange-correlation data originate, and this metallic behavior (in particular, metallic screening) is a built-in property of the exchange-correlation potential even when applied to non-

metallic systems. A generalization of this statement would imply that $V_{xc} \approx iG_1 W_{metal}$ (provided that $iG_1 W$ is a good approximation to Σ) with the appropriate W_{metal} (i.e., metallic screening). In fact, *GW* studies employing metallic screening (including ours, see Fig. 1) confirm that $V_{xc}|n\mathbf{k}\rangle \approx iG_1 W_{metal}|n\mathbf{k}\rangle$ for most electronic states $|n\mathbf{k}\rangle$.

Based on the working hypothesis that for nonhomogeneous, nonmetallic systems the largest difference to metallic behavior is the difference in *screening*, and employing $V_{xc} \approx iG_1 W_{metal}$, one arrives at the QP Hamiltonian

$$\hat{H}^{QP} \approx \hat{H}^{DFT-LDA} + iG_1(W - W_{metal}), \tag{3}$$

in which $\Delta\Sigma = iG_1(W - W_{metal})$ acts as a self-energy, yielding QP corrections [cf. Eq. (1)]. The most important change to Eq. (2) is the fact that Eq. (3) no longer evaluates the difference between the self-energies (given by $iG_1 W$ and V_{xc}), but the difference in screening: $(W - W_{metal})$. This difference is much simpler and faster to treat than the difference between self-energies. Note that on the other hand, the final accuracy of the QP band structure might be less than 0.1 eV because the entire approach, although being much more efficient from a numerical point of view, is based on the assumption that $V_{xc} \approx iG_1 W_{metal}$, meaning a further approximation in addition to the *GW* approximation. The assumption that $V_{xc} \approx iG_1 W_{metal}$ should be checked carefully for each system class. For the systems discussed in this paper (i.e., simple semiconducting matter and noble-gas systems), this assumption is fulfilled with sufficient accuracy (see Fig. 1 and Sec. IV). The assumption also holds for some organic molecules and carbon-based materials (not discussed here). Some oxides (like MgO) show deviations of up to 0.5 eV, mostly affecting the size of the fundamental gap. Many more systems will have to be tested to achieve a complete picture. There is not much experience so far on more strongly correlated materials, like e.g., *d*-electron systems.

As discussed below, the use of Eq. (3) allows for several numerical simplifications (see Sec. II D), leading to a higher efficiency than conventional *GW* calculations, allowing to tackle more complex systems. One of the most important facilitations is the use of model dielectric functions (see next section) instead of calculating the screening within the random-phase approximation.

B. Model dielectric function

The calculation of the dielectric function within the RPA, which is the common procedure within MBPT, is one of the bottlenecks of the method. A simplified evaluation of the dielectric function is an important contribution to improving the efficiency of MBPT (even without the considerations of the previous section). For this reason model dielectric functions are sometimes employed to avoid the RPA.[17–22] In the present context, the use of models is also mandatory for another reason: the key ingredient of the present theory is the difference between the correct screening of the (nonmetallic) system and its (hypothetical) metallic counterpart. This is only useful and well-defined if both types of screening result from the same approach, allowing to tune the screening from "correct" to metallic in a seamless manner. It is, however, unclear how the RPA could be used to simulate metallic behavior of a nonmetallic system. An appropriate model is therefore a necessity of the current approach.

Examples are the models proposed by Bechstedt, Enderlein, Wischnewski, and Falter and by Levine, Hybertsen, and Louie.[9–13] We have tested these models in the present context and find that they yield essentially the same results as the ones to be discussed below. In their original form, however, these models have one significant disadvantage which may hinder their application to more complex systems: they were formulated for systems that are characterized by one common dielectric constant without spatial variation. This makes it difficult to employ them for systems in which the screening shows spatial variation, like interfaces, molecules, etc.

Instead we propose a model that is based on a combination of localized and delocalized quantities. The system may consist of N atoms (at positions $\boldsymbol{\tau}_j$) in a (periodically repeated) unit cell or supercell of volume V, with reciprocal lattice vectors $\mathbf{G}$. To each atom we attribute a static charge-density response $\chi^{(j)}$ (see below) and an effective volume V_j. The dielectric function of the whole system is then obtained as

$$\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \delta_{\mathbf{G},\mathbf{G}'} + \frac{1}{|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|} \sum_{j=1}^{N} \frac{V_j}{V} \chi^{(j)}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}). \tag{4}$$

The volume attributed to each atom controls the weight which the atom contributes to the response. The transformation from the charge-density response to the dielectric function further involves a convolution with the Coulomb interaction, i.e., the multiplication by $1/(|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|)$ in Eq. (4). Note that we work with a symmetrized dielectric function.[23,24]

It was suggested by Bechstedt *et al.* to describe the charge-density response of a (homogeneous) system with dielectric constant ϵ_∞ by[9]

$$f(Q;\bar{\rho},\epsilon_\infty) = \left[\frac{1}{\epsilon_\infty - 1} + \frac{Q^2}{q_{TF}^2(\bar{\rho})} + \frac{Q^4}{\omega_P^2(\bar{\rho})} \right]^{-1}, \tag{5}$$

where the Thomas-Fermi wave number q_{TF} and plasma frequency ω_P depend on the average electron density $\bar{\rho}$. Equation (5) is related to the Lindhard dielectric function. In combination with Eq. (4) (for homogeneous systems, disregarding the summation over atoms), $\chi = Q^2 f$ would describe the dielectric function. In particular, for $Q \to 0$ one would correctly obtain $\epsilon(Q) \to \epsilon_\infty$ (if $\epsilon_\infty < \infty$, i.e., for nonmetals) or $\epsilon(Q) \to 1 + q_{TF}^2/Q^2$ (if $\epsilon_\infty = \infty$, i.e., for metals), respectively. The generalization to nonhomogeneous systems is less clear. While the large-Q behavior, which reflects the short-range reaction of electronic charge to external fields on the subatomic length scale, appears realistic for nonhomogeneous systems as well, the realistic incorporation of atomic-length-scale charge-density variation and of local fields [i.e., off-diagonal matrix elements of $\chi^{(j)}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$] is less clear.

