

Natural Charge-Transfer Analysis: Eliminating Spurious Charge-Transfer States in Time-Dependent Density Functional Theory via Diabatization, with Application to Projection-Based Embedding

Kevin Carter-Fenk,* Christopher J. Mundy, and John M. Herbert*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 4195–4210

Read Online

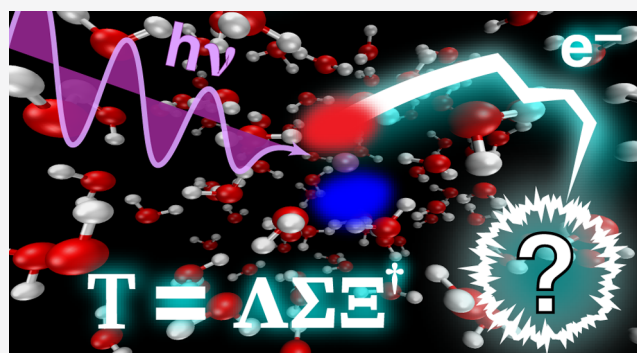
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: For many types of vertical excitation energies, linear-response time-dependent density functional theory (LR-TDDFT) offers a useful degree of accuracy combined with unrivaled computational efficiency, although charge-transfer excitation energies are often systematically and dramatically underestimated, especially for large systems and those that contain explicit solvent. As a result, low-energy electronic spectra of solution-phase chromophores often contain tens to hundreds of spurious charge-transfer states, making LR-TDDFT needlessly expensive in bulk solution. Intensity borrowing by these spurious states can affect intensities of the valence excitations, altering electronic bandshapes. At higher excitation energies, it is difficult to distinguish spurious charge-transfer states from genuine charge-transfer-to-solvent (CTTS) excitations. In this work, we introduce an automated diabaticization that enables fast and effective screening of the CTTS acceptor space in bulk solution. Our procedure introduces “natural charge-transfer orbitals” that provide a means to isolate orbitals that are most likely to participate in a CTTS excitation. Projection of these orbitals onto solvent-centered virtual orbitals provides a criterion for defining the most important solvent molecules in a given excitation and be used as an automated subspace selection algorithm for projection-based embedding of a high-level description of the CTTS state in a lower-level description of its environment. We apply this method to an *ab initio* molecular dynamics trajectory of $\text{I}^-(\text{aq})$ and report the lowest-energy CTTS band in the absorption spectrum. Our results are in excellent agreement with the experiment, and only one-third of the water molecules in the $\text{I}^-(\text{H}_2\text{O})_{96}$ simulation cell need to be described with LR-TDDFT to obtain excitation energies that are converged to <0.1 eV. The tools introduced herein will improve the accuracy, efficiency, and usability of LR-TDDFT in solution-phase environments.



1. INTRODUCTION

Charge-transfer excitations are critical to the function of dye-sensitized solar cells,^{1,2} solar cells based on organic semiconductors,³ charge transport at interfaces,⁴ and photoprotective mechanisms in plant biology.^{5–7} Characterizing charge displacement in a predictive way could also be of considerable benefit to the design of photocatalytic materials capable of achieving charge separation in the context of solar energy storage.⁸ In the solution phase, there exists a special class of charge-transfer phenomena known as charge-transfer-to-solvent (CTTS) excitations, characterized by photodetachment of an electron from a parent chromophore that stimulates rapid solvent rearrangement to produce a solvated electron.^{9–20} Due to the intimate participation of the solvent, CTTS states provide a useful experimental handle for investigating the local solvation structure.^{21–24} CTTS excitation of solutes such as $\text{I}^-(\text{aq})$ and $\text{CN}^-(\text{aq})$ is a standard means to generate hydrated electrons in the laboratory.^{25,26} Aside from its exotic nature as a quantum-mechanical

solute,^{27–30} $e^-(\text{aq})$ is a byproduct of water radiolysis^{26,31} and is frequently implicated in DNA damage by ionizing radiation.^{31–35}

Modeling CTTS excitations presents many challenges for quantum chemistry, the most apparent of which is the system size required to model the solvent-supported $e^-(\text{aq})$ species.³⁶ Configuration interaction with single excitations (CIS) is an affordable excited-state method but lacks the dynamical correlation necessary for quantitative accuracy. CT excitations are systematically overestimated by the CIS theory³⁷ because they are dominated by one single substitution between donor and acceptor orbitals with little spatial overlap, meaning that

Received: April 26, 2021

Published: June 30, 2021



the only singly excited determinant that contributes to the excited-state wave function is the one that shifts the charge. In this limit, there can be no orbital relaxation in the absence of double excitations.

Linear-response time-dependent density functional theory^{38–40} (LR-TDDFT) does describe dynamical electron correlation and achieves a statistical accuracy of 0.2–0.3 eV for many types of localized, valence excitation energies,⁴¹ but it often badly underestimates CT excitation energies.^{42–53} When using LR-TDDFT to compute the electronic spectroscopy of a hydrated chromophore, spurious low-energy CT states inevitably appear in the spectrum.^{47–49} This behavior originates in the lack of any penalty for forming a charge-separated excited state, within the linear response of semilocal density functionals, as a result of incorrect asymptotic behavior of the exchange–correlation potential.^{42–44} In a sufficiently large system, there is no reason why there cannot exist well-separated occupied and virtual orbitals whose energy levels differ by a gap corresponding to an optical photon (for example), and in the absence of the aforementioned exchange penalty, such coincidences manifest as spurious CT excitations at optical wavelengths. Stated differently, for semilocal DFT, the orbital Hessian that defines the LR-TDDFT eigenvalue problem assumes a block structure when the donor (ψ_D) and acceptor (ψ_A) orbitals are well-separated, and the corresponding $\psi_D \rightarrow \psi_A$ excitation energy reduces to the orbital eigenvalue difference, $\Delta E_{D \rightarrow A} = \epsilon_A - \epsilon_D$.⁴³ In large systems, it is almost inevitable that some of these spurious CT states will be quasi-degenerate with the genuine bright states of interest, leading to spurious intensity borrowing that depletes intensity from the genuine states.⁴⁸ Furthermore, because the cost of LR-TDDFT scales as $n_{\text{states}} \times \mathcal{O}(n_{\text{basis}}^4)$, the presence of these spurious states significantly increases the cost of solution-phase LR-TDDFT, as a dense manifold of CT states must be computed to reach the first bright state.

Range-separated hybrid (RSH) functionals can address these problems^{54–61} by pushing the spurious CT states out of the spectral range of interest. The consequences for genuine CTTS states are unclear, however. Another problem that limits the routine application of LR-TDDFT to aqueous chromophores is that the use of finite model systems leads to anomalously low-energy virtual orbitals on the vacuum-exposed water molecules at the edges of a hydrated cluster model. These result in spurious excitations from the chromophore to the surface of the solvent cluster.⁴⁸ The same phenomenon afflicts hybrid quantum mechanics/molecular mechanics (QM/MM) calculations, resulting in excitations from the chromophore to the QM/MM boundary.⁴⁸ The present state of affairs is that the most affordable workhorse method for excited-state quantum chemistry calculations cannot be routinely applied to solution-phase chromophores in explicit solvent.

There have been some efforts to avoid spurious CT states in LR-TDDFT calculations based on the use of “absolutely localized” molecular orbitals (ALMOs).^{62–67} This approach provides spectra that are free of CT states by construction and therefore sacrifices the ability to describe physically meaningful CT states. It has been proposed to allow these back into the spectrum based on a radial cutoff.⁶⁶ However, the CTTS excitations considered here involve water molecules as much as 10 Å away from the parent chromophore and the requisite distance cutoff would encompass essentially the entire system, including numerous water molecules that are spectators with

respect to the CTTS transition. This would reintroduce the spurious CT states.

The present work introduces an alternative. Using a pseudo-diabatization of the ground-state density matrix originally developed for electron-transfer problems,⁶⁸ we introduce a form of “natural” CT analysis based on diagonalization of the first-order donor–acceptor coupling matrix. The resulting natural CT populations provide a means for automated selection of the water molecules that are needed to describe the CTTS state, affording a black-box way to define the high-level subspace for projection-based embedding.^{69–71} We demonstrate that this approach can be used to isolate the physically meaningful CTTS states from the spurious CT states while preserving the valence excitation spectrum of the chromophore. The algorithm is stable with respect to the choice of the basis set and density functional and is used herein to obtain quantitative agreement with the experimental CTTS spectrum of $\Gamma^-(\text{aq})$.

2. THEORY

In what follows, indices i, j, k, \dots refer to occupied MOs; a, b, c, \dots refer to virtual MOs; and r, s, t, \dots index arbitrary MOs. Atomic orbital (AO) basis functions are indicated with Greek indices μ, ν, λ, \dots . Orbitals are denoted ϕ and ψ , whereas Φ and Ψ indicate many-electron wave functions. Spin indices are omitted for simplicity. To contextualize this work, Section 2.1 provides an overview of projection-based embedding and Section 2.2 outlines how to apply embedding with dielectric continuum boundary conditions, which will be used for application to $\Gamma^-(\text{aq})$. The localization scheme that we use is described in Section 2.3 and the diabaticization procedure in Section 2.4, following which we introduce natural CT orbitals in Section 2.5. We then describe how to connect this construction to the selection of solvent molecules, in Section 2.6, which affords an automated partition for projection-based embedding. Finally, Section 2.7 describes our approach to range separation in LR-TDDFT.

2.1. Projection-Based Embedding. There are many different flavors of QM/QM embedding methods,⁷² ranging in complexity from the subtractive “ONIOM” approach⁷³ to inverse schemes.^{74,75} Methods of intermediate complexity include frozen density embedding,^{76–78} which is closely related to subsystem DFT,⁷⁹ and projection-based embedding.^{69–71} Underlying each of these methods is the idea of applying a higher level of QM theory to a small region of interest while the environment is described at a lower level of theory.

The present work is based on the projection-operator embedding formalism,^{69–71} which provides a simple and formally exact solution to the subspace partitioning problem that does not require inversion of the Kohn–Sham equations nor approximate construction of nonadditive kinetic energy functionals. In the projection-based approach, one partitions the (super)system of interest into a high-level subsystem and its environment at the level of atoms, which must be mapped in some way onto MOs, thus defining density matrices for the subsystem (γ^A) and the environment (γ^B). For the mapping from atoms to MOs, we employ the “subsystem-projected AO decomposition” or “SPADE” procedure,⁸⁰ as described in Section 2.3.

In its original formulation,⁶⁹ projection-based embedding amounts to the following expression for the energy of the AUB supersystem, with density matrix $\gamma^{A \oplus B}$:

$$E = E_{\text{low}}[\gamma^{A\oplus B}] + E_{\text{high}}[\tilde{\gamma}^A] - E_{\text{low}}[\gamma^A] + \text{tr}[(\tilde{\gamma}^A - \gamma^A)\mathbf{v}_{\text{embed}}] + \mu \text{tr}(\tilde{\gamma}^A \mathbf{P}^B) \quad (1)$$

Here, the density matrix $\tilde{\gamma}^A$ reflects subsystem A computed at the higher level of theory. The first three terms in this energy expression are analogous to the ONIOM scheme,⁷³ wherein the low-level approximation for the subsystem ($E_{\text{low}}[\gamma^A]$) is subtracted from the low-level description of the supersystem ($E_{\text{low}}[\gamma^{A\oplus B}]$) and replaced by the high-level treatment of the subsystem ($E_{\text{high}}[\tilde{\gamma}^A]$). Writing the Fock matrix as $\mathbf{F} = \mathbf{h} + \mathbf{g}[\mathbf{P}]$, where \mathbf{g} represents the two-electron contributions built from density matrix \mathbf{P} , the embedding potential $\mathbf{v}_{\text{embed}}$ in eq 1 is defined as

$$\mathbf{v}_{\text{embed}} = \mathbf{g}[\gamma^{A\oplus B}] - \mathbf{g}[\gamma^A] \quad (2)$$

This is constructed solely from density matrices obtained at the lower level of theory. The final ingredient in eq 1 is \mathbf{P}^B , the matrix representation of the projection operator

$$\hat{P}^B = \sum_{i \in B}^{\text{occ}} |\phi_i^B\rangle\langle\phi_i^B| \quad (3)$$

When combined with a sizable level-shift parameter μ , the effect of $\mu\hat{P}^B$ in eq 1 is to ensure that the occupied MOs of the environment B cannot hybridize with those of subsystem A. This maintains orthogonality between the two subsystems, such that there is no nonadditive kinetic energy.⁶⁹

In practice, the need for an empirical level-shift parameter can be eliminated by projecting the variational degrees of freedom from subsystem B out of the Fock matrix for A

$$\tilde{\mathbf{F}}^{A\text{-in-}B} = (\mathbf{1} - \mathbf{S}\mathbf{y}^B)\mathbf{F}(\mathbf{1} - \mathbf{y}^B\mathbf{S}) \quad (4)$$

(Here, \mathbf{S} denotes the AO overlap matrix.) This eliminates the need for a level-shift term in eq 1, so that the supersystem energy is given by

$$E = E_{\text{low}}[\gamma^{A\oplus B}] + E_{\text{high}}[\tilde{\gamma}^A] - E_{\text{low}}[\gamma^A] + \text{tr}[(\tilde{\gamma}^A - \gamma^A)\mathbf{v}_{\text{embed}}] \quad (5)$$

eliminating the μ parameter. All of the projection-based embedding calculations presented in this work employ the “parameter-free” approach in eqs 4 and 5.

