106° Congresso Nazionale SIF - September 2020

A Unified Approach to Enhanced Sampling

Michele Invernizzi[1,2], Pablo M. Piaggi[3], Michele Parrinello[1,4]

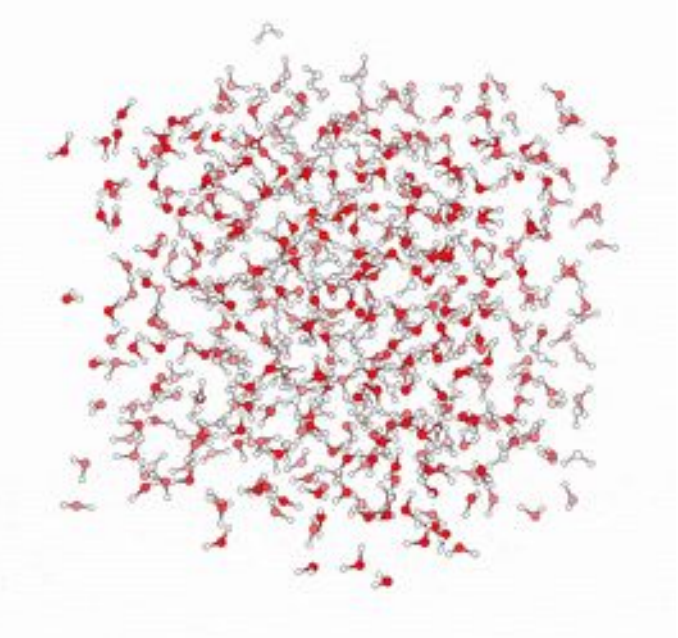
1. ETH zürich
Università della Svizzera italiana
2. MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH
FNSNF
SWISS NATIONAL SCIENCE FOUNDATION
3. PRINCETON UNIVERSITY
4. iit ISTITUTO ITALIANO DI TECNOLOGIA

Topic

Computational methods for molecular simulations

Outline

1. Enhanced sampling
2. Unified perspective
3. The method in action



Invernizzi, Piaggi, and Parrinello. “A Unified Approach to Enhanced Sampling.” arXiv:2007.03055 (2020)

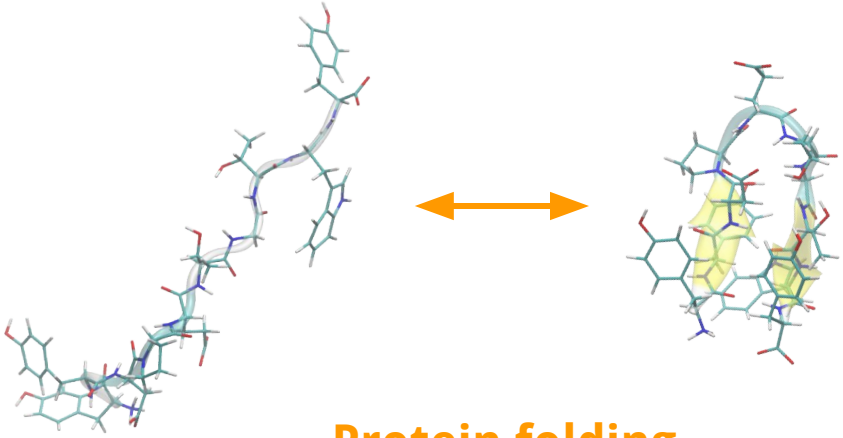
1. Enhanced sampling

tempering and collective variables approaches

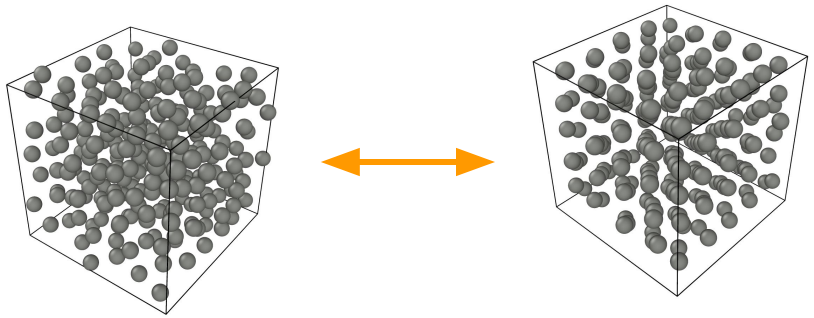
The sampling problem

Atomistic simulations are a great tool to **describe and predict** phenomena in physics, chemistry, biology, and material science. However, many phenomena are out of reach

