Tracing potential energy surfaces of electronic excitations via their transition origins: application to Oxirane

Jian-Hao Li, T. J. Zuehlke, M. C. Payne, and N. D. M. Hine

TCM Group, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom

(Dated: December 16, 2014)

We show that the transition origins of electronic excitations identified by quantified natural transition orbital (QNTO) analysis can be employed to locate crossings among potential energy surfaces across a wide range of molecular geometries. By tracing the switching of transition origins of adiabatic potential surfaces as the geometry changes, conical intersections can be located to a small geometrical region. The transition vectors for analysing transition origins are provided by linear response time-dependent density functional theory (TDDFT) calculations under the Tamm-Dancoff approximation. We study the photochemical CO ring opening of oxirane as an example and show that the results corroborate the traditional Gomer-Noyes mechanism derived experimentally. The state-specific information for the reaction also agrees well with that given by previous theoretical work using TDDFT surface-hopping dynamics that was validated by high-quality quantum Monte Carlo calculations. We also show that QNTO can be useful for considerably larger and more complex systems: by projecting the excitations to those of a reference oxirane molecule, the approach is able to identify and analyse specific excitations of a trans-2,3-Diphenyloxirane molecule.

I. Introduction

Kohn-Sham density functional theory (DFT)\(^\text{11}\) can accurately reproduce many properties of molecular and solid state systems\(^\text{5–7}\), as compared to experimental methods and higher-level theory, e.g. molecular structures, vibrational frequencies, atomisation energies and ionisation potentials. Meanwhile, its time-dependent extension, TDDFT\(^\text{12–15}\), particularly in the linear response (LR) formalism\(^\text{16}\), can deliver properties of excited states and optical absorption spectra of molecular systems. DFT and TDDFT calculations are increasingly able to treat large, complex biological systems, such as proteins, enzymes, and nucleic acids, due to the increasing availability of large computational facilities in recent years\(^\text{17–24}\).

However, there remain serious obstacles to such uses: the intrinsic cubic scaling of traditional KS-DFT with respect to the number of electrons; the explosion of the size of the available configurational phase space associated with large, flexible molecules; and the challenges of accurate electrostatics of solvated systems\(^\text{25–27}\). When excited state properties are considered, further challenges come into play such as the proliferation of spurious charge transfer excitations, and the difficulty of identifying connected excited states on a potential energy surface.

Recent advances in linear-scaling (LS)-DFT\(^\text{17,23}\) have made the study of very large molecular systems much more feasible. LS-DFT relies on the exponential decay of the density matrix, \(\rho(r, r')\), with respect to \(|r - r'|\) in a system with a finite band-gap\(^\text{25,27}\). In such systems, the density matrix, expressed using local orbitals, can be truncated by value or by range to render it local. In this work we use the ONETEP code\(^\text{23}\), where a minimal number of local optimisable functions, called nonorthogonal generalised Wannier functions (NGWFs), are optimised in situ to best represent the density matrix of the system. Recently, Linear-Response TDDFT Tamm-Dancoff approximation (TDA) calculations have also become available in ONETEP, using a similar density matrix formalism\(^\text{28}\). Such methods can provide valuable knowledge about the electronic excitations of very large molecular systems such as proteins and DNA segments. However, the low symmetry and complex geometry of biological systems exacerbates the aforementioned problem of tracing the potential energy surface of specific excited states. This work therefore focuses on quantified natural transition orbital (QNTO) analysis, recently proposed by Li et al.\(^\text{29}\), which offers progress towards this goal.

QNTO analysis rewrites the transition vectors of an electronic excitation identified by a theoretical calculation, e.g. TDDFT, in terms of the natural transition orbitals (NTOs)\(^\text{11,30,31}\). The projection of hole (NTO\(_H\)) and electron (NTO\(_E\)) orbitals of the nth NTO pair onto a standard orbital basis set such as natural bond orbitals (NBOs)\(^\text{32–34}\) then helps quantify the transition origins of the NTOs of the excitation. Alongside the absorption energy and oscillator strength, knowledge of these transition origins helps provide a physical interpretation of the role of a given excitation dominated by single excitation character. Similar analysis may also be developed for single de-excitations, which is not as important in generic biological systems where de-excitations are usually negligible,\(^\text{11}\) signified by the good approximation of TDDFT/TDA to TDDFT. In addition, even in large and complex systems, electronic excitations may involve numerous electronic transitions between valence and conduction orbitals; using NTOs as the orbital basis, the number of dominant electronic transitions is often reduced to one or two, helping clarify what transition components contribute to the electronic excitations. QNTO analysis can help to assess the variation of transition origins determined by the different theoretical methods available in an electronic excitation calculation. For example, this approach has been used to assess the per-
performance of different exchange-correlation functionals in predicting transition origins by comparing to a high-level theoretical calculation. The variations of transition origins caused by the presence of different molecular environments and/or different molecular geometries can also be investigated. For example, QNTO has been used to study the role of the DNA backbone in mediating the transition origins of electronic excitations of B-DNA.

In the present paper, we introduce theoretical background to Linear-Response calculations in ONETEP (Sec. II), and show that QNTO analysis can be implemented with a non-orthogonal basis set such as the NGWFs used in ONETEP (Sec. III). In Sec. IV we then demonstrate two applications of LR-TDDFT/TDA/QNTO: (1) studying the photo-driven ring opening reaction of oxirane, and (2) identifying and comparing similar excitations between trans-2,3-Diphenyloxirane and oxirane. We demonstrate that QNTO analysis can be readily used to locate state crossings that are otherwise unclear from plotting potential energy curves. The results are consistent with previous experimental and theoretical studies, showing that the present method is a feasible and reliable approach for investigating excited state related reactions.

II. LR-TDDFT in ONETEP

Linear-scaling DFT as implemented in ONETEP relies on a reformulation of the single-electron density matrix in terms of local orbitals, as follows

\[ \rho(r, r') = \sum_p f_p \psi_p(r) \psi_p^*(r') = \phi_\alpha(r) K^{\alpha\beta} \phi_\beta^*(r') \]  

where \( \psi_p(r) \) are single-particle eigenstates with occupation \( f_p \), while \( \phi_\alpha(r) \) are spatially-localised nonorthogonal orbitals. \( K^{\alpha\beta} \) is a generalised matrix representation of the occupancies, known as the density kernel. Note that implicit summation over repeated Greek indices is used throughout this work. The density kernel is defined by \( K^{\alpha\beta} = \langle \phi^\alpha | \rho_{KS} | \phi^\beta \rangle \), with \( \rho_{KS} \) being the KS single-particle density operator. The nonorthogonal orbitals \( \{ \phi_\alpha \} \) and their duals \( \{ \phi^\alpha \} \) form a biorthonormal basis set such that \( \langle \phi_\alpha | \phi^\beta \rangle = \delta_{\alpha\beta} \).

