
Peer reviewed version

Link to published version (if available):
10.1021/acs.jctc.6b01065

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via ACS Publications at http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b01065. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Embedded mean-field theory with block-orthogonalized partitioning

Feizhi Ding,† Frederick R. Manby,‡ and Thomas F. Miller III*,†

†Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
‡Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, UK

E-mail: tfm@caltech.edu

Abstract

Embedded mean-field theory (EMFT) provides a simple, flexible framework for describing subsystems at different levels of mean-field theory. Subsystems are defined by partitioning a one-particle basis set, with a natural choice being the atomic orbital (AO) basis. Although generally well behaved, EMFT with AO partitioning can exhibit unphysical collapse of the self-consistent solution. To avoid this issue, we introduce subsystem partitioning of a block-orthogonalized (BO) basis set; this eliminates the unphysical collapse without significantly increasing computational cost. We also investigate a non-self-consistent implementation of EMFT, in which the density matrix is obtained using BO partitioning, and the final energy evaluated using AO partitioning; this density-corrected EMFT approach is found to yield more accurate energies than BO partitioning while also avoiding issues of unphysical collapse. Using these refined implementations of EMFT, previously proposed descriptions of the exact-exchange coupling between subsystems are compared: although the EX1 coupling scheme is slightly
more accurate than EX0, the small improvement does not merit its substantially greater computational cost.

**Introduction**

A central challenge in theoretical chemistry is the development of efficient electronic structure methods to enable the accurate description of properties and dynamics for complex systems. Quantum embedding has emerged as a powerful strategy in which the total system is divided into a small, chemically active region described by a high level of theory, while the surrounding region is treated at a lower level of theory. Popular realizations of the approach include QM/MM, \(^1\)–\(^4\) ONIOM, \(^5\)–\(^8\) fragment molecular orbital, \(^9\)–\(^{14}\) and density-based embedding methods. \(^15\)–\(^{30}\) Despite their successes, embedding methods can exhibit drawbacks that include the need to specify fixed particle number and spin-state for each subsystem and uncontrolled errors associated with subsystem interactions, which has inspired the recent development of embedding strategies that are formally exact in the description of subsystem interactions\(^{24}\)–\(^{36}\) and allow for particle-number fluctuations between subsystems via their description as open quantum systems.\(^{34}\)–\(^{36}\)

Recently, we introduced embedded mean-field theory (EMFT),\(^{36}\) which achieves these goals by describing embedded subsystems at different levels of mean-field theory, without the need to specify or fix the particle number or spin-state for each subsystem. EMFT is simple, parameter-free, and yields straightforward formulations of nuclear gradients and response properties. Moreover, the method was demonstrated to be accurate over a wide range of benchmark systems and chemical applications, including applications that involve subsystem partitioning across conjugated bonding networks.

A key aspect of EMFT is that subsystems are partitioned at the level of a one-particle basis, and the method was originally implemented with partitioning in terms of the atomic orbital (AO) basis.\(^{36}\) Although in many cases this works well, we show here that it can also lead to unphysical collapse of the self-consistent EMFT solution, which
is associated with mismatch in the functional form of the energy expressions between the high- and low-level theories. To address this issue, we introduce an alternative partitioning scheme based on a basis set that is block orthogonalized (BO) with respect to the two subsystems. EMFT with BO partitioning is tested on a range of systems for which AO partitioning fails, and the new choice of partitioning is shown to eliminate the problem of unphysical collapse of the EMFT solution while retaining the appealing features of the original method. We additionally introduce a "density corrected EMFT" scheme in which the energy is evaluated using AO partitioning with the density matrix optimized using BO partitioning. This generally yields better accuracy than EMFT with BO partitioning while avoiding the unphysical collapse of AO partitioning. These refined implementations of EMFT are used to investigate various schemes for describing exact-exchange coupling between the embedded subsystems.

**Embedded Mean-Field Theory**

We first review the embedded mean-field theory (EMFT), which allows for the embedding of one mean-field theory in another. EMFT employs a partitioning of the one-particle density matrix in terms of a one-particle basis set, such as the AO basis,

\[
D = \begin{bmatrix}
D_{AA} & D_{AB} \\
D_{BA} & D_{BB}
\end{bmatrix},
\]

(1)

where \(D_{AA}\) and \(D_{BB}\) denote the density-matrix blocks that belong to subsystems A and B, respectively.

Given two different mean-field theories for describing subsystems A and B, the
EMFT ground-state energy is obtained by minimizing the energy functional

\[
E[D] = E^\text{Low}[D] + E^\text{High}[D;AA] - E^\text{Low}[D;AA]
\]

\[
= \text{Tr}(DH_0) + G^\text{Low}[D] + (G^\text{High}[D;AA] - G^\text{Low}[D;AA])
\]

while enforcing the idempotency and normalization constraints for the total density matrix, such that for closed-shell systems,

\[
DSD = 2D,
\]

\[
\text{Tr}(DS) = N_e.
\]

Here, \(G^\text{High}\) and \(G^\text{Low}\) represent the two-electron energy calculated at the high- and low-level of the mean-field theories, respectively; \(H_0\) is the one-electron Hamiltonian, \(S\) is the basis overlap matrix, and \(N_e\) is the total number of electrons. This functional minimization leads to the usual self-consistent field (SCF) equation

\[
FC = SC\varepsilon,
\]

where \(F\) is the EMFT Fock matrix and is defined in the usual way as the derivative of the EMFT energy given in Eq. (2),

\[
F = \frac{\partial E[D]}{\partial D},
\]

\(C\) is the molecular orbital (MO) coefficient matrix, and \(\varepsilon\) is the diagonal matrix containing the canonical MO eigenvalues.

Because EMFT is itself a mean-field theory, gradients and other response theories retain their usual mean-field form. For example, the gradient with respect to nuclear displacements is given by

\[
\frac{\partial E[D]}{\partial x} = \text{Tr}(DF^{(x)}) - \text{Tr}(WS^{(x)}),
\]
where $\mathbf{W}$ is the energy-weighted density and $\mathbf{F}^{(x)}$ and $\mathbf{S}^{(x)}$ are the nuclear derivatives of the Fock and overlap matrices, respectively.

