PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 17533

An improved guess for the variational calculation of charge-transfer excitations in large systems†

Nicola Bogo, *[a] Zeyi Zhang, [bc] Martin Head-Gordon[bc] and Christopher J. Stein [ad]

Ab initio quantum-chemical methods that perform well for computing the electronic ground state are not straightforwardly transferable to electronically excited states, particularly in large molecular systems. Wave function theory offers high accuracy, but is often prohibitively expensive. Methods based on time-dependent density functional theory (TD-DFT) are crucially sensitive to the chosen exchange–correlation functional (XCF) parameterization, and system-specific tuning protocols were therefore proposed to address the method's robustness. Methods based on the variational relaxation of the excited-state electron density showcased promising results for the calculation of charge-transfer excitations, but the complex shape of the electronic hypersurface makes convergence to a specific excited state much more difficult than for the ground state when standard variational techniques are applied. We address the latter aspect by providing suitable initial guesses, which we obtain by two separate constrained algorithms. Combined with the squared-gradient minimization algorithm for all-electrons relaxation in a freeze-and-release scheme (FRZ-SGM), we demonstrate that orbital-optimized density functional theory (OO-DFT) calculations can reliably converge to the charge-transfer states of interest even for large molecular systems. We test the FRZ-SGM method on a phenothiazine-anthraquinone CT excitation in a supramolecular Pd(II) coordination cage complex as a function of the cage conformation. This compound has been studied experimentally prior to our work. We compare this freeze-and-release scheme to two XCF reparameterizations, which were recently proposed as low-cost TD-DFT-based alternatives to variational methods. Two dye–semiconductor complexes, which were previously investigated in the context of photovoltaic applications, serve as a second example to investigate the convergence and stability of the FRZ-SGM approach. Our results demonstrate that FRZ-SGM provides reliable convergence for charge-transfer excited states and avoids variational collapse to lower-lying electronic states, whereas time-dependent DFT calculations with an adequate tuning procedure for the range-separation parameter provide a computationally efficient initial estimate of the corresponding energies, with a computational cost comparable to that of configuration-interaction singles (CIS) calculations.

Received 18th May 2025,
Accepted 21st July 2025

DOI: 10.1039/d5cp01867f

rsc.li/pccp

1 Introduction

Over the last decade, the resolution of photo-lithographic techniques for silicon semiconductor fabrication has approached the size of molecules, and the physical limits of optical technology have made further miniaturization increasingly challenging.[1,2] A bottom-up approach where molecular components are assembled into molecular devices,[3,4] like biological photosynthetic systems,[5,6] would allow further miniaturization in chip design, paving the way for a disruptive innovation in semiconductor technology: nanostructured carbon-based semiconductors. However, the cost of organic building blocks equipped with extended π electrons systems, which are delocalized and prone to participate in conduction, the development costs of novel synthesis and characterization techniques, and ultimately, the impact of trial and error on semiconductor design amount to a significant barrier in the development of these materials.[7–9] In this contribution, we improve an efficient method to compute electronically excited states of supramolecules, and we apply it to the study of charge-transfer excitations in a Pd(II) supramolecular coordination cage[10] and

[a] *Department of Chemistry and Catalysis Research Center, TUM School of Natural Sciences, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany. E-mail: christopher.stein@tum.de*
[b] *Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley, California 94720, USA*
[c] *Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[d] *Atomistic Modeling Center, Munich Data Science Institute, Technical University of Munich, Garching 85748, Germany*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5cp01867f

two dye–semiconductor complexes.[11] In this section, we motivate our choice to develop efficient and robust electronic structure theory methods for the calculation of charge-transfer excitations in supramolecular systems and introduce our diverse test examples.

A rational strategy for the development of carbon-based semiconductor technology implies the study of charge transfer (CT) in biological systems, its reproduction in bio-inspired energy materials, and their technological application. Charge transfer (CT) refers to the process in which an electron moves from an electron-rich donor (D) moiety to an electron-deficient acceptor (A). The donor and acceptor groups can belong to the same molecule (intramolecular CT), or be located on different molecular units (intermolecular CT). On top of that, light absorption can trigger the formation of CT excitations (photo-induced CT). The distance R_{DA} of the nuclear center of mass of the donor and acceptor moieties is a common measure for the extent of charge separation.[12] The fundamental understanding of charge-transfer processes enabled by theoretical work is fundamental, forging guiding principles for molecular design and testing them with computer simulations. Given this premise, excited-state electronic structure methods play a crucial role in the study and discovery of charge transport in molecular systems.[13–15] Since their introduction in the 1990s, tools based on density functional theory (DFT) have achieved predictive power in the simulation of ground-state (GS) chemistry[16] and are routinely used to develop new chemicals.[17,18] Likewise, excited-state electronic-structure methods based on DFT, like time-dependent density functional theory[19] (TD-DFT), gained popularity due to the straightforward simulation of light absorption features in small molecules. In a TD-DFT calculation, an approximation of the response function of the system under investigation is computed. The success of TD-DFT resides in the application of efficient approximations that lower its computational demand at the cost of introducing small errors. With a focus on economic simulations, the most popular TD-DFT methods expand the response function in a Taylor series and truncate it to the linear term, applying the Tamm–Dancoff approximation[20] (TDA) to recast the resulting electronic structure problem to the well-known configuration-interaction singles[21,22] (CIS) method. This method, which we name TDA from here on, is plagued by systematic errors[23,24] for common (hybrid) exchange–correlation functionals (XCFs) with fixed parameters. This error can be reduced by system-specific tuning protocols or with functionals that are higher up in the Jacob's ladder classification, such as double-hybrid functionals. When a range-separated hybrid XCF is employed, the range separation parameter ω can be iteratively adjusted such that the HOMO and LUMO energies match the ionization potential (IP) and electron affinity (EA), respectively.[25] In double-hybrid XCFs, such as *e.g.* B2PLYP,[26] parametrization schemes for the range-separation parameter also exist to tune the functional to match the IP and EA. Yet another alternative is the local variation of the fraction of exact exchange according to a local mixing function in so-called local hybrid XCFs (*e.g.* LH20t[27]), which can be implemented efficiently,[28] but leads to a non-trivial parametrization problem. All of these options can significantly reduce the error in TDA calculations with the tuning of range-separated hybrid functionals standing out as an efficient and straightforward procedure.

In our previous publication,[24] we demonstrated how orbital-optimized DFT (OO-DFT) methods constitute the most accurate low-scaling electronic structure approximation for the calculation of inter-molecular CT excitations and do not require special adjustments of the range-separation parameter. OO-DFT methods compute the ground state and an excited state variationally in two separate self-consistent field (SCF) calculations. For this reason, this approach is also referred to as ΔSCF for excited states, and its theoretical foundations were proven under several formulations.[29–31] In contrast to the ground state, excited-state energies belong to saddle points of the electronic hypersurface, making numerical convergence more challenging. Indeed, we relied on outlier detection rules in our previous publication[24] to exclude data points flawed by poor convergence in our benchmark study, where we subsequently tested the accuracy of OO-DFT methods. We introduced a dataset of accurate intermolecular charge-transfer (ICT) excitation energies[24] in donor–acceptor molecular dimers at various separation distances (R_{DA}) and found that OO-DFT methods perform well, but struggle to converge the electronic structure at short R_{DA} (3.5–5 Å) because of strong changes in the polarization of the excited state MOs. Following our study, Schmerwitz *et al.*[32] analyzed the performance their freeze-and-release optimization algorithm on our R_{DA}-dataset. The freeze-and-release algorithm is a 2-step optimization algorithm, which combines an iterative constrained (''frozen'') optimization step with a subsequent relaxation of all degrees of freedom. The authors were able to obtain a smooth potential energy surface for the ICT excitation also at short donor–acceptor distances, demonstrating the importance of a well-behaved guess for converging OO-DFT calculations. In this contribution, we introduce a new strategy for the refinement of the MO guess and combine it to the OO-DFT algorithms implemented in the Q-Chem software package,[33] testing it on the ICT excitations in the tetrafluoroethylene-ethylene dimer system from the R_{DA}-dataset, that we consider a good benchmark for convergence of OO-DFT methods. We implement and analyze a constrained optimization algorithm to refine the MO guess used in OO-DFT calculations using the efficient geometric direct minimization (GDM) algorithm,[34] akin to the approach introduced by Schmerwitz *et al.*[32] Moreover, we investigate an alternative guess for ICT excitations obtained from the absolutely-localized molecular orbitals[35,36] (ALMO) method and compare the maximum overlap method to an initial guess[37] (IMOM) to the squared-gradient minimization[38,39] (SGM) algorithm on these refined guesses.

TDA is certainly of fundamental importance for the calculation of light absorption properties, and is by far the most popular method for the calculation of electronic excitations in molecular systems. The OO-DFT methods we apply in this work also require an initial TDA calculation to set up the OO-DFT calculation. Nevertheless, results obtained with the TDA method depend significantly on the chosen XCF approximation

and should therefore be validated for a given compound class by comparing them to reference data. On top of that, CT excitations in organic systems tend to have a very small transition dipole moment (TDM) and therefore a small absorption coefficient compared to the bright transitions that dominate the spectrum. Theory-to-theory validation is therefore the only suitable strategy to validate computational methods for CT excitations,[24,40] but a direct assessment is often prohibited by the steep computational scaling of accurate wavefunction methods. For this reason, we demonstrated the accuracy that can be achieved with OO-DFT methods in an extensive benchmark for small systems in our previous work,[24] and we confront technical challenges regarding the stable convergence towards the desired CT target states for large molecular systems in the present study. That the accuracy that we demonstrated for small systems does not deteriorate for the prototypical large systems studied in this work is a conjecture we make based on experience with ground-state DFT.

In Section 2, we review the current state-of-the-art methods for variational calculations of electronically excited states. Our implementation of an efficient partitioning of the MO matrix **C** is discussed in Section 3.2. Convergence and accuracy of the SGM and IMOM methods using the ground-state MO guess, ALMO guess, and refined MO guess through constrained optimization with frozen excited electron and hole MOs (Frozen guess or ''FRZ guess'' in the following) are tested on the tetrafluoroethylene-ethylene dimer in Section 4.1. The guess refinement algorithm is combined with the SGM algorithm for the calculation of CT excitations in large systems. We test our method on CT excitations on two classes of supramolecular systems: a Pd(II) coordination cage and two dye–semiconductor model materials. The supramolecular cage we studied was previously investigated by Frank *et al.*, who combined spectroelectrochemistry and time-resolved spectroscopy experiments to investigate the photoinduced relaxation upon excitation.[10] Upon dissolution of the cage components in a reaction medium, intercalated double-cage complexes were formed by self-assembly. Two banana-shaped ligands, one bearing a phenothiazine (PTZ) dye and one an anthraquinone (ANQ) dye, were employed for the preparation of several intercalated cage dimers. Each cage dimer is composed of four Pd(II) metal centers and eight organic ligands that coordinate the metal with a pyridyl group. The cage dimers observed were either homoleptic, solely composed of the PTZ and ANQ ligands, respectively $[Pd_4(PTZ)_8]^{8+}$ and $[Pd_4(ANQ)_8]^{8+}$ or heteroleptic with a statistical mixture of PTZ and ANQ ligands ($[Pd_4(ANQ)_n(PTZ)_{8-n}]^{8+}$). The homoleptic intercalated cage dimers proved stable towards repetitive oxidation/reduction cycles, and the changes in the visible light absorption spectrum after reduction and oxidation steps were investigated. Transient absorption (TA) of heteroleptic $[Pd_4(ANQ)_n(PTZ)_{8-n}]^{8+}$ cage upon 385 nm excitation was also recorded. 100 ps after the excitation pulse, the TA shows similar features to the spectral changes obtained by oxidizing and reducing the homoleptic donor and acceptor complexes, respectively, giving experimental evidence of a charge-transfer excitation formed in the supramolecular structure.

