Density matrix embedding: A strong-coupling quantum embedding theory

Gerald Knizia[1] and Garnet Kin-Lic Chan[1]

Department of Chemistry, Princeton University, Princeton NJ 08540

We extend our density matrix embedding theory (DMET) [Phys. Rev. Lett. **109** 186404 (2012)] from lattice models to the full chemical Hamiltonian. DMET allows the many-body embedding of arbitrary fragments of a quantum system, even when such fragments are open systems and strongly coupled to their environment (e.g., by covalent bonds). In DMET, empirical approaches to strong coupling, such as link atoms or boundary regions, are replaced by a small, rigorous quantum bath designed to reproduce the entanglement between a fragment and its environment. We describe the theory and demonstrate its feasibility in strongly correlated hydrogen ring and grid models; these are not only beyond the scope of traditional embeddings, but even challenge conventional quantum chemistry methods themselves. We find that DMET correctly describes the notoriously difficult symmetric dissociation of a 4×3 hydrogen atom grid, even when the treated fragments are as small as single hydrogen atoms. We expect that DMET will open up new ways of treating of complex strongly coupled, strongly correlated systems in terms of their individual fragments.

Embedding techniques are powerful tools for enabling high-level many-body treatments on system sizes they cannot normally reach. They work by dividing a chemical system into fragments, each of which is handled individually; the interaction with the other fragments—the environment—is treated in a simplified way. In this communication we are concerned with embeddings for fragments which are *strongly* coupled to their environment, for example via covalent bonds.

A particular embedding is characterized by the precise manner in which the environment, and its influence on the fragment, are represented. To date, most techniques represent the environment through a one-particle embedding potential v. For example, in QM/MM methods, v is obtained through electrostatics or polarization interactions,[1,2] while in *ab-initio* DFT embedding, v is the derivative of the non-additive energy functional.[3–6] However, an embedding potential, regardless of how it is obtained, cannot represent the effect of the environment on the many-body fragment state when the coupling is *strong*.[7] This is illustrated by the simple example of embedding a hydrogen atom A within a hydrogen molecule AB. If hydrogen atom B is represented by an embedding potential v, then hydrogen atom A appears as a closed system with a single electron; thus, any wavefunction description of the fragment, regardless of the choice of v, provides no information on electron correlation.

This failure of potential based embedding is rooted in the fact that the fragments are *open* systems that are entangled with their environment. Is it possible *even in principle* to formulate an embedding description of an open fragment? This question has far reaching consequences; an affirmative answer would imply, for example, that one could in principle exactly calculate the properties of a bulk diamond crystal by treating a single embedded carbon atom. Here we argue that this is in fact the case. The key is to represent the environment not by a potential, but rather through a quantum *bath* that reproduces the entanglement of the embedded fragment with the full environment. In existing embedding approaches, empirical link atoms or boundary regions[8] can be thought of as baths, but here we show that an *exact* bath, that exactly reproduces *all* many-body environment effects, can in fact be defined. In realistic systems, the construction of this exact bath is not practical. However, as we show below, this point of view naturally leads to a practical embedding method which we call *density matrix embedding theory* (DMET).

We have previously introduced DMET in the context of fermionic lattice models[9] where it showed excellent performance; in particular, in comparison to the more complex dynamical mean-field theory (DMFT). Here we describe the extension and modification of DMET to treat inhomogeneous systems and the full chemical Hamiltonian with long-range interactions, namely

$$H = \sum_{rs} h_s^r E_r^s + \frac{1}{2} \sum_{rstu} V_{tu}^{rs} E_{rs}^{tu}. \tag{1}$$

We note especially that DMET is defined through simple linear algebra and avoids the numerical issues of inverse problems associated with constructing potential embeddings, as found in *ab-initio* DFT embedding (see also Ref. 10, another solution to the inverse problem).

We use the following notation: The full quantum system Q is spanned by an orthogonal one-particle basis (e.g., an symmetrically orthogonalized atomic orbital basis). The orthogonal basis functions are indexed by r,s,t,u (e.g., Eq. (1)) and referred to as *sites*. Q is divided into sets of sites called *fragments*, such that each site occurs in exactly one fragment. Fragments are handled sequentially, and for each fragment A, the union of the other fragments is treated as environment B. $|A|$ is the number of sites in set A.

