JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Fock-Matrix Corrections in Density Functional Theory and Use in Embedded Mean-Field Theory

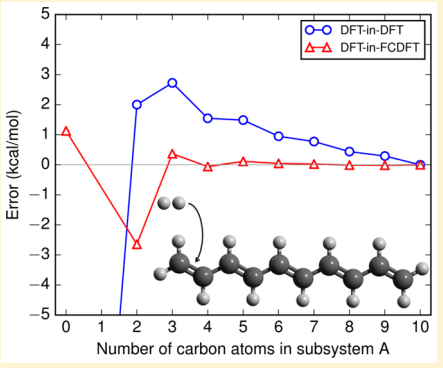
Kaito Miyamoto,[†] Thomas F. Miller III,[‡] and Frederick R. Manby*[,†]

[†]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom

[‡]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

Supporting Information

ABSTRACT: We introduce Fock-corrected density functional theory (FCDFT), a semiempirical minimal-basis method part way between density-functional tight binding (DFTB) and DFT. FCDFT contains DFTB-like Fock-matrix contributions calculated using simple pairwise formulas and Slater−Koster transformations, but it also contains the full Kohn−Sham treatment of Coulombic electrostatics. The resulting method is better suited than either minimal-basis DFT or DFTB for modeling the low-level subsystem in embedded mean-field theory (EMFT), improving upon the former by correcting for basis-set incompleteness and upon the latter by properly accounting for electrostatics. EMFT calculations using DFT-in-FCDFT have much smaller errors in orbital energies, dipole moments, and reaction energies than our previous DFT-in-DFT calculations.



INTRODUCTION

Multiscale embedding theories have made significant contributions to the understanding of chemical processes in a wide variety of large, complex systems. In particular, the celebrated combination of quantum mechanics (QM) and molecular mechanics (MM) in QM/MM revolutionized modeling of catalysis, reactivity, and binding in biomolecular systems.[1−6] The ONIOM method offers the flexibility to combine different QM methods as well as QM and MM,[7−11] and the polarizable continuum model (PCM) provides an important framework for modeling chemical reactions in solution.[12−16]

Despite the successes of these methods, each suffers from potential drawbacks and limitations, arising from the nature of the interaction between system and environment and from the quality of the description of the environment itself. For instance, PCM methods do not capture the structural or dynamical detail of the solvent that are important in some cases, and in ONIOM, the high-level treatment is typically performed without polarization by the environment. Potential problems in QM/MM calculations are numerous, but many have been addressed through corrections and improvements, for example, overpolarization of the QM region by neighboring MM point charges,[17] lack of exchange-repulsion effects on MM atoms,[18] lack of polarizability in the MM environment,[19] and issues with partitioning across covalent bonds.[20,21]

In order to reduce the computational cost for QM calculations while avoiding some of these complications, many researchers have made efforts to investigate quantum embedding theories, where accurate but costly QM methods are embedded in an environment modeled using low-level QM methods.[22−45] Advantages of these theories have been shown by applying them to problems in surface chemistry,[46−49] electrochemistry,[50,51] enzymology,[52] and photochemistry.[53−60]

At present, however, most quantum embedding theories impose the limitation that the number of electrons (and spin state) in subsystems has to be fixed in advance, preventing electron transfer between subsystems or number or spin fluctuations. This limitation may cause problems when considering delocalized systems, such as molecules with conjugated double bonds, surface chemistry on metals and semiconductors, and defects.[61−64] The problem becomes more serious when molecular dynamics simulations are carried out because, for example, changing bond alternations at the boundary between subsystems can result in an apparently reasonable *a priori* choice of particle number and spin state becoming invalid during the trajectory. Attempts have been made to address this issue, for example, in the partition density functional theory of Elliott et al.[28] and the potential functional embedding theory of Huang and Carter.[26] The issue of properly handling particle number and spin fluctuations in open quantum subsystems is fully addressed (at the mean-field level) in the density-matrix embedding theory of Chan and co-workers[37,40,65] and in our recently introduced embedded mean-field theory (EMFT).[61]

EMFT is a quantum embedding formalism based on partitioning of the mean-field one-particle density matrix.[61] In EMFT, a high-level mean-field method is applied in subsystem A, and a lower-level, in subsystem B; the theory is strikingly simple and parameter-free and has no restriction with respect to electron numbers (or spin state): subsystems A and B in EMFT are simply specified by disjoint subsets of atomic orbitals.

Received: July 11, 2016
Published: October 17, 2016

Downloaded via INDIAN INST OF TECHNOLOGY KANPUR on April 22, 2019 at 09:47:52 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications © 2016 American Chemical Society

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

Since EMFT is itself a mean-field theory, properties such as total energy and gradients are obtained by small modifications of existing mean-field theory codes. Numerical assessments show that EMFT has advantages in terms of flexibility in setting boundary conditions and applicability to a variety of chemical systems.

For the greatest efficiency, it is attractive to try minimal basis sets for the description of subsystem B. Although in many cases this works well, we encountered problems when handling processes where charges are formed in subsystem A. This problem may result in part from unphysical charge flow due to the combination of different theories; in particular, the difference of basis sets in the two subsystems appears to play the dominant role. This also causes problems with the description of dipole moments and orbital energies, which are important for predicting electrochemical and photochemical properties.

In early studies of EMFT, we attempted to couple Kohn–Sham density functional theory (DFT)[66,67] with various forms of density functional tight binding (DFTB[68] and its self-consistent charge extension, SCC-DFTB[69]). This turned out not to be successful, and our analysis suggested that the primary cause of the problem was the incompatibility of the treatment of electrostatics in the two approaches.

In this article, we propose a semiempirical electronic structure theory that lies somewhere between DFTB and DFT and is more suitable for EMFT calculations: it improves on minimal-basis DFT through use of Fock-matrix corrections to compensate for basis set deficiencies, and it improves on DFTB through inclusion of Coulombic electrostatics. The key issue is not the central role of electrostatics in the physical systems studied here but the problems that arise from incompatible treatments of electrostatics when different methods are juxtaposed in embedding calculations.

Another prominent attempt to start from a minimal-basis mean-field calculation and add corrections to account for various shortcomings is the HF-3c approach of Grimme et al.[70] This method achieves impressive accuracy in applications on biomolecular systems,[70] organic crystals,[71] and supramolecular host–guest complexes.[72] However, the energy corrections applied in HF-3c do not address the fundamental problem with juxtaposing minimal and more complete basis sets because the corrections to do not change the model Hamiltonian, so they do not address inaccuracies in the electronic density that arise from incompatibility between different levels of theory.

THEORY

Fock-Corrected Density Functional Theory. We will describe our semiempirical electronic structure theory, FCDFT, starting out from the sum of Kohn–Sham orbital eigenvalues

$$\sum_{i\in\mathrm{occ}} \epsilon_i = T_s[\rho] + E_{\mathrm{ext}}[\rho] + 2J[\rho] + \int V_{\mathrm{xc}}[\rho(\mathbf{r})]\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{1}$$

where T_s, E_{ext}, and J are the noninteracting kinetic, electron–nuclear potential, and electronic Coulomb energies, respectively, V_{xc} is the exchange-correlation potential, and ρ is the electronic density. This quantity is related to the Kohn–Sham total energy through double-counting corrections for the Coulomb and exchange-correlation contributions as follows

$$E[\rho] = \sum_{i\in\mathrm{occ}} \epsilon_i - J[\rho] - \int V_{\mathrm{xc}}[\rho(\mathbf{r})]\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[\rho] + E_{\mathrm{nuc}} \tag{2}$$

Here, E_{xc} is the exchange-correlation energy and E_{nuc} is the internuclear repulsion.

The computational cost of evaluating eq 2 can be reduced by using a minimal basis set, but this comes at great loss in accuracy. With FCDFT, we borrow two key components of DFTB to correct for the basis set deficiency: first, we introduce a semiempirical correction term L, to be defined below. Second, we approximate double-counting terms associated with this correction by a sum of pairwise functions of interatomic distances, U_{cor}. The total FCDFT energy for a minimal-basis density $\tilde{\rho}$ is then written

$$E^{\mathrm{FCDFT}}[\tilde{\rho}] = E^{\mathrm{DFT}}[\tilde{\rho}] + L[\tilde{\rho}] + U_{\mathrm{cor}} \tag{3}$$

or in terms of the minimal-basis one-particle density matrix $\tilde{\mathbf{D}}$ as

$$E^{\mathrm{FCDFT}}[\tilde{\mathbf{D}}] = E^{\mathrm{DFT}}[\tilde{\mathbf{D}}] + \mathrm{tr}\,\tilde{\mathbf{D}}\tilde{\mathbf{L}} + U_{\mathrm{cor}} \tag{4}$$

where $\tilde{\mathbf{L}}$ is a minimal-basis semiempirical correction matrix.

