JCTC Journal of Chemical Theory and Computation

Article

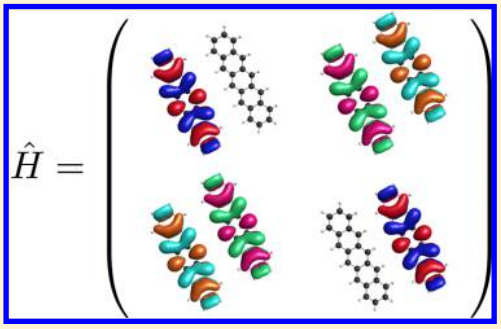
pubs.acs.org/JCTC

Ab Initio Implementation of the Frenkel–Davydov Exciton Model: A Naturally Parallelizable Approach to Computing Collective Excitations in Crystals and Aggregates

Adrian F. Morrison, Zhi-Qiang You, and John M. Herbert*

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

ABSTRACT: A fragment-based method for computing vertical excitation energies of molecular clusters is introduced based on an ab initio implementation of a Frenkel–Davydov exciton model consisting of singly excited monomer basis states. Our strategy is to construct and diagonalize the exact Hartree–Fock Hamiltonian in such a basis. Matrix elements between nonorthogonal determinants are computed via the corresponding orbital transformation and the resulting generalized eigenvalue problem is solved to determine collective excitation energies and wave functions. The basis may be expanded to include higher-lying fragment excited states in order to account for interfragment polarization effects. Absolute errors of $\lesssim$0.1 eV (relative to supersystem methods) are achievable for systems such as water clusters and crystalline arrays of organic chromophores such as pentacene and napthalenediimide. Preliminary tests for a nine-chromophore subunit of an organic nanotube suggest that it is possible to target the optically bright state, even when it is a high-lying excitation, by using carefully selected basis states. The highly parallel nature of this method provides a foundation for further developments to treat collective excitations in large molecular assemblies.



I. BACKGROUND

Quantum chemical calculations of excited-state properties have played an important role in numerous fields of modern chemical research, such as solar energy conversion,[1−3] nanomaterials,[4,5] and more. However, the potential of quantum chemistry as a tool that can dramatically benefit the research and development of novel materials is only beginning to be realized. The challenge for excited-state methods, as with much of quantum chemistry, is the highly nonlinear scaling of the computational cost; even the cheapest excited-state methods formally scale as $O(N^4)$ with system size.[6] Assemblies of electronically coupled chromophores (such as molecular crystals, or the light harvesting complex in the photosynthetic reaction center) pose special problems in this respect, insofar as the excited states may be delocalized over multiple chromophores. Such cases require especially large quantum-chemical model systems.

On the computational side, it has been the case for some time that single-threaded CPU performance has essentially reached an asymptotic limit. Today, Moore's Law is realized by improvements in concurrent multithreaded performance by the continued addition of processor cores to computer systems. Modern supercomputers include of tens of thousands of CPU cores, and even an average workstation may have dozens, and these numbers are only increasing. To continue to push the boundaries of quantum chemistry research, algorithms must be designed to scale efficiently across these massively parallel, peta-scale architectures. The simplest way to do so is to exploit the "embarrassing parallelizability" of an algorithm whose effort can be subdivided into a large number of completely independent processes.

Here, we introduce a novel method for computing excitation energies for systems of weakly interacting fragments such as liquids or molecular aggregates. The method is based on the molecular exciton model that was first introduced by Frenkel in 1931,[7] who described the excited states of solids as superpositions of excitation waves. This idea was further expanded by Davydov in 1964,[8] who wrote the Hamiltonian for a molecular crystal as

$$\hat{H} = \sum_n \hat{H}_n + \frac{1}{2}\sum_{m<n} \hat{V}_{mn} \tag{1}$$

where $\hat{H}_n$ is the Hamiltonian operator for the molecule located at lattice site n, and $\hat{V}_{mn}$ is the coupling between sites. Energy is variationally minimized and solutions to the requisite secular equation have the form

$$\Psi_I = \sum_n \sum_i k_n^i \phi_n^i \prod_{m\neq n} \phi_m \tag{2}$$

Here, Ψ_I is the Ith excited state of the supersystem and ϕ_n^i is the ith singly-excited eigenstate of the nth monomer unit, whereas ϕ_m is the wave function for monomer m in its ground state.

It is important to make clear the physics that distinguishes excitonic behavior, in no small part because "exciton" is a confusing term in the literature. We take "exciton" to mean an

Received: August 24, 2014
Published: November 13, 2014

   dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366–5376

excited state of a collection of chromophores that is delocalized across one or more of them but which can nevertheless be represented as a linear combination of excited states that are spatially localized on particular subunits. This representation is the fundamental *ansatz* that we make in putting forth an "exciton model". Note that cases where the excitation *is* actually localized on one subunit emerge naturally as a special case. Molecular (as opposed to solid-state) exciton theory has been applied in the past to explain the spectroscopic behavior of oligomeric systems of interest in biology and photosynthesis[9] and, more recently, in characterizing the aggregation of light harvesting complexes.[10]

Historically, the coupling matrix elements between direct-product basis states were computed as Coulomb integrals over transition densities,[11] or even more simply within a dipole approximation. Some improvements beyond this have been made, as summarized in ref 12, but electronic overlap and exchange effects are still typically neglected. More recently, new approaches have been developed that construct and diagonalize an effective Hamiltonian projected onto an excitonic basis,[13,14] with the aim of reducing computational cost as compared to construction and diagonalization of the full Hamiltonian for the aggregate.

As an alternative, we present a fully ab initio implementation of the original Frenkel−Davydov exciton model. The full Hartree−Fock Hamiltonian is diagonalized in a basis of singly excited fragments, as in eq 2, so that the Coulomb and exchange interactions from Hartree−Fock theory are treated exactly, as is the nontrivial overlap of the fragment wave functions. The basis states are constructed by forming direct products of configuration state functions computed from N_{frgm} independent fragment self-consistent field (SCF) calculations and results in a basis whose dimension is at least $N_{\mathrm{frgm}} + 1$, including the direct-product ground state. Fragment states used to construct the basis are adiabatic in the sense that they are determined in the absence of interfragment interactions. The basis space can be expanded to include higher-lying fragment excitations, which increases the flexibility of the direct-product *ansatz*, or to include charge-transfer basis states to explore the mixing of neutral and charge-transfer excitons. To the best of our knowledge, our approach is the first to include exact Hartree−Fock exchange interactions in the excitonic coupling. These prove to be crucial for accuracy at short intermolecular distances.

In its current implementation, the aggregate CPU time required for our method scales more steeply, as a function of system size, than does the CPU time required for traditional supersystem methods such as configuration interaction singles (CIS). However, the overwhelming computational bottleneck in our approach is calculation of individual matrix elements of the exciton Hamiltonian, and these concurrent tasks are embarrassingly parallelizable. As such, the required wall time should be reduced in nearly direct proportion to the number of available CPU cores. We provide examples of several exciton calculations for relevant systems of interest where this combination of parallelism and a priori simplification of the problem based on chemical intuition results in significantly reduced wall times and memory requirements, relative to a traditional supersystem calculation, while maintaining accuracy of $\lesssim$0.1 eV with respect to supersystem CIS results.

II. THEORY

A. Direct-Product Configuration State Function Basis.

The strategy of our ab initio fragment exciton approach is to construct and diagonalize the Hamiltonian in an excitonic basis made up of direct products of fragment configuration state functions (CSFs). The fragment CSFs are computed from ground- and excited-state SCF calculations for the independent fragments. CSFs on different fragments are generally non-orthogonal and may be comprised of multiple determinants. The notation we use is as follows: for fragment M, Ψ_M indicates a CSF, Φ_M indicates a Slater determinant, ϕ_p^M is a molecular orbital (MO), and n_M indicates the total number of occupied orbitals, including both α and β spin states. The total number of occupied orbitals in the system of N_{frgm} fragments is denoted by N.