First, let us review the the exact formal bath that exactly embeds a given fragment A. Let Q be the full quantum system, and $|\Psi\rangle$ be an eigenstate of Q. $|\Psi\rangle$ may expanded as

$$|\Psi\rangle = \sum_{ij} \psi_{ij} |\alpha_i\rangle |\beta_j\rangle \tag{2}$$

where $|\alpha_i\rangle$ and $|\beta_i\rangle$ are states in the Fock space spanned by the fragment A and the environment B, respectively. Simple algebra shows that

$$\begin{aligned}|\Psi\rangle &= \sum_{ij} \psi_{ij} |\alpha_i\rangle |\beta_j\rangle = \sum_i |\alpha_i\rangle \left(\sum_j \psi_{ij} |\beta_j\rangle \right) \\ &= \sum_i |\alpha_i\rangle |\tilde{\chi}_i\rangle = \sum_{ii'} \psi_{ii'} |\alpha_i\rangle |\chi_{i'}\rangle \end{aligned} \tag{3}$$

where $|\tilde{\chi}_i\rangle = \sum_j \psi_{ij} |\beta_j\rangle$, and $|\chi_i\rangle$ is the orthogonalized set of $|\tilde{\chi}_i\rangle$ states. (This rewriting is closely related to the Schmidt

decomposition of quantum information theory[11]). Note that in the last line, although $|\chi_{i'}\rangle$ are states in the environment B, there are only M_A of them: the dimension of the Fock space of A. This set of special environment states defines the bath. We see that (i) no matter how large the environment, in a given state $|\Psi\rangle$, a fragment A can only be entangled with M_A environment states in B. Thus the entanglement effect of the environment is *fully represented by a bath of the same size as the fragment* it is embedding. This combination of the fragment with its bath, we refer to as the *embedded* system. (ii) If $|\Psi\rangle$ is an eigenfunction of a Hamiltonian H (that acts in the full system Q), it is also an eigenfunction of a projected Hamiltonian defined in only the embedded system, $H' = PHP$, where

$$P = \sum_{ii'} |\alpha_i\rangle |\chi_{i'}\rangle \langle \alpha_i| \langle \chi_{i'}|. \tag{4}$$

H' is the *embedding Hamiltonian*.

We thus have established that fragments can, in principle, always be exactly embedded by baths no larger than the fragments themselves. While exact, this result is purely formal, because to construct the bath states $\{|\chi_i\rangle\}$ we require knowledge of the state $|\Psi\rangle$ of the full system Q. However, it naturally suggests a practical approximation: construct the bath from an approximate state of Q, $|\Phi\rangle$, and use this approximate bath in a subsequent high-level treatment of the embedded fragments. This is the combination we aim for in DMET.

A simple choice for $|\Phi\rangle$ is a Slater determinant (for example, as obtained from a mean-field treatment of the full system). For a Slater determinant, the associated bath and embedding Hamiltonian are particularly simple: they can be obtained from single-particle linear-algebra rather than the many-particle decomposition in Eq. (3), as we now show. First note that, for any $|\Psi\rangle$, the fragment many-body states $|\alpha_i\rangle$ in Eq. (3) live in the Fock space spanned by the one-particle fragment sites $\mathscr{F}(|i\rangle)$, $i \in A$. In the special case where $|\Psi\rangle = |\Phi\rangle$ is a determinant, also the bath states $|\chi_i\rangle$ live in a Fock space defined by one-particle bath orbitals $\{|b\rangle\}$ (at most $|A|$), multiplied by a common core determinant. This is seen as follows. Let $|p\rangle = \sum_r C_p^r |r\rangle$ denote the N occupied orbitals of $|\Phi\rangle$, where $p = 1 \ldots N$, $r \in Q$. Let

$$S_{pq} = \sum_{i \in A} \langle p|i\rangle \langle i|q\rangle \tag{5}$$

define the overlap matrix S of the orbitals *projected onto the sites of fragment A*. Then the eigenvectors of S define a rotation of the occupied orbitals $|p\rangle \to |\tilde{p}\rangle$ which divides them into two sets: a set of $N - |A|$ occupied orbitals with zero eigenvalues, and thus *without* any component on the fragment sites, and a set of $|A|$ occupied orbitals with non-zero eigenvalues, which have overlap with the fragment sites. We call the former "pure environment orbitals", and the latter "entangled orbitals". Projecting the entangled orbitals $\tilde{p} = 1 \ldots |A|$ onto the environment sites B and normalizing then yields a set of bath orbitals $|b\rangle$ of the same number as fragment sites,

$$\{|b\rangle\} = \left\{ \frac{\sum_{j \in B} |j\rangle \langle j|\tilde{p}\rangle}{\|\sum_{j \in B} |j\rangle \langle j|\tilde{p}\rangle\|}; \ \tilde{p} = 1 \ldots A \right\}. \tag{6}$$