Application of LR-TDDFT to a reference state obtained by minimizing the energy expression in eq 5 is still subject to the appearance of spurious CT states in solvated systems, as the entire virtual space of $A\oplus B$ is used as the active space. ALMO-based methods restrict this active space, at the expense of eliminating all of the CT states, such that CTTS states are completely excised and would need to be built back in.^{67,81} Bottom-up localization methods, such as the concentric localization scheme of Claudino and Mayhall,⁸² can be used to truncate the virtual space, but this does not establish a satisfactory search criterion to identify orbitals relevant to long-range CT excitations. Moreover, if solvent-centered virtual orbitals that contribute to the CTTS state are treated at a lower level of theory (e.g., with a semilocal functional), then any effort to describe chromophore-centered donor orbitals at a higher level of theory (e.g., using an RSH functional) is moot. It is therefore imperative to describe both the donor and acceptor subspaces using asymptotically correct functionals. Selecting solvent molecules for the acceptor space in an affordable, black-box manner requires an answer to the

question: where does charge localize after electron photo-detachment? This question is addressed by the diabaticization scheme that is introduced below.

2.2. Multilayer Embedding with Implicit Solvent.

Effective modeling of bulk solution is imperative for description of excitation spectra in the condensed phase, and CTTS states require especially careful attention in this regard. Gas-phase spectra⁸³ and LR-TDDFT calculations⁸⁴ of $\Gamma(\text{H}_2\text{O})_n$ clusters both suggest that the CTTS excitation energy increases as water molecules are added to the system, so to avoid underestimating these excitation energies, the simulation cell needs to be large enough to encompass the radial extent of the CTTS acceptor orbital. QM/MM models of $\Gamma(\text{aq})$ also suggest that long-range electrostatic polarization also plays a key role in the CTTS process,⁸⁵ as it does for other solvated anions such as $e^-(\text{aq})$.^{27–29} In this work, we present a novel multilayer embedding approach wherein the explicit QM system, which constitutes the entire simulation cell from an *ab initio* molecular dynamics (MD) trajectory, is partitioned into high- and low-level regions of the explicit solvent for application of projection-based embedding, and then, this entire atomistic system is further subject to dielectric continuum boundary conditions.

This multilayer embedding is accomplished by solving the generalized Poisson equation

$$\hat{\nabla} \cdot [\epsilon(\mathbf{r})\hat{\nabla}\varphi(\mathbf{r})] = -4\pi[\tilde{\gamma}^A(\mathbf{r}) \oplus \gamma^B(\mathbf{r})] \quad (6)$$

using the combined density $\tilde{\gamma}^A(\mathbf{r}) \oplus \gamma^B(\mathbf{r})$ from the embedding calculation, where $\epsilon(\mathbf{r})$ constitutes a sharp dielectric interface (cavity) surrounding the atomistic region.⁸⁶ In practice, this self-consistent reaction-field (SCRf) problem is implemented using a polarizable continuum formalism,⁸⁶ although direct solution of eq 6 in three-dimensional space could also be used.^{87–90} The result in either case is a one-electron potential \mathbf{v}_{SCRf} added to the Fock matrix for subsystem A:

$$\mathbf{F}^A = \mathbf{h} + \mathbf{v}_{\text{SCRf}}[\tilde{\gamma}^A \oplus \gamma^B] + \mathbf{g}[\tilde{\gamma}^A] + \mathbf{v}_{\text{embed}} \quad (7)$$

Inserting this into eq 4 facilitates self-consistent optimization of $\tilde{\gamma}^A$ within the embedding potential of both the explicit and implicit environments. Rather than obtaining \mathbf{v}_{SCRf} exclusively from the density optimized at the lower level of theory, eq 7 allows the continuum to respond to changes in the high-level subsystem density matrix and thus represents a fully self-consistent implementation of continuum boundary conditions for projection-based embedding.

2.3. SPADE Localization. Early versions of projection-based embedding used the Pipek–Mezey localization algorithm⁹¹ to assign MOs to atoms, based on the atomic population associated with each MO.⁷¹ More recently, the SPADE algorithm⁸⁰ has been introduced as a means to ensure a consistent partition of the MO space as the atoms are displaced, to make embedding usable for geometry optimizations and *ab initio* MD simulations. In contrast to the Pipek–Mezey orbitals, the SPADE orbitals are not spatially localized in the usual sense but are instead subsystem-localized. Spatial locality is not required for projection-based embedding; instead, the only requirement is that the total density be partitioned as a sum of subsystem densities.

The SPADE procedure begins with a transformation of the occupied MO coefficients \mathbf{C}_{occ} into the Löwdin-orthogonalized basis

$$\bar{\mathbf{C}}_{\text{occ}} = \mathbf{S}^{1/2} \mathbf{C}_{\text{occ}} \quad (8)$$

The MOs are then projected onto AO basis functions that are tagged to atoms belonging to the high-level subsystem A

$$\mathbf{C}_{\mu i}^{A, \text{occ}} = \sum_{\mu \in A} |\mu\rangle \langle \mu | \bar{\mathbf{C}}_{\mu i}^{\text{occ}} \quad (9)$$

Singular value decomposition of this coefficient matrix

$$\mathbf{C}^{A, \text{occ}} = \mathbf{L}^A \sum^A (\mathbf{R}^A)^\dagger \quad (10)$$

reveals a partition based on the kernel (null space) of \mathbf{R}^A . Specifically, let us write

$$\mathbf{R}^A = \mathbf{R}^{A, \text{span}} \oplus \mathbf{R}^{A, \text{null}} \quad (11)$$

where $\mathbf{R}^{A, \text{span}}$ denotes the first n_{basis}^A columns of the matrix \mathbf{R}^A , which together constitute the span of the right singular vectors. Projection onto this span localizes the MOs to subsystem A . The remaining vectors $\mathbf{R}^{A, \text{null}}$ represent the null space, and projection with $\mathbf{R}^{A, \text{null}}$ puts the remaining MOs in subsystem B

$$\mathbf{C}_{\text{SPADE}}^A = \mathbf{C}^{A, \text{occ}} \mathbf{R}^{A, \text{span}} \quad (12a)$$

$$\mathbf{C}_{\text{SPADE}}^B = \mathbf{C}^{A, \text{occ}} \mathbf{R}^{A, \text{null}} \quad (12b)$$

This same decomposition can be applied to the virtual MOs⁸² and will be used later in this work.

2.4. Diabatization Scheme. The projection-operator diabatization (POD) method⁶⁸ partitions the atoms into a donor set (D) and an acceptor set (A), and then uses the fact that the Gaussian basis functions are tagged to atoms to effect a corresponding partition of the one-electron Hamiltonian

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\text{DD}} & \mathbf{H}_{\text{DA}} \\ \mathbf{H}_{\text{AD}} & \mathbf{H}_{\text{AA}} \end{bmatrix} \quad (13)$$

This matrix and each of its sub-blocks are then transformed to the Löwdin-orthogonalized basis

$$\bar{\mathbf{H}} = \mathbf{S}^{-1/2} \mathbf{H} \mathbf{S}^{-1/2} = \begin{bmatrix} \bar{\mathbf{H}}_{\text{DD}} & \bar{\mathbf{H}}_{\text{DA}} \\ \bar{\mathbf{H}}_{\text{AD}} & \bar{\mathbf{H}}_{\text{AA}} \end{bmatrix} \quad (14)$$

The sub-blocks of $\bar{\mathbf{H}}$ are diagonalized by solving the subsystem eigenvalue equations

$$\bar{\mathbf{H}}_{\text{DD}} \mathbf{U} = \boldsymbol{\varepsilon}_{\text{D}} \mathbf{U} \quad (15a)$$

$$\bar{\mathbf{H}}_{\text{AA}} \mathbf{V} = \boldsymbol{\varepsilon}_{\text{A}} \mathbf{V} \quad (15b)$$

This affords diabatic MO coefficients \mathbf{U} (spanning the donor space) and \mathbf{V} (spanning the acceptor space). The usual POD approach terminates by rotating the off-diagonal blocks of the one-electron Hamiltonian into the diabatic representation

$$\begin{aligned} \mathbf{U}^\dagger \bar{\mathbf{H}}_{\text{DA}} \mathbf{V} &= \tilde{\mathbf{H}}_{\text{DA}} \\ \mathbf{V}^\dagger \bar{\mathbf{H}}_{\text{AD}} \mathbf{U} &= \tilde{\mathbf{H}}_{\text{AD}} \end{aligned} \quad (16)$$

The full one-electron Hamiltonian in the diabatic representation is

$$\tilde{\mathbf{H}} = \begin{bmatrix} \boldsymbol{\varepsilon}_{\text{D}} & \tilde{\mathbf{H}}_{\text{DA}} \\ \tilde{\mathbf{H}}_{\text{AD}} & \boldsymbol{\varepsilon}_{\text{A}} \end{bmatrix} \quad (17)$$

which would be diagonal but for the electron-transfer coupling matrix, $\tilde{\mathbf{H}}_{\text{DA}}$. The latter provides the couplings that appear in the electron-transfer rate theory.^{92–97}

2.5. Diabatic Minimal-Basis Construction. The first step in constructing a basis that isolates the physically relevant CTTS excitations is to realize that the structure of the POD Hamiltonian in eq 17 allows for the assembly of a first-order electron-transfer density matrix in a manner that is reminiscent of the frozen natural orbital approach based on the Møller–Plesset perturbation theory.^{98,99} The POD Hamiltonian is separable into a noninteracting diabatic part \mathbf{H}_0 and a coupling matrix \mathbf{W}

$$\tilde{\mathbf{H}} = \underbrace{\begin{bmatrix} \boldsymbol{\varepsilon}_{\text{D}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\varepsilon}_{\text{A}} \end{bmatrix}}_{\mathbf{H}_0} + \underbrace{\begin{bmatrix} \mathbf{0} & \tilde{\mathbf{H}}_{\text{DA}} \\ \tilde{\mathbf{H}}_{\text{AD}} & \mathbf{0} \end{bmatrix}}_{\mathbf{W}} \quad (18)$$

The corresponding MO coefficient matrix is block-localized

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \quad (19)$$

In this representation, the electron-transfer couplings are immediately available as matrix elements between orbitals catalogued as “donors” and those classified as “acceptors”.

Starting from the zeroth-order, one-electron Hamiltonian \mathbf{H}_0 , the first-order CT corrections to the diabatic wave functions are

$$|\Psi_{e_{\text{D}} \rightarrow h_{\text{A}}^*}^{(1)}\rangle = \sum_{i \in \text{D}} \sum_{a \in \text{A}} T_{ia} |\Psi_i^{(0)}\rangle \quad (20)$$

where $|\Psi_i^{(0)}\rangle$ is an eigenfunction of \hat{H}_0 with eigenvalue ε_i , which is contained in the diagonal matrix $\boldsymbol{\varepsilon}_{\text{D}}$. The mixing coefficients T_{ia} are given by

$$T_{ia} = \frac{\langle \Psi_i^{(0)} | \hat{W} | \Psi_a^{(0)} \rangle}{\varepsilon_a - \varepsilon_i} \quad (21)$$

with eigenvalues ε_a that are elements of $\boldsymbol{\varepsilon}_{\text{A}}$. The amplitude matrix \mathbf{T} has dimensions $n_{\text{occ}}^{\text{D}} \times n_{\text{vir}}^{\text{A}}$, where $n_{\text{occ}}^{\text{D}}$ is the number of occupied MOs on the donor and $n_{\text{vir}}^{\text{A}}$ is the number of virtual orbitals on the acceptor. The matrices $\mathbf{T}\mathbf{T}^\dagger$ and $\mathbf{T}^\dagger\mathbf{T}$ are the occupied (donor) and virtual (acceptor) blocks of the first-order CT density matrix.

For convenience, let us denote $\boldsymbol{\Gamma} = \mathbf{T}^\dagger\mathbf{T}$ or

$$\Gamma_{ab} = \sum_{i \in \text{D}} T_{ai}^* T_{ib} \quad (22)$$

The eigenvalue problem

$$\boldsymbol{\Gamma} \boldsymbol{\Xi} = \boldsymbol{\lambda} \boldsymbol{\Xi} \quad (23)$$

defines a set of natural CT orbitals $\boldsymbol{\Xi}$, with populations λ_a that are elements of the diagonal matrix $\boldsymbol{\lambda}$. In analogy to natural orbital truncation of correlated wave-function methods,^{98,99} where eigenvalues of the first-order density matrix are used to truncate the virtual space, the natural CT populations $\{\lambda_a\}$ can be used as a metric to truncate the virtual space. Similar approaches have been developed recently for nonorthogonal configuration interaction to choose the set of corresponding orbitals that most efficiently recovers correlation.^{100–103}

Perhaps a better analogy is the definition of the natural transition orbitals (NTOs) in a CIS or LR-TDDFT calculation.^{104–107} If \mathbf{X} represents the $n_{\text{occ}} \times n_{\text{vir}}$ matrix of CIS coefficients, then diagonalization of the particle and hole density matrices¹⁰⁷ ($\mathbf{X}\mathbf{X}^\dagger$ and $\mathbf{X}^\dagger\mathbf{X}$, respectively) generates NTOs for the occupied and virtual spaces, respectively.¹⁰⁴ Equivalently, these are the left and right singular vectors of

X ,¹⁰⁵ whose singular values demonstrate that the rank of the excitation space is exactly n_{occ} which is considerably smaller than the dimension of the excitation space, which is $n_{\text{occ}} n_{\text{vir}}$. Analogously, the span of Ξ in eq 23 has rank n_{occ}^D , which automatically sets an upper bound on the size of the CTTS acceptor space. The eigenvalues $\{\lambda_a\}$ quantify the role of each virtual orbital in serving as an acceptor for a CT excitation between donor and acceptor spaces defined by atomic partition.

