REVIEWS OF MODERN PHYSICS, VOLUME 80, JANUARY–MARCH 2008

Orbital-dependent density functionals: Theory and applications

Stephan Kümmel*

Physikalisches Institut, Universität Bayreuth, D-95440 Bayreuth, Germany

Leeor Kronik†

Department of Materials and Interfaces, Weizmann Institute of Science, Rehovoth 76100, Israel

(Published 2 January 2008)

This review provides a perspective on the use of orbital-dependent functionals, which is currently considered one of the most promising avenues in modern density-functional theory. The focus here is on four major themes: the motivation for orbital-dependent functionals in terms of limitations of semilocal functionals; the optimized effective potential as a rigorous approach to incorporating orbital-dependent functionals within the Kohn-Sham framework; the rationale behind and advantages and limitations of four popular classes of orbital-dependent functionals; and the use of orbital-dependent functionals for predicting excited-state properties. For each of these issues, both formal and practical aspects are assessed.

DOI: 10.1103/RevModPhys.80.3 PACS number(s): 31.15.E−, 71.15.Mb, 31.70.Hq, 31.10.+z

CONTENTS

I. INTRODUCTION

A. The Kohn-Sham equation

Density-functional theory (DFT) is an approach to the many-electron problem in which the electron density, rather than the many-electron wave function, plays the central role (Parr and Yang, 1989; Dreizler and Gross, 1990; Kohn, 1999). In recent years, DFT has become the method of choice for electronic-structure calculations across an unusually wide variety of fields, from organic chemistry (Koch and Holthausen, 2001) to condensed matter physics (Martin, 2004). There are two main reasons for the success of DFT (Kohn, 1999): First, DFT offers the only currently known practical method of fully quantum-mechanical calculations for systems with many hundreds or even thousands of electrons. Second, it enhances our understanding by relying on relatively simple, physically accessible quantities that are easily visualized even for very large systems (Geerlings *et al.*, 2003).

The central tenet of DFT is the Hohenberg-Kohn theorem (Hohenberg and Kohn, 1964), which states that the ground-state density $n(\vec{r})$ of a system of interacting electrons in some external potential $v_{\text{ext}}(\vec{r})$ determines this potential uniquely up to a constant. The importance of this theorem lies in the fact that (within the Born-Oppenheimer approximation) any atomic, molecular, or condensed matter system is completely described by the number of electrons in it [which is simply an integral of $n(\vec{r})$ over space] and by the types and positions of the nuclei in it. However, the effect of the latter two quantities on the electron system is described completely by an ion-electron attraction potential,

$$v_{\text{ion}}(\vec{r}) = -e^2 \sum_a \frac{Z_a}{|\vec{r} - \vec{R}_a|}, \tag{1}$$

---

*stephan.kuemmel@uni-bayreuth.de
†leeor.kronik@weizmann.ac.il

where e is the absolute value of the electron charge, Z_a and $\vec{R}_a$ are the atomic number and position, respectively, of the ath atom, and the sum extends over all nuclei in the system. $v_{\text{ion}}(\vec{r})$ is external to the electron system, i.e., uniquely determined by the ground-state density. This means that the latter quantity contains all the information needed to describe the system completely and uniquely.

An important variational principle is associated with the Hohenberg-Kohn theorem (Hohenberg and Kohn, 1964; Levy, 1982; Lieb, 1982). The ground-state density can be found by minimizing a total energy expression of the form

$$E_{\text{tot}} = \int v_{\text{ext}}(\vec{r})n(\vec{r})d^3r + F[n] \tag{2}$$

with respect to the density, where $F[\cdot]$ is a universal (i.e., system-independent) functional of the density.

Unfortunately, explicit variation of Eq. (2) has not become a practical computational device. While the existence and uniqueness of $F[\cdot]$ are guaranteed, its form is not known, and sufficiently accurate explicit approximations have not yet emerged. Fortunately, Kohn and Sham (1965) used Eq. (2) to construct a more practical set of equations. They considered a fictitious system of noninteracting electrons and postulated the existence of an external potential $v_{\text{KS}}(\vec{r})$, such that the ground-state density obtained from the fictitious, noninteracting-electron system is equal to the ground-state density of the given real, interacting-electron system. If such a potential exists, then solving a set of one-electron Schrödinger equations of the form

$$\left(-\frac{\hbar^2\nabla^2}{2m} + v_{\text{KS}}(\vec{r})\right)\varphi_i(\vec{r}) = \varepsilon_i\varphi_i(\vec{r}), \tag{3}$$

where $\hbar$ is the reduced Planck constant, m is the electron mass, and ε_i and φ_i are known as the Kohn-Sham energies and orbitals, respectively, with the density calculated as a sum over filled orbitals,

$$n(\vec{r}) = \sum_{i,\text{occupied}} |\varphi_i(\vec{r})|^2, \tag{4}$$

should yield the correct ground-state density of the original system. One can then choose to express the total energy as follows:

$$E_{\text{tot}} = -\frac{\hbar^2}{2m}\sum_i \langle\varphi_i|\nabla^2|\varphi_i\rangle + \int v_{\text{ext}}(\vec{r})n(\vec{r})d^3r + \frac{e^2}{2}\int\int\frac{n(\vec{r}_1)n(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}d^3r_1d^3r_2 + E_{\text{xc}}[n], \tag{5}$$

where, in light of Eq. (2),

$$E_{\text{xc}}[n] \equiv F[n] + \frac{\hbar^2}{2m}\sum_i \langle\varphi_i|\nabla^2|\varphi_i\rangle - \frac{e^2}{2}\int\int\frac{n(\vec{r}_1)n(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}d^3r_1d^3r_2. \tag{6}$$

Equation (5) expresses the total energy as a sum of the kinetic energy $T_s[n]$ for a noninteracting-electron system, the electron–external-potential interaction energy [e.g., the electron-ion interaction energy, if $v_{\text{ion}}(\vec{r})$ of Eq. (1) is used as $v_{\text{ext}}(\vec{r})$], the electron-electron classical (Hartree) Coulomb repulsion $E_H[n]$, and an "everything else" term $E_{\text{xc}}[n]$, known as the exchange-correlation energy. This last term must include all nonclassical electron interactions, namely, Pauli exchange, electron correlation, as well as the difference between the kinetic energy of the interacting- and noninteracting-electron systems (despite not being mentioned explicitly in the term exchange-correlation energy).

A key observation here is that the Hohenberg-Kohn theorem is valid for any electron interaction, including zero. It is therefore applicable to a noninteracting system just as it is applicable to an interacting-electron system. Thus the Kohn-Sham orbitals, being uniquely defined by $v_{\text{KS}}(\vec{r})$, which is itself uniquely defined by $n(\vec{r})$, are themselves functionals of the density. With this in mind, variation of Eq. (5) with respect to $n(\vec{r})$ leads to the condition (Kohn and Sham, 1965)

$$v_{\text{KS}}(\vec{r}) = v_{\text{ext}}(\vec{r}) + \int\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}d^3r' + v_{\text{xc}}([n];\vec{r}), \tag{7}$$

where

$$v_{\text{xc}}([n];\vec{r}) \equiv \frac{\delta E_{\text{xc}}}{\delta n(\vec{r})}. \tag{8}$$

Equation (3), with $v_{\text{KS}}(\vec{r})$ given by Eq. (7) and $n(\vec{r})$ given by Eq. (4), is known as the Kohn-Sham equation. Its beauty lies in the fact that, for any given system, it is an exact mapping of the original, often numerically intractable, N-electron Schrödinger equation into a numerically manageable set of N self-consistent one-electron equations, with all the pertinent many-body physics lumped into a single additive potential $v_{\text{xc}}(\vec{r})$. For later use in this paper, we quote a generalization to spin-polarized form (von Barth and Hedin, 1972; Rajagopal and Callaway, 1973),

$$\left(-\frac{\hbar^2\nabla^2}{2m} + v_{\text{ext}}(\vec{r}) + e^2\int\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}d^3r' + v_{\text{xc},\sigma}([n_\uparrow,n_\downarrow];\vec{r})\right)\varphi_{i,\sigma}(\vec{r}) = \varepsilon_{i,\sigma}\varphi_{i,\sigma}(\vec{r}), \tag{9}$$

where $\sigma=\uparrow$ or $\downarrow$ is the electron spin z-component index,

$$n_\sigma(\vec{r}) = \sum_{i=1}^{N_\sigma}|\varphi_{i,\sigma}(\vec{r})|^2 \tag{10}$$

with $n(\vec{r})=n_\uparrow(\vec{r})+n_\downarrow(\vec{r})$, N_σ is the number of filled orbitals of spin σ, and

$$v_{\text{xc},\sigma}([n_\uparrow,n_\downarrow];\vec{r}) \equiv \frac{\delta E_{\text{xc}}}{\delta n_\sigma(\vec{r})}. \tag{11}$$

As Kohn-Sham eigenvalues and orbitals are considered below, we take the opportunity to make some general remarks about these quantities. The first thing to note is that their physical interpretation is not as straightforward as the interpretation of the density or total energy. This stems from the fact that the (occupied) Kohn-Sham orbitals and their eigenvalues were originally introduced as mere mathematical devices to be used to obtain the correct ground-state density. Therefore, in the early days of density-functional theory, it was believed that no physical meaning should be attached to them. However, later research clarified that they do have a well-defined physical meaning as zeroth-order excitation energies (Görling, 1996; Filippi *et al.*, 1997) and may be used to construct correlation functionals from a perturbation expansion (Görling and Levy, 1993); see Sec. III.D.2. Still, Kohn-Sham eigenvalues need to be interpreted carefully (Perdew and Zunger, 1981; Chong *et al.*, 2002).

In Hartree-Fock theory, Koopmans' theorem (Koopmans, 1934) identifies the *i*th occupied eigenvalue as the energy required to remove an electron from that orbital without perturbing the rest of the system (i.e., unrelaxed). This theorem is generally not valid in DFT. Instead, it is replaced by a weaker condition, known as Janak's theorem (Janak, 1978), stating that the *i*th occupied eigenvalue is equal to the partial derivative of the total energy with respect to a continuous (fractional) occupation of that orbital. Only the highest occupied eigenvalue can be assigned a rigorous physical meaning (Levy *et al.*, 1984; Almbladh and von Barth, 1985): for this special case, the partial derivative, to which the eigenvalue is equal, is exactly the chemical potential of the system.

From this perspective, one might be tempted to believe that there is less physical meaning in Kohn-Sham than in Hartree-Fock eigenvalues. However, interpreting Kohn-Sham eigenvalues and orbitals as if they were one-electron energies and orbitals often results in surprisingly accurate results. For example, for the ideal uniform electron gas, where both excitation energies and Kohn-Sham eigenvalues can be computed exactly, a significant deviation between the two is apparent only far away from the Fermi level (Jones and Gunnarsson, 1989). More insight into this intriguing behavior was given by Chong *et al.* (2002), who showed that Kohn-Sham eigenvalues are good approximations to *relaxed* electron removal energies [see Gritsenko and Baerends (2002) for an extension of the conclusion to spin-polarized systems]. Thus, they provide more, rather than less, information than Hartree-Fock eigenvalues. To demonstrate this, Chong *et al.* computed accurate exchange-correlation potentials, reconstructed from charge densities determined by wave-function-based methods, for several small molecules. They found that for higher-lying (outer valence) electrons, the average difference between the Kohn-Sham eigenvalues and the experimental ionization energies was only ~0.1 eV. More substantial differences were found only for lower-lying electrons, just as in the uniform electron gas. Thus, it is appropriate to interpret Kohn-Sham eigenvalues and orbitals as tools in the traditional qualitative molecular orbital analysis that is often used in chemistry (Kohn *et al.*, 1996; Baerends and Gritsenko, 1997; Stowasser and Hoffmann, 1999). In this view, the Kohn-Sham potential is a suitable one-electron potential that carries all correlation information within it, and Kohn-Sham orbitals are one-electron orbitals that are *density* optimal (as opposed to Hartree-Fock orbitals, which are *energy* optimal).

Despite the elegance of the Kohn-Sham equation, it would remain useless in practice without adequate approximations for the exchange-correlation functional $E_{\mathrm{xc}}[n_\uparrow(\vec{r}),n_\downarrow(\vec{r})]$. The hope, however, is that inclusion of a major portion of the kinetic energy in terms of the Kohn-Sham orbitals will make approximate expressions for the exchange-correlation energy that are explicit functionals of the density easier to construct. These approximate expressions should be useful in yielding quantitative, experimentally verifiable information for at least some systems. In the following section, we review the major triumphs, but also the major failures, of such explicit approximations. This will motivate and introduce the concept of orbital-dependent exchange-correlation functionals, which is the primary focus of this paper.

B. Explicit density functionals and their limitations

An appealingly simple approximation for the exchange-correlation energy functional, already suggested by Hohenberg and Kohn (1964), is the local-density approximation (LDA) [or, in its spin-polarized form, the local-spin-density approximation (LSDA)]. In this approximation, we assume that at each point in space the exchange-correlation energy per particle is given by its value for a homogeneous electron gas, $e_{\mathrm{xc}}(n_\uparrow(\vec{r}),n_\downarrow(\vec{r}))$, namely,

$$E_{\mathrm{xc}}^{\mathrm{LSDA}}[n_\uparrow,n_\downarrow]=\int n(\vec{r})e_{\mathrm{xc}}(n_\uparrow(\vec{r}),n_\downarrow(\vec{r}))d^3r. \tag{12}$$

This approximation has several advantages. First, $e_{\mathrm{xc}}(n_\uparrow(\vec{r}),n_\downarrow(\vec{r}))$ is unique because of the theoretical existence of a system with uniform spin densities. Second, e_{xc} was known approximately already when Hohenberg, Kohn, and Sham were formulating DFT, and has since been computed more accurately using Monte Carlo methods for unpolarized and fully polarized electron gases (Ceperley and Alder, 1980). Furthermore, several useful analytical parametrizations of the Monte Carlo results, which also take into account other known limits and/or scaling laws of e_{xc}, have been given (Vosko *et al.*, 1980; Perdew and Zunger, 1981; Perdew and Wang, 1992). Last but not least, the exchange-correlation potential, given by Eq. (11), becomes a simple function, rather than a functional, of $n_\uparrow(\vec{r})$ and $n_\downarrow(\vec{r})$.

The difficulty with the L(S)DA is that in real systems the density is clearly not uniform. More often than not, it actually exhibits rapid changes in space. The L(S)DA was therefore expected to be of limited value in providing an accurate description of the electron interaction

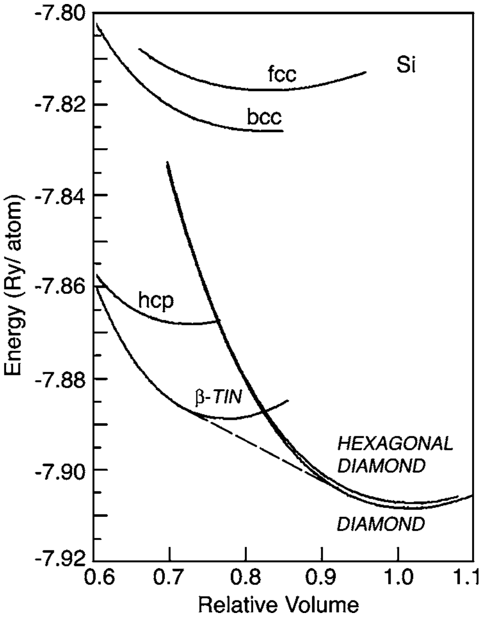

FIG. 1. Energy vs relative volume (with respect to the experimental one) for several crystal structures of silicon: diamond, hexagonal diamond, β-tin, hexagonal close packed (hcp), body-centered cubic (bcc), and face-centered cubic (fcc). Also shown is the tangent line to the diamond and β-tin structures, from which the transition volumes and pressure can be computed. From Yin and Cohen, 1980.

(Kohn and Sham, 1965). Nevertheless, it has often been found to provide surprisingly accurate predictions of experimental results. The literature on L(S)DA applications is far too vast to review here, but one striking example is given by the pioneering calculations of solid-solid phase transformations in silicon by Yin and Cohen (1980). They used the LDA to compute the total energy, as a function of volume, for several representative phases of silicon, as shown in Fig. 1. They were able to identify the cubic diamond phase as the lowest-energy one, in agreement with experiment, obtain a lattice parameter in excellent agreement with experiment, predict the high-pressure phase transformation to the β-tin structure, again in agreement with experiment, and even compute the transition pressure and volumes.

One important reason for the success of the L(S)DA well outside its formal range of validity is provided by the concept of the exchange-correlation hole $n_{\mathrm{xc}}(\vec{r},\vec{r}\,')$. The latter can be thought of as the difference between the average density of electrons at $\vec{r}\,'$, given an electron at $\vec{r}$, and the normal probability density $n(\vec{r}\,')$ at $\vec{r}\,'$ (Gunnarson and Lundqvist, 1976), i.e.,

$$n_{\mathrm{xc}}(\vec{r},\vec{r}\,') \equiv \rho_2(\vec{r},\vec{r}\,')/n(\vec{r}) - n(\vec{r}\,'), \tag{13}$$

where $\rho_2(\vec{r},\vec{r})$ is the electron pair density, i.e., the probability density to simultaneously find an electron at $\vec{r}$ and an electron $\vec{r}\,'$. It can then be shown that (i) the exchange-correlation hole obeys an important sum rule—its integral over all space is exactly −1 (to account for the electron already at $\vec{r}$); (ii) the exchange-correlation energy can be expressed as the Coulomb interaction of the density with the (coupling-constant-averaged; see Sec. III.C.1) exchange-correlation hole and depends only on its spherical average (Gunnarson and Lundqvist, 1976). The approximate exchange-correlation hole afforded by the L(S)DA has been found to be in poor agreement with the exact one for systems deviating substantially from homogeneity (e.g., atoms). Nevertheless, its integral over all space is still exactly −1, and, moreover, its spherical average is often still a reasonable approximation of the exact spherically averaged exchange-correlation hole (Gunnarson *et al.*, 1979; Hood *et al.*, 1998). In fact, it is a subtle cancellation of errors between the approximate exchange and correlation holes that allows for the successes of the L(S)DA, an issue elaborated in Secs. III.C and III.D.

While the above reasoning and numerical example are compelling, the L(S)DA is by no means a panacea. Even when its predictions are qualitatively acceptable (which is not always the case—see below), the L(S)DA is far from perfect quantitatively. The L(S)DA tends to overestimate the bonding strength, resulting in bond lengths that typically are too short by several percent. Therefore, the L(S)DA tends to overly favor close-packed structures. It also provides an absolute error of molecular atomization energies of the order of 1 eV—much larger than the desired chemical accuracy of roughly 0.05 eV (Becke, 1992a, 1992b; Perdew and Kurth, 2003).

Many of these quantitative failures of the L(S)DA are remedied using the generalized gradient approximation (GGA) for the exchange-correlation energy (Perdew and Kurth, 2003),

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow(\vec{r}),n_\downarrow(\vec{r})] = \int f(n_\uparrow(\vec{r}),n_\downarrow(\vec{r}),\vec{\nabla}n_\uparrow(\vec{r}),\vec{\nabla}n_\downarrow(\vec{r}))d^3r. \tag{14}$$

The GGA is often dubbed a semilocal approximation of the exchange-correlation energy. It is no longer strictly local like the L(S)DA, but includes information on deviations from homogeneity only by considering the gradients of the spin-polarized charge densities. It is called a generalized gradient approximation because the function $f(\cdot)$ in Eq. (14) is not obtained from a simple gradient expansion of E_{xc}. This is because, unlike the L(S)DA, which is derived from a possible (though highly idealized) physical system, a simple gradient expansion is not so derived. Therefore, it does not obey many of the formal properties of the exact E_{xc} that the L(S)DA does, and so its accuracy is often poor (Ma and Brueckner, 1968; Langreth and Perdew, 1980; Perdew and Kurth, 2003). Instead, $f(\cdot)$ in Eq. (14) is constructed to reproduce the exact result in certain limits, e.g., the slowly varying and rapidly varying limits, and obeys many important properties of the exact functional, e.g., scaling laws (Levy, 1991). There are also empirical avenues for determining $f(\cdot)$. Use of such constructed functionals (Becke, 1986, 1988; Perdew and Wang, 1986; Lee *et al.*, 1988; Perdew, 1991; Perdew *et al.*, 1992; Perdew, Burke, and Ernzerhof, 1996; Perdew and Kurth, 2003) offers a significantly improved quantitative agreement with experiment. It corrects (often slightly overcorrects) for the L(S)DA overestimation of bonding strength, does not overly favor close-packed structures, and provides an absolute error of molecular atomization energies of the order of 0.3 eV. This is still above the desired chemical accuracy, but much better than the L(S)DA (Becke, 1992a, 1992b; Perdew and Kurth, 2003).

Two simple yet instructive examples of these general trends are afforded by bulk sodium and bulk iron. Sodium is often dubbed the quintessential simple metal and is probably the real system closest to the idealization of the uniform electron gas. However, even there the GGA outperforms the LDA. The LDA underestimates the lattice parameter by more than 3%; GGA [in the Perdew, Burke, and Ernzerhof form (PBE) (1996)] by less than 1%. The LDA overestimates the bulk modulus by over 30%; the GGA by only ~10% (Kronik *et al.*, 2001). For bulk iron, the LDA incorrectly predicts the paramagnetic fcc phase to be more stable than the ferromagnetic bcc phase (by ~5 mRy). The GGA, however, correctly predicts the latter phase to be more stable (by ~20 mRy) (Zhu *et al.*, 1992).

Unfortunately, use of the GGA is not a panacea either. Indeed, the GGA often improves quantitatively upon the LDA [but not always; see, e.g., Stampfl and van der Walle (1999); Yan, Perdew, Kurth, Fiolhais, and Almeida (2000); Mattsson and Jennison (2002); Yu and Scheffler (2004); and Alfe and Gillan (2006)]. However, both suffer from serious formal deficiencies that can result in qualitatively incorrect computations. Here we focus on two important deficiencies that are at the heart of many problems that occur with L(S)DA and GGA functionals: the presence of self-interaction and the absence of a derivative discontinuity in the exchange-correlation energy.

In the total energy expression of Eq. (5), we have somewhat arbitrarily selected to partition the consequences of electron-electron interaction into a classical Coulomb part and an everything else exchange-correlation part. As a result, the Kohn-Sham equation (9) has an explicit classical electron-electron repulsion potential $e^2\int d^3r' n(\vec{r}\,')/|\vec{r}-\vec{r}\,'|$ known as the Hartree potential. Effectively, this expression views the quantum charge probability distribution as a distribution of continuous classical charge. This, however, means that it is inherently in error, because each electron is repelled from the total charge in the system, including a spurious repulsion from itself. This effect is known as the self-interaction error and was already noted by Fermi and Amaldi (1934) in the context of Thomas-Fermi theory—a forerunner of DFT that suggested a semiclassical explicit density functional for the kinetic energy and neglected correlation altogether (Fermi, 1927; Thomas, 1927). The spurious repulsion is particularly easy to see in one-electron systems, where clearly there should be no electron-electron repulsion at all, and yet the Hartree term is not zero.

Because the electron-electron interaction is, in principle, handled exactly in DFT, whatever error we are making in the Hartree term must be completely canceled out by the exchange-correlation term. Unfortunately, complete error cancellation is guaranteed for the exact exchange-correlation functional, but only partial cancellation is obtained in either the L(S)DA or the GGA. This seemingly innocuous issue has far-reaching consequences and is one of the most fundamental problems in DFT.

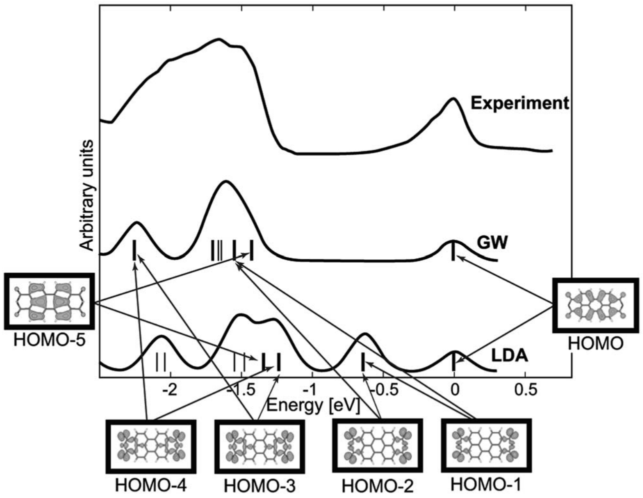

FIG. 2. Gas phase photoelectron spectrum of the outer valence electrons of PTCDA, compared with LDA and *GW* outer valence eigenvalue spectra. Calculated spectra were rigidly shifted so that the highest occupied level is at zero and convolved with a 0.1-eV-wide Gaussian to account for the experimental resolution and for broadening effects. The unbroadened eigenvalues are denoted by vertical lines. Also given is a top view of the spatial distribution of orbitals associated with selected energy levels (superimposed on a schematic view of the molecule), as assigned by *GW* and LDA calculations. The orbital numbering shown is the LDA one, with arrows pointing to the position of the same orbital in the different spectra. HOMO indicates the highest occupied molecular orbital. Adapted from Dori *et al.*, 2006.

The possible errors caused by self-interaction are well demonstrated in, e.g., the work of Dori *et al.* (2006). They computed the valence electronic structure of gas-phase 3,4,9,10-perylene tetracarboxylic dianhydride (PTCDA)—an organic molecule comprising a perylene core and anhydride end groups, whose structure is shown in Fig. 2. The same figure shows the high-energy end of the eigenvalue spectra computed using the LDA and using many-body perturbation theory in the *GW* approximation (Hedin, 1965; Hybertsen and Louie, 1986), as well as an experimental ultraviolet photoelectron spectrum. It is readily observed that, whereas the *GW* calculation agrees reasonably well with experiment, the LDA one does not. As explained above, the ionization energies should be well described by the eigenvalues coming from the exact Kohn-Sham Hamiltonian for these outer valence electrons. Clearly, they are not at all well described by the eigenvalues of the approximate Kohn-Sham Hamiltonian. Insight into the root of this problem is afforded by the detailed orbital analysis shown in Fig. 2. It demonstrates that Kohn-Sham orbitals delocalized over the perylene core (e.g., HOMO and HOMO-5 in the figure) yield LDA energies that are very close to the *GW* ones, i.e., in agreement with experiment. However, Kohn-Sham orbitals localized on the anhydride groups (e.g., HOMO-1 through HOMO-4 in the figure) yield LDA energies that are significantly (well over an eV) too high, resulting in a distorted spectrum. This is a clear self-interaction effect. For localized orbitals, the spurious self-interaction destabilizes the orbital and diminishes its binding energy. For delocalized orbitals, the self-interaction is much smaller because, at each point in space, the spurious same-electron contribution to the Coulomb potential comes from a greater average distance. GGA calculations (not shown here for brevity) suffer from the same problem for the same reason.

Figure 2 demonstrates a general principle. Because the L(S)DA or GGA may fail in the quantitative treatment of localized states due to the self-interaction error, but both normally do much better for delocalized states, they may paint a *qualitatively* distorted picture of the electronic structure whenever localized states interact significantly with delocalized states. Such a scenario arises naturally in studies of defects in solids (Perdew and Zunger, 1981), where, e.g., the LSDA fails to predict hole localization on an Al center in quartz (Pacchioni *et al.*, 2000; Lægsgaard and Stokbro, 2001). Similar problems arise in stoichiometric compounds, e.g., in describing the interaction of Cu 3*d* and O 2*p* electrons in the superconductor $YBa_2Cu_3O_7$ (Temmerman *et al.*, 2001), and even in elemental solids, e.g., in describing the interplay between localized 4*f* electrons and itinerant *s*-*d* electrons in rare-earth metals (Strange *et al.*, 1999).

Self-interaction errors have also been associated with many significant failures of the LDA and/or GGA for predicting chemical trends (see, e.g., Lundberg and Siegbahn, 2005, and references therein). Here we give two examples of particularly egregious qualitative failures. The self-interaction error causes an incorrect long-range behavior of the Kohn-Sham potential (Perdew and Zunger, 1981; van Leeuwen and Baerends, 1994). At a large distance r from the nucleus (or nuclei), the density decays exponentially and therefore so does the LDA and GGA exchange-correlation potential. Thus, the long-range behavior is dominated by the ionic and Hartree potential terms, which, for an ion of charge Q, sum to $-Q/r$. However, an electron at large r should not interact with its own charge, and the correct potential decay should be $-(Q+1)/r$, i.e., the exchange-correlation potential should decay as $\sim -1/r$. This results in "insufficient binding" that leads, e.g., to erroneous LDA and/or GGA predictions for the instability of many atomic and small molecular anions that are experimentally known to be stable (Shore *et al.*, 1977; Schwartz, 1978; Cole and Perdew, 1982; Rösch and Trickey, 1997). We note that this insufficient binding arises also for GGA functionals constructed such that the exchange-correlation potential decays as $1/r^2$ rather than exponentially (Becke, 1988).

Self-interaction errors often lead to qualitatively incorrect predictions for the dissociation behavior of radical ions, with $H_2{}^+$ and $He_2{}^+$ being notable examples (Bally and Sastry, 1997). As the distance between the two H (or He) atoms is increased from its equilibrium value, elementary physical considerations (as well as detailed quantum chemistry calculations) tell us that the binding energy should increase monotonically with interatomic distance, reaching zero at infinity. However, in L(S)DA or GGA computations, the binding energy displays a spurious maximum with respect to the interatomic distance and then actually decreases instead of increasing. This can be interpreted as a self-interaction error: the system-destabilizing self-interaction decreases with increasing electron delocalization over the larger interatomic separation (Zhang and Yang, 1998). In this view, then L(S)DA or GGA correctly predict that the $H_2{}^+$ molecule dissociates into $H^{+0.5}$-$H^{+0.5}$, but they erroneously predict this configuration to be far more stable than the H-H^+ configuration, with which it should actually be degenerate in an exact theory—a serious failure This problem, which is readily understood for the dimers discussed here, occurs also when the dissociation of more complex systems is studied, and thus limits the practical usefulness of DFT.

A different important formal property of the exact exchange-correlation functional is the possibility of functional derivative discontinuities. To understand how these arise, consider first the natural definition of the chemical potential μ in DFT,

$$\mu \equiv \delta E_{\mathrm{tot}}/\delta n(\vec{r}), \tag{15}$$

i.e., the chemical potential is the functional derivative of the total energy with respect to the charge density, on the domain of densities with arbitrary particle number, where μ arises as a Lagrange multiplier that enforces electron number conservation (Perdew *et al.*, 1982). Note that equilibrium implies that μ must be position independent when evaluated at the ground-state density. Perdew *et al.* (1982) noted that, if a system is comprised of two well-separated neutral atoms X and Y, with different chemical potentials $\mu_X > \mu_Y$, then, if E_{tot} is differentiable with respect to particle number, a density variation that shifts some infinitesimal charge from atom X to atom Y will lower the overall energy, although the system is already at stable equilibrium. This paradox is resolved only if the chemical potential exhibits a *discontinuity*: its absolute value should equal the ionization potential I if the integer number of electrons is approached from below, but it should equal the electron affinity A if the number of electrons is approached from above. Because I_X is always larger than A_Y for any atoms X and Y, the energy minimizes at the correct configuration, albeit in a nonanalytical way.

Besides providing the above *reductio ad absurdum* argument, Perdew *et al.* (1982) also provided a constructive proof of the possible discontinuity of the chemical potential. They noted that Eq. (15) is meaningful for charge variations that do not conserve the particle number only if the Hohenberg-Kohn functional, $F[\cdot]$ of Eq. (2), is extended to densities integrating to a noninteger number of particles. This generalization was achieved by considering densities obtained from a statistical mixture of N-electron and $(N+1)$-electron pure states. The total energy versus particle number curve was then shown to be a series of linear segments, for any system. Thus, the E_{tot} versus N curve has to be continuous, but its derivative, namely, the chemical potential, does not.

Of the four terms comprising the total energy expression of Eq. (5), the functional derivative of the external potential and Hartree terms, with respect to $n(\vec{r})$, is clearly continuous. The derivative discontinuities at integer particle numbers can arise, in principle, from either the noninteracting kinetic energy or the exchange-correlation energy (the latter then corresponding to particle number discontinuities in v_{xc}). Special cases in which one of these two terms is differentiable can be constructed, but in general both are expected to exhibit a derivative discontinuity (Perdew and Levy, 1983). This has, in fact, been proved explicitly by considering the general relation between the exact exchange-correlation potential and the electron self-energy (Sham and Schlüter, 1983). Furthermore, explicit computations of the derivative discontinuity show that it can be rather large. For example, by reconstructing the exchange-correlation potential from the self-energy computed within the GW approximation, Godby *et al.* (1986) deduced a discontinuity of 0.58 eV for bulk Si. This is indeed a sizable portion of the overall discontinuity of the chemical potential in Si, i.e., its forbidden gap ($\sim$1.1 eV), with the remaining difference coming from the discontinuity in the noninteracting kinetic energy term. For molecules, where $I-A$ can be much larger, reconstruction of the exchange-correlation potential from the charge density obtained from wave-function-based quantum chemistry calculations yields derivative discontinuities that are easily several eV (Chan, 1999; Allen and Tozer, 2002).

Because the L(S)DA and GGA exchange-correlation potentials are expressions that are continuous in the density and its gradient, these potentials do not exhibit any particle number discontinuity, exposing yet another formal deficiency of these approximations. However, one can question whether this is of any practical consequence. Superficially, one could argue that, because real systems always have an integer number of particles, derivative discontinuities should be irrelevant for the computation of any ground-state property. This reasoning is specious, and derivative discontinuities need to be assessed when developing or testing exchange-correlation functionals (Tozer, 1998; Tozer and Handy, 1998).

To understand why derivative discontinuities matter, consider that the limit of well-separated atoms given in the above *reductio ad absurdum* argument should be approached smoothly by gradually increasing the interatomic distance in the diatomic molecule comprising these two atoms. One can therefore expect the absence of derivative discontinuities to result in spurious charge transfer already at finite separation. For example, a spurious prediction of the LDA for the dissociation of LiH into $\text{Li}^{+0.25}$-$\text{H}^{-0.25}$ was already pointed out by Perdew *et al.* (1982) as due to the absence of a derivative discontinuity. In fact, the lack of degeneracy between $\text{H}^{+0.5}$-$\text{H}^{+0.5}$ and H-H^{+} in the LDA description of the dissociation of $\text{H}_2{}^{+}$, which we interpreted above as a self-interaction issue (Zhang and Yang, 1998), can be rationalized equally well as a derivative discontinuity issue (Chan, 1999). Generally, the lack of derivative discontinuities results in an insufficient barrier to partial charge transfer between weakly overlapping subunits. This makes an accurate description of long-range charge transfer states with the LDA or GGA problematic (Tozer, 2003; Dreuw and Head-Gordon, 2004; Neugebauer *et al.*, 2006; and see comments at the end of Sec. II.C). Similarly, the absence of a proper derivative discontinuity results in excess charge transfer in extended molecular systems polarized by an external electric field (Kümmel *et al.*, 2004) and in spurious fractional charge transfer between weakly interacting subunits under bias (Baer and Neuhauser 2005, 2006; Baer *et al.*, 2006), including failure to capture Coulomb blockade effects (Koentopp *et al.*, 2006).

A different arena in which the derivative discontinuity is clearly manifested is the value of the highest occupied Kohn-Sham orbital. Being rigorously equal to the chemical potential, it should equal $-I$ or $-A$, depending on whether the integer number of particles is approached from above or below. It has been shown that if the Kohn-Sham potential vanishes at large distances, the $-I$ limit should be obtained (Levy *et al.*, 1984; Perdew and Levy, 1997). However, the L(S)DA or GGA averages over the discontinuity (Perdew and Levy, 1983; Tozer, 1998). Thus, the highest occupied Kohn-Sham eigenvalue given in the L(S)DA or GGA, even when they approximate the exchange-correlation hole well, differs from $-I$ by approximately half the derivative discontinuity (Allen and Tozer, 2002). As explained in the preceding section, Kohn-Sham eigenvalue differences are approximations to excitation energies. Therefore, even when computed L(S)DA or GGA spectra agree quite well with experimental photoelectron spectra (i.e., self-interaction is not a major issue), they only do so to within a rigid shift (Akola *et al.*, 2000; Kronik *et al.*, 2002; Moseler *et al.*, 2003). For finite systems, I, and therefore the shift, can be assessed from a simple total energy difference between the N- and $(N-1)$-electron systems. In an extended system, one cannot simply add or remove a single electron. Thus, one often has to go outside the realm of DFT and resort to GW calculations (Hybertsen and Louie, 1986) or to direct comparison with experiment (Segev *et al.*, 2006) to assess the rigid shift.

From a unified perspective, freedom from self-interaction and the existence of derivative discontinuities are related requirements. Both are a manifestation of the principle of integer preference (Perdew, 1990). According to this principle, in collections of separated objects, there is a preference for locating an integer number of electrons on each object. In a sense, the derivative discontinuities lead to the manifestation of this principle in the potential, and freedom from self-interaction is its manifestation in the energy.

C. The promise of orbital-dependent functionals

In the preceding section, we showed that approximate forms for exchange correlation that are explicit density functionals have turned out to be useful for a wide variety of problems in chemistry and physics. Nevertheless, they still possess formal deficiencies that can be linked directly to specific failures for many systems of practical interest. How is further progress to be made? One natural and potentially highly advantageous idea is to use Kohn-Sham orbitals, themselves being functionals of the density, as ingredients in approximate exchange-correlation functionals that are *implicit* density functionals (Grabo *et al.*, 2000; Engel, 2003).

Four strong arguments can be made in favor of such constructs. First, use of orbitals is precisely what made DFT so successful to begin with. It is the definition of a noninteracting kinetic energy density functional, written in terms of one-electron orbitals, that leads to the Kohn-Sham equations (Kohn and Sham, 1965). Without these equations, DFT ceases to be a practical tool, at least at present.

Second, using orbitals allows for much more flexibility in the construction of functionals and thus for more accurate functionals. This is best seen in terms of the Jacob's ladder for exchange-correlation functionals (Perdew and Schmidt, 2001). In this picture, the ladder of DFT reaches up to the heaven of ultimate accuracy. But in order to ascend the ladder, both developers and users of DFT need to climb its rungs. The lowest one is given by the simplest functional, the LDA, which depends on the density alone. The next rung, corresponding to the more accurate GGA functionals, is reached by using not only the density but also its gradient as ingredients in the construction of the functional. A yet higher rung is reached by also making use of the Laplacian of the density. Within this philosophy, each additional ingredient should lead to a better functional. By employing the orbitals themselves, the set of ingredients is extended enormously, with corresponding high hopes for a far more accurate exchange-correlation approximation.

Third, functionals that are orbital dependent can be used to compensate for self-interaction on a per orbital basis (Perdew and Zunger, 1981). Moreover, because the addition of an infinitesimal charge can occupy a previously empty orbital, orbital-dependent functionals possess a natural mechanism that allows for derivative discontinuities, i.e., particle number discontinuities in the exchange-correlation potential. Thus, both major deficiencies emphasized above can be eliminated, or at least greatly reduced.

Last, but far from least, a fundamental piece of many-electron physics, namely, the exchange energy, is formulated naturally in terms of orbitals. It is therefore hoped that, in addition to the Hartree term, the exchange term will also be an explicit ingredient in the total energy expression of Eq. (5). Any approximation would then be only in the correlation term. This would be very important, given that an unsatisfactory description of the exchange energy was previously pointed out as a major source of errors in explicit density functionals (Gunnarson and Jones, 1985). This is most easily seen for the H_2^+ molecule discussed above, which having only one electron is correlation-free. In this case, an exact-exchange term in place of E_{xc} in Eq. (5) leads to the exact energy functional, removing both self-interaction and derivative discontinuity errors in one fell swoop.

This review article is dedicated to a survey of the successes, pitfalls, and remaining challenges in the development and applications of such orbital-dependent exchange-correlation functionals. It is arranged as follows. In the next section, we discuss the difficulties associated with determining a Kohn-Sham potential consistent with an exchange-correlation functional that is an explicit functional of the Kohn-Sham orbitals and thus only an implicit functional of the density. Next, we review the principles, successes, and failures of four classes of orbital-dependent functionals. Last, we discuss the role of orbital-dependent functionals in the computation of excited-state properties in general and in time-dependent DFT in particular.

II. THE OPTIMIZED EFFECTIVE POTENTIAL

A. Optimized effective potential formalism

The above discussion suggests that using the Kohn-Sham orbitals explicitly in the expression for E_{xc} is an attractive idea. However, in attempting to use such functionals, one is immediately confronted with a difficulty. How does one actually evaluate the functional derivative of Eq. (11), which is needed to obtain $v_{xc,\sigma}(\vec{r})$? This is a trivial task for explicit functionals, but what should one do in the absence of explicit expressions for either the dependence of E_{xc} on the density or the mapping between orbitals and density?

The first step toward solving this problem was taken even before DFT in its modern form had come into existence. Sharp and Horton (1953) minimized a Hartree-Fock-like expression for the total energy under the constraint that the potential appearing in the single-particle equations be a *local* one. They found an integral equation for the potential and interpreted their result as defining an approximation to Hartree-Fock theory. Therefore, they called their potential the optimized effective potential (OEP).

It took 23 years (Talman and Shadwick, 1976) for the OEP equation to be solved numerically for the first time, and it took another six years to conclude that the Sharp-Horton potential is actually the exact-exchange potential of Kohn-Sham theory (Sahni *et al.*, 1982). From this perspective, "Optimized Effective Potential," or the synonymously used term "Optimized Potential Method" may appear as somewhat superfluous special names. However, it is instructive to convince oneself that the Sharp-Horton condition

$$\delta E_{\rm tot}/\delta v_{\rm KS}(\vec{r}) = 0 \tag{16}$$

is equivalent to the Hohenberg-Kohn variational principle (Sahni *et al.*, 1982). By virtue of the chain rule for functional derivatives,

$$0 = \frac{\delta E_{\rm tot}}{\delta v_{\rm KS}(\vec{r})} = \int \frac{\delta E_{\rm tot}}{\delta n(\vec{r}\,')} \frac{\delta n(\vec{r}\,')}{\delta v_{\rm KS}(\vec{r})} d^3r'. \tag{17}$$

The second factor under the integral is the static Kohn-Sham (i.e., noninteracting) response function

$$\chi_{\rm KS}(\vec{r},\vec{r}\,') = \delta n(\vec{r})/\delta v_{\rm KS}(\vec{r}\,'). \tag{18}$$

Multiplying by $\chi_{\rm KS}^{-1}$ and integrating, i.e., acting with the inverse response operator on Eq. (17), yields

$$0 = \delta E_{\rm tot}/\delta n(\vec{r}), \tag{19}$$

which is the Hohenberg-Kohn variational principle. Note that here, unlike in Eq. (15), the functional derivative is defined on the domain of densities with fixed particle number and, therefore, Eqs. (15) and (19) are consistent with each other (Perdew *et al.*, 1982).