=> **rare events**, that happen on **macroscopic timescales**



Protein folding

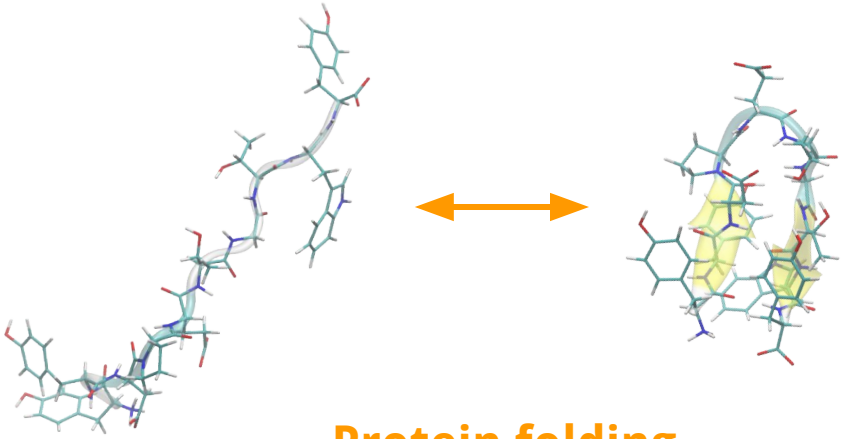


Phase transitions

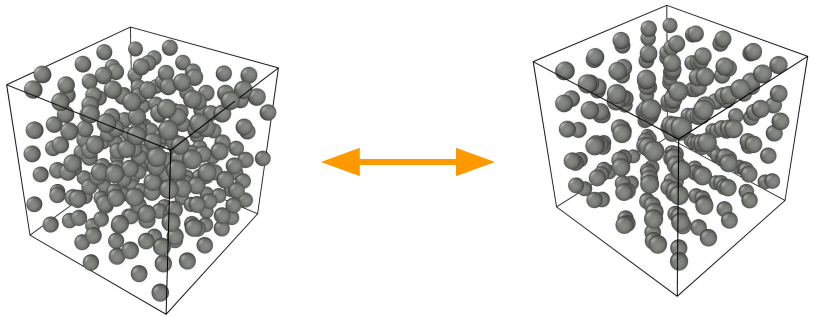
The sampling problem

The simulations remains stuck in one of the **metastable** states

Might take years of computation to observe just one **transition** event!



Protein folding



Phase transitions

The sampling problem

Many **enhanced sampling** methods have been proposed to overcome this problem

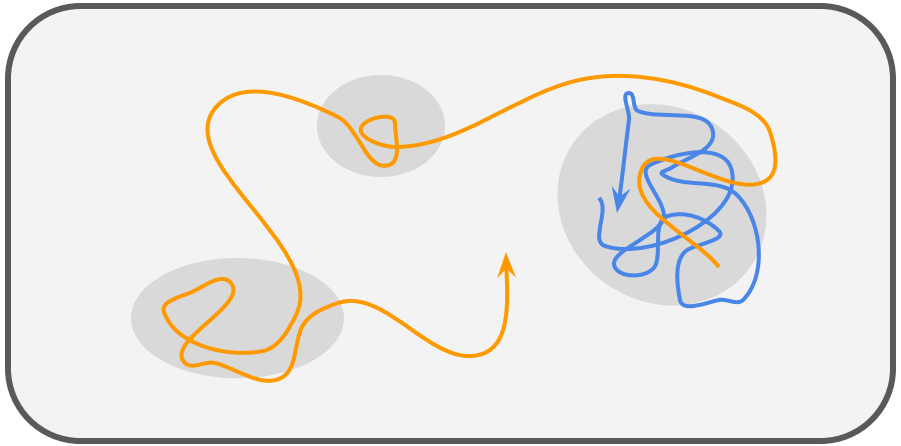
Two main families, with complementary approaches:

- Tempering methods
- Collective variables methods

We propose a unifying perspective **and a novel** general method **that allows to combine them and opens up to new possibilities**

