COMPUTING CATALYTIC REACTION TIMES AND PATHS WITH MACHINE LEARNING AND RARE EVENTS SAMPLING METHODS

THOMAS PIGEON

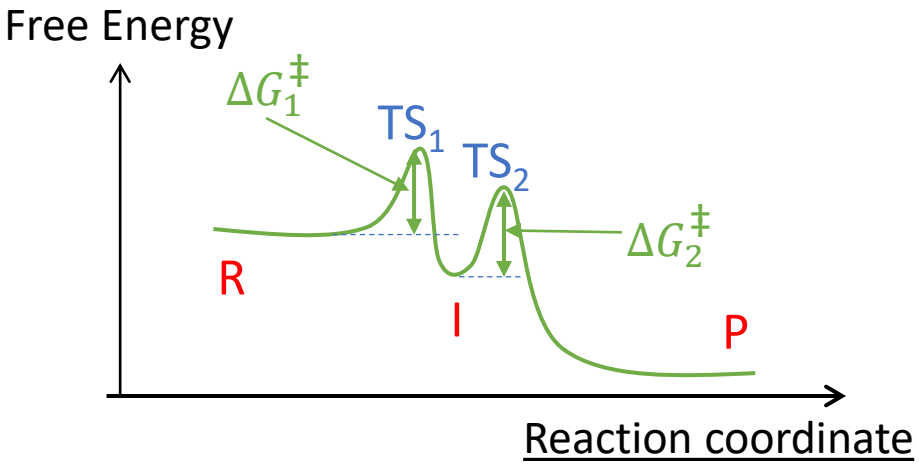
Pascal Raybaud[1], Manuel Corral Valero[1],
Ani Anciaux-Sedrakian[2], Maxime Moreaud[2]
Gabriel Stoltz[3], Tony Lelièvre[3]

[1] IFPEN: Catalysis, Biocatalysis and Separation
[2] IFPEN: Digital Science and Technology
[3] CERMICS, Ecole des Ponts ParisTech, and Equipe-projet MATHERIALS, Inria Paris,

INTRODUCTION: CHARACTERIZING REACTION MECHANISMS

Targets : compute reaction rates (average times)

identify reaction paths (how atoms rearranges themselves)



Different methods exist:

- Transition State Theory (TST)
  → Need to estimate free energy
- **Directly from the time evolution of the system ?**
  → **Need Molecular Dynamics (MD)**

INTRODUCTION: STANDARD MOLECULAR DYNAMICS

Simulates the dynamic of the system by adding a thermostat to newton equations of motion

ex. Langevin formalism[1]

NVE ensemble NVT ensemble

$$\begin{cases} dq_t = M^{-1} p_t dt \\ dp_t = -\nabla V(q_t) dt - \gamma p_t dt + \sqrt{2\gamma M k_B T} dW_t \end{cases}$$

Preserves energy — Dissipate energy — Provides energy

Newton equation — Langevin part

Not efficient for the simulation of rare events due to high energy barriers and entropic bottlenecks

Time scales: integration time step : $\sim 10^{-15} s$ rare event rate $\sim 10^{-9} s^{-1}$ to $10^{3} s^{-1}$

MD based approaches to overcome barriers:

→ biased MD such as Metadynamics[2], Blue-Moon sampling[3] …

Dynamics is lost but we can estimate free energy

→ rare events sampling methods such as Adaptive multi-level splitting[4]

Dynamics preserved and rates can be “directly” be computed

[1] P. Langevin P. (1908), Comptes-Rendus de l'Académie des Sciences, 146, 530

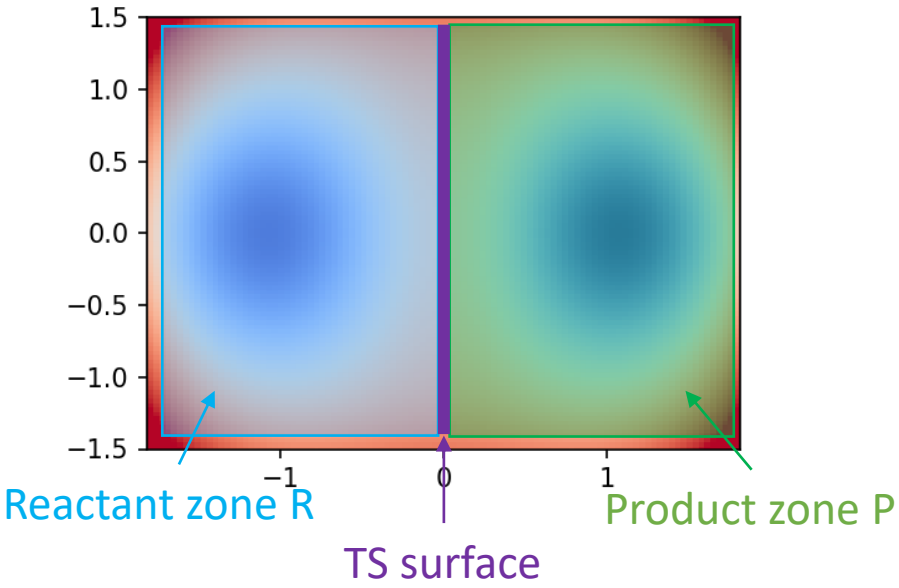
[2] A. Laio, M. Parrinello, (2002) PNAS, 99, 20, 12562

[3] E. A. Carter, G. Ciccotti, J. T. Hynes, R. Kapral, (1989). Chem. Phys. Lett., 156, 5, 472

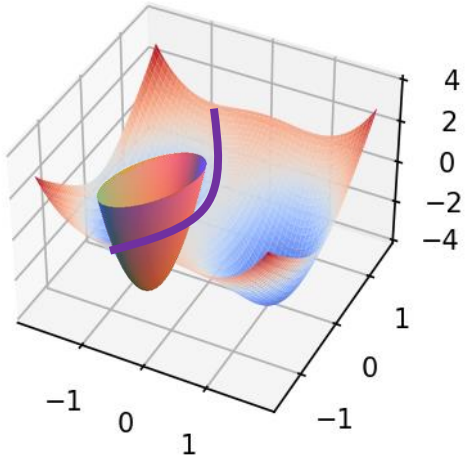
[4] Cérou, F., & Guyader, A. (2007) Stoch. Anal. Appl., 25, 2, 417

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES

Transition State Theory



2-dimensional potential



Rate = probability of being in TS with respect to R × frequency of decomposition to P

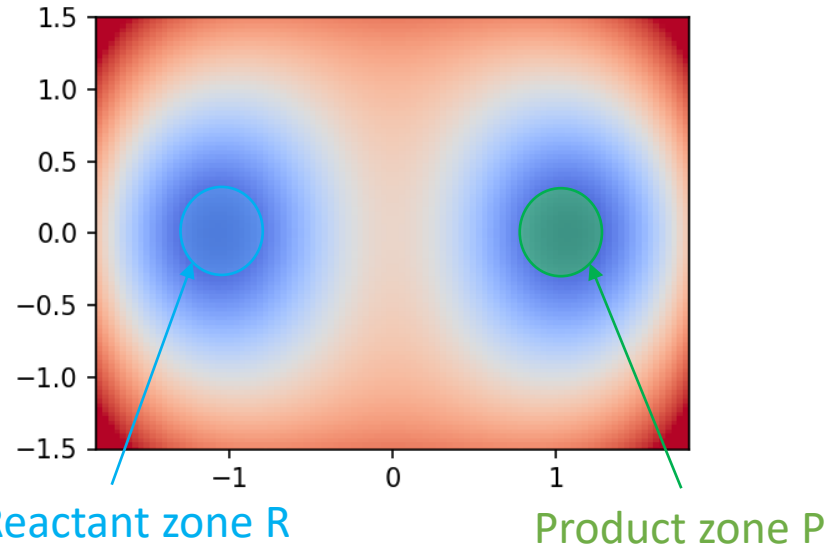
$$\boldsymbol{k}^{TST} = p(TS \mid R)\ \phi_{TS \to P} \qquad \boldsymbol{k}^{hTST} = e^{-\frac{\Delta G^{\ddagger}}{k_B T}} \frac{k_B T}{h}$$

Sensitive to the TS definition

TST overestimates rates (κ)

hTST poorly captures entropy

Hill relation



Rate = probability of reaching P before R starting from ∂R × frequency of exits of R