The span of Ξ establishes an initial size for the virtual (acceptor) space that is the same size as the occupied (donor) space, resulting in an excitation space that is $n_{\text{occ}}^D \times n_{\text{occ}}^D$. An excitation space of this size should be sufficient for a purely CT excitation, but a state whose character is mixed may not be well described by a single pair of natural CT orbitals. In these cases, we recommend an iterative construction of T. At iteration n , the idea is to compute

$$T_{ia}^{(n)} = \frac{\langle \Psi_i | \hat{W} | \Psi_{a,\text{fz}}^{(n-1)} \rangle}{\epsilon_a - \epsilon_i} \quad (24)$$

where $|\Psi_{a,\text{fz}}^{(n-1)}\rangle$ denotes a (frozen) virtual orbital that was excluded from the excitation space during iteration $n-1$. Each successive diagonalization of $\mathbf{T}^{(n)} = (\mathbf{T}^{(n)})^\dagger \mathbf{T}^{(n)}$ (eq 23) yields a further set of n_{occ}^D ancillary CTTS acceptor orbitals (secondary, tertiary, etc.). This procedure can be repeated until a threshold percentage of the CT acceptor space is encompassed within the active orbitals. Each iteration increases the size of the CT acceptor space by n_{occ}^D , and in the limit of a 100% threshold, the entire CT acceptor space will be included in the active orbitals.

Although the diabatic MO coefficient matrix could be used directly in the screening procedure, we instead introduce a bijective mapping between the diabatic orbitals and the SPADE-localized MOs that have been projected onto the diabatic subsystems. The mapping is achieved through application of the following set of operators to the diabatic MO coefficients:

$$\hat{D}_o = \sum_{i \in D}^{\text{occ}} |\psi_i^{\text{SPADE}}\rangle \langle \phi_i^{\text{diab}}| \quad (25a)$$

$$\hat{A}_o = \sum_{j \in A}^{\text{occ}} |\psi_j^{\text{SPADE}}\rangle \langle \phi_j^{\text{diab}}| \quad (25b)$$

$$\hat{D}_v = \sum_{a \in D}^{\text{vir}} |\psi_a^{\text{SPADE}}\rangle \langle \phi_a^{\text{diab}}| \quad (25c)$$

$$\hat{A}_v = \sum_{b \in A}^{\text{vir}} |\psi_b^{\text{SPADE}}\rangle \langle \phi_b^{\text{diab}}| \quad (25d)$$

This allows the selected orbitals to be expressed directly in the SPADE representation. We find that the SPADE orbitals yield better oscillator strengths than the diabatic subspace MOs, allowing this method to be applied independent of any subsequent embedding scheme if only the CT spectrum is desired. A proof of the bijection between diabatic subspace MOs and the SPADE orbitals is provided in Appendix A.

Finally, let us arrange the MO coefficient matrix by columns in the format

$$\mathbf{C}_{\text{diab}}^{\text{SPADE}} = (\mathbf{C}_{\text{occ}}^A | \mathbf{C}_{\text{occ}}^D | \mathbf{C}_{\text{vir}}^A, \text{CTTS} | \mathbf{C}_{\text{vir}}^D | \mathbf{C}_{\text{vir}}^A, \text{fz}) \quad (26)$$

active

Note that the indicated “active” set of CT donor/acceptor orbitals results from diabaticization, not from an LR-TDDFT calculation, meaning that each CT excitation involving these orbitals may be taken to be a physically relevant CT state although these states need not appear in the low-energy part of the spectrum. To gauge which are the energetically important ones, an LR-TDDFT calculation can be performed using the very small set of active orbitals that are indicated in eq 26. The dimension of the excitation space for this calculation is $n_{\text{occ}}^D \times n_{\text{occ}}^D$ as compared to $n_{\text{occ}} \times n_{\text{vir}}$ (where $n_{\text{vir}} \gg n_{\text{occ}}^D$) for the full excitation space. This minimal-basis LR-TDDFT calculation represents a trivial amount of overhead, enabling rapid and accurate representation of low-lying CTTS states.

2.6. Selection of Relevant Solvent Orbitals. A wave function for the CTTS state can be constructed from the minimal-basis LR-TDDFT calculation suggested above, according to

$$|\Psi_{\text{CTTS}}\rangle = \sum_{i \in D} \sum_{\bar{a} \in A} c_{i\bar{a}} |\Psi_i^{\bar{a}}\rangle \quad (27)$$

The acceptor orbital is

$$|\Phi_{\text{CTTS}}^{\bar{a}}\rangle = \sum_{i \in D} \sum_{\bar{a} \in A} c_{i\bar{a}} |\phi_{\bar{a}}\rangle \quad (28)$$

In these expressions, \bar{a} indicates an active acceptor orbital belonging to the solvent. The acceptor orbital in eq 28 is then projected onto every individual solvent molecule using a projection operator

$$\hat{Q} = \sum_{\mu \in R} |\mu\rangle \langle \mu| \quad (29)$$

for fragment R , where $\{|\mu\rangle\}$ is the Löwdin-orthogonalized AO basis. The result is a fragment-blocked representation of the CTTS acceptor orbital, which can be written in matrix form as

$$\mathbf{Q}\mathbf{C}_{\text{CTTS}} = \begin{bmatrix} \mathbf{C}_R \\ \mathbf{C}_S \\ \vdots \\ \mathbf{C}_N \end{bmatrix} \quad (30)$$

where indices R, S, \dots denote molecular fragments.

The probability of finding the CTTS acceptor orbital on the R th solvent molecule is

$$p_R = \frac{\sum_a \sum_{\mu \in R} C_{R,\mu a}^* C_{R,\mu a}}{\left[\sum_R \left(\sum_a \sum_{\mu \in R} C_{R,\mu a}^* C_{R,\mu a} \right)^2 \right]^{1/2}} \quad (31)$$

Values of p_R assign each solvent molecule a probability that weights its importance in the description of the CTTS acceptor space. These probabilities are similar to natural populations, as they represent the lone eigenvalue of each projected CTTS MO density, and should therefore be quite stable with respect to changes in the AO basis set. The probability vector \mathbf{p} is sorted in descending order, allowing for a cumulative sum over normalized probabilities (representing fractional CTTS acceptor populations) to act as a threshold for the partition. Based on this threshold, the number of “important” charge-accepting solvent molecules can be determined, for use as the high-level subspace in projection-based embedding.

This concludes the natural CT analysis (NCTA) approach for pruning the LR-TDDFT manifold of spurious CT states. Figure 1 provides a flowchart of the procedure.

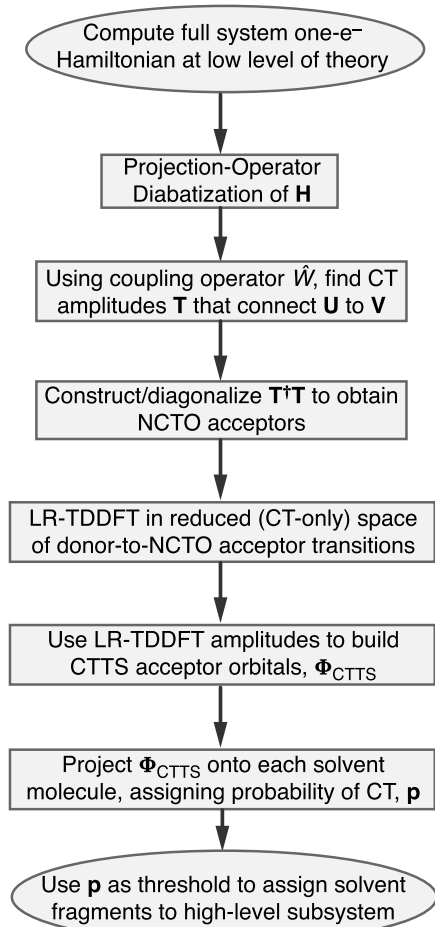


Figure 1. Flowchart summarizing the NCTA procedure.

2.7. Screened RSH Functionals. Systematic underestimation of CT excitation energies is largely solved in LR-TDDFT by application of long-range corrected (LRC) functionals,^{52–61} which are often deployed using an “optimal tuning” procedure based on the ionization energy (IE) criterion, $IE = -\epsilon_{\text{HOMO}}$.⁵⁸ Application of this criterion, however, exhibits a troubling dependence on system size,^{60,108–116} and a different approach is pursued here.

Range separation consists of a partition of the electron–electron Coulomb operator r_{12}^{-1} into short-range (SR) and long-range (LR) components. A widely used form for this partition is^{117,118}

$$\frac{1}{r_{12}} = \frac{1 - [\alpha + \beta \text{erf}(\omega r_{12})]}{r_{12}^{\text{SR}}} + \frac{\alpha + \beta \text{erf}(\omega r_{12})}{r_{12}^{\text{LR}}} \quad (32)$$

and a generic expression for the corresponding RSH functional is¹¹⁸

$$E_{\text{xc}}^{\text{RSH}} = \alpha E_{\text{x,HF}}^{\text{SR}} + (\alpha + \beta) E_{\text{x,HF}}^{\text{LR}} + (1 - \alpha) E_{\text{x,GGA}}^{\text{SR}} + (1 - \alpha - \beta) E_{\text{x,GGA}}^{\text{LR}} + E_{\text{c}}^{\text{GGA}} \quad (33)$$

where

$$E_{\text{x,GGA}}^{\text{LR}} = E_{\text{x}}^{\text{GGA}} - E_{\text{x,GGA}}^{\text{SR}} \quad (34)$$

While it has become common to identify the difference between a generalized gradient approximation (GGA) exchange functional ($E_{\text{x}}^{\text{GGA}}$) and its SR counterpart ($E_{\text{x,GGA}}^{\text{SR}}$) as the “LR” part of the exchange interaction, as is done in eq 34, this is something of a misnomer insofar as the reason for introducing range separation in the first place is that semilocal functionals do not exhibit genuine LR exchange.¹¹⁸ GGA exchange decays with the density, and operationally $E_{\text{x,GGA}}^{\text{LR}}$ might more appropriately be called “middle-range” exchange.

The parameter α in eq 33 controls the fraction of SR Hartree–Fock exchange (SR-HFX), $E_{\text{x,HF}}^{\text{SR}}$. The LR component of the Coulomb operator in eq 32 decays asymptotically as $\sim(\alpha + \beta)/r_{12}$; therefore, we consider that proper range separation requires $\alpha + \beta = 1$, at least in the gas phase. This means that β is not an independent parameter once the fraction of SR-HFX is chosen. (However, some RSH functionals violate this constraint, notably CAM-B3LYP.¹¹⁷) According to eq 33, any functional that satisfies $\alpha + \beta = 1$ is equivalent to using 100% HFX in the asymptotic limit, $r_{12} \rightarrow \infty$. We have suggested that these should be called LRC functionals,^{52,55,56,107,118} e.g., LRC- ω PBE if $E_{\text{x,GGA}}^{\text{SR}}$ is ω PBE,⁵⁶ the SR counterpart of the semilocal Perdew–Burke–Ernzerhof (PBE) exchange.^{56,119} This nomenclature distinguishes the LRC approach from functionals such as HSE,^{120,121} which also employs ω PBE exchange but omits $E_{\text{x,HF}}^{\text{LR}}$.

In a condensed medium whose dielectric constant is ϵ , one may argue that the proper asymptotic behavior should instead be $\sim(\epsilon r_{12})^{-1}$, leading to an alternative constraint $\alpha + \beta = 1/\epsilon$. Functionals satisfying the latter constraint have been called screened RSH (sRSH) or screened LRC (sLRC) functionals^{118,122} and have been applied recently to predict fundamental gaps and optical gaps of molecular solids.^{118,123–125}

The optimal-tuning procedure introduced by Baer and co-workers⁵⁸ selects the range-separation parameter ω in eq 32 to satisfy the condition

$$\epsilon_{\text{HOMO}}(\omega) = -IE(\omega) \quad (35)$$

However, when this tuning procedure is applied to negatively charged water clusters $(\text{H}_2\text{O})_n^-$, which are finite models of the hydrated electron, one finds that the tuned value of ω changes dramatically as the size of the water cluster increases,¹¹² portending problems for condensed-phase applications. As an alternative, in the present work, we employ a “global density-dependent” (GDD) tuning procedure,¹¹⁰ in which a self-consistent DFT calculation is first performed using a functional such as LRC- ω PBE and then the value of ω is set to an optimal value

$$\omega_{\text{GDD}} = C \langle d_x^2 \rangle^{-1/2} \quad (36)$$

defined by the average distance $\langle d_x^2 \rangle^{1/2}$ between an electron in the outer regions of a molecule and the exchange hole in the region of valence MOs. Our group has successfully used this procedure in other contexts,^{116,126–128} as a means to avoid the hassle associated with tuning based on eq 35 and also as a means to mitigate the dependence of ω -tuning on system size.¹¹⁶

The quantity C in eq 36 is an empirical parameter and has been optimized for use with LRC- ω PBE in several different basis sets.^{110,126} However, we find that the optimal value of C is sensitive to the fraction of SR-HFX that is used in the self-

consistent DFT calculation, thus eq 36 was reparameterized for use with $\alpha = 0.40$ because the corresponding hybrid PBE functional, which we call PBEh(40), has been demonstrated to afford good electronic energy levels for neat liquid water¹²⁹ and for $e^-(aq)$.¹³⁰

To generate a value of C that is appropriate for use with sLRC functionals with $\epsilon = 78.39$, we followed the least-squares fitting procedure of ref 110, except that instead of fitting to the gas-phase IE-tuning condition for small molecules, we instead fit to results from IE-tuned sLRC-PBEh(40). The fitting was performed under dielectric boundary conditions to capture changes in the density induced by a dielectric continuum; see Section 3 for details. Optimized values $C = 0.335$ and 0.470 were determined for the def2-SVP and def2-TZVPP basis sets,¹³¹ respectively. These values are considerably smaller than the values $C \approx 0.9$ that have been optimized previously for use with LRC- ω PBE(30) and LRC- ω PBEh(20),^{110,126} but this is consistent with the larger fraction of SR-HFX in the present case. (A smaller value of C means a smaller value of ω_{GDD} and thus a longer length scale for activating LR-HFX. This is presumably compensated by the larger admixture of SR-HFX.) The optimal values of C increase slightly when gas-phase tuning ($\alpha + \beta = 1$) is used instead, e.g., we determined $C = 0.560$ for use with LRC- ω PBEh(40)/def2-TZVPP. Only the sLRC- ω PBEh(40) values are used in the present work.

