JCTC Journal of Chemical Theory and Computation

Editors' Choice

pubs.acs.org/JCTC Article

Importance of Orbital Invariance in Quantifying Electron−Hole Separation and Exciton Size

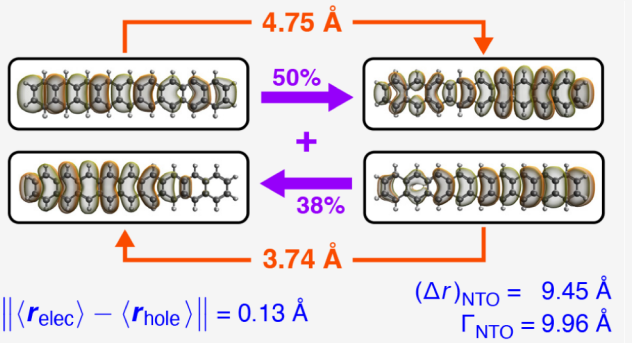
John M. Herbert* and Aniket Mandal

Cite This: J. Chem. Theory Comput. 2024, 20, 9446−9463

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A fundamental tenet of quantum mechanics is that properties should be independent of representation. In self-consistent field methods such as density functional theory, this manifests as a requirement that properties be invariant with respect to unitary transformations of the occupied molecular orbitals and (separately) the unoccupied molecular orbitals. Various *ad hoc* measures of excited-state charge separation that are commonly used to analyze time-dependent density-functional calculations violate this requirement, as they are based on incoherent averages of excitation amplitudes rather than expectation values involving coherent superpositions. As a result, these metrics afford markedly different values in various common representations, including canonical molecular orbitals, Boys-localized orbitals, and natural orbitals. Numerical values can be unstable with respect to basis-set expansion and may afford nonsensical results in the presence of extremely diffuse basis functions. In contrast, metrics based on well-defined expectation values are stable, representation-invariant, and physically interpretable. Use of natural transition orbitals improves the stability of the incoherent averages, but numerical values can only be interpreted as expectation value in the absence of superposition. To satisfy this condition, the particle and hole density matrices must each be dominated by a single eigenvector so that the transition density is well described by a single pair of natural transition orbitals. Counterexamples are readily found where this is not the case. Our results suggest that *ad hoc* charge-transfer diagnostics should be replaced by rigorous expectation values, which are no more expensive to compute.



1. INTRODUCTION

Time-dependent density functional theory (TD-DFT),[1] the most widely used quantum chemistry technique for computing electronic excitation energies, has a simple and conceptually pleasing one-electron interpretation in terms of a particle (the excited electron) and a hole.[2] Excited-state properties can be decomposed into contributions from each of these quasiparticles. In particular, the particle and hole densities can be used to define a mean electron−hole distance ($d_{e\text{-}h}$) in the excited state,[2−8]

$$d_{e\text{-}h} = \|\langle \mathbf{r}_{elec} - \mathbf{r}_{hole}\rangle\| \quad (1)$$

Over the years, a variety of *ad hoc* charge-transfer (CT) diagnostics have been introduced in order to quantify electron−hole separation in TD-DFT calculations,[9−21] as reviewed recently.[2] The original purpose of these analysis tools was to alert the user to the presence of a charge-separated state,[9,20−25] for which TD-DFT with conventional semilocal and hybrid functionals may fail badly.[1,26−30] Remarkably, however, the simple expectation value in eq 1 seems not to have been considered as a diagnostic of CT character in TD-DFT calculations until recently,[31,32] where it was borrowed from careful excited-state analyses in wave function theory.[4−7]

The problem of spurious CT states in TD-DFT is substantially ameliorated using range-separated and long-range corrected density functionals,[30,33−41] yet measures of electron−hole separation and exciton size remain important tools for analysis of excited-state calculations.[42−62] In the present work, we demonstrate that some widely used charge-separation metrics are not invariant under unitary transformations of the occupied or virtual molecular orbitals (MOs). This implies that their numerical values depend on the choice of representation. This has been pointed out before,[11,57,63] but the present work explores the numerical and conceptual implications of this result for the first time. Dependence on the choice of orbitals is not an esoteric point; we will demonstrate substantial numerical differences (including nonsensical results) using several common MO representations.

Received: August 19, 2024
Revised: September 19, 2024
Accepted: September 25, 2024
Published: October 16, 2024

As an antidote to these problems, we suggest orbital-invariant (and thus representation-independent) charge-displacement metrics that attempt to model the same charge-separation physics as earlier *ad hoc* definitions. These invariant metrics are based on well-defined expectation values with respect to the transition density, following a paradigm laid out for wave function calculations by Plasser, Dreuw, and coworkers.[3−8] Herein, we demonstrate that these measures can be used to the same ends as previous CT metrics, namely, to diagnose problems with TD-DFT or to quantify electron−hole separation in a physically meaningful way.

2. THEORY

2.1. TD-DFT. Linear-response TD-DFT calculations involve solution of a non-Hermitian pseudoeigenvalue problem,[2]

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \tag{2}$$

for excitation energies ω and amplitudes $\{\mathbf{x}, \mathbf{y}\}$. Matrices $\mathbf{A}$ and $\mathbf{B}$ are orbital Hessians, as discussed elsewhere.[1,64] Transition amplitudes x_{ia} and y_{ia} indicate excitation and de-excitation, respectively, between an occupied MO ψ_i and an unoccupied (virtual) MO ψ_a.

Importantly, TD-DFT excitation energies and excited-state properties are invariant to unitary transformations of either the occupied or the virtual MOs, so at this point it does not need to be stipulated which MOs (canonical or otherwise) we mean by ψ_i and ψ_a. This allows the freedom to use localized MOs,[65−72] or other transformations that are favorable for visualization or interpretative purposes,[2] without affecting observables.

The de-excitation amplitudes y_{ia} are often neglected, resulting in a Hermitian eigenvalue problem

$$\mathbf{A}\mathbf{x} = \omega\mathbf{x} \tag{3}$$

This is known as the Tamm-Dancoff approximation (TDA)[1,2] and it is typically quite accurate for small-molecule excitation energies,[73] although larger errors have been noted in solids.[74] Importantly, the TDA decouples the excitation energy problem from the ground-state stability problem[1,75] and this is sometimes necessary in photochemical simulations.[76−78] On the other hand, the TDA violates gauge invariance and thus introduces ambiguity for oscillator strengths,[79] which are typically more accurate in full linear response theory.[80] In any case, the CT metrics introduced below can be used with either eq 2 or eq 3, although all calculations reported here are full linear response.

2.2. Conventional CT Metrics. A recent review[2] provides an overview of *ad hoc* CT metrics for TD-DFT calculations, the most common of which are introduced here. Historically, the first of these diagnostics (called Λ) was introduced by Tozer and coworkers,[9,22−25] who defined a metric

$$\Lambda = \frac{\sum_{ia} \kappa_{ia}^2 O_{ia}}{\sum_{jb} \kappa_{jb}^2} \tag{4}$$

where

$$\kappa_{ia} = x_{ia} + y_{ia} \tag{5}$$

and

$$O_{ia} = \int |\psi_i(\mathbf{r})||\psi_a(\mathbf{r})| d\mathbf{r} \tag{6}$$

The integral in eq 6 quantifies spatial overlap between occupied and virtual MOs.

The normalization condition for TD-DFT is[1]

$$\sum_{ia} (x_{ia}^2 - y_{ia}^2) = 1 \tag{7}$$

so the denominator in eq 4 generally differs from unity unless the TDA is invoked,[81] although it does normalize Λ such that $0 \leq \Lambda \leq 1$. Larger values indicate a localized excitation whereas $\Lambda = 0$ indicates complete spatial separation of the excited electron and the hole. Several puzzling failures in early TD-DFT calculations were explained by using Λ to measure "hidden" CT character,[24,25,82] although certain types of charge separation elude this particular metric.[22,63,83]

The Λ metric has undeniable intuitive appeal as an amplitude-weighted sum of donor−acceptor overlaps O_{ia}, yet it is an incoherent average that does not correspond to the expectation value of any measurable (or even potentially measurable) physical quantity. A proper expectation value is expressible as the trace of an operator with a density matrix, as shown in Appendix A. From another point of view, the definition of Λ involves squaring the amplitudes, then constructing an *incoherent* superposition of the quantities O_{ia} associated with $\psi_i \rightarrow \psi_a$ excitation, rather than evaluating an average using coherent superposition (i.e., an expectation value). A consequence is that Λ is not invariant to rotations of the occupied and virtual MOs, even within the TDA. As such, its numerical value depends on the choice of representation.

To the best of our knowledge, the Λ metric has only ever been evaluated using canonical MOs (CMOs) but this choice is arbitrary. Nevertheless, the value Λ_{CMO} in the CMO representation does provide a useful (albeit functional-specific) threshold, beyond which TD-DFT excitation energies should not be taken seriously unless range separation is used.[9,22] For interpretative purposes, a drawback of this dimensionless metric is that it does not quantify electron−hole separation in a way that is easily interpretable or visualizable.

To remedy this, other CT metrics have been suggested that have dimensions of length and are intended to be interpreted as quantitative measures of charge separation. A seemingly obvious choice is to replace O_{ia} in eq 4 with

$$\mathbf{R}_{ia} = \langle \psi_i | \hat{\mathbf{r}} | \psi_i \rangle - \langle \psi_a | \hat{\mathbf{r}} | \psi_a \rangle \tag{8}$$

which is the vector displacement between the centroids of orbitals ψ_i and ψ_a. Using the scalar norm $\|\mathbf{R}_{ia}\|$ of this displacement vector results in a charge-separation metric that has been called Δr,[10,11,17] defined as

$$\Delta r = \frac{\sum_{ia} \kappa_{ia}^2 \|\mathbf{R}_{ia}\|}{\sum_{jb} \kappa_{jb}^2} \tag{9}$$

This amounts to an incoherent average of the charge displacements associated with $\psi_i \rightarrow \psi_a$ excitation. Like Λ, the definition of Δr fails to satisfy orbital invariance so its numerical value depends upon the choice of MOs.

Assuming that the MOs transforms like irreducible representations of the molecular point group, then $\Delta r = 0$ for any centrosymmetric system because the charge displacements $\mathbf{R}_{ia}$ must preserve inversion symmetry in that case. From a symmetry perspective, this is a feature and not a bug. However, it does mean that charge-resonance states[2,3,41,84] (involving counterbalancing contributions from forward and backward CT) cannot be detected by Δr. To obtain a nonvanishing metric for centrosymmetric systems, an "electron displacement" measure

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

$$\Gamma = \Delta r + \Delta\sigma \tag{10}$$

has been suggested.[11] It combines Δr from eq 9 with

$$\Delta\sigma = \frac{\sum_{ia} \kappa_{ia}^2 \sigma_{ia}}{\sum_{jb} \kappa_{jb}^2} \tag{11}$$

which is the weighted sum of the quantities

$$\sigma_{ia} = |\langle\psi_i|\hat{r}^2|\psi_i\rangle - \langle\psi_a|\hat{r}^2|\psi_a\rangle|^{1/2} \tag{12}$$

Here, $\langle\psi_r|\hat{r}^2|\psi_r\rangle$ is the second moment of orbital ψ_r, so σ_{ia} is a measure of the disparity between the size of the excited electron and the size of the hole, for the excitation $\psi_i \rightarrow \psi_a$.

Both Δr and $\Delta\sigma$ involve incoherent sums, rather than constructing a transition density and evaluating a proper expectation value for the excited state in question. As such, these quantities fail to preserve orbital invariance, and this failure is inherited by Γ in eq 10. Nevertheless, Γ has been used to define a critical threshold or "trust radius" for TD-DFT excitation energies,[10] which is similar to the manner in which Λ has been used but with an ostensibly more direct physical interpretation, since Γ has units of length. It was observed in ref 11 that Γ has a different numerical value in the CMO basis as compared to the basis of natural transition orbitals (NTOs), although no explanation was provided. (See Appendix B for a brief overview of NTOs and ref 2 for more complete discussion.). Furthermore, Δr was seen to be sensitive to the presence of diffuse basis functions.[11] The present work will provide a physical and mathematical explanation for these observations.

2.3. Invariant Measures of CT. Lack of orbital invariance in metrics such as Λ, Δr, $\Delta\sigma$, and Γ is problematic in applications, as shown by examples that are discussed in Section 4. As such, we propose alternative means to measure charge displacement that are based on proper expectation values and thus invariant to unitary transformations of the MOs. Some of these were introduced previously by Plasser, Dreuw, and coworkers,[4−7] and can be defined also for correlated wave function models.[4]

Density matrices for the excited electron and the hole (eq A3), constructed from the amplitudes $\mathbf{x}$ and $\mathbf{y}$, are introduced in Appendix A. In position space, these density matrices correspond to (unrelaxed) difference densities $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$, which together afford the difference density with respect to the ground state:

$$\Delta\rho(\mathbf{r}) = \Delta\rho_{\text{elec}}(\mathbf{r}) + \Delta\rho_{\text{hole}}(\mathbf{r}) \tag{13}$$

From the density matrices corresponding to $\Delta\rho_{\text{elec}}$ and $\Delta\rho_{\text{hole}}$, expectation values of $\mathbf{r}_{\text{elec}}$ and $\mathbf{r}_{\text{hole}}$ can be evaluated as discussed in Appendix A.

Invariant measures of electron−hole separation and exciton size are obtained from expectation values of the particle and hole coordinates. For example, we define[2]

$$d_{\text{e-h}}^{\pm} = \left\| \int [\Delta\rho_{\text{elec}}(\mathbf{r}) \pm \Delta\rho_{\text{hole}}(\mathbf{r})]\hat{\mathbf{r}}\, d\mathbf{r} \right\| \tag{14}$$

More succinctly, this is

$$d_{\text{e-h}}^{\pm} = \|\langle\mathbf{r}_{\text{elec}}\rangle \pm \langle\mathbf{r}_{\text{hole}}\rangle\| \tag{15}$$

where

$$\langle x_{\text{elec}}\rangle = \int x\Delta\rho_{\text{elec}}(\mathbf{r})d\mathbf{r} \tag{16}$$

is the x component of $\langle\mathbf{r}_{\text{elec}}\rangle$, for example. The quantity $d_{\text{e-h}}^{-}$ is the distance between the centroids of $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$, or in other words, the norm of the expectation value of the intracule coordinate $\mathbf{r}_{\text{elec}}-\mathbf{r}_{\text{hole}}$.[85] It is the same quantity introduced in eq 1, and we henceforth omit the superscript and define

$$d_{\text{e-h}} \equiv d_{\text{e-h}}^{-} \tag{17}$$

The quantity $d_{\text{e-h}}^{+}$ in eq 15 is the expectation value of the extracule coordinate, $\mathbf{r}_{\text{elec}} + \mathbf{r}_{\text{hole}}$.[85] The intracule and extracule coordinates have sometimes been taken to be quasiparticle coordinates for excitons in conjugated polymers,[86−88] but the extracule coordinate will not concern us here.

As elaborated in ref 2, the quantity $d_{\text{e-h}}$ is essentially the same as a metric called "D_{CT}",[13] which is widely used to analyze TD-DFT calculations[13−16,21,42−50] but was introduced in a manner that obscures its connection to the particle and hole density matrices. As such, we prefer the straightforward definition in eq 15 and the intuitive nomenclature of "$d_{\text{e-h}}$".[2] More complicated extensions of D_{CT} have been suggested,[16] though the need for them is unclear to us.

