Nanosecond Photodynamics Simulations of a *cis-trans* Isomerization are Enabled by Machine Learning

Jingbai Li[1], Patrick Reiser[2], André Eberhard[3], Pascal Friederich[2,3]*, and Steven A. Lopez[1]*
[1] Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA 02115, USA
[2] Institute of Nanotechnology, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3] Institute of Theoretical Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany
* Correspondence to pascal.friederich@kit.edu, s.lopez@northeastern.edu

Abstract
Photochemical reactions are being increasingly used to construct complex molecular architectures with mild and straightforward reaction conditions. Computational techniques are increasingly important to understand the reactivities and chemoselectivities of photochemical isomerization reactions because they offer molecular bonding information along the excited-state(s) of photodynamics. These photodynamics simulations are resource-intensive and are typically limited to 1–10 picoseconds and 1,000 trajectories due to high computational cost. Most organic photochemical reactions have excited-state lifetimes exceeding 1 picosecond, which places them outside possible computational studies. Westermeyr *et al.* demonstrated that a machine learning approach could significantly lengthen photodynamics simulation times for a model system, methylenimmonium cation ($CH_2NH_2^+$).
We have developed a Python-based code, Python Rapid Artificial Intelligence *Ab Initio* Molecular Dynamics (PyRAI^{2}MD), to accomplish the unprecedented 10 ns *cis-trans* photodynamics of *trans*-hexafluoro-2-butene (CF_3–CH=CH–CF_3) in 3.5 days. The same simulation would take approximately 58 years with ground-truth multiconfigurational dynamics. We proposed an innovative scheme combining Wigner sampling, geometrical interpolations, and short-time quantum chemical trajectories to effectively sample the initial data, facilitating the adaptive sampling to generate an informative and data-efficient training set with 6,232 data points. Our neural networks achieved chemical accuracy (mean absolute error of 0.032 eV). Our 4,814 trajectories reproduced the S_1 half-life (60.5 fs), the photochemical product ratio (*trans*: *cis* = 2.3: 1), and autonomously discovered a pathway towards a carbene. The neural networks have also shown the capability of generalizing the full potential energy surface with chemically incomplete data (*trans* → *cis* but not *cis* → *trans* pathways) that may offer future automated photochemical reaction discoveries.

1. Introduction
Academic and industrial chemists have turned to photochemistry as a sustainable technique to construct highly strained molecular architectures[1], photoswitches[2], and organic photovoltaics[3], solar fuel materials[4-5], which are characterized by mild conditions and high atom economy. Photochemical reactions typically consist of a series of molecular transformations that occur in excited molecules after light absorption. Unlike ground state processes–where a chemist can use spectroscopy and crystal structures to glean structural information, excited-state light-driven reactions involve short-lived femto- to picosecond (10^{-15} to 10^{-12} s) molecular excited states and reactive intermediates. This ultrafast process usually involves relaxation to the excited-state

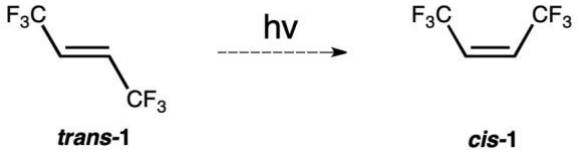
minimum for fluorescence or radiationless transition to ground-state through a state crossing point or seam, which determines the chemoselectivity of a photochemical reaction.[6-7]

The origin of chemoselectivity and stereoselectivity in organic photochemical reactions is challenging because of the short-lived molecular excited states. Quantum chemical calculations offer insight into the bonding changes that occur along a reaction coordinate and non-adiabatic molecular dynamics (NAMD) simulations to gain mechanistic insights and develop structure-reactivity relationships in complex photochemical reactions. The nuclei-electron coupling and time-dependent terms in NAMD increase the complexity of Hamiltonian and thus computation time. Multiple methods developed in the last two decades simplified the time-dependent molecular wave functions, *e.g.*, *ab initio* multiple spawning (AIMS)[8-9] and fewest switches surface hopping (FSSH).[10-12] However, computing high dimensional PESs with the requisite multiconfigurational methods is hugely resource-intensive. For example, most quantum chemical software packages encode an upper limit to active space of 16 electrons and 16 orbitals for complete active space self-consistent field (CASSCF) calculations. A typical NAMD experiment requires thousands of such calculations, resulting in a maximum simulation time on the order of 1 picosecond, at a computational cost of approximately 10^1–10^4 wall-clock hours. The realities of the low quantum yield of organic photochemical reactions and long excited-state lifetimes result in prohibitively expensive NAMD simulations. These compounding problems severely limit the application of NAMD simulations to only the most efficient organic photochemical reactions.

An increasing number of studies reported that fitting machine learning (ML) potentials could substantially accelerate the NAMD simulation. Aleotti *et al.* have parameterized *ad hoc* force fields for a 10 ps dynamic simulation of azobenzene.[13] Westermeyr *et al.* have trained multilayer feedforward neural networks (NNs) to enable 1 nanosecond simulations of methylenimmonium cation ($CH_2NH_2^+$) in 59 days.[14] Recently they have extended the application of the NNs model in predicting excited-states electronic properties for other small molecules, such as SO_2 and CSH_2 employing a deep continuous-filter convolutional-layer neural network, SchNet, combined with Sharc.[15] Beyond the proof-of-concept models, applying ML-based NAMD on more complex molecules is challenging because of the exponential increase of conformational space arising from the degrees of freedom of flexible bonds and functional groups. Training ML models becomes increasingly difficult with the rapidly growing training dataset, including atom-wise molecular features.

We will address the need for our combined ML and NAMD approach (ML-NAMD) that expands the scope of organic photochemical reactions. We will demonstrate the high accuracy in the trained NNs and low computational cost in the ML-NAMD simulations. Our trajectory statistics will highlight the ML-NAMD predicted photochemical product distribution in an excellent agreement with quantum chemical results. The ML-NAMD has shown the capability of extrapolating the full potential energy surface even with restricted data. Our initial ML-NAMD simulations focused on the first such photodynamics of hexafluoro-2-butene (**1** in Scheme 1), which is mechanistically straightforward because it has a typical $\pi\pi^*$ vertical excitation from HOMO to LUMO (the π and π^*orbital). Hexafluoro-2-butene is a nontoxic and not flammable industrial working fluid used as a refrigerant and as a foam-blowing agent.[16]

Scheme 1. The photochemical *cis-trans* isomerization of hexafluoro-2-butene (**1**)



2. Theoretical background

2.1 Quantum chemical method

We used OpenMolcas 19.11[17] to prepare CASSCF calculations of ***trans*-1** and ***cis*-1** with an active space of 2 π-electrons and 2 orbitals from the bonding and antibonding π-orbital of the C=C bond denoted CAS(2,2). We used the cc-pVDZ basis set for all atoms and averaged the ground-state and the first excited-state in equal weight. We optimized ***trans*-1** and ***cis*-1** to local minima, in which a vibrational analysis confirmed only positive frequencies. We located two minimum energy crossing points (MECPs) corresponding to the *trans*-to-*cis* (**MECP-*trans*-1**) and the reverse reaction (**MECP-*cis*-1**). The converged orbitals and optimized geometries are available in the Supporting Information.

2.2 Reference data

Our NAMD simulations at CASSCF(2,2)/cc-pVDZ provided references for state populations evolution and photochemical production distribution. We used the fewest switches surface hopping algorithm (FSSH) implemented in OpenMolcas 19.11.[17] We sampled 750 initial non-equilibrium geometries and associated velocities near the global minimum of ***trans*-1** with Wigner distribution at 300K. The NAMD trajectories propagated at 300K (Nosé-Hoover thermostat[18]) from the S_1 Franck-Condon region for ~500 fs and 0.5 fs timesteps (2,000 a.u. with 20 a.u. timesteps). 703 of 750 reference trajectories arrived in the ground-state within 500 fs, whereas the others stayed in the excited state. We characterized the 703 trajectories and removed 19 trajectories that involved structures with unphysically long bonds due to incorrectly converged CASSCF wavefunctions and averaged the state population of 684 trajectories.

2.3 Initial training set generation

Accurate NNs predictions require robust training data; thus, the training data must include a broad sampling of possible non-equilibrium geometries. Westermeyr *et al.* generated their training data set by leveraging a vibrational mode scan to generate a set of non-equilibrium structures for $CH_2NH_2^+$.[14] This report describes the isomerization photodynamics of a significantly more complex structure, **1**. The additional complexity demanded that we innovate training set generation because of the competing reality of the overwhelming computational cost of CASSCF calculations on an exponentially increasing number of structures in the training set. Figure 1 illustrates the workflow to automatically generate compact yet relevant training data for ML-accelerated excited-state dynamics.

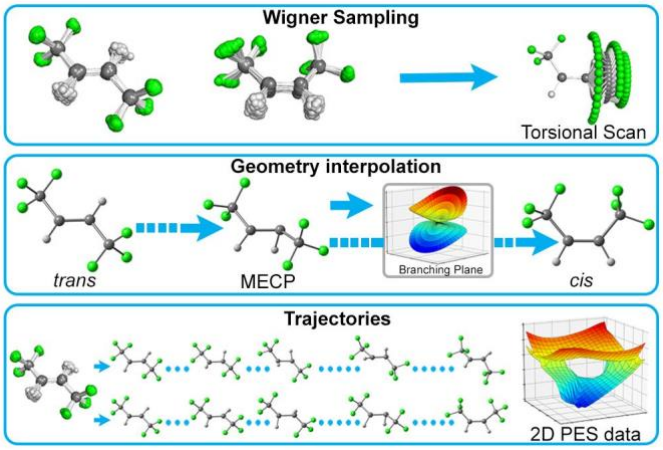


Figure 1. Three initial training set generation approaches in PyRAI^{2}MD. *Wigner Sampling* explores the conformational space near the equilibrium geometries. A subsequent torsional scan connects the sampled reactants and products. *Geometry Interpolation* collects the data along with reaction coordinates from reactant to product. *Trajectories* samples the data from quantum chemical trajectories. The 2D PES data can also supplement the data sampling. Detailed information about generation of the initial training set is available in the Supporting Information.

The initial training set generation scheme combines *Wigner Sampling*, *Geometrical Interpolation*, and *Trajectories*. The *Wigner Sampling* approach generates training data by sampling reactant and product structures and including a torsional scan component to simulate the *cis*-*trans* isomerization. The *Geometry Interpolation* approach accesses the reaction coordinate diagram by the optimized reactant, product, and crossing point geometries. It systematically varies the reactant geometry to that of the minimum energy crossing point (MECP) and then to that of the product in equal increments of Z-matrix coordinate parameters. Further interpolation at the crossing point covers the branching plane where the non-adiabatic couplings (NACs) have large variations. We expanded the range of sampled structures in the training data by applying the Wigner sampled geometrical perturbations to the interpolated reaction coordinate diagram. The *Trajectories* approach involves generating 10–200 Wigner sampled initial conditions and propagating the continued NAMD trajectories in 50–100 fs with quantum chemical calculation. We screened the trajectories by examining the last geometries and selected 132 trajectories corresponding to *trans*-to-*cis* isomerization. The trajectories toward other reaction pathways were filtered out. We included every 10th snapshot of the first 50 fs of 132 reference trajectories. The 2D PES data (Figure 1) were used as a test set for the ML model, which will be discussed in section **4.3**. The total number of data points in our initial training set is 4,961.

2.4 Adaptive sampling

Adaptive sampling[19] is one of the most efficient methods to expand the initial training set. Based on uncertainty quantification, the method samples structures in the unexplored regions of the potential energy surfaces. The adaptive sampling propagated 129 NAMD trajectories in 500 fs at 300K with the same initial conditions of the quantum chemical reference trajectories. In each trajectory, two sets of NNs predict the energies, force, and NACs in each time step, and the standard deviation between the predictions quantifies the prediction uncertainty. The most

uncertain points found in the conformational space were selected for quantum chemical calculations with CASSCF(2,2)/cc-pVDZ. The adaptive sampling stopped after 32 iterations of sampling, recalculating and retraining. 81% (104) trajectories have converged, and we did not observe further improvement in the subsequent 50 iterations. Our final training data set has 6,232 data points (*i.e.*, 1,271 points were added through adaptive sampling). Detailed information about the adaptive sampling is available in the supporting information.

2.5 Forces and non-adiabatic couplings

Fully quantum chemical NAMD trajectory propagation requires millions of CASSCF calculations to determine the energies, forces, and NACs at each timestep. Westermeyr *et al.* [15] describes the particular challenge of predicting forces and NACs because they have independent vector components. The direction of forces and NACs are rotationally covariant (*i.e*., they depend on the molecule's orientation). Figure 2 shows a rotation of geometry associated with the covariantly revolved force vectors.

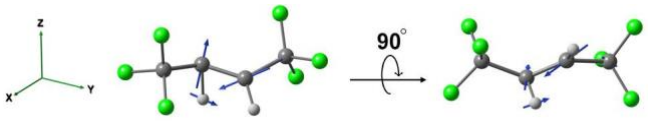


Figure 2. 3D representation of the rotational covariance of forces and NACs. The blue arrows represent the same force vectors. The geometry is rotated by 90° that leads to different x,y, and z components of force vectors in each orientation.

However, typical ML representations (*e.g.*, ML predictions are scalar) are rotationally invariant and have no information about the global orientation of the molecule. Thus, we implemented efficient and differentiable NNs that predict forces according to the first derivatives of the energy with respect to nuclear coordinates. We used the first-order derivatives of a non-physical, anti-derivative of NACs, to predict NACs. The FSSH algorithm computes the non-adiabatic transition probability using NACs. NACs depend on two electronic state wavefunctions (*i.e.*, state *i* and *j* in Eq. 1), where the phases of CASSCF wavefunctions do not necessarily cancel out. OpenMolcas computes wavefunctions with an arbitrary phase (*i.e.*, sign). The phase change can give the opposite NACs values per Eq. 2. The sign of NACs frequently and randomly alternates along the isomerization reaction coordinate (Figure 3b), which is undesired when training ML models.

$$E_i = \frac{\left\langle \Psi_i \left| \frac{\partial H}{\partial R} \right| \Psi_j \right\rangle}{E_i - E_j} \qquad \text{(Eq 1)}$$

$$\left\langle -\Psi_i \left| \frac{\partial H}{\partial R} \right| \Psi_i \right\rangle = -\left\langle \Psi_i \left| \frac{\partial H}{\partial R} \right| \Psi_i \right\rangle \qquad \text{(Eq 2)}$$

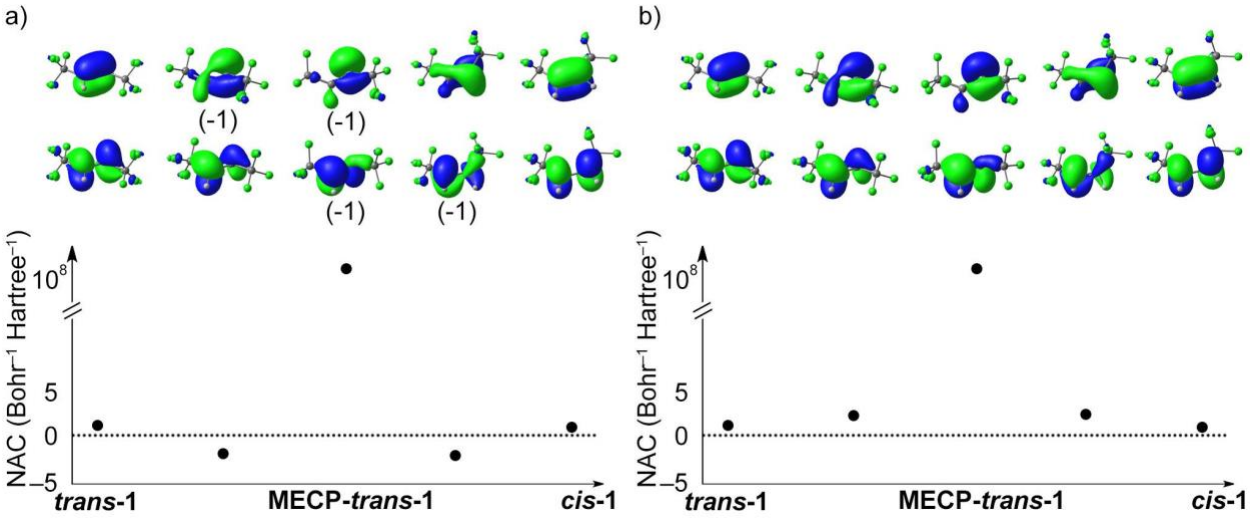


Figure 3. NAC representation using π- and π*-orbitals as a function of interpolated geometries from ***trans*-1** to **MECP-*trans*-1** to ***cis*-1**. (a) The phases are randomly assigned, resulting in an unsmooth NACs function. The –1 indicates the undesired phase changes if we do not apply corrections. (b) The phases of electronic state wavefunctions are corrected, leading to a smooth NACs function. The NACs in the equilibrium geometry of ***trans*-1** are the reference. Detailed information about the phase correction scheme is available in the Supporting Information.

Figure 3a demonstrates the alternating signs of the NACs. We employed a phase correction scheme[14, 20-21] based on geometrical interpolation to correct the sign by tracking and correcting the phase of the interacting states of each interpolated geometry (Figure 3b). The phase-corrected NACs become the reference for subsequent trajectory points. The discontinuous nature of NACs near the surface crossings make NNs learning extremely difficult; instead of training with the full expression of NACs, we trained NNs with only the numerator, known as configuration interaction (CI) derivatives. The ML-NAMD evaluates NACs on-the-fly along a given trajectory based on the state energy differences and predicted CI derivative terms.

2.6 Machine learning model

We have implemented multilayer feedforward NNs as the primary ML model in our ML-NAMD approach. We used an inverse distance-based[14-15] feature representation to predict energies, forces, and NACs. **1** has 12 atoms ($N = 12$), which lead to 66 unique entries in the inverse distance matrix. The ML model predicts two energy values (ground- and excited-state; $k = 2$) and uses their derivatives with respect to the input coordinates to compute the $k*3N = 72$ forces component. The $3N = 36$ NAC components for each data point (conformer) were predicted using derivatives of N atom-specific, physically meaningless scalar potentials with respect to their input coordinates. We chose a leaky soft plus activation function for hidden layers. We have optimized the hyperparameters to predict energies, forces, and NACs with a grid-search over 864 NNs. The training was done using the Adam optimizer[22] with a stepwise decrease of the initial learning rate from 10^{-3} to 10^{-5} on validation error plateaus (energies and forces: 2,700 epochs; NACs: 1,700 epochs). To evaluate the prediction uncertainty in NAMD simulation, we used the standard deviation between the NNs. We picked two sets of the most accurate and efficient NNs (each set

has one NN for energies and forces together, and one for NACs) with distinct architectures (Table 1). We implemented all NNs using the TensorFlow/Keras (v2.3) packages[23] for Python.

Table 1. Selected hyperparameters of NNs trained on 6,232 data points of **1**. The supporting information provides additional details on hyperparameters and training statistics.

	Energies, forces		NACs	
Hyperparameters	NN1	NN2	NN1	NN2
Activation function	Leaky soft plus		Leaky soft plus	
Hidden layers	4	5	4	5
Neurons/layer	400	300	300	300
Batch size	64	64	64	64

3. PyRAI^{2}MD

We have developed a Python Rapid Artificial Intelligence *Ab Initio* Molecular Dynamics program (PyRAI^{2}MD) to integrate our ML models and the NAMD algorithm. It is an open-source code to enable ML-NAMD simulations for unprecedented molecular complexity and timescales. Figure 4 illustrates the computational architecture of PyRAI^{2}MD, which requires ML and NAMD kernels.

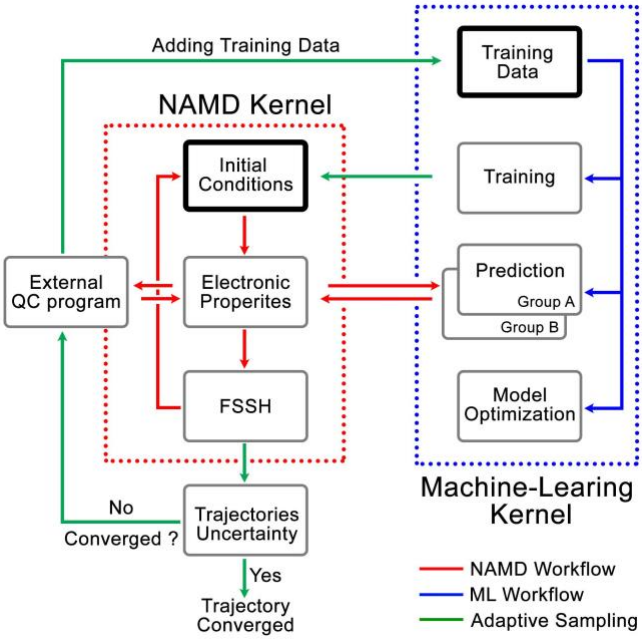


Figure 4. The computational architecture of PyRAI^{2}MD. The bold boxes are the initial stage in NAMD and ML kernel.

The workflow in the NAMD kernel starts with the Wigner sampled initial conditions. The following procedures, red arrows in Figure 4, iteratively compute the energies, forces, NACs, and surface hopping probability, propagating the NAMD trajectory. We generalized the input and output format

of computing energies, forces, and NACs in the NAMD kernel. This feature enables efficient communication with our chosen external quantum chemical program, OpenMolcas 19.11,[17] and the internal ML kernel. The workflow in ML kernel reads a training set or an existing model. It provides a convenient interface to train models and make predictions marked by blue arrows in Figure 4. The green arrows in Figure 4 illustrate the adaptive sampling workflow incorporating the NAMD and ML kernels. It first loads two trained ML models to run multiple trajectories in the NAMD kernel. The trajectories explore the unsampled conformational space while the two models evaluate the prediction uncertainty, as described in section **2.3** and the Supporting Information. The sampled data were added into the initial set to train a new model.

4. Results and Discussion

4.1 Performance of ML potential

We have trained four NNs using the training set of 6,232 data points with PyRAI^{2}MD and summarized the results in Table 2. The predictions in the energies, forces, and NACs of ***trans*-1** take 0.00945 seconds on a single CPU, whereas the equivalent computations at CASSCF(2,2)/cc-pVDZ require 336 seconds on a single CPU. This represents an NNs acceleration of 37,307-fold relative to ground-truth CASSCF calculations. We have evaluated the mean absolute error (MAE) and coefficient of determination (R^2) in predicted energies, forces, and NACs to measure the accuracy of NNs (Table 2).

Table 2. The mean absolute errors and R^2 in NNs predicted energies, forces, and NACs, trained with 6,232 data points.

	Energies, eV	Forces, eV Å^{-1}	NACs, Å^{-1}
MAE_{NN1}	0.032	0.18	0.009
MAE_{NN2}	0.032	0.17	0.008
R^2_{NN1}	0.9999	0.9249	0.5976
R^2_{NN2}	0.9999	0.9415	0.6092

The MAE of predicted energies is 0.032 eV in both NNs, which meets the 'chemical accuracy' threshold (1 kcal mol^{-1} = 0.043 eV). The R^2 values are 0.9999 in NN1 and NN2, suggesting an almost linear correlation between the NNs prediction and QC reference. The MAE and R^2 in force calculation are 0.17–0.18 eV Å^{-1} and 0.9249–0.9415, respectively. These values are consistent with previous reports on SO_2 and $CH_2NH_2^+$ photochemistry.[14] The MAE and R^2 in NACs are 0.008–0.009 Å^{-1} and 0.5976–0.6092, respectively. The low R^2 relative to energies and forces suggests possible overfitting of NACs (the training MAE of NN1 and NN2 are both 0.0004 Å^{-1} and the training R^2 are 0.9664 and 0.9689 for NN1 and NN2, respectively) due to the indefinite term of the antiderivative of NACs and unavailable training data for it. We validated the NNs on the reaction coordinate diagram from ***trans*-1** to ***cis*-1**. Figure 5a and 5b show the predicted S_0, S_1 energies and the S_0–S_1 energy differences along with the reaction coordinate. The red and grey shades correspond to the prediction errors, where the error regions multiplied by 50 for clarity.

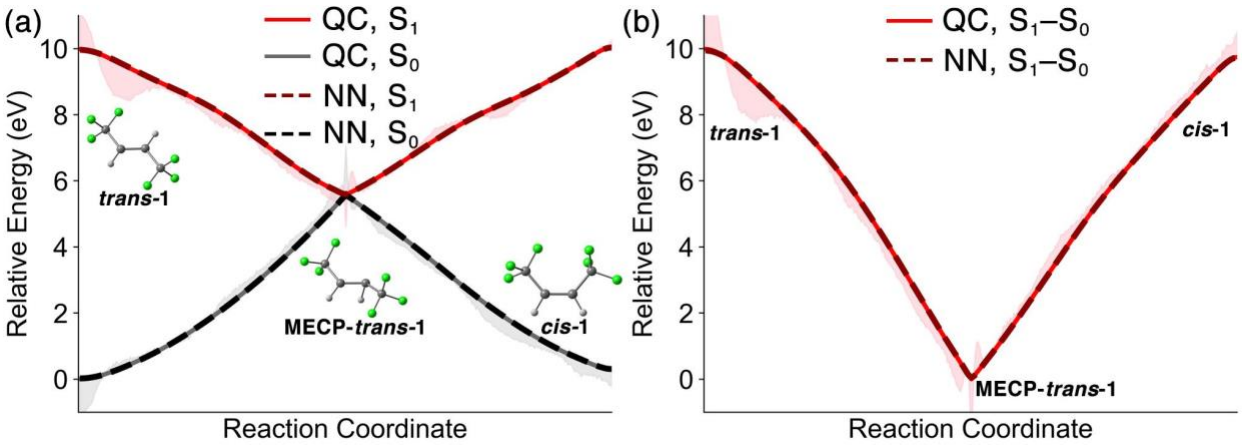


Figure 5. Quantum chemical *vs.* NNs predicted reaction coordinate diagram**.** The reaction coordinates were interpolated from optimized ***trans*-1** to ***cis*-1** via **MECP-*trans*-1** with CASSCF(2,2)/cc-pVDZ. (a) The S_0 (grey) and S_1 (red) energies were computed with CASSCF(2,2)/cc-pVDZ. The black and dark red dashed lines plot the NNs predicted S_0 and S_1 energies, respectively. The grey and red shades are upscaled 50-fold to highlight the small errors (E_{NNs} – E_{QC}) in S_0 and S_1 energies. The overestimations are below the curve, and the underestimations are above the curve. (b) The red line and dark red dashed lines plot the quantum chemical, and NNs predicted S_0–S_1 energy differences.

The overall MAE between NNs and QC energies are $5.7 \cdot 10^{-3}$ eV and $5.3 \cdot 10^{-3}$ eV for the S_0 and S_1 states. The maximum absolute errors of predicted S_0 and S_1 energies are $3.0 \cdot 10^{-3}$ eV and $2.2 \cdot 10^{-3}$ eV at **MECP-*trans*-1** and ***trans*-1**, respectively. Figure 5b shows that the largest absolute error of the S_1–S_0 energy difference is 0.043 eV near the ***trans*-1** local minimum. The NNs overestimate the S_0–S_1 energy difference by 0.050 eV at **MECP-*trans*-1**, which is still sufficiently small to promote the surface hopping.

4.2 10 nanoseconds ML-NAMD photodynamics simulation

For 500 fs simulations with 0.5 fs time step, our NNs predicted 4,814 trajectories using Wigner sampled initial conditions at 300K with an average cost of 28 seconds, while the CASSCF(2,2)/cc-pVDZ calculations required 40 hours. We compared the state population of the ML and QC trajectories in Figure 6.

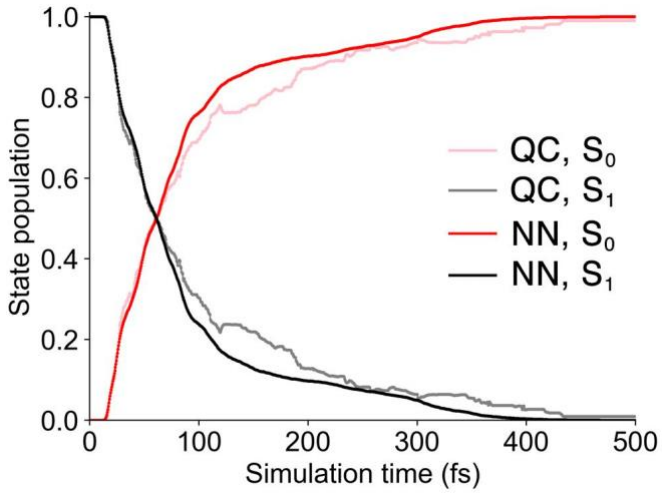


Figure 6. Quantum chemical *vs.* NNs predicted trajectory population in 500 fs NAMD simulations. The quantum chemical trajectory populations at CASSCF(2,2)/cc-pVDZ are the average over 684 trajectories. The NNs trajectory populations are the average over 4,814 trajectories predicted by NNs.

The NNs and QC calculations predict virtually identical S_1 half-life of 60.5 and 60.0 fs, respectively. The MAE between the QC and NNs predicted populations is 3.0 %. The maximum deviation (9.4 %) appeared in 80–200 fs; there is a shoulder in this range, suggesting a statistically inadequate number of QC trajectories. In 0–80 fs and 200–500 fs, the MAE is 1.3 % to 2.0 %. These suggest high accuracy of our NNs reproducing the surface hopping events. We then turned to NNs predicted photochemical product distribution starting from the Franck-Condon region of ***trans*-1.** Figure 7 shows a comparison of 430 QC and 3,470 ML-NAMD trajectories and their energetic dependence on the ∠H-C-C-H and ∠C-C-C-C dihedral angles.

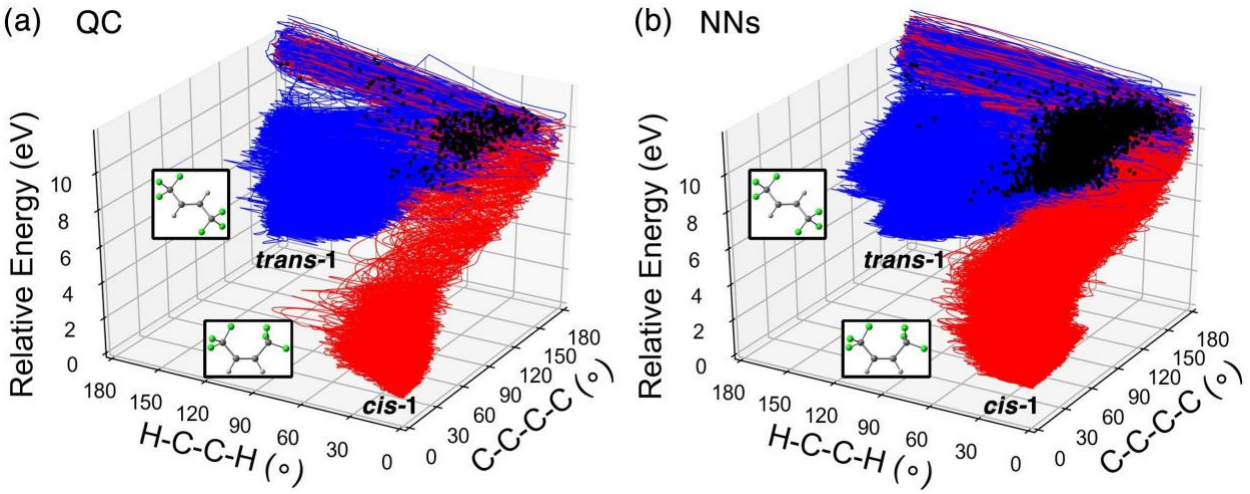


Figure 7. (a) The selected 430 reference trajectories with CASSCF(2,2)/cc-pVDZ in 500 fs simulations at 300K. (b) The 3,470 NNs predicted trajectories using Wigner sampled initial conditions at 300K in 500 fs simulations. The black dots represent the last surface hopping point in each trajectory.

Figure 7a illustrates the trace of the selected reference trajectories that bifurcated toward ***cis*-1** and reverted to ***trans*-1**. All trajectories started with a narrow spreading in the range of ∠C-C-C-C = 150–180° and ∠H-C-C-H = 0–60° (upper right corner). 135 surface hopping events turned the trajectories along the direction of the ∠H-C-C-H axis toward ***cis*-1**. 356 trajectories took the same path returning to ***trans*-1**. The *trans*: *cis* ratio is 2.6: 1 with CASSCF(2,2)/cc-pVDZ. The NNs predicted trajectories in Figure 7b recreated the topology of the reference. The NNs predicted 1,056 trajectories transformed to ***cis*-1** and 2,414 trajectories back to ***trans*-1**, resulting in a corresponding ratio of 2.3: 1. In the reference trajectories, we observed a few discontinuities in the total energy. It is because the CASSCF calculation on discrete geometries cannot guarantee correctly converged active space in NAMD simulations. Our NNs are analytic functions of energy that avoid this problem, thereby producing smooth trajectory traces.

We analyzed the crossing region and hopping points as identified by the QC and ML-NAMD trajectories. Figure 8 shows the spatial distribution of the last S_1/S_0 surface hopping point in each trajectory as well as examples of surface hopping geometries that define the S_1/S_0 crossing seam.

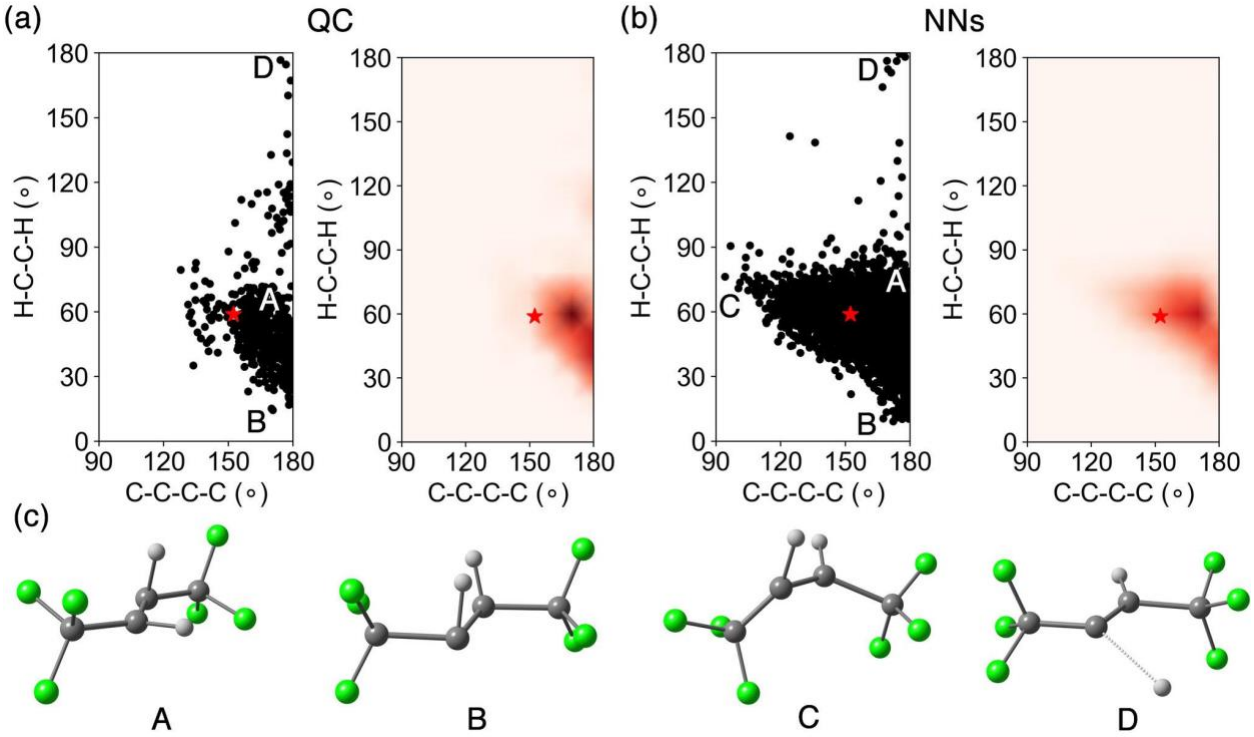


Figure 8. Spatial distribution of the latest S_1/S_0 surface hopping points in (a) CASSCF(2,2)/cc-pVDZ and (b) NNs predicted trajectories in 500 fs simulation. The left panel illustrates the position of the surface hopping points and the right panel shows the probability density of the surface hopping points. The probability density magnitude is defined as the number of surface hopping points in an interval of 10° divided by the total number of surface hopping points. The magnitude ranges from 0.00 to 0.12 and is colored from light pink to dark red. The red star marks **MECP-*trans*-1**. (c) Four surface hopping points that define the S_1/S_0 crossing seam.

Figure 8a projects the latest S_1/S_0 surface hopping points from the CASSCF(2,2)/cc-pVDZ trajectories into the 2D conformational space of **1**. The dense areas represent the S_1/S_0 crossing region as defined by the two reaction coordinates. In the right panel of Figure 8a, the surface hopping points accumulated at ∠C-C-C-C = 170° and ∠H-C-C-H = 60°, which has a larger ∠C-C-C-C angel than **MECP-*trans*-1**. Beside the hopping points near the **MECP-*trans*-1**, Figure 8c shows two pyramidalized surface hopping geometries at **A** (∠C-C-C-C = 180° and ∠H-C-C-H = 80°) and **B** (∠C-C-C-C = 180° and ∠H-C-C-H = 20°). We noted C-H stretching surface hoppings at **D** (∠C-C-C-C = 180° and ∠H-C-C-H = 180°); the C–H distance is 2.24Å. Figure 8b depicts the distribution of the S_1/S_0 surface hopping points by NNs prediction. The NNs not only replicated the probability density region of surface hopping points but also and provided the increased statistical significance of our crossing region analysis. The NNs located a twisting surface hopping geometry at **C** (∠C-C-C-C = 90° and ∠H-C-C-H = 80°). The probability density map (Figure 8b, right panel) suggests it is a relatively rare surface hopping event as few points were located at **C**. The NNs were able to detect the rare event because of the substantially increased number of trajectories (4,814) with minimal computational cost.

Based on the performance and state-of-the-art trajectories statistics in 500 fs ML-NAMD simulations, we decided to test the limits of our newly established ML-NAMD model. We ran 18 trajectories over an unprecedented 10 ns simulation time with a time step of 0.5 fs. We measured the average populations of the 10 ns ML-NAMD trajectories and showed the results in Figure 9.

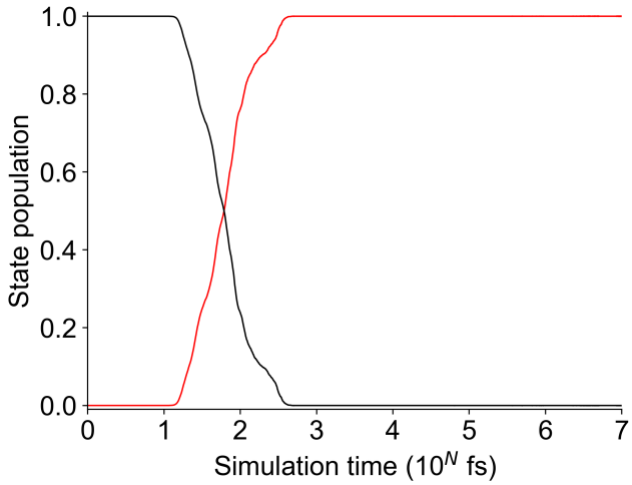


Figure 9. The state population in 10 ns ML-NAMD trajectories. The population analysis averaged 4,832 trajectories in the first 500 fs and 18 trajectories up to 10 ns. The x-axis converts simulation time into a logarithm of 10.

In the first 500 fs, we include the earlier cited 4,814 trajectories for the average population. From 500 fs to 10 ns, we averaged the state populations of the 18 trajectories. The S_1/S_0 transition completes within 1 ps. The flat S_0 population curve after 10^3 fs indicates all trajectories stayed in S_0 up to 10 ns, without surface hopping up to the S_1 state. The complete separation of S_1 and S_0 populations suggests that the NNs predicted reliable non-adiabatic coupling in full 10 ns simulations. We anticipate that these PyRAI^{2}MD-enabled simulations will pave the way to statistical analyses of currently inaccessible photodynamics simulations. The total $2 \cdot 10^7$ iterations

were accomplished in an average of 3 days 12 hours on a single CPU. A 10 ns simulation with the QC method would otherwise require approximately 58 years of simulation time. We have included the ML-NAMD trajectory movie in the supporting information.

4.3 Photochemical reaction discovery with ML-NAMD

In this section, we demonstrate an example of using ML-NAMD to discover unexpected or unknown photochemical reactions with the dramatically increased number of possible trajectories enabled by the ML acceleration. Our CASSCF(2,2)/cc-pVDZ trajectories have revealed that ***trans*-1** can undergo an intramolecular hydrogen abstraction to form hexafluoro-2-butene carbene, **2** in addition to the major *cis-trans* isomerization reaction. This pathway is consistent with prior theoretical studies on ethylene intramolecular hydrogen migration reactions.[24-26] Figure 10a shows this reaction pathway, Figure 10b and 10c show 208 QC and 1,344 ML-NAMD trajectories toward **2**. A close look at the structural features of this pathway is shown in Figure 10d with 5 snapshots from ***trans*-1-FC** to **2**, including the surface hopping point.

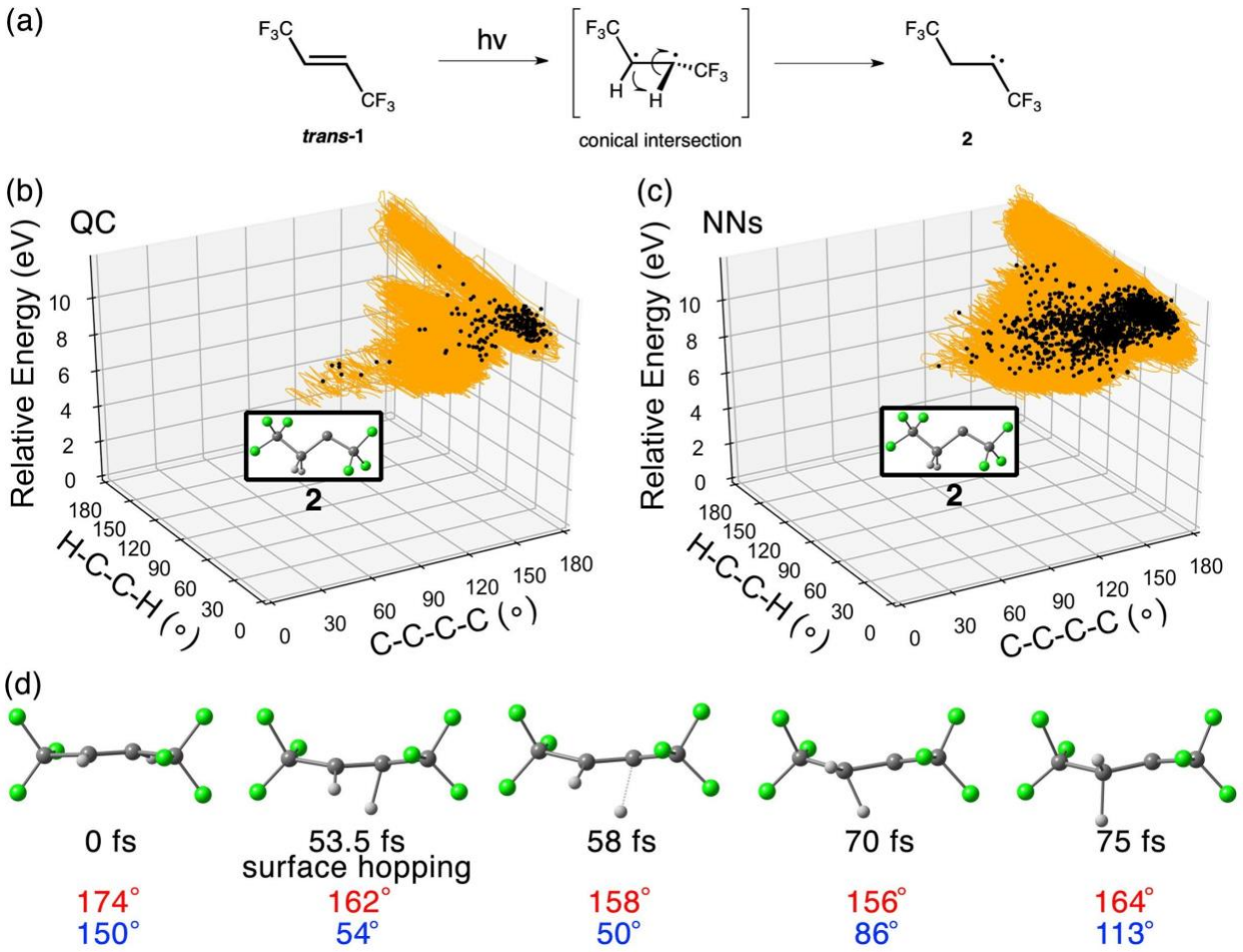


Figure 10. (a) Formation of hexafluoro-2-butene carbene (**2**) via a possible conical intersection. (b) The selected 208 reference trajectories with CASSCF(2,2)/cc-pVDZ in 500 fs NAMD simulation. (c) The 1,344 NNs predicted trajectories using Wigner sampled initial conditions at 300K in 500 fs NAMD simulation. The black dots represent the last surface hopping points in each

trajectory. (d) Snapshots of hexafluoro-2-butene carbene formation in CASSCF(2,2)/cc-pVDZ trajectories. The ∠H-C-C-H and ∠C-C-C-C angles are in blue and red, respectively.

210 CASSCF(2,2)/cc-pVDZ trajectories formed **2** in a *trans*: carbene ratio of 1.7: 1. The trajectories in Figure 10b display that the transformation shared the path with *trans*-to-*cis* isomerization from the Franck-Condon region to the S_1/S_0 crossing region. The S_1/S_0 surface hoppings happened in the area of ∠C-C-C-C = 150–180° and ∠H-C-C-H = 0–60°. In figure 10d, the hydrogen migrated to the other carbon atom from 58 to 70 fs. The ∠H-C-C-H angle, 50° at 58 fs, increased to 86° at 70 fs. The C=C double bond transformed to a C-C single bond**,** resulting in a nonplanar geometry (∠C-C-C-C = 164° and ∠H-C-C-H = 113° at 75 fs). To test whether the adaptive learning method can "rediscover" the alternative carbene pathway without prior knowledge or human bias, we trained NNs with an initial training set which only included the *trans*-to-*cis* isomerization of **1** as described in section **2.3**. The adaptive sampling strategy explored the conformational space and found 105 geometries (1.6% in the final training set) that correspond to the hydrogen migration reaction that was intentionally left out in the initial training set. Figure 10b shows the 1,344 trajectories toward carbene by NN. The predicted *trans*: carbene ratio of 1.8: 1 almost exactly matched the CASSCF(2,2)/cc-pVDZ results. The exceptional agreement indicates that the NNs have fully revealed the reactivity and chemoselectivity of ***trans*-1** after efficiently sampling the conformational space.

Finally, we compared the topologies of the S_0 and S_1 PESs predicted by the newly trained NNs with QC calculations to determine the predictive efficacy of the NNs on important surfaces. The scan constrains the two dihedral angles of **1**, ∠C-C-C-C, and ∠H-C-C-H (Scheme 1), ranging from 0 to 180° with a step size of 10°. The 2D PES presented separate areas of *trans*-to-*cis* and *cis*-to-*trans* paths; the initial training set of the NNs had no information about the *cis*-to-*trans* reaction except for the sampled data enclosed the ***trans*-1** and ***cis*-1.**

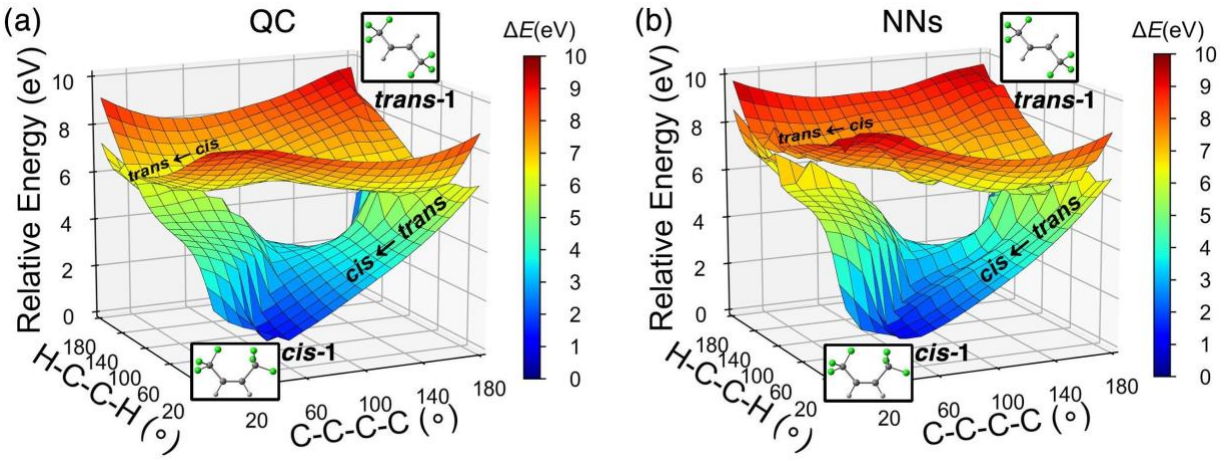


Figure 11. (a) 2D potential energy surface of **1** in S_1 and vertical S_0 surface with CASSCF(2,2)/cc-pVDZ. (b) NNs predicted potential energy surfaces on the same geometries. The *trans*-to-*cis* path varies the ∠H-C-C-H and ∠C-C-C-C from 180° to 0° via the crossing region (∠H-C-C-H < ∠C-C-C-C near the right corner). The *cis*-to-*trans* path alters the ∠H-C-C-H and ∠C-C-C-C in the opposite direction, from 0° to 180° passing through a crossing region (∠H-C-C-H > ∠C-C-C-C

near the left corner). Detailed discussions on the 2D PES are available in the Supporting Information.

Figure 11a demonstrates the optimized S_1 and S_0 surface (vertical emission from S_1) of **1** with CASSCF(2,2)/cc-pVDZ. It shows two crossing regions corresponding to the *trans*-to-*cis* (∠H-C-C-H < ∠C-C-C-C) and *cis*-to-*trans* (∠H-C-C-H > ∠C-C-C-C) isomerization. The diagonal (∠H-C-C-H = ∠C-C-C-C) represents the avoided-crossing in the twisting model of **1**. The 2D PES has a 19 x 19 grid points, which the NNs have never seen in the training. Nonetheless, the NNs predicted S_0 and S_1 surfaces resemble the full 2D PES surprisingly well. The NNs overestimated the S_0 and S_1 surface with an MAE of 0.32 and 0.42 eV, respectively. The MAE of S_1–S_0 energy difference is 0.35 eV. These significant energetic discrepancies are because the data at the *cis*-to-*trans* S_1/S_0 crossing region (∠H-C-C-H = 90–180° and ∠C-C-C-C = 0–90°) were absent. At this region, the MAEs of S_0 and S_1 energies and their differences are 0.50, 0.82, and 0.43 eV, respectively. As a result, the NNs shifted the *cis*-to-*trans* crossing area above the *trans*-to-*cis* crossing area (Figure 11b), which decreases the barrier in between (Figure 11a). It suggests the NNs would bias the surface hopping structure in the *cis*-to-*trans* path toward that in the *trans*-to-*cis* path. This can be easily improved by including the 2D PES data in the training set.

5. Concluding Remarks

We have developed an ML-NAMD approach using neural networks that overcome present limits in machine learning photodynamics simulations, namely, the simulation time, the size of the molecule, and the complexity of the photoreactions. The ML-NAMD provides a *priori* prediction for long-lived excited molecules that would leverage the current scope of discovering new photoswitches and photosensitizers. As the first demonstration of the ML-NAMD approach, we have accomplished the unprecedented 10 ns simulations of the photodynamics of ***trans*-1** using our PyRAI^{2}MD code. We constructed highly efficient NNs that accelerated the computations of energies, forces, and NACs 37,307-fold compared to quantum chemical calculations with CASSCF(2,2)/cc-pVDZ. We developed a composite generation scheme to effectively sample the initial training set, which combines the Wigner sampling, geometrical interpolations, and short-time quantum chemical trajectories. We implemented an automatic workflow to expand the initial training set by adaptively sampling the unexplored conformational space using a query of committee model. We obtained a final training set with 6,232 data points, with which we trained a final set of NNs, with a prediction error in energy of 0.032 eV, achieving chemical accuracy. The forces and NACs errors were 0.17–0.18 eV $Å^{-1}$ and 0.008–0.009 $Å^{-1}$respectively. We validated our NNs on the reaction coordinate diagram of *trans*-to-*cis* isomerization of **1.** The predicted S_0 and S_1 energy and their difference are sufficient to describe surface hopping. The maximum absolute error of S_0, S_1 energies and their difference are 0.030, 0.022, and 0.050 eV.

The 10 ns ML-NAMD simulations used a single CPU and only took 3.5 days for each of the 18 trajectories, which would have approximately required 58 years for running quantum chemical calculations with CASSCF(2,2)/cc-pVDZ. We propagated 4,814 ML-NAMD trajectories with Wigner sampled initial conditions at 300K, starting from the S_1 Franck-Condon region in 500 fs simulations. The predicted S_1 half-life is 60.5 fs, and the *trans*: *cis* ratio is 2.3: 1. They agree with the results of 684 CASSCF(2,2)/cc-pVDZ trajectories, where the S_1 half-life is 60 fs, and the *trans:*

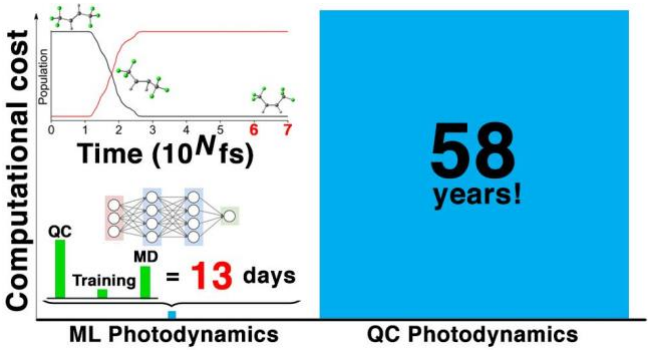
cis ratio is 2.6: 1. The NNs also predicted occasional surface hopping events near the avoided-crossing in the twisting model of **1**.

