PLUMED Masterclass

21.3: Umbrella sampling

Giovanni Bussi
SISSA - Trieste
bussi@sissa.it

@BussiGio

SISSA

PLUMED

open-source
freely-available
C++ library

- enhanced-sampling methods
- free-energy methods
- analysis MD data

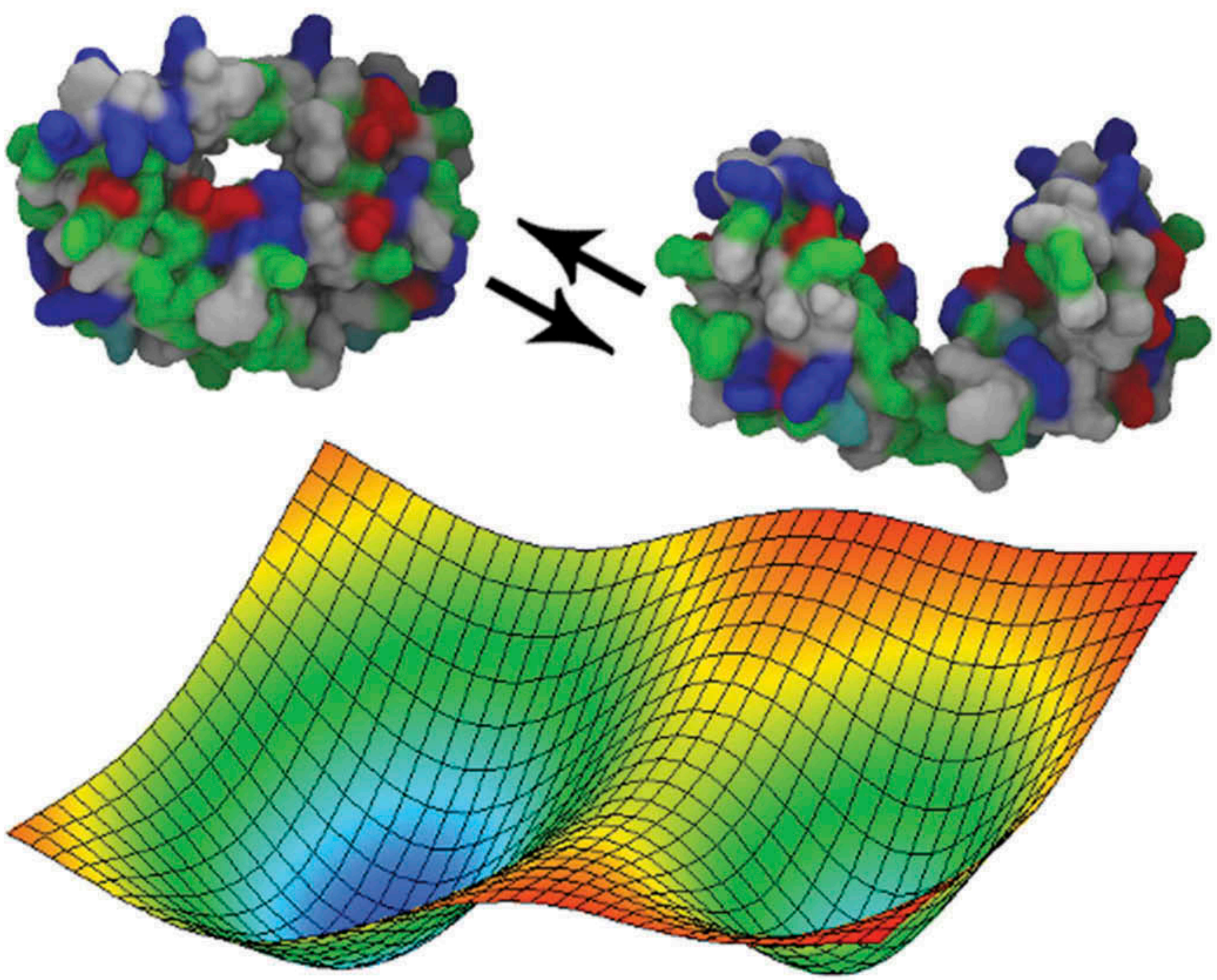
www.plumed.org

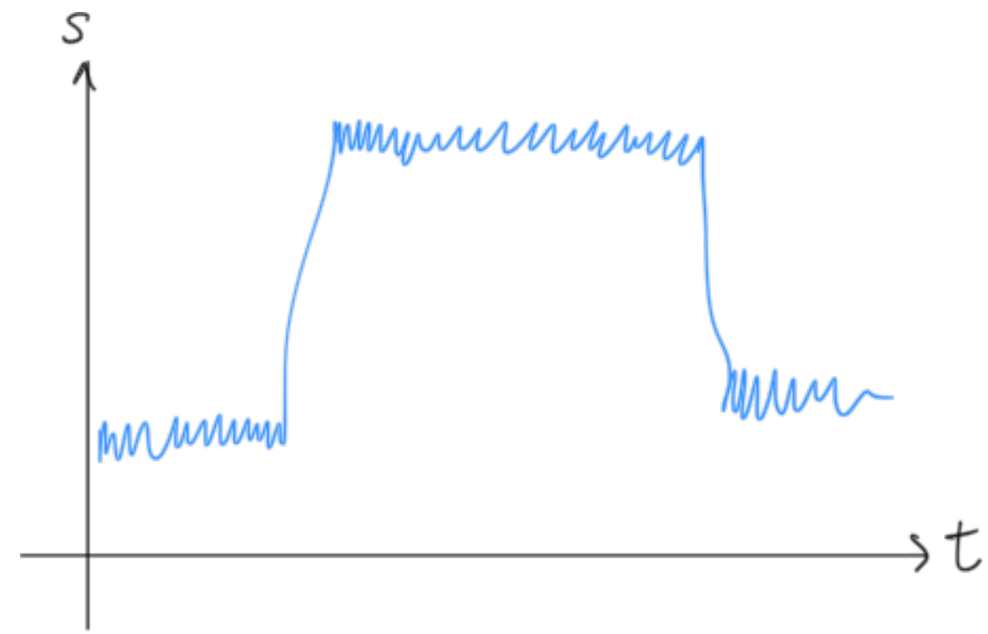
@plumed_org

Bonomi *et al.* CPC 2009
Tribello *et al.* CPC 2014

Class	Topic	Lecture I	Lecture II	Instructor
21.1	PLUMED syntax and analysis	January 18, 2021	January 25, 2021	M. Bonomi
21.2	Statistical errors in MD	February 1, 2021	February 8, 2021	G. Tribello
21.3	Umbrella sampling	February 15, 2021	February 22, 2021	G. Bussi
21.4	Metadynamics	March 1, 2021	March 8, 2021	M. Bonomi
21.5	Replica exchange methods	March 15, 2021	March 22, 2021	G. Bussi
21.6	Dimensionality reduction	April 12, 2021	April 19, 2021	G. Tribello
21.7	Performance optimization	April 26, 2021	May 3, 2021	M. Bonomi
21.8	Poster session	May 10, 2021		