The results obtained with our OO-DFT methods on the PTZ-ANQ CT excitation in a Pd(II) coordination cage are compared to the TDA prediction obtained with recently proposed reparametrizations of the LRC-ωPBE[41] XCF for the TDA method and experimental measurements in Section 4.2. Our calculations on the stacked conformation of a heteroleptic cage monomer confirm a PTZ-ANQ CT excitation, and we compared its energy to the PTZ-Pd(II) ligand-to-metal CT (LMCT) state. However, a direct comparison of the computed energy to the experimental recordings is not possible, since the formation of a CT state 100 ps after light excitation is proven experimentally, but the energy of such a state cannot be measured in TA experiments. Excited-state dynamic simulations and excited-state absorption would be required to unambiguously prove the formation of the CT excitation, but no methodology is available to date for systems of this size. We consider OO-DFT methods excellent candidates for dynamic simulations of large systems, and ongoing work in our lab is aimed at making such calculations feasible. Finally, CT excitations of two dye–TiO_2 complexes were investigated in Section 4.3. These systems are chemically rather different than the coordination cages previously investigated, because the electron acceptor is a semiconductor material. We were therefore interested in testing the performance of our method also for such a system. The dyes we investigate—which we refer to as JK2 and D102 for simplicity—are highly relevant from an application point of view as a potential material in solar cells, and have been studied both experimentally[42–45] and in computer simulations.[11,46,47] However, CT excitations typically form weak bands compared to local (valence) states, which dominate the UV-Vis spectrum. Therefore, we are unable to assign contributions to the experimental absorption spectrum to dye–TiO_2 CT excitations and we do not expect direct formation of the CT state by light absorption. Although we cannot directly compare our calculations to experimental results, we compared to other theoretical studies and demonstrate the stable convergence of our method also for this chemically distinct class of large heterogeneous systems.

2 Theoretical background

A rigorous introduction to TD-DFT is beyond the scope of our study, so we point the reader to an exhaustive introduction,[48] and focus instead on the variational methods we employed. In the following, Greek indices denote the Gaussian AO basis functions in our notation, and Latin characters denote the MOs. We use indices j, k for occupied MOs, a, b for virtual MOs, and m, n for general MOs, independent of occupation. The superscript (i) indicates the iteration number in variational electronic-structure calculations. For ground-state calculations, the convergence of the electronic structure is accelerated with two alternative approaches: direct optimization (DO) and extrapolation methods. DO methods apply a unitary transformation **U** to the MO matrix **C** at every step

$$\mathbf{C}^{(i+1)} = \mathbf{C}^{(i)}\mathbf{U}. \quad (1)$$

The unitary transformation is parametrized by taking the exponent of an antisymmetric matrix Δ

$$\mathbf{U} = e^{\Delta}. \tag{2}$$

This way, the optimization is performed directly on the elements of Δ, instead of applying small variations on the $\{c_{i\mu}\}$ elements while constraining the columns of $\mathbf{C}$ to stay orthonormal. The energy minimization involves the evaluation of the gradient of the electronic energy with respect to the upper triangular elements of Δ. Suitable minimization algorithms are *e.g.* Broyden–Fletcher–Goldfarb–Shanno (BFGS) which is used in the implementation of the geometric direct minimization[34] (GDM) algorithm in Q-Chem. The elements of Δ are the orbital rotations denoted by the letter θ. $\theta_{mn}^{(i)}$ mixes MOs m and n from the i-th iteration to construct $\mathbf{C}$ in iteration $i+1$, and only rotations mixing occupied and virtual MOs $\theta_{ja}^{(i)}$ (off-diagonal blocks in Δ) affect the total electronic energy.

Conversely, extrapolation methods, such as the direct inversion of the iterative subspace (DIIS) method from Pulay,[49,50] leverage the mathematical properties of the Fock $\mathbf{F}^{(i)}$ and density $\mathbf{P}^{(i)}$ matrices at convergence to construct an error vector. The commutator between density and Fock matrix vanishes at convergence, such that

$$\mathbf{e}^{(i)} = \mathbf{S}\mathbf{P}^{(i)}\mathbf{F}^{(i)} - \mathbf{F}^{(i)}\mathbf{P}^{(i)}\mathbf{S} \tag{3}$$

is a suitable error vector where the overlap matrix is denoted as $\mathbf{S}$. In DIIS, the density matrix is extrapolated with information from previous iterations. As a consequence, it is not guaranteed that the closest local minimum is found for a given initial guess since the extrapolation can lead to tunneling through the local basin. In case of ground-state calculations, the combined DIIS-GDM algorithm accelerates convergence by using DIIS in the first iterations until the DIIS error is below a given threshold and refining the DIIS guess with GDM to tight convergence. Excited states, however, are saddle points instead of minima. An adaptation of the algorithms to a variational optimization of excited states is, therefore, not straightforward.

To mitigate the complexity of direct saddle-point optimization of the electronic hypersurface, a constraint is applied to the density in constrained DFT[51] (CDFT) methods. Constraints can impose site occupation[52] conditions or can be applied directly to the density difference between ground and excited state.[53] These methods hold great potential when enough information is available about the target excited state (*e.g.* information about site occupancy), but in this work, we use constraints to compute an improved guess density. In our calculations, the electronic energy is hence computed by optimization of all degrees of freedom of the electronic energy without applying constraints.

To compute specific excited states, Gilbert *et al.* proposed an adaptation of direct optimization or extrapolation algorithms,[54] where the Fock matrix is constructed from the orbitals that overlap maximally with a reference configuration rather than following the Aufbau principle. In this maximum overlap method (MOM), the overlap of the occupied MO set belonging to the i-th iteration to a reference MO set $\mathbf{C}^{\mathrm{ref}}$

$$\mathbf{O} = (\mathbf{C}^{\mathrm{ref}})^{\dagger}\mathbf{S}\mathbf{C}^{(i)} \tag{4}$$

is computed at every iteration, where the element O_{jk} quantifies the overlap between the j-th MO in the reference set and the k-th MO in the MO set belonging to the i-th iteration. The projection of the k-th MO onto the reference is

$$p_k = \left(\sum_{j}^{n} O_{jk}\right)^{\frac{1}{2}} \tag{5}$$

and the MOs are occupied starting from the highest occupation, to ensure the electronic structure is computed for the desired electron configuration. As reference MOs, either the initial orbitals from a ground-state optimization can be used in the "maximum overlap to an initial guess" (IMOM) method,[37] or they can be updated for each iteration step (MOM). The IMOM method is generally very fast to converge to the targeted excited state. However, when the provided guess is far from the correct targeted state, convergence is challenging, and undesired stationary points might be reached. This happens because the overlap of the occupied MO set with a reference set is maximized at every iteration in the MOM method. If at least one occupied-virtual pair has a remarkably different shape in the target stationary state, quantitative mixing of this pair can lead the algorithm into a flat region of the hypersurface, which does not represent a physical state. The special case of CT excitations was investigated recently by Schmerwitz *et al.*,[32] who showcased how the MOM method frequently fails to converge to the targeted stationary points of the energy hypersurface when coupled to a direct-optimization algorithm. The lack of convergence is the result of strong polarization in the excited-state density: over the iterations, the hole and particle MOs are mixed quantitatively, and the optimization algorithm cannot exit these strong mixing regions.

Hait and Head-Gordon proposed an alternative approach exploiting DO algorithms,[38,39] taking the square of the energy derivative with respect to occupied-virtual orbital rotations

$$\Lambda = \left|\nabla_{\vec{\theta}} E\right|^2 = \sum_{ja}\left|\frac{\partial E}{\partial \theta_{ja}}\right|^2 \tag{6}$$

and feed this as the objective function to the DO algorithm. The resulting algorithm, named squared gradient minimization (SGM), which is usually coupled to standard line-search schemes, can converge on the saddle points of the electronic energy hypersurface. Still, the manipulated objective function is characterized by undesired minima and cusps,[55] so it is again sensitive to the quality of the provided guess. Moreover, in the SGM algorithm, the gradient is reduced quadratically when taking its square, meaning the condition number for the convergence must also be squared, requiring very tight convergence. Despite these aspects, the SGM algorithm has been proven to be very successful *e.g.* for the simulation of core-electrons spectroscopies[56] because ground or ionized states

provide good guess functions for the orbital optimization, meaning the SGM algorithm starts in the appropriate quadratic well and can converge flawlessly.

Finally, Selenius *et al.* proposed an efficient diagonal preconditioner for DO algorithms[40]

$$\frac{\partial^2 E}{\partial \kappa_{mn}{}^2} \approx -2(\varepsilon_m - \varepsilon_n)(f_m - f_n), \tag{7}$$

where $\varepsilon_{m/n}$ are the orbital energies and $f_{m/n}$ are the occupation numbers of orbitals m and n, respectively. This definition inverts the gradient along all the rotations mixing an occupied orbital with a lower-lying unoccupied orbital, ensuring the optimization is started by going "uphill" along the correct rotations that lead to the desired saddle point on the electronic hypersurface. This idea had already been employed for the construction of appropriate objective functions through manipulation of the gradient, employing efficient approximations of the inverse Hessian used in direct optimization algorithms. This approach can be understood as an adaptation of the popular eigenvector-following algorithm for saddle-point search that is used for transition-state (TS) calculations to the electronic problem. A principal difference between the two applications is that only first-order saddle points are of interest for a TS search, whereas also higher-order saddle points are of interest for excited-state electronic structure. This poses additional challenges for the variational calculations of electronic excitations. Like for TS calculations, eigenvector-following algorithms are sensitive to the quality of the initial guess. Furthermore, molecular orbitals can change their energetic order over the DO iterations, leading to undesired consequences: firstly, energy terms belonging to the excited orbital directly enter the minimized electronic energy, whereas virtual orbitals don't, and this can guide the optimization algorithm into the aforementioned strong mixing regions with occupied-virtual rotations of 45°. Secondly, when the energetic order of occupied and virtual MOs is swapped, the target saddle point order is changed, and the objective function must be altered accordingly to avoid convergence to the wrong stationary point.