Here we propose to model the charge-density response attributed to each atom by

$$\text{atom with metallic response: } \chi^{(j)}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \sqrt{f(|\mathbf{q}+\mathbf{G}|;\bar{\rho}_j,\infty)f(|\mathbf{q}+\mathbf{G}'|;\bar{\rho}_j,\infty)} \cdot |\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'| \cdot e^{-\gamma_j(\mathbf{G}'-\mathbf{G})^2} e^{i(\mathbf{G}'-\mathbf{G})\boldsymbol{\tau}_j}, \quad (6)$$

$$\text{atom with nonmetallic response: } \chi^{(j)}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \sqrt{f(|\mathbf{q}+\mathbf{G}|;\bar{\rho}_j,\epsilon_j)f(|\mathbf{q}+\mathbf{G}'|;\bar{\rho}_j,\epsilon_j)} \cdot (\mathbf{q}+\mathbf{G})(\mathbf{q}+\mathbf{G}') \cdot e^{-\gamma_j(\mathbf{G}'-\mathbf{G})^2} e^{i(\mathbf{G}'-\mathbf{G})\boldsymbol{\tau}_j}. \quad (7)$$

In both cases, the factor $\sqrt{f(|\mathbf{q}+\mathbf{G}|)f(|\mathbf{q}+\mathbf{G}'|)}$ is a reasonable average of the large-Q behavior in directions $(\mathbf{q}+\mathbf{G})$ and $(\mathbf{q}+\mathbf{G}')$. The phase factor for each atom results from the position of the atom within the unit cell or supercell. The factor $\exp[-\gamma_j(\mathbf{G}'-\mathbf{G})^2]$ describes the spatial extent of the charge density of atom j. Without this factor (or with $\gamma_j \to 0$), the model describes a sharp point-charge-density response at position $\boldsymbol{\tau}_j$. With $\gamma_j \to \infty$ all local fields would be switched off, turning the model into that of a homogeneous system again. For a nonzero, finite value of γ_j, the factor $\exp[i(\mathbf{G}'-\mathbf{G})\boldsymbol{\tau}_j]\exp[-\gamma_j(\mathbf{G}'-\mathbf{G})^2]$ is the Fourier transform of a Gaussian-shaped charge density $\sim\exp[-(\mathbf{r}-\boldsymbol{\tau}_j)^2/(4\gamma_j)]$ centered at $\boldsymbol{\tau}_j$. In short, this means that the charge-density response is neither perfectly local (i.e., exactly at $\boldsymbol{\tau}_j$) nor completely delocalized (except for a truly homogeneously system, to be characterized by $\gamma_j \to \infty$). Instead, the charge-density response of an atom originates from its charge density (or at least from that of the polarizable electronic states), and its spatial form is included in the model. Correspondingly, $2\sqrt{\gamma_j}$ approximates the radius of the atom. It should be noted that the term $\exp[-\gamma_j(\mathbf{G}'-\mathbf{G})^2]\cdot\exp[i(\mathbf{G}'-\mathbf{G})\boldsymbol{\tau}_j]$ corresponds to the factor $\rho(\mathbf{G}-\mathbf{G}')/\rho(0)$ (i.e., Fourier transform of the charge density) in the model by Bechstedt *et al.*[9] In our model the charge density of the entire system is approximated by a composition of atomic contributions with simplified shape.

A particular role is played by the factors $|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|$ (for metallic response) and $(\mathbf{q}+\mathbf{G})\cdot(\mathbf{q}+\mathbf{G}')$ (for nonmetallic response) in Eqs. (6) and (7). These factors reflect the qualitatively different origin of the response of metallic and nonmetallic systems. For metals, long-range charge fluctuations and displacements are possible, resulting from intraband transitions near the Fermi level. Such displacements lead to charge accumulation at some atoms and charge depletion at others. Here our model assumes that such charge accumulation or depletion would basically show the same spatial structure as the original charge density of the atom (modeled by $\exp[-(\mathbf{r}-\boldsymbol{\tau}_j)^2/(4\gamma_j)]$), i.e., $\delta\rho_j(\mathbf{r}) \sim \rho_j(\mathbf{r})$.

The charge-density response of a nonmetal, on the other hand, is of completely different origin. Here the response to an external field is mainly given by a short-range displacement of charge density from one side of the atom to the other, i.e., by a polarization of the atom. In many cases, this polarizability is dominated by transitions from s orbitals to p orbitals or vice versa. The spatial structure of such $s \leftrightarrow p$ polarizability is given by a factor $(\mathbf{r}-\boldsymbol{\tau}_j)\cdot(\mathbf{r}'-\boldsymbol{\tau}_j)$, leading to a factor of $(\mathbf{q}+\mathbf{G})\cdot(\mathbf{q}+\mathbf{G}')$ in reciprocal space. Again, the additional factor $\exp[-(\mathbf{r}-\boldsymbol{\tau}_j)^2/(4\gamma_j)]$ (or its reciprocal-space counterpart) reflects the fact that the response comes from the whole atom (including some spatial extent) rather than from a single point.

The model can also be generalized to the case of anisotropic response, e.g., if an atom is embedded in a nonisotropic chemical environment, like in molecules, at surfaces, in sp^2-bonded carbon, in atomic monolayers on a substrate, or similar. The same holds for materials with an anisotropic dielectric-constant tensor. In both cases, the response of each atom should be modeled with a direction-dependent dielectric-constant parameter $\boldsymbol{\epsilon}_j(\hat{\mathbf{q}})$. Since such a situation can be expressed in terms of the three principal axes $\mathbf{n}^{(k)}$ and corresponding principal values $\boldsymbol{\epsilon}^{(k)}$ of a 3×3 tensor (see next section for details), a straight-forward generalization of Eq. (7) is possible:

$$\text{atom with nonmetallic response: } \chi^{(j)}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \sum_{k=1}^{3} \sqrt{f(|\mathbf{q}+\mathbf{G}|;\bar{\rho}_j,\epsilon_j^{(k)})f(|\mathbf{q}+\mathbf{G}'|;\bar{\rho}_j,\epsilon_j^{(k)})} \cdot [(\mathbf{q}+\mathbf{G})\mathbf{n}_j^{(k)}] \cdot [\mathbf{n}_j^{(k)}(\mathbf{q}+\mathbf{G}')] \cdot e^{-\gamma_j(\mathbf{G}'-\mathbf{G})^2} e^{i(\mathbf{G}'-\mathbf{G})\boldsymbol{\tau}_j}. \quad (8)$$

One especially useful feature of the model proposed in this work is the possibility to combine metallic and nonmetallic response in one system. This is particularly relevant for adsorbates on metallic substrates, for metal-insulator interfaces, etc. Here our model simply allows to attribute metallic parameters [i.e., Eq. (6)] to some atoms and nonmetallic parameters [i.e., Eqs. (7) or Eq. (8)] to others. Note that the differentiation between atoms with isolated and metallic response might be difficult or impossible in some cases, e.g., if electronic states of a metal extend very far into an insulator.