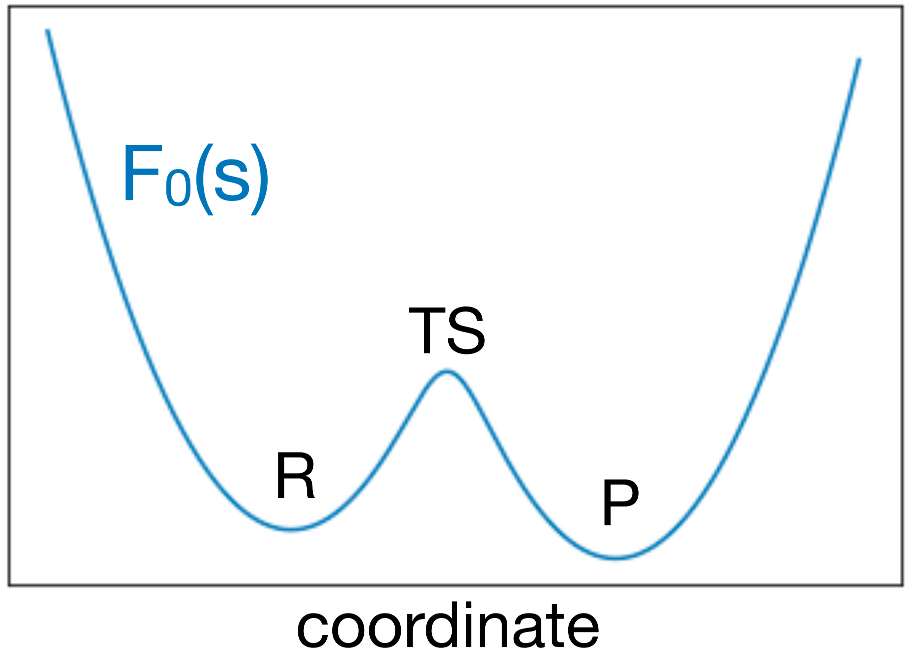
Rare events



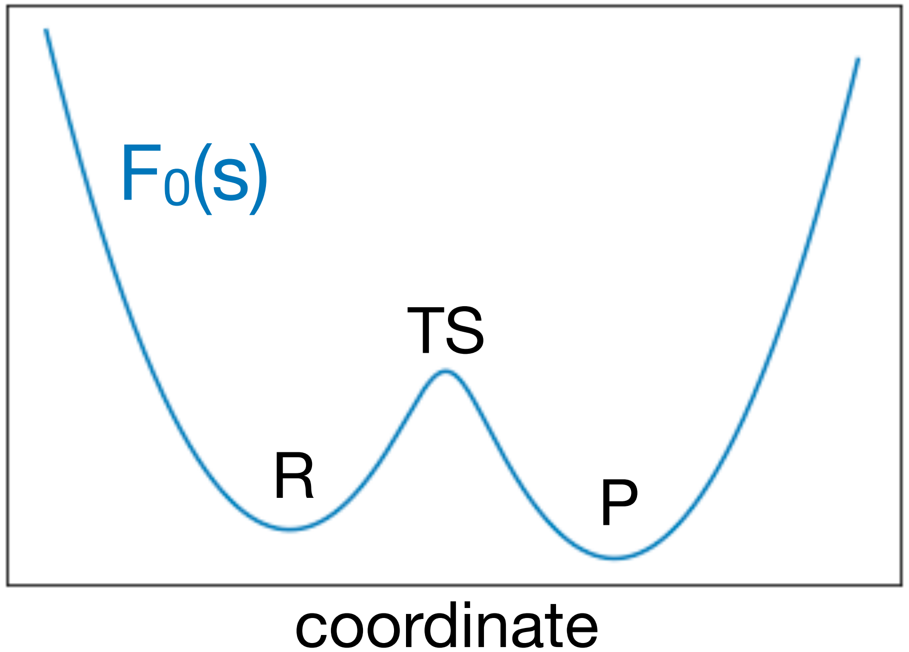


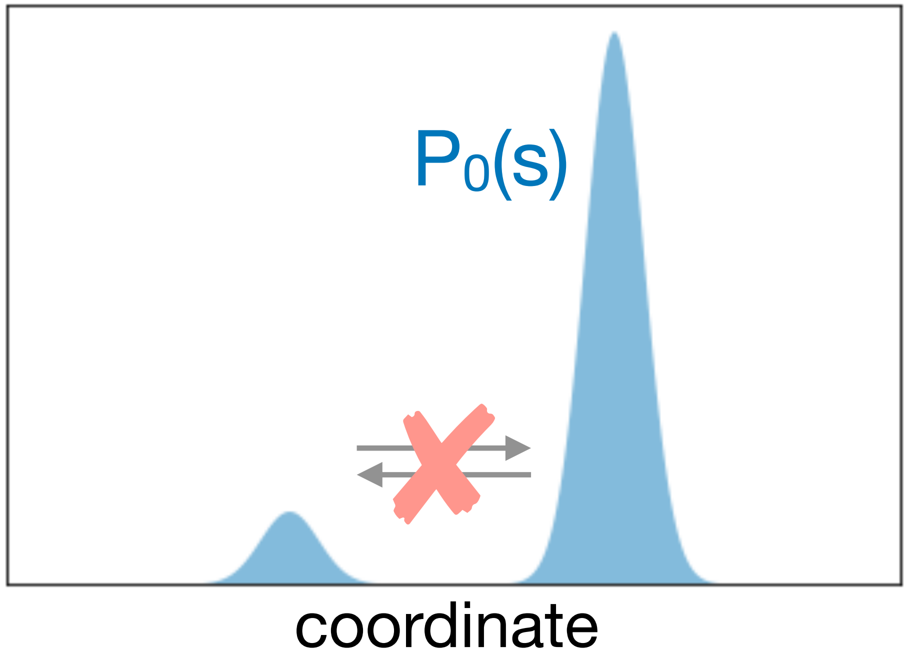
Camilloni and Pietrucci, Adv in Phys X (2018)