3. COMPUTATIONAL DETAILS

The $I^-(aq)$ calculations reported below are based on structures extracted from a periodic *ab initio* MD simulation performed using CP2K v. 6.1,¹³² for a $14.4 \times 14.4 \times 14.4 \text{ \AA}^3$ simulation cell consisting of $I^-(H_2O)_{96}$. The functional used in these simulations is PBEh(40)-rVV10, which includes the rVV10 nonlocal correlation functional,¹³³ which is a revision of the VV10 functional¹³⁴ that is designed for better efficiency in plane-wave basis sets. We use the parameter $b = 5.3$ in rVV10, as suggested in ref 135, because this functional has been demonstrated to yield a Kohn–Sham gap within <0.1 eV of the experimental band gap for liquid water.^{135,136} The TZV2P-GTH basis sets were used for water and DZVP-MOLOPT-SR-GTH basis sets for iodine.¹³⁷ The temperature was maintained at $T = 298$ K using canonical sampling with velocity rescaling.¹³⁸

Following 20 ps of equilibration, 30 snapshots were taken from the last 300 fs of dynamics, which were then used to obtain the $I^-(aq)$ spectra presented in Section 4.3. Although these snapshots are likely highly correlated, given the short production time scale, we will see that the number of water molecules included in the high-level subsystem fluctuates quite dramatically over just 300 fs. Since the primary purpose of our analysis in this initial report is to gauge the capacity of NCTA to prune the manifold of spurious low-lying CT states, these snapshots suffice.

As our analysis incurs a change of boundary conditions (from periodic boundaries in the CP2K simulation to a finite cluster with dielectric boundary conditions for the LR-TDDFT calculations), we collocate atoms belonging to each water molecule about the molecular center-of-mass prior to the LR-TDDFT calculations. In principle, this change of boundary conditions could introduce discontinuities when a water molecule is shifted by an entire box length, but in practice, the calculations presented here are not affected in any noticeable way, as will be apparent when we examine excitation energies along the MD trajectory. All LR-TDDFT calculations

were performed using a locally modified version of Q-Chem v. 5.3,¹³⁹ in which the theory described in Section 2 has been implemented.

To compare the results from our NCTA protocol to full-system LR-TDDFT methods, we must be able to identify the physically relevant CT states in the latter calculation without appeal to the analysis method introduced herein, in order to assess the veracity of the new technique. We accomplish this via NTO analysis, which allows us to find optically allowed CTTS states by searching the LR-TDDFT spectrum for states that feature a particle/hole excitation from the chromophore to the solvent. This entails computing (and sifting through) a large number of spurious states when semilocal functionals are employed, but NTOs provide a convenient way to identify the CTTS character and thus form the foundation of the benchmarking presented herein.

CT excitations are the focal point of the present work, and for quantitative accuracy, it will be important to incorporate the response of the dielectric medium in the LR-TDDFT excitation energies. For this, we use a nonequilibrium formulation dielectric continuum theory,⁸⁶ applying both the perturbative linear response (ptLR) and perturbative state-specific (ptSS) corrections,^{140,141} in conjunction with the Marcus–Brady–Carr partition of the fast and slow polarization charges.^{86,140} Whereas ptSS is a more conventional approach to electronic polarization in the excited state, it has been argued that the ptLR correction represents a form of solute–continuum dispersion interaction,⁸⁶ and the combination of ptLR and ptSS approaches has been found to yield good solvatochromic shifts at the LR-TDDFT level.¹⁴⁰ All SCRf calculations employ the integral equation form of the polarizable continuum model (IEF-PCM),^{86,142–144} with a solute/continuum interface that is defined by a solvent-accessible surface (SAS) with a solvent probe radius of 1.4 \AA .⁸⁶ The ptLR + ptSS correction modifies the LR-TDDFT excitation energies by 0.1 – 0.2 eV, which is small but not negligible on the scale of the agreement with the experiment that we report in Section 4.3.

4. RESULTS AND DISCUSSION

4.1. Stability with Respect to the Basis Set and Density Functional. As with any scheme that attempts to partition the virtual space into localized subsystems, a great deal of caution and testing is warranted regarding stability with respect to expansion of the AO basis set. We also wish to examine the extent to which the results depend on the choice of density functional approximation. In this section, we report calculations on a hydrated acrolein molecule, using a $C_3H_4O(H_2O)_{24}$ cluster whose geometry is taken from ref 67. We test the PBE,¹⁴⁵ SCAN,¹⁴⁶ and B3LYP^{147,148} functionals as examples of a GGA, a meta-GGA, and a global hybrid functional, respectively, to assess what low-level approximation might be used in projection-based embedding calculations. Convergence behavior with respect to the basis set is assessed using def2-SVPD, def2-TZVPD, and def2-QZVPD.^{131,149} All calculations reported in this section use a dense Euler–MacLaurin–Lebedev quadrature grid for the DFT calculations,¹⁵⁰ consisting of 99 radial points and 590 angular points per atom. Water molecules involved in the CTTS transition were selected based on a threshold that retains 99.8% of the CTTS acceptor orbital.

Results for nine combinations of functional and basis sets are shown in Figure 2. In each case, the water molecules selected

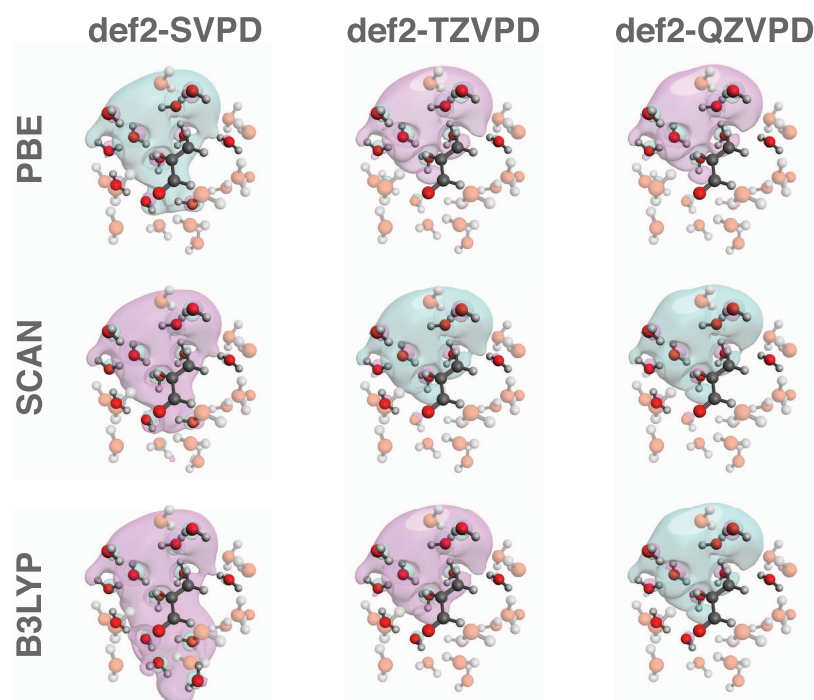


Figure 2. Basis set and functional dependence of the automated diabatization scheme for nuclear partitioning, as applied to $C_3H_4O(H_2O)_{24}$. The opaque atoms include the chromophore (C_3H_4O) along with the important water molecules as selected based on a threshold that retains 99.8% of the CTTS acceptor orbital. That orbital is shown, using an isocontour of 0.01 au, with green and purple representing different signs.

by our automated scheme appear only in the region of space highlighted by the CTTS acceptor orbital. The excited state obtained using the double- ζ basis set is more diffuse than the others, particularly in the case of B3LYP, but in all cases the CTTS acceptor orbital and the solvent selection converge at the triple- ζ level are unchanged in the quadruple- ζ basis set. Once this limit is reached (at the triple- ζ level), the partitions obtained using various density functionals differ by no more than one water molecule. The PBE functional selects one fewer water molecule as compared to SCAN, which selects the same high-level subsystem as B3LYP, suggesting that the partition is relatively stable regardless of density functional approximation.

For a more quantitative comparison, Table S1 reports the percentage of natural CT populations per solvent molecule for the lowest-lying CTTS state. By weighting the CTTS excitation space based on the eigenvalues of a first-order density matrix, we achieve a stable partition that does not change with the density functional *ansatz* except minimally at the double- ζ level, where the density itself is not yet converged.

4.2. Isolating the Physically Relevant CTTS State in Hydrated Acrolein. We next aim to quantify the accuracy of the high-level subsystem selection in the context of projection-based embedding. The CT problem in LR-TDDFT is especially pronounced in the $C_3H_4O(H_2O)_{24}$ test system introduced above. Even using B3LYP, the lowest $\pi \rightarrow \pi^*$ excitation of the chromophore is obscured by 62 CT states at lower energies and thus the ${}^1\pi\pi^*$ bright state appears as S_{64} in our calculations. (There is also a ${}^1n\pi^*$ state at low energy.) This is consistent with results reported for the same system in ref 67, where ALMO-based embedding was used to remove the spurious states. This comes at the price of also removing physically meaningful CTTS excitations, however.

At the B3LYP/jun-cc-pVDZ level, the first meaningful CTTS excitation in this system appears as state S_{11} , and while

not quite as difficult to find as the valence $\pi \rightarrow \pi^*$ transition, it is but one among tens of spurious states, making it difficult to isolate and analyze. We have confirmed through NTO analysis (Figure S1) that the S_{11} state is the first bright state featuring a particle/hole excitation from acrolein to the surrounding solvent. In the following, we explore the capacity of our method to prune the other CT states from the excitation space to test the hypothesis that this approach should retain the physically relevant and spectroscopically interesting CTTS states. All calculations in this section use the jun-cc-pVDZ basis set.¹⁵¹

To benchmark our partition of the relevant solvent molecules, we use the rCAM-B3LYP functional,¹⁵² a reparameterization of CAM-B3LYP¹¹⁷ intended to reduce many-electron self-interaction error. Figure 3 compares excitation energies computed using rCAM-B3LYP-in-B3LYP embedding to results obtained by applying rCAM-B3LYP to the full $C_3H_4O(H_2O)_{24}$ system. B3LYP is used as the low-level method in these calculations, so Figure 3 also shows the results of B3LYP applied to the entire system, for comparison. In contrast to rCAM-B3LYP, the B3LYP functional underestimates the $\pi \rightarrow \pi^*$ excitation energy by ≈ 0.6 eV and underestimates the CTTS excitation by almost 2.5 eV relative to rCAM-B3LYP results. The latter observation is unsurprising given the too-soft asymptotic decay of the B3LYP exchange–correlation potential, leading to underestimation of CT excitation energies.⁴⁴ This problem will worsen as the number of explicit water molecules is increased.⁴⁸

As the number of water molecules that are included in the high-level subsystem is increased, embedding results for the $\pi\pi^*$ state move somewhat closer to the rCAM-B3LYP benchmark, converging to 6.84 eV versus the benchmark value of 6.76 eV; see Figure 3a. In contrast, embedding results for the CTTS state converge precisely to the benchmark value.

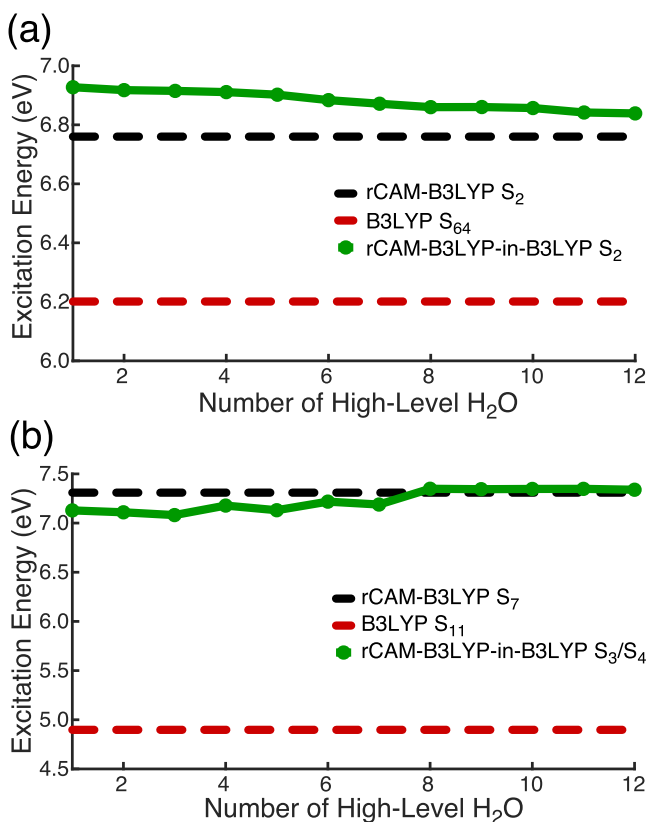


Figure 3. Convergence of excitation energies with respect to the number of water molecules in the high-level subsystem, for (a) the valence $\pi\pi^*$ state and (b) the lowest-lying CTTS state of the hydrated acrolein cluster $C_3H_4O(H_2O)_{24}$, using rCAM-B3LYP-in-B3LYP embedding. The benchmark in either case is a full-system calculation using rCAM-B3LYP. The full-system B3LYP result is also shown for comparison.