In such cases our model might not be applicable.

For the construction of W_{metal}, finally, we simply take metallic response of all atoms [i.e., Eq. (6)]. Apparently, for metals (or metallic regions) the dielectric function is the same in both cases.

Note that the assumptions of the present model, like, e.g., the Gaussian-shaped atomic charge-density distribution or the homogeneous response of Eq. (5), might not be appropriate for all systems. In such situations more elaborate expressions can easily be implemented in the model.

C. Determination of the model parameters

The determination of the parameters is a particular task. Fortunately, the final use of the model for *differences* between metallic and nonmetallic screening makes the entire approach insensitive to the actual choice of the parameters V_j, $\bar{\rho}_j$, and γ_j. Within this work, we simply attribute a realistic volume to each atom (for silicon, e.g., we choose $V_j=20$ Å^3, which is the volume per atom of bulk Si), as well as a realistic valence electron number (for silicon, apparently $N_j=4$). The average electron density $\bar{\rho}_j=N_j/V_j$ defines the Thomas-Fermi wave number $q_{TF,j}$ and plasma frequency $\omega_{P,j}$ for this atom, to be used in $f(Q)$. The parameter γ_j is obtained from least-squares fitting of the atomic charge density by a Gaussian function. These parameters are used for both the metallic and the nonmetallic response of atom j.

For nonmetals one needs the dielectric-constant parameter ϵ_j (or the principal axes and values of the corresponding tensor for anisotropic situations). Such values can either be taken from experiment or they are calculated for the particular system. One possibility is given by the evaluation of the small-$\mathbf{q}$ limit of $\epsilon_{\mathbf{G}=0,\,\mathbf{G}'=0}(\mathbf{q})$ from the electrical-dipole operator applied to the interband transitions of the system, leading to a 3×3 tensor of

$$\epsilon_{ab}=\delta_{ab}+\sum_{v}^{occ}\sum_{c}^{empty}\sum_{\mathbf{k}}\frac{\langle v\mathbf{k}|\hat{p}_a|c\mathbf{k}\rangle\langle v\mathbf{k}|\hat{p}_b|c\mathbf{k}\rangle^*}{(E_{c\mathbf{k}}-E_{v\mathbf{k}})^3} \tag{9}$$

$(a,b=x,y,z)$ from which the principal axes and values can be evaluated. Note that Eq. (9) does not contain local-field effects. However, since the resulting ϵ_j enter our model *before* the inversion of $\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ (which then leads to the local-field effects), the employment of local-field-free parameters is not a problem but rather a requirement of the model.

In many systems the responses of the various atoms will differ from each other, leading to the question of distributing the results of Eq. (9) over the individual atoms. Here we propose to employ an atom-centered local-orbital basis for the calculation of the electronic states, $|n\mathbf{k}\rangle$. Such a basis allows to decompose the dipole matrix elements $\langle v\mathbf{k}|\hat{p}_a|c\mathbf{k}\rangle$ into individual contributions of each atom j, i.e., one can focus on atom j and switch off the dipole strength of all other atoms. In this case Eq. (9) yields an individual result for each atom alone, allowing to find out the individual parameters of each atom in the system.

D. Numerical efficiency

As an important consequence of the above findings, the calculation of *GW*-like band structures via $\Delta\Sigma$ allows for a tremendous gain in numerical efficiency in four respects: (i) the use of model dielectric functions, (ii) small basis-set requirement, (iii) small band-summation requirement, and (iv) weak influence of dynamical effects. As mentioned, the underlying reason for all four issues is that the current approach calculates QP corrections to DFT-LDA as a *perturbation*. The four issues of efficiency (i)–(iv) deserve detailed discussion.

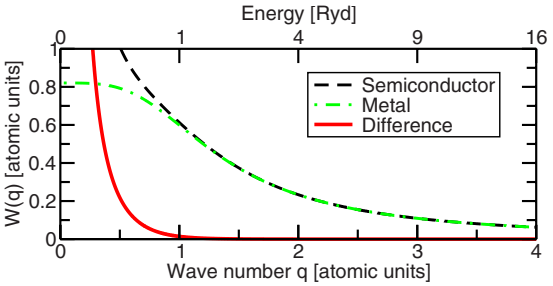

FIG. 2. (Color online) Reciprocal-space behavior of the screened Coulomb interaction $W(q)=\epsilon^{-1}(q)/q^2$ as resulting from Eq. (5) for a homogeneous semiconductor (with silicon parameters), for the corresponding metal, and their difference. At small q, $W(q)\to 1/(\epsilon_\infty\cdot q^2)$ and $W_{\text{metal}}(q)\to 1/q_{TF}^2$. At large q, both $W(q)$ and $W_{metal}(q)$ behave like $1/q^2$.

(i) The advantage of working with a dielectric model function rather than employing the random-phase approximation has already been pointed out in the last section. In particular, the perturbative idea of Eqs. (1) and (3) *requires* a model even beyond the issue of numerical efficiency.