ONETEP employs a minimal number of NGWFs, each expressed in terms of periodic bandwidth limited delta functions (psincs) centred on the grid points of the simulation cell, with a spacing determined by a kinetic energy cutoff akin to that in a plane-wave code. NGWFs are localised to atom-centred spherical regions defined by a cutoff \( R_\sigma \). The total energy calculation proceeds by minimising the total energy simultaneously with respect to the matrix elements of the density kernel and the expansion coefficients of the NGWFs in terms of the psinc functions, via nested loops of conjugate gradients optimisation. NGWFs are initialised to pseudoatomic orbitals which solve the pseudopotential for a free atom. We confine our attention to closed-shell systems (even number of electrons) with non-fractional occupancy, and the total electronic energy of the real system is written as

\[ E_0 = 2 \text{Tr} \left[ K_{\{v\}} H^{(v)} \right] + E_{DC} \]  

where \( H^{(v)}_{\alpha\beta} = \langle \phi_\alpha | \hat{H}_{\text{proj}} | \phi_\beta \rangle \) and \( K^{\alpha\beta}_{\{v\}} = \langle \phi^\alpha | \hat{\rho}^{(v)} | \phi^\beta \rangle \) with \( \hat{\rho}^{(v)} = \sum_{\nu} |\psi_\nu\rangle \langle \psi_\nu| \). The factor 2 accounts for the spin degeneracy and \( E_{DC} \) is the standard “double-counting” correction. An FFT box technique is employed to allow one to construct matrix elements and density components in \( O(1) \) computational effort independent of system size.

As described in detail in Ref., the minimisation is carried under the constraint of conserved total electron number,

\[ N = 2 \text{Tr} \left[ K_{\{v\}} S^{(v)} \right] \]  

where \( S_{\alpha\beta}^{(v)} = \langle \phi_\alpha | \phi_\beta \rangle \), and also the constraint of idempotent density matrix,

\[ K_{\{v\}} = K_{\{v\}} S^{(v)} K_{\{v\}} \]  

which is equivalent to requiring the orthonormal orbitals have occupancy of 0 or 1. An important advantage of optimising the NGWFs in situ is that basis set superposition error (BSSE) is eliminated. Sophisticated algorithms have been implemented so that a high performance can be achieved running on parallel computers.

Because the NGWFs are optimised in situ to describe the occupied part of the density matrix, they do not in general provide a good description of unoccupied states. Therefore, we instead generate and optimise in situ a second set of NGWFs, denoted as \( \chi_\alpha \), obtained by minimising the bandstructure energy associated with a chosen range of low-energy conduction band states, having first projected out the valence states. The bandstructure energy to be minimised is written as

\[ E = 2 \text{Tr} \left[ K_{\{c\}} H^{(c)} \right] \]  

where \( H^{(c)}_{\alpha\beta} = \langle \chi_\alpha | \hat{H}_{\text{proj}} | \chi_\beta \rangle \) and \( K^{\alpha\beta}_{\{c\}} = \langle \chi^\alpha | \hat{\rho}^{(c)} | \chi^\beta \rangle \). A full description of this procedure can be found in Ref.

Recently Zuchlsdorff et al. proposed a formulation for linear-scaling LR-TDDFT/TDA (hereafter simply denoted as TDDFT) calculations, which uses the conduction and valence NGWFs as a basis for describing the response densities of electronic excitations. This approach adapts the Casida formulation for finding transition vectors, \( \mathbf{X}_T \), and excitation energies \( \omega_T \) by using the
where \( S \) is a joint set, \( \{ \alpha \} = \{ \phi_\alpha \} \cup \{ \chi_\alpha \} \). Using this representation, and the corresponding overlap matrix \( S^{(j)} \), in place of \( \{ \alpha \} \) and \( S^{(c)} \), means that we no longer need the explicit representation of the conduction space density operator. Instead, the conduction space projector can be replaced by subtracting the valence density operator from the identity operator within the subspace of the full joint NGWF representation. This defines \( K^{(j)} \) through \( K^{(j)} = \langle \alpha | \beta \rangle | \beta \rangle = \langle \alpha | \bar{I} - \bar{V} | \beta \rangle \) giving

\[
K^{(j)} = S^{(j)} - S^{(j)\dagger} \bar{V} K^{(c)} \bar{V} S^{(j)} \tag{9}
\]

where \( S^{(j)}_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle \) and \( S^{(j)}_{\alpha\beta} = S^{(j)}_{\beta\alpha} \) is the inverse of \( S^{(j)} \).

We also note that in the TDDFT the KS-system is linked to the real-system by

\[
\sum_{cv} \psi^{s}_v(r) \psi^{c}_v(r) \langle \Psi_0 | \hat{a}^\dagger_v \hat{a}_v | \Psi_I \rangle = \sum_{cv} \psi^{s}_v(r) \psi^{c}_v(r) x^{s}_cv \tag{10}
\]

where \( | \Psi_0 \rangle \) and \( | \Psi_I \rangle \) are the ‘exact’ many-body wavefunctions that are obtained from diagonalising the exact Hamiltonian in the configuration space. The linear dependence among the products \( \{ \psi^{s}_v(r) \psi^{c}_v(r) \} \) becomes likely as the orbital basis approaches completion\cite{West2003} and the simple assignment \( \langle \Psi_0 | \hat{a}^\dagger_v \hat{a}_v | \Psi_I \rangle = x^{s}_cv \) may not be justified. Nevertheless, this assignment would become reasonable\cite{West2003} if the number of orbitals involved in \( \langle \Psi_0 | \hat{a}^\dagger_v \hat{a}_v | \Psi_I \rangle \) is not too large, e.g. for low-lying excitations, and therefore the study of transition origins becomes possible via directly analysing \( x^{s}_cv \) or \( K^{(1)} \) from the TDDFT calculation.

### III. Quantified natural transition orbital (QTO) analysis

In this section, we first introduce the basic formulation of QTO analysis in terms of an orthonormal orbital basis such as KS-orbitals. We then detail how it is implemented in ONETEP based on a non-orthogonal NGWF basis. The standard notation \( ^t \) is used throughout to denote conjugate transpose.