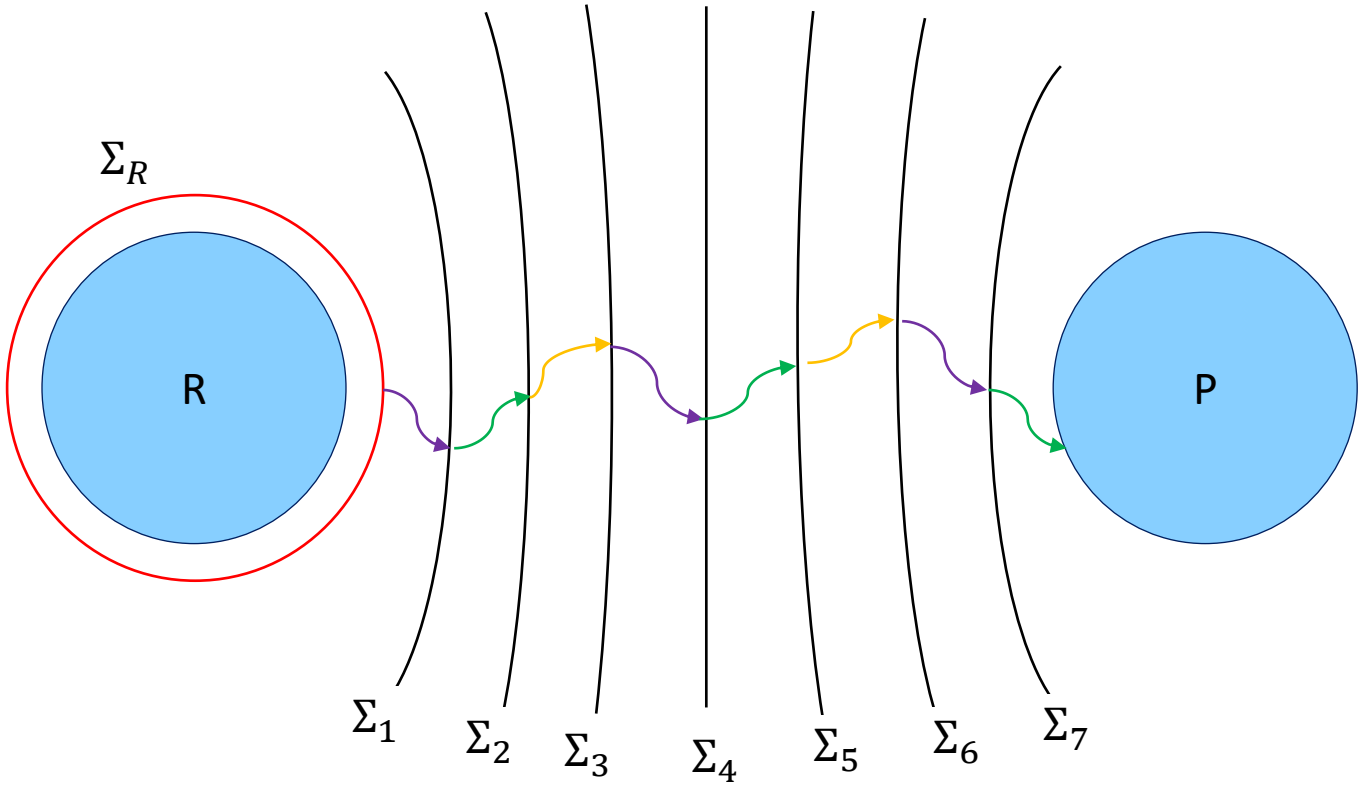
$$\boldsymbol{k}^{Hill} = p_{R \to P}(\partial R)\ \phi_R$$

Not extremely sensitive to the definition of R and P

[1] H. Eyring, (1935). J. Chem. Phys., 3, 2, 107
[2] P. Hänggi, P. Talkner, M. Borkovec, (1990) Rev. Mod. Phys., 62, 2, 251
[2] T. Hill, (2012) Free energy transduction in biology: The steady-state kinetic and thermodynamic formalism. Elsevier Science and Technology Books

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES

What is a Multilevel Splitting estimator:



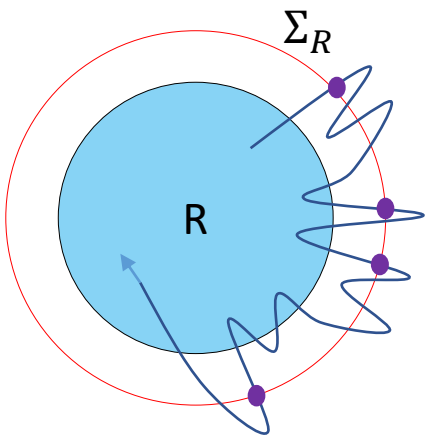
$$p_{R\to\Sigma_1}(\Sigma_R)p_{R\to\Sigma_2}(\Sigma_1)p_{R\to\Sigma_3}(\Sigma_2)p_{R\to\Sigma_4}(\Sigma_3)p_{R\to\Sigma_5}(\Sigma_4)p_{R\to\Sigma_6}(\Sigma_5)p_{R\to\Sigma_7}(\Sigma_6)p_{R\to P}(\Sigma_7) = p_{R\to P}\ (\Sigma_R)$$

How to place Σ_i and compute $p_{\mathrm{R}\to\Sigma_{i+1}}(\Sigma_i)$?

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES

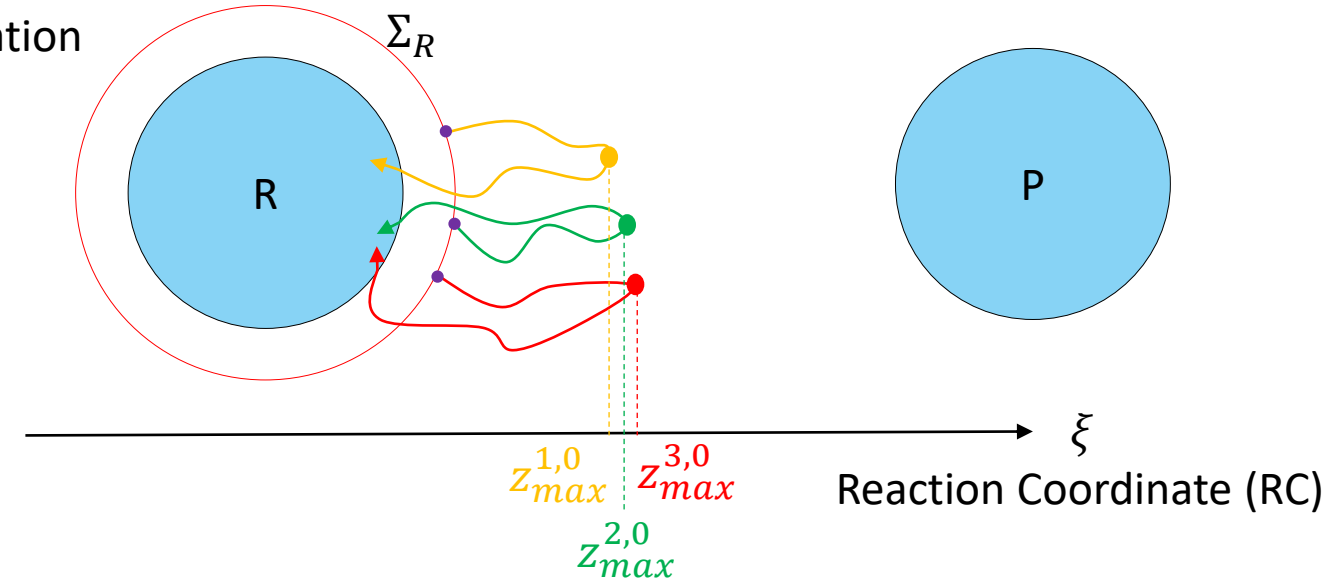
- AMS aims at estimating $p_{\Sigma \to P}$[1,2]. It can be split in 3 steps:
  1. Generating initial conditions on Σ and estimate $t_{R-\Sigma-R} = \frac{1}{\phi_R}$
  2. Initialize N replicas by running an unbiased dynamics until it reaches R or P. Set p = 1. Classify all the replicas by increasing ξ_{max}.
  3. Apply the AMS loop until all replicas have reached P.

