

Rare events in deep potential simulations

Pablo Piaggi

CIC nanoGUNE and Ikerbasque
San Sebastian, Spain



Deep Modeling for
Molecular Simulation
Princeton University
June 27, 2024

About San Sebastian



Cite This: ACS Nano 2019, 13, 12254–12256

www.acsnano.org

San Sebastian, a City of (Nano)Science and Technology



CFM
Centro de Física de Materiales
Materials Physics Center



Universidad
del País Vasco
Euskal Herriko
Unibertsitatea

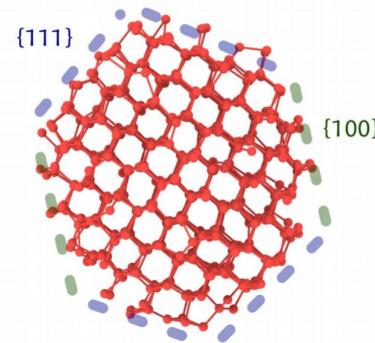
The timescale problem in molecular simulations

$$m\ddot{\mathbf{R}} = \mathbf{F} = -\nabla_{\mathbf{R}}U$$

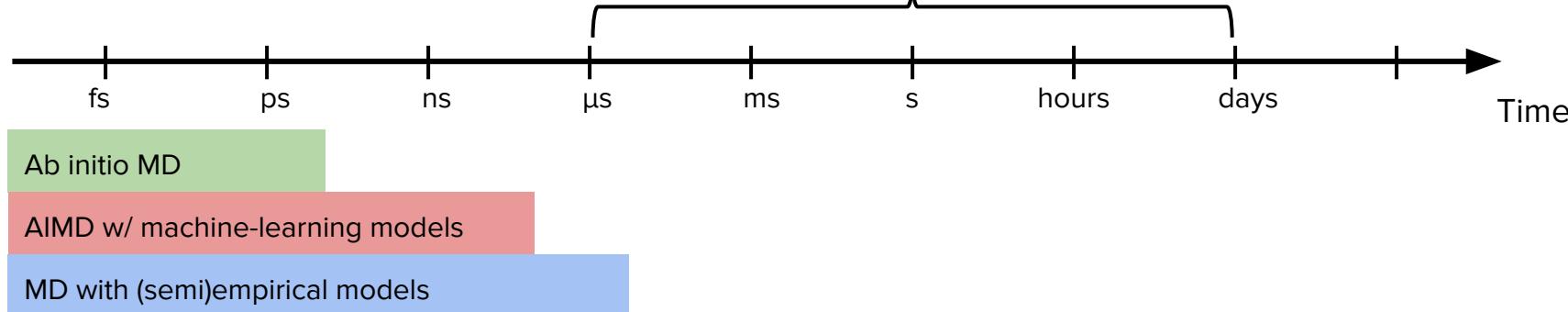
$$\mathbf{R}(t + \Delta t) = \mathbf{R}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2m}\mathbf{F}(t)\Delta t^2$$

Integration time step
(fs)

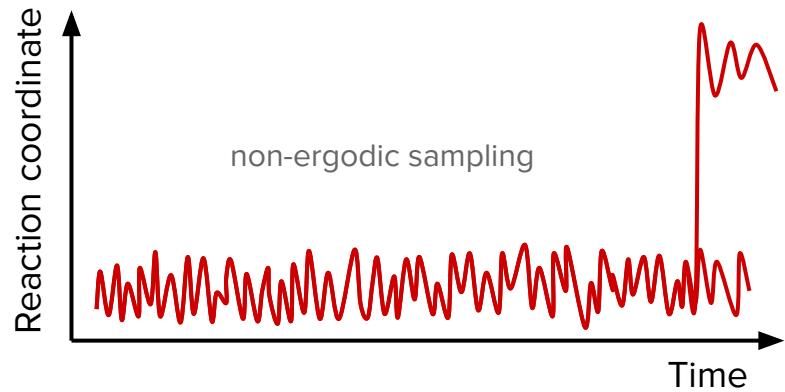
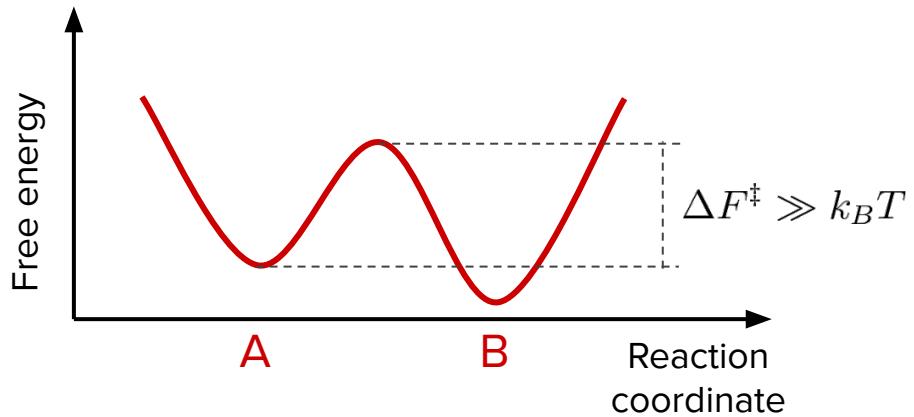
Bond vibration



Homogeneous nucleation



Rare event scenario



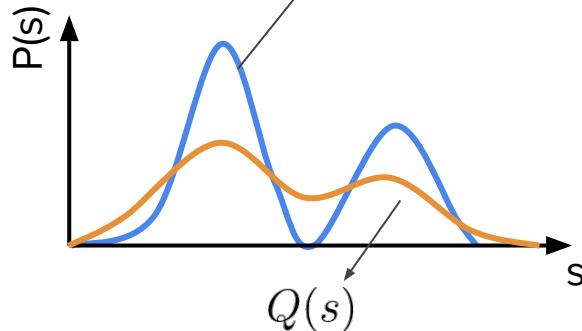
Impossible to obtain ergodic sampling in standard MD simulations under the rare event scenario → In order to calculate properties or study the transition mechanism one must accelerate the transitions

Classes of rare-event or enhanced-sampling methods

Collective-variable-based methods

Enhance the fluctuations of a slow mode of the system

$$P(s) \propto \int d\mathbf{R} e^{-\beta U(\mathbf{R})} \delta(s - s(\mathbf{R}))$$

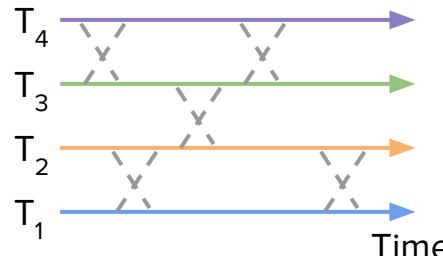


Examples: metadynamics, OPES, umbrella sampling

Generalized-ensemble methods

Combine different thermodynamic conditions or Hamiltonians in the same simulation

