Extending density functional embedding theory for covalently bonded systems

Kuang Yu<sup>a,1</sup> and Emily A. Carter<sup>b,2</sup>

<sup>a</sup>Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544-5263; and <sup>b</sup>School of Engineering and Applied Science, Princeton University, Princeton, NJ 08544-5263

Contributed by Emily A. Carter, November 1, 2017 (sent for review July 18, 2017; reviewed by Garnet Chan and Troy Van-Voorhis)

Quantum embedding theory aims to provide an efficient solution to obtain accurate electronic energies for systems too large for full-scale, high-level quantum calculations. It adopts a hierarchical approach that divides the total system into a small embedded region and a larger environment, using different levels of theory to describe each part. Previously, we developed a density-based quantum embedding theory called density functional embedding theory (DFET), which achieved considerable success in metals and semiconductors. In this work, we extend DFET into a density-matrix-based nonlocal form, enabling DFET to study the stronger quantum couplings between covalently bonded subsystems. We name this theory density-matrix functional embedding theory (DMFET), and we demonstrate its performance in several test examples that resemble various real applications in both chemistry and biochemistry. DMFET gives excellent results in all cases tested thus far, including predicting isomerization energies, proton transfer energies, and highest occupied molecular orbital–lowest unoccupied molecular orbital gaps for local chromophores. Here, we show that DMFET systematically improves the quality of the results compared with the widely used state-of-the-art methods, such as the simple capped cluster model or the widely used ONIOM method.

Significance

Quantum mechanics (QM) simulations increasingly help researchers understand phenomena in chemistry, geoscience, material science, and biology. Unfortunately, the exact solution of the relevant QM equations is not available; consequently approximate yet accurate QM methods must be developed. In many situations, e.g., heterogeneous catalysis and enzymatic reactions, although the entire system is too large to treat with high-level QM, the relevant phenomena happen in a confined space. Quantum embedding theories can study exactly such cases by using highly accurate but expensive QM methods to simulate only the region of interest, while using fast but less accurate approximations to treat the surroundings. In this article, we present a unique, accurate quantum embedding theory applicable to covalently bonded systems.

Author contributions: K.Y. and E.A.C. designed research; K.Y. performed research; K.Y. and E.A.C. analyzed data; and K.Y. and E.A.C. wrote the paper.

Reviewers: G.C., California Institute of Technology; and T.V.-V., Massachusetts Institute of Technology.

The authors declare no conflict of interest.

Published under the PNAS license.

<sup>1</sup>Present address: DEShaw Research, New York, NY 10036.

<sup>2</sup>To whom correspondence should be addressed. Email: eac@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712611114/-/DCSupplemental.
dangling bond states that disrupt the electronic structure of the embedded region and cause severe convergence issues for embedded self-consistent field (SCF) calculations. This reflects a problem that is universal for essentially all OFET schemes: the local embedding potential utilized in OFET is often unable to capture the correct quantum couplings in a covalent bonding environment. An embedding potential is, in a sense, analogous to pseudopotentials used to describe core electrons and nuclei. In both cases, we partition a full quantum system into two parts (i.e., embedded region/environment or valence electrons/core electrons plus nuclei) and aim to replace one part using an effective Hamiltonian. Inspired by the idea of nonlocal pseudopotentials, here we aim to extend the DFET into a nonlocal form and examine its performance to describe typical covalently bonded systems.

The uniqueness of the local embedding potential in DFET essentially is rooted in the Hohenberg–Kohn theorem (10, 18), which states that there is a one-to-one correspondence between the local external potential and the density. Similarly, in density-matrix functional theory, Gilbert proved that there is a one-to-one correspondence between the nonlocal external potential and the density matrix (19). We therefore need to use the one-particle density matrix as the target to reproduce, upon introduction of the embedding potential’s nonlocal form, leading to a density-matrix-based embedding method denoted here as density-matrix functional embedding theory (DMFET).

Several other embedding theories have been developed in recent years that utilize the density matrix as the working variable. In particular, Manby et al. (20) devised a density-matrix–based level-shifting projector in the embedded region Hamiltonian to enforce the orthogonality between the embedded region and the environment molecular orbitals (MOs). Meanwhile, Knizia and Chan (21) developed the density-matrix embedding theory (DMET), which was further studied and developed by Bulik et al. (22). Both embedding schemes have successfully dealt with strong quantum couplings; however, they both need to explicitly orthogonalize subsystem orbitals. Here, we try to develop a conceptually simpler and easy-to-implement alternative algorithm that avoids explicit orbital manipulations, while still achieving relatively high accuracy in typical covalently bonded systems. We show that with a single optimized effective potential (OEP) process, we can uniquely partition the total density matrix and obtain the corresponding nonlocal embedding potential simultaneously. We then perform accurate CW/DFT embedding calculations to study chemical reactions, utilizing the nonlocal embedding potential obtained by DMFET. Finally, we call attention to recent related work in the density-matrix embedding field: the bootstrap embedding theory of Welborn et al. (23). The bootstrap embedding scheme aims to optimize the wavefunction of each subsystem consistently using a high-level theory, so it is most conceptually similar to fragmentation methods such as partition DFT (6). By comparison, our work pursues a hybrid CW/DFT description of the system, as we focus on one particularly interesting region to treat with high-accuracy methods and simplify the description of the environment to that of DFT. Although less consistent theoretically across the entire system, the hybrid CW/DFT approach greatly reduces the computational cost and is still accurate enough for many applications.

In the following sections, we first briefly review the theoretical framework of DFET and discuss the necessary extensions needed to formulate DMFET. We then introduce numerical details of the implementation, followed by several proof-of-concept test cases resembling real chemical and biochemical applications. After we demonstrate DMFET’s capability for covalent embedding, we summarize our work and offer concluding remarks in the final section.