1. Initial conditions and flux



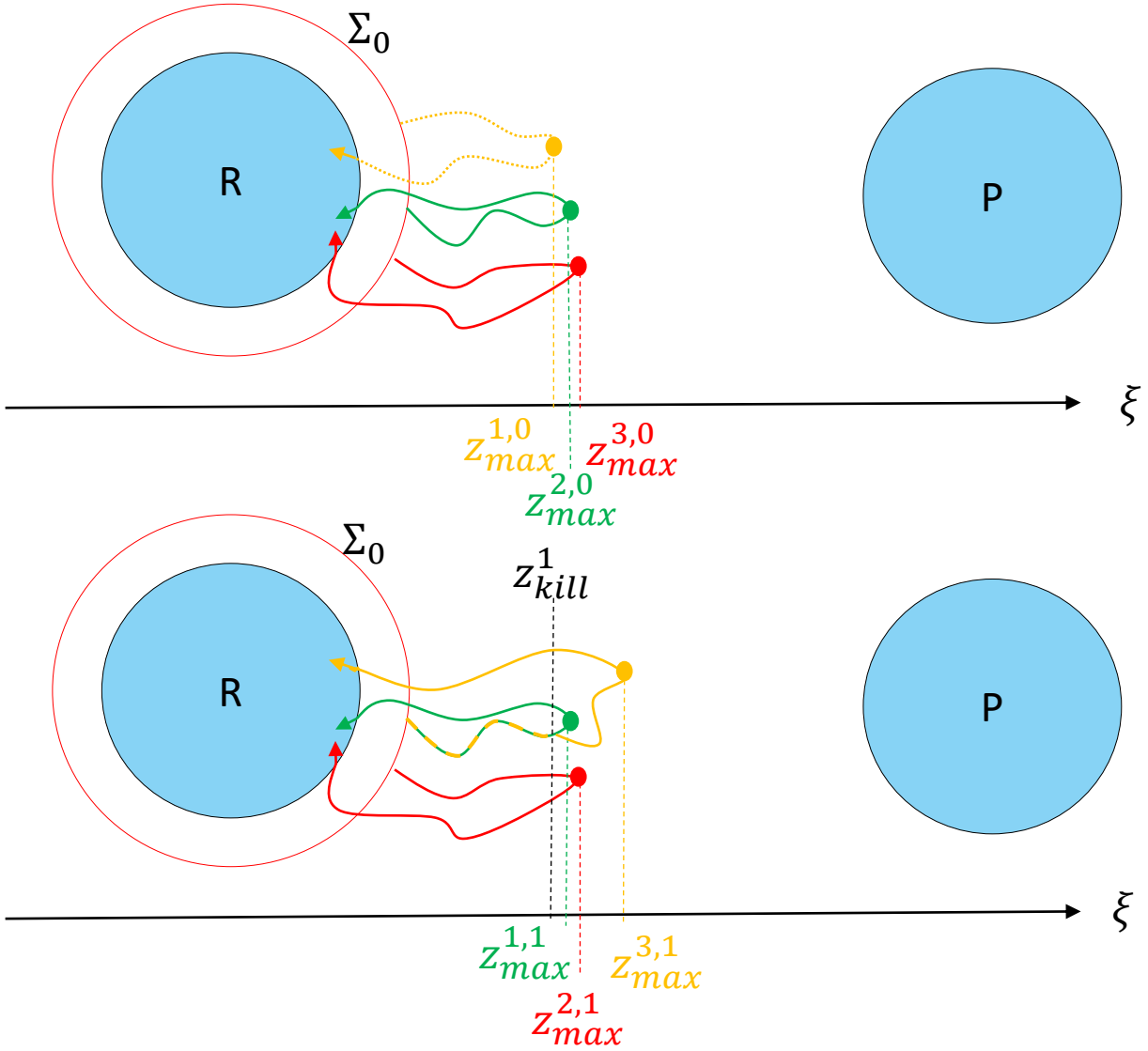
MD for Initial conditions.

2. Initialization



[1] F. Cérou, A. Guyader, (2007) Stoch. Anal. and Appl. 25, 2, 417.
[2] L. J. S. Lopes, T. Lelièvre, (2019) J. Comput. Chem. 40, 1198

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES



3. AMS iterations: $i \geq 0$

a) Save the smallest ($z_{max}^{1,i}$) as z_{kill}^{i+1} and delete all the trajectories that did not "go above" z_{kill}^{i+1}

b) Randomly select one trajectory within the remaining ones. Copy it until it reaches z_{kill}^{i+1} and continue it until it reaches R or P.

c) Classify all the replicas by increasing z_{max}.

$$\tilde{p} = \prod_{i=0}^{i_{max}} \tilde{p}_{\Sigma_{z_{kill}^{i}} \to \Sigma_{z_{kill}^{i+1}}} = \left(1 - \frac{1}{N}\right)^{i_{max}}$$

Unbiased estimator: $\mathbb{E}[\tilde{p}] = p_{R-P}(\Sigma_R)$

Variance depends on RC: $\mathrm{Var}[\tilde{p}] = f(\xi)$

Implemented with VASP software[1,2]

[1] G. Kresse, J. Hafner, (1993) J. Phys. Rev. B, 47, 558–561.
[2] G. Kresse, D. Joubert, (1999) Phys. Rev. B,59, 1758–1775.

II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE

Multistate problem

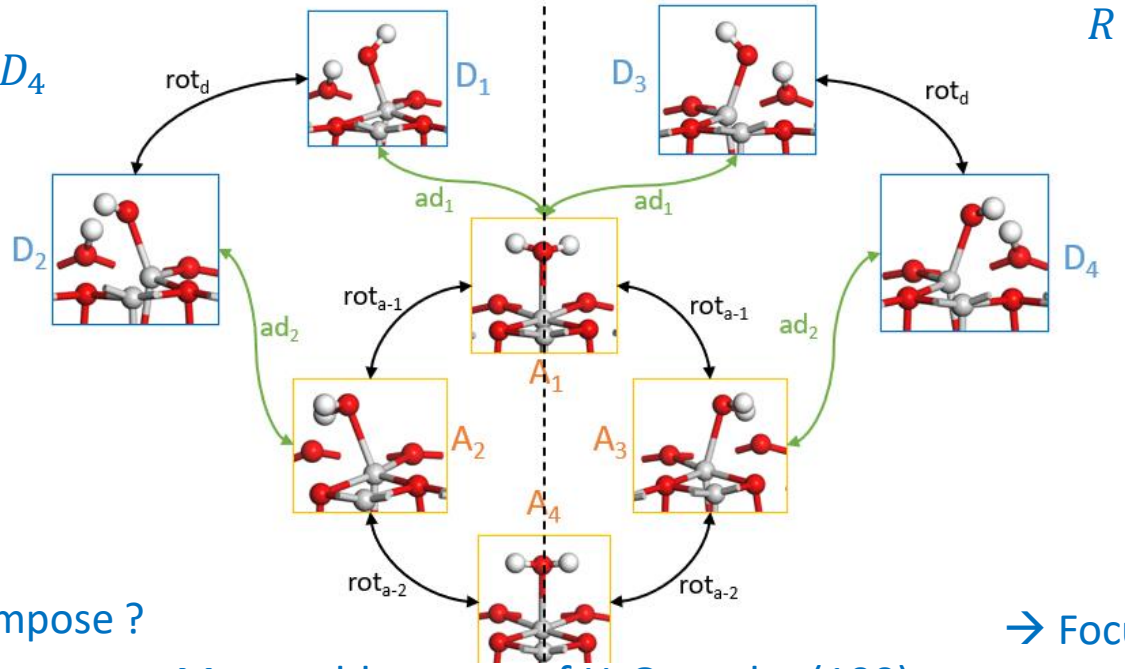
With
$R = A_1$
$\Sigma_R = \Sigma_{A_1}$
$P = A_2A_3 \cup A_4 \cup D_1D_3 \cup D_2D_4$

AMS can sample :

$A_1 \rightarrow A_2A_3$
$A_1 \rightarrow A_4$
$A_1 \rightarrow D_1D_3$
$A_1 \rightarrow D_2D_4$

→ Answers how A_1 can decompose ?

The most probable transition will be sampled, with precision conditioned by ξ



Metastable states of H_2O on the (100) surface of γ-alumina

With
$R = A_1 \cup A_2A_3 \cup A_4 \cup D_2D_4$
$\Sigma_R = \Sigma_{A_1}$
$P = D_1D_3$

AMS can sample :

$A_1 \rightarrow D_1D_3$

→ Focus specifically on one event

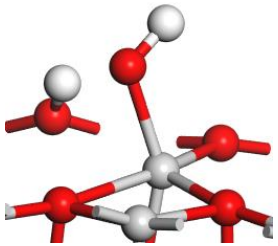
Quality of the sampling depends on ξ

II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE

Method:

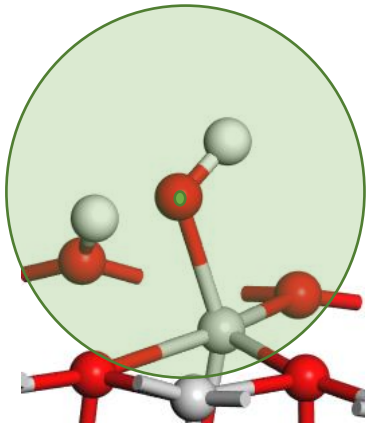
1. Identify the various metastable states (intermediates)
   → dissociated (D_i) or associated (A_i)
2. Run short dynamics in these states to sample Potential Energy Surface (PES) around the minima
3. SOAP[1] atom centered descriptors to numerically encode the structure for training the MLCV.

