JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

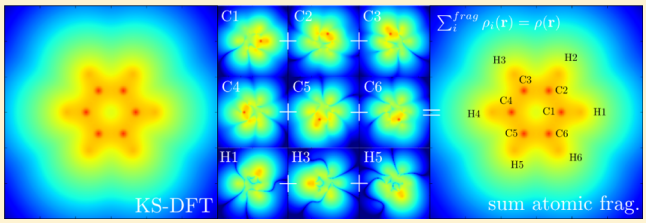
# Frozen Density Embedding with External Orthogonality in Delocalized Covalent Systems

Dhabih V. Chulhai and Lasse Jensen*

Department of Chemistry, The Pennsylvania State University, 104 Chemistry Building, University Park, Pennsylvania 16802, United States

Supporting Information

**ABSTRACT:** Frozen density embedding (FDE) has become a popular subsystem density functional theory (DFT) method for systems with weakly overlapping charge densities. The failure of this method for strongly interacting and covalent systems is due to the approximate kinetic energy density functional (KEDF), although the need for approximate KEDFs may be eliminated if each subsystem's Kohn−Sham (KS) orbitals are orthogonal to the other, termed external orthogonality (EO). We present an implementation of EO into the FDE framework within the Amsterdam density functional program package, using the level-shift projection operator method. We generalize this method to remove the need for orbital localization schemes and to include multiple subsystems, and we show that the exact KS-DFT energies and densities may be reproduced through iterative freeze-and-thaw cycles for a number of systems, including a charge delocalized benzene molecule starting from atomic subsystems. Finally, we examine the possibility of a truncated basis for systems with and without charge delocalization, and found that subsystems require a basis that allows them to correctly describe the supermolecular delocalized orbitals.




## 1. INTRODUCTION

A combination of the calculation expense of large systems and the fact that most chemistry is centered around a smaller subsystem has led to the development of subsystem methods.[1] In these methods, each subsystem may be treated either with the same or different theoretical descriptions. Among the most common subsystem approaches are combined quantum mechical/molecular mechanics methods,[2−4] which enable one to focus on the region of interest using quantum mechanics while having an approximate classical description of the environment. Another approach that has gained popularity recently is subsystem density functional theory (subsystem DFT),[5−8] in which each subsystem is treated using Kohn−Sham (KS) DFT.

Subsystem DFT allows for intuitive partitioning of the supermolecular system via the real-space electron density $\rho(\mathbf{r})$. This supermolecular system density is divided into the region of interest (the active/embedded subsystem, hereafter, $\rho_I(\mathbf{r})$) and the environment (hereafter, $\rho_{II}(\mathbf{r})$). The total energy functional is then minimized under the constraint of a fixed number of electrons for each subsystem. The most common implementation of subsystem DFT is frozen density embedding (FDE),[9] in which one subsystem ($\rho_{II}(\mathbf{r})$) is kept frozen while the total energy is minimized with respect to changes in the other subsystem density ($\rho_I(\mathbf{r})$). Minimization of total energy with respect to the supermolecular density is achieved through freeze-and-thaw cycles, where the roles of subsystems I and II are iteratively interchanged.

FDE is exact in principle, although approximations for the exchange-correlation (XC) functional and the kinetic energy density functional (KEDF) have to be made for practical calculations. In particular, the KEDF contributes to the embedding potential via the nonadditive kinetic potential (NAKP) defined as

$$\nu_T[\rho_I, \rho_{II}; \mathbf{r}] = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} - \frac{\delta T_s[\rho_I]}{\delta \rho_I(\mathbf{r})} \quad (1)$$

where $T_s$ is the KEDF. Available approximations to the KEDF are acceptable only for FDE subsystems with weakly overlapping densities, but they break down in the region of strongly overlapping densities and subsystems with covalent character.[10−12] These failures are generally attributed to the nonexactness of available KEDFs, although it has been shown that intersubsystem KS orbitals' orthogonality is needed to ensure that the density sum relationship of subsystems ($\rho(\mathbf{r}) = \rho_I(\mathbf{r}) + \rho_{II}(\mathbf{r})$) is upheld.[13] Therefore, even for the exact KEDF, the resulting NAKP may still lead to incorrect results unless subsystem KS orbitals are both internally (with respect to orbitals of the same subsystem) and externally (with respect to orbitals of the other subsystem) orthogonal. In addition, external orthogonality (EO) ensures that $\nu_T = 0$ and therefore removes the dependence on approximate KEDFs.[14]

Methods that enforce EO between subsystems have been in use for decades, such as in frozen-core approximations,[15] the

Received: March 28, 2015





DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Phillips−Kleinman pseudopotential approach,[16] and the methods of Stoll and co-workers,[17] Mata and co-workers,[18] and Henderson.[19] Manby, Miller, and co-workers recently introduced EO into subsystem DFT and wave function theory in DFT embedding through the use of a level-shift projection operator.[20−22] This method requires that the supermolecular KS-DFT results are known *a priori* and requires further MO localization schemes. For both covalent and noncovalent systems, they showed that this method reproduces the supermolecular KS-DFT energies exactly in the supermolecular basis[20,22] and accurately in a truncated basis.[21] Hoffmann and co-workers implemented a method where EO is included as a constraint to the coupled KS Lagrangian,[23] which consequently removes the need for an initial supermolecular calculation. For noncovalent systems of varying interaction strengths, they showed that the supermolecular KS-DFT energies, densities, and potential energy curves are exactly reproduced by their subsystem formalism in the supermolecular basis and are well-approximated in an extended monomer basis.

In this article, we present a flexible implementation of EO into the FDE framework[24,25] in the Amsterdam density functional (ADF) program suite.[26−28] We generalize the Miller and co-workers projection operator method[20,21] to include any starting subsystem KS orbitals represented in any basis set. This allows one, in the spirit of conventional FDE, to start with monomer subsystem KS orbitals and iterate through successive freeze-and-thaw cycles to achieve converged results, similar to that of the constrained Lagrangian approach.[23] We show that this implementation exactly reproduces the KS-DFT energies and densities for a number of systems of varying density overlaps. This projection operator method is more easily extended to describe multiple subsystems, and we show that the exact KS-DFT density of a charge delocalized benzene molecule can be obtained from isolated atomic starting subsystems. We also examine the effects of basis set truncation for a number of covalent systems, including systems with large charge delocalization, in order to assess the viability of a truncated basis in subsystem DFT for such systems.