In the end, ΔSCF methods are based on local optimization algorithms, and the underlying problem with variational calculations of excited-state energies resides in the complexity of the energy hypersurface. Rapid convergence to the desired stationary point hence requires a suitable initial guess. Therefore, guess-refinement algorithms are desirable to reduce the number of iterations and ensure convergence to the correct electronic configuration in OO-DFT calculations. Schmerwitz *et al.* made use of constrained optimization in their variational excited-state methods[32,57,58] implemented in the GPAW software package[59] with broad success and proved the reliability of their approach by recomputing the outlier data points from our R_{DA}-dataset. This motivated us to implement our variant of a constrained optimization algorithm and test the performance of the IMOM and SGM algorithms on the refined guess. Details about the constrained optimization algorithm are discussed in Section 3.2.

An alternative method for the calculation of inter-molecular excitations is the absolutely-localized molecular orbitals (ALMO) method. ALMO enables the variational calculation of molecular fragments by partitioning the Gaussian basis $\{\varphi_\mu\}$ spanning the rows of the **C** matrix into localized subsets. In the simple case of two fragments, named A and B, the Gaussian basis is partitioned into the localized bases $\{\varphi_\nu\}^A$ and $\{\varphi_\kappa\}^B$ belonging to fragments A and B respectively. The MO matrix **C** is therefore enforced to maintain the block structure

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_A & 0 \\ 0 & \mathbf{C}_B \end{bmatrix} \tag{8}$$

throughout the calculation, where $\mathbf{C}_A$ and $\mathbf{C}_B$ are the MO matrices of fragments A and B, respectively. The ALMO method is based on the equations of the locally-projected SCF for molecular interactions.[35,36] This strategy enables the straightforward calculation of energies and properties by expanding the MOs of each fragment in terms of the atomic orbitals of the same fragment, resulting in orthonormal MOs for each fragment (whereas MO sets belonging to different fragments are not mutually orthonormal). Charge transfer, where the MOs of the acceptor fragment accommodate part of the electron density from the donor fragment, cannot be represented. In contrast, the transfer of integer charges—*i.e.* a full electron in the case of ICT—can be computed straightforwardly by treating the donor and acceptor monomers as cation and anion, respectively. In addition to the ALMO method, where we keep the above constraint during the variational optimization, we also evaluate the performance of the IMOM and SGM algorithms on the guess MOs provided by the ALMO method for the tetrafluoroethylene-ethylene model system in Section 4.1.

Finally, in Section 4.2 we compare OO-DFT methods to reparametrizations of the LRC-ωPBE[41] XCF in TD-DFT. A system-specific tuning of the range-separation parameter ω proved effective in combination with TD-DFT by recovering properties of the exact XCF with semi-empirical RSH XCFs.[60] Two alternative schemes to determine a suitable ω in LRC-ωPBE have recently been proposed. The global density-dependent (GDD) tuning[61,62] adjusts ω by computing the distance $d_{\mathbf{x}}$ between an electron in the outer regions of a molecule and the exchange hole in the localized valence orbitals

$$\omega_{GDD} = C\langle d_{\mathbf{x}}{}^2\rangle^{-1/2}, \tag{9}$$

where C is an empirical constant determined for each XCF. Automatic GDD tuning is implemented in Q-Chem *via* ground-state density calculation for a system of interest computed with an initial guess for ω to determine $d_{\mathbf{x}}$. The optimal ω_{GDD} can then be directly plugged into the expression for the LRC-ωPBE XCF to be used in subsequent calculations. GDD tuning is size-extensive and has successfully been applied to the calculation of intramolecular charge-transfer excitations and extended aromatic systems.[62] In contrast, a common problem in alternative tuning strategies comes from rather abrupt changes in the optimal range-separation parameter for similar systems,

e.g. different conformations of the same chemical compound, which results in discontinuous properties.

Yan *et al.*[63] proposed an alternative empirical XCF reparametrization designed for the description of CT states. From a preliminary TD-DFT calculation using a related XCF (*e.g.* the PBE0[64] XCF is used for tuning LRC-ωPBE[41]), and the average excitation distance D_{CT} is computed for several CT excited states, respectively. The ω parameter for the final calculations is then obtained from the maximum D_{CT} in the set

$$\omega^* = \frac{2}{D_{CT}^{max}}, \tag{10}$$

where the 2 in the numerator is an empirical constant. The calculation of absorption spectra in push–pull organic dyads with LRC-ω*PBE showed promising results,[63] improving intramolecular CT excitations while preserving the accurate prediction of valence states.

3 Methods

3.1 Computational settings

All OO-DFT and most TDA calculations presented in this work are performed with a development version of Q-Chem[33] using the def2-TZVP[65] basis set. TDA calculations generally employ the LRC-ωPBE[41] XCF and for the large PTZ-ANQ molecular dimers we adjusted the range-separation parameter ω according to tuning procedures based on properties of the exchange hole,[61,62] or the excited-state density.[63] The guess refinement method introduced in Section 3.2 is combined with the IMOM method[37] and the SGM algorithm,[38,39] performing a line search at each iteration. Corresponding OO-DFT calculations with the ALMO-SGM and FRZ-SGM methods are performed with the ωB97X-D XCF,[66] and the def2-ECP effective core potential[67,68] is used for the Pd ions. The different choice of functionals for TDA and OO-DFT methods was a conscious decision ensuring that each method employs their respective best available functional. In Section 4.3, we compute the low-lying excitations in the two dye–TiO_2 complexes with the simplified TDA[69] (sTDA) method implemented in Orca,[70] and recompute the corresponding excitations with the FRZ-SGM method and the ωB97X-D XCF. Here, we also used the ωB97X-D XCF for the sTDA calculation since we did not aim at a method comparison but rather required the results of the sTDA calculation to set up the subsequent OO-DFT calculation.

3.2 Constrained optimization algorithm

ICT excitations involve major changes in the electron density because, ultimately, an ion pair is formed. The resulting Coulombic attraction changes the electrostatic problem substantially, and any initial guess that takes this into account should be a good starting point for OO-DFT calculations of CT states. The ALMO method we described in the previous section achieves this if the fragments are chosen in their respective ionic states. An alternative initial guess can be generated by freezing the electron and hole orbitals during a variational optimization as suggested by Schmerwitz *et al.*[32] We implemented this initial guess by

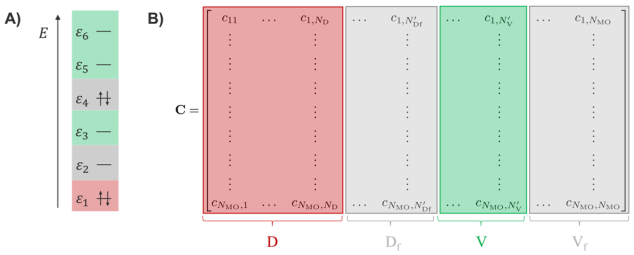

Fig. 1 **C** matrix partitioning for frozen electron-and-hole constrained optimization. Panel A shows the example of a double excitation from MO no. 2, with the energy ε_2, to MO no. 4, with the energy ε_4, on an arbitrary energy axis E. The **C** matrix is given in the general form in panel B. Active doubly-occupied vectors are highlighted in red, active virtual orbitals in green, and all frozen vectors in grey. A detailed discussion of the index notation used in panel B and the **C** matrix reordering procedure is given in the main text.

rearranging the molecular orbital coefficient matrix **C**, thereby effectively modifying the Aufbau principle and subsequently excluding the occupied-virtual rotations of the frozen electron and hole orbitals from the electronic gradient. Hence, the resulting calculation is a constrained minimization including all occupied-virtual rotations for MOs that are not directly involved in the excitation.

We start from a partitioning of the orbital coefficient matrix **C** of a restricted Hartree–Fock (RHF) calculation but note that this approach can be extended straightforwardly to the unrestricted case and Kohn–Sham theory. All **C** elements are real numbers and **C** is ordered according to the block structure shown in Fig. 1.

Given an atomic orbital (AO) basis of size N_{basis}, the **C** matrix of dimension $N_{MO}(\leq N_{basis})$ is ordered into blocks. Active doubly-occupied vectors are in the first block D, indexed from 1 to N_D. A block D_f with N_{Df} frozen doubly-occupied orbitals follows, indexed from N_D+ 1 to $N_D + N_{Df} = N'_{Df}$. Similarly, virtual orbitals are partitioned into the V block containing active virtual orbitals $(N'_{Df}+1,\ldots,N'_{Df}+N_V=N'_V)$ and a block V_f containing frozen virtual orbitals $(N'_V+N_{Vf}+1,\ldots,N_{MOs})$. Note that for ICT excitations the D_f and N_f blocks are simple column vectors. At the beginning of the calculation, occupied (red) and virtual (green) orbitals are reordered to move the frozen vectors (in grey) to the end of each block. In the simple example shown in panel A of Fig. 1, a double excitation is computed by exciting one α and one β electron from MO no. 2 to MO no. 4, and the MOs are sorted in the order 1-4-3-5-6-2. The D_f and V_f blocks contain only MOs no. 4 and 2, respectively. The block structure emerging for the **C** matrix by partitioning into active and frozen parts is depicted in panel B of Fig. 1, using the general notation introduced above. Once the matrix is reordered to have the MOs involved in the excitation at the end of the occupied and virtual block, the constrained optimization just requires input on the number of holes and electrons to constrain, and in principle, multiple electron–hole pairs can be considered.