H
O
Al



Cartesian coordinates
(vector of 3N lines)

Select central atom



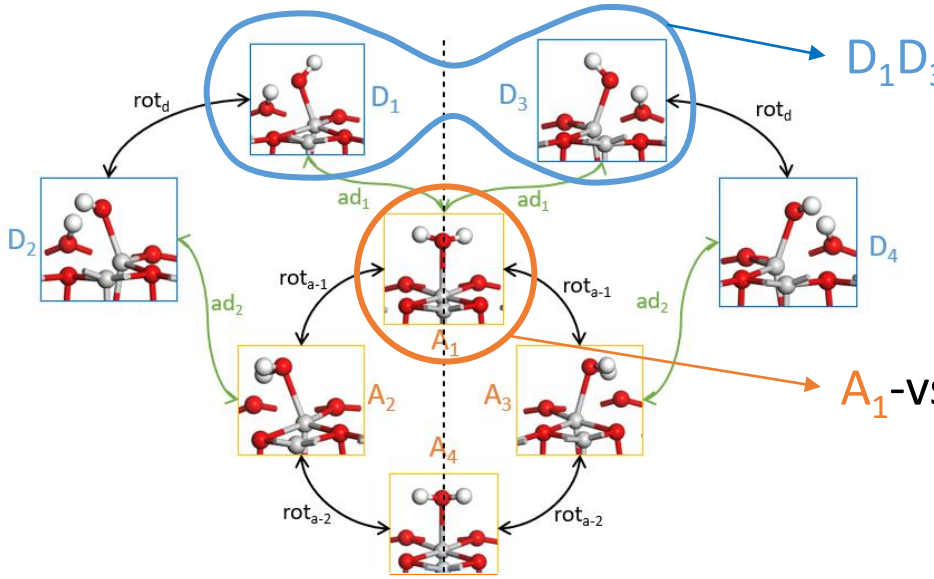
Atom centered descriptor of the structure
(vector of ~ 10^3-10^4 lines)

[1] A. P. Bartók, R. Kondor, .. G. Csányi, (2013) Phys. Rev. B, 87, 18, 184115.

II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE

SVM classifiers separate two sets of points by the highest margin plane[1]

SOAP-SVM CV : classifier decision function (f_X): algebraic distance to the plane.



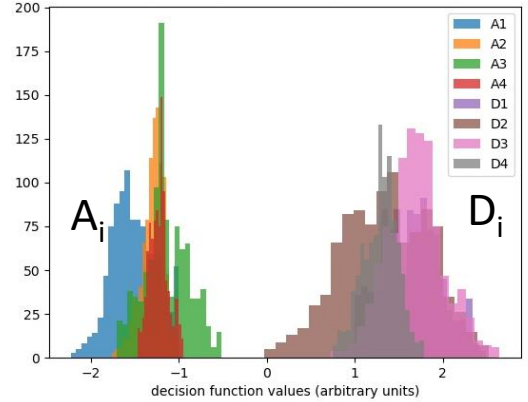
D$_1$D$_3$-vs-rest classifier

A$_1$-vs-rest classifier

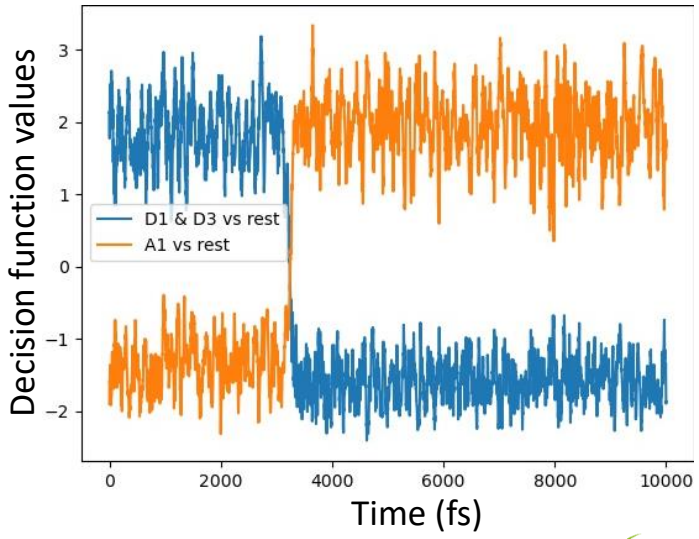
Classifier decision function interpretation:

$$f_X(\boldsymbol{q}) \in (-\infty, -1] \Leftrightarrow \boldsymbol{q} \in X$$

Histogram of SVM classifier decision function values

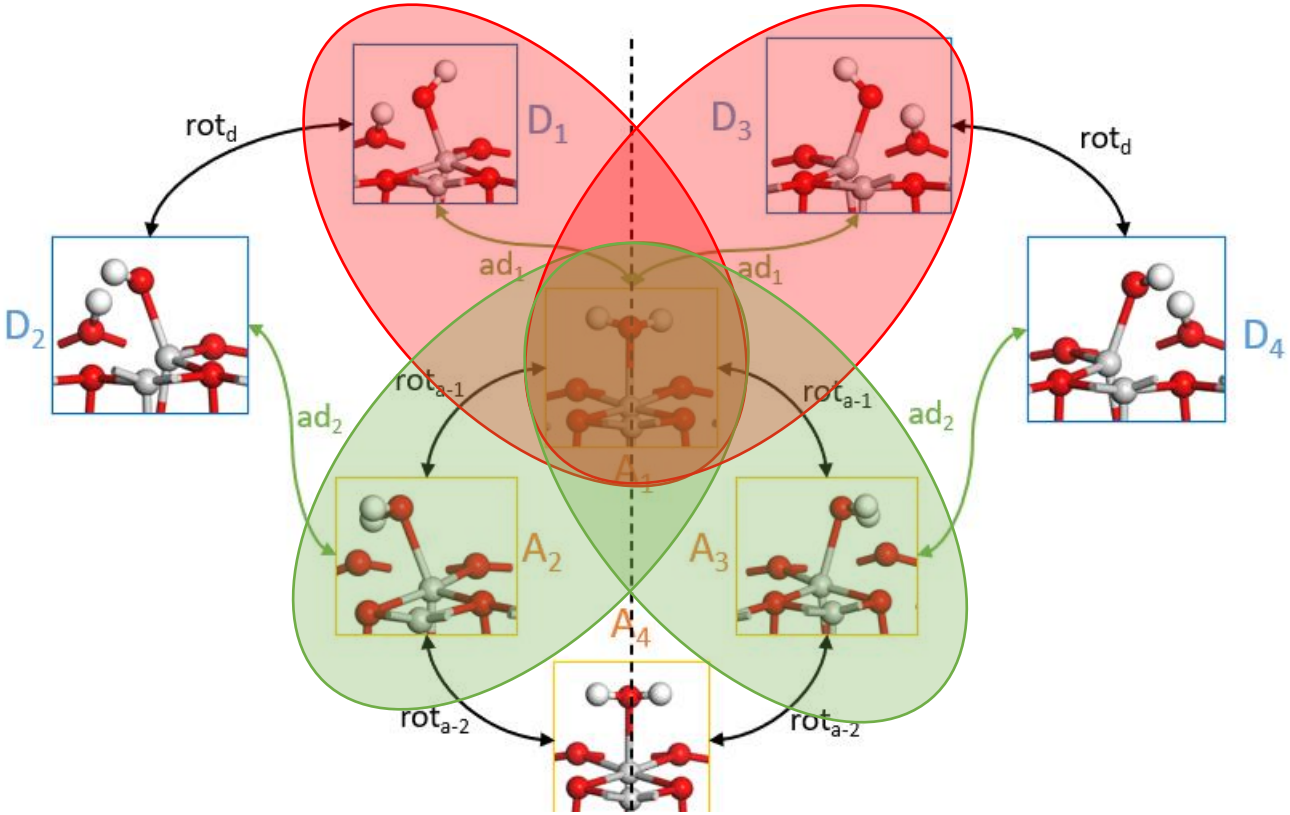


Training set : A1 = 0, D1 = 1



[1] K. P. Murphy, (2022) Probabilistic Machine Learning: An introduction; MIT Press: , 2022

