

A silhouette of a person climbing a steep, rocky mountain peak. The sun is rising in the background, creating a bright glow and long shadows. The mountain's surface is rugged and textured. The sky is a clear, pale blue.

# Studying rare events in molecular dynamics simulations

## From free energies to mechanistic insights

Dr. Kristof Bal

Han-sur-Lesse Winterschool in Theoretical Chemistry and Spectroscopy

29 November – 3 December 2021

## Format of these lectures

Classic lecture with slides

- Figures & formulas
- Lots of references for further exploration

“Hands on”

- **Example X:** simulation setups + output (PLUMED + LAMMPS/CP2K)
- Some preliminary analysis as a Jupyter Notebook
- Head start for those who wish to dive deeper

<https://github.com/kbal/winterschool2021>



I'm biased and don't have time!

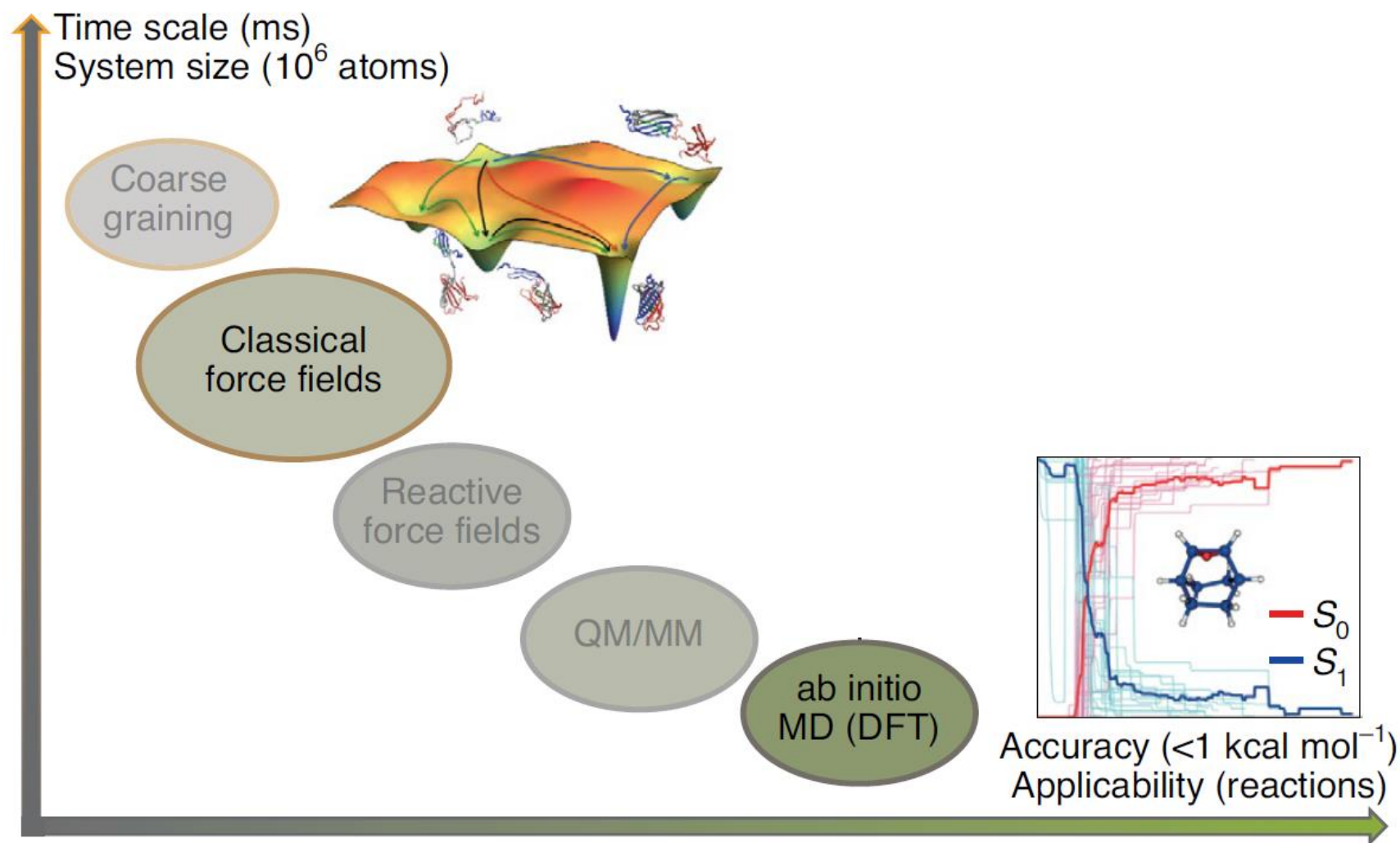
**Path not taken:** relevant topics shamelessly avoided



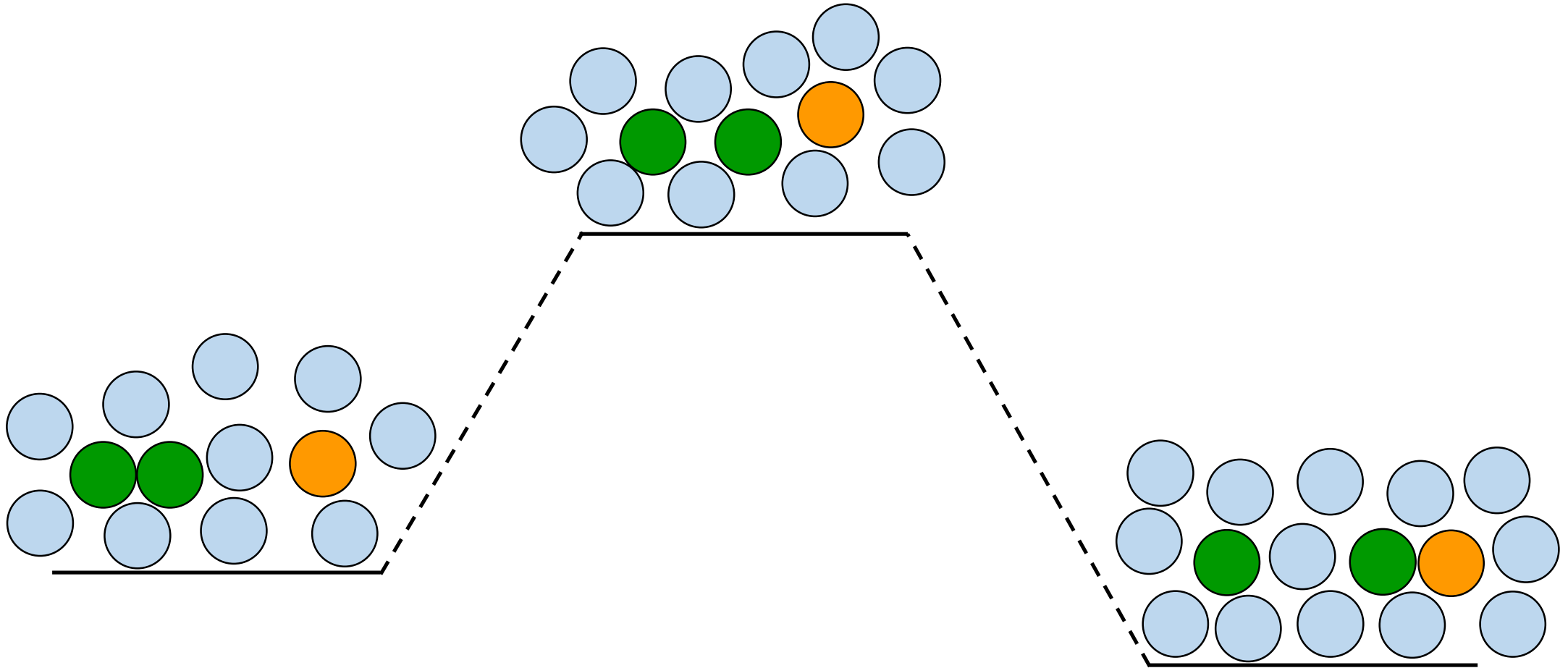
# Setting the scene

Thermodynamic sampling in molecular simulations

Assumption: we've built a microscopic model and know how to compute energies/forces as a function of system coordinates  $\mathbf{R}$ .

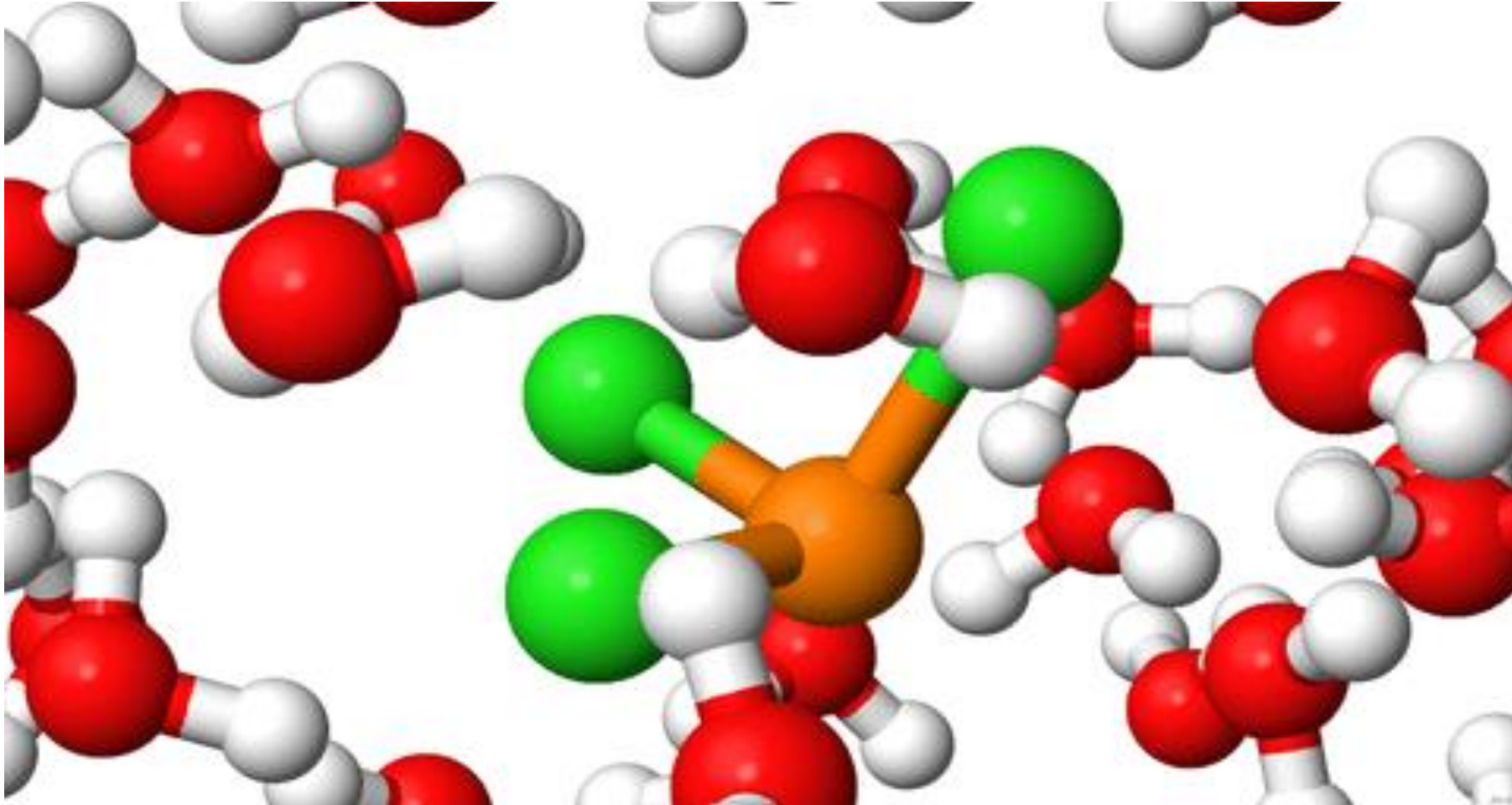


We pick an appropriate level of theory, identify relevant structures, and compute properties. That's it!



Or not!

Whenever we're modelling soft matter, we have no single "representative" configuration anymore



Statistical mechanics: we can express any macroscopic quantity of interest as averages over all microscopic configurations

$$\langle O \rangle = \frac{1}{Z} \sum_i O_i e^{-\beta E_i} \text{ with } Z = \sum_i e^{-\beta E_i}$$

Or, in a classical system:

$$\langle O \rangle = \frac{1}{Z} \int d\mathbf{R} O(\mathbf{R}) e^{-\beta E(\mathbf{R})} \text{ with } Z = \int d\mathbf{R} e^{-\beta E(\mathbf{R})}$$

Impossible to evaluate numerically in any reasonable  $N$ -atom system!



Idea: Can we draw a meaningful sample of  $n$  configurations so that

$$\langle O \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_i O_i$$

That is, generate a subset of all configurations that, on average, has the same properties/distribution as the full ensemble?

Historically, the two computer simulation approaches that aim to do this are **molecular dynamics (MD)** and **Monte Carlo (MC)** simulations.

# Path not taken: (Metropolis) Monte Carlo

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

## Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,  
*Los Alamos Scientific Laboratory, Los Alamos, New Mexico*

AND

EDWARD TELLER,\* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

Molecular dynamics (MD) = integrate equations of motion in small time steps  $\Delta t$ .

Simplest form only requires forces  $\mathbf{f}$ , samples “NVE” ensemble

$$\mathbf{f} = m\mathbf{a}$$

In practice we'll use modified equations of motion to sample isothermal (canonical), isobaric, ... ensembles.

**Path not taken:** discussion of ensembles, finite size effects, model construction

In the ergodic limit, time averages are equivalent to spatial averages.

$$\langle O \rangle = \frac{1}{Z} \int d\mathbf{R} O(\mathbf{R}) e^{-\beta E(\mathbf{R})} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt O(t)$$

That is, we assume the simulated trajectory is long enough to visit all configurations of interest.