Theories and Implementation Details

**DFET.** In the original DFET (10), a DFT calculation on the total system is first performed to find the density \( \rho_{\text{tot}} \). The total system then is partitioned into two parts, denoted as subsystem A (embedded region) and subsystem B (environment), respectively. The electron numbers in each subsystem also are specified and kept fixed in all subsequent calculations:

\[
\begin{align*}
\int \rho_A(\vec{r}) d\vec{r} &= n_A \\
\int \rho_B(\vec{r}) d\vec{r} &= n_B.
\end{align*}
\]

Here, \( \rho_A \) and \( \rho_B \) are the densities of the corresponding subsystems. In principle, this partition is arbitrary, and one specific partition is not “more physical” than others. To remove this arbitrariness, we apply an extra constraint that the embedding potential of subsystem A must be identical to the embedding potential of subsystem B:

\[
V_{\text{emb}}^{A\rightarrow B} = V_{\text{emb}}^{B\rightarrow A} = V_{\text{emb}}.
\]

We then search for a common embedding potential \( V_{\text{emb}} \), under which the summation of the subsystem densities exactly reproduces the total density:

\[
\rho_{\text{tot}} = \rho_A[V_{\text{emb}}] + \rho_B[V_{\text{emb}}].
\]

Here, the subsystem densities are computed by solving the embedded Kohn–Sham (KS) equations:

\[
\begin{align*}
-\frac{1}{2} \nabla^2 + V_{\text{emb}}(\vec{r}) + \phi_A(\vec{r}) &= e_A^i \phi_A(\vec{r}) \\
-\frac{1}{2} \nabla^2 + V_{\text{emb}}(\vec{r}) + \phi_B(\vec{r}) &= e_B^i \phi_B(\vec{r})
\end{align*}
\]

\[
\sum_i (\phi_A^i)^2 = \rho_A \\
\sum_i (\phi_B^i)^2 = \rho_B.
\]

In practice, this common embedding potential is found by maximizing the following extended Wu–Yang OEP expression (24):

\[
W[V_{\text{emb}}] = E_A[\rho_A] + E_B[\rho_B] + \int (\rho_A + \rho_B - \rho_{\text{tot}}) \cdot V_{\text{emb}} d\vec{r}.
\]

It can be proven that the derivative of the extended Wu–Yang functional is simply \( \rho_A + \rho_B - \rho_{\text{tot}} \) (10); thus, Eq. 3 is naturally satisfied when \( W \) reaches its maximum.

**DMFET.** As mentioned in the Introduction, the local form of DFET’s embedding potential is insufficient to capture the quantum interactions in covalent bonds. Two major problems arise in applying local DFET to covalently bonded systems. First, it has long been known that exchange interactions are critical components of covalent bonding (25, 26) and, unlike the DFET embedding potential, the exchange potential of the environment orbitals acting on the embedded region wavefunction is inherently nonlocal. Second, cleavage of covalent bonds produces two open quantum systems, with dangling bond states that are not well described with single-determinant wavefunctions. In this work, we deal with both challenges by introducing two modifications to DFET that comprise the key components of our DMFET method.

E10862 | www.pnas.org/cgi/doi/10.1073/pnas.1712611114

Yu and Carter
To deal with the first problem, it is natural to extend the local embedding potential \( V_{\text{emb}}(\mathbf{r}) \) to a two-index nonlocal form: \( V_{\text{emb}}(\mathbf{r}, \mathbf{r}') \). In this situation, the off-diagonal elements of the density matrix must be included explicitly in the theory, as these elements directly respond to the off-diagonal components of the nonlocal potential.

Therefore, the first modification in DMFET is to replace \( V_{\text{emb}}(\mathbf{r}) \) using \( V_{\text{emb}}(\mathbf{r}, \mathbf{r}') \) and replace the local density \( n(\mathbf{r}) \) using the one-particle density matrix \( D(\mathbf{r}, \mathbf{r}') \). We then search for a common nonlocal embedding potential \( V_{\text{emb}}(\mathbf{r}, \mathbf{r}') \), under which the density matrix of the total system can be exactly reproduced by summation of the subsystem density matrices:

\[
D_{\text{tot}}(\mathbf{r}, \mathbf{r}') = D_A(\mathbf{r}, \mathbf{r}') + D_B(\mathbf{r}, \mathbf{r}').
\]  

The nonlocal embedding potential generated by fitting density matrices significantly stabilizes the embedded SCF calculations numerically compared with local DFET. However, we find that this simple extension alone faces a fundamental challenge, namely that such a nonlocal embedding potential that satisfies Eq. 7 may not exist. In fact, if we consider two bonded hydrogen atoms with a minimal basis set (i.e., one atomic orbital centered on each of the two nuclei), it can be proved that a common nonlocal embedding potential does not exist. More generally, we can show that the standard extended Wu–Yang OEP scheme is not likely to find the desired nonlocal embedding potential for any homonuclear diatomic molecule with any reasonable basis set (see SI Analytical Analysis of Homonuclear Diatomic Molecules for details). The analysis in the Supporting Information indicates that this newly emerged existence problem is related to the fact that DMFET enforces a much stronger constraint compared with DFET. All of the elements of the total density matrix must be reproduced in DMFET, whereas in DFET, only the total (local) density must be reproduced.

In fact, the existence of the embedding potential in DMFET is connected with the second open system problem mentioned above. In the extreme case of a homonuclear diatomic molecule, if the two bonding electrons are partitioned equally between the two subsystems, the DMFET formalism results in symmetric electron distributions for the embedded subsystems. This requires a very strong embedding potential that delocalizes the dangling bond orbital of the subsystem to imitate the truly symmetric covalent bond orbital in the molecule. As we show in SI Analytical Analysis of Homonuclear Diatomic Molecules, this embedding potential may not exist. Or, even if it exists in a more general situation, it would lead to extremely problematic electronic structures difficult to describe using single determinants. A seemingly straightforward solution is to partition the electron density unequally (i.e., assign both bonding electrons to one subsystem), generating two well-defined closed-shell subsystems. However, this leads to charged subsystems that carry a long-range electrostatic potential that is difficult to capture from the density (matrix) matching scheme: because the electron density (matrix) is mostly localized on the molecule, it is not sensitive to any long-range behavior of the potential. To address these issues, we introduce a second modification to DMFET: a flexible capping scheme that is compatible with the essential idea of DMFET.

