PHYSICAL REVIEW B **78**, 121201(R) (2008)

Dielectric properties and excitons for extended systems from hybrid functionals

Joachim Paier, Martijn Marsman, and Georg Kresse
Faculty of Physics, Universität Wien and Center for Computational Materials Science, Sensengasse 8/12, A-1090 Wien, Austria
(Received 5 March 2008; revised manuscript received 1 August 2008; published 15 September 2008)

We have calculated static and dynamic response properties for several semiconducting and insulating solids using hybrid functionals, which admix a small fraction of nonlocal Fock exchange to an otherwise semilocal density functional. The calculated static and dynamic properties are clearly improved compared to conventional semilocal density functionals; in particular the oscillator strength at low energy excitations is well described.

DOI: 10.1103/PhysRevB.78.121201 PACS number(s): 71.10.−w, 71.15.Qe, 71.35.−y, 78.20.Bh

The accurate calculation of optical-absorption spectra is a long-standing challenge to computational physics. Although the *GW* approximation followed by a subsequent solution of the Bethe-Salpeter equation (BSE) is now routine, only very few absorption spectra have been published to date, one reason being that the approach is rather involved since it requires the diagonalization of a large two-particle matrix.[1] Combined with the slow k-point convergence of optical spectra, calculations are time consuming even by today's standards. Kohn-Sham density-functional theory (KS-DFT) on the other hand—which in its time-dependent formulation[2] is, at least in principle, applicable to the calculation of absorption spectra—is well known for its underestimation of band gaps. As a result the absorption spectra calculated using standard local or semilocal functionals are unsatisfactory, significantly redshifted, and lack typical excitonic features. This is related to the absence of an attractive electron-hole term in the present semilocal exchange-correlation functionals. Remedies for some aspects of these problems have been suggested, among them orbital-dependent density functionals,[3] such as the exact-exchange optimized effective potential method.[4] An approach that has recently attracted considerable interest is hybrid functionals mixing typically 25% of the nonlocal Hartree-Fock exchange term to the otherwise semilocal functional.[5–7] Although very successful for molecules, their application to extended systems has been rather limited until the recent suggestion of Heyd *et al.*[8] to use a screened exchange term. This allows for a more efficient treatment of large molecules and periodic systems, since the nonlocal part of the exchange becomes fairly short ranged.[9,10] Remarkably, this hybrid functional yields truly impressive band gaps, which are often close to the band gaps calculated using the *GW* method.[11,12] The question arises whether these improvements carry over to the description of optical-absorption spectra and related properties such as the static dielectric constants, which are notoriously overestimated using semilocal functionals.[13] The main achievement of the present work is to show that hybrid functionals indeed substantially improve the description of static and dynamic screening properties.

The key property of the HSE (Heyd, Scuseria, Ernzerhof) functional is the separation of the exchange energy into a short-range nonlocal and orbital-dependent exchange term $E_{x-\mathrm{nl}}^{\mathrm{SR}}$ and the remainder treated by a semilocal approximation[8,14]

$$E_x^{\mathrm{HSE}} = E_{x\text{-}l}^{\mathrm{PBE}} - 1/4E_{x\text{-}l}^{\mathrm{SR,PBE}}(\mu) + 1/4E_{x\text{-}\mathrm{nl}}^{\mathrm{SR}}(\mu). \quad (1)$$

$E_{x\text{-}l}^{\mathrm{PBE}}$ is the exchange energy of the PBE (Perdew, Burke, Ernzerhof) functional,[15] and $E_{x\text{-}l}^{\mathrm{SR,PBE}}(\mu)$ is the semilocal density functional corresponding to $E_{x\text{-}\mathrm{nl}}^{\mathrm{SR}}(\mu)$. The resulting one-electron equation contains 1/4 of a Hartree-Fock-type term, and the functional thus belongs to the class of generalized Kohn-Sham functionals.[16] The parameter μ is empirically set to 0.2–0.3 $\mathrm{Å}^{-1}$ in the nonlocal and semilocal part, and it defines a length scale for the separation. Note that the interaction range ($2/\mu \approx 10$ Å) is typically over a few nearest neighbors and, therefore, only short ranged compared to the bare Fock exchange potential. In the present work, we have used $\mu = 0.3$ $\mathrm{Å}^{-1}$ except when otherwise noted.