#### A. Natural transition orbitals

The NTO basis

\[
\begin{align*}
| \psi^{(Nc)}_n \rangle &= \sum_{c} U_{cn} | \psi^{c}_n \rangle \\
| \psi^{(Nv)}_n \rangle &= \sum_{v} V_{vn} | \psi^{v}_n \rangle 
\end{align*}
\]

is an orbital basis that makes \( \Sigma = U^\dagger X_c V \) diagonal, with non-negative real numbers \( \sqrt{\lambda_n} \) on the diagonal. In this representation the transition density operator is simplified to require summation over just one index:

\[
\rho^{(1)} = \sum_{n=1}^{N_N} | \psi^{(Nc)}_n \rangle \sqrt{\lambda_n} | \psi^{(Nv)}_n \rangle \tag{12}
\]

\( N_N \) denotes the number of non-zero values \( \sqrt{\lambda_n} \), which are ranked by magnitude. \( \sqrt{\lambda_n} \) are called the singular values from the singular value decomposition (SVD) of \( X_c = U \Sigma V^\dagger \). We also see that the \( U \) and \( V \) respectively diagonalise \( X_c X_c^\dagger \) and \( X_v X_v^\dagger \),

\[
\begin{align*}
U^\dagger X_c X_c^\dagger U &= \Sigma \Sigma^\dagger \\
V^\dagger X_v X_v^\dagger V &= \Sigma^\dagger \Sigma 
\end{align*}
\]

with the same eigenvalue set \( \{ \lambda_n \} \), where \( X^\dagger_c X_c \) and \( X^\dagger_v X_v \) are density matrices for the electron and hole orbitals, respectively.
In TDDFT calculations in ONETEP, the local orbitals (NGWFs) are not an orthonormal basis. If the NTOs are expressed in terms of NGWFs as

$$|\psi_n^{(\text{NGWF})}\rangle = |\zeta_n\rangle U_n^\alpha$$
$$|\psi_n^{(\text{NTO})}\rangle = |\phi_n\rangle V_n^\alpha$$  (14)

$X_I$ is then rewritten as $X_I = U^\dagger S^{(v)} K_{(1)} S^{(c)} \nabla \nabla^\dagger$ and the target for decomposition becomes

$$K_{(1)} = U \Sigma V^\dagger$$  (15)

If we hope to directly employ an SVD algorithm for orthonormal basis, we would have to first construct $MK_{(1)} N^\dagger = (MU) \Sigma (N V)^\dagger$, where $M^\dagger M = S^{(c)}$ and $N^\dagger N = S^{(v)}$. Once the singular vector matrices $M U$ and $N V$ are obtained, we would then need to calculate $M^{-1}$ and $N^{-1}$ in order to solve for $U$ and $V$.

However, since the square root and inverse of a sparse matrix are in general much less sparse, their calculation would be very costly when dealing with large matrices. As an alternative approach, we can instead choose to diagonalise the $X_I^T X_I$ and $X_I^T X_I$ as shown in Eq. 13.

Since $V S^{(v)} = V^{-1}$ and $U S^{(c)} = U^{-1}$, the eigenvalue equations can be derived as

$$(S^{(c)} K_{(1)} S^{(v)}) (S^{(c)} K_{(1)} S^{(v)}) \nabla = (\Sigma \Sigma^T) S^{(c)} U$$
$$(S^{(v)} K_{(1)}^T S^{(c)} K_{(1)} S^{(v)}) V = (\Sigma^T \Sigma) S^{(c)} \nabla$$  (16)

Note that in the SVD of $X_I = U \Sigma V^\dagger$, the relative phase between each pair of left (NTO-H) and right (NTO-H) singular vectors are fixed: if the left vector is multiplied by $(-1)$, the right vector must be multiplied by $(-1)$ to remain a solution. Hence, the relative phase between each pair of eigenvectors needs to be determined after solving Eq. 16. The following relation from Eq. 15 can be used:

$$K_{(1)}^{\alpha \beta} = \sum_{n=1}^{N_X} \sqrt{\lambda_n U_n^\alpha V_n^\beta}$$  (17)

If we define a matrix $d_{n \pm}$ as

$$d_{n \pm}^{\alpha \beta} = K_{(1)}^{\alpha \beta} - \left( \sum_{m=1}^{n-1} \sqrt{\lambda_m U_m^\alpha V_m^\beta} \pm \sqrt{\lambda_n U_n^\alpha V_n^\beta} \right)$$  (18)

there is the relation:

$$\text{Tr} \left[ d_{n \pm} S^{(v)} d_{n \mp}^\dagger S^{(c)} \right] < \text{Tr} \left[ d_{n \pm} S^{(v)} d_{n \mp}^\dagger S^{(c)} \right]$$  (19)

We may start from NTO1 and proceed iteratively through higher NTOs, checking which of the associated two eigenvectors obtained in Eq. 16 denoted as $U_n$ and $V_n$, produces the smaller value in Eq. 19. If not, we simply multiply the whole vector $V_1^\beta$ (or $U_1^\alpha$) by $(-1)$ to obtain a correct pair of $V_1^\beta$ and $U_1^\alpha$. As this is a separate operation from the calculation of $U_1^\alpha$ (Eq. 16), we can first solve for a number of eigenvalues $n$, where $n \leq N_X$, and subsequently define a number $m$, where $(m \leq n)$ of these where the associated singular vectors are of interest. This can reduce the computational effort compared to directly running an SVD algorithm given that many excitations will be dominated by just NTO1 or a small number of NTO$n$ pairs.

**B. Interpretation of the transition origins of NTO$n$ electron-hole pairs**

After obtaining the NTO-based transition vector element to any NTO$n$, $n \leq N_X$, next we want to interpret the transition origins of each NTO$n$ electron-hole pair from the projection to a standard orbital set. In the present work, we use the natural bond orbitals (NBOs)\(^{22}\) as the standard orbital set. By such a projection the composition of NTO$n$-H and NTO$n$-E can be quantitatively defined, and, for example, the precise contribution of an orbital of interest to an excitation can be computed. We may be interested in the hole-orbital involved in a $^1\pi\pi^*$ excitation of a molecule, and can compute its quantitative involvement in an excitation of the same molecule embedded in a chemical/physical environment with/without covalent bond formed.

Note that the transition vector, $X_I$, can be based on different sets of reference orbitals in different calculations. Rewriting the transition vector using the unique NTO basis removes this reference dependence. Projection of NTOs onto a common orbital set such as NBOs then provides a unification of interpretative orbital set for these NTOs. For example, excitations in two systems of different environments and/or geometries can be compared. Using the root mean square deviation, $\sigma$, of the projection coefficients with respect to a common standard orbital set, two NTOs can be compared\(^{22}\). The transition origins of each excitation can be seen as its fingerprint, helping to identify excitations of similar character. The potential energy surfaces of excited states can then be constructed. We also note that the similarity of two NTO$n$-H(E) orbitals can be quantified by the magnitude of projection from one to the other. The projection sign can be arbitrary, as switching the overall phases of both NTO$n$-H and NTO$n$-E orbitals at the same time will result in them remaining a singular vector pair in Eq. 13. We will use this method for comparing two NTO$n$-H(E) orbitals in the present work.