## 2. THEORY

The FDE Kohn−Sham equation (sometimes called the Kohn−Sham equation with constrained electron density[9]) is given as

$$\left[-\frac{\nabla^2}{2} + \nu_{\mathrm{eff}}^{\mathrm{KS}}[\rho_{\mathrm{I}}; \mathbf{r}] + \nu_{\mathrm{eff}}^{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}; \mathbf{r}]\right]\phi_i^{\mathrm{I}}(\mathbf{r}) = \epsilon_i\phi_i^{\mathrm{I}}(\mathbf{r}); \quad i = 1, ..., N_{\mathrm{I}} \tag{2}$$

where $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$ are the embedded (I) and frozen (II) subsystem densities (with $N_{\mathrm{I}}$ and $N_{\mathrm{II}}$ numbers of electrons), respectively, $\phi^{\mathrm{I}}$ are the KS orbitals for subsystem I, $\nu_{\mathrm{eff}}^{\mathrm{KS}}$ is the KS effective potential for subsystem I, and $\nu_{\mathrm{eff}}^{\mathrm{emb}}$ is the effective embedding potential due to the frozen $\rho_{\mathrm{II}}$, written as

$$\nu_{\mathrm{eff}}^{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}; \mathbf{r}] = \nu_{\mathrm{II}}^{\mathrm{nuc}}(\mathbf{r}) + \int \frac{\rho_{\mathrm{II}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}\mathbf{r}' + V_{\mathrm{XC}}^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}; \mathbf{r}] + \nu_{\mathrm{T}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}; \mathbf{r}] \tag{3}$$

where the first and second terms on the right-hand side are the nuclear and electron Coulombic potentials, respectively, $V_{\mathrm{XC}}^{\mathrm{nadd}}$ is the nonadditive contribution to the XC potential, and $\nu_{\mathrm{T}}$ is the NAKP defined eariler.

The NAKP is present because the KS orbitals of subsystem II are not necessarily orthogonal to those of subsystem I. In the case where these subsystems are mutually orthogonal, then $\nu_{\mathrm{T}} = 0$ and we do not require approximate KEDFs. In this work, we enforce EO by implementing the level-shift projection operator introduced in ref 20. We expand on this method by allowing for any starting subsystem MOs, any level of basis set overlap between subsystems, and iterative freeze-and-thaw cycles. This removes the need for a starting supermolecular calculation and MO localization. We modify the Fock matrix of the embedded system in the AO basis using the projection operator as

$$f_{\alpha\beta}^{\mathrm{EO}} = f_{\alpha\beta}^{\mathrm{FDE}(\nu_{\mathrm{T}}=0)} + \mu P_{\alpha\beta}^{\mathrm{II}} \tag{4}$$

where $\alpha$ and $\beta$ are the AO indices of subsystem I. The first term on the right-hand side is the conventional FDE Fock matrix without an NAKP contribution, and $\mu\mathbf{P}^{\mathrm{II}}$ is a term that raises the energy of the $i$th orbital in subsystem II to $\epsilon_i^{\mathrm{II}} + \mu$, where $\mu$ is a scaling parameter and $\mathbf{P}^{\mathrm{II}}$ is the projection operator of subsystem II. The orbitals are forced to be perfectly orthogonal in the limit $\mu \to \infty$; however, a large number ($10^4 < \mu < 10^8$) is usually used in practical applications. The projection operator is defined as

$$\begin{aligned} P_{\alpha\beta}^{\mathrm{II}} &= \langle\chi_\alpha^{\mathrm{I}}|\left\{\sum_{i\in\mathrm{II}} |\phi_i^{\mathrm{II}}\rangle\langle\phi_i^{\mathrm{II}}|\right\}|\chi_\beta^{\mathrm{I}}\rangle \\ &= \sum_{\tau,\upsilon} \langle\chi_\alpha^{\mathrm{I}}|\chi_\tau^{\mathrm{II}}\rangle\gamma_{\tau\upsilon}^{\mathrm{II}}\langle\chi_\upsilon^{\mathrm{II}}|\chi_\beta^{\mathrm{I}}\rangle \\ &= [\mathbf{S}^{\mathrm{I,II}}\gamma^{\mathrm{II}}\mathbf{S}^{\mathrm{II,I}}]_{\alpha\beta} \end{aligned} \tag{5}$$

In this equation, $\{\chi^{\mathrm{I}}\}$ and $\{\chi^{\mathrm{II}}\}$ are the set of AOs describing subsystems I (with AO indices $\alpha$ and $\beta$) and II (with AO indices $\tau$ and $\upsilon$), respectively, $\{\phi^{\mathrm{II}}\}$ is the set of KS orbitals of subsystem II (with density matrix $\gamma^{\mathrm{II}}$), and $\mathbf{S}^{\mathrm{I,II}}/\mathbf{S}^{\mathrm{II,I}}$ are the overlap matrices between the AOs of the two subsystems. The density matrix $\gamma^{\mathrm{II}}$ is symmetric, and $(\mathbf{S}^{\mathrm{I,II}})^{\mathrm{T}} = \mathbf{S}^{\mathrm{II,I}}$, which ensures that the projection operator, and therefore the modified Fock matrix, is Hermitian. Since $\{\chi^{\mathrm{I}}\}$ and $\{\chi^{\mathrm{II}}\}$ may span any subset of the supermolecular basis set $\{\chi^{\mathrm{S}}\}$, we will refer to the case where $\{\chi^{\mathrm{I}}\} \cap \{\chi^{\mathrm{II}}\} = 0$ as the monomer basis implementation (FDE(EO,m)), $\{\chi^{\mathrm{I}}\} = \{\chi^{\mathrm{II}}\} = \{\chi^{\mathrm{S}}\}$ as the supermolecular basis implementation (FDE(EO,s)), and $\{\chi^{\mathrm{I}}\} \cap \{\chi^{\mathrm{II}}\} = \{\chi^{\mathrm{O}}\} \neq \{\chi^{\mathrm{S}}\}$ as an extended monomer implementation (FDE(EO,e), where $\{\chi^{\mathrm{O}}\} \neq 0$ describes the overlap region).

We also extend this projection operator to describe multiple frozen subsystems, where $\rho_{\mathrm{II}}(\mathbf{r}) = \sum_{i\in\mathrm{II}}\rho_i(\mathbf{r})$. In such a calculation, each frozen subsystem will be composed of its own set of KS orbitals, spanning their own AO basis, and we rewrite eq 4 as

$$f_{\alpha\beta}^{\mathrm{EO}} = f_{\alpha\beta}^{\mathrm{FDE}(\nu_{\mathrm{T}}=0)} + \mu\left(\sum_{i\in\mathrm{II}} P_{\alpha\beta}^{i}\right) \tag{6}$$

This ensures that subsystem A is mutually orthogonal to each frozen subsystem. Note that this does not necessarily ensure orthogonality between different frozen subsystems. Mutual orthogonality between all subsystems is achieved through freeze-and-thaw cycles, ensuring that each subsystem is orthogonal to all others.



DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Energy Errors for FDE(cft) Using the TF NAKP and EO Methods

| system | method | KE/$E_h$ | XC/$E_h$ | Coulomb/$E_h$ | total/$E_h$ |
|---|---|---|---|---|---|
| $H_2O-H_2O$ | TF | −0.005 475 74 | 0.002 046 65 | −0.001 073 68 | −0.004 502 77 |
| | EO | −0.000 000 03 | 0.000 000 01 | 0.000 000 05 | 0.000 000 03 |
| $F-H-F^-$ | TF | 0.031 478 52 | 0.000 197 19 | −0.059 215 52 | −0.027 539 81 |
| | EO | 0.000 000 49 | −0.000 000 06 | −0.000 000 18 | 0.000 000 25 |
| $BH_3NH_3$ | TF | −0.044 359 39 | −0.015 477 93 | 0.042 474 37 | −0.017 362 95 |
| | EO | 0.000 000 20 | −0.000 000 01 | 0.000 000 04 | 0.000 000 23 |
| $C_2H_6$ | TF[a] | −3.811 613 27 | 0.246 893 16 | 4.540 666 11 | 0.975 946 00 |
| | EO | 0.000 000 06 | 0.000 000 01 | 0.000 000 05 | 0.000 000 12 |

[a]These values were obtained after 50 freeze-and-thaw cycles.