To determine the static and dynamic dielectric functions we first calculate the occupied and a few virtual unoccupied orbitals (one-electron states) using the HSE functional and the plane-wave projector augmented wave (PAW) code VASP.[17] For the evaluation of the dielectric matrices we essentially use time-dependent generalized Kohn-Sham density-functional theory (TD-DFT) in its linear-response formulation, where the full polarizability χ is given by the usual Dyson equation

$$\begin{aligned}\chi &= [1 - \chi_0(v + f_{\mathrm{xc}})]^{-1}\chi_0 \\ &= \chi_0 + \chi_0(v + f_{\mathrm{xc}})\chi_0 + \chi_0(v + f_{\mathrm{xc}})\chi_0(v + f_{\mathrm{xc}})\chi_0 + \ldots .\end{aligned} \quad (2)$$

The matrix χ_0 is the independent-particle polarizability and v the Coulomb kernel $1/|\mathbf{r}-\mathbf{r}'|$. In the second line, we have written the inverse as a geometrical series, which casts the Dyson equation in a simple transparent form: The difference between χ and χ_0 is the inclusion of the response of the electrons to the induced changes in the potential $(v+f_{\mathrm{xc}})\chi_0$ and a summation of these diagrams up to infinity. We note that χ_0 and χ are frequency-dependent matrices that are expanded in a plane-wave basis set in the present implementation.[18] For local adiabatic functionals, the kernel f_{xc} is simply the second variation in the exchange-correlation energy with respect to the density

$$f_{\mathrm{xc}\text{-}l}(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}')\frac{\delta^2 E_{\mathrm{xc}\text{-}l}[n(\mathbf{r})]}{\delta n(\mathbf{r})\,\delta n(\mathbf{r}')}. \quad (3)$$

Due to the locality of this term, it lacks any electrostatic interactions between electrons and holes. To include the ef-

fective exchange-correlation kernel from the nonlocal exchange term we follow the recent suggestions of Reining and co-workers[19–22] and recast the exact four-orbital electron-hole interactions deriving from the nonlocal exchange term into an effective nonlocal frequency-dependent kernel $f_{x\text{-nl}}$, defined through (nanoquanta kernel)

$$f_{x\text{-nl}} = \chi_0^{-1} T \chi_0^{-1} \leftrightarrow \chi_0 f_{x\text{-nl}} \chi_0 = T. \tag{4}$$

The matrix T describes how the *local* electronic charge density changes upon changes in the nonlocal exchange potential induced by an external *local* potential. In other words, the response function $T(\mathbf{r}', \mathbf{r}, \omega)$ captures the following process: change in external potential at $\mathbf{r}\rightarrow$ linear response of orbitals $\rightarrow$ change in nonlocal exchange potential $\rightarrow$ linear response of orbitals, and resultant charge density change at $\mathbf{r}'$. Note that the equivalent terms for the Hartree potential and the local exchange-correlation potential have been accounted for by the term $\chi_0(v+f_{\text{xc-}l})\chi_0$ in Eq. (2). It is straightforward to derive explicit expressions for the matrix T, and the results have been published in a series of recent papers.[19–22] Since the calculation of the T matrix remains a rather time-consuming task involving the summation over millions of two-particle interactions, we neglect the frequency dependence of $f_{x\text{-nl}}$ and determine it only once at zero frequency. We also include the two-electron terms related to the coupling between resonant and antiresonant parts of the excitation.[1,23]

An alternative well established approach to calculate frequency-dependent response functions for density or hybrid functionals is to solve Casida's equation.[24] The corresponding matrix equation in a four-point formulation is essentially equivalent to the well-known Bethe-Salpeter equation, with the screened interaction W replaced with one-quarter of the nonlocal screened exchange term with a fixed screening length μ. We note that this analogy immediately suggests that excitonic effects are *approximately* captured by hybrid functionals. Due to the time-consuming diagonalization of the two-particle electron-hole problem, we found this approach to be more expensive than the TD-DFT method, but emphasize that, for coarse k-point grids, both methods yield practically identical results in agreement with previous work.[21]

In Fig. 1 we show the imaginary part of the head of the test-charge test-charge dielectric function $\epsilon=(1+v\chi)^{-1}$ for Si, GaAs, SiC, and C calculated using Eq. (2) and compare with experimental absorption spectra. Also shown are the results for the independent-particle approximation $\epsilon_{\text{IP}}=1+v\chi_0$, where many-electron effects have been entirely neglected. For the present calculations $32\times32\times32$ k points and eight empty orbitals were used, except for the weakly screened materials ZnO and LiF, where 36 empty orbitals had to be used in order to obtain sufficiently converged results. Since a straightforward calculation would be too expensive in the HSE case, we have performed the calculations summing over many shifted $8\times8\times8$ grids, where the shifts were systematically varied to mimic the larger grid. For the independent-particle approximation the results are essentially exact, but for the full TD-HSE case the nonlocal exchange kernel $f_{x\text{-nl}}$ is limited, in each direction, to an interaction