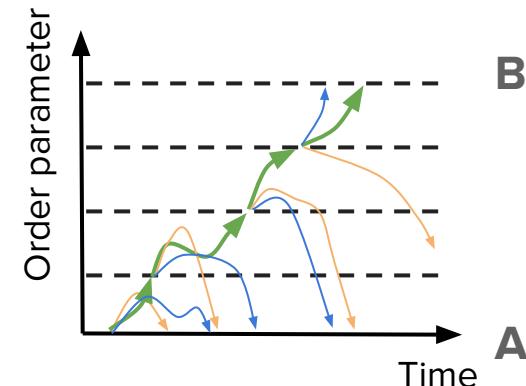
$$P(\mathbf{X}) \propto e^{-\sum_i \beta_i E(\mathbf{R}_i)}$$



Example: parallel tempering

Path-sampling methods

Sample the space of possible paths and find those that connect A and B

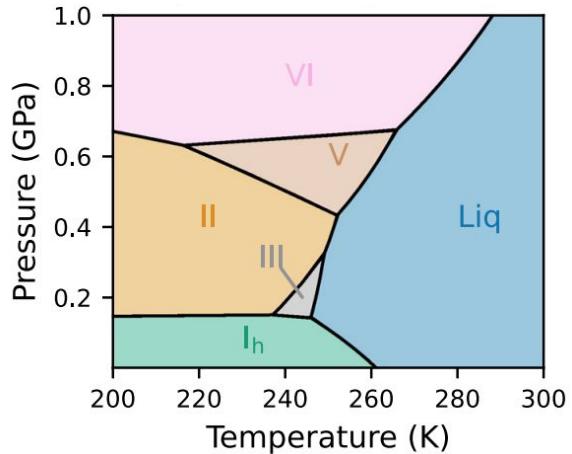


Example: forward-flux sampling

Use cases

Thermodynamic properties

Phase diagrams, melting temperatures, chemical potentials

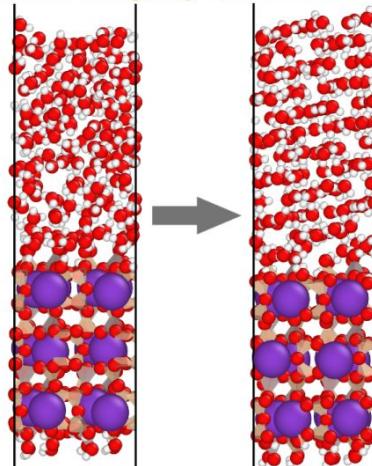


Bore, Piaggi, Car and Paesani, J. Chem. Phys. (2022)
Bore and Paesani, Nat. Commun. (2023)

Training of deep potentials

Accurate training of deep potentials in rare-event scenarios

Biased sampling - ice formation

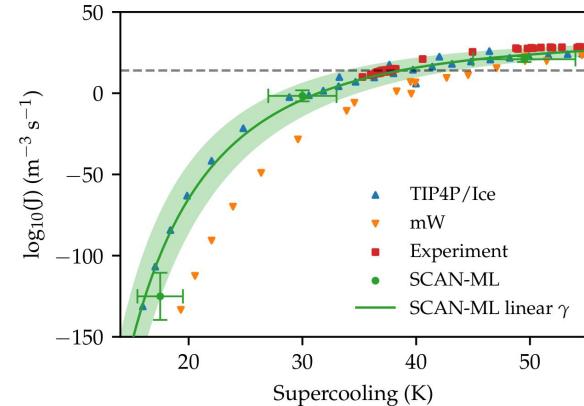


Water to ice transformation at (100),
(010), and (001) surfaces

Piaggi, Selloni, Panagiotopoulos, Car, and
Debenedetti, Faraday Discussions 249, 98 (2023)

Kinetics

Recover kinetics not accessible to direct MD

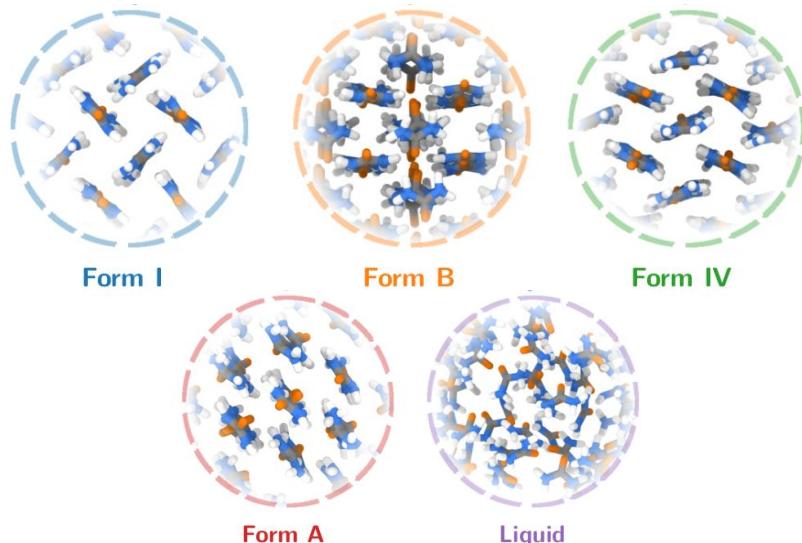


Piaggi, Weis, Panagiotopoulos, Debenedetti, and Car,
PNAS 119, 33 (2022)

Use cases

Exploration

Discovery, crystal structure prediction

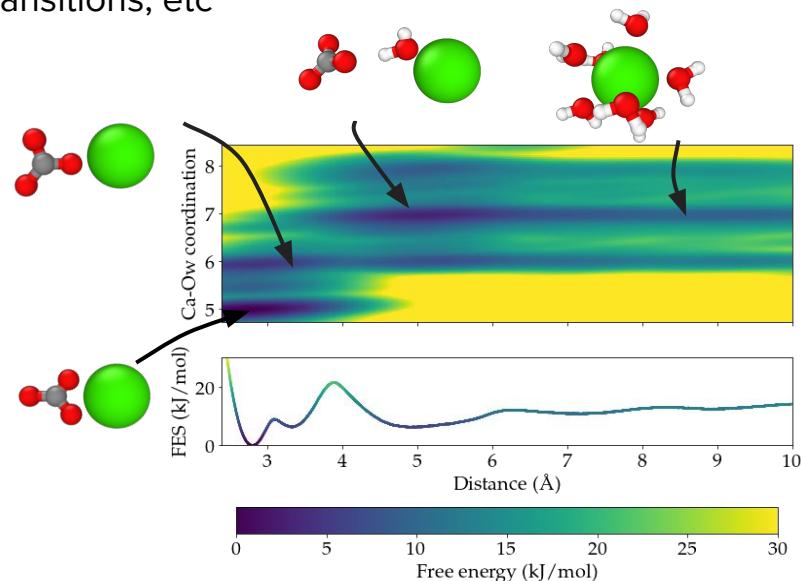


Piaggi, Valsson, and Parrinello, Phys. Rev. Lett. 119, 015701 (2017)

Piaggi and Parrinello, PNAS 115 (41), 10251 (2018)

Free Energy Surfaces (FES)

Study mechanisms of chemical reactions, phase transitions, etc



Piaggi, Gale, and Raiteri, in preparation (2024)