The energy optimization is then performed with a second-order method, which requires gradients and previous steps to compute an approximation of the inverse Hessian. In our implementation, we leverage the efficient GDM algorithm[34]

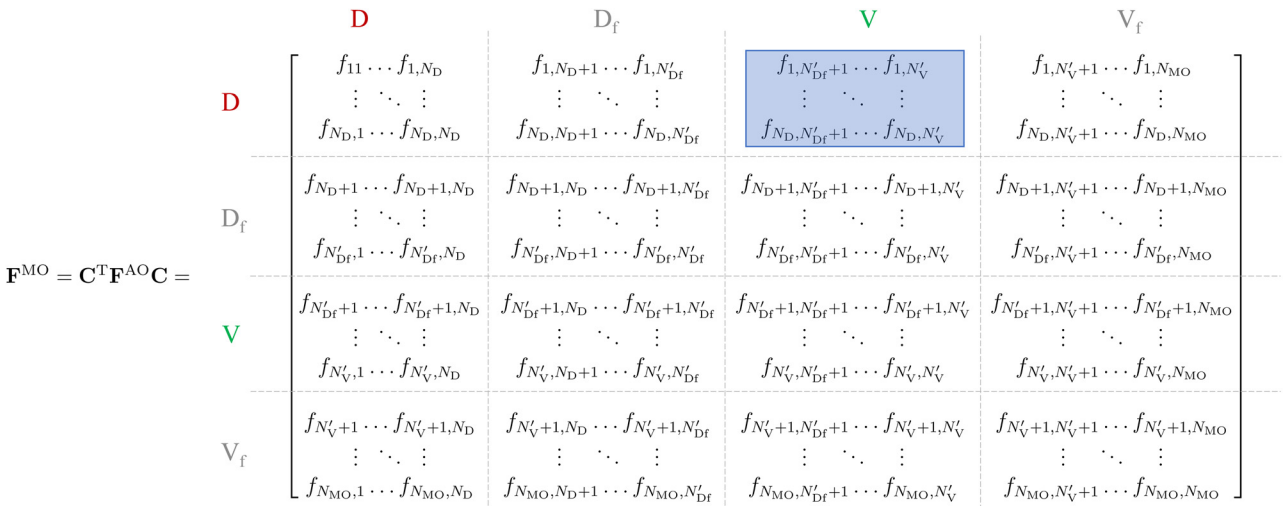

Fig. 2 **F** matrix block structure arising from the **C** matrix partitioning introduced in Fig. 1. The Fock matrix has 16 blocks, six in the upper triangle. Frozen blocks do not contribute to the gradient, such that only the DV block from the upper triangle (highlighted in blue) contributes to the gradient.

implemented in Q-Chem. The **C** matrix reordering results in the partitioning of the Fock matrix depicted in Fig. 2. The Fock matrix is computed in the MO basis. The energy gradient with respect to the MO rotations is given by the off-diagonal elements of the symmetric Fock matrix. In the presence of frozen blocks D_f and V_f, the Fock matrix is partitioned into 16 blocks, six belonging to the upper triangle. Only the DV block in the upper triangle (highlighted in blue in Fig. 2) enters the gradient, as frozen vectors do not contribute in this constrained optimization. Similarly, gradients and steps from previous iterations are stored only for the free active vectors and are transformed by taking the corresponding blocks from the transformation matrix given the step Δ.

We evaluate the performance of this refined guess strategy (termed FR) combined with the IMOM and SGM algorithms on the lowest-lying ICT excitation of the tetrafluoroethylene-ethylene dimer at various donor–acceptor displacements in the following Section 4.1.

4 Results and discussion

4.1 Convergence study: tetrafluoroethylene-ethylene dimer

In our previous work on ICT excitations,[24] we observed that especially for short donor–acceptor distances R_{DA} = 3.5–5 Å, convergence to the desired ICT states is unreliable and variational collapse to different excited states or even the ground state occurs frequently. In contrast, Schmerwitz[32] *et al.* demonstrated that for the example of the tetrafluoroethylene-ethylene dimer system, their freeze-and-release algorithm reliably converges to the lowest ICT state. Here, we analyze how far this is due to an improved initial guess rather than the optimization algorithm. We therefore combined initial guesses provided by the FRZ method presented in the previous section and the ALMO guess with IMOM and SGM optimization for the same dimer system. In our calculations, IMOM is coupled to the DIIS extrapolation algorithm, and the SGM algorithm performs a line search at each iteration. All results on the tetrafluoroethylene-ethylene dimer are converged to 10^{-5} H. We measure the ICT character of the states by the average electron–hole distance parameter D_{CT} (panel B of Fig. 3), which can be readily computed from a grid representation of the ground- and excited-state electron densities and were introduced by Le Bahers *et al.* in ref. 12.

The energies of the lowest ICT states as a function of the donor–acceptor distance are shown in panel A of Fig. 3. The orange and green dotted lines result from calculations where the ALMO and FRZ constraints are sustained, respectively, whereas the purple line corresponds to a subsequent variational optimization with SGM (FRZ-SGM). We note that starting SGM from the ALMO guess yields identical results in this example. For large donor–acceptor distances, all methods presented here reliably converge to an ICT state that shows the expected $1/R_{DA}$ behaviour. Evidently, even at short distances, the ALMO and FRZ constraints provide excellent guesses that are only minimally optimized during the subsequent unconstrained variational optimization. At long donor–acceptor distances, this discrepancy is invisible on the scale of the graph. The CT descriptor in panel B confirms convergence to an ICT state.

Different results are obtained when these initial guesses are combined with a subsequent IMOM optimization. Starting from a ground-state (gray *x*-markers), FRZ guess (blue dots), or ALMO (pink triangles), panel A of Fig. 3 shows that for short donor–acceptor distances, IMOM optimization converges to states that are significantly lower in energy than the initial guesses or the SGM results described above. Predictably, the discrepancy is worst on average for the ground state guess, but also the refined guess strategies cannot guarantee convergence to the targeted ICT state. Indeed, and in line with our previous work,[24] panel B demonstrates that these calculations converged to lower-lying non-ICT states, because D_{CT} quantifies how far the electron density is shifted upon excitation, and its value is <2 Å.

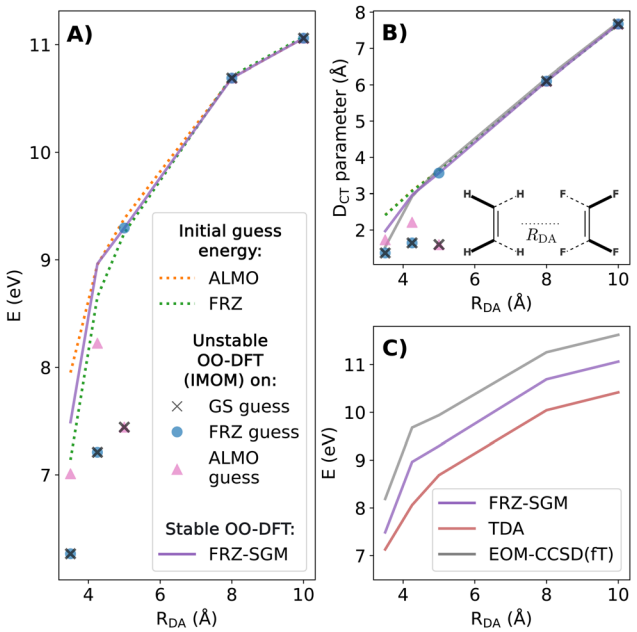

Fig. 3 Lowest-lying ICT excitation in the tetrafluoroethylene-ethylene dimer as a function of donor–acceptor distance R_{DA}. The convergence of several OO-DFT methods and initial guess MOs is analyzed in panel A, and the accuracy of the FRZ-SGM method is compared against TDA and the EOM-CCSD(fT)[71]/cc-pVTZ[72] reference in panel C. The density-based CT descriptor D_{CT} is plotted in panel B, alongside a sketch of the system investigated in the inset. The markers show results for IMOM calculations computed with the ground-state MO guess orbitals (grey x markers), ALMO guess orbitals (pink triangles), and frozen hole-and-electron guess orbitals (blue dots). The dashed orange lines correspond to the energy and descriptors computed on the ALMO density, and the dashed green lines belong to the frozen hole-and-electron density. The purple line is obtained by using the FRZ guess orbitals to initiate the SGM algorithm with line search (FRZ-SGM). In panel C, the red line is the TDA result, and the reference (grey line) is computed with EOM-CCSD(fT) and the cc-pVTZ basis.

We hence conclude that ALMO and FRZ provide excellent initial guesses for ICT states, and SGM provides a reliable convergence to the targeted states once the constraints are lifted in a variational optimization. We further compare the OO-DFT results to TD-DFT calculations (red line) and an accurate equations-of-motion coupled-cluster (EOM-CC) reference calculation (gray line) in panel C of Fig. 3. The reference densities (panel B in Fig. 3) were computed with EOM-CCSD, and the energies (panel C) were adjusted by perturbative triples correction with the EOM-CCSD(fT) method. All DFT methods are red-shifted by an offset, but as Table 1 shows, the OO-DFT methods yield a much smaller mean error (0.64 eV) than TDA (1.27 eV). The improvement achieved by the OO-DFT method is attributed to the accuracy of the range-separated hybrid XCF, combined with an algorithm that ensures convergence to the targeted state. The approximation of weakly interacting fragments introduced in the ALMO method is invalid for short donor–acceptor separations, such that the stabilization achieved upon variational optimization for this situation with significant orbital overlap between the fragments is not captured. The statistical errors for the constrained ALMO results

Table 1 Mean signed error (MSE) and mean signed variance (MSV) on the ICT excitation energy for the tetrafluoroethylene-ethylene dimer system R_{DA} scan for each excited-state DFT method. All calculations are performed with the LRC-ωPBE XCF and the def2-TZVP basis. All values reported in eV unit

Method	Mean error (eV)	Mean variance (eV)
TDA	−1.27	0.043
ALMO	−0.53	0.031
FRZ guess	−0.77	0.062
FRZ-SGM	−0.64	0.005

can hence look better than the results of the OO-DFT methods (here: FRZ-SGM) but the former does not correspond to saddle points on the electronic hypersurface and are therefore not valid excited states. The TDA method gives the largest mean error and mean variance in Table 1. We observe, however, a very small variance for FRZ-SGM, meaning that the energies deviate from the reference energies by a rather constant shift. A systematic offset can be beneficial compared to a statistical error in the prediction of deactivation paths if all excited states are affected by the same shift, resulting in fortuitous error compensation. Interestingly, the FRZ energy (green dotted line) lies lower than the target stationary points at short distances, indicating that the saddle-point search must proceed uphill along the electron–hole MO rotation, which is possible since we optimize the squared gradient. Furthermore, the D_{CT} descriptor (panel B) shows how FRZ-SGM is the only method correctly describing the electron–hole interaction at 3.5 Å separation, deviating from the linear trend just as the reference method.

We conclude that when started with a suitable initial guess, the SGM method converges reliably to the desired stationary point on the excited-state surface. In combination with a suitable range-separated hybrid XCF, OO-DFT methods yield a small mean error and variance with respect to highly accurate reference calculations.