How biased sampling could help: umbrella sampling

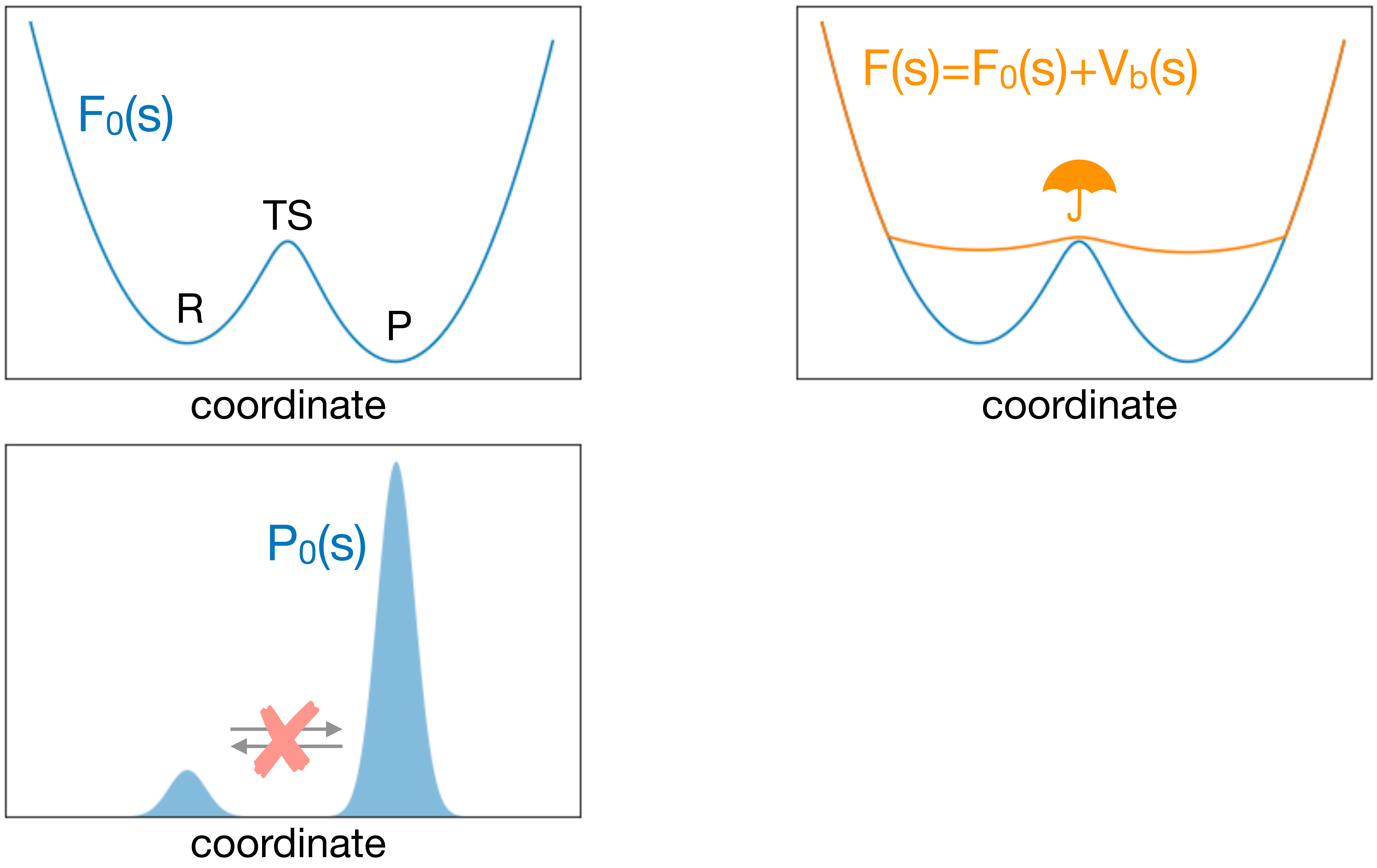


How biased sampling could help: umbrella sampling

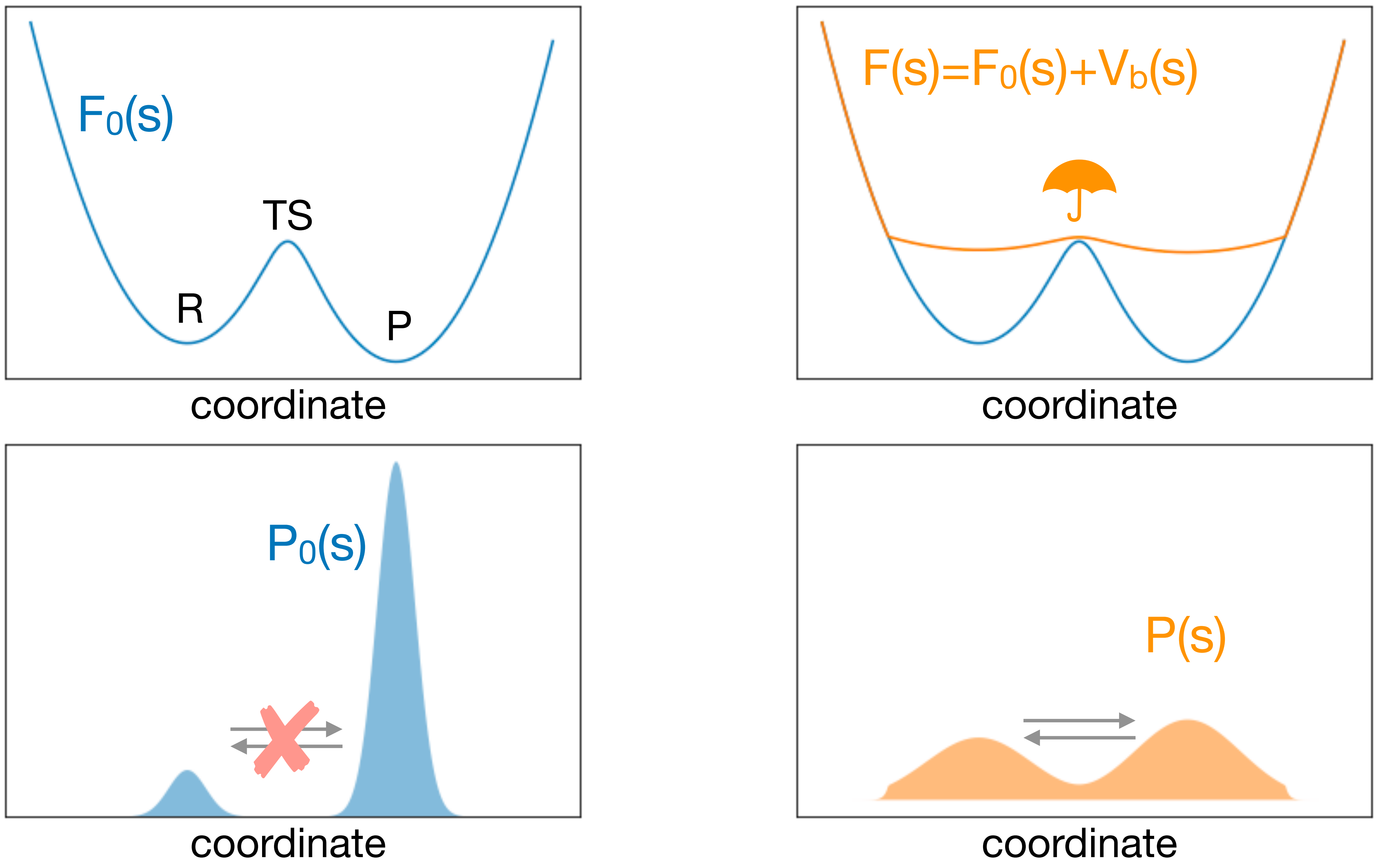




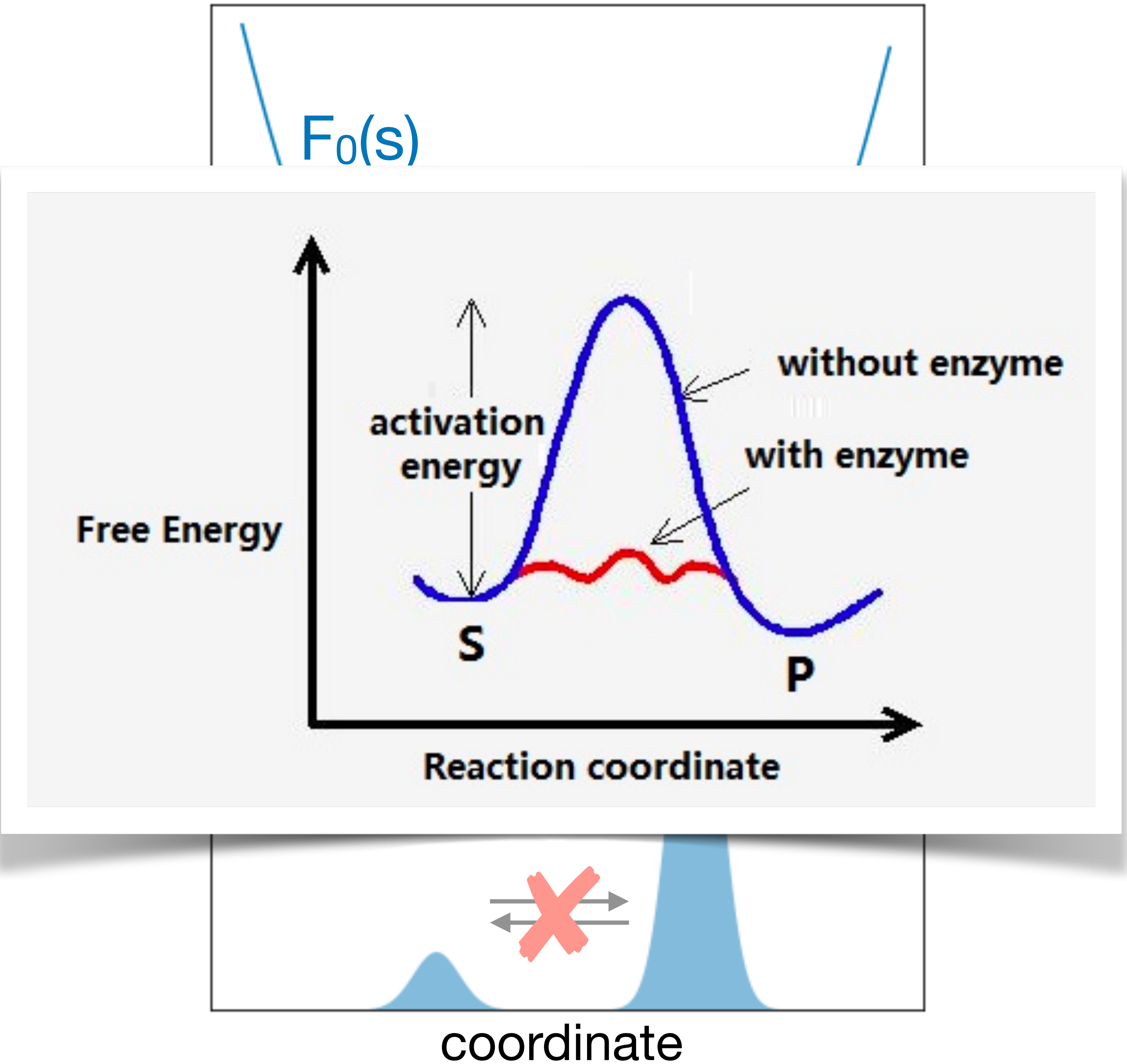
How biased sampling could help: umbrella sampling