Additional advantages of EMFT emerge from the fact that the number of electrons in each subsystem is determined directly from minimization of the energy functional in Eq. (2). The first is that charge flow between the subsystems is naturally included, and the second is that there is no need for the user to specify the number of electrons or spin state for each subsystem; beyond the definition of the two mean-field theories and the list of atomic orbitals in each subsystem, there are no additional parameters. Finally, we note that the individual subsystems in EMFT are treated as open quantum systems, such that the entanglement between them is fully included at the mean-field level.\textsuperscript{36}

The EMFT framework is general, enabling the embedding of any mean-field theory in any other. Natural choices for the description of subsystem A might include DFT with hybrid, double-hybrid, or range-separated hybrid functionals, as well as self-consistent random-phase approximations. Subsystem B might be described by more computationally efficient DFT methods or tight-binding.\textsuperscript{37} The efficiency of EMFT depends on the relative cost of the high- and low-level method and the relative size of subsystems A and B. Assuming that the low-level theory is negligible in cost with respect to the high-level theory, the cost of the combined EMFT calculation will be comparable to the cost of the high-level calculation on subsystem A; alternatively, in the limiting case for which subsystem B is very large and the cost of the low-level method is not negligible, the cost of the combined EMFT calculation will be comparable to the cost of the low-level calculation on subsystem B. In both of these regimes, the cost of ONIOM calculations behave similarly.

**EMFT with exact exchange**

For applications of EMFT in which the two mean-field theories are DFT, with a hybrid exchange-correlation (xc) functional corresponding to the high-level theory (labeled as
and a pure functional corresponding to the low-level theory (labeled as $E^{(2)}_{xc}$), the energy expression is

$$E[D] = \text{Tr}[DH_0] + E_H[D] + E^{(2)}_{xc}[D] + c_x E_{EX}[D^{AA}] - E^{(2)}_{xc}[D^{AA}],$$

(8)

where $E_H[D]$ is the Coulomb energy, $E_{EX}[D^{AA}]$ is the exact exchange (EX) energy for the high-level region, and $c_x$ is the fraction of the exact exchange used to define the hybrid functional.

The exact exchange energy $E_{EX}[D^{AA}]$ can be expressed in three different schemes, depending on the amount of the subsystem exchange interaction to be included. The simplest one is the EX0 scheme, where only exchange interaction within subsystem A is considered. For a closed-shell system, the EX0 energy is expressed as

$$E_{EX0} = -\frac{1}{4} \sum_{\mu\nu\lambda\kappa \in A} (\mu\kappa|\nu\lambda) D_{\mu\nu}^{AA} D_{\kappa\lambda}^{AA}.$$

(9)

A slightly more complicated model for the exact exchange (EX1) symmetrically averages the exchange interaction between the two subsystems, such that

$$E_{EX1} = -\frac{1}{4} \sum_{\mu\nu\lambda\kappa \in A} (\mu\kappa|\nu\lambda) D_{\mu\nu}^{AA} D_{\kappa\lambda}^{AA} - \frac{1}{4} \sum_{\mu\nu \in A, \lambda\kappa \notin A} (\mu\kappa|\nu\lambda) D_{\mu\nu}^{AA} D_{\kappa\lambda}.$$

(10)

A third model, EX2, includes the full exchange interaction between the subsystems, such that

$$E_{EX2} = -\frac{1}{4} \sum_{\mu\nu\lambda\kappa \in A} (\mu\kappa|\nu\lambda) D_{\mu\nu}^{AA} D_{\kappa\lambda}^{AA} - \frac{1}{2} \sum_{\mu\nu \in A, \lambda\kappa \notin A} (\mu\kappa|\nu\lambda) D_{\mu\nu}^{AA} D_{\kappa\lambda}.$$

(11)

In terms of the computational efficiency, the formal scaling of evaluating the exchange energy and Fock matrix contribution for the EX0 scheme is $O(N_A^4)$, compared to the formal scaling of $O(N_A^2 N^2)$ for the EX1 and EX2 schemes, where $N_A$ and $N$ are the number of the basis functions for subsystem A and for the whole system. For the test cases considered in this work, we find that the EX2 model consistently performs
worse than either EX0 or EX1 (see Figures S1-S4 in Supporting Information), but at a cost that is similar to EX1 (and much higher than EX0). Therefore, in the following discussion we do not include the results of EX2, and only focus on the comparison between the EX0 and EX1 schemes.

Collapse of the self-consistent EMFT solution with AO partitioning

The (non-orthogonal) AO basis is a natural and convenient choice of one-particle basis sets to employ for subsystem partitioning in EMFT. However, for applications of EMFT that employ AO partitioning, we have identified cases in which a mismatch between the high- and low-level theories can lead to an unphysical collapse of the self-consistent EMFT solution.

![Figure 1: Schematic representation for the partitioning of pentacene. For two different choices for the size of subsystem A (n = 2, 6), the atoms in the highlighted region represent subsystem A, while the remaining atoms constitute subsystem B.](image)

To illustrate this effect, Table 1 presents the EMFT results for pentacene using AO partitioning (columns labeled "AO"). Two different sizes of subsystem A (i.e., two-carbon and six-carbon embedding, see Figure 1) are shown to yield qualitatively similar results. The B3LYP/6-31G* is used as the high-level DFT method, and the LDA/6-
31G* is used as the low-level DFT method. The EX0 scheme is used for modeling exact exchange interactions. As seen from Table 1, EMFT with AO partitioning converges to an unphysical solution with a much lower energy than either the high- or low-level DFT energies for the whole system. Population analysis reveals that the trace of the diagonal blocks of the density matrix, $\text{Tr}(D^{AA}S^{AA})$ and $\text{Tr}(D^{BB}S^{BB})$, are unphysically large. Since the total number of electrons is conserved by the normalization condition,

$$N_e = \text{Tr}(DS) = \text{Tr}(D^{AA}S^{AA}) + \text{Tr}(D^{AB}S^{BA}) + \text{Tr}(D^{BA}S^{AB}) + \text{Tr}(D^{BB}S^{BB}),$$  \hspace{1cm} (12)

the large population in subsystems A and B is accompanied by large, negative values in the coupling terms of the population, $\text{Tr}(D^{AB}S^{BA}) = \text{Tr}(D^{BA}S^{AB})$. The magnitude of the dipole moment obtained using AO partitioning (which should be zero due to the symmetry of the molecule) is also unphysically large.

Table 1: Comparison of EMFT results for pentacene obtained using AO partitioning ("AO"), BO partitioning ("BO"), and the density-corrected ("DC") implementation.\textsuperscript{a} The labels $n = 2$ and $n = 6$ indicate the number of carbon atoms in subsystem A, as illustrated in Figure 1.