Our NNs have shown intelligence of self-navigating the learning process and speculating the unseen PES with restricted training data. It autonomously collected 105 data points relevant to hexafluoro-2-butene carbene during the adaptive sampling. The NNs reproduced a *trans*: carbene ratio of 1.8: 1 in an excellent agreement with the ratio of 1.7: 1 with CASSCF(2,2)/cc-pVDZ. Moreover, the NNs intentionally trained for *trans*-to-*cis* reaction resembled the full 2D PES of **1**. They replicated the topography of the S_1 and S_0 surfaces without learning the reverse *cis-to-trans* reaction. Currently, we are actively testing additional molecular features, such as bond angles and dihedral angles, and working on approximate NACs prediction and NACs free surface hopping algorithms.

Acknowledgment
J.L. and S.A.L. acknowledges Biruk Abreha, Fatemah Mukadum, and Marcus Schwarting for helpful discussion on NNs. J.L. and S.A.L. acknowledge the Office of Naval Research (ONR N00014-12-1-0828) for funding this research. J.L. and S.A.L. appreciate the assistance from the Northeastern Research Computing Team and access to the computing resources of the Discovery cluster. P.F. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 795206. P.R. acknowledges the funding by the Virtual Materials Design (VirtMat) initiative at Karlsruhe Institute of Technology.

Supporting Information
The Supporting Information includes details about the machine learning model, initial training set generation, NACs phase corrections, NNs training and adaptive sampling, computational cost of ML-NAMD, QC *vs.* NNs 2D PES, data availability, QC calculations of hexafluoro-2-butene, and Cartesian coordinates of optimized geometries.

TOC:



References

1. Karkas, M. D.; Porco, J. A., Jr.; Stephenson, C. R., Photochemical Approaches to Complex Chemotypes: Applications in Natural Product Synthesis. *Chem. Rev.* **2016,** *116* (17), 9683-9747.
2. Gonzalez, A.; Kengmana, E. S.; Fonseca, M. V.; Han, G. G. D., Solid-state photoswitching molecules: structural design for isomerization in condensed phase. *Materials Today Advances* **2020,** *6*.
3. Cox, J. M.; Mileson, B.; Sadagopan, A.; Lopez, S. A., Molecular Recognition and Band Alignment in 3D Covalent Organic Frameworks for Cocrystalline Organic Photovoltaics. *The Journal of Physical Chemistry C* **2020,** *124* (17), 9126-9133.
4. Calbo, J.; Weston, C. E.; White, A. J.; Rzepa, H. S.; Contreras-Garcia, J.; Fuchter, M. J., Tuning Azoheteroarene Photoswitch Performance through Heteroaryl Design. *J. Am. Chem. Soc.* **2017,** *139* (3), 1261-1274.
5. Saydjari, A. K.; Weis, P.; Wu, S., Spanning the Solar Spectrum: Azopolymer Solar Thermal Fuels for Simultaneous UV and Visible Light Storage. *Adv. Energy Mater.* **2017,** *7* (3).
6. Li, J.; Lopez, S. A., Multiconfigurational Calculations and Nonadiabatic Molecular Dynamics Explain Tricyclooctadiene Photochemical Chemoselectivity. *J. Phys. Chem. A* **2020,** *124* (38), 7623-7632.
7. Cox, J. M.; Lopez, S. A., Multiconfigurational dynamics explain photochemical reactivity and torquoselectivity towards fluorinated polyacetylenes. *Journal of Materials Chemistry C* **2020,** *8* (31), 10880-10888.
8. Ben-Nun, M.; Martínez, T. J., Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem. *J. Chem. Phys.* **1998,** *108* (17), 7244-7257.
9. Ben-Nun, M.; Quenneville, J.; Martínez, T. J., Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics. *The Journal of Physical Chemistry A* **2000,** *104* (22), 5161-5175.
10. Hammes-Schiffer, S.; Tully, J. C., Proton transfer in solution: Molecular dynamics with quantum transitions. *J. Chem. Phys.* **1994,** *101* (6), 4657-4667.
11. Tully, J. C., Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990,** *93* (2), 1061-1071.
12. Tully, J. C.; Preston, R. K., Trajectory Surface Hopping Approach to Nonadiabatic Molecular Collisions: The Reaction of H+ with D2. *J. Chem. Phys.* **1971,** *55* (2), 562-572.
13. Aleotti, F.; Soprani, L.; Nenov, A.; Berardi, R.; Arcioni, A.; Zannoni, C.; Garavelli, M., Multidimensional Potential Energy Surfaces Resolved at the RASPT2 Level for Accurate Photoinduced Isomerization Dynamics of Azobenzene. *J. Chem. Theory Comput.* **2019,** *15* (12), 6813-6823.
14. Westermayr, J.; Gastegger, M.; Menger, M.; Mai, S.; Gonzalez, L.; Marquetand, P., Machine learning enables long time scale molecular photodynamics simulations. *Chem. Sci.* **2019,** *10* (35), 8100-8107.
15. Westermayr, J.; Gastegger, M.; Marquetand, P., Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics. *J Phys Chem Lett* **2020,** *11* (10), 3828-3834.
16. Intergovernmental Panel on Climate Change. (2014). Anthropogenic and Natural Radiative Forcing. In *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 659-740. Cambridge University Press.

17. Fdez Galvan, I.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; Chibotaru, L. F.; Creutzberg, J.; Dattani, N.; Delcey, M. G.; Dong, S. S.; Dreuw, A.; Freitag, L.; Frutos, L. M.; Gagliardi, L.; Gendron, F.; Giussani, A.; Gonzalez, L.; Grell, G.; Guo, M.; Hoyer, C. E.; Johansson, M.; Keller, S.; Knecht, S.; Kovacevic, G.; Kallman, E.; Li Manni, G.; Lundberg, M.; Ma, Y.; Mai, S.; Malhado, J. P.; Malmqvist, P. A.; Marquetand, P.; Mewes, S. A.; Norell, J.; Olivucci, M.; Oppel, M.; Phung, Q. M.; Pierloot, K.; Plasser, F.; Reiher, M.; Sand, A. M.; Schapiro, I.; Sharma, P.; Stein, C. J.; Sorensen, L. K.; Truhlar, D. G.; Ugandi, M.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Weser, O.; Wesolowski, T. A.; Widmark, P. O.; Wouters, S.; Zech, A.; Zobel, J. P.; Lindh, R., OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019,** *15* (11), 5925-5964.
18. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L., Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996,** *87* (5), 1117-1157.
19. Behler, J., Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem* **2015,** *115* (16), 1032-1050.
20. Akimov, A. V., A Simple Phase Correction Makes a Big Difference in Nonadiabatic Molecular Dynamics. *J Phys Chem Lett* **2018,** *9* (20), 6096-6102.
21. Plasser, F.; Ruckenbauer, M.; Mai, S.; Oppel, M.; Marquetand, P.; Gonzalez, L., Efficient and Flexible Computation of Many-Electron Wave Function Overlaps. *J. Chem. Theory Comput.* **2016,** *12* (3), 1207-1219.
22. Diederik P. Kingma, a. J. B., Adam: A Method for Stochastic Optimization. **2017**.
23. Martín Abadi, A. A., Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. **2016**, Software available from tensorflow.org.
24. Ben-Nun, M.; Martínez, T. J., Photodynamics of ethylene: ab initio studies of conical intersections. *Chem. Phys.* **2000,** *259* (2-3), 237-248.
25. Ohmine, I., Mechanisms of nonadiabatic transitions in photoisomerization processes of conjugated molecules: Role of hydrogen migrations. *J. Chem. Phys.* **1985,** *83* (5), 2348-2362.
26. Zhou, W.; Mandal, A.; Huo, P., Quasi-Diabatic Scheme for Nonadiabatic On-the-Fly Simulations. *J Phys Chem Lett* **2019,** *10* (22), 7062-7070.