

**Machine learning facilitated search of global reaction coordinate for photoisomerization reaction with nonadiabatic molecular dynamics simulation**

Goran Giudetti, Madhubani Mukherjee, Shobhit Srivastava, Sraddha Agrawal

Proposal for CSCI-596

## **Abstract**

It is proposed to study photoisomerization retinal via non-adiabatic molecular dynamics calculations. Application of machine learning based protocol will enable the determination of the reaction coordinate. Whereas most of the quantum chemical calculations assess the adiabatic potential surfaces with fixed nuclei many chemically relevant processes, such as isomerization, proceed nonadiabatically and involve a concerted motion of electrons and nuclei. Very recent experiments involving femtosecond optical pump and x-ray probe pulses enabled study of the evolution of electronic states and the motion of the nuclei during isomerization reactions of cyclohexadiene for the first time. Similar techniques will likely be applied for the study of different systems of increasing complexity. The interpretation of the results requires the development of the computational techniques . Due to the multistate and multidimensional nature of isomerization reactions computation of the relevant potential energy surfaces (PES) and the reaction coordinate is very costly. In this work, a neural network based machine learning technique is proposed to identify the most important internal coordinate by rank ordering all the internal coordinates. The reaction path and the transition state will also be validated by the use of the probability analysis.

## Introduction

Photochemical isomerization reactions are extremely important in engineering of photovoltaics, photo-switches and molecular motors<sup>1</sup>. These reactions proceed in femto- to picosecond timescale and occurs through a crossing between electronically excited and ground states of the molecules. Figure 1 shows the formation of photo-product after excitation to an electronically excited state and through the crossing between ground and excited state of the molecule.

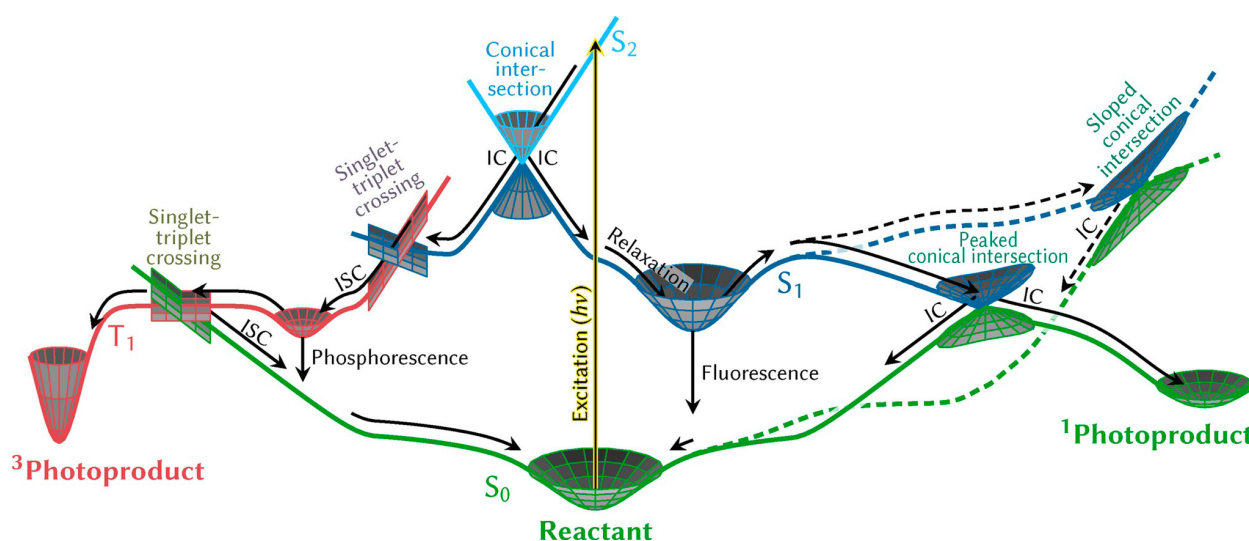


Figure 1. Photochemical processes following excitation to an electronically excited state

Simultaneous measurement of the evolution of electronic states and the motion of the nuclei remains challenging, which pose difficulty in the study of dynamics of photochemical reactions. Recently the photoinduced reaction of the ring opening of cyclohexadiene has been studied via near edge x-ray absorption fine structure (NEXAFS) spectra corresponding to core-to-valence excitations as well as with time resolved x-ray scattering<sup>2,3</sup>. These two kinds of experiments provide complementary information on the reaction.