Today, the term OEP is used in a broad sense to denote generally the exchange-correlation potential of any orbital-dependent functional. Derivations of the OEP equation that directly make use of Eq. (11), and, in that sense, are in the spirit of modern DFT, were published in the early 1990s (Shaginyan, 1993; Görling and Levy, 1994) elaborated in later works, e.g., by Grabo and Gross (1995) and Grabo *et al.* (1997, 2000) and by Engel (2003) and recently extended to noncollinear spin-polarized DFT (Sharma *et al.*, 2007) and current-spin-density-functional theory (Pittalis *et al.*, 2006; Rohra and Görling, 2006). The central idea behind the derivation of the (standard) OEP equation is that (11) can be rewritten by applying the chain rule for functional derivatives twice,

$$\begin{aligned} v_{{\rm xc},\sigma}(\vec{r}) &= \frac{\delta E_{\rm xc}[\{\varphi_{j,\tau}\}]}{\delta n_\sigma(\vec{r})} \\ &= \sum_{\alpha=\uparrow,\downarrow} \sum_{i=1}^{N_\alpha} \int \frac{\delta E_{\rm xc}[\{\varphi_{j,\tau}\}]}{\delta \varphi_{i,\alpha}(\vec{r}\,')} \frac{\delta \varphi_{i,\alpha}(\vec{r}\,')}{\delta n_\sigma(\vec{r})} d^3r' + \text{c.c.} \\ &= \sum_{\alpha,\beta=\uparrow,\downarrow} \sum_{i=1}^{N_\sigma} \int\!\!\int \frac{\delta E_{\rm xc}[\{\varphi_{j,\tau}\}]}{\delta \varphi_{i,\alpha}(\vec{r}\,')} \frac{\delta \varphi_{i,\alpha}(\vec{r}\,')}{\delta v_{{\rm KS},\beta}(\vec{r}\,'')} \\ &\quad\times \frac{\delta v_{{\rm KS},\beta}(\vec{r}\,'')}{\delta n_\sigma(\vec{r})} d^3r'\, d^3r'' + \text{c.c.}, \end{aligned} \tag{20}$$

where greek indices denote spin polarization and c.c. denotes the complex conjugate of the preceding term. (Note that, in order to keep the notation as general as possible, we allow for complex orbitals.) The second chain rule is required because obviously one cannot evaluate $\delta\varphi_{i,\alpha}(\vec{r}\,')/\delta n_\sigma(\vec{r})$ by calculating $\delta n_\sigma(\vec{r})/\delta\varphi_{i,\alpha}(\vec{r}\,') = \varphi_{i,\alpha}^*(\vec{r}\,')$ and taking the inverse.

Equation (20) can be simplified by noting that the third factor under the integral is the inverse $\chi_{{\rm KS},\sigma}^{-1}(\vec{r},\vec{r}\,')$ of (the spin-dependent version of) Eq. (18),

$$\chi_{{\rm KS},\sigma\beta}(\vec{r},\vec{r}\,') = \delta_{\sigma,\beta} \frac{\delta n_\sigma(\vec{r})}{\delta v_{{\rm KS},\beta}(\vec{r}\,')}. \tag{21}$$

The index β on the left-hand side of Eq. (21) is omitted henceforth because the noninteracting response function is diagonal with respect to the spin index. The second factor under the integral in Eq. (20) can be evaluated by first-order perturbation theory,

$$\frac{\delta \varphi_{i,\alpha}(\vec{r}\,')}{\delta v_{{\rm KS},\sigma}(\vec{r})} = \delta_{\alpha,\sigma} \sum_{\substack{j=1 \\ j\neq i}}^{\infty} \frac{\varphi_{j,\sigma}(\vec{r}\,')\varphi_{j,\sigma}^*(\vec{r})}{\varepsilon_{i,\sigma} - \varepsilon_{j,\sigma}} \varphi_{i,\sigma}(\vec{r}). \tag{22}$$

The same argument can be used to write the response function as

$$\chi_{{\rm KS},\sigma}(\vec{r},\vec{r}\,') = \sum_{k=1}^{N_\sigma} \sum_{\substack{j=1 \\ j\neq i}}^{\infty} \frac{\varphi_{k,\sigma}^*(\vec{r})\varphi_{k,\sigma}(\vec{r}\,')\varphi_{j,\sigma}(\vec{r})\varphi_{j,\sigma}^*(\vec{r}\,')}{\varepsilon_{k,\sigma} - \varepsilon_{j,\sigma}} + \text{c.c.} \tag{23}$$

Finally, the functional derivative $\delta E_{\rm xc}[\{\varphi_{j,\tau}\}]/\delta\varphi_{i,\sigma}(\vec{r}\,')$ can be evaluated straightforwardly for a given orbital-dependent functional $E_{\rm xc}[\{\varphi_{j,\sigma}\}]$. Thus, by multiplying Eq. (20) with $\chi_{{\rm KS},\sigma}(\vec{r}\,',\vec{r})$ from the right, integrating over $\vec{r}\,'$, and making use of Eqs. (21)–(23), one obtains the OEP equation in its celebrated form,

$$\begin{aligned} &\sum_{i=1}^{N_\sigma} \int \varphi_{i,\sigma}^*(\vec{r}\,')[v_{{\rm xc},\sigma}(\vec{r}\,') - u_{{\rm xc}i,\sigma}(\vec{r}\,')] \\ &\qquad\times G_{{\rm KS}i,\sigma}(\vec{r}\,',\vec{r})\varphi_{i,\sigma}(\vec{r}) d^3r' + \text{c.c.} = 0, \end{aligned} \tag{24}$$

where $G_{{\rm KS}i,\sigma}(\vec{r}\,',\vec{r})$ denotes the Kohn-Sham Green's function,

$$G_{{\rm KS}i,\sigma}(\vec{r}\,',\vec{r}) = \sum_{\substack{j=1 \\ j\neq i}}^{\infty} \frac{\varphi_{j,\sigma}(\vec{r}\,')\varphi_{j,\sigma}^*(\vec{r})}{\varepsilon_{i,\sigma} - \varepsilon_{j,\sigma}} \tag{25}$$

and

$$u_{{\rm xc}i,\sigma}(\vec{r}) = \frac{1}{\varphi_{i,\sigma}^*(\vec{r})} \frac{\delta E_{\rm xc}[\{\varphi_{j,\tau}\}]}{\delta \varphi_{i,\sigma}(\vec{r})}. \tag{26}$$

An alternative derivation of the OEP equation from a somewhat different point of view is based on the observation [first made by Krieger *et al.* (1992b) and further explained by Grabo *et al.* (1997, 2000)] that the OEP integral equation (24) can formally be written in the simple form

$$\sum_{i=1}^{N_\sigma} \psi_{i,\sigma}^*(\vec{r})\varphi_{i,\sigma}(\vec{r}) + \text{c.c.} = 0, \tag{27}$$

with the orbital shift

$$\psi^*_{i,\sigma}(\vec{r}) = -\sum_{\substack{j=1\\ j\neq i}}^{\infty} \frac{\langle \varphi_{i\sigma}|u_{\mathrm{xc}i,\sigma} - v_{\mathrm{xc},\sigma}|\varphi_{j\sigma}\rangle}{\varepsilon_{i,\sigma} - \varepsilon_{j,\sigma}} \varphi^*_{j,\sigma}(\vec{r}). \tag{28}$$

(Krieger *et al.* did not use the ψ_i defined above, but the diverging functions $p_i = -\psi_i/\varphi_i$.) Obviously, $\psi^*_{i,\sigma}(\vec{r})$ has a structure well known from first-order perturbation theory: It is the (negative of the) first-order change that the ith Kohn-Sham orbital undergoes when the Kohn-Sham potential in Eq. (3) is replaced by $u_{\mathrm{xc}i,\sigma}(\vec{r})$. As elaborated in Sec. II.B, $u_{\mathrm{xc}i,\sigma}(\vec{r})$ can be interpreted as an orbital-specific potential corresponding to a given orbital-dependent expression for the exchange-correlation energy. The OEP equation states that to first order the density does not change when the Kohn-Sham equations are subjected to perturbations $\Delta v_{i,\sigma}(\vec{r}) = u_{\mathrm{xc}i,\sigma}(\vec{r}) - v_{\mathrm{xc},\sigma}(\vec{r})$. Thus, an alternative derivation (Kümmel and Perdew, 2003b) of the OEP equation (27) can be based directly on perturbation theory arguments, making use of the fact that the Kohn-Sham energy functional and the energy functional defined using the orbital-specific potentials (cf. Sec. II.B) agree to first order.

At first sight, this second derivation may be thought of as having little merit apart from the fact that its arguments stay closer to the spirit of DFT because they focus on the density. However, the perturbation theory arguments link OEP theory to theories using orbital-specific potentials (see Sec. II.B), and the orbital shifts $\psi^*_{i,\sigma}(\vec{r})$ play a prominent role in computational schemes that are used for actual OEP calculations (see Sec. II.D). Therefore, this alternative perspective on the OEP equation is also of practical importance.

The usefulness of the orbital shifts defined in Eq. (28) is made clear by the following line of argument (Krieger *et al.*, 1992b; Grabo *et al.*, 1997, 2000). Because the Kohn-Sham orbitals form a complete set, i.e.,

$$\delta(\vec{r} - \vec{r}\,') = \sum_{j=1}^{\infty} \varphi_{j,\sigma}(\vec{r}\,')\varphi^*_{j,\sigma}(\vec{r}), \tag{29}$$

straightforward evaluation shows that $G_{\mathrm{KS}i,\sigma}(\vec{r}\,',\vec{r})$ satisfies

$$[h_{\mathrm{KS},\sigma}(\vec{r}) - \varepsilon_{i,\sigma}]G_{\mathrm{KS}i,\sigma}(\vec{r}\,',\vec{r}) = -[\delta(\vec{r}\,' - \vec{r}) - \varphi_{i,\sigma}(\vec{r}\,')\varphi^*_{i,\sigma}(\vec{r})], \tag{30}$$

where $h_{\mathrm{KS},\sigma}(\vec{r})$ is the Kohn-Sham Hamiltonian, i.e., the operator on the left-hand side of Eq. (3). Thus, rewriting the right-hand side of Eq. (28) in terms of the Green's function Eq. (25) and then applying the operator $[h_{\mathrm{KS},\sigma}(\vec{r}) - \varepsilon_{i,\sigma}]$ to both sides, one obtains

$$\begin{aligned}[h_{\mathrm{KS},\sigma}(\vec{r}) - \varepsilon_{i,\sigma}]\psi^*_{i,\sigma}(\vec{r}) = -[&v_{\mathrm{xc},\sigma}(\vec{r}) - u_{\mathrm{xc}i,\sigma}(\vec{r}) \\ &- (\bar{v}_{\mathrm{xc}i,\sigma} - \bar{u}_{\mathrm{xc}i,\sigma})]\varphi^*_{i,\sigma}(\vec{r}),\end{aligned} \tag{31}$$

where

$$\bar{v}_{\mathrm{xc}i,\sigma} = \int \varphi^*_{i,\sigma}(\vec{r})v_{\mathrm{xc},\sigma}(\vec{r})\varphi_{i,\sigma}(\vec{r})d^3r \tag{32}$$

and

$$\bar{u}_{\mathrm{xc}i,\sigma} = \int \varphi^*_{i,\sigma}(\vec{r})u_{\mathrm{xc}i,\sigma}(\vec{r})\varphi_{i,\sigma}(\vec{r})d^3r \tag{33}$$

are orbital averages. Equation (31) can also be obtained as a direct consequence of the perturbation theory arguments mentioned above (Kümmel and Perdew, 2003b).

The orbital shifts satisfy partial differential equations that look like Kohn-Sham equations with an additional inhomogeneity on the right-hand side. From the definition of $\psi^*_{i,\sigma}$, it follows that each orbital is orthogonal to its orbital shift, $\int\psi^*_{i,\sigma}(\vec{r})\varphi_{i,\sigma}(\vec{r})d^3r = 0$, or, in the case of degeneracies, $\int\psi^*_{i,\sigma}(\vec{r})\varphi_{j,\sigma}(\vec{r})d^3r = 0$ for all pairs (i,j) with $\varepsilon_{i,\sigma} = \varepsilon_{j,\sigma}$. Together with appropriate boundary conditions [e.g., for finite systems $\psi_{i,\sigma}(\vec{r})$ must vanish for large distances from the system's center], Eq. (31) uniquely determines the orbital shifts. Solving Eq. (31) for $v_{\mathrm{KS},\sigma}(\vec{r})\psi^*_{i,\sigma}(\vec{r})$, inserting the result into Eq. (27) multiplied by $v_{\mathrm{KS},\sigma}(\vec{r})$, and solving for $v_{\mathrm{xc},\sigma}(\vec{r})$ yields

$$\begin{aligned} v_{\mathrm{xc},\sigma}(\vec{r}) = \frac{1}{2n_\sigma(\vec{r})}\sum_{i=1}^{N_\sigma}\Bigg\{&|\varphi_{i,\sigma}(\vec{r})|^2[u_{\mathrm{xc}i,\sigma}(\vec{r}) + (\bar{v}_{\mathrm{xc}i,\sigma} - \bar{u}_{\mathrm{xc}i,\sigma})] \\ &+ \varphi_{i,\sigma}(\vec{r})\left[\left(\frac{\hbar^2}{2m}\nabla^2 + \varepsilon_{i,\sigma}\right)\psi^*_{i,\sigma}(\vec{r})\right]\Bigg\} + \mathrm{c.c.} \end{aligned} \tag{34}$$

Using the Kohn-Sham equation to replace $\varepsilon_{i,\sigma}\psi^*_{i,\sigma}(\vec{r})$ in the last term on the right-hand side of Eq. (34), followed by using Eq. (27) to get rid of the ensuing v_{KS}-containing term and $\vec{\nabla}\cdot\{\vec{\nabla}[\sum_{i=1}^{N_\sigma}\varphi_{i,\sigma}(\vec{r})\psi^*_{i,\sigma}(\vec{r}) + \mathrm{c.c.}]\} = 0$, Eq. (34) simplifies further to

$$\begin{aligned} v_{\mathrm{xc},\sigma}(\vec{r}) = \frac{1}{2n_\sigma(\vec{r})}\sum_{i=1}^{N_\sigma}\Bigg\{&|\varphi_{i,\sigma}(\vec{r})|^2[u_{\mathrm{xc}i,\sigma}(\vec{r}) + (\bar{v}_{\mathrm{xc}i,\sigma} - \bar{u}_{\mathrm{xc}i,\sigma})] \\ &- \frac{\hbar^2}{m}\vec{\nabla}\cdot[\psi^*_{i,\sigma}(\vec{r})\vec{\nabla}\varphi_{i,\sigma}(\vec{r})]\Bigg\} + \mathrm{c.c.} \end{aligned} \tag{35}$$

At first sight, this expression does not look simpler than Eq. (24) or Eq. (27). It is still an integral equation because of the $\bar{v}_{\mathrm{xc}i,\sigma}$ terms appearing on the right-hand side. The complications that arise due to the sum over all orbitals, which enter into Eq. (24) via $G_{\mathrm{KS}i,\sigma}(\vec{r}\,',\vec{r})$, are also still there. They are simply "buried" in the orbital-shift terms. However, as elaborated below, the OEP equation in its rewritten form, Eq. (35), is of great practical usefulness for deriving approximations to the OEP as well as for numerical solution schemes.

Exchange-correlation potentials determined from Eq. (35) are not just harder to obtain mathematically than those determined from explicit functionals. They also possess a peculiar feature in their long-range asymptotic behavior. Because any potential, including $v_{\mathrm{xc},\sigma}(\vec{r})$, is fixed only up to an additive constant, one is at liberty to choose one of the constants $\bar{v}_{\mathrm{xc}i,\sigma} - \bar{u}_{\mathrm{xc}i,\sigma}$ freely. A reasonable choice is to make the potential vanish at infinity. Because in a local potential each orbital decays exponentially, with the exponent being proportional to the orbital's corresponding eigenvalue, the density and sum over orbitals on the right-hand side of Eq. (35) are dominated in the long-range limit by the term corresponding

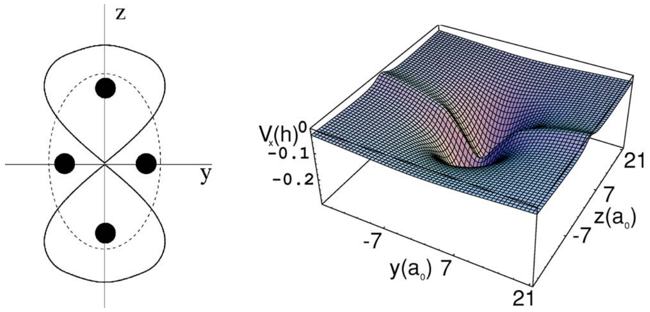

FIG. 3. (Color online) Evidence for a counterintuitive property of the exchange-only OEP. Left: Schematic depiction of the nuclear positions and the structure of the two valence orbitals in a cluster of four sodium atoms. In the pseudopotential approximation (see Sec. II.E) used here, the lower and higher orbitals are of s and p character, respectively. Right: Exchange-only OEP potential landscape for this cluster in the y-z plane, showing $v_x(\vec{r})$ (in hartrees) as the height at each spatial point (y,z) for a fixed $x=0$. The pronounced ridge in the OEP potential reflects the nonvanishing asymptotic constant on the nodal plane of the p orbital.

to the highest occupied orbital. (For a detailed discussion of the asymptotic behavior, see, e.g., Kreibich *et al.*, 1999; Della Sala and Görling, 2002a.) If $u_{\mathrm{xc}N_\sigma,\sigma}(\vec{r})$ vanishes asymptotically—a condition that is satisfied for all functionals for which the OEP has been constructed so far—then the value of $v_{\mathrm{xc},\sigma}(\vec{r})$ for large distances will be the difference of $\bar{v}_{\mathrm{xc}N_\sigma,\sigma}$ and $\bar{u}_{\mathrm{xc}N_\sigma,\sigma}$. Thus, the condition that the potential vanish asymptotically requires that

$$\bar{v}_{\mathrm{xc}N_\sigma,\sigma} - \bar{u}_{\mathrm{xc}N_\sigma,\sigma} = 0. \tag{36}$$

However, the highest occupied orbitals in open-shell systems typically have nodal surfaces extending to infinity. On such a nodal surface, the density is never dominated by the highest occupied orbital. Thus, the arguments of the preceding paragraph apply to all regions of space except the nodal surfaces of the highest occupied orbital. On the nodal surface (which is a set of zero measure), the potential asymptotically goes to $\bar{v}_{\mathrm{xc}N_\sigma-1,\sigma} - \bar{u}_{\mathrm{xc}N_\sigma-1,\sigma}$, where $N_\sigma-1$ is the index of the highest occupied orbital that does not vanish on the nodal surface. As a consequence, the exchange-correlation potential can tend to a nonzero value on the nodal surface even if Eq. (36) is satisfied (Della Sala and Görling, 2002a). In other words, the Kohn-Sham potential can go to different asymptotic values in different regions of space. This has been confirmed numerically for the bare exchange potential in both the localized Hartree-Fock approximation (see Sec. II.C) to the OEP (Della Sala and Görling, 2002a) and in full OEP calculations (Kümmel and Perdew, 2003b).

Figure 3 shows an example of how the nonvanishing asymptotic constant manifests itself in an actual calculation. A neutral cluster of four sodium atoms in its ground state has two valence orbitals, each occupied by two electrons. The lower valence orbital is of s character; the higher one is the highest occupied one and is of p character. Therefore, it has a nodal plane extending to infinity. In the figure, this is the x-y plane. If one looks at the potential in the y-z plane, the potential asymptotically tends to zero everywhere except for points on the y axis. On this axis, the potential takes a nonvanishing, positive asymptotic value, which appears as a ridge in the potential landscape. This ridge can have quite a noticeable influence on Kohn-Sham eigenvalues (Della Sala and Görling, 2002a, 2002b), in particular unoccupied ones, and thus may influence, e.g., the calculation of excited states in time-dependent DFT.

One may be tempted to call this counterintuitive feature "nonphysical." One may also argue that, so far, the nonvanishing constants have been shown to exist only in exchange-only calculations and that inclusion of correlation may change the picture. However, based on the form of the presently used self-energy expressions, Della Sala and Görling (2002b) have argued that the sum of the exchange and correlation potentials will also show nonvanishing asymptotic constants. Furthermore, one can argue on general grounds that the Kohn-Sham potential leads to a physical observable—the density—and to quantities that can physically be interpreted in an approximate way—the Kohn-Sham eigenvalues. However, it is not an observable itself. It is a mathematical construct tailored to capture many-body physics in a single-particle picture. Therefore, it can have features that a physical potential cannot have (Kümmel and Perdew, 2003b).

B. Kohn-Sham versus orbital-specific potentials

The preceding section has shown that, from a mathematical point of view, orbital-dependent functionals are well within Kohn-Sham theory. However, from a practical point of view one may doubt the utility of the concept, as employing the OEP requires considerable additional effort compared to the use of explicit density functionals like those obtained within the L(S)DA or GGA framework. It is therefore a natural idea to try to avoid the need for the OEP. This is especially so given that this goal appears to be readily achieved naturally by minimizing the total energy with respect to the *orbitals* rather than the density. We elucidate this by an example.

The exact exchange energy is defined in terms of the Fock integral,

$$E_x = -\frac{e^2}{2}\sum_{\sigma=\uparrow,\downarrow}\sum_{i,j=1}^{N_\sigma}\int\int\frac{\varphi_{i,\sigma}^*(\vec{r})\varphi_{j,\sigma}^*(\vec{r}\,')\varphi_{j,\sigma}(\vec{r})\varphi_{i,\sigma}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} \times d^3r'\,d^3r. \tag{37}$$

In Hartree-Fock theory, this integral is one part of the ground-state energy. The latter is found by unconstrained minimization of the total energy with respect to a Slater determinant built from single-particle orbitals. The minimizing orbitals are obtained by self-consistently solving a set of single-particle equations in which the potential is the orbital-specific integral operator corresponding to Eq. (26). Inserting the Hartree-Fock orbitals into Eq. (37) yields the Hartree-Fock exchange energy and defines exact exchange in wave-function-based approaches.

In Kohn-Sham theory, exact exchange is defined differently. As discussed above Eq. (7), the Kohn-Sham potential must be local for the Hohenberg-Kohn theorem to apply to the system of noninteracting electrons. Without this locality, the Kohn-Sham framework breaks down. Therefore, the natural definition of the exchange energy within the Kohn-Sham framework also uses Eq. (37), but with a different set of orbitals. The orbitals used to evaluate the expression must by definition be Kohn-Sham orbitals, i.e., they must stem from a local, multiplicative potential that is the same for all orbitals. Since these orbitals are, in general, different from the Hartree-Fock orbitals, the exact exchange energy in Kohn-Sham theory is different from the Hartree-Fock exchange energy. Consequently, the correlation energy, defined as the difference between the total energy (which is the same in all exact theories) and the total energy in the exchange-only approximation, must also be different in the two approaches. See Gross, Petersilka, and Grabo (1996) for a detailed discussion, and note that the exact exchange energy of Kohn-Sham theory is often defined by evaluating the Fock integral of Eq. (37) with the Kohn-Sham orbitals yielding the exact density. Both definitions of exchange and correlation are consistent and rigorously defined concepts. They do not contradict each other because the value of the *total* energy is the same in both. In a nutshell, an orbital-dependent expression for the exchange-correlation energy is well defined only if one specifies exactly what orbitals are to be used to evaluate the expression in addition to specifying the expression itself.

It is important to note that analogs of the procedures outlined above for the exchange energy can be defined for any (approximate) orbital-dependent expression for the exchange-correlation energy: For any given orbital-dependent expression, one can define two functionals. The first is obtained by direct variation with respect to single-particle orbitals. It leads, in general, to single-particle equations in which the potential is different for each orbital and can be nonmultiplicative. The energy obtained in this way is the lowest possible one. In contrast, the Kohn-Sham potential corresponding to an orbital-dependent expression for E_{xc} is a local multiplicative potential. It defines a different, second functional: the Kohn-Sham functional.

In general, this difference has important consequences. By virtue of the Hohenberg-Kohn theorem, the Kohn-Sham orbitals are functionals of the density (see Sec. I). This is not only an abstract statement—numerical procedures for constructing the Kohn-Sham orbitals from a given density have been developed (Aryasetiawan and Stott, 1988; Görling, 1992; Wang and Parr, 1993; Umrigar and Gonze, 1994; van Leeuwen and Baerends, 1994; Zhao *et al.*, 1994; Peirs *et al.*, 2003). Thus, for a given density, the Kohn-Sham orbitals are uniquely fixed and any functional depending on them is a density functional. But the orbitals that are obtained from the equations with the orbital-specific potentials are not uniquely fixed by the density: To find them, one must additionally specify an orbital-dependent expression for the energy that is to be minimized. Therefore, expressions depending on the latter type of orbital in general, and the expression for the noninteracting kinetic energy T_s [cf. Eq. (5)] in particular, are not uniquely defined by the density in the orbital-specific framework. The Kohn-Sham theorem, which provides a safe mathematical foundation as outlined in Sec. I, does not apply when one is using orbital-specific potentials.

For completeness, we note that the above differences have led to some confusion and erroneous statements in the literature, e.g., by Nesbet (1998, 2001, 2004). These articles should be seen in light of subsequent publications (e.g., Gal, 2000; Holas and March, 2001, 2002; Kümmel and Perdew, 2003b; Lindgren and Salomonson, 2003; Liu and Ayers, 2004).

Despite possible formal difficulties associated with orbital-specific potentials, the idea of using them is actually as old as DFT itself (Kohn and Sham, 1965). For self-interaction-corrected functionals (see Sec. III.B), it was proposed by Perdew and Zunger (1981); for hybrid functionals (cf. Sec. III.C) by Becke (1993a, 1993b); for meta-GGA functionals (cf. Sec. III.A) by Neumann *et al.* (1996); etc. From a practical point of view, the important question is whether using orbital-specific potentials or the Kohn-Sham potential makes a physically meaningful difference in actual calculations.

The answer depends on which quantities are of interest. Generally speaking, there is hope that ground-state total energies and total densities can be expected to be similar in the two approaches. This statement is based on the observation that the corresponding energy functionals agree to first order in the density, as discussed in Sec. II.A in the context of Eq. (27). A more rigorous statement is that, when the second- and higher-order contributions in the perturbation expansion discussed in the paragraph following Eqs. (27) and (28) are small, the total energies and densities obtained from the two approaches are expected to be similar (Kümmel and Perdew, 2003b). For the exchange-only case, this statement has also been discussed by Ivanov and Levy (2003) and numerous calculations have confirmed the closeness of the energies obtained numerically (see, e.g., Talman and Shadwick, 1976; Krieger *et al.*, 1992a; Görling, 1999). Therefore, many ground-state properties are very similar in OEP and orbital-dependent potential calculations if the orbital-dependent component of the employed functional is just the Fock integral. However, one can also find quantities where clear differences are seen, as discussed below.

Whether the closeness of ground-state energies that has been observed in the comparison of Hartree-Fock and exchange-only OEP calculations will also hold for correlation functionals with possibly rather involved dependencies on the orbitals is an open question, as there is no rigorous way to estimate whether agreement to first order implies sufficiently close agreement. The OEP approach is the conceptually rigorous way to stay within the Kohn-Sham scheme. But since the orbital-specific potentials can be easier to handle computationally, can have advantages for the description of localized states, as discussed above and below, and can still be seen as protected by the Hohenberg-Kohn theorem [see, e.g., the discussions by Levy (1979) and Perdew and Zunger (1981)], they are frequently employed. How large the differences between the two approaches are should be checked for each new class of orbital-dependent functionals by comparing the respective results for a number of characteristic systems.

Turning to quantities other than the total energy, one classical example of profound differences between the orbital-specific potential and the Kohn-Sham approach is the unoccupied eigenvalues. Referring once more to the prominent example of exchange-only theory, it is well known that only very few, if any, bound unoccupied states are found in Hartree-Fock calculations for finite systems. On the other hand, the exchange-only Kohn-Sham potential leads to a Rydberg series and, therefore, yields a single-particle spectrum that is accessible to a much larger extent to physical interpretation (see, e.g., Grabo *et al.*, 2000, for a detailed discussion). Here it is the eigenvalues and orbitals coming from the local Kohn-Sham potential that lend themselves to physical interpretation, and not the ones from orbital-specific potentials. Another example of the importance of the choice of potential is afforded by the calculation of nuclear magnetic resonance (NMR) shielding constants, chemical shifts, magnetizabilities, and rotational g tensors of main group molecules. Using several hybrid functionals, discussed in Sec. III.C (Wilson and Tozer, 2001; Arbuznikov and Kaupp, 2004; Hieringer *et al.*, 2004; Lutnaes *et al.*, 2006; Teale *et al.*, 2007), as well as a meta-GGA functional, discussed in Sec. III.A (Arbuznikov and Kaupp, 2003), it was consistently found that results obtained with a local potential are a significant improvement over those obtained with orbital-specific potentials. However, there are also examples in which the use of orbital-specific potentials can be beneficial. As pointed out by Rinke (2007), it is not unusual that the Hartree-Fock and exchange only OEP eigenvalues obtained for d electrons in solids are quite different. For example, the Hartree-Fock d-electron binding energy in ZnO is 7.8 eV (Fuchs *et al.*, 2007), whereas the corresponding exchange-only OEP energy is 5.12 eV (Rinke *et al.*, 2005). But here it is the eigenvalue from the orbital-specific potentials, i.e., Hartree-Fock, that is in better agreement with available experimental data. It seems that, when localization phenomena play a role, orbital-specific potentials are advantageous. Other examples of this are discussed in greater detail in the context of the self-interaction correction (see Sec. III.B).

On a more technical level, further differences between orbital-specific and local potentials can be found. Some of them can be seen as advantages of the orbital-specific approach and some as advantages of the OEP. For example, a major advantage of the orbital-specific potential approach is simplicity, in the sense that only the $u_{\mathrm{xc}i,\sigma}$ need to be calculated, without further worries about the OEP equation. However, orbital-specific potentials can have the disadvantage of leading to nonorthogonality of the orbitals, thus requiring the use of off-diagonal Lagrangian multipliers to enforce orbital orthogonality (Perdew and Zunger, 1981; Harrison *et al.*, 1983; Svane and Gunnarsson, 1990; Goedecker and Umrigar, 1997). In terms of understanding computational results, a local potential can be advantageous because it can easily be plotted, analyzed, and used to gain a deeper understanding of the observed effects (see, e.g., van Gisbergen *et al.*, 1999; Kümmel *et al.*, 2004). Also, it is easier to take advantage of parallel programming techniques when the potential is local (Stathopoulos *et al.*, 2000; Kronik *et al.*, 2006). To propagate the Kohn-Sham orbitals in time as in nonperturbative time-dependent calculations (see Sec. IV.B.2), a local potential can have advantages because it allows for more efficient propagator splitting schemes. Because the pros and cons of the two approaches are typically weighted differently for different E_{xc} functionals, we discuss typical potential choices further in the context of each of the functionals in Sec. III.

The discussion of Kohn-Sham versus orbital-specific potentials has been presented so far entirely in black and white—either one did an OEP calculation and was within the Kohn-Sham framework or one did not and was not. Before concluding this discussion, it is important to paint a few shades of gray. As discussed in Sec. I.A, the essence of the Kohn-Sham scheme is in finding a noninteracting model system, where a single Slater determinant yields the same ground-state density as that of the real system. Seidl *et al.* (1996) pointed out that one may also attempt to map the real system to an *interacting* model system that can still be represented by a single Slater determinant. This may be beneficial because a judiciously chosen interacting model system may already include many of the exchange and correlation effects, resulting in a smaller energy component that still needs to be approximated.

Such an alternate mapping is achieved in practice by defining an energy functional of the orbitals that comprise the Slater determinant $S[\{\varphi_i\}]$, and defining an associated energy density functional $F^s[n]$, obtained from the Slater determinant that minimizes $S[\cdot]$ while yielding the density n, i.e.,

$$F^s[n] \equiv \min_{\{\varphi_i\}\to n(\vec{r})} S[\{\varphi_i\}]. \tag{38}$$

The minimizing orbitals $\{\varphi_i\}$ comprise the model system and play a role similar to those of Kohn-Sham orbitals. One can then recast the Hohenberg-Kohn total energy expression of Eq. (2) in the form

$$E_{\mathrm{tot}} = \int v_{\mathrm{ext}}(\vec{r}) n(\vec{r}) d^3r + F^s[n] + R^s[n], \tag{39}$$

where $R^s[n]$ is a remainder energy functional. In a manner entirely analogous to the transition from Eq. (5) to the Kohn-Sham equation, one obtains the generalized Kohn-Sham equation (Seidl *et al.*, 1996)

$$(\hat{O}^S[\{\varphi_i\}] + v_{\mathrm{ext}}(\vec{r}) + v_R(\vec{r}))\varphi_j(\vec{r}) = \epsilon_j\varphi_j(\vec{r}), \tag{40}$$

where

$$v_R(\vec{r}) \equiv \frac{\delta R^s[n]}{\delta n(\vec{r})} \tag{41}$$

is the remainder potential and $\hat{O}^S[\{\varphi_i\}]$ is a generally orbital-specific operator that depends on the choice of $S[\cdot]$, but not on $v_{\mathrm{ext}}(\vec{r})$ or $v_R(\vec{r})$.

Two special cases help in clarifying the physical content of the generalized Kohn-Sham equation (Seidl *et al.*, 1996). If $S[\{\varphi_i\}]$ is chosen as the Slater-determinant expectation value of the kinetic energy operator [i.e., the first term on the right side of Eq. (5)], then $\hat{O}^S[\{\varphi_i\}]$ is the single-particle kinetic energy operator and Eq. (40) reduces to the usual Kohn-Sham equation. If, however, $S[\{\varphi_i\}]$ is chosen as the operator corresponding to the Slater-determinant expectation value of the sum of the kinetic energy and electron-repulsion energy operators, then $\hat{O}^S[\{\varphi_i\}]$ is the sum of the single-particle kinetic energy operator and the Hartree-Fock operator. Equation (40) then becomes a Hartree-Fock-Kohn-Sham equation. Unlike the original Hartree-Fock equation, it is in principle exact. The Hartree-Fock approximation is obtained from it by (crudely) approximating the remainder potential as zero. A third choice for $S[\{\varphi_i\}]$, the screened exchange approach, is discussed in Sec. IV.A.

The possibility of bringing the Hartree-Fock scheme (and other orbital-specific potentials) under the DFT umbrella within a generalized Kohn-Sham scheme may naively seem to make superfluous the distinction emphasized above between the OEP and orbital-specific potentials. This is not the case. Each choice of $S[\{\varphi_i\}]$ implies a different remainder energy functional $R^s[n]$, which is as unknown and likely to be as complicated as the Kohn-Sham exchange-correlation functional $E_{\mathrm{xc}}[n]$. It is important to note that $R^s[n]$ is not equal to the portion of $E_{\mathrm{xc}}[n]$ left outside the interacting model system. An obvious example is that $R^s[n]$ for the Hartree-Fock-Kohn-Sham scheme is not the same as the Kohn-Sham correlation energy $E_c[n]$. Practical orbital-dependent functionals, e.g., meta-GGA, self-interaction-corrected and hybrid functionals, to name a few of the approaches discussed in Sec. III, are all based on different approximations developed explicitly for $E_{\mathrm{xc}}[n]$, not for $R^s[n]$. Viewing the use of an orbital-specific potential as equivalent to employing the generalized Kohn-Sham equation does not remove the fact that one left the conventional Kohn-Sham world by not using the OEP. Instead, one merely recasts the problem in terms of differences between $R^s[n]$ and $E_{\mathrm{xc}}[n]$.

C. Approximate forms

In the preceding sections, we have explained why the OEP is attractive conceptually. However, actually solving the OEP equation is a highly nontrivial task. Inspection of Eq. (24) immediately reveals some of the complications. It is an integral equation involving the full set of occupied and unoccupied Kohn-Sham orbitals and eigenvalues, which has to be solved self-consistently with the Kohn-Sham equation, resulting in a complicated integro-differential set of equations. Since this is a formidable task, the first paper on the OEP by Sharp and Horton (1953) already proposed solving the equation in an approximate way. Following their reasoning, a series of approximations based on the Kohn-Sham Green's function, Eq. (25), has been developed over the years.

Sharp and Horton (1953) observed that, if the different eigenvalues $\varepsilon_{j,\sigma}$ appearing in the denominator of the sum in Eq. (25) are replaced by just one number $\tilde{\varepsilon}_\sigma$ (which in some vague way is interpreted as a mean value), then one obtains an approximate Green's function

$$\tilde{G}_{\mathrm{KS}i,\sigma}(\vec{r}\,',\vec{r}) = \frac{1}{\varepsilon_{i,\sigma} - \tilde{\varepsilon}_\sigma}\sum_{\substack{j=1\\ j\neq i}}^{\infty} \varphi_{j,\sigma}(\vec{r}\,')\varphi^*_{j,\sigma}(\vec{r}) \tag{42}$$

$$= \frac{1}{\varepsilon_{i,\sigma} - \tilde{\varepsilon}_\sigma}[\delta(\vec{r}-\vec{r}\,') - \varphi_{i,\sigma}(\vec{r}\,')\varphi^*_{i,\sigma}(\vec{r})]. \tag{43}$$

In the second step, the closure relation Eq. (29) has been used.

If one neglects the second term in the brackets involving $\varphi_{i,\sigma}$ and makes the further approximation that $\varepsilon_{i,\sigma} - \tilde{\varepsilon}_\sigma$ is replaced by yet another vaguely defined mean value $\Delta\varepsilon_\sigma$, then the resulting approximate Green's function reduces the OEP equation (24) to the simple form

$$\sum_{i=1}^{N_\sigma}\int \varphi^*_{i,\sigma}(\vec{r}\,')[v_{\mathrm{xc},\sigma}(\vec{r}\,') - u_{\mathrm{xc}i,\sigma}(\vec{r}\,')]\frac{\delta(\vec{r}-\vec{r}\,')}{\Delta\varepsilon_\sigma}\varphi_{i,\sigma}(\vec{r})d^3r' + \mathrm{c.c.} = 0. \tag{44}$$

This can readily be solved for $v_{\mathrm{xc},\sigma}$ by making use of Eq. (10). The result is an approximation that represents the OEP as an orbital-weighted average of the orbital-dependent potentials,

$$v^{\mathrm{Sl}}_{\mathrm{xc},\sigma}(\vec{r}) = \frac{1}{2n_\sigma(\vec{r})}\sum_{i=1}^{N_\sigma}|\varphi_{i,\sigma}(\vec{r})|^2[u_{\mathrm{xc}i,\sigma}(\vec{r}) + u^*_{\mathrm{xc}i,\sigma}(\vec{r})]. \tag{45}$$

This approximation is commonly referred to as the Slater potential (Slater, 1951). It is appealing because of its simplicity, but, since the terms that have been neglected turn out to be quite important, it is hardly ever used in calculations today. It should also be noted that, although the literature refers mostly to the exchange energy, the procedures outlined above and below are equally valid for any other orbital-dependent functional.

Krieger *et al.* (1990) were the first to point out that a better approximation than the Slater potential can be derived from Eq. (43) by making the same approximations for the energy denominators as above, but without neglecting the second term in the brackets on the right-hand side of Eq. (43). The calculation is analogous to the previous one with just one additional term. It yields what has become known as the Krieger-Li-Iafrate (KLI) approximation,

$$v^{\mathrm{KLI}}_{\mathrm{xc},\sigma}(\vec{r}) = \frac{1}{2n_\sigma(\vec{r})}\sum_{i=1}^{N_\sigma}|\varphi_{i,\sigma}(\vec{r})|^2[u_{\mathrm{xc}i,\sigma}(\vec{r}) + (\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}i,\sigma} - \bar{u}_{\mathrm{xc}i,\sigma})] + \mathrm{c.c.} \tag{46}$$

Obviously, Eq. (46) is still an integral equation because of the orbital-averaged terms $\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}i,\sigma}$ on its right-hand side. However, Krieger *et al.* (1990) pointed out that this integral equation can be solved with moderate effort. Multiplying Eq. (46) by $|\varphi_{j,\sigma}(\vec{r})|^2$ and integrating over all space yields

$$\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}j,\sigma} = \bar{v}^{\mathrm{SI}}_{\mathrm{xc}j,\sigma} + \sum_{i=1}^{N_\sigma-1} m_{ji,\sigma}[\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}i,\sigma} - \tfrac{1}{2}(\bar{u}_{\mathrm{xc}i,\sigma} + \bar{u}^{*}_{\mathrm{xc}i,\sigma})], \tag{47}$$

where

$$\bar{v}^{\mathrm{SI}}_{\mathrm{xc}j,\sigma} = \int \frac{|\varphi_{j,\sigma}(\vec{r})|^2}{2n_\sigma(\vec{r})}\sum_{i=1}^{N_\sigma}|\varphi_{i,\sigma}(\vec{r})|^2[u_{\mathrm{xc}i,\sigma}(\vec{r}) + u^{*}_{\mathrm{xc}i,\sigma}(\vec{r})]d^3r \tag{48}$$

is the orbital average of the Slater potential, and

$$m_{ji,\sigma} = \int \frac{|\varphi_{j,\sigma}(\vec{r})|^2|\varphi_{i,\sigma}(\vec{r})|^2}{n_\sigma(\vec{r})}d^3r. \tag{49}$$

Thus, Eq. (47) can be recast as the set of linear equations

$$\sum_{i=1}^{N_\sigma-1}(\delta_{ji} - m_{ji,\sigma})[\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}i,\sigma} - \tfrac{1}{2}(\bar{u}_{\mathrm{xc}i,\sigma} + \bar{u}^{*}_{\mathrm{xc}i,\sigma})] = [\bar{v}^{\mathrm{SI}}_{\mathrm{xc}j,\sigma} - \tfrac{1}{2}(\bar{u}_{\mathrm{xc}j,\sigma} + \bar{u}^{*}_{\mathrm{xc}j,\sigma})] \tag{50}$$

for the unknowns $\bar{v}^{\mathrm{KLI}}_{\mathrm{xc}i,\sigma} - \frac{1}{2}(\bar{u}_{\mathrm{xc}i,\sigma} + \bar{u}^{*}_{\mathrm{xc}i,\sigma})$, which can be solved either as a matrix equation, as written, or implicitly by iterating Eqs. (46) and (32) to self-consistency. Note that the term corresponding to the highest occupied orbital can be omitted in Eq. (50) because of Eq. (36).

Thus, the KLI potential is considerably simpler to calculate than the full OEP, and as shown below yields total energies and Kohn-Sham eigenvalues that are in close agreement with the OEP values in many cases. It can in fact be shown to be exact for one-orbital systems and for the homogeneous electron gas. However, in the above derivation, replacing the eigenvalue differences by one global constant seems unmotivated, and therefore the success of the KLI approximation comes as a surprise.

Krieger *et al.* (1992b) have pointed out a reason for the success of the KLI approximation that amounts to an alternative derivation of Eq. (46). Its starting point is the observation that the only difference between Eq. (46) and the true OEP, Eq. (35), is that the orbital-shift term has been neglected in the KLI potential. But if one multiplies this difference term by $n_\sigma(\vec{r})$ and integrates over all space, then the Gauss divergence theorem shows that the integral will vanish for all finite systems since the orbitals vanish at infinity,

$$-\int n_\sigma(\vec{r})\frac{1}{2n_\sigma(\vec{r})}\frac{\hbar^2}{m}\vec{\nabla}\cdot[\psi^{*}_{i,\sigma}(\vec{r})\vec{\nabla}\varphi_{i,\sigma}(\vec{r})]d^3r + \mathrm{c.c.} = 0. \tag{51}$$

Grabo *et al.* (2000) have generalized this argument by showing that the density-averaged integral of the difference term vanishes even without the assumption of asymptotically decaying orbitals. From this point of view, the KLI potential appears as a kind of mean-field approximation to the full OEP and its success becomes understandable. However, it must be noted that this argument implicitly assumes that the other terms, i.e., all contributions to the potential other than the one that is being averaged over in Eq. (51), are similar in the OEP and KLI potential. This generally is the case when the orbitals that are used to evaluate Eqs. (35) and (46) are similar. But in self-consistent OEP and KLI calculations, the missing term can lead to substantial differences (Körzdörfer and Kümmel, 2006; also see Sec. III.B). As an aside, we mention that Krieger *et al.* (1992b) pointed out an approximation that is even more accurate than Eq. (46), but which has never come into widespread use (see Sec. IV.B).