Tempering methods

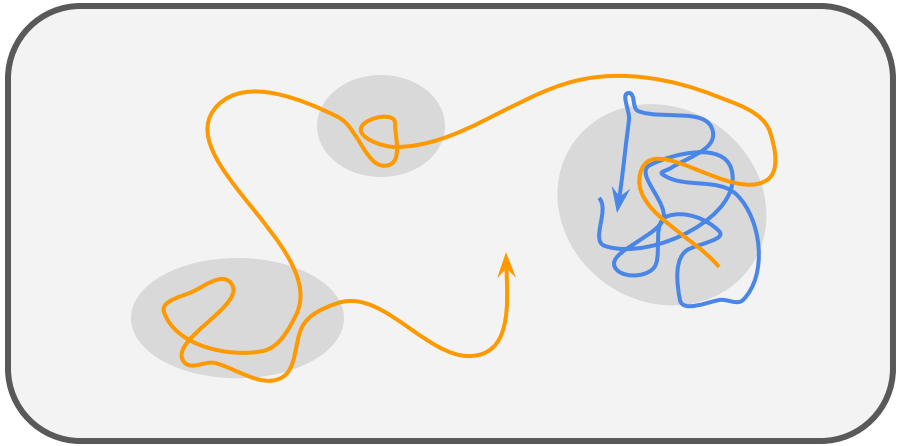
At higher temperature the system explores a larger portion of phase space, and escapes metastable basins **=>** combine **different temperatures** to enhance the sampling



- *Parallel tempering*
- *Simulated tempering*
- *Replica exchange*
- *Multicanonical methods*
- *...*

Tempering methods

At higher temperature the system explores a larger portion of phase space, and escapes metastable basins => combine **different temperatures** to enhance the sampling

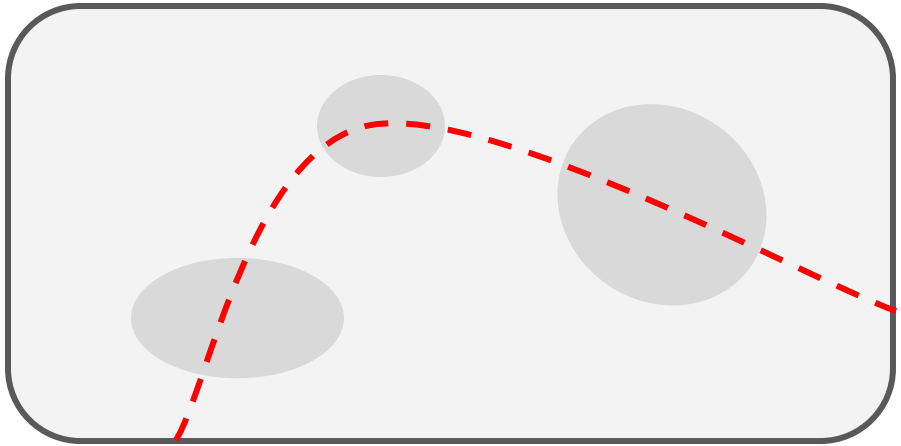


- *Parallel tempering*
- *Simulated tempering*
- *Replica exchange*
- *Multicanonical methods*
- *...*

Expanded ensembles not only in temperature

Collective variables methods

Identify an order parameter or **collective variable (CV)** that describes the process

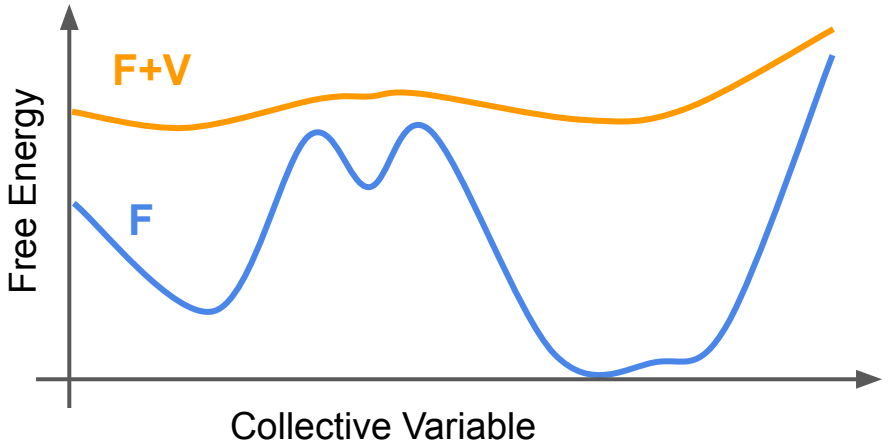


- *Umbrella sampling*
- *Metadynamics*
- *Variationally enhanced sampling*
- *...*

Collective variables methods

Identify an order parameter or **collective variable (CV)** that describes the process