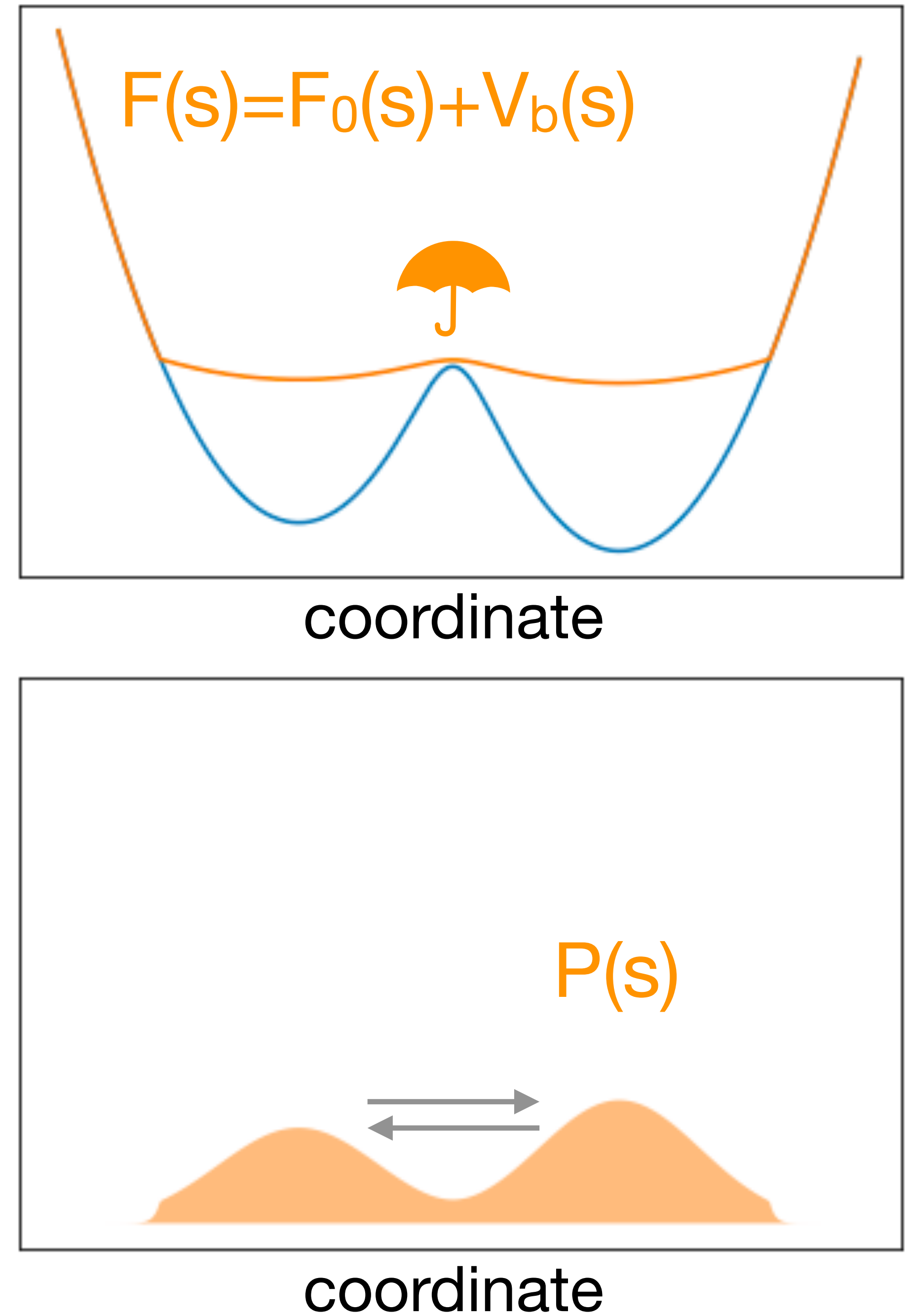


How biased sampling could help: umbrella sampling



How biased sampling could help: umbrella sampling





Disfavor reactant and product (i.e. favor transition state), like a biological enzyme

Unbiasing the result

how much the bias V_b affects the probability

number of times state j was visited

$$P(t_1, \dots, t_M) \propto \prod_{j=1}^{M} \left(e^{-\beta V_b(s_j)} p_j \right)^{t_j}$$

product over all states

unbiased probability of observing state j

Bussi and Tribello, arXiv (2018)

Unbiasing the result

how much the bias V_b affects the probability

number of times state *j* was visited

$$P(t_1,\dots,t_M) \propto \prod_{j=1}^{M} \left(e^{-\beta V_b(s_j)} p_j\right)^{t_j}$$

product over all states

unbiased probability of observing state *j*

Find p that maximises the likelihood

$$w(s) \propto e^{+\beta V_b(s)}$$

Averages can be computed by reweighing

$$\langle A \rangle = \frac{\sum_j A_j w(s_j)}{\sum_j w(s_j)}$$

Bussi and Tribello, arXiv (2018)

Unbiasing the result

how much the bias V_b affects the probability

number of times state *j* was visited

$$P(t_1,\dots,t_M) \propto \prod_{j=1}^{M} \left(e^{-\beta V_b(s_j)} p_j\right)^{t_j}$$

product over all states

unbiased probability of observing state *j*

Weight more if system went there even if disfavored

Find p that maximises the likelihood

$$w(s) \propto e^{+\beta V_b(s)}$$

Averages can be computed by reweighing

$$\langle A \rangle = \frac{\sum_j A_j w(s_j)}{\sum_j w(s_j)}$$

Bussi and Tribello, arXiv (2018)