Table 2. Integrated Absolute Density Difference ($\Delta^{abs}$, in $e$) for FDE(TF,s,cft), FDE(EO,s,2ft), and FDE(EO,s,cft) Methods

| system | active | TF | TF + cft | EO | EO + 2ft | EO + cft |
|---|---|---|---|---|---|---|
| $H_2O-H_2O$ | *I* | 0.1019 | 0.0371 | 0.1168 | 0.0001 | 0.0000 |
| | *II* | 0.1166 | | 0.0784 | 0.0000 | |
| $F-H-F^-$ | F−H | 0.5736 | 0.1154 | 0.5936 | 0.0021 | 0.0000 |
| | $F^-$ | 0.7174 | | 0.4544 | 0.0021 | |
| $BH_3NH_3$ | $BH_3$ | 0.7433 | 0.4724 | 0.8357 | 0.0110 | 0.0000 |
| | $NH_3$ | 0.6423 | | 0.3693 | 0.0033 | |
| $C_2H_6$ | $CH_3^+$ | 1.7171 | 4.0116[a] | 1.8818 | 0.0711 | 0.0000 |
| | $CH_3^-$ | 1.3109 | 3.9072[a] | 0.7973 | 0.0145 | |

[a]These values were not converged after 50 freeze-and-thaw cycles.

## 3. COMPUTATIONAL DETAILS

All calculations presented in this work were performed using a local version of the Amsterdam density functional (ADF) program package,[26−28] with the Becke−Perdew (BP86) exchange-correlation (XC) potential.[29,30] The QZ4P basis set was used for the water dimer, $FHF^-$, $BH_3NH_3$, ethane, and benzene systems, whereas the TZ2P basis set was used for all other systems. All geometries were first optimized using the BP86 XC functional. The Thomas−Fermi (TF) local density approximation to the NAKP[31,32] was used for all conventional FDE calculations. For all external orthogonality calculations, the subsystems' KS orbitals are orthogonalized using a level-shift parameter of $\mu = 10^6$ $E_h$. This was found to be reasonable since values $>10^7$ $E_h$ generally lead to numerical instabilities and values $<10^2$ $E_h$ were insufficient at achieving external orthogonality.[20] All density manipulations were performed numerically on a cubic grid of 0.1 Å spacing, with the densities calculated using pyADF.[33] Other ADF specifications include the following: integration threshold = 8.0; self-consistent field energy convergence criterion = $10^{-8}$.

The typical workflow using our implementation of EO into the FDE framework (FDE(EO)) within ADF is as follows: (1) Identify the subsystems of the supermolecular system. (2) Determine which atoms constitute the overlap region. (3) Calculate (and save) the KS orbitals of each subsystem, using ghost or empty AOs for atoms that are not a part of this subsystem but are included in the overlap region. (4) Perform an FDE(EO) calculation using the KS orbital information for each starting subsystem.

## 4. RESULTS AND DISCUSSION

**4.1. Exact KS-DFT Results.** We show that the EO method within the FDE framework with converged freeze-and-thaw (cft) cycles exactly reproduces the supermolecular Kohn−Sham results by comparing the energies and densities of four systems that vary in their amount of density overlap. These systems were previously used to explore the effectiveness of a reconstructed NAKP[12] and include hydrogen-bonded systems (water dimer and $FHF^-$), a donor−acceptor system ($BH_3NH_3$), and a covalent system (ethane). In Table 1, we show the energy differences between FDE with converged freeze-and-thaw iterations (using the TF NAKP and EO) and that of the supermolecular KS-DFT calculation for these systems. For these methods, the total XC energy is taken as the sum of the XC energy of the individual subsystems plus that of the nonadditive XC energy ($E_{XC}[\rho_I\rho_{II}] = E_{XC}[\rho_I] + E_{XC}[\rho_{II}] + E_{XC}^{nadd}[\rho_I\rho_{II}]$). By definition, this is the XC energy of the sum of densities $\rho_I + \rho_{II}$. The total noninteracting KE for the EO method is taken as the sum of the noninteracting KE of the individual subsystems ($T_S^{EO}[\rho_I\rho_{II}] = T_S[\{\phi^I\}] + T_S[\{\phi^{II}\}]$), whereas for the TF method, the nonadditive KE is also added ($T_S^{TF}[\rho_I\rho_{II}] = T_S[\{\phi^I\}] + T_S[\{\phi^{II}\}] + T_S^{nadd}[\rho_I\rho_{II}]$). For the TF method, the error in the KE varies from −5.4 $mE_h$ in the water dimer to −3.8 $E_h$ in the case of ethane. This is expected since the TF (and generally most NAKPs) approximation is accurate only in the limit of weakly overlapping densities. This leads to incorrect densities and consequently large errors in the XC and Coulomb energies for systems with a stronger covalent character. However, fortuitous cancellation of errors often results in a total energy error that is always smaller than that of the noninteracting KE. When EO is enforced, all energies are numerically identical (within 0.5 $\mu E_h$) to those of the KS-DFT results. The projection operator DFT-in-DFT embedding was shown to produce sub-$\mu E_h$ energy errors for systems where the subsystems were obtained from localization of the MOs of the supermolecular system,[20] whereas these results show that the same level of accuracy is obtained from isolated starting subsystems using the iterative freeze-and-thaw method.

We compare the density error between the FDE methods and that of the KS-DFT density using the integrated absolute density difference[12,14,34−36] ($\Delta^{abs} = \int|\rho^{KS}(\mathbf{r}) - \rho^{FDE}(\mathbf{r})|\,d\mathbf{r}$) for the four systems in Table 2. We include results for regular FDE and FDE(cft) for both methods, as well as FDE with two freeze-and-thaw cycles (2ft) for the EO method. For the water



DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

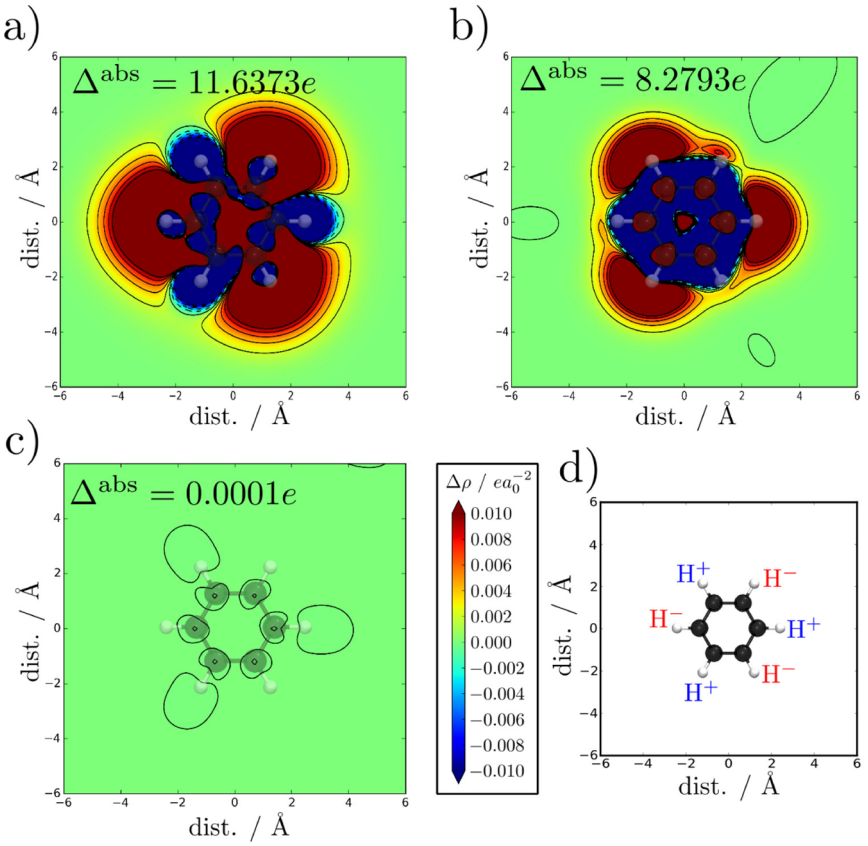



**Figure 1.** Density difference (integrated over the perpendicular axis not shown) for a benzene molecule, using (a) the sum of subsystem densities, (b) FDE(TF,s,30ft), and (c) FDE(EO,s,30ft). The molecular coordinates and location of the $H^-$ and $H^+$ subsystems are shown in (d).

dimer, subsystem I is the water molecule whose hydrogen atom is involved in a hydrogen bond with the oxygen lone pair of subsystem II. The $\Delta^{abs}$ values for the FDE(TF,s) method starting with each subsystem I and II are 0.10 and 0.12 *e*, respectively. These are both acceptable values for $\Delta^{abs}$,[12,36] and we will therefore use $\Delta^{abs} < 0.15$ *e* as our $\Delta^{abs}$ threshold. Our results with the TF NAKP closely resemble those from an earlier study[12] with the PW91k NAKP,[37] and we will therefore not examine them in detail. In summary, the TF NAKP can describe only the water dimer (both regular FDE and FDE(cft) approaches) and $FHF^-$ (only the FDE(cft) approach) systems, but it fails for the $BH_3NH_3$ and ethane systems. In addition, the iterative freeze-and-thaw method failed to converge for the strongly interacting (covalent) ethane system,[10] and the results shown in Table 1 were taken after 50 cycles. Using the bare FDE method with EO, that is, without freeze-and-thaw, we found that the $\Delta^{abs}$ values were comparable to those obtained with the TF NAKP. This is expected since the error is dominated by the frozen subsystem. However, we observe numerically identical densities in the FDE(cft) calculations. We also include the results from the FDE(2ft) method to show that this convergence of density is rapid, requiring an average of 16 iterations for the ethane molecule and 7 iterations for the other systems with a convergence criterion of $10^{-8}$ $E_h$. For all systems, FDE(EO,2ft) was sufficient to yield acceptable density errors with $\Delta^{abs}$ comparable to (and often significantly better than) those of the water dimer with FDE(TF,ctf). As the results indicate, the supermolecular KS-DFT densities are exactly obtained from iterative freeze-and-thaw calculations with EO and yield acceptable density errors for covalent systems in as few as two freeze-and-thaw cycles.

We also show that the EO method allows for multiple subsystems and any arbitrary partitioning of the density by reconstructing the exact KS-DFT density of benzene from atomic subsystems. In Figure 1, we show the density difference integrated over the perpendicular axis ($\Delta\rho = \int(\rho^{KS} - \rho^{FDE})\,dz$, in units of $ea_0^{-2}$) of a benzene molecule for the FDE(TF,s,30ft) and FDE(EO,s,30ft) methods, each starting from atomic subsystem densities. In order to construct closed-shell subsystems, the six H atomic subsystems are instead represented as three $H^-$ and three $H^+$ ions, with their respective positions given in Figure 1d. These positions are reflected as the large positive and negative regions in the difference density in Figure 1a, with a calculated $\Delta^{abs} = 11.6373$ *e*. The TF NAKP improves only slightly upon the sum of subsystems description, with $\Delta^{abs} = 8.2793$ *e* (Figure 1b). This model is not expected to be qualitatively correct for such a system, and it seems to push the electron density toward the carbon and hydrogen nucleii and away from the benzene ring. With the EO method, however, we clearly see (Figure 1c) that the density is identical to that of the KS-DFT density, with a $\Delta^{abs}$ value of 0.0001 *e*. The benzene system shown here was converged only to $\sim 10^{-7}$ $E_h$ after 50 freeze-and-thaw iterations, which may be due to the delocalized nature of the system and the number of subsystems used. These results show the robustness of our implementation of the EO method and that it allows one to arbitrarily (limited to closed-shell subsystems in the current implementation) divide any, even delocalized, supermolecular system.

**4.2. Effects of Basis Set Truncation.** To reduce the cost of the FDE(EO) method, it is therefore desirable to truncate the number of AOs in the overlap region of each subsystem.



DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

We study the effects of different levels of a truncated basis set on a series of alkene compounds, with 6−20 carbon atoms and 1−9 conjugated double bonds. The compounds with one double bond are of the general form $CH_3(CH_2)_nCH{=}CH(CH_2)_nCH_3$ ($n$ = 1, 2, 3, 4, 5), where each subsystem is $CH_3(CH_2)_nCH$. These systems are hereafter referred to as the alkane-like systems. The compounds with multiple conjugated double bonds are of the general forms $CH_2(CH)_nCH_2$ ($n$ = 4, 8, 12, 16) or $CH_3(CH)_nCH_3$ ($n$ = 6, 10, 14, 18), with identical starting subsystems of the form $CH_2(CH)_n$ ($n$ = 2, 4, 6, 8) or $CH_3(CH)_n$ ($n$ = 3, 5, 7, 9), respectively. The optimized geometry of these systems are included in the Supporting Information. We chose these systems since they enable one to cut across a covalent (double) bond to obtain two identical closed-shell starting subsystems. These hydrocarbon chains allow for the exploration of EO in a truncated basis for systems with localized (in the alkane-like systems) and delocalized (in the conjugated alkenes) electrons.