Add a **bias potential V** that flatters the **free energy F**

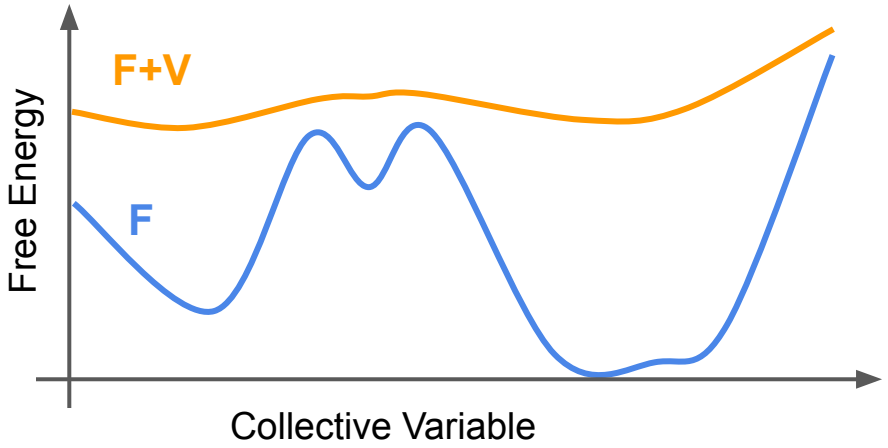


- *Umbrella sampling*
- *Metadynamics*
- *Variationally enhanced sampling*
- *...*

Collective variables methods

Identify an order parameter or **collective variable (CV)** that describes the process

Add a **bias potential V** that flatters the **free energy F**



- *Umbrella sampling*
- *Metadynamics*
- *Variationally enhanced sampling*
- *...*

More efficient than tempering, but CV might be hard to find!

2. Unified perspective

reaching a target distribution

The target distribution

- The **equilibrium Boltzmann distribution** is hard to sample
- Enhanced sampling methods explicitly or implicitly aim at sampling a different **target distribution**
- Via some **reweighting** technique it is then possible to retrieve statistics on $P(\mathbf{x})$

$$P(\mathbf{x}) = \frac{e^{-u(\mathbf{x})}}{Z}$$

$$p^{tg}(\mathbf{x})$$

We develop a biasing scheme that can be used to sample the target distributions of both tempering and collective variables methos

The OPES method

On-the-fly Probability Enhanced Sampling (OPES):

- We can sample any $p^{tg}(\mathbf{x})$ by adding the proper bias potential

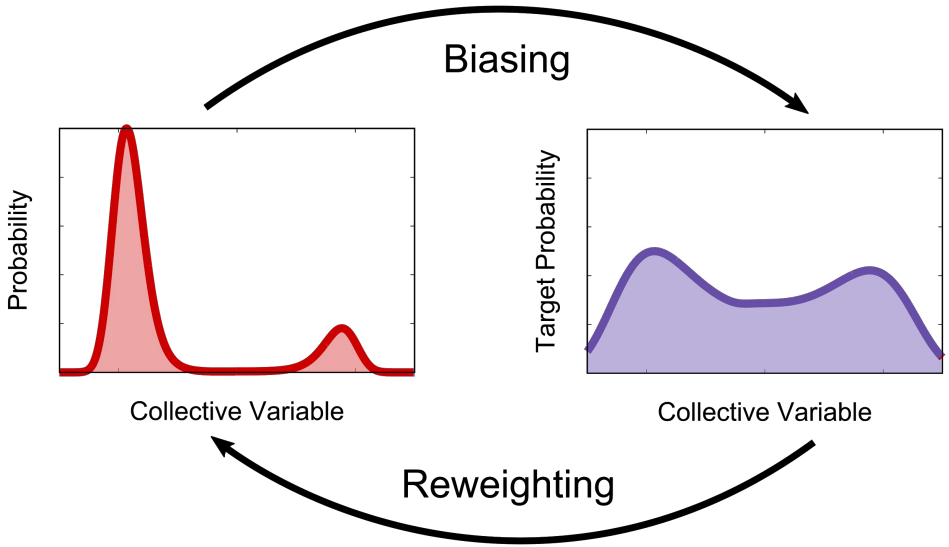
$$v(\mathbf{x}) = -\log \frac{p^{tg}(\mathbf{x})}{P(\mathbf{x})}$$

- Since $P(\mathbf{x})$ is unknown we set up an iterative scheme based on on-the-fly reweighting and adiabatic evolution of the bias