In practice, we average over finite simulations

$$\langle O \rangle \approx \frac{\Delta t}{t} \sum_i O(t_i)$$



Suppose a two state ( $A$  and  $B$ ) system. The probability of observing state  $A$  is

$$P_A = \frac{Z_A}{Z} = \frac{1}{Z} \int_A d\mathbf{R} e^{-\beta E(\mathbf{R})}$$

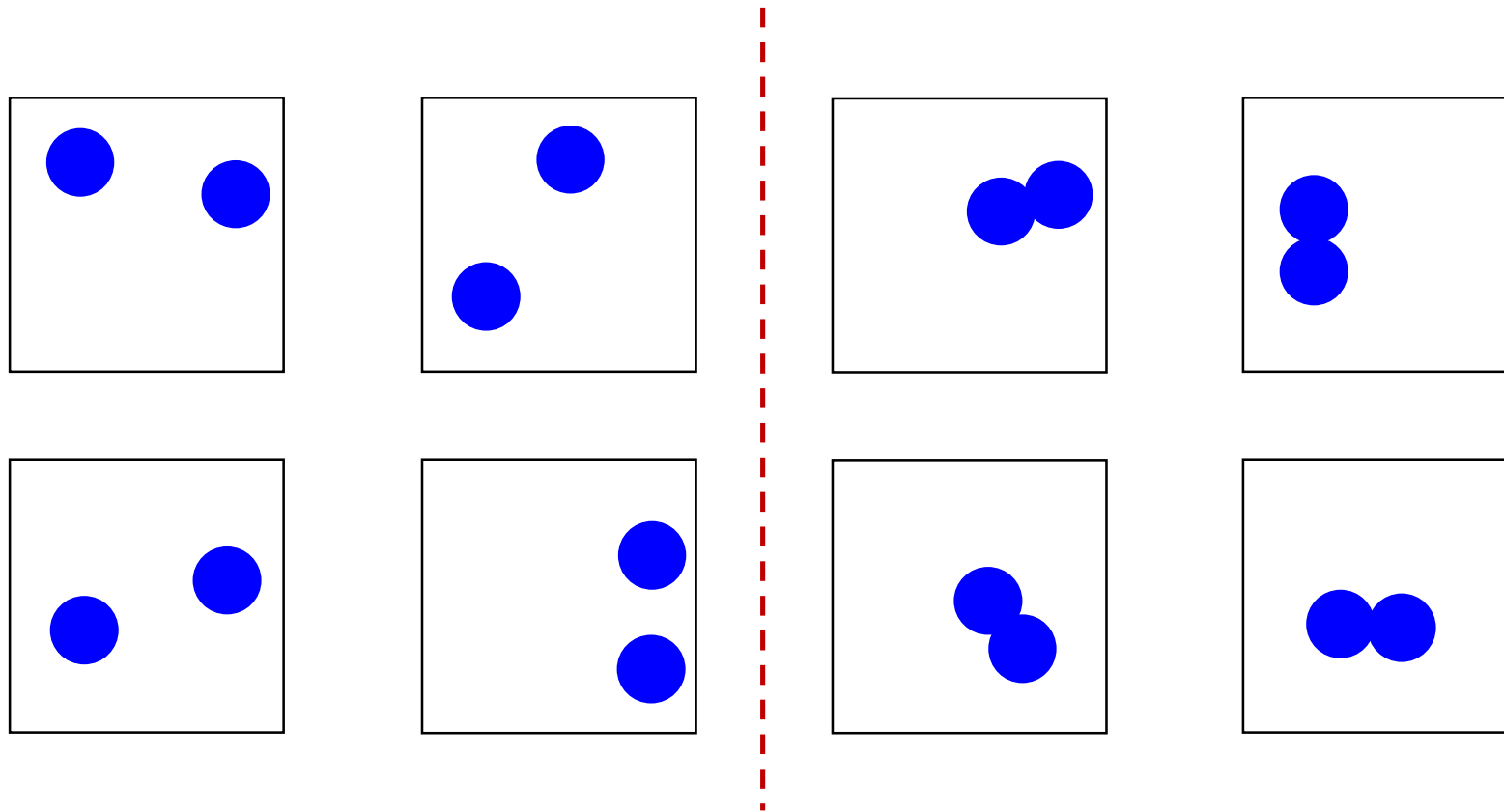
So that the equilibrium constant for the process  $A \leftrightarrow B$  is

$$K = \frac{P_B}{P_A} = \frac{Z_B}{Z_A}$$

This also means we get the free energy difference of states:

$$\Delta F = -\frac{1}{\beta} \ln K = -\frac{1}{\beta} \ln \frac{Z_B}{Z_A}$$

In an MD simulation,  $P_A \propto (\text{\#observed } A \text{ configurations})$



state  $A$

state  $B$

How to easily distinguish states in a simulation with a high-dimensional configuration space  $\mathbf{R}$ ?

→ Introduce small number of collective variables (CVs)  $\mathbf{s}(\mathbf{R})$  that capture main features of interest

We can now define the marginal probability density  $P(\mathbf{s})$  of the CVs

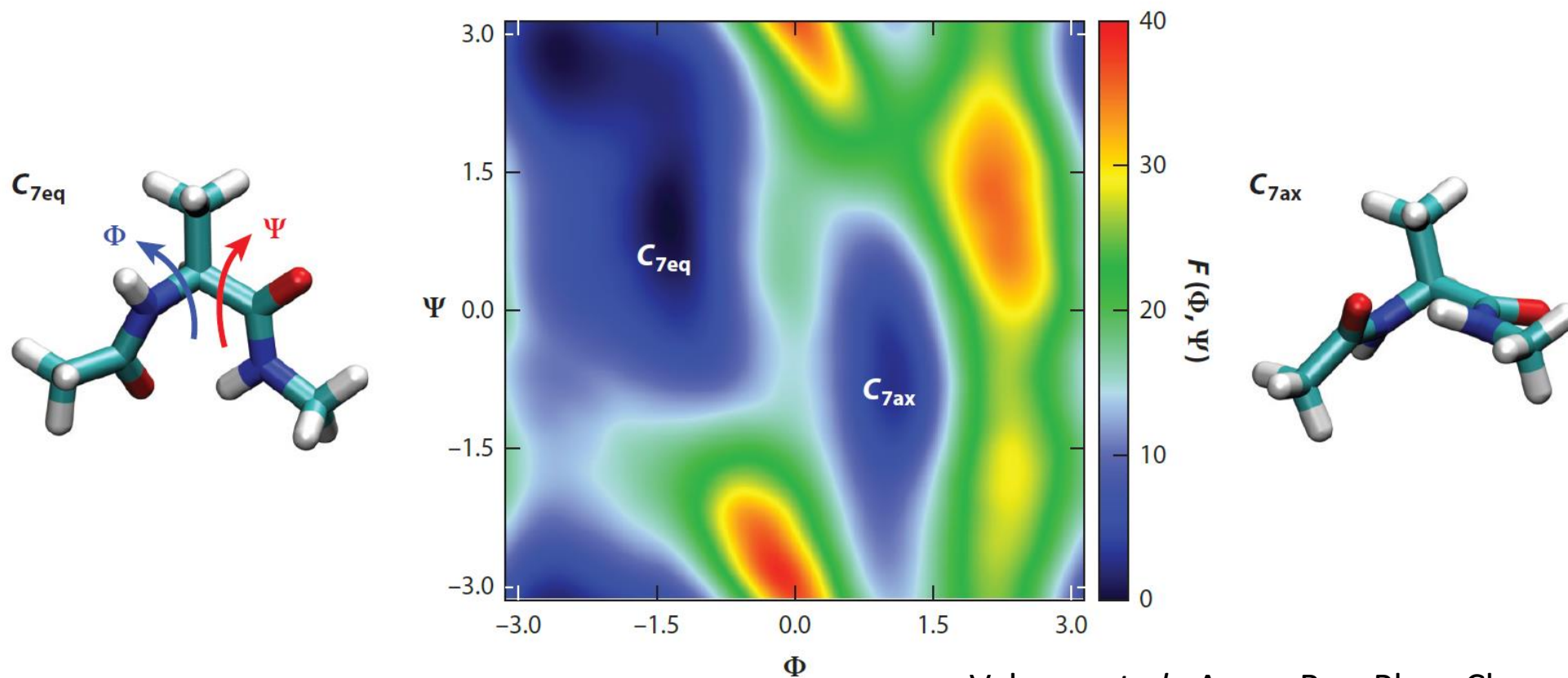
$$P(\mathbf{s}) = \frac{1}{Z} \int d\mathbf{R} \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] e^{-\beta E(\mathbf{R})}$$

Or, in practice: the sampled histogram of the CVs

$$P(\mathbf{s}) = \langle \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] \rangle = H(\mathbf{s})$$

Then, the free energy surface (FES) is:

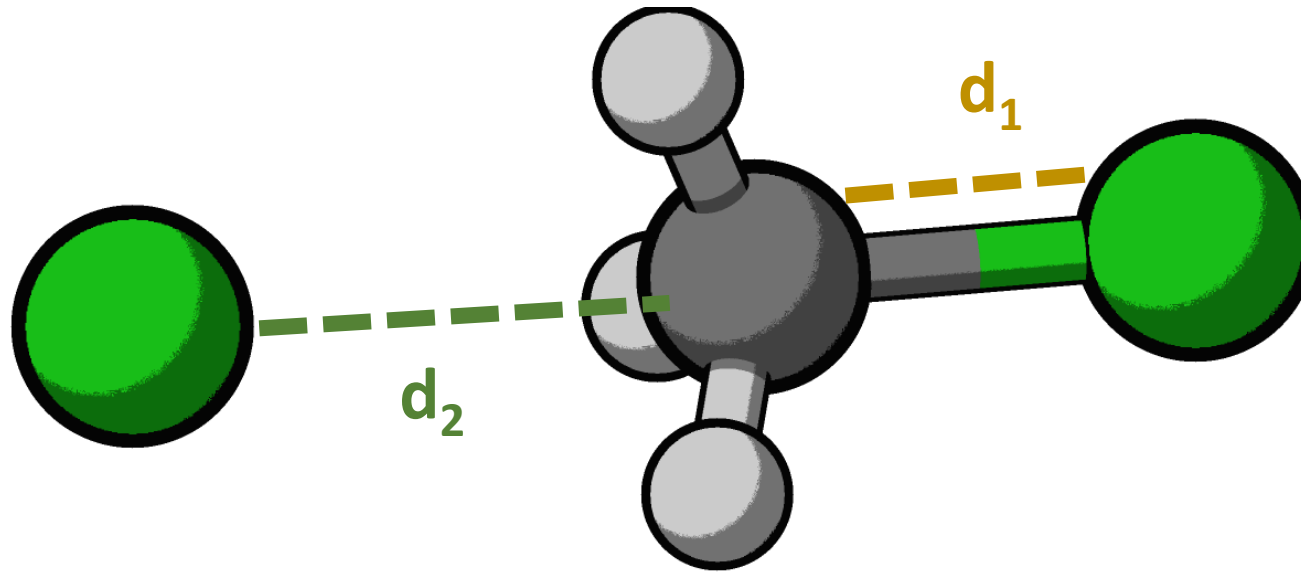
$$F(\mathbf{s}) = -\frac{1}{\beta} \ln P(\mathbf{s})$$





How easy is it to accumulate a histogram  $H(\mathbf{s})$ ?

**Example 1:** Let's take a simple chemical reaction





Many systems are dominated by **rare events**.

