Wave Function Frozen-Density Embedding: Coupled Excitations

Sebastian Höfener† and Lucas Visscher*

Amsterdam Center for Multiscale Modelling (ACMM), VU University Amsterdam, Theoretical Chemistry Section, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

ABSTRACT: We report quasi-ab initio correlated ground-state and excitation-energy calculations for agglomerates consisting of several molecules with total system sizes of up to more than one hundred atoms using a combination of a density-fitted, approximate second-order coupled-cluster singles and doubles (RICC2) method and frozen-density embedding (FDE), denoted RICC2-in-RICC2. Working equations are presented for CC2 ground-state energies and approximate coupled excitation energies, which are a necessary prerequisite for investigations of potential energy surfaces (PESs) of both ground and excited states. The approach is applicable to all systems that can be decomposed into interacting individual molecules for which the RICC2 calculation itself is feasible. Because of the absence of exact exchange in the formalism and the possibility to efficiently evaluate Coulomb coupling integrals using density fitting or a dipole approximation, the coupling step is insignificant in terms of computation time.

I. INTRODUCTION

In many fields of science, transfer processes of energy as well as matter are of fundamental interest. Quantum chemistry is a useful tool to study for instance energy or electron transfer from a photoexcited molecule to a molecule in its ground state. Among the most prominent examples is photosystem II in photosynthesis, where special chlorophyll and carotenoid antennae molecules collect photons, followed by a series of transfer processes, eventually arriving at the reaction center in which water oxidation takes place.

Such photophysical processes are commonly called resonance energy transfers (RET) or electronic energy transfers (EET). An approximate dipole–dipole description of EET was already achieved by Förster in the late 1940s1 and is well-suited for long distances. An important extension to the Förster model is the Dexter formulation,2 which includes higher multipolar terms as well as exchange effects being important for short distances with significant overlap of the monomer electron clouds. Together, these models form the basis for most qualitative discussions on energy transfer between distinguishable subsystems.

In more detailed models, quantum chemistry is used in the calculations of EET rate constants. One target is the excitonic coupling constants between the donor and acceptor transition, which are directly related to the energy difference between the coupled stationary electronic states of two chromophores. Although supermolecular treatments are feasible for small model systems, a problem often arises in the identification and proper description of the excitonic states, which is needed to extract these constants. This problem can be remediated by the use of fragment or subsystem approaches3,4 in which the basis of localized states is defined a priori.

Several methods have been proposed that can be characterized by specifying both the level of detail with which the interaction between the subsystems is described5 and the electronic structure method(s) that are used for the individual systems. The simplest approach is to obtain transition densities for noninteracting systems and calculate the electrostatic coupling between localized states without resorting to the often too crude dipole approximation.6,7 This approach can be extended by including the bridge and solvent effects that contribute to the coupling8,9 and by including the polarization of subsystems by their environment.10,11 Utilizing the fact that the electron density alone is sufficient to describe the electron energy,12 a formally exact theory can be defined by treating the interaction between subsystems in the framework of density functional theory (DFT).13 This frozen density embedding (FDE) theory of Wesolowski and Warshel14 is nowadays the basis for many subsystem DFT methods.15 The first extension of the FDE method to excited states in terms of time-dependent DFT (TDDFT) was presented by Casida and Wesolowski16 and focuses on states of only one subsystem. The calculation of these states that are only perturbed but not mixed by states of the environment is nowadays called FDEu to distinguish it from the general subsystem TDDFT or coupled FDE-TDDFT (FDEc) approach that was formulated later by Neugebauer.17,18 This FDEc formalism is very efficient and allows calculation of excitonic couplings for large aggregates of pigments.19 As shown in our previous work,20 the FDEc formalism can also be cast in the form of a quasi-energy response formalism to allow for a coupled-cluster treatment of one or more subsystems. This theoretical work forms the basis...
for our current paper in which we report the implementation of the coupling terms between the subsystem excited states, also making the step from FDEu to FDEC for subsystems entirely treated by coupled-cluster type methods.

Prerequisite to the coupling step is the calculation of local excited states, which we will do via response theory, calculating their energy as the sum of the ground state energy and an excitation energy. An alternative is to calculate the energy of excited states directly, as done in FDE context by Daday,28 and define multiple embedding potentials (one for the ground state and one for each excited state that is considered). This direct optimization of excited states formally provides access to only a subset of the excited states,22 although in applications24 this appears to be a minor limitation. A more practical problem with this approach is that the generation of state-specific embedding potentials can be cumbersome and that the orbitals for different excited states are nonorthogonal. This complication can be avoided by introducing a linearization approximation.25

These orbital orthogonality issues do not appear in the linear response approach as only the ground state energy functional and its derivatives are needed. All state-specific response of the environment is contained in the (frequency-dependent) response kernel. As such effects can be localized and most available density functional approximations for the exchange-correlation and kinetic energy are only semilocal and frequency-independent, calculations are needed to assess the accuracy of the method. In our previous work on coupled-cluster embedding, CC-in-DFT25 as well as CC-in-CC50 uncoupled excitation energies were obtained, but the couplings of the subsystems were not included. These important couplings are presented in the current work and demonstrated for prototypical examples of the splitting of the energy levels of weakly interacting chromophores (ethylene, benzene, porphyrin). Additionally, we also provide an example of the calculation of interaction energies between two subsystems of which the density is calculated using coupled cluster energy. Together, these implementations make the study of excimer complexes possible as they give access to the relevant potential energy surfaces (PES).