(ii) The basis-set requirements for $(W-W_{metal})$ are much weaker than for W alone for two reasons: (1) Within full *GW* the bare-exchange contribution requires a large basis for convergence. In our present approach, on the other hand, the bare-exchange effects are the same in *GW* and GW_{metal} and thus cancel each other. (2) While both W and W_{metal} are structured in real space, their *difference* is a rather smooth function and converges with very few basis functions. This is illustrated by Fig. 2 which shows (neglecting local-field effects) the screened Coulomb interaction in reciprocal space, i.e., $W(q)=\epsilon^{-1}(q)/q^2$ from Eq. (5), for a homogeneous semiconductor (with silicon parameters), for the corresponding metal, and their difference. It is evident that the difference $W-W_{metal}$ is close to zero for wave numbers (or corresponding energies) beyond 1 Ry (in the case of Si), and convergence of MBPT is achieved much faster than from W or W_{metal} alone. Figure 3 a shows representative gap energies of Si as a function of the plane-wave basis size used for W. Full *GW* requires about 60 plane waves (~5 Ry cutoff) for reasonable accuracy, while the LDA+*GdW* data are already converged with 9 plane waves (~1.4 Ry). Similarly, basis-set convergence for Ar requires cutoff energies of about 15 Ry for GWA, but only about 2 Ry for LDA+*GdW*.

(iii) The band summation in $\Delta\Sigma$ is less demanding than in a full *GW* calculation because the influence of the higher conduction bands (via G_1) is weak. This is shown in Fig. 3(b). The use of about as many conduction bands as valence bands is sufficient for LDA+*GdW* (at least for states near the gap), while conventional *GW* calculations are usually performed with at least about 10 times more conduction than

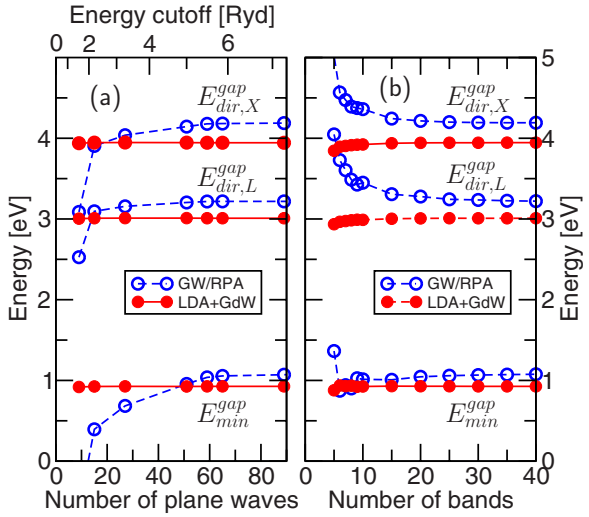

FIG. 3. (Color online) Gap energies of Si (indirect minimum gap, direct gap at L, and direct gap at X), calculated within a full GW calculation (employing RPA screening) and within the present LDA+GdW scheme. (a) Dependence of the gap energies on the plane-wave basis representation of W [or $(W-W_{metal})$, respectively], as controlled by the energy cutoff (upper axis). (b) Dependence of the gap energies on the number of bands considered in the band summation inside the self-energy operator. (Note that this does not refer to the number of bands considered in the calculation of the RPA screening.)

valence bands. This behavior again results from the smooth spatial structure of $(W-W_{metal})$. To summarize statements (ii) and (iii), Fig. 1 includes data from 1.4 Ry cutoff (2.0 Ry for Ar) and four conduction bands in G_1 as asterisks (∗, "fast"). The agreement with the converged LDA+GdW data shown by (●) is sufficient, except for higher-energy states.

(iv) Within conventional GW calculations, the correlation part of Σ (which is subject to dynamical effects) can be as large as 5–10 eV. Our $\Delta\Sigma$, on the other hand, is much smaller in magnitude and thus much less sensitive to dynamical effects. This allows us to treat $\Delta\Sigma$ on the level of the static Coulomb-hole plus screened-exchange (COHSEX) approximation,[25] which we employ in all LDA+GdW calculations in this paper. Note that this does not apply to the GW/RPA and GW/Metal reference calculations in this paper, all of which include dynamical effects by using a plasmon-pole approximation. In those cases, the generalization of the static model of Sec. II B to a dynamic dielectric function is realized by evaluating the f-sum rule.[26]

Owing to the above-mentioned simplifications and approximations, an enormous speed up of up to two orders of magnitude can be achieved, depending on the nature of the particular system. As a rule of thumb, the avoidance of the RPA (which takes about half of the computational effort of conventional MBPT) already means about a factor of two in terms of efficiency. Furthermore, the present approach requires less plane waves (by about a factor of 10, see above) and less bands (by about a factor of 5, see above) for convergence. The calculation of the three-center integrals $\int \psi^*_{m\mathbf{k}}(\mathbf{r})\exp[-i(\mathbf{q}+\mathbf{G})\mathbf{r}]\psi_{n,\mathbf{k}+\mathbf{q}}(\mathbf{r})d^3r$, which usually consumes more than 90% of the time of conventional MBPT, is now faster by a factor of 10–50. Test calculations of the band structure of, e.g., carbon nanotubes of 64 atoms take about 50 CPU h, only. The scaling with system size is the same as in conventional MBPT, i.e., $\sim N^4$ for most problems (with N being the number of atoms). Combined with parallelization, calculations for several hundred atoms should be easily feasible.

III. RESULTS FOR BULK SILICON AND ARGON

The QP corrections to DFT-LDA for bulk Si and Ar ar compiled in Fig. 1. The LDA+GdW calculations for Si were performed for the experimental lattice constant (5.43 Å), with ϵ_j=12.0, 6 special $\mathbf{q}$ points (136 in the full Brillouin zone), 8 bands for the band summation, and 9 $\mathbf{G}$ vectors (from a cutoff of 1.4 Ry). The LDA+GdW calculations for Ar were performed for the experimental lattice constant (5.26 Å), with ϵ_j=1.6, 2 special $\mathbf{q}$ points (32 in the full Brillouin zone), 8 bands for the band summation, and 15 $\mathbf{G}$ vectors (from a cutoff of 2.0 Ry). As discussed above, full GW calculations using a metallic W yield QP corrections close to zero, opening the possibility of perturbative LDA+GdW as proposed in Sec. II. In fact, the LDA+GdW data are close to those of a full GW calculation employing correct, nonmetallic screening from RPA (open circles). There are, however, some deviations (related to the LDA+GdW method as such, and also to the dielectric model function). For Si, for example, the lowest valence bands observes very small QP corrections within GW/RPA, but significant negative QP corrections within LDA+GdW. This was already observed by Gygi and Baldereschi[5] and might result from the neglect of dynamical effects within the present COHSEX approximation. The slight underestimation of the fundamental gap of Si has also been observed for other dielectric model functions in conventional GW calculations.[27] Furthermore, the QP corrections for the conduction bands appear to be less accurately reproduced by LDA+GdW than for the upper valence bands. Additional deviations are observed for the fast LDA+GdW approach (at minimal basis-set and band-summation specification), in particular for the higher conduction bands. These details notwithstanding, we can conclude that LDA+GdW yields sufficient accuracy if one is interested in states near the fundamental gap. Furthermore, systematic deviations between LDA+GdW and GW/RPA can be expected to be similar in bulk systems and other, more complicated systems of the same material (like, e.g., surfaces, nanostructured systems, interfaces, etc.). LDA+GdW will allow for systematic comparison between the spectral data of such systems.