Consider a system of two chromophores, A and B, and a direct product state $|\Psi_A^*\Psi_B\rangle$ involving excitation of fragment A. The ground-state CSF for fragment B,

$$|\Psi_B\rangle = |\Phi_B\rangle = \frac{1}{\sqrt{n_B}}\left|\phi_1^B\phi_2^B\cdots\phi_{n_B}^B\right| \tag{3}$$

is a single Slater determinant constructed from a set of SCF MOs. The spin-adapted CSF for the excited fragment A,

$$|\Psi_A^*\rangle = \sum_{ia}\sum_{\sigma=\alpha,\beta} C_\sigma^{ia}|\Phi_A^{ia}\rangle \tag{4}$$

is a linear combination of singly substituted determinants $|\Phi_A^{ia}\rangle$ weighted by CI coefficients C_σ^{ia} that are computed for the spin-restricted singlet or triplet state of each isolated fragment. We use restricted SCF calculations for the fragments, so the spatial orbitals and CI coefficients are identical, up to a possible sign change, for α and β spin, but we must treat both spin components explicitly in order to account for spin coupling between states. The overall excitonic basis state can then be written

$$|\Psi_A^*\Psi_B\rangle = \frac{1}{\sqrt{N}}\left|\sum_{ia\sigma} C_\sigma^{ia}\phi_a^A\Big(\prod_{j\neq i}^{n_A-1}\phi_j^A\Big)\phi_1^B\phi_2^B\cdots\phi_{n_B}^B\right| \tag{5}$$

For clarity, the notation in eq 5 only includes two fragments, A and B. More generally, basis states $|\Psi_A^*\Psi_B\Psi_C\cdots\rangle$ will simply append ground-state (MOs) $\phi_1^C\phi_2^C\cdots\phi_{n_C}^C$, etc., to the Slater determinant in eq 5. In this way, every basis state includes all of the occupied MOs for the entire system. This stands in contrast to traditional exciton theory that considers only pairwise Coulomb interactions of the excited sites and neglects the overlap of the remaining fragments. In order to include nonpairwise additivity in the exchange interactions, and therefore enforce antisymmetry in the excitonic wave functions, all MOs *must* be included. However, the fragment MOs computed from independent SCF calculations are not orthogonal therefore the excitonic basis states are not orthogonal. Therefore, overlap matrices must be computed, in addition to matrix elements of the Hamiltonian, a generalized eigenvalue problem must be solved. These matrix elements between CSFs are

$$\langle\Psi_A^*\Psi_B|\Psi_A\Psi_B^*\rangle = \sum_{ia\sigma}\sum_{kb\tau} C_\sigma^{ia}C_\tau^{kb}\langle\Phi_A^{ia}\Phi_B|\Phi_A\Phi_B^{kb}\rangle \tag{6a}$$

$$\langle\Psi_A^*\Psi_B|\hat{H}|\Psi_A\Psi_B^*\rangle = \sum_{ia\sigma}\sum_{kb\tau} C_\sigma^{ia}C_\tau^{kb}\langle\Phi_A^{ia}\Phi_B|\hat{H}|\Phi_A\Phi_B^{kb}\rangle \tag{6b}$$

dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366−5376

In our spin-restricted implementation there are only two distinct spin multiplicity terms for each pair of transitions. The sign of the coupling term is derived from the spin of the fragment: positive for singlet and negative for triplet. This, in turn, determines the spin of target state.

B. Natural Transition Orbital Representation. In order to minimize the number of terms in eq 6, we transform the fragment excited states into the natural transition orbital (NTO) basis.[15−17] The $n_M^{\mathrm{occ}} \times n_M^{\mathrm{virt}}$ single-particle transition density matrix, $\mathbf{T}$, couples occupied and virtual orbitals with coefficients C^{ia} from a CIS or time-dependent density functional theory (TDDFT) calculation, the latter performed within the Tamm−Dancoff approximation (TDA).[18] NTOs are computed from a singular value decomposition of $\mathbf{T}$:[16]

$$\mathbf{\Lambda} = \mathbf{OTN}^{\dagger} \tag{7}$$

Here, $\mathbf{O}$ and $\mathbf{N}$ are separate unitary transformations of the canonical occupied and virtual MOs, respectively, which transform these orbitals into a set of (state-specific) paired hole and particle NTOs. This transformation reduces the dimension of the CI expansion to no more than n_M^{occ} particle-hole excitations and the diagonal matrix $\mathbf{\Lambda}$ contains the coefficients of these excitations in the NTO basis. This results in no more than n_M terms for a single excited state and no more than $n_A n_B$ terms in eq 6. Similarly, only $n_A n_B$ appear in the more general expression for $\langle \Psi_A^* \Psi_B \Psi_C \cdots | \hat{H} | \Psi_A \Psi_B^* \Psi_C \cdots \rangle$.

In practice, the number of significant NTOs needed for the CSF expansions depends on the size and complexity of the fragment; for small molecules, we typically find that only one or two Λ_i are significant. The NTO expansions may therefore be truncated at a specified fraction of the excitation amplitude (norm of $\mathbf{T}$) in order to reduce the length of the summations in eq 6.

C. Corresponding Orbital Transformation. To compute matrix elements between nonorthogonal Slater determinants, we turn to the corresponding orbital transformation of Amos and Hall.[19] For a given term in eq 6, the two sets of spin orbitals associated with the bra and ket will be denoted $\mathbf{l}$ and $\mathbf{r}$, respectively. The sets are Schmidt-orthogonalized among themselves and expanded in a common atomic orbital (AO) basis, $\{\chi\}$:

$$\begin{aligned} \mathbf{l} &= \chi \mathbf{L} \\ \mathbf{r} &= \chi \mathbf{R} \end{aligned} \tag{8}$$

Let $\mathbf{S}^{\mathrm{LR}} = \mathbf{L}^{\dagger}\mathbf{SR}$ denote the (nondiagonal) overlap matrix between the left ($\mathbf{L}$) and right ($\mathbf{R}$) sets of orbitals, where $\mathbf{S}$ denotes the AO overlap matrix. We can then apply left and right unitary transformations that diagonalize $\mathbf{S}^{\mathrm{LR}}$ but leave the original Slater determinants unchanged, except possibly for a phase that is equal to the determinant of $\mathbf{U}^{\dagger}$ or $\mathbf{V}^{\dagger}$:

$$\mathbf{U}^{\dagger}\mathbf{L}^{\dagger}\mathbf{SRV} = \tilde{\mathbf{L}}^{\dagger}\mathbf{S}\tilde{\mathbf{R}} = \tilde{\mathbf{s}}^{\mathrm{LR}} \tag{9}$$

Here, $\tilde{\mathbf{s}}^{\mathrm{LR}}$ is diagonal and the matrices $\mathbf{U}$ and $\mathbf{V}$ are computed from the singular value decomposition of $\mathbf{S}^{\mathrm{LR}}$.

The corresponding orbitals transformation leads to a set of generalized Slater−Condon rules[20] that may be used to compute the Hamiltonian and overlap matrix elements in terms of the AO basis. The overlap matrix element is

$$\Xi^{\mathrm{LR}} = \langle \Phi_A^{ia} \Phi_B | \Phi_A \Phi_B^{kb} \rangle = \det(\mathbf{U}) \det(\mathbf{V}^{\dagger}) \prod_i^N \tilde{s}_{ii}^{\mathrm{LR}} \tag{10}$$

Note that the transformations $\mathbf{L}$ and $\mathbf{R}$ depend, implicitly, on the MOs ϕ_i^A, ϕ_a^A, ϕ_k^B, and ϕ_b^B that are involved in the excitation, but for brevity the indices in question are subsumed into the "LR" in Ξ^{LR}. The corresponding matrix element of $\hat{H}$ can be written

$$\langle \Phi_A^{ia} \Phi_B | \hat{H} | \Phi_A \Phi_B^{kb} \rangle = \Omega_1^{\mathrm{LR}} + \Omega_2^{\mathrm{LR}} \tag{11}$$

where

$$\Omega_1^{\mathrm{LR}} = \sum_i^N \sum_{\lambda\rho} \tilde{R}_{\rho i} \left(\frac{\Xi^{\mathrm{LR}}}{\tilde{s}_{ii}^{\mathrm{LR}}} \right) \tilde{L}_{i\lambda}^{\dagger} \langle \lambda | \hat{h} | \rho \rangle \tag{12}$$

and

$$\Omega_2^{\mathrm{LR}} = \frac{1}{2} \sum_{ij} \sum_{\lambda\rho\mu\nu} \tilde{R}_{\rho i} \tilde{R}_{\nu j} \left(\frac{\Xi^{\mathrm{LR}}}{\tilde{s}_{ii}^{\mathrm{LR}} \tilde{s}_{jj}^{\mathrm{LR}}} \right) \tilde{L}_{i\lambda}^{\dagger} \tilde{L}_{j\mu}^{\dagger} \langle \lambda\mu || \rho\nu \rangle \tag{13}$$

The interaction terms for α and β electrons are computed explicitly, so

$$\Omega_1^{\mathrm{LR}} = {}^{\alpha}\Omega_1^{\mathrm{LR}} + {}^{\beta}\Omega_1^{\mathrm{LR}} \tag{14}$$

and

$$\Omega_2^{\mathrm{LR}} = {}^{\alpha\alpha}\Omega_2^{\mathrm{LR}} + {}^{\beta\beta}\Omega_2^{\mathrm{LR}} + {}^{\alpha\beta}\Omega_2^{\mathrm{LR}} + {}^{\beta\alpha}\Omega_2^{\mathrm{LR}} \tag{15}$$

where the spin coupling terms in eq 15 cause the antisymmetrized AO integrals, $\langle \lambda\mu || \rho\nu \rangle$ in eq 13, to be replaced with with Coulomb integrals $\langle \lambda\mu | \rho\nu \rangle$. This fundamentally comes from the vanishing of the exchange contributions to the generalized density matrices in the $\alpha\beta$ case, reflecting the lack of exchange interaction between α and β densities.