The article is organized as follows: we first present the calculation of the ground-state energy followed by the treatment of coupled excitations. After providing the computational details, we apply the new methods and investigate a representative ground state PES as well as selected excitation couplings between different excited states of the prototypical chromophores of C2H4, benzene, and porphyrin. The article closes with conclusions and outlook.

II. THEORY

To calculate the energy of an excited state with linear-response theory, we need to obtain the ground state and excitation energies relative to the ground state. We start with presenting ground-state energies followed by the description of the treatment of coupled excitations within the framework of coupled-cluster (CC) subsystem DFT (sDFT).

A. Ground-State Energy. The fundamental ansatz in sDFT is to divide the supermolecular Lagrangian, e.g., in case of two subsystems into three parts,

$$L = L_g + E_{int} + L_{exc}$$

where $E_{int}$ is the interaction contribution and $L_g$ and $L_{exc}$ are the two subsystem energy functionals. In the following, we will discuss the special case of coupled-cluster wave functions

$$|CC⟩ = e^T[HF]$$

where $T$ is the cluster operator and $|HF⟩$ is the reference Hartree–Fock wave function. In the case of a vacuum, the Lagrange functional is conveniently obtained using the similarity-transformed Hamiltonian, e.g., for subsystem I:

$$L_i = ⟨HF|e^{-T}He^{-T}|HF⟩ + \sum_{\mu} \bar{t}_\mu (\mu e^{-T}He^{-T}|HF⟩)$$

where $\mu$ indicates all excitations from all excitation levels (singles, doubles, etc.) and $T$ denote the Lagrangian multipliers. The subscript I attached to the Hamiltonian indicates that only electronic and nuclei coordinates corresponding to subsystem I are included.

In sDFT, the interaction energy is expressed in terms of subsystem densities used in classical Coulomb interaction expressions as well as nonadditive exchange-correlation $E_{exc}$ and kinetic energy $E_{kin}$ density functionals

$$E_{exc} [\rho_1, \rho_{II}] = \int \rho_1 (r) \bar{v}^{II}_{exc} (r) dr + \int \rho_{II} (r) \bar{v}^I_{exc} (r) dr$$

$$+ E^{II}_{kin} \int \rho_1 (r) \rho_{II} (r') \frac{dr'}{|r - r'|}$$

$$+ E^{nadd}_{exc} [\rho_1, \rho_{II}] + E^{nadd}_{kin} [\rho_1, \rho_{II}]$$

(4)

It is important to note that, because of the sDFT ansatz, there is no explicit embedding-potential contribution in the Hamiltonian $H$, e.g., in eq 3, because these contributions arise from the interaction energy expression in eq 4.

If coupled-cluster methods are used to treat the subsystems, coupled-cluster densities are used in the above equation, which can be expressed as

$$\rho(r)_{CC} = \langle \Lambda e^{-T} \Delta (r) e^T | HF⟩$$

$$= \langle HF | \sum_{\mu} \bar{t}_\mu (\mu e^{-T} \Delta (r) e^T | HF⟩$$

(5)

where $\Delta (r)$ the density operator at point $r$. In this so-called "unrelaxed" coupled-cluster formalism, the deviation of the Hartree–Fock density is entirely parametrized by the values of the excitation amplitudes in the cluster operator $T$. The Hartree–Fock density itself can, however, be relaxed with respect to the environment in freeze–thaw cycles.31,32 For interaction energy contributions with linear dependence on each subsystem density, this leads to the following contribution

$$\int \rho_1 (r) \bar{v}_{exc} (r) dr = \langle \Lambda e^{-T} \bar{v}_{exc} e^T | HF⟩$$

(6)

if subsystem I is treated at the coupled-cluster level of theory. It is not possible to express the exchange-correlation and nonadditive kinetic contribution in a similar (operator-based) manner due to their nonlinear dependence on the density. For these nonadditive contributions, the (unrelaxed) coupled-cluster electron density (and density gradient) has to be evaluated on a grid from which the energy can be calculated with the usual numerical integration techniques of DFT. In the equations, we do not consider a separate functional for errors made in the description of correlation energy in the CC calculation,33 as no practical approximate functional to improve the CC energy is available.