The quite reliable LDA+GdW band structures can be employed to yield reasonable optical spectra, as well. To this end we include electron-hole interaction on the level of the Bethe-Salpeter equation.[2] The interaction kernel is calculated with the same (nonmetallic) dielectric model function and same basis as the band structure. One exception is the unscreened exchange interaction between electrons and holes (originating from the Hartree potential) which may require a larger energy cutoff than the screened interaction and must be treated separately. Its calculation is relatively cheap and

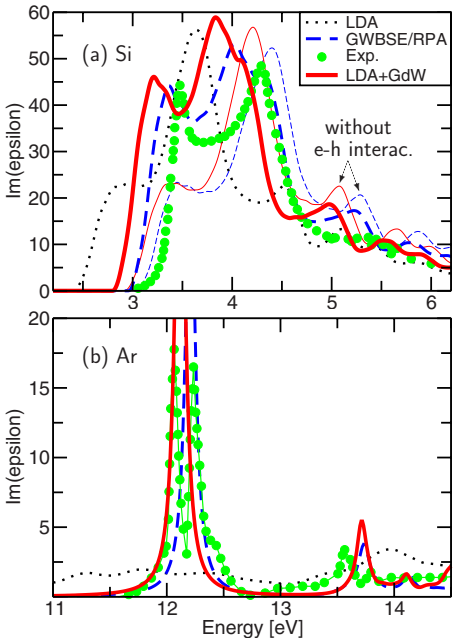

FIG. 4. (Color online) Optical spectrum (imaginary part of macroscopic dielectric function) of bulk Si and of bulk Ar. In the case of Si, the thin blue dashed (thin red solid) line shows the QP-corrected GW-RPA (LDA+*GdW*) free-particle interband spectrum without electron-hole interaction. Experimental data are from Refs. 28–30 (Si) and from Ref. 31 (Ar). Note that spin-orbit coupling (leading to the measured double-peak structure of the Ar exciton) is not included in our calculations.

does not affect the efficiency of our approach.

Figure 4 shows the macroscopic imaginary dielectric function $\epsilon_2(\omega)$ of bulk Si and Ar (resulting from three valence and four conduction bands at 500 **k** points on an asymmetric mesh). The dotted curve displays the spectrum from the uncorrelated LDA interband transitions, which is usually qualitatively and quantitatively wrong (in particular for insulators, like Ar). The dashed lines are reference data from a full *GW*+BSE/RPA calculation, which can be considered as the state-of-the-art approach to $\epsilon_2(\omega)$. The solid lines display our current results, including the drastic numerical simplification (fast) as outlined above. In comparison with the *GW*+BSE/RPA data and with experiment, the LDA+*GdW* data are very gratifying. They correctly yield the two characteristic peaks (at 3–3.5 eV and at 4–4.5 eV) of the Si spectrum. For completeness, the thin blue dashed line and the thin red solid line show the free-particle interband spectrum without electron-hole interaction for Si, as resulting from the full *GW* calculation (within RPA) and from our LDA+*GdW* band-structure calculation. The comparison to the electron-hole correlated BSE spectra shows that the influence of excitonic effects is the same in both theories (GWBSE/RPA and LDA+*GdW*), and the differences between the two BSE spectra are due to differences in the underlying QP band structures (cf. Fig. 1). For Ar, we obtain an exciton peak at 12.2 eV (*GW*+BSE/RPA) or 12.1 eV (LDA+*GdW*), respectively. In experiment, the spin-orbit interaction (neglected in our present work) splits the exciton into two peaks at 12.0 and 12.2 eV. Furthermore, a second excitonic peak is found near 13.5 eV. In the case of Ar, the differences between the two BSE spectra are very small. Here the excitonic binding energy is slightly larger in LDA+*GdW* (2.3 eV) than in GWBSE/RPA (2.0 eV), leading to a slightly lower exciton energy in LDA+*GdW* (12.1 eV) compared to GWBSE/RPA (12.2 eV). In the independent-particle QP band structure, the fundamental gap is slightly larger in LDA+*GdW* (14.4 eV) than in GW/RPA (14.2 eV). Note that this close agreement of LDA+*GdW* with the GWBSE/RPA reference data is better than can be expected from the method.

Most importantly, the differences between LDA+*GdW* and *GW*+BSE/RPA are not significantly larger than the deviations from experiment, thus advertising LDA+*GdW* as a useful alternative. Compared to the LDA interband spectrum, a tremendous improvement of explanatory power is achieved.

IV. RESULTS FOR MORE COMPLEX SYSTEMS

We have tested the LDA+*GdW* approach for a number of inhomogeneous systems, starting from the bulk materials (Si and Ar) discussed above.

A. Si(111)-(2×1) surface and silane molecule

Based on the experience with bulk silicon, we investigate two prototypical systems of silicon in reduced dimensions, i.e., the SiH_4 molecule and the Pandey-chain terminated Si(111)-(2×1) surface. Both systems have been intensively studied in theory and experiment (see, e.g., Refs. 35 and 36 and references therein). Here we focus on their electronic structure within the present LDA+*GdW* approach.