We compare the energy ($|\Delta E| = |E^{\text{KS-DFT}} - E^{\text{EO}}|$), density ($\Delta^{\text{abs}}$), and dipoles ($|\Delta\mu| = |\mu^{\text{KS-DFT}} - \mu^{\text{EO}}|$) to those of the supermolecular KS-DFT results in Figures 2, 3, and S1 in the Supporting Information, respectively. These figures show the results after 5 freeze-and-thaw iterations (5ft); we found that increasing to 10 iterations provided no significant improvements on these properties. In these figures, m denotes the FDE(EO,m) method and e$x$ denotes a level of the extended monomer basis expansion, where $x$ denotes the number of carbon atoms and its associated hydrogen(s) whose AOs are included in the overlap region. We show an example of this truncation naming convention for one subsystem of $C_{16}H_{20}$ in Figure 4. It should be clear from this that e3 denotes the supermolecular basis in the six carbon chains, e4 is the supermolecular basis in the eight carbon chains, and so on.

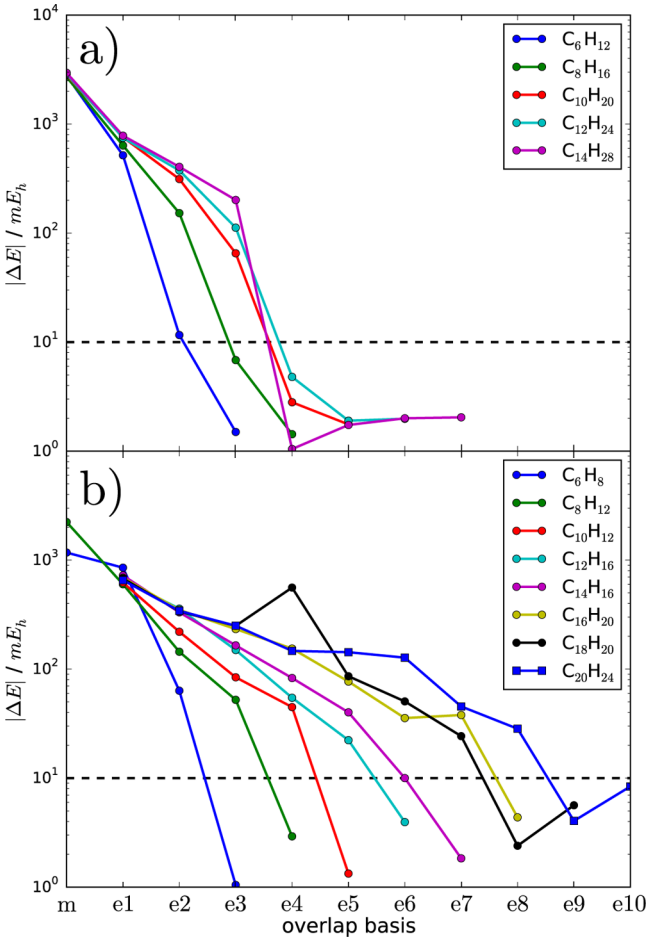



**Figure 2.** $|\Delta E|$ in the FDE(EO,5ft) method with different levels of basis set overlap for (a) alkane-like systems and (b) conjugated alkenes. The dashed line is the energy difference threshold of 10 m$E_h$.

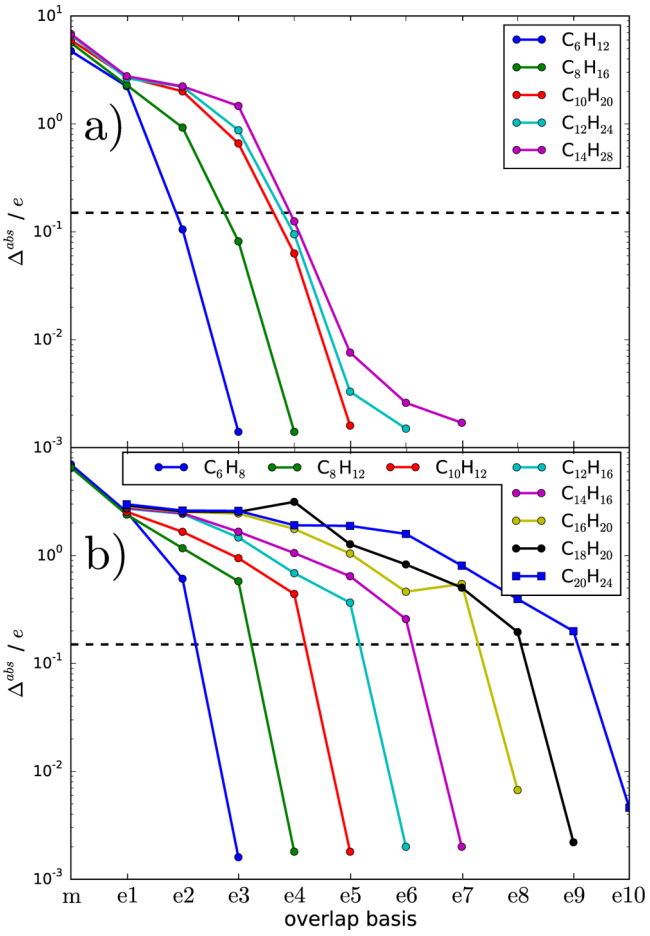



**Figure 3.** $\Delta^{\text{abs}}$ for the FDE(EO,5ft) method with different levels of basis set overlap for (a) alkane-like systems and (b) conjugated alkenes. The dashed line is the $\Delta^{\text{abs}}$ threshold of 0.15 $e$.

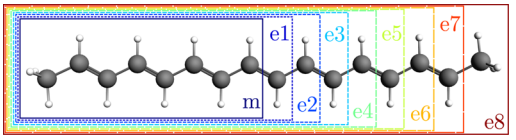



**Figure 4.** Example of the basis set truncation used for one subsystem of $C_{16}H_{20}$.

On the basis of the (acceptable) total energy errors obtained with the FDE(TF,cft) method for the water dimer and $FHF^-$ systems (in Table 1), we will consider values of $|\Delta E| < 10$ m$E_h$ as our energy difference threshold. This is represented as the dashed line in Figure 2. We observe that only the supermolecular basis yields acceptable energies for the conjugated alkenes (with the exception of e8 for $C_{18}H_{20}$ and e9 for $C_{20}H_{24}$), whereas different levels of the extended monomer basis yield acceptable energies for the alkane-like systems. In addition to the supermolecular basis, the e3 basis results in


DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

acceptable energies for $C_8H_{16}$; the e4, for $C_{10}H_{20}$; e4 and e5, for $C_{12}H_{24}$; and e4, e5, and e6, for $C_{14}H_{28}$. These results indicate that an e4 level of truncation (which corresponds to including AOs from all nucleii <5.7 Å away for these systems) is required to obtained acceptable energies in alkane-like systems (without electron delocalization). However, for the delocalized conjugated alkenes, only the supermolecular basis results in acceptable energies.