We generate NBOs by applying the NBO 5.\(^{22}\) program as a postprocessing step after the ONETEP calculation. A relatively large atomic orbital (AO) basis set is used in ONETEP for generating NBOs, which reduces the electron spillage to a low level (typically 0.006%). In the systems studied here, the HOMO energy deviation is less than 0.03 eV when compared to a calculation using
standard in-situ optimised NGWFs.

As an example we can examine the 2nd excitation of an oxirane molecule using TDDFT/QNTO. The geometry of the oxirane molecule, with CCO angle=100°, is taken from a scan along the CCO angle as will be discussed in greater detail in Sec. IV. There are 9 valence (occupied) KS-orbitals in the calculation, and therefore the HOMO(LUMO) is indexed 9(10) etc. The transition density operator expressed in the KS-orbital basis results from summation over several pairs of orbitals:

$$\hat{\rho}^{(1)}_{1=2} = |\psi_{10}| 0.72 |\psi_8| + |\psi_{10}| 0.12 |\psi_9| + |\psi_{10}| 0.10 |\psi_6| + |\psi_{10}| 0.01 |\psi_9| + |\psi_{14}| 0.01 |\psi_8| + |\psi_{11}| 0.01 |\psi_8| + \ldots$$

(20)

By contrast, in the NTO basis it is mostly accounted for by a single NTO pair:

$$\hat{\rho}^{(1)}_{1=2} = |\psi_{1}^{(Nc)}| 0.98 |\psi_{2}^{(Nc)}| + |\psi_{2}^{(Nc)}| 0.13 |\psi_{2}^{(Nv)}| + |\psi_{2}^{(Nc)}| 0.09 |\psi_{3}^{(Nv)}| + |\psi_{4}^{(Nv)}| 0.07 |\psi_{4}^{(Nv)}| + \ldots$$

(21)

The transition density operator is dominated by the NTO1 transition, which accounts for about 97% of the total transition density, whereas in the KS-orbital basis the largest component, $|\psi_{10}| 0.72 |\psi_8|$, only accounts for about 51% of the density. In Fig. 1 we plot the NTO1-H and NTO1-E orbitals as well as the most dominant NBOs for each.

In all calculations a cut-off Coulomb approach has been employed to avoid the influence of periodic images from neighbouring simulation cells. A minimal set of 1 NGWF for H atom and 4 NGWFs for C, N, and O atoms is used. All calculations are well-converged with respect to the variational parameters of the localisation radii; suitable values are 10 bohr for valence NGWFs and 16 bohr for conduction NGWFs. The kinetic energy cutoff for the grid spacing is well-converged at 1000 eV. Norm-conserving pseudopotentials are used to replace core electrons. The (A)LDA functional in the Perdew-Zunger parameterisation is used throughout.

FIG. 1. The NTO1-H and NTO1-E orbitals for the 2nd excitation of an oxirane molecule with CCO=100°. The 1st dominant NBO for each NTO1-H(E) and the projection coefficient is shown below. Isovalue=0.05 e/bohr. The NBO labels specify bonding (BD) and anti-bonding (BD*) and denote which pairs of atoms (eg C1 and O3) are involved in a given bond.

IV. Results and Discussion

In this section, we discuss two applications of QNTO. In subsection A, we focus on tracing the potential energy surfaces associated with the ring opening process of oxirane. In subsection B, we use QNTO to locate and compare similar excitations between trans-2,3-Diphenyloxirane and those of oxirane.

FIG. 2. Gomer-Noyes mechanism. The underlying snapshots are from the CCO angle scans discussed in the text.

FIG. 3. Constrained optimisation scan for different CCO angles followed by TDLDA/TDA calculation. All internal coordinates during the scan are optimised except for the CCO bond angle that is fixed in each step.
Each pair of CCO angles form a block in the map, which is divided into 9 cells for different pairing of excitations. The first and second number in a cell is the projection result (%) for NTO-H and NTO-L, respectively. A cell with yellow background indicates that both numbers are ≥ 1/2.

| CCO angle | Excitation number | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   |
|-----------|-------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| 60        | 0.9997            | 0.9999 | 0.9999 | 0.9997 | 0.9997 | 0.9985 | 0.9992 | 0.9961 |
| 65        | 0.9997            | 0.9999 | 0.9999 | 0.9980 | 0.9994 | 0.9986 | 0.9976 | 0.9980 |
| 70        | 0.9995            | 0.9999 | 0.9999 | 0.9976 | 0.9992 | 0.9916 | 0.9900 | 0.9983 |
| 75        | 0.9997            | 0.9998 | 0.9999 | 0.9998 | 0.9964 | 0.9989 | 0.8737 | 0.8832 | 0.9473 |
| 80        | 0.9999            | 0.9999 | 0.9999 | 0.9921 | 0.9420 | 0.9874 | 0.7348 |
| 85        | 0.9999            | 0.9998 | 0.9992 | 0.8237 | 0.7862 | 0.9972 | 0.9978 | 0.9560 |
| 90        | 0.9999            | 0.8798 | 0.8477 | 0.9971 | 0.9977 | 0.9142 | 0.9858 | 0.9589 |
| 95        | 0.9999            | 0.9860 | 0.9963 | 0.9997 | 0.9954 | 0.9809 | 0.9996 |
| 100       | 0.9999            | 0.9931 | 0.9994 | 0.9966 | 0.9997 | 0.9859 | 0.9734 | 0.9877 |
| 105       | 0.9999            | 0.9863 | 0.9996 | 0.9997 | 0.7387 | 0.7487 | 0.9853 |
| 110       | 0.9989            | 0.9568 | 0.9810 | 0.9067 | 0.9952 | 0.8161 | 0.7883 | 0.8233 |
| 115       | 0.9997            | 0.9959 | 0.9992 | 0.9983 | 0.9993 | 0.9969 | 0.9965 | 0.9877 |

### A. Photoinduced ring opening of oxirane

The Gomer-Noyes mechanism, as depicted in Fig. 2, was postulated and experimentally confirmed as a way to explain the photoinduced ring opening process of the oxirane molecule. More discussions about this textbook molecule of stereochemistry can be found, for example, in Ref. 49. In particular, the detailed state-specific information of the Gomer-Noyes mechanism has been examined in some detail through TD-DFT excited state molecular dynamics (MD) simulations validated by high-quality quantum Monte Carlo calculations, and several states have been found to be involved in mediating the ring-opening reaction.

Here, we perform a constrained optimisation scan along the CCO bond angle of oxirane. We relax all internal coordinates except for a fixed CCO bond angle, which is set to values in the range 65° to 110°. For each oxirane geometry in the scan, a TD-DFT calculation is performed followed by QNTQ analysis to obtain the transition origins of the first few low-lying excitations.

The potential energy curves of the low-lying singlet and triplet states are plotted in Fig. 3. For reference we also show the negative HOMO energy and the value of the HOMO-LUMO gap at each angle. Note that in TD-DFT, use of the TDA avoids the problems associated with singlet and triplet instabilities. The former can occur for systems which have strong static correlation, signified by the vanishing HOMO-LUMO gap, whereas the latter can occur for systems which have an unstable closed-shell ground state, reflected by a negative triplet excitation energy within the TD-DFT/TDA. A caveat is that the calculation quality would deteriorate for those excitations of energies higher than the negative HOMO energy. As this tends to be underestimated by the TD-DFT/TDA, one may obtain fewer bound states than are present in the real system.

We note that from the excitation energies alone there
For the first three excitations, the smallest have also been plotted, with isovalue=0 surfaces have switched their characters of excitations can be represented by the transition density operator for these systems is dominated by the NTO1 component (Eq. 21), the character of excitations is from the 3rd excitation of the oxirane geometry at CCO=90°, which still captures about 72% of the density contribution. Thus, all the excitations on the oxirane geometries from the scan are dominated by the NTO1 transition, and when a switch of the NTO1 transition origins occurs, it shows that the excitation changes its character.

To put this in a form suitable for locating CX points, we construct a map (Fig. 4) showing the projections of the NTO1 pairs for each of the first three excitations (S1, S2, S3) at each geometry onto their equivalents at different geometries, forming a 3 × 3 block for each pair of geometries. We refer to the geometry listed down the row of the map as the ‘System’ (Sys) and compare it to a ‘Reference’ (Ref) geometry along the columns. To obtain a meaningful projection while accounting for the geometry change, the NTO1-H(E) orbitals are translated from the Sys-geometry to the Ref-geometry and renormalised before being used for projection.

The optimised atom-centred NGWFs |φα,s⟩ from the calculation in the Sys-geometry are translated to the corresponding atom centres in the Ref-geometry to produce NGWFs |φα,s→r⟩. The NTO1-H orbital (Eq. 14) is then re-expressed in terms of these using the original coefficients Vα1,r

\[ |\psi_1^{(Nv)}⟩ = |φ_{α,s→r}⟩ V_{1,s}^{α} . \]  

The translated orbital is then renormalised:

\[ |\psi_1^{(Nv)}⟩ = \frac{|\psi_1^{(Nv)}⟩}{\sqrt{\langle \psi_1^{(Nv)} | V_{1,s}^{α} | \psi_1^{(Nv)}⟩}} , \]  

and each 3 × 3 block of projections is calculated as

\[ \langle \psi_1^{(Nv)} | J_1 | ψ_1^{(Nv)}⟩; \langle ψ_1^{(Nv)} | J_2 | ψ_1^{(Nv)}⟩; \langle ψ_1^{(Nv)} | J_3 | ψ_1^{(Nv)}⟩ \]  

for I = 1, 2, 3 and J = 1, 2, 3.

In Fig. 5 one can immediately see which cells involve pairs of excitations with similar transition origins. We highlight those that have both projections greater than 1/√2, indicating that the density contribution to one
This can be due to the fact that a CX between S caused by a character change of the ground state itself. The latter would be passed between the two excitations. The former case can indicate that there has been one or more CXs all excited states, from one block to another. The former method can help us locate those reaction-facilitating CX points, whose precise locations, if one is interested, may further be searched for using other high-level theoretical methods.59 Fig. 5 shows an example where the switch of dominant NTO character cannot be seen from the plot of NTO1-H(E) orbitals, a switch of the dominating electron orbitals is observed between CCO=65° – 66°. For more details see the text.

Such examples of switches of excitation which show subtly different behaviour. We also note that due to the scheme of TDDFT and the approximations used, the behaviour of TDDFT around CXs can be incorrect, e.g. having too rapid energy variation. TDDFT has, however, been able to offer a reasonable description of oxirane photochemistry55 and we stress that the current method can help us locate those reaction-facilitating CX points, whose precise locations, if one is interested, may further be searched for using other high-level theoretical methods.59 Fig. 8 shows an example where the switch of excitation character previously shown from the coarser map to occur between 60° and 65° is narrowed down to occurring abruptly between 62° and 63°. The corresponding potential energy curves for the two states involved, and the NTO1-H and NTO1-E orbitals before and after the switch, are plotted in Fig. 9 showing that...
FIG. 10. NBOs corresponding to Table I. The lone pair on the oxygen (LP/O3) accounts for a large proportion of the NTO1-H for all three excitations. The 1st excitation is represented by NBOs denoted BD*/C2 H7, BD*/C1 H4, BD*/C1 H6, and BD*/C2 H5, whereas the 2nd excitation is represented by the same set of NBOs each with a reduced contribution, plus the more important BD*/C1 H2. The 3rd excitation is represented by BD*/C1 O3 and BD*/C2 O3.

We note that the exact point where the domination switches occur can be somewhat affected by the choice of reference orbitals used, although this does not in general prevent us from tracing down the CX to a small region in the scanned coordinate. In the language of diabatic configuration[22], the situation can be analyzed by considering the matrix elements of a Hamiltonian \( \hat{H} \) acting on a reduced subspace consisting of just two such configurations, denoted as \( |\Phi_1\rangle \) and \( |\Phi_2\rangle \). The switch of domination of two diabatic configurations as a given geometry path is traversed indicates that the surface \( H_{11} - H_{22} = 0 \) has been crossed, though the exact shape of the surface \( H_{11} - H_{22} = 0 \) will be affected by the reference diabatic configurations used.

For a CX between \( S_1 \) and \( S_0 \), on the other hand, it would be reflected in the map by the fact that a whole column of excitations change transition character at the same time, because the ground state itself on which excitations are based has changed character. We are able to locate two such points in Fig. 6 between 107° and 108°, and between 111° and 112°, both related to the change of the ground state character expected from the Gomer-Noyes mechanism.