We have designed the Fock-correction matrix $\tilde{\mathbf{L}}$ and energy correction term U_{cor} based on standard DFTB technologies.[68,69,73−75] The latter is expressed as

$$U_{\mathrm{cor}} = \sum_{I<J} U_{IJ}(R_{IJ})$$

where I and J index a pair of atoms, R_{IJ} is the distance between them, and U_{IJ} is a short-range potential represented by a fourth-order spline function specific to the pair of atom types.[73,75]

The semiempirical matrix $\tilde{\mathbf{L}}$ has the same structure as the DFTB Hamiltonian,[68,69] i.e., atom-diagonal and off-diagonal blocks given by

$$\tilde{L}_{\mu\nu} = \begin{cases} \delta_{\mu\nu}\epsilon_\mu & \mu,\nu \in I \\ F^{\mathrm{SK}}_{\mu\nu}(\mathbf{R}_{IJ}) & \mu \in I,\ \nu \in J,\ I \neq J \end{cases} \tag{5}$$

where $\mathbf{R}_{IJ}$ is the vector between atoms I and J. The $F^{\mathrm{SK}}_{\mu\nu}(\mathbf{R}_{IJ})$ elements are computed using atom–atom interaction functions with Slater–Koster transformations.[76] The untransformed atom–atom interaction functions (up to second row elements) are given by

$$F_{\mathrm{ss}\sigma}(R) = \alpha\,\exp(-\beta R^\gamma) \tag{6}$$

$$F_{\mathrm{sp}\sigma}(R) = \alpha R\,\exp(-\beta R^\gamma) \tag{7}$$

$$F_{\mathrm{pp}\sigma}(R) = (\delta R^2 - \alpha)\,\exp(-\beta R^\gamma) \tag{8}$$

$$F_{\mathrm{pp}\pi}(R) = \alpha\,\exp(-\beta R^\gamma) \tag{9}$$

with distinct values for the parameters for each pair of atoms and for each interaction type. The polynomial prefactors reflect symmetry properties of the underlying matrix elements so that, for example, the spσ interaction exactly vanishes as the separation between atoms goes to zero. Since these functions are simple closed-form expressions, computational cost for constructing $\tilde{\mathbf{L}}$ is completely negligible. The empirical parameters ϵ in eq 5 are particular to a specific element, whereas α, β, γ, and δ from eqs 6–9 are associated with element pairs. Procedures for fitting these parameters will be discussed in the following section.

Finally, we compare FCDFT with recently proposed semiempirical theories, i.e., DFTB and HF-3c methods.[70] The main difference between FCDFT and DFTB is the handling of electrostatic interactions. FCDFT calculates the Coulomb energy using the charge density obtained using the minimal basis set, whereas DFTB starts from an assumption of neutral

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

atoms in which electrostatic interaction with a nucleus is exactly canceled by the interaction with its surrounding cloud of electronic charge. FCDFT and HF-3c both make corrections to minimal-basis *ab initio* methods, but where HF-3c adds corrections only to the total energy, FCDFT also applies a correction to the Fock matrix. Thus, there is an opportunity to improve not only energies but also the shape of molecular orbitals and orbital eigenvalues.

EMFT Using FCDFT as the Low-Level Method. We derive an energy expression for EMFT that embeds DFT into an environment modeled using FCDFT. The key idea in EMFT is the block partitioning of the density matrix as

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}^{AA} & \mathbf{D}^{AB} \\ \mathbf{D}^{BA} & \mathbf{D}^{BB} \end{pmatrix} \tag{10}$$

where $\mathbf{D}^{AA}$ and $\mathbf{D}^{BB}$ denote density-matrix blocks in subsystems A and B, respectively, and where $\mathbf{D}^{AB}$ and $\mathbf{D}^{BA}$ are blocks that couple the two subsystems. The total energy expression for EMFT is simply defined by

$$E^{\mathrm{EMFT}} = E^{\mathrm{low}}[\mathbf{D}] - E^{\mathrm{low}}[\mathbf{D}^{AA}] + E^{\mathrm{high}}[\mathbf{D}^{AA}] \tag{11}$$

where E^{low} and E^{high} represent the total energy by low- and high-level theories, respectively.

The coupling terms are treated at the FCDFT level of theory, and typically a flexible basis set will be used in subsystem A. Thus, it is necessary to calculate an **L**-matrix correction with one index drawn from a flexible basis. This can be achieved most simply by the standard nonorthogonal projection operator

$$\hat{P} = \sum_{\mu\nu \in A \cup B} |\mu\rangle S^{-1}_{\mu\nu} \langle\nu| \tag{12}$$

where $|\mu\rangle$ and $|\nu\rangle$ are components of the minimal basis set that span the whole system A∪B. Using this projector, the Fock correction for an arbitrary basis set can be written

$$L_{\alpha\beta} = \sum_{\mu\nu\sigma\lambda} S_{\alpha\mu} S^{-1}_{\mu\nu} \tilde{L}_{\nu\sigma} S^{-1}_{\sigma\lambda} S_{\lambda\beta} \tag{13}$$

where α and β index atomic orbitals for the general basis set, whereas μ, ν, σ, λ index functions in the minimal basis.

The total energy expression for the embedded DFT-in-FCDFT theory is given by

$$E^{\text{DFT-in-FCDFT}}[\mathbf{D}] = E^{\mathrm{low}}[\mathbf{D}] + (E^{\mathrm{high}}_{\mathrm{xc}}[\mathbf{D}^{AA}] - E^{\mathrm{low}}_{\mathrm{xc}}[\mathbf{D}^{AA}]) + (\mathrm{tr}\,\mathbf{DL} - \mathrm{tr}\,\mathbf{D}^{AA}\mathbf{L}) + U_{\mathrm{cor}} - U^{AA}_{\mathrm{cor}} \tag{14}$$

where $E^{\mathrm{low}}_{\mathrm{xc}}$ and $E^{\mathrm{high}}_{\mathrm{xc}}$ are the exchange-correlation energies computed by the FCDFT and DFT, respectively, and U^{AA}_{cor} is equivalent to U_{cor} for atoms in subsystem A.

The Fock matrix elements are obtained by taking the derivative of eq 14 with respect to the density-matrix elements. Hybrid functionals can be embedded into an environment modeled using FCDFT in exactly the same manner as for embedded DFT-in-DFT theory.[61] FCDFT and embedded DFT-in-FCDFT theory have been implemented in the development version of Molpro software package.[77,78]

PARAMETER FITTING

In this article, we consider two FCDFT methods for subsystem B, namely, Fock-corrected Hartree (FCH) and Fock-corrected LDA (FCLDA). In the latter, $E^{\mathrm{low}}_{\mathrm{xc}}$ is the LDA exchange-correlation energy.[79,80] In the former, $E^{\mathrm{low}}_{\mathrm{xc}} = 0$, and this is attractive as it involves no numerical quadrature at all in the environment, with the Fock correction **L** and potential U_{cor} taking the entire role of supplying an exchange-correlation contribution. The STO-3G basis set[81,82] is used for the basis functions of FCDFT. Since we use hybrid-DFT as well as pure-DFT for the high level of theory in EMFT, two parameter sets are created for each theory. As high-level DFT theories, PBE/6-31G* and B3LYP/6-311G** are chosen (for the latter, geometry optimizations were performed using B3LYP/6-31G*).[83−89] Here, VWN3 is used as the LDA correlation component of B3LYP, but VWN5 is used elsewhere.[80]

Table 1. Chemical Reactions Used for Fitting Parameters, Drawn Mainly from Refs 73 and 75

chemical reactions
$H_2 \rightarrow 2H$
$C_2H_2 + C_2H_6 \rightarrow 2C_2H_4$
$C_2H_4 + H_2 \rightarrow C_2H_6$
$C_2H_2 + H_2 \rightarrow C_2H_4$
$C_2H_2 + 2H_2 \rightarrow C_2H_6$
$3C_2H_2 \rightarrow C_6H_6$
$C_6H_6 + 6H_2 \rightarrow 3C_2H_6$
$C_2H_4 + 2CH_4 \rightarrow 2C_2H_6$
$C_2H_2 + 4CH_4 \rightarrow 3C_2H_6$
$C_2H_6 + H_2 \rightarrow 2CH_4$
$C_2H_4 + 2H_2 \rightarrow 2CH_4$
$C_2H_2 + 3H_2 \rightarrow 2CH_4$
$C_6H_6 + 9H_2 \rightarrow 6CH_4$

Table 2. Comparison of Errors in Orbital Energies for H_2 and Hydrocarbons from the G2/97 and G3/99 Test Sets[a]

	HOMO/eV		LUMO/eV	
method	MAE	σ	MAE	σ
PBE/STO-3G	1.33	0.29	4.25	2.10
FCH(PBE)/STO-3G	0.58	0.29	1.23	0.78
FCLDA(PBE)/STO-3G	0.19	0.16	0.60	0.41
B3LYP/STO-3G	1.44	0.33	5.40	2.81
FCH(B3LYP)/STO-3G	0.62	0.36	2.54	0.91
FCLDA(B3LYP)/STO-3G	0.16	0.22	1.19	1.24

[a]All errors using PBE are relative to PBE/6-31G*, and those using B3LYP are relative to B3LYP/6-311G**. MAE and σ represent the mean absolute error and standard deviation, respectively.

