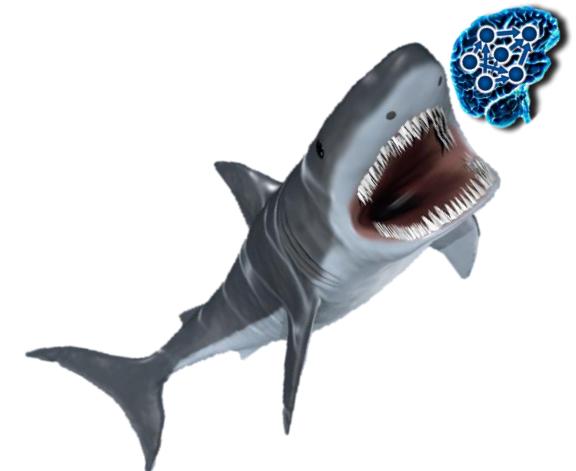




WARWICK  
THE UNIVERSITY OF WARWICK

# Machine Learning for Surface Hopping Molecular Dynamics: The Case of Excited Tyrosine

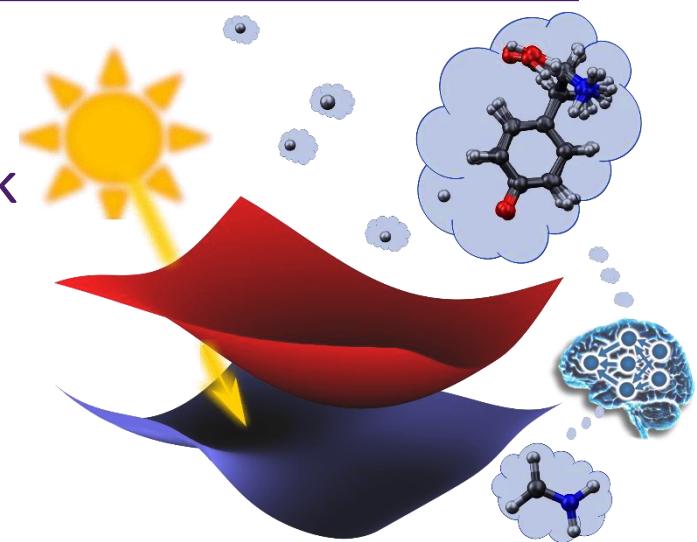


Julia Westermayr

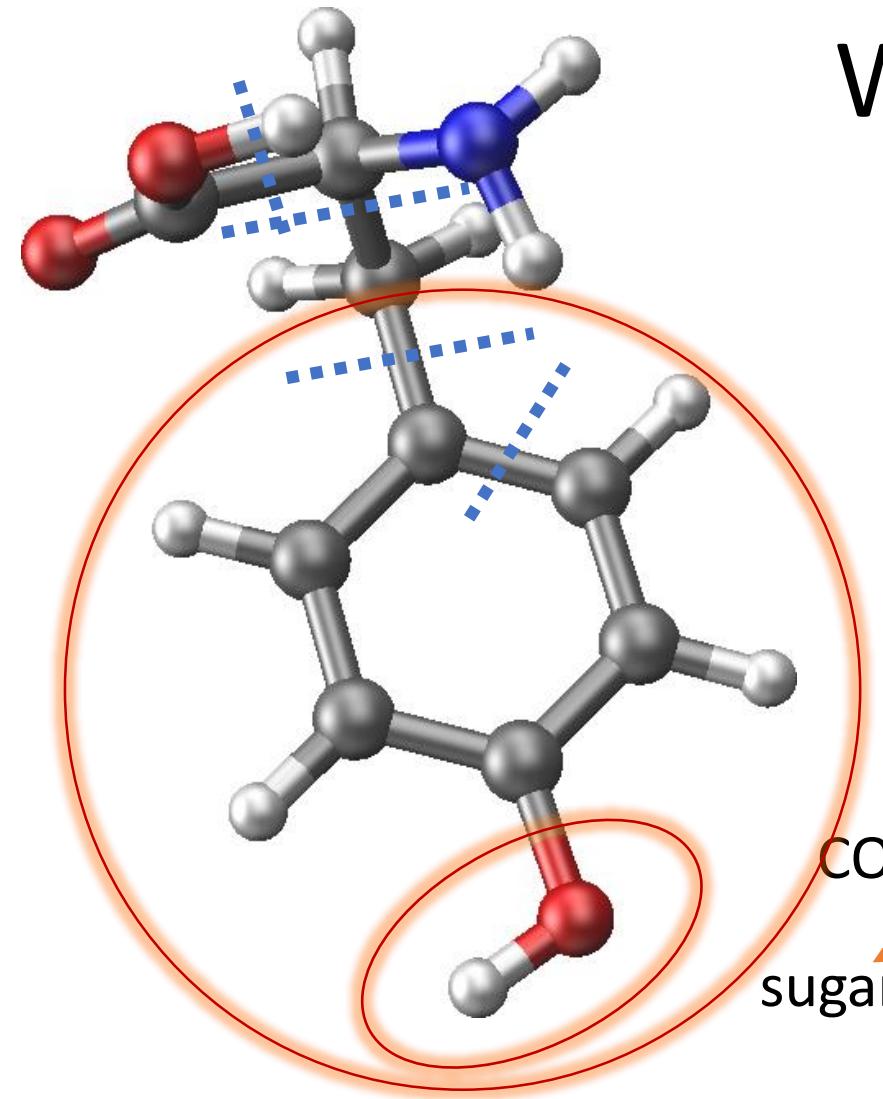
Department of Chemistry, University of Warwick

VISTA Webinar

May 26, 2021

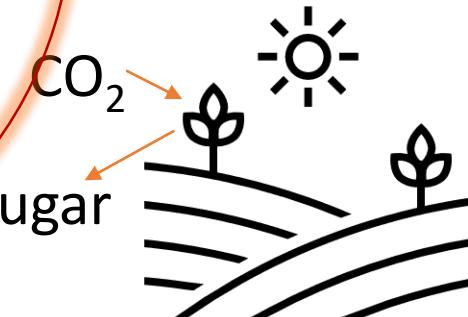


# Why are we interested in tyrosine?

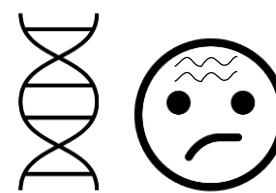


- Multifaceted photochemistry of peptides and proteins
- Experimental evidence for photodissociation after light excitation
- Experiments suggest size-dependent photodynamics
- Many theoretical studies focus on smaller chromophores

Photosynthesis



Skin alteration  
and cancer

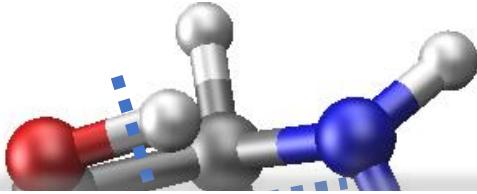


Cataracts:  
blindness



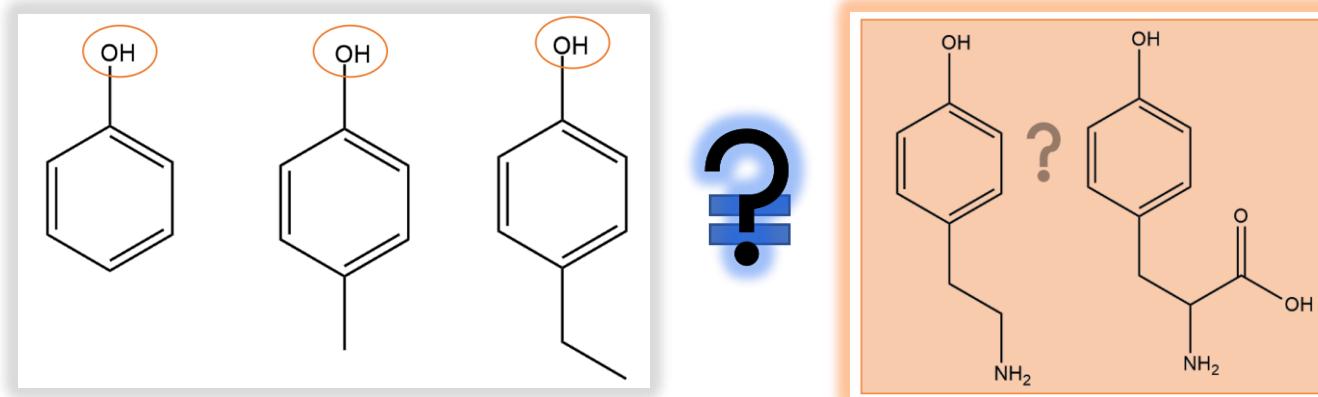
Drug  
alteration



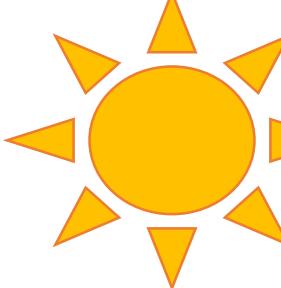


# The excited states of tyrosine

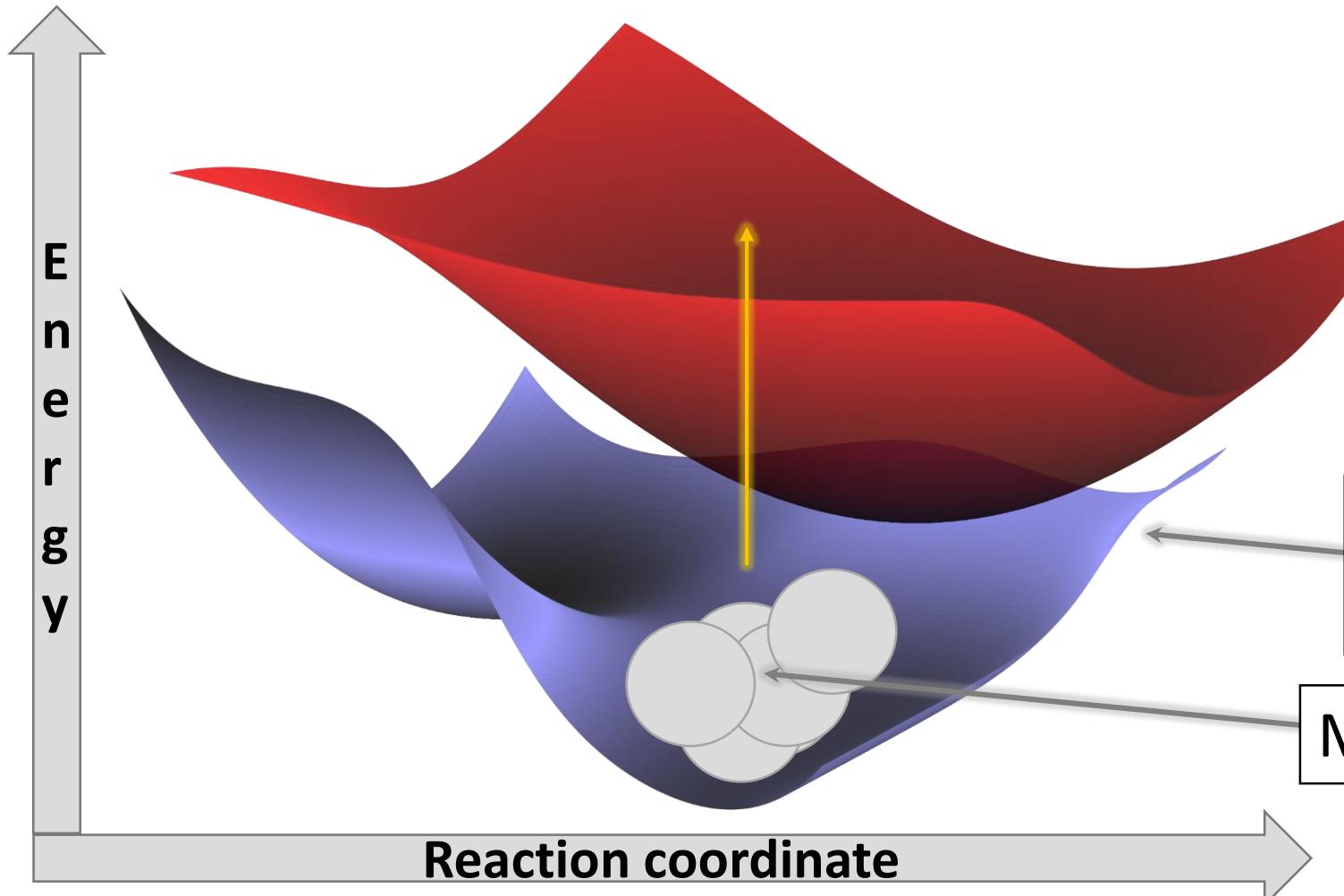
What mechanisms lead to different photoproducts in tyrosine?



Goal: complement experiments and shed light on the mechanisms of excited tyrosine



# Surface hopping molecular dynamics

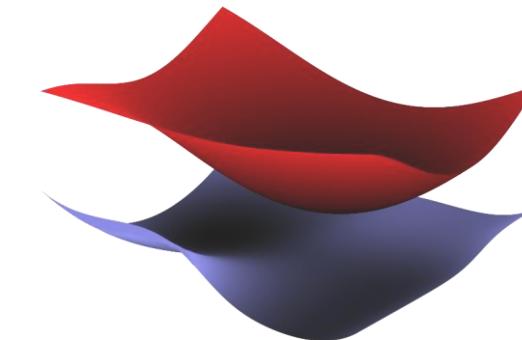
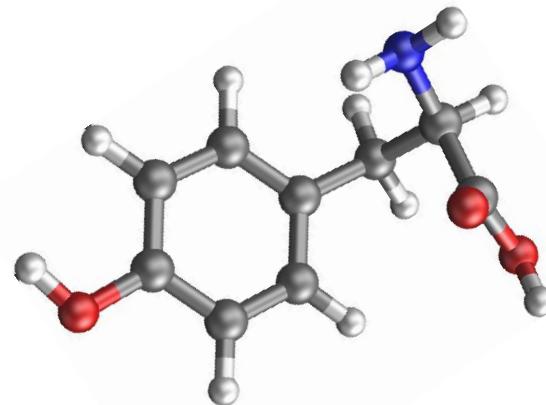


- Mixed quantum-classical
- Nonadiabatic transitions
- Many independent

Quantum chemical potential energy surfaces

Molecules: move classically

# Problem: expensive calculations of potential energy surfaces

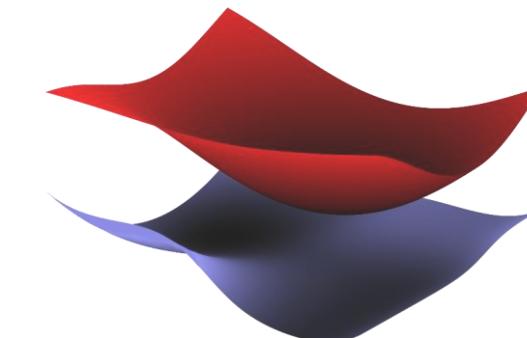
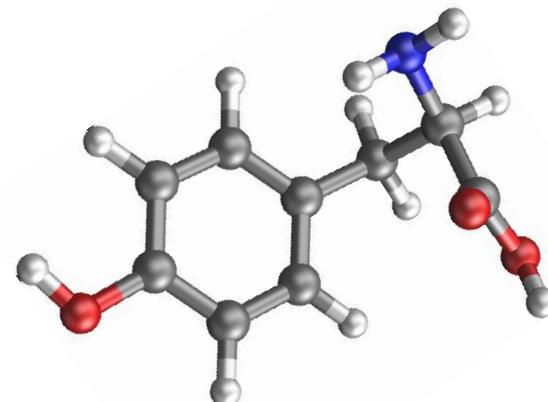


Electronic  
structure

Quantum Chemistry	Accuracy	CPU time QC 1 ps
MS-CASPT2(12,11)/ano-rcc-pVDZ	very high	~ 2,910 days <b>x 1000</b>

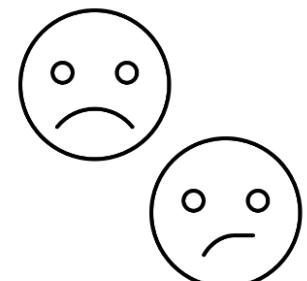


# Problem: expensive calculations of potential energy surfaces

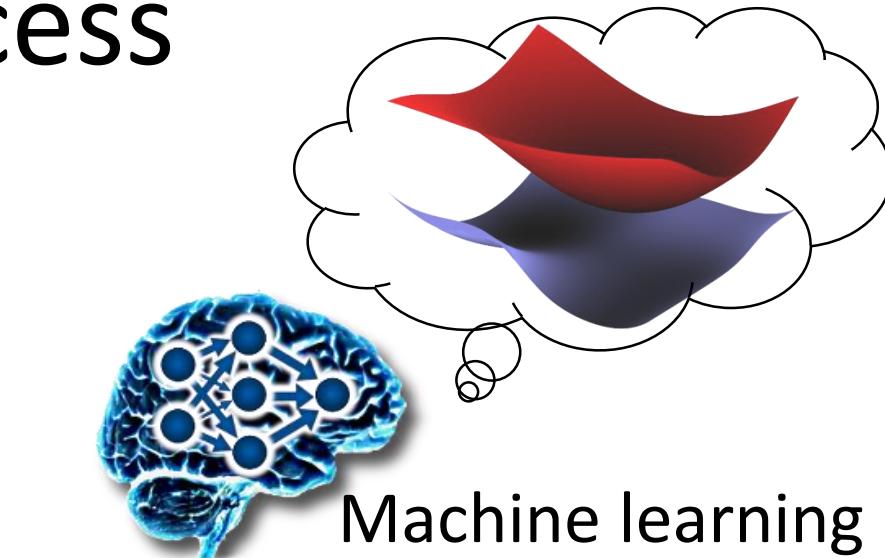
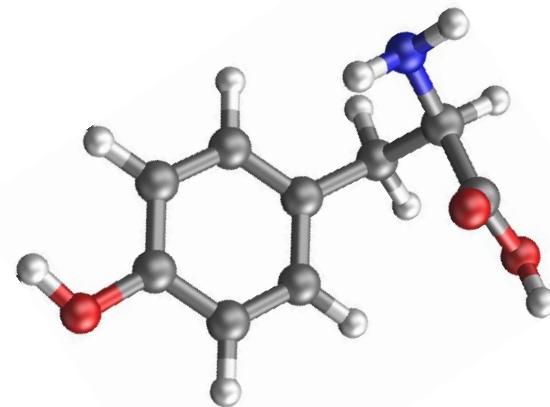


Electronic  
structure

Quantum Chemistry	Accuracy	CPU time QC 1 ps
MS-CASPT2(12,11)/ano-rcc-pVDZ	very high	$\sim 2,910 \text{ days} \times 1000$
ADC(2)/def2-SVP	not given for dissociation	$\sim 10 \text{ months} \times 1000$

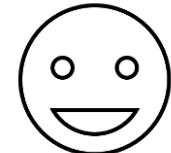


# Problem: expensive calculations of potential energy surfaces



Machine learning (ML)

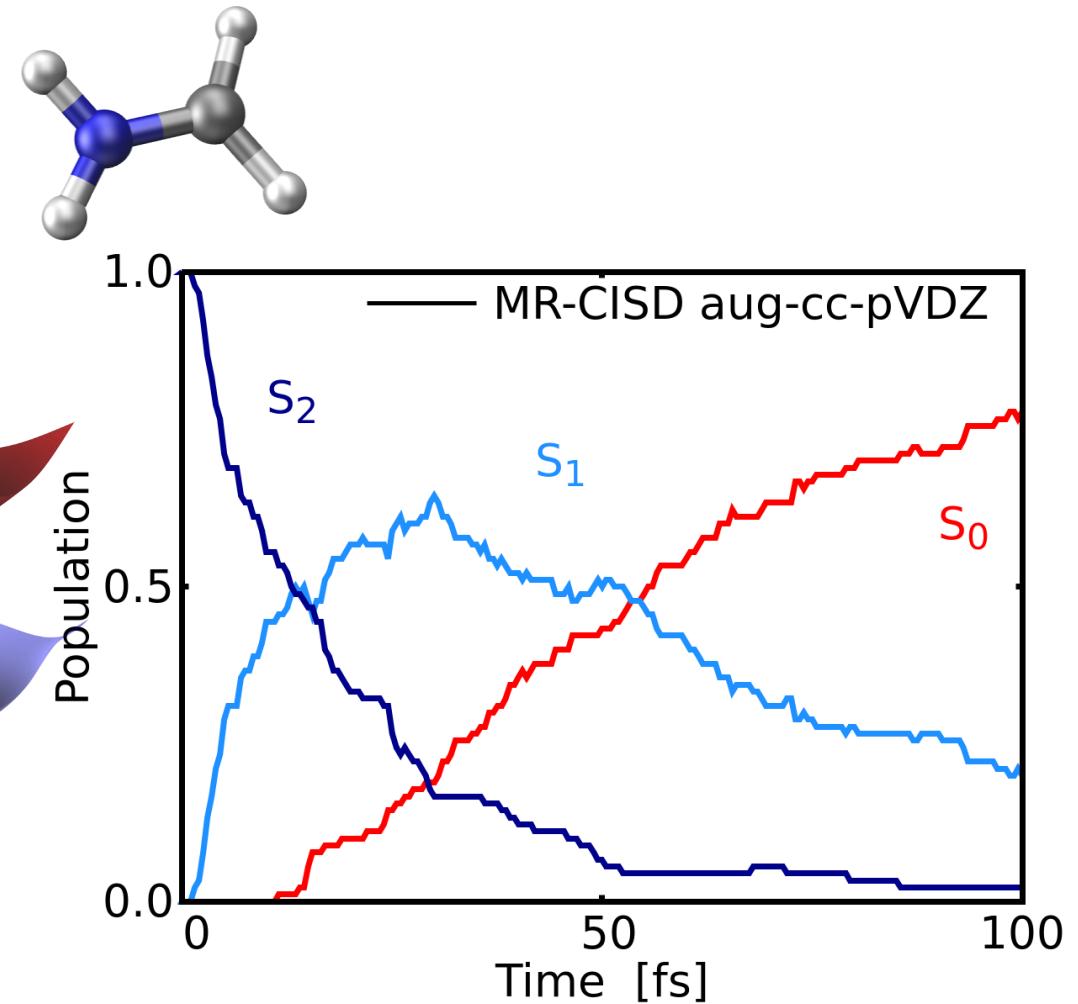
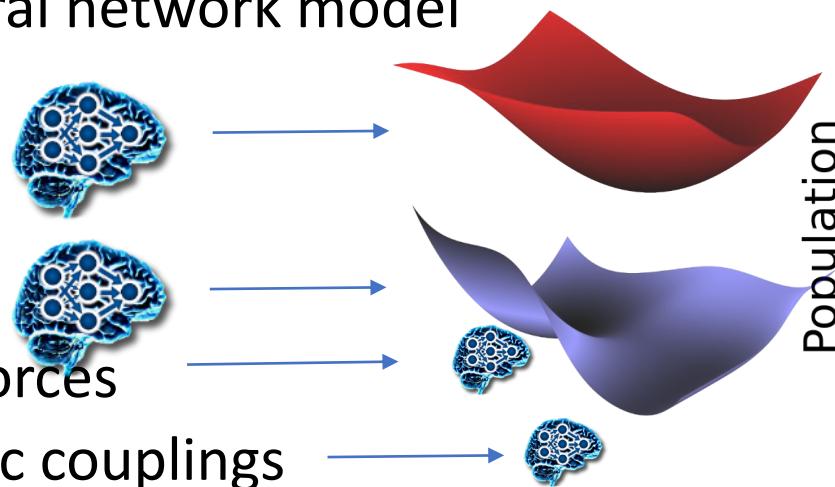
Quantum Chemistry	Accuracy	CPU time QC 1 ps
MS-CASPT2(12,11)/ano-rcc-pVDZ	very high	$\sim 2,910 \text{ days} \times 1000$
ADC(2)/def2-SVP	not given for dissociation	$\sim 10 \text{ months} \times 1000$
Machine learning	~ reference	Few min. – hours $\times 1000$



Can we use ML to enhance photodynamics simulations?