III. ACCURACY

A preliminary version of our ab initio fragment molecular exciton method has been implemented in a developer's version of the Q-Chem electronic structure program.[21,22] Unless otherwise noted, all fragment and supersystem calculations were carried out at the CIS level of theory, and "error" in excitation energies obtained for the exciton models is defined with respect to a supersystem CIS calculation using the same basis set. (Since our exciton model is based on the Hartree−Fock Hamiltonian, this is the most appropriate comparison.) A few calculations use monomer basis states computed from density functional theory (DFT), but these results will nevertheless be compared to supersystem CIS calculations because matrix elements of the exciton Hamiltonian are evaluated at the Hartree−Fock level. In this sense, DFT serves only as an alternative means to obtain basis states but does not fundamentally change the exciton model. Unless otherwise noted, the 6-31G basis set was used for all calculations except those for He_n, for which the 6-311G basis set was used.

A. Linear Helium Chains. Linear chains of He atoms (with each atom separated by the He_2 equilibrium distance of 1.581978 Å) are a potentially difficult test system as their excited states tend to be fully delocalized across the entire chain. Figure 1 plots errors in excitation energies for He_n chains (n = 2−30) computed using our exciton model. Results are shown for a variety of exciton bases.

There are two primary approximations in our model. First, the isolated fragment excited states are computed in the absence of interactions with the surrounding fragments. The severity of this approximation can be reduced by including higher-lying fragment excited states in the direct-product basis,

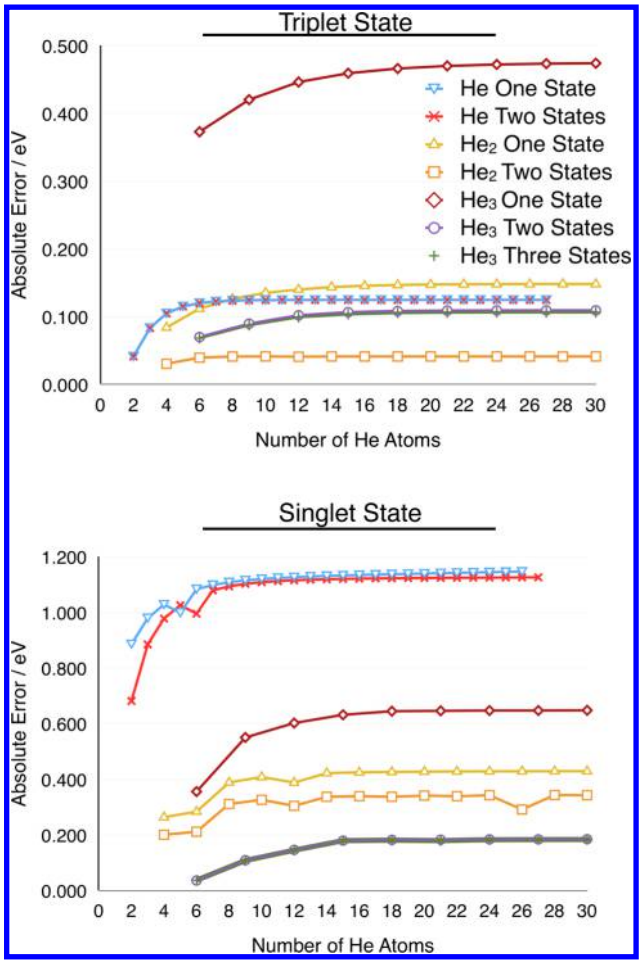


Figure 1. Absolute errors (relative to a supersystem CIS/6-311G calculation) in the excitation energy predicted by the exciton model for the lowest triplet and singlet states of He_n chains. Results are shown for several versions of the exciton model, using fragments ranging in size from He to He_3, in conjunction with an excitonic basis consisting of 1−3 excited states per fragment. CIS/6-311G calculations are used also for the fragments. The number of NTOs retained per fragment is equal to the number of He atoms per fragment, as this is sufficient to recover essentially the full norm of the transition density matrix.

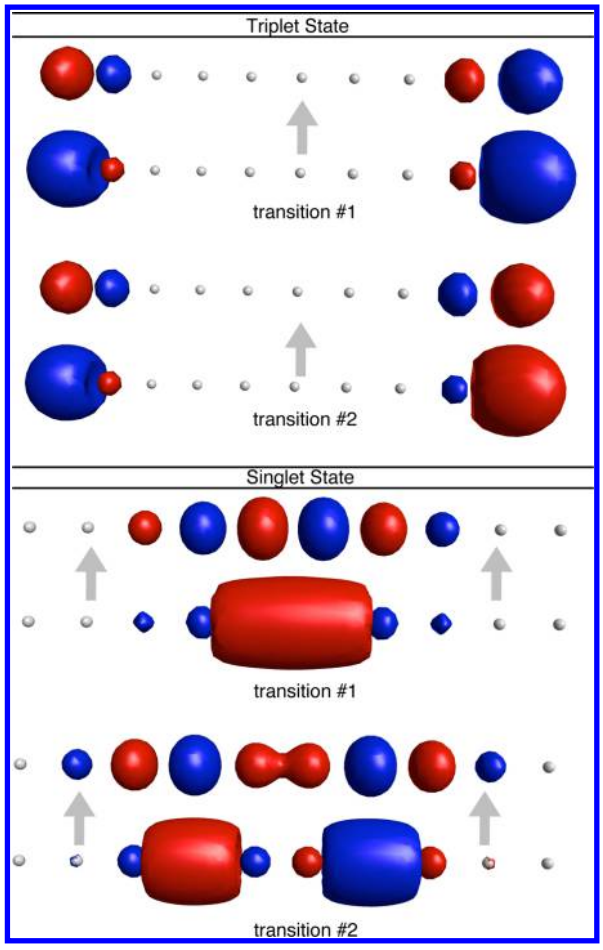


Figure 2. Plots of the two dominant NTOs for the lowest singlet and triplet excitations of a He_{10} chain. (An isosurface value of 0.05 au is used in each case.) Each pair of particle/hole NTOs for the triplet state accounts for 38.5% of the norm of the transition density matrix, while in the singlet case the two NTO pairs account for 60.3% (upper pair) and 24.1% (lower pair) of the norm.

which add variational flexibility and allow the excitonic wave functions to deform in response to interfragment perturbations. For the weakly polarizable He_n test systems, however, this approximation is not a significant source of error.

The second approximation is fundamental to the exciton model itself: we assume that each supersystem excitation can be described as a linear combination of excitations localized on individual subunits. Helium chains put this assumption to the test, because their excited states qualitatively resemble those of a particle in a one-dimensional box. Our data show that even such highly delocalized states are well described by the model, so long as the fragment size is sufficient to capture the (relatively localized) charge transfer character of the excited state. For triplet and singlet excited states, fragments made up of two and three He atoms, respectively, are adequate to achieve reasonable accuracy (Figure 1). That the description of singlets requires additional fragment states, relative to the triplet case, makes sense in view of the fact that Pauli repulsion tends to delocalize singlet excitations to a greater extent than triplet excitations. Analysis of the supersystem NTOs for He_{10} (Figure 2) shows that the overall excitation is well represented as a combination of relatively localized excitations, the sizes of which correspond roughly to the two and three He atoms per fragment that afford good results in the triplet and singlet exciton models, respectively.

B. Water Clusters. The collective nature (or lack thereof) of the excited states of water has been a subject of some debate.[23] We have used our method to compute the excitation energies of water clusters at their equilibrium geometries and the results are given in Figure 3. (Geometries are MP2/cc-pVDZ, from ref 24.) Relative to the He_n results, the excited states of $(H_2O)_n$ are *not* well represented by adiabatic fragment states, as reflected in the large errors reported in Figure 3. By including three excitations per H_2O fragment, however, we obtain excellent agreement with supersystem calculations, with errors that are generally less than 0.1 eV. We find that these results are robust with respect to the choice of AO basis set as well, and similar results are obtained in the 6-31+G* basis.