Both Δr and $d_{\text{e-h}}$ are intended to measure electron−hole separation but only the latter is invariant to unitary transformations of the MOs. Like Δr, however, $d_{\text{e-h}}$ vanishes in any centrosymmetric system and this may cause interpretative problems, e.g., for symmetric or near-symmetric arrangements of molecular chromophores.[41] (The change in dipole moment is proportional to $d_{\text{e-h}}$ and also vanishes for centrosymmetric charge-resonance states, even when the forward and reverse CT contributions are individually large.) To circumvent this limitation, the quantity $\Delta\sigma$ was added to Δr (eq 10) to indicate size disparity between the donor and acceptor orbitals. Invariant metrics along the same lines are

$$\sigma_{\text{elec}} = (\langle\mathbf{r}_{\text{elec}}\cdot\mathbf{r}_{\text{elec}}\rangle - \langle\mathbf{r}_{\text{elec}}\rangle\cdot\langle\mathbf{r}_{\text{elec}}\rangle)^{1/2} \tag{18a}$$

$$\sigma_{\text{hole}} = (\langle\mathbf{r}_{\text{hole}}\cdot\mathbf{r}_{\text{hole}}\rangle - \langle\mathbf{r}_{\text{hole}}\rangle\cdot\langle\mathbf{r}_{\text{hole}}\rangle)^{1/2} \tag{18b}$$

These are the root-mean-square (RMS) sizes of the excited electron and the hole, respectively.[2,6,7]

Two other quantities worth mentioning are the "H-index" and the "t-index" introduced by Ciofini and coworkers.[13,14] The former is essentially $(\sigma_{\text{elec}} + \sigma_{\text{hole}})/2$ but restricted to a one-dimensional donor−acceptor coordinate,[13,14,42,43] while $t = D_{\text{CT}} - H$.[13] These quantities will not be addressed directly in the present work, as they do not appear to provide new information beyond what is obtainable using the readily interpretable expectation values defined above, combined in various ways.

To that end, results presented below suggest that the charge-displacement metric

$$d_{\text{CD1}} = d_{\text{e-h}} + |\sigma_{\text{hole}} - \sigma_{\text{elec}}| \tag{19}$$

is a reasonable surrogate for Γ (eq 10) in some instances, yet one that is based on rigorous expectation values and independent of representation. The first term in eq 19 ($d_{\text{e-h}}$) can be seen as a representation-invariant alternative to Δr, while $|\sigma_{\text{hole}} - \sigma_{\text{elec}}|$ encodes any size disparity between the electron and the hole. In previous work,[1,2] it was speculated that the quantity

$$d_{\text{CD2}} = d_{\text{e-h}} - \frac{1}{2}(\sigma_{\text{hole}} + \sigma_{\text{elec}}) \tag{20}$$

might provide a useful alternative to the aforementioned $t = D_{\text{CT}} - H$, although no data were provided. The quantity

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

$$d_{CD3} = d_{e\text{-}h} + d_{exc} \tag{21}$$

was also suggested,[1,2] where

$$d_{exc} = \langle ||\mathbf{r}_{elec} - \mathbf{r}_{hole}||^2 \rangle^{1/2} \tag{22}$$

is the RMS exciton size. The metrics d_{CD1}, d_{CD2} and d_{CD3} are explored for the first time in the present work and d_{CD1} will prove to be particularly useful as an invariant replacement for Γ.

Other invariant CT metrics based upon particle and hole densities have been introduced by Etienne et al.[51−54] These include a charge-overlap index (ϕ_S),[51] defined as

$$\phi_S = \vartheta^{-1} \int [\Delta\rho_{elec}(\mathbf{r})\Delta\rho_{hole}(\mathbf{r})]^{1/2} d\mathbf{r} \tag{23}$$

where $\vartheta = (1/2)\int[\Delta\rho_{elec}(\mathbf{r}) - \Delta\rho_{hole}(\mathbf{r})]d\mathbf{r}$ is a normalizing denominator.[51] Evaluation of ϕ_S requires numerical integration over a real-space grid and for that reason it is not considered here.

In contrast, d_{CD1}, d_{exc}, $d_{e\text{-}h}$, and other expectation values are no more expensive to compute than Δr or $\Delta\sigma$, as they require the same TD-DFT amplitudes and one-electron integrals. They are even less expensive to calculate than Λ, which requires a quadratic number of numerical quadrature steps in order to evaluate the integrals O_{ia}.

3. COMPUTATIONAL METHODS

The metrics introduced in Section 2.2 have been implemented in a locally modified version of the Q-Chem program.[89] Invariant metrics (Section 2.3) were evaluated using Q-Chem's implementation of the `libwfa` library.[90] Orbital isosurface plots were rendered using IQmol[91] and VMD,[92] with isoprobability contour values computed using the OpenCubMan program.[93]

3.1. Numerics. Integrals O_{ia} (eq 6) were evaluated by numerical quadrature using a single-center Euler-MacLaurin-Lebedev (EML) grid[94] with $N_r = 300$ radial points and $N_\Omega = 302$ angular points per atom, except in the calculations reported in Section 4.1 where extremely diffuse basis functions were employed. There, a denser grid ($N_r = 400$, $N_\Omega = 434$) was used to compute the O_{ia}. Either grid is considerably denser than those that are typically used to integrate the exchange-correlation functional.[95] For the latter purpose, the SG-1 quadrature grid[96] was used except in Sections 4.1 and 4.4 where the SG-3 grid[95] was used instead. For most calculations, the integral drop tolerance was set at $\tau_{ints} = 10^{-12}\ E_h$ and the shell-pair drop tolerance was set at $\tau_{shlpr} = 10^{-12}$ a.u., although in Sections 4.1 and 4.4 the tighter values $\tau_{ints} = 10^{-14} E_h$ and $\tau_{shlpr} = 10^{-16}$ a.u. were used instead. (See ref 97 for a discussion of appropriate numerical thresholds for diffuse basis functions.) The self-consistent field (SCF) convergence criterion was set to $\tau_{SCF} = 10^{-8}\ E_h$ in the norm of the orbital gradient and eq 2 was considered to be converged when the maximum element of the residual vector fell below $\tau_{CIS} = 10^{-6}\ E_h$.

3.2. Choice of Orbitals. In introducing the various CT metrics of Section 2.2, it was not explicitly stated how the MOs $\{\psi_i\}$ and $\{\psi_a\}$ are to be defined, and we will consider several (seemingly reasonable) choices. The most obvious choice is to use the CMOs that diagonalize the Fock matrix and have well-defined one-electron energy levels. An opposite limit is to use MOs that are localized in space rather than energy. For that purpose, we will consider Boys-localized MOs,[98] computed using the iterative algorithm from ref 99. We have occasionally noticed that the resulting Boys orbitals can be sensitive to the initial set of orbitals used to seed the algorithm, especially in the highly diffuse basis sets that are considered in Section 4.1. For consistency, in all cases involving Boys orbitals the initial guess for the SCF procedure consists of a superposition of atomic densities generated *in situ*, using the target density functional and basis set.

The NTO representation (Appendix B) is now widely used for interpretative purposes.[2] NTOs are the best one-particle orbitals in the well-defined sense of minimizing configuration mixing by reducing the transition density to a minimal number of nonvanishing amplitudes, λ_i ($i = 1, \ldots, n_{occ}$).[2] Although sometimes defined using singular value decomposition of $\mathbf{x}$,[51,100−103] a more general definition when $\mathbf{y} \neq \mathbf{0}$ is that the NTOs diagonalize both $\Delta\mathbf{P}_{elec}$ and $\Delta\mathbf{P}_{hole}$. For future reference, note that the particle and hole densities in the NTO basis are

$$\Delta\rho_{elec}(\mathbf{r}) = \sum_{i=1}^{n_{occ}} \lambda_i^2 |\psi_i^{elec}(\mathbf{r})|^2 \tag{24a}$$

$$\Delta\rho_{hole}(\mathbf{r}) = -\sum_{i=1}^{n_{occ}} \lambda_i^2 |\psi_i^{hole}(\mathbf{r})|^2 \tag{24b}$$

where the orbitals ψ_i^{elec} are eigenfunctions of $\Delta\mathbf{P}_{elec}$ (particle NTOs), with eigenvalues λ_i^2, and the functions ψ_i^{hole} are eigenfunctions of $\Delta\mathbf{P}_{hole}$ (hole NTOs), with eigenvalues $-\lambda_i^2$.[81] In both eqs 24a and 24b, the number of nonzero eigenvalues is no larger than n_{occ}, the number of occupied MOs.[2] (See Appendix B.)

It has been suggested that Δr correlates with D_{CT} in the NTO basis,[17] although reasonable correlation is also observed using CMOs.[57] For visualization purposes, one typically examines only the occupied/virtual NTO pair having the largest amplitude, namely, ψ_1^{hole} and ψ_1^{elec} (assuming that the eigenvalues are ordered according to $\lambda_1^2 \geq \lambda_2^2 \geq \lambda_2^3 \cdots$). However, in present work we retain all of the NTOs so that the transformation from CMOs to NTOs is unitary, without loss of information. Changes in Λ, Γ, or other metrics from one representation to another are thus inherent to the definitions of these quantities and are not any kind of numerical artifact or approximation.

4. RESULTS AND DISCUSSION

In the following discussion, we will present several examples that reveal how the representation dependence of metrics such as Λ, Γ, Δr, and $\Delta\sigma$ negatively impacts their use as measures of CT character and/or exciton size. We first consider a catastrophic example (Section 4.1) where diffuse basis functions lead to nonsensical results for a compact valence transition. In Section 4.2, we consider how different choices for the MOs impact the use of Λ and Γ as metrics for assessing CT character, in the sense that was originally envisaged by Tozer and coworkers[9] and by Guido et al.[10,11] In Section 4.3, we consider several examples that cannot be reduced to a single NTO pair (because $\lambda_2^2 \gg 0$), which is not uncommon in conjugated polymers. These examples demonstrate that Δr and Γ may lose their simple physical interpretation in such cases. Finally, Section 4.4 considers a set of CT complexes in which the S_1 state is well described by electron transfer from the highest occupied MO (HOMO) of a donor molecule into the lowest unoccupied MO (LUMO) of its partner. These examples highlight that Γ_{NTO} does have intuitive and interpretative value in this idealized case.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

4.1. Catastrophic Example. To illustrate how badly things can go awry when orbital invariance is sacrificed, we begin with an example of the low-lying excited states of a dipeptide that has been used in previous tests of Λ and Γ.[9,11] Guido et al.[11] noted that the CMO and NTO representations afford different values of Δr and Γ for this system, but did not pursue this observation in detail. We computed excited states at the TD-CAM-B3LYP/6-31(n+,n+)G* level, meaning that n sets of diffuse functions are added to all atoms, starting from 6-31G*.[104,105] Beyond 6-31++G* (n = 1), diffuse functions are added in geometric progression, scaling the exponents by successive factors of $(3.32)^{-1}$.[104] Due to the ultradiffuse nature of these basis sets, tighter thresholds and denser quadrature grids are used for these calculations, as described in Section 3.1.

Principal NTOs for the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_3$ excitations of this dipeptide are depicted in Figure 1. The S_1 and S_2 states both have $n\pi^*$ character, involving different carbonyl groups, but are otherwise qualitatively similar to one another so only one of them is considered. The $S_0 \rightarrow S_3$ transition involves a Rydberg state. Each of these transitions is well-described by a single eigenvalue of $\Delta\mathbf{P}_{\text{elec}}$, with $\lambda_1^2 > 0.94$. For $n > 2$ sets of diffuse functions, the excitation energies change by ≤ 0.001 eV (Table 1) and additional diffuse functions also have vanishingly little impact on the nature of the NTOs, as shown in the side-by-side comparison of 6-31++G* and 6-31(9+,9+)G* results that appears in Figure S1. However, several of the CT metrics continue to change their values as additional diffuse basis functions are added.

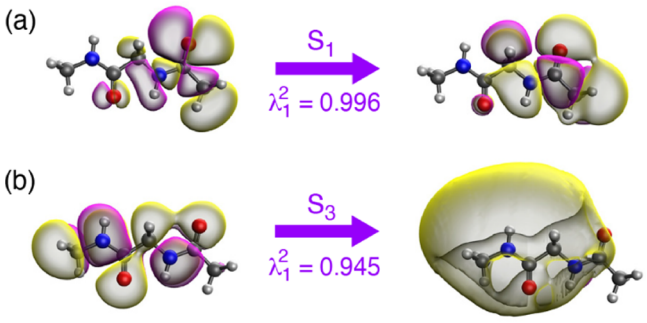


Figure 1. Principal NTO pairs for (a) $S_0 \rightarrow S_1$ excitation and (b) $S_0 \rightarrow S_3$ excitation of a dipeptide, computed at the TD-CAM-B3LYP/6-31++G* level. Isosurfaces are plotted using a contour value of 0.02 $a_0^{-3/2}$ that encapsulates at least 94% of $|\psi|^2$ in each case. For both transitions, the largest eigenvalue (λ_1^2) of $\Delta\mathbf{P}_{\text{elec}}$ is indicated.

4.1.1. CMO and Boys Representations. Numerical values of Δr, $\Delta\sigma$ and Γ are listed in Table 1, up to n = 9 sets of diffuse functions, in both the CMO and the Boys-localized representations. (See Table S2 for the Λ metric.) In the CMO representation, the value of Γ continues to grow as additional diffuse shells are added, driven mostly by growth in the value of $\Delta\sigma$ and reaching Γ_{CMO} = 13.5 Å for the S_1 state and Γ_{CMO} = 24.9 Å for S_3, for n = 9 sets of diffuse functions.

We first discuss the results in the CMO representation. Depending on basis set, the dipeptide in question has at most two (very slightly) bound virtual orbitals, with orbital energies $\varepsilon_{\text{LUMO}} = -0.15$ eV and $\varepsilon_{\text{LUMO+1}} = -4.4 \times 10^{-5}$ eV at the CAM-B3LYP/6-31(9+,9+)G* level, whereas the 6-31G* and 6-31++G* basis sets do not afford any bound virtual orbitals at all. Even the Rydberg state S_3 (Figure 1b) is relatively compact as compared to these unbound CMOs, which requires significant configuration mixing among spatially diffuse basis functions in order to generate the target state. The result is large values of σ_{ia} when $\varepsilon_a > 0$, with mixing that grows more pronounced as additional diffuse shells are added. As evidence, in the 6-31(9+,9+)G* basis set the largest CMO transition amplitude for the $S_0 \rightarrow S_1$ excitation represents only ∼25% of the transition density ($|x_{ia}|^2$ = 0.255), and for the $S_0 \rightarrow S_3$ transition $|x_{\text{HOMO,LUMO}}|^2$ = 0.364. In these cases, $\|\mathbf{y}\| \lesssim 10^{-3}$ so TDA results are quite similar.

In the Boys-localized representation, these metrics behave somewhat similarly for $S_0 \rightarrow S_1$ although their numerical values are certainly different. For $S_0 \rightarrow S_3$, however, Γ_{Boys} increases dramatically as the basis set becomes more diffuse, reaching Γ_{Boys} = 516 Å for n = 9. As compared to Γ_{CMO}, consistently larger values of Γ_{Boys} are driven by the fact that $(\Delta r)_{\text{Boys}}$ also becomes very large, e.g., $(\Delta r)_{\text{Boys}}$ = 228 Å for n = 9 as compared to $(\Delta r)_{\text{CMO}}$ = 5 Å. These outrageously large values in the Boys-localized representation prompt us to remind the reader that this is the same $S_0 \rightarrow S_3$ state that is visualized in Figure 1b for n = 1, and (with little difference) in Figure S1d for n = 9. Similar artifacts are observed in the 6-31G(n+)G* basis set, e.g., Γ_{Boys} = 413 Å for n = 9 (Table S3).