To obtain environment effects in the subsystem coupled-cluster calculation in the form of an embedding potential, we
needed to differentiate the interaction energy in eq 4 with respect to the variables of the coupled cluster wave function followed by collecting different contributions in an operator expression. For the amplitude equations, this procedure leads to

\[
\frac{\partial L}{\partial \mu} = \langle \mu e^{-T} H \bar{e}^T | \bar{F} \rangle + \frac{\partial E_{\text{int}}}{\partial \mu} \frac{\partial L}{\partial \mu} = \langle \mu e^{-T} (H + v_{\text{emb}}) \bar{e}^T | \bar{F} \rangle
\]

(7)

to give an example. Although written similar to a conventional operator, the embedding term arises due to the expansion and does not possess properties of a conventional second-quantization operator. Therefore, it cannot be added to the Hamiltonian a priori. To stress this point, this term is denoted embedding potential (instead of embedding operator).

Compared to a conventional vacuum treatment, the Lagrangian term in eq 3 is not zero because the amplitudes are optimized in the presence of the embedding potential, whereas the Hamiltonian does not contain embedding contributions. However, because of the stationary condition in eq 7, the Lagrangian contribution can conveniently be evaluated, and the subsystem energy for subsystem I, for example, thus reads

\[
E_{\text{CC},I} = \langle \text{HF} e^{-T} H e^T | \bar{F} \rangle + \sum \tilde{t}_i \langle \mu e^{-T} H e^T | \bar{F} \rangle
\]

(8)

This energy will be higher than the energy obtained in a vacuum calculation, where the difference is denoted promotion energy caused by the polarization of the system due to the environment.

1. Working Equations. In analogy to the Hamilton operator, the Fock matrix \( F \) is defined without the embedding potential

\[
F_{pq} = h_{pq} + \frac{1}{2} G_{pq} \{ D_{\text{HF},i} \}
\]

(9)

where \( h_{pq} \) denote the conventional one-electron and \( G_{pq} \) the conventional two-electron contributions, respectively. However, the embedding potential is added in the Hartree–Fock (HF) iterations, and the orbital energies thus obey

\[
\epsilon_p^e \delta_{pq} = F_{pq} + v_{\text{emb}}^q
\]

(10)

The working equations for the coupled-cluster energy with the embedding potential present in the HF iterations of subsystems \( i \) therefore read

\[
E_{\text{CC},i} = \sum h_{pq} D_{\text{HF},i} + \sum G_{pq} \{ D_{\text{HF},i} \} D_{\text{HF},i} + E_{\text{nuc}}^i - 2 \sum_{a} v_{\text{emb}}^{ia} \tilde{t}_a + \sum_{ia,jb} (t_{ia}^b + t_{ia}^b)(2(\sigma_{iajb}) - (\sigma_{ija}))
\]

\[
- \sum_{ia} \tilde{t}_i (\sum_{ck} (2t_{ick}^a - t_{ick}^a) v_{ck}^a + v_{ik}^a) + \sum_{ia,jb} P_{ij}^{ab} \tilde{v}_{ij} (\sum_{ck} v_{ck}^a - \sum_{ck} v_{ck}^a) + \sum_{ia,jb} P_{ij}^{ab} \tilde{v}_{ij} (\sum_{ck} v_{ck}^a - \sum_{ck} v_{ck}^a).
\]

(11)

A hat denotes \( T_{\text{i}} \)-transformed quantities, \( \tilde{v} = \exp(-T_{\text{i}}) \) \( \nu \exp(T_{i}) \), and \( P_{ij}^{ab} \) is a permutation operator, such that

\[
P_{ij}^{ab} X(i, a, j, b) = X(i, a, j, b) + X(j, b, i, a)
\]

(12)

where \( X \) can be any four-index entity. The treatment of the Lagrangian constraints and multipliers is discussed in prior publications.

B. Coupled Excitations. The treatment of coupled excitations in sDFT is known as FDEc to distinguish the approach from the uncoupled FDEu approach in which the response of the environment is only partially included. In both approaches, charge-transfer from one subsystem to another is excluded because only local excitations are coupled. We follow Neugebauer’s approach \(^{17} \) in which the subsystem linear-response functions are solved separately for each subsystem in the presence of a relaxed embedding potential as obtained in freeze–thaw iterations and are coupled in the final step. Although the FDEc scheme has been introduced and discussed in ref \(^{17} \), we present a full derivation in the following, because for coupled-cluster methods, a nonsymmetric eigenvalue problem has to be solved, resulting in fundamental differences in the linear-response treatment compared to variational methods.