To develop a clear understanding about the mechanism of reactions we need the potential energy surface (PES) of the molecules. PES is of fundamental importance in chemistry. The notion of a PES is based on the Born–Oppenheimer (BO) approximation which ignores the coupling between the motion of electrons and nuclei. While the BO approximation is generally accurate for molecules in their electronic ground state, it breaks down whenever two or more electronic states

come close together. Many processes involving electronically excited states such as radiationless decay, intramolecular energy and charge transfer, and most photochemical reactions are non-adiabatic, i.e., they involve two or more strongly coupled BO PESs. The most frequently used technique to simulate the dynamics of complex photochemical reactions is non-adiabatic molecular dynamics (NA-MD) using the fewest switches surface hopping (FSSH)<sup>1</sup>. In this technique the forces are computed as the gradients of single BO PESs, and the non-adiabatic effects arise from the “hopping” between the PESs<sup>4</sup>. Due to the multistate and multidimensional nature of these reactions computation of the PESs and of the reaction coordinate is highly demanding<sup>5</sup>. The significance of reaction coordinates to studies of reactive dynamics is reflected in the following aspects. Knowledge of the reaction coordinates provide the fundamental details of the underlying mechanisms of a given chemical transformation. The free energy profile (FEP) along the reaction coordinates allows to determine the activation energy and transition states, and thus the essence of the reaction dynamics<sup>6</sup>. In particular, the FEP provides a projection of the dynamics in high-dimensional space onto a few internal coordinates that allows an intuitive and immediate interpretation of a complex process. Very often the definition of a reaction path is based on intuitive considerations. Only recently, the attention was attracted to systematic approaches for selecting appropriate variables and mapping them onto multistep kinetic. The search for reaction coordinate usually involves attaining large amount of data from molecular dynamics simulation which are then analyzed in search for the reaction coordinate. This approach considers every possible coordinate as a candidate without constraining the reaction coordinates within a small set of pre-selected internal coordinates. This assumption is not necessarily true, as the identities of the correct reaction coordinates are often counter-intuitive. So care needs to be taken while selecting the reaction coordinate. The requirement to identify the reaction coordinate is a long equilibrium trajectory that samples enough transitions between stable states. However, accounting for all the internal coordinates into the reaction coordinate is not feasible except for small molecules. Machine learning based tools can be very helpful in determining the most important internal coordinate involved in the reaction mechanism. Tavadze *et al* have used a neural network based technique in search for the reaction coordinate of cis-trans isomerisation of azobenzene and shown that the C-N=N-C dihedral angle corresponds to the reaction coordinate<sup>7</sup> (Figure 2). However, choosing a single internal coordinate from machine learning as the reaction coordinate might not be sufficient, as the transition state it provides, might not be the actual transition state.