The KLI approximation has been interpreted in yet another way by relating it to Hartree-Fock theory (Nagy, 1997). By relating Hartree-Fock and Kohn-Sham exchange-only theory, one can also derive a different approximation to the OEP, the so-called localized Hartree-Fock potential (Della Sala and Görling, 2001). This same approximation to the OEP has been deduced independently in a different way under the name common energy denominator approximation (CEDA) by Gritsenko and Baerends (2001). Their derivation follows the systematics of making approximations to the Kohn-Sham Green's function. Its central idea is that the replacement of all energy denominators $\varepsilon_{i,\sigma} - \varepsilon_{j,\sigma}$ in Eq. (25) seems poorly justified for at least two reasons. First, the difference between two occupied eigenvalues is in general much smaller than the difference between an occupied eigenvalue and an unoccupied one. Therefore, instead of replacing all eigenvalue differences by one common constant $\Delta\varepsilon_\sigma$, at least a distinction between occupied-occupied and occupied-unoccupied pairs should be made. Second, since the index i in Eq. (24) refers to an occupied eigenvalue, $\varepsilon_{i,\sigma} - \varepsilon_{j,\sigma}$ is always negative when j denotes an unoccupied orbital, but can have a positive or negative sign for j occupied. Again, this suggests treating occupied-occupied and occupied-unoccupied pairs on a different footing. Thus, the CEDA only approximates the occupied-unoccupied eigenvalue differences by a common constant $\Delta\varepsilon_\sigma$ and keeps the exact differences for occupied-occupied pairs. Inserting this approximation for the Green's function into Eq. (24) leads, after some algebra, to

TABLE I. Total energy and eigenvalues (absolute values) in hartrees for the Ar atom. First line, LDA, using the parametrization of Perdew and Wang (1992); second line, OEP values from the exchange-only functional; third line, exchange-only KLI approximation (both from Kümmel and Perdew, 2003a); fourth line, Hartree-Fock values (from Heinemann *et al.*, 1990; Li *et al.*, 1993).

v_{xc}	E	ε_1	ε_2	ε_3	ε_4	ε_5
xc-LDA	525.9398	113.8001	10.7940	8.4433	0.8832	0.3822
x-OEP	526.8122	114.4522	11.1532	8.7338	1.0993	0.5908
x-KLI	526.8105	114.4279	11.1820	8.7911	1.0942	0.5893
HF	526.8175	118.6104	12.3222	9.5715	1.2774	0.5910

$$v_{\mathrm{xc},\sigma}^{\mathrm{CEDA}}(\vec{r}) = v_{\mathrm{xc},\sigma}^{\mathrm{Sl}}(\vec{r}) + \frac{1}{2n_\sigma(\vec{r})}\sum_{i,j=1}^{N_\sigma} \varphi_{i,\sigma}(\vec{r})\varphi_{j,\sigma}^{*}(\vec{r})(\bar{v}_{\mathrm{xc}ij,\sigma}^{\mathrm{CEDA}} - \bar{u}_{\mathrm{xc}ij,\sigma}) + \mathrm{c.c.}, \tag{52}$$

where $\bar{v}_{\mathrm{xc}ij,\sigma}^{\mathrm{CEDA}}$ and $\bar{u}_{\mathrm{xc}ij,\sigma}$ are matrix elements taken with the ith and jth Kohn-Sham orbitals. This equation can be solved in a similar way to that outlined above for the KLI potential (Grüning *et al.*, 2002).

As pointed out by Della Sala and Görling (2001), the CEDA for the exchange-only functional is invariant under unitary transformations of the occupied orbitals, whereas the KLI approximation is not. Comparing Eqs. (45), (46), and (52) further shows a "Russian doll" hierarchy in which the Slater potential terms are contained within those of the KLI potential, which are, in turn, contained within those of the CEDA potential. One may therefore expect that accuracy with respect to the true OEP improves upon going from the Slater to the KLI to the CEDA approximation.

Recently, another approximation to the OEP, dubbed the effective local potential by its developers, was put forward (Staroverov *et al.*, 2006b). This approximation has been reported to allow for efficient numerical implementation, and its relation to the CEDA was discussed by Izmaylov *et al.* (2007).

The above approximate forms of the OEP have been used in calculations ranging from atoms through molecules to solids, employing different numerical techniques ranging from localized basis sets to plane waves to discretizations on real-space grids. In the following, we discuss examples and summarize the main conclusions emerging from these calculations.

The KLI approximation was first extensively tested by Krieger *et al.* themselves (1990, 1992a, 1992b) for atoms and the bare exchange functional (x-KLI). The equations were discretized on a one-dimensional radial grid where angular components were averaged over analytically to obtain spherical densities. Comparison to exchange-only OEP results (x-OEP) obtained by Talman and Shadwick (1976) with the same technique showed that total energies typically agreed within millihartrees and eigenvalues within several 10^{-2} hartrees. As a typical example, we report in Table I the total energies and eigenvalues obtained with four different approaches for the argon atom. The total energies of all three exchange-only theories are close to each other, whereas larger differences are seen for all eigenvalues except for the highest occupied one ε_5. The closeness of the highest occupied eigenvalues is once more a consequence of Eq. (27) (Kümmel and Perdew, 2003b). Comparing x-KLI eigenvalues to LDA eigenvalues shows that the x-KLI eigenvalues correspond much better to binding energies, as discussed in the Introduction. Similar calculations were performed by Grabo and Gross (1995, 1997), who combined the exact exchange energy with a correlation functional based on work by Colle and Salvetti (1975, 1979); see Sec. III.D.3. For an extensive overview of atomic calculations, including relativistic ones, see Grabo *et al.* (2000). KLI calculations exploiting spherical symmetry were also reported for jellium clusters (Ullrich *et al.*, 2000) using a self-interaction-corrected LDA (see Sec. III.B).

Molecular calculations were performed by Grabo and Gross (1995, 1997) and Grabo *et al.* (1997) using the KLI approximation for exact exchange with the Colle-Salvetti functional and a fully numerical approach suitable for closed-shell diatomics, by Kim *et al.* (1999) using a three-dimensional real-space approach combining LDA pseudopotentials (cf. Sec. II.E) with an x-KLI treatment of the valence electrons, by Engel *et al.* (2000a, 2000b) using a Hylleraas-basis expansion for diatomics and the x-KLI potential, and by Della Sala and Görling (2001) and Grüning *et al.* (2002) using x-KLI and x-CEDA with Gaussian basis sets. A few typical all-electron exchange-only results are compiled in Table II. They confirm that the x-KLI approximation yields energies and eigenvalues close to the x-OEP values and they show that the CEDA and KLI approximations are of comparable accuracy for these quantities. The fact that CEDA and KLI ground-state energies are close is plausible since the off-diagonal matrix elements that make the difference between the two potentials [see Eqs. (46) and (52)] will contribute less than the diagonal matrix elements: the spatial overlap between orbitals i and j is smaller for $i \neq j$ than for $i=j$ (Della Sala and Görling, 2001). Small differences between different methods might be due to numerical inaccuracies and differences in the molecular geometries.

The KLI approximation has also been employed in several calculations for solids. Li *et al.* (1991) studied the band structure of noble-gas solids and NaCl in self-interaction-corrected LDA with the OEP and KLI approximation, using the orbital-dependent functionals

TABLE II. Total energies (absolute values) for three small molecules, in hartrees.

Molecule	E^{HF}	$E^{\mathrm{x\text{-}OEP}}$	$E^{\mathrm{x\text{-}CEDA}}$	$E^{\mathrm{x\text{-}KLI}}$
CO	112.7894[c]/112.7893[d]	112.7850[a,c]	112.7817[c,d]	112.7817[c,d]/112.7834[b]
H_2O	76.0665[c]/76.0571[d]	76.0646[a,c]	76.0631[c]/76.0535[d]	76.0628[c]/76.0532[d]
LiH	7.9874[c]	7.9870[c]	7.9870[c]	7.9870[c]/7.9868[b]

[a] From Görling (1999).
[b] From Engel *et al.* (2000b).
[c] From Della Sala and Görling (2001).
[d] From Grüning *et al.* (2002).

only within a muffin-tin sphere and the LDA in the interstitial region. Bylander and Kleinman (1995a, 1995b, 1996, 1997) investigated *sp* semiconductors in the x-KLI approximation with a plane-wave pseudopotential approach. Süle *et al.* (1999, 2000) calculated the band structure of polyethylene using x-KLI and approximations thereof. As far as comparisons with OEP results are available, the KLI approximation seems to be rather accurate for these systems.

A discussion on the physical meaning of the result obtained naturally depends crucially on the E_{xc} functionals employed. It is thus beyond the scope of this section and is addressed in Sec. III. Here we focus on the extent to which employing approximations such as KLI and CEDA in the calculation of v_{xc} can influence the computed results. To that end, we return to two of the central motivations for the use of orbital functionals: eliminating self-interaction errors and incorporating derivative discontinuities. Eliminating self-interaction errors is predominantly a question of E_{xc} and not v_{xc}, because it is determined by the cancellation of self-interaction in the Hartree term by the E_{xc} term. Therefore, as long as the same orbitals are used to evaluate both, an approximate potential (e.g., KLI or CEDA) can neither harm nor improve freedom from self-interaction.

Whether the particle number discontinuities in $v_{\mathrm{xc}}(\vec{r})$ that were introduced in Sec. I.B are correctly captured by the approximations is a more involved question that deserves careful analysis. Krieger *et al.* (1990) pointed out that the KLI approximation does include an integer discontinuity. It is instructive to convince oneself of the mechanism through which this is achieved. The relevant equations in this context are Eqs. (35), (36), and (46). From the discussion of the OEP's long-range asymptotic behavior in Sec. II.A, it is clear that, for both the OEP and the KLI approximation, Eq. (36) is crucial for ensuring that the potential decays to zero. Now assume that we have calculated the OEP or KLI potential for a finite system with N_σ electrons and have properly obeyed Eq. (36). Further, assume that we calculate the potential for a second system that is identical to the first one save for the fact that an infinitesimal amount of charge, described by an infinitesimal fractional occupancy of the previously lowest unoccupied orbital, has been added to the system. [See, e.g., Krieger *et al.* (1992a, 1992b) or Grüning *et al.* (2002) for the KLI and CEDA equations, respectively, with occupation numbers given explicitly.] Clearly, the infinitesimal charge will have a negligible influence on the self-consistent solution of the equations, so the shape of the potential will be practically the same in the first and second calculation. However, the highest occupied orbital in the second calculation is the orbital which in the first calculation was the lowest unoccupied one. Therefore, the value of $\bar{u}_{\mathrm{xc}N_\sigma,\sigma}$ in the second calculation is different from that in the first one, because N_σ now refers to the newly occupied orbital. Correspondingly, in order to keep obeying Eq. (36), the exchange-correlation potential must change by a constant. Thus, it changes discontinuously as the particle number goes through an integer value.

We further consider what happens if instead of an infinitesimal charge we add a very small but finite one, corresponding to a small but finite ϵ_σ, to the N_σ-electron system, namely, $n_\sigma(\vec{r})=\sum_{i=1}^{N_\sigma} f_i|\varphi_{i,\sigma}(\vec{r})|^2$ with $f_i=1$ for $i<N_\sigma$ and $f_{N_\sigma}=\epsilon_\sigma$. In the center of the system, ϵ_σ has a negligible influence since the density is dominated by the fully occupied orbitals. Therefore, in this region the potential of the $N_\sigma+\epsilon_\sigma$ system will have the same shape as the potential of the N_σ system, but for the reason given above it will be upshifted by a constant. However, no matter how tiny ϵ_σ is, far away from the system's center the density will eventually be dominated by the highest (fractionally) occupied orbital because the highest occupied orbital is the slowest to decay. The smaller ϵ_σ is, the further out (i.e., closer to infinity) is the point where the highest occupied orbital takes over, i.e., where $n_\sigma(\vec{r})\approx\epsilon_\sigma|\varphi_{N_\sigma,\sigma}(\vec{r})|^2$ begins to hold. But once the latter relation is close to an equality, Eq. (36) ensures for both Eqs. (35) and (46) that $v_{\mathrm{xc},\sigma}(\vec{r})$ tends to zero [provided $u_{\mathrm{xc}N_\sigma,\sigma}(\vec{r})$ does]. In the region where the highest occupied orbital takes over, a step occurs in the potential that smoothly connects the inner and outer regions. An explicit example of this behavior is shown in Fig. 4.

The above discussion shows that the mechanism responsible for the particle number discontinuity is the same for the OEP and the KLI potential. Equation (52) shows that it appears in a similar way for the CEDA potential. However, the value of the discontinuity is generally not the same in the different approaches. Because the potentials differ, so do the orbital-averaged terms, and thus the discontinuities differ. Whether the discontinuity of the potential corresponds to the true, physical discontinuity depends on the E_{xc} employed. But even

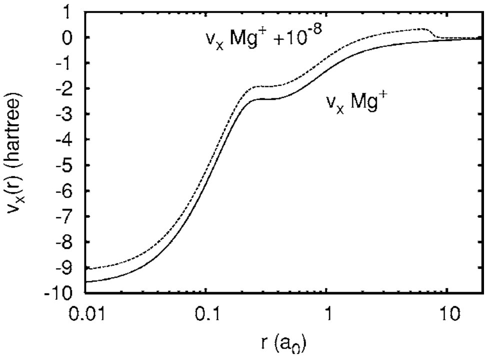

FIG. 4. Spin-up component of the exchange-only potential of the Mg cation in the KLI approximation. Solid curve, potential for Mg^+. Dashed curve, potential for Mg^+ plus a fractional (spin-up) charge of $\epsilon = 10^{-8}$. For this value of ϵ, the density starts to become dominated by the highest occupied orbital around $r = 7a_0$. Thus, addition of the minute charge shifts the potential up by a constant in the region $r < 7a_0$ and the up-shift gradually disappears around $r = 7a_0$ (see text). The result is a steplike structure in the potential. For a smaller added charge ϵ, the step would appear further out, i.e., at a larger value of r. In the limit $\epsilon \to 0$, the step reaches infinity and the derivative-discontinuity limit is reached. This type of figure was first shown by Krieger *et al.* (1992a).

given the hypothetical ultimate E_{xc}, only the OEP will yield the correct discontinuity.

The orbital-averaged terms that are responsible for the discontinuity in Eq. (35) also generate a step structure in the exchange-correlation potential at electronic shell closings. These steps have been analyzed (Gritsenko *et al.*, 1994; van Leeuwen *et al.*, 1995) by decomposing the exchange-correlation potential into screening and response parts (Gunnarsson *et al.*, 1979; Gritsenko *et al.*, 1993). In terms of Eq. (35), the long-ranged Slater term of Eq. (45) is the screening potential, and the orbital-average and orbital-shift terms together yield the short-ranged response potential. The steps at shell closings can be traced to the response part of the potential (Gritsenko *et al.*, 1994; van Leeuwen *et al.*, 1995). Therefore, approximations like the KLI and CEDA qualitatively capture the step structure, but the Slater approximation does not. This is a major deficiency of the Slater approximation, because the terms that cause the steplike features of the exchange-correlation potential also play an important role in a system's response to external perturbations. This has been demonstrated for the linear and nonlinear responses of extended molecular systems (van Gisbergen *et al.*, 1999). Because this type of application is also a major area where approximate potentials like the KLI and CEDA do not yield reliable results [NMR shieldings being another such area (Teale and Tozer, 2005)], we discuss this point in detail.

It is well known (Champagne *et al.*, 1998) that both the LDA and GGA dramatically overestimate the dipole polarizability, and even more so the hyperpolarizability, of conjugated, extended molecular systems (oligomers and polymers). Because such systems are potentially useful for nonlinear optics devices (Kanis *et al.*, 1994), efforts have been made to understand this particularly egregious failure of semilocal functionals. To this end, model systems consisting of "hydrogen chains," i.e., one-dimensional arrangements of hydrogen atoms (Champagne *et al.*, 1995a, 1995b), proved to be very useful. They mimic the important electronic features of true molecular chains, such as bond-length alternation, high and directional electron mobility along the backbone of the chain, and large response coefficients, but are much easier to handle computationally. Therefore, benchmark wave-function-based calculations for the response coefficients were possible (Champagne *et al.*, 1995a, 1995b). For development and insight purposes, these calculations are ideal for comparing with DFT results, because the uncertainties that plague comparison with experimental data (such as incomplete orientation of the chains in an experimental sample, etc.) are avoided. The wave-function-based calculations revealed that Hartree-Fock theory already provides for quite reasonable response coefficients, whereas the GGA and LDA hyperpolarizabilities are off by orders of magnitude. This may be seen as a major surprise, because for most other observables even the LDA is more reliable than Hartree-Fock theory.

van Gisbergen *et al.* (1999) calculated the response coefficients of hydrogen chains employing the exchange-only KLI potential. They found that, compared to the LDA and GGA, the x-KLI approximation greatly reduces the overestimation of the response coefficients. They were able to trace this reduction back to a qualitative difference in the behavior of the exchange-correlation potential of semilocal functionals and the x-KLI potential. When an external electrical field is applied along the backbone of the chain, the electron density is polarized, i.e., it shifts with the field. The exchange-correlation potential of the LDA simply follows the density, thereby magnifying the effect of the externally applied field by working with it. However, the studies of van Gisbergen *et al.* revealed that the x-KLI potential shows the opposite behavior. It works against the externally applied field, thus leading to a smaller polarization, i.e., smaller response coefficients. By decomposing the potential into the above-mentioned screening and response parts, they showed further that the field-counteracting term is due to the response potential. Therefore, the latter is a crucial part of Eqs. (27), (46), and (52), and its omission is a serious deficiency of the Slater approximation, Eq. (45).

Although a major improvement over the LDA and GGA, the difference between the (hyper)polarizabilities calculated with the x-KLI approximation and the wave-function-based results was still much larger than that found for Hartree-Fock theory. For example, for H_{18}, the x-KLI second hyperpolarizability is about twice as large as the corresponding Hartree-Fock value. This result raised an important question. Is this difference (a) due to the KLI approximation to the OEP, or (b) a reflection of a fundamental difference between Kohn-Sham and Hartree-Fock exchange? The latter answer would have implied that the "tiny piece" of correlation

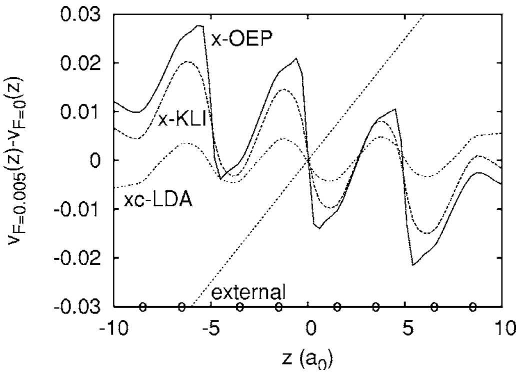

FIG. 5. Difference between the exchange potential with and without an external electric field applied to a chain of eight hydrogen atoms placed at alternating distances of $2a_0$ and $3a_0$. Atomic positions are indicated by small circles at the lower boundary. Solid line, difference between $v_x^{\mathrm{OEP}}(0,0,z)$ with an electric field $F=0.005e/a_0^2$ and with $F=0$. Dashed and double-dotted lines, corresponding differences for v_x^{KLI} and v_{xc}^{LDA}. Straight dotted line, potential corresponding to the external electric field F. Note that v_x^{OEP} and v_x^{KLI} are biased against the external field, whereas v_{xc}^{LDA} is aligned with it. The field-counteracting effect is noticeably larger for the OEP that for the KLI potential. From Kümmel *et al.*, 2004.

energy that, from the Kohn-Sham point of view, makes the difference between Hartree-Fock and x-OEP results plays a major role in charge-transfer problems—a disconcerting thought.

Analytical arguments and numerical calculations based on the x-CEDA potential (Gritsenko and Baerends, 2001; Grüning *et al.*, 2002) indicated that the KLI potential misses non-negligible parts of the response potential. Calculations also demonstrated that x-CEDA (hyper)polarizabilities are more accurate than x-KLI ones. However, they were still considerably less accurate than the Hartree-Fock values. The question was finally resolved by calculating the response coefficients from a numerical solution of the x-OEP equation (see Sec. II.D) without further approximations (Kümmel *et al.*, 2004; Kümmel and Kronik, 2006). The response coefficients obtained from the x-OEP calculations were found to be very close to the Hartree-Fock values, thus indicating that the answer to the question posed above is (a) and not (b). This is very reassuring from a conceptual point of view. At the same time, it serves as a warning because it shows that approximate potentials like KLI and CEDA do not yield reliable results for all observables. The origin of the differences can be seen in Fig. 5: The field-counteracting term is noticeably larger with x-OEP than with x-KLI, and this itself has its origin in noticeable differences in the $\bar{v}_{\mathrm{xc}i,\sigma}-\bar{u}_{\mathrm{xc}i,\sigma}$ terms. We stress that, contrary to earlier belief (see, e.g., Dreuw *et al.*, 2003), using a local Kohn-Sham potential does not hinder a correct description of long-range charge-transfer situations. The reason that Kohn-Sham DFT and time-dependent DFT can capture charge transfer correctly, as explicitly demonstrated, e.g., in the above-mentioned example and others (Tozer, 2003; Mori-Sánchez *et al.*, 2003; Maitra, 2005; Champagne *et al.*, 2006), are the step structures and discontinuities in the potential just discussed.

In summary, it can be said that the KLI and CEDA approximations, according to all experience so far, are accurate for total energies and not as (but still quite) accurate for occupied eigenvalues. However, they can fail poorly in predicting response properties. This also has important consequences for the use of orbital-dependent functionals in time-dependent DFT, where effects similar to the ones discussed above play an important role (see Sec. IV.B). Finally, it should be mentioned that recently an idea has been put forward to mimic OEP behavior with semilocal ingredients (Becke and Johnson, 2006). Whether approaches of this type can be a useful alternative to approximations like KLI and CEDA remains to be seen.

D. Numerical construction

The above discussion shows that the ability to calculate the exact OEP for a given orbital-dependent functional is needed. The first technique for solving the OEP equation was developed by Talman and Shadwick (1976) for the exact exchange functional and spherically symmetric, effectively one-dimensional systems. With this technique, they were able to show that the numerical value for the OEP energy of spherical atoms is close to the Hartree-Fock energy. Their numerical technique and computer code were later refined (Engel and Vosko, 1993) and used in many investigations of spherically symmetric systems (e.g., Krieger *et al.*, 1992a; Engel and Vosko, 1994; Grabo *et al.*, 1997, 2000). The first OEP calculations for solids were also based on an effectively one-dimensional approximation (Kotani, 1994, 1995; Kotani and Akai, 1996).

The breakthrough that made calculations for three-dimensional systems possible came with the realization that the OEP can be calculated by numerically inverting the Kohn-Sham response function (Görling and Levy, 1994, 1995a). The basic idea of the approach is straightforward: Evaluate the right-hand side of Eq. (20) explicitly via Eq. (22) and the inversion of Eq. (23). But the practical realization of this idea is quite involved. The first obstacle is that, in principle, the full spectrum of occupied and unoccupied Kohn-Sham orbitals and eigenvalues must be calculated to obtain the response function. The second and yet larger obstacle is that the response function is, strictly speaking, not invertible: Any change of the potential by just a constant will leave the density unaffected. Thus, in Eq. (21), a zero numerator can be associated with a nonzero denominator. Therefore, in order to make possible an inversion one must explicitly exclude a constant shift of the potential in the numerical representation of the response function. In a plane-wave basis representation, this can be achieved easily since a constant is part of the basis set. The resulting inversion scheme was employed to calculate, e.g., the exact-exchange OEP band structure of semiconductors and insulators (Städele *et al.*, 1997, 1999;

Magyar *et al.*, 2004) and of molecular hydrogen under pressure (Städele and Martin, 2000). Schemes that use the inversion have also been developed for the localized orbitals typically used in quantum chemistry (Görling, 1999; Ivanov *et al.*, 1999; Fernandez *et al.*, 2003a, 2003b) and other all-electron approaches (Sharma *et al.*, 2005), and successful calculations have been reported for atoms, molecules, and solids. However, excluding the constant shift is considerably more difficult with a localized basis, and finite-basis-set calculations can suffer from non-negligible basis-set errors, as discussed by Della Sala and Görling (2001), Hirata *et al.* (2001), and Hamel *et al.* (2002). In the calculations of Ivanov *et al.* (1999), these errors were large enough to make the OEP energies larger than those obtained from accurate KLI calculations (see Engel *et al.*, 2000b, for a detailed discussion), contradicting the variational character of the OEP.

A direct approach to calculating the OEP is to implement Eq. (16) itself numerically: The Kohn-Sham potential is expanded in a basis set and then the total energy is minimized with respect to the coefficients of the expansion. This idea was first used by Rose and Shore (1975) to calculate the Kohn-Sham potential. In the modern OEP context, it was first employed by Fritsche and Yuan (1998) in calculations for spherical atoms and by Yang and Wu (2002) and Mori-Sánchez *et al.* (2003) in calculations for atoms and molecules of various sizes. The appealing aspect of this approach is its directness. It completely avoids the complications that are associated with the inversion of the response function. A drawback is that the numerical burden is quite high for three-dimensional systems (Yang and Wu, 2002) and that eigenvalues are hard to determine accurately (Rohr *et al.*, 2006). In order to keep the expansion tractable, localized basis sets have been used. With these, the asymptotic behavior of the potential cannot be captured. A natural idea is then to enforce the correct asymptotics by constructing the potential as a sum of a fixed long-ranged term that is asymptotically correct and a short-ranged term that is optimized as just described. Using the Fermi-Amaldi (1934) approximation (see Sec. III.B) for the long-ranged part, it was believed that essentially the exact OEP would be obtained (Yang and Wu, 2002). However, due to the nonvanishing asymptotic constants discussed in Sec. II.A, the Fermi-Amaldi potential's isotropic asymptotic $\sim -1/r$ decay does not correspond exactly to the correct asymptotic behavior of the OEP. Similar conclusions hold for other approaches (see, e.g., Görling, 1999; Izmaylov *et al.*, 2007) that enforce $\sim -1/r$ asymptotics.

The direct approach of numerically constructing the OEP by expanding both the potential and orbitals in finite Gaussian basis sets and then minimizing the total energy with respect to the potential expansion has recently become the subject of a discussion. Staroverov *et al.* (2006a, 2006b) observed that by increasing the size of the basis set in which the potential is expanded, while keeping the basis set for the expansion of the orbitals fixed (and possibly smaller), they could find, within the chosen basis-set representation, local, multiplicative potentials that yield not the OEP energy but the corresponding orbital-specific potential energy. The latter is lower than the OEP energy. For example, for the exact-exchange functional they were able to obtain, within the chosen finite-basis representation, the Hartree-Fock energy from local potentials. This has been seen as a paradox, casting doubt on the OEP approach as such. However, we believe that this does not expose a fundamental problem with the OEP approach, because the OEP equation as such has a well-defined, unique solution. In our opinion, the problem observed by Staroverov *et al.* with their finite basis sets is a consequence of inconsistent basis-set expansions. A consistent expansion would converge the basis using the same (orthogonal) basis set for all quantities, thus avoiding the definition of different quantities in different spaces. By choosing different basis representations for potential and orbitals and a larger basis set for the potential than for the orbitals, an originally well-defined mathematical problem can be turned into an ill-defined one: When the potential is defined on a larger space than the orbitals, there are variational degrees of freedom that are not properly constrained, leading to spurious effects. These arguments are supported by calculations that solve the OEP equation on numerical grids (see, e.g., Talman and Shadwick, 1976; Engel and Vosko, 1993; Kümmel and Perdew, 2003a, 2003b). Such numerical grids can be viewed [see, e.g., Beck (2000) and Kronik *et al.* (2001)] as basis sets comprising Dirac delta functions. In such calculations, orbitals and potential are represented on the same grid, i.e., consistent expansions are used. The grids can be made as large and as tight as one desires (corresponding to the systematic, consistent use of large basis sets), without observing numerical instabilities. The calculations are accurate and do not show multiple solutions, and systematic convergence to the OEP energy is obtained. The same reliability is to be expected from other consistent basis sets. Observations related to this subject have also been made by others (Hirata *et al.*, 2001; Görling, 2006; Teale *et al.*, 2007), and schemes have been discussed to circumvent the problem in practice by converging the orbital basis set for a fixed potential basis set.

The possibility of using the same basis set for both expansions as a consistent strategy has also been considered by Staroverov *et al.* (2006b). However, as they explain, this is hardly an option in quantum chemistry program packages using Gaussian expansions since "a shared basis set is hard to balance because tight and diffuse basis functions needed for many molecular properties are bad for the potential." Therefore, what we referred to above as an inconsistent basis-set expansion is, from their perspective, not just a little nuisance that can easily be avoided, but rather a problem typical of quantum chemistry program packages when one tries to find the OEP by energy minimization: First, a basis-set representation for the orbitals is chosen and then one needs to find the Kohn-Sham potential represented in some other basis set. This observation makes clear the great practical importance of the work of Staroverov et al. (2006a, 2006b) and Izmaylov et al. (2007). Their work pointed out that great care has to be exercised in such procedures. Furthermore, they devised a strategy for obtaining approximations to the true OEP in the form of effective local potentials, as mentioned in Sec. II.C.

Yet another and completely different way of constructing the OEP is based on solving Eq. (31) numerically for the orbital shifts and using these to calculate the OEP (Krieger *et al.*, 1992b; Grabo *et al.*, 2000; Kümmel and Perdew, 2003a, 2003b). This amounts to iteratively solving a set of coupled partial differential equations instead of an integral equation. The most direct approach of this kind combines Eq. (31) with Eq. (35) and the Kohn-Sham equations into a self-consistent iteration for orbitals and orbital shifts. However, for most applications this is not the most practical way to use the orbital shifts, because in Eq. (35) one has to divide the divergence term on the right-hand side by the density, which leads to problems, e.g., in the asymptotic region of finite systems where the density practically vanishes. This problem can be avoided (at the cost of losing accuracy in the asymptotic region) by rewriting the iteration such that division by the density is not necessary (Kümmel and Perdew, 2003a, 2003b). The resulting scheme is stable and accurate and in principle is suitable for either a real-space- or a basis-set-based calculation. It has been applied, e.g., to atoms, clusters (Kümmel and Perdew, 2003a, 2003b; Mundt *et al.*, 2006; Payami, 2006), hydrogen chains (Kümmel, 2004; Kümmel *et al.*, 2004), and the metal-vacuum interface problem (Horowitz *et al.*, 2006).

For completeness, we note that additional strategies for solving the OEP equation have also been suggested. Examples include studies of the Kohn-Sham exact exchange potential as a function of sub-band filling in a quasi-two-dimensional electron gas (Goñi *et al.*, 2002; Reboredo and Proetto, 2003; Rigamonti *et al.*, 2005), a gradient search method that is related to the iterative method discussed above (Hyman *et al.*, 2000) and an extension thereof (Kosaka, 2006), and a finite-temperature formalism (Lippert *et al.*, 2006).

In summary, OEP calculations are still far more tedious and time consuming than calculations based on explicit density functionals, but they are possible and already have been used to investigate a great variety of systems. We expect that performing ground-state OEP calculations will become an increasingly routine task in the near future, even though they probably will never become as computationally inexpensive as LDA calculations.

E. Pseudopotentials

The numerical solution of the Kohn-Sham equations is a nontrivial affair even when semilocal functionals are used, and certainly when the OEP method (or any of its approximate forms) is used. One popular and successful approach for simplifying the solution of Kohn-Sham equations considerably is to employ them in conjunction with pseudopotentials (see, e.g., Cohen and Chelikowsky, 1988; Pickett, 1989; Chelikowsky and Cohen, 1992; Singh and Nordstrom, 2005). In fact, many of the plane-wave and real-space calculations discussed in the preceding sections were facilitated by the pseudopotential approximation.

The pseudopotential concept, first suggested by Fermi (1934), is based on the same reasoning that is behind the Periodic Table of elements: core electrons are chemically inert (see also Goedecker and Maschke, 1992). Thus, it is sufficient to consider valence electrons that are subject to an effective potential (i.e., the pseudopotential) that describes the screening of the nucleus by core electrons. Because of this screening, the pseudopotential (unlike the nuclear potential) is weak and binds only the valence electrons (Philips and Kleinman, 1959). Furthermore, use of appropriately constructed pseudopotentials can allow for the consideration of relativistic effects near the core, within an otherwise nonrelativistic framework (Kleinman, 1980).

To demonstrate how pseudopotentials can be constructed and employed in practice within DFT, we focus on a particularly important class of pseudopotentials known as norm-conserving pseudopotentials (Hamman *et al.*, 1979). Other constructs, notably ultrasoft pseudopotentials (Blöchl, 1990; Vanderbilt, 1990), are not discussed here because we are not aware of their use in conjunction with the OEP method.

The chemically uninteresting core region of the valence orbitals is highly oscillatory and complicates the numerical treatment of the Kohn-Sham equations greatly. We therefore construct atomic pseudo-valence-orbitals $\varphi^a_{\mathrm{ps},i}$ (where the superscript a denotes atomic and the subscript ps denotes pseudo). These should be identical to the atomic all-electron valence orbitals φ^a_i (obtained from a complete solution of the Kohn-Sham equation), outside some prescribed core radius r_c, but should be slowly varying functions inside the core region. If both $\varphi^a_{\mathrm{ps},i}$ and φ^a_i are normalized, and given their identity outside the core region, the integral of the square of the pseudo- and all-electron orbitals within the core region should be equal—a norm-conservation condition. One can choose the core-region slowly varying function in many different ways, all of which obey norm conservation. Some popular recipes are, e.g., those of Bachelet *et al.* (1982), Rappe *et al.* (1990), Troullier and Martins (1991), and variants thereof. As an example, LDA all-electron orbitals for the valence 3*s* and 3*p* states of sodium, together with their ensuing norm-conserving Troullier-Martins pseudo-orbitals, are given in Fig. 6(a) (Chelikowsky, Kronik, Vasiliev, Jain, *et al.*, 2003).

Unlike the all-electron valence orbitals, the atomic pseudo-orbitals are *not* solutions of the Kohn-Sham equations with the true nuclear potential. However, one can seek a fictitious nuclear potential—the pseudopotential—that, when screened by the pseudo-valence-electrons, would yield the all-electron Kohn-Sham eigenvalues together with the Kohn-Sham pseudo-wave-functions. This process is typically performed in two steps. First, a modified Kohn-Sham po-

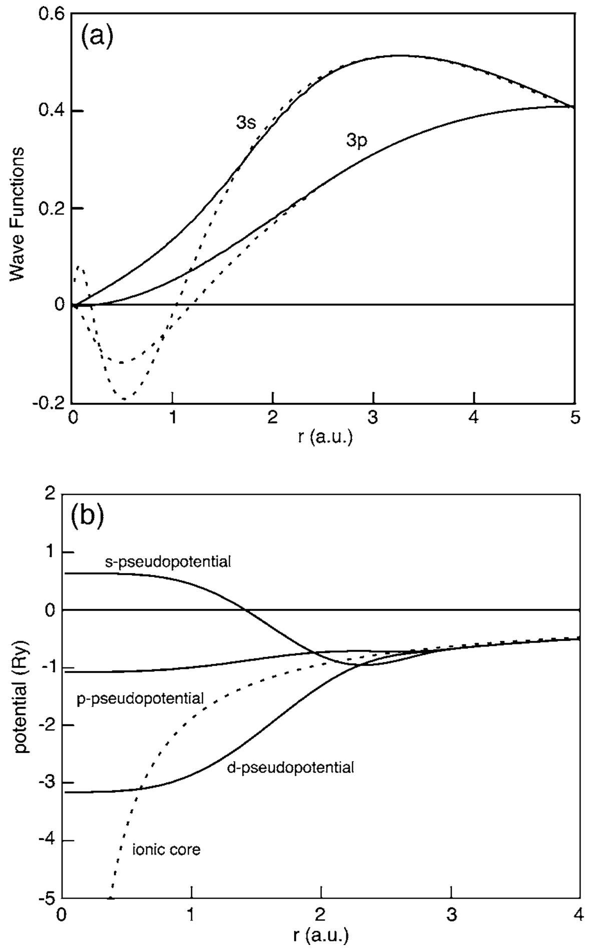

FIG. 6. An example of pseudopotential generation. (a) All-electron orbitals (dashed lines) and pseudo-orbitals (solid lines) for the sodium atom. (b) Pseudopotentials for the sodium atom (solid lines), compared to the electrostatic potential of the ionic core, $-Z_{\rm ps}/r$ (dashed line). From Chelikowsky, Kronik, Vasiliev, *et al.*, 2003.

tential $v_{\rm ps}^{a,\rm scr}$ corresponding to a screened pseudopotential, is obtained by inverting the Kohn-Sham equation (3), i.e.,

$$v_{\rm ps}^{a,\rm scr} = \varepsilon_i + \hbar^2\nabla^2\varphi^a_{{\rm ps},i}/(2m\varphi^a_{{\rm ps},i}). \tag{53}$$

Second, the bare pseudopotential is obtained by unscreening the screened pseudopotential using only the valence pseudodensity $n^a_{v,\rm ps} \equiv \Sigma_{i,\rm occup}|\varphi^a_{{\rm ps},i}(\vec{r})|^2$, where the summation is restricted to filled pseudo-orbitals. Explicitly,

$$v^a_{\rm ps} = v^{a,\rm scr}_{\rm ps} - e^2 \int \frac{n^a_{v,\rm ps}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r' - v_{\rm xc}([n^a_{v,\rm ps}];\vec{r}). \tag{54}$$

Equation (54) is reasonable only if the exchange-correlation screening can be considered as a linear combination of core and valence screening. Given the nonlinear nature of the exchange-correlation functional, this will be the case only if overlap between core and valence orbitals is negligible. If not, one can include a version of the core charge, $n^a_{c,\rm NLCC}$ (where NLCC indicates the nonlinear core correction), itself typically smoothed in the vicinity of the nucleus, in (and only in) the exchange-correlation term of Eq. (54) (Louie *et al.*, 1982). Explicitly,

$$v^a_{\rm ps} = v^{a,\rm scr}_{{\rm ps},i} - e^2 \int \frac{n^a_{v,\rm ps}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r' - v_{\rm xc}([n^a_{c,\rm NLCC} + n^a_{v,\rm ps}];\vec{r}). \tag{55}$$

As an example, pseudopotentials corresponding to the $3s$, $3p$, and $3d$ pseudo-orbitals of sodium are shown in Fig. 6(b). They are indeed nonsingular and much weaker than the original nuclear potential.

Replacing the ionic potential in Eq. (3) by the pseudopotential leads to the pseudopotential Kohn-Sham equation in the form

$$\left(-\frac{\hbar^2\nabla^2}{2m} + \hat{v}_{\rm ps} + e^2 \int \frac{n_{v,\rm ps}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r' + v_{\rm xc}([n^a_{c,\rm NLCC} + n^a_{v,\rm ps}];\vec{r})\right)\varphi_{{\rm ps},i} = \varepsilon_i\varphi_{{\rm ps},i}. \tag{56}$$

Using Eq. (56) instead of Eq. (3) simplifies the computation in two major ways: First, core electrons are no longer computed explicitly so the number of eigenvalues and eigenfunctions that need to be determined is significantly reduced. Second, all potentials and wave functions are now *soft*, i.e., slowly varying functions that require a much smaller numerical effort (e.g., fewer plane waves in a Fourier expansion, or fewer grid points in a real-space approach) even for the remaining valence electrons.

A price to be paid for this numerical simplification is that the pseudopotential operator must depend on the angular momentum because, as shown in Fig. 6(b), the pseudopotential $v^a_{\rm ps}$, generated using Eq. (54) or Eq. (55) from orbitals with different angular momenta, is different. Thus, the pseudopotential operator $\hat{v}_{ps}$ in Eq. (56) is generally of the form

$$\hat{v}_{\rm ps} = \sum_a v^{a,l}_{\rm ps}(|\vec{r}-\vec{R}_a|)P^l(\vec{r}-\vec{R}_a), \tag{57}$$

where l is the angular momentum and P^l is the projection operator corresponding to that angular momentum. Importantly, this means that the pseudopotential is *not* a local potential.

The nonlocality of the pseudopotential raises an apparent formal difficulty. A nonlocal potential is not determined uniquely by the density, but rather by the density matrix (Gilbert, 1975). Consequently, Eq. (56), unlike Eq. (3), is not formally grounded within the Hohenberg-Kohn theorem because the pseudodensity does not uniquely define a pseudopotential. In our opinion, this is actually a nonissue. The purpose of a good pseudopotential is not to recast the original many-body system in a different form, as this task has already been achieved by the Kohn-Sham construct. Instead, a good pseudopotential should merely mimic the Kohn-Sham construct. This means that, when screened with the same Hartree and exchange-correlation expressions as in the original Kohn-Sham potential, but with pseudo-orbitals only (possibly with NLCC), a good pseudopotential should yield (i) a valence electron eigenvalue spectrum as close as possible to the true Kohn-Sham one, (ii) pseudo-orbitals as close as possible to the true Kohn-Sham orbitals outside the core radius of each atom. As long as these two conditions are met, one-to-one correspondence between the pseudodensity and the pseudopotential, or lack thereof, is irrelevant because neither quantity is physical anyway.

For isolated atoms, the pseudopotential clearly meets both of the above demands rigorously, by construction. The beauty of the norm-conserving pseudopotential is that, to a large extent, these demands can be met in a controllable approximate way even when pseudopotentials generated from an atomic configuration are used in a molecular or condensed matter system. In other words, these pseudopotentials, if well constructed, are *transferable*. Norm conservation assists transferability in two ways (Hamman *et al.*, 1979): It produces the correct electrostatic behavior outside the core region, and it produces scattering properties that remain correct to first order as bonding shifts energy levels from their atomic values. In practice, construction of a pseudopotential is always a tradeoff between transferability (which improves with decreasing r_c) and softness (which improves with increasing r_c). The above-mentioned examples of popular pseudopotential recipes are all clever schemes that utilize the nonuniqueness of the pseudo-orbital and pseudopotential in the core region to optimize this tradeoff for various scenarios.

Naively, it would seem that generating norm-conserving pseudopotentials for OEP calculations is, in principle, no different from generating them for LDA or GGA calculations. It simply requires that the evaluation of the exchange-correlation term in the all-electron atomic calculations, as well as in the subsequent unscreening of the pseudopotential, be performed using the OEP equation (24), or any of its approximate forms discussed above. This was first attempted by Bylander and Kleinman (1995a, 1995b), who constructed norm-conserving pseudopotentials within the KLI approximation for the exact-exchange functional. They found that pseudopotential construction for OEP calculations involves an additional difficulty: For well-constructed norm-conserving pseudopotentials within a semilocal functional [e.g., Fig. 6(b)], a trivial outcome of the Gauss theorem is that for $r > r_c$ all pseudopotential components converge rapidly to $-Z_{\mathrm{ps}}/r$, where Z_{ps} is the difference between the nuclear charge Z_a and the core-electron charge Z_c. (The very small residual difference is due to Coulomb multipoles in the core region.) For norm-conserving pseudopotentials within OEP, this is, unfortunately, no longer the case. This is demonstrated in Fig. 7, which shows $v_{\mathrm{ps}} + Z_{\mathrm{ps}}/r$ for the p component of an Al pseudopotential with exact-exchange OEP (Engel *et al.*, 2001). Clearly, a small, but non-negligible, spurious tail is apparent. The tail arises because, whenever an orbital-dependent functional is used, changes in the core-electron density affect the Kohn-Sham potential not only electrostatically through the Hartree term, but also quantum mechanically through the exchange-correlation term. Because the latter term can be highly nonlocal, equal valence wave functions outside the core region need not imply equal potentials outside it. The spurious tail is problematic because it implies spurious interatomic potentials and forces. These may result in very significant and uncontrolled errors for, e.g., bond lengths and binding energies. Worse, the spurious forces can be either attractive or repulsive, depending on the circumstances.