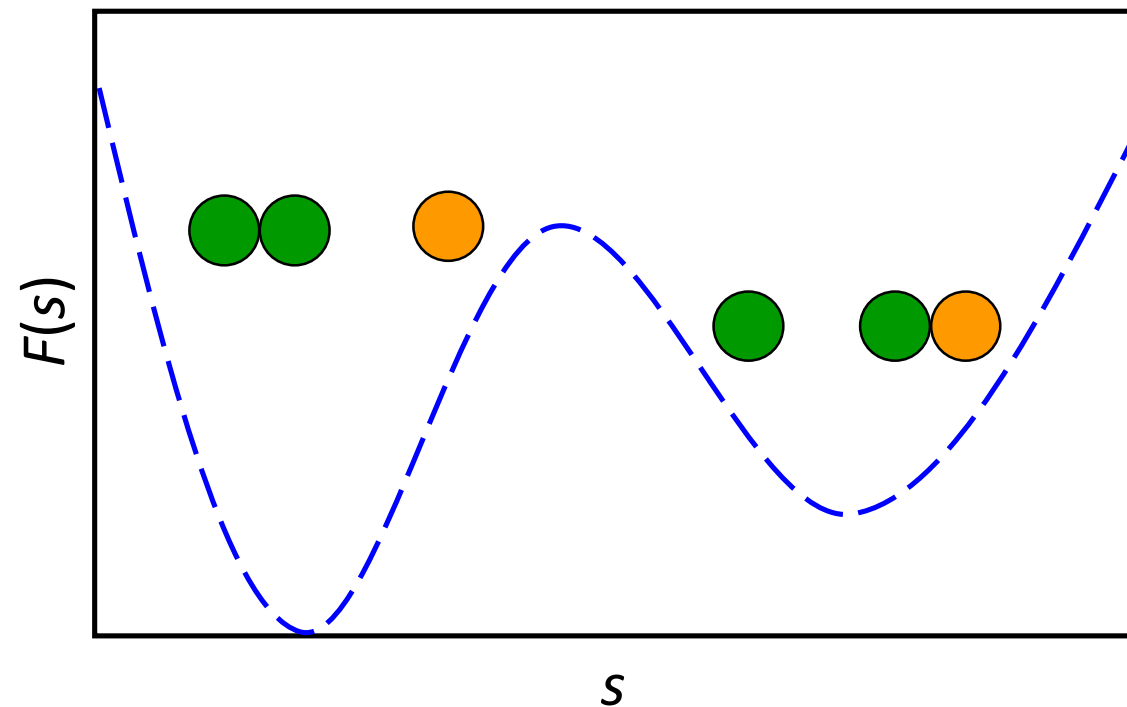
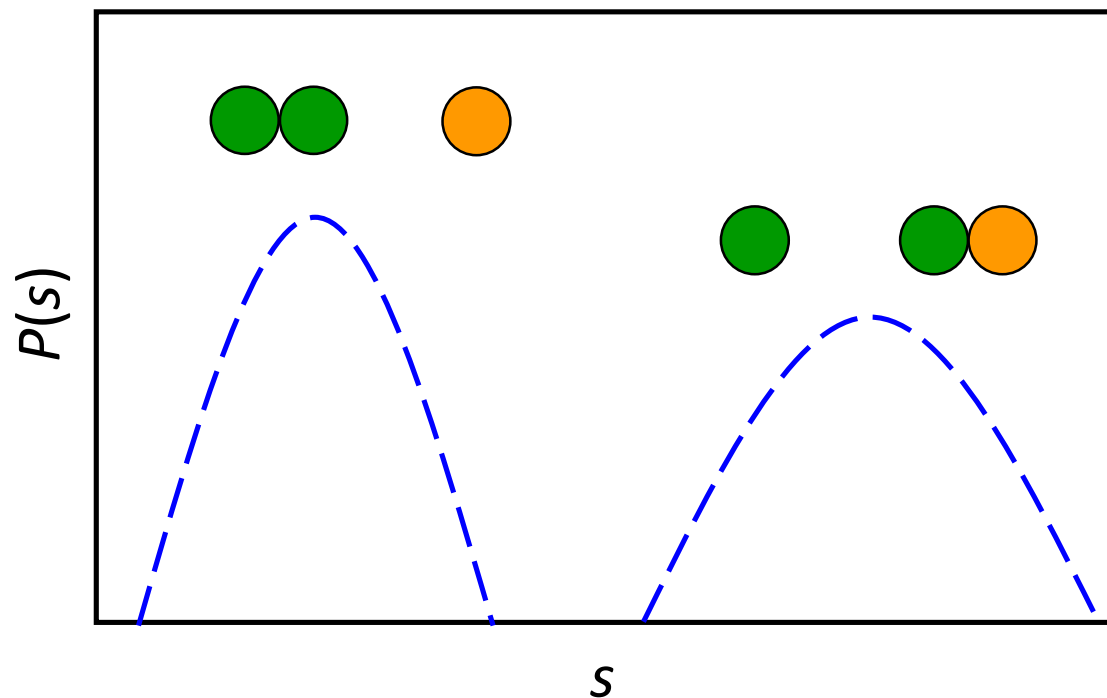
Processes of interest such as

- Chemical reactions,
- Conformational transitions,
- Phase transitions,
- ...

occur on time scales much longer than atomic motion.

We'll assume we're always looking at this type of processes.

A sharply peaked  $P(\mathbf{s})$  is our enemy



We can only properly sample if it is “flat” enough.

**Question:** can we make  $H(\mathbf{s})$  flatter, but still get correct  $P(\mathbf{s})$ ?



# Biased simulations

Metadynamics as a case study

Laio & Parrinello: if we want the system to explore a larger configuration space, we have to discourage it from staying where it is right now.

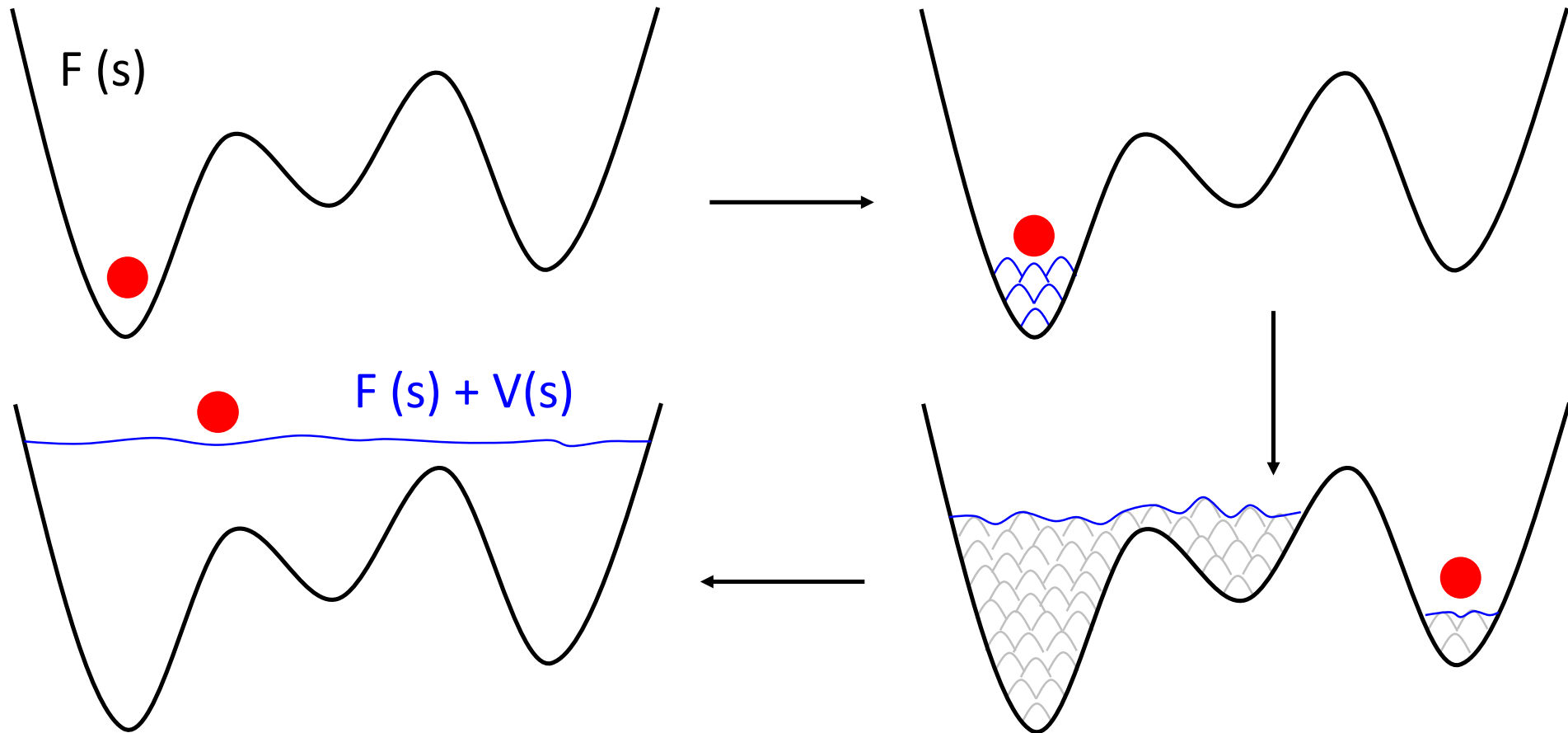
So: add a repulsive potential that pushes the system “somewhere else”. They chose this potential to be Gaussian-shaped.

$$G(\mathbf{s}, \mathbf{s}') = w e^{-(\mathbf{s}-\mathbf{s}')^2/\sigma^2}$$

Add Gaussian bias applied to  $\mathbf{s}$ , centred at the current CV values  $\mathbf{s}'$ , at regular intervals.

This way, we accumulate a bias that compensates  $F(\mathbf{s})$ :

$$V(\mathbf{s}, t) = \sum_i w_i e^{-(\mathbf{s} - \mathbf{s}(t))^2 / \sigma^2}$$



Of course, any average in such biased simulation will be off.

For a bias  $V(\mathbf{s})$ , averages can be reweighted as

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

Or, the relation between the sampled histogram and underlying probability density is:

$$P(\mathbf{s}) = H(\mathbf{s}) e^{\beta V(\mathbf{s})}$$



So, our goal is to design a bias  $V(\mathbf{s})$  that makes  $H(\mathbf{s})$  sufficiently flat

**SLOWLY SAMPLE  
 $H(\mathbf{s})$  AS A DIRECT  
ESTIMATOR OF  $P(\mathbf{s})$**



**SAMPLE  
FLATTENED  $H(\mathbf{s})$  AND  
OBTAIN  $P(\mathbf{s})$   
THROUGH REWEIGHTING**



Slightly more complicated with metadynamics because  $V = V(\mathbf{s}, t)$ .

Original metadynamics generates a bias of the form:

$$V(\mathbf{s}, t) = -F(\mathbf{s}) + c(t)$$

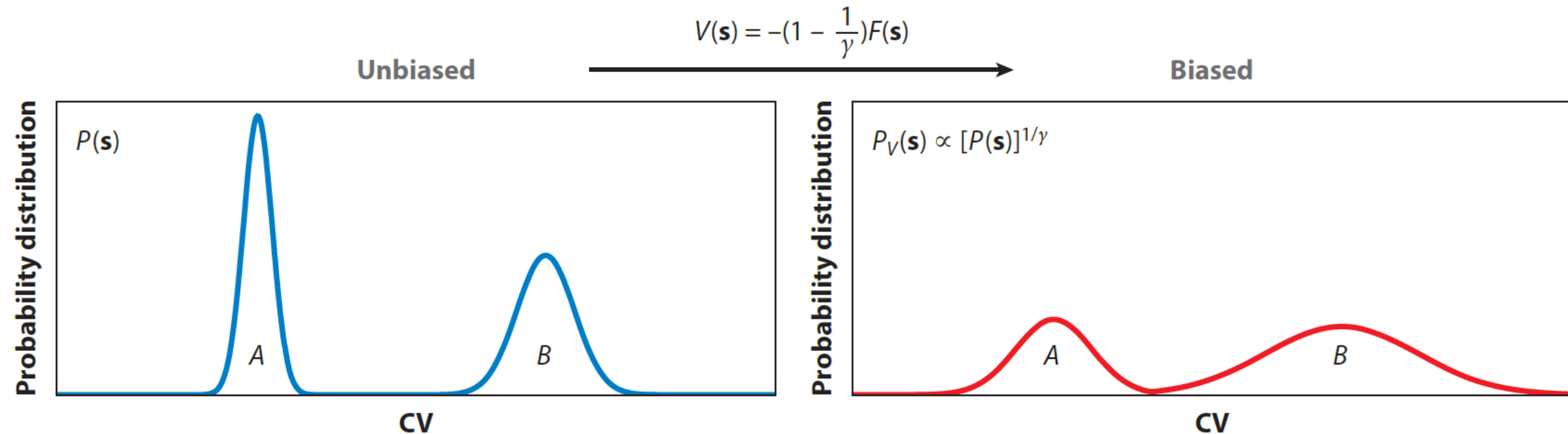
In **well-tempered metadynamics** (WTMetaD), new hills get slower over time, by setting  $w_i = w e^{-\beta V(s(t_i))/(\gamma-1)}$ .

The bias then evolves to

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

with  $c(t)$  asymptotically converging.

The **bias factor**  $\gamma$  controls histogram flattening:  $H(s) = [P(s)]^{1/\gamma}$



$\gamma = 1$ : unbiased MD

$\gamma \rightarrow \infty$ : complete flattening (old school metaD, but doesn't converge)

**Example 2:** Revisit  $S_N2$  example with WTMetaD

Practical implementation = apply extra force during MD

Chain rule: force applied on  $\mathbf{R}$  ( $-\nabla V = -\partial V / \partial \mathbf{R}$ ) is multiple of

- Derivative of summed Gaussian w/r/t  $\mathbf{s}$  ( $\partial V / \partial \mathbf{s}$ )
- Derivative of  $\mathbf{s}$  w/r/t  $\mathbf{R}$  ( $\partial \mathbf{s} / \partial \mathbf{R}$ )

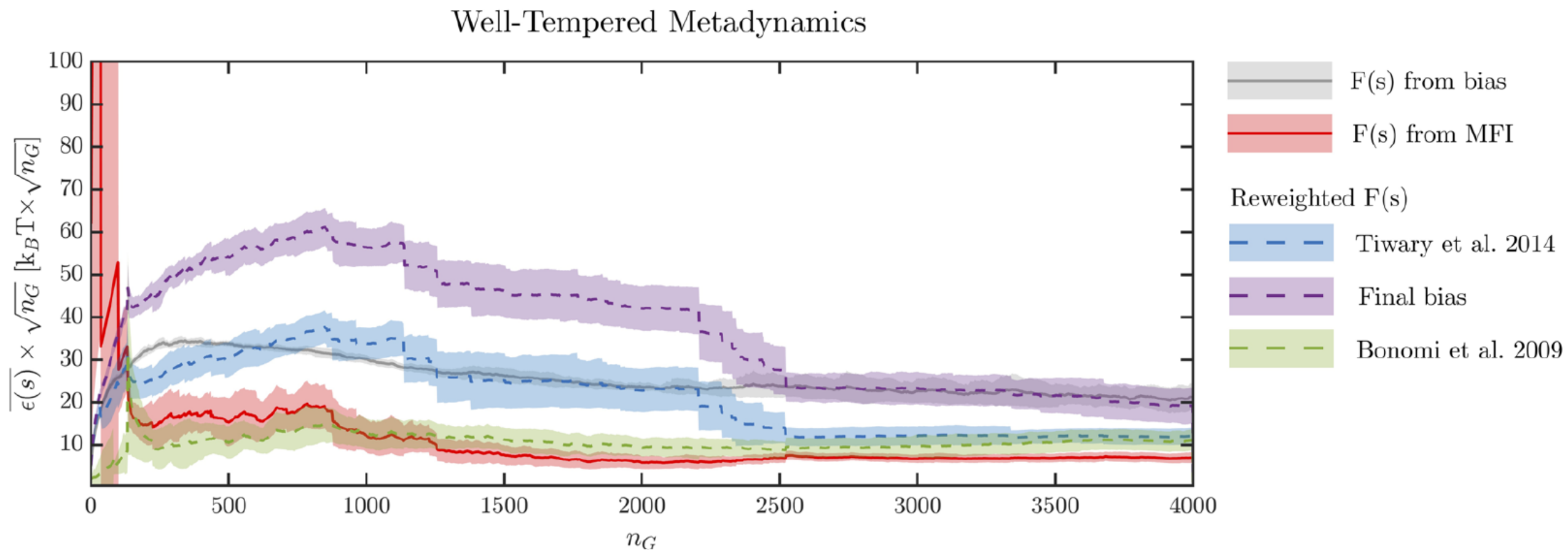
Bias itself is fairly easy, but complexity of CVs can be quite high.