In this article, we create parameters for hydrocarbons. To determine the parameters, orbital energies, potential energy curves, atomization energies, and reaction energies are used for fitting. Since our aim for this study is to develop low-level methods suitable for use in EMFT, reference orbital energies and potential energy curves are obtained using the high-level DFT theories described above. Reference reaction energies are obtained using CCSD(T)/6-311+G(3df,2p),[90,91] and experimental atomization energies[92] are used as reference values.

We determine parameters by minimizing Δ defined by

$$\Delta = \frac{a\Delta^{\mathrm{OE}} + b\Delta^{\mathrm{PEC}} + c\Delta^{\mathrm{AE}} + d\Delta^{\mathrm{RXN}}}{a + b + c + d} \tag{15}$$

where a, b, c, and d are weight factors and Δ^{OE}, Δ^{PEC}, Δ^{AE}, and Δ^{RXN} are the root-mean-square errors of the orbital energies, potential energy curves, atomization energies per valence electron pair, and reaction energies, respectively. All parameters are determined from properties of just six molecules: H_2, CH_4, C_2H_2,

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

Table 3. Comparison of Errors in Atomization Energies E^{AE} (kcal/mol), Reaction Energies E^{RXN} (kcal/mol), and Dipole Moment Vectors μ (Debye)[a]

	E^{AE}		E^{RXN}		μ	
method	MAE	MAX	MAE	MAX	MAE	MAX
B3LYP/6-311G**	9.8	−22.7	2.3	10.4	0.01	0.09
B3LYP/6-31G*	4.4	−13.4	4.5	19.2	0.03	0.16
PBE/6-31G*	28.6	85.2	6.5	31.5	0.02	0.11
B3LYP/STO-3G	218.0	496.1	19.8	61.5	0.04	0.27
PBE/STO-3G	265.8	570.6	20.0	73.9	0.03	0.24
DFTB2/MIO	56.6	118.8	5.9	15.4	0.04	0.40
DFTB3/3OB	3.3	−11.7	3.8	11.8	0.04	0.38
FCH(PBE)/STO-3G	13.8	−42.9	7.1	20.8	0.15	1.06
FCH(B3LYP)/STO-3G	14.7	−49.9	7.3	22.2	0.21	1.49
FCLDA(PBE)/STO-3G	2.6	−15.2	2.6	8.5	0.03	0.24
FCLDA(B3LYP)/STO-3G	7.4	−49.0	4.5	24.7	0.08	0.58

[a]Errors in E^{AE} are evaluated relative to experiment[92] (with harmonic corrections; see text), errors in E^{RXN} are relative to G3B3 theory,[93] and errors in μ are relative to B3LYP/6-311+G(3df,2p) results. All DFT calculations are carried out using Gaussian 09,[96] and all DFTB calculations are performed using DFTB+.[99,100] Detailed data are provided in the Supporting Information.

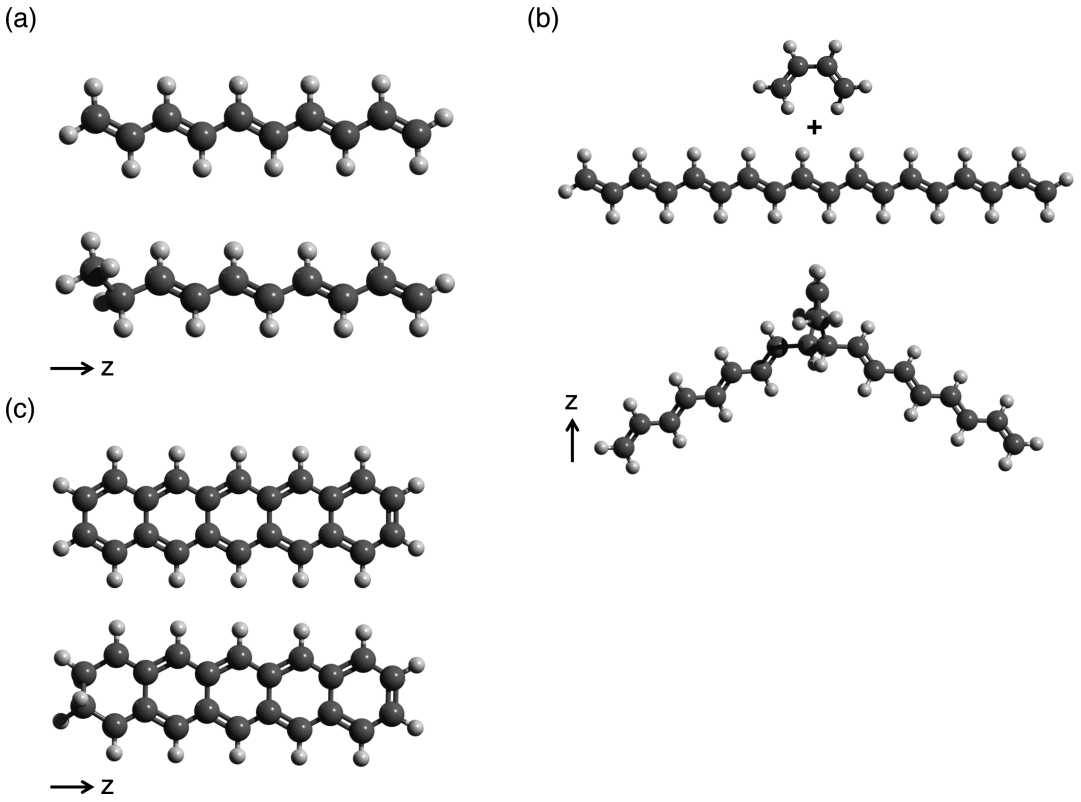


Figure 1. Geometries of reactants and products used in EMFT tests: (a) terminal hydrogenation of decapentaene, (b) Diels−Alder reaction of butadiene and an 18-carbon polyene, and (c) terminal hydrogenation of pentacene. Geometries were optimized using the 6-31G* basis set with the corresponding reference DFT approximation, using Gaussian 09.[96]

C_2H_4, C_2H_6, and C_6H_6. For Δ^{OE}, all valence occupied orbitals were included, and for parametrizing against PBE/6-31G*, the lowest unoccupied molecular orbital was also included. To determine Δ^{PEC}, the H−H bond of H_2, the C−H bond of CH_4, and the C−C bonds of C_2H_2, C_2H_4, and C_2H_6 are elongated on meshes of 8 points from 1.3, 1.8, 2.0, 2.2, and 2.6 bohr, respectively, and with spacing of 0.1 bohr. The point closest to the equilibrium value R_e is replaced by three points with values R_e and $R_e \pm 0.004$ bohr; this effectively adds in a contribution from harmonic frequencies to our objective function. The set of reactions for evaluation of Δ^{RXN} is shown in Table 1.

When FCDFT atomization energies are calculated, we add the carbon atom spin-polarization energy (E^{spin}) as an extra fitting parameter since $\tilde{\mathbf{L}}$ does not take account of this energy contribution. Thus, the FCDFT atomization energy for C_nH_m is obtained by

$$E^{AE} = n(E^{C} + E^{spin}) + mE^{H} - E^{mol} \quad (16)$$

where E^{C}, E^{H}, and E^{mol} are the total energies of the carbon atom, hydrogen atom, and the target molecule. Our reference values for atomization energies were obtained by correcting experimental values for zero-point and enthalpic contributions