Rewriting $|\Phi\rangle$ in terms of the rotated orbitals $|\tilde{p}\rangle$, and expressing each $|\tilde{p}\rangle$ in terms of its fragment, bath, and pure environment components, we see that the many-body bath states $|\chi_i\rangle$

span the same space as $\mathscr{F}(|b\rangle) \otimes \det(e_1 e_2 \ldots e_{N-|A|})$, where $\det(e_1 e_2 \ldots e_{N-|A|})$ is the determinant of pure environment orbitals. In other words, when split across a fragment and environment, a determinant $|\Phi\rangle$ appears as a CAS-CI (complete active space configuration interaction) expansion in a half-filled active *embedding* basis of fragment plus bath orbitals, $\{|i\rangle\} \oplus \{|b\rangle\}$, with a core determinant of pure environment orbitals $|e\rangle$.

We next construct the embedding Hamiltonian corresponding to the Slater determinant $|\Phi\rangle$. Formally, this is defined from the many-body projection Eq. (4), but for the case of a Slater determinant, it can be constructed by a simple change of single-particle basis. We define the embedded Hamiltonian H' by projecting H into the space $\mathscr{F}(|i\rangle) \otimes \mathscr{F}(|b\rangle) \otimes \det(e_1 e_2 \ldots e_{N-|A|})$. This is equivalent to transforming H into the active space of the embedding basis (fragment plus bath orbitals), and including a core contribution from the pure environment determinant, $\det(e_1 e_2 \ldots e_{N-|A|})$. Denoting the embedding basis by labels v, w, x, y, and its representation in terms of the full system sites by $|x\rangle = B_x^r |r\rangle$, we find

$$H' = \sum_{vw} h_w^v E_v^w + \frac{1}{2} \sum_{vwxy} V_{xy}^{vw} E_{vw}^{xy}, \tag{7}$$

$$h_w^v = B_v^r h_s^r B_w^s + f_{vw}^{\text{core}} \tag{8}$$

$$V_{xy}^{vw} = B_v^r B_s^w V_{tu}^{rs} B_x^t B_y^u. \tag{9}$$

If the full system $|\Phi\rangle$ was obtained from a Hartree-Fock calculation, then carrying out a Hartree-Fock calculation in this embedded system with the embedding Hamiltonian H' yields a mean-field Fock operator f' for which $|\Phi\rangle$ is an eigenstate.

In Fig. 1 we numerically demonstrate the exactness of the above mean-field embedding by the following process: (i) a Hartree-Fock calculation is performed on a molecule, (ii) the molecule is split into arbitrary groups of atoms as fragments. For each fragment, an embedding is constructed (bath orbitals and H') and a Hartree-Fock calculation is run on the embedded system, (iii) the system is reassembled by adding up the energies and electrons located on the various fragments. We note that only the electrons and energy contributions associated with the fragment sites should be considered when reassembling the system, not the contributions associated exclusively with bath or pure environment orbitals, as this would lead to double counting. To that end, we define the energy of a fragment A using the one- and two-particle density matrices, γ, Γ, with at least one index in fragment A,

$$E_A = \sum_{i \in A, s} \gamma_s^i h_i^s + \frac{1}{2} \sum_{i \in A, stu} \Gamma_{tu}^{is} V_{is}^{tu}. \tag{10}$$

Note that γ and Γ include the contributions from the pure environment orbitals (the core determinant). As we see from Fig. 1, the embedding allows an arbitrary fragmentation into open fragments to be exactly reassembled (note the fractional electron numbers!), recovering the exact Hartree-Fock energy of the wave function $|\Phi\rangle$ used to contruct it.