4.2 Charge-transfer excitations in a donor–acceptor Pd coordination cage

Since our goal is to calculate ICT in large supramolecular systems, we require scalable excited-state electronic-structure methods and identified FRZ-SGM as a suitable approach. In this section, we focus on a supramolecular cage structure, where four banana-shaped ligands coordinate two square-planar Pd ions.[73] The Clever group pioneered the design of such supramolecular coordination cage complexes by integrating dyes into the structural elements of the supramolecule, and investigated an interlocked double-cage with electron-rich phenothiazine (PTZ) and electron-deficient anthraquinone (ANQ) dyes in the walls.[10] The ligands used for the preparation of the Pd cage are displayed in Fig. 4 (panel A, note that the R residue in the PTZ ligand is a methyl group in our calculations). The formation of a PTZ-ANQ charge-transfer excitation upon irradiation with a 385 nm photon was investigated by time-resolved spectroscopy and spectro-electrochemistry. We consider this a very relevant system to compute with scalable electronic-structure methods, due to the presence of multiple CT excitations involving

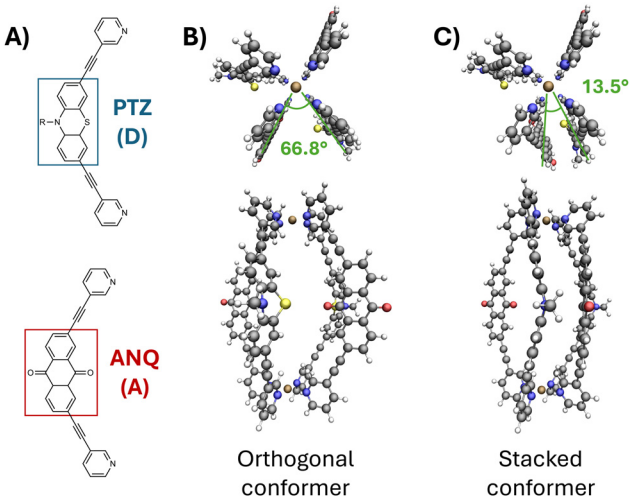

Fig. 4 Panel A shows the structures of the dye–bearing ligands with dye cores highlighted in both the donor (D) and acceptor (A) ligand. In our calculations, R is a methyl group. Panels B and C: stable $[Pd_2(PTZ)_2(ANQ)_2]_{trans}^{4+}$ coordination cage conformations obtained by optimizations with ωB97X-D, def2-SVP basis set, and def2-ECP effective core potentials. Optimization of an initial guess with 90° angles between the ligands leads to the orthogonal conformer in panel B, whereas a stable stacked conformer is obtained by reducing the D–A dihedral and is shown in panel C. All structures are displayed along the Pd–Pd axis (axial view, top) and from the side (bottom).

the organic dyes and the metal ions, which are affected by changes in the cage conformation. For simplicity, we do not compute the interlocked double-cage structure and focus on a single coordination cage with two PTZ and two ANQ dyes in *trans* conformation, $[Pd_2(PTZ)_2(ANQ)_2]_{trans}^{4+}$. The PTZ and ANQ ligands were built using Avogadro[74] 1.2 and were positioned on perpendicular planes to saturate the coordination sphere of two square-planar Pd ions. In the following, we refer to PTZ as the "donor" ligand D and ANQ as the "acceptor" ligand A. The cage structure was optimized with ωB97X-D in a def2-SVP basis, and def2-ECP effective core potentials. Optimization of the "orthogonal" guess structure relaxed to a local minimum, with well-separated dyes (Fig. 4, panel B). An alternative conformation (Fig. 4, panel C) was prepared by slightly tilting the dihedral angle defined by the dye planes and reoptimizing the structure, where the PTZ and ANQ ligands participated in π−π stacking interaction. The electronic energy of the stacked and orthogonal conformers differs by 0.61 kcal mol^{-1}, so we assume both conformers to be accessible at room temperature. In the first section of our results, TD-DFT and OO-DFT methods are compared by computing the charge-transfer excitations in the D–A dimer extracted from each cage conformation, neglecting the Pd ions and the dye pair not involved in the excitation. Additional conformations interpolating the orthogonal and stacked forms were constructed by manipulation of the *Z*-matrix using Molden,[75] producing seven intermediate conformations, for a total of nine conformers. The D–A dihedral angles are measured by fitting a plane through each dye; since the PTZ ligand is not fully planar, the measured dihedral is 66.8° in the orthogonal conformer, whereas the stacked conformer is characterized by a D–A dihedral of 13.5°. The dihedral angles for the intermediates measured in this way are 31.3, 36.3, 41.3, 46.3, 51.3, 56.3, and 61.3°. In the second section of the results, low-lying charge-transfer excitations of the whole Pd cage system are computed for a set of interpolated structures connecting the stacked and orthogonal conformers. These interpolated structures were obtained using the Atomistic Simulation Environment,[76,77] employing the Image-Dependent Pair Potential[77] (IDPP) algorithm.

Fig. 5 shows the TDA results obtained with the PBE0[64] XCF for the lowest-lying ICT excitation of the D–A dimer in the conformations of the PTZ-ANQ cage, as a function of the D–A dihedral angle (and corresponding distance) discussed above. We define the donor–acceptor distance R_{DA} as the distance between nuclei of the dye centers highlighted in panel A of Fig. 4. The D_{CT} magnitude is color-coded, and the D_{CT} is additionally shown separately in the inset. The excitation energy is fitted with the asymptotic expression

$$E = a - \frac{b}{R_{DA}}, \tag{11}$$

where a and b are fitting parameters, and a can be identified with $IP_D - EA_A$, according to Mulliken's formula.[78] D_{CT} is fitted linearly as in

$$D_{CT} = kR_{DA} + q, \tag{12}$$

where k and q are fitting parameters. The fitted parameters are reported in the respective Figures. Data points marked with an orange square are outliers excluded from both fits. An isosurface plot of the density difference w.r.t. the electronic ground state is displayed at the bottom of Fig. 5 for all conformations, using an isovalue −0.0005 for the red surface and isovalue 0.0005 for the blue surface. Overall, the results obtained with

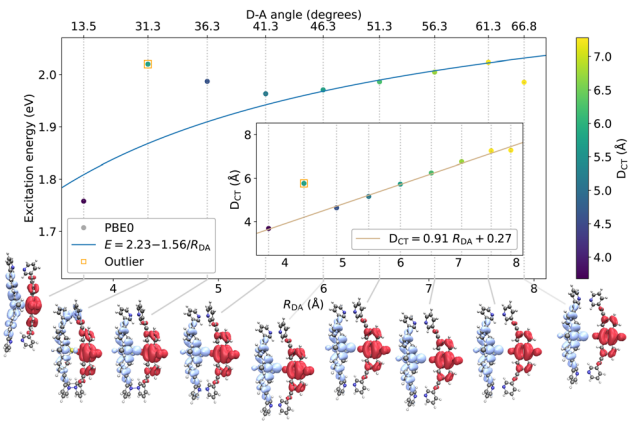

Fig. 5 Lowest-lying ICT excitation energies of several structures with varying dihedral angle between donor and acceptor obtained by interpolating the orthogonal and stacked structures. The density difference between the ground and ICT excited state is shown at the bottom as an isosurface plot (isovalue 0.0005 for the blue surface and −0.0005 for the red surface). The value of D_{CT} is plotted in the inset and is used to color-code the scatter plots. The energies are fitted with a horizontal asymptote, and the D_{CT} plot is fitted with a linear expression. Data points marked with an orange square are identified as outliers and excluded from both fits.

Table 2 Results of the fit over the energy expression in eqn (11) for the lowest-lying ICT excitation in the PTZ-ANQ dimer over the D–A dihedral scan for different methods and XCFs

Method	XCF	$a = IP_D - EA_A$ (eV)	b (eV Å^{-1})
TDA	PBE0[64]	2.23	1.56
TDA	LRC-ω_{GDD}PBE[61,62]	4.64	8.02
TDA	LRC-$(2/D_{CT})$PBE[63]	5.38	15.11
ALMO-SGM	ωB97X-D[66]	4.18	6.24
FRZ-SGM	ωB97X-D[66]	4.22	5.87

PBE0 demonstrate that the global hybrid XCF captures only a fraction of the hole–electron attraction, blue-shifting the ICT excitation energy by only 0.2 eV over the dihedral scan. The fit shows that $a = IP_D - EA_A$ is significantly underestimated compared to the more accurate methods discussed below, which is expected from a global-hybrid XCF approximation.[79,80] Values for $IP_D - EA_A$ are reported in Table 2 for the DFT methods investigated here, alongside the values for b in eqn (11). The incomplete description of the electron–hole attraction by the global hybrid XCF affects all data points in the R_{DA} scan, and the resulting curve is not as steep as other methods, evident by the smaller b.

Fig. 6 shows the results of the dihedral scan for the remaining electronic-structure methods in Table 2. Panel A is obtained by applying the GDD reparametrization scheme, tuning ω_{GDD} for each conformation in the scan. The resulting curve is smooth, with an asymptote that is blue-shifted by more than 2 eV compared to the non-range-separated XCF. The data point corresponding to 56.3 dihedral is excluded from the fit, since the calculated D_{CT} value strongly deviates from the expected linear trend.

Results obtained from the empirical parametrization[63] suggested by Yan *et al.* are shown in panel B of Fig. 6. The 40 lowest-lying singlet excited states are computed with the PBE0 XCF (results reported in Fig. 5), and the state with the maximum D_{CT} is used to compute ω^*. The resulting curve is steeper than for the previous results and results in the largest $IP_D - EA_A$ and b obtained for all methods. The data points corresponding to the stacked conformer were removed from the fit, but the obtained energies scatter slightly around the fitted trend. To better understand the reason for the smoother curve obtained with the ω_{GDD}-tuning procedure as compared to the tuning based on the maximum D_{CT}, we report the respective range-separation parameters in Table 3. The negligible variation of ω_{GDD} along the scan explains the smooth and monotonous curve obtained with this tuning procedure. On the contrary, major changes of D_{CT}^{max} result in strong variations of ω^* for neighboring data points in the scan. Notably, for large donor–acceptor separations, the results are almost identical. Future work might reveal if empirical tuning based on the D_{CT} descriptor in the asymptotic limit can generally improve the results. In conclusion, for systems where conformational changes can significantly affect the energy of CT excitations, size-consistent reparametrizations like ω_{GDD} are recommended for accurate TD-DFT energetics of ICT states.