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

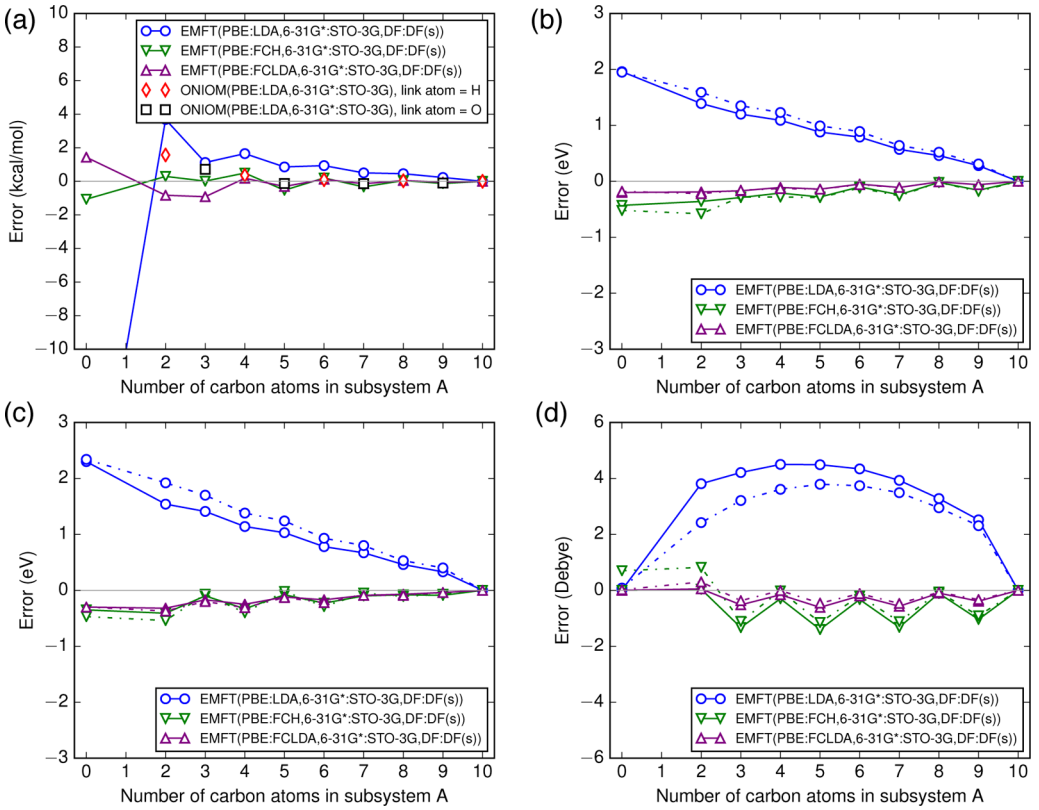


Figure 2. Errors in embedding calculations using PBE/6-31G* in subsystem A for terminal hydrogenation of decapentaene relative to PBE/6-31G* on the whole system. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with PBE/6-31G*) for each panel are (a) −32.8 kcal mol^{-1}; (b) −4.53 eV (reactant), −4.53 eV (product); (c) −2.51 eV (reactant), −2.16 eV (product); and (d) 0.00 D (reactant), −1.01 D (product).

in the harmonic approximation using B3LYP/6-31G*; for this calculation, frequencies are scaled by 0.96.[93]

First, parameters for the hydrogen atom and H−H bond are determined, followed by all other parameters. The pairwise functions that define U_{cor} are parametrized for interatomic distances ranging from a little shorter than the shortest chemically relevant equilibrium bond length to a maximum value taken from a widely used DFTB parametrization.[75] If shorter range values are necessary, these functions can be connected to analytical repulsive potentials, as proposed in ref 73. Weight factors for the orbital energies and potential energy curves were $a = 1$ and $b = 10$. To optimize parameters for FCH, we used $c = 1$ and $d = 0$; for FCLDA, we used the reverse, i.e., $c = 0$ and $d = 1$. E^{spin} for FCLDA was determined only after determining all other parameters and only by using Δ^{AE}, whereas that for FCH was determined simultaneously with the other parameters. The objective functional was minimized using the Broyden−Fletcher−Goldfarb−Shanno (BFGS) algorithm[94] as implemented in the SciPy library.[95] All geometries for FCDFT calculations are obtained by the high-level DFT theories described above using Gaussian 09.[96] All parameters are provided in the Supporting Information.

RESULTS

FCDFT. We examine the performance of FCDFT by comparing basic properties with *ab initio* and semiempirical electronic structure theories. In this section, the high level of theory that is used to determine the parameter set is given in parentheses, so, for example, FCH(PBE) denotes Fock-corrected Hartree with parameters based on PBE/6-31G*.

Although a key point of the present parametrization of FCDFT is to improve on orbital energies to increase compatibility in EMFT calculations, we must also ensure that it has reasonable accuracy for other properties. We first assess equilibrium geometries and harmonic frequencies of small molecules (H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6) using FCH and FCLDA; all derivatives were computed using finite differences. Errors from the high-level DFT theories (PBE/6-31G* or B3LYP/6-311G**) are about the same as those by the corresponding low-level theories (PBE/STO-3G or B3LYP/STO-3G), i.e., mean-absolute errors of less than 0.01 Å for the equilibrium bond lengths and average errors of approximately 5−7% in the frequencies. It is particularly reassuring that the errors in geometries are small. Detailed data are provided in the Supporting Information. We also tested reliability on four small molecules not in the training set (propane, propene, 1,3-butadiene, and toluene) and find similar accuracy. Since we did not include bond angles in our training set, we find that the agreement is slightly less good for these quantities, with average errors of 1−2° using FCLDA and around 4° using FCH.

Next, we examine error distributions in HOMO and LUMO energies for H_2 and 37 closed-shell hydrocarbons from the G2/97[97] and G3/99[98] test sets. Table 2 shows mean-absolute errors (MAE) and standard deviations of errors (σ) in the HOMO and LUMO energies relative to reference DFT results. It is clear that both FCH and FCLDA, parametrized for either PBE/6-31G* or B3LYP/6-311G**, improve considerably over

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

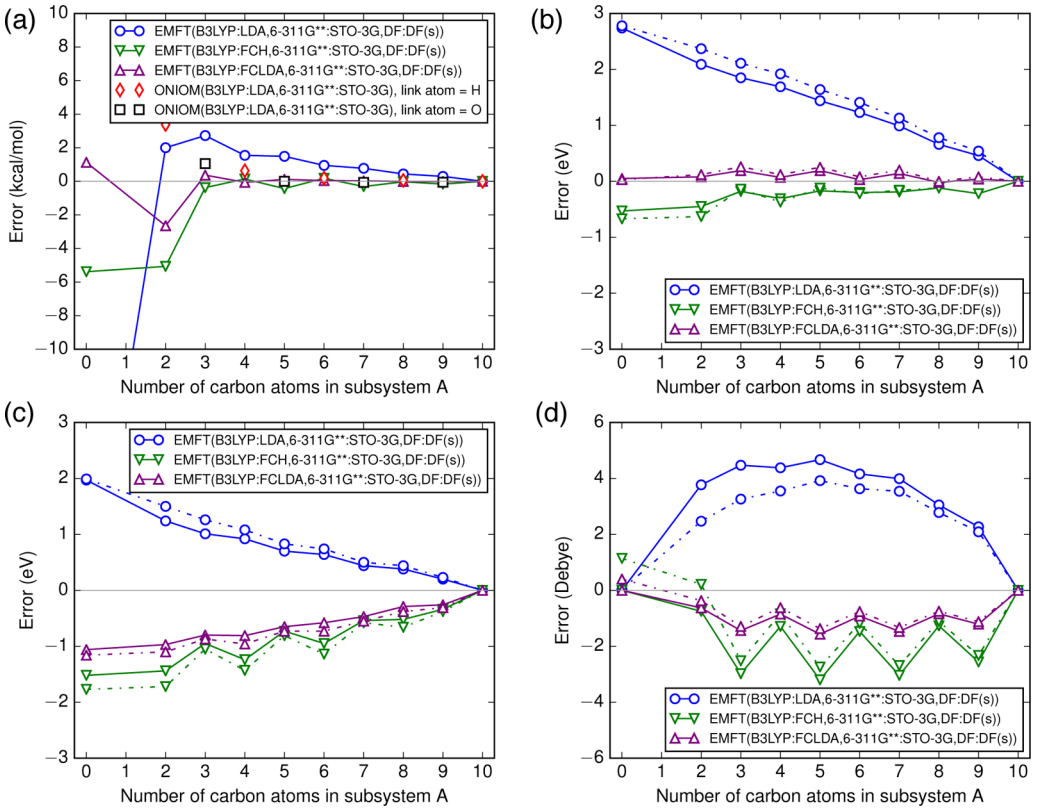


Figure 3. Errors in embedding calculations using B3LYP/6-311G** in subsystem A for terminal hydrogenation of decapentaene relative to B3LYP/6-311G** on the whole system. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with B3LYP/6-311G**) for each panel are (a) −29.4 kcal mol^{-1}; (b) −5.41 eV (reactant), −5.45 eV (product); (c) −2.11 eV (reactant), −1.73 eV (product); and (d) 0.00 D (reactant), −0.93 D (product).

results obtained with either functional in a minimal basis. It can further be seen that the improvement in FCLDA surpasses that of FCH, leading to almost an order of magnitude reduction in errors in the HOMO energy.

The MAEs for the LUMO energies, for both uncorrected and corrected minimal-basis methods, are 2−3 times larger than those of the HOMO energies. Both FCDFT correction schemes reduce the LUMO error very significantly compared to PBE/STO-3G for the PBE parametrization, but the improvement is less pronounced when trying to approximate B3LYP/6-311G**.