Collective-variable-based methods

General framework

Add a bias potential $V(\mathbf{R})$

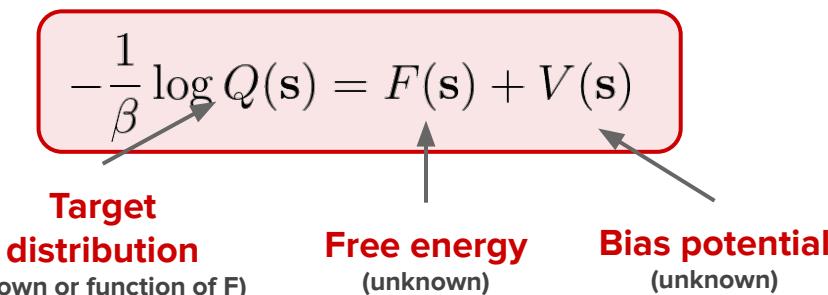
$$U'(\mathbf{R}) = U(\mathbf{R}) + V(\mathbf{R})$$

$$V(\mathbf{R}) = V(\mathbf{s}(\mathbf{R})) \rightarrow \text{Collective variables (CV) or order parameter}$$

New distribution:

$$Q(\mathbf{s}) \propto P(\mathbf{s}) e^{-\beta V(\mathbf{s})}$$

$$F(\mathbf{s}) = -\frac{1}{\beta} \log P(\mathbf{s})$$



Reweighting: Calculation of observables

$$\langle O(\mathbf{R}) \rangle = \frac{\langle O(\mathbf{R}) e^{\beta V(s(\mathbf{R}))} \rangle_V}{\langle e^{\beta V(s(\mathbf{R}))} \rangle_V}$$

Only if the bias potential is static!

Not so obvious cases:

$$F(\mathbf{s}) = -\frac{1}{\beta} \ln \langle \delta(\mathbf{s} - \mathbf{s}(\mathbf{R})) e^{\beta V} \rangle_V$$

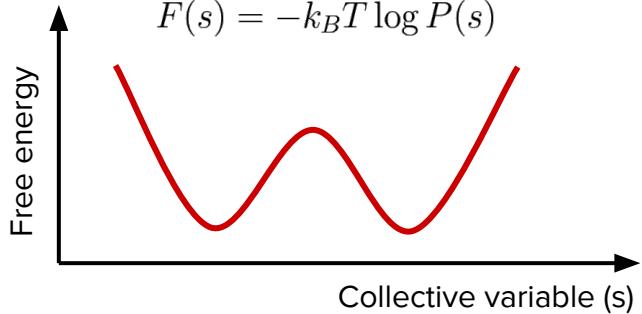
$$\Delta F = -\frac{1}{\beta} \ln \left(\frac{\langle H(s - s^*) e^{\beta V} \rangle_V}{\langle 1 - H(s - s^*) e^{\beta V} \rangle_V} \right)$$

Torrie and Valleau, J. Comp. Phys. (1977)

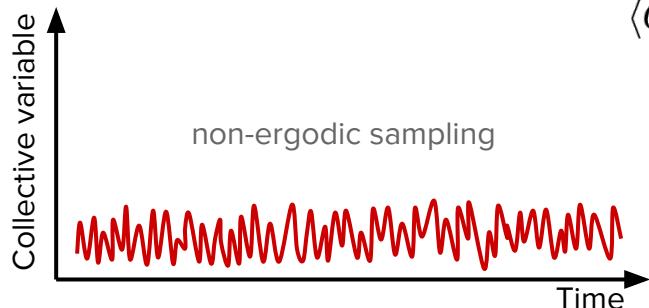
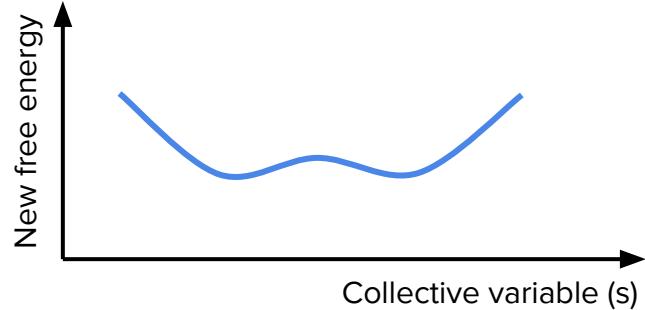
Valsson et al, Annu. Rev. Phys. Chem. 67, 159-184 (2016)

Collective-variable-based methods

Summary of the strategy

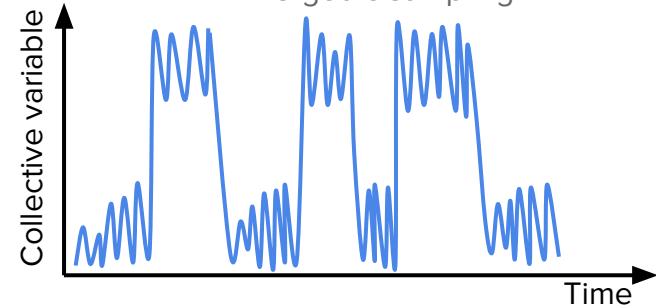


add bias potential $V(s)$



non-ergodic sampling

$$\langle O(\mathbf{R}) \rangle = \frac{\langle O(\mathbf{R}) e^{\beta V(s(\mathbf{R}))} \rangle_V}{\langle e^{\beta V(s(\mathbf{R}))} \rangle_V}$$



Collective-variable-based methods

Well-tempered Metadynamics



https://people.sissa.it/~laio/Research/Res_metadynamics.php

Add a bias potential:

$$V(\mathbf{s}, t) = \sum_{k=1}^n W e^{-||\mathbf{s}-\mathbf{s}_k||^2} \exp \left[-\frac{1}{\gamma - 1} \beta V_{k-1}(\mathbf{s}_k) \right]$$

Converges to:

$$V(\mathbf{s}, t) = - \left(1 - \frac{1}{\gamma} \right) F(\mathbf{s}) + c(t),$$

$F'(\mathbf{s}) = F(\mathbf{s})/\gamma$
well-tempered distribution
with bias factor γ

$c(t)$ always increases and
has to be subtracted for
reweighting!