# Proof of concept

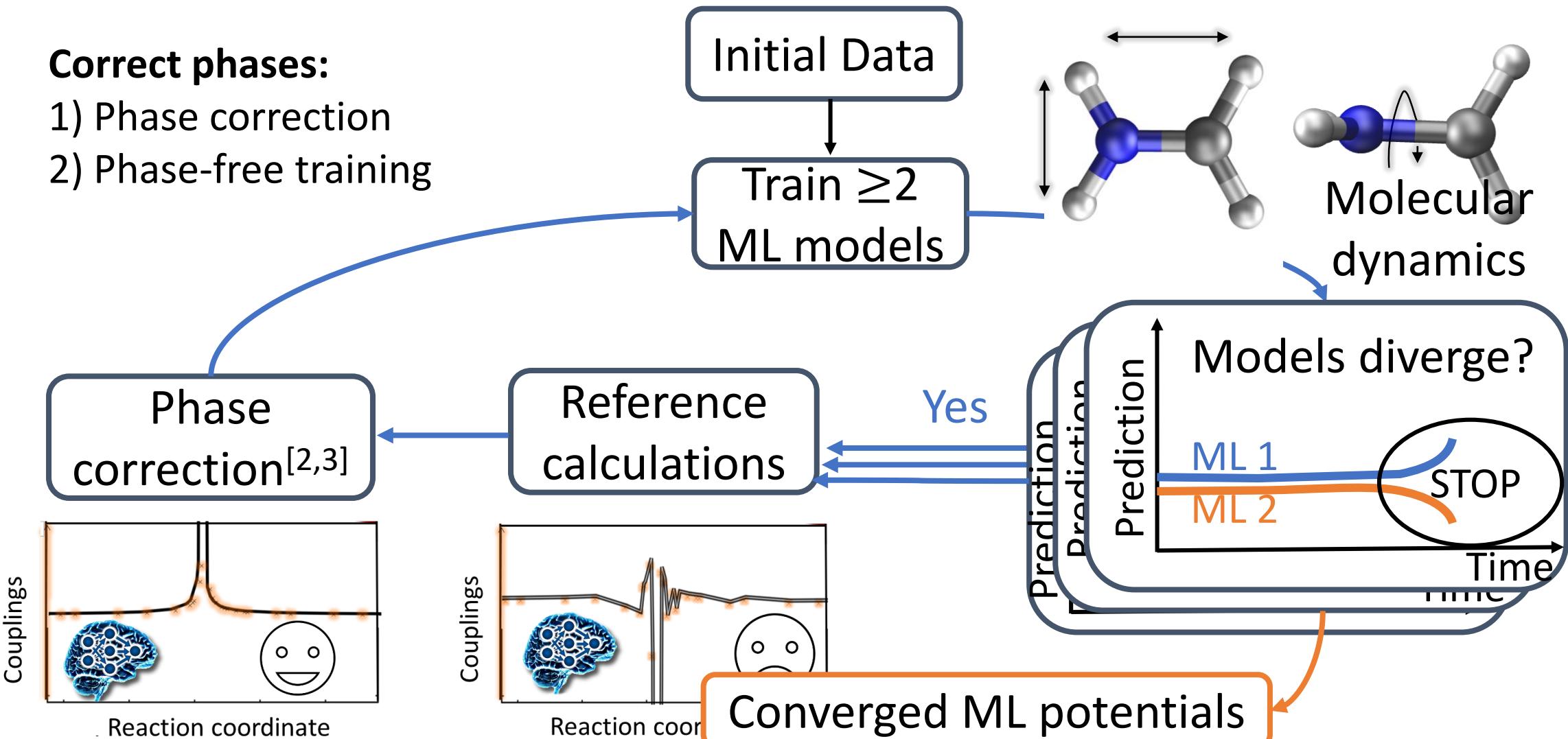
- Methylenimmonium cation,  $\text{CH}_2\text{NH}_2^+$
- Ultrafast population transfer
- MR-CISD/CAS(6,4)/aug-cc-pVDZ
- Multi-state neural network model
- 3 singlet states
  - Energies + forces
  - Nonadiabatic couplings
- Requirement: comprehensive, accurate potentials with few data points



# Adaptive sampling<sup>[1]</sup> for excited states<sup>[2]</sup>

**Correct phases:**

- 1) Phase correction
- 2) Phase-free training

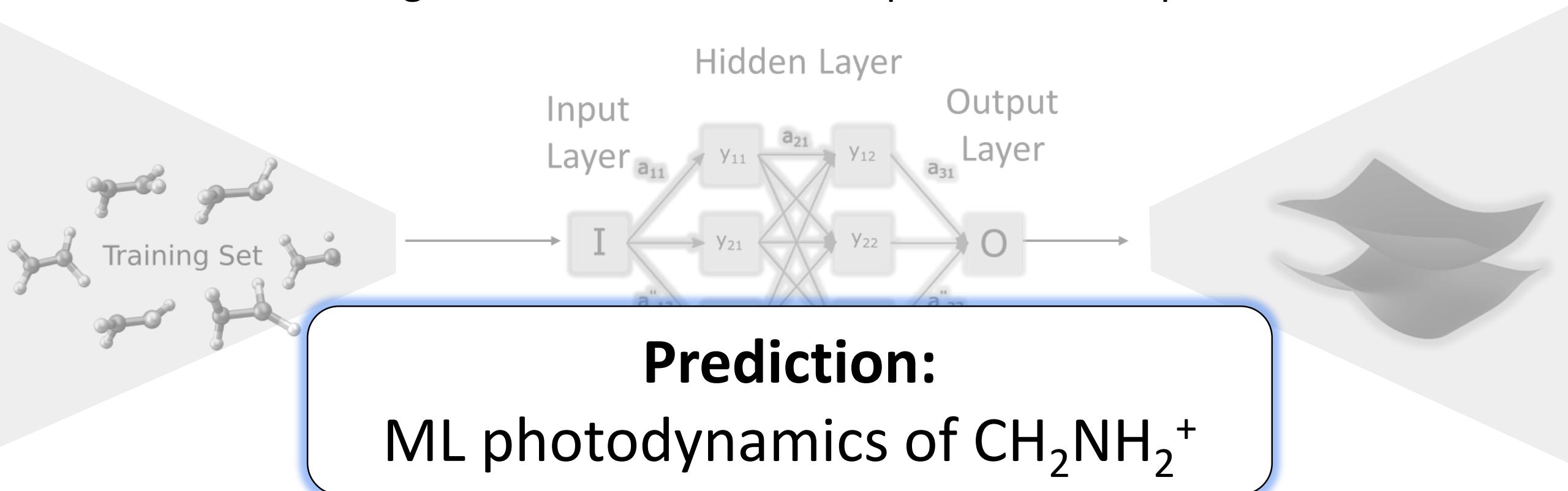


[1] J. Behler, *Int. J. Quantum Chem.* 115, 1032–1050 (2015) [2] J. Westermayr, et al. *Chem. Sci.*, 10, 8100-8107 (2019).

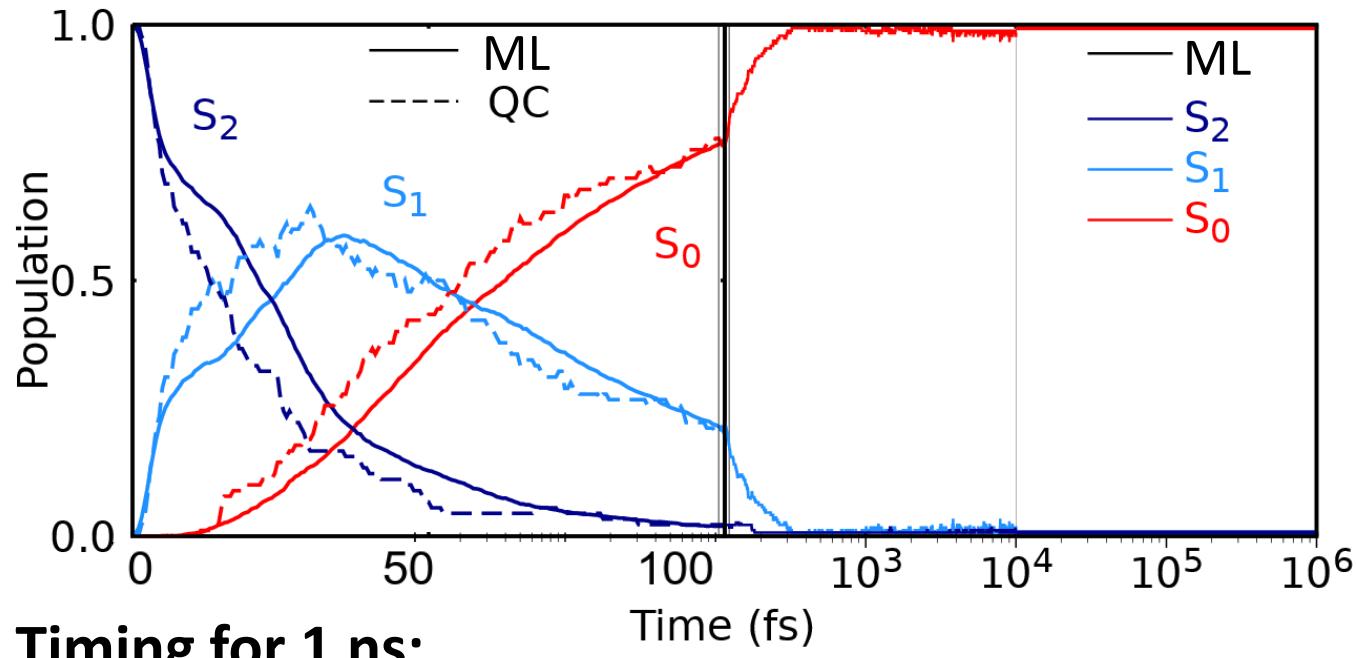
[3] A. V. Akimov, *J. Phys. Chem. Lett.*, 9(20), 6096-6102 (2018)

# ML photoynamics simulations

Finding a relation between an input and an output



# Machine Learning Surface Hopping Molecular Dynamics



**Timing for 1 ns:**

ML: ~118 times faster

(~ 59 days; QC: ~19 years on the same CPU)   QC: MR-CISD/aug-cc-pVDZ

## Goal for ML:

- Reproduce excited state dynamics of QC (100 fs) ✓
- Better statistics ✓
- Long excited-state dynamics ✓

# SchNarc approach for photodynamics<sup>[1]</sup>

Validated for different systems with singlet and triplet states



UNIVERSITY  
OF VIENNA



Technische  
Universität  
Berlin



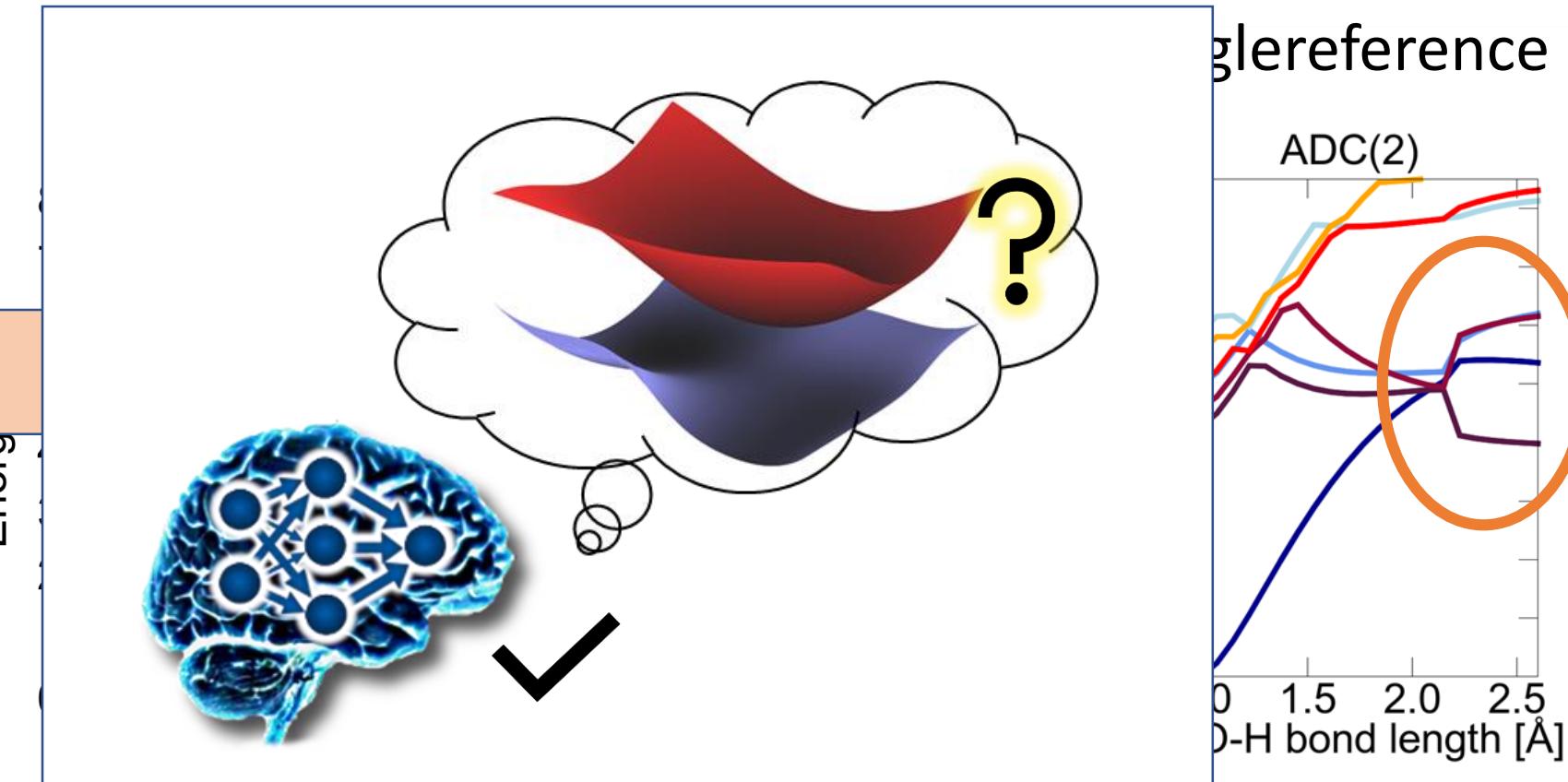
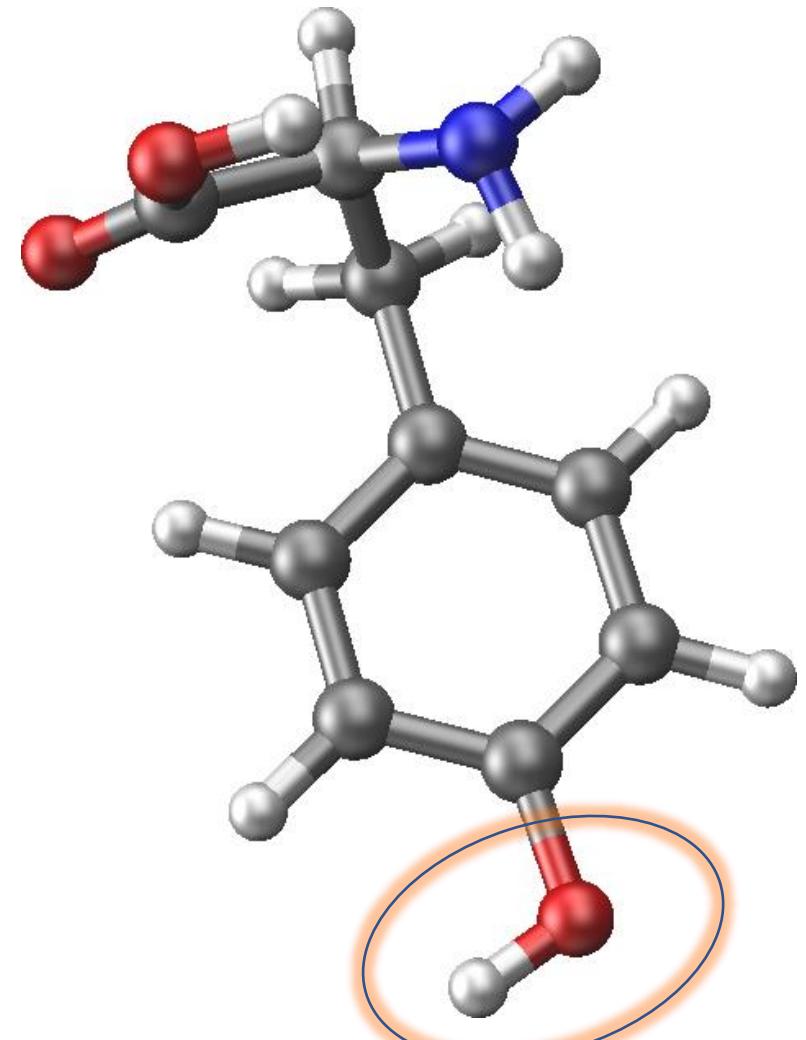
Application:  
The excited states of tyrosine



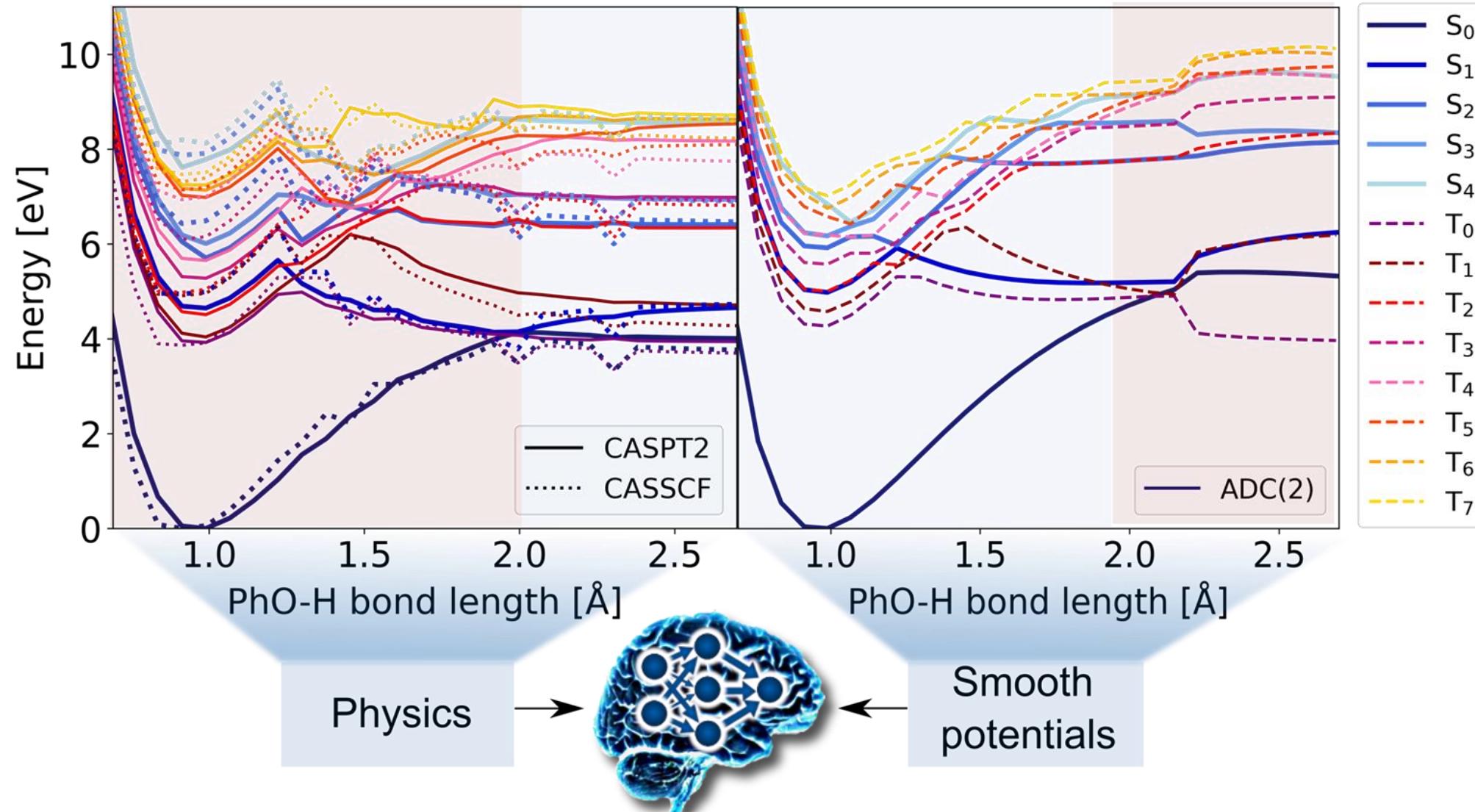
[1] <https://sharc-md.org/> [2] K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller., *J. Chem. Theory Comput.*, 2019, **15**(1):448–455. [3] J. Westermayr, M. Gastegger, P. Marquetand, *J. Phys. Chem. Lett.* **11**(10), 3828-3834 (2020).