Let's first see the case of collective variables enhanced sampling using OPES

The OPES method: CVs

$$v(s) = -\log \frac{p^{tg}(s)}{P(s)}$$



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

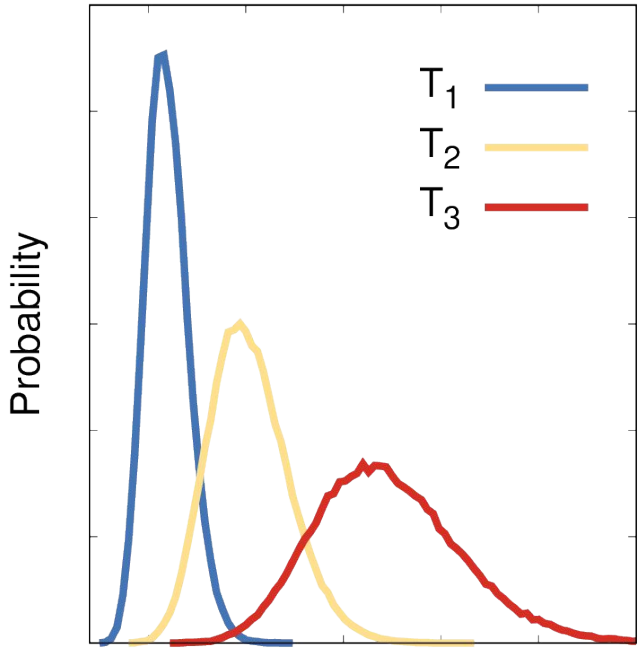
- Collective variable $s = s(\mathbf{x})$
- Target distribution is chosen by requiring its **marginal** $p^{tg}(s)$ to be easy-to-sample
- We only need to reconstruct the probability along the CV, we use a weighted kernel density estimation

Invernizzi, and Parrinello. "Rethinking Metadynamics: from bias potentials to probability distributions." **J. Phys. Chem. Lett. 11.7 (2020)**

The OPES method: Tempering

The target distribution is a combination of different temperatures

$$p^{tg}(\mathbf{x}) = \frac{1}{N} \sum_{T} P_T(\mathbf{x})$$



The OPES method: Tempering

The target distribution is a combination of different temperatures

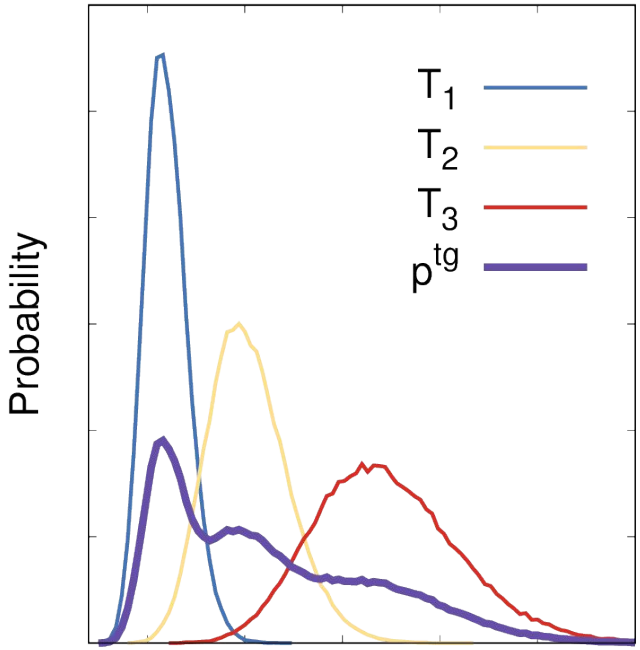
$$p^{tg}(\mathbf{x}) = \frac{1}{N} \sum_T P_T(\mathbf{x})$$

The bias can be written as

$$v(\mathbf{x}) = -\log \left[\frac{1}{N} \sum_T e^{-\Delta u_T(\mathbf{x}) + \Delta F(T)} \right]$$

Where the free energy difference as a function of temperature is iteratively estimated via reweighting

$$\Delta F(T) = -\log \left[\frac{\left\langle e^{-\Delta u_T + v} \right\rangle_{p^{tg}}}{\left\langle e^{v} \right\rangle_{p^{tg}}} \right]$$