We now return to the DMET. Recall that here we still construct the embedding based on a mean-field $|\Phi\rangle$ of the full system Q, but use it to embed high level, calculations on the
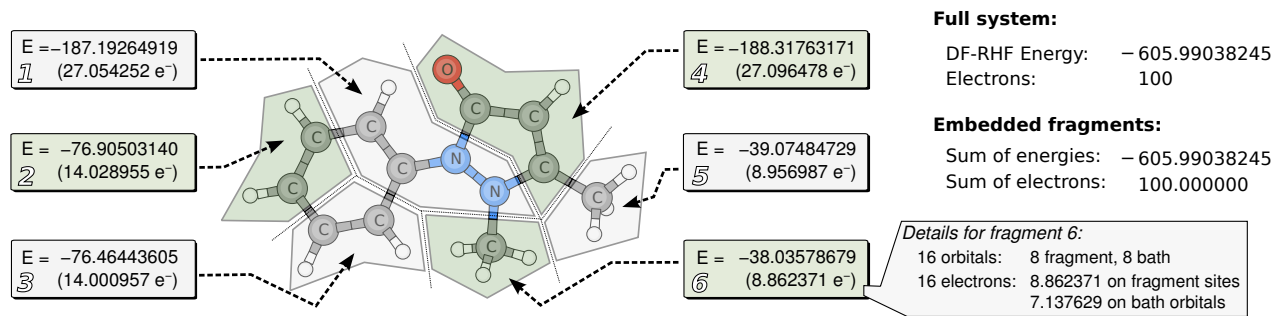
FIG. 1. Example for HF in HF embedding: A molecule is split into random atomic fragments. For each fragment, an embedding is constructed, and the embedded system is treated with Hartree-Fock. The left panel shows the obtained distribution of energy and electrons amongst the fragments. The right inset shows how the open system embedding (e.g., a fractional electron number) is realized by coupling to a bath.

fragments, rather than Hartree-Fock. For each fragment A, the high level calculation on the embedded system (fragment sites, bath orbitals, and Hamiltonian H') yields a correlated $|\Psi_A\rangle$. Here we require an additional self-consistency cycle to ensure a consistent fragment description by the mean-field $|\Phi\rangle$ and the high-level embedded $|\Psi_A\rangle$. Although the choice of self-consistency condition is not unique, we note that a mean-field state is characterized by its one-particle density matrix $\langle\Phi|a_r^\dagger a_s|\Phi\rangle$, and thus it is convenient to enforce consistency at the level of the density matrices by minimizing the metric

$$\Delta = \sum_A \sum_{rs\in A} \| \langle\Phi|a_r^\dagger a_s|\Phi\rangle - \langle\Psi_A|a_r^\dagger a_s|\Psi_A\rangle \|^2. \quad (11)$$

Here we define the density matrix difference only over (intra)fragment sites for simplicity, but very similar results are obtained by defining the difference over fragment plus bath sites for each embedding, as we did in our earlier work.[9] $|\Phi\rangle$ is then obtained from a mean-field Fock operator f augmented by a set of one-particle operators u_A for each fragment, with $u_{rs}^A = 0$ for $r \notin A$ or $s \notin A$. The u_A are chosen to minimize Δ and capture the correlation effects on the one-particle density matrix. The embedded Hamiltonian is also augmented with the correlation operators on the fragments other than the one currently being considered, projected into the embedding basis. The DMET self-consistency cycle is thus:

1. The full system is treated at the Hartree-Fock level using the Fock operator $f + \sum_A u_A$ to determine $|\Phi\rangle$. Initially all u_A are zero.

2. For each fragment, $|\Phi\rangle$ is used to construct an embedding basis. The embedding Hamiltonian for fragment A, H'_A, is obtained by projecting $H + \sum_{A'\neq A} u_{A'}$ into the embedding basis following Eq. (7). The embedded fragment's state $|\Psi_A\rangle$ is calculated at a correlated level, for example, with full configuration interaction (FCI).

3. For each fragment A, we adjust the correlation operator $u_A = \sum_{rs\in A} u_{rs}^A E_r^s$ to minimize the difference between the Hartree-Fock one-particle density matrix and the correlated one-particle density matrix, $\Delta_A = \sum_{rs\in A} \| \langle\Phi|a_r^\dagger a_s|\Phi\rangle - \langle\Psi_A|a_r^\dagger a_s|\Psi_A\rangle \|^2$.