Starting point for the derivation is the exact supermolecular coupled-cluster eigenvalue problem

\[
\lambda F_0 \Omega = \lambda \Omega \Omega_0
\]

(13)

where \( \lambda \) are the left eigenvectors, \( \Omega \) are the right eigenvectors of the supermolecular Jacobian \( F \), and \( \Omega \) is a diagonal matrix containing the eigenvalues. Following the sDFT ansatz, \( F \) is written as the subsystem Jacobians \( J \) and coupling \( A \)

\[
F = \left[ \begin{array}{cc} J_{\text{I,II}} & 0 \\ 0 & J_{\text{II,II}} \end{array} \right] + \left[ \begin{array}{cc} 0 & A_{\text{I,II}} \\ A_{\text{II,II}} & 0 \end{array} \right]
\]

(14)

where no approximation is introduced. We can similarly express the excitation vectors \( \Omega \) in subsystem excitation vectors \( R \) and a rotation matrix \( \bar{\Omega} \)

\[
\Omega = R \bar{\Omega} = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix} \begin{pmatrix} \bar{\Omega}_{\text{I,II}} \\ \bar{\Omega}_{\text{II,II}} \end{pmatrix}
\]

(15)

where \( R_i \) are chosen as the right-hand eigenvectors of subsystem \( i \),

\[
J_{\text{i}} R_i = \omega_i R_i
\]

(16)

Introducing a similar (exact) partitioning of the left eigenvectors into \( L \) and \( \lambda \), respectively, yields

\[
\lambda L(J + A) R_0 = \lambda \Omega \Omega_0
\]

(17)

where each matrix consists of two subsystem blocks on the diagonal and two coupling blocks on the off diagonal according to the equations above.

Up to this point, no approximation has been introduced in the coupling step except for the neglect of intersystem charge-transfer excitations. The method is characterized by the intrinsic sDFT ansatz of nonorthogonal intersubsystem orbitals and overlapping subsystem densities in which all effects of nonorthogonality can be formally accounted for by the nonadditive kinetic energy and exchange correlation functionals and their density derivatives. In practice, one needs to be careful with strongly overlapping systems, as the available density functional approximations for these terms are known to yield unreliable results in such cases. For weakly overlapping systems in which the exchange contribution to the couplings is small, the sDFT approach is known to work rather well. Still,
direct implementation of eq 17 is not desirable because it basically yields a reformulation of the supermolecular response equations that is costly to solve. Important is to first solve a subspace, i.e., reduced eigenvalue problem, and limit the number of excited states that need to be coupled. This ensures an efficient and general treatment in which the cost of the calculation primarily depends on the calculations that are carried out for the individual subsystems. We will discuss the realization of this approach in the framework of coupled-cluster theory in the following.

1. Approximate Coupled Excitations: Subspace Transformation. As shown by Neugebauer for TDDFT, the key step to achieve a computationally efficient procedure is to first transform the full eigenvalue problem to a reduced one. This step is a bit more complicated in coupled cluster theory because the Jacobian is not symmetric and has different left and right eigenvectors that are not orthogonal among themselves, but only among the two sets,

$$\mathbf{L}_R \mathbf{Q}_R = \mathbf{1}$$

(18)

which is denoted a biorthonormal basis. To obtain the exact diagonal representation and have access to properties such as oscillator strengths, both left and right eigenvectors need to be determined. A more efficient treatment is possible if only excitation energies and no further properties are to be calculated. One can then avoid the calculation of the left eigenvectors and use only the right eigenvectors for the transformation, replacing $\mathbf{L}_R$ in eq 17 with $\mathbf{L}_R^T$.

$$\mathbf{Q}_R^T (\mathbf{J} + \mathbf{A}) \mathbf{Q}_R = \mathbf{\Omega}_R \mathbf{Q}_R^T \mathbf{Q}_R$$

(19)

This ansatz comes at the cost of losing the diagonal representation and requires solving a nonsymmetric, generalized eigenvalue problem to obtain the excitation energies in the coupled equation

$$(\mathbf{\omega} \mathbf{S}_R + \mathbf{C}) \mathbf{q} = \mathbf{\Omega} \mathbf{S}_R \mathbf{q}$$

(20)

where $\mathbf{S}_R$ is the overlap of the solution vectors $\mathbf{R}$, $\mathbf{\omega}$ is a diagonal matrix with the eigenvalues of the uncoupled subsystems, and $\mathbf{C}$ has the following structure

$$\begin{pmatrix}
0 & \mathbf{C}_{\mathbf{I}\mathbf{I}}^T \\
\mathbf{C}^T_{\mathbf{II}\mathbf{I}} & 0
\end{pmatrix} = \begin{pmatrix}
0 & \mathbf{R}_{\mathbf{II}}^T \mathbf{A} \mathbf{R}_{\mathbf{II}} \\
\mathbf{R}_{\mathbf{I}}^T \mathbf{A} \mathbf{R}_{\mathbf{I}} & 0
\end{pmatrix}$$

(21)

The more complicated representation in eq 20 is advantageous compared to the simpler eq 17 in terms of computational efficiency because the matrix dimension is much smaller. The number of required excitations per subsystem can typically be limited to the lowest 10 or 20 so that the total number of excitations has a linear scaling with system size with a modest prefactor.