In this capping scheme, we change the common embedding potential assumption in DFET, replacing Eq. 2 with the following relationship:

\[
V_{\text{emb}}^A(\mathbf{r}, \mathbf{r}') = V_{\text{emb}}^A(\mathbf{r}, \mathbf{r}') - \Delta V(\mathbf{r}, \mathbf{r}').
\]  

Here, \( \Delta V \) is a potential that is predetermined before the optimization process, and, as we can easily see, the common embedding potential assumption is restored when \( \Delta V = 0 \). We can prove that the embedding potentials for both subsystems \( [V_{\text{emb}}^A(\mathbf{r}, \mathbf{r}') \text{ and } V_{\text{emb}}^B(\mathbf{r}, \mathbf{r}') ] \) are still uniquely defined if \( \Delta V \) is fixed, and thus the density-matrix partition \( (D_A \text{ and } D_B) \) is also unique. The technical details of the proof can be found in SI Proof of Uniqueness for DMFET Embedding Potential and are a natural generalization of Hohenberg and Kohn’s (18) and Gilbert’s (19) works.

In practice, we can use the \( \Delta V \) term to implement capping potentials for the subsystems under the theoretical framework of DMFET. In this work, we always assign both bonding electrons to the embedded region (subsystem A) when we cleave a covalent bond, leaving a cation in the environment (subsystem B). The embedding potentials for both subsystems therefore can be written equivalently as:

\[
\begin{align*}
V_{\text{emb}}^A &= V_{\text{emb}} + V_{\text{cap}}^A, \\
V_{\text{emb}}^B &= V_{\text{emb}} + V_{\text{cap}}^B,
\end{align*}
\]  

assuming \( \Delta V = V_{\text{cap}}^A - V_{\text{cap}}^B \). Here, \( V_{\text{cap}} = V(q) \) is the potential of a positive point charge placed to cap the dangling bonds of the embedded region, and \( V_{\text{cap}} = V(-q) \) is the potential of a negative point charge used to cap the environment. For subsystem A, this setup is very similar to the link-atom approach in QM/MM calculations (27), in which the dangling bonds are typically saturated with hydrogen atoms. However, a conceptual difference exists in that we do not equally partition the bonding electrons here: the capping potential is a positive point charge instead of a hydrogen atom for subsystem A. Consequently, subsystem B, which also needs to be treated using DFT in DMFET, is capped with a counter point charge, which is not necessary in QM/MM.

We note that this choice of \( \Delta V \) is neither unique nor rigorously optimal. In principle, there is no constraint on the form of the capping potentials, so we can design more flexible nonlocal pseudopotentials to suit different bonding situations. For example, Huang et al. (28) recently proposed an extension of DFET for covalently bonded systems by modifying the interfacial atoms (i.e., replacing a C atom with B-H groups). While they still aim to reproduce the local density instead of the density matrix, their treatment of the interfacial atoms can be considered as a special case of our theory with \( V_{\text{cap}} = V_{\text{cap}}(\text{B-H}) = V_{\text{nuc}}(C) \). There are many more possible choices; how to systematically optimize the capping potential in different environments we leave to future research. Here, we mainly focus on single bonds in typical organic molecules and find that a point charge potential is already adequate in these cases. Compared with conventional DFET, this capping potential plays a crucial role in localizing the subsystem density distributions and in stabilizing the dangling bond states. DMFET therefore is numerically much more robust compared with DFET for covalently bonded systems when facilitated by the capping potential.

The common part of the embedding potential (i.e., the \( V_{\text{emb}} \) term in Eq. 9) then is obtained by maximizing the nonlocal version of the extended Wu–Yang OEP:

\[
W[V_{\text{emb}}] = E_A[D_A] + E_B[D_B] + \int (D_A + D_B - D_{\text{tot}}) \cdot V_{\text{emb}} \, d\mathbf{r}.
\]  

Again, \( D_A \) and \( D_B \) are obtained via the embedded region and the environment calculations using the embedding potential in Eq. 9. It is trivial to prove that the derivative of \( W[V_{\text{emb}}] \) with respect to \( V_{\text{emb}} \) is \( (D_A + D_B - D_{\text{tot}}) \), so Eq. 7 is satisfied when \( V_{\text{emb}} \) is an extremum of \( W[V_{\text{emb}}] \). As with DFET, once we have obtained the embedding potential \( V_{\text{emb}} = V_{\text{emb}} + V_{\text{cap}} \) at the inexpensive DFT level of theory, we then can compute the energy at a more costly hybrid CW/DFT level, using the following formula:

Yu and Carter

PNAS | Published online December 4, 2017 | E10863
Here,  \( E_{\text{tot}} \) is the DFT energy of the total system, and  \( E_{A} \) and  \( E_{\text{emb}} \) are the embedded region energies including the density-matrix–potential interaction term (\( E_A = E_A^X + D_A^{\text{emb}} \)),  \( X = \text{CW, DFT} \). This energy formula has a direct connection with that of the popular ONIOM (29) method, which is equivalent to Eq. 11 with zero embedding potential acting on the embedded region  \( V_{\text{emb}} = 0 \). In covalently bonded cases, bare ONIOM without any capping potential suffers from dangling bond artifacts, yielding unreasonable results. Therefore, we compare DMFET with “cap-ONIOM,” which is the standard ONIOM with capping H-atoms, in the following section. The energy of this cap-ONIOM can be written in our notation as follows:

\[
E_{\text{ONIOM}}^{\text{CW/DFT}} = E_{\text{tot}}^{\text{DFT}} + E_{A}^{\text{CW}} - E_{A}^{\text{DFT}} - E_{\text{emb}}^{\text{DFT}}. \tag{12}
\]

We refer to cap-ONIOM as ONIOM hereafter for simplicity, since all following ONIOM calculations in this work are performed with the capping potential. All of the methods we denote with a “cap-” prefix [e.g., cap-HF (HF, Hartree–Fock)] refer only to a single nonequilibrium calculation for the capped embedded region (i.e., where  \( V_{\text{emb}} = 0 \) and  \( V_{\text{emb}} = V_{\text{cap}} \)). Results computed with Eqs. 11 and 12 demonstrate that the common part of the embedding potential  \( (V_{\text{emb}}) \) derived in DMFET does significantly reduce the error in ONIOM.