We examine the performance of FCDFT for calculating atomization energies E^{AE}, reaction energies E^{RXN}, and dipole moments μ by comparing them with a variety of electronic structure methods. Mean absolute errors (MAE) and maximum errors (MAX) for those properties are summarized in Table 3.

We first look at the errors in the atomization energies (E^{AE}). For DFT calculations, both MAE and MAX are very strongly dependent on the basis set, with enormous errors when the minimal STO-3G set is used. The errors can be reduced significantly by the use of double-ζ basis sets with polarization functions. For the chosen test set, the B3LYP/6-311G** gives 2 times larger values than B3LYP/6-31G* for both MAE and MAX, indicating some cancellation of error in the latter case. Of the two tight-binding methods we tested, DFTB2[69] with the MIO[69] parameter set gives accuracy between DFT in a minimal basis set and in more flexible basis sets, but DFTB3[74] with the 3OB[75] parameter set gives the best results in this list. FCH gives similar errors regardless of parameter sets, and the accuracy of this theory is comparable to that of DFT in double-ζ plus polarization functions; the best FCDFT results have an accuracy similar to DFTB3/3OB. It is worth noting that the errors in atomization energies from FCDFT methods are considerably lower than from the corresponding reference DFT methods used to parametrize them; this is because of the high-accuracy data used in computing the Δ^{AE} and Δ^{RXN} terms in eq 15.

The clear conclusion is that minimal-basis DFT with Fock-matrix corrections produces atomization energies that are as accurate as (or better than) DFT in a flexible basis set and at far lower computational cost.

Next, we compare errors in the reaction energies (E^{RXN}). DFT results show a similar trend as that for E^{AE}; i.e., DFT in the STO-3G basis set gives the largest errors in the table, and the errors are unsurprisingly reduced drastically by the use of 6-31G*. B3LYP/6-311G** gives the best results, in contrast to the results of E^{AE}. Both DFTB methods perform well in predicting E^{RXN}, and the accuracy approaches that of B3LYP/6-311G**. The Fock-corrected methods FCH(PBE), FCH(B3LYP), and FCLDA(B3LYP) give an accuracy similar to PBE/6-31G*, whereas FCLDA(PBE) gives excellent accuracy.

Finally, we compare errors in the dipole moment vectors (μ). In this property, DFT methods give better results than semiempirical theories regardless of the basis. The accuracies of DFTB2/MIO and DFTB3/3OB are almost the same, and MAX values are around 2 times larger than DFT in the minimal basis set. Unfortunately, the FCH errors are much larger than those of the other methods. However, the addition of the LDA

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

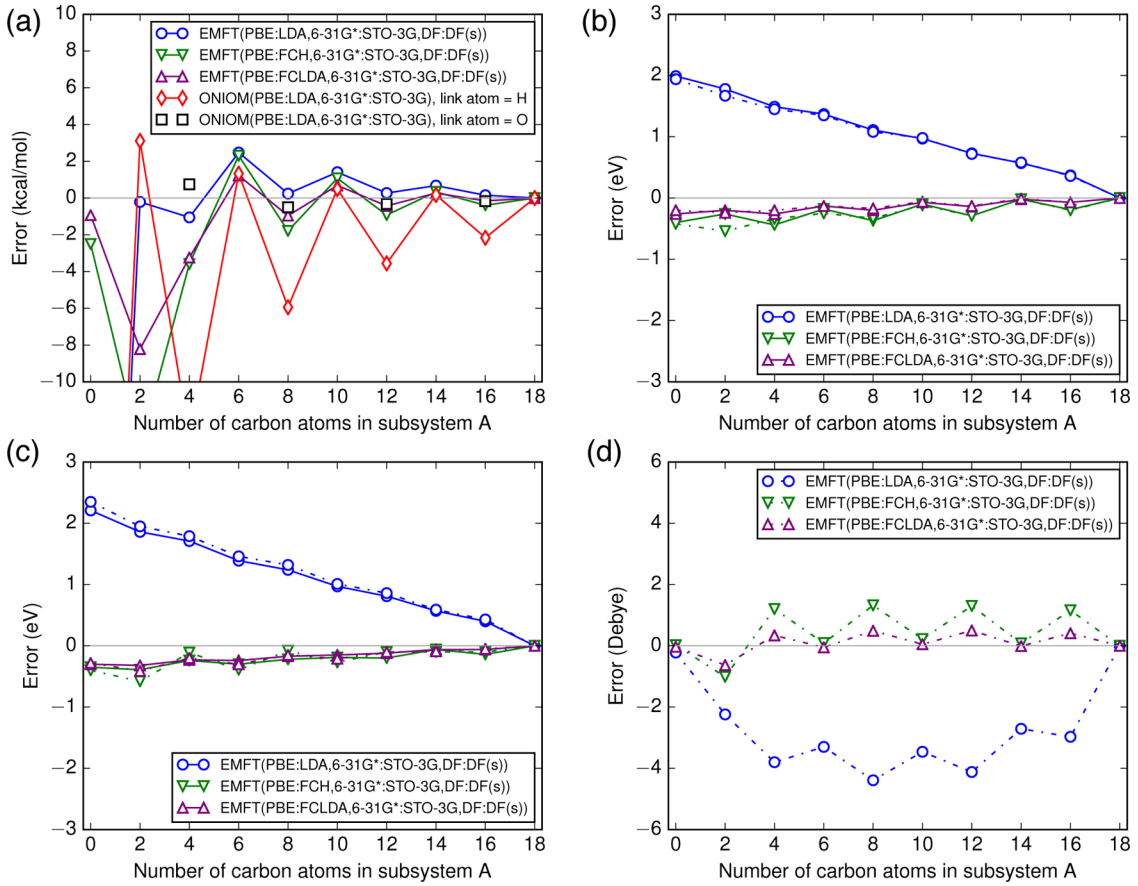


Figure 4. Errors in embedding calculations using PBE/6-31G* in subsystem A for a Diels−Alder reaction of an 18-carbon polyene chain and a butadiene relative to PBE/6-31G* on the whole system. Horizontal axes correspond to the number of carbon atoms in the 18-carbon polyene chain. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with PBE/6-31G*) for each panel are (a) −27.0 kcal mol^{-1}; (b) −4.16 eV (reactant), −4.47 eV (product); (c) −2.92 eV (reactant), −2.32 eV (product); and (d) 0.00 D (reactant), 0.76 D (product).

exchange-correlation potential reduces the errors significantly: the errors of FCLDA are comparable to minimal-basis DFT and DFTB results.

EMFT Using FCDFT as the Low-Level Method. We examine the performance of FCDFT as the low level of theory in EMFT, using hydrocarbon test systems. We investigate the subsystem-size dependence of errors in reaction energies, HOMO and LUMO energies, and dipole moments. For comparison, DFT-in-DFT calculations with EMFT[61] and ONIOM[48] are also performed. The chemical reactions studied are the terminal hydrogenation of decapentaene, Diels−Alder reaction between an 18-carbon polyene and butadiene, and terminal hydrogenation of pentacene. In each case, the conjugated electronic structure provides a stringent test on the sensitivity of embedding theories to boundary effects. Geometries of reactants and products are shown in Figure 1.

In this section, PBE/6-31G* and B3LYP/6-311G** are used in subsystem A, whereas LDA/STO-3G and FCDFT are used subsystem B. For EMFT calculations, density fitting is used to evaluate the Coulomb[101,102] and exact exchange[103] energies. We use Ahlrichs' density-fitting basis set[104,105] for PBE calculations and cc-pVTZ/JKFIT[103] for B3LYP calculations. In subsystem A, these full basis sets are used, but in subsystem B, only s-type functions are used; the effectiveness of this treatment is discussed in ref 61. The EX0 implementation is used to evaluate the exact exchange contribution in EMFT calculations.[61] This approach can be viewed as computing the exchange contribution only including four-index integrals with all four indices subsystem A (although in practice density fitting is used). In some regards, EX0 resembles the auxiliary density-matrix method approach ADMM3, in which an auxiliary density matrix used for computation of exact-exchange contributions is approximated in block-diagonal form over fragments of the system.[106]

We begin by introducing the notation used in this section, which follows that employed in our previous work.[61] Low- and high-level methods in embedding calculations are described in parentheses and are separated using a colon (:). For example, EMFT(PBE:LDA, 6-31G*:STO-3G, DF:DF(s)) indicates an EMFT calculation using PBE/6-31G*/DF and LDA/STO-3G/DF(s) as the high- and low-level theories, respectively. Similarly, EMFT(PBE:FCH, 6-31G*:STO-3G, DF:DF(s)) represents an EMFT calculation using PBE/6-31G*/DF and FCH/STO-3G/DF(s), and the FCH parameter set is fitted to PBE/6-31G*. Finally, ONIOM(B3LYP:LDA, 6-311G**:STO-3G) represents an ONIOM calculation using B3LYP/6-311G** and LDA/STO-3G as the high- and low-level theories, respectively. Here, DF represents the use of the density fitting technique and DF(s) represents the use of the subset of the density-fitting basis set which consists of only s-type functions for subsystem B. For ONIOM calculations, we use default values for link-atom bond lengths in Gaussian 09.[96]