Three comments concerning the employed approximations are appropriate at this point: (a) The use of the right eigenvectors to carry out a transformation from the left is not an approximation in itself because the resulting metric is taken into account. As long as the right eigenvectors form a complete set, one can also express the left eigenvectors in this space. Truncating the set of eigenvectors gives an error in the coupled excitation energies, but this would also be the case if truncated left and right eigenvectors would be used to define the transformation to the smaller basis. (b) In the limit of numerically zero coupling matrices $\mathbf{A}$, the transformed coupling $\mathbf{C}$ is zero, and thus the uncoupled excitation energies $\mathbf{\omega}$ are obtained as solutions for the coupled excitation energies $\mathbf{\Omega}$ in eq 20 because the (nonorthogonal) metric $\mathbf{S}$ enters the eigenvalue problem on both sides. (c) Concerning the subspace transformation, we note that it has the advantage of being very fast but at the expense of being a “true” approximation in contrast to (a) because the coupled excitations are obtained only within the subspace spanned by the lowest subsystem excitations.

2. Approximate Coupled Excitations: Matrix Elements. The matrix elements of the off-diagonal coupling blocks $\mathbf{A}$ depend on the choice of the method used: for coupled-cluster singles (CCS), only single excitations are taken into account; for CC2, the double excitation part of the solution vector also has to be used. All of the required matrix elements are obtained applying the quasienergy approach.20

Often, the lowest excitations exhibit more than 90% HOMO–LUMO character; the doubles amplitudes contribute less than 10%, whereas for higher excitations or excitations of a different kind, the doubles part is also needed or even dominates the excitation vector. However, assuming that the excitation vectors have the dominant contribution in the singles manifold for the lowest excitations, we apply a CCS-like approximation and take into account only the following contributions

$$\mathbf{A}_{\mathbf{I}\mathbf{I}}^\mathbf{I} = \mathbf{w}_{\mathbf{I}} = 2(\mathbf{i} \mathbf{a}_1^\mathbf{I} \mathbf{b}_1^\mathbf{I}) + (\mathbf{i} \mathbf{a}_1^\mathbf{I} \mathbf{w}_{\mathbf{I}}^\mathbf{I} \mathbf{b}_1^\mathbf{I})$$

(22)

where the exact exchange contribution should not be considered as the intersubsystem interaction is defined by DFT only and the intersubsystem $\mathbf{w}_{\mathbf{I}}$ is defined as

$$\mathbf{w}_{\mathbf{I}} = \mathbf{w}_{\mathbf{I}}^\mathbf{I} = \mathbf{w}_{\mathbf{I}}^\mathbf{I}$$

(23)

Additionally, we have dropped all T1 transformations. Note that these integrals involve the integrals over two different systems, which implies evaluating integrals over orbitals with nonzero overlap. The first term is the Coulomb contribution of long-range, whereas the second term represents the kernel contributions arising from the nonadditive exchange and correlation as well as kinetic energy contributions. These terms are quite similar to the ones encountered in DFT-in-DFT, which is a consequence of the employed CCS approximation for the coupling matrix. A subtlety concerns the contribution of spin-conserving double excitations that are composed of two spin-breaking single excitations on each of the subsystems. Such excitations are included in a supermolecular CC response calculation but not in a supermolecular TDDFT calculation in the adiabatic approximation. Neglect of such intermolecular spin coupling terms is therefore an approximation (relative to a supermolecular reference) in CC-based FDEC but not in DFT-based FDEC.

Although the calculation of exact four-index two-electron integrals is also available but rather slow for large-scale applications, the integrals needed are better calculated using density fitting. Here, we expand the orbital basis set only in the fitting functions of the same subsystem

$$\sum_{\mathbf{P}^\mathbf{II}} \mathbf{C}_{\mathbf{I}\mathbf{P}^\mathbf{II}} \mathbf{C}_{\mathbf{P}^\mathbf{II}^\mathbf{I}} \approx \sum_{\mathbf{P}^\mathbf{II}} \mathbf{C}_{\mathbf{I}\mathbf{P}^\mathbf{II}} \mathbf{C}_{\mathbf{P}^\mathbf{II}^\mathbf{I}}$$

(24)

$$\mathbf{C}_{\mathbf{I}\mathbf{P}^\mathbf{II}} = \sum_{\mathbf{R}} (\mathbf{i} \mathbf{a}_1^\mathbf{I} \mathbf{b}_1^\mathbf{I}) (\mathbf{R}_1 \mathbf{e}_1^\mathbf{II}) (\mathbf{Q}_1^\mathbf{II})^{-1}$$

(25)
where \((Q_1 g_1 P_1)^{-1}\) denotes the matrix inverse. Furthermore, we introduce a dipole approximation \(^{35}\) for Coulomb integrals between widely separated transition densities, see for example ref \(^{36}\), and references therein for use thereof in modern linear scaling techniques, using a subsystem gauge

\[
(i_a j_b | h_{II}) \approx \frac{1}{R_{ij}^3} \left( \bar{\mu}_{ia} \cdot \bar{R}_{ij} - \frac{3}{R_{ij}^2} \bar{\mu}_{ia} \cdot \bar{R}_{ij} \right)
\]