# Application of SchNarc

- 1) Trainingset generation: which reference method?



# Combine the best of both methods



# ML for excited states of tyrosine



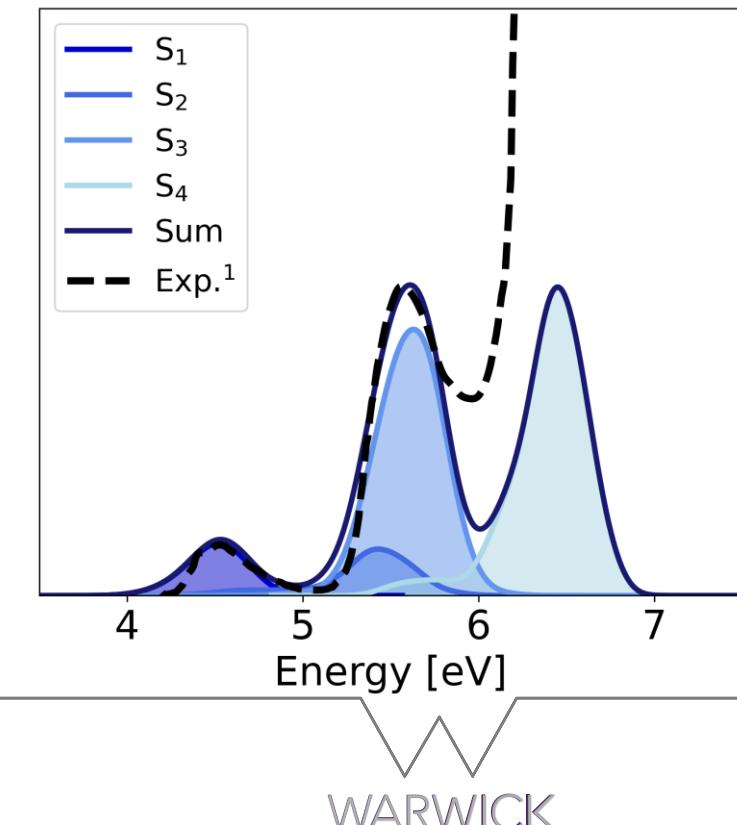
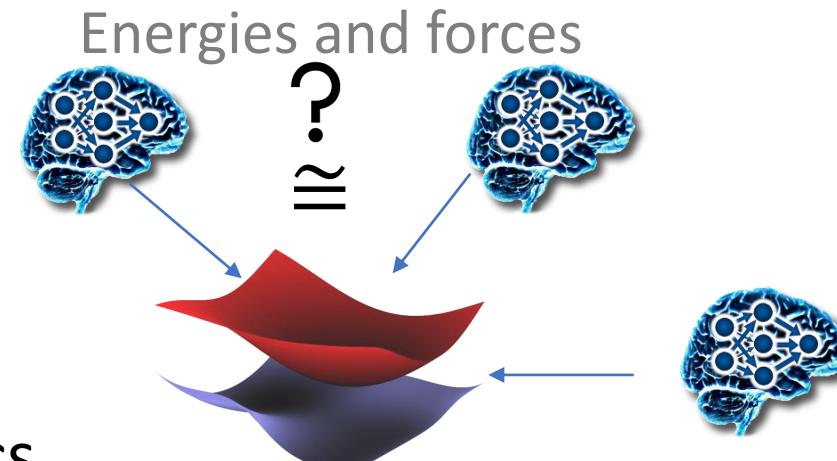
D. Vörös    L. Panzenböck    F. Jörg

**Training set generation: 17,265 data points**

- Initial sampling + adaptive sampling (6 rounds)
- Phasecorrected data
- 5 singlets, 8 triplets

**Photodynamics:**

- Excitation to  $S_4$
- 3 NNs during dynamics
- 1022 trajectories up to 10 ps



# Photodynamics of tyrosine



SHARE

REPORT



## The Roaming Atom: Straying from the Reaction Path in Formaldehyde Decomposition

D. Townsend<sup>1,2</sup>, S. A. Lahankar<sup>3</sup>, S. K. Lee<sup>1,2,3</sup>, S. D. Chambreau<sup>3</sup>, A. G. Suits<sup>1,2,3,\*</sup>, X. Zhang<sup>4</sup>, J. Rheinecker<sup>4</sup>, L. B. Harding<sup>5</sup>, J. M. Bowman<sup>4,\*</sup>

<sup>1</sup> Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, USA.

<sup>2</sup> Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, USA.

<sup>3</sup> Department of Chemistry, Wayne State University, Detroit, MI 48202, USA.

<sup>4</sup> Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, GA 30322, USA.

<sup>5</sup> Chemistry Division, Argonne National Laboratory, Argonne, IL 60439, USA.

\* To whom correspondence should be addressed. E-mail: [asuits@wayne.edu](mailto:asuits@wayne.edu); [jmbowma@emory.edu](mailto:jmbowma@emory.edu)

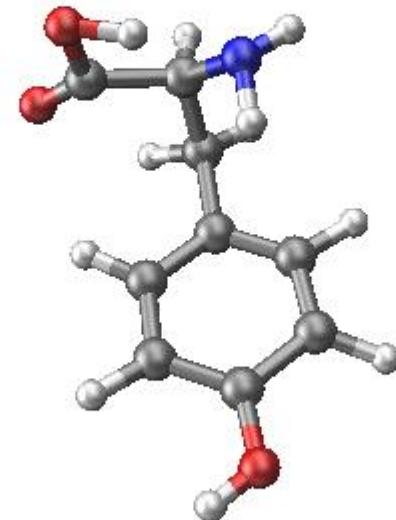
- Hide authors and affiliations

D. Townsend, *et al.*, Science 306, 5699, 1158-1161 (2004).

# Roaming in tyrosine

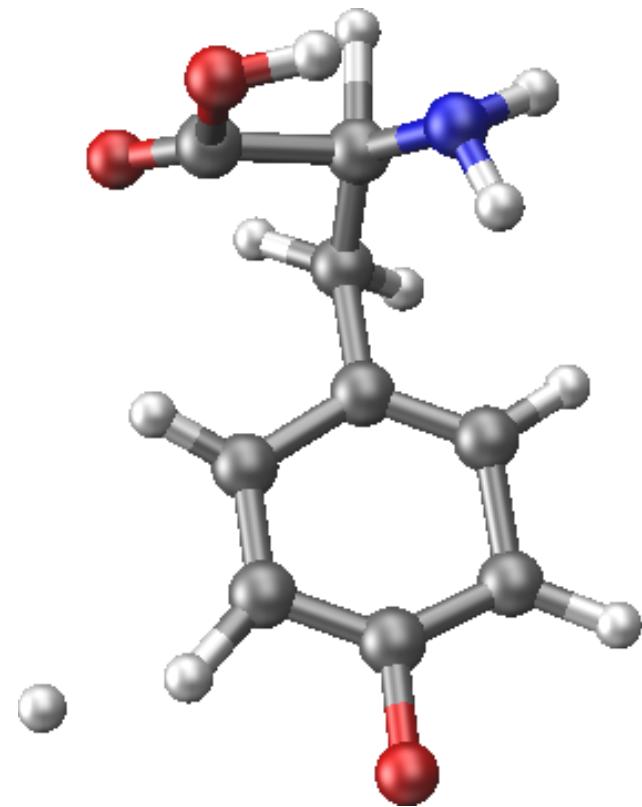
Time: 0.5 fs

- **Validated with CASPT2(12,11)/ano-rcc-pVDZ**
- Energetically favored path compared to dissociation



# Roaming in tyrosine

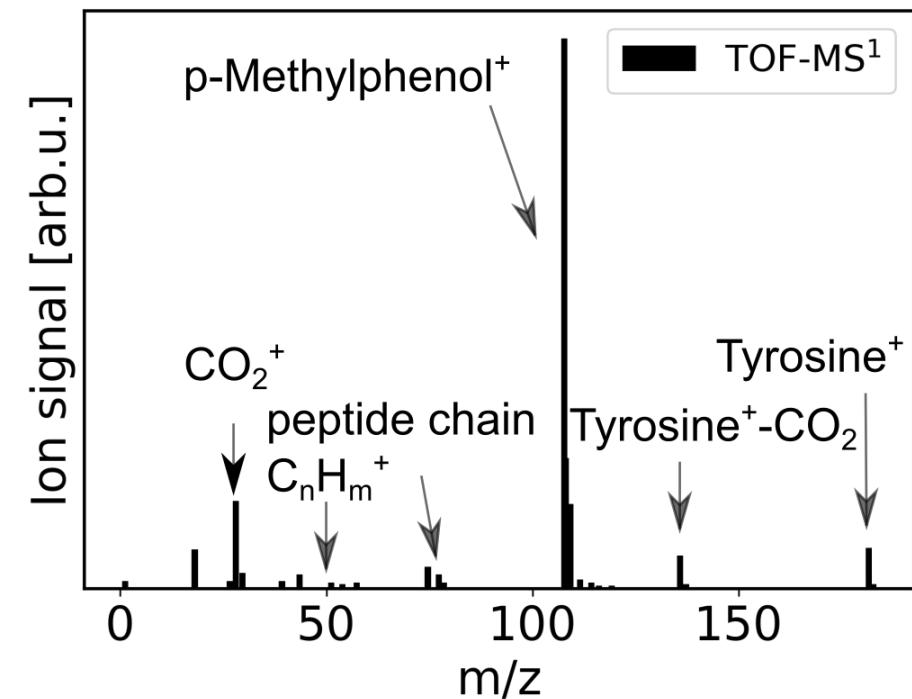
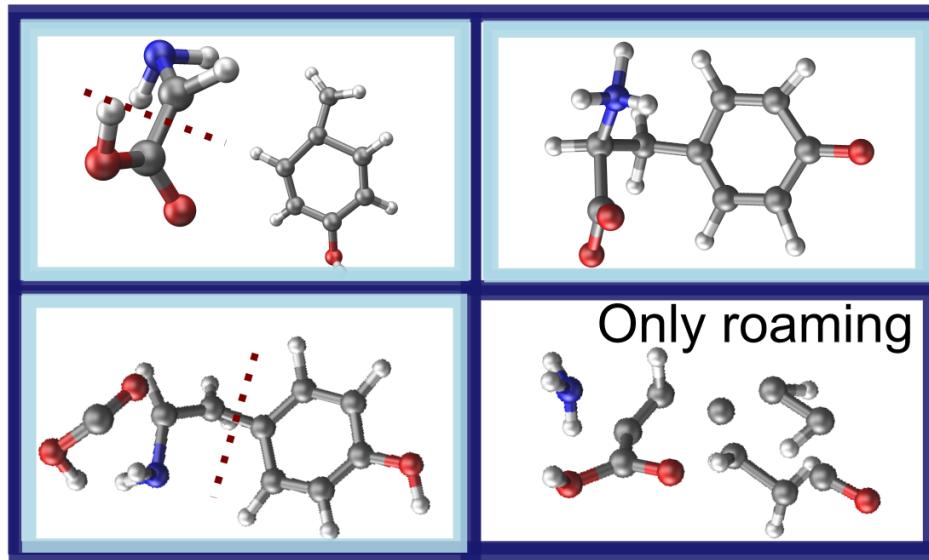
- Validated with CASPT2(12,11)/ano-rcc-pVDZ
- Energetically favored path compared to dissociation
- Found about 17% roaming
- Unsupervised ML for analysis of 1022 trajectories:
  - Extract important features from roaming and non-roaming trajectories



# Roaming in tyrosine

- **No roaming:** 10% H dissociation, 14% fragmentation
- **Roaming:** 50% H dissociation, 25% fragmentation

Clusters found with unsupervised learning:

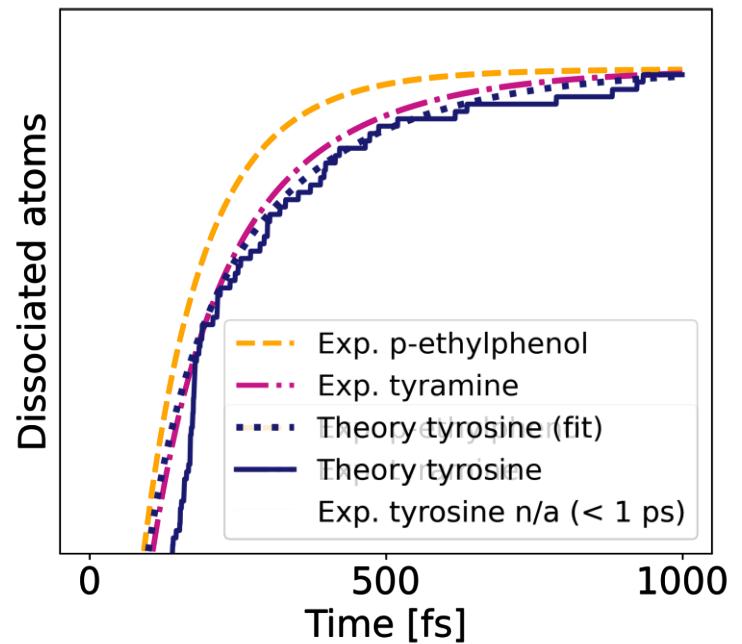
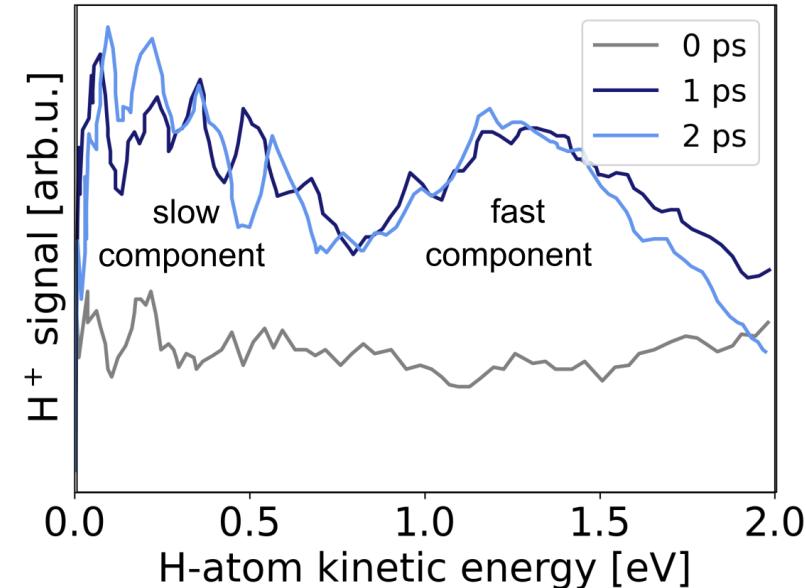


J. Westermayr *et al.*, unpublished (2021).

Mass spectrum extracted from A. Iqbal, PhD thesis, University of Warwick (2010).

# The dissociative ${}^1\pi\sigma^*$ state

- Stavros and co-workers:<sup>[1]</sup>
  - Time-resolved pump/probe spectroscopy + velocity map ion imaging
  - Two reaction channels, ultrafast time scales (< 1 ps)
  - $k$  (p-Ethylphenol) <  $k$ (tyramine) <  $k$ (tyrosine)
- Photodynamics:
  - Tyrosine:  $k_1 = 66 \pm 9$  fs,  $k_2 = 237 \pm 77$  fs
  - Fast channel: dissociation in ground state
  - Slow channel: dissociation in excited state



[1] A. Iqbal, V. G. Stavros, *J. Phys. Chem. Lett.* **1**, 2274–2278 (2010)

KER spectrum extracted from Ref. 1

J. Westermayr *et al.*, unpublished (2021).



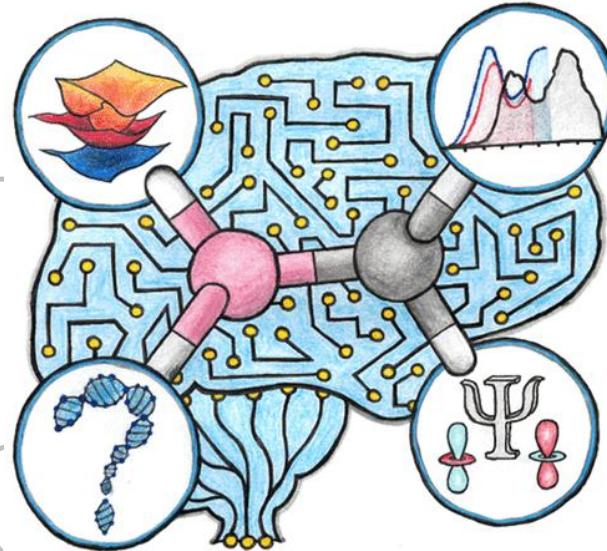
# Conclusion

## Photodynamics of tyrosine:

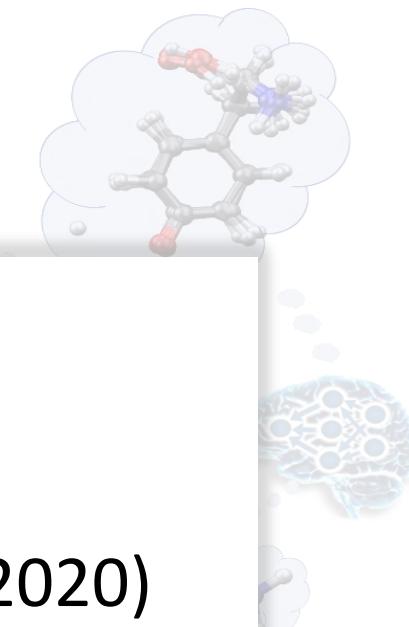
- Power of theory: complemen...
- How does this knowledge infl...
- What is the influence of an er...
- Still many questions unanswe...

## ML for photo...

- Can enab...
- Describe...
- Can help...



...de insights  
peptides and



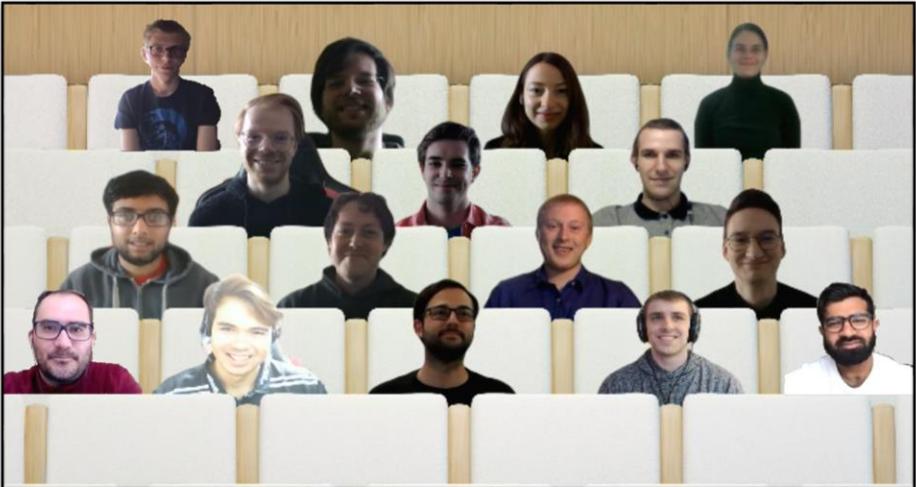
## (Extensive) review:

J. Westermayr and P. Marquetand

„Machine learning for excited states“,

*Chem. Rev.*, doi:10.1021/acs.chemrev.0c00749 (2020)

# Thank you for your attention!



- Priv.-Doz. Philipp Marquetand
- Univ.-Prof. Leticia González
- Dr. Michael Gastegger



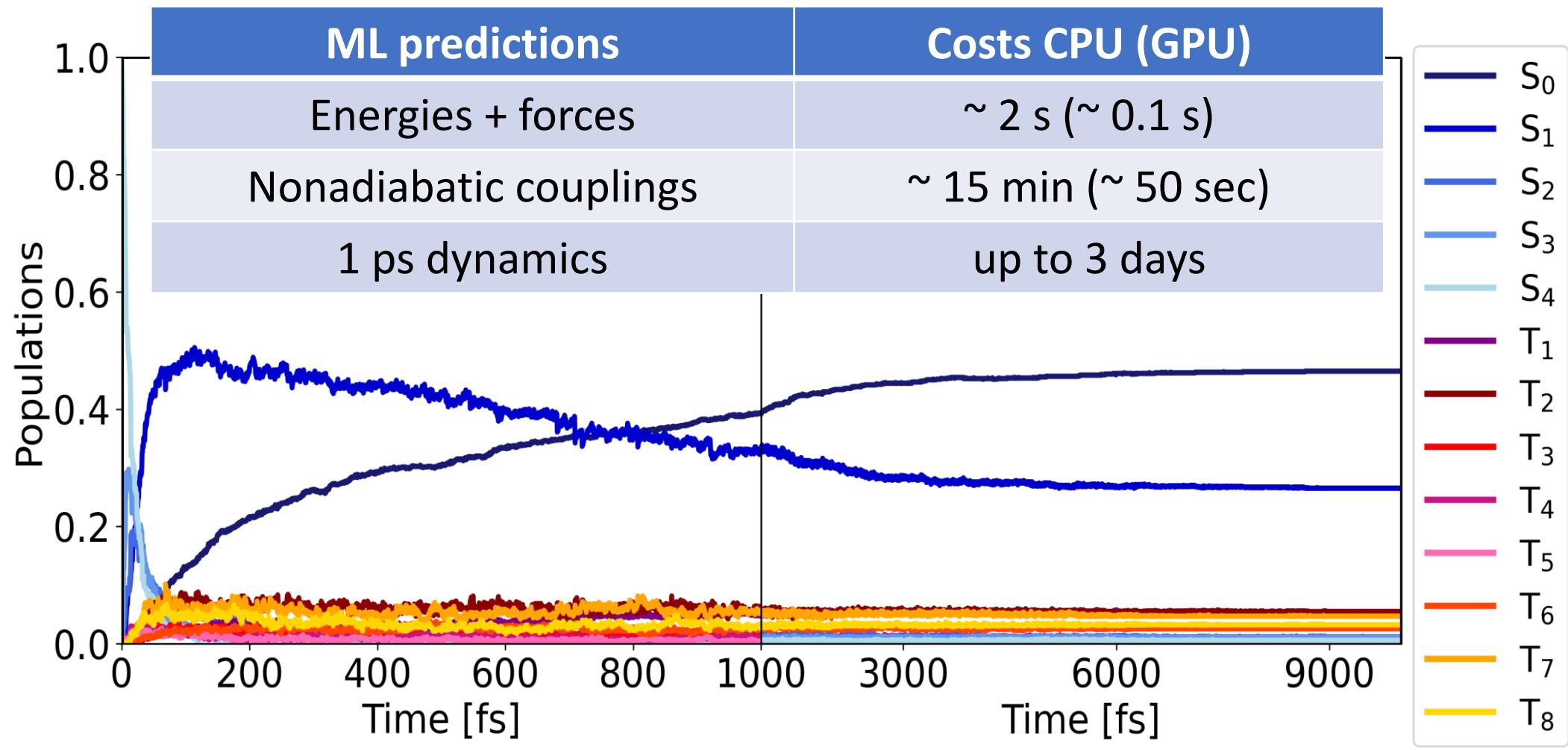
Thank you for your attention

Any Questions?