II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE



Dissociation	Hill	hTST
$k_{A_1 \to D_1 D_3} =$	$1.6\ 10^{9}\ s^{-1}$	$3.4\ 10^{11}\ s^{-1}$
$k_{D_1 D_3 \to A_1} =$	$2.3\ 10^{10}\ s^{-1}$	$1.1\ 10^{12}\ s^{-1}$

Rotation	Hill	hTST
$k_{A_1 \to A_2 A_3} =$	$3.8\ 10^{10}\ s^{-1}$	$7.6\ 10^{10}\ s^{-1}$
$k_{A_2 A_3 \to A_1} =$	$1.5\ 10^{11}\ s^{-1}$	$2.1\ 10^{12}\ s^{-1}$

hTST rates are larger

Might come from entropy estimation / recrossing

~ 2 10^6 CPU Hours

[1] TP, G. Stoltz, M. Corral-Valero, A. Anciaux-Sedrakian, M. Moreaud, T. Lelièvre, P. Raybaud (2023) J. Chem. Theory Comput. 19, 12, 3538–3550

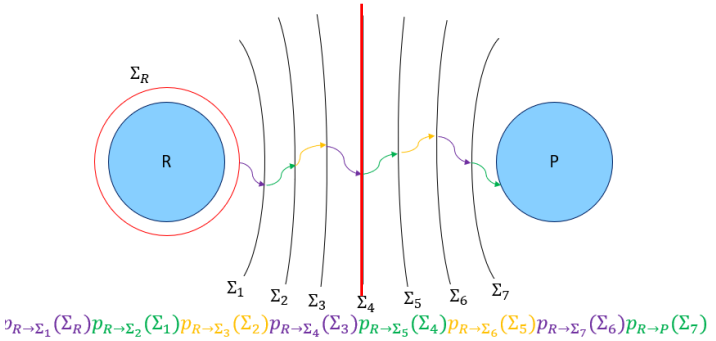
II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE

Identify TS structures

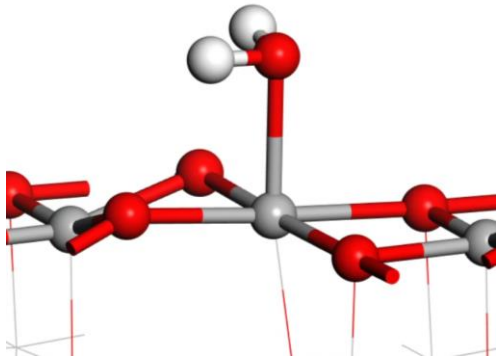
TS in the sense of committor function $p_{R\to P}$ (probability of reaching P before R)[1]

Find the level of the RC z^n_{kill} such that $p_{R\to P}\left(\Sigma_{z^n_{kill}}\right) = 0.5$

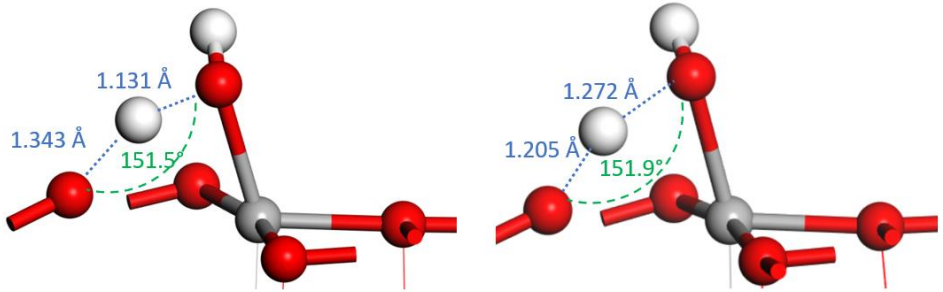
$$\prod_{i=n}^{i_{max}} \tilde{p}_{R\to\Sigma_{z^{i+1}_{kill}}}\left(\Sigma_{z^i_{kill}}\right) = 0.5$$



Along each trajectory, take the structure right after the level $\Sigma_{z^n_{kill}}$ is crossed, then find the average structure



Example for the $A_1 \to D_1D_3$ reaction



Saddle point

AMS estimated $p_{A_1\to D_1D_3} = 0.5$

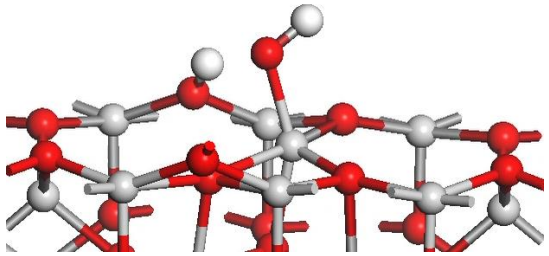
[1] E. Vanden-Eijnden, E. Transition Path Theory (2006) in Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology. p 453-493

III. REDUCING AMS COMPUTATIONAL COST WITH MACHINE LEARNING FORCE FIELD

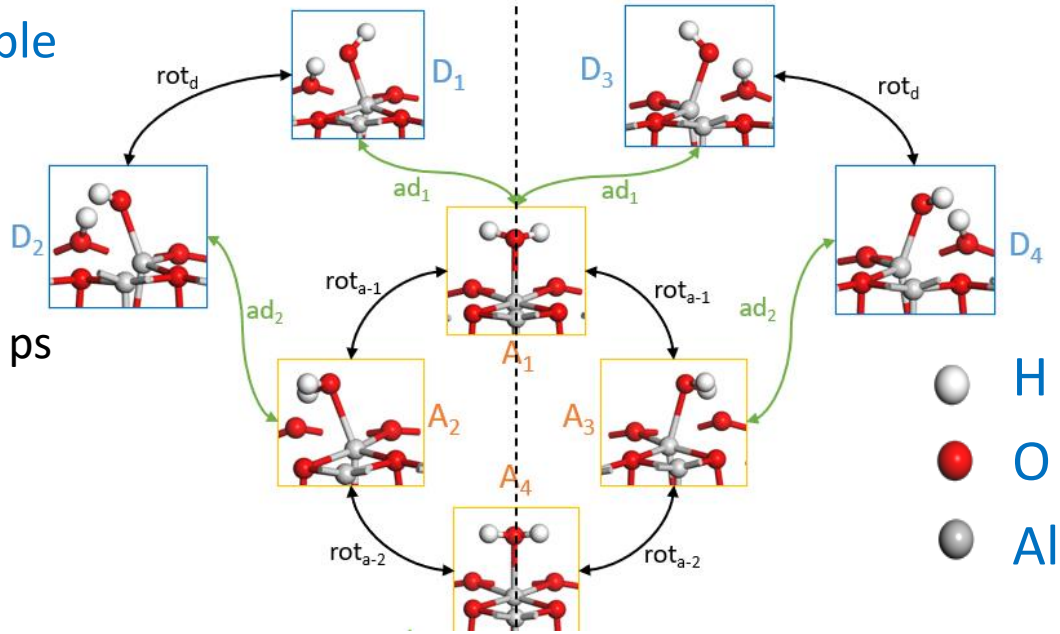
Method to train a MLFF[1,2]:

1. Identify the various metastable states (intermediates)
   → dissociated (D_i) or associated (A_i)
2. Run active learning dynamics in these states to sample Potential Energy Surface (PES) around the minima

D_1 dissociated structure



dt = 1 fs
Total time = 50 ps



Identified structures and **intuitively** plausible transitions

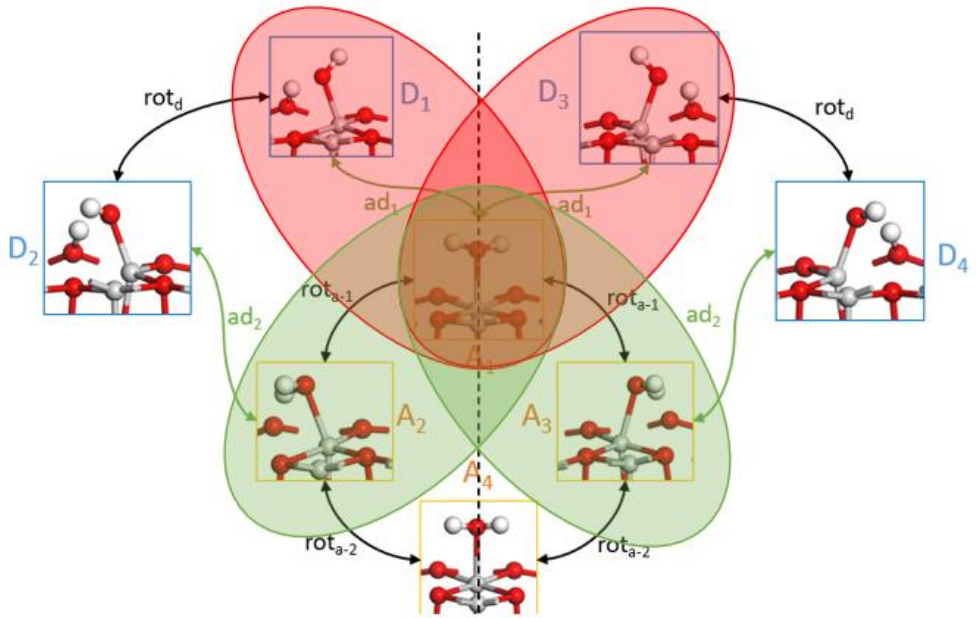
3. Concatenate the dataset (containing E, Forces, and positions)
4. AMS with Active learning
5. Re-fit the force field

→ Rate can be estimated with AMS with MLFF without active learning

[1] R. Jinnouchi, F. Karsai, G. Kresse, G. (2019) Phys. Rev. B, 100, 014105.
[2] R. Jinnouchi, K. Miwa, F. Karsai, G. Kresse, R. Asahi, (2020) J. Phys. Chem. Lett., 11, 6946–6955.