That's why we prefer standard plugins or drivers for their implementation. Have a look at PLUMED for a huge library of CVs!

Alternatives: Colvars module, SSAGES.

**Question:** if the WTMetaD bias converges to  $V(\mathbf{s}) \propto -F(\mathbf{s})$  anyway, why bother with reweighting?

**Answer:** sampled histograms (or forces) converge faster than the bias!



## Reading list

The original publication (barely any theory!):

Laio & Parrinello, “Escaping free-energy minima.” *PNAS* **99**, 12562 (2002)

Discussing all *post hoc* theory:

Valsson, Tiwary & Parrinello, “Enhancing important fluctuations: rare events and metadynamics from a conceptual viewpoint.” *Annu. Rev. Phys. Chem.* **67**, 1545 (2016)

Contemporary perspective:

Bussi & Laio, “Using metadynamics to explore complex free-energy landscapes.” *Nat. Rev. Phys.* **2**, 200 (2020).



# Collective Variables

What should they be, and where can I find them?



**I PICKED A CV A GOOD ONE, RIGHT?**

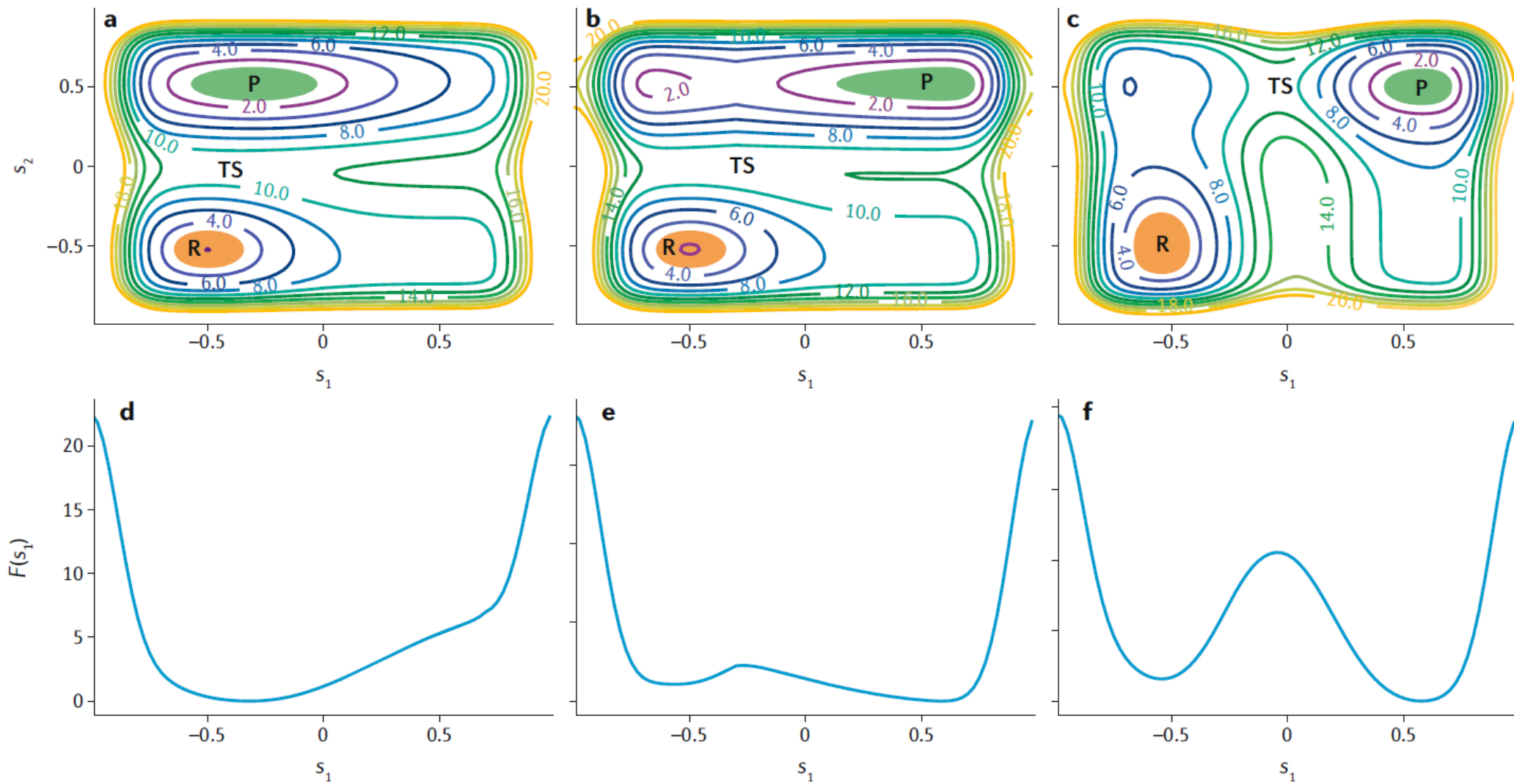
**A GOOD ONE... RIGHT?**

CVs should:

- Allow us to distinguish between all states of interest
- Preferably be physically meaningful
- Contain all “slow” modes of system – all remaining orthogonal degrees of freedom should be fast on an MD time scale
- Be small in number

### **Dependent on process**

- Chemical reactions (distances, coordination numbers)
- Conformational transitions (dihedrals, H-bond contacts, solvent)
- Phase transitions (crystallinity fingerprint, density)

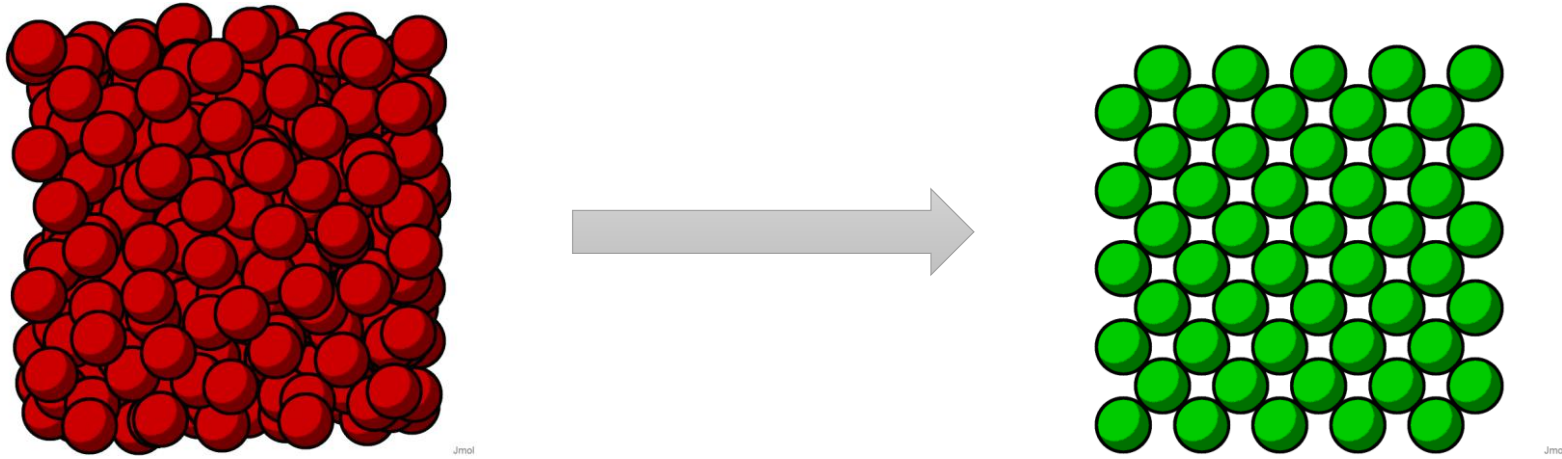


useless

order parameter

reaction coordinate

### Example 3: crystallization of liquid Na



Generic, global CV (pair entropy of system):

$$s_p = -2\pi k_B \int_0^\infty dr [g(r) \ln g(r) - g(r) + 1] r^2$$

Specific CV: Steinhardt parameter matches symmetry of coordination shell with spherical harmonics (where  $\sigma(r) \rightarrow \infty$  at large  $r$ )

$$q_{6m,i} = \frac{\sum_j \sigma(r_{ij}) Y_{6m}(\mathbf{r}_{ij})}{\sum_j \sigma(r_{ij})}$$

Then, average the components of each atom with its neighbours:

$$\bar{q}_{6m,i} = \frac{q_{6m,i} + \sum_j \sigma(r_{ij}) q_{6m,j}}{1 + \sum_j \sigma(r_{ij})}$$

Final CV: take norm of each atomic  $\bar{q}_{6m,i}$ , and mean over all atoms

## How to find good CVs?

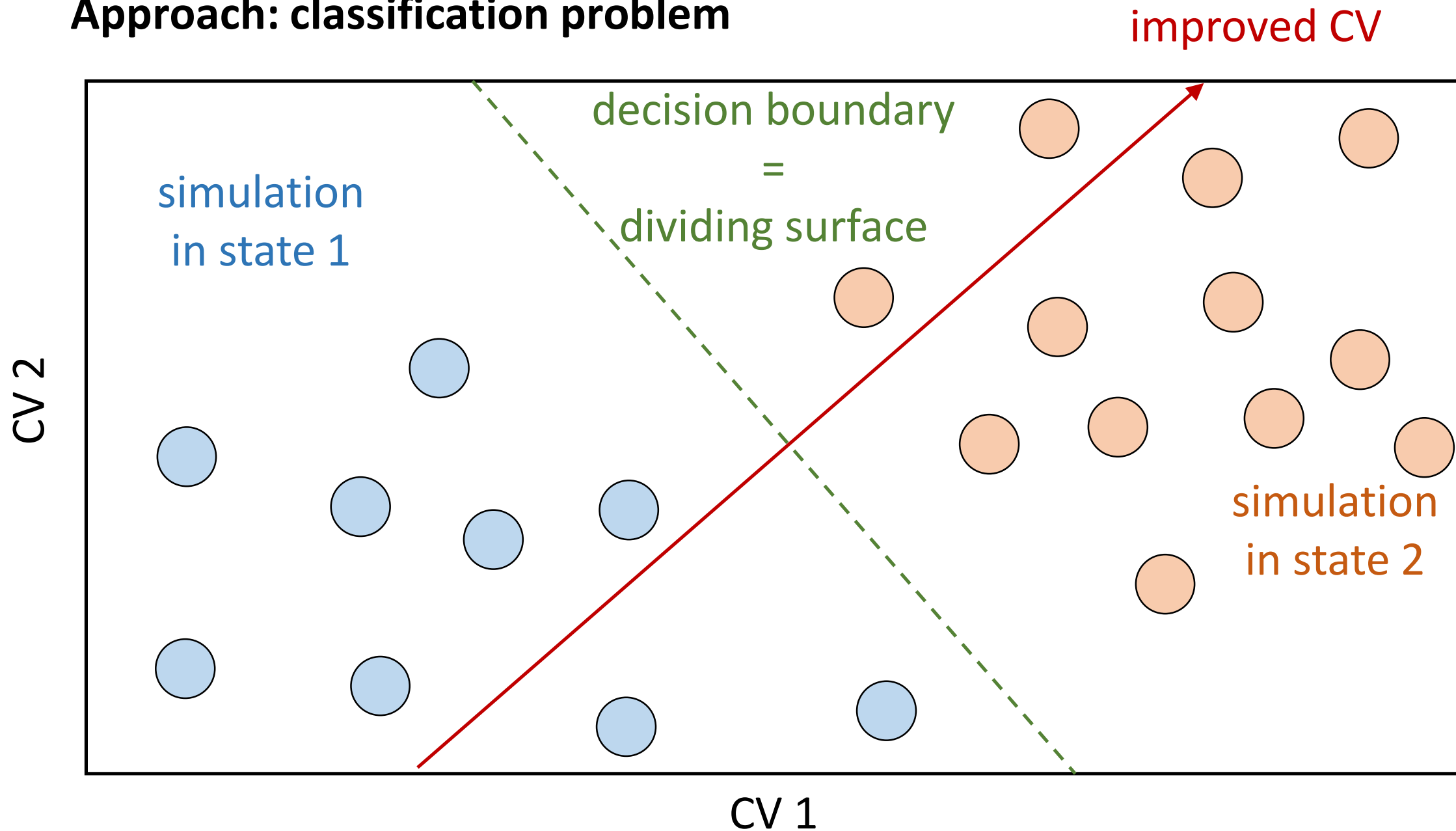
- Intuition & physical insight
- Trial & error

Kind of a chicken & egg problem...

The rise of data science & machine learning methods has introduced new, automated means of finding good CVs.