The slight discrepancy for the $\pi\pi^*$ state, which is really just a slower convergence toward the benchmark in that case, arises because our diabaticization scheme intentionally biases selection of H_2O molecules for the high-level subsystem toward description of the CTTS state. Some first-shell water molecules, which are important for an accurate description of the valence excited state, are therefore omitted. If higher accuracy is desired for localized valence excitations, this can be achieved by application of a radial cutoff to retain those nearby water molecules that affect the chromophore-localized states. A sensible procedure, in situations where both CTTS and valence excitations are of interest, is to retain the first hydration shell (as defined by a radial cutoff, say) in conjunction with automated diabaticization. We will do this for $\Gamma^-(aq)$ in Section 4.3.

The performance for CTTS excitations is the focus of the present work, and the data in Figure 3b demonstrate rapid convergence to the benchmark excitation energy of 7.31 eV. Including just 8 of 24 water molecules in the high-level subsystem affords an rCAM-B3LYP-in-B3LYP value of 7.34 eV. No further improvement is obtained with additional water molecules, but neither does this have deleterious consequences on the converged result. This demonstrates that the partition is not only accurate but also robust with respect to the inclusion of unimportant solvent molecules. Equally importantly, the embedding calculation preserves the correct state ordering, consistently predicting it to be $S_1(n\pi^*)$ followed by $S_2(\pi\pi^*)$.

By removing the solvent molecules that support spurious, low-lying CT states, the valence excitation spectrum of the chromophore is recovered with ease.

That said, even with rCAM-B3LYP, there remain problems with spurious states when gas-phase boundary conditions are used, as some of the frontier virtual orbitals on peripheral water molecules are not sufficiently destabilized by Pauli repulsion due to surrounding water molecules, as they would be in a condensed-phase environment. As a result, the gas-phase rCAM-B3LYP calculation predicts several excitations involving occupied/virtual pairs that are localized on peripheral water molecules. While RSH functionals may mitigate problems with spurious CT states, their use is not a panacea and does not always eliminate CT problems, especially in the presence of vacuum boundary conditions.^{48,49} In the present example, $C_3H_4O(H_2O)_{24}$ using rCAM-B3LYP, a CTTS state appears as S_7 , but the problem is certain to grow more severe as the cluster size increases. The natural CT partitioning scheme selects the high-level subsystem in a manner that admits far fewer of these states into the spectrum and eliminates what is otherwise a size-extensive growth in the number of CT states. The first CTTS state is isolated as S_3 (above only the $1n\pi^*$ and $1\pi\pi^*$ states) when ≤ 7 water molecules are included, at which point it shifts to S_4 as one peripheral water excitation appears in the low-energy spectrum. Isolating the physical CTTS states by removing the solvent molecules that contribute to the spurious ones not only clears up the valence excitation spectrum of the chromophore but also sidesteps the issue of spurious states associated with vacuum boundary conditions.

4.3. $\Gamma^-(aq)$ CTTS Spectrum. Aqueous iodide features one of the most well-studied CTTS spectra in condensed-phase photochemistry and is commonly used to generate hydrated electrons.²⁵ In this section, we apply our method to an *ab initio* MD trajectory of $\Gamma^-(aq)$ to study the lowest-lying CTTS band of the absorption spectrum in bulk solution. (Simulation details have been provided in Section 3.) The higher-level calculations in this section were performed at the sLRC- ω PBEh(40) level, which is a screened LRC functional with 40% SR-HFX, as described in Section 2.7. Lower-level calculations are performed using PBEh(40), consistent with the level of theory used for the periodic DFT simulation although we omit the nonlocal correlation functional rVV10 in the excited-state calculations. This is a more sophisticated and expensive level of theory than is typically used for the low-level calculations in projection-based embedding; however, we find that the density of states is simply not represented well enough to produce accurate spectra when a GGA functional is used for the low-level method. This sensitivity is understandable given that the $\Gamma^-(aq)$ CTTS band lies between the conduction and valence bands of liquid water.²⁵ If the manifold of states is misrepresented such that the CTTS band lies within the conduction band, then low-lying CTTS states may mix with states in the conduction band and yield poor excitation energies and a poor representation of the character of the CTTS state as a whole.⁸⁵ The *ab initio* MD simulations use the global hybrid PBEh(40) functional precisely because it seems to reproduce the band gap of bulk water,^{135,136} and the computational expense of the low-level part of the embedding calculations is consistent with this choice.

For these calculations, we use continuum boundary conditions (IEF-PCM and SAS cavity) around the entire $\Gamma^-(H_2O)_{96}$ unit cell, leading to a multilayer embedding scheme with sLRC- ω PBEh(40) and PBEh(40) as the high- and low-

level methods, respectively, for describing the explicit solvent and continuum boundary conditions as described in Section 2.2. The high-level subsystem is selected based on diabatization of the PBEh(40) molecular orbitals. To incorporate strongly hydrating water molecules that may not be selected automatically by our diabatization scheme, all water molecules within a 4.0 Å radial cutoff from the iodine atom are also included in the high-level subsystem. In all cases, the iodine atom was treated as the donor species and all water molecules as the acceptor. All calculations in this section employ the def2-SVP basis and the corresponding effective core potential for iodine.

Our results are superimposed with experimental data for the lowest CTTS band of $I^-(aq)$ in Figure 4. The computed

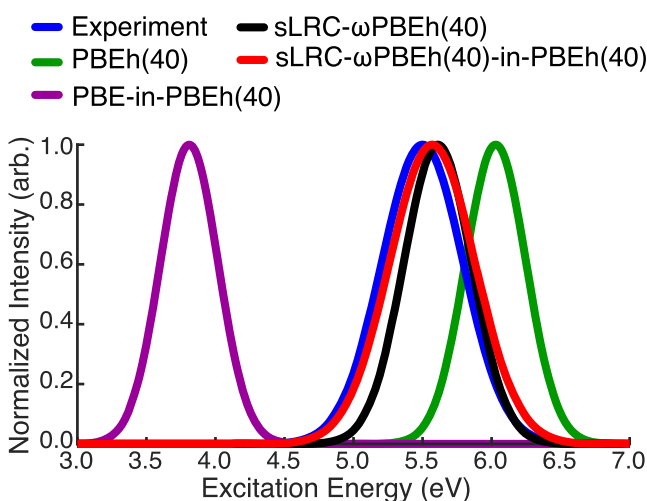


Figure 4. Comparison between experiment, full-system sLRC- ω PBEh(40), full-system PBEh(40), and projection-based embedding for the first CTTS band of $I^-(aq)$, modeled as $I^-(H_2O)_6$ with continuum boundary conditions. Experimental data from spectral deconvolution of the $I^-(aq)$ ultraviolet absorption spectrum at 298 K are taken from ref 153.

spectra are obtained from 30 snapshots taken from the condensed-phase simulation and plotted in Figure 4 as Gaussian distributions with 2σ widths, where σ is the standard deviation of the 30 vertical excitation energies. (Peak heights are normalized to unity.) The full-system sLRC- ω PBEh(40) calculation for $I^-(H_2O)_6$ is in quantitative agreement with the experiment,¹⁵³ with a computed band maximum at 5.6 versus 5.5 eV in the experiment. The embedding calculation is faithful to the full-system calculation to within $\lesssim 0.1$ eV.

In comparison, the PBEh(40) absorption band (obtained by applying the low-level method to the full system) is centered at 6.0 eV and thus overestimates the experimental CTTS band maximum by 0.5 eV. Whereas CT excitation energies are typically underestimated in gas-phase calculations using global hybrid functionals, in the condensed phase where dielectric stabilization is important, this is not always the case.¹¹⁸ Here, the behavior may be more similar to what is typically observed for localized valence excitations, where functionals with a large fraction of HFX tend to overestimate excitation energies.⁴¹

Note that SR-HFX cannot be eliminated from the low-level calculation, as demonstrated by the failure of the PBE-in-PBEh(40) results in Figure 4. Use of the GGA functional as the low-level method results in underestimation of the CTTS spectrum by nearly 2 eV despite the use of a global hybrid

functional for the high-level subsystem. While PBE-in-PBEh(40) is not a useful level of theory for this particular application, it is interesting to note that the use of the partitioning method developed here does result in the CTTS excitation appearing among the first 10 excited states in each of the 30 snapshots used to generate the PBE-in-PBEh(40) spectrum. It is remarkable even that we are able to obtain GGA results for such a large system, and this is only feasible because the density of states has been pruned by the diabatization algorithm.

Overall, we take the accuracy of sLRC- ω PBEh(40) relative to the experiment as a testament to the quality of our screened GDD tuning approach. The fact that the embedding calculation is faithful to the full-system LR-TDDFT calculation is a validation of our procedure for subsystem selection based on automatic diabatization.

4.4. Sensitivity of Excitation Energies to Subsystem Partition. Whenever embedding methods are applied, an artificial interface is introduced between the high- and low-level subsystems. This is a problem shared by QM/MM methods; however, QM/QM embedding (like any other active-space method) is also subject to discontinuities as the atoms move and the orbitals selected for the subsystem may change. In the present case, our partition based on diabatization need not select the same water molecules at every snapshot along a trajectory. Here, we quantify the magnitude and significance of the errors associated with changing the character of this high-level/low-level interface.

Results plotted in Figure 5a–c show that the errors introduced by embedding (as gauged by comparison to a calculation that applies the higher level of theory to the full system) are smaller than the thermal fluctuations in the excitation energy. This is despite the fact that the number of water molecules that is included in the high-level subsystem fluctuates quite a bit from one snapshot to the next, as shown in Figure 5d. (On average, 37 ± 12 out of 96 water molecules are included in the high-level subsystem. This number is chosen automatically by the diabatization algorithm.) Errors are slightly different for various *a posteriori* corrections to the excitation energy based on nonequilibrium dielectric continuum theory,⁸⁶ but in each case, these differences fall well within the envelope of thermal fluctuations. Average embedding errors are <0.1 eV regardless of the continuum treatment that is used.

Perturbation theory energies are strongly dependent on the zeroth-order wave function, so one might wonder whether small changes in the character of the excited states, arising from the embedding procedure, might be amplified by the pLR and/or ptSS continuum corrections that are applied to the excitation energies, but in fact these differences remain <0.1 eV. Because the errors remain small despite the application of two different perturbations, it would appear that embedding does not impose dramatic changes in the character of the excited-state wave function.

These results are promising for the future possibility of developing analytic gradients to use in conjunction with this particular partitioning scheme, which would facilitate excited-state dynamics in the condensed phase. In the immediate term, these results serve to validate the application of this cost-effective approach to analyze vertical excitation spectra of solvated chromophores in fully-QM treatments of aqueous systems.

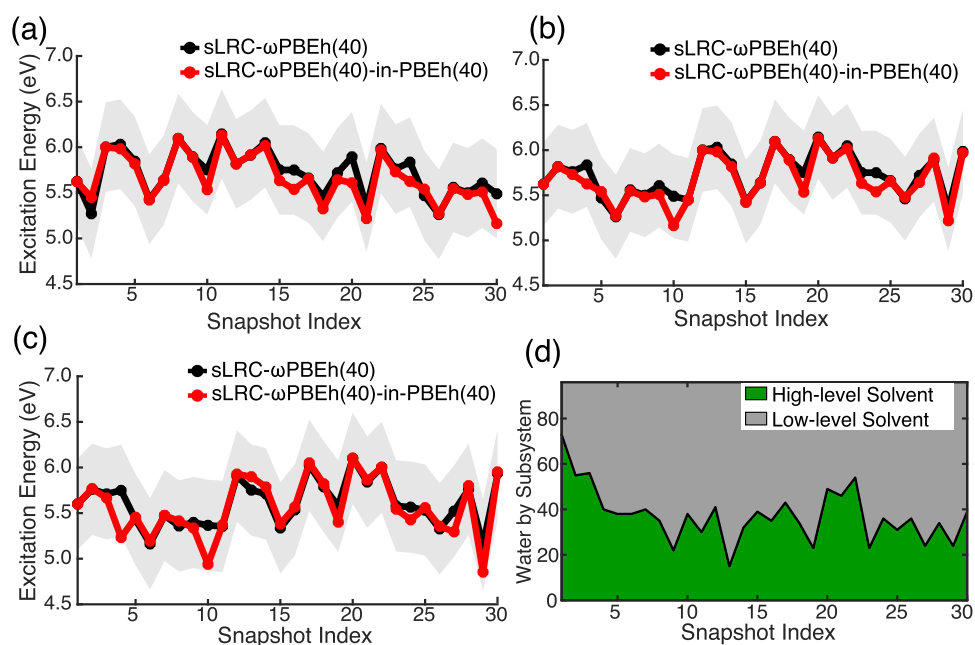


Figure 5. Lowest CTTS excitation energy of $\text{I}^-(\text{H}_2\text{O})_{96}$, comparing sLRC- ω PBEh(40)-in-PBEh(40) embedding results to full-system sLRC- ω PBEh(40) calculations: (a) zeroth-order excitation energies, based on solvent-polarized orbitals with IEF-PCM boundary conditions; (b) excitation energies computed using the ptLR (nonequilibrium continuum) correction only; (c) excitation energies obtained upon applying both the ptLR and ptSS corrections. The gray shaded region in each of (a)–(c) represents thermal fluctuations of the benchmark (full-system) excitation energy and demonstrates that errors arising from embedding are considerably smaller than these fluctuations. Panel (d) shows the number of water molecules in either subsystem, demonstrating that this number exhibits large fluctuations from one snapshot to the next.