Boys localization is most commonly applied to the occupied MOs but the algorithm is perfectly well-defined for the virtual space and we have used it to localize both the occupied and (separately) the virtual MOs. In the latter case, the unbound and spatially diffuse nature of the canonical virtual orbitals leads to

Table 1. CT Metrics for Transitions of a Dipeptide, Computed at the TD-CAM-B3LYP/6-31(n+,n+)G* Level

n	$S_0 \rightarrow S_1(n\pi^*)$							$S_0 \rightarrow S_3$(Rydberg)						
		Δr (Å)		$\Delta\sigma$ (Å)		Γ (Å)			Δr (Å)		$\Delta\sigma$ (Å)		Γ (Å)	
	ΔE (eV)	CMO	Boys	CMO	Boys	CMO	Boys	ΔE (eV)	CMO	Boys	CMO	Boys	CMO	Boys
0	5.796	1.2	3.7	0.4	0.5	1.6	4.3	7.189	2.17	5.06	0.11	0.33	2.28	5.39
1	5.828	1.5	4.5	2.2	1.1	3.6	5.6	6.258	3.49	5.86	1.65	1.86	5.14	7.73
2	5.813	1.8	3.6	4.2	1.6	6.0	5.1	6.236	4.04	9.18	3.80	4.29	7.84	13.47
3	5.812	2.3	6.9	6.2	3.7	8.5	10.6	6.236	4.48	13.58	6.24	9.11	10.72	22.69
4	5.812	2.3	6.1	7.2	3.5	9.5	9.7	6.236	4.88	17.18	8.22	14.10	13.11	31.28
5	5.812	2.3	5.3	7.3	2.8	9.6	8.1	6.236	4.53	21.71	10.24	20.35	14.78	42.06
6	5.812	2.6	5.9	8.1	3.3	10.7	9.2	6.236	5.20	26.99	12.09	26.84	17.29	53.63
7	5.812	2.3	5.9	8.7	3.8	10.9	9.7	6.236	5.17	34.51	13.97	43.12	19.14	77.64
8	5.812	2.4	7.2	9.0	6.2	11.3	13.4	6.236	5.01	36.02	15.49	53.08	20.50	89.10
9	5.812	2.3	6.8	11.2	5.6	13.5	12.5	6.236	5.22	228.29	19.64	287.90	24.87	516.19

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

Boys orbitals that are *extremely* diffuse in a few cases. For example, the Rydberg state S_3 primarily involves a pair of excitations (with opposite phase) into a single ultradiffuse virtual orbital, such that $\|\mathbf{R}_{ia}(\text{Boys})\| \approx 366$ Å and $\sigma_{ia}(\text{Boys}) \approx 462$ Å for both of the amplitudes in question. In contrast, the relatively compact $S_0 \rightarrow S_1$ excitation (Figure 1a) involves amplitudes for which $\sigma_{ia}(\text{Boys}) \lesssim 5$ Å, which is actually smaller than the corresponding values for $S_0 \rightarrow S_1$ in the CMO basis. The latter range from $\sigma_{ia}(\text{CMO}) = 7.4$ Å to $\sigma_{ia}(\text{CMO}) = 13.5$ Å.

The most diffuse exponent in the 6-31(9+,9+)G* basis set is $\zeta \approx 2.44 \times 10^{-6} a_0^{-2}$, corresponding to a half-width at half-maximum of 533 Å.[106] This means that much of the probability density for the aforementioned ultradiffuse Boys orbitals lies near the edges of the basis set, which may impair our ability to numerically integrate the functions in question even though we have used dense grids for these calculations, as described in Section 3.1. However, there is good consistency in $(\Delta r)_{\text{Boys}}$ and $(\Delta\sigma)_{\text{Boys}}$ up to $n = 8$ (Table 1), which might be considered more reliable. Even for $n = 8$, the $S_0 \rightarrow S_3$ excitation has significant contributions from amplitudes x_{ia} with $\|\mathbf{R}_{ia}(\text{Boys})\| = 49$ Å and $\sigma_{ia}(\text{Boys}) = 76$ Å. In a proper expectation value, such long-distance charge separation would cancel in a coherent superposition of $i \rightarrow a$ excitations with different phases, but this does not occur in the $\kappa_{ia}^2\|\mathbf{R}_{ia}\|$ and $\kappa_{ia}^2\sigma_{ia}$ terms that appear in the definitions of Δr and $\Delta\sigma$.

It is worth emphasizing that we are not necessarily advocating for the use of Boys orbitals in TD-DFT, although localized-orbital implementations of TD-DFT have certainly been reported[65−72] and can be used to reduce cost for large systems,[65−69] for qualitative analysis,[72] and to eliminate spurious CT states.[67,68] Similarly, we are not advocating for the use of numerous diffuse shells when one or two is enough to obtain converged properties. That said, Gaussian basis sets with up to $n = 8$ diffuse shells,[107−110] designed to approximate continuum states, have been used to simulate molecular high-harmonic generation induced by strong laser fields.[108−115] (Such calculations are typically performed using real-time Kohn−Sham theory.[1]) Thus, applications of both localized orbitals and ultradiffuse basis sets do exist in the TD-DFT literature.

More importantly, we believe that the metrics used to characterize excited states should be robust and stable in *any* basis set, so that the user need not worry that the wrong basis set might afford nonsensical results. Orbital-invariant metrics might be used to verify that amplitudes discarded in a low-cost, localized-orbital implementation of TD-DFT make negligible contributions to the transition density, and thus to expectation values. This is only possible if the metrics are compatible with orbital localization.

4.1.2. NTO Representation. In contrast to instabilities observed in the CMO and Boys representations, values of Δr and $\Delta\sigma$ (and thus Γ as well) are quite stable in the NTO representation, as shown in Table 2 for the problematic Rydberg excitation. For $n > 2$ sets of diffuse functions, there is absolutely no change in either $(\Delta r)_{\text{NTO}} = 1.71$ Å or $(\Delta\sigma)_{\text{NTO}} = 0.98$ Å. The NTO basis also affords stable results for the $S_0 \rightarrow S_1$ transition and for the Λ metric, as documented in Table S2.

This stability originates in the fact that both transitions are dominated by a single pair of principal NTOs, with $\lambda_1^2 \geq 0.99$ for $S_0 \rightarrow S_1$ and $\lambda_1^2 \geq 0.94$ for $S_0 \rightarrow S_3$, even when the basis set contains numerous diffuse shells. If $\lambda_1^2 \approx 1$ then $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$ in eq 24 can be approximated using just a single term involving $\psi_1^{\text{elec}}(\mathbf{r})$ or $\psi_1^{\text{hole}}(\mathbf{r})$. Expressions for $(\Delta r)_{\text{NTO}}$ and $(\Delta\sigma)_{\text{NTO}}$ reduce to a single term under these conditions, and use of the NTO representation approximates an expectation value that is stable with respect to additional diffuse shells.

Table 2. Exciton Parameters (in Å) for the $S_0 \rightarrow S_3$ Rydberg Transition of a Dipeptide, Computed at the TD-CAM-B3LYP/6-31(n+,n+)G* Level[a]

n	NTO		Invariant			
	Δr	$\Delta\sigma$	$d_{\text{e-h}}$	σ_{hole}	σ_{elec}	d_{CD1}
0	1.83	0.15	1.80	1.98	1.84	1.94
1	1.68	0.84	1.68	2.28	3.14	2.53
2	1.71	0.98	1.71	2.31	3.32	2.72
3	1.71	0.98	1.71	2.31	3.32	2.72
4	1.71	0.98	1.71	2.31	3.32	2.72
5	1.71	0.98	1.71	2.31	3.32	2.72
6	1.71	0.98	1.71	2.31	3.32	2.72
7	1.71	0.98	1.71	2.31	3.32	2.72
8	1.71	0.98	1.71	2.31	3.32	2.72
9	1.71	0.99	1.71	2.31	3.33	2.73

[a]Same $S_0 \rightarrow S_3$ transition as in Table 1.

Such stability is obviously desirable so it is worth noting that exciton properties such as $d_{\text{e-h}}$, d_{exc}, d_{CD1}, σ_{hole}, etc., which were introduced in Section 2.3, are inherently stable because they are formulated as expectation values. Results for a few of these invariant metrics are shown in Table 2, confirming this stability. Note also that these exciton properties are no more expensive to compute than Δr or $\Delta\sigma$ yet there is no ambiguity regarding which representation is the best choice because proper expectation values are invariant to unitary transformations of the MOs.

Finally, it is worth noting that $(\Delta r)_{\text{NTO}} \approx d_{\text{e-h}}$ (see Table 2). For a transition that is dominated by a principal NTO pair, it follows that the quantity $(\Delta r)_{\text{NTO}}$ does measure electron−hole separation. The same cannot be said for $(\Delta r)_{\text{CMO}}$ or $(\Delta r)_{\text{Boys}}$, even if $\lambda_1^2 \approx 1$. Furthermore, we find that $(\Delta\sigma)_{\text{NTO}}$ is a good approximation to $|\sigma_{\text{elec}} - \sigma_{\text{hole}}|$, which makes sense in terms of the definition of $\Delta\sigma$ (eq 11) and implies that $\Gamma_{\text{NTO}} \approx d_{\text{CD1}}$ when $\lambda_1^2 \approx 1$. This is clear in the data provided in Table 2, where $\Gamma_{\text{NTO}} = 2.69$ Å as compared to $d_{\text{CD1}} = 2.72$ Å. This provides a rationale for the stability of Γ_{NTO} that was observed but not explained in ref 11.

At the same time, this observation suggests that correspondences between proper expectation values and the numerical values of $(\Delta r)_{\text{NTO}}$, $(\Delta\sigma)_{\text{NTO}}$, and Γ_{NTO} may degrade in cases where more than one principal NTO pair is significant (i.e.,when $\lambda_2^2 \gg 0$). Such cases will be examined in Section 4.3. Before that, we consider the use of Γ and Λ for their original purpose: detecting errors in TD-DFT excitation energies by correlating those errors with a measure of CT character.

4.2. CT Diagnostics. The Λ metric (eq 4) and the Γ metric (eq 10) have often been used to detect errors in TD-DFT calculations.[9−11,22−25,82] With few exceptions,[63] it has been tacitly assumed that Λ should be evaluated in the CMO basis. The Γ metric was originally introduced with the same assumption,[10] although in that case it was quickly realized that Δr (and thus Γ) is sensitive to diffuse functions for Rydberg transitions.[11] The use of NTOs was suggested as a means to mitigate that dependence,[11] and results in Section 4.1 explain why.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

In the present section, we examine the representation-dependence of both Λ and Γ in the context of their use as CT diagnostics. We use the data set assembled by Tozer and coworkers for testing Λ;[9,116] the same data set has also been used to evaluate the diagnostic properties of Γ.[10,11] (Geometries are provided in the Supporting Information; see Tables S5 and S6 for the excitation energies.) This data set consists of 59 singlet excited states including numerous $^1n\pi^*$ and $^1\pi\pi^*$ transitions of two dipeptides and a tripeptide, the $^1B_{2u}$ and $^1B_{3u}$ states of acenes up to hexacene, the ^1B_u state of polyacetylenes $H(C_2H_2)_nH$ up to n = 5, several excited states of *N*-phenylpyrrole and 4-(*N*,*N*-dimethylamino)benzonitrile (DMABN), and finally a variety of singlet excited states of N_2, CO, H_2CO, and HCl. Benchmark excitation energies are taken from ref 9 and used to determine errors that are plotted below as functions of Λ or Γ. Following ref 9, we use the cc-pVTZ basis set for all calculations in this section, except for the molecules N_2, CO, and H_2CO that are responsible for all of the Rydberg transitions in the data set. For these molecules, we use the doubly augmented d-aug-cc-pVTZ basis set instead.[117] For brevity, we refer to these calculations as "TD-DFT/triple-ζ". We focus on the hybrid functionals B3LYP and PBE0 since errors are much larger for semilocal functionals. (See Figure S3 for a side-by-side comparison of TD-PBE and TD-PBE0 errors.)

4.2.1. Representation Dependence of Λ and Γ. Figure 2 plots TD-DFT errors as a function of Λ, computed using PBE0, where the metric is evaluated using either CMOs or else Boys-localized MOs. (See Figure 3 for TD-B3LYP errors versus Λ_{CMO}, which can be compared directly to analogous data plotted in ref 9.) Data in Figure 2a demonstrate that Λ_{CMO} correlates reasonably well with errors in excitation energies. Localized valence excitations have errors that are generally smaller than 0.5 eV in magnitude but may be positive or negative, and are furthermore characterized by $\Lambda_{CMO} \geq 0.3$. In principle, this critical value might be functional dependent although valence excitations computed using TD-PBE are also characterized by $\Lambda_{CMO} \geq 0.3$; see Figure S3b. Rydberg states have larger errors and are characterized by smaller values of the metric, e.g., $\Lambda_{CMO} < 0.3$ for the TD-PBE0 data in Figure 2a. Finally, errors for CT excitation energies grow larger as the metric Λ_{CMO} gets smaller. The DMABN molecule is a special case that is often classified as a CT state, but whose excitation is relatively accurate even for a global hybrid functional such as B3LYP.[118] The explanation is that the nominal CT state has a rather large value of Λ_{CMO}, due to the molecule's compact size, indicating that electron and hole are not well separated.[2,9,24]

Because Λ is not invariant to unitary transformations, its numerical value may change dramatically upon rotating the orbitals. A vivid demonstration comes from using Boys-localized orbitals; see Figure 2b. Nearly all transitions, including localized valence excitations, exhibit rather small values of Λ_{Boys}, such that the overall scale is quite compressed even while the theoretical limits ($0 \leq \Lambda_{Boys} \leq 1$) remain the same. In the Boys MO representation, Λ provides no diagnostic ability whatsoever, as the Rydberg and CT states have similar values of Λ_{Boys} as compared to localized excitations.

Figure 3 correlates Λ against errors in TD-B3LYP calculations, comparing the CMO and NTO representations within the same plot. For several of the CT states, Λ_{CMO} is numerically quite different from Λ_{NTO} and this is highlighted for a few of the CT transitions. A threshold value Λ_{CMO} = 0.3, suggested by Tozer and co-workers[9] for calculations involving hybrid functionals, works reasonably well in the CMO representation but rotation into the NTO representation moves a few of these transitions across the line. Furthermore,

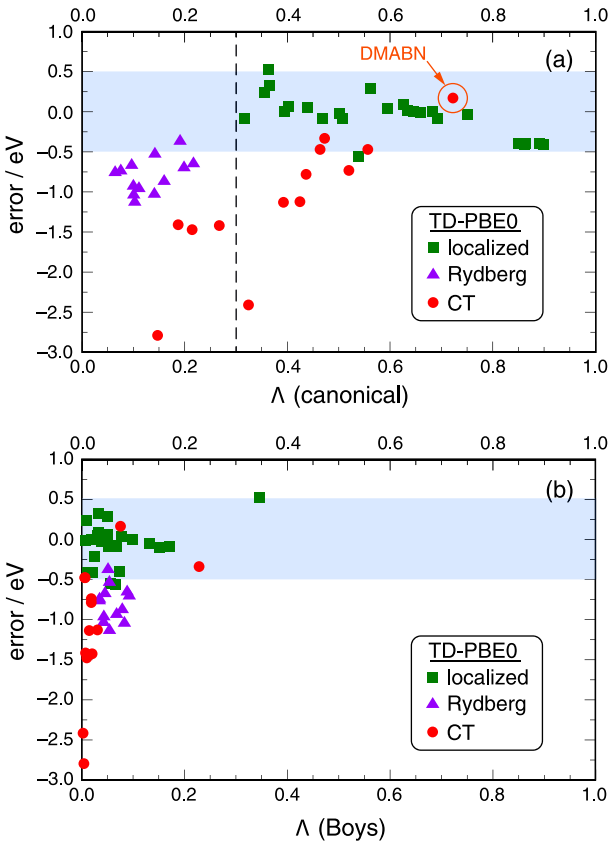


Figure 2. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-PBE0/triple-ζ level and plotted as a function of the Λ metric, which is evaluated using either (a) the CMO representation or (b) the Boys-localized MO representation. Data are partitioned into localized, Rydberg, and CT excited states and the blue shaded regions delineate where the absolute error is smaller than 0.5 eV. In (a), the data point for the CT transition in the DMABN molecule is indicated explicitly and a suggested threshold value (Λ_{CMO} = 0.3) is also indicated.