<table>
<thead>
<tr>
<th></th>
<th>$n = 2$</th>
<th>$n = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AO</td>
<td>BO</td>
</tr>
<tr>
<td>Energy (hartree)</td>
<td>7726.239013</td>
<td>-839.600160</td>
</tr>
<tr>
<td>$\text{Tr}(D^{AA}S^{AA})$</td>
<td>779.3</td>
<td>18.9</td>
</tr>
<tr>
<td>$\text{Tr}(D^{BB}S^{BB})$</td>
<td>875.9</td>
<td>127.1</td>
</tr>
<tr>
<td>$\text{Tr}(D^{AB}S^{BA})$</td>
<td>-754.6</td>
<td>0.0</td>
</tr>
<tr>
<td>Dipole (debye)</td>
<td>162.0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

\textsuperscript{a} DFT with the B3LYP functional is used as the high-level method, and DFT with the LDA functional is used as the low-level method. The 6-31G* basis is used for both subsystems A and B. The EX0 scheme is used for modeling the exact exchange interactions. The DFT energy for the full system is -846.243261 hartree at the B3LYP/6-31G* level of theory and -828.829653 hartree at the LDA/6-31G* level of theory. The dipole moment from both high- and low-level calculations is zero.

\textsuperscript{b} The electron populations are computed via partitioning of the density and overlap matrices in the AO basis for AO partitioning and density-corrected EMFT. For EMFT with BO partitioning, the populations are obtained by partitioning the corresponding matrices in the BO basis. The subsystem A overlap matrix $S^{AA}$ is the same in both AO and BO bases.

In general, it is found that unphysical collapse of the self-consistent EMFT solution with AO partitioning can occur when there is a mismatch between the functional forms for the high- and low-level theories, such as when a DFT functional (e.g. PBE) is used
as the high-level theory and the crude Hartree method is used as the low-level theory,

\[ E[D] = \text{Tr}(DH_0) + E_H[D] + E_{xc}^{PBE}[D^{AA}], \quad (13) \]

or when DFT with a hybrid exchange-correlation functional (e.g. B3LYP) is used as the high-level theory and DFT with an exchange-correlation functional without HF exchange (e.g. LDA) is used as the low-level theory,

\[ E[D] = \text{Tr}[DH_0] + E_H[D] + E_{xc}^{LDA}[D] + E_{xc}^{B3LYP}[D^{AA}] + 0.2E_{EX}[D^{AA}] - E_{xc}^{LDA}[D^{AA}]. \quad (14) \]

In these cases, each high-level energy term associated with subsystem A is not matched by a low-level counterpart, and these uncompensated energy terms can assume large negative values during the energy minimization. Because of the normalization constraint (Eq. (12)), this process is accompanied by collapse of the density into the off-diagonal blocks, as indicated by the large negative values in the coupling terms in the population expression, i.e. \( \text{Tr}(D^{AB}S^{BA}) \) and \( \text{Tr}(D^{BA}S^{AB}) \).

To further illustrate the manner in which mismatch in the high- and low-level theories leads to the unphysical collapse of the self-consistent EMFT solution, Figure 2 presents a series of EMFT calculations on pentacene with two carbons included in subsystem A, using the calculation details from Table 1. However, in Figure 2, the high-level method corresponds to a modified version of the B3LYP functional in which the fraction of HF exchange \( c_x \) increases from 0 to 0.2 (the value used in B3LYP). As seen in Figure 2a, the EMFT solution with AO partitioning behaves well for small fractions of the HF exchange, for which there is a relatively small mismatch in the high- and low-level energy functionals. The unphysical collapse in the energy occurs at approximately \( c_x = 0.04 \) (Figure 2a) and is accompanied by distortion of the density matrix to exhibit unphysically large values of the population associated with the off-diagonal blocks of the density matrix (Figure 2b).
Figure 2: The self-consistent EMFT energy (a) and off-diagonal population $\text{Tr}(D_{AB}S_{BA})$ (b) for pentacene as the function of the amount of HF exchange used in the high-level subsystem. The calculations are performed with two carbons included in subsystem A, using the calculation details from Table 1. For the high-level theory, a modified version of the B3LYP functional is employed, with the fraction of HF exchange varying between 0 and 0.2.

We also note that for some cases that exhibit unphysical collapse of the self-consistent EMFT solution, there exists a meta-stable local minimum that gives physically reasonable energies and densities, which can be converged to with the use of a good initial guess and a small SCF step size. Indeed, this strategy was employed for the EMFT results involving hybrid exchange-correlation functionals in Ref. 36.
Block-orthogonalized partitioning

Accompaniment of the unphysical collapse by large values of $\text{Tr}(D^{AB}S^{BA})$ and $\text{Tr}(D^{BA}S^{AB})$ suggests an alternative approach, in which the density matrix is partitioned with respect to a one-particle basis for which the off-diagonal blocks of the overlap matrix are strictly zero, i.e. $\tilde{S}^{AB} = \tilde{S}^{BA} = 0$. One such choice would be the fully orthogonalized AO basis, but this would defeat the efficiency gains of EMFT; since the subsystem A basis functions in the fully orthogonal basis generally have contributions from the entire set of non-orthogonal AO basis functions, evaluation of the integrals for subsystem A in the orthogonal basis would require the computation of the integrals for the entire system in the non-orthogonal AO basis.

As an alternative, we investigate partitioning with respect to a basis for which subsystem A remains unchanged from the non-orthogonal AO basis, while the remaining Hilbert space is block-orthogonalized with respect to subsystem A basis using the projection operator

$$\hat{P} = \sum_{\mu \nu \in \Lambda} |\phi^{A}_{\mu}\rangle(S^{AA})_{\mu \nu}^{-1}\langle\phi^{A}_{\nu}|,$$  \hspace{1cm} (15)

such that

$$|\tilde{\phi}^{B}_{\lambda}\rangle = (1 - \hat{P})|\phi^{B}_{\lambda}\rangle.$$  \hspace{1cm} (16)

Here $\{|\phi^{A}_{\mu}\rangle\}$ and $\{|\phi^{B}_{\lambda}\rangle\}$ are the non-orthogonal AO basis functions for subsystems A and B, respectively, $|\tilde{\phi}^{B}_{\lambda}\rangle$ are basis functions for subsystem B that are orthogonal to AO basis functions for subsystem A, and $(S^{AA})^{-1}$ is the inverse of the AO-basis overlap matrix for subsystem A.