5. CONCLUSIONS

We have reported an automated subsystem partition for use with projection-based embedding to study electronic spectroscopy in an aqueous environment at the LR-TDDFT level. Our procedure, natural CT analysis, introduces natural CT orbitals constructed as a simple extension of a projection-operator diabatization developed for electron-transfer problems.⁶⁸ These orbitals provide a minimal excitation space, within which an LR-TDDFT calculation is performed (at negligible computational cost) within a diabatic representation, furnishing a spectrum of CT excitations by construction. Any CT state of interest can then be projected onto the surrounding solvent to determine the contribution of each solvent molecule to the overall CT acceptor space. Solvent molecules that are assigned to the CT acceptor space can be included in the high-level subsystem in a subsequent projection-based embedding calculation. Other CT states, which might be artifacts of LR-TDDFT and therefore not of interest, are excluded by construction.

We have demonstrated that this approach is stable with respect to the choice of density functional approximation and rapidly convergent with respect to the choice of the AO basis set. This stability permits on-the-fly evaluation of the CT acceptor space based on the orbitals obtained from an *ab initio* MD trajectory, with the density functional that is used in the dynamics serving as the low-level method for subsequent embedding calculations of the vertical excitation spectrum. The automated partition removes spurious low-lying CT excitations (which normally plague solution-phase LR-TDDFT calculations) by removing the solvent orbitals that support those excitations. This serves to reveal the valence excitations localized on the chromophore, which may otherwise be obscured or even “dissolved” (via intensity borrowing) in a sea of spurious CT excitations.⁴⁸ At the same time, we can

preserve the appearance of physically meaningful CTTS states that have typically also been removed by other procedures designed to avoid the appearance of artificial CT transitions. Artifacts of vacuum boundary conditions are also removed automatically by our procedure.

We have applied this approach to the prediction of the low-lying CTTS band of $\text{I}^-(\text{aq})$, in conjunction with *ab initio* MD simulations. Quantitative agreement with experiment is obtained, and the errors introduced by embedding are consistently smaller than thermal fluctuations in the excitation energy itself. This suggests that the high-level/low-level partition that is used in the embedding calculations provides stable excitation energies despite the fact that the automated diabatization procedure selects different numbers of important water molecules at different snapshots along the MD trajectory.

Spurious, low-energy CT states are the bane of LR-TDDFT calculations in large systems.^{48,49,53} As noted recently in the context of organic photovoltaic materials,¹¹⁸ however, once these artificial interlopers are removed from the spectrum, it need no longer be the case that the genuine CT excitation energies are underestimated by LR-TDDFT. In condensed media, and in particular in the aqueous systems considered here, dielectric screening needs to be considered in the range-separation *ansatz*.¹²²

An ancillary result of this work is the demonstration that screened LRC functionals based on “ ω_{GDD} tuning”^{110,116,126} provide an accurate and convenient means to access condensed-phase electronic spectroscopy, eliminating the computational overhead and system size-dependence associated with the widely used optimal tuning procedure based on the IE criterion (eq 35).⁵⁸ Together, these developments open a path for simulation of aqueous-phase photochemistry by

combining *ab initio* MD with LR-TDDFT by means of projection-based embedding.

■ APPENDIX: PROOF OF BIJECTION BETWEEN DIABATIC AND SPADE REPRESENTATIONS

For a given subspace, the SPADE MOs are defined by

$$\mathbf{C}_{\text{SPADE}}^A = \mathbf{C}^A \mathbf{R}^{A,\text{span}} \quad (\text{A1})$$

where $\mathbf{R}^{A,\text{span}}$ is the span of the right-hand singular vectors from eq 10. Alternatively, this can be written as

$$C_{\mu\lambda}^{A,\text{SPADE}} = \sum_r \sum_{\mu \in A} |\mu\rangle \langle \mu| \bar{C}_{\mu r} R_{r\lambda} \quad (\text{A2})$$

where the $\bar{C}_{\mu r}$ are generic MO coefficients. For the duration of this discussion, both of these coefficients and all AO basis functions are presumed to be represented in the Löwdin-orthogonalized basis. Consider the projections $\langle \phi_{\text{diab}} | \psi_{\text{SPADE}} \rangle$ of SPADE MOs onto diabatic MOs. Elements of the corresponding overlap matrix are given by

$$O_{\sigma\lambda} = \sum_r \sum_{\mu \in A} \bar{U}_{\sigma\nu} \langle \nu | \mu \rangle \bar{C}_{\mu r} R_{r\lambda} = \sum_r \sum_{\mu \in A} \bar{U}_{\sigma\mu} \bar{C}_{\mu r} R_{r\lambda} \quad (\text{A3})$$

The orthogonal projector onto any of the diabatic MO subspaces is written in the form of a density matrix

$$\hat{P}_{\text{diab}} = \sum_{\mu \in A} |\mu\rangle U_{\mu r} U_{r\nu}^* \langle \nu| = \sum_{\mu \in A} |\mu\rangle P_{\mu\nu} \langle \nu| \quad (\text{A4})$$

where $U_{\mu r}$ are the diabatic MO coefficients for subspace A. Application of this projection operator to the SPADE MOs reveals a bijective mapping

$$\begin{aligned} P_{\text{diab}} \mathbf{C}_{\text{SPADE}}^A &= \sum_{\mu \in A} |\mu\rangle P_{\mu\nu} \langle \nu | \mu \rangle \langle \nu | \bar{C}_{\mu r} R_{r\lambda} \\ &= \sum_{\mu \in A} |\mu\rangle P_{\mu\mu} \langle \mu | \bar{C}_{\mu r} R_{r\lambda} \end{aligned} \quad (\text{A5})$$

This implies that for any column in the SPADE MO coefficient matrix (for a given diabatic subspace), the corresponding column of the diabatic MO coefficient matrix \mathbf{U} has a one-to-one and onto correspondence. This correspondence allows us to directly select SPADE MOs based on the selected column indices of the diabatic MO coefficient matrix.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00412>.

Coordinates for the $\text{C}_3\text{H}_4\text{O}(\text{H}_2\text{O})_{24}$ and $\text{I}^-(\text{H}_2\text{O})_{96}$ snapshots used in the LR-TDDFT calculations and NTO images and oscillator strength data for $\text{C}_3\text{H}_4\text{O}(\text{H}_2\text{O})_{24}$ state assignments (PDF)
Geometries (XYZ)

■ AUTHOR INFORMATION

Corresponding Authors

Kevin Carter-Fenk – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0000-0001-8302-4750; Email: carter-fenk.2@osu.edu

John M. Herbert – Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; orcid.org/0000-0002-1663-2278; Email: herbert@chemistry.ohio-state.edu

Author

Christopher J. Mundy – Physical Science Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States; Department of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States; orcid.org/0000-0003-1378-5241

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.1c00412>

Notes

The authors declare the following competing financial interest(s): J.M.H. serves on the board of directors of Q-Chem Inc.

■ ACKNOWLEDGMENTS

Work by K.C.-F. and J.M.H. was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award no. DE-SC0008550. C.J.M. is supported by the U.S. DOE, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. K.C.-F. acknowledges a Presidential Fellowship from The Ohio State University, as well as additional support from the DOE Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under contract no. DE-SC0014664. All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of DOE, ORAU, or ORISE. Calculations were performed on both Pacific Northwest National Laboratory's Institutional Computing resource and the National Energy Research Scientific Computing Center (NERSC). NERSC is a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract no. DE-AC02-05CH11231.

■ REFERENCES

- (1) Odobel, F.; le Pleux, L.; Pellegrin, Y.; Blart, E. New photovoltaic devices based on the sensitization of p-type semiconductors: Challenges and opportunities. *Acc. Chem. Res.* **2010**, *43*, 1063–1071.
- (2) Kurumisawa, Y.; Higashino, T.; Nimura, S.; Tsuji, Y.; Iiyama, H.; Imahori, H. Renaissance of fused porphyrins: Substituted methylene-bridged thiophene-fused strategy for high-performance dye-sensitized solar cells. *J. Am. Chem. Soc.* **2019**, *141*, 9910–9919.
- (3) Willems, R. E. M.; Meskers, S. C. J.; Wienk, M. M.; Janssen, R. A. J. Effect of charge-transfer state energy on charge generation efficiency via singlet fission in pentacene-fullerene solar cells. *J. Phys. Chem. C* **2019**, *123*, 10253–10261.
- (4) Zhu, X.-Y.; Yang, Q.; Muntwiler, M. Charge-transfer excitons at organic semiconductor surfaces and interfaces. *Acc. Chem. Res.* **2009**, *42*, 1779–1787.
- (5) Dreu, A.; Fleming, G. R.; Head-Gordon, M. Charge-transfer state as a possible signature of a zeaxanthin-chlorophyll dimer in the non-photochemical quenching process in green plants. *J. Phys. Chem. B* **2003**, *107*, 6500–6503.