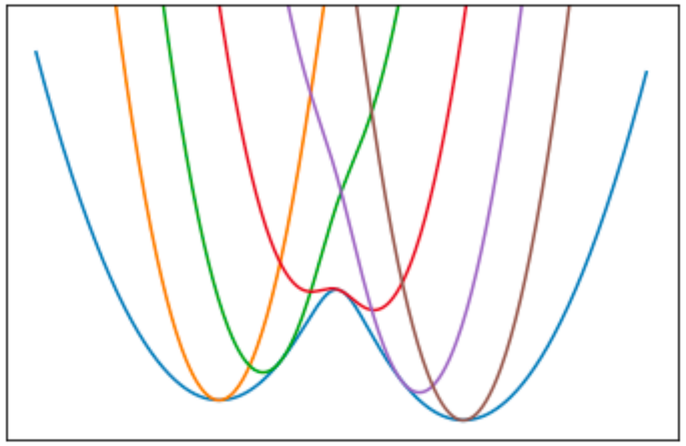
Multiple-windows umbrella sampling

Each simulation estimates the profile for a narrow range of the biased variable

stiffness

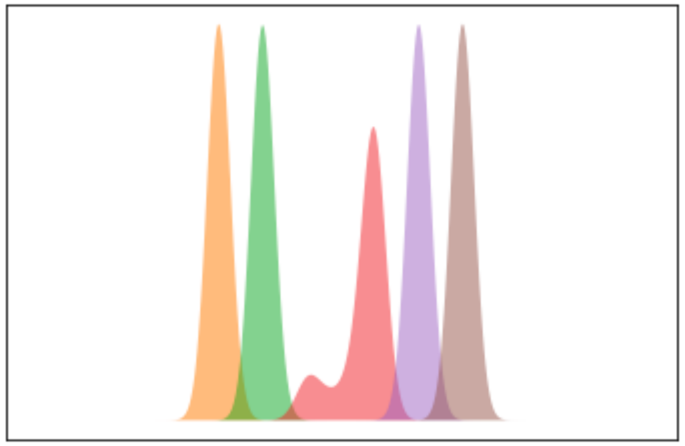
$$V_b^j(s) = \frac{k}{2}\left(s - s_j\right)^2$$

interpolate between reactant and product



coordinate

Overlapping histograms



coordinate

Combining multiple simulations: WHAM

how much the *k*-th bias affects the probability

number of times state *j* was visited in traj *k*

$$P(t_1^1, \ldots, t_M^N) \propto \prod_{j=1}^{M} \prod_{k=1}^{N} \left(c_k e^{-\beta V_b^k(s_j)} p_j \right)^{t_{kj}}$$

product over all states and all trajectories

make sure each probability is normalized

unbiased probability of observing state *j*

Bussi and Tribello, arXiv (2018)

Combining multiple simulations: WHAM

$$P(t_1^1, \ldots, t_M^N) \propto \prod_{j=1}^{M} \prod_{k=1}^{N} \left(c_k e^{-\beta V_b^k(s_j)} p_j \right)^{t_{kj}}$$

how much the *k*-th bias affects the probability

number of times state *j* was visited in traj *k*

product over all states and all trajectories

make sure each probability is normalized

unbiased probability of observing state *j*

Find p that maximises the likelihood

Bussi and Tribello, arXiv (2018)

Combining multiple simulations: WHAM

how much the k-th bias affects the probability

number of times state j was visited in traj k

$$P(t_1^1, \ldots, t_M^N) \propto \prod_{j=1}^{M} \prod_{k=1}^{N} \left(c_k e^{-\beta V_b^k(s_j)} p_j \right)^{t_{kj}}$$

product over all states and all trajectories

make sure each probability is normalized

unbiased probability of observing state j

$$c_k = \frac{1}{\sum_j e^{-\beta V_b^k(s_j)} w(s_j)}$$

$$w(s) \propto \frac{1}{\sum_k c_k e^{-\beta V_b^k(s)}}$$

Self-consistent equations

Find p that maximises the likelihood

Bussi and Tribello, arXiv (2018)

Combining multiple simulations: WHAM

how much the k-th bias affects the probability

number of times state j was visited in traj k

$$P(t_1^1, \ldots, t_M^N) \propto \prod_{j=1}^{M} \prod_{k=1}^{N} \left(c_k e^{-\beta V_b^k(s_j)} p_j \right)^{t_{kj}}$$

product over all states and all trajectories

make sure each probability is normalized

unbiased probability of observing state j

$$c_k = \frac{1}{\sum_j e^{-\beta V_b^k(s_j)} w(s_j)}$$

$$w(s) \propto \frac{1}{\sum_k c_k e^{-\beta V_b^k(s)}}$$

Self-consistent equations

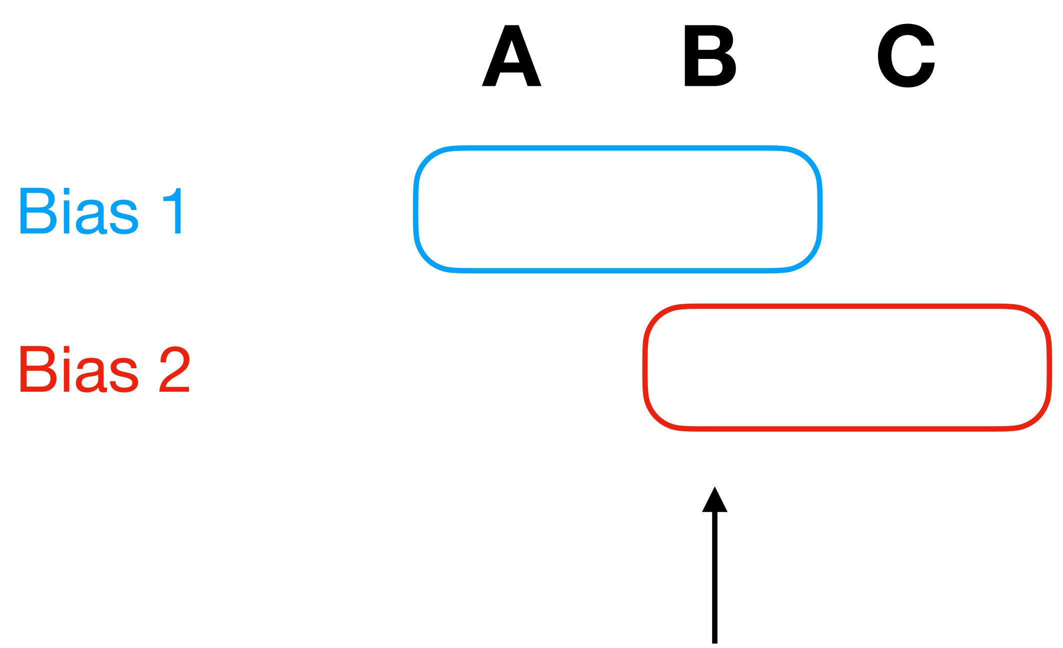
Find p that maximises the likelihood

Averages can be computed by reweighing

$$\langle A \rangle = \frac{\sum_j A_j w(s_j)}{\sum_j w(s_j)}$$

Bussi and Tribello, arXiv (2018)

Combining multiple simulations: WHAM



Twice as many samples
in B since it can be seen
in both simulations

$$c_k = \frac{1}{\sum_j e^{-\beta V_b^k(s_j)} w(s_j)}$$

$$w(s) \propto \frac{1}{\sum_k c_k e^{-\beta V_b^k(s)}}$$

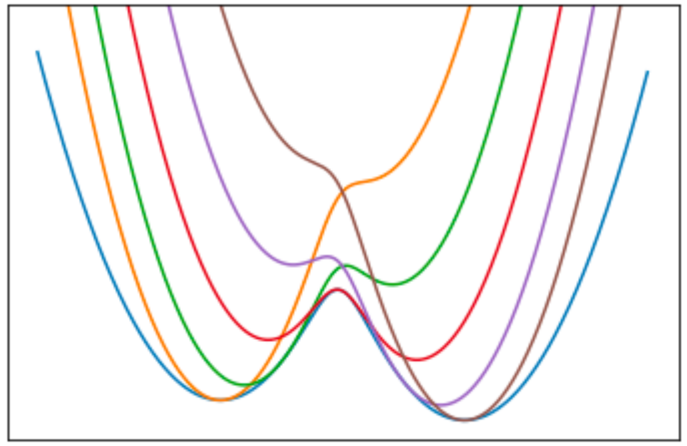
Self-consistent equations

Pay attention to stiffness

stiffness

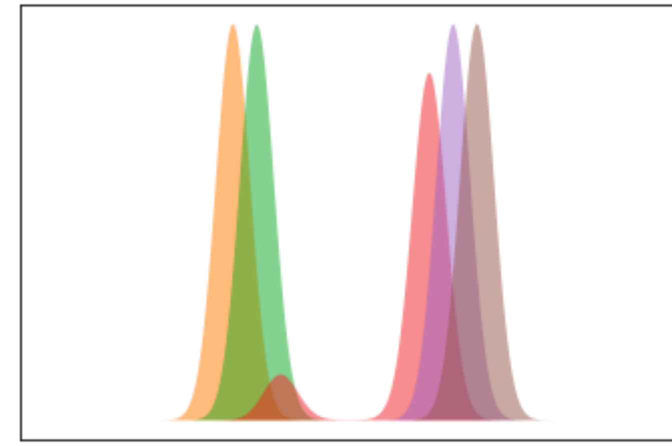
$$V_b^j(s) = \frac{k}{2}\left(s - s_j\right)^2$$