(26)

where

\[
\bar{\mu}_{ia} = \langle i| \bar{R} \cdot \mu_a \rangle \text{ and } \bar{R}_{ij} = R_{ij} - \bar{R}_i
\]

(27)

(28)

(29)

(30)

In the current work, the distance threshold for this dipole approximation was chosen to be 20 Bohr. Assuming that for well-separated systems the local exchange-correlation potential is negligible, like reported in the work by Neugebauer, \(^4\) the exchange-correlation potential can be neglected, thus making the coupling step of almost zero computational cost.

### III. COMPUTATIONAL DETAILS

In the present study, supermolecular and FDE calculations and their comparison in particular are to be studied, whereas influences such as basis-set insufficiencies are not discussed because they are expected to influence both treatments in similar ways. The presented examples serve to understand whether the coupling matrix elements used lead to small or large errors in the energy splittings of the excitations. This is also reflected in the geometries, for which we have optimized the isolated monomers and put the complexes together without reoptimization. In case of the \((C_6H_6)_{25}\) complex, geometries were obtained with the Packmol package. \(^{37}\)

The presented methods have been implemented in the KOALA program. \(^{38}\) For the ground-state investigations of the formic acid dimer, Dunning’s augmented triple-ζ basis (aug-cc-pVTZ) has been used, \(^{39,40}\) whereas for all other calculations, the Turbomole def2-SVP basis \(^{41}\) has been used.

For the calculation of all quantities, we have used a supermolecular grid of type ‘4’ \(^4\) as implemented in KOALA. For all methods used, the resolution-of-the-identity approximation was used, that is, RIJK Hartree–Fock and RI-CC2. \(^{38}\) For the calculation of the ground state interaction energies, the Grimme D3 dispersion correction has been used. \(^{43,44}\) The PBE \(^4\) functional was employed for exchange-correlation and the PW91k \(^{45,46}\) functional functional for the kinetic contribution.

The coupling matrix elements in the CCS approximation have been calculated both with and without the dipole approximation in eq 26. The usage of the right eigenvectors for the transformation was numerically tested by removing the coupling contributions, yielding the independent subsystem excitation energies to all significant digits when including the overlap.

### IV. RESULTS

In this section, we will assess the accuracy and efficiency of the WFT-in-WFT method with four illustrative examples. We start with a calculation of the strongly bound formic acid dimer with geometries taken from the S22 × 5 \(^{48}\) test set as provided in the BEGDB database. \(^{49}\) To investigate the accuracy of calculated excited state splittings, we chose the prototypical ethylene dimer, which allows for an unambiguous comparison with supermolecular data. We then discuss the efficiency of the procedure by considering an agglomerate of benzene molecules that is representative for modeling excitation coupling in a solvent model. Finally, we consider a porphyrin tetramer, an example of a system for which supermolecular calculations are impractical and where DFT and WFT approaches may yield qualitatively different results.

**A. The Formic Acid Dimer.** The ground state potential energy curve for elongating the hydrogen bonds in the formic acid dimer is displayed in Figure 1. For this type of strongly hydrogen bond system, subsystem DFT is known \(^{50–52}\) to perform rather well and, as expected, the same holds true for subsystem WFT. The results agree well with the sDFT-PW91/PW91k data of ref \(^{52}\). The small discrepancy with the supermolecular results is probably partially due to a cancellation of errors; the PW91k functional often underestimates the Pauli repulsion between the monomers and thereby compensates for missing attractive dispersion interactions. This explains why adding an explicit D3 correction leads in this case to some overestimation of the binding energy. Nevertheless, the overall agreement is satisfactory for this dimer. For molecules in which the primary interaction cannot be classified as hydrogen bonding, the currently available kinetic energy functionals \(^{14}\) are not always adequate; known problematic cases are coordination bonding, \(^{53}\) heavy elements, \(^{54}\) and of course covalent bonding \(^{55}\) in which it is very difficult to work with a density-only embedding approach, and capping methods \(^{56}\) are necessary.

**B. The C\textsubscript{2}H\textsubscript{4} Dimer.** The ethylene molecule has only one π \(\rightarrow\) π \(^*\) transition, which gets split considerably when decreasing the distance between the monomers. As shown in Figure 2 and Table I, the energy of the upper excitation energy is at all distances well-reproduced by the FDEc approach, whereas the lower state shows significant errors relative to the supermolecular treatment below 5 Å. Similar effects have been observed before, also at approximately 5 Å, for the benzaldehyde dimer \(^{17}\) treated with the DFT-in-DFT approach. These errors can be due to either the lack of explicit exchange terms in the interaction energy or the admixture of charge-transfer excitations, which are missing in the FDEc treatment.