In the case of the Si(111)-(2×1) we focus on the band structure of the Pandey-chain derived dangling-bond states.[33,34,38–43] The dangling bonds result from the lower coordination (threefold instead of fourfold) of the Pandey-chain atoms, leading to one occupied and one empty state within the bulk band gap.[35] These two bands constitute one of the most intensively studied surface electronic structures. At the *J* point of the surface Brillouin zone the two bands are well separated from the silicon bulk states and define the surface band gap. Within LDA, this gap suffers from the same type of band-gap underestimation as all semiconductor systems. Here we observe a value of 0.4 eV, much smaller than the experimental result of 0.7 eV from a combination of direct and inverse photoemission (see Table I).[33,34]

Within *GW*/RPA, the surface bands are significantly shifted and result in very good agreement with the measured data.[33–35,41,42] It is most gratifying to see that this behavior is also given by the present LDA+*GdW* approach, which yields a surface gap energy of 0.8 eV. This good agreement also holds for the absolute energetic position (with respect to the bulk band structure). Both for the occupied and for the empty band, the data from *GW*/RPA, LDA+*GdW*, and experiment all agree to within 0.1 eV.

The screening properties for this calculation have been obtained from the approach as outlined in Sec. II C, yielding individual screening properties for each atom. Here we find

TABLE I. Characteristic band-structure data for the Si(111)-(2 $\times$1) surface, which is dominated by two dangling-bond states derived from the Pandey-chain termination. At the J point of the surface Brillouin zone the related bands (occupied D_{up} and unoccupied D_{down} state) are closest to each other and define the surface band gap (Ref. 35).

[eV]	LDA	*GW*/Metal	*GW*/RPA	LDA+*GdW*	Expt.
Bulk E_{min}^{gap}	0.49	0.51	1.10	0.95	1.17[a]
$D_{up}(J)$	0.0	0.0	0.0	0.0	0.0[b]
$D_{down}(J)$	0.4	0.4	0.7	0.8	0.7[c]

[a]Reference 32.
[b]Reference 33.
[c]Reference 34.

that the charge-density response of the surface atoms is slightly larger than that of the bulklike atoms in the center of the slab. The response of the bulklike atoms agrees with that of a true bulk calculation to within 10%. At the surface, on the other hand, the smaller surface band gap, the π-conjugated nature of the Pandey chain, and the vicinity of the vacuum lead to an anisotropic response. Perpendicular to the surface, the response is reduced by about 25% (leading to a dielectric-constant parameter of about $\epsilon_j^{(\perp)}=9$ instead of the bulk value of $\epsilon=12$). Parallel to the Pandey chain, on the other hand, the response is doubled to $\epsilon_j^{(\parallel)}=24$.

The calculations were carried out for the structural model developed in Ref. 35, with a Pandey-chain buckling of 0.51 Å. Reciprocal-space summation used 24 **q** points from the surface Brillouin zone. The LDA+*GdW* approach used 66 bands for summation (i.e., twice the number of occupied bands in the present eight-layer hydrogen-terminated slab of 12 Å thickness) and 380 **G** vectors from a cutoff energy of 1.5 Ry.

As another, even more extreme case for silicon in reduced dimension, we discuss the silane molecule (SiH_4).[37,36,44] Its electronic structure is dominated by quantum confinement, leading to much larger band gaps and QP corrections than for extended semiconductors. All relevant data are compiled in Table II. Compared to the LDA data, the occupied states [i.e., the lowest occupied molecular orbital (LOMO) and (threefold degenerate) highest occupied molecular orbital (HOMO)], are lowered in energy by more than 4 eV. Here the current LDA+*GdW* approach reproduces these QP shifts to within about 0.5 eV. The lowest unoccupied molecular orbital (LUMO), on the other hand, is shifted to higher energies by 1.0 eV (*GW*/RPA) or 1.2 eV (LDA+*GdW*), respectively. Based on these reliable data for single-particle states, LDA+*GdW* also yields reasonable data for charge-neutral electron-hole excitations (see Table II). Here we take the lowest-energy singlet and triplet excitation as representative examples. While *GW*+BSE within RPA yields data in excellent agreement with experiment[37] and with other many-body and quantum-chemical methods,[36] the data from LDA+*GdW* show slightly lower excitation energies. The deviations are in the order of 0.5 eV and correspond to the differences in the band-structure energy of the HOMO state, for which LDA+*GdW* yields a slightly too high value. Nevertheless, in light of the huge QP corrections and very strong electron-hole interaction of about 5 eV in SiH_4, we consider the accuracy of LDA+*GdW* (i.e., yielding QP shifts and electron-hole binding to within 20%) extremely gratifying.

TABLE II. Spectral data of the SiH_4 molecule, which is dominated by quantum confinement and shows the typical electronic excitations of a small molecule.

[eV]	LDA	*GW*/Metal	*GW*/BSE/RPA	LDA+*GdW*	Exp. [37]
E_{LOMO}	−13.5	−14.0	−17.8	−17.3	
E_{HOMO}	−8.4	−9.0	−12.5	−11.8	−12.6
E_{LUMO}	−0.6	−0.2	0.4	0.6	
$\Omega_{triplet}$			8.0	7.6	
$\Omega_{singlet}$			9.0	8.3	8.8

Similar to the case of the Si(111)-(2$\times$1) surface, screening in SiH_4 differs significantly from that of bulk silicon. The much larger gap reduces the charge-density response strongly. Our approach of locally evaluating the density-response contribution of each atom yields an isotropic response of the silicon atom to be described by $\epsilon_j=3.75$ (and similar results for the H atoms), i.e., weaker than bulk Si by a factor of 4. Such strong reduction for chemically passivated silicon in confined geometries was already found earlier.[45]

We close this section by mentioning that for both systems, Si(111)-(2$\times$1) and SiH_4, the underlying reason for the success of the LDA+*GdW* approach is again given by the reproduction of the DFT-LDA band-structure data when metallic screening is employed in a full *GW* calculation. The corresponding data are included in Tables I and II. In particular for the Si(111)-(2$\times$1) surface, this mandatory condition for the applicability of LDA+*GdW* is nearly exactly fulfilled. For the SiH_4 molecule some difference of the order of 0.5 eV are found. Considering the massive deviation of this system from a homogeneous metal, even this agreement to within 0.5 eV is an amazing result.

The calculations were done in a cubic cell of 20 Å cell size, for a tetrahedral SiH_4 molecule with a Si-H bond length of 1.50 Å. Eight **q** points were use for **q**-point summation. A cut-off energy of 1 Ry yields 925 **G** vectors for the LDA+*GdW*.