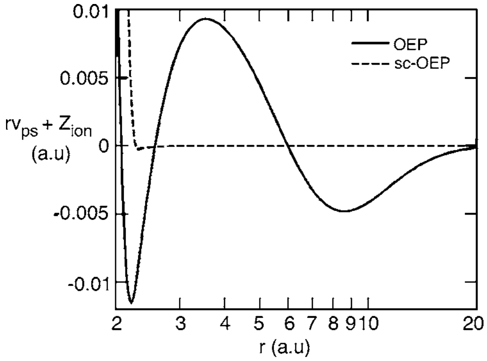

FIG. 7. Sum of the pseudopotential and ionic core potential ($v_{\mathrm{ps}} + Z_{\mathrm{ps}}/r$) for the p component of an Al pseudopotential, with exact-exchange OEP and with an exact exchange self-consistent OEP (sc-OEP) procedure elaborated below. From Engel *et al.*, 2001.

The spurious tail problem has been handled in many ways. The simplest and also earliest approach (Bylander and Kleinman, 1995a) is to live with it, with the understanding that some error is introduced. According to the studies of Engel *et al.* (2001), there are cases, e.g., the HF molecule, where the error introduced with respect to the all-electron solution really is negligible. In other cases, however (e.g., O_2), the error can be far too large to ignore (0.365 eV in the binding energy, for example). A different option is to use LDA pseudopotentials even when performing OEP calculations (Kim *et al.*, 1999; Fleszar, 2001). Clearly, this avoids spurious tails, but the price one has to pay is an inconsistency, i.e., a different screening of core and valence electrons. The detailed studies of Engel *et al.* (2001) show that, while the average error obtained with this approach is somewhat smaller than that obtained with the as-is OEP potential, large uncontrolled errors are still possible. Similarly, significant differences between LDA pseudopotentials and the pseudopotentials discussed in the following paragraph were found by Rinke *et al.* (2005). In fact, Fuchs *et al.* (1998) have shown that serious errors can arise even when using LDA pseudopotentials within a GGA calculation. That this is so even though the LDA and GGA are much closer conceptually than the LDA and exact exchange should warn against using different exchange-correlation functionals for pseudopotential generation and application.

A third option, first suggested by Bylander and Kleinman (1995b), is to filter out the spurious tail *a posteriori* using a smooth dampening function. This approach has often been used since (see, e.g., Bylander and Kleinman, 1996; Städele *et al.*, 1999; Moukara *et al.*, 2000), with various flavors of dampening functions. A conceptual difficulty with this approach is that the final, filtered pseudopotential is no longer truly self-consistent, i.e., it will not exactly yield the valence pseudo-orbitals of a single atom. This is not necessarily a serious issue in practice. For anything other than the configuration for which it was created, the pseudopotential does not yield the exact Kohn-Sham orbitals outside the core radius of each atom, but rather an approximation thereof. The practical question, then, is the quantitative extent of the approximation made by the filtering step. This aspect has also been studied by Engel *et al.* (2001). They concluded that, in principle, this approach may yield values in good agreement with the all-electron ones. However, in practice the values obtained can be very sensitive to the exact form of the dampening function, requiring careful testing of the pseudopotential in various chemical environments.

Engel *et al.* (2001) have suggested a fourth approach to avoiding the spurious tail. They set out to retain the advantage of the *a posteriori* tail elimination while avoiding its disadvantages, namely, lack of consistency between the pseudopotential and pseudo-orbitals and introduction of additional parameters. In their approach, one begins by solving the all-electron atomic problem and using it to generate a set of norm-conserving pseudo-orbitals. But then, instead of using Eq. (53) followed by Eq. (54) or Eq. (55), one constructs a modified potential in the form

$$\begin{aligned} v^{\mathrm{mod}}(\vec{r}) = v^{a}_{\mathrm{KS}}(\vec{r}) - e^2 \int \frac{n^{a}_{v,ae}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r' \\ - v_{\mathrm{xc}}([n^{a}_{c,ae} + n^{a}_{v,ae}];r) + e^2 \int \frac{n^{a}_{v,\mathrm{ps}}(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r' \\ + v_{\mathrm{xc}}([n^{a}_{v,\mathrm{ps}}];\vec{r}). \end{aligned} \tag{58}$$

In the first three terms of $v^{\mathrm{mod}}(\vec{r})$, all of the exchange-correlation term and the valence part of the Hartree term are removed from the all-electron Kohn potential. The nuclear potential is then screened only by core electrostatics and so must have the correct $-Z_{\mathrm{ps}}/r$ asymptotic behavior. The next two terms then screen this potential by the pseudo-Hartree and exchange-correlation terms to create $v^{\mathrm{mod}}(\vec{r})$. However, because its construction avoided Eq. (53), $v^{\mathrm{mod}}(\vec{r})$ is not a screened pseudopotential, in the sense that it will not generate the pseudo-orbitals from which it was created. Therefore, in the next step one solves for the atomic Kohn-Sham orbitals outside r_c, using $v^{\mathrm{mod}}(\vec{r})$ as the Kohn-Sham potential, while demanding that the all-electron eigenvalues are retained. Asymptotic decay of the all-electron orbital is used as a boundary condition at infinity and norm conservation outside r_c is required. Because $v^{\mathrm{mod}}(\vec{r})$ is different from $v^{a}_{\mathrm{KS}}(\vec{r})$, the orbitals obtained are not identical to the all-electron ones outside r_c, but the asymptotic and norm-conservation conditions usually dictate that they are very close. One can then use these orbitals to generate pseudo-orbitals and a pseudodensity yet again. These are used to update $v^{\mathrm{mod}}(\vec{r})$, find new all-electron-like orbitals outside r_c, use those to generate pseudo-orbitals and a pseudodensity, and so on until convergence. The screened pseudopotential generated from the converged pseudo-orbitals, found using Eq. (53), will then be equal to $v^{\mathrm{mod}}(\vec{r})$ outside r_c by construction. Furthermore, Eq. (58) guarantees that the unscreened pseudopotential found using Eq. (54) or Eq. (55) will be devoid of a spurious tail. This is demonstrated in Fig. 7, which also shows the p component of an Al self-consistent OEP pseudopotential.

Pseudopotentials generated using the scheme of Engel *et al.* (2001) are known as self-consistent norm-conserving OEP pseudopotentials. They succeed in eliminating the spurious tail without adjustable parameters, while retaining the consistency between pseudopotential and pseudo-orbital, by sacrificing a different property: the exact equality between the all-electron and pseudo-orbitals outside r_c. While this certainly introduces some error, numerical experimentation shows that it provides results that are at least on par with those of *a posteriori* smoothed pseudopotentials with much less uncertainty in the pseudopotential construction.

If we compare the construction of a high-quality norm-conserving pseudopotential within the OEP framework and within a semilocal functional, the following picture emerges. Most importantly, pseudopotentials can be used for accurate total energy calculations in conjunction with orbital-dependent functionals. Furthermore, like their semilocal counterparts, OEP pseudopotentials can be (and have been) constructed within a relativistic framework (Moukara *et al.*, 2000; Engel *et al.*, 2001). The bad news is that, if a spurious pseudopotential tail is to be avoided, at least one extra assumption is required and a more elaborate procedure needs to be performed. The good news, however, is that once this effort has been undertaken and an adequate OEP atomic pseudopotential is generated, it tends to be much more robust than its semilocal counterpart (Höck and Engel, 1998; Moukara *et al.*, 2000). First, due to the explicit treatment of orbitals, OEP pseudopotentials tend to be significantly less sensitive to the exact chemical configuration in which they were generated, i.e., they are more transferable. Second, because the exact-exchange term is linear in the density matrix, OEP pseudopotentials are usually sufficiently accurate even without NLCC (although that could depend, of course, on the nature of the correlation functional). Thus, although not yet in widespread use, pseudopotentials can be expected to play an important role in OEP studies of real materials.

III. FOUR CLASSES OF ORBITAL-DEPENDENT FUNCTIONALS

So far, we have discussed general aspects of using orbital-dependent functionals in density-functional theory. In the following, we discuss specific examples of orbital-dependent functionals. The particular functionals were chosen because they either reveal major fundamental concepts or have been proven to be of general practical usefulness.

A. The meta-generalized-gradient approximation

The meta-generalized-gradient-approximation (MGGA) functionals represent our first class of orbital-dependent functionals. A functional is called a MGGA if it uses the kinetic energy densities

$$\tau_\sigma(\vec{r}) = \sum_{j=1}^{N_\sigma} \frac{\hbar^2}{2m} |\vec{\nabla}\varphi_{j,\sigma}(\vec{r})|^2 \tag{59}$$

and/or the Laplacians $\nabla^2 n_\sigma$ in the construction of $E_{\text{xc}}[n]$. These quantities appear rather naturally in expansions of the exchange-correlation energy (see, e.g., Negele and Vautherin, 1972; Becke, 1998) and can be used in different ways. A popular route for MGGA development has been to devise an ansatz based on guiding principles such as hole properties, supplement it with a possibly large set of parameters, and then use the additional degrees of freedom (with respect to the GGA form) for fitting the functional to empirical data sets (see, e.g., Proynov *et al.*, 1994; Becke, 1998; Filatov and Thiel, 1998a, 1998b; Van Voorhis and Scuseria, 1998). An alternative approach has been to restrict the functional form as much as possible through enforcing known constraints on E_{xc} and related quantities (Tao *et al.*, 2007). The PKZB and TPSS functionals [named for the initials of their developers (Perdew *et al.*, 1999; Tao *et al.*, 2003)], which are mentioned in Sec. III.C.2, belong in this category.

The orbital dependence of the kinetic energy density in Eq. (59) makes the MGGA orbital dependent by definition. Yet they are close in spirit to the traditional semilocal, explicit density functionals as they use the orbitals in a mostly semilocal way. One may argue that a functional that is semilocal in the orbitals need not be semilocal in the density, as the link between orbitals and density is established by solving the Kohn-Sham equations. This link may be interpreted as introducing full nonlocality in terms of the density. Thus, a functional that is semilocal in the orbitals may be viewed as being fully nonlocal in the density from this perspective. However, experience so far seems to show that, in practice, semilocality in the orbitals is similar to semilocality in the density. Because of this similarity, MGGAs are addressed only briefly here.

An important feature of MGGA functionals is that they can achieve freedom from self-correlation errors. This means that their correlation functional is constructed such that the correlation energy vanishes for any one-electron density, and this property is typically achieved by employing iso-orbital indicators (Becke, 1998; Van Voorhis and Scuseria, 1998; Perdew *et al.*, 1999). An iso-orbital indicator is a functional expression that goes to a known, well-defined value in regions of space in which the density is dominated by one orbital. Such a functional expression usually depends on ingredients that are commonly used in density functionals, such as the density, its derivatives, and/or kinetic energy densities. Typically used iso-orbital indicators are combinations of the von Weizsäcker kinetic energy density

$$\tau_\sigma^W(\vec{r}) = \frac{\hbar^2}{8m} \frac{|\vec{\nabla} n_\sigma(\vec{r})|^2}{n_\sigma(\vec{r})} \tag{60}$$

and the orbital kinetic energy density of Eq. (59). For example, the ratio $\tau_\sigma^W(\vec{r})/\tau_\sigma(\vec{r})$ has been shown to equal unity in any iso-orbital region and be otherwise bounded between 0 and 1 (Kurth *et al.*, 1999). Thus, it is an iso-orbital indicator and, together with the condition that the density is fully spin polarized in a one-electron region, can be used as a "detector" for one-electron densities. With the help of these indicators, the condition that the correlation energy must vanish for any one-electron system can be built into MGGAs. Thus, MGGAs can satisfy the constraint of freedom from self-correlation that the LDA or GGAs do not. But the crucial problem of self-interaction is only slightly reduced in this way, because a major contribution to the self-interaction energy comes from the Hartree energy. This contribution must be canceled by the exchange part of the functional, i.e., whereas the correlation energy must vanish in a one-electron system, the exchange energy must not. Therefore, setting exchange to zero with the help of iso-orbital indicators is not an option. With respect to the question of which potential to employ in self-consistent calculations, a decision to use either orbital-specific potentials (Neumann *et al.*, 1996) or the OEP (Arbuznikov and Kaupp, 2003) must be made.

We will not go into a discussion of the performance of various MGGA functionals. However, we believe it is fair to say that, although the course of MGGA development was not a straight one, by now MGGA functionals are among or are even the most accurate semilocal functionals. Still, MGGAs are in spirit semilocal functionals in the following sense: their behavior in terms of freedom from self-interaction and the related issue of incorporating a derivative discontinuity, which were identified as two important aspects driving the development of orbital functionals, may be quantitatively different from the one of traditional (semi)local functionals—but typically it is similar qualitatively. Therefore, other classes of orbital-dependent functionals must be investigated.

B. Self-interaction correction

One of the oldest orbital-dependent functionals is the self-interaction-correction (SIC) scheme (Perdew, 1979; Perdew and Zunger, 1981). The idea behind this type of functional is straightforward. The exchange-correlation energy of the density corresponding to a single, fully occupied orbital $\varphi_{i,\sigma}(\vec{r})$ must exactly cancel that orbital's self-Hartree energy, i.e.,

$$E_H[n_{i,\sigma}] + E_{\text{xc}}[n_{i,\sigma},0] = 0, \tag{61}$$

where $n_{i,\sigma}(\vec{r}) = |\varphi_{i,\sigma}|^2$ is the orbital density. This is because the orbital density is a special case of a one-electron density, for which self-interaction is impossible. With the definition of the exact exchange energy of Eq. (37), the above equation (61) can be separated into the two conditions

$$E_H[n_{i,\sigma}] + E_x[n_{i,\sigma},0] = 0 \tag{62}$$

and

$$E_c[n_{i,\sigma},0] = 0, \tag{63}$$

i.e., in one-electron systems the Hartree self-interaction error must be canceled by the exchange part of the functional and the correlation energy must vanish. Given an approximate functional E_{xc}^{app} that is not free from self-interaction in the spirit of Eq. (61), one can define a new, self-interaction-corrected functional by

$$\begin{aligned} E_{xc}^{SIC,app}[n_\uparrow,n_\downarrow] = E_{xc}^{app}[n_\uparrow,n_\downarrow] - \sum_{i=1}^{N_\uparrow} (E_{xc}^{app}[n_{i,\uparrow},0] \\ + E_H[n_{i,\uparrow}]) - \sum_{i=1}^{N_\downarrow} (E_{xc}^{app}[0,n_{i,\downarrow}] \\ + E_H[n_{i,\downarrow}]). \end{aligned} \tag{64}$$

The idea behind this expression is that the self-interaction energy terms are subtracted on an orbital-by-orbital basis. If applied to the "ultimate functional," the correction of Eq. (64) vanishes as required. The intuitive plausibility of the SIC scheme is one of its most favorable features but also one of its major drawbacks: you can love it or leave it, but you cannot improve it systematically since it is an *ad hoc* scheme.

The SIC scheme can be applied to any given functional approximation, but it has predominantly been used to correct LSDA. One aspect, recognized early on (Harrison *et al.*, 1983), that makes the SIC scheme attractive is that it yields orbital energies that correspond more closely to binding energies and can be used to calculate excitations (Perdew and Zunger, 1981; Pederson *et al.*, 1985). In this context it should be noted, however, that one has to be careful about which of the quantities appearing in the SIC scheme are interpreted as orbital energies. Quantities that have been used for that purpose are, e.g., the eigenvalues of the matrix of Lagrangian multipliers that enforce orbital orthogonality (see Sec. II.B) or approximations thereof. But recently, Vydrov *et al.* (2007) have argued that it is not the eigenvalues of that matrix but rather its diagonal elements that are most suitable for physical interpretation.

One reason why the SIC orbital energies can be more physical than LSDA eigenvalues is that, with the removal of the self-energy error, the SIC potential shows the correct asymptotic decay $-e^2/r$ (cf. Sec. I.B). Another major improvement of SIC over LSDA is that, being explicitly orbital dependent, the SIC functional exhibits a derivative discontinuity. For a more detailed discussion of these features of the SIC scheme, see Perdew (1990) and Perdew and Ernzerhof (1998).

A significant difficulty with the SIC scheme is that the total SIC energy is not invariant under transformation of the orbitals, i.e., a unitary transformation in the space of occupied orbitals can change the energy even though the total density remains unchanged. This is a consequence of the SIC energy expression depending explicitly on the individual orbital densities. The lowest energy is typically found for the maximally localized orbitals because these maximize the correction terms on the right-hand side of Eq. (64). As an example, consider a periodic solid. The eigenfunctions of a periodic Hamiltonian are Bloch orbitals, and for these, the SIC vanishes (Perdew and Zunger, 1981). But when the Bloch orbitals are transformed to Wannier orbitals and the latter are used to evaluate Eq. (64), the SIC can have a major effect (Perdew and Zunger, 1981; Heaton *et al.*, 1982, 1983).

Because it makes such a big difference which orbitals are used to evaluate the SIC energy, it is important which potential(s) are used in self-consistent SIC calculations and how the relation between potential, orbitals, and energy is set up. The idea of Perdew and Zunger (1981) was to find the absolute minimum of the SIC energy using orbital-specific potentials. Since each orbital is subjected to a different Hamiltonian, the orbitals that are obtained in this way are generally not orthogonal. Orbital orthogonality can be enforced by using off-diagonal Lagrangian multipliers. However, Perdew and Zunger noted that, in their atomic calculations, the orbital-specific potentials lead to only slightly nonorthogonal orbitals and results changed little when an additional Gram-Schmidt orthogonalization was performed after each iteration. Recently, the effect of the orbital-specific potentials has been mimicked by effectively including SIC-like terms into a pseudopotential scheme (Vogel *et al.*, 1996; Stampfl *et al.*, 2000; Filippetti and Spaldin, 2003; Filippetti *et al.*, 2004).

Shortly after the SIC procedure was suggested, SIC calculations for molecules were performed (Pederson *et al.*, 1984, 1985). But, as discussed below, these calculations deviated from the original Perdew and Zunger concept. It was many years before schemes that followed the approach of direct minimization of energy—under the additional constraint of orbital orthogonality—were implemented and used for molecules (Goedecker and Umrigar, 1997; Polo *et al.*, 2002a, 2002b; Vydrov and Scuseria, 2004). One reason for this may be that, from a computational point of view, the orbital-specific potentials were frequently considered to be a drawback. They yield localized SIC valence orbitals that correspond to the localized bonds and lone pairs of chemical intuition (Goedecker and Umrigar, 1997), which is good in terms of interpreting results, but hinders the use of symmetry arguments to facilitate calculations. Typically, it is easier to calculate the (delocalized) orbitals that are consistent with the symmetry of the Hamiltonian (Pederson *et al.*, 1984).

One may therefore consider using the orbitals that are obtained from an OEP calculation to evaluate Eq. (64). However, it has been argued (Perdew and Zunger, 1981; Kümmel and Perdew, 2003c) that the resulting scheme is not size consistent. Size consistency is a universal principle in electronic-structure theory. Expressed simply, it

TABLE III. Comparison of self-consistent SIC calculations evaluating Eq. (64) directly with KLI orbitals (KLI) and with localized orbitals (loc). Shown are the total energy E, the lowest unoccupied orbital energy ε_{N+1}, the highest occupied one ε_N, and their difference Δ, in hartrees. Note that for Be and LiH there is no difference within the computed numerical accuracy, whereas noticeable differences are seen for CH_4 and H_2O, where localized orbitals are found. From Garza *et al.* (2000).

System	E^{loc}	$\varepsilon_{N+1}^{\text{loc}}$	$\varepsilon_N^{\text{loc}}$	Δ^{loc}	E^{KLI}	$\varepsilon_{N+1}^{\text{KLI}}$	$\varepsilon_N^{\text{KLI}}$	Δ^{KLI}
Be	−14.5783	−0.308	−0.179	0.130	−14.5783	−0.308	−0.179	0.130
LiH	−7.9911	−0.301	−0.151	0.150	−7.9911	−0.301	−0.151	0.150
CH_4	−40.3624	−0.564	−0.174	0.390	−40.1426	−0.480	−0.146	0.334
H_2O	−76.2816	−0.518	−0.216	0.302	−76.1154	−0.466	−0.196	0.269

means that if a large system consists of two subsystems that do not interact with each other, it must not make a difference whether the total energy is calculated for the whole system or as the sum of the energies of the two subsystems. The paradigm example for a non-size-consistent functional is that of Fermi and Amaldi (1934), which in DFT terms may be written as

$$E_{\text{xc}}^{\text{FA}}[n] = -NE_H\left[\frac{n}{N}\right], \tag{65}$$

where N is the total number of electrons. Equation (65) can be viewed as the earliest attempt at a self-interaction correction. On evaluating it, e.g., for a density that is the sum of two well-separated one-electron densities, the size consistency violation becomes evident: When Eq. (65) is applied to each of the one-electron systems separately, $E_{\text{xc}}^{\text{FA}}+E_H$ vanishes, as it should. But when it is applied to the combined system, i.e., when the total density is inserted into Eq. (65), it does not vanish. Size consistency is typically violated by functionals that depend on integrated properties such as the total number of electrons. Therefore, it is violated in SIC approaches that stress computational simplicity, such as the Fermi-Amaldi approach or the averaged density SIC (Legrand *et al.*, 2002). As discussed, e.g., in the latter paper, such approaches can work reliably when all orbitals are delocalized over a common global scale, as encountered, e.g., in electronically simple systems like sodium clusters. But, as demonstrated by Ciofini *et al.* (2003), they fail when localization effects are non-negligible. Similarly, the SIC scheme using Kohn-Sham orbitals from the exact SIC OEP directly has been considered to be non-size-consistent, because when applied to one atom, where all Kohn-Sham orbitals are localized, it delivers a properly self-interaction-corrected energy. But, applied to a periodic lattice of atoms separated by large lattice constants, where all Kohn-Sham orbitals are delocalized (Bloch-like), it produces no self-interaction correction to the energy of an atom, since the correction given by the sum in Eq. (64) will then vanish on a per atom basis (Perdew and Zunger, 1981; and see Körzdörfer and Kümmel, 2006, for a discussion of a subtlety in this argument).

An explicit manifestation of the differences that result from using localized or delocalized orbitals can be seen in the work of Garza *et al.* (2000). They calculated the KLI potential corresponding to the LSDA SIC energy expression for several atoms and molecules. Their calculations showed that the KLI orbitals are delocalized molecular orbitals, and that in some cases, the energy from the SIC scheme that directly used the KLI orbitals was considerably higher than the energy obtained with localized orbitals. Table III shows that, for instance, it makes no difference which orbitals are used to calculate the total energy and orbital energies of LiH, but it makes a noticeable difference for CH_4.

Steps to merge the two, quite different, above-mentioned approaches of orbital-specific potentials and OEP—and to combine their respective benefits—were taken rather early. Harrison *et al.* (1983) and Pederson *et al.* (1984, 1985) proposed calculating orbitals from a common, unified Hamiltonian in a first step, and then subjecting these orbitals to a specific localizing transformation in a second step. Thus, a SIC calculation of this type works with two sets of orbitals: The molecular orbitals (or Bloch-like orbitals in solids) that have the full symmetry of the Hamiltonian, and a second set of localized orbitals that are used to evaluate the SIC energy expression. Garza *et al.* (2000) suggested identifying the symmetric orbitals with the Kohn-Sham orbitals; effectively, this concept brings the SIC scheme back under the umbrella of Kohn-Sham theory. The computational procedure—calculate the molecular or Bloch orbitals from the OEP, perform the unitary transformation that achieves maximal localization of these orbitals, then use the obtained maximally localized orbitals to evaluate the SIC expressions—corresponds to a legitimate way of defining an implicit density functional using the Kohn-Sham orbitals (Sec. II.B). The studies of Garza *et al.* (2000), Patchkovskii *et al.* (2001), and Patchkovskii and Ziegler (2002) are close in spirit to this approach. However, they make additional approximations by using the KLI form for the OEP and a localization procedure that is not necessarily the maximally localizing one, but is established and reasonable (Foster and Boys, 1960). In the context of SIC, localization transformations have been discussed, for example by Edmiston and Ruedenberg (1963) and Fois *et al.* (1993).

There is an extensive body of literature on applications of the SIC scheme, and we focus here on some hallmark examples. In the solid-state community, SIC has successfully been used for describing systems that can be characterized as strongly correlated, where semilocal functionals almost invariably fail; see Sec. I.B. This important feature of the scheme is related to its ability to localize orbitals, as demonstrated, for example, in the pioneering band-structure calculations of transition-metal oxides by Svane and Gunnarsson (1990). Later solid-state investigations included studies of the energy differences between divalent and trivalent states in the rare earths (Strange *et al.*, 1999), studies of the electronic structure of high-T_c superconductors (Temmerman *et al.*, 2001), and studies of dilute magnetic semiconductors (Filippetti *et al.*, 2004; Schulthess *et al.*, 2005). These successes of the SIC scheme in solid-state applications are pitted against less favorable results for molecules. Goedecker and Umrigar (1997) confirmed the Perdew and Zunger result that, for atoms, the LDA with SIC yields more realistic results than the LDA, but showed that, for a set of small organic molecules, the bond lengths obtained from the LDA with SIC are considerably less realistic than the ones obtained from the pure LDA. Furthermore, while the SIC significantly improves orbital energies (Pederson *et al.*, 1985), the energy barriers to chemical reactions (Johnson *et al.*, 1994; Csonka and Johnson, 1998; Patchkovskii and Ziegler, 2002), the description of molecular transport (Toher *et al.*, 2005), and many nuclear magnetic resonance properties (Patchkovskii and Ziegler, 2002), net reaction energies are not much improved relative to the LSDA (Patchkovskii and Ziegler, 2002). Finally, a systematic study (Vydrov and Scuseria, 2004) of thermochemical properties such as enthalpies of formation for a test set of small organic molecules showed that the LDA with SIC is an improvement over the LDA, but still considerably less accurate than a typical GGA. Yet more discomforting, if the SIC is applied to any other functional than the LDA, e.g., to typical GGAs, the agreement with experimental data can be improved in certain cases (Blake and Metiu, 1999), but on average is worse in comparison to the uncorrected functional (Vydrov and Scuseria, 2004).

From these findings one can conclude that, on the one hand, the SIC scheme shows features that are quite desirable in a functional, but, on the other hand, it is again a functional that is far from being a panacea. A clue to the less-than-perfect performance of SIC schemes in questions of energetics is the observation that part of the self-interaction error of semilocal exchange functionals is needed to model correlation effects, in particular static correlation (Cremer, 2001; Cremer *et al.*, 2002; Polo *et al.*, 2002a, 2002b; Gräfenstein *et al.*, 2004). The connection between semilocal exchange and static correlation is explored in greater detail in Sec. III.C. However, we note that the improved reaction barriers that SIC yields can be rationalized by this observation: Because transition states typically involve less static correlation than ground states, taking static correlation out via the SIC can improve their description. But at the same time it can worsen the ground-state description.

At first glance one may be puzzled that Eq. (64), which appears to be intuitively plausible and correct, can actually introduce something wrong into the theory. However, one must be aware of the fact that, while one can easily quantify the self-interaction error of a given density-functional approximation in a one-electron system, there is no similarly easy way of doing the same in a many-electron system (Ruzsinszky *et al.*, 2007; Vydrov *et al.*, 2007). This observation is at the bottom of why the self-interaction problem for many-electron systems, despite appearing so simple at first, is so hard to solve in practice. By defining the SIC of Eq. (64), one implicitly has also defined one particular way of turning the one-electron equations (61)–(63) into a definition of a self-interaction error in a many-electron system. This is by no means the only way to define a SIC, and alternatives have been proposed (Cortona, 1986; Whitehead, 1997; Lundin and Eriksson, 2001; Unger, 2001; Novàk *et al.*, 2003; Perdew *et al.*, 2005; Vydrov, Scuseria, Perdew, *et al.*, 2006). While a detailed discussion of these different schemes is beyond the scope of the present paper, the conclusion which seems to emerge from these studies is that one must take into account the fact that the self-interaction of semilocal functionals mimics static correlation. Thus, it has been suggested that, in order to obtain a SIC scheme that yields more realistic results for equilibrium properties, the self-interaction error in a many-electron system should be scaled down instead of removed completely (Vydrov, Scuseria, Perdew, *et al.*, 2006). Of course, any modified version of the SIC scheme must still obey the one-electron equations (61)–(63). One possible way to achieve this may be the use of iso-orbital indicators (Perdew *et al.*, 2005; Vydrov, Scuseria, Perdew, *et al.*, 2006), discussed in Sec. III.A. However, the practical success of such schemes is unclear (Ruzsinszky *et al.*, 2006, 2007), and it should not be forgotten that the ultimate self-interaction correction could be the use of energy density functionals that employ exact exchange with compatible correlation (see Sec. III.D).

Referring to the Jacob's ladder of density functionals once more, one may conclude that the SIC scheme does not fit into the successive rungs picture. It is more akin to an additional rope of unpredictable strength dangling besides the ladder. Depending on the problem one wishes to solve, this rope may be a clever way up, or it may fail, leading to a bad fall.

C. Hybrid functionals

The physical principles upon which hybrid functionals are based are best understood within the framework of a fundamental DFT concept known as the adiabatic connection theorem (Harris and Jones, 1974; Gunnarson and Lundqvist, 1976; Langreth and Perdew, 1977; Levy, 1983; Harris, 1984). We therefore begin our review of hybrid functionals with a brief overview of this important theorem.

1. Adiabatic connection and coupling-constant integration

As explained in Sec. I.A, the Kohn-Sham equations stem from an exact mapping between the physical interacting-electron system and a fictitious noninteracting-electron system. This mapping can also be viewed in terms of an “adiabatic” turn-on of the electron-electron repulsion, as follows. Consider a many-electron Hamiltonian in the form

$$\hat{H}_\lambda = \hat{T} + \hat{V}_{\text{ext},\lambda} + \lambda \hat{V}_{\text{ee}}, \tag{66}$$

where $\hat{T} \equiv -(\hbar/2m)\Sigma_i \nabla_i^2$ and $\hat{V}_{ee} \equiv \Sigma_{i<j} e^2/|\vec{r}_i - \vec{r}_j|$ are the usual many-electron kinetic and electron-electron energy operators, respectively, $\hat{V}_{\text{ext},\lambda} \equiv \Sigma_i v_{\text{ext},\lambda}(\vec{r}_i)$ is an external potential operator, and $0 \leqslant \lambda \leqslant 1$. We choose $v_{\text{ext},1}(\vec{r})$ to be $v_{\text{ion}}(\vec{r})$ of Eq. (1), so that $\hat{H}_1$ is the exact, physical many-electron Hamiltonian. For any other λ, we choose $v_{\text{ext},\lambda}(\vec{r})$ such that it yields the same ground-state density as that of the physical system. The Kohn-Sham Hamiltonian is then a special case of Eq. (66) with $\lambda = 0$, i.e., $v_{\text{ext},0}(\vec{r}) = v_{\text{KS}}(\vec{r})$.

The Hohenberg-Kohn theorem guarantees that for any choice of λ one can extend Eq. (2) in the form

$$E_{\text{tot},\lambda} = \int v_{\text{ext},\lambda}(\vec{r}) n(\vec{r}) d^3 r + F_\lambda[n]. \tag{67}$$

If we denote by Ψ_λ the normalized, antisymmetric many-electron wave function that minimizes the expectation value of $\hat{T} + \lambda \hat{V}_{ee}$ while yielding a given $n(\vec{r})$, then (Levy, 1982, 1983)

$$F_\lambda[n] = \langle \Psi_\lambda | \hat{T} + \lambda \hat{V}_{\text{ee}} | \Psi_\lambda \rangle. \tag{68}$$

Using Eq. (68), we note that the first term on the right-hand side of Eq. (6) is simply $F_1[n]$. The second term is exactly $F_0[n]$, because by construction the Kohn-Sham orbitals minimize the noninteracting kinetic energy density. Taken together,

$$\begin{aligned} E_{\text{xc}}[n] &= F_1[n] - F_0[n] - E_H[n] \\ &= \int_0^1 \frac{\partial F_\lambda[n]}{\partial \lambda} d\lambda - E_H[n] \\ &\equiv \int_0^1 E_{\text{xc},\lambda}[n] d\lambda, \end{aligned} \tag{69}$$

where the last step serves as the definition of a new functional, namely, the coupling-constant integrand $E_{\text{xc},\lambda}[n]$. Using the Hellmann-Feynman theorem to evaluate the derivative of Eq. (68), we readily obtain

$$E_{\text{xc},\lambda}[n] = \langle \Psi_\lambda | \hat{V}_{\text{ee}} | \Psi_\lambda \rangle - E_H[n]. \tag{70}$$

$E_{\text{xc},\lambda}[n]$ does not involve a kinetic energy term. Thus, expressing $E_{\text{xc}}[n]$ via Eq. (69) can be viewed as circumventing the problem of an explicit evaluation of the kinetic energy contribution, at the cost of having to evaluate the potential energy contributions of all adiabatic connection states between the Kohn-Sham system and the physical one.

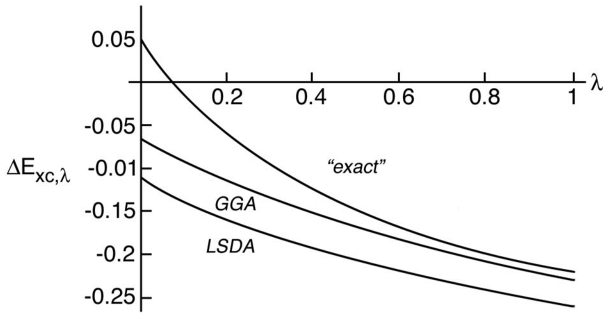

FIG. 8. $\Delta E_{\text{xc},\lambda} \equiv E_{\text{xc},\lambda}|_{\text{N}_2} - 2\, E_{\text{xc},\lambda}|_{\text{N}}$, in hartrees, as a function of λ. The three curves shown are a sketch of the trend expected for the exact functional and the results of the LSDA and GGA calculations. Adapted from Ernzerhof *et al.*, 1997. Reprinted with permission of Wiley & Sons, Inc.

Several exact properties of $E_{\text{xc},\lambda}[n]$ that are important for the functional construction arguments below are known. First, at the noninteracting Kohn-Sham limit $\lambda = 0$, correlation is zero by definition and Ψ_λ reduces to Ψ_0, the Slater determinant composed of the ground-state Kohn-Sham orbitals. Thus, the exchange part of $E_{\text{xc},\lambda}[n]$ of Eq. (70) is

$$E_x[n] = E_{\text{xc},0}[n] = \langle \Psi_0 | \hat{V}_{\text{ee}} | \Psi_0 \rangle - E_H[n] \tag{71}$$

and does not depend on λ. This reflects the fact that exchange is a consequence of the antisymmetry requirement, which is independent of the coupling strength. $E_x[n]$ sets the lower end of the coupling-constant integrand (or adiabatic connection curve) $E_{\text{xc},\lambda}$ and is given by the Kohn-Sham exact-exchange energy; see Eq. (37) and Sec. II.B.

Second, at the physical system limit $\lambda = 1$, Ψ_λ is the true many-electron wave function. Thus, $E_{\text{xc},1}[n]$ is similar to $E_{\text{xc}}[n]$ of Eq. (6) in the sense of accounting for all electron-repulsion interactions beyond simple electrostatics, but unlike $E_{\text{xc}}[n]$ it does not include kinetic energy corrections. Third, using variational arguments one can show that $E_{\text{xc},\lambda}[n]$ is a monotonically decreasing function of λ (Levy and Perdew, 1985). Fourth, scaling arguments show that $E_{\text{xc},\lambda}[n]$ is related to $E_{\text{xc}}[n]$ by (Levy *et al.*, 1996)

$$E_{\text{xc},\lambda}[n] = \frac{d}{d\lambda} \{ \lambda^2 E_{\text{xc}}[n(\vec{r}/\lambda)/\lambda^3] \}. \tag{72}$$

To gain more insight into $E_{\text{xc},\lambda}[n]$, consider the example of the N_2 molecule (Ernzerhof *et al.*, 1997). The dependence of $\Delta E_{\text{xc},\lambda}[n] \equiv E_{\text{xc},\lambda}[n]|_{N_2} - 2\, E_{\text{xc},\lambda}[n]|_N$ on λ is shown in Fig. 8. This is the exchange-correlation contribution to the binding energy. Recall that, in light of Eq. (69), the physically interesting $\Delta E_{\text{xc}}[n]$ is simply the area under the $\Delta E_{\text{xc},\lambda}[n]$ curve between $\lambda = 0$ and 1. Comparison of the LSDA and the (Perdew-Wang) GGA approximations for $\Delta E_{\text{xc},\lambda}[n]$, both obtained using Eq. (72), shows that the GGA compensates for the overbinding of the LDA (see Sec. I.B) by an essentially uniform shift of the $\Delta E_{\mathrm{xc},\lambda}[n]$ curve. However, the GGA curve must still be wrong. While it is expected to be a reasonable approximation at the physical limit of $\lambda=1$, the semilocal exchange of the GGA does not approximate well the nonlocal exact exchange at $\lambda=0$. Evaluation of the exact-exchange integrals using the GGA Kohn-Sham orbitals (under the assumption that they do not differ drastically from the exact ones) shows that the GGA-calculated $\Delta E_{\mathrm{xc},0}[n]$ is far too negative. This immediately explains why the GGA still overbinds the N_2 molecule.

2. The hybrid approach

Becke (1993a) realized that the adiabatic connection formalism can serve as a practical tool for functional construction. He suggested that the area under the $E_{\mathrm{xc},\lambda}[n]$ curve, namely, $E_{\mathrm{xc}}[n]$, can be approximated by a trapezoidal rule integration between $\lambda=0$ and 1, with exact exchange used for the former and the LSDA for the latter, i.e.,

$$E_{\mathrm{xc}} = 0.5E_x^{\mathrm{exact}} + 0.5E_{\mathrm{xc},1}^{\mathrm{LSDA}}[n]. \qquad (73)$$

The improvement in accuracy afforded by Eq. (73), if any, with respect to modern GGAs is usually too modest to warrant its practical use. The significance of this equation lies in drawing attention to the potential benefits of functionals that combine a fixed fraction of exact exchange with a fixed fraction of semilocal exchange and correlation. Becke (1993a) has dubbed such functionals hybrid functionals, viewing them as a true mixture of the Hartree-Fock and DFT approaches. We emphasize, however, that, as also recognized by Becke (1993a), a functional with a fractional contribution of exact exchange is just a special case of an orbital-dependent functional. As such, it is well within DFT and there is nothing inherently hybrid about it. The term is still appropriate because, in the context of hybrid functionals, one does not use the OEP equation for the Kohn-Sham potential. Instead, one uses a sum of the Kohn-Sham potential stemming from the semilocal portion of the functional and a fraction of the orbital-specific Fock potential operator coming from the fractional exact-exchange part. In general, this usually works partly because of the first-order similarity in densities between the orbital-specific and OEP potentials (cf. Sec. II.B, but also note the differences discussed there) and partly because any remaining error can be buried in the coefficients of the exchange and semilocal terms, which are determined empirically in many cases (see below).

A significant improvement over the simple half-and-half approach of Eq. (73) can be obtained by improving the approximate integration of $E_{\mathrm{xc},\lambda}[n]$. Given the monotonic decrease of $E_{\mathrm{xc},\lambda}[n]$, the mean-value theorem guarantees that for any system one can always find a coefficient b such that

$$E_{\mathrm{xc}} = bE_x^{\mathrm{exact}} + (1-b)E_{\mathrm{xc},1}. \qquad (74)$$

Practical use of Eq. (74) requires an approximate semilocal form for $E_{\mathrm{xc},1}$, but using an approximation for E_{xc} is more convenient. This can be achieved by using $E_{\mathrm{xc},1} = E_x^{\mathrm{exact}} + E_{c,1}$ in Eq. (74), which leads to the identity $E_c = (1-b)E_{c,1}$. Thus, the practically useful form of Eq. (74) is

$$E_{\mathrm{xc}}^{\mathrm{hyb}} = bE_x^{\mathrm{exact}} + (1-b)E_x^{\mathrm{app}} + E_c^{\mathrm{app}}, \qquad (75)$$

where the superscripts hyb and app denote the hybrid functional and the approximate (typically semilocal) functional from which it is constructed.

Interestingly, Eq. (75) mixes the approximate and exact exchange linearly, but the approximate correlation is taken as is. This can also be motivated physically. In quantum chemistry, it is common to distinguish between two types of correlation (Sinanoğlu, 1964). *Dynamical* correlation is essentially due to the reduction expected in the many-electron wave function when two electrons approach one another. Because of screening, it is inherently a short-range effect. From a wave-function perspective, dynamical correlation is related to the cusp condition and results from linear combinations of many Slater determinants. *Static* correlation arises from degeneracies or near degeneracies of several Slater determinants, i.e., it is an effect driven by just a few Slater determinants. In molecules, an important special case of static correlation is the left-right correlation, which accounts for the degeneracy of electrons localized on equivalent atoms, and which is needed if electrons are to localize on separate atoms as the molecule dissociates.

Semilocal correlation is inherently short range and cannot account for long-range static correlation. Indeed, by comparing LDA or GGA correlation energies to wave-function-based calculations for first-row dimers, Mok *et al.* (1996) concluded that semilocal correlation agrees closely with dynamical correlation. For several first-row dimers, Gritsenko *et al.* (1997) compared wave-function-based calculations with both exact Kohn-Sham solutions reconstructed from the wave-function calculations and LDA and GGA calculations. They reached similar conclusions, but further showed that the difference between the semilocal GGA exchange and the exact Kohn-Sham exchange arises from the effective inclusion of static correlation in the GGA exchange. Static correlation is known to be significant in multiply bonded molecules such as the triply bonded N_2 molecule of Fig. 8. The extra contribution from static correlation explains, then, why the GGA value for $E_{\mathrm{xc},0}$ in Fig. 8 is much more negative than the exact one.

The above argument can be rationalized further using exchange-hole considerations. If a reference electron is placed close to one nucleus of a dimer, the exchange hole is delocalized over both nuclear centers, irrespective of their separation. This delocalization at large distances is not physical and is canceled by a similar nonphysical delocalization in the exact-correlation hole. Together, the two holes construct a localized exchange-correlation hole near the reference position. In semilocal functionals, this localization is inherent, so that the semilocal exchange must contain much of the static correlation [in fact, an empirical semilocal exchange form, parametrized against static correlation, has been suggested (Handy and Cohen, 2001; Molawi *et al.*, 2002)]. Thus, dynamical correlation is well modeled by semilocal correlation and is better left alone. But semilocal exchange differs from exact exchange by static correlation, and therefore a linear mix of the two, as in Eq. (75) for the approximation of the adiabatic connection integral of Eq. (69), is reasonable.

What value of mixing parameter b should one choose and what improvement over the LSDA or GGA, if any, is obtained? Becke (1996), who suggested the practical usage of Eq. (75), determined b empirically by least-squares fitting to the atomization energies, ionization potentials, and proton affinities of a canonical set of mostly organic molecules, based on first- and second-row atoms, known as the (modified) G2 set. This is one of several such sets, known as Gn ($n=1$–3, where a larger index n implies a richer set of data), used frequently in the quantum chemistry community for testing the performance of first-principles approaches and parametrizing semiempirical methods (Pople *et al.*, 1989; Curtiss *et al.*, 1990, 1997, 1999, 2000). Becke found the optimal value of b to be 0.16 or 0.28, depending on the type of semiempirical correlation functional used for the fit. These values resulted in mean atomization errors of only 0.13 or 0.09 eV, respectively, for the G2 set, proving that Eq. (75) is in principle a useful form.