We select an $(H_2O)_7$ cluster from our test set as an illustrative example. As shown by the NTOs in Figure 4, the excitation primarily involves transitions localized on monomers 1 and 3 with minimal intermolecular excitation transfer. Comparing this picture to the coefficients of the exciton eigenvector computed with our method, as provided in Table 1, we find that the excitonic wave function is dominated by basis

dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366−5376

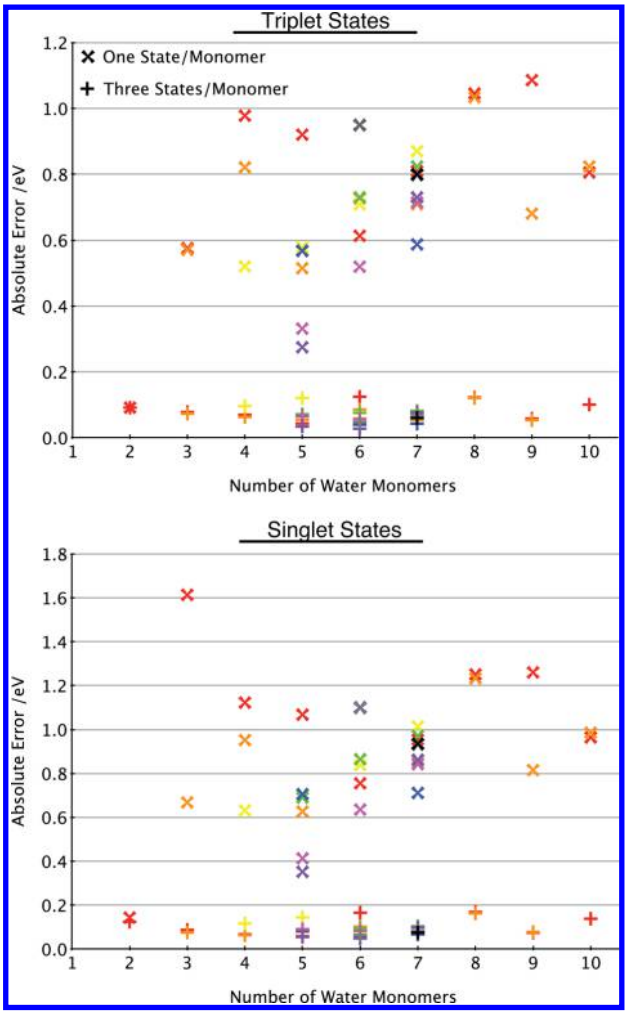


Figure 3. Absolute errors for the lowest triplet and singlet excitation energies for various isomers of water clusters, relative to supersystem CIS/6-31G excitation energies. The various colors refer to different cluster isomers.

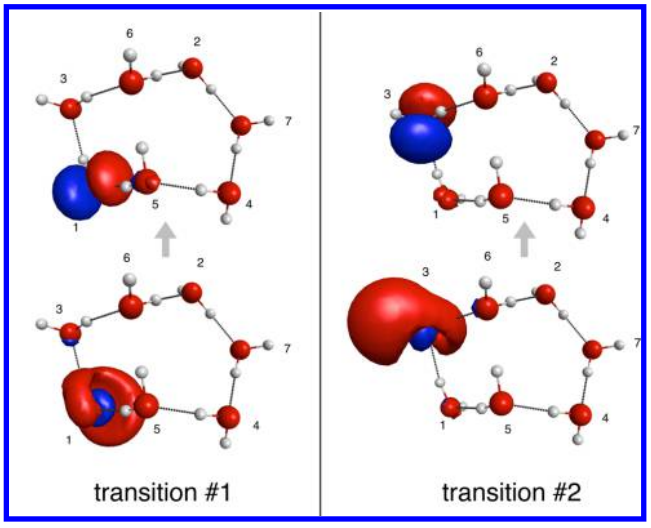


Figure 4. Plots of the two dominant NTOs for the first triplet excitation of an $(H_2O)_7$ cluster. The NTO particle/hole pairs account for 76% and 14% of the overall transition, respectively. (The isosurface value for all plots is 0.05 au.) Numbers indicate the monomer index of the adjacent molecule.

states involving excitation of precisely these two water monomers. This behavior is typical and shows how the exciton approximation still captures the fundamentals of the supersystem excited state for realistic molecular clusters. Note that the relative magnitude of the dominant states are reversed when compared to the corresponding localized NTO amplitudes. This is due to neglect of intermolecular polarization, and largely disappears when higher-lying monomer excited states are included in the excitonic basis.

C. Alternative Basis States. An exciton approach is essentially a variational energy minimization in a strategically chosen trial basis. Accordingly, determination of the fragment orbitals for the basis states need not be limited to a particular quantum chemistry method, so long as they accurately represent the excited state(s) of the fragment. Kohn−Sham orbitals and TDA-TDDFT coefficients C^{ia} offer a basis that includes some intrafragment electron correlation. It is also well-known that Kohn−Sham orbitals are more representative of electronic excitations as compared to Hartree−Fock orbitals,[25] which bear more resemblance to ionized states. Indeed, we find that NTO expansions are generally more compact for TDDFT excited states as compared to their CIS counterparts, which proves useful in reducing computational time when the monomer units are large (see Section III.D). Here, we use the B3LYP functional to compute monomer basis states for water clusters. The results, shown in Table 2, demonstrate that the B3LYP-based approach is slightly more accurate for water clusters, as compared to the same exciton model constructed from CIS monomer wave functions. (This conclusion is not true for all of the systems that we have explored, however; see below.)

For systems comprised of polar monomers, another way to augment the adiabatic approximation is to compute monomer wave functions in the presence of some classical representation of the electrostatic environment of the supersystem, for example, by embedding the monomer calculations in a field of atom-centered point charges on the other monomers. The variational "XPol" (explicit polarization) approach of Xie et al.[26] uses a self-consistent charge embedding procedure to accomplish this, and here, we use XPol in conjunction with "ChElPG" charges that are fit to reproduce the molecular electrostatic potential outside of the van der Waals region. (See refs 27 and 28 for details of the combined XPol + ChElPG algorithm.) The polarized MOs generated by the XPol procedure are then used to perform CIS calculations on the fragments.

Results in Table 2 show that the use of XPol wave functions significantly enhances the accuracy of the exciton model for water clusters; in particular, use of a single XPol CIS state per monomer is as accurate as three adiabatic (gas-phase) states per monomer. This combination presents a highly appealing route for economical excited-state calculations in clusters, providing accuracy within 0.2 eV of supersystem calculations with only a minimal basis.

D. Results for Large Systems. We next consider whether comparable accuracy is maintained in larger systems. Table 3 shows errors in the exciton model for several snapshots of a $(H_2O)_{57}$ cluster and a $(H_2O)_{117}$ cluster that were extracted from simulations of liquid water. The errors for these systems tend to be less than similar calculations in Table 2 including one excitation per fragment. The water molecules from these simulations are generally further apart than those in equilibrium (i.e., cluster) geometries with fewer hydrogen bonds between

dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366−5376

Table 1. Amplitudes for the First Triplet Excitation of the $(H_2O)_7$ Cluster Shown in Figure 4

	monomer						
	1	2	3	4	5	6	7
	one state per monomer						
	0.581	−0.000	0.813	−0.001	−0.009	−0.009	0.000
	three states per monomer						
state 1	0.832	−0.000	0.078	−0.001	−0.010	−0.000	0.000
state 2	0.034	0.000	−0.000	−0.000	0.003	0.000	−0.000
state 3	0.421	−0.000	−0.353	−0.000	0.020	−0.000	0.000

