

# Part and whole in wavefunction/DFT embedding

Thomas Dresselhaus<sup>1</sup> · Johannes Neugebauer<sup>1</sup>

Received: 5 March 2015 / Accepted: 30 June 2015 / Published online: 24 July 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Density functional theory (DFT)-based subsystem and embedding methods have found a widespread use in Quantum Chemistry. The combination of correlated wavefunction (WF) methods and density-based embedding methods is particularly promising for accurate descriptions of complex systems. Here, we address some conceptual issues in such WF/DFT methods concerning (1) the interpretation of wavefunctions and densities in these frameworks, (2) the interpretation of subsystem and supersystem density changes as a polarization of the electronic system, and (3) non-orthogonality effects between wavefunctions of ground and excited states in WF/DFT embedding making use of state-specific embedding potentials. Illustrative examples are provided to analyze the significance of these issues in practical calculations. We find that physically reasonable subsystem densities and subsystem dipole moments can often be obtained from subsystem calculations making use of typical setups in practice, in spite of formal issues that are in principle in conflict with such interpretations. For excited-state WF/DFT calculations, we demonstrate that orthogonality violations due to state-specific embedding potentials are usually small.

**Keywords** Density-based embedding · Hybrid methods · Wavefunction/DFT embedding

## 1 Introduction

Hybrid methods play an increasingly important role for simulations of complex systems in Theoretical and Computational Chemistry and Physics [1–6]. Often, accurate quantum chemical methods are combined with less computationally demanding, empirical or semiempirical methods [3, 4]. A different approach is taken in density-based embedding schemes derived from frozen-density embedding (FDE) [7] and in the related partition DFT procedure [8, 9]. FDE theory can be derived rigorously in the context of density functional theory (DFT). It is based on subsystem density functional theory (sDFT) [10–12], where an FDE-like embedding potential appears in the effective Kohn–Sham equations for each of the subsystems, due to the presence of the other subsystems’ densities. Recent reviews are provided in Refs. [5, 13, 14]. Already in 1998, the density-embedding idea was transferred to the context of wavefunction(WF)/DFT embedding by Carter and coworkers [15] in a hybrid-method fashion. Later, Wesolowski [16] provided a more formal theoretical basis for WF/DFT embedding from a pure density functional theory viewpoint. The two strategies are very similar in the resulting working equations. But the different perspectives on WF/DFT embedding (either as a pragmatic hybrid method or as a rigorous DFT approach with non-Kohn–Sham reference systems) can lead to intriguing differences concerning formal aspects and the interpretation of the results. In this feature article, we want to compare the different viewpoints and to elaborate on some of the open conceptual issues. Since some of these issues also affect sDFT, we will not restrict our discussion to WF/DFT embedding, but also consider the DFT/DFT (sDFT) case at some points.

After introducing the theoretical background leading to the different points of view in Sect. 2 for sDFT and in Sect.

✉ Johannes Neugebauer  
j.neugebauer@uni-muenster.de

<sup>1</sup> Theoretische Organische Chemie, Organisch-Chemisches Institut und Center for Multiscale Theory and Computation, Westfälische Wilhelms-Universität Münster, Corrensstraße 40, 48149 Münster, Germany

3 for WF/DFT, we discuss formal implications and their consequences for the interpretation of results obtained with density-based embedding methods in Sect. 4. Illustrative calculations are presented in Sect. 5 to assess the relevance of these theoretical issues for practical calculations. We conclude in Sect. 6.

## 2 Background: sDFT versus DFT/DFT

The basic idea of sDFT is to split the total electron density into a set of subsystem electron densities. Each subsystem electron density is then represented in terms of a noninteracting reference system, i.e., through a set of Kohn–Sham-like orbitals. For simplicity, we will consider a two-partitioning case here with an active system  $A$  with an associated density  $\rho_A(\mathbf{r})$  and an environmental system  $B$  with a density  $\rho_B(\mathbf{r})$ , so that the total density is simply  $\rho(\mathbf{r}) = \rho_A(\mathbf{r}) + \rho_B(\mathbf{r})$ . Then, the sDFT energy functional can be written as a bifunctional [11, 12] (for further details, see also the recent reviews in Refs. [5, 14]),

$$E_{v_{\text{ext}}}^{\text{sDFT}}[\rho_A, \rho_B] = T_s[\rho_A] + T_s[\rho_B] + J[\rho_A + \rho_B] + E_{\text{xc}}[\rho_A + \rho_B] + V_{\text{ext}}[\rho_A + \rho_B] + T_s^{\text{nad}}[\rho_A, \rho_B] \quad (1)$$

where the index  $v_{\text{ext}}$  refers to the specific total external potential  $v_{\text{ext}}(\mathbf{r})$  in this system. The noninteracting kinetic energy functional  $T_s[\rho]$ , the Coulomb energy functional  $J[\rho]$ , and the exchange–correlation energy functional  $E_{\text{xc}}[\rho]$  are defined as in the context of Kohn–Sham DFT [17]. The nonadditive kinetic energy functional  $T_s^{\text{nad}}[\rho_A, \rho_B]$  is defined as,

$$T_s^{\text{nad}}[\rho_A, \rho_B] = T_s[\rho_A + \rho_B] - T_s[\rho_A] - T_s[\rho_B], \quad (2)$$

and  $V_{\text{ext}}[\rho_A + \rho_B]$  can be written explicitly as,

$$V_{\text{ext}}[\rho_A + \rho_B] = \int [\rho_A(\mathbf{r}) + \rho_B(\mathbf{r})] v_{\text{ext}}(\mathbf{r}) d\mathbf{r}. \quad (3)$$

The total external potential is usually just the electrostatic potential of the atomic nuclei, which can be assigned to either subsystem  $A$  or subsystem  $B$ . Hence, the external potential can be split up into  $v_{\text{ext}}(\mathbf{r}) = v_{\text{ext}}^A(\mathbf{r}) + v_{\text{ext}}^B(\mathbf{r})$ , where  $v_{\text{ext}}^A(\mathbf{r})$  denotes the electrostatic potential of the nuclei in system  $A$  only [and correspondingly for  $v_{\text{ext}}^B(\mathbf{r})$ ].

Minimization of  $E_{v_{\text{ext}}}^{\text{sDFT}}[\rho_A, \rho_B]$  with respect to (w.r.t.) a subsystem density, represented in terms of a noninteracting reference system, and assuming all other subsystem densities fixed, leads to a set of Kohn–Sham-like equations known as Kohn–Sham equations with constrained electron density (KSCED) [12]. In these equations, an extra embedding potential of the form

$$v_{\text{emb}}^A[\rho_A, \rho_B; v_{\text{ext}}^B](\mathbf{r}) = \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{\text{ext}}^B(\mathbf{r}) + \left. \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_A+\rho_B} - \left. \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_A} + \left. \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_A+\rho_B} - \left. \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_A}, \quad (4)$$

appears for system  $A$  (and correspondingly for system  $B$ ). Approximate forms are needed both for the exchange–correlation contributions and for the nonadditive kinetic energy functional in practical calculations using sDFT. The ground-state energy of the system can be found by minimizing the energy w.r.t. all subsystem densities (e.g., in so-called freeze-and-thaw cycles [12]).

The connection to hybrid-energy schemes becomes apparent if we split the energy functionals in Eq. (1). The Coulomb energy can easily be partitioned into,

$$J[\rho_A + \rho_B] = J[\rho_A] + J[\rho_B] + J_{\text{int}}[\rho_A, \rho_B], \quad (5)$$

where

$$J_{\text{int}}[\rho_A, \rho_B] = \int \int \frac{\rho_A(\mathbf{r})\rho_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (6)$$

$V_{\text{ext}}$  can be split into,

$$V_{\text{ext}}[\rho_A + \rho_B] = \int \rho_A(\mathbf{r})v_{\text{ext}}^A(\mathbf{r})d\mathbf{r} + \int \rho_B(\mathbf{r})v_{\text{ext}}^B(\mathbf{r})d\mathbf{r} + \int \rho_A(\mathbf{r})v_{\text{ext}}^B(\mathbf{r})d\mathbf{r} + \int \rho_B(\mathbf{r})v_{\text{ext}}^A(\mathbf{r})d\mathbf{r} = V_{\text{ext}}^A[\rho_A] + V_{\text{ext}}^B[\rho_B] + V_{\text{ext}}^B[\rho_A] + V_{\text{ext}}^A[\rho_B], \quad (7)$$

where the last line introduces short-hand notations for the different contributions. Finally, also the exchange–correlation energy functional can be split up,

$$E_{\text{xc}}[\rho_A + \rho_B] = E_{\text{xc}}[\rho_A] + E_{\text{xc}}[\rho_B] + E_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B],$$

where the nonadditive exchange–correlation energy functional  $E_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B]$  is defined analogously to the nonadditive kinetic energy functional in Eq. (2). Noting that the Kohn–Sham energy functional for a certain external potential  $v_{\text{ext}}^K(\mathbf{r})$  is given as,

$$E_{v_{\text{ext}}^K}^{\text{KS}}[\rho_K] = T_s[\rho_K] + J[\rho_K] + E_{\text{xc}}[\rho_K] + V_{\text{ext}}^K[\rho_K], \quad (9)$$

we can rewrite Eq. (1) as,

$$E_{v_{\text{ext}}}^{\text{sDFT}}[\rho_A, \rho_B] = E_{v_{\text{ext}}^A}^{\text{KS}}[\rho_A] + E_{v_{\text{ext}}^B}^{\text{KS}}[\rho_B] + J_{\text{int}}[\rho_A, \rho_B] + V_{\text{ext}}^A[\rho_B] + V_{\text{ext}}^B[\rho_A] + E_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B] + T_s^{\text{nad}}[\rho_A, \rho_B] \quad (10)$$

$$= E_{v_{\text{ext}}^A}^{\text{KS}}[\rho_A] + E_{v_{\text{ext}}^B}^{\text{KS}}[\rho_B] + E_{A \leftrightarrow B}^{\text{OFDFT}}[\rho_A, \rho_B], \quad (11)$$

where the last line defines the interaction-energy functional  $E_{A \leftrightarrow B}^{\text{OFDFT}}[\rho_A, \rho_B]$ . The superscript OFDFT stands for orbital-free DFT and indicates that this functional will be represented in terms of orbital-free density functional approximations in practice, which concerns in particular the term  $T_s^{\text{nad}}[\rho_A, \rho_B]$ . But also this functional can be considered exact in principle and can be understood as a difference of the corresponding energy functionals for the total system and the subsystems,

$$E_{A \leftrightarrow B}^{\text{OFDFT}}[\rho_A, \rho_B] = E_{v_{\text{ext}}}^{\text{OFDFT}}[\rho_A + \rho_B] - E_{v_{\text{ext}}^A}^{\text{OFDFT}}[\rho_A] - E_{v_{\text{ext}}^B}^{\text{OFDFT}}[\rho_B] \quad (12)$$