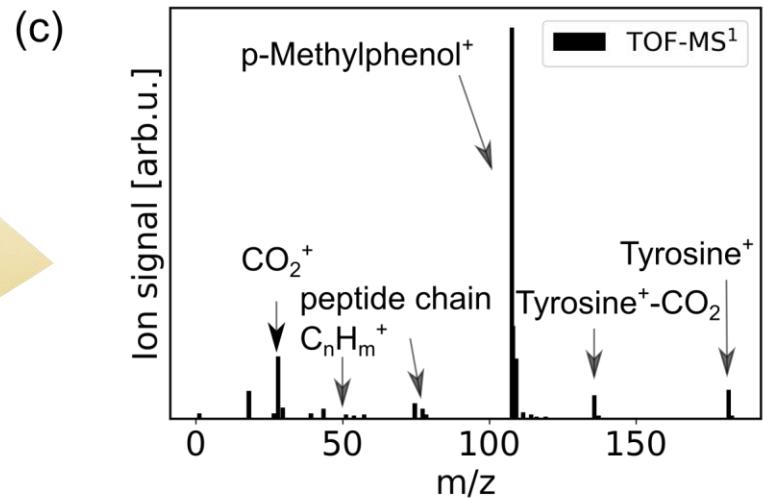
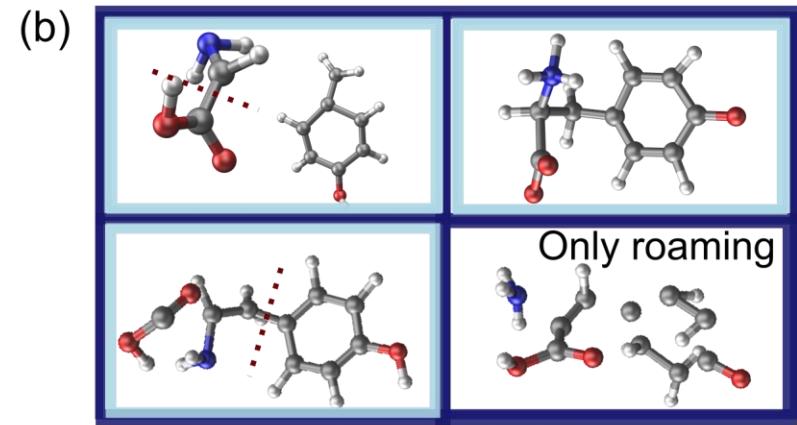
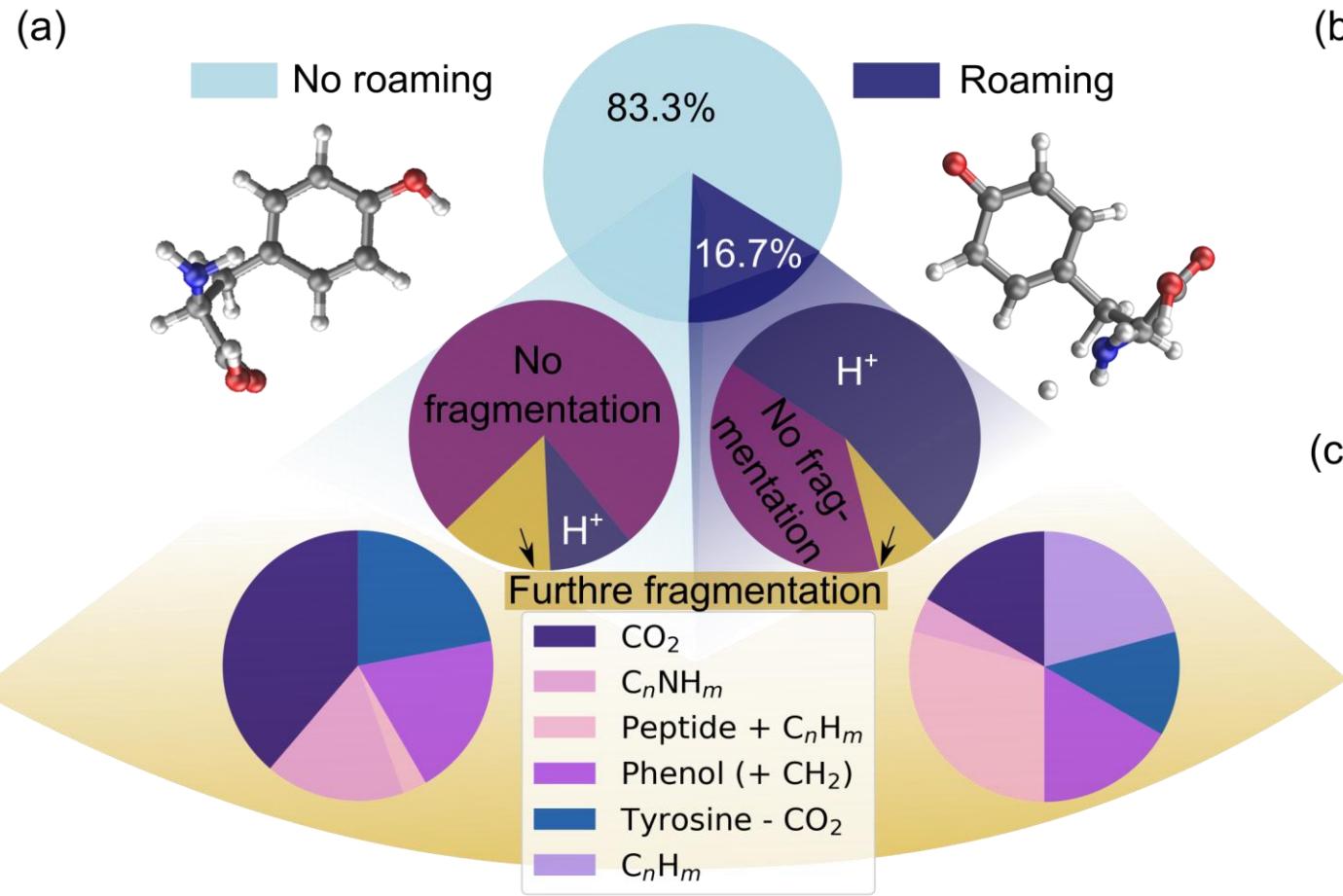


W  
A  
R  
W  
I  
C  
K

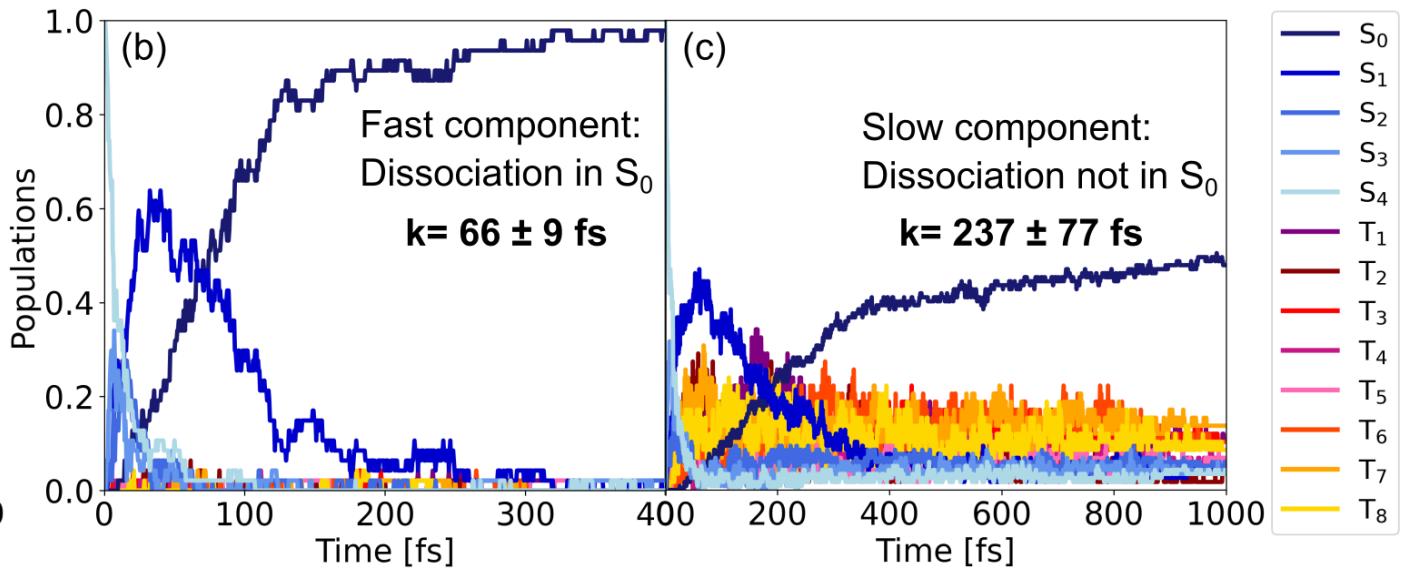
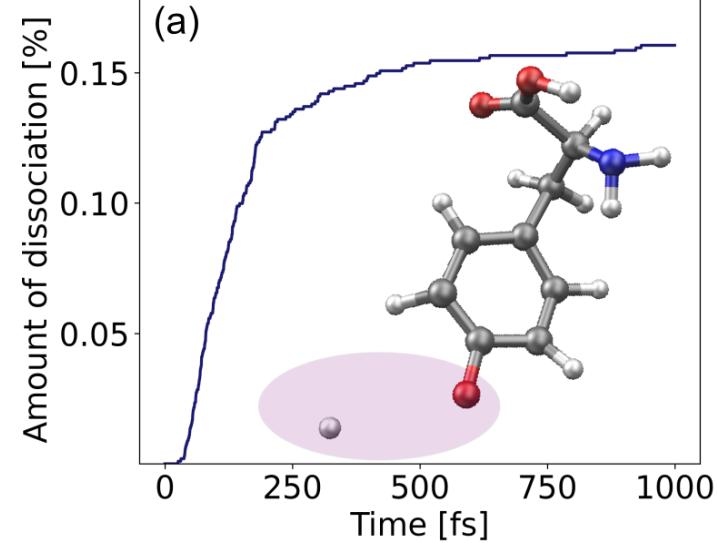
# Populations