Similar results are obtained for density difference (Figure 3), with a threshold value of $\Delta^{abs} < 0.15$. Only the supermolecular calculations results in acceptable density errors in the conjugated alkene systems. For the alkane-like systems, a number of truncated basis yielded acceptable densities: the e2 for $C_6H_{12}$; e3 for $C_8H_{16}$; e4 for $C_{10}H_{20}$; e4 and e5 for $C_{12}H_{24}$; and e4, e5, and e6 for $C_{14}H_{28}$. For dipoles (Figure S1), with an acceptable threshold of $|\Delta\mu| < 0.25$ D, the same level of basis set truncation yields acceptable for the alkane-like systems, whereas only the supermolecular basis (and e4 truncation for $C_{10}H_{12}$) yields acceptable dipole errors for the conjugated alkenes. We note that e7 for $C_{16}H_{20}$ and e4 for $C_{18}H_{20}$ appear as outliers in the trend of improving energy, density, and dipole errors with increasing levels of basis set truncation. The reason for this is uncertain and may be due to the quality of the starting subsystem. We point out that most of the conjugated alkenes' starting subsystems were converged to ~0.1 $E_h$ after 100 SCF iterations, yet this does not seem to affect the results of the FDE(EO,5ft) calculations (except for the two truncations mentioned above).

The large errors in some levels of the extended basis truncation may be attributed to the $\mu$-dependence of these truncated systems, as was explored in ref 21, where $|\Delta E|$ changes of 2 orders of magnitude and greater were observed for $\mu$ values ranging between $10^2$ and $10^8$. In that paper, it was found that projecting only the MOs that are localized within the AO space of the active subsystem removes the $\mu$-dependence and subsequently led to a reduction in the energy error of the truncated systems. In Figure 5, we plot the energy differences of the different levels of basis set truncations for the two 10-carbon ($C_{10}H_{20}$ and $C_{10}H_{12}$) systems. We observe no strong dependence of $|\Delta E|$ on the $\mu$-parameter for any of the extended monomer basis truncation, with very little change between $\mu$ values of $10^4$ to $10^8$. The only exception to this is in the monomer basis truncations, for which there are no MO localization counterpart since there are no common AOs describing both subsystems. We will continue to explore the reasons for the large errors seen in some levels of the extended monomer basis truncation, particularly as it relates to the results seen for the conjugated alkenes, in the remainder of this article.

The failure of an FDE(EO) method with an extended monomer truncated basis for the conjugated alkenes may be attributed to three factors: (1) insufficient overlap basis to achieve orthogonality between subsystems, (2) insufficient (mutual) polarization of subsystems, and (3) insufficient AO space for subsystems to accurately reproduce supermolecular MOs. We propose that the extent of subsystem MO orthogonality obtained may be assessed through the value of the overlap energy, defined as

$$E^{ovrlp}[\phi^I, \phi^{II}] = \mu \mathrm{tr}(\gamma^I \mathbf{P}^{II}) \quad (7)$$

This value was shown to be the first-order perturbative correction to the total energy of the system,[20] and we will use it as a measure of the extent of orthogonality attained between subsystem MOs. Mathematically, this value measures

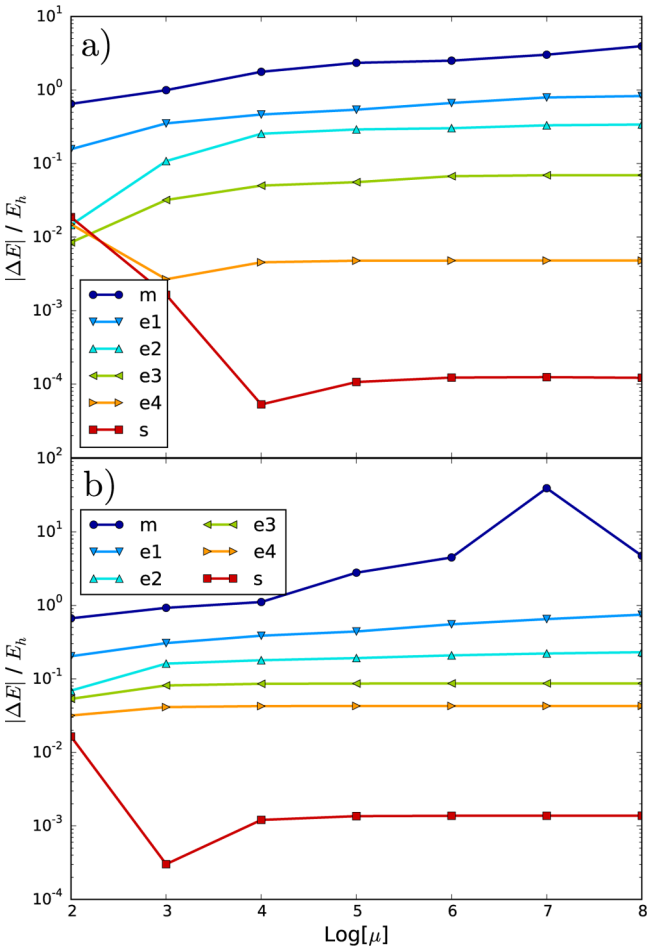



**Figure 5.** Dependence of the $|\Delta E|$ on the $\mu$-parameter for different levels of truncations in (a) $C_{10}H_{20}$ and (b) $C_{10}H_{12}$.

the overlap between the subsystem MOs and is exactly zero for perfectly orthogonal MOs. We explore the overlap energy at successive freeze-and-thaw iterations for an alkane-like system ($C_{14}H_{28}$ in Figure 6a) and a conjugated alkene ($C_{18}H_{20}$ in Figure 6b). In this figure, a half-cycle corresponds to the thawing of the initially frozen subsystem, and an $E^{ovrlp}$ value of $10^{-6}$ $E_h$ is effectively zero with $\mu = 10^6$ $E_h$. In the $C_{14}H_{28}$ system, we found that the e4, e5, and e6 truncations were sufficient at reproducing acceptable energy, density, and dipole errors (Figures 2, 3, and S1), but only the e5 and e6 truncations were suffcient at producing perfectly orthogonal MOs. The e4 truncation was sufficient to enforce orthogonality of one subsystem MOs with respect to the other, but not the other way around. Note that, although the initial subsystems are identical, we choose one subsystem to be the starting active subsystem, which can lead to asymmetrical results. For the $C_{18}H_{20}$ system, the e7 and e8 truncations were sufficient at achieving orthogonality between subsystem MOs. This suggests that a lack of EO is not the reason for large energy, density, and dipole errors observed for this molecule at these truncation levels. We can also rule out insufficient polarization since increasing to 10 freeze-and-thaw cycles did not improve on these errors. Therefore, we can conclude that the inability of the e7 and e8 truncations of this system to reproduce the KS-DFT energy, density, and dipole is due to an inability of the spatially limited subsystem basis at describing the delocalized supermolecular MOs.


DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

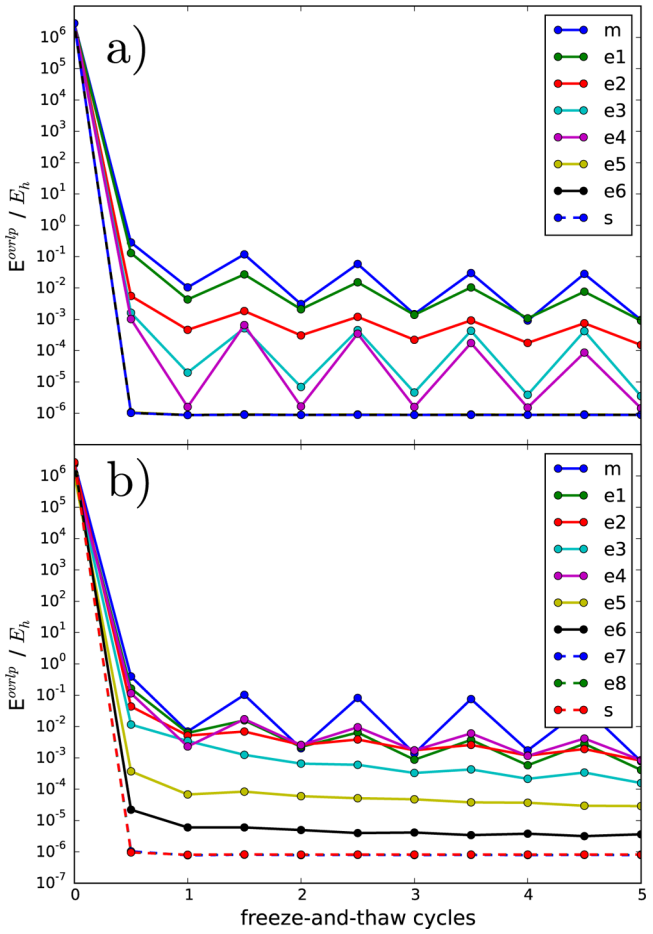



**Figure 6.** $E^{ovrlp}$ at successive freeze-and-thaw cycles for (a) $C_{14}H_{28}$ and (b) $C_{18}H_{20}$.

There are other ways in which one may truncate the basis set of the overlap region while maintaining the spatial distributions of AOs. We examined the effects of reducing the quality of overlapping AOs, from TZ2P to DZP, while maintaining a TZ2P description of remaining AOs. This was done by simply changing the basis type of the ghost atoms in the overlap region. We show these results in Figure 7, along with the results of supermolecular basis and the extended monomer basis that is one level below the supermolecular basis (that is, e2 for $C_6H_8$, e3 for $C_8H_{12}$, e4 for $C_{10}H_{12}$, and so on). This DZP truncation yields acceptable energy and density errors for all systems, and large dipole errors were observed only for the $C_{18}H_{20}$ and $C_{20}H_{24}$ systems. The dipole errors for these alkene chains are particularly sensitive to the chain length because they rely on the addition of large dipole vectors of the individual subsystems to yield the (almost) zero dipole vector of the supermolecular system. These results clearly show that reducing the quality of the overlap basis is a better basis set truncation strategy in delocalized systems rather than just including the set of nearby AOs.

Additionally, the DZP truncation is much more efficient in reducing the number of AOs when compared to that with the extended monomer strategy. The DZP truncation resulted in a 25% reduction of the supermolecular AOs for all systems, whereas the extended monomer basis resulted in an 8% reduction for the $C_{20}H_{24}$ system to a maximum of 20% reduction for the $C_6H_8$ system (see Figure S2 in the Supporting Information). We note that a further reduction to a SZ basis set

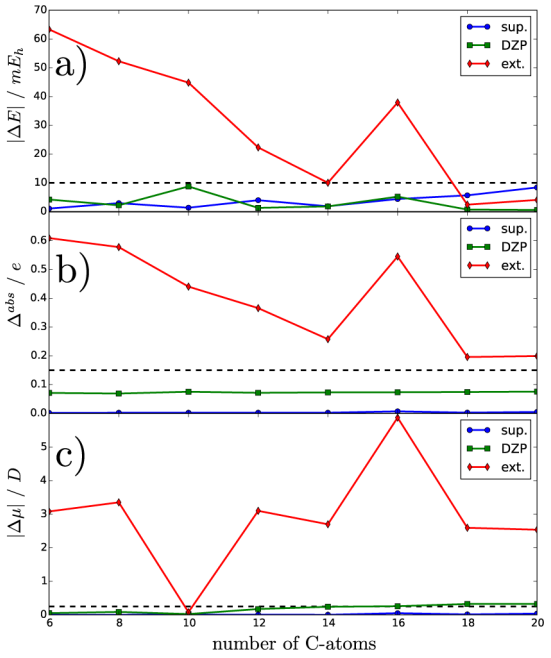



**Figure 7.** Results from the supermolecular basis (sup.), truncated basis with DZP overlap (DZP), and the extended monomer basis that is one level below that of the supermolecular basis (ext.) showing the (a) absolute energy difference, (b) integrated density difference, and (c) absolute dipole difference for the conjugated alkenes.

for the overlap region resulted in errors that were comparable to those with the extended monomer strategy. This reduction in AOs results in a significant reduction in computation time (Figure S3 in the Supporting Information) and is more efficient than the ineffective extended monomer-type truncation. However, we point out that these methods are still more expensive than the supermolecular KS-DFT calculations. The effectiveness of these embedding methods will therefore depend on eliminating the need for iterative freeze-and-thaw cycles and/or increasing the number of basis functions truncated. Olsen and co-workers[38] have implemented a strategy that eliminates the need for mutual polarization through iterative freeze-and-thaw cycles by using a classical polarizable force field. Additionally, one may employ the dual truncation strategy of reducing the number of atoms in the overlap region as well as the quality of these overlap AOs when working with noncovalent systems or covalent systems with localized charge distributions.

## 5. CONCLUSIONS

In this article, we presented an EO implementation into the FDE framework within ADF. This method is a generalized form of the level-shifting projection operator method of Miller and co-workers, extended to include any (and multiple) starting subsystem KS orbitals. We showed that the KS-DFT densities and energies are exactly reproduced with iterative freeze-and-thaw cycles for a number of systems with a range of overlapping density strengths. We also showed that the exact density of a charge delocalized system (benzene) is exactly reconstructed from atomic subsystems. Finally, we examined the limits of a truncated basis for such a method and found that, while including only the nearest AOs works well for alkane-like systems, only the supermolecular basis reproduced the exact supermolecular KS-DFT results in charge delocalized systems.


DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

For such systems, we found that reducing the quality of the overlap basis is a more effective and efficient strategy.

## ■ ASSOCIATED CONTENT

### Supporting Information

All system (and subsystem) geometries as well as data from the truncated basis analyses. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00293.

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: jensen@chem.psu.edu.



### Funding

This work was supported by NSF award no. CHE-1362825. We acknowledge support received from Research Computing and Cyberinfrastructure, a unit of Information Technology Services at Penn State.


### Notes

The authors declare no competing financial interest.

## ■ REFERENCES


(1) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632−672.

(2) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(3) Lin, H.; Truhlar, D. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(4) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(5) Senatore, G.; Subbaswamy, K. R. Density Dependence of the Dielectric Constant of Rare-Gas Crystals. *Phys. Rev. B* **1986**, *34*, 5754−5757.

(6) Johnson, M. D.; Subbaswamy, K. R.; Senatore, G. Hyperpolarizabilities of Alkali Halide Crystals Using the Local-Density Approximation. *Phys. Rev. B* **1987**, *36*, 9202−9211.