The form of the sDFT energy expression in Eq. (11) shows the typical separation into energy components for different subsystems as well as an interaction energy, which is well known in the case of (additive) hybrid-energy schemes [3]. Hence, one may consider this form also as a DFT/DFT hybrid method, especially since the approximations needed in practice for the exchange–correlation contributions may be chosen differently for subsystem *A*, *B*, and in the interaction-energy term. For example, one could think of pragmatically using a hybrid XC functional for subsystem *A*, a computationally simpler generalized gradient approximation (GGA) for subsystem *B*, and the local-density approximation (LDA) for  $E_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B]$ .

Hybrid-energy schemes can also be defined in a subtractive fashion, e.g., in the spirit of the ONIOM method [4, 18]. The sDFT energy functional can be brought into a corresponding form by using Eq. (12) and writing,

$$E_{v_{\text{ext}}}^{\text{sDFT}}[\rho_A, \rho_B] = E_{v_{\text{ext}}}^{\text{OFDFT}}[\rho_A + \rho_B] + \left( E_{v_{\text{ext}}^A}^{\text{KS}}[\rho_A] - E_{v_{\text{ext}}^A}^{\text{OFDFT}}[\rho_A] \right) + \left( E_{v_{\text{ext}}^B}^{\text{KS}}[\rho_B] - E_{v_{\text{ext}}^B}^{\text{OFDFT}}[\rho_B] \right). \quad (13)$$

Again, this has to be understood as an alternative to KS-DFT in principle. Only at the point where approximations are chosen in practice, differences will appear. A comparison of Eqs. (1), (11), and (13) indicates that sDFT may be regarded either as an alternative DFT formulation, which in principle can lead to the exact ground-state energy, or as a DFT/DFT hybrid-energy method, which leaves ample room for using different approximations for different subsystems (and their interaction energy) in practice.

The hybrid-method viewpoint on sDFT becomes fully apparent if we abandon the idea of using the energy expression in a true energy minimization with respect to the subsystem densities. Instead, pragmatic hybrid methods may simply use separate calculations to obtain the different energy contributions and charge densities for the subsystems.

For example, a DFT/DFT hybrid energy could in practice be calculated as,

$$E^{\text{DFT/DFT}} = E_{A+B}^{\text{OFDFT}} + \left( E_A^{\text{KS}} - E_A^{\text{OFDFT}} \right) + \left( E_B^{\text{KS}} - E_B^{\text{OFDFT}} \right), \quad (14)$$

where subscripts refer to the corresponding (sub-)systems. This implies that several independent calculations of OFDFT or KS-DFT type (using approximate functionals) may have to be carried out. Further details would have to be given to fully specify the computational approach. For example,  $E_A^{\text{KS}}$  could be, in the simplest case, the result of an isolated system Kohn–Sham calculation on subsystem *A*. Or, it could be obtained from a Kohn–Sham calculation employing an additional embedding potential, leading to a so-called electronic (or Hamiltonian) coupling scheme. Concerning a suitable form of such a potential, one could resort to a form similar to the embedding potential in the KSCED equations. But maybe, one would choose a potential evaluated for a given set of densities  $\rho_A$  and  $\rho_B$  instead of consistently employing a potential functional of these two densities. The resulting working equations may appear to be very similar, but not all energies and potentials may be obtained fully self-consistent in the general hybrid DFT/DFT case.

We would like to note that the expression in Eq. (14), which bears a similar structure as the sDFT energy functional in Eq. (13), is not exactly what one would expect for a KSDFT/OFDFT hybrid method in the style of the ONIOM approach. There, typically only one local energy correction for the active region is included using a more accurate method. In the present case, now referring to typical approximate versions of KS-DFT and OFDFT, one would consider KS-DFT as the more accurate method. Hence, the expected form of a KSDFT/OFDFT hybrid energy would be,

$$E_{A+B}^{\text{KSDFT/OFDFT}} = E_{A+B}^{\text{OFDFT}} + \left( E_A^{\text{KS}} - E_A^{\text{OFDFT}} \right). \quad (15)$$

This indicates that several degrees of freedom exist in the setup of sDFT or DFT/DFT hybrid calculations in practice, concerning in particular the choice of approximate energy functionals and the way in which the densities used in the final energy calculations are obtained (e.g., under the influence of consistent, non-consistent, or no embedding potentials at all). The last aspect also concerns the choice of free variables if a DFT/DFT hybrid-energy expression is considered a true energy functional (see also the discussion in Ref. [14]): Either, we could choose the subsystem densities  $\rho_A(\mathbf{r})$  and  $\rho_B(\mathbf{r})$  as functions to be optimized, as in the original sDFT method; or, for an energy functional with a partitioning as in Eq. (15), it might seem more plausible to consider  $\rho_A(\mathbf{r})$  and the total density  $\rho(\mathbf{r})$  as free variables, i.e.,

$$E_{v_{\text{ext}}}^{\text{KSDFT/OFDFT}}[\rho_A, \rho] = E_{v_{\text{ext}}}^{\text{OFDFT}}[\rho] + \left( E_{v_{\text{ext}}^A}^{\text{KS}}[\rho_A] - E_{v_{\text{ext}}^A}^{\text{OFDFT}}[\rho_A] \right). \quad (16)$$

### 3 Different views on WF/DFT hybrid methods

Just like sDFT and methods related to it, also combinations of WF and DFT methods can be understood either from a pragmatic multilevel (or hybrid method) perspective or from a fundamental DFT point of view. These different points of view reflect very different motivations for using WF/DFT combinations. One viewpoint assumes that an accurate correlated WF method is needed to describe a certain active subsystem by means of Quantum Chemistry, which is embedded in a larger environment. Since a full WF treatment of active subsystem and environment is often computationally not feasible, one makes the attempt to use a simpler description of the environment in terms of a density functional theory approximation. We will call this perspective on WF/DFT methods the *hybrid-method point of view*. Related to this, a motivation could be to just describe changes in the active subsystem caused by the modulation of its properties due to the environment. The properties of the environment itself are assumed to be of no interest in this context. Then, the effect of the environment may be considered as a perturbation in the energy and wavefunction of the active system. DFT in this *perturbation-theory point of view* would only be needed to define an (approximate) perturbation operator.

The *fundamental DFT point of view* is quite different. It starts from the observation that the use of a noninteracting Kohn–Sham reference system may lead to certain difficulties in practical calculations, which may be circumvented if a part of the total electron density is represented in terms of an interacting system of electrons. This part—typically the active center—can then be expressed through a many-body wavefunction, while a standard Kohn–Sham-like system is used for the remainder (environment).

The reason why we introduce these different approaches as different *points of view* is that the resulting working equations may appear very similar (or even identical), but can be thought of as being derived in completely different theoretical frameworks. Hence, how the results of a WF/DFT calculation may be interpreted will depend on the point of view one is taking on how the working equations were derived. This will be explained in the following.

#### 3.1 The hybrid-method point of view

The extension of the discussion in Sect. 2 to WF/DFT embedding is straightforward, although even more possible combinations/approximations arise in practice. We will, in the following, only indicate the structure of the energy expression, but not indicate explicitly at each time whether an actual energy functional is constructed and subsequently used in a consistent variational optimization. Rather, we will include examples from the literature to indicate which

strategies are in use for calculating the different energy components. The reader is referred to the original work by Carter and coworkers [15, 19, 20] as well as to recent reviews [1, 5, 14] for additional details.