In contrast to the purely local, density-based DFET, DMFET uses information from the off-diagonal elements of the density matrix, which applies more constraints on the subsystem wavefunctions. In particular, if we use single-determinant methods (e.g., HF or KS-DFT) in DMFET, all of the subsystem and total density matrices in Eq. 7 are idempotent by construction. We can further prove (see SI Wavefunction Orthogonality in DMFET) that if the summation of two idempotent density matrices is also idempotent, then the MOs used to construct these two density matrices must be orthogonal to each other. Hence, by optimizing Eq. 10 at the HF or KS-DFT level of theory, we effectively enforce the orthogonality between the occupied orbitals of the two subsystems. This good property does not exist in the DFET-derived subsystem orbitals and is a direct consequence of introducing the extra information contained in the nonlocal density-matrix elements. Because of this dual orthogonality property, we can also use the maximum component of the MO overlap matrix (\( \max[C_{A,B}^{\text{SC}}] \)) with  \( C_{A,B} \) being the column-based MO vectors (as an additional in addition to max\[D_A + D_B - D_{\text{tot}}\]) to check the convergence of the extended Wu–Yang OEP. Note that the virtual orbitals of the embedded region are not necessarily orthogonal to the environment orbitals in DMFET; thus, the embedded region CW wavefunction is not exactly orthogonal with the environment wavefunction. In this sense, DMFET is less rigorous compared with other methods that explicitly construct a full orthogonal Fock space [e.g., DMET (21) and the Manby et al. (20) projector-based approach].

Because of the orthogonality between the two subsystems and the idempotencies of both subsystem density matrices, DMFET at mean-field levels of theory (e.g., HF or DFT) is effectively a localized molecular orbital (LMO) method. Mathematically, DMFET simply rotates the fully occupied MOs of the total system to construct two new sets of MOs localized on each subsystem, respectively. Compared with other LMO methods, which utilize somewhat arbitrary metrics [e.g., maximization of self-Coulomb energies (30)] to localize orbitals, DMFET is theoretically more satisfying as its uniqueness can be rigorously proved. Moreover, the localized MOs in DMFET are the natural eigenstates of the embedding potential simultaneously generated in the DMFET procedure. By contrast, in other LMO methods, several extra OEP calculations would be needed to obtain the embedding potential for every subsystem. Moreover, the DMFET formalism can be generalized to other non–single-determinant methods, in which more physics beyond simple LMOs can be included. To illustrate this point, we show how a finite smearing scheme helps the description of localized excited states in one of our test cases (see HOMO-LUMO Gaps).

Because an extended OEP calculation must be performed in DMFET, its computational cost is about one to two orders of magnitude higher than that of a regular DFT calculation. At first glance, it is unclear whether DMFET is more efficient than a simple ONIOM calculation with a larger embedded region. However, in three-dimensional (3D) materials (e.g., bulk semiconductors with covalent bonds), the size of the embedded region grows as  \( r^3 \), and the computational cost of the embedded CW (ECW) method grows even faster (e.g., for coupled-cluster theory with single, double, and perturbative triple excitations [CCSD(T)], it is \( (r^9) = r^{3} \)), which quickly becomes intractable. In these cases, DMFET provides a viable approach to significantly reduce the size of the poor-scaling calculations via introducing a linear prefactor in front of the DFT computational cost. Therefore, although it may be less favorable for simpler, lower-dimensional materials such as organic polymers, DMFET is a very efficient tool in studies of 3D systems.

Implementation and Numerical Details. Most of our previous DFET calculations were performed in a plane-wave (PW) basis set, as it corresponds to a 3D uniform grid that naturally depicts the 3D density distribution. However, density-matrix–based embedding methods are more straightforward to implement with Gaussian-type orbitals (GTOs) instead of PW basis sets. In a PW representation, the dimension of the full density matrix is too high to be affordable for any practical applications. The number of GTO basis functions is much smaller, so we use only the GTO basis set throughout this study for proof-of-concept purposes. For a medium-size GTO basis, DMFET is actually faster than DFET because DMFET works directly on matrix elements, so it does not require the real-space integrations (e.g., \( \int \rho \cdot V \text{d}r \)) needed in DFET. It also avoids the step of assembling the density matrix to compute the total real-space electron density. However, PW bases are still relevant because in extended system calculations, PWs behave much better than GTOs and are the more natural choice when a system has translational symmetry. A possible implementation of DMFET in a PW basis set would be to first project the PW density matrix onto a GTO basis set (31) and then perform the extended OEP in the GTO representation.

We now describe three types of test cases used in this study, including the isomerization energy of an alkene modeling part of the retinal molecule, proton transfer reactions between two carboxylic acids, and the highest occupied MO–lowest unoccupied MO (HOMO-LUMO) gap of Cu\textsuperscript{2+} coordinated with organic ligands. Unless explicitly specified, the def2-SVP basis set (32) was used in all calculations to expand the wavefunctions. We used the B3LYP (33–36) exchange-correlation functional for all subsystem calculations in the embedding potential construction process. For the embedded region, we used either HF or CISD (CI with single and double excitations) embedded in a B3LYP environment, denoted as HF/B3LYP or CISD/B3LYP, respectively. As introduced above, in DMFET, we always assign both bonding electrons to the embedded region and place a negative capping point charge at the embedded region atom position in the environment calculations. For the embedded region calculation, we cap the dangling bond, using a positive point charge, which is placed on the original bonding axis, at a distance roughly matching the natural bond lengths of the C-H bond. This bond length was selected to be 1.084 Å in the first alkene isomerization test case (see Alkene Isomerization) and was set to be 1.090 Å in the other
interaction between the two carbon atoms (and their associated hydrogen atoms) marked with red numbers 1 and 2 in Fig. 1 (referred to as C1 and C2 hereafter). In the ONIOM and DMFET calculations, we define the terminal ethenyl group that contains C1 as the environment, while the remaining part of the molecule is taken as the embedded region. The environment interacts with the embedded region via both the single covalent bond and the intramolecular nonbonding interactions between C1 and C2 and their associated hydrogen atoms. We observe that although ONIOM performs reasonably well around the (trans) energy minimum, it has a large error in this repulsive (cis) geometry, which is corrected by DMFET (Table 1).