After locating the regions where two adiabatic excited state curves switch their characters, we can plot the adiabatic singlet states again, with the curves for the various excitations ‘re-connected’ based on their transition origins, as shown in Fig. 9. State labels of \( ^1(n,3p_\pi) \) and \( ^1(n,3\sigma) \) based on the symmetry at the CCO=60° geometry and propagated to other geometries based on the connection of the states.

It then becomes clear that population of the \( S_1' \) state at CCO=60° can directly lead to a crossing between \( S_0 \) and \( S_1 \). The energy decreases monotonically along with the widening CCO angle, which will result in a rapid ring opening reaction. This is consistent with the previous excited states MD study[23] which found that the population of the \( ^1(n,3p_\pi) \) state would facilitate the oc-
TABLE II. Decomposition of NTO1-H(E) into NBOs for the first three excitations of oxirane at CCO=60°. N indicates the excitation number. Proj denotes the projection result. NBO type gives information on the NBO feature. For each excitation, the first 6 ranked NBOs of lone pair (LP), bonding (BD), and anti-bonding (BD*) types have been listed. The highly basis-dependent Rydberg-type NBOs (RY*) that accounts for most of these excitations are not listed. Note that the projection magnitude of the 6th-ranked NBO has gone down to below 0.1 for all cases. Total means the total density contribution (square sum) of all the 6 NBOs listed. These NBOs are plotted in Fig. 10.

| N  | NTO1-H   | NTO1-E   |
|----|---------|---------|
|    | NBO type | Proj    | NBO type | Proj    |
| 1  | LP O3    | 0.894   | BD* C2 H7 | -0.168 |
|    | BD C2 H7 | -0.215  | BD* C1 H4 | -0.167 |
|    | BD C2 H5 | 0.213   | BD* C1 H6 | -0.165 |
|    | BD C1 H6 | 0.199   | BD* C2 H5 | -0.164 |
|    | BD C1 H4 | -0.198  | BD* C1 O3 | 0.055  |
|    | BD* C2 H5| 0.080   | BD* C2 O3 | -0.048 |
|    | Total    | 0.976   | Total     | 0.116  |
| 2  | LP O3    | -0.894  | BD* C1 C2 | -0.210 |
|    | BD C2 H7 | 0.215   | BD* C1 H4 | 0.124  |
|    | BD C2 H5 | -0.213  | BD* C2 H5 | -0.123 |
|    | BD C1 H6 | -0.200  | BD* C1 H6 | 0.121  |
|    | BD C1 H4 | 0.198   | BD* C2 H7 | -0.119 |
|    | BD* C2 H5| -0.080  | BD* C1 O3 | 0.019  |
|    | Total    | 0.976   | Total     | 0.104  |
| 3  | LP O3    | 0.890   | BD* C1 O3 | -0.306 |
|    | BD C2 H7 | -0.219  | BD* C2 O3 | 0.270  |
|    | BD C2 H5 | 0.218   | BD* C2 H7 | -0.077 |
|    | BD C1 H6 | 0.203   | BD* C1 H4 | -0.076 |
|    | BD C1 H4 | -0.202  | BD* C1 H6 | -0.067 |
|    | BD* C2 H5| 0.080   | BD* C2 H5 | -0.065 |
|    | Total    | 0.977   | Total     | 0.187  |

Occurrence of ring opening process. If the 1(n,3s) state is populated instead, it has to overcome an energy barrier to a widening of CCO angle before meeting a crossing with the 1(n,3p₂) state which can then lead to the ring opening of oxirane. Thus, this chemical process is more difficult and would produce a lower quantum yield. We identify that the ground state is re-populated from the 1(n,3s) state after oxirane passing the critical point between 107° and 112°.

Meanwhile, passing the critical point between 111° and 112° leads to another change of ground state character, which corresponds to a proton transfer process producing acetaldehyde, the product in the 3rd stage of the Gomer-Noyes mechanism (Fig. 2). This is also consistent with the MD study finding that the proton shift occurs after the oxirane has been de-excited to the ground state. Under thermally elevated conditions this proton shift will likely occur at lower CCO angles: here we address only the quasistatic constrained geometries. A follow-up reaction breaking the C-C bond, the final stage in Fig. 2 would then be possible while remaining in the ground state, given an appropriate combination of internal motion.

To demonstrate the chemical composition of these transitions, a decomposition of NTO1-H and NTO1-E into Natural Bond Orbitals for the first three excitations at CCO=60° is shown in Table II. The AO basis used for generating NBOs is fairly large, producing 525 NBOs, but these are mostly of Rydberg (RY*) type. In practice, the lone-pair (LP) and bonding (BD) NBOs describe the majority of the NTO1-H excitation, which in all three cases is well-described by just five main NBOs. For the NTO1-E orbitals of the three excitations, however, a substantial contribution comes from Ry* orbitals, which are highly dependent on the AO basis set used in generating NBOs. The LP and BD NBOs for NTO1-H, and the most relevant BD* NBOs for NTO1-E are shown in Fig. 10.

B. Trans-2,3-Diphenyloxirane

In this subsection we focus on a significantly larger system, the trans-2,3-Diphenyloxirane molecule, which contains the same CCO subsystem as oxirane as its central motif. We show that we can use the QNTO method to compare the excitations with those of an oxirane molecule. As would be expected, molecules with different physical/chemical environments can have distinct properties. For example, aryl substitution of oxirane can lead to a different ring opening mechanism. The fact that the trans-2,3-Diphenyloxirane, referred to as the system (Sys) molecule, is formed by two aryl substitutions on oxirane makes it a good example for examining the influence of chemical environment on the excitations of the core oxirane moiety using TDDFT/QNTO.

The structure of the Sys-molecule is first fully optimised without constraints. A equivalent geometry for the oxirane molecule, referred to as the reference (Ref) molecule, is then extracted from the optimised Sys-molecule by replacing the two aryl substituents with hydrogen atoms and optimising the C-H bond lengths, at fixed bond angles. The widespread problem that the energy of charge-transfer excitations is underestimated by local density functionals such as the LDA means in this case that we expect to find large numbers of charge-transfer excitations in the energy window of interest. In
practice we find we need to calculate up to 100 excitations of the Sys-molecule so that the highest (100th) excitation energy calculated, 7.74 eV, exceeds the energy of the 8th excitation of the Ref-molecule (7.61 eV), which was the highest Ref-excitation studied in the current work.