In analogy to the subtractive approach in Eq. (15), we can write

$$E_{A+B}^{\text{WF/DFT}} = E_{A+B}^{\text{DFT}} + (E_A^{\text{WF}} - E_A^{\text{DFT}}). \quad (17)$$

For this to be consistent, “DFT” should refer to the same type of DFT method in the first and third term on the right-hand side in all practical (approximate) formulations of the method. In additive schemes of the type,

$$E_{A+B}^{\text{WF/DFT}} = E_A^{\text{WF}} + E_B^{\text{DFT}} + E_{A\leftrightarrow B}^{\text{DFT}'}, \quad (18)$$

by contrast, it is much more common to choose the approximation (DFT’) for the interaction energy  $E_{A\leftrightarrow B}^{\text{DFT}'}$  independently from the approximation for  $E_B^{\text{DFT}}$ . In fact, as we have seen above, sDFT uses KSDFT for  $E_B^{\text{DFT}}$ , but OFDFT for  $E_{A\leftrightarrow B}^{\text{DFT}'}$ . The corresponding expression in WF/DFT is (see, e.g., Ref. [21]),

$$E_{A+B}^{\text{WF/DFT}} = E_A^{\text{WF}} + E_B^{\text{KSDFT}} + E_{A\leftrightarrow B}^{\text{OFDFT}}. \quad (19)$$

In Carter’s original work [15, 19], such an expression (without selecting KS- or OFDFT at this point) was explicitly considered as an energy functional and used to derive an extra embedding potential appearing in the WF calculation on subsystem  $A$  as a functional derivative of the interaction term with respect to  $\rho_A$ , assuming  $\rho_B$  fixed. Since localized phenomena in periodic systems were studied by Carter, treating  $\rho_B$  explicitly (in the presence of  $\rho_A$ ) was avoided using a subtractive ansatz for the actual energy calculation as [19],

$$E_{A+B}^{\text{WF/DFT(Carter)}} = E_{A+B}^{\text{DFT}} + (E_A^{\text{WF}} - E_A^{\text{DFT}}). \quad (20)$$

Here,  $E_A^{\text{DFT}}$  was calculated for the converged density obtained from the embedded wavefunction calculation on system  $A$ . While this is straightforward for the electron–nucleus interaction energy, the electron–electron Coulomb energy, and the exchange–correlation energy (once a density-based approximation has been selected), calculating the kinetic energy poses a challenge here: The Kohn–Sham expression cannot be used directly, as no Kohn–Sham orbitals for this system are determined. Possible solutions include calculating this contribution (1) from an explicitly density-dependent kinetic energy functional [turning “DFT” in Eq. (20) into OFDFT] [19], (2) by approximately evaluating  $T_s$  with the Hartree–Fock (HF) orbitals underlying the correlated wavefunction calculation [22] [turning “DFT” itself into a HF–DFT hybrid-energy expression],

or (3) by reconstructing an effective Kohn–Sham potential and the corresponding orbitals that reproduce the desired density, so that  $T_s$  can be calculated from those orbitals [turning “DFT” into an optimized effective potential (OEP) KSDFT], an idea mentioned explicitly already in Ref. [22]. Other options and additional approximations are compared in Ref. [20].

### 3.2 Perturbation-theory point of view

The actual research question in many applications of hybrid schemes is not related to properties of the entire system ( $A + B$ ), but rather to how the properties of system  $A$  are changed by the presence of system  $B$ , compared to its properties as an isolated system. When talking about properties related to the electronic structure of  $A$  in an sDFT framework, this change is caused by the embedding potential given in Eq. (4).

Note that  $v_{\text{emb}}^A$  depends on both  $\rho_A$  and  $\rho_B$ , which can change during the (iterative) optimization of the total sDFT energy in f&t cycles. This means that also  $v_{\text{emb}}^A$  needs to be updated iteratively in sDFT calculations. But assuming that it is possible to create a reasonable guess for  $v_{\text{emb}}^A$  with fixed approximations for  $\rho_A$  and  $\rho_B$  (e.g., isolated-molecule densities), we would have a fixed additional embedding potential  $\tilde{v}_{\text{emb}}^{(A)}$  that may be regarded as an approximate perturbation operator for system  $A$  [23]. Then, one could try to find approximate eigenfunctions and the corresponding properties for the modified system- $A$  Hamiltonian,

$$\hat{H}_A \longrightarrow \hat{H}_A + \sum_{i=1}^{N_A} \tilde{v}_{\text{emb}}^A(\mathbf{r}_i), \quad (21)$$

where  $N_A$  is the number of electrons in system  $A$ . With this definition of the perturbation operator, a first-order change in the ground-state energy of system  $A$  could be calculated using standard perturbation theory as,

$$\Delta E_A^{(1)} \approx \left\langle \Psi_A^{(0)} \left| \sum_{i=1}^{N_A} \tilde{v}_{\text{emb}}^A(\mathbf{r}_i) \right| \Psi_A^{(0)} \right\rangle, \quad (22)$$

where  $\Psi_A^{(0)}$  is the (approximate) ground-state wavefunction of the isolated system  $A$ . Similarly, the first-order change in the wavefunction would give access to modified properties of system  $A$ . But at that point, questions about the interpretation of subsystem properties in a larger environment may arise (see Sect. 4.1).

While we are not aware of actual applications of precisely this perturbation-theory strategy, there are studies making use of similar strategies in which the active-system wavefunctions are determined under the influence of such an approximate embedding operator for extracting local properties (e.g., in Ref. [24]).

### 3.3 DFT point of view

As shown by Wesolowski in 2008 [16], a rigorous combination of WF and DFT methods can be achieved in the following way: instead of an independent-particle, single-determinantal KS system, we can represent the electron density of system  $A$  by a multi-determinantal wavefunction  $\Psi_A^{\text{MD}}$  of an interacting system of electrons. If the wavefunction ansatz can lead to the exact solution, we can set up an energy functional as,

$$\begin{aligned} E_{\text{vext}}^{\text{WF/DFT}}[\Psi_A^{\text{MD}}, \rho_B] = & \langle \Psi_A^{\text{MD}} | \hat{T} + \hat{V}_{\text{ee}} | \Psi_A^{\text{MD}} \rangle \\ & + V_{\text{ext}}^A[\rho_A] + T_s[\rho_B] + J[\rho_B] \\ & + E_{\text{xc}}[\rho_B] + V_{\text{ext}}^B[\rho_B] \\ & + V_{\text{ext}}^A[\rho_B] + V_{\text{ext}}^B[\rho_A] \\ & + J_{\text{int}}[\rho_A, \rho_B] + T_s^{\text{nad}}[\rho_A, \rho_B] \\ & + E_{\text{xc}}^{\text{nad}}[\rho_A, \rho_B], \end{aligned} \quad (23)$$

where  $\rho_A$  is the density associated with  $\Psi_A^{\text{MD}}$ . Full minimization of this functional w.r.t.  $\Psi_A^{\text{MD}}$  and  $\rho_B$  will then lead to the true ground-state energy of the system. If the wavefunction search space does not include the exact wavefunction, then formally an additional term has to be included to reach the exact ground-state energy [16]. This is due to the fact that  $\min_{\Psi_A^{\text{MD}} \rightarrow \rho_A} \langle \Psi_A^{\text{MD}} | \hat{T} + \hat{V}_{\text{ee}} | \Psi_A^{\text{MD}} \rangle$  is not exactly equal to  $T[\rho_A] + V_{\text{ee}}[\rho_A]$  in that case. If the energy is variationally minimized w.r.t.  $\rho_A$  (represented in terms of  $\Psi_A^{\text{MD}}$ ), the same form of the embedding potential as derived in the hybrid-scheme point of view arises in the equations that determine  $\Psi_A^{\text{MD}}$ —apart from the extra term in case of non-exact wavefunctions. Note the technical similarity of this approach to multiconfiguration-DFT methods [25–27].

## 4 Implications

The different points of view have some interesting implications concerning the realm of application of the methods—in principle and in practice—as well as the interpretation of results obtained from calculations based on sDFT and WF/DFT embedding methods. Generally speaking, the greatest advantage of WF/DFT embedding can be expected from the practical ability to overcome the shortcomings of approximate DFT methods currently in use, while still being able to describe a large system quantum mechanically. One could also argue that WF/DFT embedding can be seen as an improvement in QM/MM methods which does not rely on any kind of empirical parametrization of a force field (apart from parameters in the approximate functionals), and which automatically comes with a (at least potentially) consistent electronic/Hamiltonian coupling between the subsystems. Consequently, the requirement of

an extra density functional correcting for the non-exactness of the WF method as derived in Ref. [16], though formally correct, may appear counterintuitive when considered from a hybrid-method perspective: One sets out to overcome the limitations of approximate DFT methods, but then reintroduces an additional density functional to correct the accurate wavefunction method. This apparent paradox is easily reconciled when acknowledging the different points of view—one that is motivated by computational practice and the other one by formal DFT considerations. Several additional questions arise due to the fact that wavefunctions and densities of different systems are combined for a description of an entire (larger) supersystem, as will be discussed in the following.

#### 4.1 Accessible quantities and interpretation

When interpreting results from WF/DFT calculations, it is important to keep in mind which quantities are accessible through such embedding procedures. When adopting the perturbation-theory point of view, we only have access to the properties of the active system, modified by the environment; no properties of the total system are available. For specific applications, this drawback can pragmatically be considered an advantage. For example, local excitations can be described, which are free from contaminations like overdelocalization or mixing with orbitals of the environment [28]. In case of the hybrid-method point of view, total energies of the entire system are accessible, as well as subsystem energies and interaction energies between the subsystems. Besides these energetic quantities, also the subsystem densities and, as their sum, the total electron density of the system are available from these calculations. Note that approximate schemes of “subtractive” type [see Eq. (17)] may yield a total density  $\rho = \rho_A + \rho_B$  and a local correction for  $\rho_A$  [19], which can introduce inconsistencies in the interpretation of the total density [14]. If the total density is available, then naturally, also all properties that can be derived from it are accessible through such a calculation. In these respects, the hybrid-method and the DFT point of view agree.