Idea: increase the probability to observe rare but interesting configurations (e.g. transition states) by discouraging frequently visited configurations

Lai and Parrinello, PNAS 99, (2002)
Barducci et al, Phys. Rev. Lett. 100, (2008)
Valsson et al, Ann. Rev. Phys. Chem. 67, 159-184 (2016)

Collective-variable-based methods

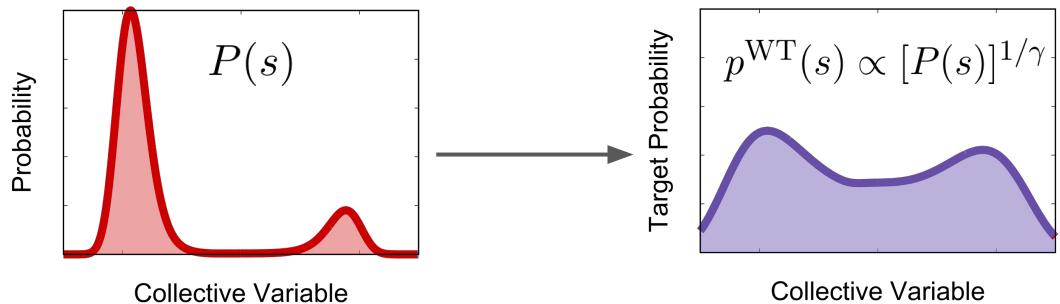
On-the-fly Probability Enhanced Sampling (OPES)

Targets the well-tempered distribution

$$V(s) = -\frac{1}{\beta} \log \frac{p^{\text{WT}}(s)}{P(s)}$$
$$V(s) = (1 - 1/\gamma) \frac{1}{\beta} \log P(s)$$

Iterative weighted KDE

$$P_n(s) = \frac{\sum_k^n w_k G(s, s_k)}{\sum_k^n w_k}$$



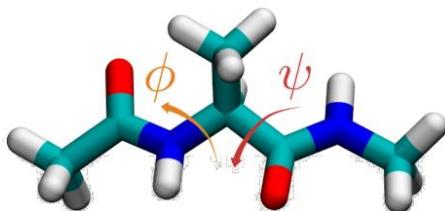
$$w_k = e^{\beta V_{k-1}(\mathbf{s}_k)}$$

$G(\mathbf{s}, \mathbf{s}_k)$: Gaussian kernel

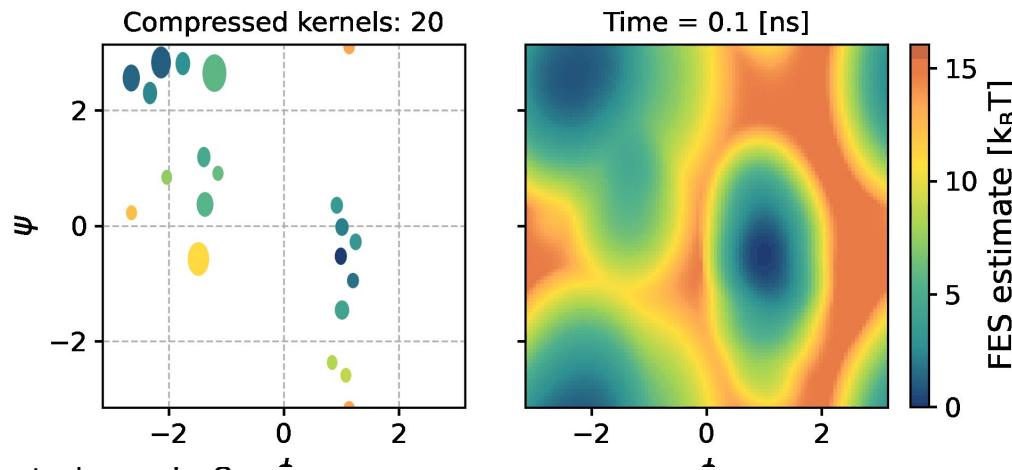
Collective-variable-based methods

On-the-fly Probability Enhanced Sampling (OPES)

Example: Alanine Dipeptide



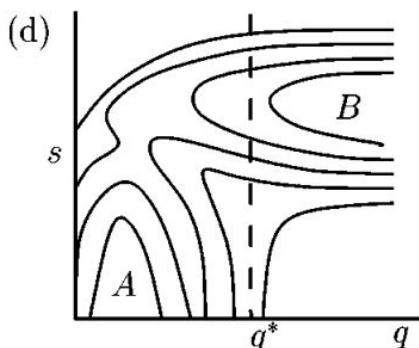
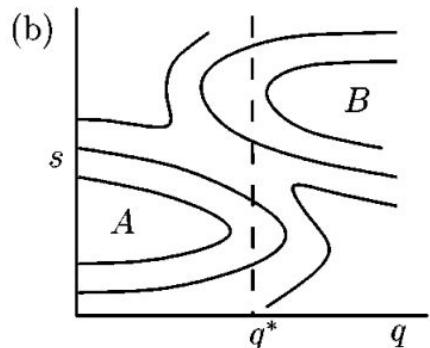
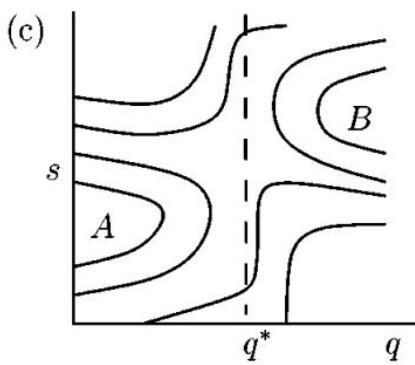
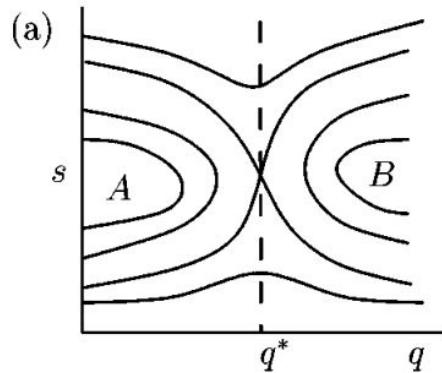
Conformational changes characterized by two CVs, namely, dihedral angles



Why is OPES becoming more popular than metadynamics?

- Very easy to use
- Fast convergence
- Virtually absent $c(t) \rightarrow$ stationary bias \rightarrow easy reweighting
- Kernel compression to avoid slowdown in the simulation

What is a good collective variable?



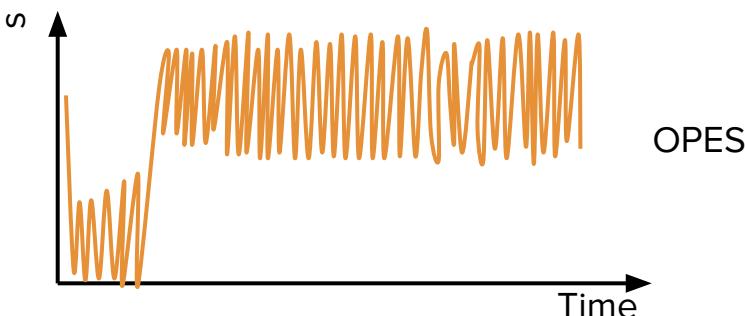
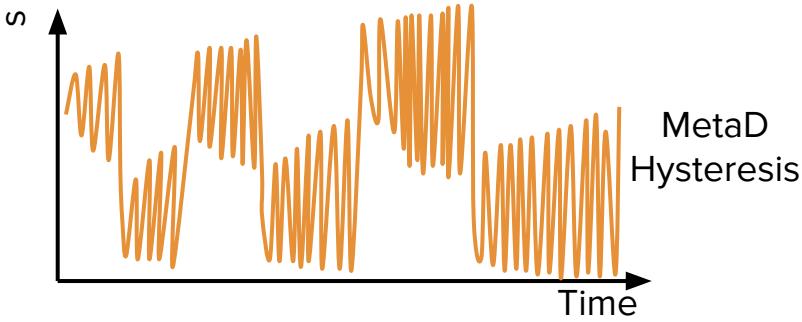
Good collective variables:

- distinguish well between metastable states
- Do not result in free energy barriers along other possible coordinates
- Describe the transition process faithfully
- Have committor distributions peaked at $p_A = 1/2$.