Usually, one “learns” a good CV as function of a (large) basis set of hand-picked simple CVs

## Approach: classification problem



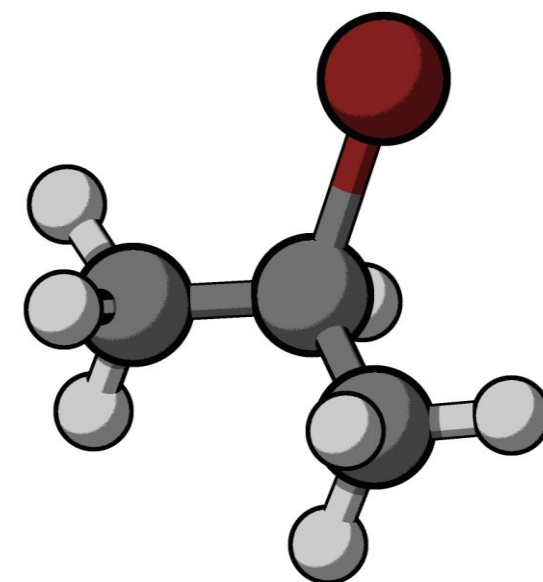
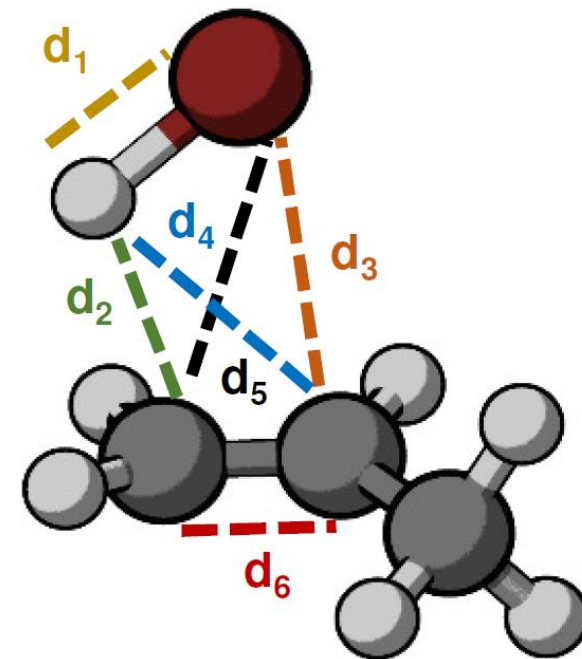
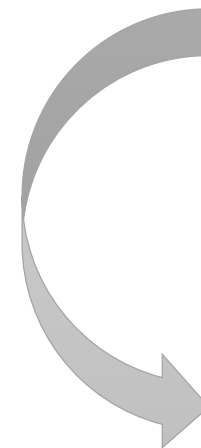


One can use:

- linear classifiers, e.g., (harmonic) linear discriminant analysis
- nonlinear classifiers, e.g., neural networks

**Example 4:** Apply HLDA to hydrobromination reaction

$$\text{Final CV: } \chi = 0.661d_1 - 0.656d_2 - 0.328d_3 + 0.011d_4 - 0.021d_5 + 0.157d_6$$

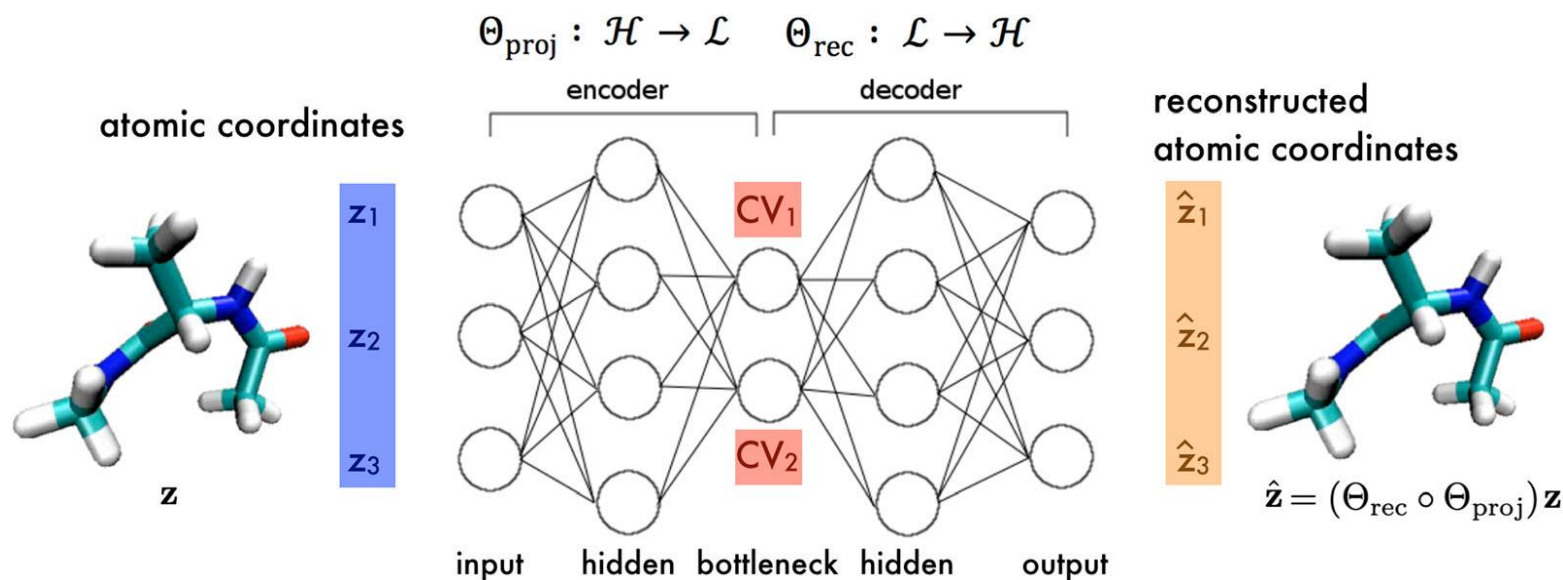


## Approach: dimensionality reduction techniques

“Compress” large number of CVs and perform sampling in lower dimensional space.

- Linear, e.g., principle component analysis (PCA)
- Nonlinear, e.g., autoencoders (neural networks)

iterative training of  
autoencoder + sampling using  
embedding layer as CVs



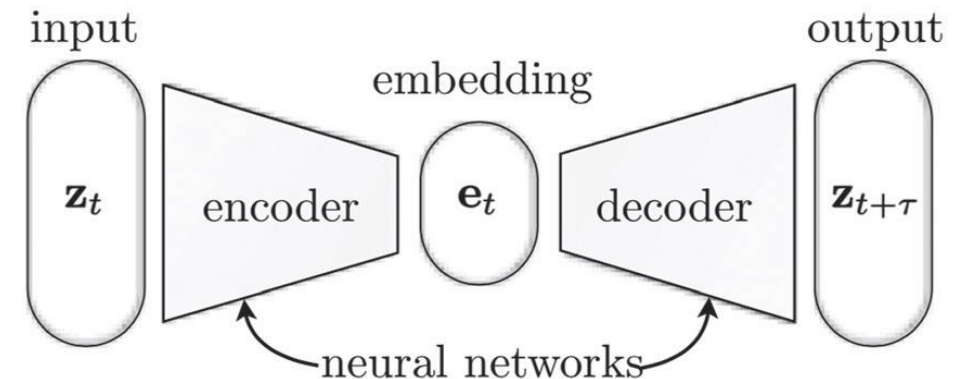
## Approach: signal procession techniques

Use time correlation of CVs to identify slowest modes in system, linear (TICA) or nonlinear (time-lagged autoencoders)

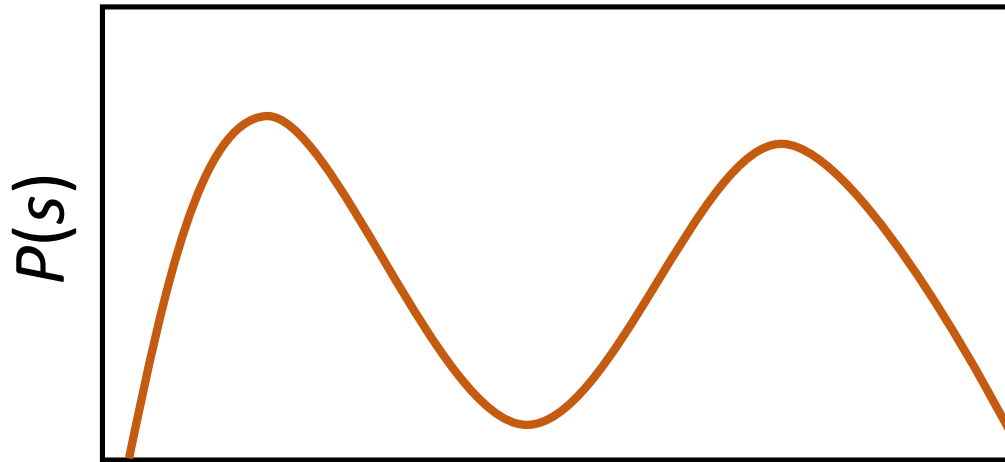
Linear approach: Optimal CV  $\chi = b_1 s_1 + b_2 s_2 + \dots$  from eigenvalue problem using covariance matrix  $\mathbf{C}$

$$\mathbf{C}(t + \tau)\mathbf{b} = \mathbf{C}(t)\lambda\mathbf{b}$$

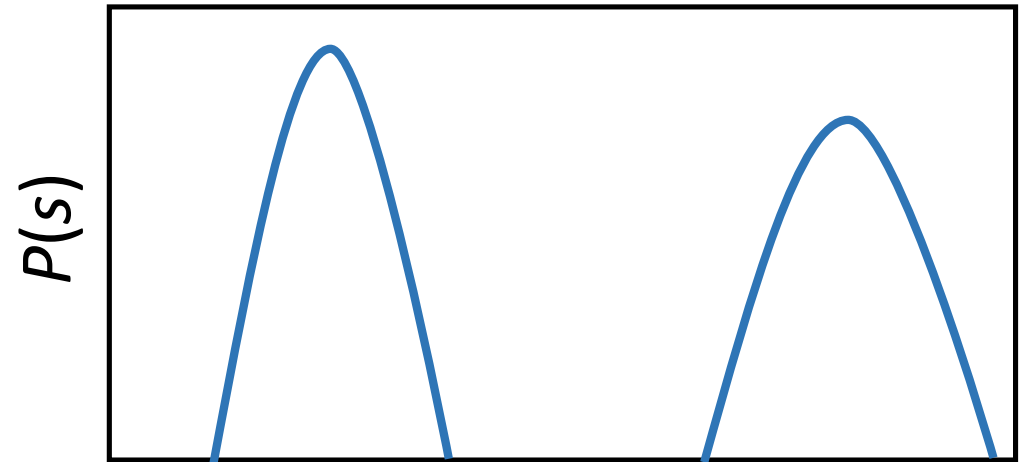
neural network: time-lagged autoencoder



## Approach: optimize the shape of $P(s)$



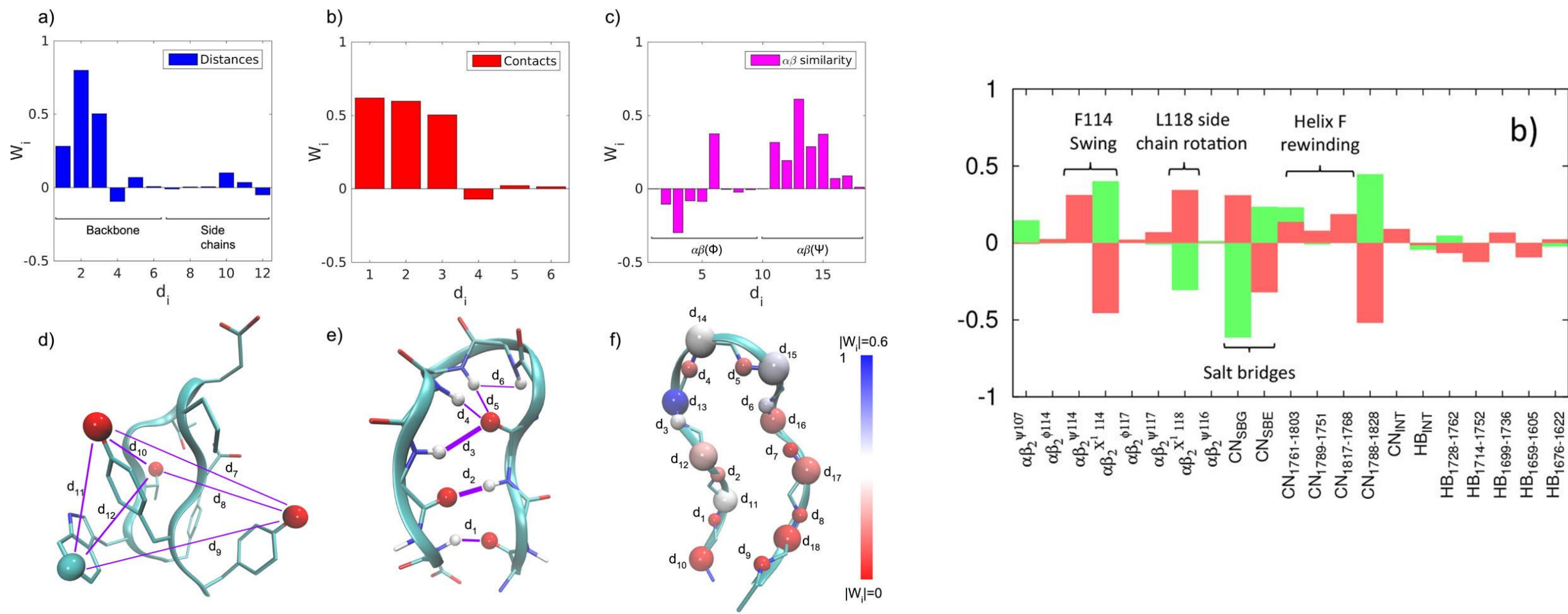
bad CV



good CV

Usually iterative strategy. First bias initial guess CV to get sample, then optimize CV by changing expansion coefficients  $\chi = b_1 s_1 + b_2 s_2 + \dots$  in order to maximize spectral gap, and so on.