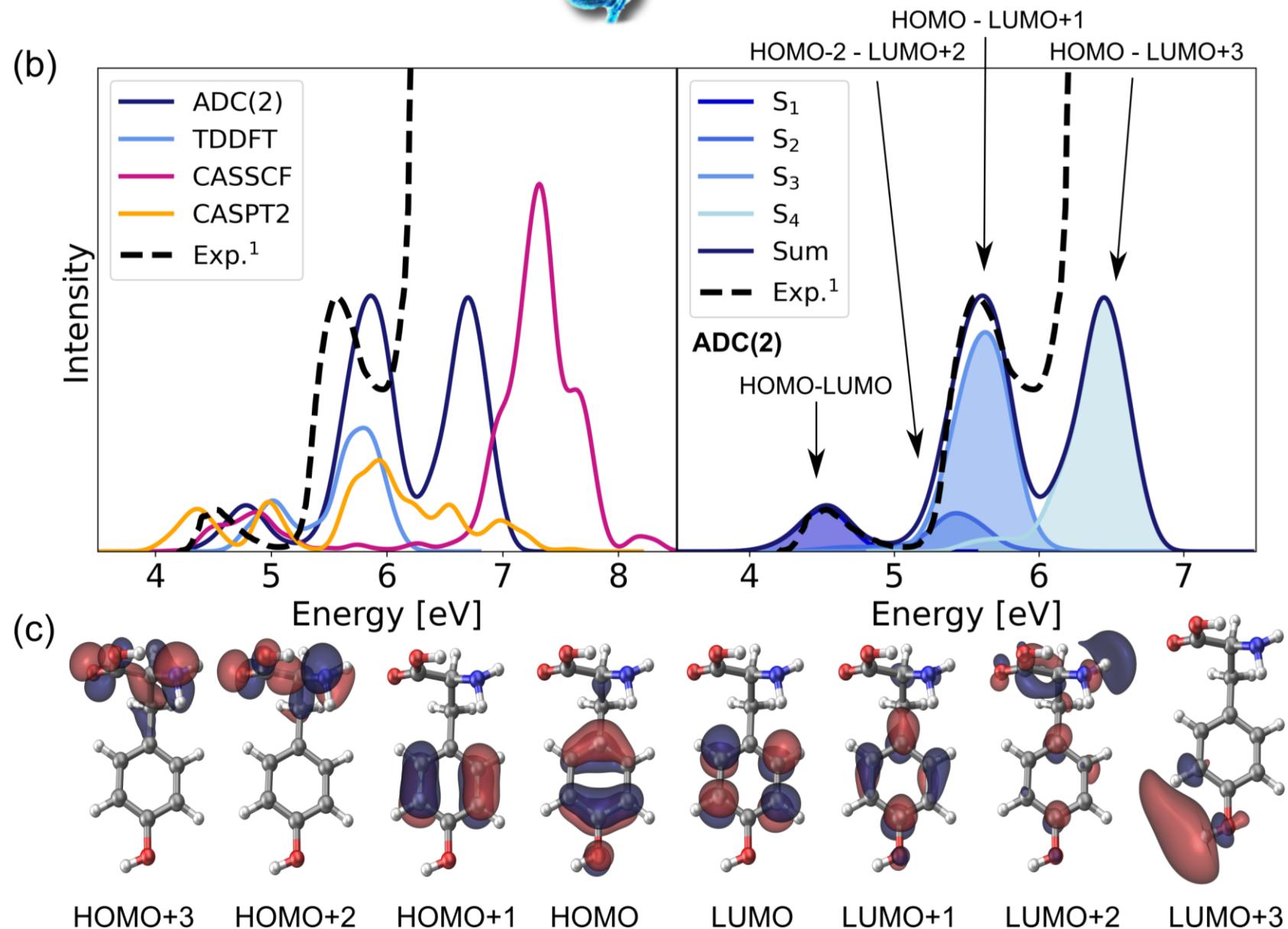
# Roaming in tyrosine



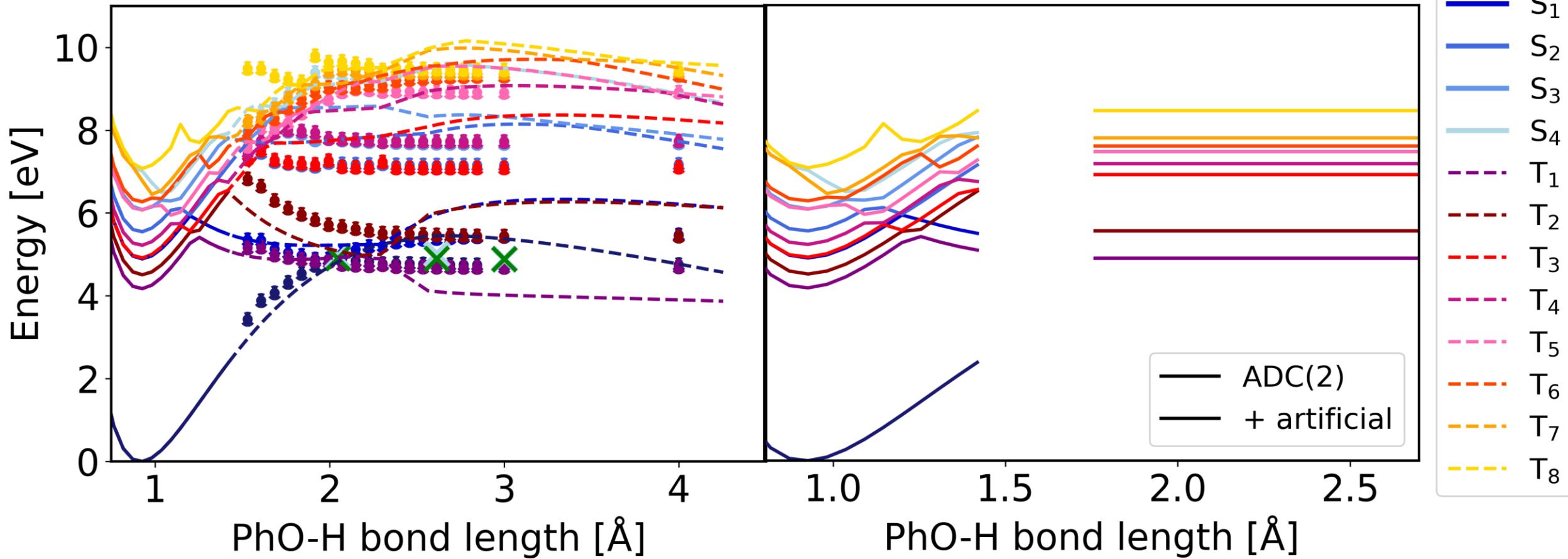
# Kinetics



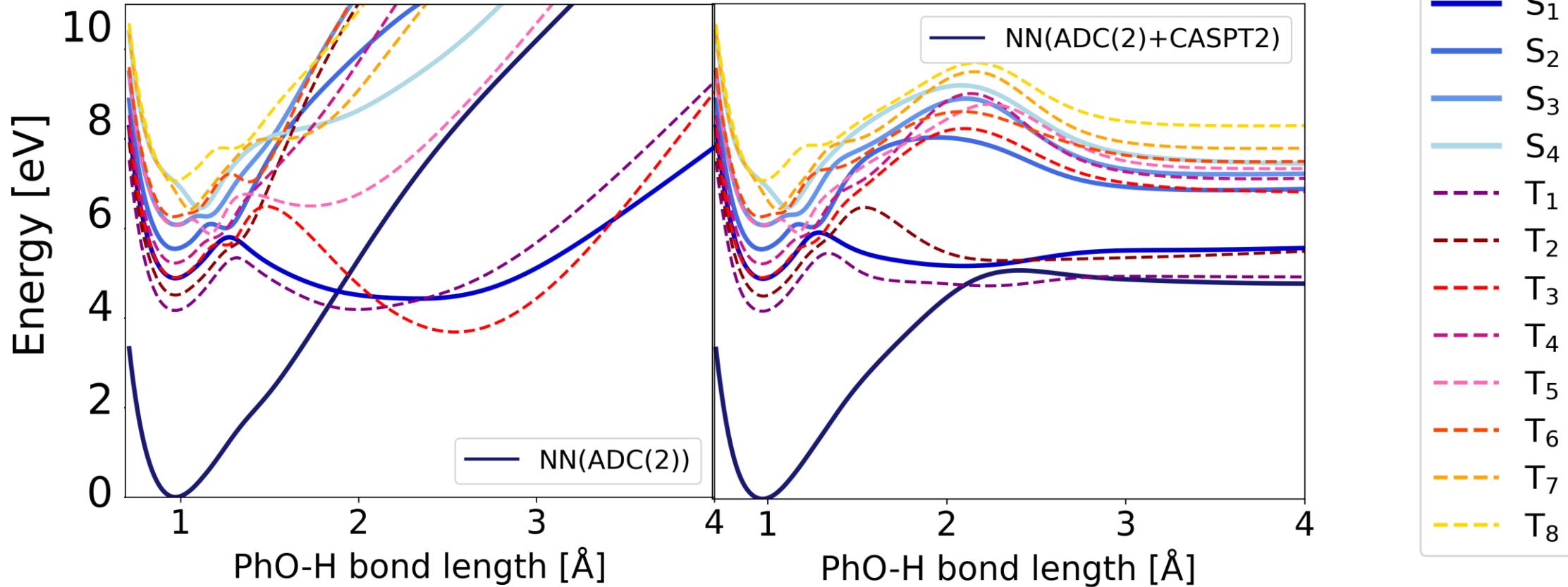
# Comparison of methods



# Artificial data

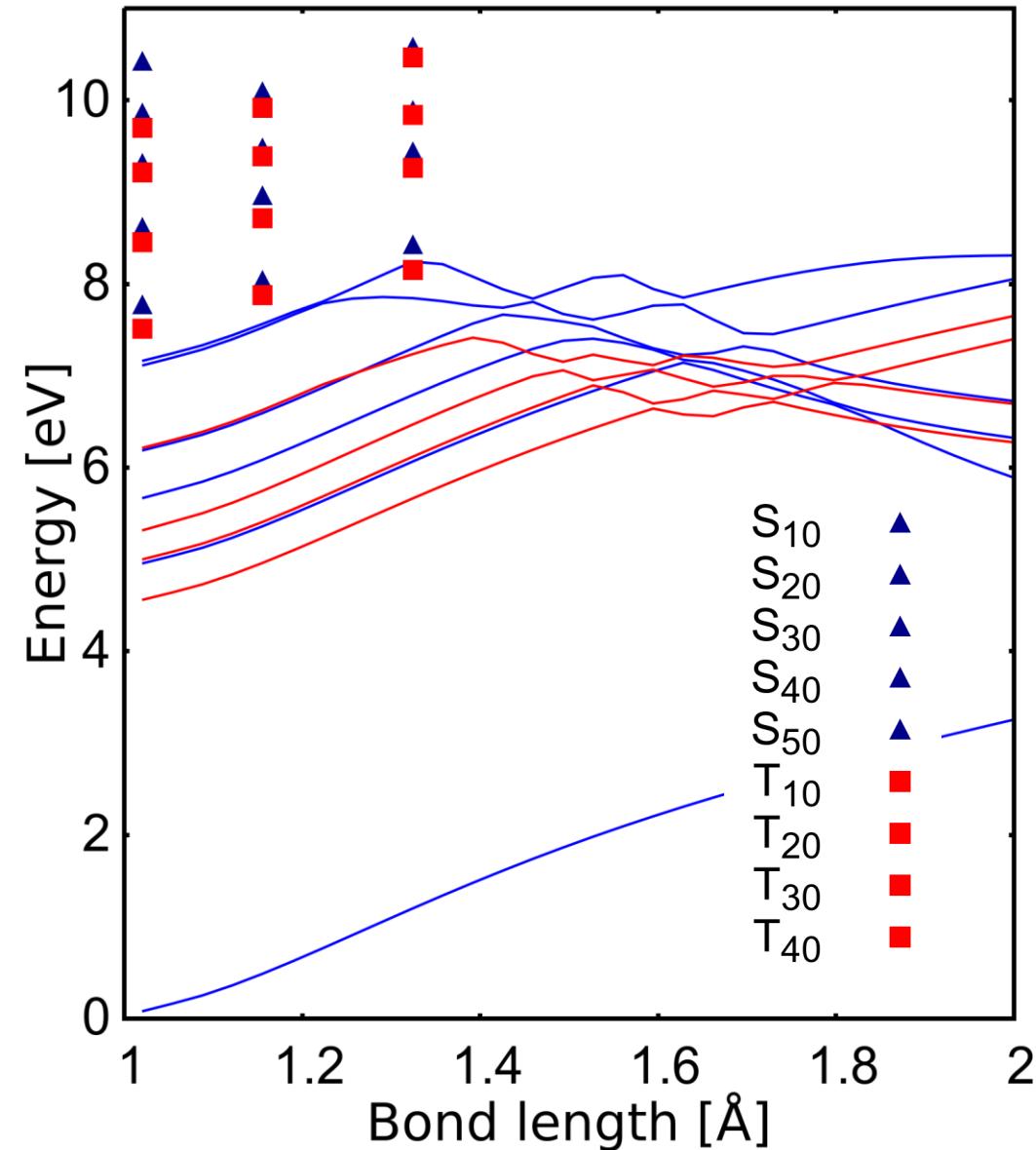


# ML predictions

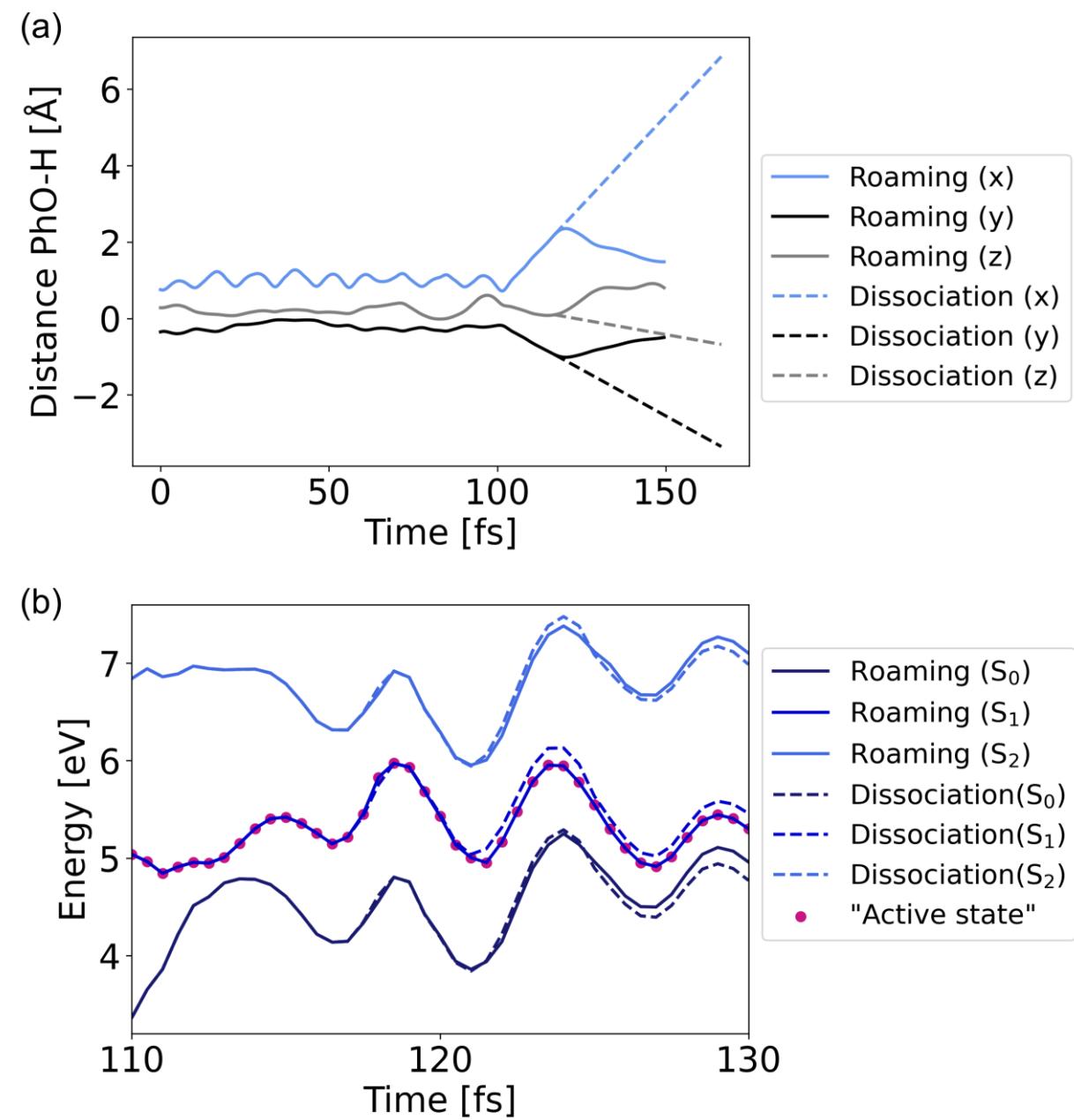


# Phase correction

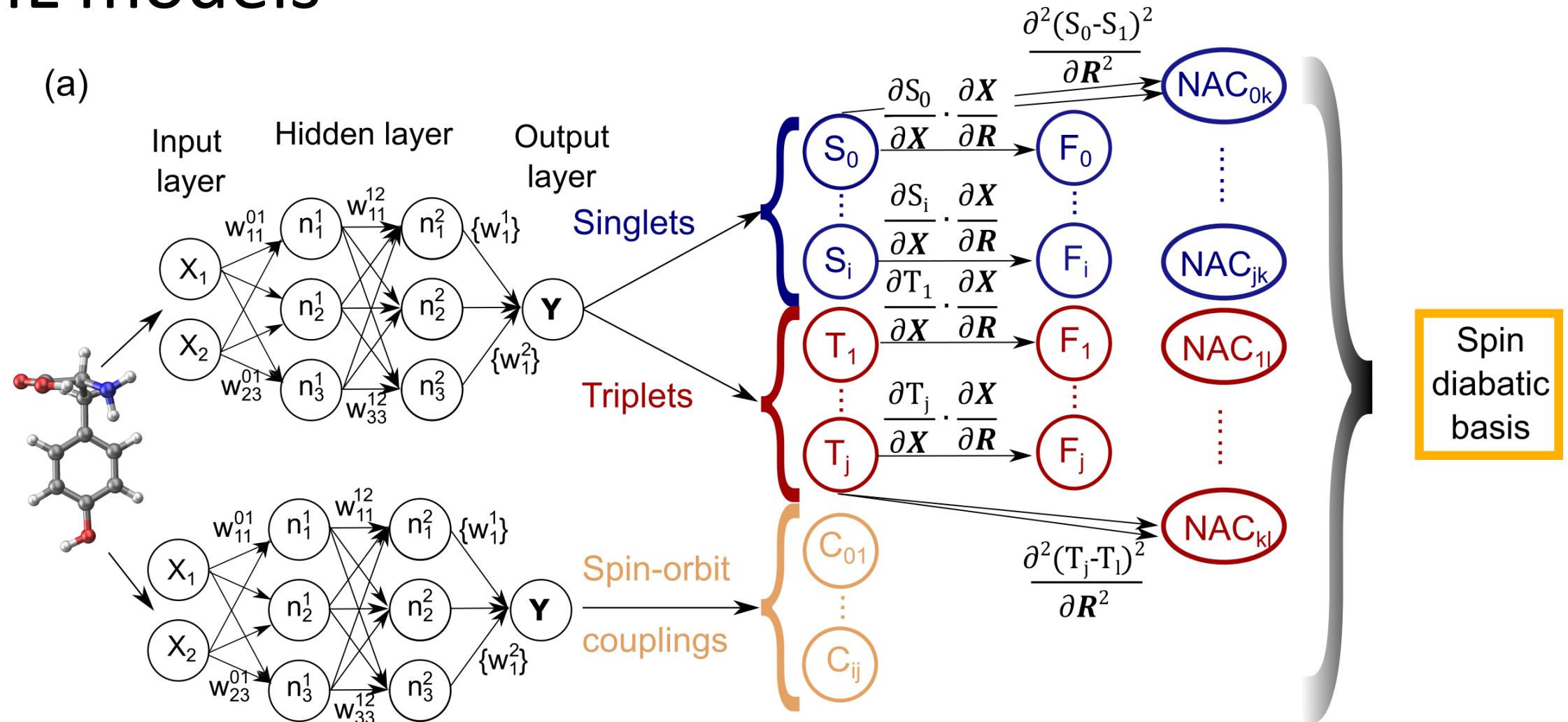
- Phase correction
- Initial ~16,500 data points
- All phase corrected (using more states)
- Additional data points were phasecorrected with ML



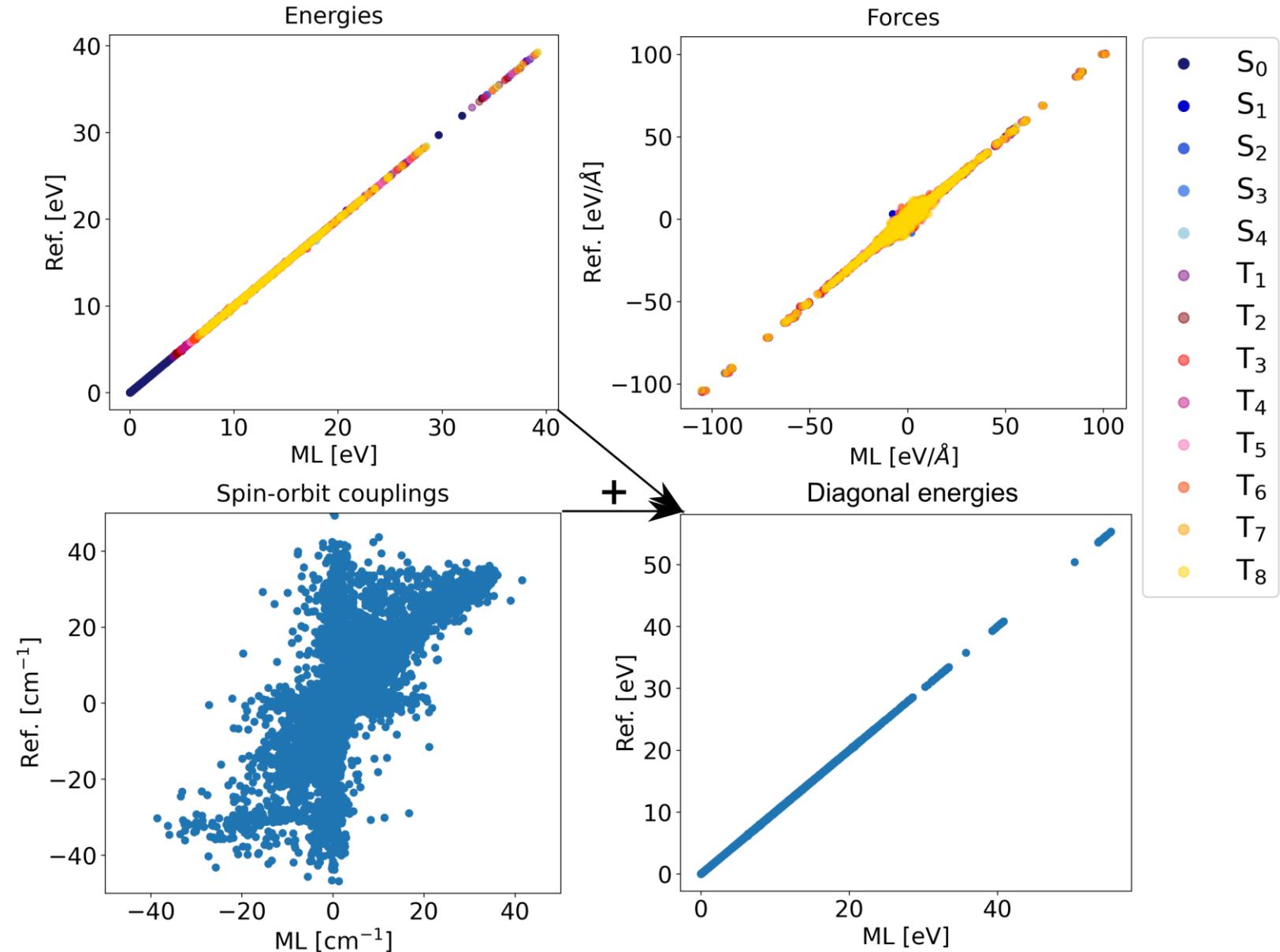
# Validation of roaming



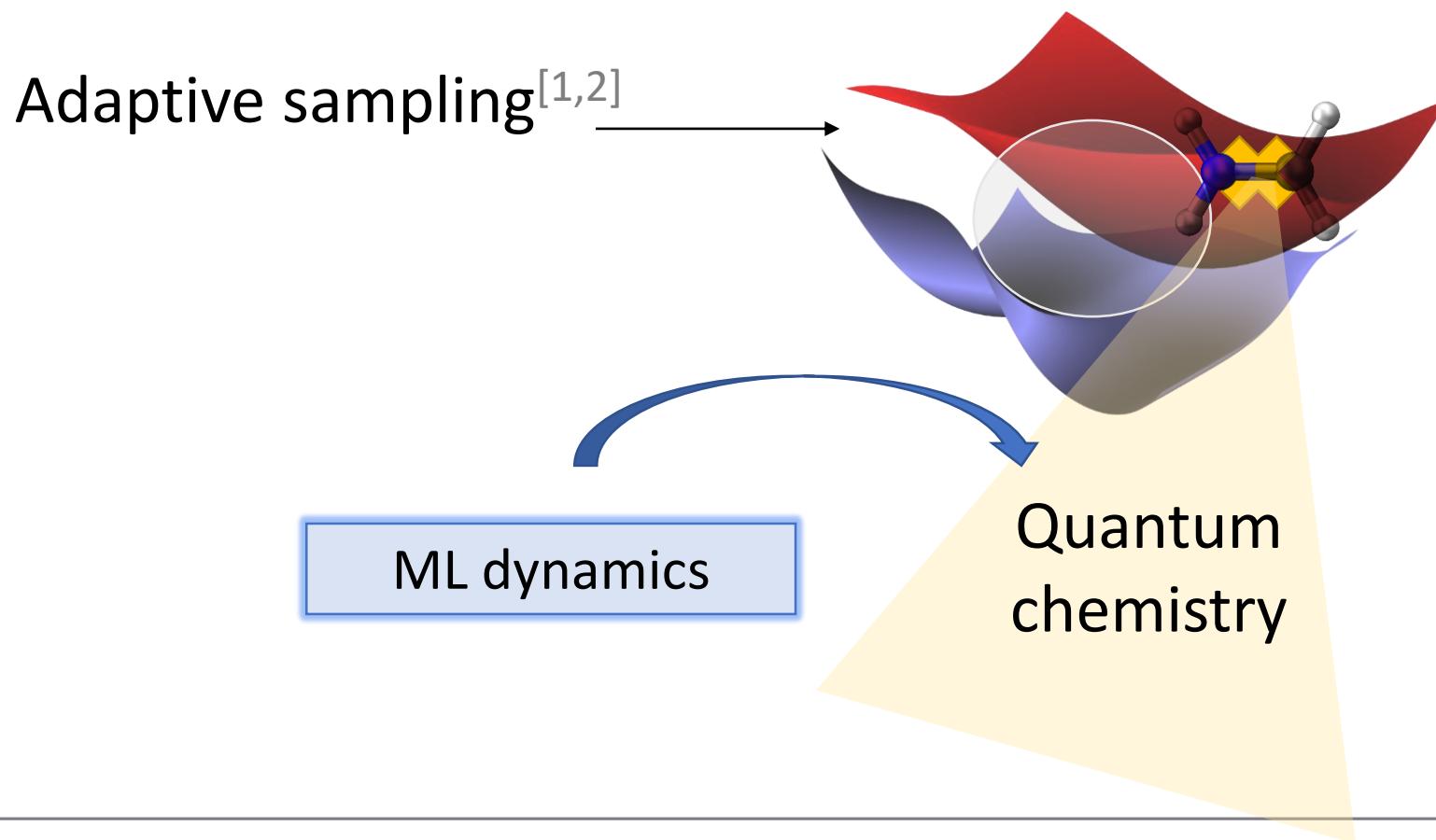
# ML models



# ML models



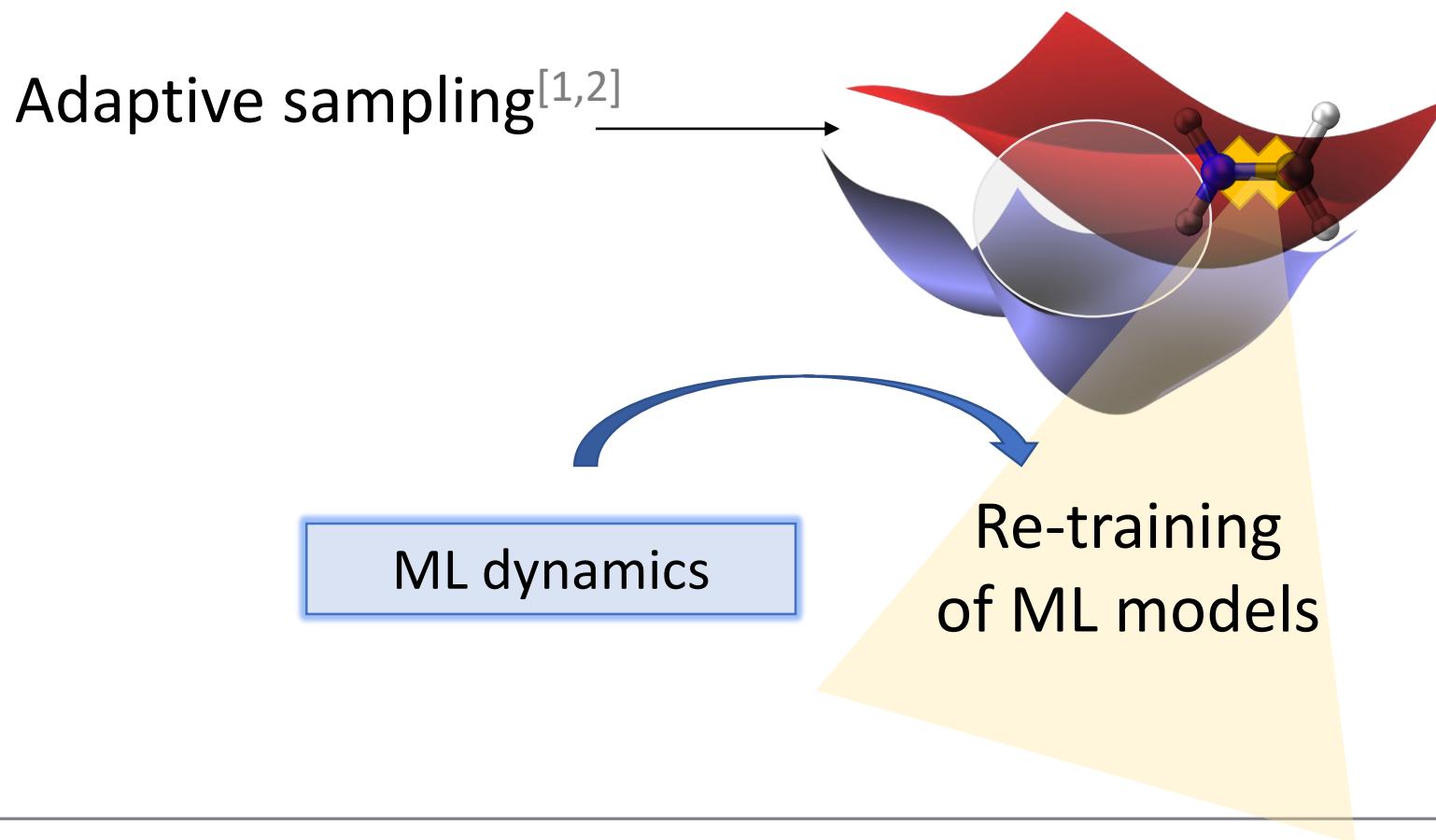
# Training Set Generation



[1] M. Gastegger, J. Behler, P. Marquetand *Chem. Sci.* **8**, 6924 (2017)

[2] J. Behler, *Int. J. Quantum Chem.* **115**, 1032–1050 (2015)

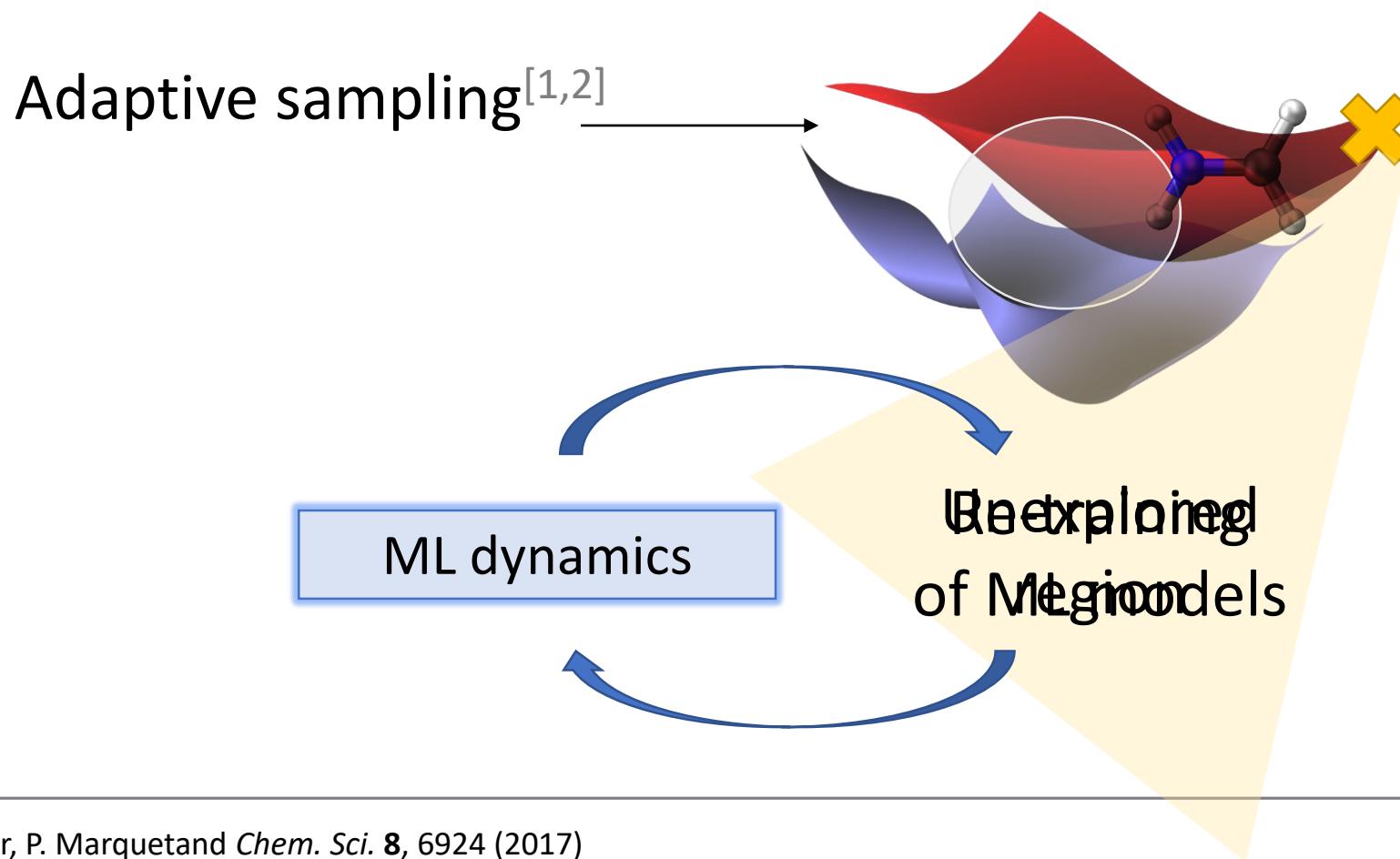
# Training Set Generation



[1] M. Gastegger, J. Behler, P. Marquetand *Chem. Sci.* **8**, 6924 (2017)