Why finding good CVs is important?

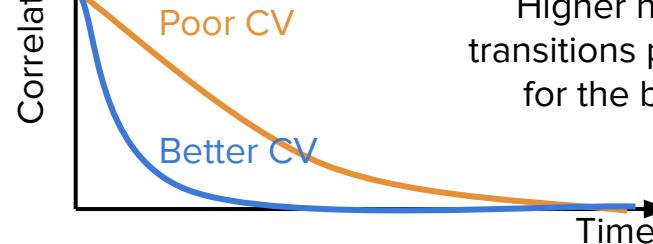
CV suboptimality

Suboptimal CVs are not able to drive reversible transitions effectively

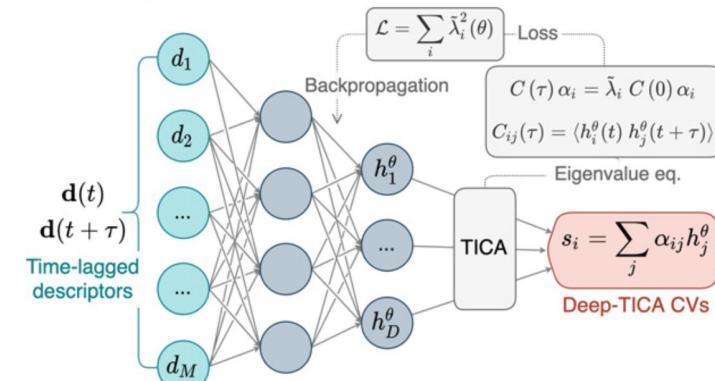


Good CVs improve performance

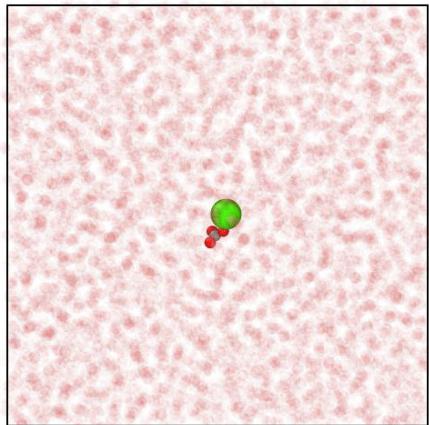
Higher number of transitions per unit time for the better CV



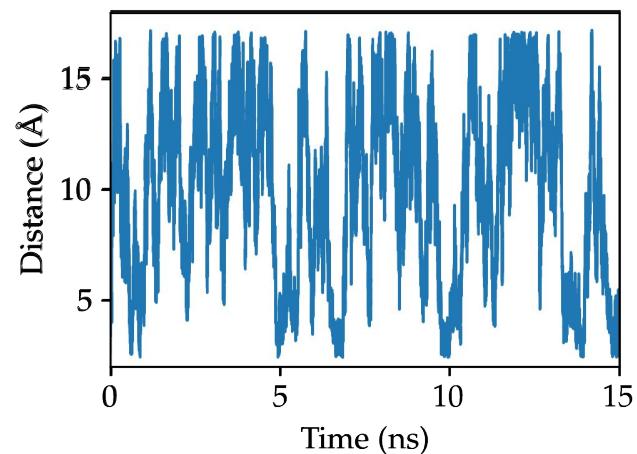
Data-driven CVs



Example 1: Calcium carbonate ion pair association in water



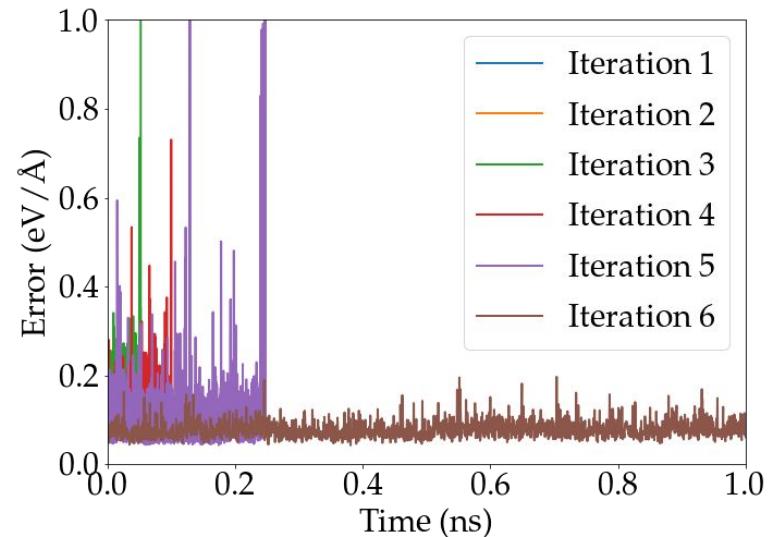
Deep potential model based on the SCAN DFT functional



OPES simulation using the distance between Ca and C as collective variable

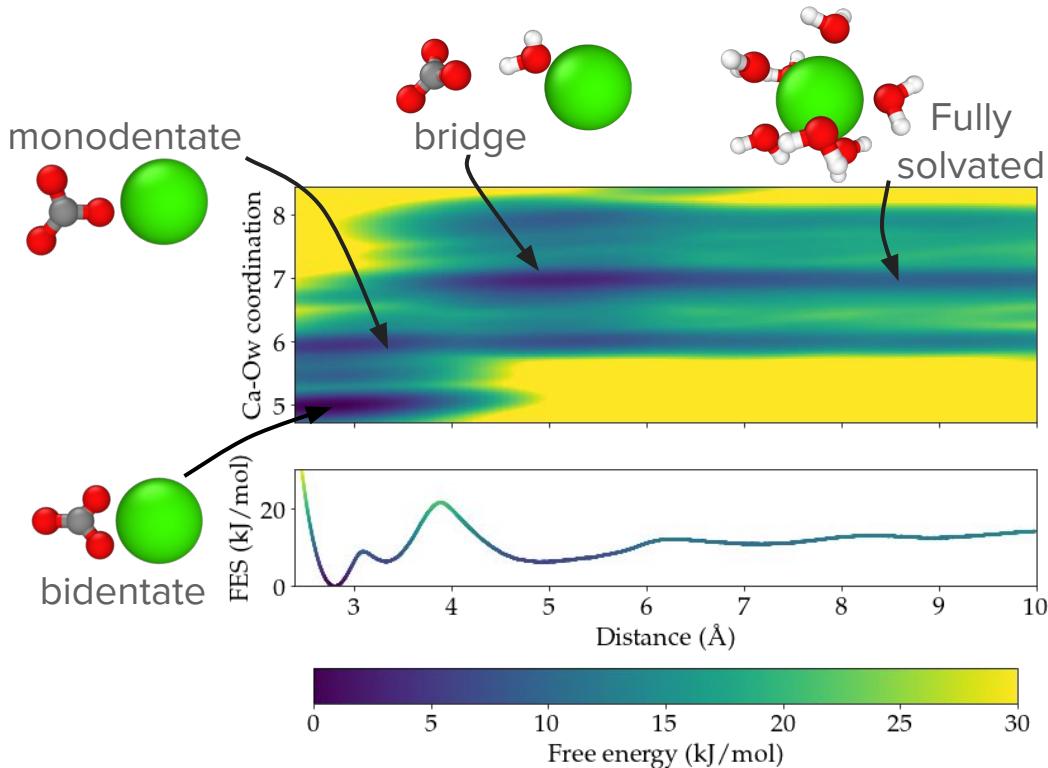
Active learning

Evolution of errors during enhanced sampling simulations for different iterations