We first compute \( \Delta E \) for the full molecule at the HF level of theory and use it as a reference to benchmark the results we obtained from HF/B3LYP ONIOM, HF/B3LYP DMFET, and cap-HF (the \( E^\text{H\text{F}}_{A}^{\text{cap}} \) term in Eq. 12). All of the results are listed in Table 1. Compared to the HF reference result, DFT-B3LYP yields an isomerization energy lower by about 15 kJ/mol, which is the deviation we hope to eliminate using the hybrid methods. In the simplest cap-HF calculation, the environment ethenyl group is truncated and replaced by a capping positive point charge. The truncated embedded region is essentially symmetric with respect to cis-trans isomerization, so we obtained a near-zero \( \Delta E \) for cap-HF (\( E^\text{H\text{F}}_{A}^{\text{cap}} \)). The remaining 2.36 kJ/mol of energy difference comes mainly from intramolecular basis set superposition error (BSSE). Similarly, the difference in \( \Delta E_{A}^{\text{DFT}} \left( V_{\text{cap}}^{A} \right) \) is also near zero, so the correction term in ONIOM (\( E^\text{H\text{F}}_{A}^{\text{cap}} - E^\text{DFT}_{A}^{\text{cap}} \)) is very small. Consequently, the ONIOM approach essentially does not improve the underlying DFT-B3LYP results, thus losing its advantage in this situation. By contrast, the DMFET embedding potential restores the asymmetry in the original molecule, thus reducing the DFT-B3LYP deviation from HF by over 50%.

After the advantage of DMFET at the HF/B3LYP level was verified, we then carried out CISD/B3LYP calculations to include dynamic correlation effects in the embedded region, as per ECW theory (8, 9). We found that DMFET at the CISD/B3LYP level overcorrects the error by 5.2 kJ/mol in DFT-B3LYP, compared with the full-system CISD benchmark. This problem is most likely related to the dispersion interaction between the C1 and C2 groups, which is not accounted for at the HF/B3LYP level. Our DMFET method solved for the environment orbitals at the B3LYP level of theory and then as usual kept them frozen in the subsequent ECW calculations. Nonlocal correlations between the environment and the embedded region thus are not included in the embedding potential at this level of theory, so DMFET could not capture the dispersion attraction between the

### Test Examples

#### Alkene Isomerization

The first test case we examined using DMFET is the cis-trans isomerization of 2-methyl-1,3,5-hexatriene. We chose this molecule as it models the core part of retinal (Fig. 1), the cis-trans isomerization of which plays a central role in human dark vision. In this study, we focus on predicting the energy difference between the cis and trans conformations of 2-methyl-1,3,5-hexatriene (\( \Delta E = E_{\text{cis}} - E_{\text{trans}} \)).

We first optimize the structure of the trans conformation at the HF/def2-SVP level of theory, and then we rotate the central double bond by 180° (Fig. 1) to create the cis conformation. In this way, we create a cis geometry that has a strong repulsive

---

**Table 1. The cis–trans isomerization energies of 2-methyl-1,3,5-hexatriene**

| Methods | \( \Delta E \), kJ/mol | Deviation, kJ/mol |
|---------|------------------------|------------------|
| HF      | 85.12                  |                  |
| DFT-B3LYP | 69.80                | −15.3            |
| ONIOM (HF/B3LYP) | 70.89             | −14.2            |
| DMFET (HF/B3LYP) | 77.92            | −7.2             |
| cap-HF   | −2.36                  | −87.5            |
| CISD     | 74.52                  |                  |
| DFT-B3LYP+D3 | 66.35             | −8.2             |
| ONIOM (CISD/B3LYP+D3) | 62.99           | −11.5            |
| DMFET (CISD/B3LYP+D3) | 73.68          | −0.8             |
| cap-CISD | −4.18                  | −78.7            |

Values in bold are the optimal ones obtained from DMFET.
two subsystems. The easiest solution to this problem is to incorporate dispersion into the lower-level method used to derive the nonlocal embedding potential and use DMFET to correct the errors that come from other sources (electrostatics, exchange, bonding interactions, etc.). We do so hereby employing the dispersion-corrected DFT-B3LYP-D3 method developed by Grimme et al. (42), the results of which are shown in Table 1. We see the same trend here as found in the HF/B3LYP calculations: cap-CISD completely misses the isomerization energy and CISD/B3LYP-D3-level ONIOM even introduces a qualitatively erroneous correction on top of B3LYP-D3. DMFET, on the other hand, corrects most of the errors and also generates a result that is in excellent agreement with the benchmark CISD calculation on the total system. We further repeated the calculations using a Dunning-style cc-pVDZ basis set (the size of cc-pVDZ is the same as def2-SVP, but the latter basis set is parameterized more specifically for DFT, while the former one performs better in CW calculations) (43), the results of which can be found in Table S1. These results are very similar despite numerical differences, thus proving that our conclusions are not very dependent on basis sets.

Proton Transfer Reactions. In this section, we examine the performance of DMFET in generating a continuous potential energy surface for chemical reactions. We use proton transfer reactions as our examples, since the proton transfer process exists universally in Brønsted acid/base catalyzed reactions. In biology, proton transfer reactions particularly are the central reactions in many molecular machines such as the proton pump, which a living cell uses to create an electrochemical potential gradient across a membrane to store energy. In these systems, proton transfer often occurs in a complicated environment with many nonbonding interactions with the solvent, as well as bonding interactions with large biomolecule backbones (i.e., proteins, nucleic acids, etc.). In many situations, a full QM description is required because these environment interactions are neither negligible nor easily approximated at the classical level. However, self-interaction and delocalization errors make conventional DFT methods inappropriate for describing charge-transfer reactions. The hybrid CW/DFT embedding schemes therefore have great potential in these situations by exploiting both the accuracy of CW to describe chemical reactions including charge transfer and the computational speed of DFT for the environment.