Table 2. Mean Unsigned Errors[a] (MUEs, in eV) for $S_0 \rightarrow S_1$ and $S_0 \rightarrow T_1$ Excitation Energies for Small Water Clusters Computed Using Various Exciton Models

		no. of water molecules								
basis	state	2	3	4	5	6	7	8	9	10
HF/6-31G (1 state)	T_1	0.09	0.57	0.77	0.54	0.77	0.65	1.06	0.88	0.82
	S_1	0.14	1.14	0.90	0.65	0.91	0.79	1.25	1.04	0.98
HF/6-31G (2 states)	T_1	0.09	0.57	0.71	0.53	0.76	0.65	1.04	0.87	0.81
	S_1	0.12	0.08	0.08	0.08	0.07	0.13	0.11	0.07	0.12
HF/6-31G (3 states)	T_1	0.09	0.08	0.08	0.06	0.06	0.13	0.09	0.06	0.10
	S_1	0.12	0.08	0.08	0.09	0.09	0.14	0.12	0.08	0.14
HF/6-31+G* (1 state)	T_1	0.21	0.51	0.65	0.52	0.68	0.66	0.96	0.77	0.75
	S_1	0.27	0.61	0.77	0.65	0.82	0.80	1.15	0.92	0.94
HF/6-31+G* (2 states)	T_1	0.20	0.50	0.63	0.51	0.66	0.65	0.94	0.75	0.74
	S_1	0.15	0.18	0.21	0.21	0.22	0.22	0.38	0.24	0.33
HF/6-31+G* (3 states)	T_1	0.13	0.13	0.15	0.14	0.15	0.15	0.21	0.14	0.21
	S_1	0.16	0.19	0.26	0.28	0.32	0.36	0.52	0.43	0.60
B3LYP/6-31G (1 state)	T_1	0.01	0.48	0.65	0.44	0.64	0.55	1.00	0.52	0.66
	S_1	0.05	1.01	0.74	0.52	0.74	0.63	1.11	0.64	0.78
B3LYP/6-31G (2 states)	T_1	0.01	0.47	0.64	0.44	0.63	0.54	0.98	0.51	0.65
	S_1	0.04	0.47	0.06	0.04	0.06	0.13	0.28	0.27	0.02
B3LYP/6-31G (3 states)	T_1	0.01	0.02	0.06	0.04	0.05	0.12	0.28	0.27	0.03
	S_1	0.04	0.47	0.06	0.04	0.06	0.12	0.28	0.27	0.02
XPol-HF/6-31G (1 state)	T_1	0.06	0.06	0.07	0.07	0.06	0.06	0.14	0.08	0.10
	S_1	0.11	0.06	0.07	0.08	0.07	0.07	0.16	0.08	0.11
XPol-HF/6-31+G* (1 state)	T_1	0.10	0.17	0.65	0.52	0.68	0.66	0.98	0.77	0.75
	S_1	0.13	0.20	0.77	0.65	0.82	0.80	1.19	0.92	0.94
XPol-HF/6-31+G* (2 state)	T_1	0.10	0.17	0.24	0.22	0.27	0.25	0.57	0.29	0.47
	S_1	0.10	0.20	0.25	0.24	0.27	0.28	0.38	0.34	0.32
XPol-HF/6-31+G* (3 state)	T_1	0.07	0.12	0.16	0.14	0.16	0.16	0.27	0.18	0.22
	S_1	0.10	0.17	0.20	0.20	0.23	0.21	0.38	0.23	0.31

[a]With respect to a supersystem CIS calculation using the same AO basis set.

Table 3. Absolute Errors[a] (in eV) in $S_0 \rightarrow S_1$ and $S_0 \rightarrow T_1$ Excitation Energies for Large Water Clusters Extracted from a Simulation

	$(H_2O)_{57}$		$(H_2O)_{117}$	
snapshot	T_1	S_1	T_1	S_1
1	0.154	0.232	0.054	0.148
2	0.106	0.177	0.049	0.084
3	0.246	0.323	0.046	0.159
4	0.181	0.261		
5	0.296	0.178		

[a]With respect to a supersystem CIS/6-31G calculation, using a one-state Hartree−Fock basis.

adjacent molecules. This results in weaker interfragment interactions such that the unperturbed fragment MOs better represent the supersystem than they do in the case of geometry-optimized clusters. Furthermore, increasing the number of fragments effectively increases the variational space and stabilizes the excited states, as can be seen by the decrease in error going from the $(H_2O)_{53}$ to $(H_2O)_{117}$.

Tests for clusters of conjugated organic chromophores are shown in Table 4. These include clusters of pentacene, $C_{22}H_{14}$, a material that has been widely discussed in the context of singlet fission,[29] along with clusters of a methylated napthalenediimide (NDI) chromophore that forms the basic building block of a self-assembling nanotube.[4,5] These structures are shown in Figure 5. (The pentacene cluster is obtained from the crystal structure of the "LT" polymorph reported in ref 30. Structural parameters for the NDI nanotube were obtained from ref 5.) We find that the exciton model affords excellent agreement with supersystem calculations for these systems, with errors as low as 0.01 eV; however, the accuracy is highly sensitive to the choice of fragment MOs included in the basis. As compared to previous examples, the NTO expansions for these more complex molecular fragments include a greater number of significant terms that must be included in eq 6 for the matrix elements. Whereas in previous

dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366−5376

Table 4. Absolute Errors[a] in the $S_0 \rightarrow T_1$ Excitation Energy for Two Systems Composed of Larger Monomers

cluster	exciton basis	threshold (%)	error (eV)
$(NDI)_2$	HF/6-31G	95	0.014
$(NDI)_2$	HF/6-31G	85	0.142
$(NDI)_2$	HF/6-31+G*	85	0.129
$(NDI)_2$	B3LYP/6-31G	90	0.282
$(NDI)_6$	HF/6-31G	90	0.079
$(NDI)_6$	B3LYP/6-31G	90	0.327
$(NDI)_9$	HF/6-31G (bright state)	85	0.108[b]
$(pentacene)_2$	HF/6-31G	95	0.010
$(pentacene)_2$	HF/6-31G	85	0.110
$(pentacene)_2$	HF/6-31+G*	85	0.114
$(pentacene)_2$	B3LYP/6-31G	85	0.357
$(pentacene)_6$	HF/6-31G	90	0.044
$(pentacene)_6$	B3LYP/6-31G	85	0.364

[a]With respect to a supersystem CIS calculation in the same AO basis set, using one state per monomer. [b]Bright state error, relative to supersystem S_{30}.

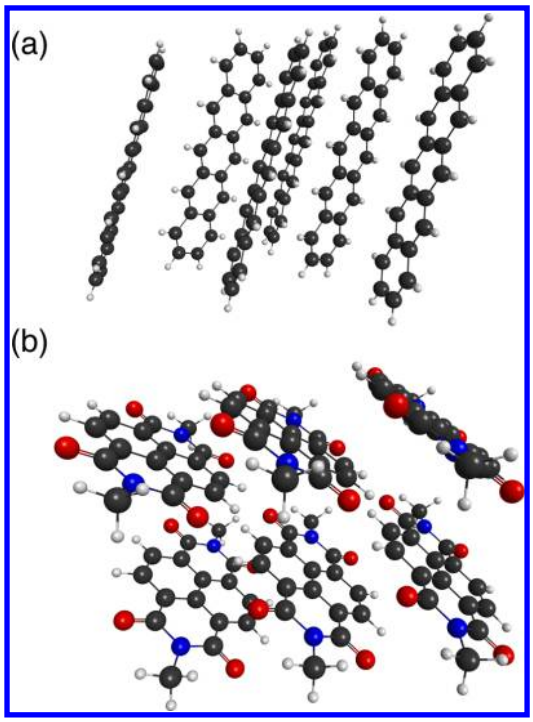


Figure 5. Structures of (a) $(pentacene)_6$ and (b) $(NDI)_6$.

examples we included a fixed number of NTOs, in these examples, it is prudent to truncate the expansions at some fixed fraction of the excitation amplitude. We have found that a threshold of 85% of the norm of **T** provides a reasonable balance of performance and accuracy.

The use of fragment MOs from DFT calculations, in this case calculated with the B3LYP functional, is another means of reducing the computational cost since their NTO expansions tend to include fewer terms. For the examples in Table 4, however, we find that the use of B3LYP MOs in the exciton model degrades the accuracy of the excitation energies somewhat, so the errors of ∼0.3 eV arise in the case of B3LYP orbitals whereas with Hartree−Fock MOs the errors were ≲0.1 eV. It should be recalled that "error" here is defined with respect to a supersystem CIS calculation; further testing is required to understand whether these larger B3LYP errors are related to this choice of supersystem reference calculation, to the B3LYP functional in particular, or to the use of Kohn−Sham MOs in general.