Eq. (16) defines a transformation from the non-orthogonal AO basis set $\{|\phi_{\alpha}\rangle\} \equiv \{|\phi^{A}_{\mu}\rangle\} \cup \{|\phi^{B}_{\lambda}\rangle\}$ to the block-orthogonalized (BO) basis set $\{|	ilde{\phi}_{\alpha}\rangle\} \equiv \{|\phi^{A}_{\mu}\rangle\} \cup \{|\tilde{\phi}^{B}_{\lambda}\rangle\}$, such that

$$|	ilde{\phi}_{\alpha}\rangle = \sum_{\beta} |\phi_{\beta}\rangle U_{\beta \alpha},$$  \hspace{1cm} (17)
where $U$ is the transformation matrix with blocked structure,

$$U = \begin{bmatrix} I^{AA} & -P^{AB} \\ 0 & I^{BB} \end{bmatrix}. \tag{18}$$

Here, $I^{AA}$ and $I^{BB}$ are identity matrices with the same dimension as the numbers of basis functions in subsystems A and B, respectively, and $P^{AB}$ is the projection matrix

$$P^{AB} = (S^{AA})^{-1}S^{AB}, \tag{19}$$

where $S^{AB}$ is the overlap of the AO basis functions between subsystems A and B. It is straightforward to confirm that this transformation matrix block-diagonalizes the overlap matrix,

$$\tilde{S} = U^T S U = \begin{bmatrix} S^{AA} & 0 \\ 0 & \tilde{S}^{BB} \end{bmatrix}. \tag{20}$$

Note that $S^{AA}$ is the same in both AO and BO bases.

Subsystem partitioning with respect to the BO basis ensures that $\text{Tr}(\tilde{D}^{AB}\tilde{S}^{BA})$ and $\text{Tr}(\tilde{D}^{BA}\tilde{S}^{AB})$ are strictly zero, such that the normalization condition, Eq. (12), simplifies to

$$N_e = \text{Tr}(\tilde{D}\tilde{S}) = \text{Tr}(\tilde{D}^{AA}S^{AA}) + \text{Tr}(\tilde{D}^{BB}\tilde{S}^{BB}). \tag{21}$$

All electrons are thus associated with the diagonal blocks of the density matrix. According to the idempotency condition, Eq. (3), the total density matrix must be positive semi-definite, which also implies that the diagonal blocks of the density matrix are positive semi-definite.\textsuperscript{38} This result, along with the normalization condition in Eq. (21),

\[ \]
leads to the following implicit constraints on the density matrix,

$$0 \leq \text{Tr}(\tilde{D}^{AA}S^{AA}) \leq N_e, \text{ and } 0 \leq \text{Tr}(\tilde{D}^{BB}\tilde{S}^{BB}) \leq N_e.$$  (22)

Using BO partitioning, the EMFT energy functional is

$$E[\tilde{D}] = \text{Tr}(\tilde{D}\tilde{H}_0) + G^{\text{Low}}[\tilde{D}] + (G^{\text{High}}[\tilde{D}^{AA}] - G^{\text{Low}}[\tilde{D}^{AA}]),$$  (23)

where $\tilde{D}^{AA}$ is the subsystem A block of the total density matrix $\tilde{D}$ in the BO basis. The EMFT solution can be obtained by minimizing this energy functional with the idempotency (Eq. (3)) and normalization (Eq. (4)) constraints. Unlike EMFT with AO partitioning, unphysical collapse of the self-consistent EMFT solution is avoided with BO partitioning, because of the implicit constraints on the density matrix in Eq. (22).

The energy expression in Eq. (23) still exhibits an unbalanced cancellation of terms between $G^{\text{High}}[\tilde{D}^{AA}]$ and $G^{\text{Low}}[\tilde{D}^{AA}]$, which will be seen in the calculated results to still cause minor artifacts in both the energies and the densities (Table 1); however, the BO partitioning prevents the massive flow of population into the off-diagonal matrix elements, which is the hallmark of the dramatic unphysical collapse observed with AO partitioning.

We emphasize that for conventional SCF calculations performed over the full system (such as full DFT or HF calculations), the BO basis and the AO basis yield identical results. However, for EMFT calculations, the BO basis ensures that Eq. (22) is obeyed, such that the trace of the diagonal blocks of the density matrix assume physically reasonable ranges of values for the embedded subsystems.

With the optimized density obtained from the minimization of the EMFT energy functional in the BO basis, Eq. (23), two methods can be employed for the evaluation of the final EMFT energy. The most straightforward method is to simply evaluate the energy using the functional expression in the BO basis (Eq. (23)). Alternatively, akin to the density-corrected DFT approach,\textsuperscript{39–46} the EMFT energy can be calculated using Eq. (2) after transforming the density matrix back to the original, non-orthogonal AO
basis. We call the second alternative the "density-corrected EMFT" approach.

Implementation details for BO partitioning

The major difference between AO and BO partitioning lies in the construction of the EMFT Fock matrix, which in the BO basis is expressed as

\[
\tilde{F}[\tilde{D}] = \tilde{H}_0 + \tilde{G}^{\text{Low}}[\tilde{D}] + \tilde{G}^{\text{High}}[\tilde{D}^{AA}] - \tilde{G}^{\text{Low}}[\tilde{D}^{AA}],
\]

(24)

where \( \tilde{G}^{\text{High}} \) and \( \tilde{G}^{\text{Low}} \) represent the two-electron contributions to the Fock matrix from the high- and low-level of theories. Naive calculation of Eq. (24), however, would require the transformation of the two-electron integrals, which formally scales as \( O(N^5) \). To avoid this, we employ the following strategy to construct the Fock matrix in the BO basis.

First, all matrices associated with the full system (i.e. \( \tilde{H}_0 \) and \( \tilde{G}^{\text{Low}}[D] \) in Eq. (24)) are computed in the original, non-orthogonal AO basis, and then transformed back to the BO basis, using

\[
\tilde{H}_0 = U^T H_0 U, \quad \text{and} \quad \tilde{G}^{\text{Low}}[\tilde{D}] = U^T G^{\text{Low}}[D] U.
\]

(25)

The density matrix \( D \) used in Eq. (25) is obtained by transforming the density matrix from the BO basis to the AO basis, using

\[
D = U \tilde{D} U^T.
\]

(26)