As a proof-of-principle study, we investigate simple proton transfer reactions between a formate anion and a functionalized

![Fig. 2. (A) CF₃- and (B) Bz(OH)₂-functionalized systems and the two proton transfer reactions studied in this work. The respective areas of the molecules inside the blue boxes are treated as environments, and the respective areas outside the blue boxes are treated as embedded regions.](image)

![Fig. 3. (A–D) HF/DFT-B3LYP results for proton transfer reactions. (A) Energy curves for the CF₃-functionalized system; (B) energy curves for the Bz(OH)₂-functionalized system; (C) relative energy errors (%) of DMFET and ONIOM for the CF₃-functionalized system; (D) relative energy errors (%) of DMFET and ONIOM for the Bz(OH)₂-functionalized system.](image)
carboxylate anion. The reactions are schematically plotted in Fig. 2, where the functional groups within the blue boxes [i.e., -CF$_3$ and -Bz(OH)$_3$] are treated as the environments. We first use the HF/def2-SVP method to fully optimize the structures with the proton attached to the formate, obtaining one geometry defined as "geometry I." We then manually move the proton to the other carboxylate anion, optimizing its position while keeping all other atoms fixed. The resulting structure is defined as "geometry II." The proton transfer reaction path then is defined as a linear interpolation between geometry I and geometry II:

$$\tilde{R}(x) = (1 - x) \cdot \tilde{R}_I + x \cdot \tilde{R}_II.$$  \[13\]

Here, \(\tilde{R}\) represents the atom positions, and \(x\) is the reaction coordinate defining the molecular structure.

In the test case shown in Fig. 2A, full-system HF and CIsD generate similar results, whereas full-system DFT-B3LYP severely underestimates the curvature of the potential energy surface (Fig. S2). These two systems thus should form a good test for the DMFET methods.

While scanning along this reaction path, we observed that the original DMFET procedure could not converge the embedding potentials to a precision high enough to generate a smooth potential energy curve. For example, we found that two different initial guesses yielded two qualitatively similar but numerically different embedding potentials, leading to \(\sim 1.6\) kJ/mol of uncertainty in the final DMFET energy. This error is not large but it is comparable to the physical energy difference between DMFET and ONIOM in this test case. The minimum region of the extended Wu–Yang OEP is so flat that simply increasing the convergence threshold leads to many more OEP steps, thereby increasing the computational cost significantly. We therefore adopted an alternative solution in which we assume that the embedding potential \((V(x))\) varies slowly with respect to geometry changes \((\partial V/\partial x \approx 0)\), so the following approximation holds:

![Graph](image1)

**Fig. 4.** (A and B) Energy curves (A) and relative errors (B) of the CF$_3$-functionalized system at the CIsD/DFT-B3LYP level of theory.

![Graph](image2)

**Fig. 5.** Elements of the subsystem density matrices for the CF$_3$-functionalized system of Fig. 2A, plotted as heat maps (computed at \(x = 0.5\), halfway along the proton transfer pathway). All GTO basis functions are sorted and grouped according to the respective atom groups to which they belong. All atom groups are labeled at the top of the heat maps and partitioned by the red dashed lines. The atom groups are arranged from left to right according to the structure previously shown in Fig. 2A, with the embedded region depicted in A and the environment in B.
Fig. 6. (A) CF$_3$- and (B) Bz(OH)$_2$-functionalized systems for HOMO-LUMO gap tests of the embedded region. The functional groups inside the blue boxes are treated as the environment in DMFET.

\[
\frac{dE}{dx} = \frac{\partial E}{\partial \mathbf{V}} \left( \frac{\partial \mathbf{V}}{\partial x} \right) \approx \frac{\partial E}{\partial \mathbf{V}} \cdot \frac{\partial \mathbf{V}}{\partial x} \approx \frac{\partial E}{\partial \mathbf{V}}.
\]

Under this approximation, we can compute the energy difference, instead of the energies, between two points \( (x, x + \Delta x) \), thus utilizing a fixed embedding potential derived at the middle point \( (x + \frac{1}{2} \Delta x) \):

\[
\begin{align*}
E(x + \Delta x; V(x + \Delta x)) - E(x; V(x)) & \approx E\left(x + \Delta x; V\left(x + \frac{1}{2} \Delta x\right)\right) \\
& - E\left(x; V\left(x + \frac{1}{2} \Delta x\right)\right).
\end{align*}
\]

In this way, the numerical noise in \( V(x) \) is largely canceled, as a consistent embedding potential is used for all of the energy calculations in Eq. 15. The energy curve then is derived by integrating the energy differences along the reaction coordinate \( x \). While this approximation works well in the current test case (as we show below), it may fail for other systems with rapidly changing embedding potentials. In cases where smooth, high-quality potential energy surfaces are needed, further research is required to improve the convergence quality in DMFET. However, we emphasize that the error caused by the OEP convergence issue (~1.6 kJ/mol) is relatively small for most applications, so this numerical issue is unlikely to be problematic for most chemical problems of interest.

Using the alternative solution stated above, the resulted DMFET potential energy surfaces at the HF/DFT-B3LYP level of theory are plotted in Fig. 3, along with the ONIOM and cap-HF results, as well as the total HF results as reference.

We see that the cap-HF method strongly underestimates the acidity difference between the two carboxylate groups due to its lack of environmental influence beyond a point charge. ONIOM describes the electronic withdrawing/donating effects of the environment functional groups at the DFT-B3LYP level, thereby achieving much better results. DMFET further refines the environmental effects on the electronic structure of the embedded region, thus significantly reducing the remaining error in ONIOM. We further performed a more expensive CISD/DFT-B3LYP-level calculation on the smaller CF$_3$-functionalized system, the results of which are shown in Fig. 4. A similar trend is observed at the CW (CISD) level of theory: DMFET outperforms ONIOM in the entire range even though ONIOM already provides a decent description of the proton transfer energies. We thus conclude that the DMFET embedding potential does successfully describe most of the interactions between the embedded region and the environment in this test example.