A difference arises, however, concerning the question of the wavefunction. Considering WF/DFT a hybrid method, we will usually assume that the wavefunction part really can be interpreted as a wavefunction of the active system under the influence of the environment (similar to a perturbation-theory point of view). As a consequence, properties of the active system could be calculated as expectation values of the wavefunction obtained under the influence of  $v_{\text{emb}}^A$ . It is also clear that such interpretations are limited to cases where the wavefunction still shows similarity to the isolated-molecule case, i.e., where the effect of  $v_{\text{emb}}^A$  is small. Otherwise, one is in danger of leaving the realm of

application of such a hybrid method. In the strict DFT perspective, however, the correlated wavefunction is merely an auxiliary mathematical object to represent the electron density, without an own physical meaning. This is similar to the strict interpretation of a KS wavefunction and of KS orbitals, which have a physical meaning only through their electron density. But since KS wavefunctions are often approximately used to calculate expectation values of operators (e.g., of  $\hat{S}^2$ ), it is conceivable that the WF/DFT community will make continued (and successful) use of the correlated wavefunction in the interpretation of properties of the active system.

#### 4.2 Ability to describe strongly coupled systems

Density-embedding schemes making use of an embedding potential have been criticized for not being able to describe the environment effect on the “many-body fragment state” of the embedded fragment [29] if the coupling is strong. This was explained by the fact that the embedded fragment is an open quantum system, entangled with its environment. Reference [29] also correctly points out (however, only in a footnote) that this argument only holds if the active system is described with an explicit many-body wavefunction method and not for pure DFT descriptions. But even in case of a WF/DFT description, at least the DFT point of view is not affected by this argument, since then the wavefunction is merely an auxiliary object for representing the density. We note in passing that the application of density matrix embedding theory [29], an embedding method which solves this problem by using a set of bath states instead of an effective embedding potential, requires the knowledge of an (approximate) wavefunction for the total system and thus, strictly speaking, cannot be combined with a DFT treatment of the environment.

From a practical perspective, several possible approaches for embedding in strongly interacting systems have been presented in the context of density-based embedding approaches [20, 30–33]. These are sometimes called “exact” to underline that they can in principle numerically reproduce the effect of an exact embedding potential in the limit of exact exchange–correlation functionals, infinite basis sets, and infinitely high precision in numerical procedures. Even with limitations in practical applications, these embedding schemes still can offer accurate results for various systems, including strongly coupled ones (e.g., those connected through covalent bonds). For some special cases, even exact analytical solutions are available [34, 35].

#### 4.3 Orthogonality between different states

Khait and Hoffmann [36] provided an explicit theoretical basis for WF/DFT calculations of excited states based on

earlier work by Perdew and Levy [37], which showed that every extremum of the energy functional is associated with a corresponding stationary-state density. A formal difficulty is that the opposite is not always true, so that only a subset of excited states is covered by this argument. Khait and Hoffmann argued that different electronic states of the active system would lead to different environment densities and embedding potentials that differ not only in the  $\rho_A$ -dependent part (active system's density), but also in the part depending on the environment density. The analogue in QM/MM schemes would be polarizable force fields, in which the environmental potential also reacts to the active system's density. A first application of such state-specific embedding potentials for excited-state WF/DFT calculations has been presented in Ref. [21]. A question that has been brought up in that study concerns the orthogonality between wavefunctions for different electronic states of the embedded system. A discussion of the special case where the environment density is frozen (and identical) in all electronic states has been presented in Ref. [38]. But since electronic excitations in general can affect the entire system, we will be concerned here with the more general case in which electron densities of all subsystems may change.

Technically, the problem may seem similar to the problem of non-orthogonality in state-specific multi-configuration self-consistent field (MCSCF) calculations, where the underlying orbitals are optimized for each state separately and thus are non-orthogonal. Similarly, if the embedding potential in a WF/DFT calculation is state specific, the orbitals obtained in the wavefunction calculation will be different for different states, again leading to non-orthogonality. The important conceptual difference, however, is the following: In state-specific MCSCF, we are dealing with two different electronic states of the same isolated system, which should be strictly orthogonal in principle. In WF/DFT with state-specific embedding potentials, by contrast, orthogonality is strictly required only for total states, not for wavefunctions describing one part of the total system. This can easily be seen for the limiting case of CASSCF with an active space including all (occupied and virtual) orbitals for the active subsystem, which is equivalent to a full configuration interaction (FCI) calculation with an obsolete orbital optimization step: In case of the isolated active system, the results will, for a given one-electron basis, be invariant with respect to orbital rotations, and all (nondegenerate) states will be orthogonal. But for the embedded system, the change in the embedding potential for different electronic states will still change the one-electron integrals and thus lead to non-orthogonal wavefunctions for the active part in state-specific FCI/DFT embedding.

Analyzing non-orthogonality for the supersystem is problematic, since the wavefunction of the environment (described by DFT) is entirely unknown, which, as

a consequence, means that also the total wavefunction is unknown. And if we assume the strict DFT point of view, then even the WF for the embedded part is only an auxiliary quantity and may not be interpreted as a true wavefunction describing the embedded electronic system.

For practical purposes, however, it is interesting to explore how large the non-orthogonality effects for the active system actually are, and how quantities like transition moments between different electronic states react to this effect. This will be tested in Sect. 5.5.

#### 4.4 Polarization versus relaxation

Fully consistent variational WF/DFT calculations, just like subsystem DFT calculations, require an iterative update of all subsystem densities/wavefunctions under the influence of the embedding potential. This procedure has to be continued until all densities are consistent with the resulting total potentials, including the embedding contribution. From the fundamental DFT point of view, the resulting subsystem properties have only an auxiliary function, and only properties of the total system may be interpreted. This also means that one can interpret the density change of the total system (compared to that of the sum of isolated subsystem densities) as a polarization effect. But there is no strict justification for referring to a *subsystem* density change as a physical polarization effect. Rather, this can be called more technically a relaxation effect [39, 40].

On the practical side, a natural, though not the only possible choice is to start such iterative calculations from the isolated subsystems which constitute the total system. Hence, it is in fact tempting to interpret the changes in properties of the subsystems due to the iterative process as a physical consequence of the interaction between the subsystems. This can be motivated with the perturbation-theory point of view and often works quite well in practice, since many properties have a well-defined local character. Prototypical examples are local excitation energies [see, e.g., Refs. [21, 22, 24, 41] (hybrid WF/DFT approaches), and Refs. [28, 42, 43] (embedded linear-response time-dependent DFT approaches)]. A question that arises, however, is whether changes in densities or related integrated descriptors like dipole and higher multipole moments of individual subsystems may be interpreted in practice. This is often done in actual calculations, and the results usually correlate well with expectations based on experimental findings. For example, the change of the dipole moment of water in water has been addressed in Refs. [44, 45].

Fundamentally speaking, the well-known problem with such interpretations is the following: If  $\rho_1, \rho_2$  are solutions of the subsystem problems adding up to the correct total density, then also

$$\tilde{\rho}_1(\mathbf{r}) = \rho_1(\mathbf{r}) + \delta\rho(\mathbf{r}) \quad (24)$$

$$\tilde{\rho}_2(\mathbf{r}) = \rho_2(\mathbf{r}) - \delta\rho(\mathbf{r}) \quad (25)$$

are suitable solutions (as long as they are noninteracting  $v$ -representable). This has been discussed in the literature at several points [13, 20, 35, 46]. These two sets of densities would give rise to different subsystem dipole moments (see also the discussion in Ref. [40]). In fact, if the subsystems are not neutral, but the total system is, then the subsystem's dipole moments are origin dependent, while the total dipole moment is not (similar problems may occur for higher multipole moments even in the neutral case). A fundamental solution to this ambiguity has been formulated by the requirement of a unique embedding potential for all subsystems, which relies on potential reconstruction techniques [20]. A similar approach is used in partition DFT [8, 9], which yields a unique partitioning into subsystems with fractional electron numbers.

In a sDFT calculation, there is in principle no preference for any specific partitioning of the density into subsystem densities in the limit of using exact functionals. However, as Wesolowski and coworkers demonstrate in Ref. [40], this is not true in practical calculations which employ approximate functionals to evaluate  $T_s^{\text{nad}}$ . If one minimizes the energy with respect to the subsystem densities, the use of approximations for  $T_s^{\text{nad}}[\rho_A, \rho_B]$  leads to an unphysical redistribution of the subsystem densities once the exact total density has been found. The reason is that the approximate  $T_s^{\text{nad}}[\rho_A, \rho_B]$  is the only term in the sDFT energy functional which depends on the partitioning of  $\rho = \rho_A + \rho_B$  into its contributions  $\rho_A$  and  $\rho_B$ . Note that in certain cases, not only the approximate  $T_s^{\text{nad}}[\rho_A, \rho_B]$ , but also approximations to  $E_{\text{xc}}[\rho_A, \rho_B]$  can cause such a redistribution. This can happen (1) if the intra-subsystem exchange–correlation approximation in sDFT is chosen differently from the nonadditive one [e.g., if orbital-dependent approximations are applied, as is done in Ref. [40] for the corresponding embedding potential], (2) if different exchange–correlation approximations for different subsystems are pragmatically employed (as in Ref. [47]), or (3) in approximate WF/DFT schemes, in which necessarily exchange and correlation effects are treated differently for the different systems. In practice, however, sDFT and WF/DFT calculations usually start out from isolated subsystem densities and lead to embedded subsystem densities which still show similarity to the starting points. They will, in general, not lead to the global minimum of the energy functional because (1) limitations in the basis set prevent an easy redistribution of density between the subsystems and (2) the starting point of the calculation may lead to local minima which correspond to “perturbed” isolated subsystem densities. Wesolowski and coworkers mention that the electronic polarization is dominant in case of charged, non-covalently linked

subsystems [40]. But there are also cases of neutral subsystems for which the electrostatic component is dominant [28, 47]. Some simple cases will be studied in Sect. 5. This discussion is getting even more difficult if excited states are considered, where it can be interesting to determine the *differential* polarization of a total system or a subsystem between ground and excited states. Also here we note that one typically starts from isolated-molecule excitations, which are modulated by a comparatively small effect of the environment. The analysis in Ref. [21] indicates that the electrostatic terms in fact dominate for the excitations studied there, which can be taken as a justification for using the term *differential polarization*.