Second, all matrices associated with subsystem A (i.e. \( \tilde{G}^{\text{High}}[\tilde{D}^{AA}] \) and \( \tilde{G}^{\text{Low}}[\tilde{D}^{AA}] \) in Eq. (24)) are computed directly in the original, non-orthogonal AO basis in subsystem A, but using the subsystem A block of the density matrix in the BO basis. For example, for the embedding case in which a local density approximation (LDA) exchange-correlation functional is employed for the low-level of theory, \( \tilde{G}^{\text{Low}}[\tilde{D}^{AA}] \) is
defined as
\[
\tilde{G}_{\mu\nu}^{\text{Low}}[\tilde{D}^{AA}] = \begin{cases} 
\int \phi_{\mu}(r)\phi_{\nu}(r)v_{xc}^{\text{LDA}}[\tilde{\rho}_A(r)]d\mathbf{r}, & \mu, \nu \in A \\
0, & \text{otherwise}
\end{cases}, \tag{27}
\]

where \(\phi_{\mu}(r)\) and \(\phi_{\nu}(r)\) are the BO basis functions in subsystem A (which are the same as the AO basis functions in subsystem A); \(v_{xc}^{\text{LDA}}\) is the exchange-correlation potential for the LDA functional; \(\tilde{\rho}_A(r)\) is defined as
\[
\tilde{\rho}_A(r) = \sum_{\mu\nu \in A} \tilde{D}_{\mu\nu}^{AA}\phi_{\mu}(r)\phi_{\nu}(r). \tag{28}
\]

As seen in Eq. (27) and Eq. (28), the evaluation of \(\tilde{G}_{\mu\nu}^{\text{Low}}[\tilde{D}^{AA}]\) only involves the basis functions in subsystem A; since these subsystem A basis functions are the same in both AO and BO representations, no transformation of \(\tilde{G}_{\mu\nu}^{\text{Low}}[\tilde{D}^{AA}]\) from the AO basis to the BO basis is required. The same simplification holds for \(\tilde{G}_{\mu\nu}^{\text{High}}[\tilde{D}^{AA}]\).

Once the Fock matrix in the BO basis is constructed, the same SCF procedure as used in EMFT with AO partitioning can be employed to achieve convergence of the density matrix,
\[
\tilde{F}[\tilde{D}]\tilde{C} = \tilde{S}\tilde{C}\tilde{\varepsilon}, \tag{29}
\]
where tildes indicate quantities in the BO basis that correspond to the terms in Eq. (5). As discussed earlier, for EMFT with BO partitioning, the final total energy is evaluated in the BO basis according to Eq. (23); for density-corrected EMFT, the total energy is computed using AO partitioning (Eq. (2)), after transforming the density matrix that was obtained with BO partitioning back to the AO basis.

In terms of efficiency, EMFT with BO partitioning (or with the density-corrected scheme) involves essentially the same computational cost as EMFT with AO partitioning. The additional operations include the one-time construction of the transformation matrix defined in Eq. (18), and the matrix transformations, Eqs. (25) and (26), dur-
ing the SCF iterations. The cost of these operation is negligible compared to that of
the two-electron integral calculations and the numerical quadrature for the exchange-
correlation potential, which dominate the EMFT calculations.

Results

Computational details

All calculations reported here are performed using the entos molecular simulation pack-
age.\textsuperscript{47} EMFT calculations with both AO and BO partitioning are performed using the
standard SCF algorithms, including the use of the superposition of atomic densities
as the initial guess, and the direct inversion in the iterative subspace (DIIS) tech-
niques\textsuperscript{48–50} for accelerating the SCF convergence. The convergence criteria are chosen
such that the SCF stops when the change in energy is below $10^{-8}$ hartree and the
norm of the orbital gradient is smaller than $10^{-6}$ hartree. Calculations are reported
for benchmark reactions (Figure 3), including the terminal hydrogenation of pentacene,
the Diels-Alder reaction between the 1,3-butadiene and conjugated octadecanonene ,
and the deprotonation of decanoic acid, as well as the formation of a Stone-Wales de-
fect in a graphene sheet. For all test cases, we use the same geometries and subsystem
definitions as in Ref. 36.
Figure 3: Geometry and schematic representation of the subsystem partitioning for the reactants in (a) the terminal hydrogenation of pentacene, (b) the Diels-Alder reaction between the 1,3-butadiene and conjugated octadecanonene, and (c) the deprotonation of decanoic acid. The atoms in the highlighted region are described by the high-level mean-field method, and the remaining atoms are described by the low-level mean-field method.

All EMFT calculations reported here describe DFT-in-DFT embedding, in which both the high- and low-level mean-field theories correspond to DFT. The notation "PBE-in-LDA" indicates that the PBE exchange-correlation functional is employed for the high-level subsystem, while the local density approximation is used for the low-level subsystem; likewise, "B3LYP-in-LDA" indicates the use of the hybrid B3LYP functional for the high-level subsystem. Density fitting is employed in all calculations for the evaluation of the Coulomb and exact exchange integrals. The Ahlrichs Coulomb Fitting basis is used for the PBE-in-LDA calculations, and the cc-pVDZ/JKFIT basis is used for the B3LYP-in-LDA embedding. When both the high- and low-level regions are described with the same AO basis in EMFT (i.e., "same-basis embedding"), we employ the 6-31G* basis set and the full density-fitting basis. Alternatively, for "mixed-basis embedding", the high-level subsystem is described with the 6-31G* basis set and the full density-fitting basis, while the low-level subsystem is described using the minimal STO-3G basis set and only the s-type functions of the corresponding density-fitting basis. Both the AO and density-fitting functions are implemented as spherical Gaussians. All calculations are closed shell and employ spin-restricted orbitals. No
linear dependencies in the basis sets are observed for the calculations reported here; for applications of EMFT to larger systems, any numerical issues related to linear dependencies in the basis set may be handled in the usual way, for example, using the canonical orthogonalization of the AO basis.\textsuperscript{55}

Elimination of the unphysical collapse

We now return to the case of pentacene (Table 1 and Figure 2) that illustrated the unphysical collapse of the self-consistent EMFT solution with AO partitioning. Columns labeled "BO" in Table 1 show the results of EMFT with BO partitioning, which avoids the unphysical collapse and provides a total energy that lies in between the high- and low-level DFT energies. For both cases with two and six carbon atoms included in subsystem A, the calculated dipole moments are close to zero (though not completely vanishing due to the unbalanced description of different mean-field methods for subsystems A and B). Elimination of the unphysical collapse using BO partitioning is further demonstrated by the population analysis, where both $\text{Tr}(\tilde{D}^{AA}S^{AA})$ and $\text{Tr}(\tilde{D}^{BB}\tilde{S}^{BB})$ assume values within the range of $[0, N_e]$, in contrast with the catastrophic results obtained with AO partitioning. Nonetheless, the populations obtained from BO partitioning significantly deviate from the conventional DFT results at the high level (B3LYP/6-31G*) of theory. For example, at the subsystem size $n = 2$, EMFT with BO partitioning gives the trace of the subsystem A density, $\text{Tr}(\tilde{D}^{AA}S^{AA}) = 18.9$, compared to $\text{Tr}(D^{AA}S^{AA}) = 13.2$ obtained from partitioning the conventional DFT density matrix in the AO basis. This discrepancy will be discussed below. For the case with two carbon atoms included in subsystem A, Figure 2 illustrates that EMFT with BO partitioning avoids unphysical collapse upon increasing the fraction of HF exchange in the high-level functional.