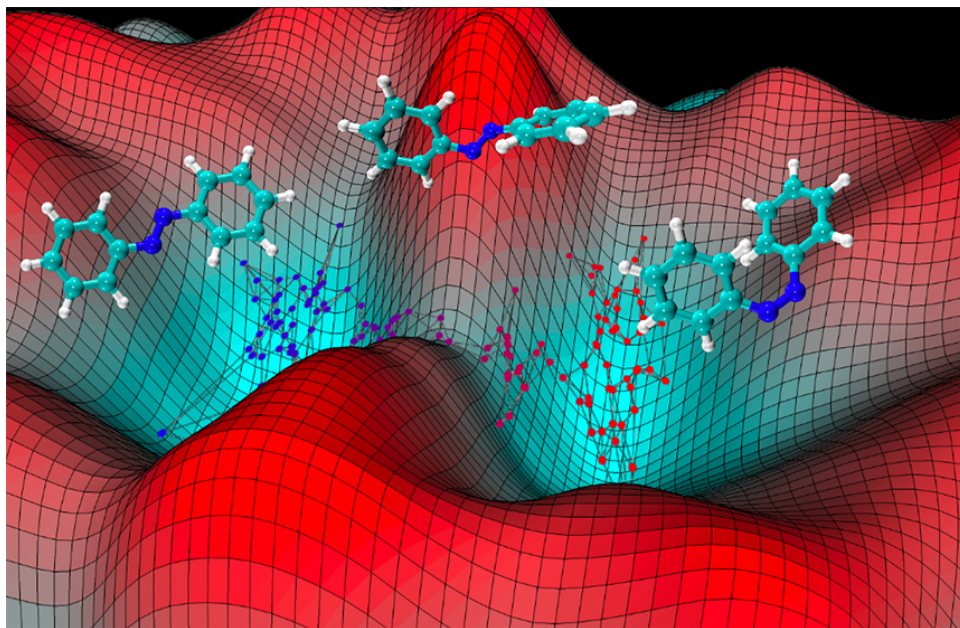


Figure 2. Multidimensional potential energy surface for photoisomerization of azobenzene

For a transition between two stable basins of PES, the reactant and the product, a trajectory initiating from a given configuration will commit to one of the basins<sup>8</sup>. Transition state was defined using this concept of splitting probability by Pratt *et al*<sup>9</sup>. Transition states are the states with a splitting probability of 0.5<sup>6</sup>, i.e. the states with equal probability to relax to the reactant and the product. Therefore, the splitting probability can be used to test whether a configuration is a transition state or not. The splitting probability was later termed as the commitment probability or committor<sup>10</sup>. In this work, committor is used as a parameter that will tell us if the transition state computed using neural network is a true transition state or not.

## Proposal

The aim of this proposal is to study photoinduced cis-trans isomerization reactions of retinal (Figure 1) using a neural network based machine learning to find the reaction coordinate by rank ordering the importance of all the internal degrees of freedom of the molecule. First, several non-adiabatic molecular dynamics (NA-MD) trajectories with fewest switch surface hopping technique will be calculated starting from different initial configurations of the reactant molecule.

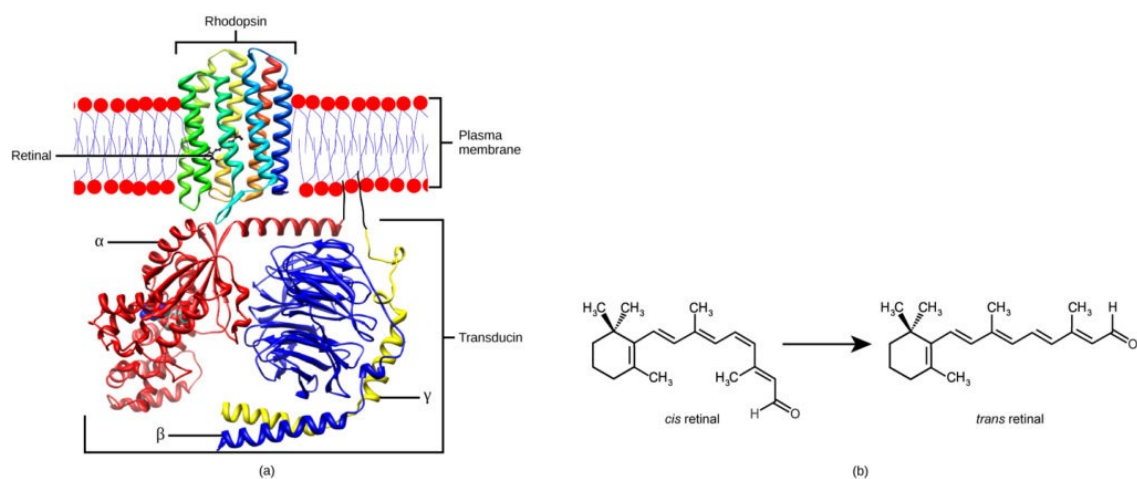


Figure 3. Proposed systems in the study

As photo-isomerization occurs through a crossing between the ground and excited state potential energy surfaces, the energy of HOMO at each geometry at different time steps in the NA-MD simulation will be monitored. From there one would know the mutual information between the HOMO energies at each geometry at each time step and different internal coordinates (attributes) at each time step. Mutual information tells how each internal coordinate is correlated to the energy of HOMO. Higher value of mutual information will assign higher rank to the particular internal coordinate. Different internal coordinates may remain correlated. To remove the redundancy in the degrees of freedom, a principal component analysis (PCA) needs to be performed, which will enable to merge any correlations and further reduce the dimensionality of the attributes. Definition of mutual information and PCA are given in the theoretical background section. Before applying PCA, one needs to first reduce the dimensionality by choosing only a few attributes that have the highest mutual information values. Some of these attributes may contain redundant internal coordinates, which will be discovered through PCA. Calculating the mutual information and filtering using principle component analysis provides those internal coordinates involved in the reaction mechanisms and their degree of participation (correlation) to the potential energy surface.

From the obtained reduced dimension set of internal coordinates, a network (Figure 4) with nodes and edges (described in the theoretical background section) representing the potential energy surface as a function of the coordinates will be built. Then the nodes corresponding to two optimized structures for local minima (reactant and product) and the shortest path between minima need to be found. The Networkx package need to be used to find the shortest path between the nodes. After finding the shortest path, a list of nodes that passed from the starting node to the final node needs to be created.

Once the shortest path i.e. the reaction coordinate has been obtained, all the features (bond length, bond angle etc.) and the energies corresponding to each node will be obtained; putting these features together will provide the reaction coordinate and the change in the energy via photoisomerization. The maximum of the HOMO energy along this path corresponds to the transition between cis and trans isomers. The structure at this transition point defines the intermediate state for the photoisomerisation. The path of least action along the HOMO energy, defines the reaction profile of the cis-trans isomerization reaction. A schematic of the workflow is shown in Figure 5.

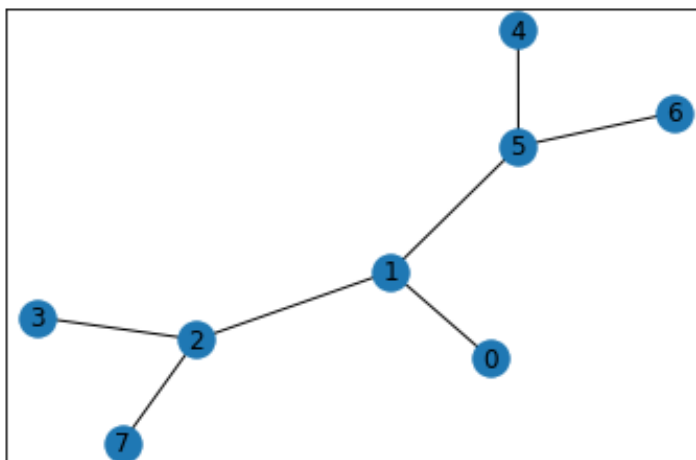


Figure 4. A graph with nodes and edges; each node contains a label

The transition state will be verified using the committor value as a parameter. Accordingly, some NA-MD trajectories starting from the transition state geometry will be generated. It will be checked if the final configuration belongs to the reactant or the product basin. The selection will be done by comparing the values of the internal coordinate at the end of the trajectory with that of the reactant and the product. For example, in the case of photoinduced cis-trans

isomerization reaction, the value of the dihedral angle, is 0° in the cis (reactant) configuration and 180° in the trans (product) configuration. Next, the committor probability will be computed. If it is close to 0.5, the transition state is verified. If it is not, another internal coordinate, which was obtained to have the second highest in rank according to the mutual information value, will be included. The linear combination of the two internal coordinates will be passed for computation of the energy profile and the transition state obtained from the corresponding energy profile will be used to compute the committor value again. This process will be repeated until the value of the committor is found to be 0.5.

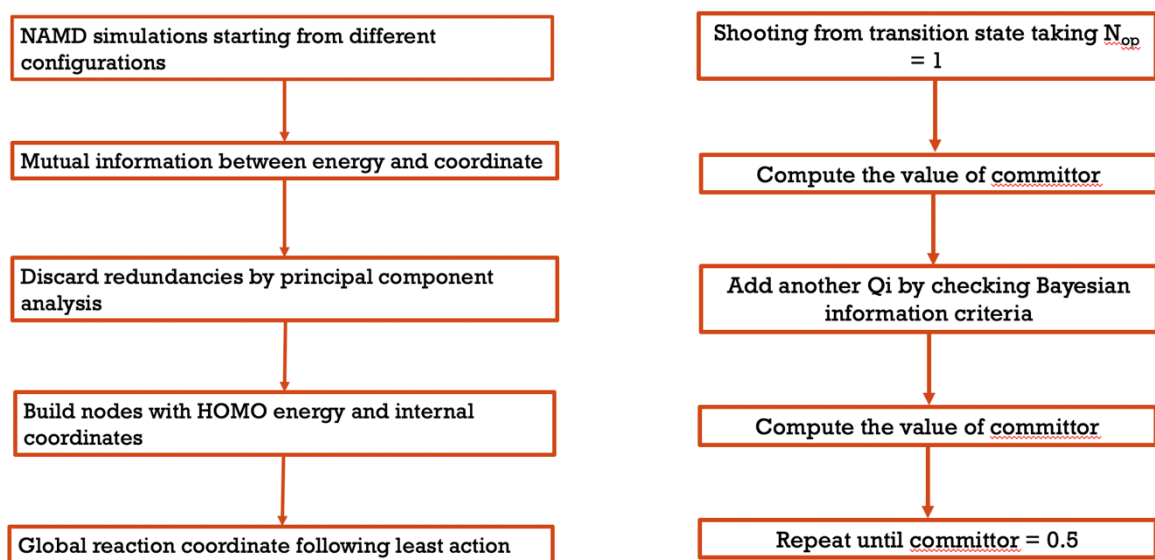


Figure 5. Flowchart to find the reaction coordinate and validate the transition state

## Theoretical Background

### A. Fewest switches surface hopping

In mixed quantum–classical dynamics, the electronic and nuclear degrees of freedom are separated – electrons evolve according to the time-dependent electronic Schrödinger equation while the nuclei are treated classically. For each nuclear configuration,  $R$ , the time-dependent electronic Schrödinger equation determines the time-evolution of the electronic wavefunction according to

$$i \frac{\partial}{\partial t} |\Psi(t, R)\rangle = \hat{H}_{el}(t, R) |\Psi(t, R)\rangle$$

The time-dependent electronic Hamiltonian is



$$H_{el}(t|R) = H_{el}(R) + \hat{V}_{ext}(t)$$

Where  $H_{el}(R)$  is the time-independent electronic Hamiltonian and  $\hat{V}_{ext}(t)$  is a time-dependent external potential.

The nuclear degrees of freedom evolve according to classical Newton's equations of motion, where  $F_j(t)$ ,  $m_j$ , and  $\ddot{R}_j(t)$  denote the force, mass, and acceleration, respectively, of the  $j^{\text{th}}$  nuclear degree of freedom at time  $t$ .

$$F_j(t) = m_j \ddot{R}_j(t)$$

Fewest switches surface hopping (FSSH) is a popular method for computing forces on the nuclei  $F_j(t)$ . The forces are computed as the gradient of single BO PES, and the non-adiabatic effects arise from the trajectory hopping between the PESs in a stochastic manner. Once the forces  $F_j(t)$  have been determined, Newton's equations of motion, may be integrated on the fly.