- (6) Ahn, T. K.; Avenson, T. J.; Ballottari, M.; Cheng, Y.; Niyogi, K. K.; Bassi, R.; Fleming, G. R. Architecture of a charge-transfer state regulating light harvesting in a plant antenna protein. *Science* **2008**, *320*, 794–797.
- (7) Cupellini, L.; Calvani, D.; Jacquemin, D.; Mennucci, B. Charge transfer from the carotenoid can quench chlorophyll excitation in antenna complexes of plants. *Nat. Commun.* **2020**, *11*, No. 662.
- (8) Hammarström, L. Accumulative charge separation for solar fuels production: Coupling light-induced single electron transfer to multielectron catalysis. *Acc. Chem. Res.* **2015**, *48*, 840–850.
- (9) Staib, A.; Borgis, D. Reaction pathways in the photodetachment of an electron from aqueous chloride: A quantum molecular dynamics study. *J. Chem. Phys.* **1996**, *104*, 9027–9039.
- (10) Sheu, W.-S.; Rossky, P. J. Electronic and solvent relaxation dynamics of a photoexcited aqueous halide. *J. Phys. Chem. A* **1996**, *100*, 1295–1302.
- (11) Kloepfer, J. A.; Vilchiz, V. H.; Lenchenkov, V. A.; Bradforth, S. E. Femtosecond dynamics of photodetachment of the iodide anion in solution: Resonant excitation into the charge-transfer-to-solvent state. *Chem. Phys. Lett.* **1998**, *298*, 120–128.
- (12) Barthel, E. R.; Martini, I. B.; Keszler, E.; Schwartz, B. J. Solvent effects on the ultrafast dynamics and spectroscopy of the charge-transfer-to-solvent reaction of sodide. *J. Chem. Phys.* **2003**, *118*, 5916–5931.
- (13) Iglev, H.; Trifonov, A.; Thaller, A.; Buchvarov, I.; Fiebig, T.; Laubereau, A. Photoionization dynamics of an aqueous iodide solution: The temperature dependence. *Chem. Phys. Lett.* **2005**, *403*, 198–204.
- (14) Moskun, A. C.; Bradforth, S. E.; Thøgersen, J.; Keiding, S. Absence of a signature of aqueous $I(^2P_{1/2})$ after 200-nm photodetachment of $I^-(aq)$. *J. Phys. Chem. A* **2006**, *110*, 10947–10955.
- (15) Shoshana, O.; Lustres, J. L. P.; Ernsting, N. P.; Ruhman, S. Mapping CTTs dynamics of Na^- in tetrahydrofuran with ultrafast multichannel pump-probe spectroscopy. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2599–2609.
- (16) Tang, Y.; Suzuki, Y.; Shen, H.; Sekiguchi, K.; Kurahashi, N.; Nishizawa, K.; Zuo, P.; Suzuki, T. Time-resolved photoelectron spectroscopy of bulk liquids at ultra-low kinetic energy. *Chem. Phys. Lett.* **2010**, *494*, 111–116.
- (17) Lübcke, A.; Buchner, F.; Heine, N.; Hertel, I. V.; Schultz, T. Time-resolved photoelectron spectroscopy of solvated electrons in aqueous NaI solution. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14629–14634.
- (18) Pham, V.-T.; Penfold, T. J.; van der Veen, R. M.; Lima, F.; Nahhas, A. E.; Johnson, S. L.; Beaud, P.; Abela, R.; Bressler, C.; Tavernelli, I.; Milne, C. J.; Chergui, M. Probing the transition from hydrophilic to hydrophobic solvation with atomic scale resolution. *J. Am. Chem. Soc.* **2011**, *133*, 12740–12748.
- (19) Messina, F.; Bräm, O.; Cannizzo, A.; Chergui, M. Real-time observation of the charge transfer to solvent dynamics. *Nat. Commun.* **2013**, *4*, No. 2119.
- (20) Nowakowski, P. J.; Woods, D. A.; Verlet, J. R. R. Charge transfer to solvent dynamics at the ambient water/air interface. *J. Phys. Chem. Lett.* **2016**, *7*, 4079–4085.
- (21) Shoji, M.; Kaniwa, K.; Hiranuma, Y.; Maselli, O.; Mafuné, F. Solvation structure of I^- and Na^+ on the surface of NaI aqueous solution studied by photodetachment spectroscopy in combination with mass spectrometry. *J. Phys. Chem. A* **2011**, *115*, 2148–2154.
- (22) Rizzuto, A. M.; Irgen-Gioro, S.; Eftekhari-Bafrooei, A.; Saykally, R. J. Broadband deep UV spectra of interfacial aqueous iodide. *J. Phys. Chem. Lett.* **2016**, *7*, 3882–3885.
- (23) Mizuno, H.; Rizzuto, A. M.; Saykally, R. J. Charge-transfer-to-solvent spectrum of thiocyanate at the air/water interface measured by broadband deep ultraviolet electronic sum frequency generation spectroscopy. *J. Phys. Chem. Lett.* **2018**, *9*, 4753–4757.
- (24) Bhattacharyya, D.; Mizuno, H.; Rizzuto, A. M.; Zhang, Y.; Saykally, R. J.; Bradforth, S. E. New insights into the charge-transfer-to-solvent spectrum of aqueous iodide: Surface versus bulk. *J. Phys. Chem. Lett.* **2020**, *11*, 1656–1661.
- (25) Chen, X.; Bradforth, S. E. The ultrafast dynamics of photodetachment. *Annu. Rev. Phys. Chem.* **2008**, *59*, 203–231.
- (26) Herbert, J. M.; Coons, M. P. The hydrated electron. *Annu. Rev. Phys. Chem.* **2017**, *68*, 447–472.
- (27) Jacobson, L. D.; Herbert, J. M. Polarization-bound quasi-continuum states are responsible for the “blue tail” in the optical absorption spectrum of the aqueous electron. *J. Am. Chem. Soc.* **2010**, *132*, 10000–10002.
- (28) Jacobson, L. D.; Herbert, J. M. A one-electron model for the aqueous electron that includes many-body electron-water polarization: Bulk equilibrium structure, vertical electron binding energy, and optical absorption spectrum. *J. Chem. Phys.* **2010**, *133*, No. 154506.
- (29) Herbert, J. M.; Jacobson, L. D. Nature’s most squishy ion: The important role of solvent polarization in the description of the hydrated electron. *Int. Rev. Phys. Chem.* **2011**, *30*, 1–48.
- (30) Herbert, J. M. Structure of the aqueous electron. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20538–20565.
- (31) Alizadeh, E.; Sanche, L. Precursors of solvated electrons in radiobiological physics and chemistry. *Chem. Rev.* **2012**, *112*, 5578–5602.
- (32) Boudaïffa, B.; Cloutier, P.; Hunting, D.; Huels, M. A.; Sanche, L. Resonance formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science* **2000**, *287*, 1658–1661.
- (33) Falcone, J. M.; Becker, D.; Sevilla, M. D.; Swarts, S. G. Products of the reactions of the dry and aqueous electron with hydrated DNA: Hydrogen and 5,6-dihydropyrimidines. *Radiat. Phys. Chem.* **2005**, *72*, 257–264.
- (34) Alizadeh, E.; Sanz, A. G.; García, G.; Sanche, L. Radiation damage to DNA: The indirect effect of low-energy electrons. *J. Phys. Chem. Lett.* **2013**, *4*, 820–825.
- (35) Alizadeh, E.; Orlando, T. M.; Sanche, L. Biomolecular damage induced by ionizing radiation: The direct and indirect effects of low-energy electrons on DNA. *Annu. Rev. Phys. Chem.* **2015**, *66*, 379–398.
- (36) Herbert, J. M. The Quantum Chemistry of Loosely-bound Electrons. In *Reviews in Computational Chemistry*; Parill, A. L.; Lipkowitz, K., Eds.; Wiley-VCH: Hoboken, NJ, 2015; Vol. 28, Chapter 8, pp 391–517.
- (37) Subotnik, J. E. Configuration interaction singles has a large systematic bias against charge-transfer states. *J. Chem. Phys.* **2011**, *135*, No. 071104.
- (38) Furche, F. On the density matrix based approach to time-dependent density functional response theory. *J. Chem. Phys.* **2001**, *114*, 5982–5992.
- (39) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.
- (40) Elliott, P.; Furche, F.; Burke, K. Excited states from time-dependent density functional theory. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Cundari, T. R., Eds.; Wiley-VCH: New York, 2009; Vol. 26, Chapter 3, pp 91–165.
- (41) Laurent, A. D.; Jacquemin, D. TD-DFT benchmarks: A review. *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.
- (42) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. Does density functional theory contribute to the understanding of excited states of unsaturated organic compounds? *Mol. Phys.* **1999**, *97*, 859–868.
- (43) Casida, M. E.; Gutierrez, F.; Guan, J.; Gadea, F.-X.; Salahub, D.; Daudey, J.-P. Charge-transfer correction for improved time-dependent local density approximation excited-state potential energy curves: Analysis within the two-level model with illustration for H_2 and LiH . *J. Chem. Phys.* **2000**, *114*, 7062–7071.
- (44) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* **2003**, *119*, 2943–2946.
- (45) Dreuw, A.; Head-Gordon, M. Failure of time-dependent density functional theory for long-range charge-transfer excited-states:

The zincbacteriochlorin-bacteriochlorin and bacteriochlorophyll-spheroidene complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

(46) Bernasconi, L.; Sprik, M.; Hutter, J. Hartree-Fock exchange in time dependent density functional theory: Application to charge transfer excitations in solvated molecular systems. *Chem. Phys. Lett.* **2004**, *394*, 141–146.

(47) Neugebauer, J.; Gritsenko, O.; Baerends, E. J. Assessment of a simple correction for the long-range charge-transfer problem in time-dependent density functional theory. *J. Chem. Phys.* **2006**, *124*, No. 214102.

(48) Lange, A.; Herbert, J. M. Simple methods to reduce charge-transfer contamination in time-dependent density-functional calculations of clusters and liquids. *J. Chem. Theory Comput.* **2007**, *3*, 1680–1690.

(49) Isborn, C. M.; Mar, B. D.; Curchod, B. F. E.; Tavernelli, I.; Martínez, T. J. The charge transfer problem in density functional theory calculations of aqueously solvated molecules. *J. Phys. Chem. B* **2013**, *117*, 12189–12201.

(50) Magyar, R. J.; Tretiak, S. Dependence of spurious charge-transfer excited states on orbital exchange in TDDFT: Large molecules and clusters. *J. Chem. Theory Comput.* **2007**, *3*, 976–987.

(51) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.* **2008**, *128*, No. 044118.

(52) Lange, A. W.; Rohrdanz, M. A.; Herbert, J. M. Charge-transfer excited states in a π -stacked adenine dimer, as predicted using long-range-corrected time-dependent density functional theory. *J. Phys. Chem. B* **2008**, *112*, 6304–6308. Erratum: *J. Phys. Chem. B* **2008**, *112*, 7345.

(53) Lange, A. W.; Herbert, J. M. Both intra- and interstrand charge-transfer excited states in B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states. *J. Am. Chem. Soc.* **2009**, *131*, 3913–3922.

(54) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range corrected time-dependent density functional theory. *J. Chem. Phys.* **2004**, *120*, 8425–8433.

(55) Rohrdanz, M. A.; Herbert, J. M. Simultaneous benchmarking of ground- and excited-state properties with long-range-corrected density functional theory. *J. Chem. Phys.* **2008**, *129*, No. 034107.

(56) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.* **2009**, *130*, No. 054112.

(57) Stein, T.; Kronik, L.; Baer, R. Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory. *J. Am. Chem. Soc.* **2009**, *131*, 2818–2820.

(58) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

(59) Zhang, C.-R.; Sears, J. S.; Yang, B.; Aziz, S. G.; Coropceanu, V.; Brédas, J.-L. Theoretical study of the local and charge-transfer excitations in model complexes of pentacene- C_{60} using tuned range-separated hybrid functionals. *J. Chem. Theory Comput.* **2014**, *10*, 2379–2388.

(60) Yang, B.; Yi, Y.; Zhang, C.-R.; Aziz, S. G.; Coropceanu, V.; Brédas, J.-L. Impact of electron delocalization on the nature of the charge-transfer states in model pentacene/ C_{60} interfaces: A density functional theory study. *J. Phys. Chem. C* **2014**, *118*, 27648–27656.

(61) Kümmel, S. Charge-transfer excitations: A challenge for time-dependent density functional theory that has been met. *Adv. Energy Mater.* **2017**, *7*, No. 1700440.

(62) Closser, K. D.; Ge, Q.; Mao, Y.; Shao, Y.; Head-Gordon, M. Superposition of fragment excitations for excited states of large clusters with application to helium clusters. *J. Chem. Theory Comput.* **2015**, *11*, 5791–5803.

(63) Liu, J.; Herbert, J. M. An efficient and accurate approximation to time-dependent density functional theory for systems of weakly coupled monomers. *J. Chem. Phys.* **2015**, *143*, No. 034106.

(64) Liu, J.; Herbert, J. M. Local excitation approximations to time-dependent density functional theory for excitation energies in solution. *J. Chem. Theory Comput.* **2016**, *12*, 157–166.

(65) Herbert, J. M.; Zhang, X.; Morrison, A. F.; Liu, J. Beyond time-dependent density functional theory using only single excitations: Methods for computational studies of excited states in complex systems. *Acc. Chem. Res.* **2016**, *49*, 931–941.

(66) Ge, Q.; Mao, Y.; White, A. F.; Epifanovsky, E.; Closser, K. D.; Head-Gordon, M. Simulating the absorption spectra of helium clusters ($N = 70, 150, 231, 300$) using a charge transfer correction to superposition of fragment single excitations. *J. Chem. Phys.* **2017**, *146*, No. 044111.

(67) Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. Absolutely localized projection-based embedding for excited states. *J. Chem. Theory Comput.* **2020**, *16*, 385–398.

(68) Kondov, I.; Čížek, M.; Benesch, C.; Wang, H.; Thoss, M. Quantum dynamics of photoinduced electron-transfer reactions in dye-semiconductor systems: First-principles description and application to coumarin 343- TiO_2 . *J. Phys. Chem. C* **2007**, *111*, 11970–11981.

(69) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

(70) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, No. 18A507.

(71) Lee, S. J. R.; Welborn, M.; Manby, F. R.; Miller, T. F. Projection-based wavefunction-in-DFT embedding. *Acc. Chem. Res.* **2019**, *52*, 1359–1368.

(72) Wasserman, A.; Pavanello, M. Quantum embedding electronic structure methods. *Int. J. Quantum Chem.* **2020**, *120*, No. e26495.

(73) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z. F.; Liu, F. Y.; Li, H. B.; Ding, L. N.; Morokuma, K. The ONIOM method and its applications. *Chem. Rev.* **2015**, *115*, 5678–5796.

(74) Libisch, F.; Huang, C.; Carter, E. A. Embedded correlated wavefunction schemes: Theory and applications. *Acc. Chem. Res.* **2014**, *47*, 2768–2775.

(75) Yu, K.; Krauter, C. M.; Dieterich, J. M.; Carter, E. A. Density and Potential Functional Embedding: Theory and Practice. In *Fragmentation: Toward Accurate Calculations on Complex Molecular Systems*; Gordon, M. S., Ed.; Wiley: Hoboken, 2017; Chapter 2, pp 81–118.

(76) Wesolowski, T. One-electron Equations for Embedded Electron Density: Challenge for Theory and Practical Payoffs in Multi-level Modelling of Complex Polyatomic Systems. In *Computational Chemistry: Reviews of Current Trends: Volume 10*; Leszczynski, J., Ed.; World Scientific: Hackensack, NJ, 2006.

(77) Gomes, A. S. P.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2012**, *108*, 222–277.

(78) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* **2015**, *115*, 5891–5928.

(79) Jacob, C. R.; Neugebauer, J. Subsystem density-functional theory. *WIREs Comput. Mol. Sci.* **2014**, *4*, 325–362.

(80) Claudino, D.; Mayhall, N. J. Automatic partition of orbital spaces based on singular value decomposition in the context of embedding theories. *J. Chem. Theory Comput.* **2019**, *15*, 1053–1064.

(81) Chulhai, D. V.; Goodpaster, J. D. Improved accuracy and efficiency in quantum embedding through absolute localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.

(82) Claudino, D.; Mayhall, N. J. Simple and efficient truncation of virtual spaces in embedded wave functions via concentric localization. *J. Chem. Theory Comput.* **2019**, *15*, 6085–6096.