Finally, we also investigated the performance of the ALMO-SGM and FRZ-SGM OO-DFT methods. Tight convergence of the

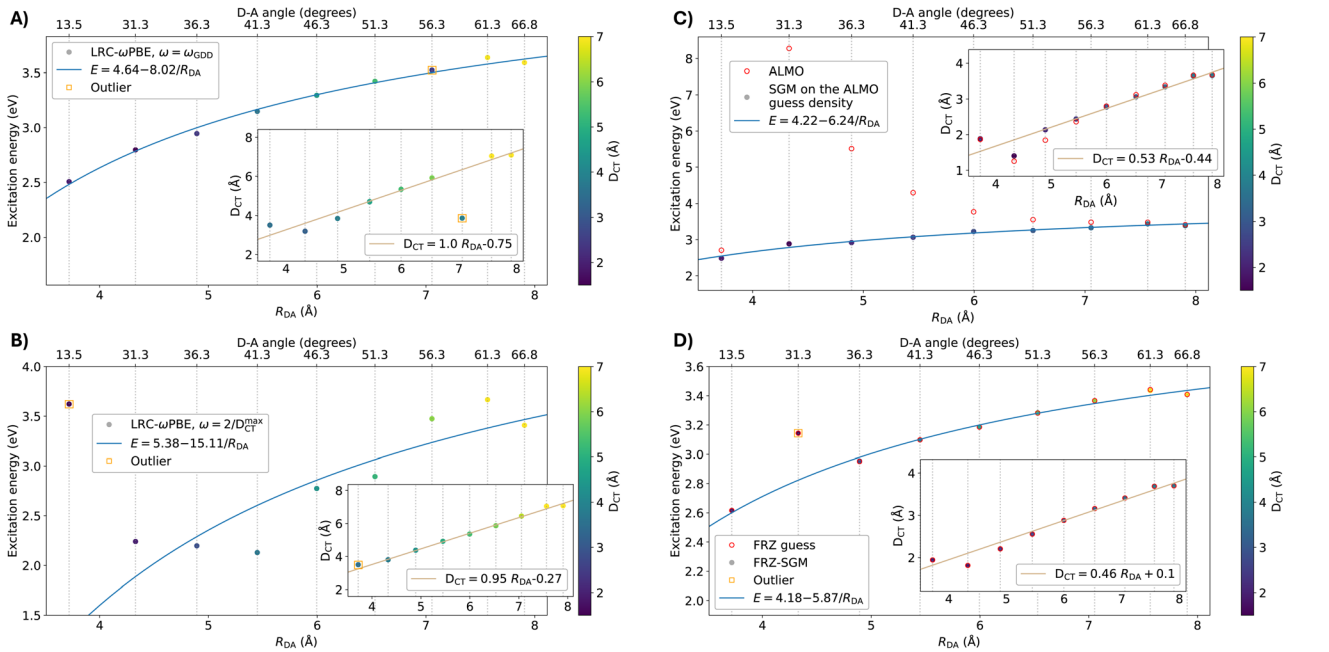

Fig. 6 Lowest-lying ICT excitation energies of several structures with varying dihedral angle between donor and acceptor obtained by interpolating the orthogonal and stacked structures. The ICT excitation energies are computed with TDA, using ω_{GDD} or a D_{CT}-based tuning in panels A and B, respectively. The results from the OO-DFT methods ALMO-SGM and FRZ-SGM are shown in panels C and D, respectively. In the plots of the OO-DFT methods, the initial guess energy and D_{CT} are shown as red circles for each structure. The value of D_{CT} is plotted in the inset and is used to color-code the scatter plots. The energy plot is fitted with a horizontal asymptote, and the D_{CT} plot is fitted with a linear expression. Data points marked with an orange square are identified as outliers and excluded from both fits.

Table 3 ω_{GDD} [61] values obtained applying the general density-dependent reparametrization, ω^* obtained using the parametrization introduced by Yan *et al.*,[63] and D_{CT}^{max} over the 40 lowest-lying roots computed with the PBE0 XCF for the ANQ-PTZ dimer for the same conformations as in Fig. 5 and 6

D–A dihedral θ (°)	ω_{GDD} (Å^{-1})	$\omega^* = 2/D_{CT}^{max}$ (Å^{-1})	D_{CT}^{max} (Å)
13.5	0.242	0.467	4.28
31.3	0.240	0.152	13.12
36.3	0.240	0.153	12.82
41.3	0.242	0.142	14.12
46.3	0.239	0.188	10.63
51.3	0.237	0.187	10.72
56.3	0.235	0.229	8.74
61.3	0.234	0.237	8.45
66.8	0.233	0.215	9.31

SGM algorithm can be challenging, because converging the gradient to a sufficient degree requires very tight convergence of the squared gradient. In our calculations, we used a rather loose convergence criterion of 10^{-4}, compared to the threshold of 10^{-5} used for single-point calculations. The results are displayed in panels C and D of Fig. 6 for the ALMO-SGM and FRZ-SGM methods, respectively. The guess energy and D_{CT} descriptors are shown as red circles, whereas the fully converged results are displayed as solid dots. The ALMO guess energy increases when bringing the D and A monomers closer, in contrast to the expected asymptotic trend. Clearly, however, the ALMO density is a suitable guess for further optimization with the SGM method. This is demonstrated by the inset of Fig. 6 panel C, where the results obtained from the guess density are close to the converged results and by the fact that the SGM optimization relaxed smoothly all data points, stabilizing some by several eVs (note the larger energy axis range as compared to the other panels). Conversely, the constrained optimization (FRZ method, Fig. 6 panel D) produces both energies and densities very close to the target stationary state. The asymptotic trend is well reproduced in the dihedral scan for both OO-DFT methods, with the sole exception of the data point corresponding to a dihedral angle of 31.3°, which was found to be problematic also for other methods. Evidently, the density obtained by the constrained FRZ optimization is already very close to the target state, meaning most of the density relaxation involves MOs that are not directly involved in the excitation, but are influenced by the charge separation in the CT state through electrostatic interaction. The ALMO-SGM and FRZ-SGM methods lead to almost identical values for $IP_D - EA_A$ in the fit (rows 4 and 5 in Table 2), within 0.5 eV of the results obtained with the LRC-ω_{GDD} PBE reparametrization. As a conclusion, we recommend the FRZ-SGM OO-DFT method for the calculation of ICT excitations in large D–A dimers, and the LRC-ω_{GDD}PBE reparametrization for TD-DFT calculations.

We then applied the FRZ-SGM method to the calculation of D–A CT excitations in the full PTZ-ANQ Pd coordination cage. The Pd coordination environment is conserved in the structures interpolated as described above and we computed the CT excitation energies for nine structures of the D–A dihedral scan. The D–A dihedral in Fig. 7 is obtained from planes fitted only through the dye cores (secondary *x*-axis). By analogy, the R_{DA} separation (primary *x*-axis) is computed as the distance of the center of nuclear charge of the D and A dye cores. The resulting dihedral spans a wider range compared to ligand-only study in Fig. 6, from −9.6° in the stacked conformer to 77.6° in the orthogonal conformer. The FRZ-SGM method converged reliably to the lowest-lying CT excitation involving the D and A dyes (circles). For the full cage, we decided to fit the calculated ICT energies to a shifted expression

$$E = a - \frac{b}{R_{DA} - c}, \tag{13}$$

where the *c* parameter describes a shift of the *x*-axis intercept. We specifically set c = 3.4 Å which is still smaller than the D–A separation R_{DA} measured in the stacked conformer. The IP-EA asymptote is lower by 0.5 eV compared to the D–A ligand-only calculations, so we expect that the CT excitation in the stacked conformation is accessible with a 385 nm pump pulse, as demonstrated in ref. 10. This means confinement of the D and A monomers in the supramolecular cage enables the formation of PTZ-ANQ CT excitations, but the cage environment in the stacked conformation slightly reduces the energy of the PTZ-ANQ CT excitation. The D_{CT} plot in the inset of Fig. 7 shows the expected monotonous linear increase with the D–A separation. No data points had to be excluded from the energy and D_{CT} fits.

We also computed ''pure'' LMCT excitations (red star marker in Fig. 7) for the stacked and orthogonal conformers, which we unambiguously identified based on a visualization of the difference density. Keeping in mind all uncertainties of our method, our calculations predict LMCT excitations in the same energy range as the D–A CT excitations, both in stacked and orthogonal conformations. Hence, we cannot rule out excited-state relaxation to the LMCT excitations following the experimentally applied 385 nm pulse.

Finally, we estimate the likelihood of inter-cage CT within the interlocked cage dimer. Starting from the crystallographic data reported in ref. 10, we isolated an ANQ-ANQ dimer in a stacked conformation. We then swapped one ANQ ligand for a PTZ ligand, aligning the N atoms to the Pd coordination sites in a π-stacked conformation. The process we followed to generate this stacked conformation as an approximation to the interlocked cage structures is summarized in Fig. 8. The ALMO-SGM calculation shows the presence of a charge-transfer excitation at 3.81 eV, confirming the presence of inter-cage excitations in the energy range of intra-cage CT excitations. However, to confirm the assignment of the PTZ-ANQ excitation to an inter-cage CT would require the computationally prohibitively expensive calculation of the full interlocked cage dimer, since in our model calculation, this ICT state would not be accessible with a 385 nm pulse. In conclusion, we confirm the presence of low-lying D–A CT and LMCT excitations in the supramolecular Pd coordination cage compounds. We find inter-cage CT excitations in an energy range compatible with the cage

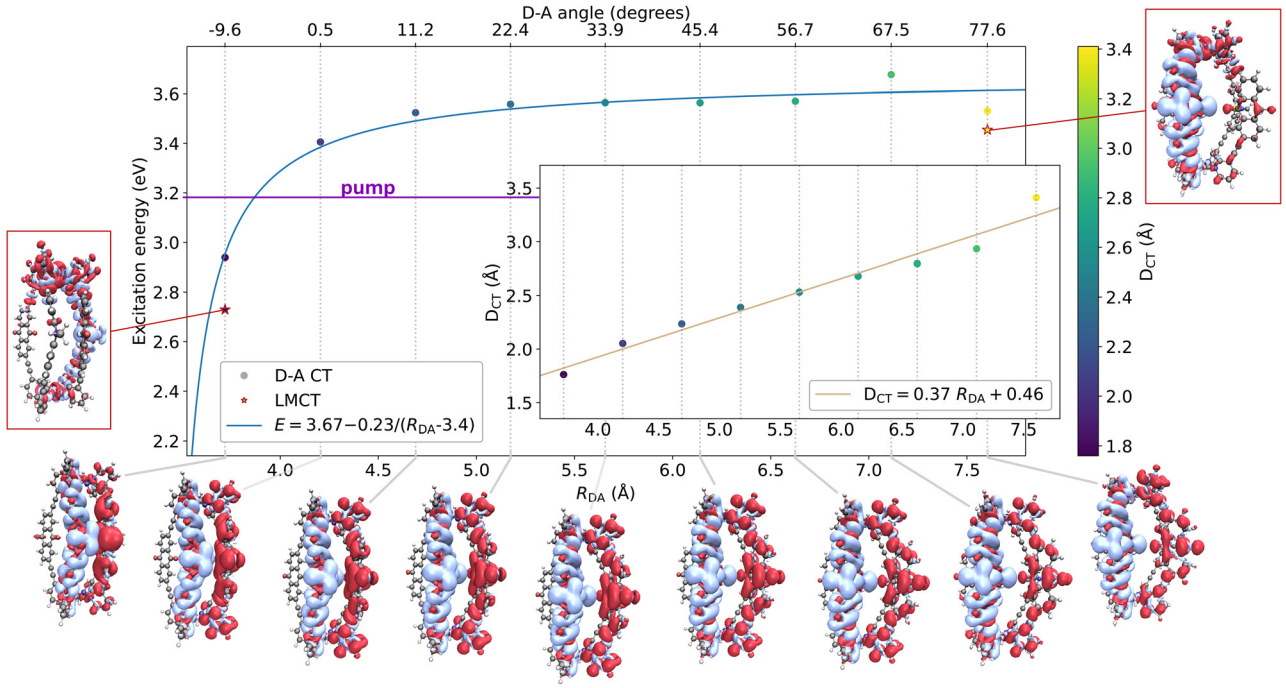

Fig. 7 D–A CT excitations calculated with FRZ-SGM for the full cage structure as a function of the donor–acceptor distance. The density difference between the ground and ICT excited state is shown at the bottom of each panel as an isosurface plot (isovalue 0.0005 for the blue surface and −0.0005 for the red surface). The energies are fitted with a shifted asymptote (eqn (13)). The D–A CT excitations are marked with dots, and additionally calculated a ligand-to-metal CT excitation are marked with a red-framed star. Inset: D_{CT} values plotted against the donor–acceptor separation distance R_{DA} and fitted with a linear expression.