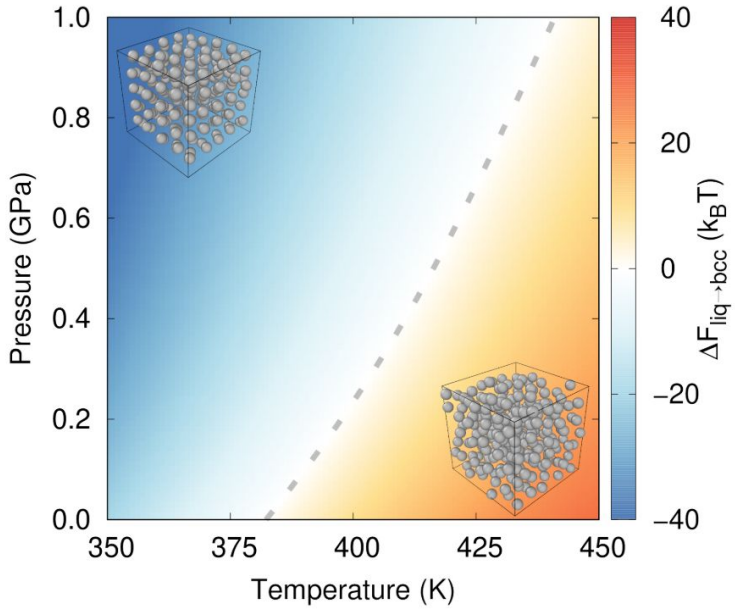


3. The method in action

a phase diagram from a single simulation

Sodium phase diagram

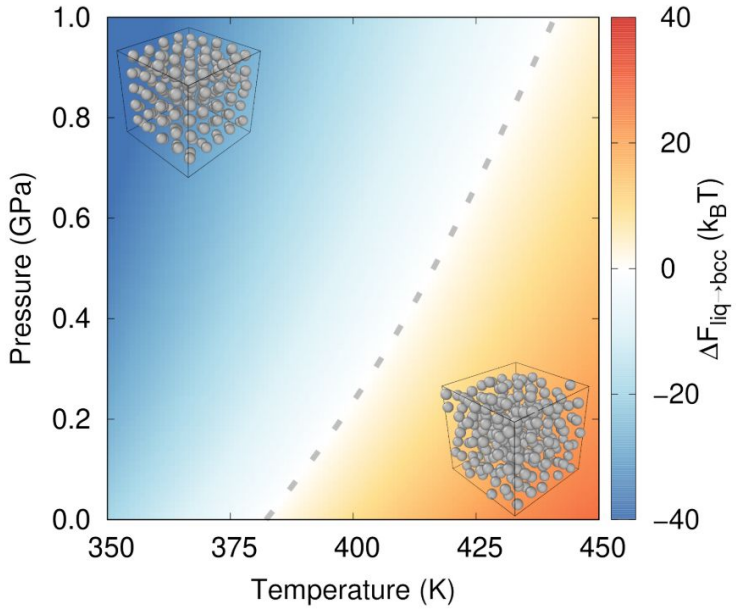
We want to calculate the phase diagram of sodium around its first order **solid-liquid** phase transition.



Sodium phase diagram

We want to calculate the phase diagram of sodium around its first order **solid-liquid** phase transition.

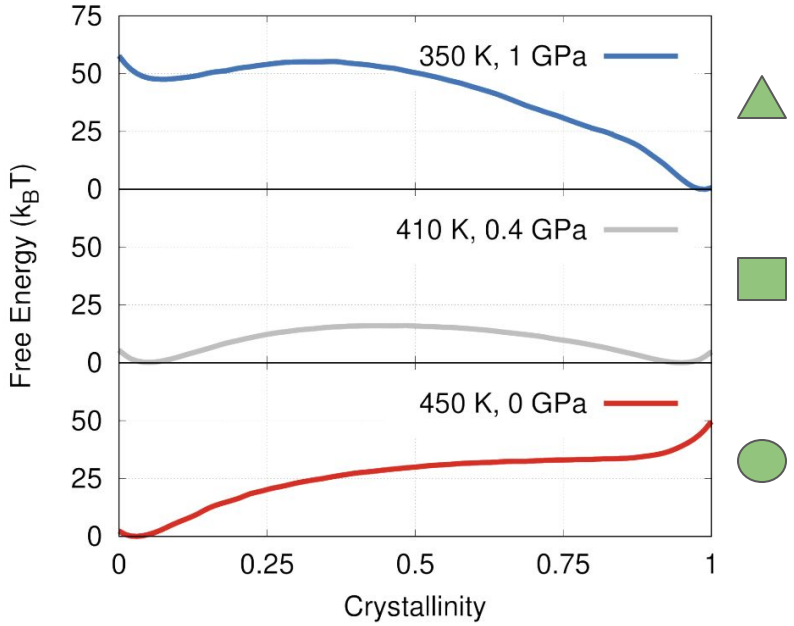
We combine tempering in **temperature** and **pressure** with the biasing of a **crystallization** collective variable[1].