interpolate between reactant and product

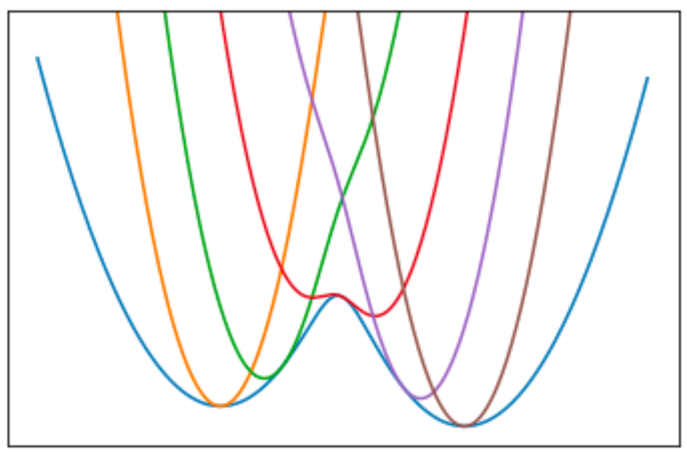


coordinate

Too soft:
Not sampling TS

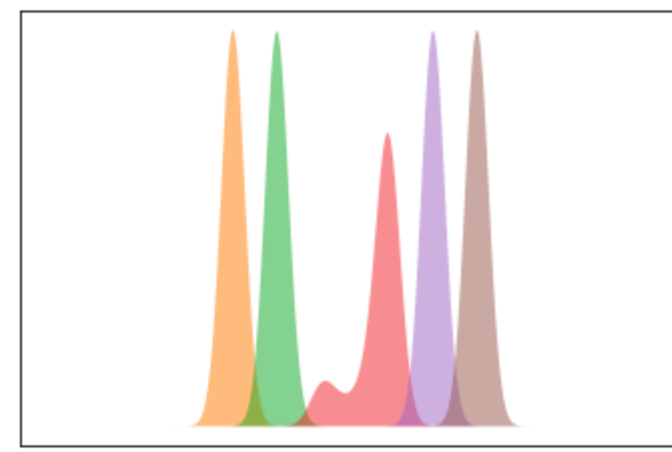


coordinate

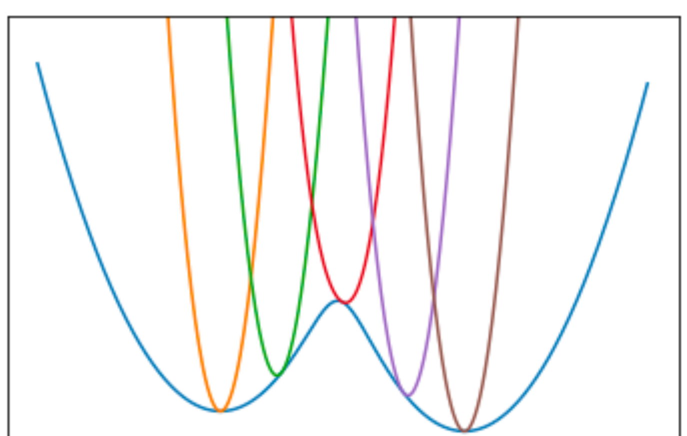


coordinate

Overlapping histograms

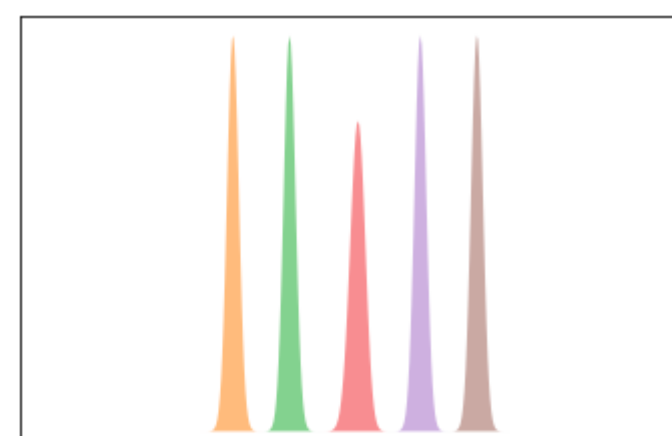


coordinate



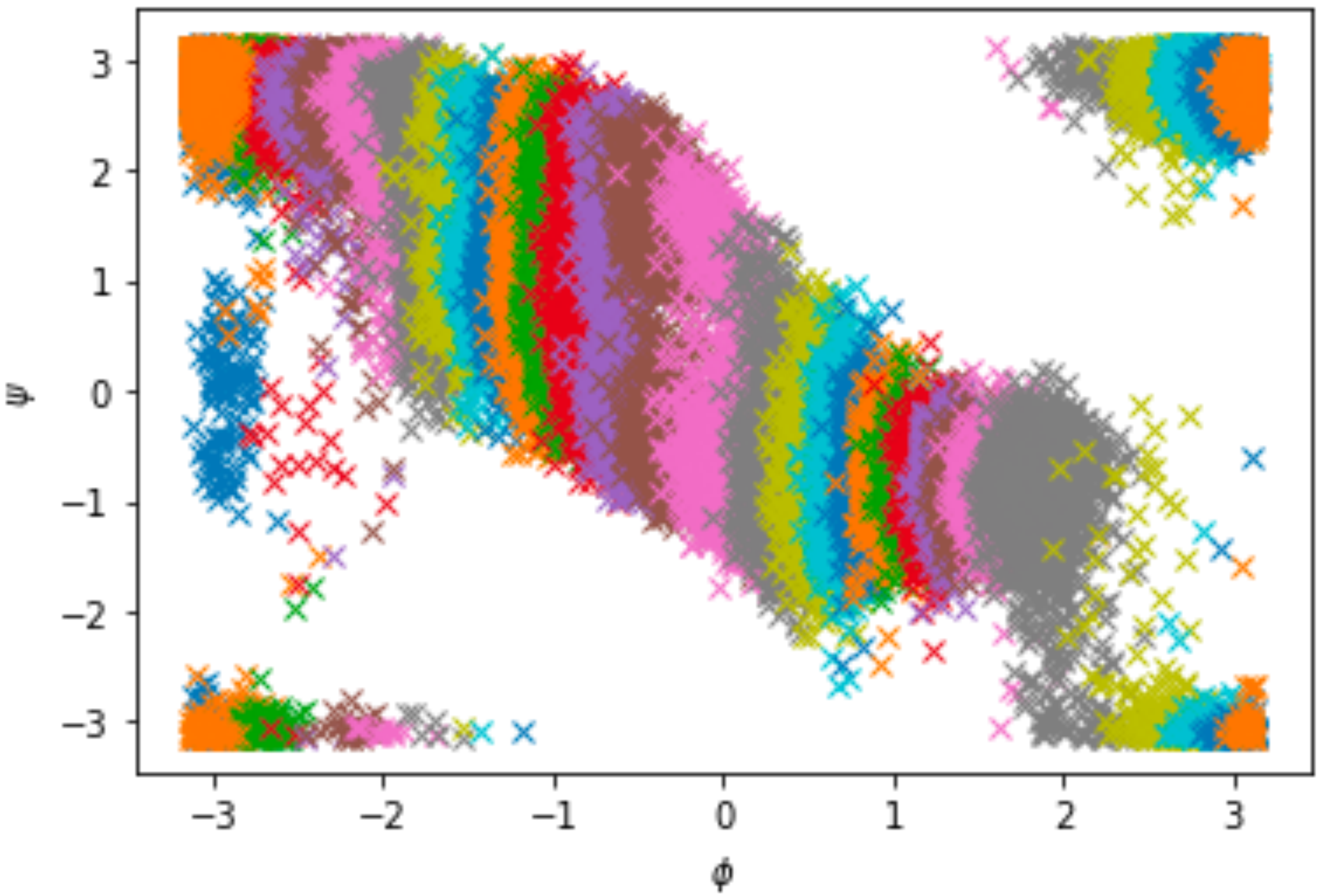
coordinate

Too stiff:
No overlap!