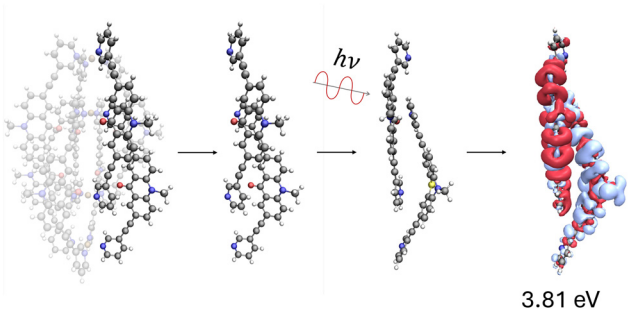

Fig. 8 Starting from the crystal structure in ref. 10 an ANQ molecular dimer is isolated with monomers coming from two different cages of the interlocked structure. One of the ANQ ligands (acceptor) is then swapped for a PTZ (donor) ligand, preserving the local N-Pd coordination environment. A calculation of the PTZ-ANQ ICT excitation in this conformation is then performed with the ALMO-SGM method. The calculated excitation energy of 3.81 eV, is in the same energy range predicted for CT excitations in the single cage-derived donor–acceptor conformations (see Fig. 7). The predicted ICT excitation energy lies above the energy of the pump pulse used in ref. 10.

conformations, but cannot exclude the formation of LMCT excitations in the excited-state deactivation process.

4.3 Charge-transfer excitations in dye–TiO_2 complexes

We also tested the FRZ-SGM method on CT excitations in prototypical dye–semiconductor complexes of technological interest for their application in dye-sensitized solar cells (DSSC). The injection of an excited electron from the organic dyes was previously simulated for the D102-TiO_2 and JK2-TiO_2 dye–semiconductor complexes by Gemeri *et al.* using TD-DFT.[11] We used the molecular structures of the dye–TiO_2 complexes reported in ref. 11. These systems are chemically different from the supramolecular cages since the charge acceptor is a semiconductor modeled as a cluster. The Orca software package[70] was used to perform a preliminary simplified TDA (sTDA) calculation of the low-lying excitations in the dye–semiconductor complexes, using the ωB97X-D XCF and def2-TZVP basis. Low-lying CT excitations were then optimized using the FRZ-SGM method and the same XCF and basis set. The sTDA and FRZ-SGM results are compared in Table 4.

Table 4 Excitation energies for low-lying charge-transfer excitations in D102-TiO_2 and JK2-TiO_2 dye–semiconductor complexes, computed with the sTDA method in Orca and the FRZ-SGM method in Q-Chem. All energies are given in eV and D_{CT} is given in Å

Dye	sTDA	FRZ-SGM	D_{CT}	Class
D102	3.48	5.45	3.21	Dye–TiO_2
JK2	2.84	2.56	6.01	Dye–spacer
JK2	4.03	3.17	5.90	Dye–spacer
JK2	4.61	3.50	7.45	Dye–TiO_2
JK2	4.69	3.57	7.41	Dye–TiO_2
JK2	4.65	3.23	6.59	Dye–TiO_2
JK2	4.75	3.76	7.48	Dye–TiO_2
JK2	4.80	3.82	7.52	Dye–TiO_2
JK2	4.95	3.72	7.33	Dye–TiO_2
JK2	4.99	3.91	7.61	Dye–TiO_2

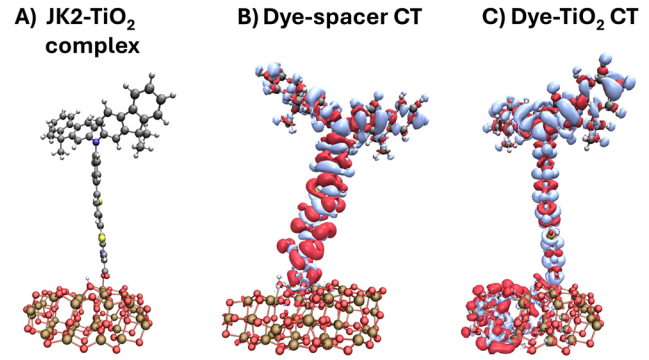

Fig. 9 Structure of the JK2-TiO_2 dye–semiconductor complex (panel A). Isosurface plot of the density differences for dye–spacer (panel B), and dye–TiO_2 (panel C) CT excitations of the JK2-TiO_2 complex computed with the FRZ-SGM method. The isosurface plot is obtained by using an isovalue of 0.0005 for the blue surface and −0.0005 for the red surface.

One direct dye–TiO_2 excitation was calculated with the FRZ-SGM method for the D102-TiO_2 complex, which appears strongly blue-shifted compared to the sTDA result. On the contrary, all excitations computed with FRZ-SGM in the JK2-TiO_2 complex are slightly red-shifted in comparison to the sTDA result. The JK2 ligand is characterized by a long aromatic "spacer" connecting the dye to the semiconductor. By inspecting the density differences, we classified all CT excitations computed in Table 4 as "dye–spacer" or "dye–TiO_2". The JK2-semiconductor complex shows two groups of CT states: dye–spacer excitations reside in the visible spectrum, and dye–TiO_2 excitations, resulting in the direct injection of a dye electron in the semiconductor band, populate the near-UV spectral region. Fig. 9 shows prototypical density difference plots for both groups of CT states.

5 Conclusions

We assessed initial guess densities for intermolecular charge-transfer excitations in the tetrafluoroethylene-ethylene dimer, a challenging system for convergence also used by Schmerwitz *et al.*[32] We tested all initial guesses with the IMOM and SGM orbital-optimized methods. The frequently applied ground-state orbital initial guess led to severe convergence problems for the CT excitations studied. In contrast, starting the variational optimization from localized orbitals obtained from the ALMO method for the ionized D^+–A^- system led to stable convergence to the target CT state. Unfortunately, this strategy cannot be applied to general intramolecular CT states since ALMO localization is not possible for fragments connected by covalent bonds. As a remedy, we have developed a refinement procedure that selectively optimizes the ground-state orbitals not involved in the excitation, while keeping the orbitals involved in the excitation frozen (FRZ guess). This approach is analogous to the constrained-optimization strategy introduced by Schmerwitz *et al.*[32] Coupling the FRZ guess with the SGM algorithm ("FRZ-SGM" method) yields very accurate ICT excitation energies, and the absolutely localized molecular orbital guess paired with SGM ("ALMO-SGM") performs equally well. We therefore recommend ALMO-SGM when the system can be cleanly partitioned into donor and acceptor fragments that are not connected by a covalent bond, whereas FRZ-SGM is the method of choice for intramolecular CT states. We further compared these two OO-DFT methods to system-specific reparametrizations of the LRC-ωPBE XCF in the TDA calculation of ICT for several conformers of a donor–acceptor ligand system which was cut out from a supramolecular cage structure. Both the ALMO and FRZ guess densities were shown to be adequate, with the FRZ guess being closer to the target energy. The global density-dependent tuning[61,62] of the range-separation parameter ω_{GDD} also performed very well for these systems. In contrast, the parametrization based on the excited-state density descriptor D_{CT}[63] results in irregular ω-values and ICT energies. We hence recommend the size-extensive LRC-ω_{GDD}PBE method for the calculation of ICT excitations with TDA or suggest the D_{CT}-based tuning procedure to only take the asymptotic value at large distances into account.

The promising FRZ-SGM method was used for the calculation of CT excitations in three large supramolecular systems. Our calculations of PTZ-ANQ excitations in the full coordination cage investigated experimentally by Frank *et al.*[10] predicted excitation energies in the visible range for the stacked conformation, as well as LMCT excitations. To confirm the true character of the experimentally observed excitations, further calculations of the intercalated double-cage dimer are required but are computationally very demanding even for the scalable methods identified in this work. For a quantitative comparison with experiment, adequate treatment of the dynamics and solvation effects is required, but beyond the scope of this work. Nevertheless, our calculated 2.9–3.5 eV window is consistent with the excitation energy pulse of 3.2 eV in the study from Frank *et al.*, thereby providing qualitative validation. Finally, CT excitations in two covalently bonded dye–semiconductor complexes, D102-TiO_2 and JK2-TiO_2,[11] were computed with the sTDA and FRZ-SGM methods. FRZ-SGM reliably converged on the CT excitations also for these systems.

We demonstrate that, when combined with a robust guess refinement strategy, OO-DFT methods are an adequate electronic-structure method for the study of photoinduced phenomena in large supramolecular systems. Being grounded in a rigorous theoretical framework, this method holds the promise to allow for an accuracy that is comparable to ground-state DFT with adequate functional development. With the stable convergence to a desired target state that we demonstrate here, OO-DFT is poised to be the method of choice for the study of excited states in large molecular systems. Necessary extensions for a more direct comparison with experiment, such as the implementation of non-adiabatic coupling vectors and combinations with QM/MM and solvation models, are currently underway in our groups. These developments will lead to an OO-DFT framework that allows for the simulation of excited-state dynamics and a study of photo-deexcitation processes for such large molecular systems that are currently out of reach. TD-DFT calculations in the Tamm–Dancoff approximation

proved to be accurate enough for qualitative initial investigations or to set up subsequent OO-DFT calculations if the system-specific optimization of the range-separation parameter followed the global density-dependent tuning procedure. Further development is desirable for the convergence acceleration of the SGM algorithm.

Conflicts of interest

M. H. G. is a part-owner of Q-Chem Inc., whose software was used for all developments and calculations reported here.

Data availability

Structures required to reproduce all data in this study are available in the ESI.† The FRZ initial guess implementation will be part of a future release of the Q-Chem electronic structure package.

Acknowledgements

N. B. thanks Gabriel Jonathan Flür for performing initial calculations on the dye–TiO_2 systems. Financial support by the DFG under Germany's Excellence Strategy – EXC 2089/1-390776260 (e-conversion) is gratefully acknowledged. A research stay of N. B. at UC Berkeley was funded by the Deutsche Forschungsgemeinschaft (DFG) through the Research Training Group "Confinement Controlled Chemistry" (GRK 2376). Work at Berkeley was funded by the DOE Office of Science, Basic Energy Science (BES) Program, Chemical Sciences, Geosciences and Biosciences Division under Contract no. DE-AC02-05CH11231, through the Atomic, Molecular, and Optical Sciences program.