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

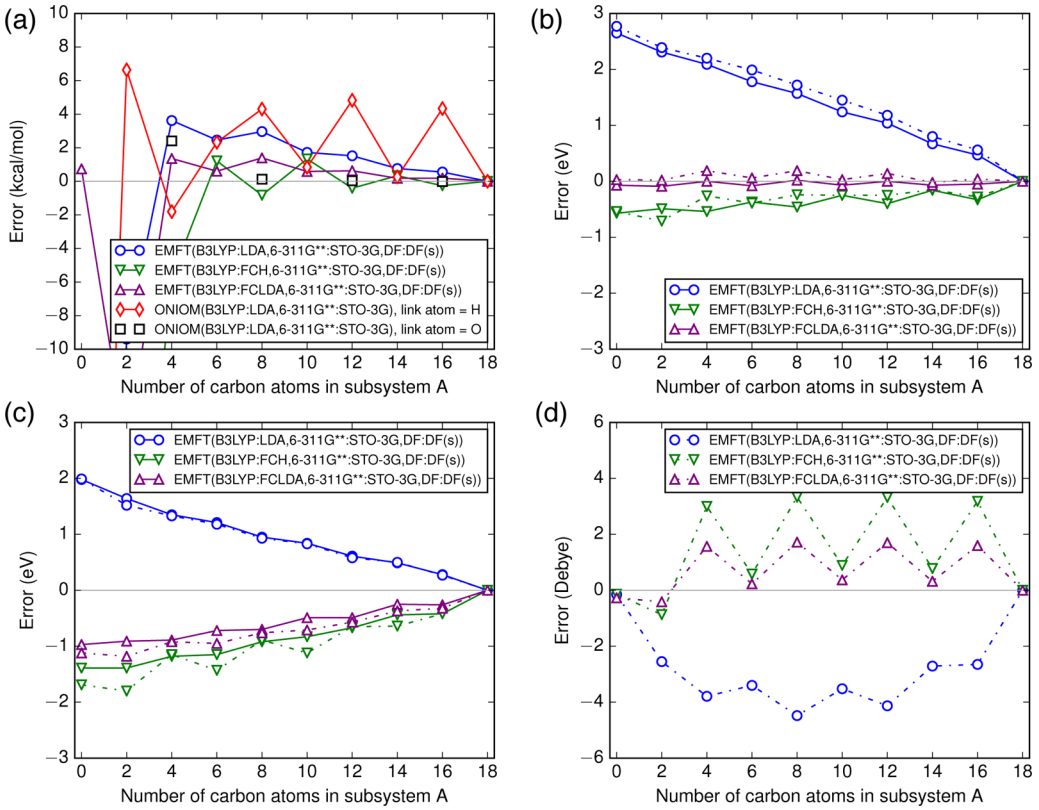


Figure 5. Errors in embedding calculations using B3LYP/6-311G** in subsystem A for a Diels−Alder reaction of an 18-carbon polyene chain and a butadiene relative to B3LYP/6-311G** on the whole system. Horizontal axes correspond to the number of carbon atoms in the 18-carbon polyene chain. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with B3LYP/6-311G**) for each panel are (a) −16.9 kcal mol^{-1}; (b) −4.94 eV (reactant), −5.39 eV (product); (c) −2.59 eV (reactant), −1.88 eV (product); and (d) 0.00 D (reactant), 0.68 D (product).

Throughout this section, unless otherwise noted, subsystem A is described by the number of carbon atoms in the region, and associated hydrogen atoms are also included. Zero carbon atoms indicates the result at the low level of theory. All EMFT calculations were carried out using the development version of the Molpro software package and ONIOM calculations were performed by Gaussian 09.

Terminal Hydrogenation of Decapentaene. Embedding calculations are performed for hydrogenation of the terminal double bond of decapentaene, and subsystem A is extended from the reaction center along the polyene chain. Figure 2 shows results of embedding calculations using the PBE/6-31G* as the high level of theory in subsystem A. In panel (a), we compare errors in the reaction energy. For ONIOM calculations, an oxygen atom is used as the link atom when a double bond crosses a boundary between two subsystems, whereas a hydrogen atom is used to terminate a single bond. It is worth reiterating that EMFT does not require such choices to be made because the partitioning is made strictly at the level of subsets of the atomic-orbital basis.

Low-level calculations over the whole system (corresponding to zero carbon atoms in subsystem A) show that both parametrizations of FCDFT predict the reaction energy with good accuracy (errors less than 2 kcal/mol), whereas LDA/STO-3G gives a significantly larger error of −24 kcal/mol. All embedding theories show rapid convergence with respect to the size of subsystem A. Interestingly, ONIOM provides accurate reaction energies even when a double bond crosses the boundary, provided that an appropriate link atom is chosen. Only EMFT calculations using FCDFT as the low-level theory produce errors below 1 kcal/mol, even when subsystem A contains just the two reacting carbon atoms.

Panels (b) and (c) of Figure 2 compare errors in the HOMO and LUMO energies of EMFT calculations. Embedding in low-level DFT (LDA/STO-3G) gives errors in the HOMO and LUMO energy of up to 2 eV, but the errors using FCDFT in subsystem B are much smaller. Although the errors decrease monotonically with the size of subsystem A regardless of method, the combination EMFT(PBE:FCLDA, 6-31G*:STO-3G, DF:DF(s)) shows the smallest error at each point.

Panel (d) compares errors in dipole moments. Although the low level of DFT theory (LDA/STO-3G) itself predicts reasonable dipole moments for both the reactant and product, EMFT calculations using this theory in subsystem B yield large errors of 2−4 debye for any partitioning choice. FCH is less accurate than LDA in estimating the dipole moment of the product, but errors in EMFT calculations using FCH for subsystem B are greatly reduced. EMFT using FCLDA in subsystem B gives the smallest errors at any point, typically around half of the size of errors using FCH and an order of magnitude smaller than using LDA/STO-3G.

Embedding calculations using the B3LYP/6-311G** in subsystem A show similar trends to those using PBE/6-31G* for reaction energies and HOMO energies, as shown in Figure 3. All embedding methods show rapid convergence of the reaction energy; errors in the HOMO energy decrease monotonically

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

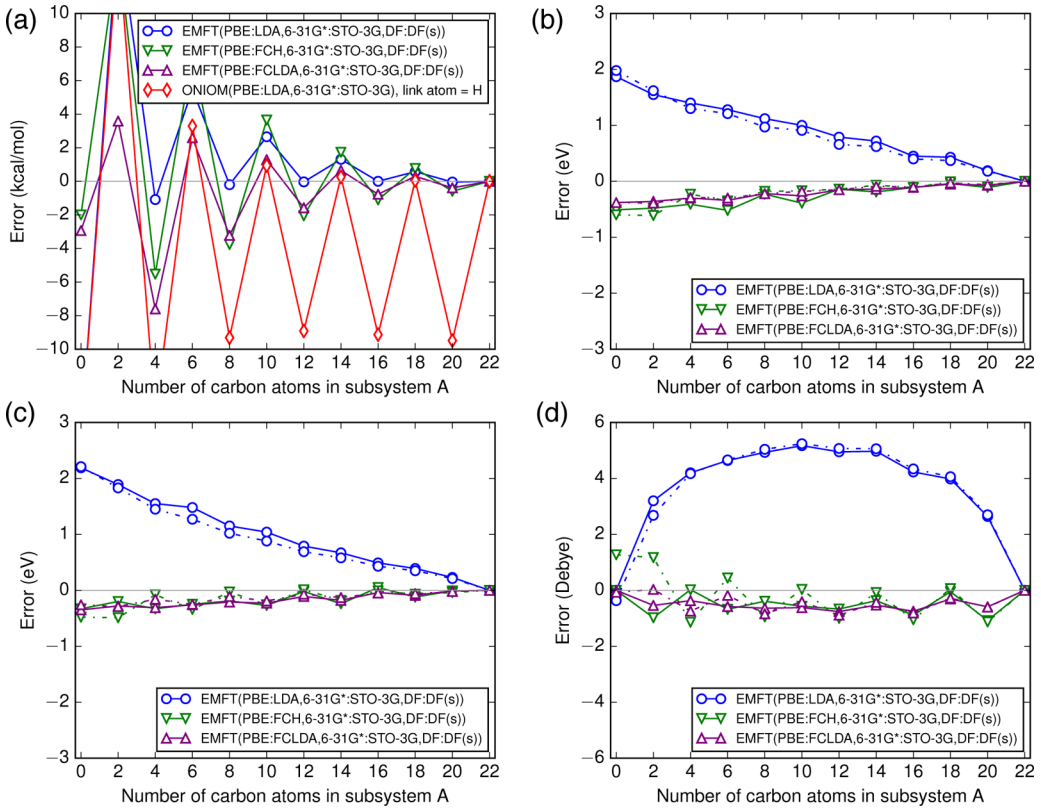


Figure 6. Errors in embedding calculations using PBE/6-31G* in subsystem A for a pentacene hydrogenation relative to PBE/6-31G* on the whole system. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with PBE/6-31G*) for each panel are (a) 21.6 kcal mol^{-1}; (b) −4.12 eV (reactant), −3.51 eV (product); (c) −2.96 eV (reactant), −3.30 eV (product); and (d) 0.00 D (reactant), −1.10 D (product).

