Finding optimal reaction coordinates via kinetic rate minimization

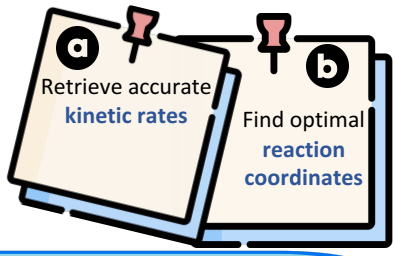
Line Mouaffac[1], Karen Palacio-Rodriguez[2], Fabio Pietrucci[1]

[1]Sorbonne Université, Musée National d'Histoire Naturelle, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Paris, France
[2]Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankurt am Main, Germany

1 INTRODUCTION

While **collective variables** (CV) are ubiquitously used to model physico-chemical transformations, finding **optimal reaction coordinates** (RC) that yield accurate thermodynamic and kinetic properties is a **fundamental challenge** in the field of atomistic simulations. Projecting onto an accurate RC helps to **accurately and efficiently** estimate **reaction rates** and identify the common properties of transition states. Customarily, optimal RCs are defined as monotonic one-to-one functions of **the committor** [1].We are proposing an original computationally affordable method to optimize reaction coordinates [2]. It is based on the fact that sub-optimal RCs yield kinetic rates higher than optimal ones.

OUR AIMS



2 THEORETICAL FRAMEWORK

In 2016, Zhang et. [3] investigated the repercussions on the reaction rate k of **projecting** the full dynamics onto a **low-dimensional space** defined by the CV

Full dynamics $\{\vec{R_i}\}_{i=1,\dots,N}$ (3N dimensions) → Projection over CV → **Effective dynamics** $q(\{\vec{R_i}\})$ (1 dimension)

The rate of the full dynamics is less than or equal than the one computed using effective dynamics.
The kinetic rate can be used as an optimization criterion for CVs

$$k^{true}_{A\to B} \leq k^{model}_{A\to B}$$

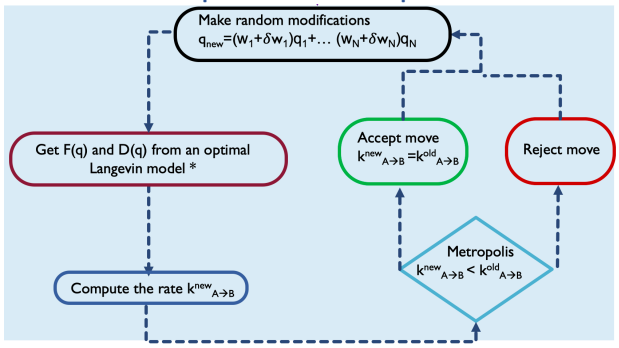
The free energy F(q) and diffusion D(q) are obtained by training an overdamped Langevin model via **likelihood maximization** [4].

$$\dot{q} = -\beta D(q)\frac{\partial F(q)}{\partial q} + \frac{\partial D(q)}{\partial q} + \sqrt{2D(q)}\eta(t)$$

Free energy (**thermodynamics**) — Diffusion profile (**kinetics**)

3 COMPUTATIONAL METHODS

1. Create a pool of collective variables
2. Generate an initial **random linear combination** of CVs
   $q_{old}=w_1q_1+w_2q_2+\dots w_Nq_N$
3. Perform **Metropolis-Monte Carlo optimization** scheme

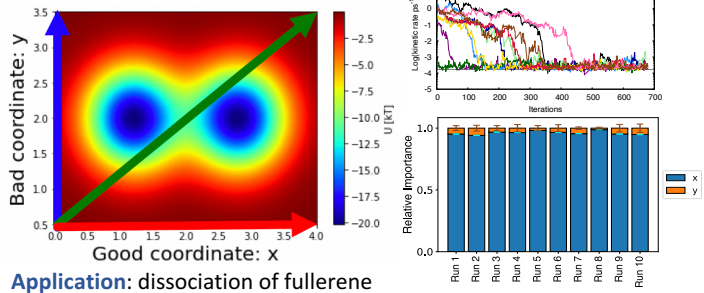


4. Get rate and optimal linear combination after convergence

$$\frac{1}{k_{A\to B}} = \int_{q_0}^{b} dy \frac{e^{\beta F(y)}}{D(y)} \int_a^y dz e^{-\beta F(z)}$$

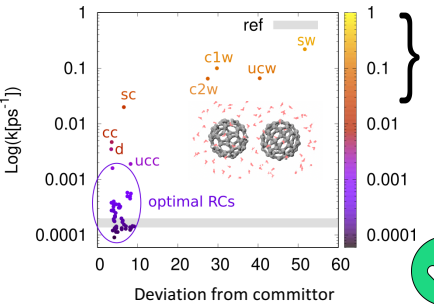
4 METHOD VALIDATION

Benchmark: 2D double well

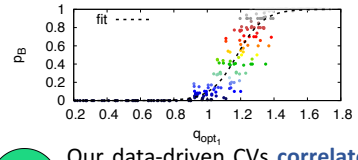


Application: dissociation of fullerene dimer in water

Pool of heuristic CVs



sw: entropy **sc**: carbon entropy
ucw: VdW Energy. **d**: distance
cc: C-C- contacts **c1w**: C-O contacts

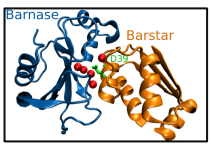


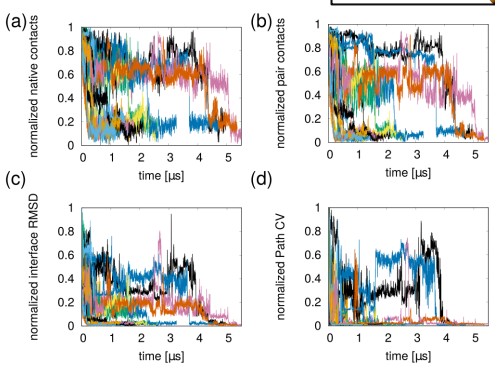
Our data-driven CVs **correlate well with the committor** using only a fraction of the computational cost

5 METHOD PERSPECTIVES

Application: predict dissociation rates of protein-protein complexes in water

Brute force association trajectories provided by DE Shaw [5]





We developed an efficient & **computationally affordable** method that automates the **optimization of RCs**

Our method could be applied to find **dissociation rates** of proteins and find **optimal RCs** that describe the transition

REFERENCES

[1] J. Lu and E. Vanden-Eijnden *JCP*, 2014.
[2] L. Mouaffac, K. Palacio-Rodriguez, and F. Pietrucci *JCTC*, 2023.
[3] W. Zhang, C. Hartmann, and C. Schütte *Faraday discuss.*, 2016.
[4] K. Palacio-Rodriguez and F. Pietrucci *JCTC*, 2022.
[5] A. Pan, D. Jacobson, K. Yatsenko, D. Sritharan, T. Weinreich, D.Shaw *PNAS (2019)*