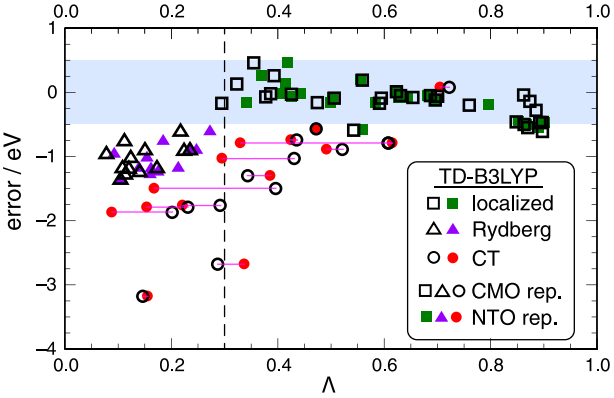


Figure 3. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-B3LYP/triple-ζ level and plotted as a function of either Λ_{CMO} (open symbols) or else Λ_{NTO} (filled symbols). The blue shaded region delineates where the absolute error is smaller than 0.5 eV. A threshold value Λ_{CMO} = 0.3, suggested in ref 9 for hybrid functionals, is indicated. Thin horizontal lines (in magenta) connect Λ_{CMO} and Λ_{NTO} for the same transition in several cases.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

not all of the shifts between Λ_{CMO} and Λ_{NTO} are in the same direction.

Errors are correlated against Γ_{CMO} and Γ_{Boys} in Figure 4, using the same data set. As with Λ_{CMO}, the metric Γ_{CMO} does a

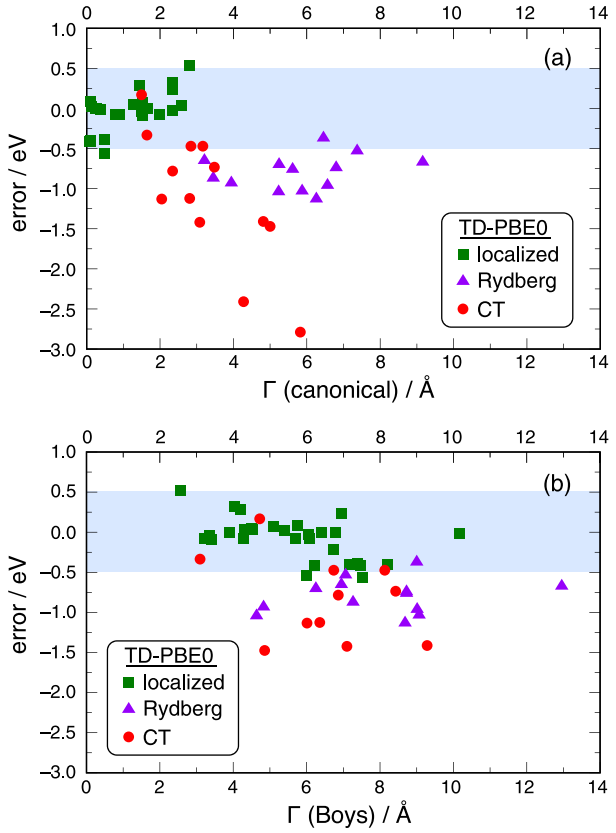


Figure 4. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-PBE0/triple-ζ level and plotted as a function of (a) Γ_{CMO} or (b) Γ_{Boys}. Blue shaded regions delineate where the absolute error is smaller than 0.5 eV.

reasonably good job of separating the localized excitations from the CT and Rydberg transitions, with errors in the CT transitions that increase in rough proportion to the value of Γ_{CMO}. Here, the Boys-localized basis does not compress the data in the same way that it does with Λ. Part of this difference may stem from the fact that unlike Λ, which is bounded, Γ is not bounded except by the basis-set size. In addition, Boys-localized MOs need not (and typically do not) transform as irreducible representations of the molecular point group, which means that $(\Delta r)_{Boys}$ need not (and typically does not) vanish in the presence of inversion symmetry. This increases the value of Γ_{Boys} relative to Γ_{CMO} since $(\Delta r)_{CMO} = 0$ in centrosymmetric molecules. In any case, the transformation to Boys orbitals once again destroys the utility of the metric, as Γ_{Boys} cannot separate the Rydberg and CT transitions from the localized excitations.

Guido et al.[11] suggest using Γ_{NTO} rather than Γ_{CMO} as the metric, due to instabilities in $(\Delta r)_{CMO}$ for Rydberg excitations. Figure 5 plots errors in TD-B3LYP excitation energies as a function of both quantities. (They are plotted separately in Figure S5. See Tables S7 and S8 for the data set.) Both representations manage to separate localized excitations from the others, yet the fact that the numerical value of Γ is sensitive to the choice of MOs is problematic if one wants to interpret its value as a quantitative measure of charge displacement, or a

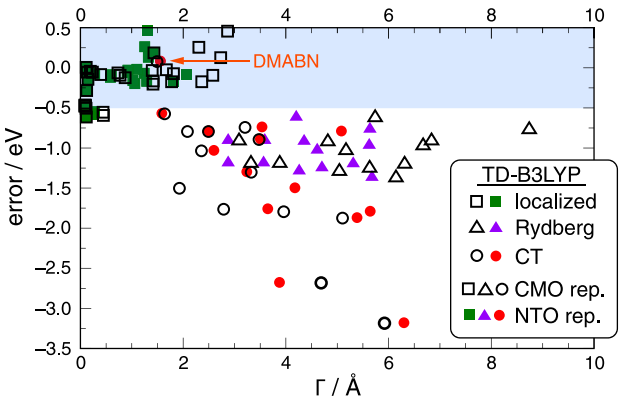


Figure 5. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-B3LYP/triple-ζ level and plotted as a function of either Γ_{CMO} (open symbols) or Γ_{NTO} (filled symbols). The blue shaded region delineates where the absolute error is smaller than 0.5 eV and the data point for DMABN's CT transition is indicated explicitly.

"trust radius" for TD-DFT.[11] The absolute difference $|\Gamma_{CMO} - \Gamma_{NTO}|$ averages 0.60 ± 0.76 Å but there are significant outliers, and $\Gamma_{CMO} - \Gamma_{NTO}$ does not have a consistent sign. For the localized excitations, the largest difference is $|\Gamma_{CMO} - \Gamma_{NTO}|$=1.4 Å (for the tripeptide) while for the Rydberg and CT states the largest differences are 3.25 Å (for a Rydberg state of H_2CO) and 2.6 Å (for a $\pi\pi^*$ state of the tripeptide), respectively.

4.2.2. Invariant Metrics. In contrast, invariant metrics do not depend on the choice of MOs and these are the only metrics that can be interpreted as genuine physical properties of the excitonic wave function. Perhaps the simplest such property is $d_{e\text{-}h}$, the mean electron−hole separation, but the present data set contains centrosymmetric molecules for which $d_{e\text{-}h} \equiv 0$. Even for molecules lacking inversion symmetry, typical values for localized valence excitations are $d_{e\text{-}h} \lesssim 1.5$ Å, with most values <1 Å, and this is not much different from values obtained for Rydberg excitations ($d_{e\text{-}h} < 2$ Å). This is readily apparent in Figure 6a, which plots errors in Tozer's data set versus $d_{e\text{-}h}$. Similarly, d_{exc} cannot distinguish between valence excitations and Rydberg or CT transitions as shown in Figure 6b.

Analogous plots of errors versus $d_{e\text{-}h}$ and d_{exc} for other functionals can be found elsewhere.[119] These results indicate that d_{exc} cannot distinguish localized excitations from either Rydberg or CT states, while $d_{e\text{-}h}$ cannot discriminate between localized and Rydberg excitations. For these reasons, we discount both $d_{e\text{-}h}$ and d_{exc} as diagnostics for CT character. Nevertheless, they remain physically interpretable measures of electron−hole separation and exciton size, respectively.

More useful as diagnostics are the metrics d_{CD1}, d_{CD2}, and d_{CD3} that were defined in eqs 19−21, and especially d_{CD1}. TD-DFT errors are plotted against these quantities in Figure 7; see Tables S9 and S10 for the numerical data. The d_{CD1} metric does the best job of separating the localized excitations from the two other types of transitions. The distribution of d_{CD1} values (Figure 7a) closely resembles that of Γ_{NTO} in Figure 5, and from these data one might infer a trust radius of about 2 Å. Indeed, that is the value put forward previously based on analysis of Γ_{NTO}.[11]

As compared to the error distribution versus $d_{e\text{-}h}$ (Figure 7a), the additional term $|\sigma_{hole} - \sigma_{elec}|$ that is included in the definition of d_{CD1} functions to separate the Rydberg states from the localized excitations, because valence excitations are characterized by similar sizes for the electron and the hole

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

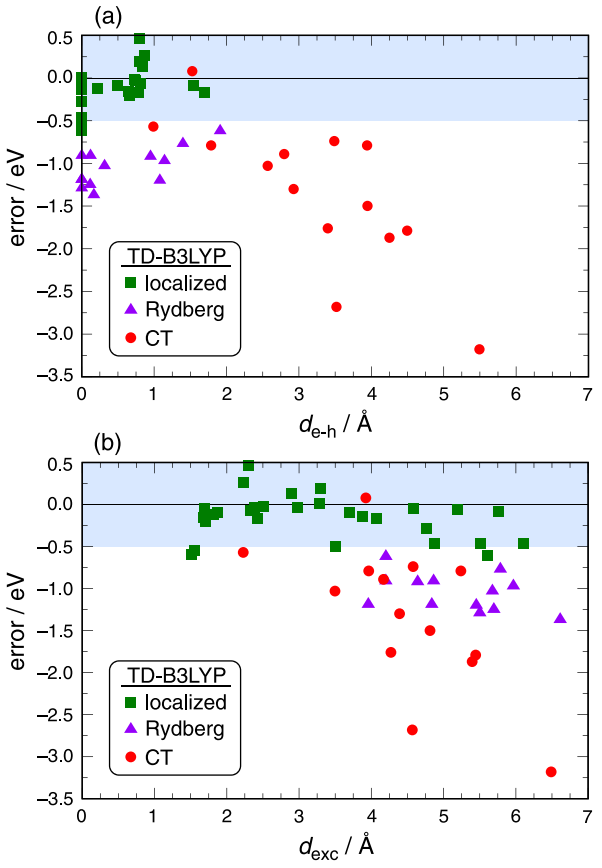


Figure 6. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-B3LYP/triple-ζ level and plotted as a function of the invariant measures (a) $d_{e\text{-}h}$ and (b) d_{exc}. Blue shaded regions delineate where the absolute error is smaller than 0.5 eV.

($\sigma_{elec} \approx \sigma_{hole}$) whereas Rydberg states involve excitation from a compact occupied MO into a relatively diffuse virtual MO ($\sigma_{elec} \gg \sigma_{hole}$). Adding d_{exc} to $d_{e\text{-}h}$, which defines the quantity d_{CD3} that was suggested in previous work (eq 21),[1,2] does not separate the localized excitations to nearly the same extent (Figure 7c) because d_{exc} does not separate the localized excitations. Lastly, d_{CD2} (eq 20) exhibits both positive and negative values and the CT states tend to have $d_{CD2} > 0$, driven by relatively large values of $d_{e\text{-}h}$. However, both localized and Rydberg excitations exhibit negative values of d_{CD2} so this metric does not discriminate between these two types of states.

4.3. Failure of the Single-NTO Approximation. In Section 4.1 we saw that $(\Delta r)_{NTO} \approx d_{e\text{-}h}$ and $\Gamma_{NTO} \approx d_{CD1}$ when $\lambda_1^2 \approx 1$. It follows that $(\Delta\sigma)_{NTO} \approx |\sigma_{elec} - \sigma_{hole}|$ in that case. In other words, these metrics approximate rigorous expectation values if the transition in question is dominated by a single eigenvalue of $\Delta\mathbf{P}_{elec}$, and then only when the metrics are evaluated in the NTO representation. The criterion $\lambda_1^2 \approx 1$ is satisfied by many (though not all) of the transitions in the data set used in Section 4.2, which has also been used elsewhere to calibrate TD-DFT errors versus metrics such as Γ or Λ.[9−11] The distribution of errors versus Γ_{NTO} (Figure 5) strongly resembles the distribution versus d_{CD1} (Figure 7a) for the transitions in that data set.

Although NTOs provide the most compact basis in which to visualize an excitation, it cannot be assumed that an arbitrary transition is dominated by a single pair of NTOs. From Tozer's data set that was used in Section 4.2, the $^1B_{3u}$ state of the linear acenes is a good example where there are two significant eigenvalues of $\Delta\mathbf{P}_{elec}$ with comparable magnitudes. Since the eigenvalues λ_i^2 are related directly to natural occupation numbers,[2,102] this indicates unresolvable multiconfigurational character in the excited state.[2] This phenomenon, which is also known as excited-state entanglement,[120] can occur even when the ground state is comfortably single-reference and implies that individual MOs alone are no longer sufficient to characterize the excited state in question. This section will explore several such examples, including linear acenes (Section 4.3.1), conjugated polymers (Section 4.3.2), and an example with multiple electronically coupled chromophore units (Section 4.3.3).

4.3.1. Linear Acenes. The $^1B_{2u}(^1L_a)$ and $^1B_{3u}(^1L_b)$ states of linear acene molecules[121] have attracted considerable interest because ^1L_b is accurately described by global hybrid functionals such as B3LYP but ^1L_a is not.[83,122] The latter exhibits ionic character that is not detected by standard metrics such as Λ_{CMO}.[63,83]

For each of the acenes in Tozer's data set (naphthalene through hexacene), we find that ^1L_a is well described by one NTO pair but ^1L_b requires two pairs. Representative results for hexacene are shown in Figure 8, including principal NTOs and the metrics Γ_{NTO} and d_{CD1}. For the ^1L_a state with $\lambda_1^2 = 0.977$, $\Gamma_{NTO} \approx d_{CD1}$ to within 0.004 Å but the difference is more substantial for the ^1L_b state, where $\lambda_1^2 = 0.555$. In the latter case, Γ_{NTO}=0.28 Å but d_{CD1} = 0.06 Å.

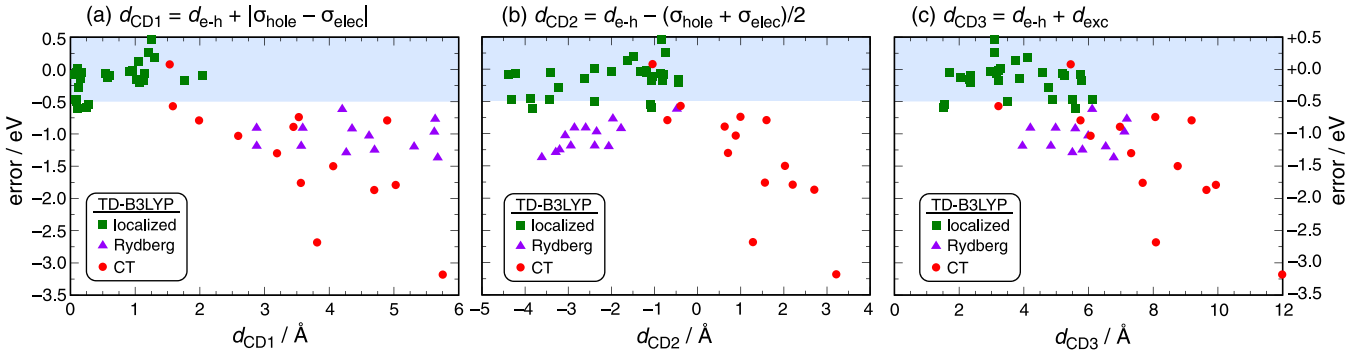


Figure 7. Errors in vertical excitation energies for Tozer's data set,[9] computed at the TD-B3LYP/triple-ζ level and plotted as a function of the invariant measures (a) d_{CD1}, (b) d_{CD2}, and (c) d_{CD3}. The vertical scale is the same in each panel but the horizontal scales differ. Blue shaded regions delineate where the absolute error is smaller than 0.5 eV.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

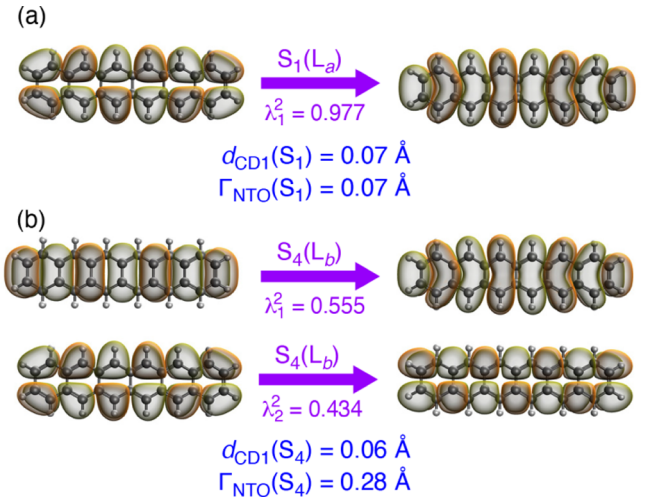


Figure 8. Principal NTOs for (a) the S_1 ($^1B_{2u}$ or ^1L_a) state and (b) the S_4 ($^1B_{3u}$ or ^1L_b) state of hexacene, computed at the TD-B3LYP/cc-pVTZ level. Values for certain metrics are also shown.