with respect to the size of subsystem A. More importantly, EMFT calculations using FCDFT give the smallest errors for both properties. Dipole moments (panel (d)) are improved through use of FCDFT, but to a lesser extent than when PBE/6-31G* is used in subsystem A. The LUMO energies are changed significantly by use of FCDFT, but the correction is overestimated. Embedding in FCH produces errors of similar magnitude to embedding in LDA/STO-3G, but they are of opposite sign, and FCLDA produces errors of around half the size.

Diels−Alder Reaction between an 18-Carbon Polyene and Butadiene. In this example, a Diels−Alder reaction takes place across the central pair of carbon atoms in the polyene (see Figure 1b). For the embedding calculations, subsystem A is symmetrically extended from this reaction center.

Figure 4 summarizes the results of embedding calculations using PBE/6-31G* as the high level of theory. Panel (a) shows the errors in the reaction energies. In the first set of ONIOM calculations (red diamonds), hydrogen atoms were used as link atoms; the second set (black squares) involved only partitions across double bonds, and oxygen was used as the link atom. ONIOM calculations using hydrogen link atoms show significant errors when the boundaries cross the double bonds, although use of oxygen atoms reduces this error dramatically. A fundamental issue for ONIOM is that in a dynamical process the conjugation pattern is not always known *a priori*, and this choice of link atoms cannot always be made in advance.

LDA/STO-3G (for the whole system) gives a huge error of −55 kcal/mol in the reaction energy, but FCDFT methods reduce this error to around 2 kcal/mol. For small numbers of carbon atoms in subsystem A, EMFT using LDA/STO-3G in subsystem B performs well, but convergence for larger subsystems is somewhat improved overall through use of FCDFT. The trends of errors in HOMO energy, LUMO energy, and dipole moment (panels (b)−(d) of Figure 4) are similar to those found for decapentaene hydrogenation, with dramatic reduction of errors on moving from minimal-basis DFT to FCDFT in subsystem B.

The behavior of EMFT with B3LYP/6-311G** in subsystem A and the various low-level models for subsystem B is shown in Figure 5. Here again, the broad conclusion is that embedding in FCDFT improves over embedding in minimal-basis DFT, producing stable convergence of the reaction energy and greatly improved accuracy in the HOMO energy. Again, the improvements in the LUMO energy and dipole moment are significant but not as great as when subsystem A is treated with PBE/6-31G*.

Hydrogenation of a Pentacene. Finally, we examine hydrogenation of a terminal C−C bond in pentacene. Geometries for the reactant and product are shown in panel (c) of Figure 1, and the results using PBE/6-31G* and B3LYP/6-311G** as the high-level methods are shown in Figures 6 and 7, respectively. Although these calculations reveal patterns similar to the first two examples, there are some notable differences. First, it can be seen that the need to partition across a strongly delocalized conjugated system presents great difficulties for ONIOM but not for EMFT. Second, it can be seen that EMFT with FCH in subsystem B does not perform particularly well,

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

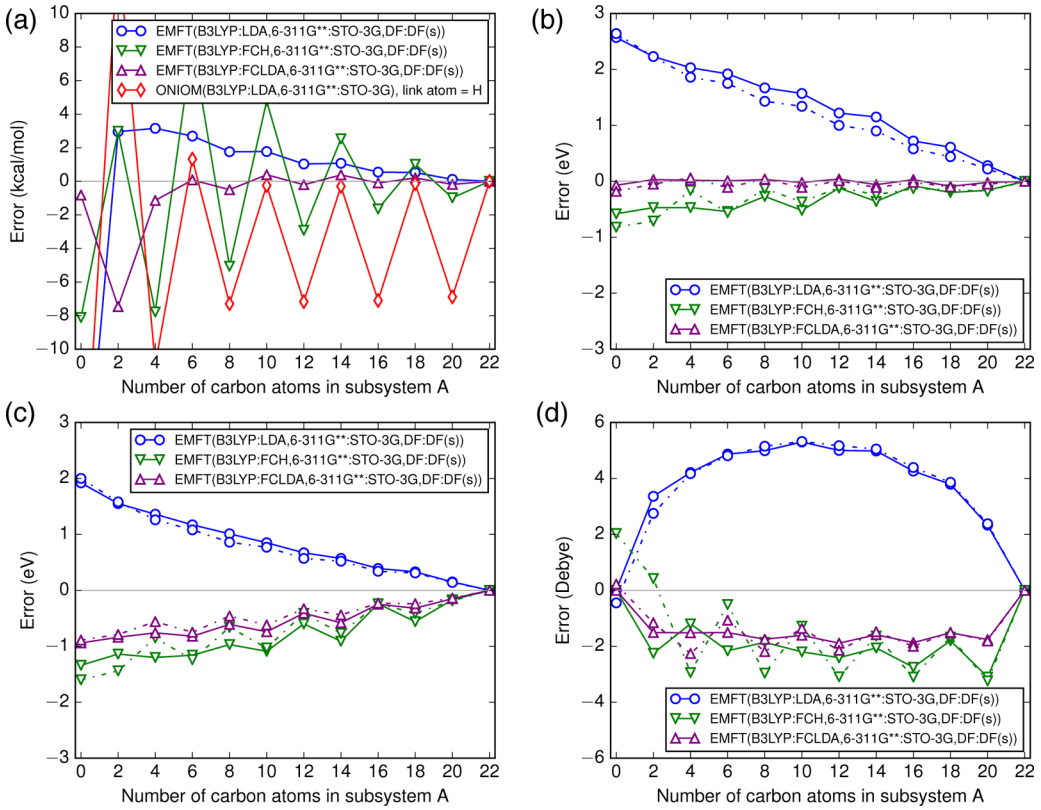


Figure 7. Errors in embedding calculations using B3LYP/6-311G** in subsystem A for a pentacene hydrogenation relative to B3LYP/6-311G** on the whole system. The panels show errors in (a) reaction energy, (b) HOMO energy, (c) LUMO energy, and (d) dipole moment along the *z*-axis. Solid and dashed lines in panels (b)−(d) indicate reactant and product, respectively. Geometries and the orientation of the *z*-axis are shown in Figure 1. Reference values (computed with B3LYP/6-311G**) for each panel are (a) 28.1 kcal mol^{-1}; (b) −4.86 eV (reactant), −4.22 eV (product); (c) −2.67 eV (reactant), −3.06 eV (product); and (d) 0.00 D (reactant), −1.00 D (product).

although FCLDA offers a clear improvement over minimal-basis DFT. As before, EMFT HOMO energies are greatly improved by Fock corrections, and LUMO energies and dipole moments are more significantly improved when the subsystem A method is PBE/6-31G* than when it is B3LYP/6-311G**.

CONCLUSIONS

A semiempirical minimal-basis electronic structure method called FCDFT has been proposed. Semiempirical parameters for FCDFT are determined to reproduce a given reference DFT approximation using a small number of molecules as a training set. Numerical assessments show that FCDFT methods give HOMO and LUMO energies closer to those of the reference DFT calculations than uncorrected minimal-basis DFT. Also, FCDFT predicts atomization energies and reaction energies of hydrocarbons with accuracy comparable to that of PBE/6-31G*. FCLDA consistently outperforms FCH, in which all exchange-correlation effects are handled by the semiempirical Fock-matrix correction, and this is particularly so for dipole moments.

EMFT calculations that use FCDFT for the low-level subsystem show better convergence in reaction energies for our three test reactions than corresponding calculations that use uncorrected minimal-basis DFT. Use of FCDFT greatly alleviates errors in the dipole moments and HOMO energies, as well as improves LUMO energies.