(7) Cortona, P. Self-consistently Determined Properties of Solids without Band-Structure Calculations. *Phys. Rev. B* **1991**, *44*, 8454−8458.

(8) Jacob, C. R.; Neugebauer, J. Subsystem Density-Functional Theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 325−362.

(9) Wesolowski, T. A.; Warshel, A. Frozen Density Functional Approach for ab Initio Calculations of Solvated Molecules. *J. Phys. Chem.* **1993**, *97*, 8050−8053.

(10) Götz, A. W.; Beyhan, S. M.; Visscher, L. Performance of Kinetic Energy Functionals for Interaction Energies in a Subsystem Formulation of Density Functional Theory. *J. Chem. Theory Comput.* **2009**, *5*, 3161−3174.

(11) Fux, S.; Kiewisch, K.; Jacob, C. R.; Neugebauer, J.; Reiher, M. Analysis of Electron Density Distributions from Subsystem Density Functional Theory Applied to Coordination Bonds. *Chem. Phys. Lett.* **2008**, *461*, 353−359.

(12) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate Frozen-Density Embedding Potentials as a First Step towards a Subsystem Description of Covalent Bonds. *J. Chem. Phys.* **2010**, *132*, 164101.

(13) Khait, Y. G.; Hoffmann, M. R. On the Orthogonality of Orbitals in Subsystem Kohn−Sham Density Functional Theory. *Annu. Rep. Comput. Chem.* **2012**, *8*, 53−70.

(14) Bernard, Y. A.; Dulak, M.; Kaminski, J. W.; Wesolowski, T. A. The Energy-Differences Based Exact Criterion for Testing Approximations to the Functional for the Kinetic Energy of Non-interacting Electrons. *J. Phys. A: Math. Gen.* **2008**, *41*, 055302.

(15) Lykos, P. G.; Parr, R. G. On the Pi-Electron Approximation and Its Possible Refinement. *J. Chem. Phys.* **1956**, *24*, 1166−1173.

(16) Phillips, J. C.; Kleinman, L. New Method for Calculating Wave Functions in Crystals and Molecules. *Phys. Rev.* **1959**, *116*, 287−294.

(17) Stoll, H.; Paulus, B.; Fulde, P. On the Accuracy of Correlation-Energy Expansions in Terms of Local Increments. *J. Chem. Phys.* **2005**, *123*, 144108.

(18) Mata, R. A.; Werner, H.-J.; Schütz, M. Correlation Regions Within a Localized Molecular Orbital Approach. *J. Chem. Phys.* **2008**, *128*, 144106.

(19) Henderson, T. M. Embedding Wave Function Theory in Density Functional Theory. *J. Chem. Phys.* **2006**, *125*, 014105.

(20) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.

(21) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F. Accurate Basis Set Truncation for Wavefunction Embedding. *J. Chem. Phys.* **2013**, *139*, 024103.

(22) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F. Accurate and Systematically Improvable Density Functional Theory Embedding for Correlated Wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.

(23) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. Density Differences in Embedding Theory with External Orbital Orthogonality. *J. Phys. Chem. A* **2014**, *118*, 9182−9200.

(24) Neugebauer, J.; Jacob, C. R.; Wesolowski, T. A.; Baerends, E. J. An Explicit Quantum Chemical Method for Modeling Large Solvation Shells Applied to Aminocoumarin C151. *J. Phys. Chem. A* **2005**, *109*, 7805−7814.

(25) Jacob, C. R.; Neugebauer, J.; Visscher, L. A Flexible Implementation of Frozen-Density Embedding for Use in Multilevel Simulations. *J. Comput. Chem.* **2008**, *29*, 1011−1018.

(26) Baerends, E.; Ziegler, T.; Autschbach, J.; Bashford, D.; Bérces, A.; Bickelhaupt, F.; Bo, C.; Boerrigter, P.; Cavallo, L.; Chong, D.; Deng, L.; Dickson, R.; Ellis, D.; van Faassen, M.; Fan, L.; Fischer, T.; Guerra, C. F.; Franchini, M.; Ghysels, A.; Giammona, A.; van Gisbergen, S.; Götz, A.; Groeneveld, J.; Gritsenko, O.; Grüning, M.; Gusarov, S.; Harris, F.; van den Hoek, P.; Jacob, C.; Jacobsen, H.; Jensen, L.; Kaminski, J.; van Kessel, G.; Kootstra, F.; Kovalenko, A.; Krykunov, M.; van Lenthe, E.; McCormack, D.; Michalak, A.; Mitoraj, M.; Morton, S.; Neugebauer, J.; Nicu, V.; Noodleman, L.; Osinga, V.; Patchkovskii, S.; Pavanello, M.; Philipsen, P.; Post, D.; Pye, C.; Ravenek, W.; Rodriguez, J.; Ros, P.; Schipper, P.; Schreckenbach, G.; Seldenthuis, J.; Seth, M.; Snijders, J.; Solà, M.; Swart, M.; Swerhone, D.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T.; van Wezenbeek, E.; Wiesenekker, G.; Wolff, S.; Woo, T.; Yakovlev, A. *Amsterdam Density Functional*; Scientific Computing & Modelling: Amsterdam, The Netherlands, 2013. http://www.scm.com.

(27) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931−967.

(28) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391−403.

(29) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymtotic-Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(30) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822−8824.

(31) Thomas, L. H. The Calculation of Atomic Fields. *Math. Proc. Cambridge Philos. Soc.* **1927**, *23*, 542−548.

(32) Fermi, E. A. Statistical Method for the Determination of Some Atomic Properties and the Application of This Method to the Theory of the Periodic System of Elements. *Z. Phys.* **1928**, *48*, 9.

(33) Jacob, C. R.; Beyhan, S. M.; Bulo, R. E.; Gomes, A. S. P.; Götz, A. W.; Kiewisch, K.; Sikkema, J.; Visscher, L. PyADF—A Scripting




DOI: 10.1021/acs.jctc.5b00293
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Framework for Multiscale Quantum Chemistry. *J. Comput. Chem.* **2011**, *32*, 2328−2338.
(34) Jacob, C. R.; Visscher, L. A Subsystem Density-Functional Theory Approach for the Quantum Chemical Treatment of Proteins. *J. Chem. Phys.* **2008**, *128*, 155102.
(35) Beyhan, S. M.; Götz, A. W.; Jacob, C. R.; Visscher, L. The Weak Covalent Bond in NgAuF (Ng=Ar, Kr, Xe): A Challenge for Subsystem Density Functional Theory. *J. Chem. Phys.* **2010**, *132*, 044114.
(36) Kiewisch, K.; Eickerling, G.; Reiher, M.; Neugebauer, J. Topological Analysis of Electron Densities from Kohn−Sham and Subsystem Density Functional Theory. *J. Chem. Phys.* **2008**, *128*, 044114.
(37) Lembarki, A.; Chermette, H. Obtaining a Gradient-Corrected Kinetic-Energy Functional from the Perdew−Wang Exchange Functional. *Phys. Rev. A* **1994**, *50*, 5328−5331.
(38) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344−5355.



DOI: 10.1021/acs.jctc.5b00293
*J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX