Incremental embedding: A density matrix embedding scheme for molecules

Hong-Zhou Ye, Matthew Welborn, Nathan D. Ricke, and Troy Van Voorhis

---

Articles you may be interested in

Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science
The Journal of Chemical Physics **149**, 180901 (2018); 10.1063/1.5052551

Curing basis-set convergence of wave-function theory using density-functional theory: A systematically improvable approach
The Journal of Chemical Physics **149**, 194301 (2018); 10.1063/1.5052714

Variational reduced density matrix method in the doubly occupied configuration interaction space using three-particle N-representability conditions
The Journal of Chemical Physics **149**, 194105 (2018); 10.1063/1.5056247

Communication: A mean field platform for excited state quantum chemistry
The Journal of Chemical Physics **149**, 081101 (2018); 10.1063/1.5045056

Matrix product state formulation of the multiconfiguration time-dependent Hartree theory
The Journal of Chemical Physics **149**, 194114 (2018); 10.1063/1.5051498

Quantum system partitioning at the single-particle level
The Journal of Chemical Physics **149**, 184104 (2018); 10.1063/1.5055942

Incremental embedding: A density matrix embedding scheme for molecules

Hong-Zhou Ye, Matthew Welborn,[a] Nathan D. Ricke, and Troy Van Voorhis[b]
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The idea of using fragment embedding to circumvent the high computational scaling of accurate electronic structure methods while retaining high accuracy has been a long-standing goal for quantum chemists. Traditional fragment embedding methods mainly focus on systems composed of weakly correlated parts and are insufficient when division across chemical bonds is unavoidable. Recently, density matrix embedding theory and other methods based on the Schmidt decomposition have emerged as a fresh approach to this problem. Despite their success on model systems, these methods can prove difficult for realistic systems because they rely on either a rigid, non-overlapping partition of the system or a specification of some special sites (i.e., "edge" and "center" sites), neither of which is well-defined in general for real molecules. In this work, we present a new Schmidt decomposition-based embedding scheme called *incremental embedding* that allows the combination of arbitrary overlapping fragments without the knowledge of edge sites. This method forms a convergent hierarchy in the sense that higher accuracy can be obtained by using fragments involving more sites. The computational scaling for the first few levels is lower than that of most correlated wave function methods. We present results for several small molecules in atom-centered Gaussian basis sets and demonstrate that incremental embedding converges quickly with fragment size and recovers most static correlation in small basis sets even when truncated at the second lowest level. *Published by AIP Publishing.* https://doi.org/10.1063/1.5053992

I. INTRODUCTION

The fast and accurate calculation of the quantum mechanical properties of molecules and materials is one of the major unsolved problems in quantum chemistry. The computational cost of most accurate electronic structure methods rises sharply with system size, limiting their applications to small systems and/or moderate-sized basis sets.[1–6] This scaling challenge can be potentially circumvented by fragment embedding, where the system is divided into smaller fragments, and the computationally involved, high-level theory is only required for each individual fragment. The complicated interaction between the fragment and its large-sized surroundings is then approximated by the interaction with an effective bath that mimics the rest of the system. The idea of fragment embedding serves as the basis for many methods, including fragment molecular orbital theory[7–10] (localized molecular orbital-based embedding), subsystem density functional theory (DFT)[11–14] (density-based embedding), and dynamic mean-field theory[15–19] (local Green's function-based embedding), to name a few.

A major challenge to the development of a general fragment-based method is the treatment of chemical bonds between fragments. Recently, Schmidt decomposition[20–22] has

been used for embedding fragments that are strongly correlated with a bath, which occurs when embedding fragments across chemical bonds. For each fragment, the Schmidt decomposition transforms the rest of the system into an entangled, effective bath which is of the *same* dimension as the fragment. A low-dimensional embedding Hamiltonian is then constructed in the Schmidt space and solved accurately therein. In practice, a high-level calculation such as FCI (full configuration interaction[23]), DMRG (density matrix renormalization group[24,25]), or CCSD (coupled-cluster singles and doubles[26,27]) is embedded in a low-level bath (usually mean-field, e.g., Hartree-Fock[23]) to recover the electron correlation missing at the mean-field level.

In order to optimize the embedding, some matching conditions are usually imposed. So far there have been two main classes. In the first class, DMET (density matrix embedding theory[28,29]) and DET (density embedding theory[30,31]), one uses rigid, non-overlapping fragments and matches the one-particle density matrix (1PDM) between the fragment and the bath. Mathematically this is achieved by adding to the low-level bath an effective one-particle potential, which changes both the low- and high-level 1PDMs. This effective potential is then tuned to satisfy the matching condition. This approach has shown good performance on model systems such as the Hubbard model and atomic rings/chains, even in the strong correlation domain.[28–33] As with many fragment embedding methods, however, the restriction to non-overlapping fragments results in persistent edge effects and slow convergence with fragment size.[32,34] In BE (Bootstrap

---
[a]Current address: Department of Chemistry, California Institute of Technology, Pasadena, CA 91125, USA.
[b]Electronic mail: tvan@mit.edu

Embedding[35]), one instead uses overlapping fragments and requires that in the overlapping region the "edge" sites from one fragment agree on density matrix elements with the "center" sites from another fragment. As long as one can make a clear distinction between the edge sites (usually on the boundary of a fragment) and the center sites (usually the most embedded part), this scheme reduces the edge effects and leads to faster convergence as demonstrated on the Hubbard model.[35]

Though successful for model systems, these methods encounter difficulties with real molecules. On the one hand, the need for a rigid, non-overlapping partition of the system makes DMET/DET ambiguous when high symmetries such as translational invariance are lost. As a consequence, applications to realistic systems have been so far restricted to atomic rings,[29,32] chains,[29] or simple polymers[31] in small basis sets where one atom or monomer with several basis functions can be considered as a fragment. One attempt at modeling real molecules was made by Wouters *et al.*[32] who performed DMET calculations for the potential energy surface of a symmetric S_N2 reaction, $C_{12}H_{25}F\cdots F^- \rightarrow C_{12}H_{25}F\cdots F^-$. However, the DMET result was less accurate than that of the full-system CCSD even with a fragment as large as four CH_2 groups. BE, on the other hand, can do overlapping fragments but requires a clear definition of edge and center sites, which is also ambiguous in systems lacking certain symmetries. This was demonstrated by Ricke *et al.*[36] who performed a number of BE calculations on the 2D Hubbard model with fragments of different shapes and choices of center sites. According to their report, the combination that gives the best energetics is not always intuitive.

We present here a scheme, incremental embedding, that allows the combination of arbitrary overlapping fragments *without* the knowledge of edge sites. As a proof of concept, we test this new fragment embedding scheme on several molecules in atom-centered Gaussian basis sets. Numerical results suggest that this method converges quickly with fragment size at equilibrium geometry and recovers most static correlation in minimal basis sets, but the performance deteriorates in either bond dissociation limit or larger basis sets. We show that this arises due to the nature of the HF bath and point out some possible solutions.

This article is organized as follows. In Sec. II, we give the theoretical background through briefly reviewing the Schmidt decomposition and existing Schmidt-space fragment embedding methods. In Sec. III, we formulate our theory of incremental embedding based on a new concept, Schmidt reduction, introduced therein. In Sec. IV, we give the computational details. Then in Sec. V, we present numerical results on several molecular systems as a proof of concept. In Sec. VI, we discuss a potential problem of using HF as the bath. Finally in Sec. VII, we conclude this work by pointing out several future directions.

II. BACKGROUND

In the following, terminologies from the lattice model are used for the formal discussion. All results can be adapted to realistic systems by replacing the "site basis" with the appropriate one-particle basis in the corresponding scenario (see Sec. IV for details). Throughout the discussion, we explicitly distinguish a general wave function (Ψ) from a mean-field wave function (Φ). We also distinguish the second-quantized form of an operator (e.g., $\hat{H}$) from its matrix representation in some single-determinantal basis (i.e., the first-quantized form, $\boldsymbol{H}$). The latter is useful when dealing with the Schmidt decomposition of correlated wave functions (see Sec. III A for details).

A. Schmidt decomposition

Suppose the system consists of two parts, the fragment (which we assume to be the minority) and the environment such that the Hilbert space for the whole system is a direct product of the two subsystems, i.e., $\mathcal{H} = \mathcal{H}_f \otimes \mathcal{H}_e$. Any state $|\Psi\rangle \in \mathcal{H}$ therefore has the following tensor product decomposition:

$$|\Psi\rangle = \sum_i^{\dim \mathcal{H}_f} \sum_j^{\dim \mathcal{H}_e} \Psi_{ij} |f_i\rangle \otimes |e_j\rangle, \tag{1}$$

where $|f_i\rangle \in \mathcal{H}_f$ and $|e_j\rangle \in \mathcal{H}_e$ are the (many-body) basis states that span the fragment and the environment, respectively. Equation (1) can be brought to the Schmidt decomposed form

$$|\Psi\rangle = \sum_p^{\dim \mathcal{H}_f} \lambda_p |f_p\rangle \otimes |b_p\rangle \tag{2}$$

by a singular value decomposition on the coefficient matrix $\boldsymbol{\Psi}$, where $|b_p\rangle$'s are the so-called entangled bath states and λ_p's are the singular values. The benefit of the Schmidt decomposition is that the length of the expansion is limited by the number of linearly independent fragment states (those with nonzero λ_p). Thus the decomposed form has a manageable length no matter how large the original system is, as long as the fragment is not too large. Equation (2) is exact in the sense that if $|\Psi\rangle$ is a ground state of $\hat{H}$ at some level of theory, the embedding Hamiltonian

$$\hat{H}_{\text{emb}} = \hat{P}\hat{H}\hat{P} \tag{3}$$

obtained by projecting $\hat{H}$ onto the Schmidt space with the operator

$$\hat{P} = \sum_{pq}^{\dim \mathcal{H}_f} |f_p\rangle\langle f_p| \otimes |b_q\rangle\langle b_q| \tag{4}$$

shares the same ground state as $\hat{H}$ at the same level of theory.

For a general wave function $|\Psi\rangle$, the bath states in Eq. (2) are highly complex many-body states and can be obtained only for very small systems. Meanwhile, $\hat{H}_{\text{emb}}$ contains many-body interaction terms even if the unprojected Hamiltonian has only one- and two-body operators. However, if one starts with a single-determinant wave function $|\Phi\rangle$, the otherwise complicated many-body decomposition can be performed in terms of a linear algebra transformation on the site basis.[29] Specifically, for a fragment composed of N_{frag} sites, $\hat{H}_{\text{emb}}$ can be written as

$$\hat{H}_{\text{emb}} = \sum_{pq}^{2N_{\text{frag}}} \tilde{h}_{pq} c_p^\dagger c_q + \sum_{pqrs}^{2N_{\text{frag}}} \tilde{V}_{prqs} c_p^\dagger c_q^\dagger c_s c_r, \tag{5}$$

where the summation goes over N_{frag} fragment sites and the same number of bath sites, and $\tilde{V}_{pqrs}$ is in chemists' notation.[23] $\hat{H}_{\text{emb}}$ in Eq. (5) contains only two-body interactions and therefore can be solved by the aforementioned accurate quantum chemical methods such as FCI, DMRG, and CCSD. For this reason, nearly all Schmidt-space fragment embeddings—including this work—use a HF bath, with only a few exceptions.[37] In order to optimize the embedding, one often needs to impose some flavor of matching conditions. So far there have been two main classes: DMET/DET matches the fragment and the bath using rigid, non-overlapping fragments and BE matches the fragment to fragment when they overlap. We will review both classes in Subsections II B and II C.

Before that, let us put special emphasis on two potential problems underlying the HF bath. First, if some site is unentangled with all other sites in the bath wave function, it gives a vanishing singular value in Eq. (2) and is therefore redundant for the embedding. This basis set degradation problem is in fact not rare for realistic systems with a mean-field bath, to which we will get back in Sec. VI. Second, for a K-site, $2N$-electron system there is an upper bound, $N_{\text{bath}}^{\text{max}}$, on the number of linearly independent bath sites, i.e., $N_{\text{bath}} = \min\{N_{\text{frag}}, N_{\text{bath}}^{\text{max}}\}$, where the bound is determined by the number of electrons and holes in the system,

$$N_{\text{bath}}^{\text{max}} = \min\{N, K - N\}. \tag{6}$$

For example, for a 10-site, 4-electron system, there are at most 2 linearly independent bath sites even for fragments of 3 or more sites. Consequently, one needs to use a fragment of 8 sites in order to obtain the exact solution. This is worse compared to the half-filled situation (i.e., $K = 10$, $N = 5$), where the embedding is exact when using a 5-site fragment. In general, the exact embedding with a HF bath requires a fragment of size $K - N_{\text{bath}}^{\text{max}}$. This number is minimized when the system is half-filled ($N = K$) and increases for either lower ($N < K$) or higher ($N > K$) filling. These observations suggest that DMET (and related embedding methods) converges fastest with respect to fragment size when the system is half-filled. In practice, however, a large-sized basis set is often needed in order to include as much dynamic correlation as possible. This will hence make the system far from half-filling and therefore deteriorate the performance of embedding with a HF bath when using a relatively small fragment. We will demonstrate this using numerical results in Sec. V C.

B. DMET and DET

In 2012, Knizia and Chan[28] proposed the idea to embed a high-level theory in a HF bath. To optimize the embedding, a one-particle effective potential $\hat{v}_{\text{eff}}$ is added to the HF bath,

$$|\Phi(\hat{v}_{\text{eff}})\rangle = \arg\min_{\Phi}\langle\Phi|\hat{H} + \hat{v}_{\text{eff}}|\Phi\rangle, \tag{7}$$

where $|\Phi\rangle$ is restricted to be a single Slater determinant. The bath 1PDM is then made to match the fragment 1PDM by tuning this effective potential,

$$\langle\Phi(\hat{v}_{\text{eff}})|c_p^\dagger c_q|\Phi(\hat{v}_{\text{eff}})\rangle = \langle\Psi(\hat{v}_{\text{eff}})|c_p^\dagger c_q|\Psi(\hat{v}_{\text{eff}})\rangle, \tag{8}$$

where $|\Psi\rangle$ denotes the (correlated) embedding wave function, and indices p and q go over fragment sites only. Note that in generating $\hat{H}_{\text{emb}}$ [Eq. (3)], $\hat{v}_{\text{eff}}$ is involved only in the bath part. Therefore, $|\Psi\rangle$ gains its dependency on $\hat{v}_{\text{eff}}$ only through the bath. As demonstrated by Tsuchimochi *et al.*, the matching condition in Eq. (8) is not always exactly satisfiable[37] and therefore is optimized in a least-squares sense.[32] Nevertheless, once the matching is achieved, the total energy can be expressed as a sum of fragment energies,

$$E_{\text{DMET}} = \sum_A (E_A)_{\text{DMET}} = \sum_A \langle\Psi_A|\hat{E}_A|\Psi_A\rangle, \tag{9}$$

where the summation goes over all (non-overlapping) fragments $\{A\}$; $\{\hat{E}_A\}$ partitions $\hat{H}$,

$$\hat{H} = \sum_A \hat{E}_A, \tag{10}$$

such that each of them only involves terms that belong to the fragment as well as half the interactions between the fragment and the bath to avoid double counting. DMET is extremely powerful for strongly correlated model systems such as the Hubbard model[28,33] and atomic rings[29,32] where unique partitions of the system are obvious due to the high symmetry. Two years later, Bulik *et al.*[30] simplified the matching condition in Eq. (8) by requiring only the diagonal of the 1PDM to be matched, i.e.,

$$\langle\Phi(\hat{v}_{\text{eff}})|c_p^\dagger c_p|\Phi(\hat{v}_{\text{eff}})\rangle = \langle\Psi(\hat{v}_{\text{eff}})|c_p^\dagger c_p|\Psi(\hat{v}_{\text{eff}})\rangle, \tag{11}$$

where p goes over fragment sites only. This variant is named DET and has shown performance that is comparable with the original DMET.[30,31]

When applying DMET/DET to realistic systems, the absence of high symmetry makes an unambiguous non-overlapping partition of the system difficult or even impossible. Different partitions often lead to different results, and there is no apparent way to evaluate the quality of those different choices. In the most general scenario, perhaps the best one can do is to adopt the following one-site embedding scheme:

$$E_{\text{1-site}} = \sum_p^K (E_p)_{\text{DMET}}, \tag{12}$$

where index p goes over all sites and $(E_p)_{\text{DMET}}$ is simply the "site energy." Equation (12) is free of the ambiguity problem by construction, but the generalization to fragments of larger size is not obvious in the context of DMET. We will see in Sec. III how this scheme could be improved systematically by incremental embedding.

C. Bootstrap embedding

In addition to the ambiguity of fragment partition, the restriction to non-overlapping fragments also results in slow convergence with fragment size due to the persistent edge effects. Recently, Welborn *et al.*[35] proposed the BE scheme in order to eliminate the edge effects in certain situations. The motivation is that when several fragments overlap, the edge sites in one fragment might be the center of another. Therefore by matching properties such as 1PDM and/or 2PDM of the former to those of the latter, one can expect improving

the description of the edge sites without deteriorating the center sites. Mathematically this is formulated as a constrained optimization. Suppose fragment A overlaps partially with fragment B, and A's center sites $\mathcal{C}(A)$ are edge sites of B. Then the matching condition between A and B is satisfied by making the following Lagrangian stationary:

$$
\begin{aligned}
\mathcal{L}[\Psi_B; \boldsymbol{\lambda}, \boldsymbol{\Lambda}] = &\langle \Psi_B | \hat{H}_{\text{emb},B} | \Psi_B \rangle \\
&+ \sum_{pq \in \mathcal{C}(A)} \lambda_{pq}^B (\langle \Psi_B | c_p^\dagger c_q | \Psi_B \rangle - P_{pq}^A) \\
&+ \sum_{pqrs \in \mathcal{C}(A)} \Lambda_{prqs}^B (\langle \Psi_B | c_p^\dagger c_q^\dagger c_s c_r | \Psi_B \rangle - \Gamma_{prqs}^A) \\
&+ \mathcal{E}_B (\langle \Psi_B | \Psi_B \rangle - 1),
\end{aligned} \tag{13}
$$

which can be transformed to an eigenvalue problem of a dressed Hamiltonian,

$$
\hat{H}_{\text{emb},B}^{\text{eff}} = \hat{H}_{\text{emb},B} + \sum_{i,j \in \mathcal{C}(A)} \lambda_{pq}^B c_p^\dagger c_q + \sum_{pqrs \in \mathcal{C}(A)} \Lambda_{prqs}^B c_p^\dagger c_q^\dagger c_s c_r. \tag{14}
$$

In other words, the Lagrange multipliers play the role of a constraint potential $\hat{v}_c$ to satisfy the matching conditions. Similar to the direct optimization method in DFT,[38,39] Ricke *et al.* proved that $\mathcal{L}$ has a negative semi-definite Hessian, making its stationary point a maximum and rendering Eq. (14) numerically favorable.[36] Results on model systems have suggested that this method indeed leads to faster convergence with fragment size.[35]

Despite its success on model systems, the generalization of BE to a general molecular system is challenging. This is because the distinction between the edge and the center sites is often vague unless certain symmetries, such as translational invariance, are present. Nevertheless, the idea of matching among overlapping fragments is of significant importance. We will see that it provides—after being combined with the technique we are going to introduce in Sec. III A—a path toward realistic systems.

III. THEORY

A. Schmidt reduction

We first introduce a tool that enables us to encode the information from a larger embedding space to a smaller one. Suppose we start with a wave function $|\Psi\rangle$ (that does not have to be mean-field), perform the Schmidt decomposition with an m-site fragment, and obtain the embedding Hamiltonian $\hat{H}_m$ in the resulting $2m$-electron, $2m$-site space (assuming all sites are entangled). Then we solve for its ground state $|\Psi_m\rangle$, perform a second Schmidt decomposition involving $n < m$ sites, and obtain a new embedding Hamiltonian $\hat{H}_{m \to n}$ in the resulting $2n$-electron, $2n$-site space. Overall, the process can be summarized as

$$
|\Psi\rangle \xrightarrow[\hat{H}]{m\text{-site SD}} \hat{H}_m \xrightarrow{\text{FCI}} |\Psi_m\rangle \xrightarrow[\hat{H}_m]{n\text{-site SD}} \hat{H}_{m \to n}. \tag{15}
$$

We call this process a *Schmidt reduction* (SR) from m sites to n sites, or $m \to n$ SR for short. Due to the exact nature of the

Schmidt decomposition, $\hat{H}_m$ and $\hat{H}_{m \to n}$ share the same ground state even though the latter is represented with a vector space of much smaller dimension.

If one starts with a HF wave function $|\Phi\rangle$, $\hat{H}_m$ has a simple form involving only one- and two-body interactions as explained in Sec. II A. However, $|\Psi_m\rangle$ is correlated and so are the individual many-body bath states that make up the Schmidt bath. This renders $\hat{H}_{m \to n}$ complicated and awkward to deal with in the site basis: the second-quantized form of $\hat{H}_{m \to n}$ will contain many-body interaction terms. To that end, we represent $\hat{H}_{m \to n}$ by its matrix representation in the Schmidt bases (i.e., the first-quantized form); these bases are simply CI states whose orbitals are the fragment and bath sites of $|\Phi\rangle$. For instance, $\hat{H}_{m \to 1}$ can be elegantly represented by a 4×4 matrix $\boldsymbol{H}_{m \to 1}$ in the complete one-site Schmidt basis,

$$
| \, \rangle \otimes |\tilde{\uparrow\downarrow}\rangle, |\uparrow\rangle \otimes |\tilde{\downarrow}\rangle, |\downarrow\rangle \otimes |\tilde{\uparrow}\rangle, |\uparrow\downarrow\rangle \otimes |\tilde{\,}\rangle, \tag{16}
$$

where the tilde signs over bath states emphasize that these states are many-body states instead of one-electron states.

One of the important consequences of Eq. (15) is that it suggests a way to match overlapping fragments: reduce each of them onto the n sites where they all overlap and require that the properties of all fragments agree on these common sites. This provides a powerful set of matching conditions that are not based on the discrimination of edge and center sites. In the following, we introduce one realization of this idea, incremental embedding.

B. Incremental embedding from two sites to one site

In this work, we introduce *Incremental Embedding* (henceforth abbreviated as IE) as a novel quantum embedding method that exploits the strength of the Schmidt reduction. The goal is to construct a one-site effective Hamiltonian $\boldsymbol{H}_p$ for any given site p such that it contains correlation at the level of m-site embeddings ($m \geq 2$). In this section, we focus on the lowest order, $m = 2$. The resulting theory is named IE from two sites to one site, or $2 \to 1$ IE for short. The generalization to $m > 2$ will be presented in Sec. III D.

Suppose we have already determined the HF wave function $|\Phi\rangle$ for a system described by the following K-site Hamiltonian:

$$
\hat{H} = \sum_{pq}^K h_{pq} a_p^\dagger a_q + \sum_{pqrs}^K V_{prqs} a_p^\dagger a_q^\dagger a_s a_r, \tag{17}
$$

where V_{prqs} is in chemists' notation.[23] Then a Schmidt decomposition of $|\Phi\rangle$ on site p gives the mean-field approximation to $\boldsymbol{H}_p$,

$$
|\Phi\rangle \xrightarrow[\hat{H}]{\text{SD on } p} \boldsymbol{H}_p^0, \tag{18}
$$

which is merely the matrix representation of $\hat{H}$ in the one-site Schmidt basis of site p derived from the mean-field bath. A better approximation can be obtained by the following $2 \to 1$ SR:

$$
|\Phi\rangle \xrightarrow[\hat{H}]{\text{SD on } (p,q)} \hat{H}_{pq} \xrightarrow{\text{FCI}} |\Psi_{pq}\rangle \xrightarrow[\hat{H}_{pq}]{\text{SD on } p} \boldsymbol{H}_p^{\,q}, \tag{19}
$$

where $q \neq p$ could be any other site. According to the exact nature of SR, $\boldsymbol{H}_p^{\,q}$ contains the correlation between p and q at

the level of two-site embedding and thus is an improvement over H_p^0. However, approximating H_p by H_p^q is problematic: different choices of q in general give different H_p^q's and it is hard to determine which choice is better than others.

Rather than picking one particular H_p^q to replace H_p, we instead choose to accumulate contributions from all of them. A simple sum of H_p^q over several sites $q \neq p$ leads to a severe double-counting problem since each matrix alone is a representation of the full Hamiltonian $\hat{H}$. To this end, we propose to first divide the Hamiltonian into pieces, obtain the SR matrix representation for each piece using the two-site fragment that is most relevant to it, and then assemble them to construct an approximate H_p that is free of double-counting.

Specifically, consider the following partition of Hamiltonian:

$$\hat{H} = \sum_{q \neq p}^{K} \hat{I}_p^q, \tag{20}$$

where we require each $\hat{I}_p^q$ includes two types of terms from $\hat{H}$: (i) terms that belong to the fragment (p, q), including h_{pp}, h_{qq}, h_{pq}, h_{qp} for one-electron terms, and

$$V_{pppp}, V_{qqqq}, V_{pppq}, V_{ppqq}, V_{pqpq}, V_{pqqq}$$

for two-electron terms (including other terms that are related by permutational symmetry), and (ii) terms that are interactions between q and other sites, including h_{qr}, h_{rq} for one-electron terms, and

$$V_{ppqr}, V_{pqpr}, \ldots, V_{pqrs}, V_{prqs}, \ldots, V_{qrst}, V_{qsrt}, V_{qtrs}$$

for two-electron terms (including other terms that are related by permutational symmetry), where $r \neq s \neq t$ are sites outside fragment (p, q). To avoid double-counting, a factor of $\frac{1}{l}$ is attached to each term that is shared by l fragments. For example, h_{qr} and h_{rq} are halved since they appear in both $\hat{I}_p^q$ and $\hat{I}_p^r$. In this way, Eq. (20) distributes $\hat{H}$ *evenly* to all two-site fragments anchored by p. With this partition in hand, we define the *incremental Hamiltonian* from two sites to one site as

$$\Delta I_p^{2 \to 1} = \sum_{q \neq p}^{K} I_p^q - (I_p^q)^0, \tag{21}$$

where I_p^q is the matrix representation of $\hat{I}_p^q$ in the one-site Schmidt basis of site p obtained by the following $(p, q) \to p$ SR:

$$|\Phi\rangle \xrightarrow[\hat{H}]{\text{SD on } (p,q)} \hat{H}_{pq} \xrightarrow{\text{FCI}} |\Psi_{pq}\rangle \xrightarrow[\hat{I}_p^q]{\text{SD on } p} I_p^q \tag{22}$$

and $(I_p^q)^0$ is its mean-field counterpart

$$|\Phi\rangle \xrightarrow[\hat{I}_p^q]{\text{SD on } p} (I_p^q)^0. \tag{23}$$

The physical meaning of $\Delta I_p^{2 \to 1}$ is clear; it accumulates the correlations between site p and all other sites that are missing at the mean-field level. Adding this correction to H_p^0, we have a better approximation to H_p,

$$H_p^{2 \to 1} = H_p^0 + \Delta I_p^{2 \to 1}. \tag{24}$$

One can expect $H_p^{2 \to 1}$ to be of the quality of two-site embeddings since each piece of the Hamiltonian is improved by the embedding calculation involving the most relevant two-site fragment. This is also schematically illustrated in Fig. 1.

Once we obtain the effective Hamiltonians $\{H_p^{2 \to 1}\}$ for all sites, we can readily determine their ground states $\{u_p^{2 \to 1}\}$ (as the eigenvectors of the lowest eigenvalue) and compute the site densities $\{P_{pp}\}$ (*vide infra*). In general, these site densities do not add up to the correct number of electrons because each $H_p^{2 \to 1}$ is generated from multiple fragments of different chemical potentials. This violation in the conservation of particle number can be fixed by introducing a global chemical potential μ, which is determined by solving,

$$\sum_p^K P_{pp}(\mu) = N, \tag{25}$$

where the μ-dependent site densities are obtained from solving $\{H_p^{2 \to 1} + \mu D\}$ (as opposed to the bare Hamiltonians); $D = \text{diag}(0, 1, 1, 2)$ in the Schmidt basis shown in Eq. (16). We note that it is also possible to apply a set of site-specific chemical potentials $\{\mu_p\}$ to tune the population for each site. This is useful when IE is performed in a non-self-consistent manner (see Sec. III F).

C. Expectation values in 2 → 1 IE

In Sec. III B, we discussed how a one-site effective Hamiltonian can be constructed from successively improving the mean-field description by incorporating the correlation with every other site. Once done, the ground state for each site is approximated by a four-dimensional vector, $\{u_p^{2 \to 1}\}$, in the one-site Schmidt basis [Eq. (16)]. The ground state expectation value of any given operator can then be evaluated by summing contributions from each site. Let us take the total energy as an example, whose corresponding operator is the Hamiltonian $\hat{H}$. First we obtain a partition of $\hat{H}$ over all sites,

$$\hat{H} = \sum_p^K \hat{E}_p, \tag{26}$$

which is a special case of Eq. (10) with each fragment involving only one site. For each site p, we further partition $\hat{E}_p$ by

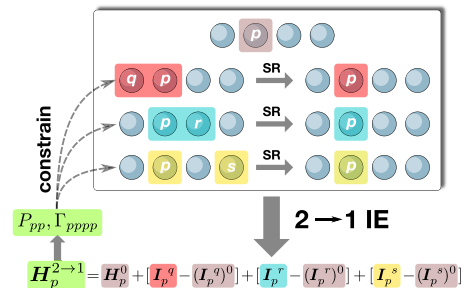
FIG. 1. Schematic illustration of 2 → 1 IE in a four-site lattice model. The mean-field approximation H_p^0 (brown) is improved by incremental Hamiltonians from three two-site embedding calculations: (p, q) (red), (p, r) (blue), and (p, s) (yellow). The site density P_{pp} and pair-density Γ_{pppp} derived from $H_p^{2 \to 1}$ (green) can in return be used to constrain the embedding calculations.

distributing it evenly to all relevant two-site fragments in a way similar to Eq. (20),

$$\hat{E}_p = \sum_{q \neq p}^{K} \hat{E}_p{}^q. \tag{27}$$

With this partition, a matrix representation of $\hat{E}_p$ in the Schmidt basis of $2 \rightarrow 1$ IE can be constructed by the same procedure described in Eqs. (21)–(24),

$$\boldsymbol{E}_p^{2\rightarrow1} = \boldsymbol{E}_p^0 + \Delta\boldsymbol{E}_p^{2\rightarrow1}, \tag{28}$$

which gives the site energy for p at the level of $2 \rightarrow 1$ IE,

$$E_p^{2\rightarrow1} = (\boldsymbol{u}_p^{2\rightarrow1})^\dagger \boldsymbol{E}_p^{2\rightarrow1} \boldsymbol{u}_p^{2\rightarrow1}. \tag{29}$$

Finally, the total energy for $2 \rightarrow 1$ IE is simply a sum of all site energies,

$$E^{2\rightarrow1} = \sum_p^{K} E_p^{2\rightarrow1}. \tag{30}$$

Equation (30) can be viewed as an unambiguous generalization of the one-site DMET energy in Eq. (12) to two-site fragments. The generalization to an arbitrary number of fragment sites is presented in Sec. III D.

Note that the process above is greatly simplified, if the operator of which expectation values are taken involves only one site, say p. In that scenario, one can bypass the partition and summation steps [Eqs. (26)–(28)] and obtain the matrix representation of that operator using any fragments involving p. Examples of this type include the diagonal elements of 1PDM (site densities, $\{P_{pp}\}$) and 2PDM (pair-densities, $\{\Gamma_{pppp}\}$).

D. Generalization to fragments of arbitrary size

In this section, we generalize IE to fragments of arbitrary size, in a way that is similar to the method of increments commonly used in local correlation methods.[40–43] The end results can be summarized in the following recursive formula for $m \geq 3$:

$$\boldsymbol{H}_p^{m\rightarrow1} = \boldsymbol{H}_p^{(m-1)\rightarrow1} + c_{m\rightarrow1}\Delta\boldsymbol{I}_p^{m\rightarrow1}, \tag{31}$$

where $c_{m\rightarrow1}$ is an appropriate constant that ensures the series of equations terminating appropriately (*vide infra*); $\Delta\boldsymbol{I}_p^{m\rightarrow1}$ is the incremental Hamiltonian from m sites to one site,

$$\Delta\boldsymbol{I}_p^{3\rightarrow1} = \sum_{r>q\neq p}^{K} \boldsymbol{I}_p^{qr} - \boldsymbol{I}_p^{q} - \boldsymbol{I}_p^{r}$$

$$\Delta\boldsymbol{I}_p^{4\rightarrow1} = \sum_{s>r>q\neq p}^{K} \boldsymbol{I}_p^{qrs} - \boldsymbol{I}_p^{qr} - \boldsymbol{I}_p^{rs} - \boldsymbol{I}_p^{qs} + \boldsymbol{I}_p^{p} + \boldsymbol{I}_p^{r} + \boldsymbol{I}_p^{s} \cdots, \tag{32}$$

where terms like $\boldsymbol{I}_p^{qr}$ and $\boldsymbol{I}_p^{qrs}$ are matrix representations of the sum of relevant pieces of $\hat{H}$ as defined in Eq. (20). For example, $\boldsymbol{I}_p^{qr}$ can be obtained by the following $3 \rightarrow 1$ SR:

$$|\Phi\rangle \xrightarrow[\hat{H}]{\text{SD on } (p,q,r)} \hat{H}_{pqr} \xrightarrow{\text{FCI}} |\Psi_{pqr}\rangle \xrightarrow[\hat{I}_p{}^q + \hat{I}_p{}^r]{\text{SD on } p} \boldsymbol{I}_p^{qr}. \tag{33}$$

The physical meaning of $\Delta\boldsymbol{I}_p^{m\rightarrow1}$ is also straightforward: it is the correction from m-site embedding calculations that are not included in any $(m-1)$-site embedding calculations.

The constant coefficients $\{c_{m\rightarrow1}\}$ arise due to the difference between traditional incremental methods (such as the aforementioned local correlation methods) and Schmidt decomposition-based embedding methods. In local correlation methods, a hierarchy similar to Eq. (31), but with $c_{m\rightarrow1} \equiv 1$ for all m, can be derived, which terminates when all sites are involved (i.e., $m = K$). In that situation, the local correlation method is exact in the sense that it is equivalent to applying the same correlation method to all sites. In Schmidt-space embedding methods, on the other hand, the embedding is exact when using any fragment composed of $K - N_{\text{bath}}^{\max}$ sites, as explained in Sec. II A. For this reason, we require the one-site Hamiltonian $\boldsymbol{H}_p^{m\rightarrow1}$ constructed according to Eq. (31) be exact when $m = K - N_{\text{bath}}^{\max}$, from which the following expressions for $\{c_{m\rightarrow1}\}$ can be derived:

$$c_{m\rightarrow1} = \binom{K-m}{K-N_{\text{bath}}^{\max}-m} \Big/ \binom{K-2}{K-N_{\text{bath}}^{\max}-2}. \tag{34}$$

Note that when $N_{\text{bath}}^{\max} = 0$ (i.e., no bath states), Eq. (34) gives $c_{m\rightarrow1} \equiv 1$ and hence formally reduces to traditional local correlation methods. Moreover, Eq. (34) gives $c_{2\rightarrow1} = 1$ for $m = 2$, which is also consistent with the $2 \rightarrow 1$ IE result as discussed above [Eq. (24)].

In order to generalize the energy evaluation scheme, we need to generalize Eq. (28) to multiple sites. To that end, a recursive formula similar to Eq. (31) can be derived for $\boldsymbol{E}_p$,

$$\boldsymbol{E}_p^{m\rightarrow1} = \boldsymbol{E}_p^{(m-1)\rightarrow1} + c_{m\rightarrow1}\Delta\boldsymbol{E}_p^{m\rightarrow1}, \tag{35}$$

where $\{c_{m\rightarrow1}\}$ is the same set of coefficients given by Eq. (34). With this in hand, the site energy

$$E_p^{m\rightarrow1} = (\boldsymbol{u}_p^{m\rightarrow1})^\dagger \boldsymbol{E}_p^{m\rightarrow1} \boldsymbol{u}_p^{m\rightarrow1} \tag{36}$$

and the total energy

$$E^{m\rightarrow1} = \sum_p^{K} E_p^{m\rightarrow1} \tag{37}$$

for $m \rightarrow 1$ IE can be straightforwardly evaluated, where $\boldsymbol{u}_p^{m\rightarrow1}$ is the lowest eigenvector of $\boldsymbol{H}_p^{m\rightarrow1}$ (with a proper chemical potential). Equation (37) can be viewed as an unambiguous generalization of the one-site DMET energy in Eq. (12) to fragments composed of an arbitrary number of sites.

E. Matching conditions

So far we have not touched one of the most powerful ingredients in embedding calculations—the matching condition. From the discussion above, constructing $\boldsymbol{H}_p^{m\rightarrow1}$ requires embedding calculations for all m-site fragments involving site p. Without any constraints, these overlapping fragments in general will have different site densities and pair-densities on their common site, with the only exception where the exact bath is used as opposed to the mean-field approximation. This observation indicates that one can optimize these embedding calculations by forcing the match to happen.

Suppose we have obtained $\{H_p^{m \to 1}\}$ for all sites. Solving them under an appropriate chemical potential, we can compute $\{P_{pp}\}$ and $\{\Gamma_{pppp}\}$ as described in Sec. III C. These values are our current best estimation of the diagonal elements of the exact 1PDM and 2PDM. Naturally, we can require the site densities and pair-densities in all fragment calculations match them. Mathematically, the problem of constraining certain density matrix elements to given values can be formulated as a constrained optimization and has already been addressed in BE [Eqs. (13) and (14)].[35,36] Here, we adapt the method to IE. Suppose we want to constrain both site densities and pair-densities for, e.g., a two-site fragment (p, q). We can achieve this by introducing the following Lagrangian:

$$\mathcal{L}_{pq}[\Psi_{pq}; \boldsymbol{\lambda}, \boldsymbol{\Lambda}] = \langle \hat{H}_{pq} \rangle_{pq} + \sum_{r=p,q} [\lambda_r (\langle a_r^\dagger a_r \rangle_{pq} - P_{rr})$$
$$+ \Lambda_r (\langle a_r^\dagger a_r^\dagger a_r a_r \rangle_{pq} - \Gamma_{rrrr})]$$
$$+ \mathcal{E}_{pq}(\langle \Psi_{pq} | \Psi_{pq} \rangle - 1), \tag{38}$$

where $\hat{H}_{pq}$ is the embedding Hamiltonian; $\{P_{rr}\}$ and $\{\Gamma_{rrrr}\}$ are the target values; $\langle \cdots \rangle_{pq}$ is short for $\langle \Psi_{pq} | \cdots | \Psi_{pq} \rangle$. Making $\mathcal{L}_{pq}$ stationary leads to the following eigenvalue equation:

$$(\hat{H}_{pq} + \hat{v}_{pq}^{\text{c}}) |\Psi_{pq}\rangle = \mathcal{E}_{pq} |\Psi_{pq}\rangle, \tag{39}$$

where $\hat{H}_{pq}$ is dressed by a constraint potential,

$$\hat{v}_{pq}^{\text{c}} = \sum_{r=p,q} (\lambda_r a_r^\dagger a_r + \Lambda_r a_r^\dagger a_r^\dagger a_r a_r). \tag{40}$$

Equations (39) and (40) enable us to apply the desired constraints to fragment calculations readily in a ground state formalism. Note that the constraint potential only exists in obtaining $|\Psi_{pq}\rangle$ and should not be included in other steps of SR.

In practice, Eqs. (38)–(40) can be solved by using either gradient- or Hessian-based numerical optimization methods. An algorithm based on the Newton-Raphson method[44] is given in the supplementary material.

F. Density optimization

Once the matching conditions are imposed to each fragment in IE, one can construct a new set of $\{H_p^{m \to 1}\}$ and recompute the site densities and pair-densities. In general, these values are of better quality compared to the old estimation, due to the embedding being optimized by the matching conditions. In return, these new densities can be used to constrain further embedding calculations which will generate $\{H_p^{m \to 1}\}$ of even better quality. This process can be repeated until self-consistency is reached, making the theory a closed loop (Fig. 1).

The discussion above immediately suggests an algorithm to optimize the densities in IE self-consistently. Here, we state it for $2 \to 1$ IE for the sake of simplicity, and the generalization to larger fragments should be straightforward.

1. Solve the HF wave function $|\Phi\rangle$ for the whole system; obtain the mean-field Hamiltonians $\{H_p^0\}$ for all sites.
2. Obtain some guess densities $\{P_{pp}^{(0)}\}$ and $\{\Gamma_{pppp}^{(0)}\}$ (e.g., HF).
3. Perform embedding calculations for all two-site

fragments; in each calculation, constrain the site densities and pair-densities of the fragment sites to match $\{P_{pp}^{(0)}\}$ and $\{\Gamma_{pppp}^{(0)}\}$, respectively, according to Eqs. (38)–(40).
4. For each site p, Schmidt reduces all p-involved fragments to p and computes $H_p^{2 \to 1}$ according to Eq. (24).
5. Diagonalize $\{H_p^{2 \to 1}\}$ under an appropriate global chemical potential μ determined by solving Eq. (25); obtain the ground states $\{u_p^{2 \to 1}\}$; and recompute $\{P_{pp}\}$ and $\{\Gamma_{pppp}\}$.
6. If the new densities do not match $\{P_{pp}^{(0)}\}$ and $\{\Gamma_{pppp}^{(0)}\}$, go back to step 2 with the new guess densities; otherwise, the density optimization is converged, and the ground state expectation values of desired operators can be computed using $\{u_p^{2 \to 1}\}$.

We note the reader that this iterative procedure of optimizing densities is different from the bath optimization in DMET[28,29,32] as described in Sec. II B [Eqs. (7) and (8)]. Like BE, the mean-field bath wave function $|\Phi\rangle$ is fixed and not subject to any change in IE; the density matching and optimization happen only in the high-level calculations between overlapping fragments. The neglect of bath response would be a problem in certain cases (e.g., electronic phase transition in infinite systems) but is a good approximation to the ground states of molecules.

In addition to the self-consistent version, we note that IE can also be formulated as a non-self-consistent theory. In terms of the algorithm, the main difference lies in step 5: instead of solving for a global chemical potential, one determines a set of site-specific chemical potentials $\{\mu_p\}$ such that for each site the population matches the guess densities, i.e.,

$$P_{pp}(\mu_p) = P_{pp}^0. \tag{41}$$

In other words, the densities are not optimized and IE is used in a one-shot style, similar to the $G_0 W_0$ method.[45–47] This approximation would be useful when (i) the system is large and hence full self-consistency is expensive and (ii) the quality of the guess densities is reasonable. We will examine the performance of this approximation in Sec. V.

G. Computational scaling

We end this section by briefly discussing the computational scaling. For $m \to 1$ IE, the total work is dominated by the embedding calculations of $\binom{K}{m} \sim O(K^m)$ m-site fragments. Symmetries can effectively reduce this number by a constant factor, as fragments related by symmetry operations will give the same one-site Hamiltonian and hence need to be evaluated only once. For each fragment, there are two potential rate-limiting steps: (i) the basis transformation of integrals (including the partitioned Hamiltonian $\{\hat{I}_p{}^q\}$) from the site basis to the Schmidt basis and (ii) the high-level calculation (in this paper, FCI). For small fragments $(m \ll K)$, the former dominates with a $O(K^4)$ scaling, which will exceed the aforementioned $O(K^m)$ scaling if $m < 4$. Fortunately, the basis transformation needs to be performed only once, due to the mean-field bath not being optimized in IE. This feature makes the basis transformation step usually negligible, especially in the self-consistent version. Under these conditions, $m \to 1$

IE has a scaling of $O(K^m)$, which is lower than most accurate quantum chemical methods [$> O(K^5)$] if the incremental expansion [Eq. (31)] can be truncated at a small m.

IV. COMPUTATIONAL DETAILS

In the following computational work, we will examine the performance of IE using several small molecules. The symmetrically orthogonalized atomic orbitals[23] (SOAOs) are used as site basis for the hydrogen ring, while localized molecular orbitals (LMOs) of the Foster-Boys style[48] are used in all other cases. We note that the entire MO space (i.e., both occupied and virtual MOs) is subject to the orbital localization to generate the LMOs. The necessary atomic integrals are generated by Psi4.[49] The Foster-Boys localization is performed in Q-Chem.[50] Molecular geometries are also optimized in Q-Chem at the B3LYP[51]/cc-pVTZ[52] level and can be found in the supplementary material. All embedding calculations, including IE, BE, and one-site DMET, are performed using the electronic structure program, frankenstein,[53] developed by one of the authors. Spin-restricted HF (RHF) and FCI are used as bath and high-level solvers, respectively. In self-consistent IE, both site densities and pair-densities are self-consistently determined based on an RHF guess. In the non-self-consistent version, only the site densities are constrained to the RHF values due to the bad quality of the mean-field pair-densities. For one-site DMET, we abandon the self-consistency and also constrain the site densities to the RHF values. For all systems tested in this work, the exact solutions are accessible and obtained by the Block DMRG code.[54–58]

V. RESULTS

A. Symmetric stretching of the hydrogen ring model

We select the minimal-basis hydrogen ring model as our first example for several reasons. First it can be viewed as the simplest generalization of the Hubbard model toward real molecules, covering both a weakly correlated domain (near equilibrium geometry) and a strongly correlated domain (dissociation limit). Second, it is an "easy" case for Schmidt-space embedding with a HF bath according to our discussion in Sec. II A because the system is at half-filling. Last but not least, the cyclic symmetry makes all sites (which are symmetrically orthogonalized 1s orbitals in this case) equivalent. This not only renders BE applicable for comparison but also tremendously reduces the computational work for IE so that the trend of convergence with fragment size can be examined thoroughly.

In Fig. 2, the energy errors per atom in the symmetric stretching of STO-3G H_{10} are plotted for IE and BE, respectively. The non-self-consistent version of IE is used since the site densities are completely determined by the cyclic symmetry. For BE, fragments involving two and three adjacent sites are used but only in the latter is there the distinction of center and edge sites. In that case, the pair-densities of edge sites are made to match that of the center site. Results of 1-site DMET are not included since BE(2) is equivalent to non-self-consistent 2-site DMET for this special case. Due to
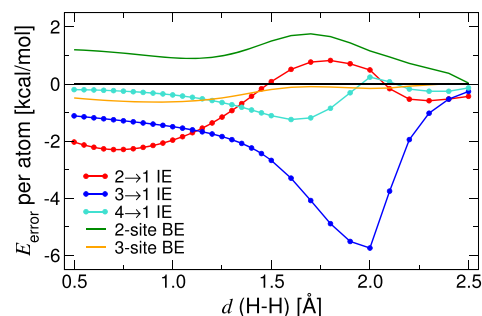

FIG. 2. Total energy error per atom (in kcal/mol) of the symmetric stretching of STO-3G H_{10}.

the half-filled configuration, both methods become exact when using fragments composed of five sites. Overall, both methods are very accurate even at the 2-site level. The error is consistently small ($<$2 kcal/mol/ per atom) for all geometries tested here. Near the equilibrium position ($\sim$ 0.95 Å), systematic improvements are observed for both methods when fragments of larger size are used. In the dissociation limit, correct asymptotic behavior is recovered in all cases even though the HF bath is spin-restricted. At intermediate geometries (1.5 $\sim$ 2.0 Å), however, the convergence with fragment size is not monotonic for IE: the 3 $\rightarrow$ 1 level suffers from severe over-correlation and is worse than 2 $\rightarrow$ 1; this over-correlation is only ameliorated by corrections from the 4 $\rightarrow$ 1 level. Quite the contrary, BE continues to reduce the error by using a larger fragment and shows better accuracy compared to IE with the same fragment size. Nevertheless, the performance loss of IE compared to BE is somewhat expected in this specific example since the model is Hubbard-like and therefore optimal for the latter.

B. Single bond breaking

Now we consider two real molecules, CH_4 and C_2H_6 in their minimal basis (STO-3G), to which BE is no longer applicable. Both molecules show merely a small deviation from the ideal half-filled configuration. Therefore, we can still expect good performance from the embedding calculations. Specifically we are interested in the energetics of the following two single-bond breaking processes:

$$\begin{aligned} CH_4 &\longrightarrow CH_3 \cdot + H \cdot, \\ C_2H_6 &\longrightarrow CH_3 \cdot + CH_3 \cdot. \end{aligned} \tag{42}$$
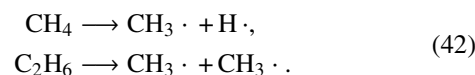
First we consider the PESs obtained by non-self-consistent IE as shown in Fig. 3, along with non-self-consistent, one-site DMET results for comparison. In both methods, the site densities are constrained to the RHF values, whose quality is high near equilibrium geometry but deteriorates quickly as the bond is stretched (see Fig. S1 in the supplementary material). If the error is mainly density-driven, we should expect good accuracy at equilibrium geometries as well as increasing error along the dissociation processes. This is indeed the case for one-site DMET (green), as can be seen from the growing gap between the embedding solution and the exact one in Fig. 3. 2 $\rightarrow$ 1 IE (red) shows a similar trend when approaching the dissociation limit but recovers only a limited amount of correlation energies at an equilibrium position. On the contrary,
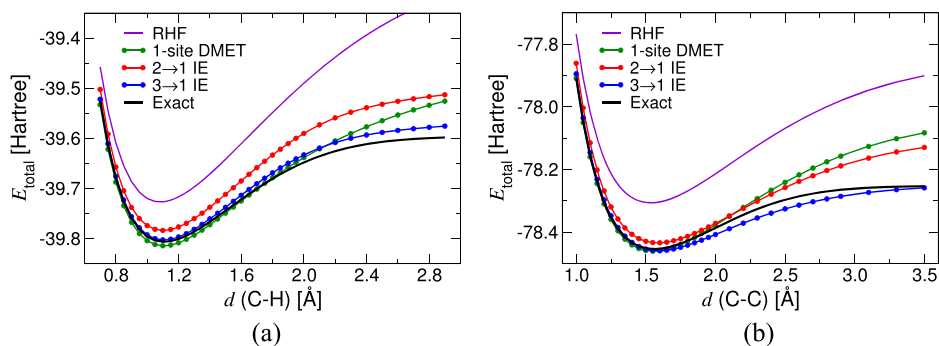
FIG. 3. Potential energy surfaces of (a) stretching one C–H bond of a methane molecule and (b) symmetrically dissociating an ethane molecule into two methyl radicals, predicted by the non-self-consistent IE and one-site DMET. The STO-3G basis set is used in both cases.
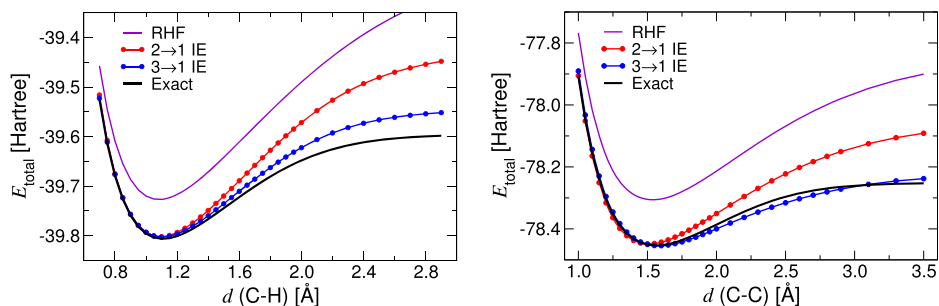

FIG. 4. The same calculation as shown in Fig. 3, with site densities and pair-densities self-consistently determined by IE.

$3 \rightarrow 1$ IE (blue) predicts equilibrium energies of very high accuracy (almost overlap the exact solution) and also improves the asymptotic behaviors significantly. This indicates that IE could effectively mitigate the sensitivity to the quality of the underlying approximate densities by using fragments of larger size.

The performance of non-self-consistent IE—especially the under-correlation at equilibrium geometry at the $2 \rightarrow 1$ level—is an indication that the error of this method is not purely density-driven. To confirm this inference, we repeat the IE calculations above but self-consistently determine the site densities and pair-densities. The results are presented in Fig. 4. By comparing it to Fig. 3, one can clearly see that imposing self-consistency significantly improves the results at equilibrium geometry at the $2 \rightarrow 1$ level and keeps the high accuracy of $3 \rightarrow 1$ IE at the same time. In the dissociation limit, however, imposing self-consistency has opposite effects: the PESs are shifted upwards slightly at both levels (though $3 \rightarrow 1$ has a much smaller amplitude) compared to the non-self-consistent results, making the under-correlation

problem even more severe therein. A scrutiny on the comparison of the site densities and pair-densities obtained by all these methods (Fig. S1 in the supplementary material) shows that the change from Figs. 3 to 4 is not density-driven, as the self-consistent densities and pair-densities are consistently worse than the non-self-consistent counterparts. Nevertheless, the results of $3 \rightarrow 1$ IE seem to be stable, especially near equilibrium geometries. In those cases, one can safely abandon the self-consistency condition without losing much accuracy.

C. Correlation energies at equilibrium geometry

As a final example, we investigate the effect of larger basis sets. As mentioned in Sec. II A, any deviation from half-filling deteriorates the performance of Schmidt-space embeddings using HF bath wave functions. In practice, however, large basis sets are often essential to recover the dynamic part of electron correlation. It is thus of significant importance to examine how IE behaves in large basis sets. In Fig. 5, we present in terms of bar plot the error of total energies for
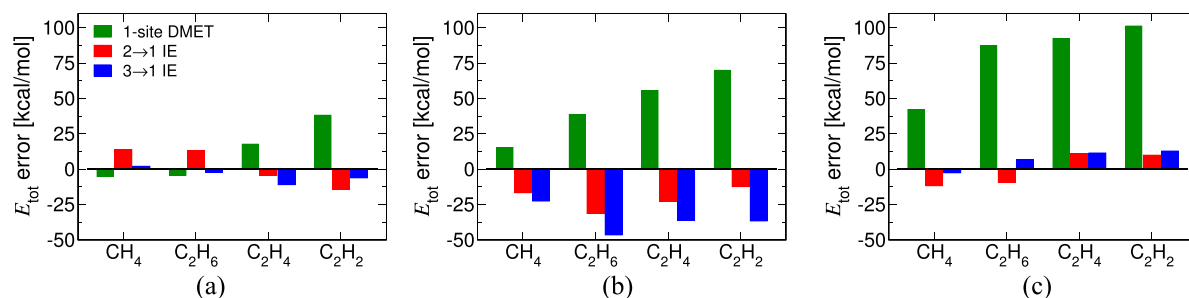

FIG. 5. Equilibrium geometry total energy errors (in kcal/mol) obtained by the non-self-consistent IE for several molecules in different basis sets: (a) STO-3G, (b) 3-21G, and (c) 6-311G. One-site DMET results are also included for comparison.

four molecules at equilibrium geometries predicted by non-self-consistent IE, along with one-site DMET for comparison. Three basis sets of increasing size are used: STO-3G, 3-21G, and 6-311G. For the minimal basis, all molecules are close to being half-filled. IE shows a consistent improvement with fragment size in all cases except for C_2H_4, and the errors of the $3 \rightarrow 1$ energies are within several kcal/mol's. One-site DMET, on the other hand, gives good results for CH_4 and C_2H_6 (as already seen in Fig. 3) but under-correlates badly for the other two unsaturated molecules. Since the densities are the same for all three cases, these results again confirm our conjecture that the errors of Schmidt-space embedding methods are not density-driven: compared to one-site DMET, IE successfully captures the entanglement among different sites, especially when fragments of larger size are used.

When the basis size is increased, the deviation from half-filling renders the under-correlation problem of one-site DMET even worse, as can be seen from the increasing heights of the green bars in Figs. 5(b) and 5(c). In terms of absolute values, one can see clearly that this is because the correlation energies recovered by one-site DMET remain unchanged or even drop slightly as the basis size gets larger, while the exact correlation energies always go up (Fig. S2 in the supplementary material). The same phenomenon appears in IE, but the trend there is more complicated. At first, switching from minimal to double-zeta basis leads to overestimation of correlation energies for all molecules. Moreover, the convergence with fragment size gets reversed: $3 \rightarrow 1$ predicts more negative numbers than $2 \rightarrow 1$ does, making it worse by going to a higher level of theory. Moving further to the triple-zeta basis, however, sets IE back on track: both $2 \rightarrow 1$ and $3 \rightarrow 1$ energies are of high accuracy. This occurs as a consequence of error cancellation: the correlation energies recovered by IE do not increase when going from 3-21G to 6-311G, which happens to cancel the over-correlation errors in the double-zeta basis accidentally (Fig. S2 in the supplementary material). This phenomenon, observed in both one-site DMET and IE, is in fact related to the basis unentanglement problem of the RHF bath. In Sec. VI, we will discuss this problem more thoroughly using a specific example.

VI. DISCUSSION

In Sec. II A, we briefly mentioned that the unentanglement in bath wave function (in this work, RHF) leads to linear dependency in the Schmidt decomposition [Eq. (2)] and effectively degrades the basis. Here we present a specific example that embodies this problem. We repeat our calculations in Sec. V A using a homemade basis set STO-3G*, obtained by adding one p_z orbital to each STO-3G hydrogen (assuming the atomic ring lies in the xy plane). The PESs obtained in the new basis are presented in Fig. 6 along with the original results for comparison. The first thing to notice is that the RHF solution remains unchanged. Population analysis suggests that the set of p orbitals are not occupied at all. This is expected by our construction of the new basis since all of the one-electron atomic integrals between the new orbitals and the original $1s$ orbitals vanish by symmetry. The only non-vanishing parts are the two-electron integrals involving an even number of
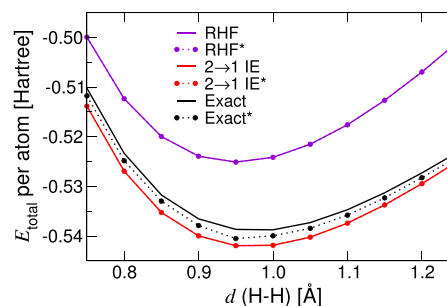

FIG. 6. A re-plot of Fig. 2 for the total energy of H_{10} near equilibrium geometry. Data labeled with an asterisk are computed in the new basis STO-3G*.

p orbitals such as $(p_z p_z | 1s 1s)$. Unfortunately, these non-trivial interactions are not captured by the mean-field wave function due to it being a one-electron theory. As a result, the Schmidt space in the new basis is exactly the same as in the original one, which explains the concurrence of two IE curves in Fig. 6. For the exact solution, however, those interactions do contribute to the total energy, and one can see a lower energy in the new basis.

This specifically designed example helps us to understand the trend shown in Fig. 5. As mentioned in Sec. V C, in terms of absolute values, the correlation energies recovered by both one-site DMET and IE cease to increase once the basis set reaches a certain size (see also Fig. S2 in the supplementary material). This observation indicates that the basis unentanglement problem of the mean-field bath could be common in realistic systems, especially when large-sized bases are used.

VII. CONCLUSION

In conclusion, we have introduced incremental embedding, a new Schmidt-space fragment embedding scheme that allows the combination of arbitrary overlapping fragments without the knowledge of edge sites. Underlying this new method is one of the key concepts introduced in this work, Schmidt reduction, which allows information to be encoded from a large embedding space to a smaller one. Based on this technique, IE constructs one-site effective Hamiltonians for all sites by hierarchically incorporating corrections from embeddings involving two-site fragments, three-site fragments, etc., to the mean-field approximation. The potential double counting problem is avoided by an elaborate application of the method of increments. This method can be viewed as an unambiguous many-site generalization of one-site DMET. It can be made either self-consistent or non-self-consistent. The computational scaling is $O(K^m)$ for the lowest few levels in the hierarchy, which are much lower than most correlated wave function theories.

Numerical simulations on small molecules in atom-centered Gaussian bases suggest that the convergence with fragment size is quick in small bases; most of the electron correlation is recovered for all molecules tested, even when truncated at the $3 \rightarrow 1$ level. Imposing self-consistency in site densities and pair-densities improves the performance of IE considerably near equilibrium geometry, through an approach that is not density-driven. For larger bases, both IE and one-site

DMET recover only a fraction of the correlation energy, which can be attributed to the more general basis unentanglement problem of the RHF bath. In summary, this work marks the first attempt of applying Schmidt-space embedding methods to realistic molecular systems using overlapping fragments.

In the future, IE can be extended in a number of directions. First of all, up to this point we restrict ourselves to FCI for the embedding Hamiltonian, which is computationally expensive and can only be applied to fragments of limited size. One can, of course, pursue other high-level solvers such as DMRG and CCSD, which have better computational scaling and can therefore be extended to fragments of larger size. Second, as for the site basis, we restrict ourselves in this work to SOAOs or LMOs by the Foster-Boys scheme for the sake of simplicity, but other choices do exist. In analogy to local correlation methods such as local MP2[40–42] and local CCSD,[43] perhaps the most straightforward way is to explore the possibilities of using other LMOs such as those given by the Pipek-Mezey scheme[59,60] and the Edmiston-Ruedenberg scheme.[61] Last but not least, the conflict between the half-filled embedding space and the maximum number of entangled sites in a HF bath calls for a better bath wave function. In this regard, the Hartree-Fock-Bogoliubov[62,63] (HFB) wave function might be a good candidate because it is (i) always half-filled in the quasi-particle space and (ii) still a mean-field theory and therefore retains the simplicity of embedding Hamiltonians in a mean-field bath.

SUPPLEMENTARY MATERIAL

See supplementary material for (i) the xyz-format geometry of all molecules, (ii) tabular data of all energy values, (iii) an algorithm of the constrained optimization outlined in Sec. III E, (iv) comparison of the site densities and pair-densities of CH_4 estimated by RHF and IE, and (v) total correlation energies of the molecules discussed in Sec. V C.

ACKNOWLEDGMENTS

[1] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[2] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).

[3] S. Chattopadhyay, R. K. Chaudhuri, and U. S. Mahapatra, J. Comput. Chem. **36**, 907 (2015).

[4] E. J. Taffet and G. D. Scholes, "The A_g^+ state falls below $3A_g^-$ at carotenoid-relevant conjugation lengths," Chem. Phys. (to be published).

[5] A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma, and P. M. Zimmerman, J. Phys. Chem. A **122**, 2714 (2018).

[6] C. Stemmle, B. Paulus, and O. Legeza, Phys. Rev. A **97**, 022505 (2018).

[7] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).

[8] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **120**, 6832 (2004).

[9] R. Z. Khaliullin, M. Head-Gordon, and A. T. Bell, J. Chem. Phys. **124**, 204105 (2006).

[10] D. G. Fedorov and K. Kitaura, J. Phys. Chem. A **111**, 6904 (2007).

[11] G. Senatore and K. R. Subbaswamy, Phys. Rev. B **34**, 5754 (1986).

[12] M. D. Johnson, K. R. Subbaswamy, and G. Senatore, Phys. Rev. B **36**, 9202 (1987).

[13] P. Cortona, Phys. Rev. B **44**, 8454 (1991).

[14] C. R. Jacob and J. Neugebauer, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 325 (2014).

[15] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).

[16] A. Georges and W. Krauth, Phys. Rev. Lett. **69**, 1240 (1992).

[17] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[18] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005).

[19] V. Turkowski, A. Kabir, N. Nayyar, and T. S. Rahman, J. Chem. Phys. **136**, 114108 (2012).

[20] I. Klich, J. Phys. A: Math. Gen. **39**, L85 (2006).

[21] I. Peschel and V. Eisler, J. Phys. A: Math. Theor. **42**, 504003 (2009).

[22] I. Peschel, Braz. J. Phys. **42**, 267 (2012).

[23] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc., Mineola, New York, 1996).

[24] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[25] F. Verstraete, V. Murg, and J. I. Cirac, Adv. Phys. **57**, 143 (2008).

[26] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[27] R. Ahlrichs and P. Scharf, *The Coupled Pair Approximation* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 1987), Vol. 67, ISBN: 9780470142936.

[28] G. Knizia and G. K.-L. Chan, Phys. Rev. Lett. **109**, 186404 (2012).

[29] G. Knizia and G. K.-L. Chan, J. Chem. Theory Comput. **9**, 1428 (2013).

[30] I. W. Bulik, G. E. Scuseria, and J. Dukelsky, Phys. Rev. B **89**, 035140 (2014).

[31] I. W. Bulik, W. Chen, and G. E. Scuseria, J. Chem. Phys. **141**, 054113 (2014).

[32] S. Wouters, C. A. Jiménez-Hoyos, Q. Sun, and G. K.-L. Chan, J. Chem. Theory Comput. **12**, 2706 (2016).

[33] B.-X. Zheng and G. K.-L. Chan, Phys. Rev. B **93**, 035126 (2016).

[34] B.-X. Zheng, J. S. Kretchmer, H. Shi, S. Zhang, and G. K.-L. Chan, Phys. Rev. B **95**, 045103 (2017).

[35] M. Welborn, T. Tsuchimochi, and T. Van Voorhis, J. Chem. Phys. **145**, 074102 (2016).

[36] N. Ricke, M. Welborn, H.-Z. Ye, and T. Van Voorhis, Mol. Phys. **115**, 2242 (2017).

[37] T. Tsuchimochi, M. Welborn, and T. Van Voorhis, J. Chem. Phys. **143**, 024107 (2015).

[38] Q. Wu and W. Yang, J. Chem. Phys. **118**, 2498 (2003).

[39] Q. Wu and T. Van Voorhis, Phys. Rev. A **72**, 024502 (2005).

[40] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).

[41] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).

[42] S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).

[43] C. Hampel and H. Werner, J. Chem. Phys. **104**, 6286 (1996).

[44] W. Press and W. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992), bk. 4, ISBN: 9780521437202.

[45] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[46] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[47] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006).

[48] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

[49] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard *et al.*, J. Chem. Theory Comput. **13**, 3185 (2017).

[50] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng *et al.*, Mol. Phys. **113**, 184 (2015).

[51] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).

[52] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[53] H.-Z. Ye, "Frankenstein embedding: A density matrix embedding theory for molecules," https://github.com/hongzhouye/frankenstein, 2017.

[54] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[55] G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).

[56] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[57] S. Sharma and G. K.-L. Chan, J. Chem. Phys. **136**, 124121 (2012).

[58] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, J. Chem. Phys. **142**, 034102 (2015).

[59] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[60] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).

[61] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

[62] N. N. Bogoliubov, V. V. Tolmachov, and D. V. Širkov, Fortschr. Phys. **6**, 605 (1958).

[63] N. N. Bogoliubov, Sov. Phys.-Usp. **67**, 236 (1959).