[2] J. Behler, *Int. J. Quantum Chem.* **115**, 1032–1050 (2015)

# Training Set Generation

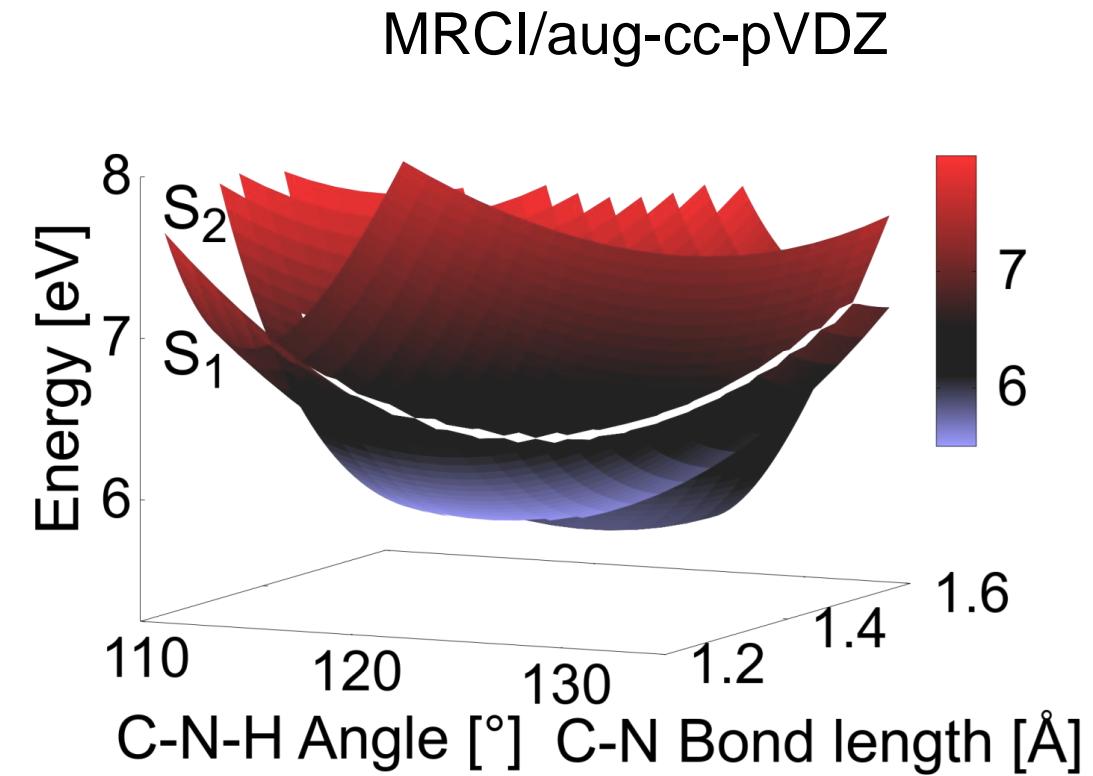
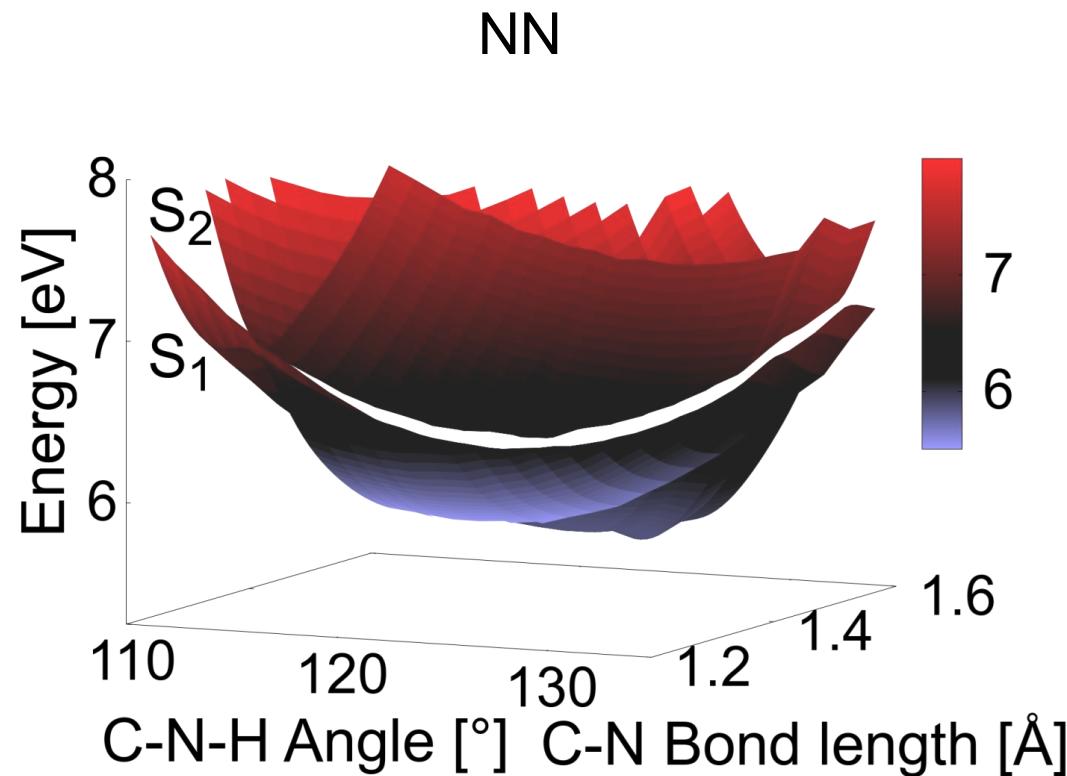


[1] M. Gastegger, J. Behler, P. Marquetand *Chem. Sci.* **8**, 6924 (2017)

[2] J. Behler, *Int. J. Quantum Chem.* **115**, 1032–1050 (2015)

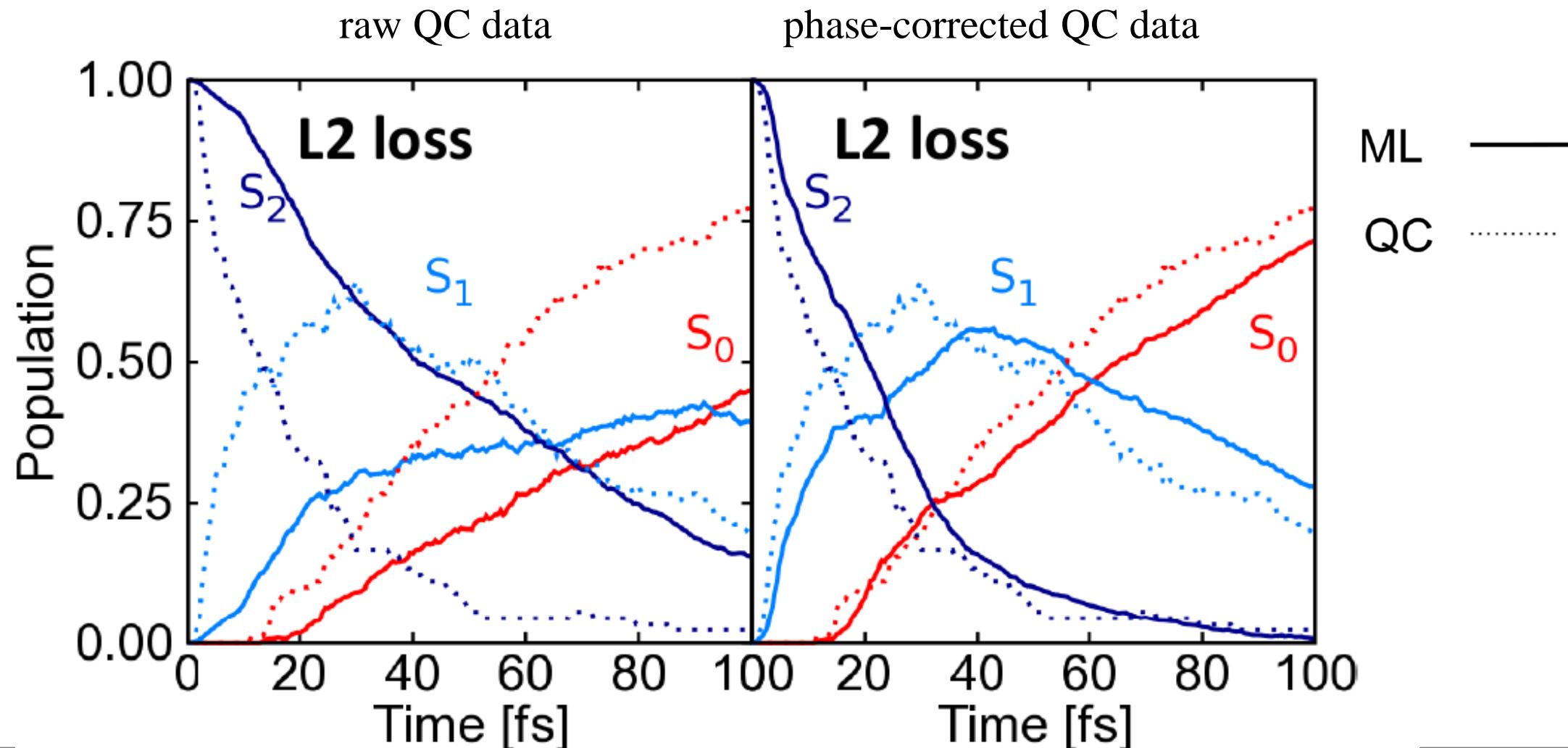
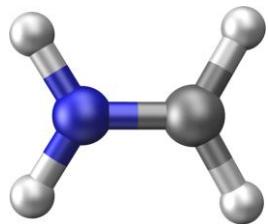
# NNs for excited states

Methylenimmonium ( $\text{CH}_2\text{NH}_2^+$ )

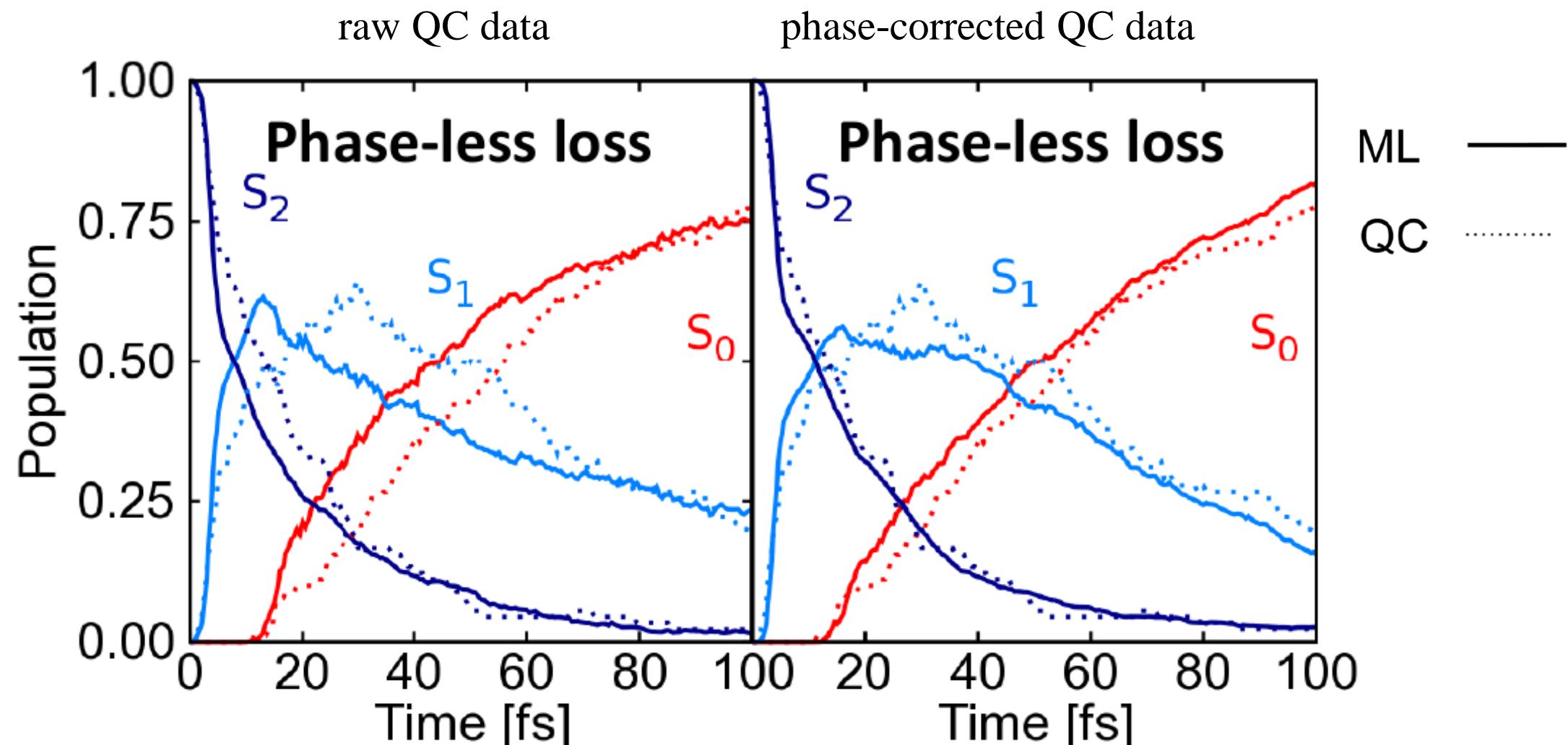


Scan around conical intersection seam

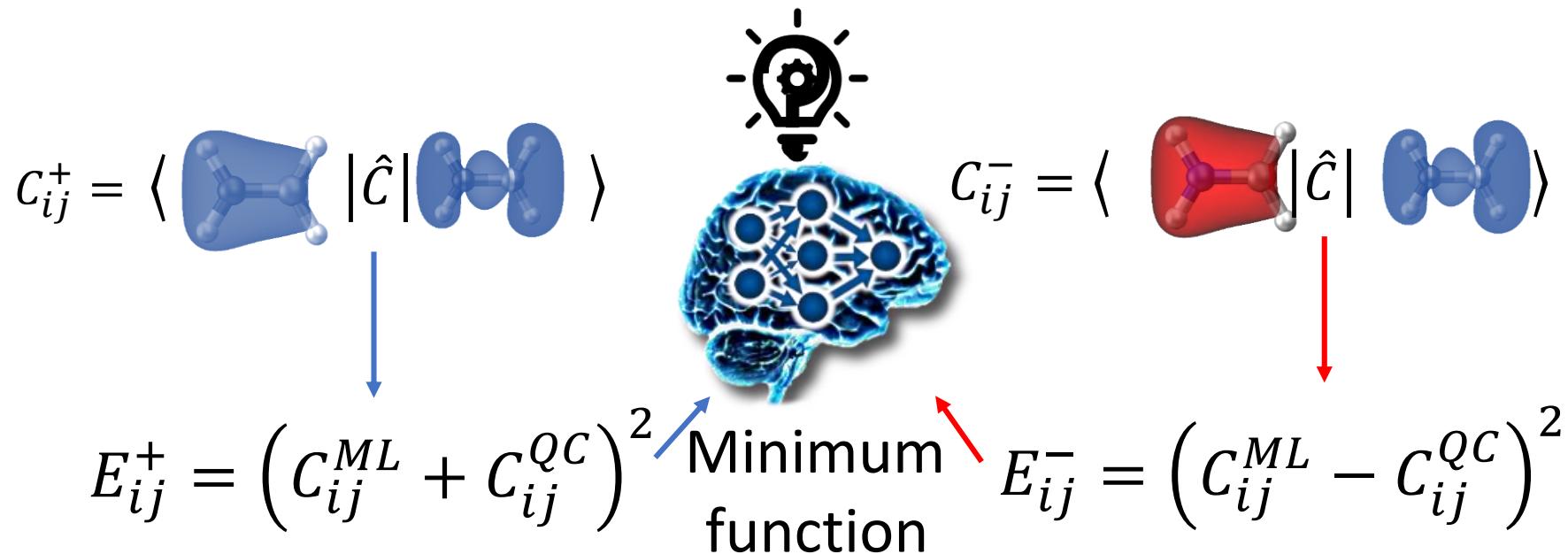
# SchNarc – phaseless loss



# SchNarc – phaseless loss

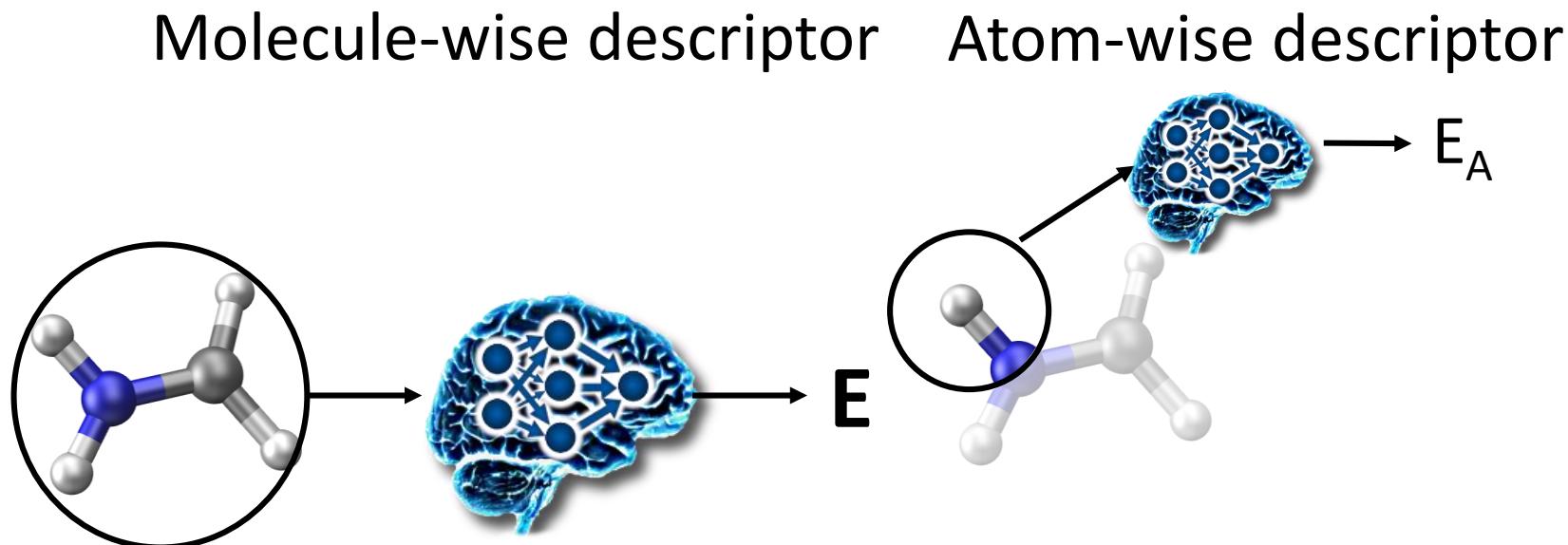


# Phase-free training algorithm



- Training independent of sign of the couplings
- Training on raw data possible
- Reduction of cost of training set generation

# Representation of molecules



# Fitting dipole moment vectors

- Absorption proportional to oscillator strength

$$f^{osc} = \frac{2}{3} \Delta E_{ij} |\mu_{ij}|^2$$

- Charge model for fitting

What we learn

$$\mu_i = \sum_a^{N_a} q_{i,a} r_a^{CM}$$

What the model builds without knowing

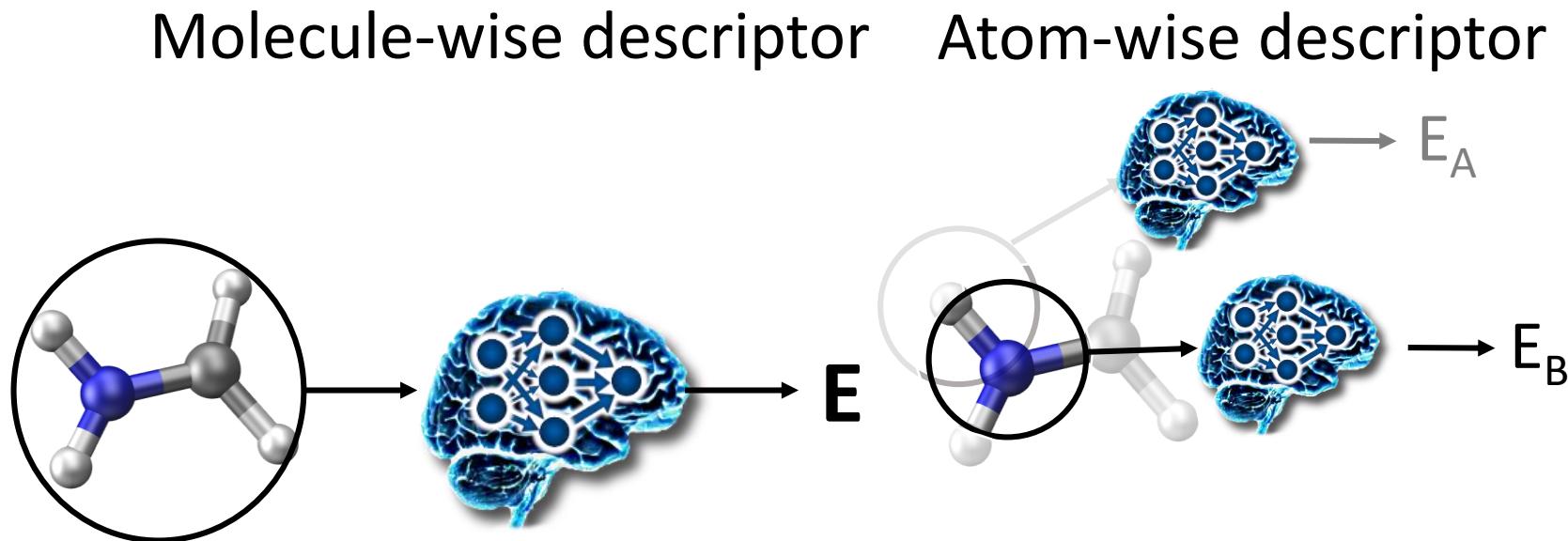
$$\mu_{ij} = \sum_a^{N_a} q_{ij,a} r_a^{CM}$$