coordinate

Checking for overlap

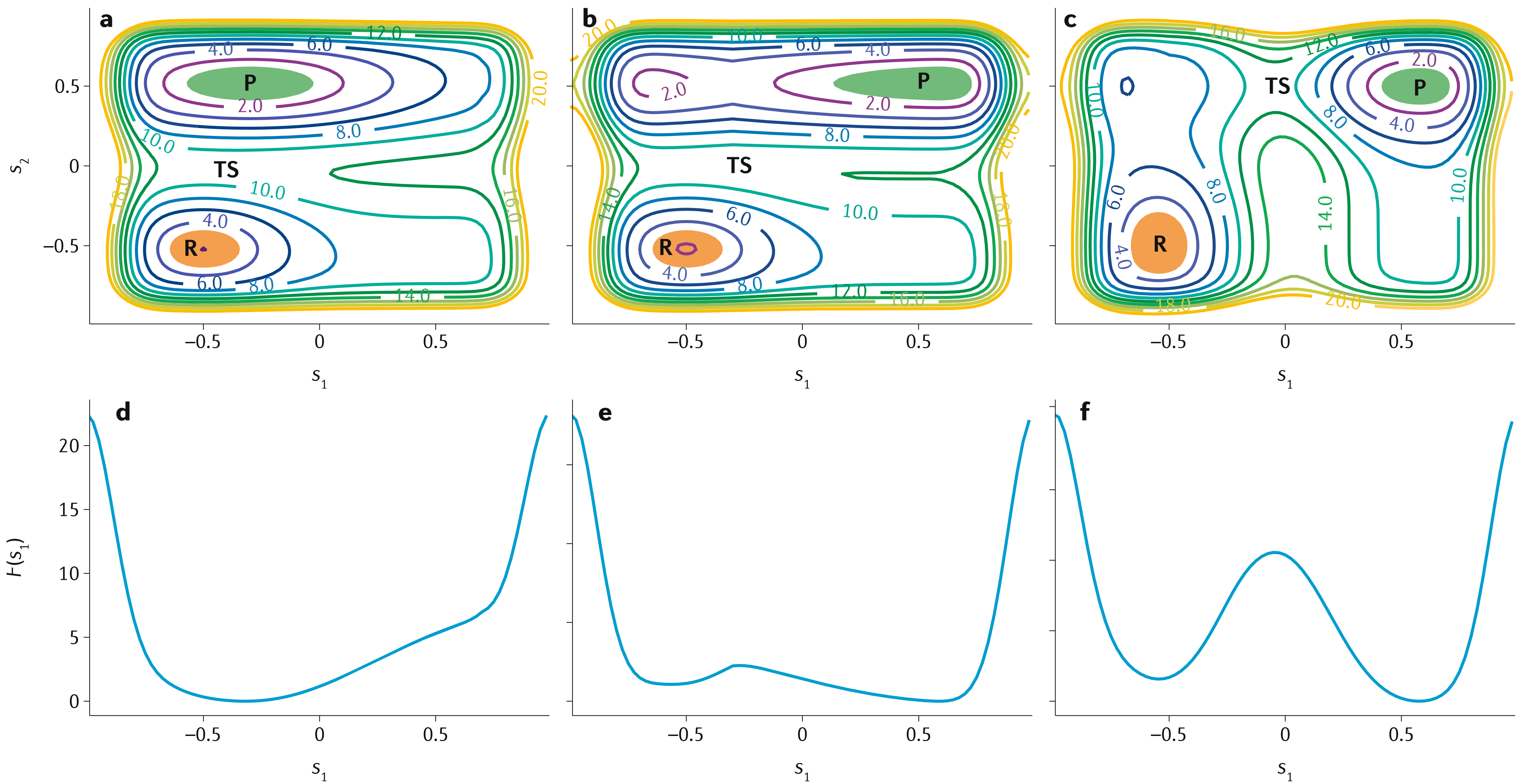


Simulations should overlap also in not-biased degrees of freedom

Very difficult to check without doing replica exchange (<u>come to masterclass 4</u>!)

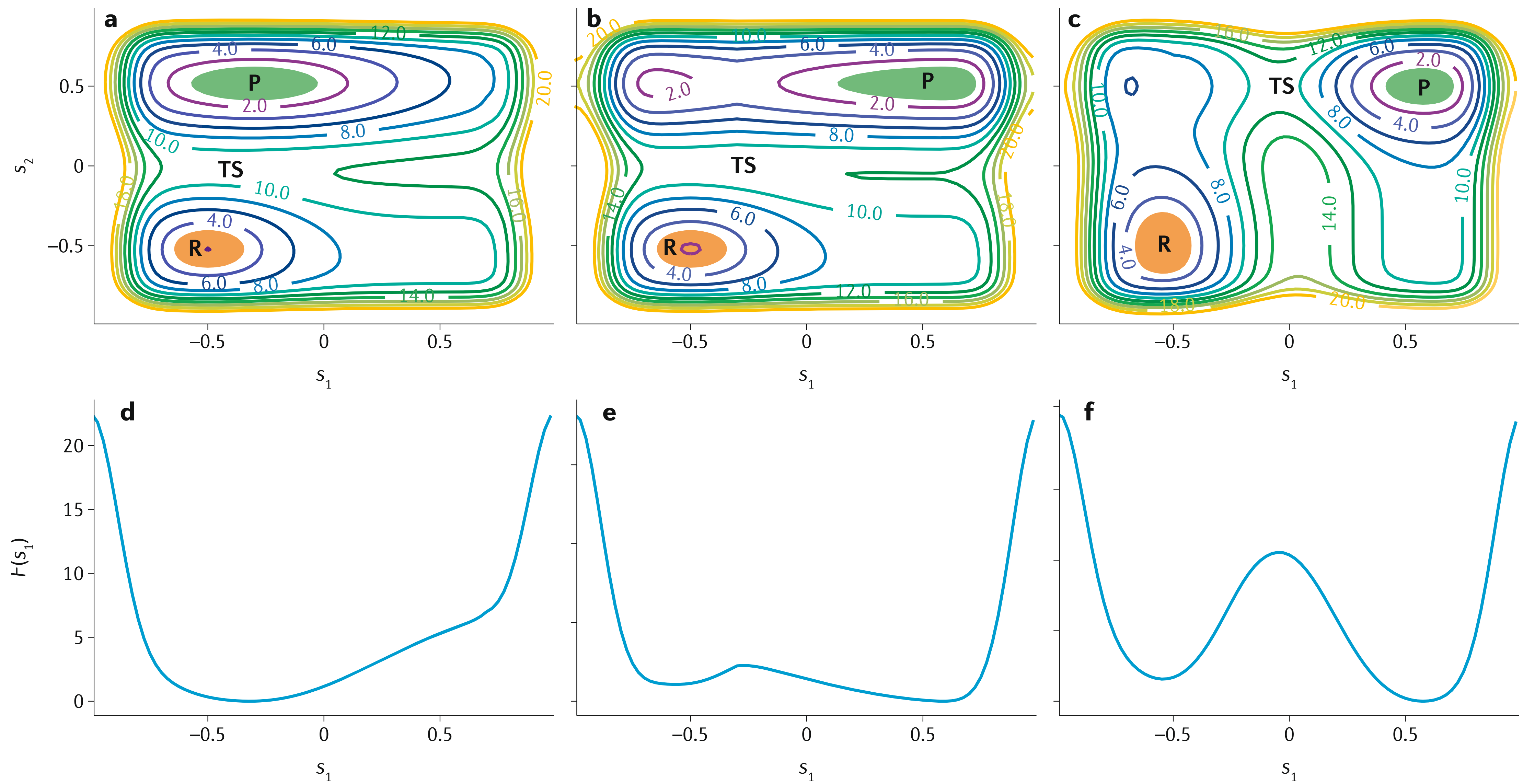
A minimal check: compare results initialising in R vs P