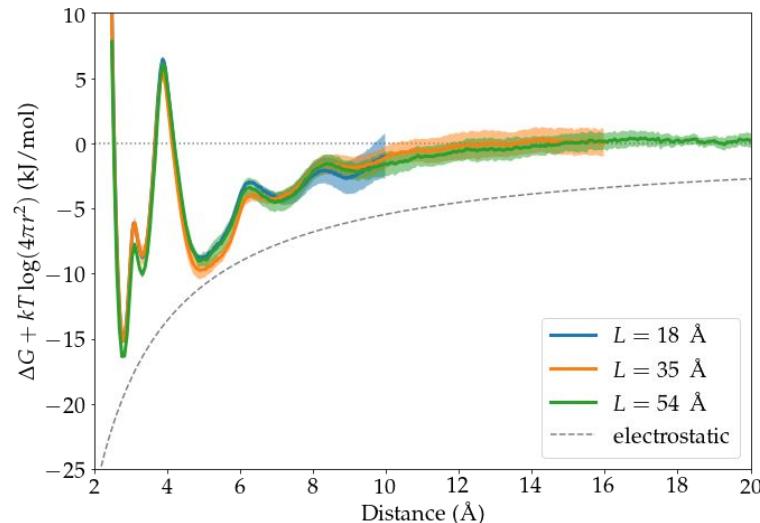


Example 1: Calcium carbonate ion pair association in water

Reweighting to compute the free energy surface (FES)



Behavior at long distances



$$\Delta G(r) = \frac{1}{4\pi\epsilon_r\epsilon_0} \frac{q_1 q_2}{r} - k_B T \log(4\pi r^2)$$

Long-range interactions are essential!

Example 2: Calculation of phase diagrams

Crystal structure-specific collective variables

Environment similarity

We start from a **definition of the density** around an atom:

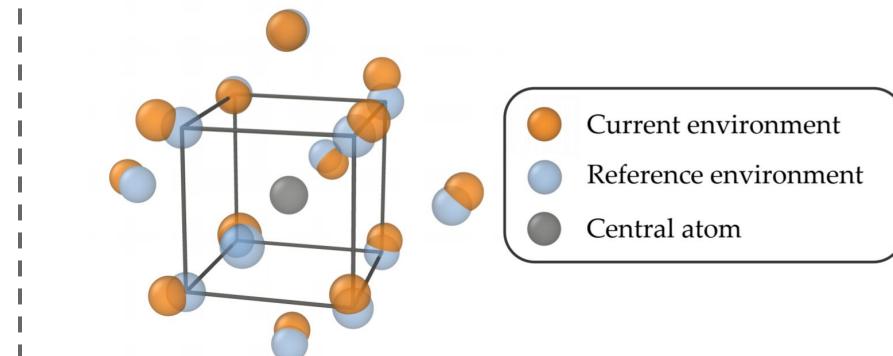
$$\rho_\chi(\mathbf{r}) = \sum_{i \in \chi} e^{-|\mathbf{r}_i - \mathbf{r}|^2 / 2\sigma^2}$$

We consider also a reference environment, and we **compare environments** using a kernel:

$$k_{\chi_0}(\chi) = \int d\mathbf{r} \rho_\chi(\mathbf{r}) \rho_{\chi_0}(\mathbf{r})$$

Analytic integration

$$k_{\chi_0}(\chi) = \sum_{i \in \chi} \sum_{j \in \chi_0} \pi^{3/2} \sigma^3 \exp \left(-\frac{|\mathbf{r}_i - \mathbf{r}_j^0|^2}{4\sigma^2} \right)$$

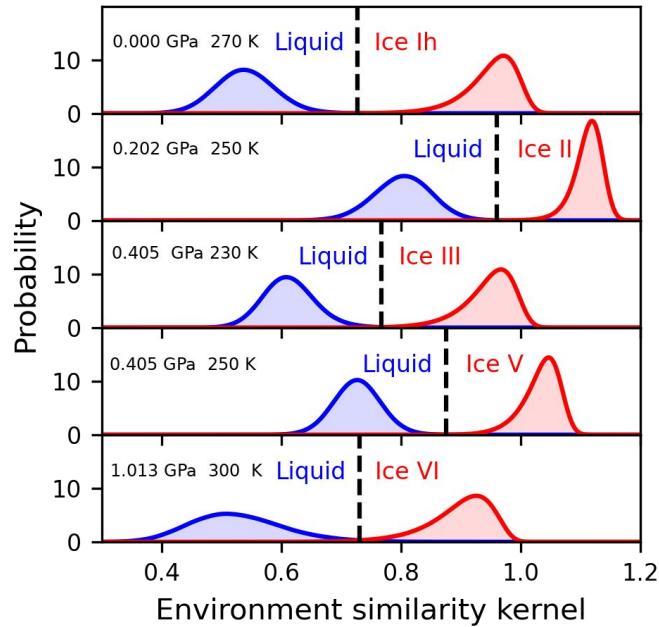
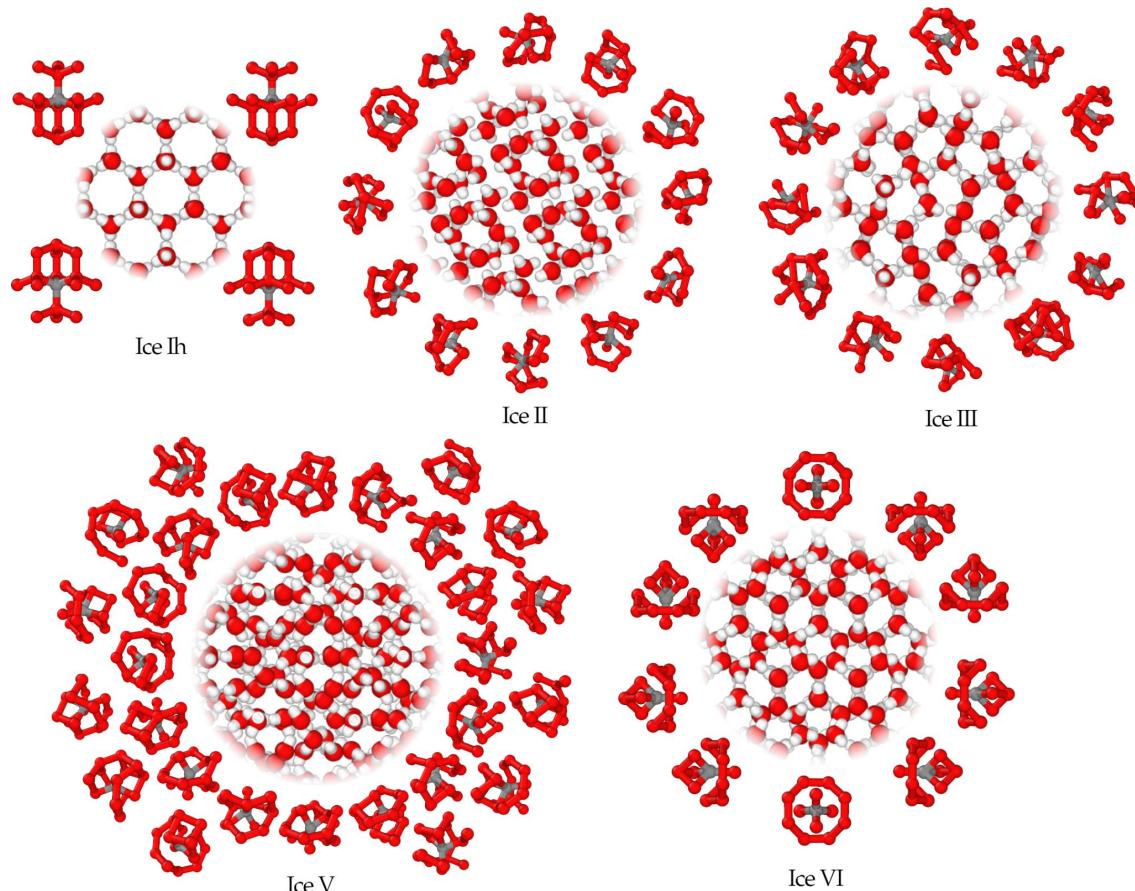