Figure 9a correlates the difference $\Gamma_{NTO} - d_{CD1}$ with the value of λ_1^2, for a data set consisting of the lowest 30 singlet transitions for each acene in the sequence from naphthalene to hexacene. A cluster of transitions with $\lambda_1^2 \approx 1$ confirms that $\Gamma_{NTO} \approx d_{CD1}$ under these circumstances. However, there are numerous states

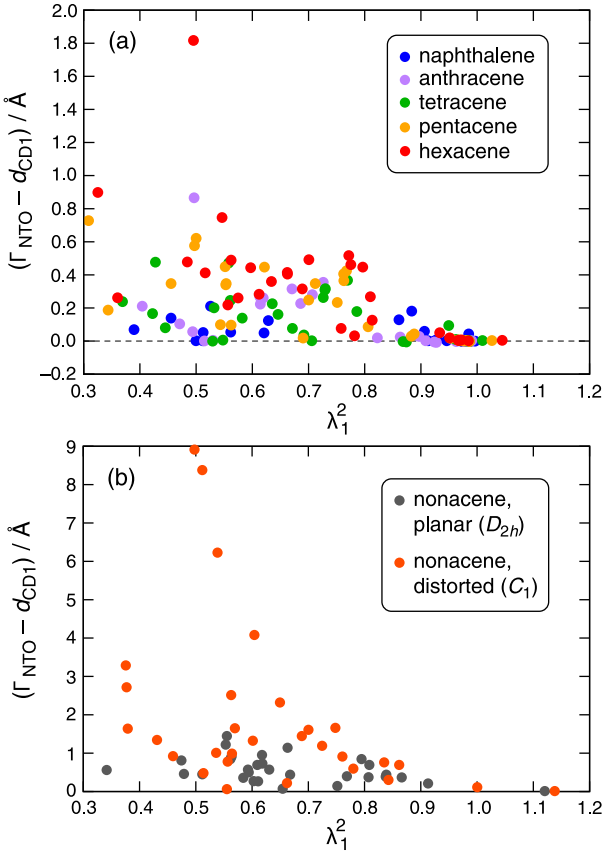


Figure 9. Difference $\Gamma_{NTO} - d_{CD1}$ versus λ_1^2 for the lowest 30 singlet excited states of linear acene molecules. (a) Planar acenes from the Tozer data set. (b) Nonacene, comparing results for a planar and a slightly distorted geometry (RMSD = 0.14 Å). All calculations were performed at the TD-B3LYP/cc-pVTZ level.

where λ_1^2 deviates significantly from unity, indicating that two or more NTO pairs are qualitatively important. (Note also that λ_1^2 may exceed unity in full linear response theory; see the discussion surrounding eq B5.) All of the geometries used to generate Figure 9a are planar and centrosymmetric (D_{2h} symmetry), so $(\Delta r)_{NTO} = 0 = d_{e\text{-}h}$ and

$$\Gamma_{NTO} - d_{CD1} \overset{D_{2h}}{=} (\Delta\sigma)_{NTO} - |\sigma_{elec} - \sigma_{hole}| \tag{25}$$

As such, values $\Gamma_{NTO} - d_{CD1} \neq 0$ reflect different estimates of how the size of the hole compares to the size of the excited electron, rather than electron−hole separation *per se*. Sometimes $(\Delta\sigma)_{NTO} \approx |\sigma_{elec} - \sigma_{hole}|$ even when λ_1^2 deviates significantly from unity but the overall trend is that the difference $\Gamma_{NTO} - d_{CD1}$ increases as λ_1^2 gets smaller and additional NTOs participate in the transition, or participate to a greater degree.

By constraining $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$ to be zero, the symmetry of these planar acenes limits the disparity between Γ_{NTO} and d_{CD1}. To understand the effect of this constraint, we lift it by examining a slightly distorted geometry for nonacene obtained by small displacements of a few atoms at one end of the molecule. The RMS displacement between this perturbed geometry (with C_1 symmetry), and a D_{2h} geometry optimized at the ωB97X-D/6-31G* level, is only 0.14 Å in the standard nuclear orientation.[96] An overlay of the two geometries (Figure S6) shows that they are essentially indistinguishable to the eye. However, the reduced symmetry does allow $d_{e\text{-}h}$ to differ from zero.

This rather benign distortion of nonacene's geometry exacerbates disparities between d_{CD1} and Γ_{NTO}, as depicted in Figure 9b for the lowest 30 singlet transitions of both the planar D_{2h} and the distorted C_1 geometries. For the D_{2h} structure, the behavior as a function of λ_1^2 resembles what was observed for smaller acenes in their D_{2h} geometries, with $|\Gamma_{NTO} - d_{CD1}| < 2.0$ Å in all cases. However, $(\Delta r)_{NTO}$ is quite large for a few excited states in the C_1 geometry, e.g., $\Gamma_{NTO}(S_{19}) = 9.96$ Å, $\Gamma_{NTO}(S_{21}) = 8.85$ Å, and $\Gamma_{NTO}(S_{29}) = 10.75$ Å (see Table S11). Because the actual electron−hole separation is much smaller ($d_{e\text{-}h} < 3$ Å for all 30 transitions), some of the differences $\Gamma_{NTO} - d_{CD1}$ are quite large in the distorted geometry.

The largest such difference, $\Gamma_{NTO} - d_{CD1} = 8.9$ Å, occurs for $S_0 \rightarrow S_{19}$ so we examine that transition in detail. The numerical difference between d_{CD1} and Γ_{NTO} in this case is driven almost entirely by the difference between $d_{e\text{-}h}$ and $(\Delta r)_{NTO}$, values for which are provided in Figure 10 alongside the principal NTOs. The $\psi_1^{hole} \rightarrow \psi_1^{elec}$ and $\psi_2^{hole} \rightarrow \psi_2^{elec}$ excitations are characterized by right-to-left and left-to-right CT, but their coherent superposition (to form $S_0 \rightarrow S_{19}$) does not displace much charge at all, with $d_{e\text{-}h} = 0.13$ Å. Because Γ_{NTO} is computed as an incoherent average, it consists of a sum of two large $\|\mathbf{R}_{ia}\|$ values corresponding to the two basis states $\psi_1^{hole} \rightarrow \psi_1^{elec}$ and $\psi_2^{hole} \rightarrow \psi_2^{elec}$. The result is $(\Delta r)_{NTO} = 9.45$ Å. This example dramatically illustrates the failure of incoherent averages as compared to proper expectation values involving coherent superpositions.

4.3.2. Poly(phenylenevinylene). Excitons in conjugated polymers sometimes require two or more π^* orbitals with different phases in order to describe the excited-state wave function, analogous to the ^1L_b state of the linear acenes (Figure 8b). Alternatively, exciton localization can create a situation in

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

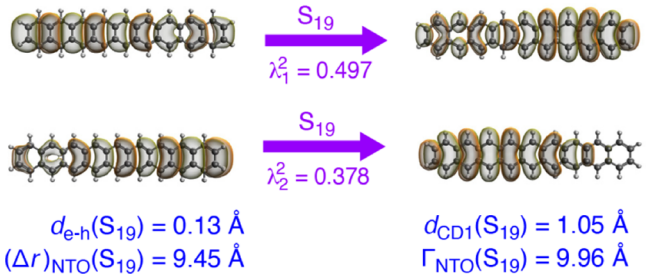


Figure 10. Principal NTOs for the $S_0 \rightarrow S_{19}$ transition of a slightly distorted (C_1) nonacene molecule, for which $(\Delta r)_{NTO}$ differs substantially from $d_{e\text{-}h}$. Calculations were performed at the TD-B3LYP/cc-pVTZ level.

which a single molecule effectively exhibits more than one chromophore, which are then electronically coupled due to their spatial proximity.[2,87] Typically, that situation also leads to more than one significant eigenvalue of $\Delta\mathbf{P}_{elec}$.[2] Poly(*p*-phenylenevinylene) or PPV chromophores exhibit these behaviors and we next examine an isomer of $(PPV)_8$ in which two of the vinylene moieties are in the *cis* configuration while the others are *trans*, which creates "kinks" in the geometry that can induce excited-state localization. An all-*trans* isomer of $(PPV)_8$ is considered in the Supporting Information (Table S13 and Figure S7) and exhibits many of the same features and trends, demonstrating that the observations presented below are not unique to the "bent" isomer.

Figure 11 depicts the principal NTO pairs for three low-energy excitations in bent $(PPV)_8$. Each transition requires more than one pair of NTOs to reach at least 80% of the norm of the transition density, i.e., so that $\lambda_1^2 + \lambda_2^2 + \cdots \geq 0.8$. For example, in the S_1 state the primary NTO pair captures only 61% of the transition density and for S_5, four NTO pairs are required to reach 80%. In the latter case, one can identify two effective chromophores within the $(PPV)_8$ molecule, on the left side and in the middle segment, whose localized excitations each require a pair of NTOs to describe and which are electronically coupled. Although this state is optically dark, the S_1 state has a large oscillator strength (see Table 3) so the need for multiple NTO pairs affects bright and dark states alike.

Metrics for a few excited states are presented in Table 3 and additional metrics can be found in Table S12. Both $d_{e\text{-}h}$ and d_{CD1} are rather small ($\lesssim$0.5 Å) for each of the ten excited states in the table, whereas Δr and Γ span a wider range. It is no longer clear that $(\Delta r)_{NTO}$ is a good approximation to $d_{e\text{-}h}$, nor is Γ_{NTO} a good approximation to d_{CD1}. This is especially evident for the $S_0 \rightarrow S_6$ transition, which has a much larger RMS exciton size (d_{exc} = 10.1 Å) as compared to the other excited states (where d_{exc} = 4.2−5.5 Å). The S_6 state represents a CT exciton whereas other excited states are Frenkel excitons, meaning superpositions of localized excitations.[2,3,41] (Frenkel excitons can nevertheless span a large distance in a conjugated molecule.) The distinction between Frenkel and CT excitons is challenging to understand based on NTOs alone but a few such excitons appear consistently in the spectra of different PPVs, where they stand out as especially large values of d_{exc} (Tables S12 and S13). Frenkel and CT excitons can be distinguished by plotting the transition density $T(\mathbf{r}_{hole}, \mathbf{r}_{elec})$, which facilitates visualization of correlations between the positions of the excited electron ($\mathbf{r}_{elec}$) and the hole ($\mathbf{r}_{hole}$).[2,41,123]

For the S_6 state in Table 3, the larger value of d_{exc} is also reflected in a larger value Γ_{NTO} = 6.45 Å, whereas $\Gamma_{NTO} < 1.7$ Å for all other states up to S_{10}. This is partly driven by a larger value of $(\Delta r)_{NTO}$ = 3.65 Å that is *not* reflected in $d_{e\text{-}h}$ = 0.1 Å. The latter is consistent with an exciton in which both the electron and the hole are delocalized over eight of the nine PPV units so there is little net change in the center of charge. Values of $(\Delta r)_{CMO}$ do not correlate at all with $d_{e\text{-}h}$, nor does Γ_{CMO} correlate with d_{CD1}.

4.3.3. Triazine Benzobisthiadiazole Propeller. The "triazine propeller" depicted in Figure 12 has been considered as a platform for optoelectronic applications.[124] Three benzobisthiadiazole ($C_6H_2N_4S_2$) substituents function as the blades of the propeller, connected by a central triazine unit. The substituents have strong dipole-allowed $^1\pi\pi^*$ transitions but

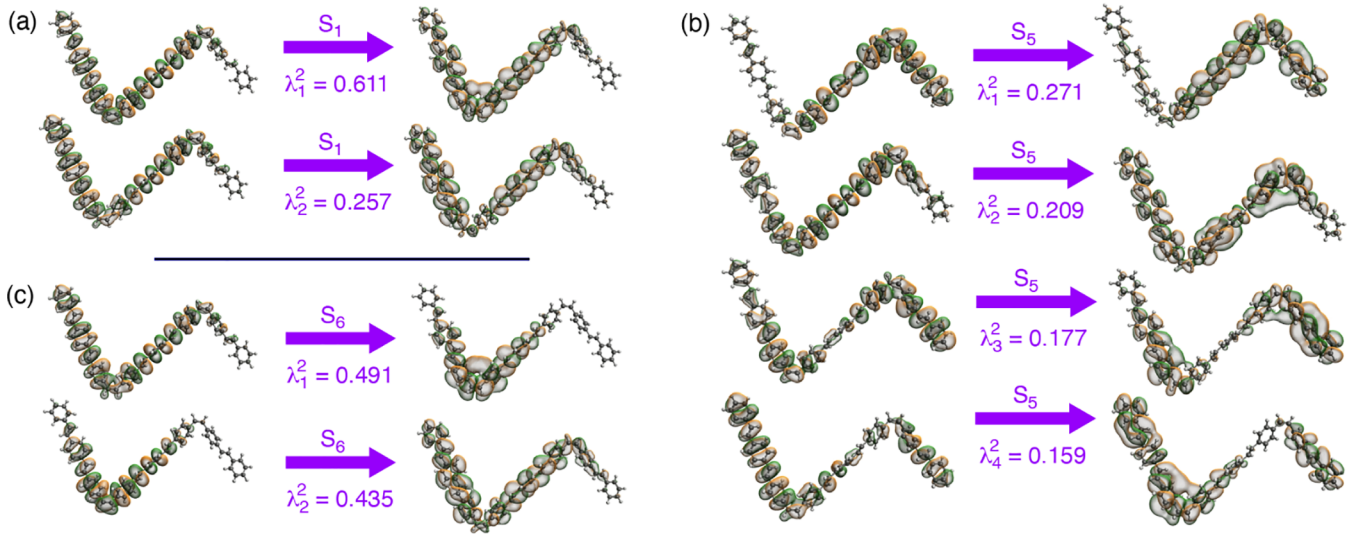


Figure 11. Principal NTO pairs for the (a) $S_0 \rightarrow S_1$, (b) $S_0 \rightarrow S_5$, and (c) $S_0 \rightarrow S_6$ transitions of a bent isomer of $(PPV)_8$. In each case, the NTOs needed to recover 80% of the transition density are shown. Calculations were performed at the TD-CAM-B3LYP/6-31+G* level and orbitals are plotted using an isocontour value of $0.02a_0^{-3/2}$, which is sufficient to capture $\gtrsim$90% of the densities $|\psi_i^{elec}(\mathbf{r})|^2$ and $|\psi_i^{hole}(\mathbf{r})|^2$.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

Table 3. Descriptors for Excited States of $(PPV)_8$ with Two *cis* Kinks[a]

				Non-Invariant Metrics (Å)				Invariant Metrics (Å)		
				CMO		NTO				
State	ΔE (eV)	Osc. Str.	λ_1^2	Δr	Γ	Δr	Γ	$d_{e\text{-}h}$	d_{CD1}	d_{exc}
S_1	3.30	2.29	0.61	0.84	1.07	0.07	0.13	0.02	0.05	5.53
S_2	3.55	1.52	0.47	2.99	4.60	0.11	0.20	0.03	0.05	5.41
S_3	3.88	1.42	0.54	6.55	7.54	0.14	0.29	0.04	0.07	5.12
S_4	4.17	0.16	0.34	3.77	5.24	0.26	0.40	0.04	0.06	5.12
S_5	4.46	0.07	0.27	3.05	4.29	0.80	1.31	0.17	0.25	4.82
S_6	4.56	0.00	0.49	3.92	5.37	3.65	6.45	0.12	0.23	10.14
S_7	4.62	0.01	0.52	4.34	5.91	0.54	1.23	0.19	0.42	4.40
S_8	4.64	0.03	0.46	6.92	9.02	1.44	1.69	0.32	0.52	5.36
S_9	4.66	0.03	0.49	4.78	6.44	0.21	1.08	0.73	0.10	4.25
S_{10}	4.67	0.01	0.41	4.81	7.08	0.29	0.83	0.09	0.20	4.30

[a]TD-CAM-B3LYP/6-31+G* level.