# Good CV = efficient sampling, but don't forget physical meaning



Mendels *et al.*, J. Chem. Phys. **149**, 194113 (2018)

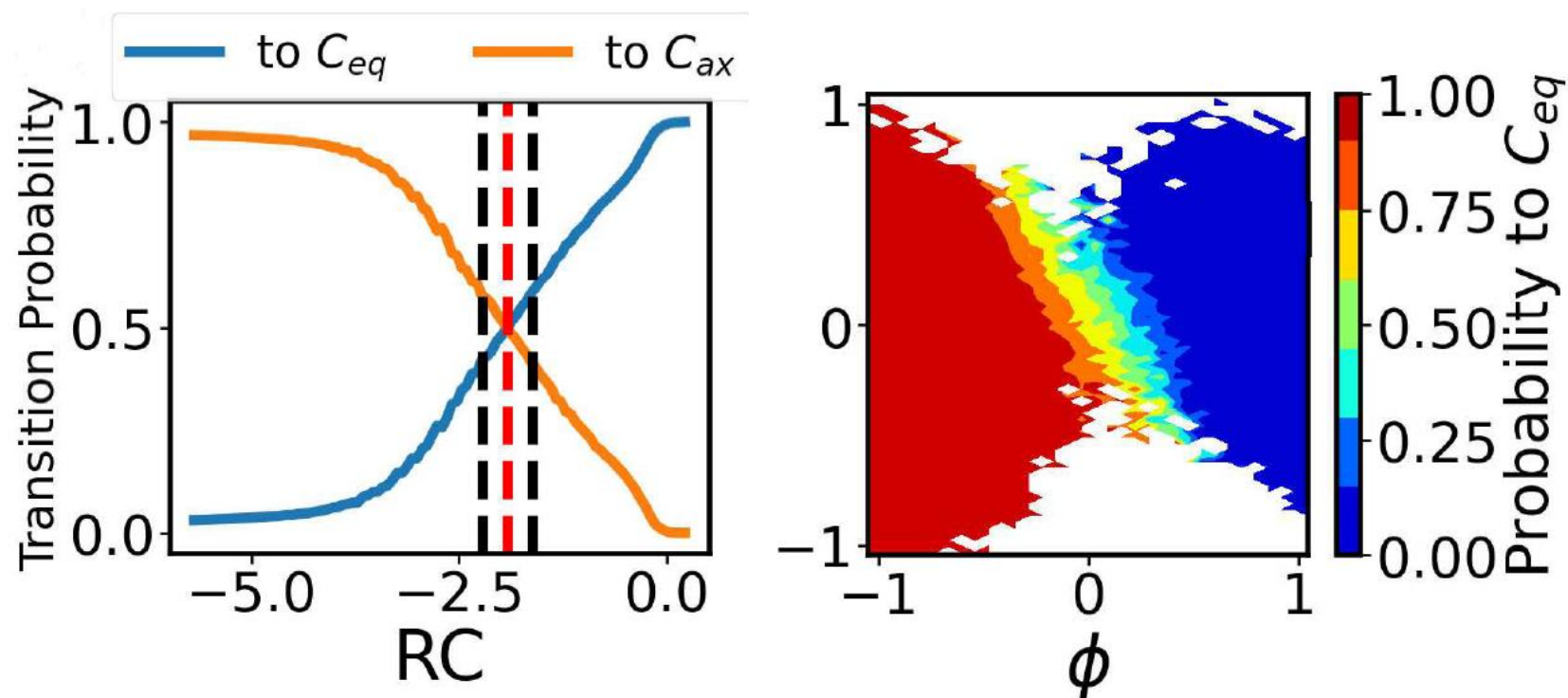
Brotzakis & Parrinello, J. Chem. Theory Comput. **15**, 1393 (2019)

## Check CV quality *a posteriori*: committor analysis

Committer  $p_X$  is fraction of trajectories that ends up in state  $X$ .

For a true dividing surface between state  $A$  and  $B$ ,  $p_A = p_B = 0.5$

Quality of a CV  $s$  with a barrier at  $s = s^*$  can be checked by calculating the committor for states on this candidate dividing surface.



## Reading list

Recent review:

Sidky, Chen & Ferguson, “Machine learning for collective variable discovery and enhanced sampling in biomolecular simulation.” *Mol. Phys.* **118**, e1737742 (2020)

More general perspective:

Glielmo *et al.*, “Unsupervised learning methods for molecular simulation data.” *Chem. Rev.* **121**, 9722 (2021).

+ Previous bibliography!

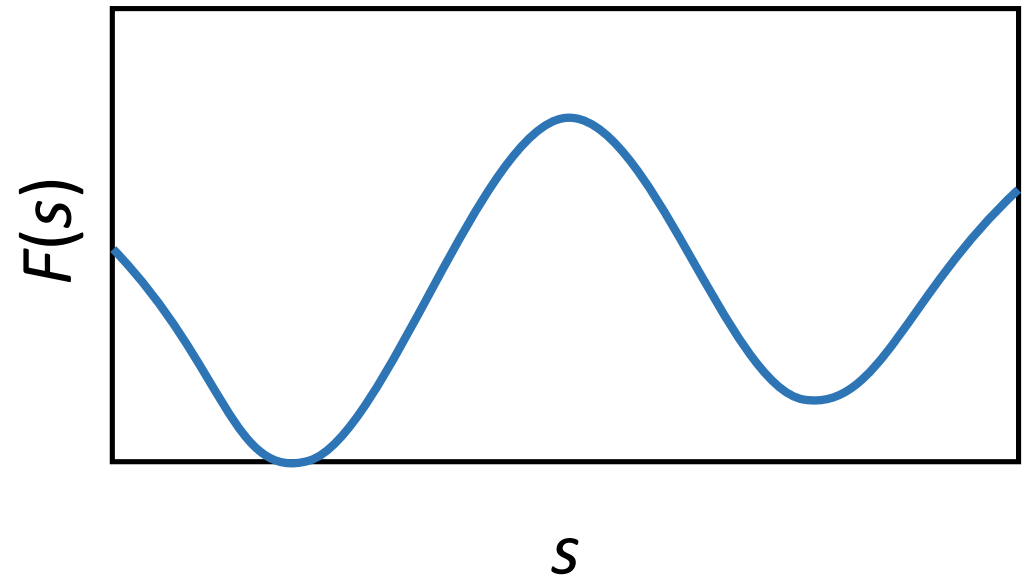
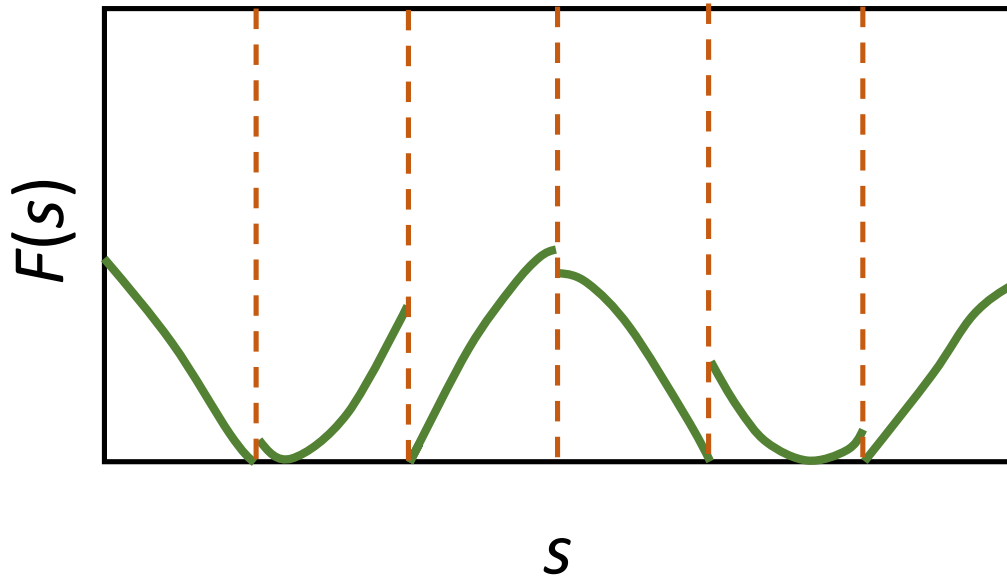
# Biasing approaches

Beyond metadynamics



## The granddaddy: umbrella sampling

1. Pick CV(s) and domain of interest
2. Divide domain in  $N$  small windows
3. Perform  $N$  simulations, confined by “umbrella” bias
4. Stitch together windows to get FES



Bias potential in window  $i$ , centred at  $\mathbf{s}_i$ , e.g.,  $V_i(\mathbf{s}) = K(\mathbf{s} - \mathbf{s}_i)^2$

Use original “umbrella reweighting” formula in each window:

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

Beware: dimensionality, grid spacing, choice of harmonic spring, ...

**General overview & more advanced topics:**

Kästner, “Umbrella sampling.” *Comput. Mol. Sci.* **1**, 932 (2011).

## Adaptive umbrella sampling: no windows

1. Unbiased MD for some time  $\tau$ , sample estimate of  $P(s)$
2. Bias potential  $V(s) = k_B T \ln P(s)$
3. Perform biased simulation using  $V(s)$
4. Update estimate of  $P(s)$  with reweighted histogram
5. Goto 2 until convergence

This core idea returns quite a lot in some form or another!

## Adaptive biasing force (ABF)

Alternative definition of FES: potential of mean force (PMF)

$$f_s(s^*) = - \left| \frac{dF}{ds} \right|_{s=s^*} = \left\langle \mathbf{f} \frac{\partial \mathbf{R}}{\partial s} + k_B T \ln \frac{\partial |J|}{\partial s} \right\rangle_{s=s^*}$$

So if we measure the average force  $f_s$  acting on  $s$  for all  $s$ , we have

$$\Delta F_{s_a \rightarrow s_b} = - \int_{s_a}^{s_b} ds f_s(s)$$

Domain divided in bins. Convenient evaluation of  $f_s$  in each bin  $i$  from time derivative

$$f_{s,i} = \left\langle \frac{d}{dt} \left( \frac{d\mathbf{R}}{ds} \cdot \mathbf{p} \right) \right\rangle_{s \in i}$$

Apply biasing force interpolated from values stored on the grid, that opposes the mean force

$$\mathbf{f}_{\text{bias}} = -f_s \nabla S$$

Biasing/sampling happen iteratively, ultimately converging  $f_s$  and  $F(s)$ .

**More reading:** Comer *et al.*, “The adaptive biasing force method: everything you always wanted to know but were afraid to ask.” *J. Phys. Chem. B* **119**, 1129 (2015)

So, ABF uses an iteratively updated  $f_s$  to bias sampling. When converged, integration of  $f_s$  yields  $F(s)$ .

*Just like adaptive umbrella sampling samples & deploys  $P(s)$ .*

Conversely, **thermodynamic integration** samples  $f_s$  in separate windows and integrates those windows directly.

*Just like conventional umbrella sampling samples  $P(s)$  in windows*

Ultimately, there's always some overlap between approaches...

## Variationally enhanced sampling (VES)

Produce arbitrary probability distribution: pick target distribution  $P_T(\mathbf{s})$  *a priori*, converged bias should lead to  $H(\mathbf{s}) = P_T(\mathbf{s})$ .

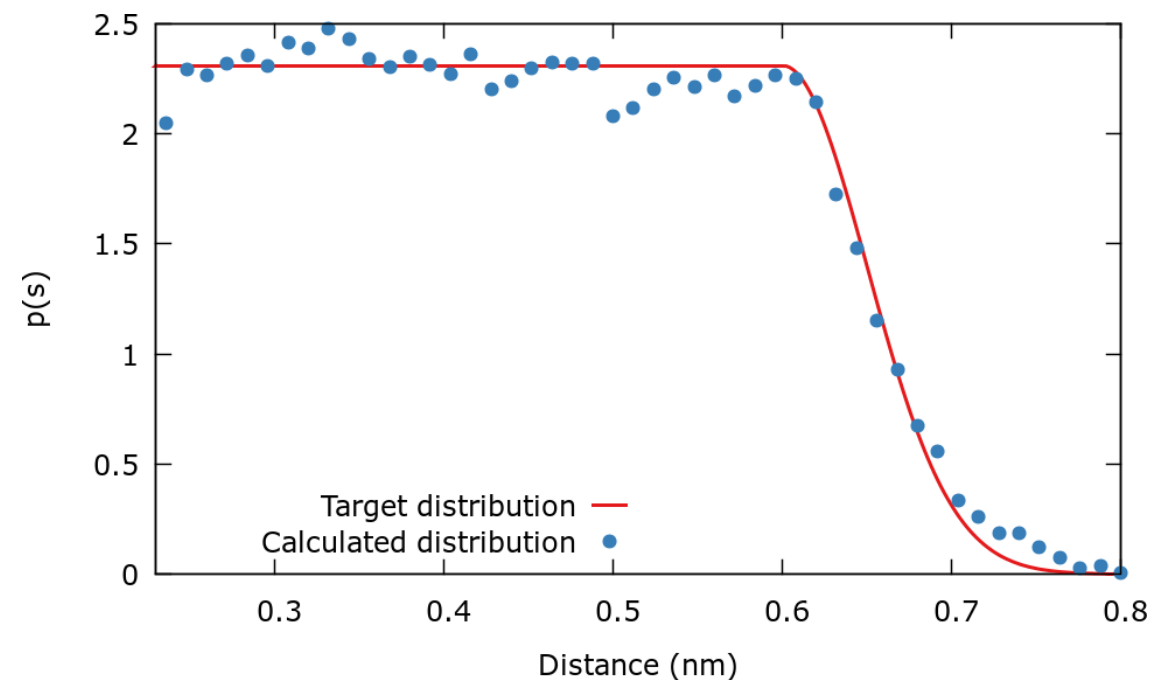
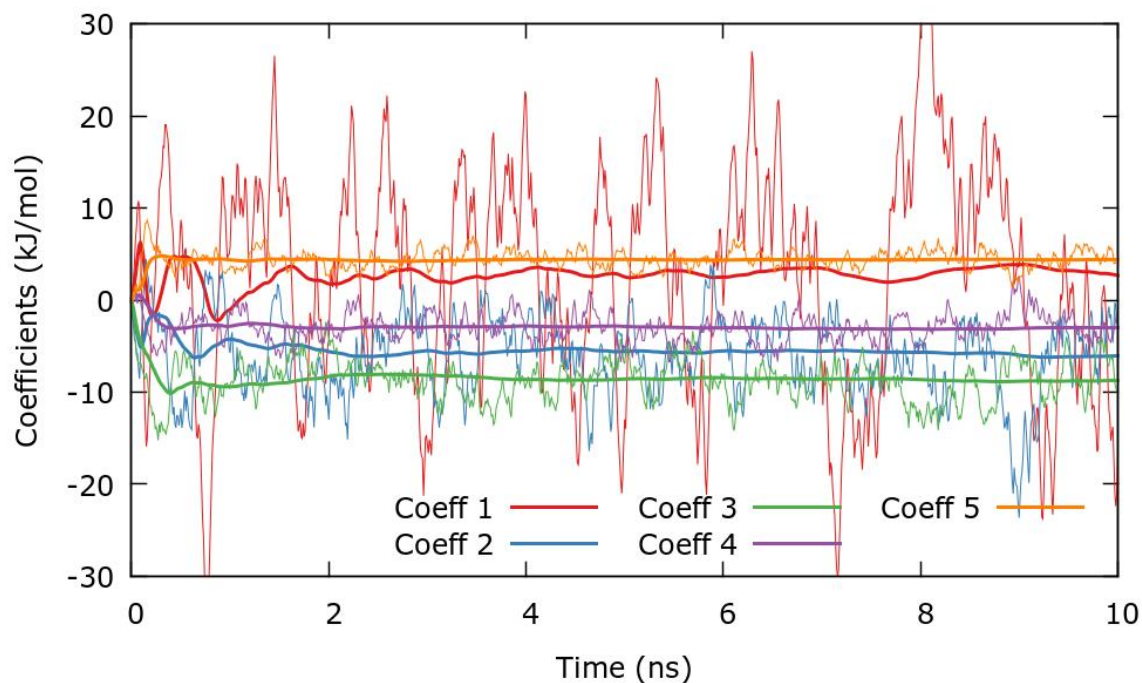
Therefore, the bias has to be