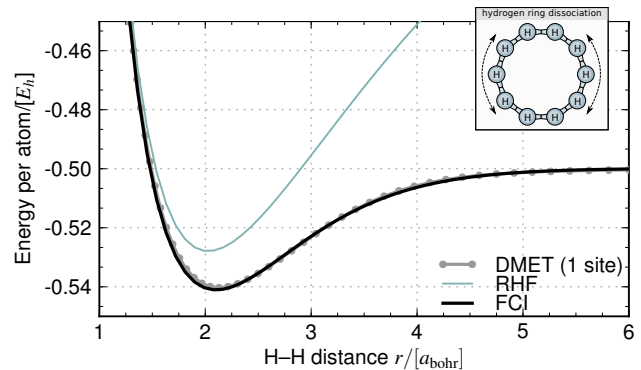4. The cycle is iterated until all u_{rs}^A converge.

FIG. 2. Symmetric stretching of a H_{10} ring: Shown are the exact results (FCI), the mean-field results (RHF), and the 1-site DMET results. In 1-site DMET, the system is fragmented into individual atoms treated with FCI, and embeddings are constructed from full system RHF.

In order to test whether the DMET provides reasonable results, we now consider simple model systems exhibiting strong correlation; namely, hydrogen rings, chains, and grids. Such systems have recently emerged as a rich testbed for new correlation methods, as the strength of the correlation can be readily tuned from weak to strong by changing the hydrogen atom spacing.[12–17] Note that such systems show high degeneracy and are difficult to handle even with full correlation treatments. Furthermore, as we argued in the introduction, any choice of fragment will be strongly coupled to the rest of the system, so potential based embeddings cannot describe them. All calculations employ a minimal hydrogen basis consisting of orthogonalized $1s$-like AO functions obtained from an underlying cc-pVTZ basis, except for the H_{50} calculation, for which we use a STO-6G basis to retain compatibility with earlier references. For the high level treatment of the embedded systems, we use FCI. All calculations are spin-adapted, and both the mean-field (restricted HF) and correlated (FCI) calculations use singlet wave functions. As DMET self consistency metric we used Eq. (11), except for the 3-site calculation on the 4×3 grid, where the previous metric[9] was used to avoid convergence problems at $r > 3.8a_{\text{bohr}}$.

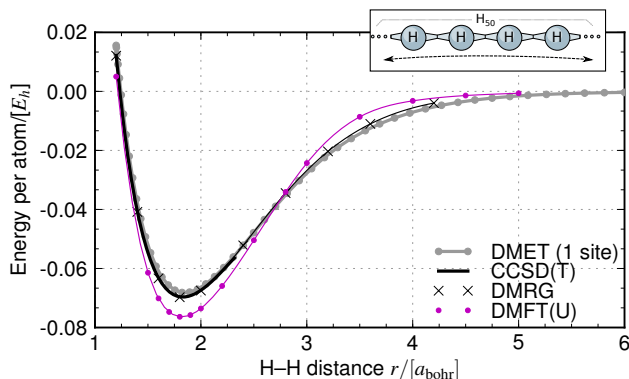As the simplest non-trivial example, we investigate the

FIG. 3. Symmetric stretching of a H_{50} chain: Shown are reference DMRG[20] and CCSD(T) data, the mean-field results (RHF), the 1-site DMET results, and the DMFT results by Lin et al.[17]
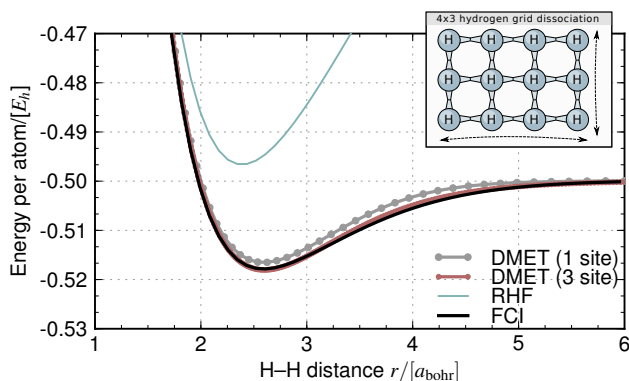

FIG. 4. Symmetric stretching of a 4×3 hydrogen grid. In the 3-site DMET, the system is fragmented into four columns of three hydrogens each.

symmetric dissociation of a ring of ten hydrogen atoms (that is, all bonds are simultaneously stretched). The system is fragmented into individual $1s$ orbitals of the hydrogen atoms, and thus each site is one fragment. The results are shown in Fig. 2. The 1-site DMET calculation almost exactly reproduces the reference FCI curve. Note that each correlated fragment calculation corresponds to a FCI calculation on only two orbitals: the H $1s$ orbital, and a single bath orbital, and is thus a trivial 3×3 matrix diagonalization.[18]