Perdew, Ernzerhof, and Burke (1996) rationalized the empirical values found for b by proposing a simple model for $E_{\mathrm{xc},\lambda}$ in the form

$$E_{\mathrm{xc},\lambda}^{\mathrm{hyb}} = E_{\mathrm{xc},\lambda}^{\mathrm{app}} + (E_x^{\mathrm{exact}} - E_x^{\mathrm{app}})(1-\lambda)^{n-1}. \tag{76}$$

If used for evaluating the adiabatic connection integral of Eq. (69), this yields $b=1/n$. Perdew *et al.* conjectured that, because for typical molecules in the data sets used for fitting b, post-Hartree-Fock fourth-order Møller-Plesset perturbation theory yields atomization errors with a small mean value, $n=4$ would be an adequate value for Eq. (76). This yields $b=0.25$, in reasonable agreement with Becke's work.

As already hinted by Becke's results, whether or not $b=0.25$ is a practically useful value has been found to be a strong function of the type of GGA used (Adamo and Barone, 1997, 1998, 1999). Fortunately, for the nonempirical PBE GGA, $b=0.25$ is found to be an excellent value (Adamo and Barone, 1999; Ernzerhof and Scuseria, 1999). For example, for the largest G3 set, the mean error obtained for the enthalpies of formation (which are closely related to the atomization energies) with the PBE GGA is 0.94 eV, whereas it is only 0.20 eV if Eq. (75) is used with the PBE GGA and $b=0.25$ (Staroverov *et al.*, 2003). Interestingly, such remarkable improvement is not achieved if the earlier nonempirical GGA of Perdew and Wang is used instead (Adamo and Barone, 1999).

Leaving fundamental arguments behind and turning to practical issues, the most common hybrid functionals in use are, in fact, not of the one-parameter form suggested by Eq. (75). Rather, they are of the following three-parameter form:

$$E_{\mathrm{xc}}^{\mathrm{hyb}} = E_{\mathrm{xc}}^{\mathrm{LSDA}} + a_0(E_x^{\mathrm{exact}} - E_x^{\mathrm{LSDA}}) + a_x \Delta E_x^{\mathrm{GGA}} + a_c \Delta E_c^{\mathrm{GGA}}, \tag{77}$$

where $\Delta E_x^{\mathrm{GGA}}$ and $\Delta E_c^{\mathrm{GGA}}$ are the exchange and correlation gradient-containing correction terms to the LSDA exchange and correlation, respectively, and a_0, a_x, and a_c are constant parameters. This hybrid functional form was suggested by Becke (1993b) shortly after his half-and-half hybrid (and, in fact, before the one-parameter functionals discussed above, which were presented first for pedagogical reasons). For two special cases—$\{a_0=b, a_x=0, a_c=0\}$, $\{a_0=b, a_x=1-b, a_c=1\}$—Eq. (77) reduces to Eq. (75) with $E_{\mathrm{xc}}^{\mathrm{app}}$ given by $E_{\mathrm{xc}}^{\mathrm{LSDA}}$ (Moreira *et al.*, 2002) or $E_{\mathrm{xc}}^{\mathrm{GGA}}$ (Becke, 1996), respectively. More generally, it allows for "optimum admixtures of exchange and correlation gradient corrections" (Becke, 1993b), i.e., one can use the greater variational freedom offered by Eq. (77) to further reduce mean errors. With a specific choice for the GGA exchange (Becke, 1988) and correlation (Perdew, 1991) gradient corrections, and fitting to the G1 data set, Becke found the optimal parameters to be $a_0=0.2$, $a_x=0.72$, and $a_c=0.81$. This functional is usually dubbed B3PW91 (Becke three-parameter Perdew-Wang 1991).

Stephens *et al.* (1994) also used Eq. (77), but with the Lee, Yang, and Parr (LYP) GGA correlation functional (Lee *et al.*, 1988), to create a functional called B3LYP. Instead of separating the LYP functional into a LSDA part and an additive gradient correction, they multiplied the entire LYP correlation by a_c and added a compensating LSDA correlation term multiplied by $1-a_c$. Furthermore, they used Becke's values for a_0, a_x, and a_c without parameter reoptimization. Both the B3PW91 and B3LYP functionals result in formation enthalpies for the G2 that are smaller than those obtained with the PBE hybrid, 0.06 and 0.05 eV, respectively (Staroverov *et al.*, 2003).

Because of its slightly better performance, B3LYP has rapidly become the *de facto* "workhorse," or default choice, for practical applications of hybrid functionals, at least within the organic chemistry community, despite its *ad hoc* construction. Its disadvantages, however, also need to be noted. First, for the larger G3 set, the B3PW91 functional actually outperforms B3LYP, with a mean formation enthalpy error of 0.09 eV versus 0.15 eV for B3LYP (Staroverov *et al.*, 2003). More seriously, the use of the LYP correlation results in a functional that does not reduce to the LSDA in the limit of a uniform electron gas. This formal deficiency—violation of one of the known limits for the exact functional—is absent with either the B3PW91 functional or with the one-parameter functionals based on either the LSDA or PBE GGA. This may translate into a very practical deficiency, i.e., errors in applications to solids and surfaces (Perdew *et al.*, 2005).

Importantly, three-parameter functionals based on Eq. (77) are still philosophically inspired by the adiabatic connection theorem, but, unlike the one-parameter functionals of Eq. (75), they are no longer rigorously rooted within it. Thus, they are best viewed as semiempirical functionals. Because they typically contain semiempirical GGA expressions for the correlation (B3LYP) and exchange (both B3PW91 and B3LYP), the effective number of fitting parameters in them is larger than three (four for B3PW91, eight for B3LYP). An obvious question, then, is whether these parameters really need to be optimized separately, or whether one can gain additional accuracy by optimizing all semiempirical parameters simultaneously within a physically reasonable approximate form, using one of the G*n* sets (Becke, 1997; Adamson *et al.*, 1998). There are serious arguments against such an approach. First, it strays even further from the spirit of the adiabatic connection theorem. Second, from a purist point of view, this actually defeats the logic of DFT as a first-principles approach that does not rely on experimental data. There also exist, however, strong arguments in favor of this approach. First, from a practical point of view, the real issue is whether these functionals possess predictive power, rather than the nature of their derivation. Second, because the exchange-correlation energy is a unique functional of the density, one can argue that the method through which this universal functional is approached is irrelevant (Becke, 1997).

Practical hybrid functionals based on simultaneous parameter optimization of fractional exact exchange and empirical GGA components were explored by Becke (1997) and have been further developed with increasingly sophisticated functional forms and parameter optimization procedures, richer experimental training sets, and training sets based additionally on computational results obtained using wave-function-based methods (see, e.g., Hamprecht *et al.*, 1998; Cohen and Handy, 2001; Wilson *et al.*, 2001; Xu and Goddard, 2004; Lynch *et al.*, 2004; Keal and Tozer, 2005). Such functionals can reduce mean errors further and may be useful for specific systems. Unfortunately, they do not result in significant breakthroughs in accuracy with respect to that of the one-parameter and three-parameter hybrids. Indeed, Becke (1997) concluded that accuracies of the order reported above for, e.g., B3LYP, are the most that can be expected by mixing a fraction of exact exchange with the LSDA and GGA components. At the same time, highly parametrized functionals run the risk of overfitting, which may result in poor performance outside the training set. For example, highly parametrized functionals, on average, perform nearly as poorly as Hartree-Fock theory in predicting $4s$-$3d$ interconfigurational excitation energies for $3d$ cations, whereas one- and three-parameter hybrids still provide for an accuracy that exceeds that of the LSDA or GGA (Holthausen, 2005).

Most of the numbers provided so far were in the context of formation energies, because the hybrid functionals surveyed so far were developed based largely on thermochemical data. Obviously, there is more to DFT than thermochemistry. The importance of hybrid functionals is that, at the very least for molecules involving first- and second-row atoms and often for other systems as well, they additionally exhibit an accuracy superior to that of the LSDA or GGA for a wide range of properties for which they were not fitted. Examples include molecular structures, vibration frequencies, electrical properties (e.g., dipole moments, polarizabilities), magnetic properties (e.g., chemical shifts in nuclear magnetic resonance experiments, hyperfine coupling constants), and more [see Koch and Holthausen (2001) for a comprehensive performance evaluation].

It is important to recognize that, although clearly of value and use, hybrid functionals are no panacea. One obvious caveat is that their performance for scenarios that are sufficiently different from those used in their original training or validation data should be assessed critically against experiment and/or wave-function-based methods before they are used for prediction. For example, B3LYP has been shown to do poorly in predicting the ground-state geometry of Jahn-Teller-distorted cations (Baboul *et al.*, 1999). For solids, it is often found that a different fraction of exact exchange is needed to reproduce experimental values of different properties, and in general a larger fraction than that typically used in molecular applications is necessary [see Corá *et al.* (2004) for a comprehensive evaluation of the performance of hybrid functionals for solid-state properties]. Furthermore, hybrid functionals often perform worse than the GGA for metals in general and for transition metals in particular (Corá *et al.*, 2004; Paier *et al.*, 2006). However, hybrid functionals are no panacea even for systems that are chemically similar to those used for fitting their parameters. Specifically, because they contain only a *fraction* of exact exchange, they are not self-interaction-free nor do they generally possess the correct derivative discontinuity. This, in principle, exposes them to many of the difficulties described for semilocal functionals in Sec. I.B. For example, they possess a spurious maximum for the dissociation curve of H_2^+ just like the LSDA or GGA (Bally and Sastry, 1997) and are unsuccessful at the description of charge-transfer systems, e.g., in the prediction of bond lengths for oligomers exhibiting large bond-length alterations (Jacquemin *et al.*, 2006).

Interestingly, because hybrid functionals do contain a certain fraction of exact exchange, together with an empirical amount of correlation, they often handle localized states well despite not being rigorously self-interaction-free. For example, a correct description of the geometry relaxation leading to localized defect states on reduced and hydroxylated $TiO_2(110)$ is achieved with the B3LYP, but not with the PBE functional (di Valentin *et al.*, 2006). In a similar vein, the eigenvalue spectrum of the PTCDA molecule of Fig. 2, already shown above to be inadequate if the LDA or GGA is used, agrees remarkably well with the *GW* calculation shown in Fig. 2 if B3LYP is used—at a fraction of the computational cost of the *GW* calculation (Dori *et al.*, 2006). As a different example, the one-parameter PBE hybrid, as is or in its screened hybrid version explained below, often provides for a good description of the electronic structure of solids with localized f electrons, e.g., CeO_2 and Ce_2O_3 (Hay *et al.*, 2006), actinide oxides (Kudin *et al.*, 2002; Prodan *et al.*, 2006), and uranium fluorides and chlorides (Peralta *et al.*, 2005).

An important arena where the hybrid functionals surveyed so far offer consistently insufficient improvement over semilocal forms is the description of transition barriers, which are crucial for predicting the kinetics of chemical reactions. This is the case even for molecules well within the original training set. Because the amount of static correlation is reduced in transition states, the percentage of semilocal exchange used to mimic it is now too large. Therefore, semilocal functionals grossly underestimate transition barriers, to the point where they often erroneously predict barrier-free reactions (Baker *et al.*, 1995). Conventional hybrid functionals increase the barrier height, but not nearly to experimental values. The reduction of static correlation in transition states means that the fraction of exact exchange needed for quantitative prediction of transition barrier heights is higher than that needed for thermochemistry—typically 40–50 % (see, e.g., Orlova and Goddard, 2002; Lynch *et al.*, 2004). Similarly, a need for a larger fraction of exact exchange has been observed in, for example, the determination of torsional potentials in π-conjugated molecules (Sancho-García *et al.*, 2003) and in the description of self-trapped holes in NaCl (Gavartin *et al.*, 2003).

Obviously, increasing the exact-exchange fraction means that the accuracy for thermochemical and other ground-state properties is diminished. Functional parametrization against a training set consisting of both thermochemical and transition-state data mitigates the errors (Lynch *et al.*, 2004; Keal and Tozer, 2005). Nevertheless, one cannot truly negotiate the inherently contradictory need for different fractions of exact exchange in different situations without introducing additional degrees of freedom.

A natural additional degree of freedom is afforded by using the kinetic energy density τ, as in meta-GGA functionals (see Sec. III.A). Indeed, this idea was suggested in the development of hybrid functionals (Becke, 1996, 1998; Schmider and Becke, 1998). A one-parameter MGGA hybrid, based on Eq. (75), using either the PKZB (Adamo *et al.*, 2000) or the TPSS (Staroverov *et al.*, 2003) nonempirical MGGA functional as the approximate functional in that equation, was suggested as a straightforward way of combining the MGGA and hybrid approaches. The optimal exact-exchange fraction turns out to be only 4.1% and 10% for PKZB and TPSS, respectively. On the one hand, this can be taken as an indication that MGGAs are already doing something right in terms of mimicking the adiabatic connection curve (Staroverov *et al.*, 2003). On the other hand, performance gain, as compared to the pure MGGA functional, is minimal (see, e.g., Zhao and Truhlar, 2006). Semiempirical hybrid meta-GGA functionals, where τ is used as an ingredient in the expression for the exchange, the correlation, or both, are a very active area of current research. Balancing the thermochemistry and transition-barrier performance of such functionals, especially for transition metal chemistry, is still difficult, but steady progress is being made (Boese and Handy, 2002; Boese and Martin, 2004; Zhao and Truhlar, 2004, 2005a, 2005b, 2006; Zhao *et al.*, 2004, 2005, 2006; Boese *et al.*, 2005; Proynov and Thakkar, 2006; Hemelsoet *et al.*, 2006).

A different interesting approach to gaining additional flexibility is to partition the exchange term into short-range (SR) and long-range (LR) parts, by means of splitting the Coulomb repulsion into a short-range and a long-range component (see Sec. III.D.3 for practical partition schemes). One can then generalize the hybridization scheme by invoking different weights for the short- and long-range exact exchange (Yanai *et al.*, 2004; Peach, Cohen, and Tozer, 2006; Peach, Helgaker, Sałek, *et al.*, 2006; Vydrov, Heyd, Krukau, *et al.*, 2006). Such generalization of, e.g., the one-parameter hybridization scheme of Eq. (75) yields

$$E_{xc}^{\text{hyb}} = aE_x^{\text{SR exact}} + (1-a)E_x^{\text{SR app}} + bE_x^{\text{LR exact}} + (1-b)E_x^{\text{LR app}} + E_c^{\text{app}}. \tag{78}$$

One useful limit of Eq. (78) is obtained by setting $b=0$, i.e., excluding the long-range portion of the exact exchange altogether (Heyd *et al.*, 2003). This is important because an accurate treatment of the long-range exchange interaction in a periodic system is possible but nontrivial. In fact, this issue was at least partly responsible for the delay between the introduction of hybrid functionals and their first applications to solid-state problems (see, e.g., Bredow and Gerson, 2000; Dovesi *et al.*, 2000; Muscat *et al.*, 2001). The PBE-based screened hybrid obtained with $b=0$ (denoted HSE after the initials of its developers) greatly facilitates calculations of solids with hybrid functionals, with an accuracy that is usually on par with or better than that obtained with the usual PBE hybrid (Heyd *et al.*, 2003; Heyd and Scuseria, 2004a, 2004b; Krukau *et al.*, 2006; Batista *et al.*, 2006; Paier *et al.*, 2006). Specifically, it has been found to provide an excellent description of the electronic structure of both semiconducting and metallic single-wall nanotubes and of graphene nanoribbons (Barone, Peralta, and Scuseria, 2005; Barone, Peralta, Wert, *et al.*, 2005; Barone *et al.*, 2006). Because treating the long-range exchange is particularly difficult for metals, where a dense sampling of the Brillouin zone is needed, HSE facilitates the study of metals in particular (Paier *et al.*, 2006). For example, a recent study of total energy differences between the semiconducting cubic diamond and the metallic β-tin phases of silicon showed that HSE outperforms any other functional and agrees quite well with diffusion Monte Carlo calculations (Batista *et al.*, 2006). Note, however, that the HSE results are not on par with standard GGA calculations for many other metals (Paier *et al.*, 2006).

In the opposite limit of Eq. (78), $b=1$, the correct asymptotic exchange behavior is recovered, and this is discussed further in Sec. III.D.3, in the context of exact exchange with compatible correlation. But generally, optimizing both a and b, as well as the nature of the short-range–long-range separation, may result in a compromise functional that allows for a different, and appropriate, effective fraction of exact exchange to emerge in different situations. Recent work (Vydrov and Scuseria, 2006; Livshits and Baer, 2007) suggests that cleverly constructed compromise hybrid functionals based on short-range–long-range partitions have made significant progress toward becoming practical tools that alleviate some of the remaining difficulties with hybrid functionals.

Finally, we note the recent work of Grimme (2006), who proposed hybridization of a perturbation theory approach, presented in Sec. III.D.2, as a practical means for obtaining accurate predictions for a wide range of properties and molecules. This demonstrates what appears to be a general theme: as new functionals emerge, their quantitative performance may be aided by deriving and optimizing a hybrid functional associated with them.

D. Exact exchange with compatible correlation

The above heading may seem puzzling and begs an obvious question: In what sense should correlation be "compatible" with exchange? From a wave-function point of view, the two are usually considered to be independent of each other and thus compatibility should not be an issue. The answer is that compatibility of exchange and correlation refers to the exchange-correlation hole properties discussed in Sec. III.C. This compatibility is most easily understood by an example of noncompatibility. Consider combining the exact exchange energy of Eq. (37) with the LSDA functional for correlation. Naively, one might expect that with a major part of the functional—the exchange part—now taken into account exactly, this functional should be better than the LSDA, in which both exchange and correlation are approximated. However, this is not the case. The crucial observation here is that, as discussed above, the exchange hole in a molecule or solid typically has a long-range component that, in equilibrium situations, is canceled to a large extent by a long-range component of the correlation hole, i.e., the combined exchange-and-correlation hole is rather local. But in the combination of exact exchange and LSDA correlation, any long-range component of the exchange-hole can never be canceled because the LSDA correlation hole is intrinsically local. In this sense, the LSDA correlation is not compatible with exact exchange. The same argument holds if the LSDA correlation is replaced by more advanced semilocal correlation functionals. For example, a combination of exact exchange with the PKZB-MGGA correlation functional (Perdew *et al.*, 1999) satisfies all known exact constraints on $E_{\mathrm{xc}}[n]$, yet the atomization energies of typical small organic molecules are considerably less accurate than the ones obtained from a typical GGA (Perdew and Schmidt, 2001).

Yet another functional that is incompatible with exact exchange is the one derived from the work of Colle and Salvetti (1975, 1979, 1990), which also underlies the LYP functional (Lee *et al.*, 1988). Colle and Salvetti devised a model wave function consisting of a Slater determinant correlated with a Jastrow-type factor building in the interelectronic cusp condition. Making a series of approximations, they derived an energy functional from this wave function and fitted a parameter related to the radius of the correlation hole against the correlation energy of the helium atom. The so-obtained Colle-Salvetti functional describes short-range correlation effects well (which is understandable to some extent because the cusp has been taken into account), but it misses important long-range correlation effects (Tao *et al.*, 2001; also see Perdew and Ernzerhof, 1998; Singh *et al.*, 1999; and Caratzoulas and Knowles, 2000 for comments on the shortcomings of the Colle-Salvetti approach).

Quite generally, a combination of exact exchange and (semi)local correlation can at best describe systems of a spatial extension so limited that *a priori* no long-range hole components can exist, such as atoms. But generally speaking it cannot describe more extended systems such as molecules or solids. Having thus explained what we understand to be a compatible correlation functional, we discuss which functionals meet, or at least in principle may meet, this requirement. As we do so, it should be realized that we are approaching the forefront of present-day density-functional research, so applications tend to be relatively few and the final word about the new functionals is far from having been spoken.

1. The hyper-generalized-gradient approximation

There are still not many functionals of the class that has been dubbed the hypergeneralized gradient approximation (HGGA) (Perdew and Schmidt, 2001) or local hybrid (Jaramillo *et al.*, 2003), and, therefore, one might debate whether a separate class is even justified, and if so, what is the most appropriate name for it. Semantics aside, there is a more important question: What physical principle do the functionals in this category have in common? In our definition, the term HGGA comprises all functionals that use a *density-dependent* combination of nonlocal Fock exchange and (semi)local exchange to model the effects of static correlation. This approach is generally reasonable because the long-range static correlation appears to be exchangelike in terms of its hole. While the fundamental idea of the HGGA is simple—combine exact exchange with an (existing) semi-local functional that models dynamical correlation and an exchangelike functional (to be developed) to model static correlation—actually deriving the functional that models the static correlation is far from easy. One major challenge is to devise the functional such that, at equilibrium geometries, the long-range parts of exchange and correlation mostly cancel each other, i.e., the functional must reduce to something that is GGA-like, whereas in stretched geometries as encountered, for example, in dissociation or transition states, true nonlocality must be achieved. If such a functional can be conceived, it will solve many of the previously discussed qualitative problems that plague (semi)local functionals—so that perhaps a panacea may come into sight on this particular horizon.

Promising attempts to derive an accurate HGGA have been made. Perdew and Schmidt (2001) sketched a concept for a HGGA that is based on a combination of MGGA ingredients and the exact exchange energy per particle. The approach can satisfy many of the constraints required for E_{xc}, but the details remain to be worked out. Jaramillo *et al.* (2003) derived a functional they called a local hybrid, because formally it resembles a hybrid functional in which the constant mixing coefficient has been replaced by a functional of the density,

$$E_{\mathrm{xc}}^{\mathrm{lochyb}} = \int n(\vec{r})\{f([n];\vec{r})e_x^{\mathrm{sloc}}([n];\vec{r}) + [1 - f([n];\vec{r})]e_x^{\mathrm{exact}}([n];\vec{r}) + e_c^{\mathrm{sloc}}([n];\vec{r})\}d^3r. \quad (79)$$

Here e_x^{sloc} is the energy per particle of a semilocal exchange functional, e_x^{exact} that of the Fock exchange functional, and e_c^{sloc} that of a semilocal correlation functional. When written as in Eq. (79), the similarity to hybrid functionals is stressed [cf. Eq. (75)]. But by rearranging terms, the exchange-correlation energy per particle of the local hybrid can also be written as

$$e_{\mathrm{xc}}^{\mathrm{lochyb}} = e_x^{\mathrm{exact}} + f(e_x^{\mathrm{sloc}} - e_x^{\mathrm{exact}}) + e_c^{\mathrm{sloc}}, \quad (80)$$

where we have dropped the functional arguments for clarity of notation. In Eq. (80), the fundamental logic of the HGGA becomes visible: The first term on the right-hand side introduces the full-exact exchange energy into the functional, the second term models the exchangelike static correlation as a combination of nonlocal and local exchange, and the third term takes into account dynamical correlation. With an appropriate (and possibly spin-dependent) choice of f, this functional form can satisfy many or all desired constraints.

In the original local hybrid functional, $f = 1 - \tau^W/\tau$ was chosen (see Sec. III.A for the definition of τ^W/τ). Thus, the local hybrid reduces to exact exchange plus semilocal correlation in iso-orbital regions, and this results in qualitatively correct dissociation curves for H_2^+ and similar systems where semilocal functionals fail badly. However, atomization energies obtained from the local hybrid underestimate considerably the experimental results and are less accurate than those from standard GGAs (Jaramillo *et al.*, 2003). This can be traced back to the fact that, on average, f is too small, i.e., in terms of Eq. (79), the local hybrid introduces too much exact exchange, or, more importantly, in terms of Eq. (80), the local hybrid does not introduce enough static correlation. This is a limitation of the exchange-correlation functional itself and therefore independent of whether an orbital-specific (Jaramillo *et al.*, 2003) or a multiplicative local (Arbuznikov *et al.*, 2006) potential is used. Recent work of Bahmann *et al.* (2007) supports the conclusion that the local hybrid with $f = 1 - \tau^W/\tau$ misses important effects of static correlation. They have shown that, if $0.48\tau^W/\tau$ is used instead of τ^W/τ [albeit with a weighted sum of the exchange-correlation energies, rather than the per-particle exchange-correlation energies as in Eq. (79)], with the LDA as the semilocal functional, the performance of the resulting functional in terms of thermochemistry is competitive with, e.g., B3LYP. However, in this way full exact exchange is not incorporated, so that this approach likely suffers from the limitations of standard hybrid functionals that motivate HGGAs in the first place (Bahmann *et al.*, 2007). Therefore, Eq. (79) lays out an interesting concept, but better functionals f have to be developed to make the concept successful in practice. Arbuznikov and Kaupp (2007) recently suggested that the weighting function for the exchange-correlation energies be chosen as a simple analytical function of the reduced density gradient $s_\sigma(\vec{r}) \equiv |\vec{\nabla}\rho_\sigma(\vec{r})|/2(3\pi^2)^{1/3}\rho_\sigma^{4/3}(\vec{r})$, such that f tends to 0 and 1 in the homogeneous and vanishing density limits, respectively. They have shown that the performance of the resulting functional in terms of thermochemistry is also competitive with B3LYP, with the added benefit of attainment of the exact-exchange limit in asymptotic regions where the density approaches zero. However, the resulting functional is still not truly self-interaction-free everywhere, and its performance for other properties, such as reaction barriers, has yet to be evaluated.

A different route to obtain a HGGA has been followed by Mori-Sánchez *et al.* (2006), similar to ideas that were originally put forward in the context of constructing hybrid functionals—which we now see as early attempts to construct HGGAs (Ernzerhof, 1996; Burke, Ernzerhof, and Perdew, 1997). They based the form of their HGGA on the following approximation to the coupling-constant integrand [cf. Eq. (69)]:

$$E_{\mathrm{xc},\lambda}[n] = f_1[n] + \frac{\lambda f_2[n]}{1 + \lambda f_3[n]}, \quad (81)$$

where $f_1[n]$, $f_2[n]$, and $f_3[n]$ are functionals of the density that must be chosen appropriately. By identifying $f_1[n]$ with the exact-exchange energy, this ansatz correctly takes into account that, in the high-density limit (corresponding to $\lambda \to 0$), the exchange-correlation energy must be dominated by exact exchange and $E_{\mathrm{xc},0} = E_x$, as discussed in Sec. III.C. The slope $dE_{\mathrm{xc},\lambda}/d\lambda$ at $\lambda = 0$ is twice the second-order correlation energy of Görling-Levy perturbation theory, and this condition can be used to fix $f_2[n]$. (See Sec. III.D.2 for the precise definition of the high-density limit and a discussion of perturbation theory.) Information from the homogeneous electron gas can be used to derive a condition for $f_3[n]$, and thus one can define a nonempirical functional based on Eq. (81). Mori-Sánchez *et al.* (2006) made further approximations, e.g., by replacing the true slope from the second-order correlation energy by the approximate slope obtained from semilocal functionals. Furthermore, because their unfitted functional turned out to be less accurate than standard functionals, they introduced some empirical parameters. After fitting these to molecular data, they were able to accurately predict atomic and molecular energies, atomization energies, and reaction barriers, thus correctly describing properties whose accurate prediction within DFT is typically mutually exclusive. Unfortunately, this suggested functional is not size consistent and its usage can therefore be problematic (cf. the discussion of size consistency in Sec. III.B).

Finally, another noteworthy HGGA was put forward by Becke (2003, 2005) under the name real-space correlation model. His functional derives from a combination of (a) modeling the above-discussed exchange-correlation hole physics with stretched H_2 as a guiding paradigm problem, and (b) theoretically guided empiricism. A key concept in his approach is to measure the extent to which the exchange hole is delocalized by reverse engineering of the Becke-Roussel (1989) exchange-hole model. This concept is used to define an exchange-correlation hole as a combination of an exchange hole and an exchange-modeled nondynamical correlation hole that is maximally deepened under the constraint of preserving elementary hole properties, like hole normalization. Taking spin dependence into account in the hole models proved to be essential (Becke, 2005). With a few empirical fit parameters, this functional yields atomization energies for the G3 set with an accuracy that is comparable to that of the best-performing hybrid functionals, and without refitting the parameters, the functional also yields very accurate reaction barriers (Dickson and Becke, 2005). So, again, quantities that so far have been hard to bring under the umbrella of a single density functional seem to be covered by a HGGA.

There are still open questions with these new functionals. For example, in the work of Becke (2003, 2005), the calculations were not done self-consistently but were based on B3LYP geometries and LSDA orbitals. Another topic needing clarification is how far the fitting to molecular data limits the usefulness of the functionals for solid-state theory. It would also be interesting to see how the HGGAs perform for response properties such as the ones discussed in Sec. II.C. But, independent of the outcome of such tests, we believe that it is already fair to say that the HGGA concept may well represent a major step forward in density-functional development.

2. Perturbation theory

While the HGGA approach is an attempt to develop models that capture the exchange-correlation effects, the orbital-dependent functionals that are based on perturbation theory follow a different philosophy. To a large extent, they are inspired by traditional many-body methods and follow the governing principle that, starting from a noninteracting zeroth-order system, the exchange-correlation energy is expanded in a Rayleigh-Schrödinger-like perturbation series, with the electron-electron interaction acting as the perturbation. In this sense, the quest for the ultimate functional takes a very different direction: The true hurdle is no longer to *develop* a functional (this follows more or less automatically from the perturbation theory expressions), but to *evaluate* it.

All perturbative schemes discussed below make use of not only the occupied Kohn-Sham orbitals but also the unoccupied ones. In terms of the Jacob's ladder picture, this is a still higher rung than the previously discussed HGGAs, and it has to be paid for by a substantial increase in computational expense. The advantages and drawbacks of perturbative approaches, if regarded as tools for application-oriented computations, can be summarized quite generally: Their main advantage is that they can achieve systematic convergence to the exact result. Their main disadvantage is that they do so by sacrificing an essential part of what made DFT attractive in the first place, namely, modest computational cost gained through insightful modeling of physical quantities such as the exchange-correlation hole or physical principles such as the adiabatic connection curve.

However, there is also another aspect to perturbation theory: In addition to (or instead of) using it for actual applications, it can serve as a tool to gain insight into properties of the exchange-correlation functional and to access special limiting cases such as the high-density limit. This point of view is stressed in the first DFT-based perturbation theory approach, the Görling-Levy perturbation theory (Görling and Levy, 1993, 1994, 1995a), which therefore is reviewed in greater detail in the following. Görling-Levy theory exploits the close connection between uniform coordinate scaling and the adiabatic connection theorem. Consider an N-particle wave function in which the spatial coordinates are uniformly scaled by a factor $\gamma>0$ while normalization is conserved, i.e.,

$$\Psi_\gamma(\vec{r}_1, \ldots, \vec{r}_N) = \gamma^{3N/2}\Psi(\gamma\vec{r}_1, \ldots, \gamma\vec{r}_N). \tag{82}$$

The density corresponding to the scaled wave function thus defined is the scaled density

$$n_\gamma(\vec{r}) = \gamma^3 n(\gamma\vec{r}). \tag{83}$$

Scaling with $\gamma>1$ yields densities that are contracted and on average higher than the unscaled density, whereas scaling with $\gamma<1$ yields densities that are spread out and on average lower than the unscaled density. The limit $\gamma\to\infty$ defines the high-density limit. By straightforward evaluation of the functionals, one can show that under this scaling the components of the Kohn-Sham energy obey the following scaling relations (for detailed discussions of coordinate scaling see, e.g., Levy and Perdew, 1985; Levy, 1991; Perdew and Kurth, 2003):

$$T_s[n_\gamma] = \gamma^2 T_s[n], \tag{84}$$

$$E_H[n_\gamma] = \gamma E_H[n], \tag{85}$$

$$E_x[n_\gamma] = \gamma E_x[n], \tag{86}$$

$$E_c[n_\gamma] = \gamma^2 E_{c,1/\gamma}[n]. \tag{87}$$

Equation (86) can be used to separate a functional into its exchange and correlation parts via coordinate scaling: $E_x[n] = \lim_{\gamma\to\infty} E_{\mathrm{xc}}[n_\gamma]/\gamma$ and $E_c[n] = E_{\mathrm{xc}}[n] - \lim_{\gamma\to\infty} E_{\mathrm{xc}}[n_\gamma]/\gamma$. Equation (87) reveals a central aspect of Görling-Levy theory. It relates the correlation energy functional evaluated with the scaled density to the coupling-constant-dependent correlation energy functional [cf. Eq. (69)] evaluated with the unscaled density, but with a coupling constant given by the inverse of the coordinate scaling factor. Thus, the high-density limit corresponds to the weak-coupling regime of the adiabatic connection.

Görling-Levy perturbation theory proceeds by employing the coupling-constant dependent Schrödinger equation [cf. Eq. (66)],

$$[\hat{T} + \lambda \hat{V}_{ee} + \hat{V}_{\mathrm{ext},\lambda}]\Psi_\lambda = E_\lambda \Psi_\lambda, \tag{88}$$

to define a perturbation expansion in the coupling constant λ. Here Ψ_λ is the ground-state wave function of the scaled Hamiltonian $\hat{H}_\lambda$, $\hat{V}_{\mathrm{ext},\lambda} = \Sigma_{i=1}^{N} v_{\mathrm{ext},\lambda}([n];\vec{r}_i)$, and

$$v_{\mathrm{ext},\lambda}([n];\vec{r}) = v_{\mathrm{KS}}([n];\vec{r}) - \lambda v_H([n];\vec{r}) - \lambda v_x([n];\vec{r}) - v_{c,\lambda}([n];\vec{r}). \tag{89}$$

The coupling-constant-dependent potential of Eq. (89) reflects the scaling relations of Eqs. (85)–(87): Whereas the classical Coulomb and the exchange potentials depend on the coupling constant linearly, the correlation potential depends on the coupling constant in an *a priori* unknown, more complicated way. For $\lambda=0$, the potential $v_{\mathrm{ext},\lambda}$ is just the Kohn-Sham potential. For $\lambda=1$, it is the true external potential of the interacting system that has $n(\vec{r})$ as its ground-state density [as discussed previously below Eq. (66)]. For a given value of the coupling constant λ, the coupling-constant-dependent correlation energy defined by

$$E_{c(\lambda)}[n] = \langle \Psi_\lambda | \hat{T} + \lambda \hat{V}_{ee} | \Psi_\lambda \rangle - \langle \Psi_0 | \hat{T} + \lambda \hat{V}_{ee} | \Psi_0 \rangle, \tag{90}$$

the wave function $\Psi_\lambda[n]$, the total energy $E_\lambda[n]$, and the correlation potential $v_{c,\lambda}([n];\vec{r}) = \delta E_{c(\lambda)}[n]/\delta n(\vec{r})$ are then expanded in powers of λ around $\lambda=0$. This amounts to a perturbation expansion that starts from the noninteracting Kohn-Sham system and takes the electron-electron interaction into account in successive orders of perturbation theory, leading to the fully interacting many-electron system. It should be noted that this perturbation expansion is similar but not identical to standard Rayleigh-Schrödinger perturbation theory, because the perturbing potential itself depends, via the correlation potential $v_{c,\lambda}([n];\vec{r})$, on the coupling constant in a nonlinear way. From the expansion one obtains expressions that are formally similar to Møller-Plesset perturbation theory. The zeroth-order term is the energy of the noninteracting Kohn-Sham system obtained from the sum of the occupied eigenvalues, the first-order term corresponds to the exact Kohn-Sham exchange energy, and the second-order term can be written as

$$E_c^{(\mathrm{GL2})} = \sum_{k=1}^{\infty} \frac{\left| \langle \Psi_0 | \hat{V}_{ee} - \sum_{i=1}^{N} [v_H(\vec{r}_i) + v_x(\vec{r}_i)] | \Psi_{0,k} \rangle \right|^2}{E_0 - E_{0,k}}, \tag{91}$$

where Ψ_0 is the ground-state wave function of the $\lambda=0$ noninteracting system, i.e., the Kohn-Sham ground-state Slater determinant (see Sec. III.C.1), and E_0 is the corresponding noninteracting ground-state energy. $\Psi_{0,k}$ are the excited Kohn-Sham Slater determinants, i.e., the energetically higher-lying eigenfunctions of the Kohn-Sham Hamiltonian having energies $E_{0,k}$. Combining Eqs. (70), (71), and (90), and invoking the Hellmann-Feynman theorem, one finds that the correlation part of the coupling-constant integrand $E_{c,\lambda}[n]$, and $E_{c(\lambda)}[n]$ of Eq. (90) are related to each other by

$$E_{c,\lambda}[n] = \frac{d}{d\lambda} E_{c(\lambda)}[n]. \tag{92}$$

Because E_x is independent of λ and because Eq. (91) defines the quadratic coefficient, i.e., to lowest order in λ, $E_{c(\lambda)}[n] = \lambda^2 E_c^{(\mathrm{GL2})}$, one further finds

$$\left.\frac{dE_{\mathrm{xc},\lambda}}{d\lambda}\right|_{\lambda=0} = \left.\frac{dE_{c,\lambda}}{d\lambda}\right|_{\lambda=0} = \left.\frac{d^2 E_{c(\lambda)}}{d\lambda^2}\right|_{\lambda=0} = 2E_c^{(\mathrm{GL2})}, \tag{93}$$

that is, the starting slope of the adiabatic connection curve is given by twice the second-order Görling-Levy energy. This relation can be used as input in functional construction (cf. Secs. III.D.1 and III.D.3), and to check functional approximations.

From what has been said so far, it is clear that, by going to successively higher orders, Görling-Levy perturbation theory can also be interpreted as defining a formally exact Kohn-Sham scheme (Görling and Levy, 1994). As with any perturbation expansion, it is then an important question whether the expansion converges and is well behaved. The LDA expression diverges in the limit $\gamma\to\infty$ (Görling and Levy, 1993), but the question of convergence in general has sparked quite some interest in recent years. Facco Bonetti *et al.* (2001, 2003) argued that the second-order correlation potential diverges for finite systems for $r\to\infty$ even if the system has a sizable gap, and this finding has been supported by Ivanov and Levy (2002). Niquet *et al.* (2003a, 2003b) disagreed with this conclusion and provided counterarguments based on the Sham-Schlüter equation. This apparent contradiction has been solved in recent work (Engel *et al.*, 2005), which pointed out that a nondiverging solution of the second-order correlation potential equation exists for finite systems with a countable spectrum (such as an atom in a cavity), but not if continuum states are present. The latter were also identified as the source of divergence in earlier work by Ivanov and Levy (2002). Suggestions have been made to overcome these difficulties (Niquet *et al.*, 2005; Engel and Jiang, 2006; Jiang and Engel, 2006), but, in any case, the evaluation of perturbation theory potentials involving unoccupied orbitals requires great care. It should also be noted that problems with divergences are not only found in the Görling-Levy expansion, but have also been reported in another DFT-based perturbative approach (Mori-Sánchez *et al.*, 2005).

Other perturbative approaches that are not necessarily as closely tied to the adiabatic connection idea as Görling-Levy theory, but also make use of Kohn-Sham orbitals, have been suggested. Typically, they stress the use of perturbation theory as a computational tool to achieve high-quality results for small systems, in the spirit of the introductory remarks at the beginning of this section. For example, Grabowski *et al.* (2002) calculated the correlation potentials for the helium isoelectronic atom series and the neon atom and demonstrated that their second-order potential is much closer to the exact correlation potential than the potentials from conventional density functionals. Grimme (2006) showed that an empirical combination of traditional second-order perturbation theory and a hybridlike functional, evaluated non-self-consistently with the Kohn-Sham orbitals, can yield accurate results for a wide range of molecules. Other studies (Bouř, 2000; Gutlé *et al.*, 2002; Hupp *et al.*, 2003; Bartlett *et al.*, 2005; Schweigert *et al.*, 2006) focused on the rapidity with which convergence of quantum-chemistry-like approaches can be achieved by a clever choice of the zeroth-order system (orbitals) or reference Hamiltonian, respectively. The relation between different DFT-based perturbation theories has been discussed by Ivanov *et al.* (2003).

With new functionals still being developed and tested (see examples above), and some theoretically proposed schemes not yet having been tested at all (Grabo *et al.*, 2000), it is at present hard to predict which direction this subfield of orbital-dependent functional theory will take. The possible benefits are large. In addition to high accuracy and systematic convergence to the correct result, the perturbative approaches offer the chance to reliably describe situations that are notoriously difficult for traditional density functionals, such as van der Waals interactions (Engel and Dreizler, 1999; Ivanov and Bartlett, 2001). This may be part of the reason that, quite generally, one sees a tendency in the field to seek an increasing overlap between DFT and quantum chemistry. This direction has also been followed, for example, in the work of Leininger *et al.* (1997) and Goll *et al.* (2006). They separated the Coulomb interaction into short- and long-ranged parts, with the hope of combining the strength of semilocal density functionals—inexpensive description of short-range dynamical correlation—with the strength of configuration-interaction or coupled-cluster theory—accurate description of long-range static correlation effects (see Sec. III.D.3). Thus, the boundaries between DFT and many-body perturbation theory start to blur. But the possible benefits of these approaches come at a heavy price, because the boundaries blur not only in terms of methodology and quality of results, but also in terms of computational expense. Much of the original attractiveness of DFT, which stemmed from the chance to gain insight into electronic structure from first principles without having to suffer through the pains of solving the many-body Schrödinger equation, is lost. So DFT-based perturbation theory appears to be a path toward merging the methods of DFT and traditional quantum chemistry both for better and for worse.

3. Range-separated functionals and other approaches

Not all orbital-dependent functionals can be categorized into one of the major classes presented above, and this section is devoted to discussing some of the functionals that have been omitted so far. It is beyond the limits of this review to go into detail and to provide coverage of all orbital-dependent functionals developed over the years. For example, we do not discuss the LDA + U (Anisimov *et al.*, 1997) and the LDA+ dynamical mean-field theory (Held *et al.*, 2003; Kotliar *et al.*, 2006), since these methods have been the topic of the just-cited recent review articles and are at or beyond the verge of what typically is considered DFT. Quite generally, we make no effort to cross the boundary to other many-body methods in this paper. However, one should note that these boundaries have become less rigid over the years, and fruitful stimulation can come from visiting the other side of the fence (see, e.g., the review by Onida *et al.*, 2002). For the sake of focus, here we shed light on only a few approaches that are of true density-functional character and either provide insight into general concepts of DFT or may serve as starting points for future density-functional development, or both.