To gain further insight into the effects of DMFET, we examined the magnitude of the subsystem density-matrix elements in the GTO matrix representation, which are shown in Fig. 5. The embedded density matrices are properly localized within the corresponding subsystems, with some small, off-diagonal elements distributed between the two bonded carbon atoms. This is an important observation as it indicates that we can use a truncated subsystem basis set (with bonded neighbors treated as ghost atoms with just a basis set and no nuclear potential or pseudopotential) instead of using the full basis set to perform the DMFET calculations without losing much accuracy. This assumption needs to be verified; if it is, the cost of DMFET can be reduced.

HOMO-LUMO Gaps. Although DMFET outperforms ONIOM in energy predictions, ONIOM is still a fairly robust method in most situations and is much cheaper than DMFET as it does not require solving for an embedding potential. However, one of ONIOM’s most important shortcomings is that it is only an a posteriori energy correction, i.e., the electronic structure of the embedded region remains the same as that of an isolated molecular fragment in the absence of an environment. In contrast, DMFET generates an embedding potential that enters the Hamiltonian of the embedded region explicitly, thus correcting the embedded wavefunction directly. In this section, we use the HOMO-LUMO gap as a test quantity to verify that DMFET does indeed improve the description of the embedded region at the wavefunction level, in addition to improving energetics. Note that we do not aim to predict the exact absorption spectrum for these systems. Instead, we are more interested in band-gap shifts caused by the functional group in the environment, which is an effect completely missing in ONIOM.

The test examples in this section, presented in Fig. 6, are similar to the test examples in the previous proton transfer studies. Here, we aim to compute the HOMO-LUMO gap of a Cu$^+$ cation bidentate bonded to a carboxylate anion ligand functionalized by either a -CF$_3$ or a -Bz(OH)$_2$ functional group. Such organometallic complexes are commonly employed in bio-inorganic chemistry, homogeneous catalysis, and materials such as metal organic frameworks. The Cu states in these two complexes are localized within the Cu$^+$-carboxylate complex, which is crucial here since DMFET is not expected to describe any delocalized states spanning both the embedded region and the environment. However, the functional group in the environment will perturb the electron wavefunction inside the embedded region and cause shifts to the HOMO-LUMO gap of the Cu$^+$ complex. We first compute the gap by running DFT-B3LYP and HF calculations on the entire system, using these results as references at the corresponding level of theory. We then construct the DMFET embedding potential at the DFT-B3LYP level, computing the HOMO-LUMO gap of the embedded region.

Fig. 7. HOMO-LUMO gap errors of the cap-DFT-B3LYP, embedded DFT-B3LYP, cap-HF, and embedded HF levels of theory, computed with a smearing width of 0.05 hartree.
at both the HF and DFT-B3LYP levels either with ("embedded") or without ("cap") the obtained embedding potential. Although for testing purposes we could have constructed a HF DMFET embedding potential and then carried out a HF/HF embedding calculation to compare with total-system HF, our ultimate goal always is to treat the environment with at least the mean-field correlation of DFT, while the embedded region will be treated with CW. Here, HF and B3LYP are sufficient to compare these HOMO-LUMO gaps in this example. These results are compared with the corresponding HF or B3LYP reference values, and the absolute errors are presented in Fig. 7.

It is evident when comparing the results of embedded B3LYP or embedded HF with cap-B3LYP or cap-HF that the common aspect of the embedding potential in DMFET is critical to improving the HOMO-LUMO gap. In this case, DMFET is indeed superior to simple ONIOM, which employs only the capping potential for the embedded region calculations.

In Fig. S3, we show the dependence of the results in Fig. 7 on the Gaussian smearing widths employed in the embedding potential construction. We found that at the DFT-B3LYP/ DFT-B3LYP level, the improvement of DMFET over ONIOM shown in Fig. 7 is observed only when we use a large smearing width (>0.03 hartree) in the DMFET calculations. With small smearing, the performance of DMFET at the DFT-B3LYP/DFT-B3LYP level is in fact worse than the cap-DFT-B3LYP calculations. We believe this trend is due to matching only the total density matrix, because the density matrix is related only to the occupied orbitals. If we use a small smearing width (<0.02 hartree) in DMFET, then the OEP optimizer does not directly optimize the LUMO because it does not contribute to the density matrix. If instead a larger smearing width (>0.03 hartree) is employed, then the LUMO is occupied by a small yet finite fractional number of electrons, with the result that the embedding potential is affected by the LUMO shape and occupancy. The LUMO occupancy is directly related to the LUMO energy in the Gaussian smearing algorithm (41), so the density matrix will be especially sensitive to the HOMO-LUMO gap. This explains why we are able to predict the gap via matching the density matrix here.

This test example clearly demonstrates the capability of DMFET to describe both occupied and virtual orbitals. In principle, DMFET could be extended to match the density matrix of any wavefunction, even those that were not originally devised for ground-state single-determinant methods. In the current test case, the relevant orbital we are interested in is localized within the embedded region and is low lying in energy. Therefore, a simple eigenenergy-based fractional occupation method is sufficient to generate accurate results. However, the problem remains as to how to develop a more robust smearing or state-averaging algorithm for more general situations. The accuracy of the dynamic correlation energy in the subsequent CW calculation depends on the quality of the virtual space of the embedded region. Therefore, an algorithm to better populate the virtual space certainly would further improve DMFET.