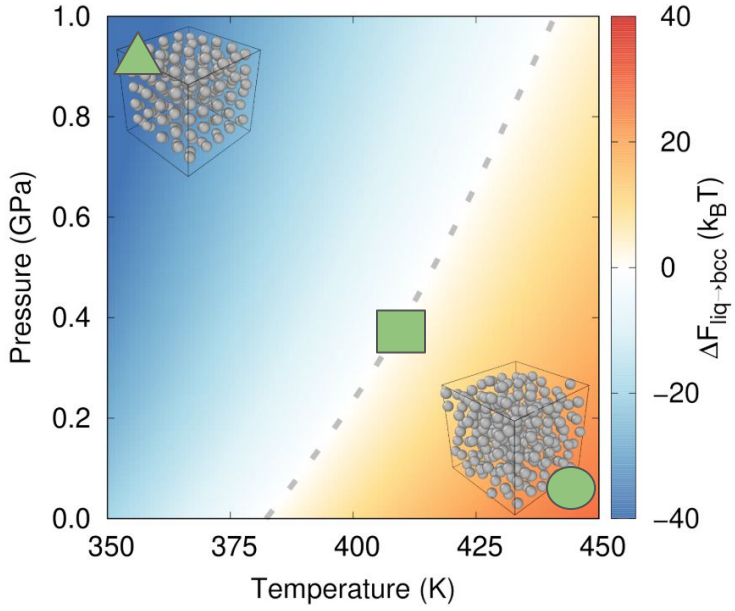


[1] Piaggi and Parrinello, J. Chem. Phys. 150, 244119 (2019)

Sodium phase diagram

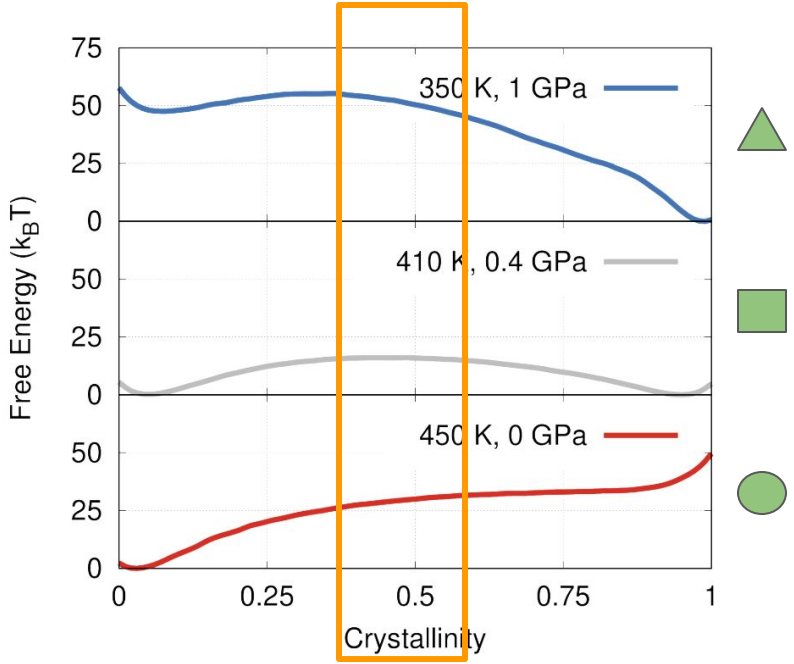
A multithermal-multibaric simulation alone would not be efficient