As second example, we choose an inhomogeneous system: the linear H_{50} chain. FCI on this system would require on the order of 10^{28} determinants, so truly exact results cannot be calculated. However, near-exact reference data can be obtained by the quantum chemistry density matrix renormalization group (DMRG).[19] We here take the data from Hachmann et al.[20] This particular system has also been the subject of a recent DMFT study by Lin et al.,[17] the results of which are shown for comparison. It is clear that the 1-site DMET once again closely reproduces the reference values. Again the fragment calculations are two orbital FCI and thus numerically trivial. It is noteworthy that in this case DMET does better than the more complex DMFT, likely due to its ability to treat long-range interactions with the bath beyond mean-field.

For a more challenging system, we now turn to a two-dimensional inhomogeneous system, the 4×3 hydrogen grid. This system is very pathological—it is non-binding at Hartree-Fock level, and converging the normally *very* robust FCI[21] required hundreds of iterations for some points. As shown in Fig. 4, even here the 1-site DMET qualitatively reproduces the FCI binding curve. The calculations remain as trivial to perform as for the one-dimensional systems. Additionally, if we embed entire columns of atoms, carrying out a 3-site DMET, the agreement between the embedded and the reference results becomes almost perfect.

The accuracy of these results may seem surprising, given that the systems are strongly correlated but the embedding is obtained from an uncorrelated, qualitatively incorrect, mean-field $|\Phi\rangle$. However, it is only the bath *states* that are determined from the mean-field theory. Once those states are determined, the embedded Hamiltonian is constructed exactly, and the coupling between the fragment sites and the bath orbitals is obtained by FCI, not mean-field. This allows the fragment state to correctly transition from the delocalized regime of weak correlation to the entangled spin regime of strong correlation. As long as the entanglement is reasonably local and does not involve the *pure* environment orbitals, we can expect good results.

The robustness and simplicity of DMET, even in the presence of strong coupling and strong correlation, makes it unique amongst current embedding approaches, and suggest that it could be useful in a wide range of applications. The next step will be to apply the theory to more realistic and larger scale chemical problems. We are now pursuing these studies.

[1] H. Lin and D. Truhlar, Theor. Chem. Acc. **117**, 185 (2007).

[2] H. M. Senn and W. Thiel, Angew. Chem. Int. Ed. **48**, 1198 (2009).

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

[4] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[5] C. Huang, M. Pavone, and E. A. Carter, J. Chem. Phys. **134**, 154110 (2011).

[6] J. D. Goodpaster, N. Ananth, F. R. Manby, and I. T. F. Miller, J. Chem. Phys. **133**, 084103 (2010).

[7] Note we are careful to write "many-body" fragment state. A potential *is* formally sufficient to represent the effect of the environment on the single-particle *density*, and this is the basis of exact DFT embedding. However, we assume here that we wish to describe the fragment with a high-level *many-body* treatment, not only at the level of the density. In this case, the environment cannot be represented by a potential.

[8] P. Slavíček and T. J. Martínez, J. Chem. Phys. **124**, 084107 (2006), provides an overview over such techniques.

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

[10] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, J. Chem. Theory Comput. **8**, 2564 (2012).

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

[12] T. Tsuchimochi and G. E. Scuseria, J. Chem. Phys **131**, 121102 (2009).

[13] A. O. Mitrushchenkov, G. Fano, R. Linguerri, and P. Palmieri, Int. J. Quantum Chem. **112**, 1606 (2012).

[14] L. Stella, C. Attaccalite, S. Sorella, and A. Rubio, Phys. Rev. B **84**, 245117 (2011).

[15] W. A. Al-Saidi, S. Zhang, and H. Krakauer, J. Chem. Phys. **127**, 144101 (2007).

[16] L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis, and G. E. Scuseria, J. Chem. Phys. **135**, 044119 (2011).

[17] N. Lin, C. A. Marianetti, A. J. Millis, and D. R. Reichman, Phys. Rev. Lett. **106**, 096402 (2011).

[18] Two orbitals span four two-electron states, of which three are singlet.

[19] G. K.-L. Chan and S. Sharma, Ann. Rev. Phys. Chem. **62**, 465 (2011).

[20] J. Hachmann, W. Cardoen, and G. K.-L. Chan, J. Chem. Phys. **125**, 144101 (2006).

[21] P. Knowles and N. Handy, Chem. Phys. Lett. **111**, 315 (1984).