E. Comparison to the Renormalized Exciton Model. Recently, Ma and co-workers[13,14] developed a "renormalized" exciton model (REM), in which an effective Hamiltonian is constructed by projecting various n-body Hamiltonians (for n = 2 or 3 subunits) onto an excitonic direct-product basis. In Table 5, we compare the results of that approach to the present

Table 5. Errors for Linear H_2 Chains[a] from the REM-CIS Method and Our Frenkel−Davydov Exciton Model

	error[b]/mHartree					
	$S_0 \rightarrow T_1$			$S_0 \rightarrow S_1$		
		this work[c]			this work[c]	
system	REM[14]	$(H_2)_2$	$(H_2)_{n/2}$	REM[14]	$(H_2)_2$	$(H_2)_{n/2}$
			1.5 R_e			
$(H_2)_6$	−8.68	69.78	39.60	−17.73	128.92	77.22
$(H_2)_8$	−13.31	76.52	29.40	−22.27	142.53	60.92
$(H_2)_{10}$	−15.86	79.70	22.14	−23.24	149.15	48.28
$(H_2)_{12}$	−17.39	81.38	17.02	−23.23	152.73	38.76
			2.0 R_e			
$(H_2)_6$	−5.58	32.01	16.64	−17.73	81.87	44.79
$(H_2)_8$	−8.53	34.42	11.23	−24.29	87.51	31.63
$(H_2)_{10}$	−10.13	35.49	7.93	−27.55	89.89	23.30
$(H_2)_{12}$	−11.08	36.03	5.83	−29.22	91.05	17.85

[a]Using the equilibrium bond length of R_e = 0.7414 Å and an intermolecular separation of either 1.5 or 2.0 times R_e. [b]With respect to a supersystem CIS calculations, STO-6G basis. [c]With basis states as indicated.

Frenkel−Davydov-type model, in the case that both techniques employ CIS/STO-6G wave functions. As with the linear helium chains, this is an extreme case due to the highly delocalized nature of the excited states, and results using our method are rather poor when the excitonic basis is constructed from H_2 dimers. Adding higher-lying excited states does not improve the situation, because it is delocalization rather than polarization that limits the accuracy in these cases. However, results are improve significantly when larger segments of the chain are used to construct the basis states, and for basis states constructed using half the chain [$(H_2)_{n/2}$ in Table 5], errors in the $S_0 \rightarrow T_1$ excitation energy are comparable to those obtained using REM-CIS, although generally larger for the $S_0 \rightarrow S_1$ excitation.

In comparison, the REM-CIS results do an impressive job of capturing these highly delocalized excitations using smaller monomer units. Note, however, that extension of the REM approach to more realistic three-dimensional systems is slightly complicated, given that the number of n-body interactions included in the REM effective Hamiltonian scales as $\binom{N}{n}$ for N fragments. Terms beyond n = 2 are often critically important to the accuracy of ground-state many-body approaches,[31] and their importance should presumably increase in the excited state, given the increased size and polarizability of the excited-state wave function. Our ab initio exciton approach, on the other hand, includes many-body Coulomb and exchange effects and is essentially a "black box" whose application to any system is straightforward.

F. Size-Consistency Considerations. Supersystem CIS calculations are rigorously size-consistent,[6] in the customary

dx.doi.org/10.1021/ct500765m | J. Chem. Theory Comput. 2014, 10, 5366−5376

supermolecular sense of the term. Nonorthogonal CI can be made to satisfy size-consistency of the total energy (meaning that the total energy for a system of well-separated fragments is equal to the sum of the fragment energies) and also size-intensivity of excitation energies (meaning that the excitation energy is unaffected by the addition of a distant and therefore noninteracting fragment).[32] These facts are straightforward to verify since off-diagonal coupling elements of the exciton Hamiltonian vanish in the limit of large interfragment separation.

Although often conflated with size-consistency, *size-extensivity* is more properly defined as a nonvanishing correlation energy per particle in the thermodynamic (infinite, periodic) limit. This property has been considered carefully by Hirata,[33] although he calls it *size-consistency*, for reasons explained in ref 33, and concludes that CIS is rigorously size-extensive provided that the HF determinant is used as the reference state. In contrast, the reference state for our Frenkel−Davydov exciton model is composed of HF MOs from multiple, independent fragment calculations; hence, the formal requirements of size-extensivity are not met. This opens the possibility of increasing, size-dependent errors for interacting systems. This issue was also pointed out by Sundholm and Head-Gordon,[32] who note that in a nonorthogonal CI calculation, an increasing [but only as $O(N^2)$] number of determinants may be required to achieve a comparable level of accuracy as system size grows, yet excitation energies remain size-intensive in the sense defined above.

Our model does not satisfy Brillouin's theorem; hence, size-inconsistency may manifest as spurious stabilization of the ground-state direct-product wave function via mixing with the excitonic basis functions. This could potentially result in an unbalanced treatment of the excited states relative to the ground state, and errors that might increase as the size of the excitonic basis increases. Since the basis functions are spin eigenstates, for a singlet ground state this stabilization appears only when computing singlet excited states. We may therefore quantify the extent of this stabilization by comparing the ground-state eigenvalue of the exciton Hamiltonian in the singlet ($S_0 \rightarrow S_n$) and triplet ($S_0 \rightarrow T_n$) excitation cases.

For linear He chains, using single He atoms as monomer units, the aforementioned stabilization is found to increase linearly with system size at the rate of 0.005 mHartree/monomer. However, errors in singlet excitation energies mirror those for triplet excitation and are essentially constant for $N >$ 10 He atoms (see Figure 1). Despite the linear growth in the ground-state stabilization per monomer, in actuality, this effect is dictated not so much by system size but rather by the size of the direct-product basis. This is particularly evident in our data for water clusters computed using the 6-31+G* basis (Table 2), where the inclusion of three excited states per monomer slightly increases the errors in excitation energy of the singlet case, relative to errors obtained using two excited states per monomer. No corresponding effect is observed for triplet excitations, and we attribute this effect to ground-state stabilization in the singlet case.

This effect can be ameliorated using XPol monomer wave functions, as is evident in the water cluster data in Table 2. XPol allows the primary orbital relaxation effects in the ground state to be included explicitly in the monomer wave functions, so that there is less of a driving force for mixing excited-state direct-product wave functions into the ground state. In general, we find that errors stemming from size-inconsistency are quite small and our exciton method performs well when compared to rigorously size-consistent supersystem CIS calculations, even for the sizable systems that we have explored.

IV. PERFORMANCE

A. Bottlenecks. The SVD step of the corresponding orbitals transformation in eq 8 formally scales as $O(N^3)$ but with a very small prefactor. In practice, the overall computational cost is dominated by contraction of two-electron integrals with a density matrix to form a Fock-like matrix:

$$F_{\lambda\rho} = \sum_{\mu\nu} \langle \lambda\mu || \rho\nu \rangle P_{\mu\nu} \tag{16}$$

For the exciton model, $P_{\mu\nu}$ is a generalized density matrix whose form is suggested by eq 13:

$$P_{\mu\nu} = \sum_{j} \tilde{R}_{\nu j} \left(\frac{\Xi^{\mathrm{LR}}}{\tilde{s}^{\mathrm{LR}}_{jj}} \right) \tilde{L}^{\dagger}_{j\mu} \tag{17}$$

The cost of the contraction in eq 16 scales as $O(N^x_{\mathrm{basis}})$ where $2 \leq x \leq 4$, depending on system size and integral thresholds.

Construction of each exciton matrix element (eq 6b) requires N^2_{NTO} contraction steps, where N_{NTO} is the number of NTO pairs retained in the CI expansion of the individual fragment excited states. We note that the primary bottleneck is not generation of the AO integrals $\langle \lambda\mu || \rho\nu \rangle$ but rather their digestion (eq 16) with numerous generalized density matrices. We are currently pursuing ways to accelerate this step.

In total, there are $N_{\mathrm{frgm}}(N_{\mathrm{frgm}} + 1)/2$ matrix elements of the exciton Hamiltonian to compute so total CPU time scales as $O(N^2_{\mathrm{frgm}} N^x_{\mathrm{basis}})$. Memory usage is favorable; each matrix element requires a few $N_{\mathrm{basis}} \times N_{\mathrm{basis}}$ arrays, although our implementation will happily exploit additional memory for efficient integral evaluation. The same is true for disk storage, requiring only the occupied and a few virtual NTOs from each fragment to be stored in the fragment (n_{basis}) basis sets.

Traditional supersystem CIS and TDDFT calculations employ Davidson iteration to diagonalize the singly-excited block of the CI Hamiltonian. The number of iterations required to reach convergence varies with the system and number of desired eigenvalues but is typically 10−30 iterations when several (but ≪30) eigenvalues are requested. For each Davidson iteration, the rate-determining step is the contraction of the AO two-electron integrals with subspace trial vectors. This is directly related to the contraction step written in eq 16, in that $P_{\mu\nu}$ is the pseudodensity matrix of Maurice and Head-Gordon.[34] Memory requirements are the same as above, but all occupied and virtual orbitals in the supersystem basis set must be stored.