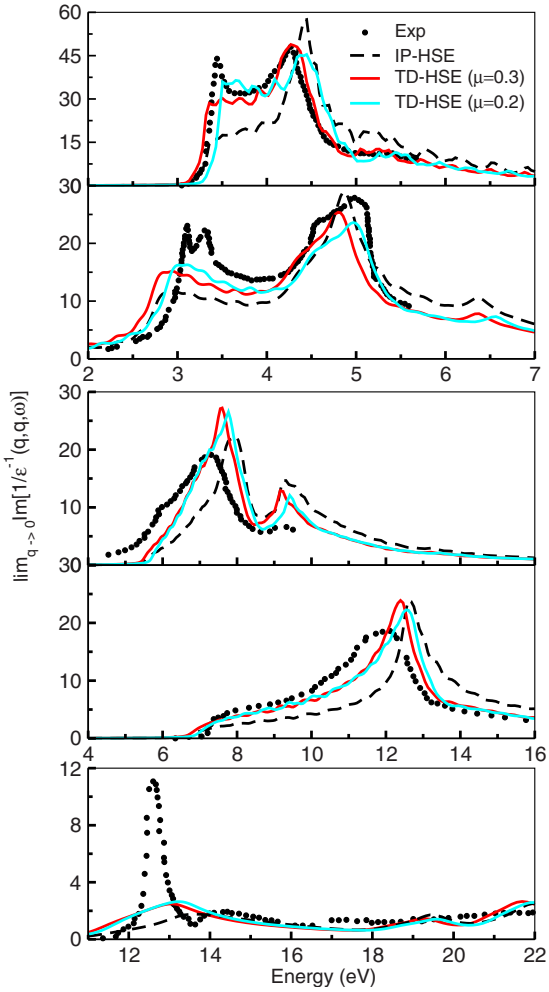

FIG. 1. (Color online) Optical-absorption spectra for Si, GaAs, SiC, C, and LiF using TD-HSE (μ=0.3 Å^{-1}) in the independent-particle approximation (dashed black line) and including electron-hole interactions [full red (dark gray) line]. TD-HSE06 results (μ=0.2 Å^{-1}) are shown using the blue (light gray) line, and experiments are shown by dots [Si (Ref. 25), GaAs (Ref. 26), SiC (Ref. 27), C (Ref. 28), and LiF (Ref. 29)].

range of eight unit cells. This is expected to be accurate since the interaction range of the HSE functional is anyway restricted.

In the independent-particle approximation (dashed line) the results are similar to those reported in literature for a scissor corrected DFT approach. Most notably at low energies the cross sections are much too weak, e.g., for Si the first peak is only visible as a flat plateau. Inclusion of many-body effects via $f_{x\text{-nl}}$ causes a redshift of the spectrum and a pronounced increase in the cross section at low energies, improving agreement with experiment [red (dark gray) line]. For Si, using a reduced screening parameter of 0.20 Å^{-1}, i.e., extending the range of the nonlocal exchange part of the HSE functional, increases the cross section of the first peak further but sacrifices the excellent position of the HSE spectrum [blue (light gray) line in Fig. 1]. Obviously, the interaction range of the HSE functional is slightly too short in order to correctly account for the electrostatic interaction in the weakly bound electron-hole pairs in bulk silicon. For

TABLE I. Ion clamped (high frequency) macroscopic dielectric constants ϵ^{∞} from TD-DFT using the LDA and the HSE (μ=0.3 Å^{-1}) hybrid functional in the independent-particle approximation ($\epsilon^{\infty}_{\mathrm{IP}}$) and including all electron-electron interactions. The HSE results have been obtained either by solving the Dyson equation or by applying a finite field and extracting the response from the change in the polarization (Refs. 30 and 31). For ZnO the dielectric constants are reported for the wurtzite structure along the *a* and *c* axes. All data are calculated at the experimental volumes.