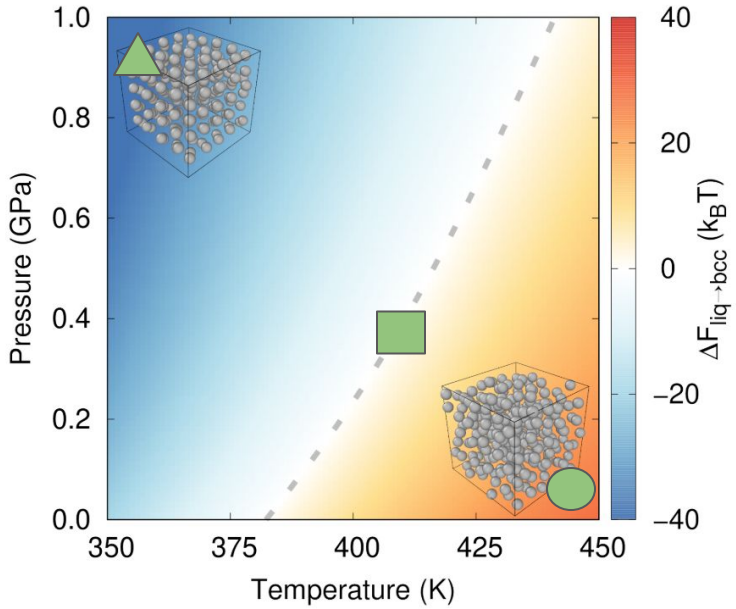




Sodium phase diagram

A multithermal-multibaric simulation alone would not be efficient

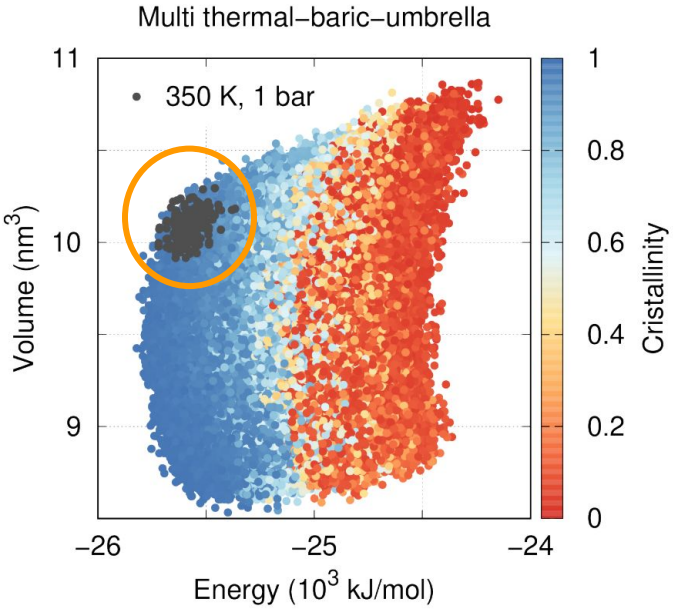




Sodium phase diagram

A standard simulation would sample only one phase and a narrow energy-volume region

Our method allows for a broad and efficient sampling



Summary

Using a target-distribution perspective we developed OPES method

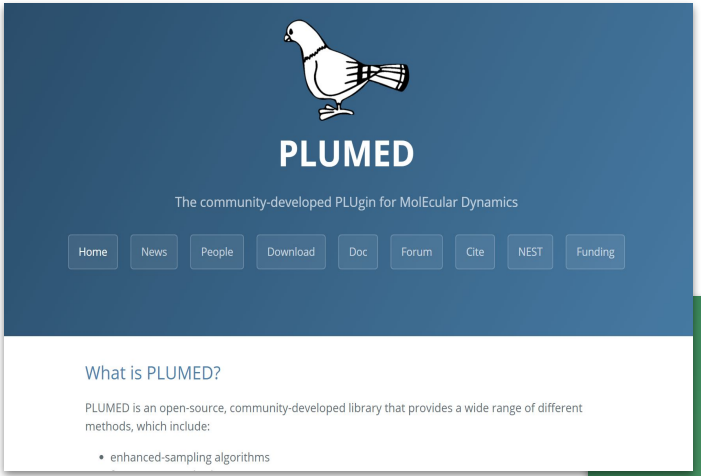
- General approach to various enhanced sampling strategies
- Efficient and robust
- Simple to use
- Scalable (no minimum number of replicas)
- Open source and portable

Available through PLUMED

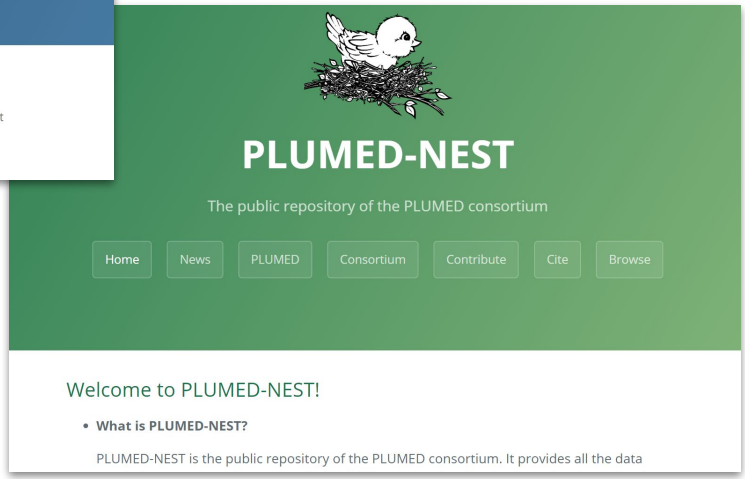
Implemented open source in PLUMED

www.plumed.org

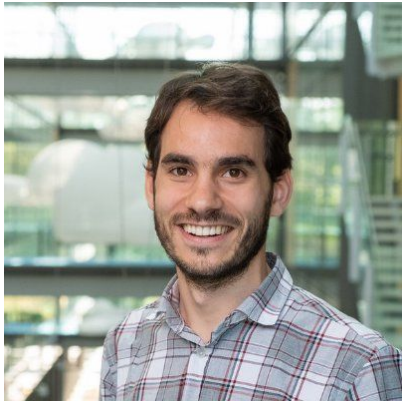
Can be used with LAMMPS, GROMACS, OpenMM, NAMD, QESPRESSO, CP2K, …



Examples on PLUMED-NEST



Acknowledgements



- Pablo M. Piaggi
- Parrinello Group