While our two FCDFT variants perform well in capturing the differences between DFT/STO-3G and PBE/6-31G*, the performance is not always as good in describing the correction to B3LYP/6-311G**. This could have contributions from the increased size of the basis in the reference calculation, but we think it is more strongly associated with the difficulty of capturing the effect of the nonlocal exchange contribution using the simple parametrization of the **L** matrix. This is supported by the fact that it is more difficult to find Fock-matrix corrections that reproduce the HOMO−LUMO gap of hybrid functionals rather than those of GGAs. This is unlikely to limit EMFT applications, where the whole motivation for a multiscale approach is reduced sensitivity to accuracy in the environment.

In this article, we showed calculations only for hydrocarbons. However, the results suggest that DFT-in-FCDFT embedding in the EMFT framework can overcome many of the leading sources of error in previously reported EMFT calculations.[61] FCDFT can be seen as a low-cost intermediate point between minimal-basis DFT and SCC-DFTB, greatly improving over the accuracy of the former and providing a treatment of electrostatics that makes it more suitable than the latter for use in EMFT calculations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00685.

All numerical results and fitted parameters (XLSX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: fred.manby@bristol.ac.uk.

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

Funding

K.M. acknowledges the support of Toyota Central R&D Laboratories, Inc. T.F.M. acknowledges support from the Office of Naval Research (ONR) under Grant No. N00014-16-1-2761 and by the U.S. Army Research Laboratory under Grant No. W911NF-12-2-0023.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(2) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.
(3) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(4) Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182−187.
(5) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(6) Karplus, M. *Angew. Chem., Int. Ed.* **2014**, *53*, 9992−10005.
(7) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct.: THEOCHEM* **1999**, *461−462*, 1−21.
(8) Vreven, T.; Morokuma, K.; Farkas, Ö; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760−769.
(9) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815−826.
(10) Chung, L. W.; Hirao, H.; Li, X.; Morokuma, K. *WIREs Comput. Mol. Sci.* **2012**, *2*, 327−350.
(11) Hratchian, H. P.; Krukau, A. V.; Parandekar, P. V.; Frisch, M. J.; Raghavachari, K. *J. Chem. Phys.* **2011**, *135*, 014105.
(12) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.
(13) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094.
(14) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(15) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999−3093.
(16) Mennucci, B. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386−404.
(17) Hu, L.; Söderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2011**, *7*, 761−777.
(18) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293−307.
(19) Brunk, E.; Rothlisberger, U. *Chem. Rev.* **2015**, *115*, 6217−6263.
(20) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(21) Sun, Q.; Chan, G. K.-L. *J. Chem. Theory Comput.* **2014**, *10*, 3784−3790.
(22) Senatore, G.; Subbaswamy, K. R. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5754−5757.
(23) Cortona, P. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *44*, 8454−8458.
(24) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050−8053.
(25) Huang, P.; Carter, E. A. *Annu. Rev. Phys. Chem.* **2008**, *59*, 261−290.
(26) Huang, C.; Carter, E. A. *J. Chem. Phys.* **2011**, *135*, 194104.
(27) Elliott, P.; Cohen, M. H.; Wasserman, A.; Burke, K. *J. Chem. Theory Comput.* **2009**, *5*, 827−833.
(28) Elliott, P.; Burke, K.; Cohen, M. H.; Wasserman, A. *Phys. Rev. A: At., Mol., Opt. Phys.* **2010**, *82*, 024501.
(29) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III *J. Chem. Phys.* **2010**, *133*, 084103.
(30) Goodpaster, J. D.; Barnes, T. A.; Miller, T. F., III *J. Chem. Phys.* **2011**, *134*, 164108.
(31) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III *J. Chem. Phys.* **2012**, *137*, 224113.
(32) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.
(33) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F., III *J. Chem. Phys.* **2013**, *139*, 024103.
(34) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III *J. Chem. Phys.* **2014**, *140*, 18A507.
(35) Bennie, S. J.; Stella, M.; Miller, T. F., III; Manby, F. R. *J. Chem. Phys.* **2015**, *143*, 024105.
(36) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632−672.
(37) Knizia, G.; Chan, G. K.-L. *Phys. Rev. Lett.* **2012**, *109*, 186404.
(38) Jacob, C. R.; Neugebauer, J. *WIREs Comput. Mol. Sci.* **2014**, *4*, 325−362.
(39) Neuhauser, D.; Baer, R.; Rabani, E. *J. Chem. Phys.* **2014**, *141*, 041102.
(40) Knizia, G.; Chan, G. K.-L. *J. Chem. Theory Comput.* **2013**, *9*, 1428−1432.
(41) Georges, A.; Kotliar, G.; Krauth, W.; Rozenberg, M. J. *Rev. Mod. Phys.* **1996**, *68*, 13−125.
(42) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959−1967.
(43) Gomes, A. S. P.; Jacob, C. R. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2012**, *108*, 222−277.
(44) Huang, C. *J. Chem. Theory Comput.* **2016**, *12*, 2224−2233.
(45) Pernal, K. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21111−21121.
(46) Yang, T.; Zhao, X.; Nagase, S. *J. Comput. Chem.* **2013**, *34*, 2223.
(47) Libisch, F.; Huang, C.; Carter, E. A. *Acc. Chem. Res.* **2014**, *47*, 2768−2775.
(48) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, *115*, 5678−5796.
(49) Stella, M.; Bennie, S. J.; Manby, F. R. *Mol. Phys.* **2015**, *113*, 1858−1864.
(50) Olsson, M. H. M.; Hong, G.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 5025−5039.
(51) Barnes, T. A.; Kaminski, J. W.; Borodin, O.; Miller, T. F., III *J. Phys. Chem. C* **2015**, *119*, 3865−3880.
(52) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. *J. Chem. Theory Comput.* **2016**, *12*, 2689−2697.
(53) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353−5362.
(54) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. *J. Chem. Phys.* **2005**, *122*, 094115.
(55) König, C.; Neugebauer, J. *ChemPhysChem* **2012**, *13*, 386−425.
(56) Neugebauer, J. *ChemPhysChem* **2009**, *10*, 3148−3173.
(57) Neugebauer, J. *Phys. Rep.* **2010**, *489*, 1−87.
(58) García-Lastra, J. M.; Wesolowski, T.; Barriuso, M. T.; Aramburu, J. A.; Moreno, M. *J. Phys.: Condens. Matter* **2006**, *18*, 1519−1534.
(59) Chiba, M.; Fedorov, D. G.; Kitaura, K. *J. Chem. Phys.* **2007**, *127*, 104108.
(60) Chiba, M.; Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2007**, *444*, 346−350.
(61) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III *J. Chem. Theory Comput.* **2015**, *11*, 568−580.
(62) Nakatsuji, H. *J. Chem. Phys.* **1987**, *87*, 4995−5001.
(63) Nakatsuji, H. *Prog. Surf. Sci.* **1997**, *54*, 1−68.
(64) Nakatsuji, H.; Nakai, H. *J. Chem. Phys.* **1993**, *98*, 2423−2436.
(65) Wouters, S.; Jiménez-Hoyos, C. A.; Sun, Q.; Chan, G. K.-L. *J. Chem. Theory Comput.* **2016**, *12*, 2706−2719.
(66) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133−A1138.
(67) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press, 1989; pp 1−197.
(68) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *51*, 12947−12957.
(69) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 7260−7268.
(70) Sure, R.; Grimme, S. *J. Comput. Chem.* **2013**, *34*, 1672−1685.
(71) Brandenburg, J. G.; Grimme, S. *Top. Curr. Chem.* **2013**, *345*, 1−23.
(72) Brandenburg, J. G.; Hochheim, M.; Bredow, T.; Grimme, S. *J. Phys. Chem. Lett.* **2014**, *5*, 4275−4284.

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822

(73) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. *J. Phys. Chem. A* **2009**, *113*, 11866−11881.
(74) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931−948.
(75) Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338−354.
(76) Slater, J. C.; Koster, G. F. *Phys. Rev.* **1954**, *94*, 1498−1524.
(77) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. *MOLPRO*, version 2012.1, a package of ab initio programs, 2012. http://www.molpro.net.
(78) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242−253.
(79) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864−B871.
(80) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(81) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657−2664.
(82) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769−2773.
(83) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(84) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.
(85) Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 213−222.
(86) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.
(87) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223−1229.
(88) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(89) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.
(90) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479−483.
(91) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221−7230.
(92) Johnson, R. D., III, Ed. *NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101*, Release 16a, 2016. http://cccbdb.nist.gov/.
(93) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650−7657.
(94) Nocedal, J.; Wright, S. J. *Numerical Optimization*; Springer, 2006; pp 136−143.
(95) *SciPy.org*. http://www.scipy.org/.
(96) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
(97) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063−1079.
(98) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374−7383.
(99) Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678−5684.
(100) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861−10873.
(101) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514−518.
(102) Dunlap, B. I. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2113−2116.
(103) Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285−4291.
(104) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283−290.
(105) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119−124.
(106) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348−2364.

DOI: 10.1021/acs.jctc.6b00685
J. Chem. Theory Comput. 2016, 12, 5811−5822