B. Argon systems

Spatially varying dielectric response is also present in metal-insulator heterostructures. As an example Fig. 5 shows the single-particle spectrum of a periodic heterostructure composed of five atomic layers (10 Å) of aluminum and one atomic layer (3 Å) of argon, stacked along the Al(001) direction. This system combines metallic screening in Al with insulating behavior in Ar, which has significant consequences on the QP energetics.[46–49] Here we focus on the local density of states (LDOS) in the Ar monolayer. The LDOS between −12 and −6 eV results from the upper valence states of Ar (3p), while the LDOS above 2 eV comes

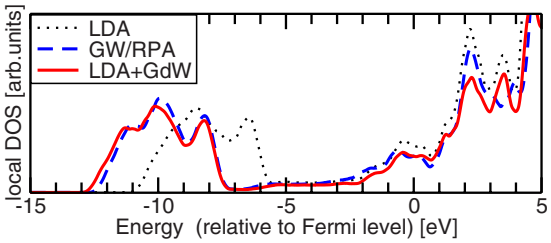

FIG. 5. (Color online) Local density of states in a monolayer of Ar, combined with a five-layer aluminum (001) slab in a heterostructure.

from the Ar conduction bands, with increasing admixture of Al states at higher energy. The LDOS inside the Ar band gap (between −6 and +2 eV) results from spill-out of Al states into the Ar layer. The most interesting feature is the rather small QP correction of the argon states, which (in *GW*/RPA) amounts to −1.7 eV(+0.2 eV) for the upper valence (lower conduction) states, yielding a total correction of 1.9 eV for the fundamental gap of Ar. In bulk Ar, on the other hand, the gap-edge states observe QP shifts of −4.1 and +2.0 eV, yielding a gap correction of 6.1 eV (cf. Fig. 1). The presence of metallic screening in the immediate neighborhood significantly weakens the QP shifts due to image-state effects,[46–49] both for holes and for electrons (by about 2 eV each). It is most gratifying to see that in our LDA+*GdW* approach (again with a plane-wave cutoff of 2 Ry), the spatial set-up of the dielectric model function [cf. Eq. (4)] reproduces these effects. Here LDA+*GdW* yields QP shifts of −1.8 and +0.2 eV for the band-edge states and a gap correction of +2.0 eV, compared to the Ar bulk data (see Fig. 1) of −3.5, +2.8, and +6.3 eV. We conclude that the LDA+*GdW* approach is a suitable method for addressing electronic properties of metal-nonmetal junctions.

The calculations were done for the experimental lattice constant of aluminum (4.05 Å), although this is in principle too small for the argon layer. The equilibrium of the argon layer between the adjacent Al layers is found at a layer-to-layer distance of 2.3 Å. All GWA and LDA+*GdW* calculations were performed with summations over 64 **q** points. The cutoff energy for LDA+*GdW* was 2.0 Ry, yielding 39 **G** vectors. The dielectric parameters were $\epsilon_j=1.6$ for the Ar atom (see Sec. III) and $\epsilon_j=\infty$ for the Al atoms.

As a last example for the potential of our method, Fig. 6 shows the exciton spectrum of noncrystalline argon. At zero temperature argon forms a periodic face-centered-cubic (fcc) lattice, which can easily be treated by MBPT (see Sec. II), leading to the results as discussed in Sec. III. For this periodic solid the exciton yields a sharp line (except for dynamical broadening effects from self-trapping or similar that are completely neglected here). This changes in the case of nonperiodic argon, like in its liquid or amorphous state. Such systems may be described by sufficiently large supercells. At present we investigate the spectra resulting from a 64 atom cell (consisting of $4\times4\times4$ fcc unit cells) and exploit its spectral features from the Γ point of the supercell, only. For the periodic fcc crystal this yields an exciton at 12.01 eV excitation energy (slightly lower than the value reported in Sec. III, which was obtained from the standard fcc unit cell

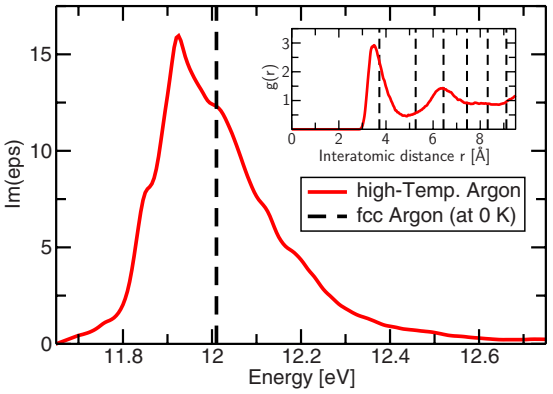

FIG. 6. (Color online) Optical spectrum of the excitons in noncrystalline argon (from molecular dynamics of 64 atoms at 300 K, at the solid-argon density, see text), resulting from the current LDA+*GdW* approach. The dashed line indicates the position of the exciton in the periodic crystal. The inset shows the Ar-Ar pair-correlation function of the MD simulation.

containing one atom, and 500 **k** points in the BSE). Within this configuration (which is computationally much more demanding than a simple one-atom-fcc calculation and is extremely demanding for the standard *GW*+BSE Hamiltonian) the spectrum of liquids or amorphous systems can be evaluated. At present we simply consider argon at its solid-state density (for comparison sake), but heated to 300 K (although this is an unrealistic high temperature for argon at this density). We perform a constant-temperature molecular-dynamics simulation (using a simple Lennard-Jones interatomic potential), leading to the Ar-Ar pair-correlation function shown in the inset of Fig. 6. Such a simulation is certainly not fully realistic in terms of describing liquid or amorphous systems; nonetheless it yields structural elements that may very well be present in liquids. The pair-correlation function clearly exhibits structures beyond harmonic vibrations (like, e.g., the vanishing of the second-nearest-neighbor peak at 5.3 Å), thus prohibiting a perturbative electron-phonon interaction treatment in the evaluation of the spectrum. Instead, our LDA+*GdW* approach (averaged over 20 snapshots of the MD run) easily allows to evaluate the spectrum. The data shown in Fig. 6 clearly demonstrate three important features: (i) the exciton line is significantly broadened, (ii) the broadening is asymmetric, leading to substantial nonzero amplitude well above the exciton energy, and (iii) the maximum of the peak is at lower energy than in the periodic system. The third feature is related to the fact that in the pair-correlation function, the first maximum also occurs at smaller distance (3.4 Å) than the fcc nearest-neighbor-distance (3.7 Å), which is a consequence of the anharmonicity of the Ar-Ar interatomic potential.