$$V(\mathbf{s}) = -F(\mathbf{s}) - k_B T \ln P_T(\mathbf{s})$$

Variationally optimize  $V(\mathbf{s})$  to minimize deviation between actual  $H(\mathbf{s})$  and target  $P_T(\mathbf{s})$ .

Typical form of bias is a linear expansion in some basis: polynomials (Chebyshev or Legendre), plane waves, ...

Instant “optimal” coefficients very noisy, but average converges.  
And reweighting is still our friend...





All discussed methods use grids and/or a basis expansion: leads to discretization and boundary effects.

Many new (variants of) methods are leveraging machine learning tricks like kernel methods or neural networks

PRL **119**, 150601 (2017)      PHYSICAL REVIEW LETTERS      week ending  
13 OCTOBER 2017

**Stochastic Neural Network Approach for Learning High-Dimensional Free Energy Surfaces**

Elia Schneider,<sup>1</sup> Luke Dai,<sup>1</sup> Robert Q. Topper,<sup>2</sup> Christof Drechsel-Grau,<sup>1</sup> and Mark E. Tuckerman<sup>1,3,4,\*</sup>

<sup>1</sup>*Department of Chemistry, New York University, New York, New York 10003, USA*

<sup>2</sup>*Department of Chemistry, The Cooper Union for the Advancement of Science and Art,  
41 Cooper Square, New York, New York 10003, USA*

<sup>3</sup>*Courant Institute of Mathematical Science, New York University, New York, New York 10003, USA*

<sup>4</sup>*NYU-ECNU Center for Computational Chemistry at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China*  
(Received 28 February 2017; published 11 October 2017)

*J. Chem. Theory Comput.* **2016**, *12*, 5100–5110

**Exploration, Sampling, And Reconstruction of Free Energy Surfaces with Gaussian Process Regression**

Letif Mones,<sup>\*,†</sup> Noam Bernstein,<sup>‡</sup> and Gábor Csányi<sup>†</sup>

<sup>†</sup>Engineering Laboratory, University of Cambridge, Cambridge, CB2 1PZ, United Kingdom

<sup>‡</sup>Naval Research Laboratory, Center for Materials Physics and Technology, Washington, D.C. 20375, United States

free energy reconstruction/FES  
representation

## Neural Network and Nearest Neighbor Algorithms for Enhancing Sampling of Molecular Dynamics

Raimondas Galvelis<sup>\*,†,‡</sup> and Yuji Sugita<sup>†,‡,§,||</sup>

<sup>†</sup>RIKEN Theoretical Molecular Science Laboratory, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

<sup>‡</sup>RIKEN Advance Institute for Computational Science, Integrated Innovation Building 7F, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

<sup>||</sup>RIKEN iTHES, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

<sup>§</sup>RIKEN Quantitative Biology Center, Integrated Innovation Building 7F, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

THE JOURNAL OF CHEMICAL PHYSICS **148**, 134108 (2018)

## Adaptive enhanced sampling by force-biasing using neural networks

Ashley Z. Guo,<sup>1,a)</sup> Emre Sevgen,<sup>1,a)</sup> Hythem Sidky,<sup>2</sup> Jonathan K. Whitmer,<sup>2</sup> Jeffrey A. Hubbell,<sup>1</sup> and Juan J. de Pablo<sup>1,3</sup>

<sup>1</sup>Institute for Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA

<sup>2</sup>Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA

<sup>3</sup>Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

(Received 26 December 2017; accepted 2 March 2018; published online 3 April 2018)

## Neural networks-based variationally enhanced sampling

Luigi Bonati<sup>a,b,c</sup>, Yue-Yu Zhang<sup>b,d</sup>, and Michele Parrinello<sup>b,c,d,e,1</sup>

PNAS | September 3, 2019 | vol. 116 | no. 36 | 17641–17647

THE JOURNAL OF CHEMICAL PHYSICS **148**, 104111 (2018)

## Learning free energy landscapes using artificial neural networks

Hythem Sidky and Jonathan K. Whitmer<sup>a)</sup>

Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA

(Received 8 December 2017; accepted 22 February 2018; published online 12 March 2018)

## Gaussian-Mixture Umbrella Sampling

Paul Maragakis,<sup>\*,†,‡</sup> Arjan van der Vaart,<sup>\*,†,§</sup> and Martin Karplus<sup>†,||</sup>

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, Center for Biological Physics, Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287, and Institut de Science et d'Ingénierie Supramoléculaires, Université Louis Pasteur, 67000 Strasbourg, France

Received: September 21, 2008; Revised Manuscript Received: January 11, 2009

neural network or kernel method  
as bias potential/force

## Rethinking Metadynamics: From Bias Potentials to Probability Distributions

Michele Invernizzi and Michele Parrinello<sup>\*</sup>



Cite This: *J. Phys. Chem. Lett.* 2020, 11, 2731–2736



Read Online

## Reweighted Jarzynski Sampling: Acceleration of Rare Events and Free Energy Calculation with a Bias Potential Learned from Nonequilibrium Work

Kristof M. Bal<sup>\*</sup>



Cite This: *J. Chem. Theory Comput.* 2021, 17, 6766–6774



Read Online

**Paths not taken:** anything not based on biasing, really!

- Thermodynamic integration
- Replica exchange/parallel tempering
- Temperature-accelerated MD (TAMD)
- ...

**Useful review for a different perspective:**

Abrams & Bussi, “Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration.”  
*Entropy* **16**, 163 (2014).

# Kinetics

Squeezing more out of our tools

Not only thermodynamics matters, but also kinetics.

We'll try to extend the previous tools & tricks to the calculation of reaction rates.

**Path not taken:** Methods like transition path and forward flux sampling.

Bolhuis *et al.*, “Throwing ropes over rough mountain passes, in the dark.” *Annu. Rev. Phys. Chem.* **53**, 291 (2002).

Hussain & Haji-Akbari, “Studying rare events using forward-flux sampling: Recent breakthroughs and future outlook.” *J. Chem. Phys.* **152**, 060901 (2020).

## Transition state theory

Rate is proportional to “equilibrium constant” of a transition state (dividing surface) and initial metastable state.

For a process  $A \rightarrow TS \rightarrow B$  we have

$$k = \omega \frac{Z_{TS}}{Z_A}$$

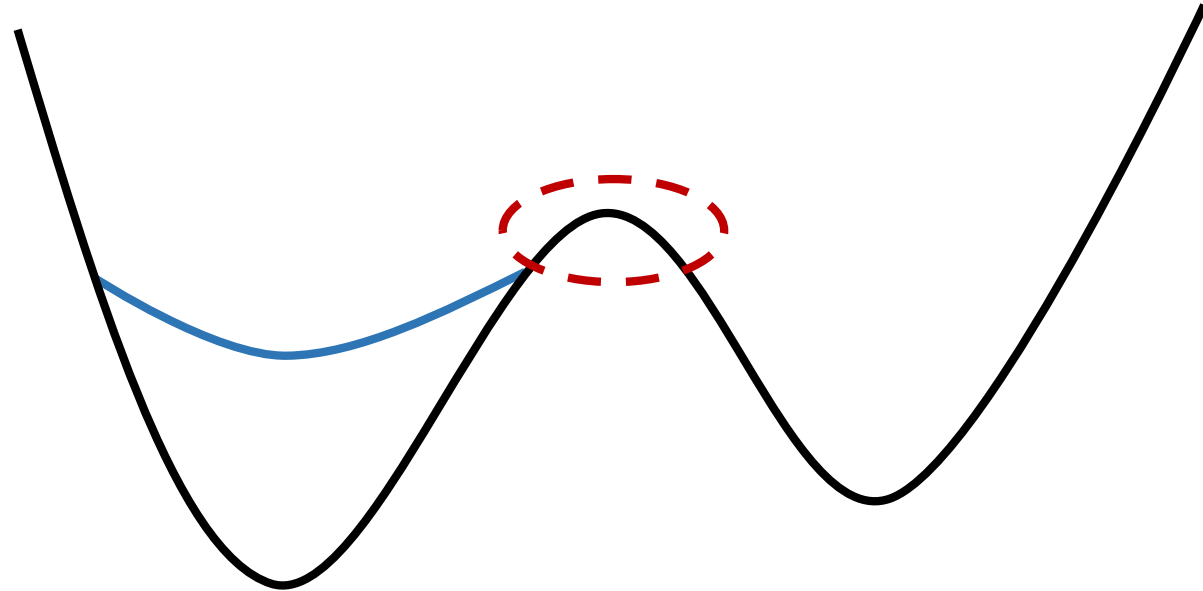
Only holds if transitions are infrequent or motion along slow mode

Voter: we can modify rates by applying a bias potential

Suppose a bias potential fills the basin without touching the TS, we accelerate the process.

Now, transitions have the modified rate

$$k^* = \omega \frac{Z_{TS}}{Z_A^*}$$



How do we recover the true rate?

We say that  $k^* = \alpha k$ , so that

$$\alpha = \frac{Z_A^*}{Z_A}$$

The only discrepancy between the two cases is the bias, so we're just reweighting again

$$\alpha = \left\langle e^{\beta V(t)} \right\rangle_b$$



Alternatively, we can consider time itself to be rescaled. The true time is actually a “hypertime.”

At time  $t_{\text{MD}} = n_{\text{MD}}\Delta t$ , the true time that has passed is  $t_{\text{hyper}} = \alpha t_{\text{MD}}$  or

$$t_{\text{hyper}} = \int_0^{t_{\text{MD}}} dt e^{\beta V(t)}$$

In a practical MD simulation we have a discrete version:

$$t_{\text{hyper}} = \Delta t \sum_i e^{\beta V(t_i)}$$

Implementation through metadynamics: deposit hills very slowly!

We call this *infrequent* metadynamics, and it's a simple & pragmatic solution.

But slow & many simulations needed for statistics.

**Paths not taken:** variational flooding, CVHD method, Voter's other methods

**Example 5:** Applying infrequent metadynamics to the  $S_N2$  example.

## Reading list

More about Voter's accelerated MD strategies:

Voter, "Extending the time scale in atomistic simulation of materials."  
*Annu. Rev. Mater. Res.* **32**, 321 (2002)

Concerns regarding infrequent metaD:

Khan *et al.*, "How fluxional reactants limit the accuracy/efficiency of infrequent metadynamics." *J. Chem. Phys* **153**, 054125 (2020)

+ everything about CV quality and the like

## Use an existing FES only

TST defines the rate as the flux through a dividing surface defined by the reaction coordinate  $\chi$

convert to "CV units"

pick  $\chi_{\text{TS}}$  that minimizes  $k$

$$k^{\text{TST}} = \frac{\langle |\nabla \chi| \rangle_{\chi=\chi_{\text{TS}}}}{\sqrt{2\pi\beta m}} e^{-\beta(F(\chi_{\text{TS}}) - F_A)}$$

1D thermal flux

$$F_A = -k_B T \ln \int_A d\chi e^{-\beta F(\chi)}$$

In chemistry we tend to use the Eyring expression:

$$k^{\text{TST}} = \frac{1}{\beta h} e^{-\beta \Delta^\ddagger F}$$

So if we stich everything together:

$$\Delta^\ddagger F = F(\chi_{\text{TS}}) + k_B T \ln \frac{\sqrt{2\pi m k_B T}}{h \langle |\nabla \chi| \rangle_{\chi=\chi_{\text{TS}}}} - F_A$$



partition function  
inside segment  $d\chi$

$P(\chi)$  is probability  
“per  $d\chi$ ”