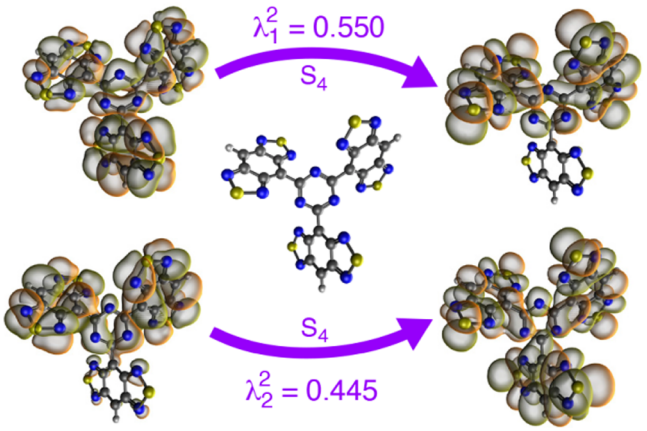


Figure 12. NTOs for the $S_0 \rightarrow S_4$ transition of a triazine benzobisthiadiazole propeller, computed at the TD-CAM-B3LYP/6-31+G* level. Isosurfaces are plotted using a contour value of 0.02 $a_0^{-3/2}$ that encapsulates at least 94% of $|\psi|^2$.

excited states on different benzobisthiadiazole units are excitonically coupled, leading to delocalization.

Each of the lowest ten singlet excited states requires at least two NTO pairs to recover 80% of the transition density, including the $S_0 \rightarrow S_4$ transition depicted in Figure 12. This is a good example of an excitation where the various NTOs exhibit rather different delocalization patterns. Whereas $\psi_1^{hole}(\mathbf{r})$ is delocalized across the entire molecule, $\psi_1^{elec}(\mathbf{r})$ is spread over only two of the three propeller blades. The opposite is true for the second principal NTO pair: $\psi_2^{hole}(\mathbf{r})$ is supported on two chromophores but $\psi_2^{elec}(\mathbf{r})$ is delocalized across all three. Both excitations have comparable weights.

Charge-separation metrics for S_1 through S_{10} are provided in Table 4 and additional metrics can be found in Table S14. For none of theses states can it be quantitatively stated that $(\Delta r)_{NTO}$ matches the value of $d_{e\text{-}h}$, or that Γ_{NTO} matches d_{CD1}. For the S_4 and S_5 states, the values are substantially different with Γ_{NTO} being more than 3 Å larger than d_{CD1}, driven by a similar disparity between $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$. Both of these states are characterized by (de)localization patterns similar to what is observed for S_4 in Figure 12, and the disparities between charge-displacement measures represent the difference between an incoherent metric (Δr or Γ) and a proper expectation value that involves a coherent superposition of orbitals.

Table 4. Descriptors for Excited States of a Triazine Benzobisthiadiazole Propeller.[a]

			Metrics (Å)			
State	ΔE (eV)	λ_1^2	$(\Delta r)_{NTO}$	Γ_{NTO}	$d_{e\text{-}h}$	d_{CD1}
S_1	2.595	0.65	0.15	0.75	0.06	0.52
S_2	2.597	0.53	0.14	0.63	0.06	0.52
S_3	2.718	0.33	0.16	0.73	0.00	0.45
S_4	3.435	0.55	3.69	4.22	0.62	1.09
S_5	3.480	0.66	4.93	5.83	1.93	2.61
S_6	3.481	0.68	2.81	3.47	2.11	2.93
S_7	3.568	0.72	2.84	3.57	1.94	2.75
S_8	3.569	0.73	3.33	4.18	1.96	2.78
S_9	3.592	0.66	0.94	1.42	0.52	0.96
S_{10}	3.909	0.50	0.25	0.64	0.01	0.46

[a]TD-CAM-B3LYP/6-31+G* level.

4.4. Electron–Hole Separation in CT Complexes.

Finally, we investigate some systems characterized by long-range electron transfer where the donor and acceptor orbitals have vanishingly little spatial overlap. It has been suggested that $(\Delta r)_{NTO}$ correlates with D_{CT},[17] meaning $d_{e\text{-}h}$, although reasonable correlation is also observed using CMOs.[57] Given what we now understand from the analysis presented herein, we can state that because long-range CT transitions are dominated by a single NTO pair, we expect excellent correspondence between $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$.

As a first example, we consider low-lying singlet excitations in a set of 29 intermolecular CT dimers assembled from common small-molecule electron donors (acenes, stilbenes, thiophenes, etc.) and acceptors (tetracyanoquinone and its fluorinated analogues).[125] In calculations at the TD-B3LYP/6-31+G* level, we find that the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions are dominated by a single NTO pair in every single case, with $\lambda_1^2(S_1) \geq 0.99$ and $\lambda_1^2(S_2) \geq 0.94$. For $S_0 \rightarrow S_3$, $\lambda_1^2 \geq 0.90$ except for a few cases involving *meso*-diphenyl tetrathia[22]annulene[2,1,2,1] (DPTTA) complexed with fluorinated tetracyanoquinone, for which $\lambda_1^2 = 0.83$. As a result, we expect that $(\Delta r)_{NTO}$ should correlate reasonably well with $d_{e\text{-}h}$ for the S_1, S_2, and S_3 states of these dimers.

That expectation is borne out by plots of $(\Delta r)_{NTO}$ versus $d_{e\text{-}h}$ in Figure 13. Deviations $(\Delta r)_{NTO} - d_{e\text{-}h}$ are strictly positive and

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

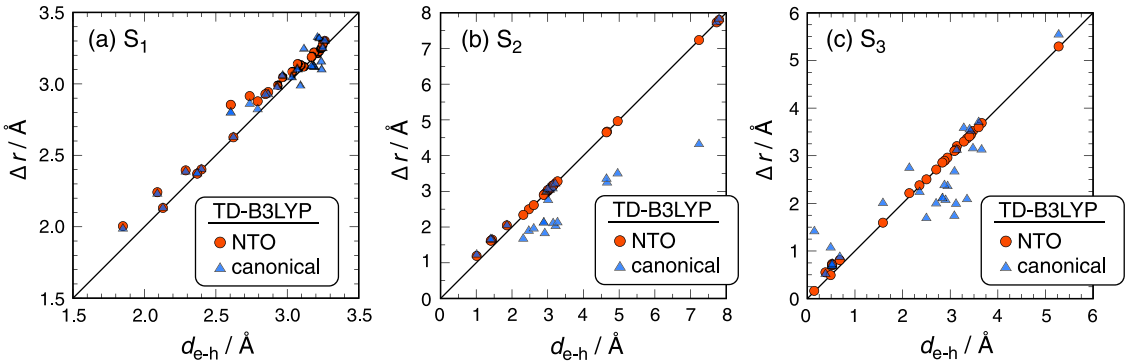


Figure 13. Correlation between Δr (in either the canonical or the NTO representation) and the invariant metric $d_{e\text{-}h}$, for excitation from S_0 to (a) S_1, (b) S_2, or (c) S_3, for a set of 29 intermolecular CT complexes[125] described at the TD-B3LYP/6-31+G* level. The length scale is different in each panel; see Figure S8 for a version in which all three panels are plotted on the same scale.

the average deviation is no larger than 0.06 Å for any of the these excited states. Despite this strong correspondence, $(\Delta r)_{NTO}$ remains numerically distinct from $d_{e\text{-}h}$. It is clear from from the S_1 data in Figure 13a that the difference is not some systematic error, despite the fact that $\lambda_1^2 \geq 0.99$ in these cases. The formula for Δr (eq 9), which involves squaring the amplitudes κ_{ia} separately from the quantity $\|\mathbf{R}_{ia}\|$ that is to be averaged, is still not an exact expectation value even in these cases. These calculations were carried out with tight thresholds and dense grids as described in Section 3.1, to answer any question of whether residual differences between $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$ might be numerical artifacts. In any case, the correlation between $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$ is much better than the correlation between $(\Delta r)_{CMO}$ and $d_{e\text{-}h}$.

Our last example involves long-range intramolecular CT in a sequence of α-$N(CH_3)_2$-ω-NO_2(phenylene)$_n$ push−pull chromophores, n = 1−8. The principal NTO pair for the largest of these molecules is depicted in Figure 14a and accounts for the entirety of the density change ($\lambda_1^2 = 1.000$), so that $(\Delta r)_{NTO}$ and $d_{e\text{-}h}$ are essentially identical. One can also watch d_{exc} converge to $d_{e\text{-}h}$ as the polymer's length increases (Figure 14b). This convergence is driven by vanishing correlations (as a function of increasing n) between $\psi_1^{hole}(\mathbf{r})$ and $\psi_1^{elec}(\mathbf{r})$, which are localized on the dimethylamino donor moiety and the nitro acceptor moiety, respectively. Notably, the $S_0 \rightarrow S_1$ excitation energy has essentially converged (as a function of length) by n = 8, which is consistent with convergence of d_{exc} to $d_{e\text{-}h}$.

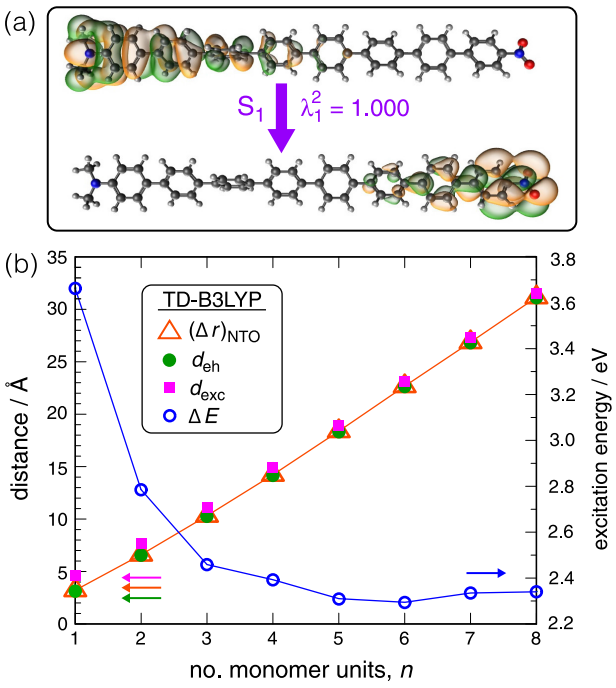


Figure 14. (a) Principal NTO pair for the $S_0 \rightarrow S_1$ transition of α-$N(CH_3)_2$-ω-NO_2(phenylene)$_8$, with isosurfaces that encapsulate 80% of $|\psi|^2$. (b) Descriptors for the $S_0 \rightarrow S_1$ transition in a sequence of α-$N(CH_3)_2$-ω-NO_2(phenylene)$_n$ chromophores. In (b), distances $(\Delta r)_{NTO}$, $d_{e\text{-}h}$, and d_{exc} should be read from the scale at left whereas excitation energies ΔE should be read from the scale on the right. Calculations were performed at the TD-B3LYP/6-31+G* level on geometries that were optimized at the ωB97X-D/6-31G* level.

5. CONCLUSIONS

The present work points out serious flaws in the definition of several commonly used CT diagnostics for TD-DFT calculations. These metrics are not independent of representation so cannot truly be said to measure electron−hole separation or exciton size. Proper measures ought to be invariant with respect to rotations of the occupied MOs and (separately) the virtual MOs, which is guaranteed for genuine expectation values but not for incoherent averages over excitation amplitudes. Metrics that fail to preserve orbital invariance can become unstable in diffuse basis sets, even when changes in the basis set do not affect excitation energies or properties. Lack of orbital invariance also introduces an arbitrary decision regarding which representation should be used to evaluate the metric in question. Different representations may afford dramatically different numerical values for putative CT diagnostics, leading to ambiguity in what is or is not classified as a CT excitation.

For small molecules, transition densities are often (but certainly not always) dominated by a principal pair of NTOs; long-range electron-transfer excitations sometimes fall into this category. In these special cases, the noninvariant metrics Δr, $\Delta\sigma$, and Γ are stable and interpretable in terms of expectation values, when evaluated in the NTO representation. Under these circumstances, $(\Delta r)_{NTO}$ does measure electron−hole separation and

$$\Gamma_{NTO} = (\Delta r)_{NTO} + (\Delta\sigma)_{NTO} \quad (26)$$

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

measures overall charge displacement, including any size disparity between the excited electron and the hole. However, it is easy to find molecules where many of the excited states cannot be described by a single pair of NTOs. These cases, involving multiconfigurational character or excited-state entanglement, cannot be properly described by incoherent superpositions of orbitals, even in the NTO basis. As a result, interpretability of the various terms in eq 26 is lost. Examples include linear acenes, conjugated polymers, and systems with multiple electronically coupled chromophores.

Fortunately, invariant metrics such as $d_{\text{e-h}}$ (mean electron−hole separation), d_{exc} (RMS exciton size), and d_{CD1} (charge displacement) are no more complicated or expensive to evaluate as compared to the aforementioned noninvariant metrics. Invariant metrics never lose their physical interpretability, are stable with respect to basis-set expansion, and do not require arbitrary choices regarding representation. For that reason, we suggest that proper expectation values should be used exclusively when characterizing excited-state wave functions and transition densities.

APPENDIX A: EXPECTATION VALUES

In practice, the `libwfa` code[90] that is used to evaluate $d_{\text{e-h}}$, σ_{elec}, and σ_{hole} relies upon treating the one-electron transition density matrix as a wave function in order to compute expectation values, as described elsewhere.[6,119] In order to motivate the introduction of NTOs in Appendix B, however, it is useful to consider the construction of density matrices for the excited electron ($\Delta\mathbf{P}_{\text{elec}}$) and for the hole ($\Delta\mathbf{P}_{\text{hole}}$). These are the matrix representations of the real-space quantities $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$ introduced in eq 1.