The NTO1-H and NTO1-E orbitals for the excitations of Sys-molecule (Sys-excitations) and Ref-molecule (Ref-excitations) are then compared to each other to see if similar excitations can be identified between the first 8 Ref-excitations and the first 100 Sys-excitations. This is achieved by constructing, for each NTO1-H(E), a renormalised orbital based on just the component of each on the local orbitals of the 5 atoms common to both geometries (the oxirane core). We denote the NTO1-H(E) orbital of a Sys-excitation as $|s\rangle$, and the corresponding core orbital cNTO1-H(E) in the form $|sc\rangle$. Similarly, we define $|r\rangle$ and $|rc\rangle$ for the Ref-molecule. The $|r\rangle$ orbitals are the NTO1-H(E) orbitals of the original reference oxirane geometry, whereas $|s\rangle$, $|sc\rangle$, and $|rc\rangle$ are constructed via a similar procedure to that used in constructing the projections of the previous section.

In particular, constructing $|s\rangle$ and $|sc\rangle$ requires a translation of the NTO1-H(E) from the Sys-molecule geometry to the Ref-molecule geometry. Using NTO1-H as an example, the set below are constructed:

$$
|s\rangle' = |\phi_{\alpha,s\rightarrow r}\rangle \tilde{V}_{1,s}$$
$$|sc\rangle' = |\phi_{\alpha,s\rightarrow r}\rangle \tilde{V}_{1,s}$$
$$|rc\rangle' = |\phi_{\alpha,r}\rangle \tilde{V}_{1,r}$$

(26)

followed by renormalisation as in Eq. 23 to obtain $|s\rangle$, $|sc\rangle$, and $|rc\rangle$, respectively. The “s→r” in Eq. 26 denotes that the orbital is translated from the Sys-molecule to the Ref-molecule. $\tilde{U}_{1}$ and $\tilde{V}_{1}$ are defined by setting respectively the coefficients $U_{1,s}$ and $V_{1,s}$ in Eq. 14 to zero for those NGWFs not centred on the core atoms; the subscript ‘s’ for Sys-excitations and ‘r’ for Ref-excitations indicate that the original NGWF coefficient matrices have been used.

By evaluating the overlap $\langle rc|sc \rangle$ between an NTON-H or NTON-E orbital of the Sys-molecule and the corresponding NTO of the Ref-molecule, one can then tell...
TABLE III. Matchings between Sys-excitations and Ref-excitations having $|\langle rc|sc \rangle| \geq 1/ \sqrt{3}$ for both the NTO1-H and NTO1-E. The first and second number in a “(.)” is the projection result (%) for NTO1-H and NTO1-E, respectively. 0.0 denotes 100. Numbers shown in boldface are the excitations plotted in Fig. [12]. Given the freedom of orbital shape outside the core region, it would be no surprising to see that the 2nd and 4th Ref-excitation match with multiple Sys-excitations.

| Ref Exc. | Sys Exc. | Sys 1 | Sys 2 | Sys 3 | Sys 4 | Sys 5 | Sys 6 | Sys 7 | Sys 8 |
|----------|----------|-------|-------|-------|-------|-------|-------|-------|-------|
| (1)      | (2)      |       |       |       |       |       |       |       |       |
| Sys 1    | (2)      | 61.38 | 90.65 | 0.00  | 0.37  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 2    | (2)      | 90.65 | 0.00  | 0.00  | 0.37  | 61.38 | 0.00  | 0.00  | 0.00  |
| Sys 3    | (2)      | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 4    | (2)      | 0.37  | 0.00  | 61.38 | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 5    | (2)      | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 6    | (2)      | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 7    | (2)      | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |
| Sys 8    | (2)      | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |

whether there is a strong overlap of the character in the core region. In addition, the overlaps $\langle sc|s \rangle$ and $\langle rc|r \rangle$ allow one to determine how large a proportion of the total excitation is accounted for by the component within the core region.

In the general cases where excitations are dominated by the NTO1, one can then say that the two excitations from different molecules are similar if the following three conditions hold: (1) a Sys-excitation must have a substantial $\langle sc|s \rangle$ for both NTO1-H and NTO1-E; (2) a Ref-excitation must have a substantial $\langle rc|r \rangle$ for both NTO1-H and NTO1-E; and (3) the magnitude of $\langle rc|sc \rangle$ must be large.

We find that of all the 100 excitations of the larger system, there are 16 whose cNTO1-H and cNTO1-E respectively match with the cNTO1-H and cNTO1-E of one or more Ref-excitations. We choose a threshold of $|\langle rc|sc \rangle| \geq 1/ \sqrt{2}$ for both cNTO1-H and cNTO1-E. Overall there are 20 matching combinations, as some excitations overlap significantly with several excitations of the reference system. The 1st, 2nd, 4th, and 8th Ref-excitation are involved in these matching pairs.

If the threshold is lowered to $|\langle rc|sc \rangle| \geq 1/ \sqrt{3}$, the number of involved Sys-excitations becomes 25 and the number of matchings becomes 35. The 3rd and 7th Ref-excitation also now match with some Sys-excitations. All the 35 excitation matchings are listed in Table III.

In Table III we see that $|\langle rc|r \rangle| \geq 1/ \sqrt{2}$ for both NTO1-H and NTO1-E of the 1st-6th and 8th Ref-excitation. As for $\langle sc|s \rangle$, of all the involved 25 Sys-excitations discussed above only the 34th Sys-excitation (hereafter denoted as Sys-34) satisfies $|\langle sc|s \rangle| \geq 1/ \sqrt{2}$ for both NTO1-H and NTO1-E. In other words, for the remaining 24 Sys-excitations there are $\geq 50\%$ density contributions of NTO1-H and/or NTO1-E which come from outside of the core region. This would be unsurprising as there remain 22 atoms outside of the core region compared to 5 atoms in the core, and the de-localisation of the density contribution for many Sys-excitations is anticipated. If we lower the threshold to $|\langle sc|s \rangle| \geq 1/ \sqrt{3}$ for both NTO1-H and NTO1-E, the Sys-37 and 44 also fall into the set. Some excitation properties of the Sys-34, 37, and 44 along with two more excitations, Sys-71 and 81, respectively for comparing with the Ref-08 and Ref-03, are listed in Table IV and the NTO1-H(E) orbitals plotted in Fig. [12]. It can be seen in Table IV that the NTO1 transition vector component, $\sqrt{\lambda_1}$, has generally decreased for the larger-sized Sys-molecule, although the NTO1 still accounts for at least 60% density contribution.

In Fig. [12] we see that the NTO1-H orbitals are essentially the same for the first 7 Ref-excitations. Moreover, the core parts of the Sys-excitations, Sys-34, 37, 44, and 81, fully coincide with the core parts of this NTO1-H orbital shared by the first 7 Ref-excitations, namely, $|\langle rc|sc \rangle| = 1.00$. The differences between Sys-excitations and Ref-excitations mainly come from the more diffuse NTO1-E orbitals.