An important question for the pragmatic interpretation of subsystem and total system properties is in this context, whether or not practical WF/DFT and subsystem DFT calculations (relying on approximate analytical rather than reconstructed embedding potentials) converge to the same resulting densities irrespective of the starting point and the way in which the densities are updated in the iterative process. Usually, it is assumed that this is in fact the case. We will present some illustrative examples also for this question in Sect. 5. It should be mentioned that different solutions can certainly be technically provoked if the starting points are chosen far from reasonable for a certain target system. As an example, consider a pair of neutral fragments that form a neutral and rather nonpolar complex. In principle, we are free to choose the numbers of electrons in the embedding calculation such that they correspond to ionic isolated systems. As long as the basis set is flexible enough, and provided that we have access to (near-)exact embedding potentials, the total density should correspond to the neutral total system (see the example of ethane represented as  $\text{CH}_3^+ \cdots \text{CH}_3^-$  in Ref. [31]). In most practical calculations, however, we will work with monomer basis sets and approximate potentials and thus may get trapped in solutions for the total density that are close to the initial guess. In fact, this can pragmatically be exploited to optimize excited charge-separated states [48–50] or other types of diabatic states [51, 52] with subsystem DFT. These examples also demonstrate that the main problem of approximations in  $T_s^{\text{nad}}[\rho_A, \rho_B]$  is that they can lead to large discrepancies in the total density (compared to KSDFT), independently of how the subsystem contributions to the density look like.

## 5 Illustrative calculations

In the following, we present illustrative calculations addressing some of the aspects discussed in the previous sections. All discussions of results obtained here have to be understood from a practical perspective, i.e., assuming approximate exchange–correlation and nonadditive kinetic energy functionals.



## 5.1 Computational details

If not stated otherwise the following programs and settings have been used for all calculations presented in this section: ADF [53, 54] was employed for all DFT-based calculations with a TZP basis [55] and the PW91 [56] functional. FDE and sDFT calculations have been carried out with the corresponding implementation [57] in ADF applying the PW91k [58] functional for the nonadditive kinetic energy; monomer basis sets have been employed if not explicitly stated otherwise. MOLCAS [59] has been used for all wavefunction calculations. WF/DFT calculations are carried out through an interface [60] integrated into the PYADF scripting framework [61]. The strategy used within that interface corresponds to the *splitSCF* scheme presented in Ref. [62]. Supersystem geometries have been optimized with TURBOMOLE [63] employing the BP86 [64, 65] functional and a def2-SVP [66] basis. Subsystem geometries were extracted from the optimized supermolecular geometries without further optimization.

## 5.2 Subsystem dipole moments

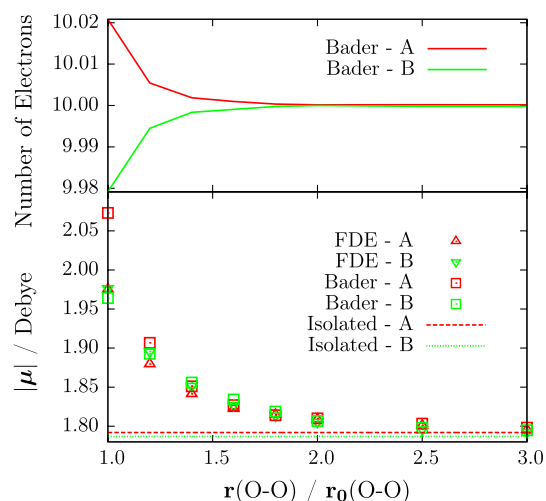
As discussed in Sect. 4.4, properties of subsystems which are part of a supersystem, are, in general, not measurable. This is a consequence of the ambiguity in the splitting of a supersystem. Nevertheless, assigning the electron density of a supersystem to molecular subsystems introduces a much smaller ambiguity than, e.g., the assignment of atomic partial charges by typical population analysis tools, as molecular subunits are typically well separated.

Here, we assess whether the dipole moment derived from a subsystem density as obtained in a sDFT calculation can be considered meaningful. It can be defined as

$$\mu^A = \sum_I^{N_{\text{nuc}}^A} Z_I \cdot (\mathbf{R}_I - \mathbf{r}_0) - \int (\mathbf{r} - \mathbf{r}_0) \cdot \rho_A(\mathbf{r}) d\mathbf{r}, \quad (26)$$

where the sum runs over all  $N_{\text{nuc}}^A$  nuclei of the active subsystem,  $Z_I$  is the charge and  $\mathbf{R}_I$  are the coordinates of one of these nuclei, and  $\mathbf{r}_0$  is the origin of the dipole moment operator. Clearly, with this definition, the dipole moment of the total system is recovered as the sum of all subsystem dipole moments for any splitting of the total density. In the following examples, we compare these subsystem dipole moments to a partitioning based on a Bader analysis (“atoms in molecules”) [67]. The Bader scheme identifies atoms solely based on the topology of a given electron density. A subsystem density can then be formulated as the sum of densities associated with the atoms of one subsystem.

As a test system, we chose the water dimer obtained from the S22 test set [68]. We varied the distance between



**Fig. 1** *Top* the number of electrons associated with the subsystems for the case of the Bader partitioning for two water molecules in a water dimer. *Bottom* subsystem dipole moments obtained from partitioning schemes based on a Bader analysis and sDFT, respectively. Data are shown for different distances  $\mathbf{r}$  between the oxygen atoms in units of the equilibrium distance  $\mathbf{r}_0$

the monomers (measured by the distance of the oxygen atoms) and calculated subsystem dipole moments with both partitioning schemes discussed above, respectively, for each distance. The sDFT data were obtained from densities of a converged f&t procedure, while for the Bader analysis the density of a converged supermolecular calculation was used.

While in sDFT the number of electrons in each subsystem is a fixed input parameter, it is a (generally non-integer) result in the Bader analysis. Since the dipole moment of a charged system is origin dependent, subsystem dipole moments based on a Bader analysis are in general also dependent on the choice of the origin, although the dependence may be weak. This problem is avoided in an analysis based on sDFT if neutral subsystems are chosen. In the Bader results reported here, we chose the center of mass of the total system as the origin for the dipole moment operator. The results are shown in Fig. 1.

Both sDFT and KSDFT predict an increase in the total dipole moment compared to the sum of dipole moments for the isolated subsystems. Compared to KSDFT, sDFT underestimates the total dipole moment at the equilibrium distance by a small amount (about 7 %), but this deviation quickly decreases for longer distances. Moreover, sDFT and the Bader analysis agree in the overall trend of increasing magnitude of the subsystem dipole moments with decreasing distance. Also the directions of the subsystem dipole vectors are in good agreement: The angles between the sDFT and Bader dipole moment vectors are smaller than 3.6 degrees for all cases considered here, and quickly decrease

to  $<1$  degree for larger distances. But it can also be seen that the dipole moment of one of the water molecules is significantly larger than that of the other one in case of the Bader partitioning at the equilibrium distance. This is at least partially caused by a small charge-transfer effect observed in the Bader scheme, as one of the water molecules has about 0.04 electrons more than the other one at equilibrium distance. At large distances, both schemes converge to the dipole moments obtained from isolated subsystem calculations. Here, also the Bader partitioning results in the expected number of 10 electrons per water molecule within the numerical accuracy of the integration grid.

As an additional test system, we considered the guanine trimer from the test set of Ref. [69]. Also for this system, we calculated the sDFT and Bader subsystem dipole moments for all monomers. The resulting dipole moment vectors for the central monomer are shown in Fig. 2. It can be seen that the sDFT and Bader subsystem dipole moments agree almost perfectly and are considerably different from the dipole moment of the isolated monomer. The magnitude of the dipole moment of the isolated central monomer is 7.48 Debye, while those resulting from sDFT (5.99 Debye) and a Bader analysis (5.95 Debye) are about 20 % smaller. The angle formed between the isolated-molecule dipole moment vector and the sDFT one is 4.2 degrees, and the angle between isolated and Bader dipole vector is 5.1 degrees.

Both schemes agree qualitatively to each other and to the intuitive understanding of chemical systems. These results are thus encouraging for subsystem properties with sDFT, especially for properties which cannot be extracted as easily from a supermolecular calculation as charges or dipole moments.

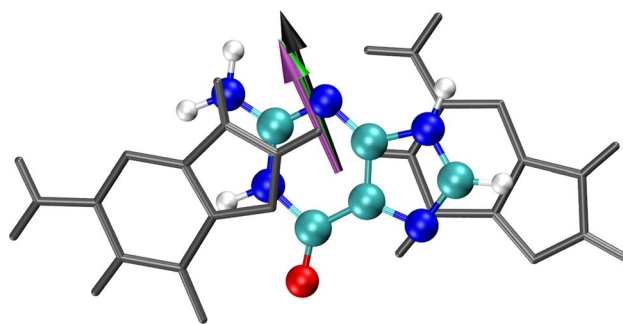
### 5.3 Polarization of a subsystem density

In the following, we provide sample calculations for assessing whether subsystem densities obtained from sDFT calculations can be considered physically meaningful. In order to do this, we will consider a case where an intuitive partitioning is possible on the basis of a supermolecular KSDFT calculation. Specifically, we look at the difference of the electron density in a HCN dimer to the sum of isolated monomer densities,

$$\delta\rho = \rho^X - (\rho_A^{\text{iso}} + \rho_B^{\text{iso}}), \quad (27)$$

where  $\rho^X$  is the density of a supermolecular ( $X = \text{super}$ ) or a sDFT ( $X = \text{sDFT}$ ) calculation and  $\rho_A^{\text{iso}}$  and  $\rho_B^{\text{iso}}$  are the densities of the isolated monomers (in the structure of the dimer). In this way, we can compare how the two different methods describe the polarization of the total system.