Consider an arbitrary operator $\hat{A}$. Its expectation value in a TD-DFT calculation is expressed as a change relative to the ground-state expectation value (A_0),

$$\langle\hat{A}\rangle = A_0 + \Delta A \tag{A1}$$

The change in $\langle\hat{A}\rangle$ upon excitation is computed according to

$$\Delta A = \text{tr}[\mathbf{A}(\Delta\mathbf{P} + \mathbf{Z})] \tag{A2}$$

where

$$\Delta\mathbf{P} = \Delta\mathbf{P}_{\text{elec}} + \Delta\mathbf{P}_{\text{hole}} \tag{A3}$$

is the unrelaxed difference density matrix and $\mathbf{Z}$ is the contribution from orbital relaxation.[49,75,126] The quantity $\Delta\mathbf{P} + \mathbf{Z}$ in eq A2 is called the *relaxed* difference density matrix.[49]

Matrix elements of the *unrelaxed* difference density matrix ($\Delta\mathbf{P}$) are given by

$$(\Delta\mathbf{P}_{\text{elec}})_{ab} = \sum_i (x_{ia}x_{ib} + y_{ia}y_{ib}) \tag{A4a}$$

$$(\Delta\mathbf{P}_{\text{hole}})_{ij} = -\sum_a (x_{ia}x_{ja} + y_{ia}y_{ja}) \tag{A4b}$$

when expressed in any orthonormal basis that preserves occupied/virtual separation.[2,127] Elsewhere, the unrelaxed density matrices defined in eq A4 have been called attachment ($\Delta\mathbf{P}_{\text{elec}}$) and detachment ($\Delta\mathbf{P}_{\text{hole}}$) density matrices,[2,128] with corresponding attachment and detachment densities $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$. In the present work, the orbital relaxation or $\mathbf{Z}$-vector contribution is omitted, in which case there is no distinction between attachment/detachment and particle/hole quantities. More generally, the particle density matrix would be constructed from those eigenvectors of $\Delta\mathbf{P} + \mathbf{Z}$ ("natural difference orbitals") that have positive eigenvalues, while the hole matrix would be constructed from eigenvectors corresponding to negative eigenvalues.[2]

Finally, the unrelaxed contribution to ΔA can be expressed in terms of particle and hole contributions based on eq A3:

$$\text{tr}[\mathbf{A}(\Delta\mathbf{P})] = (\Delta A)_{\text{elec}} + (\Delta A)_{\text{hole}} \tag{A5}$$

These two contributions are

$$(\Delta A)_{\text{elec}} = \text{tr}(\mathbf{x}\mathbf{A}\mathbf{x}^\dagger) + \text{tr}(\mathbf{y}\mathbf{A}\mathbf{y}^\dagger) \tag{A6a}$$

$$(\Delta A)_{\text{hole}} = -\text{tr}(\mathbf{x}^\dagger\mathbf{A}\mathbf{x}) - \text{tr}(\mathbf{y}^\dagger\mathbf{A}\mathbf{y}) \tag{A6b}$$

APPENDIX B: NATURAL TRANSITION ORBITALS

A thorough discussion of NTOs can be found in ref 2 and is summarized here. By definition, these are the orbitals that diagonalize $\Delta\mathbf{P}_{\text{elec}}$ and $\Delta\mathbf{P}_{\text{hole}}$ in eq A4.[2,51] Specifically, NTOs for the hole (ψ_i^{hole}) are defined by a unitary transformation $\mathbf{U}_{\text{o}}$ of the occupied MOs that diagonalizes the negative semidefinite matrix $\Delta\mathbf{P}_{\text{hole}}$. We express this transformation as[81]

$$\mathbf{U}_{\text{o}}^\dagger(\Delta\mathbf{P}_{\text{hole}})\mathbf{U}_{\text{o}} = \underbrace{\begin{pmatrix} -\lambda_1^2 & 0 & 0 & \cdots \\ 0 & -\lambda_2^2 & 0 & \cdots \\ & & \ddots & 0 \\ 0 & \cdots & 0 & -\lambda_{n_{\text{occ}}}^2 \end{pmatrix}}_{-\mathbf{\Lambda}^2} \tag{B1}$$

This notation for the eigenvalues reflects the fact that each λ_i is a singular value of the coefficient matrix $\mathbf{x}$, if $\mathbf{y} = \mathbf{0}$ as in the TDA (see below). The values λ_i^2 are connected to the natural occupation numbers of the excited-state density matrix.[2,102]

NTOs for the excited electron (ψ_i^{elec}) are defined by a unitary transformation $\mathbf{U}_{\text{v}}$ of the virtual MOs that diagonalizes the positive semidefinite matrix $\Delta\mathbf{P}_{\text{elec}}$. The values $\lambda_1^2 \geq \lambda_2^2 \geq \lambda_3^2 \geq \cdots$ in eq B1 are precisely the nonzero eigenvalues of $\Delta\mathbf{P}_{\text{elec}}$, so the transformation of the virtual space can be expressed as[81]

$$\mathbf{U}_{\text{v}}^\dagger(\Delta\mathbf{P}_{\text{elec}})\mathbf{U}_{\text{v}} = \begin{pmatrix} \mathbf{\Lambda}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{B2}$$

This illustrates that the rank of $\Delta\mathbf{P}_{\text{elec}}$ is equal to the number of occupied MOs (n_{occ}). The NTOs occur in corresponding electron/hole pairs whose contribution to the norm of the transition density is λ_i^2, hence the expressions for $\Delta\rho_{\text{elec}}(\mathbf{r})$ and $\Delta\rho_{\text{hole}}(\mathbf{r})$ in eq A4.

Note that

$$\text{tr}(\Delta\mathbf{P}_{\text{elec}}) = \sum_i^{n_{\text{occ}}} \lambda_i^2 = \sum_{ia} (|x_{ia}|^2 + |y_{ia}|^2) \tag{B3}$$

and that

$$\text{tr}(\Delta\mathbf{P}_{\text{elec}}) = -\text{tr}(\Delta\mathbf{P}_{\text{hole}}) \tag{B4}$$

The second equality in eq B3 differs from the normalization condition in eq 7, with the effect that $\text{tr}(\Delta\mathbf{P}_{\text{elec}})$ may differ from unity. Within the TDA,

$$\text{tr}(\Delta\mathbf{P}_{\text{elec}}) \overset{\text{TDA}}{=} 1 \tag{B5}$$

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

In our experience, $\|\mathbf{y}\| \lesssim 10^{-3}$ for small molecules so deviations from eq B5 in full linear response are typically small. (The y_{ia} amplitudes are somewhat larger for larger acenes.) Equation B4 holds in any case.

Within the TDA, $\Delta\mathbf{P}_{\text{elec}} = \mathbf{x}^{\dagger}\mathbf{x}$ and $\Delta\mathbf{P}_{\text{hole}} = -\mathbf{x}\mathbf{x}^{\dagger}$. In that case, it follows that $\mathbf{U}_o$ and $\mathbf{U}_v$ consist of the left and right singular values of $\mathbf{x}$, respectively. NTOs within the TDA can thus be obtained via singular value decomposition,[2]

$$\mathbf{U}_o\mathbf{x}\mathbf{U}_v^{\dagger} \overset{\text{TDA}}{=} \begin{pmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{B6}$$

consistent with earlier definitions that considered the TDA case only.[100−103] The more general definitions in eqs B1 and B2 still hold, however.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c01085.

Additional quantitative data assessing various metrics (PDF)

Coordinates for the systems considered (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

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

Aniket Mandal − *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0009-0000-8880-0122

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.4c01085

Notes

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

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under grant nos. CHE-1955282 and CHE-2402361. Calculations were performed at the Ohio Supercomputer Center.[129]

■ REFERENCES

(1) Herbert, J. M. Density-functional theory for electronic excited states. In *Theoretical and Computational Photochemistry: Fundamentals, Methods, Applications and Synergy with Experimental Approaches*, García-Iriepa, C.; Marazzi, M., Eds.;Elsevier: Amsterdam, 2023; Ch. 3, pp. 69−118.

(2) Herbert, J. M. Visualizing and characterizing excited states from time-dependent density functional theory. *Phys. Chem. Chem. Phys.* **2024**, *26*, 3755−3794.

(3) Plasser, F.; Lischka, H. Analysis of excitonic and charge transfer interactions from quantum chemical calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2777−2789.

(4) Plasser, F.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. I. Formalism. *J. Chem. Phys.* **2014**, *141*, 024106.

(5) Plasser, F.; Bäppler, S. A.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* **2014**, *141*, 024107.

(6) Plasser, F.; Thomitzni, B.; Bäppler, S. A.; Wenzel, J.; Rehn, D. R.; Wormit, M.; Dreuw, A. Statistical analysis of electronic excitation processes: Spatial location, compactness, charge transfer, and electron-hole correlation. *J. Comput. Chem.* **2015**, *36*, 1609−1620.

(7) Mewes, S. A.; Dreuw, A. Density-based descriptors and exciton analyses for visualizing and understanding the electronic structure of excited states. *Phys. Chem. Chem. Phys.* **2019**, *21*, 2843−2856.

(8) Kimber, P.; Plasser, F. Classification and analysis of molecular excited states, In *Comprehensive Computational Chemistry*, 1st ed., Yáñeez, M.; Boyd, R. J., Eds.; Elsevier: Oxford, 2024; Vol. 4, pp. 55−83.

(9) 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*, 044118.

(10) Guido, C. A.; Cortona, P.; Mennucci, B.; Adamo, C. On the metric of charge transfer molecular excitations: A simple chemical descriptor. *J. Chem. Theory Comput.* **2013**, *9*, 3118−3126.

(11) Guido, C. A.; Cortona, P.; Adamo, C. Effective electron displacements: A tool for time-dependent density functional theory computational spectroscopy. *J. Chem. Phys.* **2014**, *140*, 104101.

(12) Luzanov, A. V.; Zhikol, O. A. Excited state structural analysis: TDDFT and related models. In *Practical Aspects of Computational Chemistry I: An Overview of the Last Two Decades of Current Trends*, Leszczynski, J.; Shukla, M. K., Ed.; Springer Science + Business Media: Dordrecht, 2012; Ch. 14, pp. 415−449.

(13) Le Bahers, T.; Adamo, C.; Ciofini, I. A qualitative index of spatial extent in charge-transfer excitations. *J. Chem. Theory Comput.* **2011**, *7*, 2498−2506.

(14) Ehara, M.; Fukuda, R.; Adamo, C.; Ciofini, I. Chemically intuitive indices for charge-transfer excitation based on SAC-CI and TD-DFT calculations. *J. Comput. Chem.* **2013**, *34*, 2498−2501.

(15) Jacquemin, D.; Le Bahers, T.; Adamo, C.; Ciofini, I. What is the "best" atomic charge model to describe through-space charge-transfer excitations? *Phys. Chem. Chem. Phys.* **2012**, *14*, 5383−5388.

(16) Huet, L.; Perfetto, A.; Muniz-Miranda, F.; Campetell, M.; Adamo, C.; Ciofini, I. General density-based index to analyze charge transfer phenomena: From models to butterfly molecules. *J. Chem. Theory Comput.* **2020**, *16*, 4543−4553.

(17) Savarese, M.; Guido, C. A.; Brémond, E.; Ciofini, I.; Adamo, C. Metrics for molecular electronic excitations: A comparison between orbital- and density-based descriptors. *J. Phys. Chem. A* **2017**, *121*, 7543−7549.

(18) Wang, Z.; Liang, J.; Head-Gordon, M. Earth mover's distance as a metric to evaluate the extent of charge transfer in excitations using discretized real-space densities. *J. Chem. Theory Comput.* **2023**, *19*, 7704−7714.

(19) Fraiponts, M.; Maes, W.; Champagne, B. Earth mover's charge transfer distance: A general and robust approach for describing excited state locality. *J. Chem. Theory Comput.* **2024**, *20*, 2751−2760.

(20) Campetella, M.; Maschietto, F.; Frisch, M. J.; Scalmani, G.; Ciofini, I.; Adamo, C. Charge transfer excitations in TDDFT: A ghost-hunter index. *J. Comput. Chem.* **2017**, *38*, 2151−2156.

(21) Maschietto, F.; Campetella, M.; García, J. S.; Adamo, C.; Ciofini, I. Chasing unphysical TD-DFT excited states in transition metal complexes with a simple diagnostic tool. *J. Chem. Phys.* **2021**, *154*, 204102.

(22) Peach, M. J. G.; Le Sueur, C. R.; Ruud, K.; Guillaume, M.; Tozer, D. J. TDDFT diagnostic testing and functional assessment for triazene chromophores. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4465−4470.

(23) Peach, M. J. G.; Tozer, D. J. Illustration of a TDDFT spatial overlap diagnostic by basis function exponent scaling. *J. Mol. Struct.* **2009**, *914*, 110−114.

(24) Wiggins, P.; Williams, J. A. G.; Tozer, D. J. Excited state surfaces in density functional theory: A new twist on an old problem. *J. Chem. Phys.* **2009**, *131*, 091101.