$$\tilde{k}_{\chi_0}(\chi) = \frac{1}{n} \sum_{i \in \chi} \sum_{j \in \chi_0} \exp \left(-\frac{|\mathbf{r}_i - \mathbf{r}_j^0|^2}{4\sigma^2} \right)$$

- Per atom quantity
- Cheap to compute

Example 2: Calculation of phase diagrams

Targeting structures with multiple environments

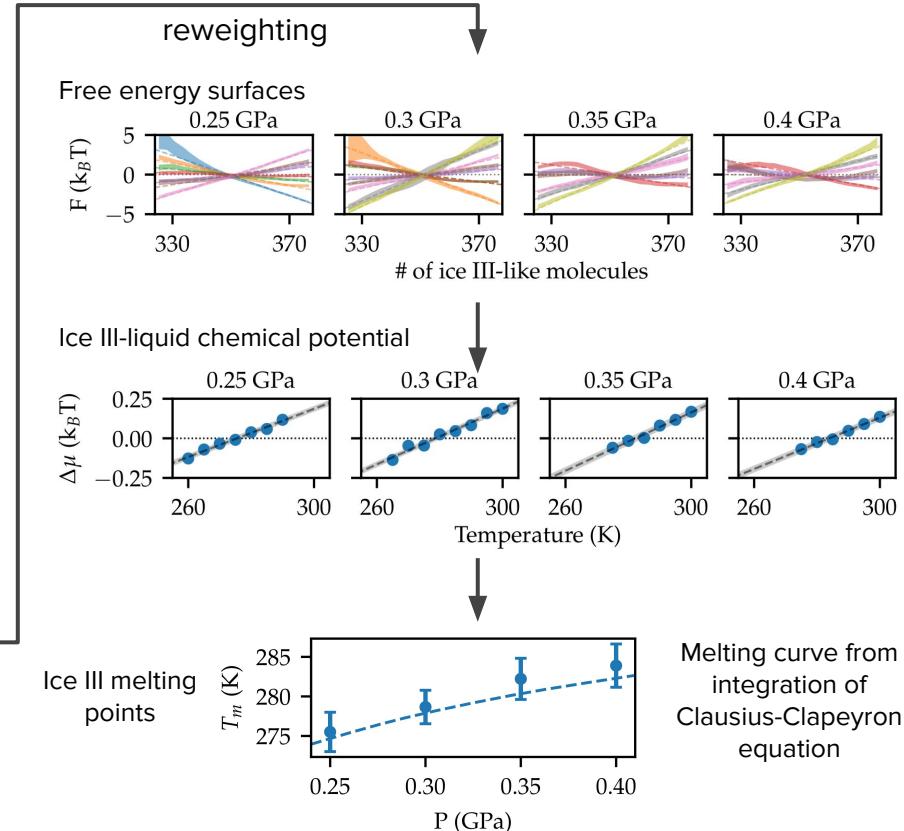
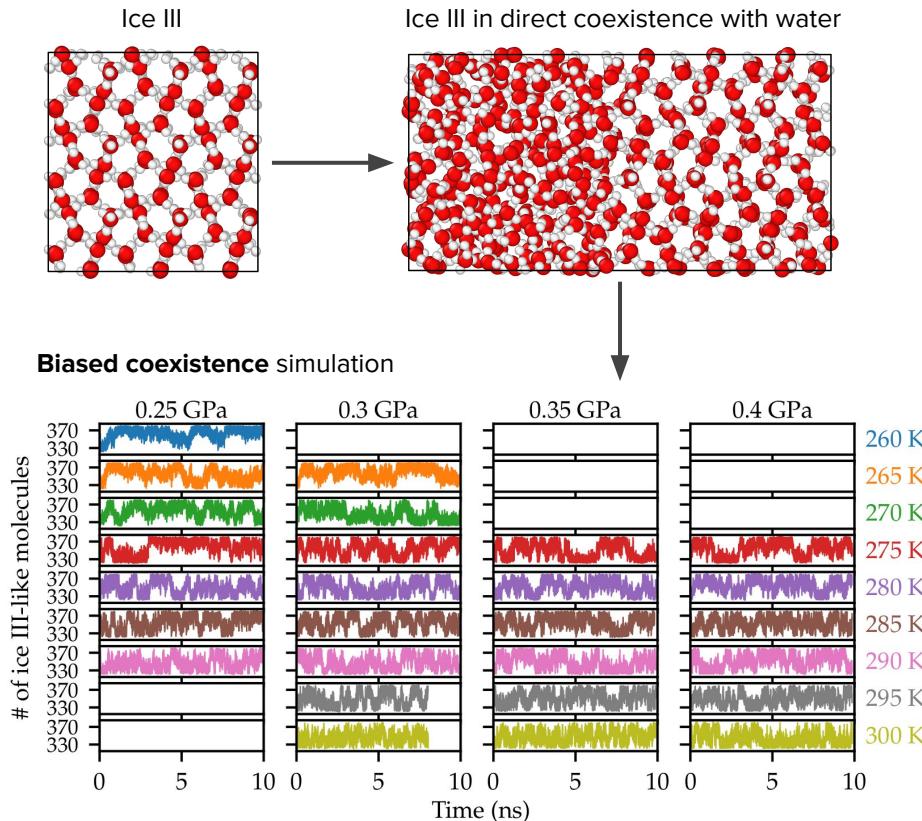