References

1 H. J. Levinson, *Jpn. J. Appl. Phys.*, 2022, **61**, SD0803.
2 H. J. Levinson, *J. Micro/Nanopatterning, Mater., Metrol.*, 2025, **24**, 011005.
3 G. Hills, C. Lau, A. Wright, S. Fuller, M. D. Bishop, T. Srimani, P. Kanhaiya, R. Ho, A. Amer and Y. Stein, *et al.*, *Nature*, 2019, **572**, 595–602.
4 A. D. Franklin, M. Luisier, S.-J. Han, G. Tulevski, C. M. Breslin, L. Gignac, M. S. Lundstrom and W. Haensch, *Nano Lett.*, 2012, **12**, 758–762.
5 R. La Rocca, K. Kato, P.-C. Tsai, Y. Nakajima, F. Akita and J.-R. Shen, *Nat. Commun.*, 2025, **16**, 4175.
6 Z. Mao, X. Li, Z. Li, L. Shen, X. Li, Y. Yang, W. Wang, T. Kuang, J.-R. Shen and G. Han, *Nat. Commun.*, 2024, **15**, 4535.
7 N.-F. Bravo, C. Cifuentes, M.-C. Ros, L.-J. Pérez and J. Portilla, *Asian J. Org. Chem.*, 2025, **14**, e202400742.
8 N. Dallaire, N. T. Boileau, I. Myers, S. Brixi, M. Ourabi, E. Raluchukwu, R. Cranston, H. R. Lamontagne, B. King and B. Ronnasi, *et al.*, *Adv. Mater.*, 2024, **36**, 2406105.
9 C. Kunkel, J. T. Margraf, K. Chen, H. Oberhofer and K. Reuter, *Nat. Commun.*, 2021, **12**, 2422.
10 M. Frank, J. Ahrens, I. Bejenke, M. Krick, D. Schwarzer and G. H. Clever, *J. Am. Chem. Soc.*, 2016, **138**, 8279–8287.
11 D. Gemeri, J. C. Tremblay, M. Pastore and H. Bahmann, *J. Chem. Phys.*, 2022, **559**, 111521.
12 T. Le Bahers, C. Adamo and I. Ciofini, *J. Chem. Theory Comput.*, 2011, **7**, 2498–2506.
13 S. Hutsch, M. Panhans and F. Ortmann, *npj Comput. Mater.*, 2022, **8**, 228.
14 O. V. Yazyev and S. G. Louie, *Nat. Mater.*, 2010, **9**, 806–809.
15 C. M. Guédon, H. Valkenier, T. Markussen, K. S. Thygesen, J. C. Hummelen and S. J. Van Der Molen, *Nat. Nanotechnol.*, 2012, **7**, 305–309.
16 N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.*, 2016, **144**, 214110.
17 V. Butera, *Phys. Chem. Chem. Phys.*, 2024, **26**, 7950–7970.
18 H. Hosseini, C. J. Herring, C. F. Nwaokorie, G. A. Sulley and M. M. Montemore, *J. Phys. Chem. C*, 2024, **128**, 18144–18157.
19 E. Runge and E. K. Gross, *Phys. Rev. Lett.*, 1984, **52**, 997.
20 S. Hirata and M. Head-Gordon, *Chem. Phys. Lett.*, 1999, **314**, 291–299.
21 J. E. D. Bene, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1971, **55**, 2236–2241.
22 J. B. Foresman, M. Head-Gordon, J. A. Pople and M. J. Frisch, *J. Phys. Chem.*, 1992, **96**, 135–149.
23 A. Dreuw, J. L. Weisman and M. Head-Gordon, *J. Chem. Phys.*, 2003, **119**, 2943–2946.
24 N. Bogo and C. J. Stein, *Phys. Chem. Chem. Phys.*, 2024, **26**, 21575–21588.
25 R. Baer, E. Livshits and U. Salzner, *Annu. Rev. Phys. Chem.*, 2010, **61**, 85–109.
26 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.
27 M. Haasler, T. M. Maier, R. Grotjahn, S. Gückel, A. V. Arbuznikov and M. Kaupp, *J. Chem. Theory Comput.*, 2020, **16**, 5645–5657.
28 H. Bahmann and M. Kaupp, *J. Chem. Theory Comput.*, 2015, **11**, 1540–1548.
29 A. Görling, *Phys. Rev. Lett.*, 1996, **54**, 3912.
30 A. Görling, *Phys. Rev. A:At., Mol., Opt. Phys.*, 1999, **59**, 3359.
31 W. Yang and P. W. Ayers, *arXiv*, 2024, preprint, arXiv:2403.04604, DOI: **10.48550/arXiv.2403.04604**.
32 Y. L. Schmerwitz, E. Selenius and G. Levi, *arXiv*, 2025, preprint, arXiv:2501.18568, DOI: **10.48550/arXiv.2501.18568**.
33 E. Epifanovsky, A. T. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons and A. L. Dempwolff, *et al.*, *J. Chem. Phys.*, 2021, **155**, 084801.
34 T. Van Voorhis and M. Head-Gordon, *Mol. Phys.*, 2002, **100**, 1713–1721.
35 H. Stoll, G. Wagenblast and H. Preuß, *Theor. Chim. Acta*, 1980, **57**, 169–178.
36 E. Gianinetti, M. Raimondi and E. Tornaghi, *Int. J. Quantum Chem.*, 1996, **60**, 157–166.
37 G. M. Barca, A. T. Gilbert and P. M. Gill, *J. Chem. Theory Comput.*, 2018, **14**, 1501–1509.

38 D. Hait and M. Head-Gordon, *J. Chem. Theory Comput.*, 2020, **16**, 1699–1710.
39 D. Hait and M. Head-Gordon, *J. Phys. Chem. Lett.*, 2021, **12**, 4517–4529.
40 E. Selenius, A. E. Sigurdarson, Y. L. Schmerwitz and G. Levi, *J. Chem. Theory Comput.*, 2024, **20**, 3809–3822.
41 M. A. Rohrdanz and J. M. Herbert, *J. Chem. Phys.*, 2008, **129**, 034107.
42 L. Schmidt-Mende, U. Bach, R. Humphry-Baker, T. Horiuchi, H. Miura, S. Ito, S. Uchida and M. Grätzel, *Adv. Mater.*, 2005, **17**, 813–815.
43 P. Pattanasattayavong, S. Rossbauer, S. Thomas, J. G. Labram, H. J. Snaith and T. D. Anthopoulos, *J. Appl. Phys.*, 2012, **112**, 074507.
44 S.-Q. Fan, B. Fang, H. Choi, S. Paik, C. Kim, B.-S. Jeong, J.-J. Kim and J. Ko, *Electrochim. Acta*, 2010, **55**, 4642–4646.
45 J.-H. Yum, S.-R. Jang, P. Walter, T. Geiger, F. Nüesch, S. Kim, J. Ko, M. Grätzel and M. K. Nazeeruddin, *Chem. Commun.*, 2007, 4680–4682.
46 S. Agrawal, T. Leijtens, E. Ronca, M. Pastore, H. Snaith and F. De Angelis, *J. Mater. Chem. A*, 2013, **1**, 14675–14685.
47 M. Marsili, E. Mosconi, F. De Angelis and P. Umari, *Phys. Rev. B*, 2017, **95**, 075415.
48 M. A. Marques, N. T. Maitra, F. M. Nogueira, E. K. Gross and A. Rubio, *Fundamentals of time-dependent density functional theory*, Springer, 2012, vol. 837.
49 P. Pulay, *Chem. Phys. Lett.*, 1980, **73**, 393–398.
50 P. Pulay, *J. Comput. Chem.*, 1982, **3**, 556–560.
51 B. Kaduk, T. Kowalczyk and T. Van Voorhis, *Chem. Rev.*, 2012, **112**, 321–370.
52 M. B. Goldey, N. P. Brawand, M. Voros and G. Galli, *J. Chem. Theory Comput.*, 2017, **13**, 2581–2590.
53 J. Kussmann, Y. Lemke, A. Weinbrenner and C. Ochsenfeld, *J. Chem. Theory Comput.*, 2024, **20**, 8461–8473.
54 A. T. Gilbert, N. A. Besley and P. M. Gill, *J. Phys. Chem. A*, 2008, **112**, 13164–13171.
55 H. G. Burton, *J. Chem. Theory Comput.*, 2022, **18**, 1512–1526.
56 D. Hait and M. Head-Gordon, *J. Phys. Chem. Lett.*, 2020, **11**, 775–786.
57 G. Levi, A. V. Ivanov and H. Jónsson, *J. Chem. Theory Comput.*, 2020, **16**, 6968–6982.
58 Y. L. Schmerwitz, G. Levi and H. Jónsson, *J. Chem. Theory Comput.*, 2023, **19**, 3634–3651.
59 J. J. Mortensen, A. H. Larsen, M. Kuisma, A. V. Ivanov, A. Taghizadeh, A. Peterson, A. Haldar, A. O. Dohn, C. Schäfer and E. Ö. Jónsson, *et al.*, *J. Chem. Phys.*, 2024, **160**, 092503.
60 G. Prokopiou, M. Hartstein, N. Govind and L. Kronik, *J. Chem. Theory Comput.*, 2022, **18**, 2331–2340.
61 M. Modrzejewski, L. Rajchel, G. Chalasinski and M. M. Szczesniak, *J. Phys. Chem. A*, 2013, **117**, 11580–11586.
62 A. Mandal and J. M. Herbert, *Chem. Phys. Lett.*, 2025, **16**, 2672–2680.
63 T. Yan, A. Bonardi, C. Adamo and I. Ciofini, *J. Chem. Theory Comput.*, 2025, **21**(4), 1892–1904.
64 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
65 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
66 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
67 S. Grimme, J. G. Brandenburg, C. Bannwarth and A. Hansen, *J. Chem. Phys.*, 2015, **143**, 054107.
68 R. Sure and S. Grimme, *J. Comput. Chem.*, 2013, **34**, 1672–1685.
69 S. Grimme, *J. Chem. Phys.*, 2013, **138**, 244104.
70 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1606.
71 A. I. Krylov, *Annu. Rev. Phys. Chem.*, 2008, **59**, 433–462.
72 T. H. Dunning Jr, *J. Chem. Phys.*, 1971, **55**, 716–723.
73 M. Han, D. M. Engelhard and G. H. Clever, *Chem. Soc. Rev.*, 2014, **43**, 1848–1860.
74 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 1–17.
75 G. Schaftenaar and J. H. Noordik, *J. Comput.-Aided Mater. Des.*, 2000, **14**, 123–134.
76 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer and C. Hargus, *et al.*, *World J. Condens. Matter Phys.*, 2017, **29**, 273002.
77 S. Smidstrup, A. Pedersen, K. Stokbro and H. Jónsson, *J. Chem. Phys.*, 2014, **140**, 214106.
78 R. S. Mulliken, *J. Am. Chem. Soc.*, 1952, **74**, 811–824.
79 B. G. Janesko, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2009, **11**, 443–454.
80 K. Sohlberg and M. E. Foster, *RSC Adv.*, 2020, **10**, 36887–36896.