A common theme in many approaches that go beyond the semilocal approximations by using orbitals is the concept of separating the electron-electron interaction into long- and short-range parts. This idea underlies, e.g., several approaches based on the random-phase approximation (RPA), or, more precisely, its density-functional version. The RPA was originally put forward (Nozières and Pines, 1958) and used (von Barth and Hedin, 1972; Vosko *et al.*, 1980) in studies of the homogeneous electron gas. The RPA density functional can be defined by starting from an exact representation of the exchange-correlation energy in the form of the adiabatic-connection fluctuation-dissipation theorem (Langreth and Perdew, 1975, 1977),

$$E_{\mathrm{xc}} = \frac{1}{2}\int d^3r \int d^3r' \frac{e^2}{|\vec{r}-\vec{r}\,'|}\int_0^1 d\lambda \times\left(-\frac{1}{\pi}\int_0^\infty d\omega \,\mathrm{Im}\, \chi_\lambda(\vec{r},\vec{r}\,',\omega) - n(\vec{r})\delta(\vec{r}-\vec{r}\,')\right), \qquad (94)$$

where $\chi_\lambda(\vec{r},\vec{r}\,',\omega)$ is the coupling-constant-dependent dynamic response function of the system, which can be expressed within time-dependent density-functional theory (Runge and Gross, 1984; Gross *et al.*, 1999),

$$\chi_\lambda(\vec{r},\vec{r}\,',\omega) = \chi_{\mathrm{KS}}(\vec{r},\vec{r}\,',\omega) + \int d^3x \int d^3y \chi_{\mathrm{KS}}(\vec{r},\vec{x},\omega) \times \left(\frac{\lambda e^2}{|\vec{x}-\vec{y}|} + f_{\mathrm{xc},(\lambda)}(\vec{x},\vec{y},\omega) \right) \chi_\lambda(\vec{y},\vec{r}\,',\omega), \tag{95}$$

and $f_{\mathrm{xc},(\lambda)}(\vec{x},\vec{y},\omega)$ is the coupling-constant-dependent exchange-correlation kernel (see Sec. IV.B). For $\lambda = 0$, χ_λ is identical to the Kohn-Sham (i.e., noninteracting) dynamic response function,

$$\chi_{\mathrm{KS}}(\vec{r},\vec{r}\,',\omega) = \lim_{\eta \to 0} \sum_{\sigma=\uparrow,\downarrow} \sum_{j,k} (f_{k,\sigma} - f_{j,\sigma}) \times \frac{\varphi_{j,\sigma}(\vec{r})\varphi^*_{k,\sigma}(\vec{r})\varphi^*_{j,\sigma}(\vec{r}\,')\varphi_{k,\sigma}(\vec{r}\,')}{\omega - (\varepsilon_{j,\sigma} - \varepsilon_{k,\sigma}) + i\eta}, \tag{96}$$

where $f_{k,\sigma}$ and $f_{j,\sigma}$ are the orbital occupation numbers (0 or 1), and j and k sum over the full spectrum of occupied and unoccupied orbitals. Different approximations to f_{xc} lead to different approximations for E_{xc} via Eqs. (94) and (95). A detailed analysis of the relation between f_{xc} and other many-body approaches has been given in the review by Onida *et al.* (2002); thus we do not discuss it here. Instead, we focus on the RPA functional, which follows from the crudest possible approximation $f_{\mathrm{xc},(\lambda)} = 0$. [Note, however, that based on Eq. (94), studies beyond the RPA have been reported. See, e.g., Dobson and Dinte, 1996; Fuchs and Gonze, 2002; and Pitarke and Perdew, 2003.]

The RPA functional has been used successfully to study jellium slabs and surfaces (Pitarke and Eguiluz, 1998, 2001; Dobson and Wang, 1999); it has been applied to atoms and molecules (Furche, 2001; Fuchs and Gonze, 2002; Fuchs *et al.*, 2003) and solids (Kotani, 1998; Niquet and Gonze, 2004). The RPA functional can describe van der Waals interactions (Dobson and Wang, 1999; Fuchs and Gonze, 2002) and generally captures long-range correlation well. However, it misses important short-range correlation effects. Therefore, it can, to some extent, be regarded as the complement of semilocal correlation functionals, which mainly capture short-range correlation. It is then a natural idea to try to merge the two approaches in order to combine the beneficial aspects of both. To this end, (semi)local functionals have been constructed with the aim of avoiding double counting of correlation effects in the combined functional (Kurth and Perdew, 1999; Yan, Perdew, and Kurth, 2000). But the results that were obtained for molecules with this approach were not encouraging (Furche, 2001), and indicate that matching the two rather different approaches is not easy.

The idea of treating the long- and short-range components of the electron-electron interaction on a different footing has also been followed under a different perspective. Leininger *et al.* (1997) split the Coulomb interaction according to

$$\frac{1}{|\vec{r}_i - \vec{r}_j|} = \frac{\mathrm{erf}(\mu|\vec{r}_i - \vec{r}_j|)}{|\vec{r}_i - \vec{r}_j|} + \frac{1 - \mathrm{erf}(\mu|\vec{r}_i - \vec{r}_j|)}{|\vec{r}_i - \vec{r}_j|}, \tag{97}$$

where $\mathrm{erf}(\cdot)$ denotes the standard error function and μ is an empirical parameter. Their central idea was to take the long-range part of the Coulomb interaction, given by the first term on the right-hand side of Eq. (97), into account through a configuration-interaction (CI) calculation, but to model the short-range part of the electron-electron interaction, represented by the second term, by a density functional (cf. Sec. III.D.2). The underlying logic is that Eq. (97) splits up the interaction in such a way that the part that is covered by the CI calculation does not diverge, i.e., it does not lead to a cusp in the wave function. Therefore, the CI calculation converges faster than for the unmodified interaction, and the short-range correlation effects, which are thus missing in the CI calculation, are brought in via the density functional. From an application point of view, the attractiveness of mixing CI with DFT is, once more, that the long-range correlation effects, which are hard to model with explicit density functionals, are rigorously taken into account. The fundamental idea of splitting up the interaction according to Eq. (97) has been developed further in subsequent work with the definition of density functionals for both components and the derivation of scaling relations and approximate density functionals, in particular for the short-range part (Toulouse *et al.*, 2004, 2005, 2006; Gori-Giorgi and Savin, 2006; Toulouse and Savin, 2006).

The splitting of Eq. (97) has also proven to be a valuable concept without quantum chemistry modeling of the long-range part. This is because it can be used to construct functionals that take the long-range part of only the exchange into account nonlocally while modeling other interaction effects with semilocal functionals. The long-range exchange is then given by a Fock integral with $\mathrm{erf}(\mu|\vec{r}_i - \vec{r}_j|)/|\vec{r}_i - \vec{r}_j|$ as the interaction. [See Iikura *et al.*, 2001; Yanai *et al.*, 2004. Also compare Eq. (102) and references given there for a different but related approach.] Such functionals have been interpreted as generalized hybrid functionals (Ángyán *et al.*, 2005; Gerber and Ángyán, 2005). The inclusion of the long-range exchange has been reported to improve the description of s-d transfer energies (Iikura *et al.*, 2001), the longitudinal polarizabilities of molecular chains (Iikura *et al.*, 2001; Sekino *et al.*, 2005), and charge-transfer excitations (Tawada *et al.*, 2004; Yanai *et al.*, 2004). Combined with the van der Waals functional by Andersson *et al.* (1996), the approach has also been reported to provide a reasonable description of van der Waals interactions (Kamiya *et al.*, 2002; Sato *et al.*, 2005a, 2005b). But, in addition to these encouraging results, it must be noted that, based on a systematic study of molecular properties, Peach, Cohen, and Tozer (2006) have cautioned that parameters that have been introduced into the functionals have a great influence on the functional performance, and parameters that lead to good results for one observable can deteriorate accuracy for others. One may speculate that a way to overcome these problems could be to turn the parameters into density functionals themselves, and then, apparently, the boundaries between the hybrid functional, HGGA, and the just-defined range-separation functional classes begin to blur. A step in this direction has been taken, for example, by Baer and Neuhauser (2005) with a functional that is motivated by a coupling-constant argument and based on the combination of nonlocal, Fock-like exchange and a local exchange-correlation functional for a range-separated interaction of the form $[1-\exp(-\mu r)]/r$.

Finally, we mention two functionals that do not fall into the range-separation category. The correlation factor model (Burke, Perdew, and Ernzerhof, 1997; Perdew and Schmidt, 2001) is an attempt to directly model the coupling-constant-averaged exchange-correlation hole. The basic idea is to multiply the exact-exchange hole by a function chosen appropriately to satisfy the hole sum rule (cf. Sec. I.B) and other constraints. Such a functional can, in principle, achieve everything a HGGA does. Whether it is possible to work out the necessary details and whether it will be computationally feasible remains to be seen. Another approach that is charming as a theoretical concept is the interaction-strength interpolation (Seidl *et al.*, 2000). Its idea is to model the coupling-constant integrand of Eq. (69) as an interpolation between the $\lambda=0$, i.e., the exact-exchange limit, and the limit of infinite coupling strength $\lambda\to\infty$, which can be well approximated by explicit density functionals (Seidl *et al.*, 1999). However, the interaction-strength interpolation needs the starting slope (i.e., the slope at $\lambda=0$) of the coupling-constant integrand curve as an input to its interpolation. As seen in Eq. (93), this means its use requires knowledge of the second-order Görling-Levy correlation energy. Therefore, the scheme is computationally demanding and has not come into widespread use.

In conclusion, it should be stressed once more that, quite generally, most of the functionals discussed in Sec. III.D are computationally much more demanding than the semilocal functionals. Thus, the compatible correlation functionals also pose computational challenges in addition to the fundamental ones that we discussed.

IV. ORBITAL-DEPENDENT FUNCTIONALS AND EXCITED STATES

A. The gap problem

All discussion and analysis presented so far have been limited to ground-state properties. This is because DFT is inherently a ground-state theory. Proof of the Hohenberg-Kohn theorem relies explicitly on the variational principle, which applies only to the ground state. In principle, because the ground-state density uniquely determines the many-body Hamiltonian, information about excited-state properties is also encoded within the ground-state density (Görling, 1996, 2005). However, in the absence of any explicit link between excited-state quantities and ground-state densities, predictions as to the former cannot be made based on the ground-state functional.

One property that is frequently related to questions about excited states is the gap of a molecule or solid. However, here it is important to distinguish between two different gaps that are relevant in this context. The fundamental gap is defined as the difference between the first ionization potential I and the first electron affinity A. In a solid, it is usually referred to as the band gap. It can be measured, for example, by using photoemission and inverse photoemission for determining I and A, respectively. The optical gap is defined as the difference between the energies of the lowest excited many-particle state and the ground state. It can be measured, for example, by determining the optical-absorption spectrum. In many solids, the fundamental and optical gaps are within ~0.1 eV. From a many-body perturbation theory perspective, this can be rationalized as being due to strong screening of the Coulomb attraction between the quasielectron and the quasihole. The quantitative similarity (but not identity) of the two gaps often results in a confusion between them. However, they are by no means the same, and in the absence of strong screening the fundamental and the optical gap can and usually will differ considerably, an issue elaborated below. The gap problem addressed in this section is the relation between the Kohn-Sham gap, defined as the difference between the lowest unoccupied and highest occupied Kohn-Sham eigenvalues, and the fundamental and optical gaps.

We first consider the fundamental gap in solids. As already hinted in Sec. I.B, I and A can be assessed from total energy differences between the $(N-1)$- or the $(N+1)$-electron systems, respectively, and the N-electron system. Thus, meaningful values for the fundamental gap can be computed rigorously within an all-ground-state theoretical framework. However, whereas these energy differences are easily calculated for finite systems, one cannot simply add or remove a single electron from an infinite system such as a solid with periodic boundary conditions. Thus, the temptation to consider Kohn-Sham gaps as an alternative in the latter case is great. This is all the more so given that, for many solids, valence- and conduction-band structure calculations based on filled and empty Kohn-Sham orbitals, respectively, often provide for surprisingly accurate agreement with experiment, to within a rigid shift (see, e.g., Chelikowsky and Cohen, 1992, and references therein). Thus, the Kohn-Sham gap has also frequently been compared to the fundamental gap. However, this comparison leads to difficult questions of its own.

It is routinely found that DFT calculations with semilocal functionals yield a Kohn-Sham gap that underestimates the fundamental one, often considerably. For example, for Si, the LSDA underestimates the fundamental gap by more than 50%. For Ge, the LSDA does not find a gap at all, misidentifying this semiconductor as a semimetal instead (see, e.g., Hybertsen and Louie, 1986). These significant differences between the Kohn-Sham and fundamental gaps are often attributed to a failure of semilocal functionals. As this appears to be a frequent source of confusion in the literature, we emphasize that in fact this difference has its source in the

Kohn-Sham approach itself. By virtue of Janak's theorem (Janak, 1978), and taking into account the derivative discontinuity of E_{xc} introduced in Sec. I.B, we know that, for an N-electron system, I and A are related to the Nth and $(N+1)$th Kohn-Sham eigenvalues via (Perdew and Levy, 1983; Sham and Schlüter, 1983)

$$I(N) = -\varepsilon_N(N-\delta), \tag{98}$$

$$A(N) = -\varepsilon_{N+1}(N+\delta), \tag{99}$$

where δ is infinitesimal. In other words, the ionization potential can be equated with the highest occupied Kohn-Sham eigenvalue for a system approaching the integer particle number from below, whereas the electron affinity can be equated with the highest (infinitesimally) occupied eigenvalue for a system approaching the integer particle number from above. The latter value, however, will differ from that obtained for the lowest unoccupied eigenvalue of a system approaching the integer particle number from below by exactly the derivative discontinuity in E_{xc}. A standard Kohn-Sham calculation, i.e., one with an integer particle number and an asymptotically vanishing exchange-correlation potential, is known to be equivalent to approaching the integer particle number from below (Levy *et al.*, 1984; Perdew and Levy, 1997). We must therefore conclude that for this standard calculation

$$I - A = \varepsilon_{N+1} - \varepsilon_N + \Delta_{\mathrm{xc}}, \tag{100}$$

where Δ_{xc} is the magnitude of the derivative discontinuity in E_{xc}. We stress that this gap problem is not a failure of semilocal functionals, which do not possess a derivative discontinuity. Even for the ultimate functional, the Kohn-Sham gap obtained from the integer-particle-number system may well differ from the fundamental gap. This is because the derivative discontinuity, even if present in the functional, is not incorporated into the Kohn-Sham gap. This analytical conclusion has been repeatedly confirmed in a variety of systems by numerical studies that reconstructed (what is believed to be) the exact Kohn-Sham potential from quantities computed in other ways (Godby *et al.*, 1986; Chan, 1999; Allen and Tozer, 2002; Grüning *et al.*, 2006a, 2006b). It is also strongly supported by a detailed analysis of the difference between the fundamental gap and the Kohn-Sham gap in terms of a GW correction (Niquet and Gonze, 2004).

The above consideration shows that a comparison of Kohn-Sham gaps to fundamental ones is inherently a comparison of apples and oranges. Thus, we believe that a search for an ultimate functional that would accurately predict fundamental gaps in addition to all ground-state properties, within the time-independent Kohn-Sham framework, is inherently doomed. It is nevertheless often hoped that orbital-dependent functionals can yield realistic fundamental gap values, perhaps via systematic cancellation of errors. In the following, we review critically the performance of the various orbital-dependent functionals discussed in the preceding section in predicting fundamental gap values.

We have already explained in Sec. III.A that MGGA functionals, although orbital dependent, are actually closer in spirit to semilocal functionals. Thus, it comes as no surprise that the hope for improved gap values does not reside with them. Indeed, Kohn-Sham gaps determined using, e.g., the TPSS MGGA functional, are found to be very similar to those obtained with other semilocal functionals (see, e.g., Barone, Peralta, Wert, *et al.*, 2005; Litimein *et al.*, 2006; Prodan *et al.*, 2006).

The hope for improved prediction of gap values via systematic error cancellation is perhaps best expressed by the example of hybrid functionals. LSDA and GGA Kohn-Sham gaps are consistently far too small and Hartree-Fock gaps are consistently far too large. Perhaps, then, a weighted sum of LSDA or GGA and Fock potentials, as in hybrid functionals, can produce reasonable gaps? This hope is further assisted by the fact that the gap predicted by hybrid functionals is not the Kohn-Sham gap because of the use of orbital-specific potentials. Therefore, the above derivative discontinuity argument does not directly apply (an issue elaborated below).

Because fundamental gaps are by no means part of the rationale and/or training set of popular hybrid functionals, the hope for quantitative gap prediction may appear unrealistic. Remarkably, however, the B3LYP functional appears to predict band gap energies that are within 10% at most of experimental values for a wide range of solids—from standard semiconductors, e.g., Si and GaAs, to wide-gap insulators, e.g., MgO and Al_2O_3 (Muscat *et al.*, 2001). For silicon, in particular, the detailed band structure predicted by B3LYP is in outstanding agreement with that obtained from much more sophisticated and computationally intensive (GW and quantum Monte Carlo) calculations (Muscat *et al.*, 2001). One should bear in mind, however, that this merry state of affairs is not universal. While the derivative discontinuity can be (and often is) substantial, there is no reason why it must be large. If it happens to be small, Eq. (100) suggests that the Kohn-Sham gap obtained with a semilocal functional can be in excellent, albeit fortuitous, agreement with experiment. For some materials, this is indeed the case. For example, both the LDA and experimental band gaps of pyrite (FeS_2) are ~1 eV (Eyert *et al.*, 1998). Use of the B3LYP functional, however, introduces a sizable and unnecessary Fock correction, resulting in a band gap of ~2 eV (Muscat *et al.*, 2001), i.e., twice the experimental value.

The practical performance of hybrid functionals with respect to the fundamental gap also depends on the type of hybrid functional used. For example, the PBE hybrid tends to overestimate band gaps for semiconductors and underestimate them for insulators (Paier *et al.*, 2006). The HSE screened hybrid does significantly better for small- and medium-gap solids (Heyd *et al.*, 2005; Paier *et al.*, 2006; Peralta *et al.*, 2006; Uddin and Scuseria, 2006). However, it still underestimates, possibly considerably, the gap of large-gap insulators (Paier *et al.*, 2006). This again indicates a subtle cancellation of errors that is different for different functionals.

A different kind of error cancellation is encountered in OEP calculations that employ the exact-exchange expression of Eq. (37), usually with LSDA correlation. In Sec. III.D we argued that this cannot be a useful general strategy for computing ground-state properties. Nevertheless, it has been consistently found that, for a variety of semiconductors, the Kohn-Sham gap obtained from such calculations is within only 0.1–0.3 eV from the experimental gap (Bylander and Kleinman, 1995a, 1995b, 1996, 1997; Städele *et al.*, 1997, 1999; Fleszar, 2001; Fernández *et al.*, 2003b; Grüning *et al.*, 2006a).

A first step toward explaining these results is to understand why the addition of 100% of the Fock exchange energy in the above-cited OEP studies has essentially the same effect on the band gap as the addition of only 20% of exact exchange in the B3LYP calculations of Muscat *et al.* (2001). Recall that in the limit of a hypothetical correlation-free electron system, the Hartree-Fock approximation is exact. The exact-exchange OEP is equally exact for the ground state of this hypothetical system. But due to the derivative discontinuity term in Eq. (100), which is Δ_x in the absence of correlation, the OEP band gap could be much smaller than the Hartree-Fock one. Such an effect was observed by Städele *et al.* (1997), who performed both a Hartree-Fock and an OEP exact-exchange calculation for Si. Additionally, they evaluated Δ_x directly, which is relatively simple for the correlation-free case (Görling and Levy, 1995b). They then found that the exact-exchange Kohn-Sham gap was 1.43 eV, the derivative discontinuity Δ_x was 5.48 eV, and their sum was equal to the Hartree-Fock band gap to within less than 3%. This immediately explains why, unlike in the hybrid functional case, adding 100% of exact exchange does not result in large band gaps. With the Kohn-Sham potential, which is local, most of the Hartree-Fock gap is reflected in the derivative discontinuity and not in the Kohn-Sham gap. The cancellation of errors here, then, is that the influence of the exact-exchange term on the Kohn-Sham gap is similar in magnitude to the unknown Δ_{xc} term that one would like to add to the exact Kohn-Sham gap [which is believed to be close to the semilocal one for standard semiconductors (Godby *et al.*, 1986)].

Alas, this cancellation is essentially fortuitous. Calculations for systems sufficiently different from standard semiconductors, such as large gap semiconductors and insulators (see, e.g., Kotani, 1995; Aulbur *et al.*, 2000; Rinke *et al.*, 2005; Grüning *et al.*, 2006a, 2006b), noble-atom solids (e.g., Magyar *et al.*, 2004; Grüning *et al.*, 2006a, 2006b) or *trans*-polyacetylene (Rohra *et al.*, 2006), reveal that the gap obtained from adding exact exchange to the LSDA within the OEP scheme is substantially larger than the LSDA gap, but falls increasingly short of the true band gap. Interestingly, for materials with very large gaps, such as GaN or ZnO, or for materials where the LSDA gap is negative, such as InN or ScN, exact-exchange calculations, while not yielding sufficiently accurate gap values themselves, were found to provide a much better starting point than the LSDA for a perturbative *GW* correction to the gap (Rinke *et al.*, 2005, 2006, 2007; Qteish *et al.*, 2006).

We note that there is a dissenting voice to this consensus on the computational facts. Sharma *et al.* (2005), using all-electron full-potential linearized augmented-plane-wave calculations [see Ambrosch-Draxl (2004) for more details of the computational approach], found Kohn-Sham gaps that were not as close to experiment as those obtained from the previous calculations, all of which were based on pseudopotentials. They interpreted this result as indicating that an explicit description of core-valence interaction is necessary for accurate determination of the Kohn-Sham band structures—a conclusion that is controversial (see, e.g., Qteish *et al.*, 2005). More work is needed to determine whether this is indeed a serious and insufficiently understood limitation of the pseudopotential approximation, or rather a reflection of limitations in the numerical approach of Sharma *et al.*

Even when use of exact-exchange OEP produces gaps in excellent agreement with experiment, recall that a correlation that is either zero or semilocal is intrinsically incompatible with the exact exchange, in the sense defined in Sec. III. The increased accuracy for gap values, then, comes at a significant cost—sacrificing, to a large extent, the predictive power for other properties such as bond lengths, binding energies, etc.

An approach that allows for prediction of gap values in solids with an accuracy that is at least on par with that of the exact-exchange OEP approach, while still offering a reasonable compromise for other properties, is the screened-exchange approach. This method was first suggested by Bylander and Kleinman (1990) on a heuristic basis. It was later placed on a rigorous footing by Seidl *et al.* (1996), who showed that it can be viewed as a special case of the generalized Kohn-Sham scheme described in Sec. II.B. In a certain sense, it can even be viewed as an approximation of *GW* theory (Yasuhara and Takada, 1991; Seidl *et al.*, 1996; van Schilfgaarde *et al.*, 1997), but here we naturally adopt the generalized Kohn-Sham point of view for introducing the approach.

In the screened exchange (sx) approach, the orbital-dependent functional $S[\{\varphi_i\}]$, defined in Sec. II.B, is chosen as

$$S[\{\varphi_i\}] = \langle\Phi|\hat{T}|\Phi\rangle + E_H[\{\varphi_i\}] + E_x^{\mathrm{sx}}[\{\varphi_i\}], \tag{101}$$

where Φ is the Slater determinant constructed from $\{\varphi_i\}$, and

$$E_x^{\mathrm{sx}}[\{\varphi_i\}] = -\frac{e^2}{2}\sum_{\sigma=\uparrow,\downarrow}\sum_{i,j=1}^{N_\sigma}\int\int d^3r'\,d^3r \times \frac{\varphi_{i,\sigma}^*(\vec{r})\varphi_{j,\sigma}^*(\vec{r}')e^{-k_{\mathrm{TF}}|\vec{r}-\vec{r}'|}\varphi_{j,\sigma}(\vec{r})\varphi_{i,\sigma}(\vec{r}')}{|\vec{r}-\vec{r}'|}. \tag{102}$$

Equation (102) is similar to the exact exchange of Eq. (37), but the bare Coulomb potential has been replaced with a Yukawa potential, i.e., the bare potential has been screened semiclassically, with $k_{\rm TF}$ being the semiclassical Thomas-Fermi screening factor [other static screening expressions are also possible—see Seidl *et al.* (1996) and Engel (1997) for details]. Comparison of Eqs. (5) and (39) shows that for this choice of $S[\{\varphi_i\}]$, the remainder functional $R^s[n]$ is given by

$$R^s[n] = E_x[n] - E_x^{\rm sx}[n] + T[n] - T_s[n] + E_c[n], \qquad (103)$$

where $E_x^{\rm sx}[n]$ is the $F^s[n]$ functional [cf. Eq. (38)] related to the choice of $S[\{\varphi_i\}]$ made in Eq. (101). For practical applications of the screened exchange approach, one usually employs the sx-LDA approximation. In this approximation, one (i) uses the usual LDA form for $E_x[n]$ and $E_c[n]$, (ii) neglects the difference between $T[n]$ and $T_s[n]$, and (iii) uses the LDA form stemming from a Yukawa electron interaction for $E_x^{\rm sx}[n]$, with a Thomas-Fermi wave vector corresponding to the average density. Adopting these approximations, we obtain the sx-LDA form of the generalized Kohn-Sham equation (40),

$$\begin{aligned}&\left(-\frac{\hbar^2\nabla^2}{2m} + v_{\rm ext}(\vec{r}) + e^2\int d^3r'\,\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} + v_{\rm xc}^{\rm sx}([n];\vec{r})\right)\varphi_i(\vec{r})\\ &\quad - e^2\sum_{j=1}^{N_\sigma}\int\frac{\varphi_j(\vec{r})e^{-k_{\rm TF}|\vec{r}-\vec{r}\,'|}\varphi_j^*(\vec{r}\,')\varphi_i(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}d^3r' = \varepsilon_i\varphi_i(\vec{r}),\end{aligned} \qquad (104)$$

where $v_{\rm xc}^{\rm sx}([n];\vec{r})$ is the functional derivative of the screened-exchange remainder functional with respect to the density.

Physically, the sx-LDA approximation is motivated by the observation that the gaps obtained from the Hartree-Fock equation are far too large because they correspond to bare exchange. The semiclassical screening of the exchange in the sx-LDA equation (104) is assumed to serve as a much better starting point for gap evaluation, with the orbital-specific nature of Eq. (104) hopefully transferring much of the fundamental gap from the derivative discontinuity of $R^s[n]$ to the generalized-Kohn-Sham gap. At the same time, the long-range properties of the exchange and correlation hole are still reasonably compatible (see also Engel, 1997), because the screened-exchange term decays at long range, giving hope for a reasonable ground-state description.

For silicon, good values for the fundamental gap (1.32 eV, versus an experimental 1.17 eV) and other principal gaps were obtained by Bylander and Kleinman (1990) using sx-LDA. Seidl *et al.* (1996) obtained similarly good values for principal gaps of Ge, GaAs, InP, and InSb, while still getting good values for the lattice constant and reasonable values for the bulk moduli. The sx-LDA approach has since been employed for determining gaps for a great variety of semiconductors and semiconductor structures (see, e.g., Asahi *et al.*, 2000; Picozzi *et al.*, 2000; Geller *et al.*, 2001; Lee and Wang, 2006; and references therein). sx-LDA has also afforded considerable improvement over the LDA in the calculation of other band-structure-related properties, such as optical properties (see, e.g., Asahi *et al.*, 1999; Rhim *et al.*, 2005) and effective masses (see, e.g., Geller *et al.*, 2001). The sx-LDA-computed fundamental gap, compared to experiment, for 24 elemental and binary semiconductors, is shown in Fig. 9 (Geller *et al.*, 2001). It is observed that the sx-LDA offers, overall, much improvement over the LDA, with substantially less scatter. Of particular note is that small-gap semiconductors are correctly described as such within the sx-LDA, whereas they are typically erroneously described as semimetals within the LDA. the sx-LDA has also been applied successfully to gap predictions in oxides (see, e.g., Asahi *et al.*, 2002; Robertson *et al.*, 2006) and nitrides (Stampfl *et al.*, 2001) and has been reported to offer a significant improvement over semilocal functionals for the energy position of vacancy levels in high-dielectric-constant oxides (see, e.g., Robertson *et al.*, 2005; Xiong *et al.*, 2005).

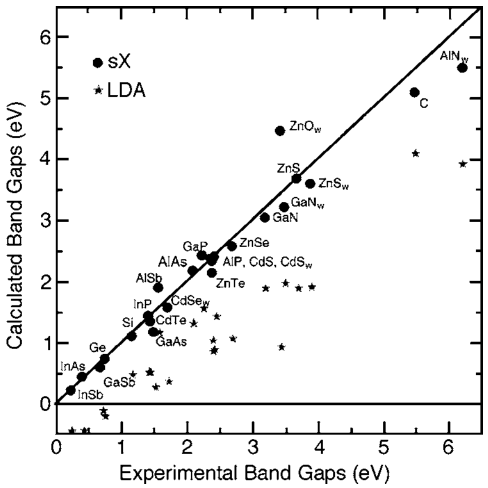

FIG. 9. Calculated vs experimental band gaps for 24 elemental and binary semiconductors. Circles, sx-LDA. Stars, LDA. Reprinted with permission from Geller *et al.*, 2001. Copyright 2001, American Institute of Physics.

Despite the many successes of the sx-LDA, there is, as always, room for caution as well. There are cases in which accuracy is clearly insufficient. Examples are ZnO and AlN (see Fig. 9), where the experimental value is overestimated or underestimated by ~1 eV, respectively. This indicates yet again that, while the shift of weight from the derivative discontinuity to the generalized Kohn-Sham gap is indeed systematic, at least for semiconductors and some oxides, it is not universal. Once again, accuracy in a new chemical environment needs to be ascertained before predictions can be made. It should also be borne in mind that, because the orbital-dependent exchange term is a short-range one, the asymptotic exchange-correlation is still governed by the LDA and thus the sx-LDA is probably far better suited for solids (where it has indeed been mostly used) than for molecules (Toulouse and Savin, 2006). Indeed, the sx-LDA has been used occasionally for atoms (see, e.g., Vukajlovic *et al.*, 1991; Yamagami *et al.*, 1994), but we are not aware of it being used for molecules. Finally, it is important to recall that the sx-LDA represents a compromise between the accuracy of the LDA for the ground state and the accurate prediction of gaps. Thus, there are cases in which the sx-LDA can differ from the LDA substantially. For example, the LDA predicts Jahn-Teller distortions and so-called negative-U effects for charged vacancies in Si, in qualitative agreement with experiment. In contrast, sx-LDA calculations erroneously predict an outward relaxation and an sp^2 hybridization for the ions surrounding the vacancy (Lento and Nieminen, 2003).

The self-interaction-corrected functionals discussed in Sec. III.B represent a different route to improving fundamental gap values. Unlike the previous approaches, SIC functionals are not based on the inclusion of an explicit Fock exchange integral in some form or other. Nevertheless, a consistent observation, already noted early on (see, e.g., Perdew and Zunger, 1981; Heaton *et al.*, 1982, 1983), is that, whenever SIC calculations improve the eigenvalue spectrum in general, they also greatly improve gap values. One example is LiCl, where the LDA gap is 5.8 eV, but the SIC-LDA gap is 10.1 eV, as compared to an experimental gap value of 9.4–9.9 eV (Heaton *et al.*, 1983). This improvement is usually rationalized as the pull-down of valence orbitals by the elimination of the self-interaction and the correct asymptotic potential "felt" by conduction orbitals. While this is undoubtedly true, it is not the whole story. As emphasized throughout, the gap of the self-interaction-free ultimate functional may still differ substantially from the correct gap. Clearly, then, the fact that practical SIC calculations employ an orbital-specific potential again has to do with mitigating the derivative discontinuity issue. Once more, there are two main concerns for practical applications. One is the possible sacrifice of accuracy in predicting ground-state properties, which was discussed in Sec. III.B. The other is the issue of systematics in gap predictions, given that the rationale behind the SIC functional does not involve gap calculations as such. Indeed, a study of various transition-metal oxides (Svane and Gunnarson, 1990) has shown that SIC LDA gaps are consistently a significant improvement over LDA ones, but for, e.g., NiO, the SIC-LDA gap still severely underestimates the experimental one.

We now turn our attention to the optical gap. One proper way of computing excited-state properties in general, and optical gaps in particular, is to use time-dependent DFT (TDDFT). This approach is discussed in the next section. However, TDDFT calculations are more complicated than DFT ones and raise difficult new questions as to the appropriate choice of functional. It is therefore tempting to invoke the concept of Görling's density-functional theory for excited states and rely on the interpretation of Kohn-Sham eigenvalues as zeroth-order excitation energies (Görling, 1996; Filippi *et al.*, 1997). Indeed, Baerends and Gritsenko (2005) pointed out that the unoccupied Kohn-Sham eigenvalues "represent an 'excited' electron interacting with $N-1$ electrons, rather than an extra electron interacting with N electrons." This is because the unoccupied orbitals are obtained from the same local potential as the occupied ones and thus both experience the effective field of $N-1$ electrons. By definition, the fundamental excitation of one electron in the presence of the other $N-1$ electrons corresponds to the optical gap. We also note that the exact Kohn-Sham gap has been conjectured to mimic the optical gap in atoms (Al-Sharif *et al.*, 1998; Savin *et al.*, 1998). In this sense, the Kohn-Sham gap may be seen as linked to the optical gap. However, calculations with the LSDA or GGA routinely find Kohn-Sham gap values that are substantially smaller than excitation gaps found experimentally (see, e.g., Hybertsen and Louie, 1986). Once again, this is not merely a consequence of using a semilocal functional because in, e.g., Si, the Kohn-Sham gap found with the exact functional (Godby *et al.*, 1986; Grüning *et al.*, 2006a) is not only well below the fundamental gap, but also well below the optical one. This shows that the zeroth-order approximation obtained by employing eigenvalue differences is generally not sufficient to predict the optical gap reliably.

Most predictions of optical gap values from ground-state DFT theory in practice are performed with hybrid functionals. Just as with fundamental gaps, optical gaps are also not part of the rationale and/or training set of popular hybrid functionals, and once again there seems to be little hope *a priori* for quantitative gap prediction with hybrid functionals. Nevertheless, results are once again often surprisingly accurate. A systematic study of 20 small- and medium-sized π-system molecules has revealed that the difference between the B3LYP and experimental absorption gap was never larger than 0.5–0.7 eV (as compared to optical gaps of ~3–7.5 eV), and often much better than that (Salzner *et al.*, 1997). As a second example, absorption spectra of semiconducting single-walled carbon nanotubes (SWNTs) can exhibit significant excitonic effects (Spataru *et al.*, 2004). Nevertheless, the one-parameter hybrid based on the TPSS MGGA yielded the first and second excitation energies of 21 semiconducting SWNTs with unusually small mean absolute errors of only 0.024 and 0.065 eV, respectively (Barone, Peralta, Wert, *et al.*, 2005).

More insight into the role of the fractional Fock exchange in obtaining accurate optical gap values is obtained from detailed comparison of the gap, calculated using a variety of functionals, with the experimental optical gap for all semiconducting nanotubes studied by Barone *et al.* This is shown in Fig. 10. The LSDA, PBE GGA, and TPSS MGGA underestimate the gap, although in this case not as badly as one might expect. The B3LYP and PBE hybrid functionals, on the other hand, overestimate the optical gap just as badly as "semilocal" functionals (including MGGA) underestimate it, because the fraction of Fock exchange in them is too large. The TPSS hybrid, as well as the HSE screened PBE hybrid, however, agree with experiment quite well. This can be rationalized by realizing that, although not constructed with optical gaps in mind, the TPSS hybrid has a smaller fraction of exact exchange (cf. Sec. III.C) and, likewise, HSE also possesses an effectively smaller fraction of exact exchange due to the screening. It is this

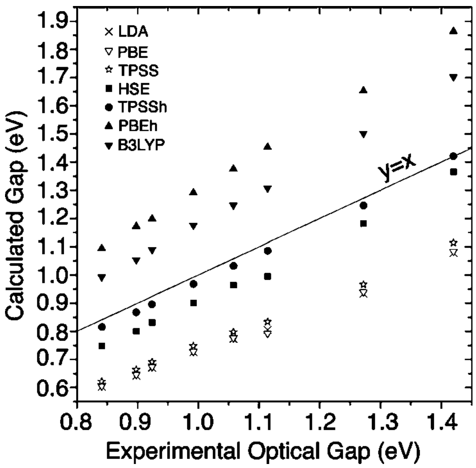

FIG. 10. Theoretical optical gap calculated from eigenvalue differences using LDA, GGA (PBE), MGGA (TPSS), PBE hybrid (PBEh), screened PBE hybrid (HSE), MGGA hybrid (TPSSh), and B3LYP, as compared to the experimental fundamental optical threshold for 21 semiconducting single-wall carbon nanotubes. Reprinted with permission from Barone, Peralta, Wert, *et al.*, 2005. Copyright 2005, American Chemical Society.

reduced fraction that allows these functionals to avoid the optical gap overestimate found with the B3LYP and hybrid PBE functionals for these systems.

It is important to realize that, if hybrid functionals are to yield the optical gap in molecules well, they cannot generally be expected to also yield the fundamental gap well. This is simply because the optical gap is smaller than the fundamental one by the exciton binding energy, Ω (see, e.g., Rohlfing and Louie, 1998), i.e., by the net attraction energy between the quasielectron and the quasihole. Unlike in many solids, in molecules Ω is often appreciable. Detailed numerical examples confirm this reasoning. In one example, the PTCDA molecule of Fig. 2 (Dori *et al.*, 2006), the fundamental gap (computed using either GW or a direct calculation of I and A within DFT based on total energy differences between the neutral species, anion, and cation) is ~5 eV. The LDA gap is only 1.5 eV. The B3LYP value is a larger 2.6 eV. This value is less than impressive with respect to the fundamental gap, but is reasonably close to the experimental optical gap of PTCDA (2.75 eV). In a second example, the (8,0) carbon nanotube (Spataru *et al.*, 2004; Barone, Peralta, Wert, *et al.*, 2005), the fundamental gap computed using GW is 2.54 eV, whereas the LDA gap is only 1.39 eV. The TPSS hybrid value for the fundamental gap is 1.67 eV. This is again unimpressive compared to the fundamental gap, but is in excellent agreement with both the experimental optical gap (1.60 eV) and the optical gap deduced from a correction of the GW gap based on a two-particle Green's function calculation of Ω (1.55 eV). This practical capability of hybrid functionals to predict optical gap trends in some classes of organic compounds is also demonstrated, for example, in recent work on gap trends in graphene nanoribbons [see Barone *et al.* (2006) and Han *et al.* (2007) for theory and experiment, respectively].

From this perspective, the successful prediction of fundamental gaps in solids by hybrid functionals may be seen as a reflection of the fact that in many solids the fundamental gaps simply do not differ substantially from the optical ones. The remaining mystery, however, is why optical gaps are well reproduced at all by hybrid functionals. We have argued above that the Kohn-Sham gap is, in a sense, a zero-order approximation of the optical gap, but clearly this does not offer a satisfactory explanation. First, we have already seen that the Kohn-Sham gap may differ substantially from the optical gap in practice. Second, the gap obtained with a hybrid functional is typically not the Kohn-Sham gap anyway, due the orbital-specific potential. Part of the story is likely that, in practice, the Kohn-Sham gap obtained with semilocal functionals still underestimates even the optical gap systematically (see, e.g., Chelikowsky, Kronik, and Vasilev 2003, as well as Fig. 10). Because mixing in a Fock potential invariably increases the gap, the semilocal functional can still serve as a reasonable starting point for non-Kohn-Sham gap "corrections" via a fraction of Fock exchange. We do caution, however, that we are not aware of any rigorous argument for systematics in quantitative agreement. Indeed, a survey of all hybrid functional results presented in this section, as well as others in the literature, shows that the degree of quantitative agreement is a strong function of both the hybrid functional(s) used and the chemical nature of the system(s) studied. And so, here also, in the absence of rigorous arguments, it is good practice to validate actual performance in a particular chemical environment, using a particular hybrid functional, before true prediction is attempted.

Finally, we note that computations of optical gaps in molecules with the SIC LDA often improve the predicted gap values with respect to the LDA for essentially the same reasons discussed above with respect to computing fundamental gaps with SIC. But once again systematics can be a serious problem. One example for the lack of systematics is found if one compares the computational results of Garza *et al.* (2000), given in Table III, with experimental optical gaps. For H_2O, it is the SIC Kohn-Sham gap (i.e., computed with the KLI approximation for the SIC OEP) that compares well with experiment. But for CH_4 it is the gap computed from localized orbitals that agrees well with experiment.

In summary, the Kohn-Sham gap does not usually compare well with experimental values for either the fundamental or the optical gap. Importantly, this is actually expected on general theoretical grounds. Nevertheless, a range of Kohn-Sham-related techniques alleviates the discrepancy between theoretical and experimental gaps, often to the point where meaningful practical predictions for gap values can be made. The rigor of the rationale behind these different approaches, the systematics of their performance, and the extent to which accuracy in ground-state properties is sacrificed, however, vary very strongly between these different approaches.

Therefore, their prudent application and judicious analysis require that these different factors be taken into account carefully.

B. The optimized effective potential in time-dependent density-functional theory

Much of the discussion in the preceding section concerned extracting, in approximate ways, excited-state information from ground-state calculations. However, there is also a rigorous way of computing excited-state properties in a density-functional framework—time-dependent DFT. Runge and Gross (1984) have shown that the densities $n(\vec{r},t)$ and $n'(\vec{r},t)$ evolving from a common initial state under the influence of two potentials $v(\vec{r},t)$ and $v'(\vec{r},t)$ are always different provided that the potentials differ by more than a purely time-dependent function. This Runge-Gross theorem is the fundamental theorem of TDDFT and it has been shown to be valid on rather general grounds (van Leeuwen, 1999). Actual TDDFT calculations are almost invariably based on the time-dependent Kohn-Sham equations

$$i\hbar\frac{\partial}{\partial t}\varphi_{j,\sigma}(\vec{r},t)=\left(-\frac{\hbar^2}{2m}\nabla^2+v_{\mathrm{KS},\sigma}(\vec{r},t)\right)\varphi_{j,\sigma}(\vec{r},t), \qquad (105)$$

where the definition of the time-dependent Kohn-Sham potential

$$v_{\mathrm{KS},\sigma}(\vec{r},t)=v_{\mathrm{ext}}(\vec{r},t)+v_H(\vec{r},t)+v_{\mathrm{xc},\sigma}(\vec{r},t) \qquad (106)$$

and the logic of mapping to a noninteracting system are in close analogy to Eq. (7) and the static Kohn-Sham approach. With the first formal difficulties (v representability) addressed (van Leeuwen, 1999), one may be tempted to believe that the time-dependent Kohn-Sham scheme in all of its aspects will be similar to the static case and calculations can proceed straightforwardly. However, this is not the case. The time dependence brings in new subtlety and raises questions that were of no concern in ground-state, i.e., static, DFT. We will not attempt to discuss the beauties and woes of TDDFT in general, since these have been the topic of numerous articles and reviews (see, e.g., Casida, 1996; Gross, Dobson, and Petersilka, 1996; van Leeuwen, 2001; Maitra, Burke, Appel, *et al.*, 2002; Marques and Gross, 2004; Burke *et al.*, 2005; Marques *et al.*, 2006; Botti *et al.*, 2007; Elliot *et al.*, 2007). Also, we do not systematically present details of how the functionals that were discussed perform if they are used in a time-dependent context. This would lead to an uninspiring shopping list of results. Instead, the focus of this section is on how orbital-dependent functionals can or cannot help to answer fundamental questions in TDDFT, and thus may improve the performance of TDDFT in typical applications.