In view of these considerations, one can straightforwardly conclude that total CPU time is minimized by the method that requires the fewest instances of the digestion step in eq 16. In general, supersystem methods are highly efficient at minimizing this step as well as the overall single-threaded CPU time. However, the scaling of this performance is limited for modern computer architectures, and below, we discuss the performance of an implementation in which calculation of the exciton matrix elements is distributed across cores.

B. Parallelization. The serial efficiency of traditional supersystem methods is highly optimized, but the potential for parallelization is fundamentally limited by their iterative

nature; each step requires the results of the previous iteration before it can proceed. The work done in each iteration can be parallelized; for example, the integrals in eq 16 can be batched into different tasks. This approach scales fairly well across the cores of a single node, but efficiency is reduced, often significantly, by latencies in communication when scaling across nodes is required.

Our method is designed to overcome both of these limitations. Iterations are limited to fragments only, and the matrix element calculations are entirely independent of one another and thus trivial to distribute across any arrangement of processors. In fact, internode parallelism could further be augmented by shared memory intranode parallelism. We have a implemented a parallel version of our method (based on the Message Passing Interface, MPI) that parallelizes matrix element computations across cores within a distributed memory model. Given $N_{\mathrm{frgm}}(N_{\mathrm{frgm}} + 1)/2$ cores, the total wall time is essentially reduced to the time to compute a single matrix element. Given this degree of parallelism, the wall time required for our exciton model will be less than that of a traditional supersystem CIS calculation if the number of terms in eq 6, each of which requires one instance of a contraction such as that shown in eq 16, is fewer than the number Davidson iterations necessary to converge the supersystem calculation. Due to the efficient scaling of our method across processors this can still be true when considering parallel performance of supersystem methods on equivalent hardware, as shown below.

Chemical insight into the nature of the electronic states of the system of interest can be applied to the fragment partitioning in order to minimize cost while maintaining accuracy. For example, the lowest triplet excitation of a gas-phase guanine−cytosine (GC) base pair is highly localized on the cytosine moiety but is significantly stabilized by hydrogen bonding to the other base. This suggests using base pairs, rather than individual nucleobases, to generate basis states for clusters or π stacks. Timing data for various clusters of GC base pairs are given in Table 6. (The GC dimer structure was taken from ref 35 and stacked to form larger systems.)

Table 6. Wall Clock Times for Parallel Computation of the Lowest Triplet Excitation of GC Base Pair Clusters

		wall time/s		
		exciton model[a]		
N_{frgm}	no. of cores	75%[b]	85%[c]	full CIS
2	3	43	158	2961
4	10	347	1130	9553
6	21	908	3175	14178
8	36	1760	6253	9886
12	75	4605	15578	24096

[a]Estimated wall time for parallel calculation on $N_{\mathrm{frgm}}(N_{\mathrm{frgm}} - 1)/2$ processers. [b]Absolute errors are 0.2 eV for this threshold. [c]Absolute errors are 0.1 eV for this threshold.

The timing data for the exciton models presented in Table 6 are estimated from the time required for the computation of a single matrix element, as the fragment SCF and CIS calculations represent negligible overhead, as does diagonalization of the exciton Hamiltonian matrix. The NWChem program[36] was used as a benchmark for the parallel supersystem calculations, which is somewhat incongruous because the exciton model is implemented in Q-Chem, which exhibits far better serial performance as compared to NWChem. However, NWChem is generally recognized as state-of-the-art for scalability (and Q-Chem's implementation of TDDFT does not scale to the number of cores indicated in Table 6); thus, the comparison presented in Table 6 is at least interesting. By taking advantage of nearly perfect parallel scaling, the exciton method can indeed outperform a traditional supersystem calculation on equivalent hardware by up to a factor of 15, with errors in excitation energies of ≈0.2 eV. Of course, for systems composed of very large monomer units, the embarrassing parallelism of the exciton model could be combined with a parallel implementation of CIS to compute the fragment excited states, and (more importantly) a parallel version of the digestion step in eq 16, in order to accelerate calculation of the individual matrix elements.

V. POSSIBLE IMPROVEMENTS AND EXTENSIONS

Results presented above suggest that our ab initio exciton model is indeed a useful starting point for efficient calculation of collective excitations in multichromophore systems. Several ways in which this basic model might be improved or extended are sketched in this section.

A. Targeted States. Thus far we have only discussed application of our ab initio exciton method to the lowest-lying excited state of a given spin symmetry. Due to the physically based nature of the model, however, we can use it to target specific higher-lying states of molecular aggregates, with potentially dramatic savings in computational cost. For example, in the $(NDI)_9$ calculations reported in ref 5, the lowest optically bright state was S_{28}, with spacings of ∼0.05 eV between lower-lying, optically dark states. TDDFT calculations reported in ref 5 used the 3-21G* basis set, but for a larger number of NDI chromophores or a higher-quality basis set, the position of the bright state would inevitably move further up the manifold of singlet excited states. The need to calculate so many eigenvalues manifests as very large storage (memory and/or disk) requirements.

Application of chemical insight into the properties of the aggregate state of interest can dictate the proper choice of fragment states to include in the exciton basis, such that the specific supersystem excitation of interest is targeted. An example is the calculation of the lowest optically bright state of $(NDI)_9$. The bright state of the NDI aggregate involves a linear combination of the monomers in their spectroscopically active excited states. Supersystem methods that utilize the traditional Davidson algorithm simply solve for the requested number of lowest-lying eigenstates, and the state of interest is determined a posteriori by the user.

In a supersystem CIS or TDDFT calculation, the contraction in eq 16 must be computed for each unconverged root at each Davidson iteration, and the resulting matrices ($F_{\lambda\rho}$ in eq 16) of dimension $N_{\mathrm{basis}} \times N_{\mathrm{basis}}$ must be stored. (Recall that N_{basis} refers to the supersystem.) For large systems composed of large monomers, this is a significant burden and parallel scaling suffers from the same limitations outlined earlier. The exciton model effectively sidesteps this increase. In order to target the bright state of the cluster, basis states are chosen to be the spectroscopically active excited states of the monomers, determined by the orientation and magnitude of the monomer transition dipole moments.

For NDI at the CIS/6-31G level of theory, the fourth CIS root corresponds to the spectroscopically bright monomer excitation. Using these fragment states as an excitonic basis will

then target the bright state of the cluster. Unlike supersystem methods, computing higher-lying excited states for the exciton model has essentially the same computational cost as computing low-lying states because diagonalization of the exciton Hamiltonian matrix is trivial. Any increase in cost for higher-lying states arises from computing a few extra excited states for the fragments. The CPU and memory requirements for the exciton calculation are effectively unchanged. The results for this exciton calculation, reported in Table 4, agree well with supersystem methods while finishing 30% faster when both calculations were run on 12 cores.

B. Correlation Corrections. As a final demonstration of the value of the distinct physical partitioning used in our method, we present (albeit without rigorous derivation in this work) an approximate means of including electron correlation effects in the excitation energies without increasing the cost of the calculation. Due to the locality of electron correlation effects, it is not unreasonable to make the approximation that their contribution to excitation energies in an aggregate are primarily due to electron correlation within fragment units and thus to neglect interfragment electron correlation. If we then consider correlation to be a weak perturbation that only minimally deforms each fragment wave function, correlation effects can also be neglected in excitonic couplings. With these two approximations we can then treat the correlation for an exciton calculation as straightforward perturb-then-diagonalize correction.

In practice, we compute the correlation energies of the ground and excited states of each fragment at the MP2 and CIS(D) levels,[37] respectively, then add this correlation correction to the diagonal elements of the exciton Hamiltonian, weighted by the corresponding overlap elements. Even given this seemingly crude set of approximations, we have found that the resulting errors for excitation energies, relative to correlated CIS(D) supersystem calculations, are not significantly worse than the agreement documented above when the uncorrelated exciton model is compared to supersystem CIS calculations.

Both sets of errors [versus CIS and versus CIS(D)] for water clusters are presented in Figure 6, and we obtain similar accuracy for organic chromophores as well. A correlation corrected calculation on the GC dimer had the same 0.1 eV absolute error relative to a supersystem CIS(D) as an equivalent uncorrected exciton calculation relative to supersystem CIS. Errors approach or exceed 1 eV for the larger systems in Figure 6, probably because only a single excited state per monomer is included in the excitonic basis. (The formalism to include more than one state per monomer at a correlated level of theory is under development in our group.) Although the correlated supersystem calculations scale as $O(N^5)$, this correction is effectively free for the exciton calculation, since the computational effort in computing excited states of any one monomer, even including correlation effects, is small in comparison to the cost of computing the matrix elements of the exciton Hamiltonian.