- $q_{i,a}$  ... atomic charge of state i
- $r_a^{CM}$  ... distance of atom a to the center of mass (CM)

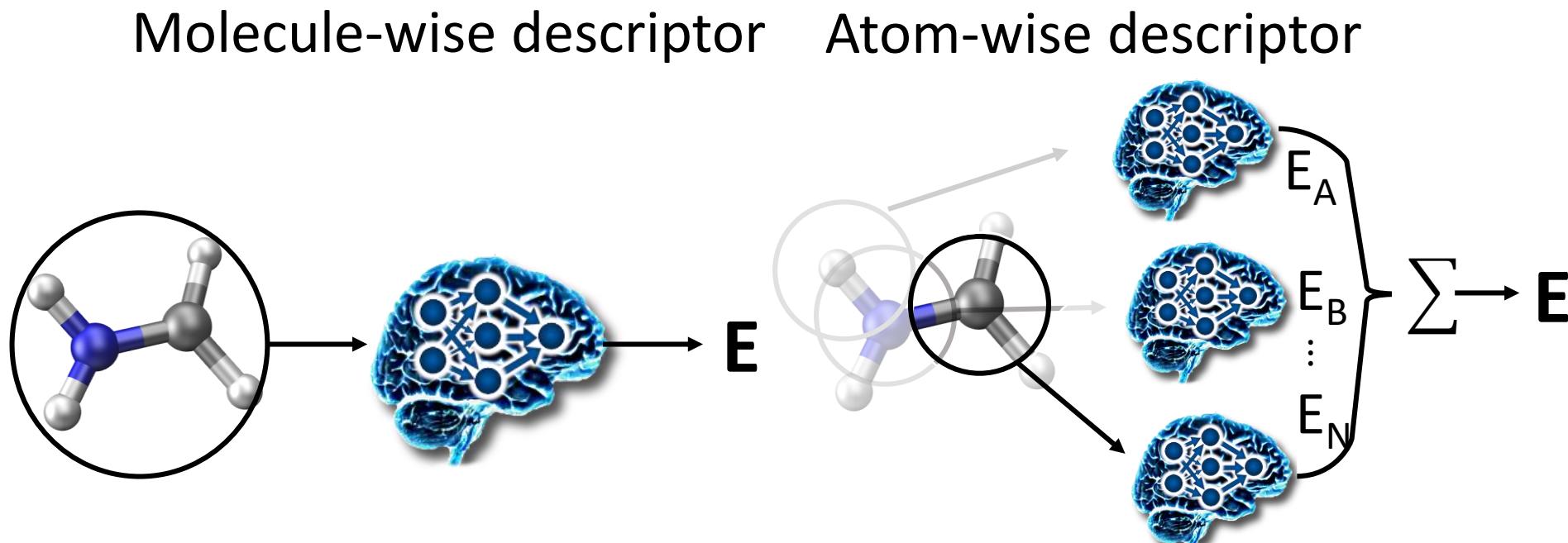
[1] Michael Gastegger, Jörg Behler, Philipp Marquetand, Chem. Sci. **8**, 6924-6935 (2017)

[2] Julia Westermayr, Philipp Marquetand, J. Chem. Phys. **153** (15), 154112 (2020)

# Representation of molecules



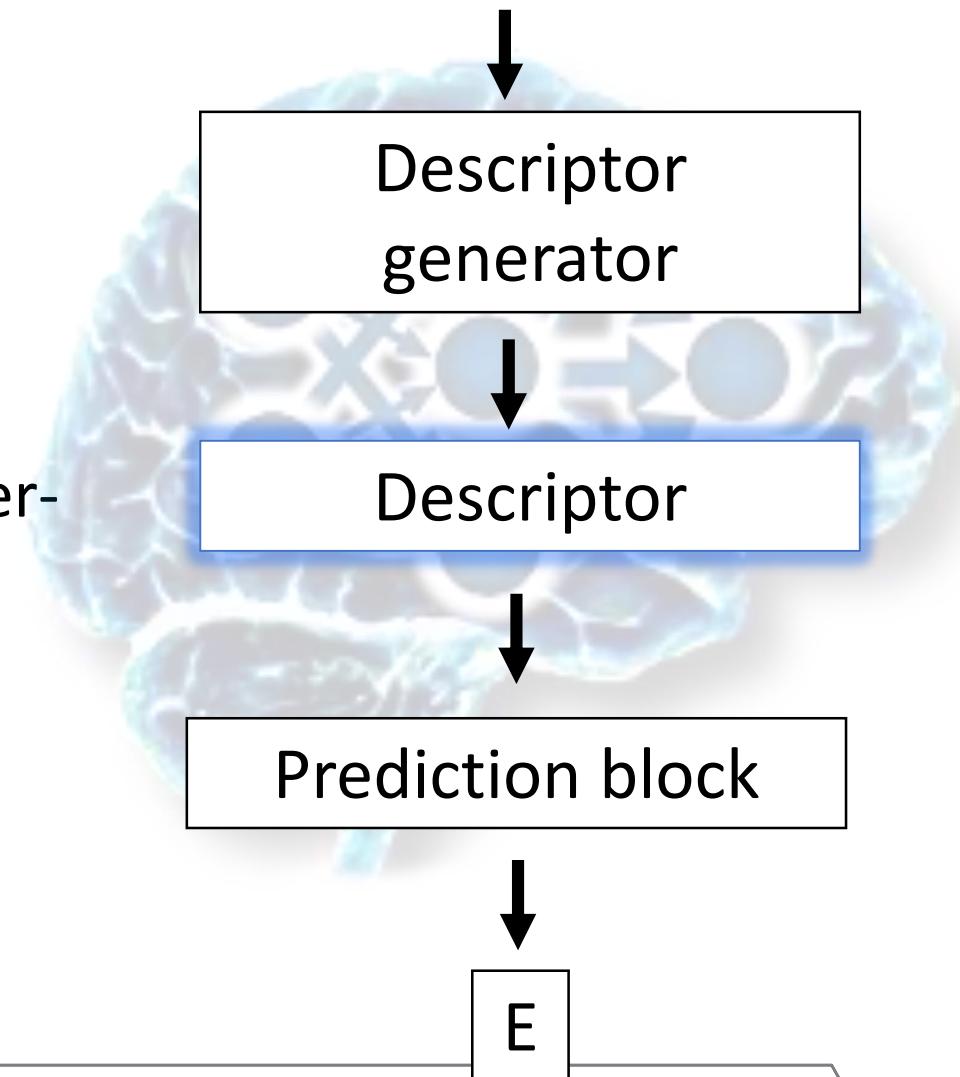
# Representation of molecules



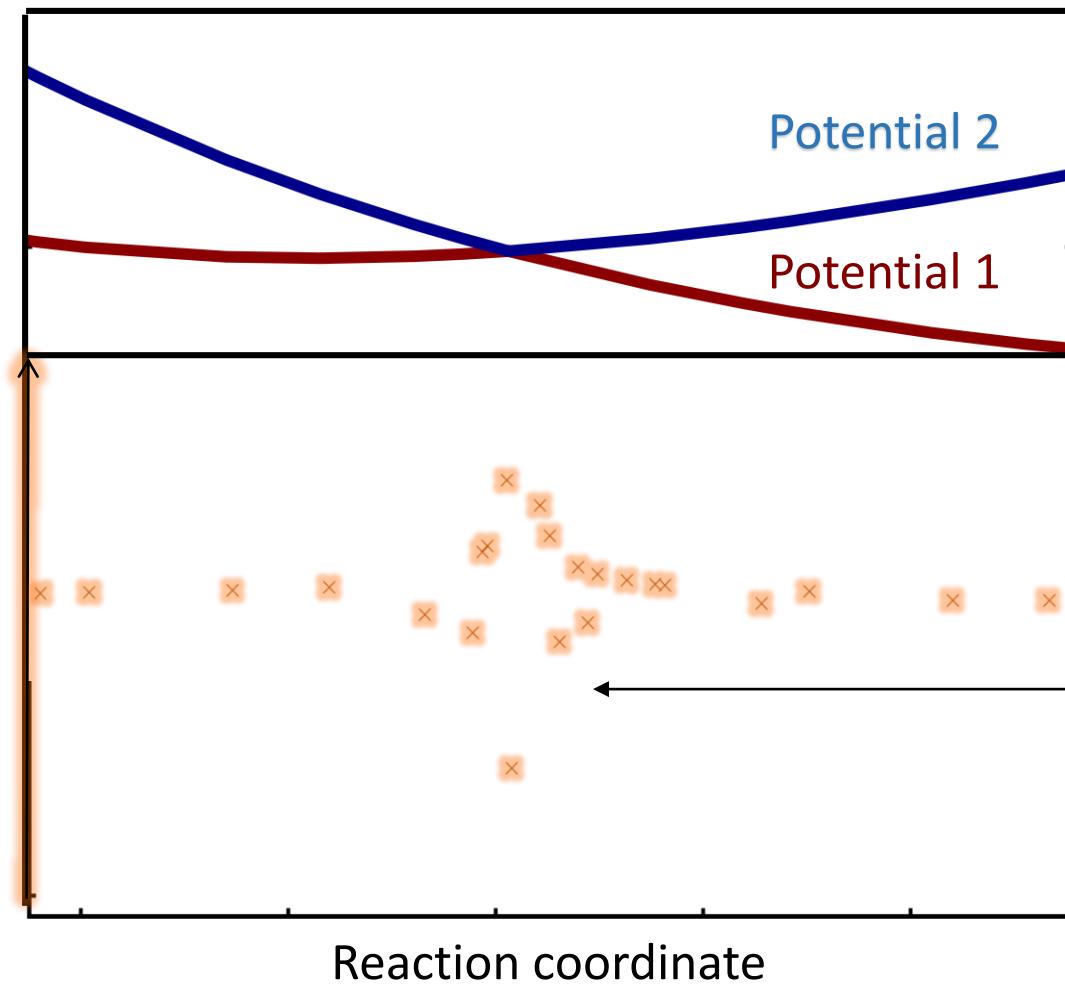
- Only for molecules of same size
- For molecules of any size and atom type
- Atomic properties

# SchNet $(Z_A \dots Z_N), (r_A \dots r_N)$

- Deep continuous-filter convolutional layer neural network
- Learns molecular representation based on interatomic distances and atom types
- Originally developed for ground-state properties



# Phase correction<sup>[1,2]</sup>



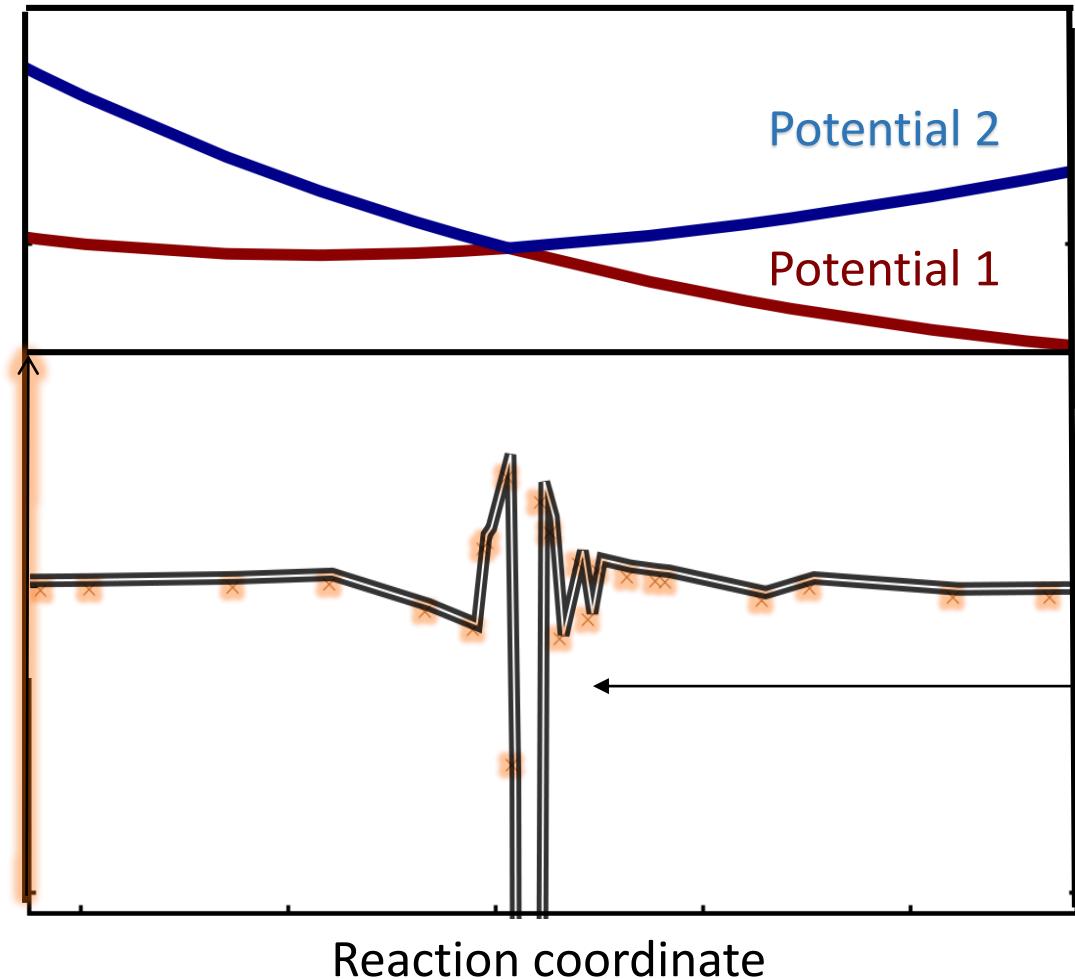
[1] J. Westermayr, et al. *Chem. Sci.*, 10, 8100-8107 (2019).

[3] A. V. Akimov, *J. Phys. Chem. Lett.*, 9(20), 6096-6102 (2018).

# Phase correction<sup>[1,2]</sup>

Machine learning potentials: smooth functions!

Solution: make couplings smooth



Couplings:  
quantum  
chemistry  
output

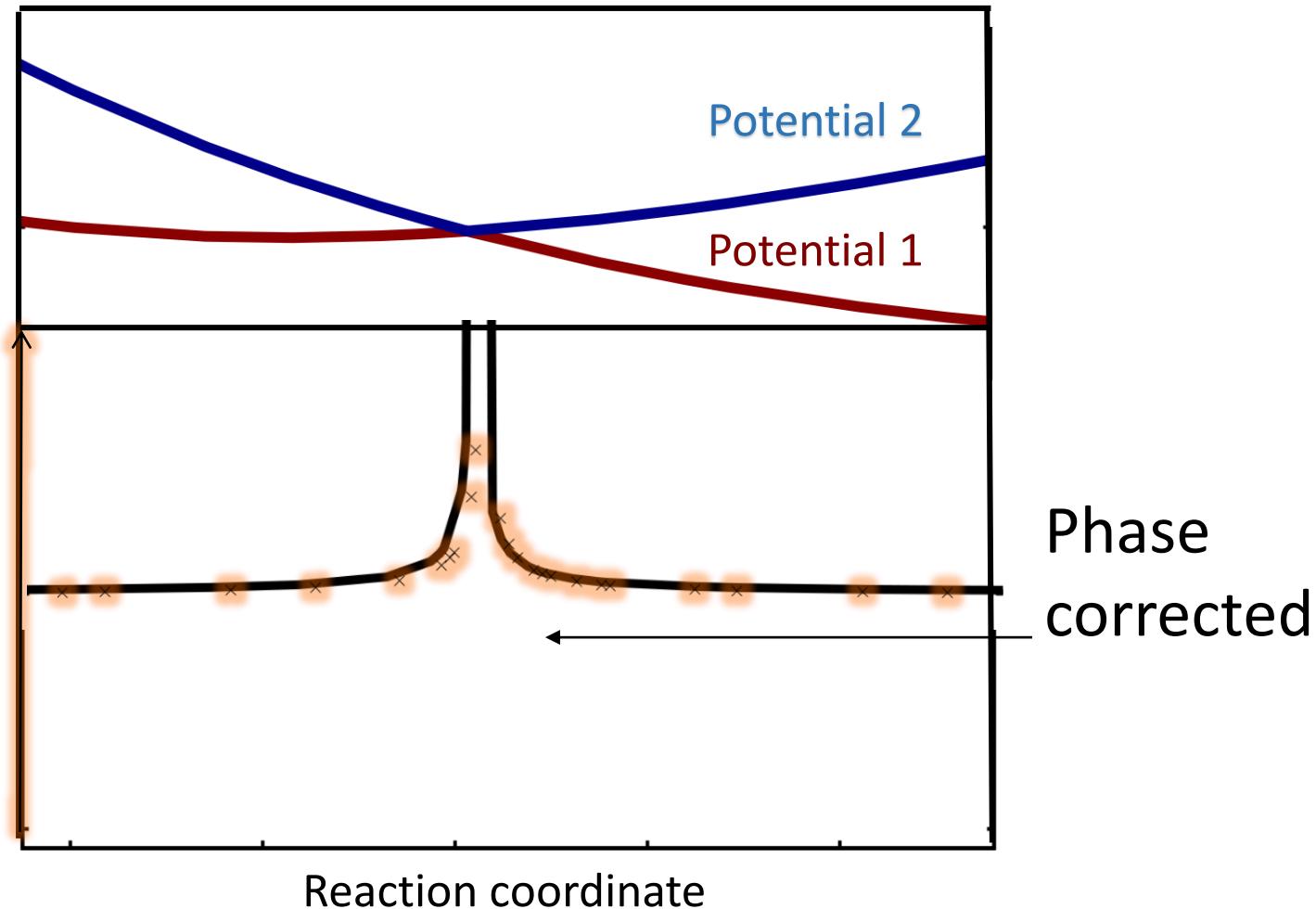
[1] J. Westermayr, et al. *Chem. Sci.*, 10, 8100-8107 (2019).

[3] A. V. Akimov, *J. Phys. Chem. Lett.*, 9(20), 6096-6102 (2018).

# Phase correction<sup>[1,2]</sup>

Phase correction:

- 1) Pre-processing of data via wave function overlap computations
- 2) Phase-free training algorithm



[1] J. Westermayr, et al. *Chem. Sci.*, 10, 8100-8107 (2019).

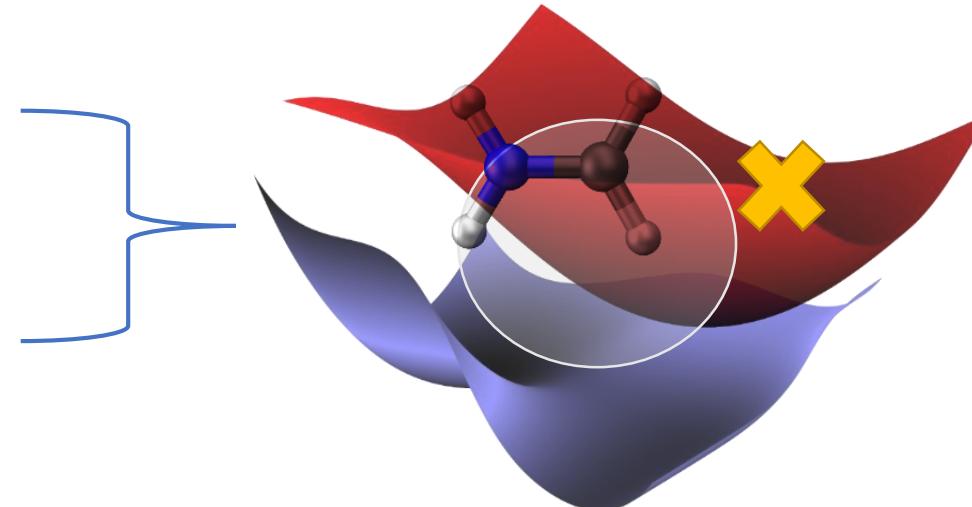
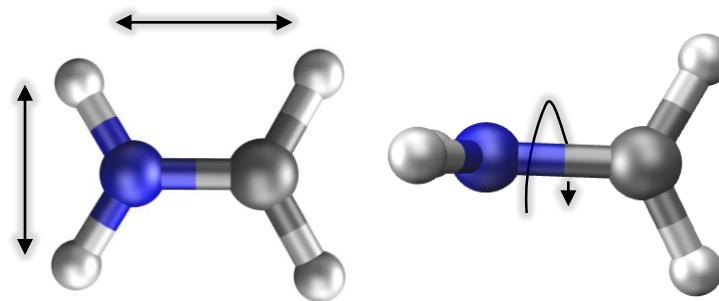
[3] A. V. Akimov, *J. Phys. Chem. Lett.*, 9(20), 6096-6102 (2018).