III. REDUCING AMS COMPUTATIONAL COST WITH MACHINE LEARNING FORCE FIELD

$A_1 \rightarrow D_1D_3$ Dissociation



with N_{rep} = 200 and M_{real} = 10

DFT

$k_{A_1 \rightarrow D_1D_3} = (1.64 \pm 1.59)\ 10^9\ s^{-1}$

MLFF

$k_{A_1 \rightarrow D_1D_3} = (2.76 \pm 3.81) 10^9\ s^{-1}$

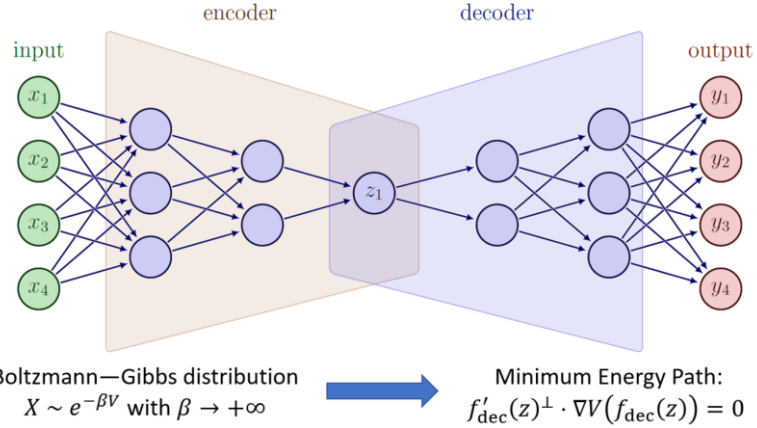
with N_{rep} = 800 and M_{real} = 10

$k_{A_1 \rightarrow D_1D_3} = (2.43 \pm 1.15)\ 10^9\ s^{-1}$

Test MAE forces =76 meV/Å
Test MAE energy = 20 meV
1000 configurations randomly drawn
from sampled reactive trajectories

CONCLUSION AND PERSPECTIVES

- hTST overestimate the DFT-MD rate estimated using AMS
- MLFF-MD and DFT-MD rates are consistent
- MLFF used in prediction mode drastically reduces de computational cost
- Current implementation of AMS with VASP limits the application of active learning
  → Restart does have an important cost for the active learning.
- Using D-optimality criterion active learning with VASP as calculator of ab-initio with ACE potential seems a good opportunity[1]
- Active learning of RC ξ can be included in the workflow[2]

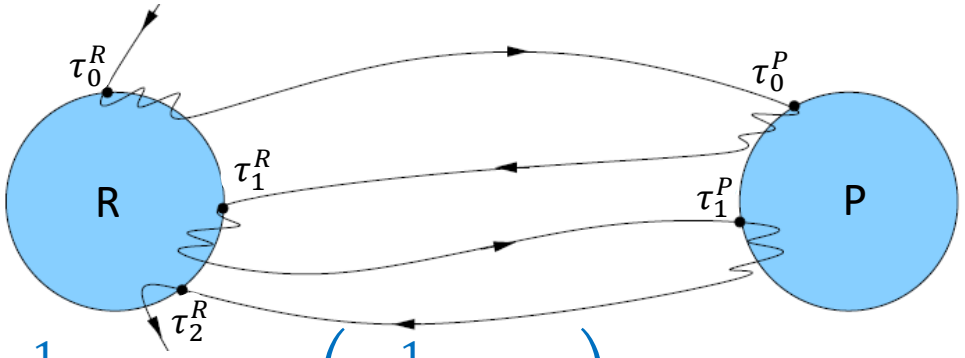


[1] Y. Lysogorskiy, A. Bochkarev, M. Mrovec, R. Drautz, (2023) Phys. Rev. Mater., 7, 4, 043801
[2] T. Lelièvre, TP, G. Stoltz, W. Zhang, (2024) J. Phys. Chem. B, 128, 11, 2607

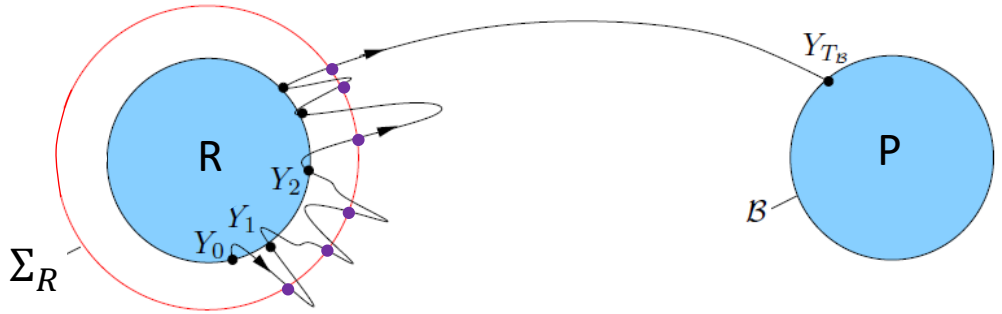
Thank you for you attention

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES

Transition time: $\frac{1}{k_{RP}} = t_{RP} = \ mean\left(\tau_i^R - \tau_i^P\right)$



We model the reaction time as: $\frac{1}{k_{RP}} = t_{RP} = \left(\frac{1}{p_{\Sigma_R - P}} - 1\right)(t_{R-\Sigma} + t_{\Sigma-R}) + t^{\dagger}_{R-\Sigma} + t_{\Sigma-P} \approx \frac{t_{R\to\Sigma\to R}}{p_{\Sigma_R\to P}} = \frac{1}{p_{\Sigma_R\to P}\,\phi_R}$



$p_{\Sigma_R - P}$: probability of reaching P before R when starting from Σ_R.

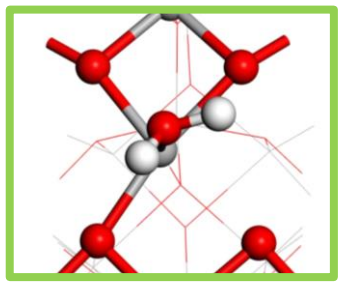
[1] Baudel, M., Guyader, A., & Lelièvre, T. (2020). On the Hill relation and the mean reaction time for metastable processes. *arXiv preprint, arXiv:2008.09790*.