In the absence of an time-dependent external potential,  $\hat{V}_{ext}(t)$ , the electronic Hamiltonian reduces to the time-independent BO Hamiltonian with eigenstates  $|\Phi_n(R)\rangle$ , and their corresponding energies  $E_n(R)$  construct the PESs. The electronic wavefunction can be expanded in this BO basis,

$$|\Psi(t|R)\rangle = \sum_n c_n(t|R) |\Phi_n(R)\rangle$$

Insertion of  $|\Psi(t|R)\rangle$  into the time-dependent Schrödinger equation yields a coupled matrix equation for the time-dependent amplitude vector  $c$

$$i\dot{c} = (H - iQ + V_{ext})c$$

where  $H_{mn} = E_{mn}\delta_{mn}$  are the BO energies,  $V_{ext,mn}(t)$  is the coupling due to an external time-dependent potential, and  $Q_{mn}$  is the first-order derivative coupling between BO states  $m$  and  $n$ ,

$$Q_{mn} = \langle \Phi_m | \frac{\partial}{\partial t} \Phi_n \rangle = \dot{R} \tau_{mn}$$

The first-order non-adiabatic derivative coupling matrix elements,  $\tau_{mn}(R)$ , account for the finite kinetic energy of the nuclei:

$$\tau_{mn}(R) = \langle \Phi_m(R) | \frac{\partial}{\partial R} \Phi_n(R) \rangle$$

At each time step along a trajectory, the instantaneous change in the state populations  $|\dot{c}_n|^2$  needs to be computed. For a sufficiently small time step  $\Delta t$ , the transition probability between states  $m$  and  $n$  is

$$g_{mn} = \frac{2\Delta t}{|c_m|^2} [\Im\{c_n^*(H_{nm} + V_{ext,nm})c_m\} - \Re\{c_n^*\dot{R}\tau_{mn}c_m\}]$$

A random number,  $\zeta \in [0,1]$ , is generated and a hop from state  $m$  to  $n$  ( $m \neq n$ ) occurs if and only if  $\zeta$  satisfies

$$\sum_{k < n} g_{mk} < \zeta < \sum_{k \leq n} g_{mk}$$

If a switch is accepted, the momenta of the nuclei are scaled to conserve the total energy. Between hops, the individual trajectories evolve on single BO PESs, according to Newton's equations of motion. Thus, for a trajectory located on BO state  $m$ , the nuclei experience forces  $F_j$  given by

$$F_j = -\frac{\partial E_m(R)}{\partial R_j}$$

The FSSH algorithm preserves several desirable features of quantum–classical dynamics: the fewest switches approach ensures trajectory splitting in regions of strong coupling; within regions of avoided crossings, the method reproduces accurate state populations, given a sufficiently large number of trajectories; total energy conservation is exact and detailed balance is approximately satisfied.

## B. Machine learning: Graph neural network

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and using data without being explicitly programmed. The process of learning begins with observations or data, such as examples, direct experience, or instruction, involves looking for patterns in the data (learning) and making better decisions in the future based on the examples that are provided.

Machine learning algorithms are often categorized as supervised or unsupervised. Supervised machine learning algorithms takes what has been learned in the past as input and predict future events taking new data. The learning algorithm can also compare its output with the correct, intended output and find errors to modify the model accordingly. Unsupervised machine learning algorithms are used when the information used to train is not labeled.

Neural network is a supervised machine learning method. A neural network works similarly to the human brain's neural network. A “neuron” in a neural network is a mathematical function that collects and classifies information according to a specific architecture. The network bears a strong resemblance to statistical methods such as curve fitting and regression analysis.

A neural network contains layers of interconnected nodes as shown in Figure 6. Each node is a perceptron (an artificial neuron) and is similar to a multiple linear regression. The perceptron feeds the signal produced by a multiple linear regression into an activation function that may be nonlinear.

In a multi-layered perceptron (MLP), perceptrons are arranged in interconnected layers. The input layer collects input patterns. The input patterns may map to the output layer.