Columns labeled "DC" in Table 1 show the results of density-corrected EMFT, which also produces reasonable values for the total energy. The dipole moments obtained using density-corrected EMFT are unchanged from those obtained with BO partitioning, since the same total density is used to calculate the dipole in both cases. Compared to EMFT
with BO partitioning, density-corrected EMFT provides an improved description of the subsystem populations. Specifically, the subsystem A population obtained using density-corrected EMFT is 13.2 for \( n = 2 \) and 39.3 for \( n = 6 \), both of which are essentially identical to the results obtained using conventional DFT at the high level (B3LYP/6-31G*) of theory.

The discrepancy for the subsystem A population obtained from EMFT with BO partitioning reflects the fact that the subsystem A block of the density matrix in the BO basis differs from that in the AO basis. This can be seen from the relation of the subsystem A block of the density matrix between the BO basis and the AO basis (from the inverse of the transformation in Eq. (26)),

\[
\tilde{D}^{AA} = D^{AA} + P^{AB}D^{BB}(P^{AB})^T + D^{AB}(P^{AB})^T + P^{AB}D^{BA},
\]

(30)

where \( P^{AB} \) is defined in Eq. (19). The subsystem A block of the density matrix in the BO basis, \( \tilde{D}^{AA} \), includes not only the subsystem A block, but also contributions from the subsystem B block and off-diagonal blocks of the density matrix in the AO basis. This "contamination" of the subsystem A density using BO partitioning leads to the qualitative discrepancy for the subsystem populations (Table 1, "BO" columns). Density-corrected EMFT, on the other hand, purifies the contamination by transforming the density matrix in the BO basis back to the AO basis and then directly applying the partitioning in the AO basis, thus producing subsystem populations that are consistent with the conventional DFT results (Table 1, "DC" columns). In later examples, we see that this contamination of the subsystem A density using BO partitioning can also manifest as larger errors for the calculated reaction energies, particularly in strongly conjugated systems. In the absence of the catastrophic errors associated with the unphysical collapse of the self-consistent EMFT solution, other sources of embedding error have been described previously,\textsuperscript{36} such as basis-set mismatch, spurious charge flow, and the scheme for modeling the exact exchange interactions; these errors typically decrease as a function of the size of subsystem A.
Hydrogenation of pentacene

We now compare the various EMFT partitioning schemes for a series of chemical reactions. The first test case is the terminal hydrogenation reaction of pentacene (Fig. 3a). This is a challenging case for embedding methods because the partitioning of the subsystems occurs across aromatic conjugation. For example, the ONIOM method requires \textit{a priori} specification of "link atoms" and the spin state for subsystems A and B; and for both singlet and triplet spin specifications, ONIOM exhibits large errors for this reaction that do not significantly decay with the increase of the subsystem A size.\textsuperscript{36} EMFT avoids such \textit{a priori} specifications and gives much better performance than ONIOM, although EMFT with AO partitioning can lead to unphysical collapse of the self-consistent solution (Table 1).

Figure 4 presents the errors in the reaction energy obtained using EMFT with AO and BO partitioning and the density-corrected implementation. For PBE-in-LDA with both same-basis (Fig. 4a) and mixed-basis (Fig. 4b) embedding, EMFT with AO partitioning does not exhibit the unphysical collapse, and all three implementations of EMFT (AO partitioning, BO partitioning and density-corrected) give almost identical results.

For B3LYP-in-LDA with same-basis embedding (Fig. 4c and d), EMFT with AO partitioning does exhibit the unphysical collapse. For cases using the EX1 (Fig. 4d) exact exchange scheme, the collapse can be prevented with the use of a good initial guess and a small SCF step, and physically reasonable meta-stable stationary solutions can be found, although they still exhibit substantial errors. For cases with either EX0 or EX1 scheme, EMFT with BO partitioning and density-corrected EMFT avoid the unphysical collapse and show rapid decay of errors with the increase of the subsystem A size, though density-corrected EMFT is more accurate than EMFT with BO partitioning. Additionally, using density-corrected EMFT (black curves in Fig. 4c and d), EX0 and EX1 give similar performance; the most substantial difference appears for the case with the smallest subsystem A size ($n = 2$), which can be attributed to the lack of subsystem exchange interactions in EX0.
For B3LYP-in-LDA with mixed-basis embedding (Fig. 4e and f), EMFT with AO partitioning does not exhibit the unphysical collapse, and density-corrected EMFT shows performance almost identical to EMFT with AO partitioning. EMFT with BO partitioning shows mildly slower convergence of error using the EX0 scheme (Fig. 4e). The two exact exchange schemes (EX0 and EX1) perform similarly, when used with the density-corrected EMFT implementation (Fig. 4e and f).

To summarize the results for the hydrogenation of pentacene (Figure 4), the unphysical collapse is only seen for EMFT with AO partitioning for the case of B3LYP-in-LDA with same-basis embedding. For cases where EMFT with AO partitioning does not exhibit the unphysical collapse, all three partitioning schemes of EMFT show similar performance. Using density-corrected EMFT, the EX0 exact exchange scheme is comparable in accuracy to EX1.
Figure 4: Performance of EMFT with AO partitioning (dashed blue lines), BO partitioning (dotted red lines) and the density-corrected implementation (solid black lines), in terms of the error in the reaction energy for the terminal hydrogenation of pentacene, as a function of the size of subsystem A for various embedding scenarios: (a) PBE-in-LDA with same-basis; (b) PBE-in-LDA with mixed-basis; (c) B3LYP-in-LDA with same-basis using EX0; (d) B3LYP-in-LDA with same-basis using EX1; (e) B3LYP-in-LDA with mixed-basis using EX0; and (f) B3LYP-in-LDA with mixed-basis using EX1. For B3LYP-in-LDA with same-basis embedding using EX0 (panel c), the AO partitioning leads to unphysical collapse of the EMFT solution and thus the results are not shown. For B3LYP-in-LDA with same-basis embedding using EX01 (panel d), the unphysical collapse can be prevented with the use of a good initial guess and a small SCF step, and physically reasonable meta-stable solutions can be found. All errors are plotted relative to the high-level DFT results: 32.5 kcal/mol for B3LYP/6-31G*, and 22.3 kcal/mol for PBE/6-31G*. See Figure S5 in Supporting Information for the plot with smaller y-axis ranges.
The Diels-Alder reaction involving a conjugated polyene