# Training Set Generation

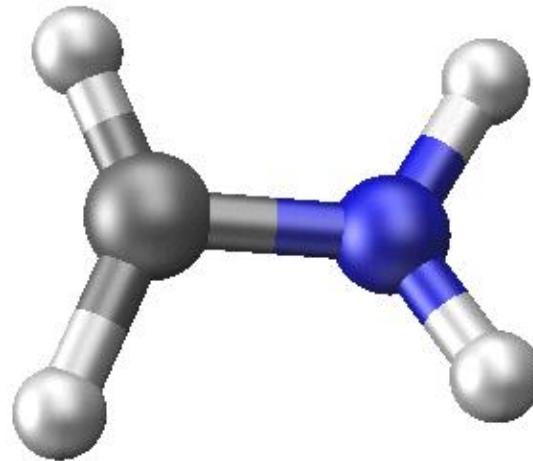
1

## Initial sampling

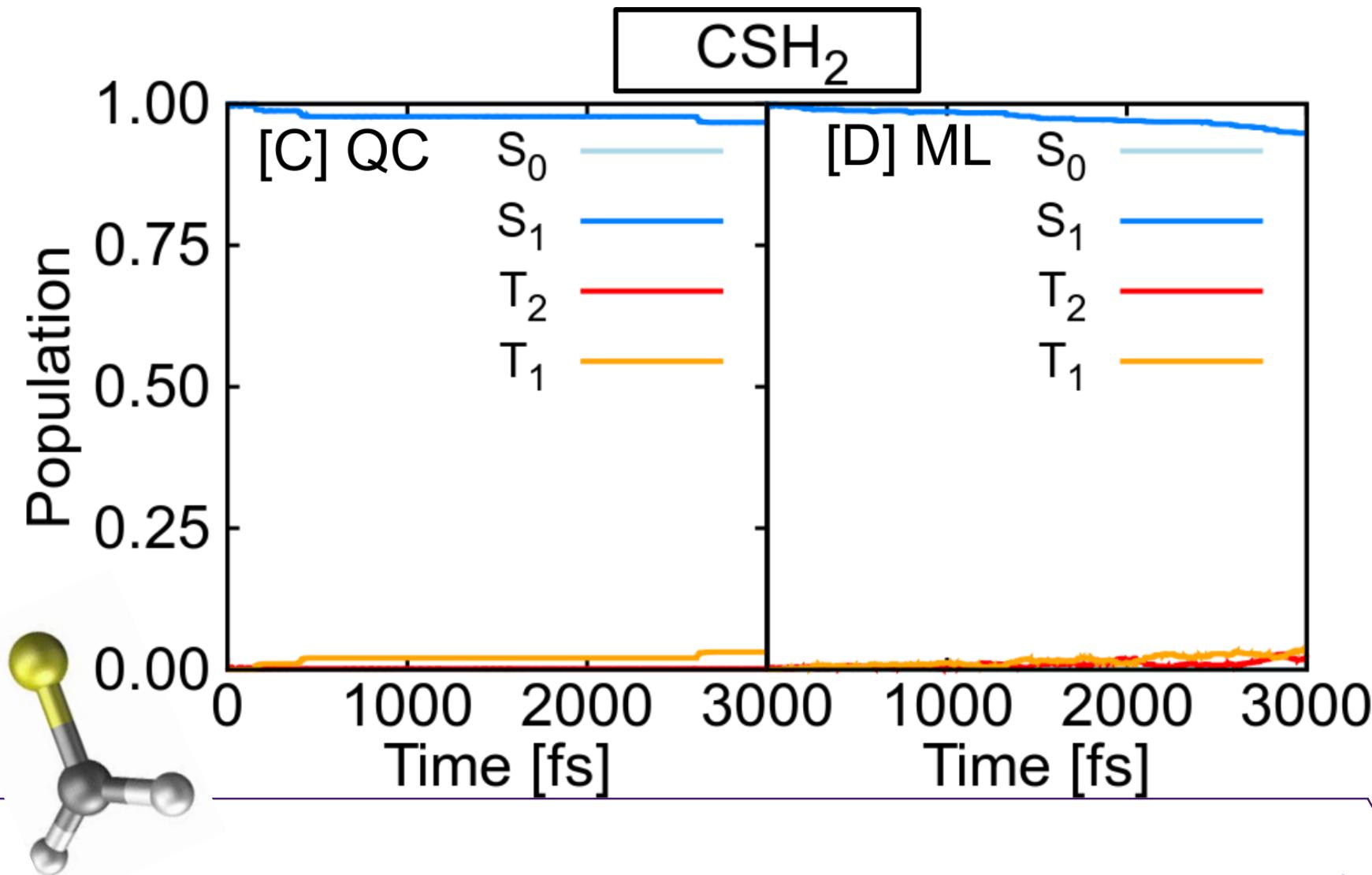
- Scan along normal modes
- Scan along dihedral angle



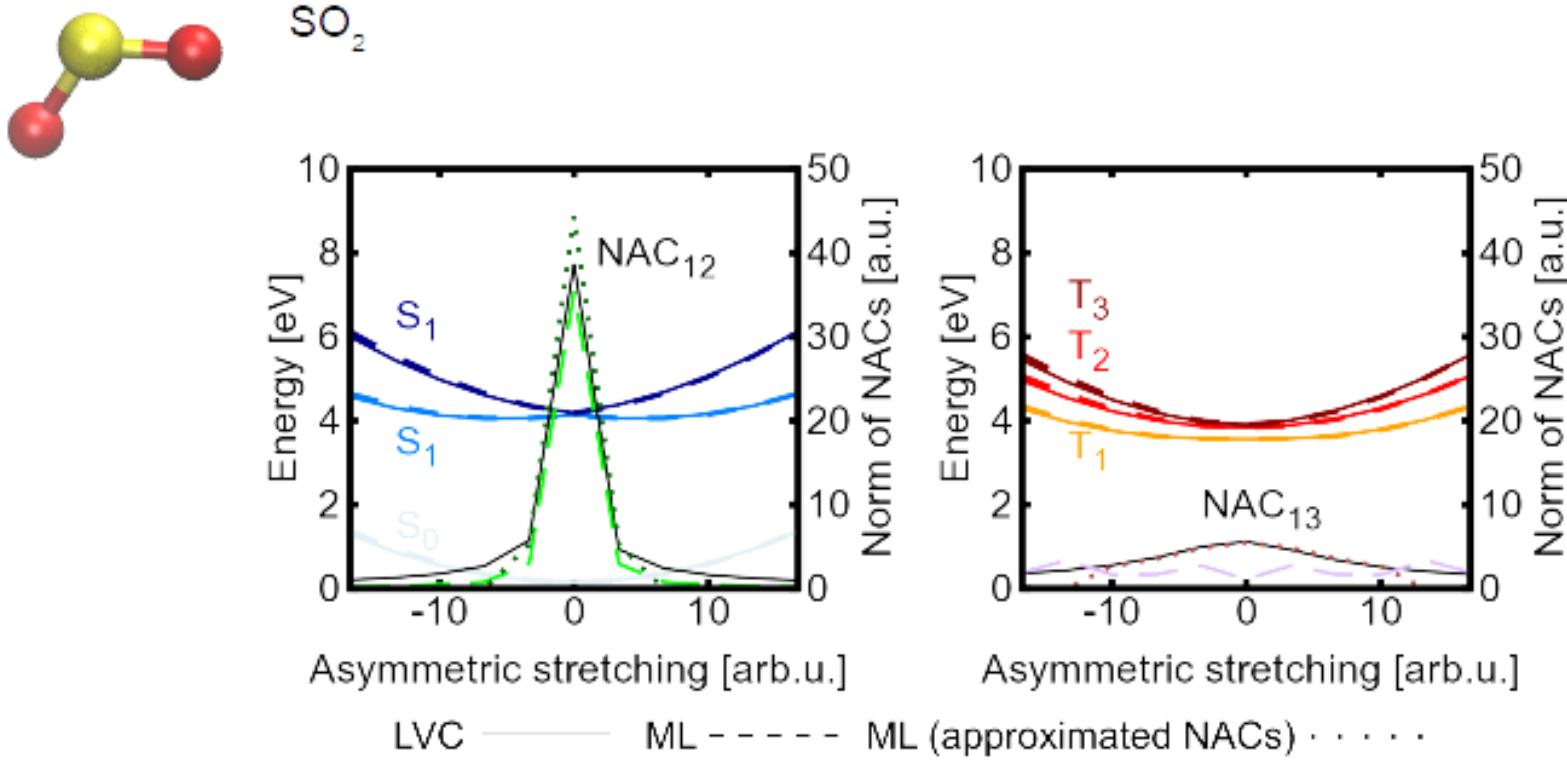
# Training Set Generation



# Slow population transfer: $\text{CSH}_2$

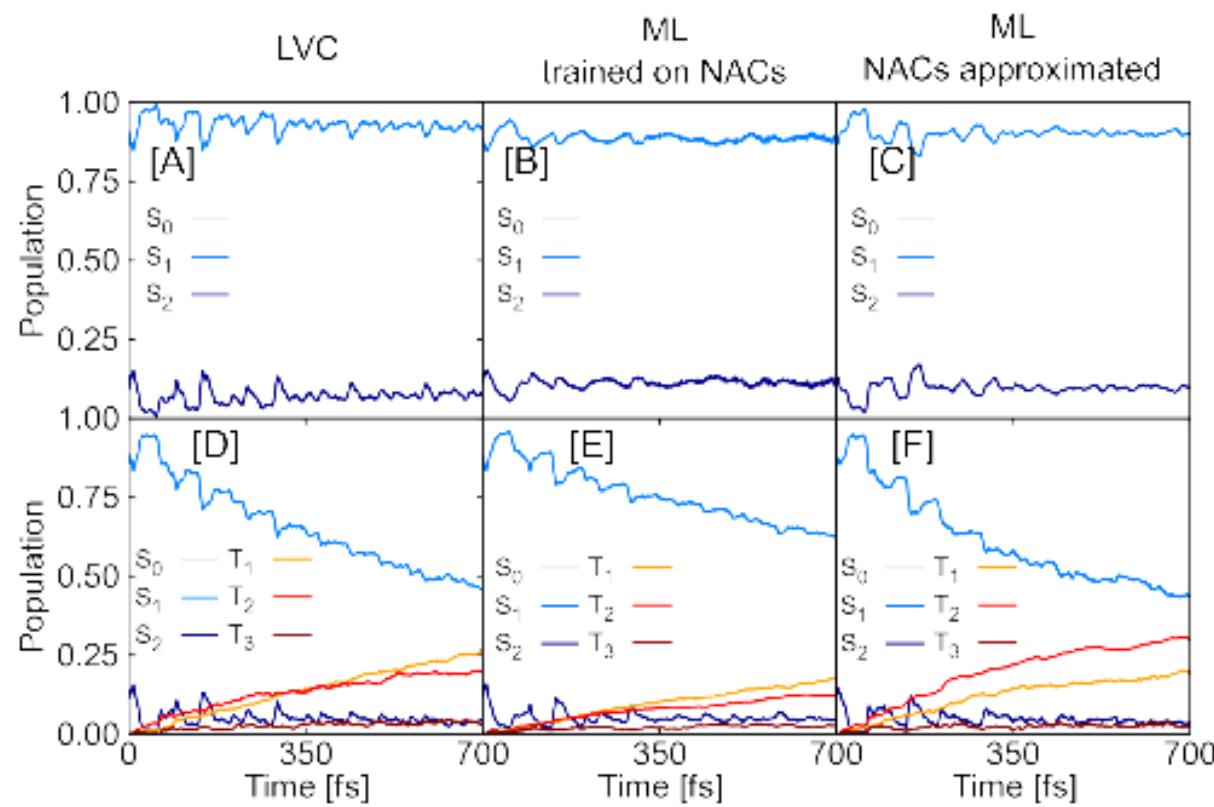


## SchNarc – approximated NACs



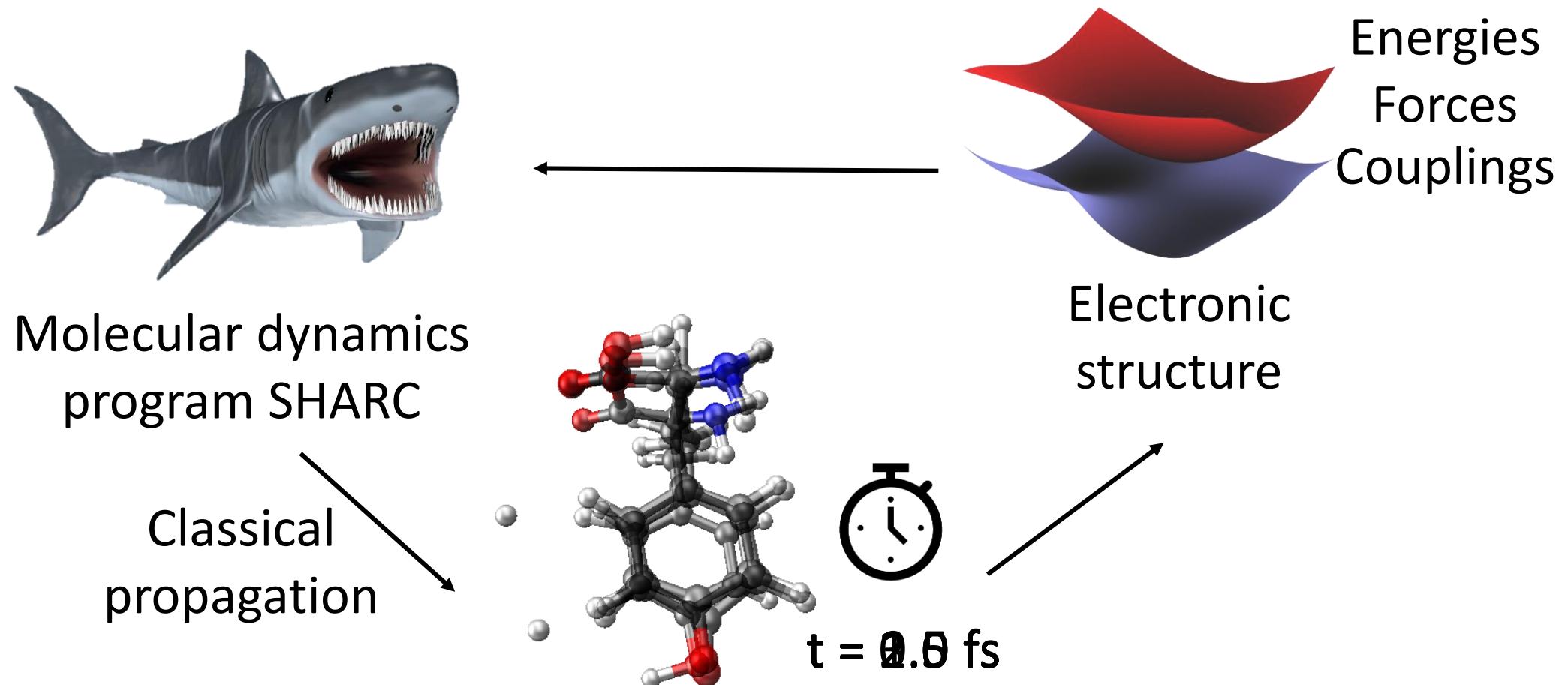
J. Westermayr, M. Gastegger, P. Marquetand, J. Phys. Chem. Lett., 11, 3828-3834 (2020).

# SchNarc – approximated NACs

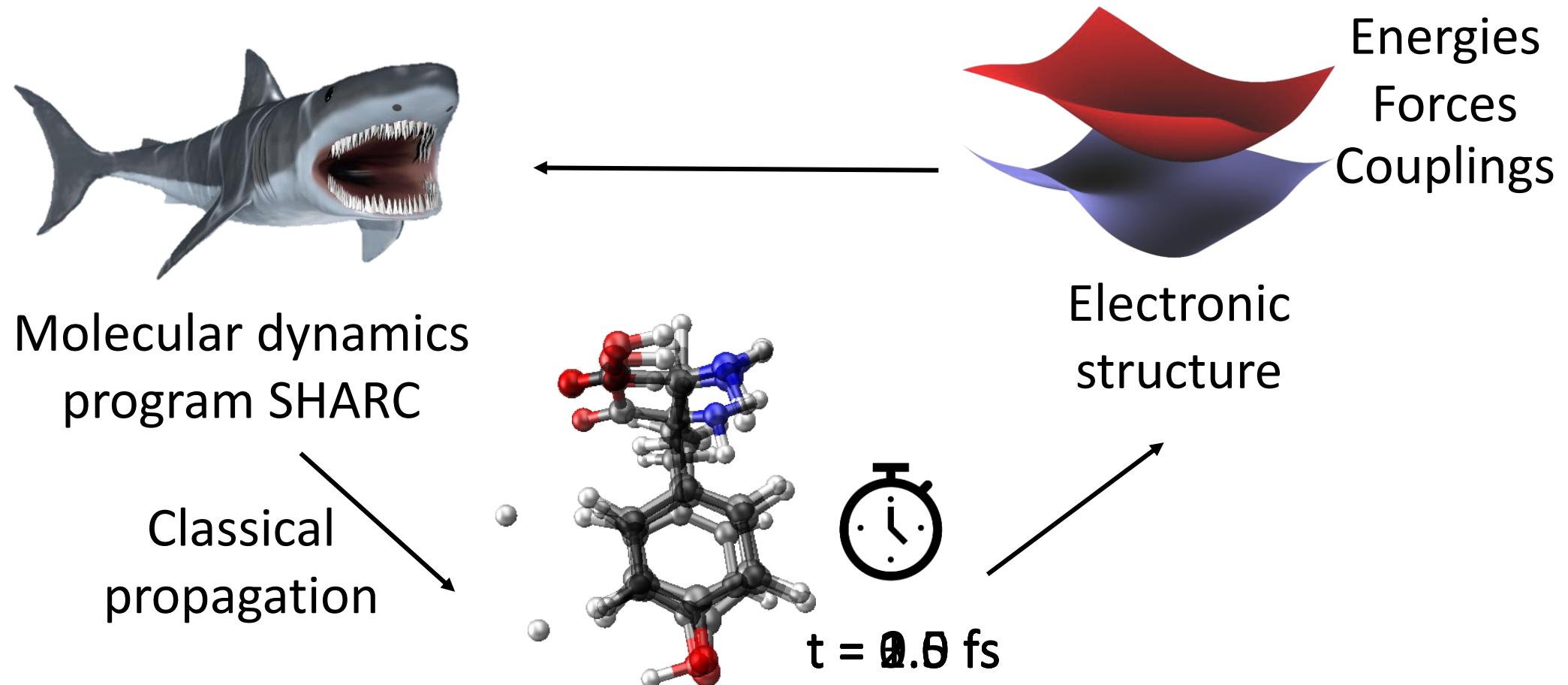


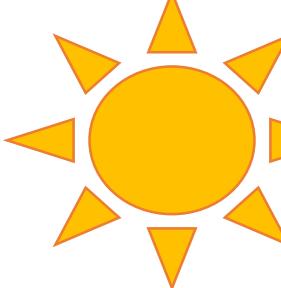
J. Westermayr, M. Gastegger, P. Marquetand, J. Phys. Chem. Lett., 11, 3828-3834 (2020).

# Surface hopping molecular dynamics

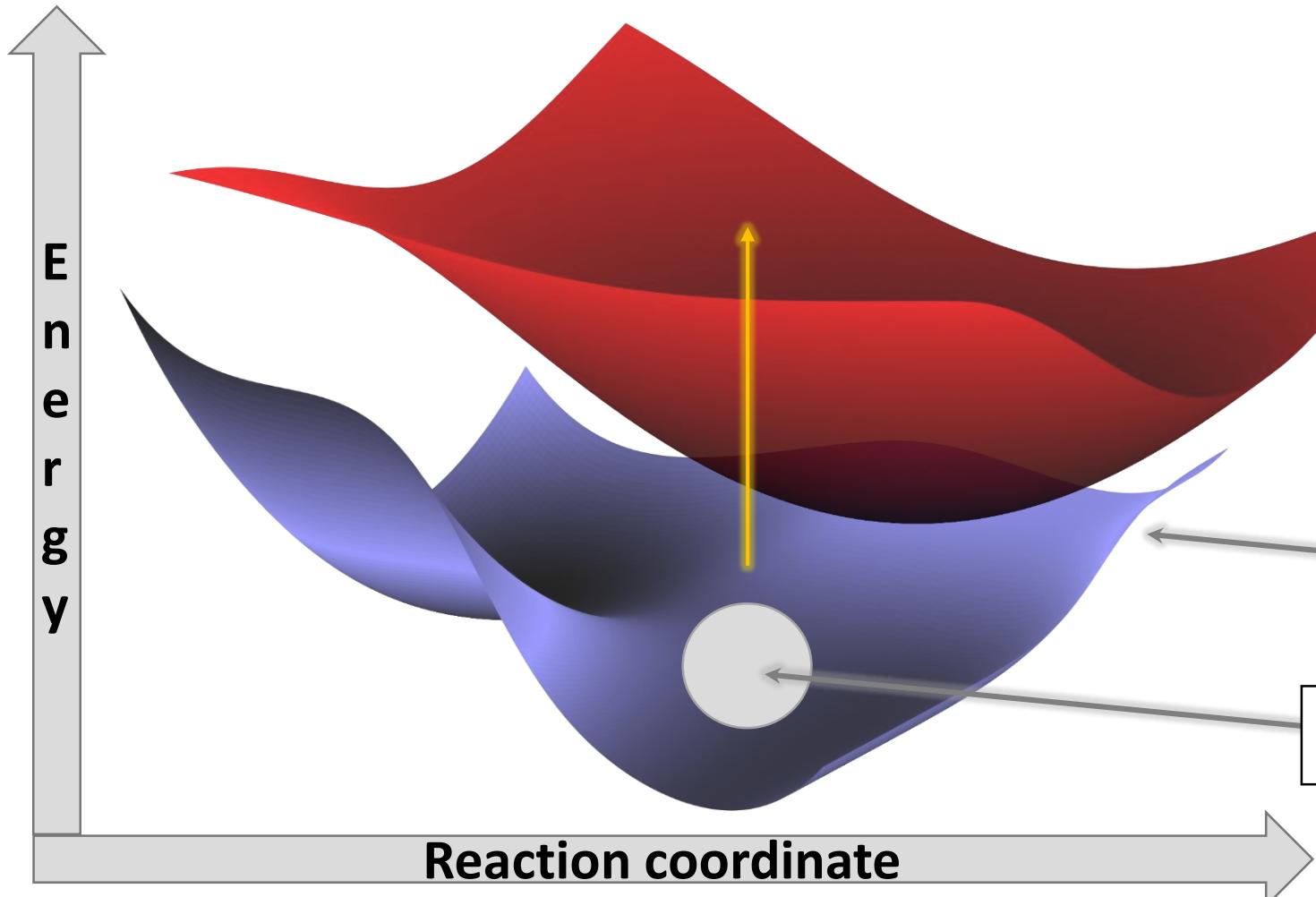


# Surface hopping molecular dynamics





# Surface hopping molecular dynamics



- Mixed quantum-classical
- Nonadiabatic transitions

Quantum chemical potential energy surfaces

Molecule: moves classically