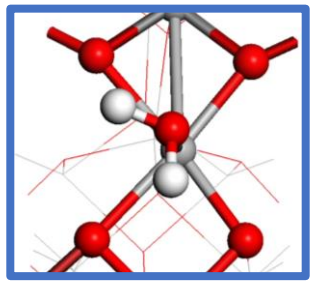
II. AIMD METHOD APPLIED TO WATER DISSOCIATION ON (100) SURFACE

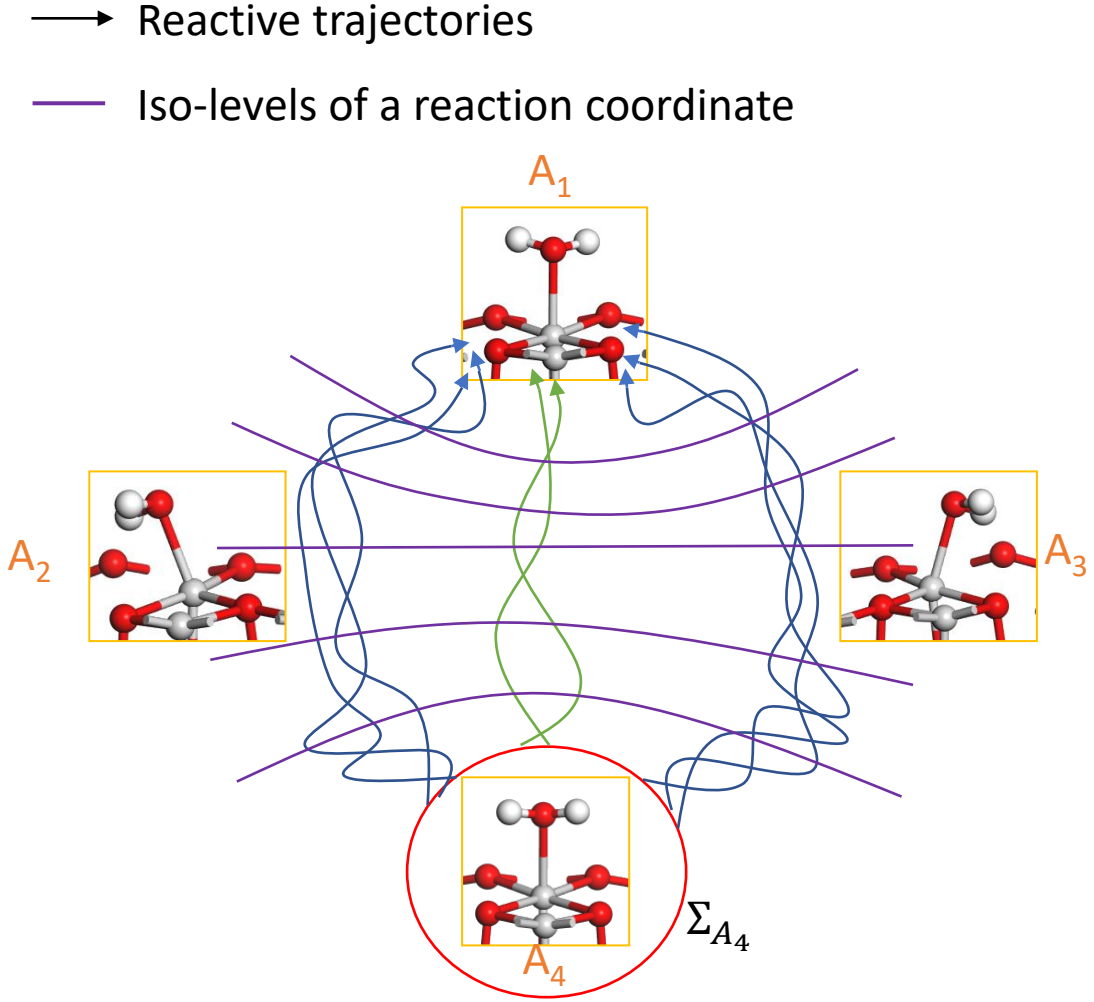
Use K-means clustering method to identify groups of trajectories.

Based on SOAP descriptor + PCA to describe 5 structures per trajectory.

5 Structures = First time trajectory cross RC iso-levels







III. REDUCING AMS COMPUTATIONAL COST WITH MACHINE LEARNING FORCE FIELD

Wall clock time	Total number of structures		
~0.5 days	2600	50 ps active learning in each metastable states A_i and D_i starting from scratch	Final force mean Bayesian error 10 to 60 meV/Å
~0.5 days	4000	50 ps active learning starting in A_1 with the merged datasets	25 meV/Å
~10 days	4010	10 A_1-vs-all AMS runs using active learning	25 meV/Å

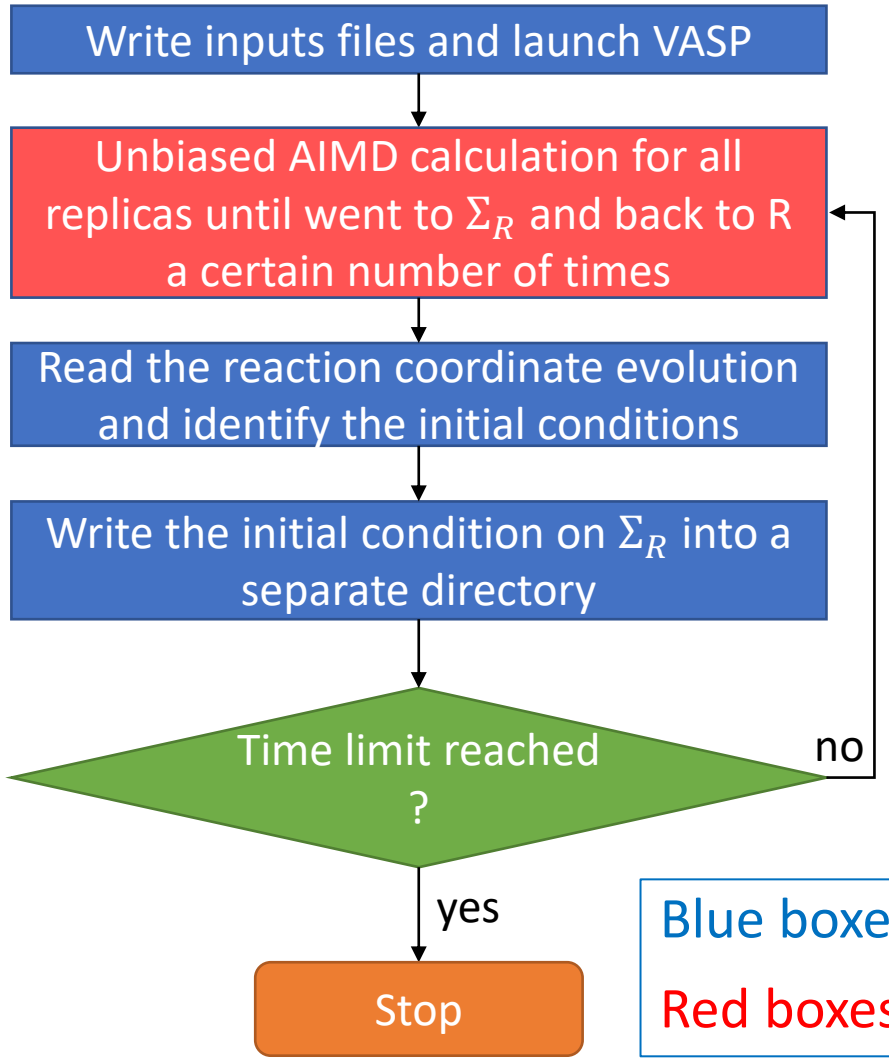
Threshold updated using the stored Bayesian errors[1,2]

→ Transitions already sampled during first two steps?

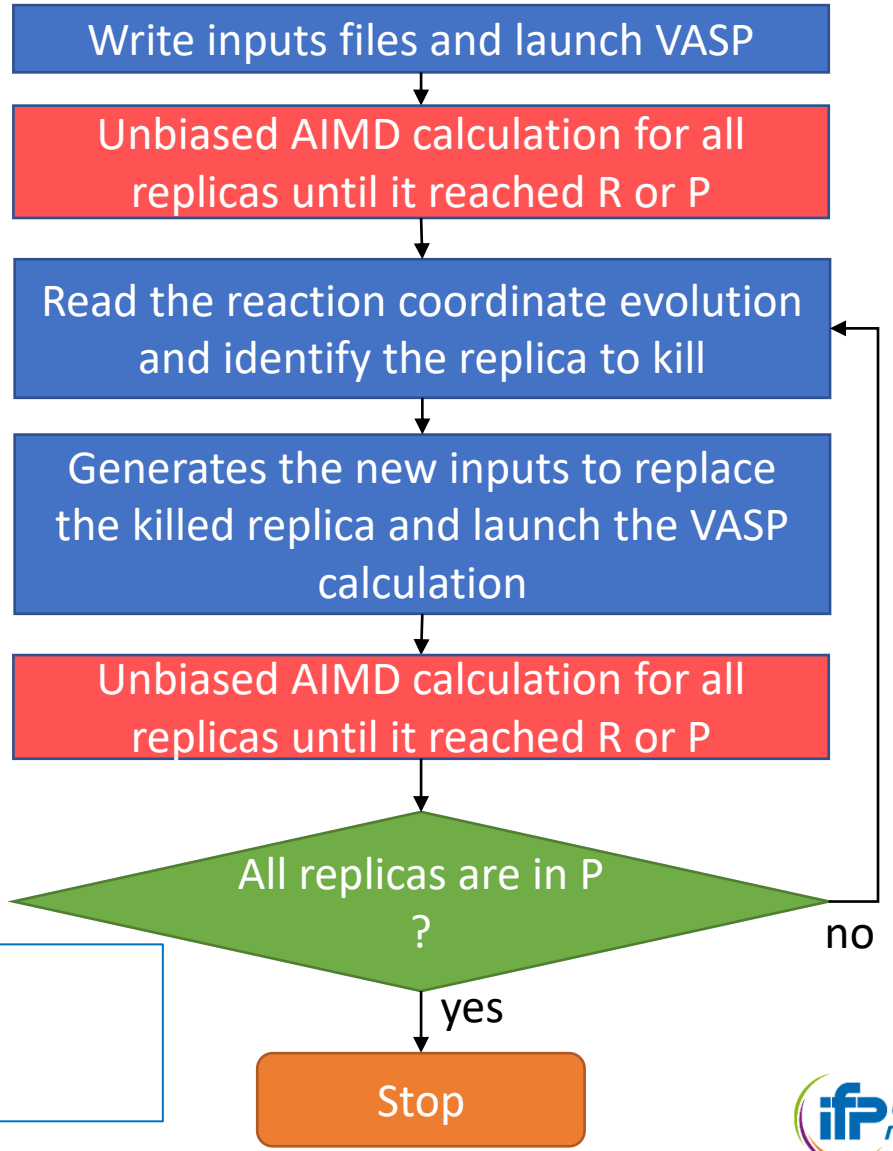
1 Jinnouchi, R.; Karsai, F.; Kresse, G. (2019) *Phys. Rev. B* 100, 014105
2 Jinnouchi, R.; Miwa, K.; Karsai, F.; Kresse, G.; Asahi, R. (2020) *J Phys. Chem. Lett.* 11, 6946