We next consider the Diels-Alder reaction between 1,3-butadiene and conjugated octadecanonene (Fig. 3b), which also involves partitioning across a conjugated bonding network. Figure 5 presents the error in the reaction energy obtained using EMFT with AO and BO partitioning and the density-corrected implementation. For PBE-in-LDA with both same-basis (Fig. 5a) and mixed-basis embedding (Fig. 5b), EMFT with AO partitioning does not exhibit the unphysical collapse, and density-corrected EMFT gives essentially the same results as EMFT with AO partitioning. For cases in which subsystem A is small, EMFT with BO partitioning differs somewhat from the other two partitioning schemes.

For B3LYP-in-LDA with same-basis embedding (Fig. 5c and d), EMFT with AO partitioning again leads to unphysical collapse of the EMFT solutions. For cases using the EX1 scheme (Fig. 5d), the unphysical collapse can be prevented and physically reasonable meta-stable solutions can be found with the use of a good initial guess and a small SCF step, leading to small errors for all subsystem sizes. Both EMFT with BO partitioning and the density-corrected implementation avoid the unphysical collapse, and using EX0 (Fig. 5c), give comparable error for cases in which subsystem A includes four or more carbon atoms. Using EX1, density-corrected EMFT is more accurate than EMFT with BO partitioning for small subsystem sizes.

For B3LYP-in-LDA with mixed-basis embedding (Fig. 5e and f), EMFT with AO partitioning does not exhibit the unphysical collapse. In these cases, density-corrected EMFT gives almost identical results to EMFT with AO partitioning, whereas EMFT with BO partitioning gives large errors for cases with two and four carbons in subsystem A.

To summarize the results for this reaction, the unphysical collapse is again only seen for EMFT with AO partitioning for the case of B3LYP-in-LDA with same-basis embedding. For cases in which EMFT with AO partitioning does not exhibit the unphysical collapse, density-corrected EMFT performs as well as EMFT with AO partitioning; and for cases in which EMFT in which AO partitioning does lead to unphysical collapse,
the performance of density-corrected EMFT remains very good. In general, for small subsystem sizes, EMFT with BO partitioning gives larger errors, due to the contamination of the subsystem A density, as described in section *Elimination of the unphysical collapse*; comparison of the results for this reaction and the hydrogenation of pentacene indicates that the magnitude of the errors for BO partitioning with small subsystem A sizes vary somewhat with the specific reaction, perhaps due to fortuitous cancellation of errors. Using density-corrected EMFT, the EX0 scheme performs similarly to EX1.
Figure 5: Performance of EMFT with AO partitioning (dashed blue lines), BO partitioning (dotted red lines) and the density-corrected implementation (solid black lines), in terms of the error in the reaction energy for the Diels-Alder reaction between the 1,3-butadiene and conjugated octadecanonene, as a function of the size of subsystem A. The x-axis indicates the number of carbon atoms from the octadecanonene chain that are included in subsystem A. Errors are obtained from the following embedding scenarios: (a) PBE-in-LDA with same-basis; (b) PBE-in-LDA with mixed-basis; (c) B3LYP-in-LDA with same-basis using EX0; (d) B3LYP-in-LDA with same-basis using EX1; (e) B3LYP-in-LDA with mixed-basis using EX0; and (f) B3LYP-in-LDA with mixed-basis using EX1. For B3LYP-in-LDA with same-basis embedding using EX0 (panel c), the AO partitioning leads to unphysical collapse of the EMFT solution and thus the results are not shown. For B3LYP-in-LDA with same-basis embedding using EX01 (panel d), the unphysical collapse can be prevented with the use of a good initial guess and a small SCF step, and physically reasonable meta-stable solutions can be found. All errors are plotted relative to the high-level DFT results: 21.1 kcal/mol for B3LYP/6-31G*, and 27.4 kcal/mol for PBE/6-31G*. See Figure S6 in Supporting Information for the plot with smaller y-axis ranges.

Deprotonation of decanoic acid

As a third example, we consider the deprotonation reaction of decanoic acid (Figure 3c). Unlike the previous examples, the subsystems are partitioned across a single covalent bond in this case. Figure 6 shows the error in the reaction energy obtained using EMFT with AO and BO partitioning and the density-corrected implementation. For this
reaction, EMFT with mixed-basis embedding (Fig. 6b, e and f) converges significantly more slowly than EMFT with same-basis embedding (Fig. 6a, c, and d), which has been identified previously. For PBE-in-LDA with both same-basis (Fig. 6a) and mixed-basis (Fig. 6b) embedding, where AO partitioning does not exhibit the unphysical collapse, all three implementations (AO and BO partitioning, and density-corrected) of EMFT give essentially the same results in terms of the reaction energies.

For B3LYP-in-LDA with same-basis embedding (Fig. 6c and d), EMFT with AO partitioning again exhibits the unphysical collapse. Using EX1 (Fig. 6d), this collapse can be prevented and physically reasonable meta-stable solutions can be found with the use of a good initial guess and a small SCF step. EMFT with BO partitioning and the density-corrected implementation avoid the unphysical collapse and give very similar results to EMFT with AO partitioning, when available. Interestingly, using EX0, EMFT with BO partitioning is somewhat more accurate for this system, although density-corrected EMFT also gives reasonable accuracy. For B3LYP-in-LDA with mixed-basis embedding (Fig. 6e and f), EMFT with AO partitioning does not exhibit the unphysical collapse, and all three implementations of EMFT show nearly identical performance.
Figure 6: Performance of EMFT with AO partitioning (dashed blue lines), BO partitioning (dotted red lines) and the density-corrected implementation (solid black lines), in terms of the error in the reaction energy for the deprotonation of decanoic acid, as a function of the size of subsystem A for various embedding scenarios: (a) PBE-in-LDA with same-basis; (b) PBE-in-LDA with mixed-basis; (c) B3LYP-in-LDA with same-basis using EX0; (d) B3LYP-in-LDA with same-basis using EX1; (e) B3LYP-in-LDA with mixed-basis using EX0; and (f) B3LYP-in-LDA with mixed-basis using EX1. For B3LYP-in-LDA with same-basis embedding using EX0 (panel c), the AO partitioning leads to unphysical collapse of the EMFT solution and thus the results are not shown. For B3LYP-in-LDA with same-basis embedding using EX01 (panel d), the unphysical collapse can be prevented with the use of a good initial guess and a small SCF step, and physically reasonable meta-stable solutions can be found. All errors are plotted relative to the high-level DFT results: 363.4 kcal/mol for B3LYP/6-31G*, and 363.1 kcal/mol for PBE/6-31G*. See Figure S7 in Supporting Information for the plot with smaller y-axis ranges.
Stone-Wales defects in a graphene sheet

We finally apply the EMFT partitioning schemes to the formation of a Stone-Wales defect in a graphene sheet (Figure 7). It was previously shown that EMFT is a promising method for studying carbon nanomaterials, consistently providing improved accuracy over the ONIOM method for such systems.