A first obvious difference between the time-dependent and the ground-state case is the appearance of the initial-state dependence in the Runge-Gross theorem. This initial-state dependence is intricately related (Maitra, Burke, Appel, *et al.*, 2002; Maitra, Burke, and Woodward, 2002) to memory effects, i.e., to the fact that the exchange-correlation potential at time t depends not only on the density at time t, but also on the density at all prior times $t'<t$ (see, e.g., Dobson and Dinte, 1996; Gross, Dobson, and Petersilka, 1996; Hessler *et al.*, 2002; Kurzweil and Baer, 2006). In the frequency domain, this nonlocality in time of the exchange-correlation kernel defined in Eq. (107) translates into a frequency dependence. A second difference is that the potential plays a more prominent role in the time-dependent case than in the static case. This statement requires elaboration. In ground-state DFT, one typically thinks about E_{xc} first and v_{xc} just follows as its functional derivative. This is a consequence of the Ritz variational principle on which the proof of the Hohenberg-Kohn theorem rests. There is no direct analog of this principle in the time-dependent case. The first derivation of the time-dependent Kohn-Sham equations (Runge and Gross, 1984) employed the stationary action principle to derive the time-dependent exchange correlation potential as a functional derivative $\delta/\delta n(\vec{r},t)$ of an exchange-correlation action functional as one would naturally define it. But it was later realized that this definition leads to the following inconsistency (see, e.g., the discussion in Gross, Dobson, and Petersilka, 1996). The exchange-correlation kernel, which plays a prominent role in the linear-response version of TDDFT (see below), is defined as the functional derivative

$$f_{\mathrm{xc},\sigma}([n_0];\vec{r},t,\vec{r}',t')=\left.\frac{\delta v_{\mathrm{xc},\sigma}([n];\vec{r},t)}{\delta n_\sigma(\vec{r}',t')}\right|_{n_0} \qquad (107)$$

and is directly related to the response function [cf. Eq. (95)]. [Note that here we are only concerned with the "real" interaction, i.e., $\lambda=1$. Therefore, in contrast to Eq. (95), there is no index for the coupling constant in Eq. (107).] Because of causality, a response function must vanish for $t'>t$, and by virtue of Eq. (95), the same must then hold for the kernel. But with both the exchange-correlation potential and the kernel defined as functional derivatives with respect to the time-dependent density, one concludes from Schwarz's lemma that the kernel must be symmetric in t and t', which contradicts causality. This paradox was resolved by van Leeuwen (1998) by defining a different action functional within the Keldysh formalism. However, comparing now with the ground-state case, one notes that (a) the action is typically less accessible to intuition than the energy and (b) observables that are typically of interest in TDDFT calculations (e.g., the time-dependent dipole moment, which leads to the frequency-dependent polarizability and thus the photoabsorption spectrum) are not calculated from the action, but from the time-dependent density itself. Therefore, one may argue that in TDDFT it is really the potential that plays the central role.

The natural next question then is how $v_{\mathrm{xc}}([n];\vec{r},t)$ is calculated, for example, for a simple case such as noninteracting fermions, i.e., the exchange-only situation. The corresponding action functional is

$$A_x[n] = -\frac{1}{2}\sum_{\substack{i,j=1\\ \sigma=\uparrow,\downarrow}}^{N_\sigma} f_{i,\sigma}f_{j,\sigma}\int_C dt\int d^3r\int d^3r' \times \frac{\varphi^*_{i,\sigma}(\vec{r}\,',\tau)\varphi_{j,\sigma}(\vec{r}\,',\tau)\varphi_{i,\sigma}(\vec{r},\tau)\varphi^*_{j,\sigma}(\vec{r},\tau)}{|\vec{r}-\vec{r}\,'|}, \quad (108)$$

where the integration is along the Keldysh contour C (van Leeuwen, 1998). This, obviously, is an orbital-dependent functional closely related to the Fock integral of Eq. (37). Thus we arrive at the general question of how the time-dependent exchange-correlation potential can be calculated for an orbital-dependent action functional such as that of Eq. (108). The answer has been given by van Leeuwen (1998), who showed that by taking the appropriate functional derivative one arrives at the following integro-differential equation for the exchange-correlation potential:

$$0 = \sum_{j=1}^{N_\sigma}\frac{i}{\hbar}\int\int \varphi^*_{j,\sigma}(\vec{r}\,',t')[v_{\mathrm{xc},\sigma}(\vec{r}\,',t') - u_{\mathrm{xc}j,\sigma}(\vec{r}\,',t')] \times \sum_{k=1}^{\infty}\varphi^*_{k,\sigma}(\vec{r},t)\varphi_{k,\sigma}(\vec{r}\,',t')\theta(t-t')\varphi_{j,\sigma}(\vec{r},t)dt'd^3r' + \mathrm{c.c.}, \quad (109)$$

where

$$u_{\mathrm{xc}j,\sigma}(\vec{r},t) = \frac{1}{\varphi^*_{j,\sigma}(\vec{r},t)}\left.\frac{\delta A_{\mathrm{xc}}[\{\varphi_{i,\tau}\}]}{\delta\varphi_{j,\sigma}(\vec{r},\tau)}\right|_{\varphi_{i,\tau}=\varphi_{i,\tau}(\vec{r},t)}. \quad (110)$$

Equation (109) is the famous time-dependent optimized effective potential equation (TDOEP), which was first derived by Ullrich *et al.* (1995) in a somewhat different way.

In analogy to procedures discussed for the static OEP equation in Sec. II.A, the TDOEP of Eq. (109) can be cast into the simpler form (Ullrich *et al.* 1995; Mundt and Kümmel, 2006)

$$\sum_{j=1}^{N_\sigma}\psi^*_{j,\sigma}(\vec{r},t)\varphi_{j,\sigma}(\vec{r},t) + \mathrm{c.c.} = g_\sigma(\vec{r},t), \quad (111)$$

where

$$g_\sigma(\vec{r},t) = \frac{i}{\hbar}\sum_{j=1}^{N_\sigma}|\varphi_{j,\sigma}(\vec{r},t)|^2\int_{-\infty}^{t}[\bar{u}_{\mathrm{xc}j,\sigma}(t') - \bar{u}^*_{\mathrm{xc}j,\sigma}(t')]dt' \quad (112)$$

vanishes for functionals depending on $\{\varphi_{i,\sigma}\}$ only through the combination $\varphi_{i,\sigma}(\vec{r},t)\varphi^*_{i,\sigma}(\vec{r},t)$, like the exact-exchange functional (Gross, Petersilka, and Grabo, 1996). The orbital average $\bar{u}_{\mathrm{xc}j,\sigma}(t)$ is defined by

$$\bar{u}_{\mathrm{xc}j,\sigma}(t) = \int \varphi^*_{j,\sigma}(\vec{r},t)u_{\mathrm{xc}j,\sigma}(\vec{r},t)\varphi_{j,\sigma}(\vec{r},t)d^3r \quad (113)$$

and

$$\psi_{j,\sigma}(\vec{r},t) = -\frac{i}{\hbar}\int\int\varphi_{j,\sigma}(\vec{r}\,',t)[v_{\mathrm{xc},\sigma}(\vec{r}\,',t) - u^*_{\mathrm{xc}j,\sigma}(\vec{r}\,',t)] \times \sum_{\substack{k=1\\ k\neq j}}^{\infty}\varphi^*_{k,\sigma}(\vec{r}\,',t)\varphi_{k,\sigma}(\vec{r},t)\theta(t-t')dt'd^3r' \quad (114)$$

is the time-dependent orbital shift [in analogy to Eq. (28)], obeying the orthogonality relation

$$\int\psi^*_{j,\sigma}(\vec{r},t)\varphi_{j,\sigma}(\vec{r},t)d^3r = 0 \quad (115)$$

and the inhomogenous time-dependent Kohn-Sham equation

$$\left(i\hbar\frac{\partial}{\partial t} - h_{\mathrm{KS},\sigma}(\vec{r},t)\right)\psi_{j,\sigma}(\vec{r},t) = \{v_{\mathrm{xc},\sigma}(\vec{r},t) - u^*_{\mathrm{xc}j,\sigma}(\vec{r},t) - [\bar{v}_{\mathrm{xc}j,\sigma}(t) - \bar{u}^*_{\mathrm{xc}j,\sigma}(t)]\}\varphi_{j,\sigma}(\vec{r},t). \quad (116)$$

The time-dependent orbital average $\bar{v}_{\mathrm{xc}j,\sigma}(t)$ is defined in analogy to Eqs. (113) and (32). Multiplying Eq. (111) by v_{KS} and using Eqs. (116) and (105), one can derive

$$v_{\mathrm{xc},\sigma}(\vec{r},t) = \frac{1}{2n_\sigma(\vec{r},t)}\sum_{j=1}^{N_\sigma}\left[|\varphi_{j,\sigma}(\vec{r},t)|^2\left(u_{\mathrm{xc}j,\sigma}(\vec{r},t) + [\bar{v}_{\mathrm{xc}j,\sigma}(t) - \bar{u}_{\mathrm{xc}j,\sigma}(t)] - \frac{\hbar^2}{m}\vec{\nabla}\cdot[\psi^*_{j,\sigma}(\vec{r},t)\vec{\nabla}\varphi_{j,\sigma}(\vec{r},t)]\right)\right] + \mathrm{c.c.} - f(\vec{r},t), \quad (117)$$

where

$$f(\vec{r},t) = \frac{1}{2n_\sigma(\vec{r},t)}\left\{-\frac{\hbar^2}{2m}\nabla^2 g(\vec{r},t) + i\hbar\frac{\partial}{\partial t}\sum_{j=1}^{N_\sigma}[\psi^*_{j,\sigma}(\vec{r},t)\varphi_{j,\sigma}(\vec{r},t) - \mathrm{c.c.}]\right\}. \quad (118)$$

In this shorter form, the TDOEP equation naturally lends itself to numerical solution attempts, discussed below.

Comparing Eqs. (109) and (110) to Eqs. (24) and (26), or alternatively, Eq. (111) to Eq. (27), or Eq. (117) to Eq. (35), one notes a close analogy between the static and time-dependent equations. Since the static OEP equation was already hard to solve, one can expect—and all evidence so far indicates that this expectation is justified—that solving the TDOEP equation will be a still harder problem. So why should one attempt to do it? We believe that the answer to this question must be given in two parts. The first part is that, indeed, maybe one does not need to solve the TDOEP equation in each and every TDDFT calculation. Instead, the problem can be sidestepped in a manner similar to that pursued in the static case, i.e., by using (time-dependent) orbital-specific potentials. Orbital-specific potentials are used, for example, in time-dependent Hartree-Fock theory. However, the analog to the static case is a far-reaching one, in the sense that the possible drawbacks of using orbital-specific potentials that were mentioned in Sec. II.B (e.g., not being on the safe grounds of Kohn-Sham theory, interpretability, computational disadvantages, etc.) also carry over to the time-dependent case. Therefore, as the second part of the answer, we argue that at the very least one needs the possibility to do TDOEP calculations for a few pertinent systems and for each new type of orbital-dependent functional. This will allow us to examine whether straying from Kohn-Sham territory by using orbital-specific potentials introduces major differences.

Understanding the question posed above in a broader sense leads one to wonder what benefits are gained from using orbital-dependent functionals in TDDFT. There are clear answers to this question. First, orbital-dependent functionals are attractive in TDDFT for some of the same reasons that they are attractive in DFT and which have been repeatedly addressed above: First, the possibility of achieving freedom from self-interaction, the related beneficial properties of the Kohn-Sham potential, and the inclusion of a derivative discontinuity. Second, orbital-dependent functionals include memory effects (Ullrich *et al.*, 1995), because the orbitals at time t depend on the density at all prior times $t' < t$ (Maitra and Burke, 2001; Maitra, Burke, Appel, *et al.*, 2002). Third, orbital-dependent functionals are part of a systematic method formally leading to the exact exchange-correlation potential via the time-dependent generalization of the Görling-Levy perturbation expansion (Görling, 1997) or the Keldysh diagrammatic method (van Leeuwen, 1996, 1998). Why and how some of these possible benefits are helpful in actual calculations is discussed in the final part of this section. In accordance with the two different domains in which TDDFT is used, we split this discussion into the linear-response and the nonperturbative regime.

1. Linear-response regime

In the typical linear-response formulation of TDDFT (Casida, 1996; Gross, Dobson, and Petersilka, 1996; Petersilka *et al.*, 1996), the true excitation energies of an interacting system can be viewed as being built from two parts: (a) the excitations of the noninteracting Kohn-Sham system, which correspond to the differences of Kohn-Sham eigenvalues; and (b) a term involving the exchange-correlation kernel that takes the interaction effects into account. Correspondingly, the exchange-correlation approximation used in an actual calculation has an influence on the results in two ways. First, the quality of the exchange-correlation potential determines how physically meaningful the Kohn-Sham eigenvalue differences are. In other words, the potential determines how far the excitations of the noninteracting system can serve as reasonable approximations to the excitations of the interacting system. The beneficial effects of orbital-dependent functionals in this respect have been discussed in the previous sections. Therefore, here we do not elaborate on them further. We do note, however, that, for example, van Gisbergen *et al.* (1998) have shown that, with eigenvalue differences calculated from highly accurate potentials reconstructed from *ab initio* densities, one can get very close to the true excitation energies (of some systems) even if the exchange-correlation kernel is still computed with one of the simple approximations.

The second way in which orbital-dependent functionals can improve TDDFT calculations is via the exchange-correlation kernel. In fact, they can have a major impact on the kernel because of their above-mentioned inherent nonlocality in time, leading to a frequency-dependent kernel. Correctly modeling or capturing the frequency dependence is an important task. While even simple, non-frequency-dependent kernels like that of the adiabatic local density approximation can provide for a realistic description of excitations in many cases [for historical examples, see, e.g., Peuckert (1978) and Zangwill and Soven (1980); for an overview including more recent results, see, e.g., Chelikowsky Kronik, and Vasiliev (2003) and the above-mentioned TDDFT review articles], the lack of frequency dependence impedes the use of TDDFT for many problems of practical interest. For example, frequency dependence is needed quite generally for describing atomic and molecular excitations accurately (Tozer and Handy, 2000), but in particular for double excitations (Maitra *et al.*, 2004) and long-range charge-transfer excitations (Maitra, 2005). It is crucial for correct excitation energies of dissociating molecules (Gritsenko *et al.*, 2000) and also relevant for excitations in solids (Lein *et al.*, 2000; Tokatly and Pankratov, 2001). [But it should also be noted that, for solids, the *spatial* nonlocality (Tatarczyk *et al.*, 2001) is crucial, for example, for describing excitonic effects (Kim and Görling, 2002a; Onida *et al.*, 2002; Reining *et al.*, 2002; Marini *et al.*, 2003)].

While much insight has been gained in these studies, the use of orbital-dependent functionals for calculating excitations is still not easy. Many of the studies just mentioned employed kernels derived from linking TDDFT with other many-body methods, i.e., they are not as computationally easy to apply as one would wish. But even for functionals where an explicit expression in terms of the orbitals is known, such as the exact-exchange energy, deriving the kernel is not easy. This is intuitively clear because calculating the functional derivative to obtain the TDOEP is not easy and for the kernel, a second functional derivative needs to be evaluated. For the exact-exchange case, the kernel has been derived (Görling, 1998; Kim and Görling, 2002b) and employed successfully (Kim and Görling, 2002a; Shigeta *et al.*, 2006). But without a compatible correlation functional to go with it, it will be of limited usefulness. Thus, as in the ground-state case, the construction of correlation functionals that are compatible with exact exchange and yet computationally feasible for general applications continues to be an important task for future work.

2. Nonperturbative excitations

Whereas most applications of TDDFT belong in the linear-response regime, it was realized early (Saalmann and Schmidt, 1996; Ullrich *et al.*, 1996; Yabana and Bertsch, 1996; Tong and Chu, 1997; Reinhard *et al.*, 1999) that TDDFT also allows us to study nonlinear, nonperturbative excitations. This is a very attractive feature of TDDFT because nonperturbative excitations as triggered, for example, by the fields of intense lasers, are of great interest due to the rapid progress in laser technology. Because the time-dependent many-electron Schrödinger equation is even harder to solve than the ground-state Schrödinger equation (and solving the latter is not particularly easy), TDDFT is highly attractive as an alternative first-principles approach allowing for the quantum-mechanical investigation of strong electronic excitations. The attractiveness of TDDFT in the nonperturbative regime can be exemplified by the observation that, even for a two-electron system, calculating from the wave function the long-ranged nonperturbative dynamics that is triggered by the low-frequency lasers typically used in strong-field experiments is a serious computational challenge (van der Hart *et al.*, 2005). In TDDFT, the computational hurdle is much lower, because the time-dependent Kohn-Sham orbitals can be propagated according to Eq. (105), i.e., within a single-particle framework. Thus, the time-dependent density that results from a nonlinear, nonperturbative excitation can be calculated via orbital propagation and summation over the time-dependent orbitals. In this way, any observable that can readily be calculated from the density, e.g., the time-dependent dipole moment, is easily accessible. [We will not go into the difficulties one encounters with observables that are not easily calculated from the density, like ionization probabilities, but we refer the reader to Lappas and van Leeuwen (1998), Petersilka and Gross (1999), and de Wijn *et al.* (2007) for a discussion of this problem.] It should be noted that orbital propagation has become a well-established technique [see, e.g., Calvayrac *et al.* (2000) and Marques *et al.* (2001, 2003) for reviews] and it does not require evaluation of the exchange-correlation kernel. This is an advantage when calculating the kernel is cumbersome, as it is for orbital-dependent functionals.

The use of orbital-dependent functionals is highly desirable in the nonperturbative regime for at least two reasons. First, if long-ranged dynamics or ionization is investigated, the potential, and in particular its long-range asymptotics, plays an even more important role than in the linear-response regime. Therefore, studying such processes requires functionals that are free from self-interaction error. As one such functional, the time-dependent exact-exchange energy has been employed (Ullrich *et al.*, 1996); as another, the time-dependent SIC scheme has been used (Tong and Chu, 1997, 1998, 2001; Ullrich *et al.*, 1998, 2000; Calvayrac *et al.*, 2000; Pohl *et al.*, 2000; Chu, 2005). The latter has become popular because of its relatively low computational demands.

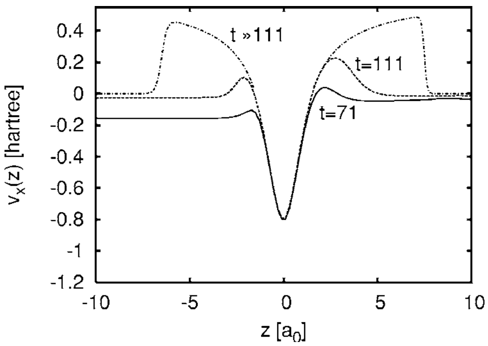

FIG. 11. An electron escapes from a (one-dimensional) Li atom. The atom is subjected to a strong, ionizing external field $F=0.06$ a.u., which draws electrons to the left. The time-dependent exchange potential (KLI approximation) of two spin-up electrons at different times as labeled (in a.u.). At time $t=0$, the atom was in its ground state. The external field was linearly ramped up over 40 a.u. and finally held constant. Note the buildup of the steplike structure that is a hallmark of the particle-number discontinuity. $t \gg 111$ denotes an extrapolation to very large times. This figure should be compared to Fig. 4. From Mundt and Kümmel, 2005.

A second aspect that makes orbital-dependent functionals attractive for nonperturbative TDDFT is the fact that they incorporate particle-number discontinuities. The importance of such discontinuities for the proper description of nonperturbative dynamics has recently been demonstrated by calculating the *exact* time-dependent exchange-correlation potential associated with an ionization process [Lein and Kümmel (2005)]. The potential was obtained by feeding the exact (within numerical limits) time-dependent density, coming out of the time-dependent two-electron Schrödinger equation, into the inverted time-dependent Kohn-Sham procedure. The exchange-correlation potential corresponding to the ionization process shows a pronounced steplike structure that resembles the steplike structure seen in Fig. 4. These numerical findings were put into perspective in further work showing on analytical grounds that discontinuities must show up in the TDOEP (Mundt and Kümmel, 2005). The relation between the steplike structures and the derivate discontniuity is stressed by the fact that the height of the step in the exact time-dependent exchange-correlation potential was shown to be close to the value of the derivative discontinuity. A similar finding was also made in exchange-only calculations. Part of one such calculation is shown in Fig. 11. Depicted are snapshots of the exchange potential during an ionization process in which an electron leaves a model lithium atom by escaping to the left. (The escape is driven by an external electrical field.) One clearly sees the buildup of steplike structures in the potential that can be interpreted as a hallmark of the particle-number discontinuity (see Fig. 4 and the discussion there).

Orbital-dependent functionals thus show very attractive features that promise to increase the accuracy of nonperturbative TDDFT calculations enormously. But these benefits come at the price of having to solve the

TDOEP equation, and here we reach the limits of present capabilities. All orbital propagation calculations that were mentioned above and that employed orbital-dependent functionals, be they SIC or exact exchange, were based on the time-dependent KLI (TDKLI) approximation (Ullrich *et al.*, 1995) and not on the true TDOEP. The TDKLI potential is obtained from Eq. (117) by neglecting the divergence term and the $f(\vec{r},t)$ term. One may hope that, just as the KLI approximation of ground-state DFT is a good approximation to the OEP in many cases, the TDKLI approximation may be a good approximation to the TDOEP in many cases. However, the response calculations for the molecular chains discussed in Sec. II.C with their pronounced differences between the KLI approximation and OEP show that TDKLI will definitely not be a useful approximation for all systems. Therefore, a scheme to solve the time-dependent OEP equation beyond the linear-response regime is highly desirable. This need is further stressed by numerical studies indicating that the TDKLI approximation violates the fundamental zero force theorem (Mundt *et al.*, 2007). As one natural candidate for a numerical TDOEP approach, a scheme has been tested in which the TDOEP is calculated from propagation of the time-dependent orbitals and orbital shifts based on Eqs. (114)–(117) (Mundt and Kümmel, 2006). However, numerical instabilities severely limit the usefulness of this scheme, at least in its present version. Therefore, only approximations to the TDOEP and not the true TDOEP itself can be employed at present in propagation calculations. This limitation defines another task for future research.

V. SUMMARY

In this paper, we have attempted to provide a broad perspective and a critical review of the use of orbital-dependent functionals within density-functional theory.

In the first section, we presented a short summary of the rationale behind and the successes of semilocal, orbital-independent functionals. We then identified two major shortcomings of such functionals, namely, the presence of spurious self-interaction and the absence of a functional derivative discontinuity. These two problems are related to each other, and we explained why both may, in principle, be solved or at least significantly ameliorated using orbital-dependent functionals.

In the second section, we presented the optimized effective potential as the rigorous approach to evaluating the functional derivative of the orbital-dependent exchange-correlation energy, which is needed for incorporating orbital-dependent functionals within the usual Kohn-Sham framework. We also discussed the differences and similarities between the OEP and the orbital-specific nonlocal potentials often employed, which arise from energy minimization with respect to orbitals rather than density. The OEP equation is, in general, a complicated integro-differential equation. We have therefore provided an in-depth discussion of popular approximate forms of it, numerical approaches for its solution, and proper pseudopotential construction within the OEP framework.

In the third section, we surveyed popular and promising practical orbital-dependent functionals. We started with the meta-GGA class of functionals, which is formally orbital dependent, but in practice close in spirit to semilocal functionals. We then discussed self-interaction-corrected functionals and hybrid functionals as two very different pragmatic approaches to overcoming the limitations of semilocal functionals. Finally, we discussed the relatively young but highly promising field of orbital-dependent functionals based on correlation that is compatible with exact exchange.

In the fourth and final section, we turned our attention to excited-state properties. First, we explained why experimental gaps differ in principle from the Kohn-Sham one. We then reviewed how and why various orbital-dependent functionals provide a partial remedy to this issue. Finally, we discussed the emerging field of orbital-dependent functionals within time-dependent Kohn-Sham theory, with an emphasis on the time-dependent optimized effective potential.

Throughout this review, we have attempted to provide a balanced perspective, in which both formal and practical aspects are presented. In particular, we tried to emphasize the practical manifestations of formal approximations, as well as the formal justification, if any, of practical shortcuts. In modification of a well known proverb, one can say that to some extent, developing density functionals is like searching for a black cat in a dark room (and with eyes blindfolded in the case of TDDFT). But due to Hohenberg and Kohn and Runge and Gross, we at least have the assurance that we are not searching in an empty room—the cat is actually there. Time and again, the developers and users of DFT and TDDFT find torches—tools to make the hunt more efficient. We strongly believe that orbital-dependent functionals represent such a torch, and that this one may keep burning for quite some time.

ACKNOWLEDGMENTS

We are grateful to R. Baer, K. Burke, J. R. Chelikowsky, T. Körzdörfer, J. E. Peralta, J. P. Perdew, P. Rinke, G. E. Scuseria, V. N. Staroverov, M. Thiele, D. J. Tozer, and O. A. Vydrov for discussions and comments. S.K. gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft and the Deutscher Akademischer Austauschdienst. L.K. acknowledges financial support from the Israel Science Foundation and the Schmidt Minerva Center for Supramolecular Chemistry.

REFERENCES

Adamo, C., and V. Barone, 1997, Chem. Phys. Lett. **274**, 242.
Adamo, C., and V. Barone, 1998, Chem. Phys. Lett. **298**, 113.
Adamo, C., and V. Barone, 1999, J. Chem. Phys. **110**, 6158.
Adamo, C., M. Ernzerhof, and G. E. Scuseria, 2000, J. Chem. Phys. **112**, 2643.

Adamson, R. D., P. M. W. Gill, and J. A. Pople, 1998, Chem. Phys. Lett. **284**, 6.

Akola, J., M. Manninen, H. Häkkinen, U. Landmann, X. Li, and L.-S. Wang, 2000, Phys. Rev. B **62**, 13216.

Alfe, D., and M. J. Gillan, 2006, J. Phys.: Condens. Matter **18**, L435.

Allen, M. J., and D. J. Tozer, 2002, Mol. Phys. **100**, 433.

Almbladh, C.-O., and von Barth, U., 1985, Phys. Rev. B **31**, 3231.

Al-Sharif, A. I, R. Resta, and C. J. Umrigar, 1998, Phys. Rev. A **57**, 2466.

Ambrosch-Draxl, C., 2004, Phys. Scr., T **T109**, 48.

Andersson, Y., D. C. Langreth, and B. I. Lundqvist, 1996, Phys. Rev. Lett. **76**, 102.

Ángyán, J. G., I. C. Gerber, A. Savin, and J. Toulouse, 2005, Phys. Rev. A **72**, 012510.

Anisimov, V. I., F. Aryasetiawan, and A. I. Lichtenstein, 1997, J. Phys.: Condens. Matter **9**, 767.

Arbuznikov, A., and M. Kaupp, 2003, Chem. Phys. Lett. **381**, 495.

Arbuznikov, A. V., and M. Kaupp, 2004, Chem. Phys. Lett. **386**, 8.

Arbuznikov, A. V., and M. Kaupp, 2007, Chem. Phys. Lett. **440**, 160.

Arbuznikov, A. V., M. Kaupp, and H. Bahmann, 2006, J. Chem. Phys. **124**, 204102.

Aryasetiawan, F., and M. J. Stott, 1988, Phys. Rev. B **38**, 2974.

Asahi, R., W. Mannstadt, and A. J. Freeman, 1999, Phys. Rev. B **59**, 7486.

Asahi, R., W. Mannstadt, and A. J. Freeman, 2000, Phys. Rev. B **62**, 2552.

Asahi, R., A. Wang, J. R. Babcock, N. L. Edleman, A. W. Metz, M. A. Lane, V. P. Dravid, C. R. Kannewurf, A. J. Freeman, and T. J. Marks, 2002, Thin Solid Films **411**, 101.

Aulbur, W. G., M. Städele, and A. Görling, 2000, Phys. Rev. B **62**, 7121.

Baboul, A. G., L. A. Curtiss, P. C. Redfern, and K. Raghavachari, 1999, J. Chem. Phys. **110**, 7650.

Bachelet, G. B., D. R. Hamman, and M. Schlüter, 1982, Phys. Rev. B **26**, 4199; **29**, 2309(E) (1984).

Baer, R., and D. Neuhauser, 2005, Phys. Rev. Lett. **94**, 043002.

Baer, R., and D. Neuhauser, 2006, J. Chem. Phys. **125**, 074709.

Baer, R., E. Livshits, and D. Neuhauser, 2006, Chem. Phys. **329**, 266.

Baerends, E. J., and O. V. Gritsenko, 1997, J. Phys. Chem. A **101**, 5383.

Baerends, E. J., and O. V. Gritsenko, 2005, J. Chem. Phys. **123**, 062202.

Bahmann, H., A. Rodenberg, A. V. Arbuznikov, and M. Kaupp, 2007, J. Chem. Phys. **126**, 011103.

Baker, J., J. Andzelm, J. M. Muir, and P. R. Taylor, 1995, Chem. Phys. Lett. **237**, 53.

Bally, T., and G. N. Sastry, 1997, J. Phys. Chem. A **101**, 7923.

Barone, V., O. Hod, and G. E. Scuseria, 2006, Nano Lett. **6**, 2748.

Barone, V., J. E. Peralta, and G. E. Scuseria, 2005, Nano Lett. **5**, 1830.

Barone, V., J. E. Peralta, M. Wert, J. Heyd, and G. E. Scuseria, 2005, Nano Lett. **5**, 1621.

Bartlett, R. J., V. F. Lotrich, and I. V. Schweigert, 2005, J. Chem. Phys. **123**, 062205.

Batista, E. R., J. Heyd, R. G. Hennig, B. P. Uberuaga, R. L. Martin, G. E. Scuseria, C. J. Umrigar, and J. W. Wilkins, 2006, Phys. Rev. B **74**, 121102.

Beck, T. L., 2000, Rev. Mod. Phys. **72**, 1041.

Becke, A. D., 1986, J. Chem. Phys. **84**, 4524.

Becke, A. D., 1988, Phys. Rev. A **38**, 3098.

Becke, A. D., 1992a, J. Chem. Phys. **96**, 2155.

Becke, A. D., 1992b, J. Chem. Phys. **97**, 9173.

Becke, A. D., 1993a, J. Chem. Phys. **98**, 1372.

Becke, A. D., 1993b, J. Chem. Phys. **98**, 5648.

Becke, A. D., 1996, J. Chem. Phys. **104**, 1040.

Becke, A. D., 1997, J. Chem. Phys. **107**, 8554.

Becke, A. D., 1998, J. Chem. Phys. **109**, 2092.

Becke, A. D., 2003, J. Chem. Phys. **119**, 2972.

Becke, A. D., 2005, J. Chem. Phys. **122**, 064101.

Becke, A. D., and E. R. Johnson, 2006, J. Chem. Phys. **124**, 221101.

Becke, A. D., and M. R. Roussel, 1989, Phys. Rev. A **39**, 3761.

Blake, N. P., and H. Metiu, 1999, J. Chem. Phys. **110**, 7457.

Blöchl, P. E., 1990, Phys. Rev. B **41**, 5414.

Boese, A. D., and N. C. Handy, 2002, J. Chem. Phys. **116**, 9559.

Boese, A. D., W. Klopper, and J. M. L. Martin, 2005, Int. J. Quantum Chem. **104**, 830.

Boese, A. D., and J. M. L. Martin, 2004, J. Chem. Phys. **121**, 3405.

Botti, S., A. Schindlmayr, R. Del Sole, and L. Reining, 2007, Rep. Prog. Phys. **70**, 357.

Bouř, P., 2000, J. Comput. Chem. **21**, 8.

Bredow, T., and A. R. Gerson, 2000, Phys. Rev. B **61**, 5194.

Burke, K., M. Ernzerhof, and J. P. Perdew, 1997, Chem. Phys. Lett. **265**, 115.

Burke, K., J. P. Perdew, and M. Ernzerhof, 1997, Int. J. Quantum Chem. **61**, 287.

Burke, K., J. Werschnik, and E. K. U. Gross, 2005, J. Chem. Phys. **123**, 062206.

Bylander, D. M., and L. Kleinman, 1990, Phys. Rev. B **41**, 7868.

Bylander, D. M., and L. Kleinman, 1995a, Phys. Rev. Lett. **74**, 3660.

Bylander, D. M., and L. Kleinman, 1995b, Phys. Rev. B **52**, 14566.

Bylander, D. M., and L. Kleinman, 1996, Phys. Rev. B **54**, 7891.

Bylander, D. M., and L. Kleinman, 1997, Phys. Rev. B **55**, 9432.

Calvayrac, F., P.-G. Reinhard, E. Suraud, and C. A. Ullrich, 2000, Phys. Rep. **337**, 493.

Caratzoulas, S., and P. J. Knowles, 2000, Mol. Phys. **98**, 1811.

Casida, M. E., 1995, Phys. Rev. A **51**, 2005.

Casida, M. E., 1996, in *Recent Developments and Applications in Density-Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam), pp. 155–192.

Ceperley, D. M., and B. J. Alder, 1980, Phys. Rev. Lett. **45**, 566.

Champagne, B., F. A. Bulat, W. Yang, S. Bonness, and B. Kirtman, 2006, J. Chem. Phys. **125**, 194114.

Champagne, B., D. H. Mosley, M. Vracko, and J.-M., André, 1995a, Phys. Rev. A **52**, 178.

Champagne, B., D. H. Mosley, M. Vracko, and J.-M., André, 1995b, Phys. Rev. A **52**, 1039.

Champagne, B., E. A. Perpéte, S. J. A. van Gisbergen, E.-J. Baerends, J. G. Snijders, C. Soubra-Ghaoui, K. A. Robins, and B. Kirtman, 1998, J. Chem. Phys. **109**, 10489.

Chan, G. K.-L., 1999, J. Chem. Phys. **110**, 4710.

Chelikowsky, J. R., and M. L. Cohen, 1992, in *Handbook on Semiconductors*, edited by P. Landsberg (Elsevier, Amsterdam), Vol. 1, p. 59.

Chelikowsky, J. R., L. Kronik, and I. Vasiliev, 2003, J. Phys.: Condens. Matter **15**, R1517.

Chelikowsky, J. R., L. Kronik, I. Vasiliev, M. Jain, and Y. Saad, 2003, in *Handbook of Numerical Analysis—Volume X: Computational Chemistry*, edited by C. Le Bris (Elsevier, Amsterdam), pp. 613–637.

Chong, D. P, Gritsenko, O. V., and Baerends, E. J., 2002, J. Chem. Phys. **116**, 1760.

Chu, S.-I, 2005, J. Chem. Phys. **123**, 062207.

Ciofini, I., H. Chermette, and C. Adamo, 2003, Chem. Phys. Lett. **380**, 12.

Cohen, A. J., and N. C. Handy, 2001, Mol. Phys. **99**, 607.

Cohen, M. L., and J. R. Chelikowsky, 1988, *Electronic Structure and Optical Properties of Semiconductors* (Springer-Verlag, Berlin).

Cole, L. A., and J. P. Perdew, 1982, Phys. Rev. A **25**, 1265.

Colle, R., and D. Salvetti, 1975, Theor. Chim. Acta **37**, 329.

Colle, R., and D. Salvetti, 1979, Theor. Chim. Acta **53**, 55.

Colle, R., and O. Salvetti, 1990, J. Chem. Phys. **93**, 534.

Corá, F., M. Alfredsson, G. Mallia, D. S. Middlemiss, W. C. Mackrodt, R. Dovesi, and R. Orlando, 2004, Struct. Bonding (Berlin) **113**, 171.

Cortona, P., 1986, Phys. Rev. A **34**, 769.

Cremer, D., 2001, Mol. Phys. **99**, 1899.

Cremer, D., M. Filatov, V. Polo, E. Kraka, and S. Shaik, 2002, Int. J. Mol. Sci. **3**, 604.

Csonka, G. I., and B. G. Johnson, 1998, Theor. Chem. Acc. **99**, 158.

Curtiss, L. A., C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, 1990, J. Chem. Phys. **93**, 2537.

Curtiss, L. A., K. Raghavachari, P. C. Redfern, and J. A. Pople, 1997, J. Chem. Phys. **106**, 1063.

Curtiss, L. A., K. Raghavachari, P. C. Redfern, and J. A. Pople, 2000, J. Chem. Phys. **112**, 7374.

Curtiss, L. A., P. C. Redfern, K. Raghavachari, and J. A. Pople, 1999, J. Chem. Phys. **109**, 42.

Della Sala, F., and A. Görling, 2001, J. Chem. Phys. **115**, 5718.

Della Sala, F., and A. Görling, 2002a, J. Chem. Phys. **116**, 5374.

Della Sala, F., and A. Görling, 2002b, Phys. Rev. Lett. **89**, 033003.

de Wijn, A. S., S. Kümmel, and M. Lein, 2007, J. Comput. Phys. **226**, 89.

Dickson, R. M., and A. D. Becke, 2005, J. Chem. Phys. **123**, 111101.

di Valentin, C., G. Pacchioni, and A. Selloni, 2006, Phys. Rev. Lett. **97**, 166803.

Dobson, J. F., M. J. Bünner, and E. K. U. Gross, 1997, Phys. Rev. Lett. **79**, 1905.

Dobson, J. F., and B. P. Dinte, 1996, Phys. Rev. Lett. **76**, 1780.

Dobson, J. F., and J. Wang, 1999, Phys. Rev. Lett. **82**, 2123.

Dori, N., M. Menon, L. Kilian, M. Sokolowski, L. Kronik, and E. Umbach, 2006, Phys. Rev. B **73**, 195208.

Dovesi, R., R. Orlando, C. Roetti, C. Pisani, and V. R. Saunders, 2000, Phys. Status Solidi B **217**, 63.

Dreizler, M., and E. K. U. Gross, 1990, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, Berlin).

Dreuw, A., and M. Head-Gordon, 2004, J. Am. Chem. Soc. **126**, 4007.

Dreuw, A., J. L. Weisman, and M. Head-Gordon, 2003, J. Chem. Phys. **119**, 2943.

Edmiston, C., and K. Ruedenberg, 1963, Rev. Mod. Phys. **35**, 457.

Elliot, P., K. Burke, and F. Furche, 2007, arXiv:cond-mat/0703590.

Engel, E., 2003, in *A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. A. L. Marques (Springer, Berlin).

Engel, E., and R. M. Dreizler, 1999, J. Comput. Chem. **20**, 31.

Engel, E., A. Höck, and R. M. Dreizler, 2000a, Phys. Rev. A **61**, 032502.

Engel, E., A. Höck, and R. M. Dreizler, 2000b, Phys. Rev. A **62**, 042502.

Engel, E., A. Höck, R. N. Schmid, R. M. Dreizler, and N. Chetty, 2001, Phys. Rev. B **64**, 125111.

Engel, E., and H. Jiang, 2006, Int. J. Quantum Chem. **106**, 3242.

Engel, E., H. Jiang, and A. Facco Bonetti, 2005, Phys. Rev. A **72**, 052503.

Engel, E., and S. H. Vosko, 1993, Phys. Rev. A **47**, 2800.

Engel, E., and S. H. Vosko, 1994, Phys. Rev. B **50**, 10 498.

Engel, G. E., 1997, Phys. Rev. Lett. **78**, 3515.

Ernzerhof, M., 1996, Chem. Phys. Lett. **263**, 499.

Ernzerhof, M., J. P. Perdew, and K. Burke, 1997, Int. J. Quantum Chem. **64**, 285.

Ernzerhof, M., and G. E. Scuseria, 1999, J. Chem. Phys. **110**, 5029.

Eyert, V., K.-H. Höck, S. Fiechter, and H. Tributsch, 1998, Phys. Rev. B **57**, 6350.

Facco Bonetti, A., E. Engel, R. N. Schmid, and R. M. Dreizler, 2001, Phys. Rev. Lett. **86**, 2241.

Facco Bonetti, A., E. Engel, R. N. Schmid, and R. M. Dreizler, 2003, Phys. Rev. Lett. **90**, 219302.

Fermi, E., 1927, Rend. Accad. Naz. Lincei **6**, 602.

Fermi, E., 1934, Nuovo Cimento **11**, 157.

Fermi, E., and E. Amaldi, 1934, Atti R. Accad. Naz. Lincei, Mem. Cl. Sci. Fis., Mat. Nat. **6**, 119.

Fernández, J. J., C. Tablero, and P. Wahnón, 2003a, Int. J. Quantum Chem. **91**, 157.

Fernández, J. J., C. Tablero, and P. Wahnón, 2003b, Comput. Mater. Sci. **28**, 274.

Filatov, M., and W. Thiel, 1998a, Phys. Rev. A **57**, 189.

Filatov, M., and W. Thiel, 1998b, Chem. Phys. Lett. **295**, 467.

Filippetti, A., and N. A. Spaldin, 2003, Phys. Rev. B **67**, 125109.

Filippetti, A., N. A. Spaldin, and S. Sanvito, 2004, Chem. Phys. Lett. **309**, 59.

Filippi, C., Umrigar, C. J., and Gonze, X., 1997, J. Chem. Phys. **107**, 9994.

Fleszar, A. 2001, Phys. Rev. B **64**, 245204.

Fois, E. S., J. I. Penman, and P. A. Madden, 1993, J. Chem. Phys. **98**, 6352.

Foster, J. M., and S. F. Boys, 1960, Rev. Mod. Phys. **32**, 300.

Fritsche, L., and J. Yuan, 1998, Phys. Rev. A **57**, 3425.

Fuchs, M., M. Bockstedte, E. Pehlke, and M. Scheffler, 1998, Phys. Rev. B **57**, 2134.

Fuchs, M., K. Burke, Y. M. Niquet, and X. Gonze, 2003, Phys. Rev. Lett. **90**, 189701.

Fuchs, M., J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, 2007, e-print arXiv:cond-mat/0604447.

Fuchs, M., and X. Gonze, 2002, Phys. Rev. B **65**, 235109.

Furche, F., 2001, Phys. Rev. B **64**, 195120.

Gál, T., 2000, Phys. Rev. A **62**, 044501.

Garza, J., J. A. Nichols, and D. A. Dixon, 2000, J. Chem. Phys. **112**, 7880.

Gavartin, J. L., P. V. Sushko, and A. L. Shluger, 2003, Phys. Rev. B **67**, 035108.

Geerlings, P., F. de Proft, and W. Langenaeker, 2003, Chem. Rev. (Washington, D.C.) **103**, 1793.

Geller, C. B., S. Picozzi, A. Continenza, R. Asahi, W. Mannstadt, A. J. Freeman, and E. Wimmer, 2001, Appl. Phys. Lett. **79**, 368.

Gerber, I. C., and J. G. Ángán, 2005, Chem. Phys. Lett. **415**, 100.

Gilbert, T. L., 1975, Phys. Rev. B **12**, 2111.

Godby, R. W., M. Schlüter, and L. J. Sham, 1986, Phys. Rev. Lett. **56**, 2415.

Goedecker, S., and K. Maschke, 1992, Phys. Rev. A **45**, 88.

Goedecker, S., and C. J. Umrigar, 1997, Phys. Rev. A **55**, 1765.

Goll, E., H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, and Andreas Savin, 2006, Chem. Phys. **239**, 276.

Goñi, A. R., U. Haboeck, C. Thomsen, K. Eberl, F. A. Reboredo, C. R. Proetto, and F. Guinea, 2002, Phys. Rev. B **65**, 121313(R).

Gori-Giorgi, P., and A. Savin, 2006, Phys. Rev. A **73**, 032506.

Görling, A., 1992, Phys. Rev. A **46**, 3753.

Görling, A., 1996, Phys. Rev. A **54**, 3912.

Görling, A., 1997, Phys. Rev. A **55**, 2630.

Görling, A., 1998, Phys. Rev. A **57**, 3433.

Görling, A., 1999, Phys. Rev. Lett. **83**, 5459.

Görling, A., 2005, J. Chem. Phys. **123**, 062203.

Görling, A., 2006, private communication at the workshop "Time dependent Density-Functional Theory: Prospects and Applications," Benasque, Spain.

Görling, A., and M. Levy, 1993, Phys. Rev. B **47**, 13 105.

Görling, A., and M. Levy, 1994, Phys. Rev. A **50**, 196.

Görling, A., and M. Levy, 1995a, Int. J. Quantum Chem. **29**, 93.

Görling, A., and M. Levy, 1995b, Phys. Rev. A **52**, 4493.

Grabo, T., and E. K. U. Gross, 1995, Chem. Phys. Lett. **240**, 141.

Grabo, T., and E. K. U. Gross, 1997, Int. J. Quantum Chem. **64**, 95.