**Example 2** shows application of formalism.

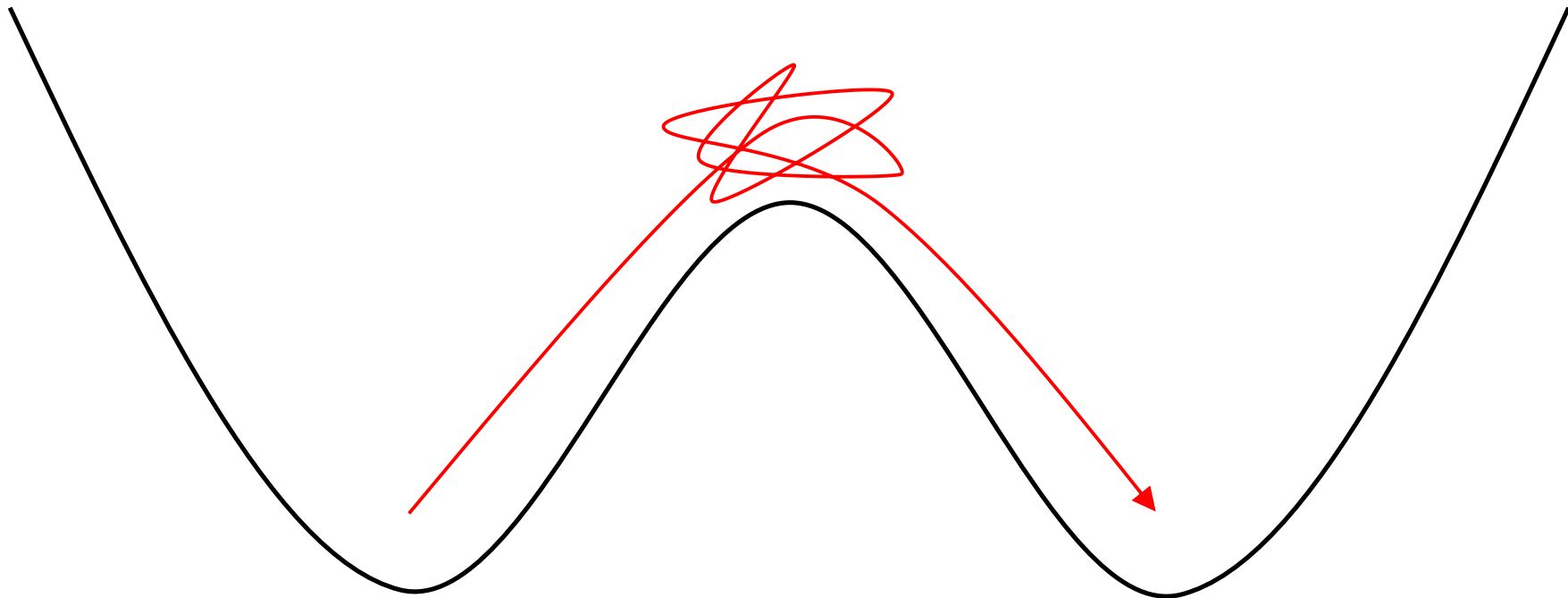
Compare relative cost with infrequent metaD! But beware

- CV quality
- Recrossings



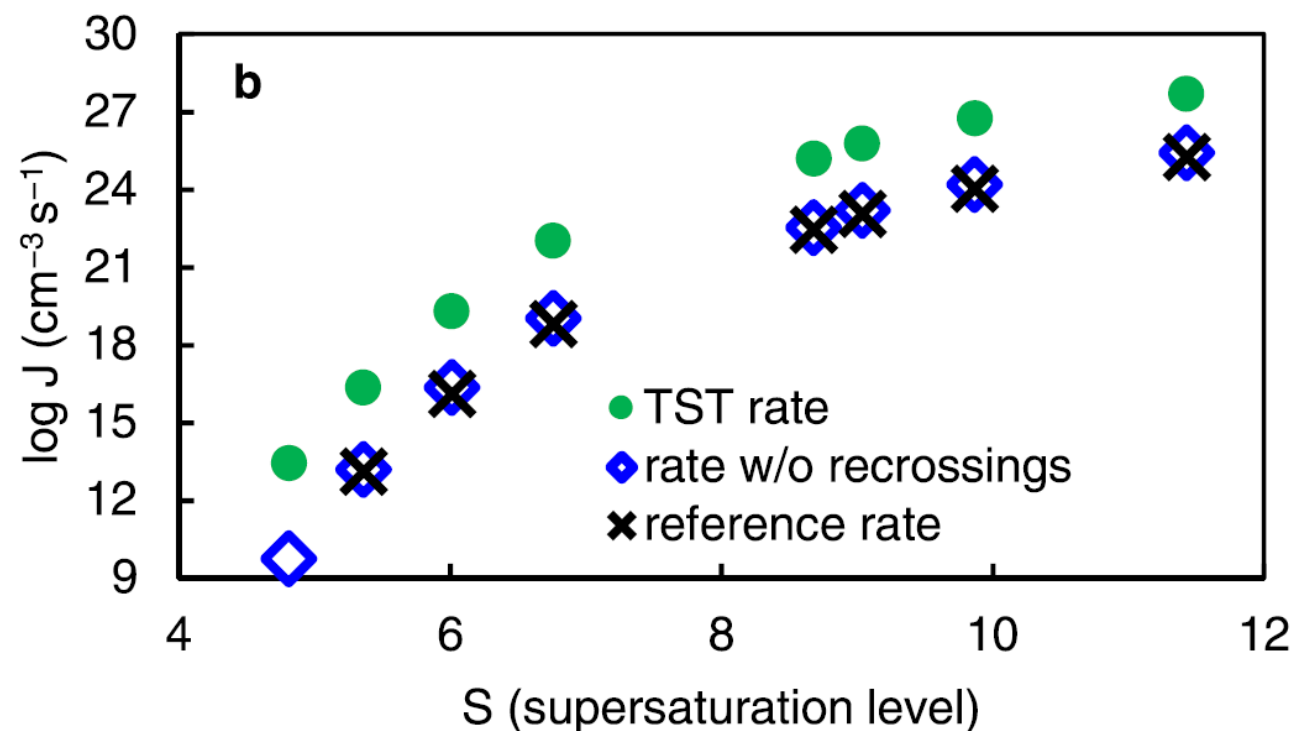
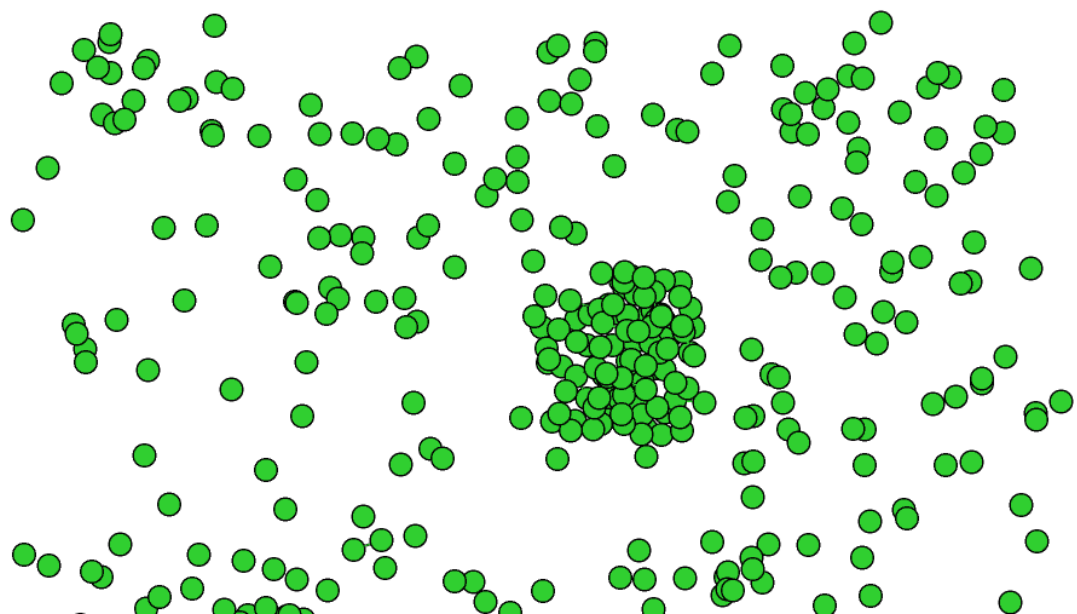
The TST rates measures **all** crossings of the TS. To account for non-effective recrossings we introduce a **transmission coefficient**  $\kappa \leq 1$

$$k = \kappa k^{\text{TST}}$$



My lazy trick to account for diffusional recrossing, tested for droplet nucleation processes:

Run a few trajectories starting at putative TS and count number of crossings before committing

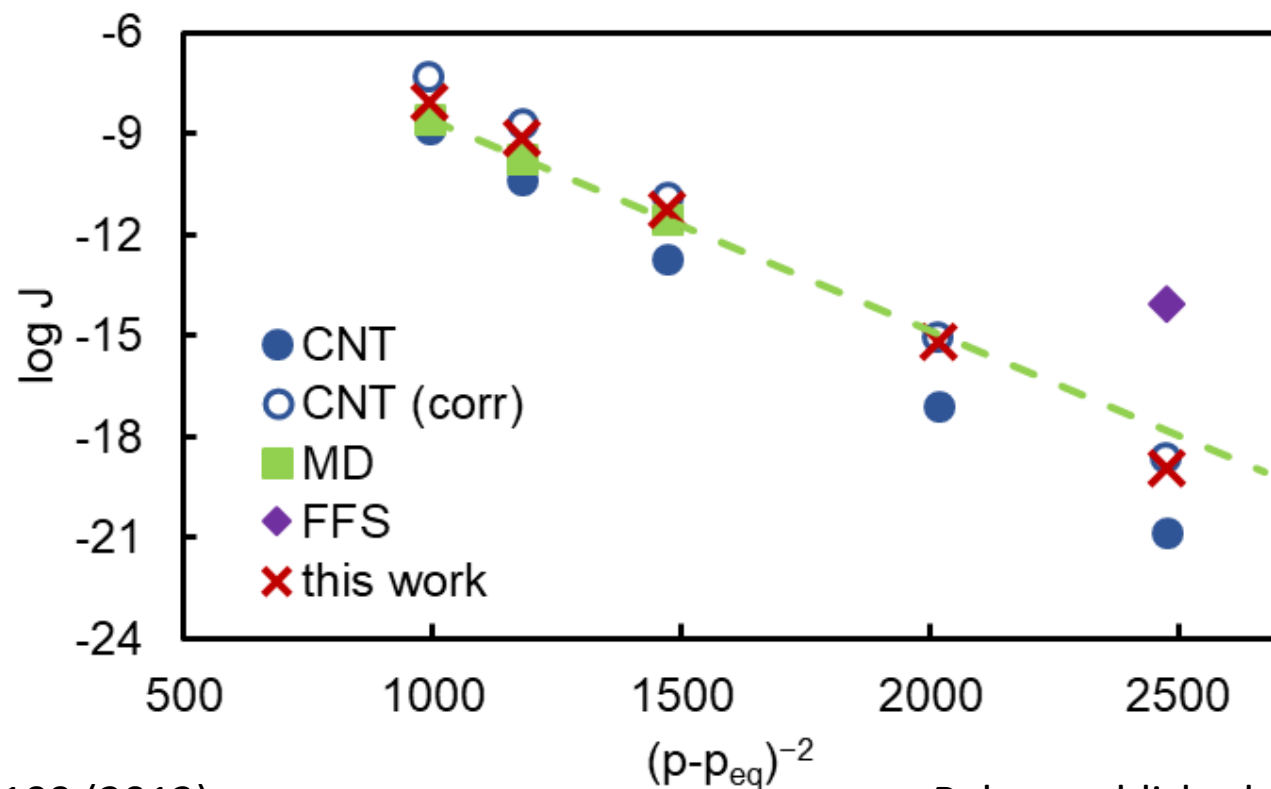
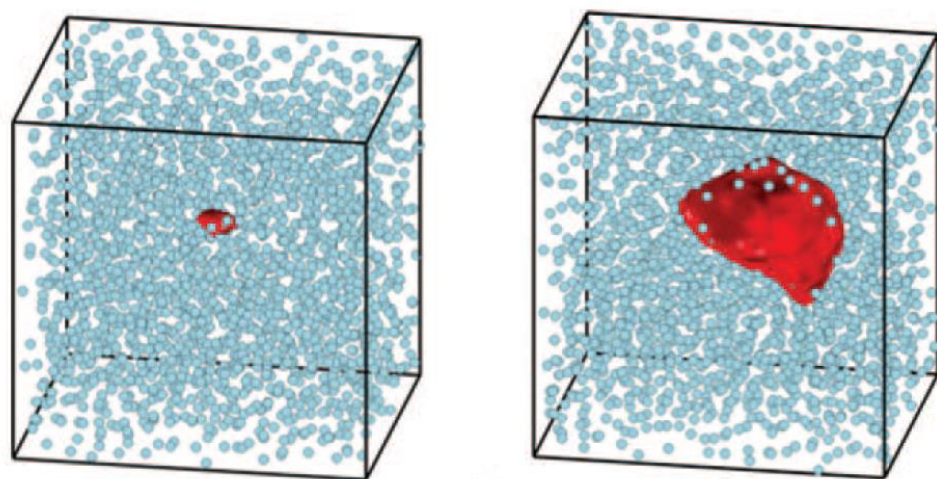




Disagreement between methods: bubble nucleation

Most methods only probe limited rate ranges and don't extrapolate consistently.

Combining a corrected TST formalism with my own biasing strategy helps spotting the “correct” approaches.



## Reading list

Vanden-Eijnden & F. A. Tal, “Transition state theory: Variational formulation, dynamical corrections, and error estimates.” *J. Chem. Phys.* **123**, 184103 (2005)

Bal *et al.*, “Free energy barriers from biased molecular dynamics simulations.” *J. Chem. Phys.* **153**, 114118 (2020)

Bal, “Nucleation rates from small scale atomistic simulations and transition state theory.” *J. Chem. Phys.* **155**, 144111 (2021)

# Fin

Conclusions

## **To conclude...**

I've given a perspective on rare event sampling, centred around free energy calculation & CV-based biasing methods.

Perhaps it can serve as a rough map for some of you.

I hope you have fun exploring!