The computational parameters correspond to those of Sec. III (in particular, $E_{\mathrm{cut}}=2.0$ Ry yields 459 plane waves), except for a slightly lower band summation with 384 bands, only (i.e., 6 bands/atom).

V. SUMMARY

In summary, we have discussed an extremely efficient modification of standard many-body perturbation theory

(*GW* method plus Bethe-Salpeter equation). Based on the observation that metallic screening in the *GW* method approximately reproduces the DFT-LDA band structure (which should be checked for each material), quasiparticle (QP) corrections to DFT-LDA are obtained in a truly perturbative approach at minimal cost, provided that the dielectric screening can be described by an appropriate model. The resulting QP band structures and optical spectra (including electron-hole interaction) are slightly less accurate than those from conventional *GW*+BSE, but they include all Coulomb-interaction effects (such as screening, electron-hole binding, etc.) in a physically correct way, allowing to systematically investigate excitations beyond DFT and beyond the computational limits of conventional MBPT.

*michael.rohlfing@uos.de

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[2] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[3] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1970).
[4] C. S. Wang and W. E. Pickett, Phys. Rev. Lett. **51**, 597 (1983).
[5] F. Gygi and A. Baldereschi, Phys. Rev. Lett. **62**, 2160 (1989).
[6] V. Fiorentini and A. Baldereschi, Phys. Rev. B **51**, 17196 (1995).
[7] S. Massidda, A. Continenza, M. Posternak, and A. Baldereschi, Phys. Rev. Lett. **74**, 2323 (1995); Phys. Rev. B **55**, 13494 (1997).
[8] A. Continenza, S. Massidda, and M. Posternak, Phys. Rev. B **60**, 15699 (1999).
[9] F. Bechstedt, R. Enderlein, and R. Wischnewski, Phys. Status Solidi B **107**, 637 (1981).
[10] Z. H. Levine and S. G. Louie, Phys. Rev. B **25**, 6310 (1982).
[11] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **37**, 2733 (1988).
[12] C. Falter, M. Klenner, and W. Ludwig, Phys. Status Solidi B **167**, 85 (1991).
[13] G. Cappellini, R. Del Sole, L. Reining, and F. Bechstedt, Phys. Rev. B **47**, 9892 (1993).
[14] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).
[15] J. E. Northrup, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987); Phys. Rev. B **39**, 8198 (1989).
[16] G. D. Mahan and B. E. Sernelius, Phys. Rev. Lett. **62**, 2718 (1989).
[17] J. E. Northrup, Phys. Rev. B **47**, 10032 (1993).
[18] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998).
[19] W. G. Schmidt, J. L. Fattebert, J. Bernholc, and F. Bechstedt, Surf. Rev. Lett. **6**, 1159 (1999).
[20] L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999).
[21] P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. Lett. **88**, 016402 (2001).
[22] J. Furthmüller, G. Cappellini, H.-Ch. Weissker, and F. Bechstedt, Phys. Rev. B **66**, 045110 (2002).
[23] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **52**, 1905 (1995).
[24] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **54**, 13759 (1996).
[25] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[26] D. L. Johnson, Phys. Rev. B **9**, 4475 (1974).
[27] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993).
[28] H. R. Philipp, J. Appl. Phys. **43**, 2835 (1972).
[29] D. E. Aspnes and J. B. Theeten, J. Electrochem. Soc. **127**, 1359 (1980).
[30] P. Lautenschlager, M. Garriga, L. Viña, and M. Cardona, Phys. Rev. B **36**, 4821 (1987).
[31] V. Saile, M. Skibowski, W. Steinmann, P. Gürtler, E. E. Koch, and A. Kozevnikov, Phys. Rev. Lett. **37**, 305 (1976); F. Sottile, M. Marsili, V. Olevano, and L. Reining, Phys. Rev. B **76**, 161103 (2007).
[32] *Semiconductors. Physics of Group IV Elements and III-V Compounds*, edited by K.-H. Hellwege and O. Madelung, Landolt-Börnstein, New Series, Group III, Vol. 17, Pt. A (Springer-Verlag, Berlin, 1982), p. 208.
[33] R. I. G. Uhrberg, G. V. Hansson, J. M. Nicholls, and S. A. Flodström, Phys. Rev. Lett. **48**, 1032 (1982).
[34] P. Perfetti, J. M. Nicholls, and B. Reihl, Phys. Rev. B **36**, 6160 (1987).
[35] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **83**, 856 (1999).
[36] J. C. Grossman, M. Rohlfing, L. Mitas, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **86**, 472 (2001).
[37] U. Itoh, Y. Toyoshima, and H. Onuki, J. Chem. Phys. **85**, 4867 (1986).
[38] F. Ciccacci, S. Selci, G. Chiarotti, and P. Chiaradia, Phys. Rev. Lett. **56**, 2411 (1986).
[39] G. Chiarotti, S. Nannarone, R. Pastore, and P. Chiaradia, Phys. Rev. B **4**, 3398 (1971).
[40] P. Chiaradia, A. Cricenti, S. Selci, and G. Chiarotti, Phys. Rev. Lett. **52**, 1145 (1984).
[41] L. Reining and R. Del Sole, Phys. Rev. Lett. **67**, 3816 (1991).
[42] J. E. Northrup, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **66**, 500 (1991).
[43] K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982).
[44] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).
[45] L.-W. Wang and A. Zunger, Phys. Rev. Lett. **73**, 1039 (1994).
[46] J. P. A. Charlesworth, R. W. Godby, and R. J. Needs, Phys. Rev. Lett. **70**, 1685 (1993).
[47] N.-P. Wang, M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. Lett. **92**, 216805 (2004).
[48] J. B. Neaton, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **97**, 216405 (2006).
[49] K. S. Thygesen and A. Rubio, Phys. Rev. Lett. **102**, 046802 (2009).