#### A simple neural network

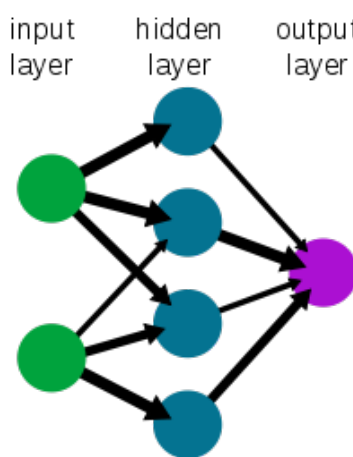


Figure 6. Schematic of a neural network

Graph Neural Network<sup>11</sup> is a type of Neural Network (GNN) which directly operates on the Graph structure. In Computer Science, a graph is a data structure consisting of two components, nodes or vertices and edges. A graph  $G$  can be well described by the set of vertices  $V$  and edges  $E$  it contains.

$$G = (V, E)$$

A typical application of GNN is node classification. Essentially, every node in the graph is associated with a label, and the goal is to predict the label of the nodes.

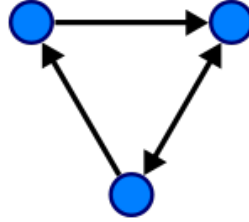


Figure 7. A graph with nodes and edges

Each node is characterized by its feature and associated with a label. Given a partially labeled graph  $G$ , the goal is to use these labeled nodes to predict the labels of the unlabeled. It learns to represent each node with a  $d$  dimensional vector which contains the information of its neighborhood nodes.

### C. Mutual information and principal component analysis

Generically consider any two random variable sets of a system,  $X$  and  $Y$ , each with its own probability distribution. In order to evaluate the correlation between these random variable sets, then one begins by measuring how similar the joint distribution  $p(X, Y)$  is to the factored distribution  $p(X)p(Y)$ . If any two variable sets are independent, then the mutual information is zero because  $p(x,y) = p(x)p(y)$  as there would be no overlap between the two distributions, where  $x \in X$  and  $y \in Y$ .

$$MI(X; Y) = \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)}$$

Principal Component Analysis (PCA) is a very common statistical procedure for reducing large datasets by examining correlations between data and generally selecting the uncorrelated variables, called principal components. To find the principal components, one needs to build a matrix of the correlations which is the matrix  $X^T X$ . In this proposal  $X$  is a matrix of all the attributes (bond length, bond angle, ..). The eigenvectors of the matrix represent the principal components of the data. The eigenvector corresponding to the largest eigenvalue will contain the largest variance in

the data. The eigenvectors are linear combinations of the original degrees of freedom of the dataset. Hence, one can take the eigenvector and determine the source of the variance from the original datasets and remove the highly correlated features which correspond to the smallest variance.

### D. Committor and optimal reaction coordinate

A model reaction coordinate is defined in terms of the order parameters, (which are internal coordinates in this proposal)  $\{Q_i\}$ , ( $i = 1, N_{op}$ ). If  $N_{op} = 2$  the simplest model reaction coordinate can be written as

$$r = \alpha_0 + \alpha_1 Q_1 + \alpha_2 Q_2$$

If  $Q_2$  depends  $Q_1$  such that  $Q_2 = \beta_{12} Q_1$  the reaction coordinate would depend on  $Q_1$  nonlinearly as

$$r = \alpha_0 + \alpha_1 Q_1 + \alpha_2 \beta_{12} Q_1$$

Using mutually orthogonal order parameters, one can eliminate nonlinear terms.  $\alpha$ 's are adjustable parameters.

To compute the reaction path, one needs to shoot from the configuration of transition state and compute the committor and half-trajectory likelihood  $L$ . Having mentioned that, let us describe the committor probability  $p_B$ , which is the probability of the trajectory originating from initial configuration to commit to the basin  $B$ .

$$p_B = \frac{1}{2}(1 + \tanh(r))$$

We can define another quantity known as the half-trajectory likelihood, which is a product over all the shooting points committed to state  $A$  and  $B$  respectively.

$$L = \prod_{x_k \rightarrow B} p_B(r) \prod_{x_k \rightarrow A} (1 - p_B(r))$$

We start by taking one internal coordinate and compute  $p_B$  and  $L$  after shooting from the obtained transition state. Then another internal coordinate is added to  $r$ . We stop once we get  $p_B = 0.5$  and  $L$  is maximized. The reaction coordinate we achieve when  $p_B = 0.5$  and maximum  $L$  is the true reaction coordinate.