---

**Figure 1.** Binding energy (kJ/mol) of the HCOOH dimer as a function of the hydrogen bond stretch (geometries taken from ref \(^{48}\)).
Figure 2. Splitting of the C2H4 dimer lowest excitation (B2g/B3u) calculated with RICC2/SVP in eV. At a short distance of 4.0 Å, the FDE approximation gives large errors compared to the supermolecular calculation due to mixing of charge-transfer character in the lowest excitation.

Table I. Excitation Energies of C2H4 Dimer at 5.0 Å Calculated with RICC2-in-RICC2 FDE and RICC2 in eV (see Fig. 2)\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>monomer</th>
<th>FDEu</th>
<th>FDEc</th>
<th>super</th>
</tr>
</thead>
<tbody>
<tr>
<td>E\textsubscript{1}</td>
<td>8.82</td>
<td>8.82</td>
<td>8.76</td>
<td>8.72</td>
</tr>
<tr>
<td>Δ</td>
<td>0</td>
<td>+0.00</td>
<td>+0.12</td>
<td>+0.18</td>
</tr>
<tr>
<td>E\textsubscript{2}</td>
<td>8.99</td>
<td>9.00</td>
<td>9.00</td>
<td>8.99</td>
</tr>
<tr>
<td>Δ</td>
<td>0</td>
<td>+0.00</td>
<td>+0.00</td>
<td>+0.01</td>
</tr>
</tbody>
</table>

\textsuperscript{a}FDEu denotes uncoupled excitation energies, FDEc denotes coupled subsystems, Δ denotes the splitting due to the coupling, E\textsubscript{1} is the lowest excitation energy (B2g/B3u), and E\textsubscript{2} is the second lowest (B3g/B2u).

For the simple ethylene system, it is possible to interpret the results using a qualitative model based on the π HOMOs and π* LUMOs of the subsystems. As is discussed by Hieringer and Görling in ref 57, these give, at long distance, rise to an (arbitrary)\textsuperscript{58} linear combination of local π → π* transitions and, at significantly higher energy, linear combinations of pure charge-transfer states. As long as these two types of states do not mix, FDEc does give correct results for the splitting. At 6 Å analysis of the supermolecular reference, CC2 calculations indicate that both the upper and lower states are almost pure combinations of local excitations, whereas at 4 Å, the contribution of charge transfer to the lower state rises above 10%. This explains the worse performance of the FDEc method for this state, whereas it remains still qualitatively valid for the upper state in which the charge-transfer character is modest at 1.5%.

C. Agglomerations: 25 Benzene Molecules. We next consider the efficiency of the procedure for coupling the lowest excitations of a large number of solvent molecules. As a test case, we chose benzene for which coupling of excited states should be reasonably well-described with the FDEc approach. Here, we investigate the ten lowest excited states for 25 molecules, respectively, giving a total of 250 coupled excitations. As we have not yet implemented the calculation of oscillator strengths for the coupled excitations, we will display in Figure 3 the spread of these excitations using an arbitrary height in the stick diagram for the FDEc excitations. The excitation energies of a single benzene molecule are also displayed for ease of reference, showing the origin of the broadened peaks in the compound system. States that carry no intensity are hardly broadened, whereas the most intense peak at 7.7 eV is broadened by approximately 0.4 eV. This is mainly due to the coupling, the peaks in FDEu are also spread due to the different environment for each benzene molecule, but this yields a much smaller broadening. Given the size of the system, it is not possible to compare with reference supermolecule data, but larger calculations in the subsystem approach are easily feasible. This calculation at coupled RICC2-in-RICC2 FDE level of theory takes approximately 15 h on 12 CPUs, where all input/output steps and other overhead of the embedding treatment is included. The code is, however, not yet optimized and speed-ups due to integral prescreening and use of local (closest neighbors only) grids instead of the currently employed supermolecular grid are certainly possible. The final step of the calculation, in which the coupled excitation energies of the supermolecular complex are determined, takes only approximately 2 min, showing the efficiency of the WFT-in-WFT approach.

D. Porphyrin. Because explicit calculation of interactions between a number of biomolecules is almost infeasible for WFT methods and DFT approaches cannot always reach the desired accuracy, WFT-in-WFT approaches can provide an interesting alternative. As a simple and well-understood model for such systems, we chose magnesium porphyrin (MgP), the computationally simplest of the metallo-porphyrins, as a representative example of a biomolecule with low-lying excitations. We will consider this molecule both with the above-described WFT-in-WFT approach as well as with the related DFT-in-DFT approach developed by Neugebauer.\textsuperscript{17} As a model system, we chose a stack of 4 MgP that are placed at exactly 4 Å from each other in a fully parallel (sandwich) conformation, as displayed in Figure 4.