We then analyze the subsystem densities. In our test system, the two monomers are clearly distinguishable for

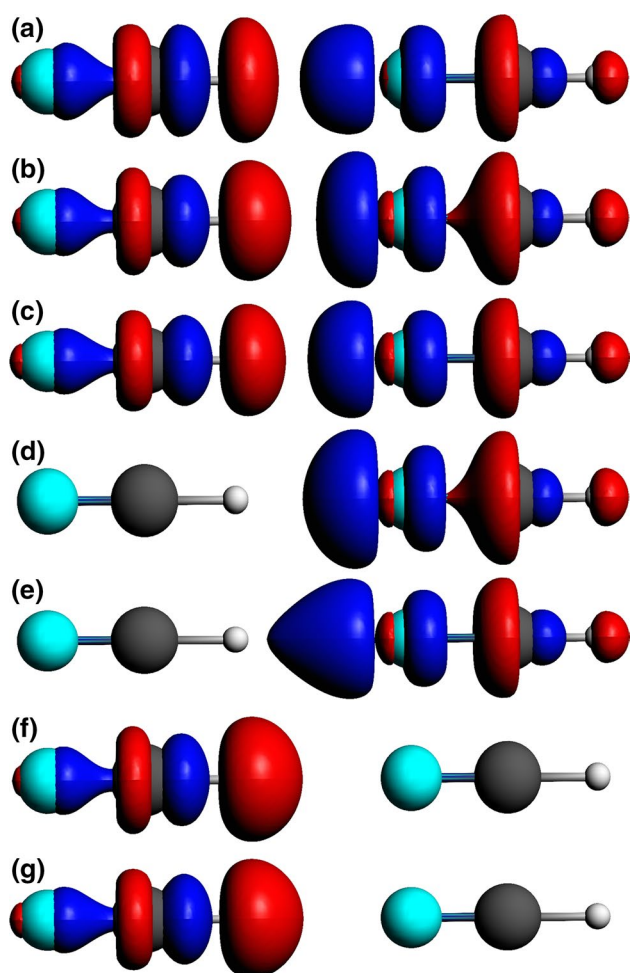


**Fig. 2** Dipole moment vector of the central monomer in the guanine trimer from Ref. [69] calculated with a subsystem density derived from a Bader analysis (*green arrow*), a sDFT calculation (*purple arrow*), and for an isolated calculation (*black arrow*). As an origin for the dipole moment vectors, we chose the center of mass of the central monomer. The environmental monomers are shown in *gray*

sufficiently large separations, since the density in the space between the monomers approaches zero. Any reasonable method to split the supermolecular density should end up with subsystem densities that resemble the results obtained from a straightforward geometrical splitting of the supermolecular density. At smaller distances, by contrast, the density in the region between the monomers cannot be split up as easily. However, for other parts of the system, an unambiguous assignment of density (and difference density) to a certain monomer is still possible.

Figure 3 shows the density change of a HCN dimer at equilibrium distance ( $R_{\text{eq}}$ ) compared to the isolated monomers. The dimer is arranged in a linear fashion, forming a hydrogen bond from N to H. For the sDFT case, also the difference of the subsystem density to the density of an isolated calculation is shown for both monomers; in addition to sDFT results employing the default monomer basis sets, we also show results obtained with a supermolecular basis.

The difference density for the total system at equilibrium distance shows the same qualitative behavior for the sDFT and supermolecular calculation. However, the agreement is not quantitative due to the approximation used for  $T_s^{\text{nad}}[\rho_A, \rho_B]$ . In panels b, d, and f of Fig. 3, an additional error source is the restriction to monomer basis sets in the subsystem calculations, which is the usual setup in multilevel calculations. By comparison of panels b and c in that figure, we see that the additional basis functions lead to a better agreement of the total density. The corresponding subsystem density plots suggest that in the calculation with a supermolecular basis, additional density is shifted from the right toward the left monomer. Qualitatively, the shown polarization of the subsystem densities obtained from the sDFT calculations match the corresponding part of the polarization of the supersystem almost perfectly for both setups. No contributions are visible on the respective other monomer. Only in the space between the subsystems,

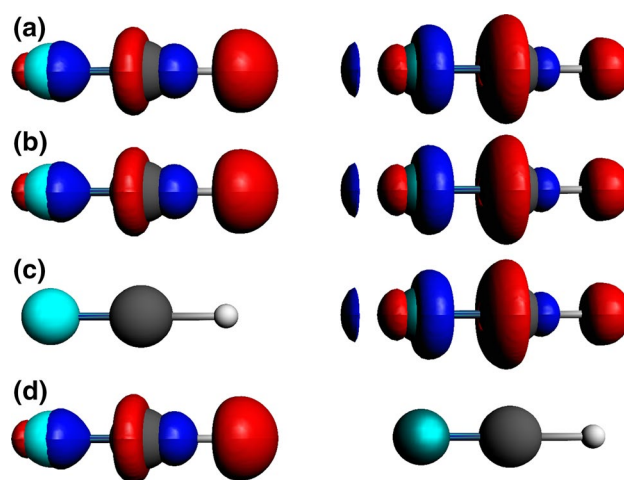


**Fig. 3** Density differences w.r.t. the densities of the isolated monomers for an HCN dimer at equilibrium distance ( $R_{\text{eq}}$ ). **a** Supermolecular density, **b, c** sDFT total density, **d, e** subsystem density of the left monomer, and **f, g** of the right monomer from an sDFT calculation. The density differences in panels **b, d**, and **f** have been obtained in sDFT calculations using monomer basis sets, the ones in panels **c, e**, and **g** in sDFT calculations with supermolecular basis sets. *Blue* positive regions, *Red* negative regions, isovalue:  $10^{-3}$

slight deviations are visible, which is attributed to the fact that sDFT allows the densities of different subsystems to overlap.

The situation for a larger distance (here:  $3R_{\text{eq}}$ ) between the monomers is depicted in Fig. 4. Here, the results based on sDFT and a supermolecular calculation agree even better (for sDFT, only calculations with monomer basis sets are shown). The polarization of the subsystem densities, which can in this case unambiguously be identified from the supermolecular density difference, are perfectly reproduced in the subsystem density difference plots.

One may argue that these results are almost trivial, but note that the systems are neutral and thus correspond to cases where it should not be taken for granted that purely



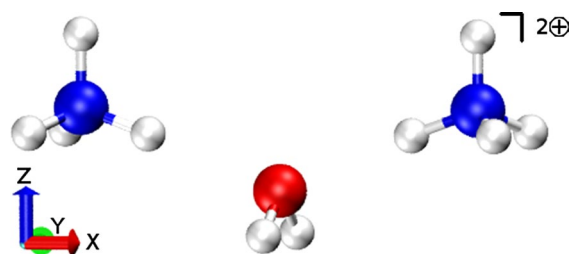
**Fig. 4** Density differences w.r.t. the densities of the isolated monomers for an HCN dimer at three times the equilibrium distance ( $3R_{\text{eq}}$ ). Note that the distance in the picture does not reflect the true distance between the monomers. **a** Supermolecular density, **b** sDFT density, **c/d** subsystem density of the *right/left* monomer from an sDFT calculation. *Blue* positive regions, *Red* negative regions, isovalue:  $10^{-4}$

electrostatic polarization effects are dominant in sDFT calculations with approximate  $T_s^{\text{nad}}[\rho_A, \rho_B]$  functionals [40]. However, due to the low-density overlap in the region between the subsystems, the non-electrostatic components are very small (especially for the larger separation).

These results demonstrate that density changes observed in typical sDFT calculations may in fact be useful for approximate analyses of physical polarization effects of a subsystem by its environment. But it has to be kept in mind that this behavior cannot be considered completely general. Often, the outcome of sDFT calculations will be influenced by practical limitations of the minimization procedure (e.g., monomer basis sets), which can trap the electronic structure at solutions close to the starting point. If large density changes are observed with flexible basis sets, this may be taken as a warning not to consider this a physical polarization effect, but rather as artificial charge-leaking or overpolarization effects [70, 71]. The background is that approximations in  $T_s^{\text{nad}}[\rho_A, \rho_B]$  do not only lead to unphysical redistributions of density between the subsystems for a given, optimum total sDFT density as discussed in detail in Ref. [40]. In addition, and maybe more importantly, deficiencies in approximations for  $T_s^{\text{nad}}[\rho_A, \rho_B]$  can lead to unphysical discrepancies between sDFT and KSDF total electron densities.

#### 5.4 Influence of the starting point for f&t calculations

To elaborate on the role of the starting point for sDFT calculations, we use a simple test case to analyze these effects during a f&t optimization. In general, iterative approaches



**Fig. 5** A water molecule symmetrically pincer by two ammonium ions

can be sensitive to the initialization of the calculation. Either the convergence behavior or, even worse, the final result of the calculation may be affected.

Our test system consists of a water molecule pincer by two ammonium ions (see Fig. 5). It is clear by symmetry that the dipole moment of the water molecule along the axis formed by the nitrogen atoms in the ammonium ions (here: the  $x$ -axis) should vanish. We performed iterative f&t calculations with three slightly different setups. In each case, the initial environmental densities are obtained from isolated calculations on the subsystems. In each setup, we carried out iterations up to full convergence.