Grabo, T., T. Kreibich, and E. K. U. Gross, 1997, Mol. Eng. **7**, 27.

Grabo, T., T. Kreibich, S. Kurth, and E. K. U. Gross, 2000, in *Strong Coulomb Correlation in Electronic Structure: Beyond the Local Density Approximation*, edited by V. Anisimov (Gordon & Breach, Tokyo).

Grabowksi, I., S. Ivanov, S. Hirata, and R. J. Bartlett, 2002, J. Chem. Phys. **116**, 4415.

Gräfenstein, J., E. Kraka, and D. Cremer, 2004, Phys. Chem. Chem. Phys. **6**, 1096.

Grimme, S., 2006, J. Chem. Phys. **124**, 034108.

Gritsenko, O. V., and E. J. Baerends, 2001, Phys. Rev. A **64**, 042506.

Gritsenko, O. V., and E. J. Baerends, 2002, J. Chem. Phys. **117**, 9154.

Gritsenko, O. V., N. A. Cordero, A. Rubio, L. C. Balbás, and J. A. Alonso, 1993, Phys. Rev. A **48**, 4197.

Gritsenko, O. V., P. R. T. Schipper, and E. J. Baerends, 1997, J. Chem. Phys. **107**, 5007.

Gritsenko, O. V., S. J. A. van Gisbergen, A. Görling, and E. J. Baerends, 2000, J. Chem. Phys. **113**, 8478.

Gritsenko, O. V., R. van Leeuwen, and E. J. Baerends, 1994, J. Chem. Phys. **101**, 8955.

Gross, E. K. U., J. F. Dobson, and M. Petersilka, 1996, in *Density Functional Theory*, edited by R. F. Nalewajski, Topics in Current Chemistry Vol. 181 (Springer Berlin).

Gross, E. K. U., T. Kreibich, M. Lein, and M. Petersilka, 1999, in *Electron Correlations and Materials Properties*, edited by A. Gonis, N. Kioussis, and M. Ciftan (Springer, Berlin).

Gross, E. K. U., M. Petersilka, and T. Grabo, 1996, in *Chemical Applications of Density Functional Theory*, edited by B. B. Laird and T. Ziegler, ACS Symposium Series 629 (American Chemical Society, Washington, D.C.).

Grüning, M., O. V. Gritsenko, and E. J. Baerends, 2002, J. Chem. Phys. **116**, 6435.

Grüning, M., A. Marini, and A. Rubio, 2006a, J. Chem. Phys. **124**, 154108.

Grüning, M., A. Marini, and A. Rubio, 2006b, Phys. Rev. B **74**, 161103.

Gunnarsson, O., and R. O. Jones, 1985, Phys. Rev. B **31**, 7588.

Gunnarsson, O., M. Jonson, and B. I. Lundqvist, 1979, Phys. Rev. B **20**, 3136.

Gunnarsson, O., and B. I. Lundqvist, 1976, Phys. Rev. B **13**, 4274; **15**, 6006(E) (1977).

Gutlé, C., J. L. Heully, J. B. Krieger, and A. Savin, 2002, Phys. Rev. A **66**, 012504.

Hamel, S., M. E. Casida, and D. R. Salahub, 2002, J. Chem. Phys. **116**, 8276.

Hamman, D. R., M. Schlüter, and C. Chiang, 1979, Phys. Rev. Lett. **43**, 1494.

Hamprecht, F. A., A. J. Cohen, D. J. Tozer, and N. C. Handy, 1998, J. Chem. Phys. **109**, 6264.

Han, M., Y. Özyilmaz, Y. Zhang, and P. Kim, 2007, Phys. Rev. Lett. **98**, 206805.

Handy, N. C., and A. J. Cohen, 2001, Mol. Phys. **99**, 403.

Harris, J., 1984, Phys. Rev. A **29**, 1648.

Harris, J., and R. O. Jones, 1974, J. Phys. F: Met. Phys. **4**, 1170.

Harrison, J. G., Heaton, R. A., and Lin, C. C., 1983, J. Phys. B **16**, 2079.

Hay, P. J., R. L. Martin, J. Uddin, and G. E. Scuseria, 2006, J. Chem. Phys. **125**, 034712 (2006).

Heaton, R. A., J. G. Harrison, and C. C. Lin, 1982, Solid State Commun. **41**, 827.

Heaton, R. A., J. G. Harrison, and C. C. Lin, 1983, Phys. Rev. B **28**, 5992.

Hedin, L., 1965, Phys. Rev. **139**, A796.

Heinemann, D., A. Ros'en, and B. Fricke, 1990, Phys. Scr. **42**, 692.

Held, K., I. A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A. K. McMahan, R. T. Scalettar, T. Pruschke, V. I. Anisimov, and D. Vollhardt, 2003, Ψ-k Newsletter1927, **56**, 65.

Hemelsoet, K., D. Moran, V. Van Speybroeck, M. Waroquier, and L. Radom, 2006, J. Phys. Chem. A **110**, 8942.

Hessler, P., N. T. Maitra, and K. Burke, 2002, J. Chem. Phys. **117**, 72.

Heyd, J., J. E. Peralta, G. E. Scuseria, and R. L. Martin, 2005, J. Chem. Phys. **123**, 174101.

Heyd, J., and G. E. Scuseria, 2004a, J. Chem. Phys. **120**, 7274.

Heyd, J., and G. E. Scuseria, 2004b, J. Chem. Phys. **121**, 1187.

Heyd, J., G. E. Scuseria, and M. Ernzerhof, 2003, J. Chem. Phys. **118**, 8207; J. Chem. Phys. **124**, 219906(E) (2006).

Hieringer, W., F. Della Sala, and A. Görling, 2004, Chem. Phys. Lett. **383**, 115.

Hirata, S., S. Ivanov, R. Bartlett, and I. Grabowski, 2005, Phys. Rev. A **71**, 032507.

Hirata, S., S. Ivanov, I. Grabowski, R. Bartlett, K. Burke, and J. Talman, 2001, J. Chem. Phys. **115**, 1635.

Höck, A., and E. Engel, 1998, Phys. Rev. A **58**, 3578.

Hohenberg, P., and W. Kohn, 1964, Phys. Rev. **136**, B864.

Holas, A., and N. H. March, 2001, Phys. Rev. A **64**, 016501.

Holas, A., and N. H. March, 2002, Phys. Rev. A **66**, 066501.

Holthausen, M. C., 2005, J. Comput. Chem. **26**, 1505.

Hood, R. Q., M. Y. Chou, A. J. Williamson, G. Rajagopal, and R. J. Needs, 1998, Phys. Rev. B **57**, 8972.

Horowitz, C. M., C. R. Proetto, and S. Rigamonti, 2006, Phys. Rev. Lett. **97**, 026802.

Hupp, T., E. Engels, and A. Görling, 2003, J. Chem. Phys. **119**, 11 591.

Hybertsen, M. S., and S. G. Louie, 1986, Phys. Rev. B **34**, 5390.

Hyman, R. A., M. D. Stiles, and A. Zangwill, 2000, Phys. Rev. B **62**, 15 521.

Iikura, H., T. Tsuneda, and K. Hirao, 2001, J. Chem. Phys. **115**, 3540.

Ivanov, S., and R. J. Bartlett, 2001, J. Chem. Phys. **114**, 1952.

Ivanov, S., S. Hirata, and R. J. Bartlett, 1999, Phys. Rev. Lett. **83**, 5455.

Ivanov, S., S. Hirata, I. Grabowski, and R. J. Bartlett, 2003, J. Chem. Phys. **118**, 461.

Ivanov S., and M. Levy, 2002, J. Chem. Phys. **116**, 6924.

Ivanov S., and M. Levy, 2003, J. Chem. Phys. **119**, 7087.

Izmaylov, A. F., V. N. Staroverov, G. E. Scuseria, E. R. Davidson, G. Stoltz, and E. Cancès, 2007, J. Chem. Phys. **126**, 084107.

Jacquemin, D., A. Femenias, H. Chermette, I. Ciofini, C. Adamo, J.-M., André, and E. A. Perpète, 2006, J. Phys. Chem. A **110**, 5952.

Janak, J. F., 1978, Phys. Rev. B **18**, 7165.

Jaramillo, J., G. E. Scuseria, and M. Ernzerhof, 2003, J. Chem. Phys. **118**, 1068.

Jiang, H., and E. Engel, 2006, J. Chem. Phys. **125**, 184108.

Johnson, B. G., C. A. Gonzales, P. M. W. Gill, and J. A. Pople, 1994, Chem. Phys. Lett. **221**, 100.

Jones, R. O., and O. Gunnarsson, 1989, Rev. Mod. Phys. **61**, 689.

Kamiya, M., T. Tsuneda, and K. Hirao, 2002, J. Chem. Phys. **117**, 6010.

Kanis, D. R., M. A. Ratner, and T. J. Marks, 1994, Chem. Rev. (Washington, D.C.) **94**, 195.

Keal, T. W., and D. J. Tozer, 2005, J. Phys. Chem. **123**, 121103.

Kim, Y.-H., and A. Görling, 2002a, Phys. Rev. Lett. **89**, 096402.

Kim, Y.-H., and A. Görling, 2002b, Phys. Rev. B **66**, 035114.

Kim, Y.-H., M. Städele, and R. M. Martin, 1999, Phys. Rev. A **60**, 3633.

Kleinman, L., 1980, Phys. Rev. B **21**, 2630.

Koch, W., and M. C. Holthausen, 2001, *A Chemist's Guide to Density Functional Theory*, 2nd ed. (Wiley-VCH, Heidelberg).

Koentopp, M., K. Burke, and F. Evers, 2006, Phys. Rev. B **73**, 121403(R).

Kohn, W., 1999, Rev. Mod. Phys. **71**, 1253.

Kohn, W., A. D. Becke, and R. G. Parr, 1996, J. Phys. Chem. **100**, 12 974.

Kohn, W., and L. Sham, 1965, Phys. Rev. **140**, A1133.

Koopmans, T., 1934, Physica (Amsterdam) **1**, 104.

Kosaka, K., 2006, J. Phys. Soc. Jpn. **75**, 014302.

Kotani, T., 1994, Phys. Rev. B **50**, 14 816.

Kotani, T., 1995, Phys. Rev. Lett. **74**, 2989.

Kotani, T., 1998, J. Phys.: Condens. Matter **10**, 9241.

Kotani, T., and H. Akai, 1996, Phys. Rev. B **54**, 16502.

Körzdörfer, T., and S. Kümmel, 2006, unpublished; see also Körzdörfer, T., 2006, Diploma thesis (University of Bayreuth).

Kotliar, G., S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, 2006, Rev. Mod. Phys. **78**, 865.

Kreibich, T., S. Kurth, T. Grabo, and E. K. U. Gross, 1999, Adv. Quantum Chem. **33**, 31.

Krieger, J. B., Y. Li, and G. J. Iafrate, 1990, Phys. Lett. A **146**, 256.

Krieger, J. B., Y. Li, and G. J. Iafrate, 1992a, Phys. Rev. A **45**, 101.

Krieger, J. B., Y. Li, and G. J. Iafrate, 1992b, Phys. Rev. A **46**, 5453.

Kronik, L., R. Fromherz, E. Ko, G. Ganteför, and J. R. Chelikowsky, 2002, Nat. Mater. **1**, 49.

Kronik, L., A. Makmal, M. L. Tiago, M. M. G. Alemany, M. Jain, X. Huang, Y. Saad, and J. R. Chelikowsky, 2006, Phys. Status Solidi B **243**, 1063.

Kronik, L., I. Vasiliev, M. Jain, and J. R. Chelikowsky, 2001, J. Chem. Phys. **115**, 4322; **115**, 8714(E) (2001).

Krukau, A. V., O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, 2006, J. Chem. Phys. **125**, 224106.

Kudin, K. N., G. E. Scuseria, and R. L. Martin, 2002, Phys. Rev. Lett. **89**, 266402.

Kümmel, S., 2004, J. Comput. Phys. **201**, 333.

Kümmel, S., and L. Kronik, 2006, Comput. Mater. Sci. **35**, 321.

Kümmel, S., L. Kronik, and J. P. Perdew, 2004, Phys. Rev. Lett. **93**, 213002.

Kümmel, S., and J. P. Perdew, 2003a, Phys. Rev. Lett. **90**, 043004.

Kümmel, S., and J. P. Perdew, 2003b, Phys. Rev. B **68**, 035103.

Kümmel, S., and J. P. Perdew, 2003c, Mol. Phys. **101**, 1363.

Kurth, S., and J. P. Perdew, 1999, Phys. Rev. B **59**, 10461.

Kurth, S., J. P. Perdew, and P. Blaha, 1999, Int. J. Quantum Chem. **75**, 889.

Kurzweil, Y., and R. Baer, 2006, Phys. Rev. B **73**, 075413.

Lægsgaard, J., and K. Stokbro, 2001, Phys. Rev. Lett. **86**, 2834.

Langreth, D. C., and J. P. Perdew, 1975, Solid State Commun. **17**, 1425.

Langreth, D. C., and J. P. Perdew, 1977, Phys. Rev. B **15**, 2884.

Langreth, D. C., and J. P. Perdew, 1980, Phys. Rev. B **21**, 5469.

Lappas, D. G., and R. van Leeuwen, 1998, J. Phys. B **31**, L249.

Lee, B., and L.-W. Wang, 2006, Phys. Rev. B **73**, 153309.

Lee, C., W. Yang, and R. G. Parr, 1988, Phys. Rev. B **37**, 785.

Legrand, C., E. Suraud, and P.-G. Reinhard, 2002, J. Phys. B **35**, 1115.

Lein, M., E. K. U. Gross, and J. P. Perdew, 2000, Phys. Rev. B **61**, 13431.

Lein, M., and S. Kümmel, 2005, Phys. Rev. Lett. **94**, 143003.

Leininger, T., H. Stoll, H.-J. Werner, and A. Savin, 1997, Chem. Phys. Lett. **275**, 151.

Lento, J., and Nieminen, R. M., 2003, J. Phys.: Condens. Matter **15**, 4387.

Levy, M., 1979, Proc. Natl. Acad. Sci. U.S.A. **76**, 6062.

Levy, M., 1982, Phys. Rev. A **26**, 1200.

Levy, M., 1983, in *Density Functional Theory*, edited by J. Keller, and J. L. Gasquez (Springer, New York).

Levy, M., 1991, Phys. Rev. A **43**, 4637.

Levy, M., N. H. March, and N. C. Handy, 1996, J. Chem. Phys. **104**, 1989.

Levy, M., and J. P. Perdew, 1985, Phys. Rev. A **32**, 2010.

Levy, M., J. P. Perdew, and V. Sahni, 1984, Phys. Rev. A **30**, 2745.

Li, Y., J. B. Krieger, and G. J. Iafrate, 1993, Phys. Rev. A **47**, 165.

Li, Y., J. B. Krieger, N. R. Norman, and G. J. Iafrate, 1991, Phys. Rev. B **44**, 10437.

Lieb, E., 1982, in *Physics as Natural Philosophy: Essays in*

Honor of Laszlo Tisza on His 75th Birthday, edited by A. Shimony, and H. Feshbach (MIT, Cambridge, MA), p. 111.

Lindgren, I., and Salomonson, S., 2003, Phys. Rev. A **67**, 056501.

Lippert, R. A., N. A. Modine, and A. F. Wright, 2006, J. Phys.: Condens. Matter **18**, 4295.

Litimein, F., B. Bouhafs, and P. Ruterana, 2006, Phys. Status Solidi A **203**, 35.

Liu, S. B., and P. W. Ayers, 2004, Phys. Rev. A **70**, 022501.

Livshits, E., and R. Baer, 2007, Phys. Chem. Chem. Phys. **9**, 2932.

Louie, S. G., S. Froyen, and M. L. Cohen, 1982, Phys. Rev. B **26**, 1738.

Lundberg, M., and P. E. M. Siegbahn, 2005, J. Chem. Phys. **122**, 224103.

Lundin, U., and O. Eriksson, 2001, Int. J. Quantum Chem. **81**, 247.

Lutnæs, O. B., A. M. Teale, T. Helgaker, and D. J. Tozer, 2006, J. Chem. Theory Comput. **2**, 827.

Lynch, B. J., P. L. Fast, M. Harris, and D. G. Truhlar, 2004, J. Phys. Chem. A **104**, 4811.

Ma, S.-K., and K. A. Brueckner, 1968, Phys. Rev. **165**, 18.

Magyar, R. J., A. Fleszar, and E. K. U. Gross, 2004, Phys. Rev. B **69**, 045111.

Maitra, N. T., 2005, J. Chem. Phys. **122**, 234104.

Maitra, N. T., and K. Burke, 2001, Phys. Rev. A **63**, 042501.

Maitra, N. T., K. Burke, H. Appel, E. K. U. Gross, and R. van Leeuwen, 2002, in *Reviews in Modern Quantum Chemistry: A Celebration of the Contributions of R. G. Parr*, edited by K. D. Sen (World Scientific, Amsterdam), pp. 1186–1225.

Maitra, N., K. Burke, and C. Woodward, 2002, Phys. Rev. Lett. **89**, 023002.

Maitra, N. T., F. Zhang, R. J. Cave, and K. Burke, 2004, J. Chem. Phys. **120**, 5932.

Marini, A., R. Del Sole, and A. Rubio, 2003, Phys. Rev. Lett. **91**, 256402.

Marques, M. A. L., A. Castro, G. F. Bertsch, and A. Rubio, 2003, Comput. Phys. Commun. **151**, 60.

Marques, M. A. L., A. Castro, and A. Rubio, 2001, J. Chem. Phys. **115**, 3006.

Marques, M. A. L., and E. K. U. Gross, 2004, Annu. Rev. Phys. Chem. **55**, 427.

Marques, M. A. L., C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross, 2006, Eds., *Time-Dependent Density Functional Theory* (Springer, Berlin).

Martin, R., 2004, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge).

Mattsson, A. E., and D. R. Jennison, 2002, Surf. Sci. **520** L611.

Mok, D. K. W., R. Neumann, and N. C. Handy, 1996, J. Phys. Chem. **100**, 6225.

Molawi, K., A. J. Cohen, and N. C. Handy, 2002, Int. J. Quantum Chem. **89**, 86.

Moreira, I. de P. R., F. Illas, and R. L. Martin, 2002, Phys. Rev. B **65**, 155102.

Mori-Sánchez, P., A. J. Cohen, and W. Yang, 2006, J. Chem. Phys. **124**, 091102.

Mori-Sánchez, P., Q. Wu, and W. Yang, 2003, J. Chem. Phys. **119**, 11 001.

Mori-Sánchez, P., Q. Wu, and W. Yang, 2005, J. Chem. Phys. **123**, 062204.

Moseler, M., B. Huber, H. Häkkinen, U. Landman, G. Wrigge, M. A. Hoffmann, and B. von Issendorff, 2003, Phys. Rev. B **68**, 165413.

Moukara, M., M. Städele, J. A. Majewski, P. Vogl, and A. Görling, 2000, J. Phys.: Condens. Matter **12**, 6783.

Mundt, M., and S. Kümmel, 2005, Phys. Rev. Lett. **95**, 203004.

Mundt, M., and S. Kümmel, 2006, Phys. Rev. A **74**, 022511.

Mundt, M., S. Kümmel, B. Huber, and M. Moseler, 2006, Phys. Rev. B **73**, 205407.

Mundt, M., S. Kümmel, R. van Leeuwen, and P.-G. Reinhard, 2007, Phys. Rev. A **75**, 050501(R).

Muscat, J., A. Wander, and N. M. Harrison, 2001, Chem. Phys. Lett. **342**, 397.

Nagy, À., 1997, Phys. Rev. A **55**, 3465.

Negele, J. W., and D. Vautherin, 1972, Phys. Rev. C **5**, 1472.

Nesbet, R. K., 1998, Phys. Rev. A **58**, R12.

Nesbet, R. K., 2001, Phys. Rev. A **65**, 010502(R).

Nesbet, R. K., 2004, Mod. Phys. Lett. B **18**, 73.

Neugebauer, J., O. V. Gritsenko, and E. J. Baerends, 2006, J. Chem. Phys. **124**, 214102.

Neumann, R., R. H. Nobes, and N. C. Handy, 1996, Mol. Phys. **87**, 1.

Niquet, Y. M., M. Fuchs, and X. Gonze, 2003a, Phys. Rev. Lett. **90**, 219301.

Niquet, Y. M., M. Fuchs, and X. Gonze, 2003b, J. Chem. Phys. **118**, 9504.

Niquet, Y. M., M. Fuchs, and X. Gonze, 2005, Int. J. Quantum Chem. **101**, 635.

Niquet, Y. M., and X. Gonze, 2004, Phys. Rev. B **70**, 245115.

Novák, P., J. Kuneš, W. E. Pickett, W. Ku, and F. R. Wagner, 2003, Phys. Rev. B **67**, 140403.

Nozières, P., and D. Pines, 1958, Phys. Rev. **111**, 442.

Onida, G., L. Reining, and A. Rubio, 2002, Rev. Mod. Phys. **74**, 601.

Orlova, G., and J. D. Goddard, 2002, Mol. Phys. **100**, 483.

Pacchioni, G., F. Frigoli, D. Ricci, and J. A. Weil, 2000, Phys. Rev. B **63**, 054102.

Paier, J., M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, 2006, J. Chem. Phys. **124**, 154709; **125**, 249901(E) (2006).

Parr, R. G., and W. Yang, 1989, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford).

Patchkovskii, S., J. Autschbach, and T. Ziegler, 2001, J. Chem. Phys. **115**, 26.

Patchkovskii, S., and T. Ziegler, 2002, J. Chem. Phys. **116**, 7806.

Payami, M., 2006, Phys. Rev. B **73**, 113106.

Peach, M. J. G., A. J. Cohen, and D. J. Tozer, 2006, Phys. Chem. Chem. Phys. **8**, 4543.

Peach, M. J. G., T. Helgaker, P. Sałek, T. W. Keal, O. B. Lutnæs, D. J. Tozer, and N. C. Handy, 2006, Phys. Chem. Chem. Phys. **8**, 558.

Pederson, M. R., R. A. Heaton, and C. C. Lin, 1984, J. Chem. Phys. **80**, 1972.

Pederson, M. R., R. A. Heaton, and C. C. Lin, 1985, J. Chem. Phys. **82**, 2688.

Peirs, K., D. Van Neck, and M. Waroquier, 2003, Phys. Rev. A **67**, 012505.

Peralta, J. E., E. R. Batista, G. E. Scuseria, and R. L. Martin, 2005, J. Chem. Theory Comput. **1**, 612.

Peralta, J. E., J. Heyd, G. E. Scuseria, and R. L. Martin, 2006, Phys. Rev. B **74**, 073101.

Perdew, J. P., 1979, Chem. Phys. Lett. **64**, 127.

Perdew, J. P., 1990, Adv. Quantum Chem. **21**, 113.

Perdew, J. P., 1991, in *Electronic Structure of Solids 91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin), p. 11.

Perdew, J. P., K. Burke, and M. Ernzerhof, 1996, Phys. Rev. Lett. **77**, 3865; **78**, 1396(E) (1997).

Perdew, J. P., J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, 1992, Phys. Rev. B **46**, 6671; **48**, 4978(E) (1993).

Perdew, J. P., and M. Ernzerhof, 1998, in *Driving out the Self-interaction Error*, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, New York).

Perdew, J. P., M. Ernzerhof, and K. Burke, 1996, J. Chem. Phys. **105**, 9982.

Perdew, J. P., and S. Kurth, 2003, in *A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. A. L. Marques (Springer, Berlin).

Perdew J. P., S. Kurth, A. Zupan, and P. Blaha, 1999, Phys. Rev. Lett. **82**, 2544.

Perdew, J. P., and M. Levy, 1983, Phys. Rev. Lett. **51**, 1884.

Perdew, J. P., and M. Levy, 1997, Phys. Rev. B **56**, 16021.

Perdew, J. P., R. G. Parr, M. Levy, and J. L. Balduz, Jr., 1982, Phys. Rev. Lett. **49**, 1691.

Perdew, J. P., A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, 2005, J. Chem. Phys. **123**, 062201.

Perdew, J. P., and K. Schmidt, 2001, in *Density Functional Theory and Its Application to Materials*, edited by V. Van Doren, C. Van Alsenoy, and P. Geerlings (AIP, New York), pp. 1–20.

Perdew, J. P., and Y. Wang, 1986, Phys. Rev. B **33**, 8800; **40**, 3399(E) (1989).

Perdew, J. P., and Y. Wang, 1992, Phys. Rev. B **45**, 13244.

Perdew, J. P., and A. Zunger, 1981, Phys. Rev. B **23**, 5048.

Petersilka, M., U. J. Gossmann, and E. K. U. Gross, 1996, Phys. Rev. Lett. **76**, 1212.

Petersilka, M., and E. K. U. Gross, 1999, Laser Phys. **9**, 105.

Peuckert, V., 1978, J. Phys. C **11**, 4945.

Phillips, J. C., and L. Kleinman, 1959, Phys. Rev. **116**, 287.

Pickett, W. E., 1989, Comput. Phys. Rep. **9**, 115.

Picozzi, S., A. Continenza, R. Asahi, W. Mannstadt, A. J. Freeman, W. Wolf, E. Wimmer, and C. B. Geller, 2000, Phys. Rev. B **61**, 4677.

Pitarke, J. M., and A. G. Eguiluz, 1998, Phys. Rev. B **57**, 6329.

Pitarke, J. M., and A. G. Eguiluz, 2001, Phys. Rev. B **63**, 045116.

Pitarke, J. M., and J. P. Perdew, 2003, Phys. Rev. B **67**, 045101.

Pittalis, S., S. Kurth, N. Helbig, and E. K. U. Gross, 2006, Phys. Rev. A **74**, 062511.

Pohl, A., P.-G. Reinhard, and E. Suraud, 2000, Phys. Rev. Lett. **84**, 5090.

Polo, V., E. Kraka, and D. Cremer, 2002a, Mol. Phys. **100**, 1771.

Polo, V., E. Kraka, and D. Cremer, 2002b, Theor. Chem. Acc. **107**, 291.

Pople, J. A., M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, 1989, J. Chem. Phys. **90**, 5622.

Prodan, I. D., G. E. Scuseria, and R. L. Martin, 2006, Phys. Rev. B **73**, 045104.

Proynov, E. I., and A. J. Thakkar, 2006, Int. J. Quantum Chem. **106**, 436.

Proynov, E. I., A. Vela, and D. R. Salahib, 1994, Chem. Phys. Lett. **230**, 419.

Qteish, A., A. I. Al-Sharif, M. Fuchs, M. Scheffler, S. Boeck, and J. Neugebauer, 2005, Phys. Rev. B **72**, 155317.

Qteish, A., P. Rinke, M. Scheffler, and J. Neugebauer, 2006, Phys. Rev. B **74**, 245208.

Rajagopal, A. K., and J. Callaway, 1973, Phys. Rev. B **7**, 1912.

Rappe, A. M., K. M. Rabe, E. Kaxiras, and J. J. Joannopoulos, 1990, Phys. Rev. B **41**, 1227.

Reboredo, F. A., and C. R. Proetto, 2003, Phys. Rev. B **67**, 115325.

Reinhard, P. G., M. Brack, F. Calvayrac, C. Kohl, S. Kümmel, E. Suraud, and C. A. Ullrich, 1999, Eur. Phys. J. D **9**, 111.

Reining, L., V. Olevano, A. Rubio, and G. Onida, 2002, Phys. Rev. Lett. **88**, 066404.

Rhim, S. H., M. Kim, A. J. Freeman, and R. Asahi, 2005, Phys. Rev. B **71**, 045202.

Rigamonti, S., C. R. Proetto, and F. A. Reboredo, 2005, Europhys. Lett. **70**, 116.

Rinke, P., 2007, private communication.

Rinke, P., A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, 2005, New J. Phys. **7**, 126.

Rinke, P., A. Qteish, J. Neugebauer, and M. Scheffler, 2007, in Psi-k Newsletter 79, http://psi-k.dl.ac.uk/newsletters/News_79/Highlight_79.pdf

Rinke, P., M. Scheffler, A. Qteish, M. Winkelnkemper, D. Bimberg, and J. Neugebauer, 2006, Appl. Phys. Lett. **89**, 161919.

Robertson, J., K. Xiong, and S. J. Clark, 2006, Phys. Status Solidi B **243**, 2054.

Robertson, J., K. Xiong, and B. Falabretti, 2005, IEEE Trans. Device Mater. Reliab. **5**, 84.

Rohlfing, M., and S. G. Louie, 1998, Phys. Rev. Lett. **80**, 3320 (1998).

Rohr, D. R., O. V. Gritsenko, and E. J. Baerends, 2006, J. Mol. Struct.: THEOCHEM **762**, 193.

Rohra, S., E. Engel, and A. Görling, 2006, Phys. Rev. B **74**, 045119.

Rohra, S., and A. Görling, 2006, Phys. Rev. Lett. **97**, 013005.

Rösch, N., and S. B. Trickey, 1997, J. Chem. Phys. **106**, 8940.

Rose, J. H., Jr., and H. B. Shore, 1975, Solid State Commun. **17**, 327.

Runge, E., and E. K. U. Gross, 1984, Phys. Rev. Lett. **52**, 997.

Ruzsinszky, A., J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, 2006, J. Chem. Phys. **125**, 194112.

Ruzsinszky, A., J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, 2007, J. Chem. Phys. **126**, 104102.

Saalmann, U., and R. Schmidt, 1996, Z. Phys. D: At., Mol. Clusters **38**, 153.

Sahni, V., Gruenebaum, J., and Perdew, J. P., 1982, Phys. Rev. B **26**, 4371.

Salzner, U., J. B. Lagowski, P. G. Pickup, and R. A. Poirier, 1997, J. Comput. Chem. **18**, 1943.

Sancho-García, J. C., J. L. Brédas, and J. Cornil, 2003, Chem. Phys. Lett. **377**, 63.

Sato, T., T. Tsuneda, and K. Hirao, 2005a, J. Chem. Phys. **123**, 104307.

Sato, T., T. Tsuneda, and K. Hirao, 2005b, Mol. Phys. **103**, 1151.

Savin, A., C. J. Umrigar, and X. Gonze, 1998, Chem. Phys. Lett. **288**, 391.

Schmider, H. L., and A. D. Becke, 1998, J. Chem. Phys. **109**, 8188.

Schulthess, T. C., W. M. Temmermann, Z. Szotek, W. H. Butler, and G. M. Stocks, 2005, Nat. Mater. **4**, 838.

Schwartz, K., 1978, Chem. Phys. Lett. **57**, 605.

Schweigert, I. V., V. F. Lotrich, and R. J. Bartlett, 2006, J. Chem. Phys. **125**, 104108.

Segev, L., A. Salomon, A. Natan, D. Cahen, L. Kronik, F. Amy, C. K. Chan, and A. Kahn, 2006, Phys. Rev. B **74**, 165323.

Seidl, A., A. Görling, P. Vogl, J. A. Majewski, and M. Levy, 1996, Phys. Rev. B **53**, 3764.

Seidl, M., J. P. Perdew, and S. Kurth, 2000, Phys. Rev. Lett. **84**, 5070.

Seidl, M., J. P. Perdew, and M. Levy, 1999, Phys. Rev. A **59**, 51.

Sekino, H., Y. Maeda, and M. Kamiya, 2005, Mol. Phys. **103**, 2183.

Shaginyan, V. R., 1993, Phys. Rev. A **47**, 1507.

Sham, L. J., and M. Schlüter, 1983, Phys. Rev. Lett. **51**, 1888.

Sharma, S., J. K. Dewhurst, and C. Ambrosch-Draxl, 2005, Phys. Rev. Lett. **95**, 136402.

Sharma, S., J. K. Dewhurst, C. Ambrosch-Draxl, S. Kurth, N. Helbig, S. Pittalis, S. Shallcross, L. Nordström, and E. K. U. Gross, 2007, Phys. Rev. Lett. **98**, 196405.

Sharp, R. T., and G. K. Horton, 1953, Phys. Rev. **90**, 317.

Shigeta, Y., K. Hirao, and S. Hirata, 2006, Phys. Rev. A **73**, 010502(R).

Shore, H. B., J. H. Rose, and E. Zaremba, 1977, Phys. Rev. B **15**, 2858.

Sinanoğlu, O., 1964, Adv. Chem. Phys. **6**, 358.

Singh, D. J., and L. Nordstrom, 2005, *Planewaves, Pseudopotentials, and the LAPW Method*, 2nd ed. (Springer-Verlag, Berlin).

Singh, R., L. Massa, and V. Sahni, 1999, Phys. Rev. A **60**, 4135.

Slater, J. C., 1951, Phys. Rev. **81**, 385.

Spataru, C. D., S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, 2004, Phys. Rev. Lett. **92**, 077402.

Städele, M., J. A. Majewski, P. Vogl, and A. Görling, 1997, Phys. Rev. Lett. **79**, 2089.

Städele, M., and R. M. Martin, 2000, Phys. Rev. Lett. **84**, 6070.

Städele, M., M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, 1999, Phys. Rev. B **59**, 10031.

Stampfl, C., W. Mannstadt, R. Asahi, and A. J. Freeman, 2001, Phys. Rev. B **63**, 155106.

Stampfl, C., and C. G. van der Walle, 1999, Phys. Rev. B **59**, 5521.

Stampfl, C., C. G. Van de Walle, D. Vogel, P. Krüger, and J. Pollmann, 2000, Phys. Rev. B **61**, R7846.

Staroverov, V. N., G. E. Scuseria, and E. R. Davidson, 2006a, J. Chem. Phys. **124**, 141103.

Staroverov, V. N., G. E. Scuseria, and E. R. Davidson, 2006b, J. Chem. Phys. **125**, 081104.

Staroverov, V. N., G. E. Scuseria, J. Tao, and J. P. Perdew, 2003, J. Chem. Phys. **119**, 12 129.

Stathopoulos, A., S. Öğüt, Y. Saad, J. R. Chelikowsky, and H. Kim, 2000, Comput. Sci. Eng. **2**, 19.

Stephens, P. J., F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, 1994, J. Phys. Chem. **98**, 11 623.

Stowasser, R., and R. Hoffmann, 1999, J. Am. Chem. Soc. **121**, 3414.

Strange, P., A. Svane, W. M. Temmerman, Z. Szotek, and H. Winter, 1999, Nature **399**, 756.

Süle, P., S. Kurth, and V. E. Van Doren, 1999, Phys. Rev. B **60**, 5429.

Süle, P., S. Kurth, and V. E. Van Doren, 2000, J. Chem. Phys. **112**, 7355.

Svane, A., and O. Gunnarsson, 1990, Phys. Rev. Lett. **65**, 1148.

Talman, J. D., and W. F. Shadwick, 1976, Phys. Rev. A **14**, 36.

Tao, J., P. Gori-Giorgi, J. P. Perdew, and R. McWeeny, 2001, Phys. Rev. A **63**, 032513.

Tao, J., J. P. Perdew, A. Ruzsinszky, G. E. Scuseria, G. I. Csonka, and V. N. Staroverov, 2007, Philos. Mag. **87**, 1071.

Tao, J., J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, 2003, Phys. Rev. Lett. **91**, 146401.

Tatarczyk, K., A. Schindlmayr, and M. Scheffler, 2001, Phys. Rev. B **63**, 235106.

Tawada, Y., T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, 2004, J. Chem. Phys. **120**, 8425.

Teale, A. M., A. J. Cohen, and D. J. Tozer, 2007, J. Chem. Phys. **126**, 074101.

Teale, A. M., and D. J. Tozer, 2005, Phys. Chem. Chem. Phys. **7**, 2991.

Temmerman, W. M., H. Winter, Z. Szotek, and A. Svane, 2001, Phys. Rev. Lett. **86**, 2435.

Thomas, L. H., 1927, Proc. Cambridge Philos. Soc. **23**, 542.

Toher, C., A. Filippetti, S. Sanvito, and K. Burke, 2005, Phys. Rev. Lett. **95**, 146402.

Tokatly, I. V., and O. Pankratov, 2001, Phys. Rev. Lett. **86**, 2078.

Tong, X.-M., and S.-I. Chu, 1997, Phys. Rev. A **55**, 3406.

Tong, X.-M., and S.-I. Chu, 1998, Phys. Rev. A **57**, 452.

Tong, X.-M., and S.-I. Chu, 2001, Phys. Rev. A **64**, 013417.

Toulouse, J., F. Colonna, and A. Savin, 2004, Phys. Rev. A **70**, 062505.

Toulouse, J., F. Colonna, and A. Savin, 2005, J. Chem. Phys. **122**, 014110.

Toulouse, J., P. Gori-Giorgi, and A. Savin, 2006, Int. J. Quantum Chem. **106**, 2026.

Toulouse, J., and A. Savin, 2006, J. Mol. Struct.: THEOCHEM **762**, 147.

Tozer, D. J., 1998, Phys. Rev. A **58**, 3524.

Tozer, D. J., 2003, J. Chem. Phys. **119**, 12697.

Tozer, D. J., and N. C. Handy, 1998, J. Chem. Phys. **108**, 2545.

Tozer, D. J., and N. C. Handy, 2000, Phys. Chem. Chem. Phys. **2**, 2117.

Troullier, N., and J. L. Martins, 1991, Phys. Rev. B **43**, 1993.

Uddin, J., and G. E. Scuseria, 2006, Phys. Rev. B **74**, 245115.

Ullrich, C. A., S. Erhard, and E. K. U. Gross, 1996, in *Super Intense Laser Atom Physics IV*, edited by H. G. Muller and M. V. Fedorov, Nato ASI Series 3/13 (Kluwer, Dordrecht), p. 267.

Ullrich, C. A., U. J. Gossmann, and E. K. U. Gross, 1995, Phys. Rev. Lett. **74**, 872.

Ullrich C. A., P.-G. Reinhard, and E. Suraud, 1998, J. Phys. B **31**, 1871.

Ullrich, C. A., P.-G. Reinhard, and E. Suraud, 2000, Phys. Rev. A **62**, 053202.

Umrigar, C. J., and X. Gonze, 1994, Phys. Rev. A **50**, 3827.

Unger, H.-J., 2001, Phys. Lett. A **284**, 124.

Vanderbilt, D., 1990, Phys. Rev. B **41**, 7892.

van der Hart, H. W., B. J. S. Doherty, J. S. Parker, and K. T. Taylor, 2005, J. Phys. B **38**, L207.

van Gisbergen, S. J. A., F. Kootstra, P. R. T. Schipper, O. V. Gritsenko, J. G. Snijders, and E. J. Baerends, 1998, Phys. Rev. A **57**, 2556.

van Gisbergen, S. J. A., P. R. T. Schipper, O. V. Gritsenko, E. J. Baerends, J. G. Snijders, B. Champagne, and B. Kirtman, 1999, Phys. Rev. Lett. **83**, 694.

van Leeuwen, R., 1996, Phys. Rev. Lett. **76**, 3610.

van Leeuwen, R., 1998, Phys. Rev. Lett. **80**, 1280.

van Leeuwen, R., 1999, Phys. Rev. Lett. **82**, 3863.

van Leeuwen, R., 2001, Int. J. Mod. Phys. B **15**, 1969.

van Leeuwen, R., and E. J. Baerends, 1994, Phys. Rev. A **49**, 2421.

van Leeuwen, R., O. V. Gritsenko, and E. J. Baerends, 1995, Z. Phys. D: At., Mol. Clusters **33**, 229.

van Schilfgaarde, M., A. Sher, and A.-B. Chen, 1997, J. Cryst. Growth **178**, 8.

van Voorhis, T., and G. E. Scuseria, 1998, J. Chem. Phys. **109**, 400.

Vogel, D., P. Krüger, and J. Pollmann, 1996, Phys. Rev. B **54**, 5495.

von Barth, U., and L. Hedin, 1972, J. Phys. C **5**, 1629.

Vosko, S. J., L. Wilk, and M. Nusair, 1980, Can. J. Phys. **58**, 1200s.

Vukajlovic, F. R., E. L. Shirley, and R. M. Martin, 1991, Phys. Rev. B **43**, 3994.

Vydrov, O. A., J. Heyd, A. V. Krukau, and G. E. Scuseria, 2006, J. Chem. Phys. **125**, 074106.

Vydrov, O. A., and G. E. Scuseria, 2004, J. Chem. Phys. **121**, 8187.

Vydrov, O. A., and G. E. Scuseria, 2006, J. Chem. Phys. **125**, 234109.

Vydrov, O. A., G. E. Scuseria, and J. P. Perdew, 2007, J. Chem. Phys. **126**, 154109.

Vydrov, O. A., G. E. Scuseria, J. P. Perdew, A. Ruzsinszky, and G. I. Csonka, 2006, J. Chem. Phys. **124**, 094108.

Wang, Y., and R. G. Parr, 1993, Phys. Rev. A **47**, R1591.

Whitehead, M. A., 1997, in *Recent Advances in Density Functional Methods, Part II*, edited by D. P. Chong (World Scientific, Singapore), and references therein.

Wilson, P. J., T. J. Bradley, and D. J. Tozer, 2001, J. Chem. Phys. **115**, 9233.

Wilson, P. J., and D. J. Tozer, 2001, Chem. Phys. Lett. **337**, 341.

Xiong, K., J. Robertson, M. C. Gibson, and S. J. Clark, 2005, Appl. Phys. Lett. **87**, 183505.

Xu, X., and W. A. Goddard III, 2004, Proc. Natl. Acad. Sci. U.S.A. **101**, 2673.

Yabana, K., and G. F. Bertsch, Phys. Rev. B **54**, 4484 (1996).

Yamagami, H., Y. Takada, H. Yasuhara, and A. Hasegawa, 1994, Phys. Rev. A **49**, 2354.

Yan, Z., J. P. Perdew, and S. Kurth, 2000, Phys. Rev. B **61**, 16430.

Yan, Z., J. P. Perdew, S. Kurth, C. Fiolhais, and L. Almeida, 2000, Phys. Rev. B **61**, 2595; **64**, 049904(E) (2001).

Yanai, T., D. P. Tew, and N. C. Handy, 2004, Chem. Phys. Lett. **393**, 51.

Yang, W., and Q. Wu, 2002, Phys. Rev. Lett. **89**, 143002.

Yasuhara, H., and Y. Takada, 1991, Phys. Rev. B **43**, 7200.

Yin, M. T., and M. L. Cohen, 1980, Phys. Rev. Lett. **45**, 1004.

Yu, D., and M. Scheffler, 2004, Phys. Rev. B **70**, 155417.

Zangwill, A., and P. Soven, 1980, Phys. Rev. A **21**, 1561.

Zhang, Y., and W. Yang, 1998, J. Chem. Phys. **109**, 2604.

Zhao, Q., R. C. Morrison, and R. G. Parr, 1994, Phys. Rev. A **50**, 2138.

Zhao, Y., B. J. Lynch, and D. G. Truhlar, 2004, J. Phys. Chem. A **108**, 2715.

Zhao, Y., N. E. Schultz, and D. G. Truhlar, 2005, J. Chem. Phys. **123**, 161103.

Zhao, Y., N. E. Schultz, and D. G. Truhlar, 2006, J. Chem. Theory Comput. **2**, 364.

Zhao, Y., and D. G. Truhlar, 2004, J. Phys. Chem. A **108**, 6908.

Zhao, Y., and D. G. Truhlar, 2005a, J. Phys. Chem. A **109**, 5656.

Zhao, Y., and D. G. Truhlar, 2005b, J. Phys. Chem. A **109**, 6624.

Zhao, Y., and D. G. Truhlar, 2006, J. Chem. Phys. **124**, 224105.

Zhu, J., X. W. Wang, and S. G. Louie, 1992, Phys. Rev. B **45**, 8887.