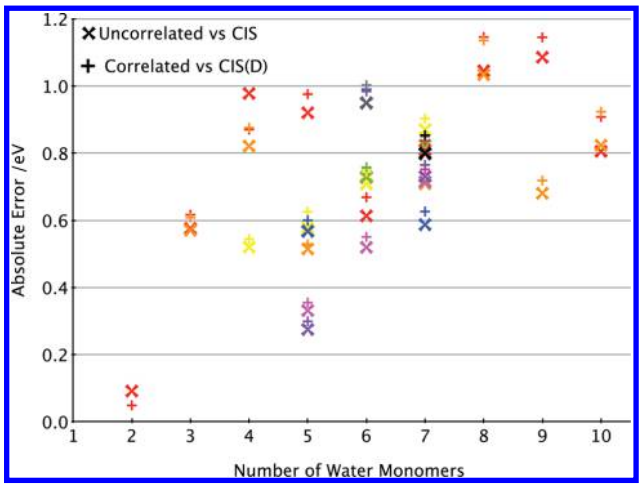


Figure 6. Absolute errors for water clusters with and without correlation corrections to the exciton model. Different colors at a given cluster size refer to different isomers, and the basis consists of a single CIS/6-31G state per monomer.

VI. CONCLUSIONS

We have introduced a novel method for computing excited states of aggregates based on an ab initio implementation of an exciton model. An excitonic basis is constructed from direct products of fragment configuration state functions and exact matrix elements of the Hartree−Fock Hamiltonian are computed using the corresponding orbitals transformation. This approximation has been shown to maintain accuracy, even for highly delocalized excitations, so long as basis states are chosen appropriately. In large systems such as water clusters and crystalline organic materials, the basis can be chosen such that excitation energies lie within ≈0.1 eV (or less) of those computed from a supersystem CIS calculation. Furthermore, the excitonic basis can be expanded with higher-lying fragment excitations in order to capture polarization effects and thereby increase the accuracy. Kohn−Sham orbitals and especially explicitly polarized (XPol) fragment orbitals perform well as an excitonic basis. A straightforward application of chemical insight can help specify the choice of the excitonic basis to increase accuracy and reduce cost. Notably, it is possible to target a specific aggregate excitation based on the choice monomer states included in the basis.

Due to its near-perfect scaling with parallel hardware, our method can outperform supersystem methods for certain systems despite the fact that, as of now, total CPU time scales unfavorably relative to supersystem methods. We believe that our approach presents a strong theoretical and computational foundation for the design of excited-state algorithms that are capable of scaling to massively parallel computer architectures while maintaining an unambiguous physically motivated strategy. Future work will focus on reducing the scaling of this method, most likely by reducing the dimension of the orbitals included in the exact Hamiltonian via distance-based charge-embedding approximations. Accurate correlation effects, hinted at in Section V.B, will be rigorously derived, along with a formal derivation of a perturbative approach applicable to an exciton Hamiltonian. Finally, we note that this method is potentially well-suited to be implemented on GPUs and other massively parallel vector accelerators. We intend to report on these developments in future publications.

AUTHOR INFORMATION

Corresponding Author

*E-mail: herbert@chemistry.ohio-state.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award No. DE-SC0008550. Calculations were performed at the Ohio Supercomputer Center under project PAA0003. J.M.H. is a Camille Dreyfus Teacher-Scholar.

■ REFERENCES

(1) Zheng, S.; Phillips, H.; Geva, E.; Dunietz, B. D. *J. Am. Chem. Soc.* **2012**, *134*, 6944−6947.
(2) Bahers, T. L.; Pauporte, T.; Laine, P. P.; Labt, F.; Adamo, C.; Ciofini, I. *J. Phys. Chem. Lett.* **2013**, *4*, 1044−1050.
(3) Oprea, I.; Panait, P.; Cimpoesu, F.; Ferbinteanu, M.; Girtu, M. A. *Materials* **2013**, *6*, 2372−2392.
(4) Shao, H.; Seifert, J.; Romano, N. C.; Gao, M.; Helmus, J. J.; Jaroniec, C. P.; Modarelli, D. A.; Parquette, J. R. *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 7688−7691.
(5) Gao, M.; Paul, S.; Schwieters, C. D.; You, Z.-Q.; Shao, H.; Herbert, J. M.; Parquette, J. R.; Jaroniec, C. P. (in preparation).
(6) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *J. Phys. Chem.* **1992**, *96*, 135−149.
(7) Frenkel, J. *Phys. Rev.* **1931**, *37*, 17−44.
(8) Davydov, A. S. *Soviet Phys.-Usp.* **1964**, *530*, 145−180.
(9) LaLonde, D. E.; Petke, J. D.; Maggiora, G. M. *J. Phys. Chem.* **1988**, *92*, 4746−4752.
(10) Lee, D.; Greenman, L.; Sarovar, M.; Whaley, K. B. *J. Phys. Chem. A* **2013**, *117*, 11072−11085.
(11) Krueger, B. P.; Scholes, G. D.; Fleming, G. R. *J. Phys. Chem. B* **1998**, *102*, 5378−5386.
(12) Tretiak, S.; Zhang, W. M.; Chernyak, V.; Mukamel, S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13003−13008.
(13) Zhang, H.; Malrieu, J.-P.; Ma, H.; Ma, J. *J. Comput. Chem.* **2012**, *33*, 34−43.
(14) Ma, Y.; Liu, Y.; Ma, H. *J. Chem. Phys.* **2012**, *136*, 024113:1−11.
(15) Luzanov, A. V.; Sukhorukov, A. A.; Umanskii, V. E. *Theor. Exp. Chem.* **1974**, *10*, 354−361.
(16) Martin, R. L. *J. Chem. Phys.* **2003**, *118*, 4775−4777.
(17) Mayer, I. *Chem. Phys. Lett.* **2007**, *437*, 284−286.
(18) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291−299.
(19) Amos, A. T.; Hall, G. G. *Proc. R. Soc. London* **1961**, *A263*, 483−493.
(20) King, H. F.; Stanton, R. E.; Kim, H.; Wyatt, R. E.; Parr, R. G. *J. Chem. Phys.* **1967**, *47*, 1936−1941.
(21) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kús, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A., Jr.; Dop, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Pieniazek, P. A.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sergueev, N.; Sharada, S. M.; Sharmaa, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Vanovschi, V.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhou, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xua, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Mol. Phys.* **2014**, DOI: 10.1080/00268976.2014.952696.
(22) Krylov, A. I.; Gill, P. M. W. *WIREs Comput. Mol. Sci.* **2013**, *3*, 317−326.
(23) Elles, C. G.; Shkrob, I. A.; Crowell, R. A.; Bradforth, S. E. *J. Chem. Phys.* **2007**, *126*, 164503:1−8.
(24) Temelso, B.; Archer, K. A.; Shields, G. C. *J. Phys. Chem. A* **2011**, *115*, 12034−12046.
(25) Baerends, E. J.; Gritsenko, O. V.; van Meer, R. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16408−16425.
(26) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108:1−9.
(27) Herbert, J. M.; Jacobson, L. D.; Lao, K. U.; Rohrdanz, M. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7679−7699.
(28) Holden, Z. C.; Richard, R. M.; Herbert, J. M. *J. Chem. Phys.* **2013**, *139*, 244108:1−13.
(29) Zeng, T.; Hoffmann, R.; Ananth, N. *J. Am. Chem. Soc.* **2014**, *136*, 5755−5764.
(30) Siegrist, T.; Besnard, C.; Haas, S.; Schiltz, M.; Pattison, P.; Chernyshov, D.; Batlogg, B.; Kloc, C. *Adv. Mater.* **2007**, *19*, 2079−2082.
(31) Richard, R. M.; Lao, K. U.; Herbert, J. M. *Acc. Chem. Res.* **2014**, *47*, 2828−2836.
(32) Sundstrom, E. J.; Head-Gordon, M. *J. Chem. Phys.* **2014**, *140*, 114103:1−11.
(33) Hirata, S. *Theor. Chem. Acc.* **2011**, *129*, 727−746.
(34) Maurice, D.; Head-Gordon, M. *Mol. Phys.* **1999**, *96*, 1533−1541.
(35) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, P.; Pulay, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 3364−3374.
(36) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, *181*, 1477−1489.
(37) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem. Phys. Lett.* **1994**, *219*, 21−29.