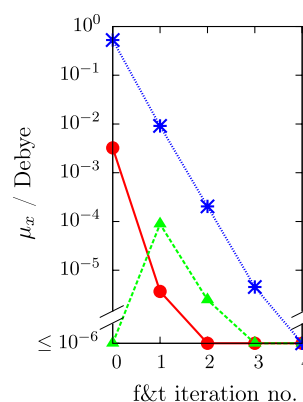
**Setup 1:** For every f&t cycle first, an FDE calculation is performed on one of the ammonium ions, then on the water molecule, and finally on the other ammonium ion. Each updated density is directly used in the calculation of the next subsystem.

**Setup 2:** For every f&t cycle, first an FDE calculation is performed on the water molecule and then on each of the ammonium ions.

**Setup 3:** The same ordering as in Setup 1 is used, but in the first f&t cycle, more precisely in the first two FDE calculations, the second ammonium ion is not part of the environment.

Setup 1 and Setup 3 induce an unphysical polarization in the first f&t cycle due to a nonsymmetric setup, which can actually be seen in the  $x$ -component of the dipole moment of the central water molecule and must be cured in the following cycles. The procedure in Setup 2, by contrast, starts with a symmetric setup. Still, a polarization into the  $x$ -direction can (and actually does) occur in further cycles, because the ammonium ions are not updated synchronously. Each of the setups could at least in principle lead to different results.

To analyze the effects, we monitor the evolution of the  $x$ -component of the dipole moment vector during the f&t cycles, as shown in Fig. 6. Comparing the first two setups, it is obvious that a faster convergence can be achieved with a better choice of the starting point. Despite the unpolarized starting point, also in Setup 2 an intermediate polarization is observable simply due to the asynchronous update of the



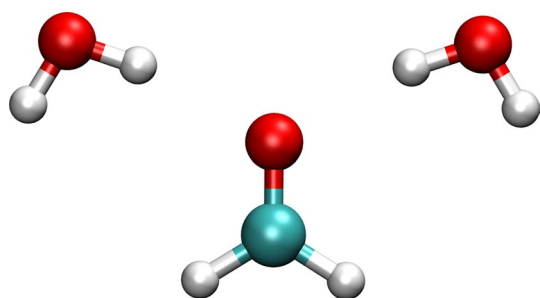
**Fig. 6**  $x$ -Component of the dipole moment for the embedded water molecule as depicted in Fig. 5 with respect to the iteration count in a f&t procedure according to the different setups described in the text. Red dots and solid line: Setup 1; green triangles and dashed line: Setup 2; blue stars and dotted line: Setup 3

ammonium ions. The behavior of Setup 3 is similar to the one of Setup 1, but much more pronounced. Nevertheless, all setups finally converge to the correct solution.

While the setups presented above only change the intermediate behavior during the f&t procedure, it is trivially conceivable that also differences in the final results can be provoked. For the present system, we can for instance compare two additional setups, in which two of the molecules are combined in one subsystem, while the third molecule represents a subsystem on its own. If the two ammonium ions are combined into one fragment, the resulting total density will be strictly symmetric. But if one ammonium ion is combined with the water molecule, then the resulting total density shows an asymmetry due to the asymmetric description of the two water–ammonium contacts.

## 5.5 Orthogonality of states

In Sect. 4.3, we have discussed that orthogonality between the electronic ground state and one or several excited states in WF/DFT embedding should be required for *total* wavefunctions rather than for the wavefunction of the active system only. This poses a formal problem, as the total wavefunctions are unavailable in WF/DFT calculations. Nevertheless, wavefunctions and properties derived from them (like transition moments) from embedded calculations are often interpreted. Hence, it is interesting to explore how severe non-orthogonality effects actually are in WF/DFT calculations with state-specific embedding potentials. Note that we use the term “state-specific” here to indicate that different *environmental* densities are employed in the setup of the embedding potentials. In addition to that, there is always a state specificity due to the dependence of the embedding potential on the *active* subsystem density [21,



**Fig. 7** Formaldehyde flanked by two water molecules in a  $C_{2v}$  symmetry

38]. In the following, we consider the  $\pi \rightarrow \pi^*$  excitation of formaldehyde in the presence of two flanking water molecules (see Fig. 7). For the energetically lower-lying  $n \rightarrow \pi^*$  excitation, the problem vanishes, because the two states involved in the transition are orthogonal by symmetry. This still holds if state-specific embedding potentials are used, at least if they have the same symmetry as the active system.

We make use of the state-interaction module [72] of MOLCAS to calculate the overlap between wavefunctions for different states of the active subsystem. In a FCI calculation, the states are, in a given basis, independent of the composition of the orbitals. Thus, when using state-specific embedding potentials, a non-orthogonality of the states can occur solely due to the environment. We performed embedded FCI calculations in a minimal basis with five f&t iterations to ensure full convergence of the subsystems w.r.t. each other. The environment was polarized w.r.t. the ground-state density or the excited-state density, respectively. The overlap of the resulting states was calculated to be  $4 \times 10^{-4}$  and is obviously very small. One could, however, argue that this result is simply a consequence of the inflexible basis.

Because of this, we looked at the same excitation using a much larger basis (cc-pVQZ) for the active subsystem. As a FCI calculation is unfeasible in this basis, we performed CASSCF calculations with a full-valence active space of 12 electrons in 10 orbitals. We performed state-specific CASSCF calculations with different environments. All embedding potentials applied were self-consistently optimized to the ground- or excited-state wavefunction density. The resulting overlaps are compiled in Table 1.

The state-specific wavefunctions for ground and excited state in the absence of an environment have an overlap of 2.66 %. While this effect is not very large, the influence of the environment on the overlaps is even much smaller than this: With a state-specific embedding potential, the overlap reduces a tiny bit to 2.59 %; if both states are described with a ground-state-like embedding potential (still containing the state specificity caused by the active system's density), the overlap is 3.35 %. An additional test employing

**Table 1** Overlap integrals between the CASSCF(12,10)/cc-pVQZ wave functions of ground and  $\pi \rightarrow \pi^*$  excited states of formaldehyde in the presence or absence of an embedding potential due to two water molecules as depicted in Fig. 7

State (environment)	G (-)	E (-)	G ( $v_{\text{emb}}^G$ )	E ( $v_{\text{emb}}^G$ )	E ( $v_{\text{emb}}^E$ )
G (-)	1.0000				
E (-)	<b>0.0266</b>	1.0000			
G ( $v_{\text{emb}}^G$ )	<i>0.9984</i>	<b>0.0474</b>	1.0000		
E ( $v_{\text{emb}}^G$ )	<b>0.0105</b>	<i>0.9951</i>	<b>0.0335</b>	1.0000	
E ( $v_{\text{emb}}^E$ )	<b>0.0026</b>	<i>0.9938</i>	<b>0.0259</b>	<i>0.9993</i>	1.0000

The orbitals of the states indicated by G (ground state) or E (excited state) have been optimized state specifically (SS) in either no environment (-), a ground-state polarized environment ( $v_{\text{emb}}^G$ ), or an excited-state polarized environment ( $v_{\text{emb}}^E$ ). The emphasis highlights whether the overlap is caused by the state specificity of orbitals (in bold values) (through the SS-CAS procedure), of  $v_{\text{emb}}$  (italic values) (through a change in the environment density), or both (bolditalic values). Note that all embedding results shown here also carry a state dependence through the dependence of  $v_{\text{emb}}$  on the density of the active system

a fixed ground-state density for the  $\rho_A$ -dependent terms in  $v_{\text{emb}}^A$  (i.e., using a linearized embedding potential) and a ground-state-like environment (not shown in Table 1) resulted in the same overlap within the precision shown here. In conclusion, the change in the ground–excited state overlap compared to the overlap for the isolated molecule in state-specific calculations is almost negligible.

The magnitude of the transition-dipole moment in the isolated case changes from 1.796 Debye when this overlap is ignored to 1.749 Debye if the states are used which result from solving the generalized eigenvalue equation in the state-interaction module of MOLCAS. This is a decrease of about 2.6 %. In a fully state-specific environment, the corresponding change is from 2.054 Debye to 2.008 Debye (2.2 % decrease). For a ground-state-like environment, the transition-dipole moment changes from 2.076 Debye to 2.014 Debye (3.0 % decrease). Hence, the change in transition-dipole moment due to orthogonalization appears mainly to be caused by the state-specific CASSCF procedure and is only mildly affected by state specificity introduced through the embedding potential. Overall, we observe only small changes in the transition moments. This may change, however, for different types of excitations, e.g., transitions with large charge-transfer character.

## 6 Conclusions and outlook

Wavefunction/DFT embedding represents a very active field of research, both concerning fundamental theoretical issues and practical implementations of this general strategy. We have outlined here that different points of view

can be adopted—either a pragmatic hybrid-method point of view, which for the active system shows some similarities to a perturbation-theory approach, or a fundamental DFT point of view, in which the wavefunction of the active part just has the role of an auxiliary object to represent the total electron density. We have addressed the issue of non-orthogonality between different states of the active system if consistent, state-specific embedding potentials are employed. This problem is only technically similar to the problem of non-orthogonal states in state-specific multi-configuration SCF calculations. But fundamentally speaking, orthogonality can only be required for total states in WF/DFT embedding calculations.