	LDA		HSE		HSE fin. field		
	$\epsilon^{\infty}_{\mathrm{IP}}$	ϵ^{∞}	$\epsilon^{\infty}_{\mathrm{IP}}$	ϵ^{∞}	$\epsilon^{\infty}_{\mathrm{IP}}$	ϵ^{∞}	Expt.
Si	14.1	13.35	10.94	11.31	10.87	11.37	11.9[a]
GaAs	14.81	13.98	10.64	10.95	10.54	11.02	11.1[a]
AlP	9.12	8.30	7.27	7.35	7.32	7.35	7.54[a]
SiC	7.29	6.96	6.17	6.43	6.15	6.44	6.52[a]
C	5.94	5.80	5.21	5.56	5.25	5.59	5.7[a]
ZnO *c*	5.31	5.15	3.50	3.71	3.57	3.77	3.78[b]
ZnO *a*	5.28	5.11	3.48	3.67	3.54	3.72	3.70[b]
LiF	2.06	2.02	1.85	1.90	1.86	1.91	1.9[c]

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

GaAs, a screening parameter of 0.20 instead of 0.30 Å^{-1} improves both the onset as well as shape of the absorption spectrum compared to experiment [blue (light gray) line in Fig. 1]. For SiC and C results agree also well with experiment. But for LiF, a large gap insulator with very weak screening, agreement with experiment remains modest, with much too weak excitonic peaks at low frequencies.

The calculations suggest that the HSE functional approximates the electrostatic interaction between a hole with the charge density ρ_h and an electron with the charge density ρ_e reasonably well. In the BSE, this electrostatic interaction is given by a term proportional to

$$\int \rho_h(\mathbf{r}) W(\mathbf{r},\mathbf{r}') \rho_e(\mathbf{r}') d^3\mathbf{r} d^3\mathbf{r}', \tag{5}$$

where $W \propto \epsilon^{-1} v$. The main achievement of the HSE functional is to approximate W reasonably well for semiconductors. For strongly localized excitons, typically found in large gap insulators, the HSE functional will approximate W by a quarter of the bare Coulomb kernel v, which is necessarily insufficient: in LiF ϵ is roughly 2 (see Table I), and at least half of the Coulomb kernel is required to approximate W. For more dispersed excitations, the long-range cutoff used in the HSE functional will become progressively more important, further reducing the amount of nonlocal exchange, and for excitations relevant to semiconductors, a cutoff around $\mu \approx 0.2$–0.3 Å^{-1} mimics the exact W quite well. But certainly no *a priori* choice of μ and no choice for the nonlocal exchange can describe all systems equally well.

We now turn to the static dielectric constants for selected semiconducting and insulating systems reported in Table I. For comparison, we present the corresponding local-density approximation (LDA) values obtained using density-functional perturbation theory (linear response),[35] as well as HSE results obtained by applying a finite field and extracting the polarizability from the change in the polarization.[30,31] It is immediately recognized that the LDA calculations overestimate the screening in the independent-particle approximation (ϵ_{IP}). The inclusion of the local-field effects via $v+f_{\mathrm{xc}}$ in Eq. (2) results in a slight reduction in the screening, which is related to the repulsive effect of the Hartree kernel v partly compensated by the attractive local exchange-correlation kernel $f_{\mathrm{xc}\text{-}l}$.

TABLE II. Band gaps for GW_0 calculations without (RPA) and with attractive electron-hole interaction (vertex corrections in W only, i.e., $GW_0^{\mathrm{TC\text{-}TC}}$). Eigenvalues were updated in G until convergence was reached, whereas W was calculated using HSE wave functions and eigenvalues. Details and experimental values are identical to Ref. 37 and references therein.

	HSE	GW_0^{RPA36}	$GW_0^{\mathrm{TC\text{-}TC}}$	Expt.
Si	1.04	1.37	1.25	1.17
GaAs	1.12	1.72	1.57	1.52
SiC	2.03	2.68	2.41	2.40
CdS	1.97	2.62	2.37	2.42
AlP	2.09	2.79	2.60	2.45
GaN	2.66	3.42	3.09	3.20
ZnO	2.12	3.07	2.55	3.44
ZnS	3.06	3.83	3.50	3.91
C	5.07	5.92	5.65	5.48
BN	5.54	6.70	6.31	6.1–6.4
MgO	6.19	8.13	7.60	7.83
LiF	11.22	14.75	14.01	14.20

For HSE, the dielectric constant is too small for the IP case, and only inclusion of $v+f_{\mathrm{xc}\text{-}l}+f_{x\text{-nl}}$ in Eq. (2) yields good agreement with experiment. This is of course related to the strong redshift of the absorption spectra and the increase in the intensity at low frequencies from the independent-particle case to the full calculation (see Fig. 1). Although, the dielectric constants remain somewhat too small, the average deviation between theory and experiment has decreased from 10% (LDA) to roughly 3% (HSE). Eventually, these results also show that the exact finite field approach and the TD-HSE approach, which involves a summation over a truncated number of virtual orbitals, yield *de facto* identical results. This validates our implementation and confirms that convergence of the TD-HSE results has been achieved.