## Expected Results

This protocol is expected to identify the reaction coordinates of the above-mentioned reactions and rank order them. This would help to understand the mechanisms of these reactions and also predict new reaction mechanisms. This knowledge is necessary to control the reactions by controlling the external parameters involved. If the reaction path predicted by this protocol matches the reaction paths for the reaction which are already known, this methodology would be applied to predict reaction paths and mechanism of unknown reactions.

## References

- (1) Nelson, T. R.; White, A. J.; Bjorgaard, J. A.; Sifain, A. E.; Zhang, Y.; Nebgen, B.; Fernandez-Alberti, S.; Mozyrsky, D.; Roitberg, A. E.; Tretiak, S. Non-Adiabatic Excited-State Molecular Dynamics: Theory and Applications for Modeling Photophysics in Extended Molecular Materials. *Chem. Rev.* 2020, *120* (4), 2215–2287. <https://doi.org/10.1021/acs.chemrev.9b00447>.
- (2) Attar, A. R.; Bhattacharjee, A.; Pemmaraju, C. D.; Schnorr, K.; Closser, K. D.; Prendergast, D.; Leone, S. R. Femtosecond X-Ray Spectroscopy of an Electrocyclic Ring-Opening Reaction. *Science* (80-. ). 2017, *356* (6333), 54–59. <https://doi.org/10.1126/science.aaj2198>.
- (3) Minitti, M. P.; Budarz, J. M.; Kirrander, A.; Robinson, J. S.; Ratner, D.; Lane, T. J.; Zhu, D.; Glowina, J. M.; Kozina, M.; Lemke, H. T.; Sikorski, M.; Feng, Y.; Nelson, S.; Saita, K.; Stankus, B.; Northey, T.; Hastings, J. B.; Weber, P. M. Imaging Molecular Motion: Femtosecond X-Ray Scattering of an Electrocyclic Chemical Reaction. *Phys. Rev. Lett.* 2015, *114* (25), 1–5. <https://doi.org/10.1103/PhysRevLett.114.255501>.
- (4) Tapavicza, E.; Bellchambers, G. D.; Vincent, J. C.; Furche, F. Ab Initio Non-Adiabatic Molecular Dynamics. *Phys. Chem. Chem. Phys.* 2013, *15* (42), 18336–18348. <https://doi.org/10.1039/c3cp51514a>.
- (5) Mai, S.; González, L. Molecular Photochemistry: Recent Developments in Theory. *Angew. Chemie - Int. Ed.* 2020, *59* (39), 16832–16846. <https://doi.org/10.1002/anie.201916381>.
- (6) Li, W.; Ma, A. Recent Developments in Methods for Identifying Reaction Coordinates. *Mol. Simul.* 2014, *40* (10–11), 784–793. <https://doi.org/10.1080/08927022.2014.907898>.
- (7) Tavadze, P.; Avendaño Franco, G.; Ren, P.; Wen, X.; Li, Y.; Lewis, J. P. A Machine-Driven Hunt for Global Reaction Coordinates of Azobenzene Photoisomerization. *J. Am. Chem. Soc.* 2018, *140* (1), 285–290. <https://doi.org/10.1021/jacs.7b10030>.
- (8) Paul, S.; Paul, T. K.; Taraphder, S. Orthogonal Order Parameters to Model the Reaction Coordinate of an Enzyme Catalyzed Reaction. *J. Mol. Graph. Model.* 2019, *90*, 18–32. <https://doi.org/10.1016/j.jmgm.2019.03.023>.
- (9) Pratt, L. R. A Statistical Method for Identifying Transition States in High Dimensional

- Problems Problems. 2000, 5045 (June 1986).
- (10) Quaytman, S. L.; Schwartz, S. D. Reaction Coordinate of an Enzymatic Reaction Revealed by Transition Path Sampling. *Proc. Natl. Acad. Sci. U. S. A.* 2007, 104 (30), 12253–12258. <https://doi.org/10.1073/pnas.0704304104>.
  - (11) A Gentle Introduction to Graph Neural Networks, Kung Hsiang, Steeve Huang