(25) Dwyer, A. D.; Tozer, D. J. Effect of chemical change on TDDFT accuracy: Orbital overlap perspective of the hydrogenation of retinal. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2816−2818.
(26) 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.
(27) 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.
(28) 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.
(29) 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.
(30) 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.
(31) Mewes, S. A.; Plasser, F.; Krylov, A.; Dreuw, A. Benchmarking excited-state calculations using exciton properties. *J. Chem. Theory Comput.* **2018**, *14*, 710−725.
(32) Mai, S.; Plasser, F.; Dorn, J.; Fumanal, M.; Daniel, C.; González, L. Quantitative wave function analysis for excited states of transition metal complexes. *Coord. Chem. Rev.* **2018**, *361*, 74−97.
(33) 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.
(34) 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*, 034107.
(35) 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*, 054112.
(36) 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.
(37) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85−109.
(38) 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.
(39) Tsuneda, T.; Hirao, K. Long-range correction for density functional theory. *WIREs Comput. Mol. Sci.* **2014**, *4*, 375−390.
(40) Kümmel, S. Charge-transfer excitations: A challenge for time-dependent density functional theory that has been met. *Adv. Energy Mater.* **2017**, *7*, 1700440.
(41) 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.
(42) Ciofini, I.; Le Bahers, T.; Adamo, C.; Odobel, F.; Jacquemin, D. Through-space charge transfer in rod-like molecules: Lessons from theory. *J. Phys. Chem. C* **2012**, *116*, 11946−11955.
(43) García, G.; Adamo, C.; Ciofini, I. Evaluating push−pull dye efficiency using TD-DFT and charge transfer indices. *Phys. Chem. Chem. Phys.* **2013**, *15*, 20210−20219.
(44) Adamo, C.; Jacquemin, D. The calculations of excited-state properties with time-dependent density functional theory. *Chem. Soc. Rev.* **2013**, *42*, 845−856.
(45) Laurent, A. D.; Adamo, C.; Jacquemin, D. Dye chemistry with time-dependent density functional theory. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14334−14356.
(46) Jacquemin, D.; Adamo, C. Computational molecular electronic spectroscopy with TD-DFT. In *Density-Functional Methods for Excited States*, Ferré, N.; Filatov, M., Eds.;Springer: Switzerland, 2016; *Topics in Current Chemistry*, Vol. *368*, pp. 347−376.
(47) Le Bahers, T.; Brémond, E.; Ciofini, I.; Adamo, C. The nature of vertical excited states of dyes containing metals for DSSC applications: Insights from TD-DFT and density based indexes. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14435−14444.
(48) Adamo, C.; Le Bahers, T.; Savarese, M.; Wilbraham, L.; García, G.; Fukuda, R.; Ehara, M.; Rega, N.; Ciofini, I. Exploring excited states using time dependent density functional theory and density-based indexes. *Coord. Chem. Rev.* **2015**, *304*−*305*, 166−178.
(49) Maschietto, F.; Campetella, M.; Frisch, M. J.; Scalmani, G.; Adamo, C.; Ciofini, I. How are the charge transfer descriptors affected by the quality of the underpinning electronic density? *J. Comput. Chem.* **2018**, *39*, 735−742.
(50) García, J. S.; Maschietto, F.; Campetella, M.; Ciofini, I. Using density based indexes and wave function methods for the description of excited states: Excited state proton transfer reactions as a test case. *J. Phys. Chem. A* **2018**, *122*, 375−382.
(51) Etienne, T.; Assfeld, X.; Monari, A. Toward a quantitative assessment of electron transitions' charge-transfer character. *J. Chem. Theory Comput.* **2014**, *10*, 3896−3905.
(52) Etienne, T.; Assfeld, X.; Monari, A. New insight into the topology of excited states through detachment/attachment density matrices-based centroids charge. *J. Chem. Theory Comput.* **2014**, *10*, 3906−3914.
(53) Etienne, T. Probing the locality of excited states with linear algebra. *J. Chem. Theory Comput.* **2015**, *11*, 1692−1699.
(54) Etienne, T. Theoretical insights into the topology of molecular excitons from single-reference excited states calculation methods. In *Excitons*. Pyshkin, S. L., Ed.; IntechOpen, 2018, pp. 31−54.
(55) Pluhar, E. A., III; Ullrich, C. A. Visualizing electronic excitations with the particle-hole map: Orbital localization and metric space analysis. *Eur. Phys. J. B* **2018**, *91*, 137.
(56) Sharifzadeh, S. Many-body perturbation theory for understanding optical excitations in organic molecules and solids. *J. Phys.: Condens. Matter* **2018**, *30*, 153002.
(57) Mester, D.; Kállay, M. Charge-transfer excitations within density functional theory: How accurate are the most recommended approaches? *J. Chem. Theory Comput.* **2022**, *18*, 1646−1662.
(58) Hall, D.; Sancho-García, J. C.; Pershin, A.; Beljonne, D.; Zysman-Colman, E.; Olivier, Y. Benchmarking DFT functionals for excited-state calculations of donor−acceptor TADF emitters: Insights on the key parameters determining reverse inter-system crossing. *J. Phys. Chem. A* **2023**, *127*, 4743−4757.
(59) Korsaye, F.-A.; Perrella, F.; Petrone, A.; Adamo, C.; Rega, N.; Ciofini, I. Monitoring density redistribution at the excited state in a dual emitting molecule: An analysis based on real-time density functional theory and density descriptors. *J. Phys. Chem. A* **2024**, *128*, 4324−4334.
(60) Selenius, E.; Sigurdarson, A. E.; Schmerwitz, Y. L. A.; Levi, G. Orbital-optimized versus time-dependent density functional calculations of intramolecular charge transfer excited states. *J. Chem. Theory Comput.* **2024**, *20*, 3809−3822.
(61) Lieberherr, A. Z.; Gori-Giorgi, P.; Giesbertz, K. J. H. Optimal transport distances to characterize electronic excitations. *J. Chem. Theory Comput.* **2024**, *20*, 5635−5642.
(62) Bogo, N.; Stein, C. J. Benchmarking DFT-based excited-state methods for intermolecular charge-transfer excitations. *Phys. Chem. Chem. Phys.* **2024**, *26*, 21575−21588.
(63) Kuritz, N.; Stein, T.; Baer, R.; Kronik, L. Charge-transfer-like $\pi \rightarrow \pi^*$excitations in time-dependent density functional theory: A conundrum and its solution. *J. Chem. Theory Comput.* **2011**, *7*, 2408−2415.
(64) 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.
(65) Miura, M.; Aoki, Y. Ab initio theory for treating local electron excitations in molecules and its performance for computing optical properties. *J. Comput. Chem.* **2009**, *30*, 2213−2230.
(66) Cui, G.; Fang, W.; Yang, W. Reformulating time-dependent density functional theory with non-orthogonal localized molecular orbitals. *Phys. Chem. Chem. Phys.* **2010**, *12*, 416−421.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

(67) 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*, 034106.
(68) 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.
(69) 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.
(70) Macetti, G.; Genoni, A. Quantum mechanics/extremely localized molecular orbital embedding strategy for excited states: Coupling to time-dependent density functional theory and equation-of-motion coupled cluster. *J. Chem. Theory Comput.* **2020**, *16*, 7490−7506.
(71) Mashkovtsev, D.; Orimoto, Y.; Aoki, Y. Fast and accurate calculation of the UV−vis spectrum with the modified local excitation approximation. *J. Chem. Theory Comput.* **2023**, *19*, 5548−5562.
(72) Sen, S.; Senjean, B.; Visscher, L. Characterization of excited states in time-dependent density functional theory using localized molecular orbitals. *J. Chem. Phys.* **2023**, *158*, 054115.
(73) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm-Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291−299.
(74) Byun, Y.-M.; Ullrich, C. A. Excitons in solids from time-dependent density-functional theory: Assessing the Tamm-Dancoff approximation. *Computation* **2017**, *5*, 9.
(75) Furche, F.; Rappoport, D. Density functional methods for excited states: Equilibrium structure and electronic spectra. In *Computational Photochemistry*, Olivucci, M., Eds.;Elsevier: Amsterdam, 2005; *Theoretical and Computational Photochemistry*, Vol. *16*, Ch. 3, pp. 93−128.
(76) Casida, M. E.; Ipatov, A.; Cordova, F. Linear-response time-dependent density functional theory for open-shell molecules. In *Time-Dependent Density Functional Theory*, Marques, M. A. L.; Ullrich, C. A.; Nogueira, F.; Rubio, A.; Burke, K.; Gross, E. K. U., Eds.;Springer-Verlag: Berlin, 2006; *Lecture Notes in Physics*, Vol. *706*, pp. 243−257.
(77) Cordova, F.; Doriol, L. J.; Ipatov, A.; Casida, M. E.; Filippi, C.; Vela, A. Troubleshooting time-dependent density-functional theory for photochemical applications: Oxirane. *J. Chem. Phys.* **2007**, *127*, 164111.
(78) Tapavicza, E.; Tavernelli, I.; Rothlisberger, U.; Filippi, C.; Casida, M. E. Mixed time-dependent density-functional theory/classical trajectory surface hopping study of oxirane photochemistry. *J. Chem. Phys.* **2008**, *129*, 124108.
(79) Furche, F. On the density matrix based approach to time-dependent density functional response theory. *J. Chem. Phys.* **2001**, *114*, 5982−5992.
(80) Garcia-Alvarez, J. C.; Gozem, S. Absorption intensities or organic molecules from electronic structure calculations versus experiments: The effect of solvation, method, basis set, and transition moment gauge. *J. Chem. Theory Comput.* **2024**, *20*, 7227−7243.
(81) Herbert, J. M. Correction: Visualizing and characterizing excited states from time-dependent density functional theory. *Phys. Chem. Chem. Phys.* **2024**, *26*, 23503−23504.
(82) Plötner, J.; Tozer, D. J.; Dreuw, A. Dependence of excited state potential energy surfaces on the spatial overlap of the Kohn−Sham orbitals and the amount of nonlocal Hartree−Fock exchange in time-dependent density functional theory. *J. Chem. Theory Comput.* **2010**, *6*, 2315−2324.
(83) Richard, R. M.; Herbert, J. M. Time-dependent density-functional description of the ^1L_a state in polycyclic aromatic hydrocarbons: Charge-transfer character in disguise? *J. Chem. Theory Comput.* **2011**, *7*, 1296−1306.
(84) Alam, B.; Jiang, H.; Zimmerman, P. M.; Herbert, J. M. State-specific solvation for restricted active space spin−flip (RAS-SF) wave functions based on the polarizable continuum formalism. *J. Chem. Phys.* **2022**, *156*, 194110.
(85) Herbert, J. M.; Harriman, J. E. Comparison of two-electron densities reconstructed from one-electron density matrices. *Int. J. Quantum Chem.* **2002**, *90*, 355−369.
(86) Barford, W.; Paiboonvorachat, N. Excitons in conjugated polymers: Wavefunctions, symmetries, and quantum numbers. *J. Chem. Phys.* **2008**, *129*, 164716.
(87) Ma, H.; Qin, T.; Troisi, A. Electronic excited states in amorphous MEH-PPV polymers from large-scale first principles calculations. *J. Chem. Theory Comput.* **2014**, *10*, 1272−1282.
(88) Mewes, S. A.; Plasser, F.; Dreuw, A. Universal exciton size in organic polymers is determined by nonlocal orbital exchange in time-dependent density functional theory. *J. Phys. Chem. Lett.* **2017**, *8*, 1205−1210.
(89) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.
(90) Plasser, F.; Krylov, A. I.; Dreuw, A. libwfa: Wavefunction analysis tools for excited and open-shell electronic states. *WIREs Comput. Mol. Sci.* **2022**, *12*, e1595.
(91) IQmol. www.iqmol.org. (accessed 9 August 2024).
(92) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(93) Haranczyk, M.; Gutowski, M. Visualization of molecular orbitals and the related electron densities. *J. Chem. Theory Comput.* **2008**, *4*, 689−693.
(94) Murray, C. W.; Handy, N. C.; Laming, G. J. Quadrature schemes for integrals of density functional theory. *Mol. Phys.* **1993**, *78*, 997−1014.
(95) Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017**, *38*, 869−882.
(96) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. A standard grid for density-functional calculations. *Chem. Phys. Lett.* **1993**, *209*, 506−512.
(97) Gray, M.; Bowling, P. E.; Herbert, J. M. Comment on "Benchmarking basis sets for density functional theory thermochemistry calculations: Why unpolarized basis sets and the polarized 6-311G family should be avoided. *J. Phys. Chem. A* **2024**, *128*, 7739−7745.
(98) Boys, S. F. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.* **1960**, *32*, 296−299.
(99) Subotnik, J. E.; Shao, Y.; Liang, W.; Head-Gordon, M. An efficient method for calculating maxima of homogeneous functions of orthogonal matrices: Applications to localized occupied orbitals. *J. Chem. Phys.* **2004**, *121*, 9220−9229.
(100) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775−4777.
(101) 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.
(102) Surján, P. R. Natural orbitals in CIS and singular-value decomposition. *Chem. Phys. Lett.* **2007**, *439*, 393−394.
(103) Nitta, H.; Kawata, I. A close inspection of the charge-transfer excitation by TDDFT with various functionals: An application of orbital- and density-based analyses. *Chem. Phys.* **2012**, *405*, 93−99.
(104) Herbert, J. M.; Head-Gordon, M. Calculation of electron detachment energies for water cluster anions: An appraisal of electronic structure methods, with application to $(H_2O)_{20}^-$ and $(H_2O)_{24}^-$. *J. Phys. Chem. A* **2005**, *109*, 5217−5229.
(105) Herbert, J. M.; Head-Gordon, M. Accuracy and limitations of second-order many-body perturbation theory for predicting vertical detachment energies of solvated-electron clusters. *Phys. Chem. Chem. Phys.* **2006**, *8*, 68−78.
(106) Herbert, J. M. The quantum chemistry of loosely-bound electrons. In *Reviews in Computational Chemistry*, Parill, A. L.; Lipkowitz, K., Eds.;Wiley-VCH: Hoboken, 2015; Vol. *28*, Ch. 8, pp. 391−517.
(107) Kaufmann, K.; Baumeister, W.; Jungen, M. Universal Gaussian basis sets for an optimum representation of Rydberg and continuum wavefunctions. *J. Phys. B: at., Mol. Opt. Phys.* **1989**, *22*, 2223−2240.

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463

(108) Coccia, E.; Mussard, B.; Labeye, M.; Caillat, J.; Taïeb, R.; Toulouse, J.; Luppi, E. Gaussian continuum basis functions for calculating high-harmonic generation spectra. *Int. J. Quantum Chem.* **2016**, *116*, 1120−1131.
(109) Labeye, M.; Zapata, F.; Coccia, E.; Véniard, V.; Toulouse, J.; Caillat, J.; Taïeb, R.; Luppi, E. Optimal basis set for electron dynamics in strong laser fields: The case of molecular ion H_2^+. *J. Chem. Theory Comput.* **2018**, *14*, 5846−5858.
(110) Morassut, C.; Coccia, E.; Luppi, E. Quantitative performance analysis and comparison of optimal-continuum Gaussian basis sets for high-harmonic generation spectra. *J. Chem. Phys.* **2023**, *159*, 124108.
(111) Luppi, E.; Head-Gordon, M. The role of Rydberg and continuum levels in computing high harmonic generation spectra of the hydrogen atom using time-dependent configuration interaction. *J. Chem. Phys.* **2013**, *139*, 164121.
(112) White, A. F.; Heide, C. J.; Saalfrank, P.; Head-Gordon, M.; Luppi, E. Computation of high-harmonic generation spectra of the hydrogen molecule using time-dependent configuration-interaction. *Mol. Phys.* **2016**, *114*, 947−956.
(113) Coccia, E.; Luppi, E. Detecting the minimum in argon high-harmonic generation spectrum using Gaussian basis sets. *Theor. Chem. Acc.* **2019**, *138*, 96.
(114) Witzorky, C.; Paramonov, G.; Bouakline, F.; Jaquet, R.; Saalfrank, P.; Klamroth, T. Gaussian-type orbital calculations for high harmonic generation in vibrating molecules: Benchmarks for H_2^+. *J. Chem. Theory Comput.* **2021**, *17*, 7353−7365.
(115) Zhu, Y.; Herbert, J. M. High harmonic spectra computed using time-dependent Kohn−Sham theory with Gaussian orbitals and a complex absorbing potential. *J. Chem. Phys.* **2022**, *156*, 204123.
(116) Geometries for Tozer's data set (introduced in ref 9) are taken from https://djtozer.webspace.durham.ac.uk/excite3 (accessed 16 August 2024).
(117) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975−2988.
(118) Jamorski, C.; Foresman, J. B.; Thilgen, C.; Lüthi, H.-P. Assessment of time-dependent density-functional theory for the calculation of critical features in the absorption spectra of a series of aromatic donor−acceptor systems. *J. Chem. Phys.* **2002**, *116*, 8761−8771.
(119) Mewes, S.; Plasser, F.; Dreuw, A. Communication: Exciton analysis in time-dependent density functional theory: How functionals shape excited-state characters. *J. Chem. Phys.* **2015**, *143*, 171101.
(120) Plasser, F. Entanglement entropy of electronic excitations. *J. Chem. Phys.* **2016**, *144*, 194107.
(121) Knippenberg, S.; Starcke, J. H.; Wormit, M.; Dreuw, A. The low-lying excited states of neutral polyacenes and their radical cations: A quantum chemical study employing the algebraic diagrammatic construction scheme of second order. *Mol. Phys.* **2010**, *108*, 2801−2813.
(122) Grimme, S.; Parac, M. Substantial errors from time-dependent density functional theory for the calculation of excited states of large π systems. *ChemPhysChem* **2003**, *4*, 292−295.
(123) Plasser, F. Visualisation of electronic excited-state correlation in real space. *ChemPhotoChem* **2019**, *3*, 702−706.
(124) Chen, T.; Zheng, L.; Yuan, J.; An, Z.; Chen, R.; Tao, Y.; Li, H.; Xie, X.; Huang, W. Understanding the control of singlet-triplet splitting for organic exciton manipulating: A combined theoretical and experimental approach. *Sci. Rep.* **2015**, *5*, 10923.
(125) Taka, A. A.; Herbert, J. M.; McCaslin, L. M. Ground-state orbital analysis predicts S_1 charge transfer in donor−acceptor materials. *J. Phys. Chem. Lett.* **2023**, *14*, 11063−11068.
(126) Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **2002**, *117*, 7433−7447.
(127) Etienne, T. A comprehensive, self-contained derivation of the one-body density matrices from single-reference excited-state calculation methods using the equation-of-motion formalism. *Int. J. Quantum Chem.* **2020**, *120*, e26110.
(128) Head-Gordon, M.; Graña, A. M.; Maurice, D.; White, C. A. Analysis of electronic transitions as the difference of electron attachment and detachment densities. *J. Phys. Chem.* **1995**, *99*, 14261−14270.
(129) Ohio Supercomputer Center. http://osc.edu/ark:/19495/f5s1ph73 (accessed 17 August 2024).

https://doi.org/10.1021/acs.jctc.4c01085
J. Chem. Theory Comput. 2024, 20, 9446−9463