It has been suggested that generalized Kohn-Sham functionals might be a good starting point for *GW* calculations.[36] But the strong excitonic effects we observe in the present work require one to take the next step with care. In the *GW* approximation W is usually constructed including the Coulomb kernel v only, neglecting f_{xc} in Eq. (2). This so-called random-phase approximation (RPA) is problematic for the HSE case since it results in a quite severe underestimation of the static screening. This suggests that *GW* calculations starting from a hybrid functional should include f_{xc}, which is usually done via vertex corrections.[22] This is clearly confirmed in Table II, where we show that straightforward GW_0

band-gap calculations starting from a HSE functional tend to overestimate the band gaps, and only the inclusion of $f_{x\text{-nl}}$ in the calculation of χ rectifies the problem. A similar observation is made for self-consistent *GW* calculations.[37] The present calculations only fail for systems with shallow *d* electrons, but here the *GW* approximation seems to suffer from fairly large self-interaction errors for local *d* electrons.[37]

In summary, we have reported the calculation of static and dynamic screening properties of extended systems using hybrid functionals. Apart from a better prediction of the precise position of the peaks in the absorption spectra, the hybrid functional also describes the scattering cross sections in the low energy region reasonably well. Even the static dielectric constants, which are profound ground-state properties, are better captured by the HSE functional. In general, the work confirms that hybrid functionals yield results superior to semilocal functionals for *semiconductors and small gap insulators*, but we have also shown that such functionals involve compromises (screening length) and are not truly as predictive as *GW*. Large gap insulators are for instance not as well described as using *GW*.

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[2] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[3] M. Petersilka, U. J. Gossmann, and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).
[4] M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10031 (1999).
[5] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[6] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[7] R. L. Martin and F. Illas, Phys. Rev. Lett. **79**, 1539 (1997).
[8] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).
[9] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **121**, 1187 (2004).
[10] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. Ángyán, J. Chem. Phys. **124**, 154709 (2006); **125**, 249901 (2006) (erratum).
[11] J. Muscat, A. Wander, and N. M. Harrison, J. Chem. Phys. **342**, 397 (2001).
[12] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005).
[13] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, Phys. Rev. B **73**, 045112 (2006).
[14] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. **125**, 224106 (2006).
[15] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[16] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, Phys. Rev. B **53**, 3764 (1996).
[17] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996); Phys. Rev. B **54**, 11169 (1996).
[18] M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[19] L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett. **88**, 066404 (2002).
[20] A. Marini, R. Del Sole, and A. Rubio, Phys. Rev. Lett. **91**, 256402 (2003).
[21] F. Sottile, V. Olevano, and L. Reining, Phys. Rev. Lett. **91**, 056402 (2003).
[22] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, Phys. Rev. Lett. **94**, 186402 (2005).
[23] V. Olevano and L. Reining, Phys. Rev. Lett. **86**, 5962 (2001).
[24] M. E. Casida, in *Recent Advances in Density Functional Methods*, edited by D. P. Chong (World Scientific, Singapore, 1995), Vol. 1.
[25] P. Lautenschlager, M. Garriga, L. Vina, and M. Cardona, Phys. Rev. B **36**, 4821 (1987).
[26] D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).
[27] S. Logothetidis and J. Petalas, J. Appl. Phys. **80**, 1768 (1996).
[28] H. R. Phillipp and E. A. Taft, Phys. Rev. **136**, A1445 (1964).
[29] D. M. Roessler and W. C. Walker, J. Opt. Soc. Am. **57**, 835 (1967).
[30] R. W. Nunes and X. Gonze, Phys. Rev. B **63**, 155107 (2001).
[31] I. Souza, J. Íñiguez, and D. Vanderbilt, Phys. Rev. Lett. **89**, 117602 (2002).
[32] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 2001).
[33] X. Wu, D. Vanderbilt, and D. R. Hamann, Phys. Rev. B **72**, 035105 (2005).
[34] F. Bechstedt and R. Del Sole, Phys. Rev. B **38**, 7710 (1988).
[35] S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).
[36] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).
[37] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).