![Reactant, Transition Structure, Product](image)

Figure 7: Partitioning of the graphene sheet for EMFT calculations for the formation of Stone-Wales defects. The atoms inside the red and blue boxes correspond to the high-level subsystems (subsystem A) with 6 and 16 carbon atoms, respectively.

Figure 8 presents EMFT results for the reaction and activation energies associated with the Stone-Wales defect formation, using B3LYP-in-LDA with mixed-basis embedding and either EX0 or EX1 for the exact-exchange coupling. As seen previously, EMFT with AO partitioning and mixed-basis embedding does not exhibit the unphysical collapse, and the resulting errors are small for both the activation and reaction energies; even with only a one-carbon-wide border around the reaction site (i.e., with only six carbon atoms in subsystem A), EMFT with AO partitioning yields results that are within a few kcal/mol of the full high-level calculation. Also as seen before, for such cases in which EMFT with AO partitioning does not exhibit the unphysical collapse, the density-corrected EMFT yields very similar results. In this highly-conjugated system, EMFT with BO partitioning exhibits larger errors, particularly when used with EX1; this is again attributed to the contamination of the subsystem A density using BO partitioning, as described in section Elimination of the unphysical collapse. Nonetheless, for EX0, EMFT with BO partitioning yields comparable results to the other partition-
ing schemes when a two-carbon-wide border is included around the reaction site (i.e., with sixteen carbon atoms in subsystem A).

Taken together, the various numerical examples indicate that EMFT with AO partitioning can exhibit unphysical collapse when same-basis embedding is used along with a pair of exchange-correlation functionals that differ substantially in functional form. However, the more likely employment of EMFT with AO partitioning and mixed-basis embedding (i.e., when a smaller basis set is used for the environment) appears to be far less susceptible to this problem. Moreover, the unphysical collapse can be strictly eliminated with the use of either EMFT with BO partitioning or with density-corrected EMFT; the former has the advantage of remaining fully self-consistent, while the latter appears to be more accurate in strongly conjugated applications, such as the Stone-Wales application shown here and in the previously discussed examples of the hydrogenation of pentacene and the Diels-Alder reaction.
Figure 8: Performance of EMFT with AO partitioning (dashed blue lines), BO partitioning (dotted red lines) and the density-corrected implementation (solid black lines) for the defect formation in a graphene sheet, as a function of the size of subsystem A, in terms of: (a) error in the reaction energy using EX0; (b) error in the activation energy using EX0; (c) error in the reaction energy using EX1; and (d) error in the activation energy using EX1. All calculations are carried out within the framework of B3LYP-in-LDA with mixed-basis embedding. 6-31G* is used as the high-level AO basis and STO-3G is employed as the low-level AO basis. All errors are plotted relative to the high-level (B3LYP/6-31G*) DFT results: 71.7 kcal/mol for the reaction energy, and 215.1 kcal/mol for the activation energy.
Conclusions

In this work, we demonstrate that EMFT with atomic-orbital (AO) partitioning can exhibit unphysical collapse of the self-consistent solution, particularly when same-basis embedding is combined with a substantial mismatch in the functional form of the high- and low-level mean-field energy expressions. To address this issue, we introduce the alternative block-orthogonalized (BO) partitioning scheme that eliminates this problem, at negligible additional computational cost. Additionally, we introduce a non-self-consistent implementation of EMFT, akin to density-corrected density functional theory,\textsuperscript{39–46} in which the density matrix is obtained via minimization of the EMFT energy with BO partitioning and then used for evaluating the EMFT energy with AO partitioning.

A diverse set of numerical tests reveals that EMFT with BO partitioning offers an advantage over AO partitioning by strictly avoiding the unphysical collapse, but for cases in which AO partitioning does not exhibit collapse, it typically provides reaction energies that are more accurate than BO partitioning. Density-corrected EMFT is seen to also avoid the problem of unphysical collapse while yielding reaction energies that are generally better than BO partitioning, especially for strongly conjugated systems; however, it should be noted that this non-self-consistent implementation of EMFT is less convenient for the calculation of gradients and other response properties.

Although the current work confirms that EMFT with AO and BO partitioning is promising for many molecular applications, future work will seek to develop partitioning schemes to achieve the combined stability and accuracy of density-corrected EMFT within a fully self-consistent EMFT framework. Additionally, we note that employment of BO partitioning is not specific to EMFT, and it may prove useful for other subsystem embedding approaches that require partitioning with respect to a one-particle basis set.

A final advantage of the refined partitioning schemes introduced here is to enable comparison of the previously introduced EX0 and EX1 descriptions of exact exchange coupling between subsystems in EMFT. Over a broad range of systems, the EX0 scheme
is generally comparable in accuracy to the EX1 scheme, but at a much lower computational cost.

Acknowledgement

This material is based upon work performed by the Joint Center for Artificial Photosynthesis, a DOE Energy Innovation Hub, supported through the Office of Science of the U.S. Department of Energy under Award Number DE-SC0004993. Additionally, T.F.M. acknowledges support from a Camille Dreyfus Teacher-Scholar Award, and F.R.M. acknowledges funding from EPSRC (EP/M013111/1).

ASSOCIATED CONTENT

Supporting Information

Additions figures are provided for the application of EMFT using the EX2 exact exchange coupling scheme, as well as versions of Figs. 4-6 with decreased y-axis ranges. Also provided are geometries in xyz format for all test molecules presented in this paper. This material is available free of charge via the Internet at http://pubs.acs.org.

References


EMFT energy vs. Fraction of HF exchange

BHxLYP-in-LDA/6-31G*