Count the number of molecules with ice-like environment

$$n_{ice} = \{ \text{number of } \chi^i : k_X(\chi^i) > \kappa \}$$

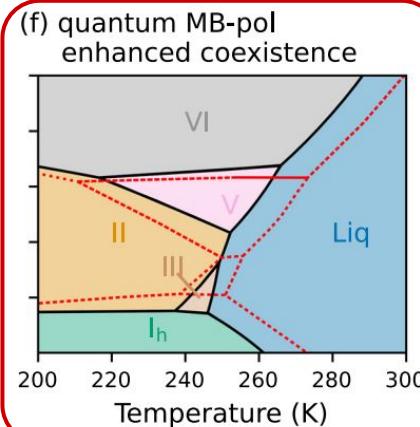
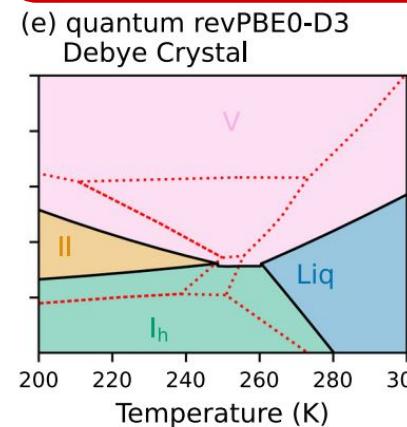
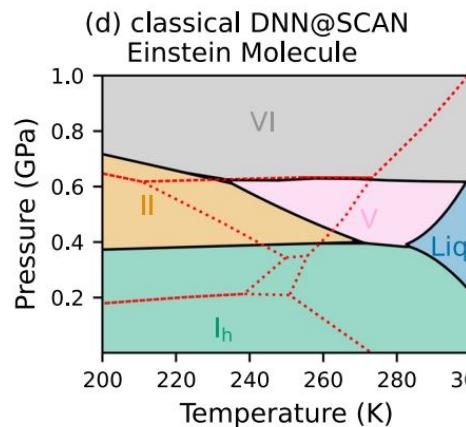
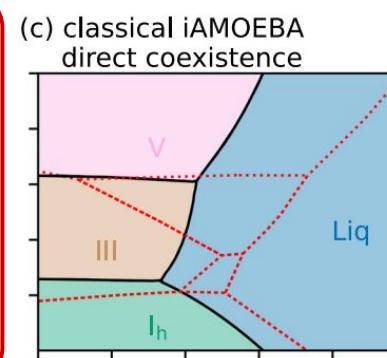
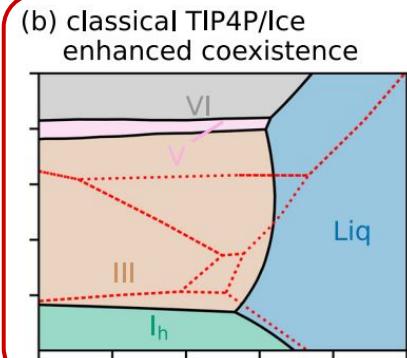
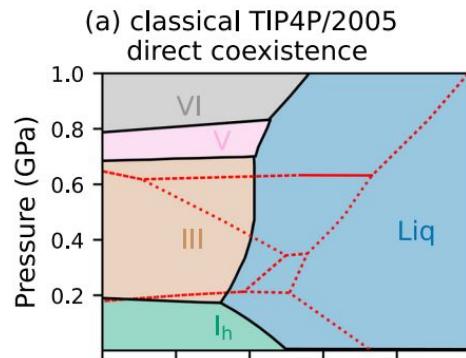
Example 2: Calculation of phase diagrams

Biased or enhanced coexistence



Bore, Piaggi, Car and Paesani, J. Chem. Phys. (2022)
Piaggi, Gartner, Car and Debenedetti, J. Chem. Phys. (2023)

Example 2: Calculation of phase diagrams



Using a deep
potential model
based on MB-pol

Excellent
agreement with
experiment!

The PLUMED enhanced sampling plugin



Home (master) Getting Started Tutorials Index of Actions Search

PLUMED

- ▶ Introduction
- ▶ Change Log
- ▶ Installation
- ▶ Getting Started
- ▶ Collective Variables
- ▶ Analysis
- ▶ Bias
- ▶ Additional Modules
- ▶ Command Line Tools
- ▶ Miscellaneous
- ▶ Tutorials
- ▶ Performances
- ▶ Index of Actions
- ▶ Todo List
- ▶ Bug List
- ▶ Bibliography

Introduction

PLUMED is a plugin that works with a large number of molecular dynamics codes ([Codes interfaced with PLUMED](#)). It can be used to analyze features of the dynamics on-the-fly or to perform a wide variety of free energy methods. PLUMED can also work as a [Command Line Tools](#) to perform analysis on trajectories saved in most of the existing formats. If PLUMED is useful for your work please read and cite [102], if you are interested in the PLUMED 1 original publication please read and cite [16].

To follow the development of PLUMED 2, you can look at the detailed [Change Log](#).

To install PLUMED, see this page: [Installation](#), while in [Getting Started](#) you can find a brief introduction on how to write your first PLUMED input file.

[Tutorials](#) are available to introduce basic as well as more advanced features of PLUMED.

About this manual

This manual has been compiled from PLUMED version 2.7.0-dev (git version: [9a30f8e](#)). Manual built on Travis CI for branch master.

Regtest results for this version can be found [here](#).

- Methods presented here can be found in PLUMED
- Interfaced with most MD engines, e.g. LAMMPS, Gromacs, etc
- Well integrated with the DeePMD-kit
- Many different biasing methods and collective variables
- Easy to program new features
- Large community
- Many tutorials and example inputs
- PLUMED NEST with examples and "recipes"

Tribello et al, Comput. Phys. Commun., 185, 2 604 (2014)
Bonomi et al, Nature methods 16 (8), 670-673 (2019)

Take-home messages

- **Enhanced sampling methods** aim at increasing the probability of observing rare but interesting phenomena in molecular simulations
- Three main ingredients: **collective variables**, **bias potential**, and **reweighting**
- These methods are useful for studying microscopic mechanisms of reactions and phase transitions, computing free energy surfaces, studying phase equilibria, crystal structure prediction, and training deep potentials.
- **PLUMED** is a open source plugin for enhanced sampling simulation interfaced with LAMMPS and the DeePMD-kit
- Try these methods out! There is a friendly community behind them ...

Thank you for your attention!
Questions?

Acknowledgments:

- Collaborators: Jack Weis, Annabella Selloni, Thanos Panagiotopoulos, Pablo Debenedetti, Roberto Car, Haiyang Niu, Michele Invernizzi, Michele Parrinello, Julian Gale, Paolo Raiteri
- Funding: U.S. Department of Energy, Office of Basic Energy Sciences, Swiss National Science Foundation, Ikerbasque (the Basque Foundation for Science)