Strictly speaking, only the Sys-34 has a significant orbital contribution of NTO1-E coming from the core region, namely $|\langle sc|s \rangle|^2 \geq 0.5$. In Table IV it can be seen that the Sys-34 has an excitation energy that is close to those of the Ref-02 and Ref-04. The oscillator strength of Sys-34 also falls in between those of the Ref-02 and Ref-04. As for the orbital shape, the NTO1-E of the Sys-34 in Fig. [12] appears to be a mix between those of the Ref-02 and Ref-04, also reflected by a substantial $|\langle rc|sc \rangle|$ for both the Ref-02 and Ref-04.

The other Sys-excitations have a high proportion of NTO1-H and/or NTO1-E on the region outside the core, so comparing these Sys-excitations with the Ref-excitations would be misleading. In fact, the excitation energy can be close, as between the Ref-04 and Sys-44, or significantly different as between the Ref-03 and Sys-81. Nevertheless, the energies of similar-core excitations between the Sys- and Ref-molecule still fall reasonably close to each other, given the wide ranges of Sys- and Ref-excitation energies, i.e. 4.85-7.4 eV for the Sys-molecule and 6.04-7.61 eV for the Ref-molecule. Hence, the present method has been able to locate similar excitations of the two molecules among numerous other excitations, by appropriately tuning the thresholds of $|\langle rc|sc \rangle|$, $\langle sc|s \rangle$, and $\langle rc|r \rangle$.

Since the LDA can be inaccurate for reproducing the transition origins of molecular excitations, as demonstrated on thymine and thymidine molecule, it would be interesting to perform a similar study using self-interaction error corrected or reduced functionals incorporating the Hartree-Fock (HF)-exchange, such as the $\omega$97X-D. In the near future, or using some technique to confine electronic transitions to a small spatial range.
TABLE IV. Excitation energy (E), oscillator strength (f), and NTO1 transition vector component of several excitations listed in Table III. Sys-excitations and Ref-excitations that have similar core-orbitals are put in the same row. We note that the Ref-02 and Ref-04 are substantially similar in their cNTO1-E orbitals to have \( \langle r_{2nd}(r_{\text{ath}}) \rangle = -0.79 \).

| E (eV) | f     | NTO1   | E (eV) | f     | NTO1   | E (eV) | f     | NTO1   |
|--------|-------|--------|--------|-------|--------|--------|-------|--------|
| Ref-01 | 6.04  | 0.0253 | Ref-02 | 6.57  | 0.0001 | Ref-03 | 6.57  | 0.9997 |
| Ref-07 | 7.46  | 0.0114 | Ref-04 | 6.62  | 0.0199 | Sys-81 | 0.9982| 0.9876 |
| Sys-37 | 6.55  | 0.0070 | Sys-34 | 6.44  | 0.0062 | Sys-71 | 0.9982| 0.8851 |
| Sys-44 | 6.75  | 0.0489 | Sys-44 | 6.75  | 0.0489 | Sys-71 | 0.9982| 0.8851 |

to counteract the de-localisation error from the LDA.

V. Conclusions and future work

In this work we have developed QNTO analysis based on a non-orthogonal basis, as implemented in the linear-scaling DFT/TDDFT code ONETEP. We use the QNTO method to analyse the transition vectors of electronic excitations of various molecular systems offered by TDDFT calculations.

We first studied the photo-induced ring-opening process of the oxirane molecule. Through a scan along the reactive CCO angle of an oxirane molecule followed by TDDFT calculations, a profile of several low-lying adiabatic excited state curves on which the excited state driven reaction may be carried can be constructed. However, these adiabatic excited state curves ranked by energies do not give us very much knowledge on how the reaction will proceed. We stress that the reaction process can be made much clearer if the excited state potential energy curves are constructed according to their excitation character instead of their energies. We demonstrated that the transition vectors of these low-lying excitations are dominated by the first NTO electron-hole pair (NTO1), and therefore the transition origins of the NTO1 can be used to represent the character of these excitations. Based on this approach, we have successfully located several special geometrical points along the scanned CCO coordinate where the transition origins of two adiabatic excited state curves switch, thus indicating that there must be a conical intersection located nearby. By reconnecting the adiabatic excited state energy points (calculated by TDDFT) based on transition origins instead of energetic ordering, we are then able to depict how the excited state-driven ring-opening process can proceed. This picture is consistent with the previous excited state molecular dynamics study that found that the population of the \( ^1(3p_z) \) state facilitates the ring-opening process, while the population of \( ^1(3s) \) state does not. We have demonstrated that the \( ^1(3p_z) \) state exhibits a monotonically-decreasing energy curve along the widening CCO angle, leading to a ring-opening reaction.

QNTO analysis has also been employed to study an aryl-substituted oxirane, the trans-2,3-Diphenyloxirane molecule. The results demonstrate that the approach can be used to identify specific excitations among a large number of other excitations of a large/complex molecular system. By comparing the excitations of this larger molecule to those of oxirane, we have shown that a small number of mutually similar excitations can be identified and compared between the two molecules. The matching conditions that define similar excitations can be tuned by the magnitudes of: (1) the projection result of the core orbitals, \( c_{\text{NTO1-H}}(E) \), between Sys-excitations and Ref-excitations; (2) the proportion of the \( c_{\text{NTO1-H}}(E) \) accounting for the whole NTO1-H(E) in the Sys-excitation; (3) the proportion of the \( c_{\text{NTO1-H}}(E) \) accounting for the whole NTO1-H(E) in the Ref-excitation. For example, a Sys-excitation can have a similar cNTO1-E to that of a Ref-excitation, but the cNTO1-E of the Sys-excitation only accounts for a tiny proportion to the whole NTO1-H(E) of Sys-excitation. In such a case, the orbital-section of the NTO1-E of Sys-excitation coming from outside of the core region cannot be ignored and the two excitations can not be assigned as similar excitations. The same approach for finding similar excitations through the transition origins would be particularly powerful when there are numerous excitations in a large molecular system. For example, those excitations that are located in a specific geometric region can be readily identified, helping to analyse the chemical properties of that region in a specific energy range of light absorption.

VI. Acknowledgement

The authors would like to thank Louis Lee for very helpful discussions. All calculations in this work were carried out using the Cambridge HPC Service under EPSRC Grant EP/J017639/1. JHL acknowledges the support of Taiwan Cambridge Scholarship. TJZ acknowledges the support of EPSRC Grant EP/J017639/1 and the ARCHER eCSE programme. NDMH acknowledges the support of the Winton Programme for the Physics of Sustainability.