I. ADAPTIVE MULTI-LEVEL SPLITTING METHOD FOR REACTION RATES

Sampling of initial conditions



Probability estimation with AMS



Blue boxes: python code

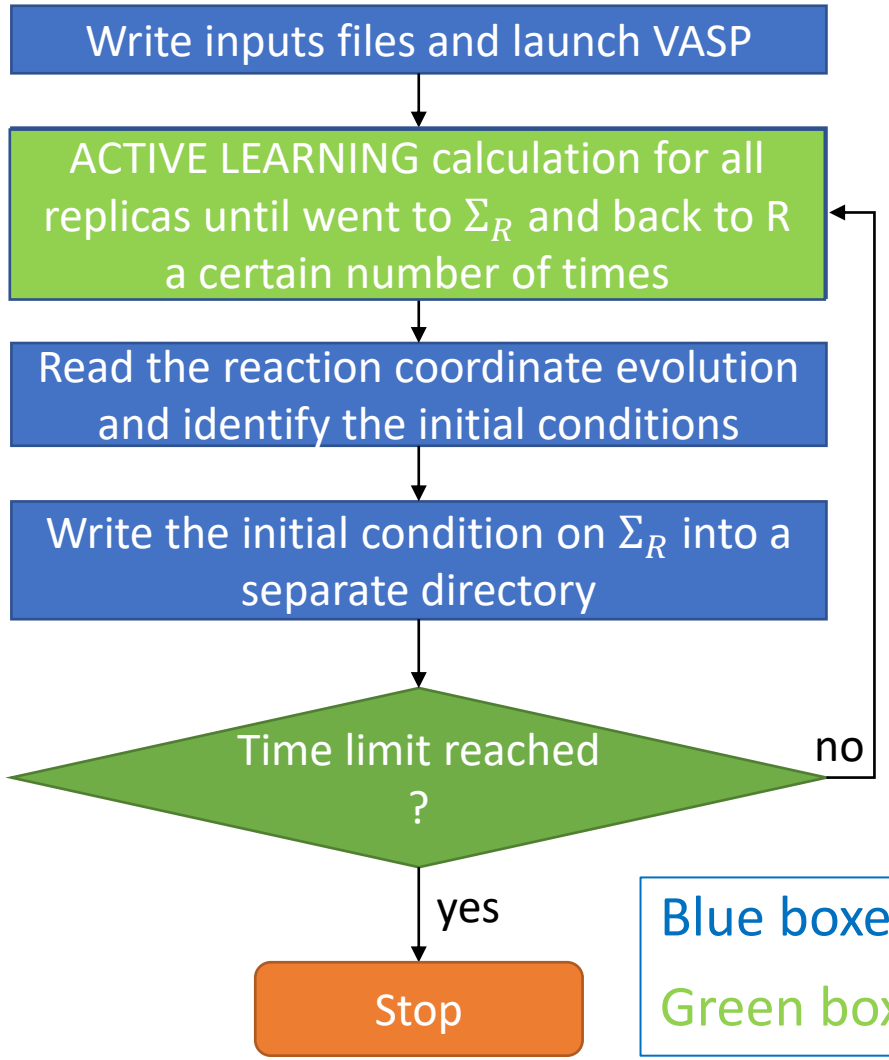
Red boxes: VASP code[1,2]

[1]Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, 47, 558–561.
[2]Kresse, G.; Joubert, D *Phys. Rev. B* **1999**, 59, 1758–1775.

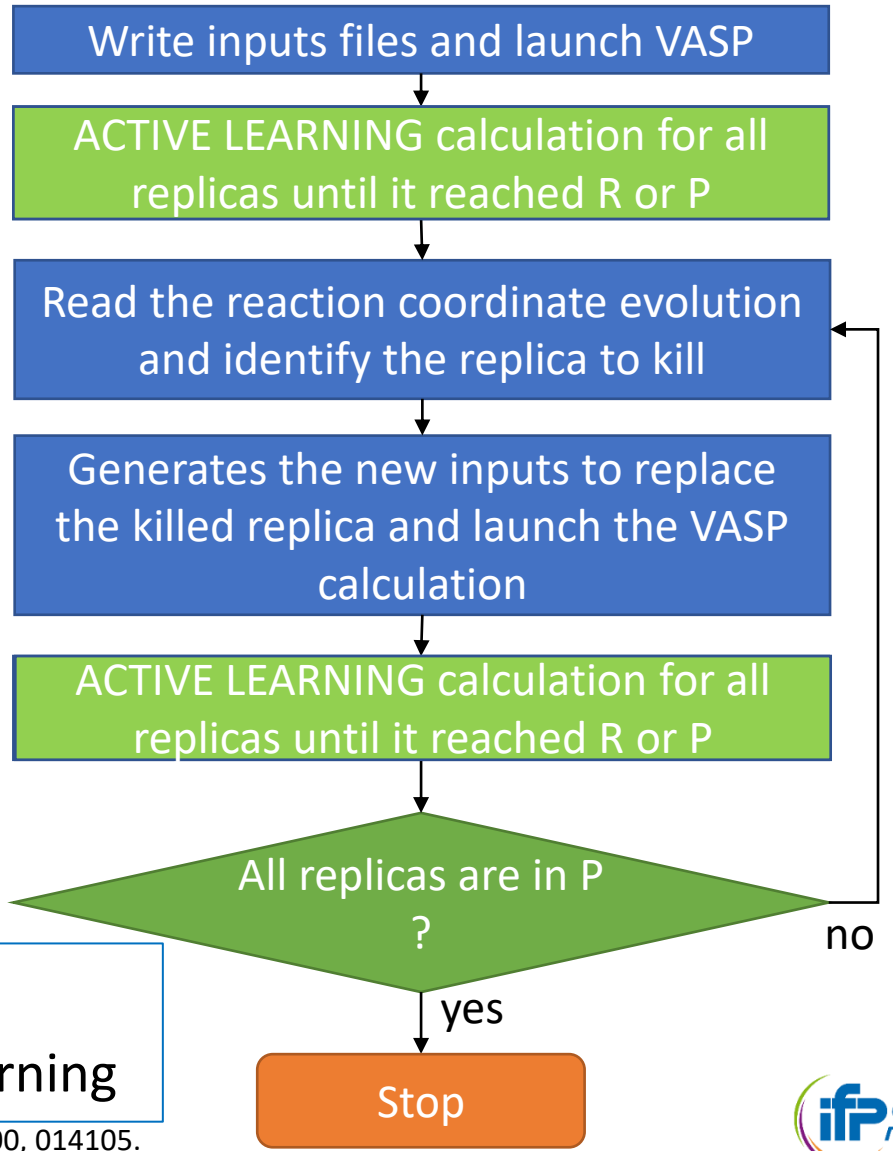
III. USING AMS WITH ACTIVE LEARNING

AMS IMPLEMENTATION WITH VASP (PLANE WAVE DFT)

Sampling of initial conditions



Probability estimation with AMS



Blue boxes: python code

Green boxes: VASP active learning

[1] Jinnouchi, R.; Karsai, F.; Kresse, G. *Phys. Rev. B* **2019,** 100, 014105.
[2] Jinnouchi, R.; Miwa, K.; Karsai, F.; Kresse, G.; Asahi, R. *The Journal of Physical Chemistry Letters* **2020**, 11, 6946–6955.