In our discussion about excited states in WF/DFT embedding, it has to be kept in mind that, in the hybrid-method point of view, only excited states localized in the active system (described by wavefunction methods) can be accessed. In the strict DFT perspective, this depends on how technically the excited states are approached. If they are calculated, as in Refs. [21, 22, 73] as differences of energies of specific states, one is effectively relying on the Perdew–Levy stationarity principle [37] for excited states as generalized by Khait and Hoffmann [36], which is weak because it does not cover all states. Furthermore, practical searches for stationary states have so far only considered excitations localized in the WF system, while it should in principle also be possible (though more difficult) to find excitations localized in the environment. Even inter-subsystem charge-transfer excitations may be accessible with a pragmatic strategy similar to the one in Ref. [50]. But it seems that resonance interactions leading to delocalized excitations over subsystem boundaries will be hard to integrate into such a formalism. They may be easier to describe in a response-theory framework, similar to the subsystem TDDFT formulation in Ref. [74]. Such a response framework has been formulated in the context of WF/DFT methods by Höfener et al. [75].

As another issue, we discussed possible reservations concerning the term “polarization” in the context of WF/DFT and subsystem DFT. As shown in previous work, there is an ambiguity in the subsystem densities and electrostatic moments [13, 20, 35, 46]. But there is no ambiguity in the total sDFT density, so that one may very well interpret total density changes compared to isolated systems as polarization effects. These results can also directly be compared to supermolecular KSDFT calculations or to experimental data (if available). The change in the density of a certain subsystem, by contrast, is ill-defined in FDE-based embedding schemes [40]. Approximations in  $T_s^{\text{nad}}[\rho_A, \rho_B]$  may, in principle, lead to a nonphysical partitioning of the total density into subsystem contributions. But more severely, they can lead to nonphysical total electron densities if the variational conditions allow. This always has to

be kept in mind when interpreting electron densities from density-based embedding and partitioning methods. In practice, however, when one starts from meaningful (usually isolated) fragment densities and avoids overpolarization effects either pragmatically through small basis sets or through special correction terms in the nonadditive kinetic energy potentials [70, 71] if necessary, subsystem densities and their changes may still often be useful for interpretation purposes.

**Acknowledgments** Funding This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center SFB 858.

#### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Libisch F, Huang C, Carter EA (2014) *Acc Chem Res* 47:2768
2. Gomes ASP, Jacob CR (2012) *Annu Rep Prog Chem Sect C Phys Chem* 108:222
3. Senn HM, Thiel W (2009) *Angew Chem Int Ed* 48:1198
4. Chung LW, Hirao H, Li X, Morokuma K (2012) *WIREs Comput Mol Sci* 2:327
5. Jacob CR, Neugebauer J (2014) *WIREs Comput Mol Sci* 4:325
6. Neugebauer J (2009) *ChemPhysChem* 10:3148
7. Wesolowski TA, Warshel A (1993) *J Phys Chem* 97:8050
8. Elliott P, Cohen MH, Wasserman A, Burke K (2009) *J Chem Theory Comput* 5:827
9. Elliott P, Burke K, Cohen MH, Wasserman A (2010) *Phys Rev A* 82:024501
10. Senatore G, Subbaswamy KR (1986) *Phys Rev B* 34:5754
11. Cortona P (1991) *Phys Rev B* 44:8454
12. Wesolowski TA, Weber J (1996) *Chem Phys Lett* 248:71
13. Wesolowski TA (2006) One-electron equations for embedded electron density: challenge for theory and practical payoffs in multi-level modeling of complex polyatomic systems. In: Leszczynski J (ed) *Computational chemistry: reviews of current trends*, vol 10. World Scientific, Singapore, pp 1–82
14. Wesolowski TA, Shedge S, Zhou X (2015) *Chem Rev* 115:5891
15. Govind N, Wang YA, da Silva AJR, Carter EA (1998) *Chem Phys Lett* 295:129
16. Wesolowski TA (2008) *Phys Rev A* 77:012504
17. Parr RG, Yang W (1989) *Density-functional theory of atoms and molecules*. Oxford University Press, New York
18. Svensson M et al (1996) *J Phys Chem* 100:19357
19. Govind N, Wang YA, Carter EA (1999) *J Chem Phys* 110:7677
20. Huang C, Pavone M, Carter EA (2011) *J Chem Phys* 134:154110
21. Daday C, König C, Valsson O, Neugebauer J, Filippi C (2013) *J Chem Theory Comput* 9:2355
22. Klüner T, Govind N, Wang YA, Carter EA (2002) *J Chem Phys* 116:42
23. We would also like to refer the reader to the discussion in Refs. [76] and [77] at this point
24. Gomes ASP, Jacob CR, Visscher L (2008) *Phys Chem Chem Phys* 10:5353
25. Gräfenstein J, Cremer D (2000) *Chem Phys Lett* 316:569
26. Weimer M, Della Sala F, Grling A (2008) *J Chem Phys* 128:144109

27. Kurzweil Y, Lawler KV, Head-Gordon M (2009) *Mol Phys* 107:2103
28. Wesolowski TA (2004) *J Am Chem Soc* 126:11444
29. Knizia G, Chan GK-L (2013) *J Chem Theory Comput* 9:1428
30. Roncero O et al (2008) *J Chem Phys* 129:184104
31. Fux S, Jacob ChR, Neugebauer J, Visscher L, Reiher M (2010) *J Chem Phys* 132:164101
32. Goodpaster JD, Ananth N, Manby FR, Miller TF III (2010) *J Chem Phys* 133:084103
33. Manby FR, Stella M, Goodpaster JD, Miller TF III (2012) *J Chem Theory Comput* 8:2564
34. Bernard YA, Dułak M, Kamiński JW, Wesolowski TA (2008) *J Phys A Math Theor* 41:055302
35. Wesolowski TA, Savin A (2013) In: Recent progress in orbital-free density functional theory, vol 6, chap 9. World Scientific, pp 275–295
36. Khait YG, Hoffmann MR (2010) *J Chem Phys* 133:044107
37. Perdew JP, Levy M (1985) *Phys Rev B* 31:6264
38. Wesolowski TA (2014) *J Chem Phys* 140:18A530
39. Zbiri M, Atanasov M, Daul C, Garcia-Lastra JM, Wesolowski TA (2004) *Chem Phys Lett* 397:441
40. Humbert-Droz M, Zhou X, Shedge SV, Wesolowski TA (2013) *Theor Chem Acc* 132:1405
41. Kanan DK, Sharifzadeh S, Carter EA (2012) *Chem Phys Lett* 519–520:18
42. Casida ME, Wesolowski TA (2004) *Int J Quantum Chem* 96:577
43. Neugebauer J, Curutchet C, Muñoz-Losa A, Mennucci B (2010) *J Chem Theory Comput* 6:1843
44. Wesolowski T, Warshel A (1994) *J Phys Chem* 98:5183
45. Jacob ChR, Jensen L, Neugebauer J, Visscher L (2006) *Phys Chem Chem Phys* 8:2349
46. de Silva P, Wesolowski TA (2012) *J Chem Phys* 137:094110
47. Jacob CR, Wesolowski TA, Visscher L (2005) *J Chem Phys* 123:174104
48. Pavanello M, Neugebauer J (2011) *J Chem Phys* 135:234103
49. Pavanello M, Voorhis TV, Visscher L, Neugebauer J (2013) *J Chem Phys* 138:054101
50. Solovyeva A, Pavanello M, Neugebauer J (2014) *J Chem Phys* 140:164103
51. Hong G, Rosta E, Warshel A (2006) *J Phys Chem B* 110:19570
52. Xiang Y, Warshel A (2008) *J Phys Chem B* 112:1007
53. Amsterdam density functional program. Theoretical chemistry. Vrije Universiteit, Amsterdam. <http://www.scm.com>
54. Te Velde G et al (2001) *J Comput Chem* 22:931
55. Van Lenthe E, Baerends EJ (2003) *J Comput Chem* 24:1142
56. Perdew JP et al (1992) *Phys Rev B* 46:6671
57. Jacob CR, Neugebauer J, Visscher L (2008) *J Comput Chem* 29:1011
58. Lembarki A, Chermette H (1994) *Phys Rev A* 50:5328
59. Aquilante F et al (2010) *J Comput Chem* 31:224
60. Dresselhaus T et al (2015) *J Chem Phys* 142:044111
61. Jacob CR et al (2011) *J Comput Chem* 32:2328
62. Dułak M, Kamiński JW, Wesolowski TA (2009) *Int J Quantum Chem* 109:1886
63. Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C (1989) *Chem Phys Lett* 162:165
64. Becke AD (1988) *Phys Rev A* 38:3098
65. Perdew JP (1986) *Phys Rev B* 33:8822
66. Weigend F, Ahlrichs R (2005) *Phys Chem Chem Phys* 7:3297
67. Bader RFW (1985) *Acc Chem Res* 18:9
68. Jurecka P, Sponer J, Cerny J, Hobza P (2006) *Phys Chem Chem Phys* 8:1985
69. Sedlak R et al (2013) *J Chem Theory Comput* 9:3364
70. Jacob CR, Beyhan SM, Visscher L (2007) *J Chem Phys* 126:234116
71. Lastra JMG, Kaminski JW, Wesolowski TA (2008) *J Chem Phys* 129:074107
72. Malmqvist P-Å, Roos BO (1989) *Chem Phys Lett* 155:189
73. Daday C, König C, Neugebauer J, Filippi C (2014) *ChemPhysChem* 15:3205
74. Neugebauer J (2007) *J Chem Phys* 126:134116
75. Höfener S, Gomes ASP, Visscher L (2012) *J Chem Phys* 136:044104
76. Wesolowski TA (2002) *Phys Rev Lett* 88:209701
77. Klüner T, Govind N, Wang YA, Carter EA (2002) *Phys Rev Lett* 88:209702