- (83) Serxner, D.; Dessent, C. E. H.; Johnson, M. A. Precursor of the Γ_{aq} charge-transfer-to-solvent (CTTS) band in $\Gamma^-(\text{H}_2\text{O})_n$ clusters. *J. Chem. Phys.* **1996**, *105*, 7231–7234.
- (84) Majumdar, D.; Kim, J.; Kim, K. S. Charge transfer to solvent (CTTS) energies of small $\text{X}^-(\text{H}_2\text{O})_{n=1-4}$ ($\text{X} = \text{F}, \text{Cl}, \text{Br}, \text{I}$) clusters: Ab initio study. *J. Chem. Phys.* **2000**, *112*, 101–105.
- (85) Bradforth, S. E.; Jungwirth, P. Excited states of iodide anions in water: A comparison of the electronic structure in clusters and in bulk solution. *J. Phys. Chem. A* **2002**, *106*, 1286–1298.
- (86) Herbert, J. M. Dielectric continuum methods for quantum chemistry. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1519.
- (87) Coons, M. P.; You, Z.-Q.; Herbert, J. M. The hydrated electron at the surface of neat liquid water appears to be indistinguishable from the bulk species. *J. Am. Chem. Soc.* **2016**, *138*, 10879–10886.
- (88) Coons, M. P.; Herbert, J. M. Quantum chemistry in arbitrary dielectric environments: Theory and implementation of nonequilibrium Poisson boundary conditions and application to compute vertical ionization energies at the air/water interface. *J. Chem. Phys.* **2018**, *148*, No. 222834. Erratum: *J. Chem. Phys.* **2019**, *151*, 189901.
- (89) Stein, C. J.; Herbert, J. M.; Head-Gordon, M. The Poisson-Boltzmann model for implicit solvation of electrolyte solutions: Quantum chemical implementation and assessment via Sechenov coefficients. *J. Chem. Phys.* **2019**, *151*, No. 224111.
- (90) Paul, S. K.; Herbert, J. M. Probing interfacial effects on ionization energies: The surprising banality of anion-water hydrogen bonding at the air/water interface. *J. Am. Chem. Soc.* DOI: 10.26434/chemrxiv.14273651, in press.
- (91) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- (92) Marcus, R. A. On the theory of oxidation-reduction reactions involving electron transfer. I. *J. Chem. Phys.* **1956**, *24*, 966–978.
- (93) Hush, N. S. Adiabatic theory of outer sphere electron-transfer reactions in solution. *Trans. Faraday Soc.* **1961**, *57*, 557–580.
- (94) Marcus, R. A. Electron transfer reactions in chemistry. Theory and experiment. *Rev. Mod. Phys.* **1993**, *65*, 599–610.
- (95) Mao, Y.; Montoya-Castillo, A.; Markland, T. E. Accurate and efficient DFT-based diabaticization for hole and electron transfer using absolutely localized molecular orbitals. *J. Chem. Phys.* **2019**, *151*, No. 164114.
- (96) Futera, Z.; Blumberger, J. Electronic couplings for charge transfer across molecule/metal and molecule/semiconductor interfaces: Performance of the projector operator-based diabaticization approach. *J. Phys. Chem. C* **2017**, *121*, 19677–19689.
- (97) Ghan, S.; Kunkel, C.; Reuter, K.; Oberhofer, H. Improved projection-operator diabaticization schemes for the calculation of electronic coupling values. *J. Chem. Theory Comput.* **2020**, *16*, 7431–7443.
- (98) Hay, P. J. On the calculation of natural orbitals by perturbation theory. *J. Chem. Phys.* **1973**, *59*, 2468–2476.
- (99) Meyer, W. PNO-CI studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. *J. Chem. Phys.* **1973**, *58*, 1017–1035.
- (100) Morrison, A. F.; You, Z.-Q.; Herbert, J. M. Ab initio implementation of the Frenkel-Davydov exciton model: A naturally parallelizable approach to computing collective excitations in crystals and aggregates. *J. Chem. Theory Comput.* **2014**, *10*, 5366–5376.
- (101) Morrison, A. F.; Herbert, J. M. Low-scaling quantum chemistry approach to excited-state properties via an ab initio exciton model: Application to excitation energy transfer in a self-assembled nanotube. *J. Phys. Chem. Lett.* **2015**, *6*, 4390–4396.
- (102) Morrison, A. F.; Herbert, J. M. Analytic derivative couplings and first-principles exciton/phonon coupling constants for an ab initio Frenkel-Davydov exciton model: Theory, implementation, and application to compute triplet exciton mobility parameters for crystalline tetracene. *J. Chem. Phys.* **2017**, *146*, No. 224110.
- (103) Kathir, R. K.; de Graaf, C.; Broer, R.; Havenith, R. W. A. Reduced common molecular orbital basis for nonorthogonal configuration interaction. *J. Chem. Theory Comput.* **2020**, *16*, 2941–2951.
- (104) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.
- (105) Mayer, I. Using singular value decomposition for a compact presentation and improved interpretation of the CIS wave functions. *Chem. Phys. Lett.* **2007**, *437*, 284–286.
- (106) Surján, P. R. Natural orbitals in CIS and singular-value decomposition. *Chem. Phys. Lett.* **2007**, *439*, 393–394.
- (107) Richard, R. M.; Herbert, J. M. Time-dependent density-functional description of the $^1\text{L}_a$ state in polycyclic aromatic hydrocarbons: Charge-transfer character in disguise. *J. Chem. Theory Comput.* **2011**, *7*, 1296–1306.
- (108) Körzdörfer, T.; Sears, J. S.; Sutton, C.; Brédas, J.-L. Long-range corrected hybrid functionals for π -conjugated systems: Dependence of the range-separation parameter on conjugation length. *J. Chem. Phys.* **2011**, *135*, No. 204107.
- (109) Karolewski, A.; Kronik, L.; Kümmel, S. Using optimally tuned range separated hybrid functionals in ground-state calculations: Consequences and caveats. *J. Chem. Phys.* **2013**, *138*, No. 204115.
- (110) Modrzejewski, M.; Rajchel, Ł.; Chalasinski, G.; Szczesniak, M. M. Density-dependent onset of the long-range exchange: A key to donor-acceptor properties. *J. Phys. Chem. A* **2013**, *117*, 11580–11586.
- (111) Körzdörfer, T.; Brédas, J.-L. Organic electronic materials: Recent advances in the DFT description of the ground and excited states using tuned range-separated hybrid functionals. *Acc. Chem. Res.* **2014**, *47*, 3284–3291.
- (112) Uhlig, F.; Herbert, J. M.; Coons, M. P.; Jungwirth, P. Optical spectroscopy of the bulk and interfacial hydrated electron from ab initio calculations. *J. Phys. Chem. A* **2014**, *118*, 7507–7515.
- (113) Garrett, K.; Vazquez, X. A. S.; Egri, S. B.; Wilmer, J.; Johnson, L. E.; Robinson, B. H.; Isborn, C. M. Optimum exchange for calculation of excitation energies and hyperpolarizabilities of organic electro-optic chromophores. *J. Chem. Theory Comput.* **2014**, *10*, 3821–3831.
- (114) de Queiroz, T. B.; Kümmel, S. Charge-transfer excitations in low-gap systems under the influence of solvation and conformational disorder: Exploring range-separation tuning. *J. Chem. Phys.* **2014**, *141*, No. 084303.
- (115) Bois, J.; Körzdörfer, T. Size-dependence of nonempirically tuned DFT starting points for G_0W_0 applied to π -conjugated molecular chains. *J. Chem. Theory Comput.* **2017**, *13*, 4962–4971.
- (116) Gray, M.; Herbert, J. M. Simplified tuning of long-range corrected density functionals for symmetry-adapted perturbation theory. *J. Chem. Phys.* DOI: 10.26434/chemrxiv.14751660, in press.
- (117) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (118) Alam, B.; Morrison, A. F.; Herbert, J. M. Charge separation and charge transfer in the low-lying excited states of pentacene. *J. Phys. Chem. C* **2020**, *124*, 24653–24666.
- (119) Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. Generalized gradient approximation model exchange holes for range-separated hybrids. *J. Chem. Phys.* **2008**, *128*, No. 194105.
- (120) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215. Erratum: *J. Chem. Phys.* **2006**, *124*, 219906.
- (121) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **2006**, *125*, No. 224106.
- (122) Kronik, L.; Kümmel, S. Dielectric screening meets optimally tuned density functionals. *Adv. Mater.* **2018**, *30*, No. 1706560.
- (123) Refaely-Abramson, S.; Sharifzadeh, S.; Jain, M.; Baer, R.; Neaton, J. B.; Kronik, L. Gap renormalization of molecular crystals from density-functional theory. *Phys. Rev. B* **2013**, *88*, No. 081204(R).
- (124) Manna, A. K.; Refaely-Abramson, A.; Reilly, A. M.; Tkatchenko, A.; Neaton, J. B.; Kronik, L. Quantitative prediction of optical absorption in molecular solids from an optimally tuned

screened range-separated hybrid functional. *J. Chem. Theory Comput.* **2018**, *14*, 2919–2929.

(125) Bhandari, S.; Cheung, M.; Geva, E.; Kronik, L.; Duneitz, B. D. Fundamental gaps of condensed-phase organic semiconductors from single-molecule calculations using polarization-consistent optimally tuned screened range-separated hybrid functionals. *J. Chem. Theory Comput.* **2018**, *14*, 6287–6294.

(126) Lao, K. U.; Herbert, J. M. Atomic orbital implementation of extended symmetry-adapted perturbation theory (XSAPT) and benchmark calculations for large supramolecular complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955–2978.

(127) Carter-Fenk, K.; Lao, K. U.; Liu, K.-Y.; Herbert, J. M. Accurate and efficient *ab initio* calculations for supramolecular complexes: Symmetry-adapted perturbation theory with many-body dispersion. *J. Phys. Chem. Lett.* **2019**, *10*, 2706–2714.

(128) Liu, K.-Y.; Carter-Fenk, K.; Herbert, J. M. Self-consistent charge embedding at very low cost, with application to symmetry-adapted perturbation theory. *J. Chem. Phys.* **2019**, *151*, No. 031102.

(129) Ambrosio, F.; Guo, Z.; Pasquarello, A. Absolute energy levels of liquid water. *J. Phys. Chem. Lett.* **2018**, *9*, 3212–3216.

(130) Ambrosio, F.; Miceli, G.; Pasquarello, A. Electronic levels of excess electrons in liquid water. *J. Phys. Chem. Lett.* **2017**, *8*, 2055–2059.

(131) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(132) Kühne, T. D.; Iannuzzi, M.; Del, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, L.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; Golze, D.; Wilhelm, J.; Chulkov, S.; Bani-Hashemian, M. H.; Weber, V.; Borštnik, U.; Taillefumier, M.; Jakobovits, A. S.; Lazzaro, A.; Pabst, H.; Müller, T.; Schade, R.; Guidon, M.; Andermatt, S.; Holmberg, N.; Schenter, G. K.; Hehn, A.; Bussy, A.; Belleflamme, F.; Tabacchi, G.; Glöß, A.; Lass, M.; Bethune, I.; Mundy, C. J.; Plessl, C.; Watkins, M.; VandeVondele, J.; Krack, M.; Hutter, J. CP2K: An electronic structure and molecular dynamics software package—Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, No. 194103.

(133) Sabatini, R.; Gorni, T.; de Gironcoli, S. Nonlocal van der Waals density functional made simple and efficient. *Phys. Rev. B* **2013**, *87*, No. 041108(R).

(134) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, No. 244103.

(135) Ambrosio, F.; Miceli, G.; Pasquarello, A. Structural, dynamical, and electronic properties of liquid water: A hybrid functional study. *J. Phys. Chem. B* **2016**, *120*, 7456–7470.

(136) Pizzochero, M.; Ambrosio, F.; Pasquarello, A. Picture of the wet electron: A localized transient state in liquid water. *Chem. Sci.* **2019**, *10*, 7442–7448.

(137) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, No. 114105.

(138) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.

(139) 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.; Kúš, 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.; Do, 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.; Khistyayev, 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.; Sharma, 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.; Zhao, 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.; Xu, 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. A.; 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. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(140) You, Z.-Q.; Mewes, J.-M.; Dreuw, A.; Herbert, J. M. Comparison of the Marcus and Pekar partitions in the context of non-equilibrium, polarizable-continuum reaction-field solvation models. *J. Chem. Phys.* **2015**, *143*, No. 204104.

(141) Mewes, J.-M.; You, Z.-Q.; Wormit, M.; Kriesche, T.; Herbert, J. M.; Dreuw, A. Experimental benchmark data and systematic evaluation of two *a posteriori*, polarizable-continuum corrections for vertical excitation energies in solution. *J. Phys. Chem. A* **2015**, *119*, 5446–5464.

(142) Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM *ab initio* level. *J. Mol. Struct.: THEOCHEM* **1999**, *464*, 211–226.

(143) Lange, A. W.; Herbert, J. M. A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.* **2010**, *133*, No. 244111.

(144) Lange, A. W.; Herbert, J. M. Symmetric versus asymmetric discretization of the integral equations in polarizable continuum solvation models. *Chem. Phys. Lett.* **2011**, *509*, 77–87.

(145) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximations made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. Erratum: *Phys. Rev. Lett.* **1997**, *78*, 1396.

(146) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, No. 036402.

(147) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(148) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(149) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*, No. 134105.

(150) Murray, C. W.; Handy, N. C.; Laming, G. J. Quadrature schemes for integrals of density functional theory. *Mol. Phys.* **1993**, *78*, 997–1014.

(151) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.

(152) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Development of exchange-correlation functionals with minimal many-electron self-interaction error. *J. Chem. Phys.* **2007**, *126*, No. 191109.

(153) Fox, M. F.; Hayon, E. Far ultraviolet solution spectroscopy of the iodide ion. *J. Chem. Soc., Faraday Trans. 1* **1977**, *73*, 1003–1016.