Conclusions

In summary, starting with the density-based DFET, we formulated a unique density-matrix–based embedding theory called DMFET to treat covalently bonded systems. We proved that the density-matrix partition between the embedded region and the environment is uniquely defined within the theoretical framework of DMFET, and thus the resulting nonlocal embedding potential also is unique. Additionally, it was proved that the resulting MOs of the embedded region and the environment are mutually orthogonal when using single-determinant methods to construct the embedding potential. This DMFET method was implemented using a GTO basis set, which is appropriate for studies of molecules. The DMFET formalism also should be extendable to simulations of condensed matter described with PW basis sets by utilizing a projection algorithm to reduce the dimensionality of the density matrix. In this work, GTO-based DMFET was tested with three examples that model a variety of important applications, from retinal isomerization to proton transfer to organometallic chemistry. We found that DMFET was much more reliable for predicting alkene isomerization and proton transfer energies compared with the simple cap and cap-ONIOM methods at the same level of theory. The embedding potential obtained in DMFET also greatly improved the HOMO-LUMO gap for localized states within the embedded region. To obtain such an improvement in the HOMO-LUMO gap, a large Gaussian smearing width was needed to incorporate some virtual orbital information within ground-state DMFET. In principle, a more flexible fractional occupancy scheme beyond Gaussian smearing is possible, designed to achieve a better description of the virtual orbitals involved in dynamical correlation. Moreover, the character of the density matrix—showing nonzero off-diagonal elements only between groups at the interface between the embedded region and the environment—indicates that we should be able to improve the efficiency of the method by using subsystem bases augmented with neighboring “ghost” basis functions rather than the entire system's basis to derive the nonlocal embedding potential. With these results, we conclude that DMFET rigorously extends embedding potential calculations to covalently bonded molecules, with great promise for investigating complicated chemical and biochemical systems and phenomena in the future.

Acknowledgments. E.A.C. thanks the US Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0002120 and the US National Science Foundation under Award 1265700 for funding this project.

1. Sun Q, Chan GKL (2016) Quantum embedding theories. Acc Chem Res 49:2705–2712.
2. Cortona P (1991) Self-consistently determined properties of solids without band-structure calculations. Phys Rev B Condens Matter 44:8454–8458.
3. Cortona P (1992) Direct determination of self-consistent total energies and charge densities of solids: A study of the cohesive properties of the alkali halides. Phys Rev B Condens Matter 46:2008–2014.
4. Wesołowski TA, Warshel A (1993) Frozen density-functional approach for ab-initio calculations of solvated molecules. J Phys Chem 97:8050–8053.
5. Neugebauer J, Lounerie MJ, Baerends EJ, Wesołowski TA (2005) The merits of the frozen-density embedding scheme to model solvatochromatic shifts. J Chem Phys 122:094115.
6. Elliott P, Burke K, Cohen MH, Wasserman A (2010) Partition density-functional theory. Phys Rev A 82:024501.
7. Yezzi AD, Neugebauer J (2014) Subsystem density-functional theory. Wiley Interdisciplinary Rev Comput Mol Sci 4:325–382.
8. Huang P, Carter EA (2008) Advances in correlated electronic structure methods for solids, surfaces, and nanostuctures. Annu Rev Phys Chem 59:261–290.
9. Libisch F, Huang C, Carter EA (2014) Embedded correlated wavefunction schemes: Theory and applications. Acc Chem Res 47:2768–2775.
10. Huang C, Pavone M, Carter EA (2011) Quantum mechanical embedding theory based on a unique embedding potential. J Chem Phys 134:154110.
19. Gilbert TL (1975) Hohenberg-Kohn theorem for nonlocal external potentials. Phys Rev B 12:2111–2120.
20. Manby FR, Stella M, Goodpaster JD, Miller TF, 3rd (2012) A simple, exact density-functional-theory embedding scheme. J Chem Theory Comput 8:2564–2568.
21. Knizia G, Chan GKL (2013) Density matrix embedding: A strong-coupling quantum embedding theory. J Chem Theory Comput 9:1428–1432.
22. Bulik IW, Scuseria GE, Dukelsky J (2014) Density matrix embedding from broken symmetry lattice mean fields. Phys Rev B 89:035140.
23. Welborn M, Tsuchimochi T, Van Voorhis T (2016) Bootstrap embedding: An internally consistent fragment-based method. J Chem Phys 145:074102.
24. Bulik IW, Scuseria GE, Dukelsky J (2014) Density matrix embedding from broken symmetry lattice mean fields. Phys Rev B 89:035140.
25. Woodrow Wilson C, Goddard WA, III (1970) Exchange kinetic energy, contra-gradient, and chemical binding. Chem Phys Lett 5:45–49.
26. Goddard WA, III, Harding LB (1978) The description of chemical bonding from ab initio calculations. Annu Rev Phys Chem 29:363–396.
27. Senn HM, Thiel W (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. Phys Chem Chem Phys 7:3297–3305.
28. Vosko SH, Wilk L, Nusair M (1980) Accurate spin-dependent electron liquid correlation energies for local spin-density calculations: A critical analysis. Can J Phys 58:1200–1211.
29. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B Condens Matter 37:785–789.
30. Becke AD (1993) Density-functional thermochemistry. 3. The role of exact exchange. J Chem Phys 98:5648–5652.
31. Sanchez-Portal D, Artacho E, Soler JM (1995) Projection of plane-wave calculations into atomic orbitals. Solid State Commun 95:685–690.
32. Weihe A, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Ru: Design and assessment of accuracy. Phys Chem Chem Phys 7:3297–3305.
33. Vosko SH, Wilk L, Nusair M (1980) Accurate spin-dependent electron liquid correlation energies for local spin-density calculations: A critical analysis. Can J Phys 58:1200–1211.
34. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B Condens Matter 37:785–789.
35. Becke AD (1993) Density-functional thermochemistry. 3. The role of exact exchange. J Chem Phys 98:5648–5652.
36. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields. J Phys Chem 98:11623–11627.
37. Vallev M, et al. (2010) NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. Comput Phys Commun 181:1477–1489.
38. Byrd RH, Lu PH, Nocedal J, Zhu CY (1995) A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput 16:1190–1208.
39. Zhu C, Byrd RH, Lu P, Nocedal J (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large scale bound constrained optimization. ACM Trans Math Softw 23:550–560.
40. Morales JL, Nocedal J (2011) Remark on algorithm 778: L-BFGS-B: Fortran subroutines for large scale bound constrained optimization. ACM Trans Math Softw 38:1–4.
41. Warren RW, Dunlap BI (1996) Fractional occupation numbers and density functional energy gradients within the linear combination of Gaussian-type orbitals approach. Chem Phys Lett 262:384–392.
42. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132:154104.
43. Dunning TH (1989) Gaussian-basis sets for use in correlated molecular calculations. 1. The atoms boron through neon and hydrogen. J Chem Phys 90:1007–1023.