Since the early work of Gouterman,\textsuperscript{59} who introduced the famous 4-orbital model, there has been a lot of work on both
freebase porphyrin, and the system nowadays serves as a benchmark molecule for excited state calculations. As discussed in great detail by Baerends in ref 60, the spectrum can be qualitatively well-understood using TD-DFT, but care needs to be taken to avoid the spuriously low charge-transfer states that may appear in TD-DFT. More accurate results are possible with WFT techniques that are now well-applicable for one porphyrin molecule, and recent benchmarks are available.61,62

For the related MgP, multireference perturbation calculations were reported in 1999 by both the Hirao group63 and the Roos group64 and give qualitatively the same picture: two $^1E_{1u}$ transitions in which an electron is excited from the $1a_{1u}$ HOMO and the $4a_{2u}$ HOMO-1 to the degenerate $e_{1g}$ LUMOs. In both calculations, the calculated values are slightly below the measured absorption maximum for substituted MgPs, but the intensity pattern agrees quite well with the first excitation being identified as the low-intensity Q-band and the next excitation the more intense B-band. TD-DFT with the SAOP functional gets the position of these bands also quite well,60 but one has to be careful with CT states that may appear in the same energy range.

In our PBE TD-DFT calculation, we see one additional $^1E_{1u}$ state below the intense B-band as well as a number of symmetry-forbidden transitions. The CC2 calculations on the monomer give a weak $^1E_{1u}$ transition at 2.50 eV followed by an intense $^1E_{1u}$ transition at 3.60 eV. In comparison with CASPT2, we see that the agreement with experimental data is comparable, albeit that the CC2 values are higher than the band maxima, whereas the CASPT2 values were lower. In comparison with TD-DFT, the intensity of the B-band is considerably higher with, in the DFT case, some of the intensity being contained in the fourth $^1E_{1u}$ transition at 3.60 eV. This difference in intensity has marked consequences for splitting of the coupled excitations.

As intensities are not yet available for the WFT-in-WFT calculation, we first display the DFT-in-DFT data in Figure 5. From this figure, we see that the coupling leads to a splitting of the excitations with a collection of almost all intensities in the highest excitation. This is governed by the high symmetry of our model system in which the other coupled excitations are symmetry forbidden. For the WFT-in-WFT couplings, we present the results in Table II, which contains both the results

Figure 4. Stack of 4 porphyrin molecules as calculated with coupled RICC2-in-RICC2 FDE (FDEc).

Figure 5. DFT-in-DFT calculations of the porphyrin tetramer. Monomer (full black), uncoupled dimer (dashed black), coupled dimer (purple), and coupled tetramer (blue) are shown with a Gaussian broadening. The monomer and dimer intensities were scaled by a factor of 4 and 2, respectively, to match the tetramer intensities.
from the full calculations in which 20 excitations per monomer are coupled as well as calculations in which only the Q- and B-band excitations are considered. From this data, we see that the larger intensity calculated by the WFT approach leads to a much larger splitting in the coupled excitations. Similar effects have been reported by König and Neugebauer when changing the amount of exact exchange in hybrid TDDFT calculations increased the magnitude of the transition densities. Whereas the DFT splitting of the B and Q bands leads to peaks in the interval 3.00 to 3.85 eV, the larger intensities of the WFT coupling only B and Q bands yields similar splittings compared to coupling of all lowest 20 subsystem excitations. The Q bands show the largest oscillator strengths of approximately 1.16 for the isolated monomer at RICC2 level of theory. All excitation energies are given in eV.

V. SUMMARY AND CONCLUSIONS

We have reported an implementation of the ground-state energy as well as coupled excitations at the level of RICC2-in-RICC2 FDE (denoted FDEc) using the efficient CCS approximation for the equations and the dipole approximation for the Coulomb matrix elements. This yields an efficient hybrid WFT/DFT method that can be applied easily to molecular agglomerations. Further computational gains can even be achieved with further approximations, such as the use of local grids, whereas in the present work, supramolecular grids have been used.

For excitation energies and splittings, the coupled FDE method yields sufficiently accurate splittings and excitations compared to the supermolecular calculation if charge-transfer excitations from one monomer to another are unimportant. In such cases, the use of a WFT treatment of the subsystem can be used to ensure sufficient accuracy in the prediction of the correct order of excited states and, importantly, also their relative intensities.

When charge-transfer between subsystems becomes important, such as for ethylene molecules separated by less than 5.0 Å, significant errors may arise, and the challenge of developing a general treatment that includes such states remains. For the special case of interaction between the ground state and one CT excited state, suitable FDE methods are available and can provide further insight into the best way to describe such excitations without sacrificing the computational simplicity of the frozen density ansatz.

AUTHOR INFORMATION

Corresponding Author
E-mail: lvisscher@vu.nl.

Present Address
Institut für Physikalische Chemie, Karlsruher Institut für Technologie, D-76131 Karlsruhe, Germany

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

S.H. has been supported by the European Commission under a Marie-Curie Intra-European Fellowship (Contract No. PIEF-GA-2010-274224).

REFERENCES