Reaction coordinates: good ones vs bad ones



Bussi and Laio, Nat Rev Phys (2020)

Reaction coordinates: good ones vs bad ones



	R ≈ P	R ≈ TS ≠ P or vice versa	R ≠ TS ≠ P
Analysis			
Enhanced sampling			

Bussi and Laio, Nat Rev Phys (2020)

Error calculations

time series → **analysis**

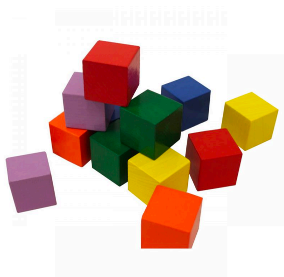
A

Efron and Tibshirani, 1993

Error calculations

time series → **analysis**

A



blocks - minimize correlations

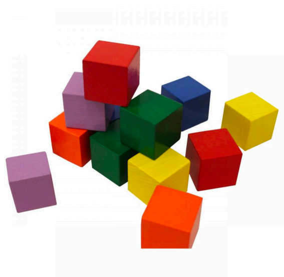
0 1 2 3 4 5 6 7 8 9

Efron and Tibshirani, 1993

Error calculations

time series → **analysis**

A



blocks - minimize correlations

0	1	2	3	4	5	6	7	8	9

bootstrap samples

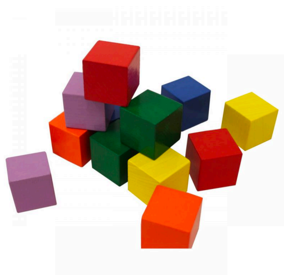
0	5	0	5	2	8	8	4	9	0	1
1	1	9	8	1	9	5	9	4	6	1
2	7	2	4	2	3	9	4	8	8	4
...	...	...	...	...	...	...	...	...	...	...
198	5	2	3	2	4	5	0	6	8	5
199	6	5	6	6	2	5	3	2	5	2

Efron and Tibshirani, 1993

Error calculations

time series → **analysis**

A

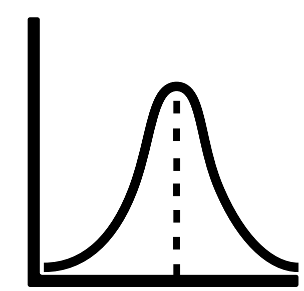


blocks - minimize correlations

0	1	2	3	4	5	6	7	8	9

bootstrap samples

0	5	0	5	2	8	8	4	9	0	1	A_0
1	1	9	8	1	9	5	9	4	6	1	A_1
2	7	2	4	2	3	9	4	8	8	4	A_2
...	...	...	...	...	...	...	...	...	...	...	
198	5	2	3	2	4	5	0	6	8	5	A_{198}
199	6	5	6	6	2	5	3	2	5	2	A_{199}



Efron and Tibshirani, 1993

Useful references on WHAM

Ferrenberg and Swendsen, PRL (1989)
Initial formulation

Kumar et al JCC (1992)
First example in biomolecular simulations

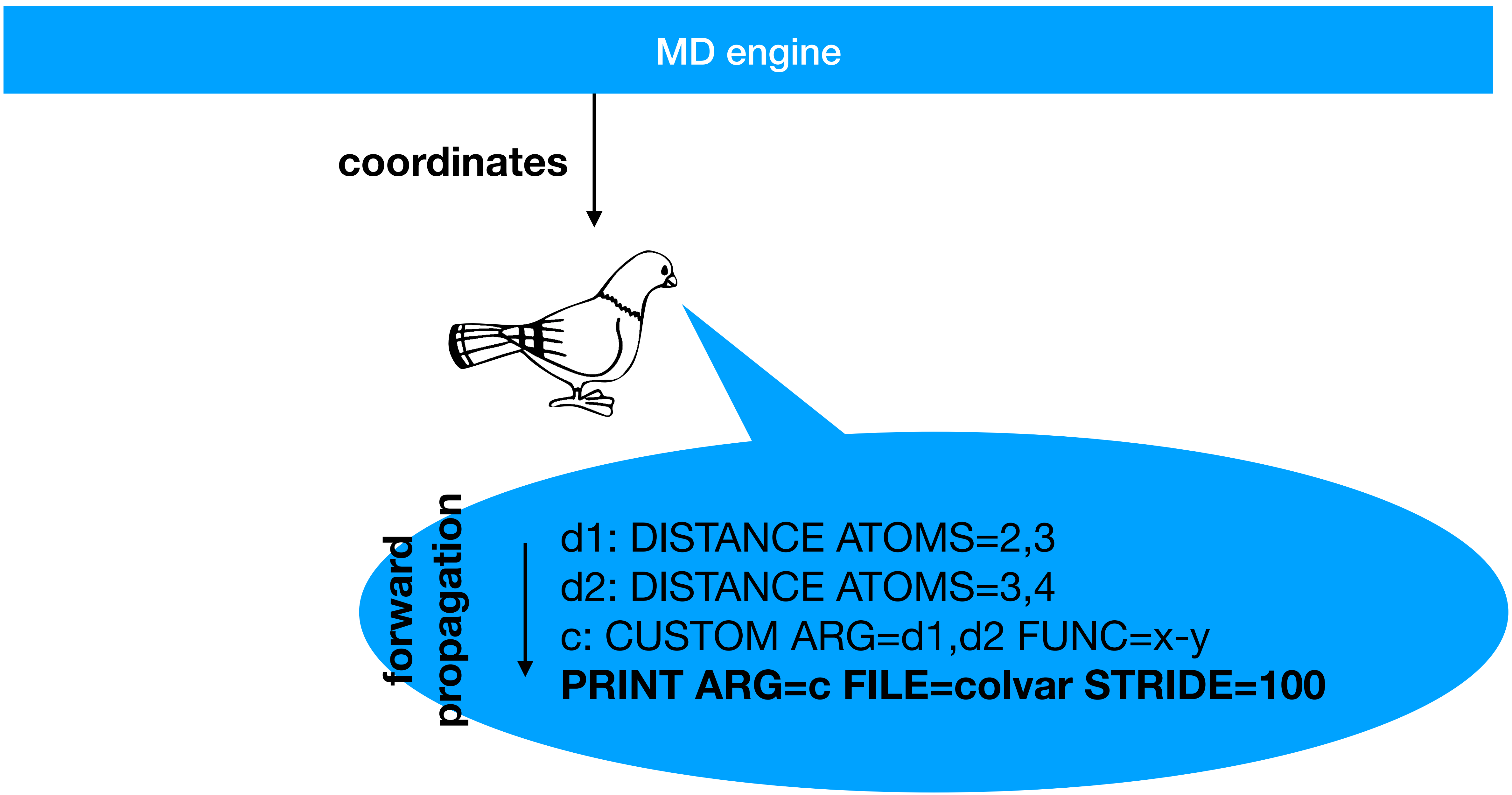
Souaille and Roux CPC (2001)
Weights for individual frames

Shirts and Chodera JCP (2008)
MBAR, can be shown to be equivalent to binless WHAM

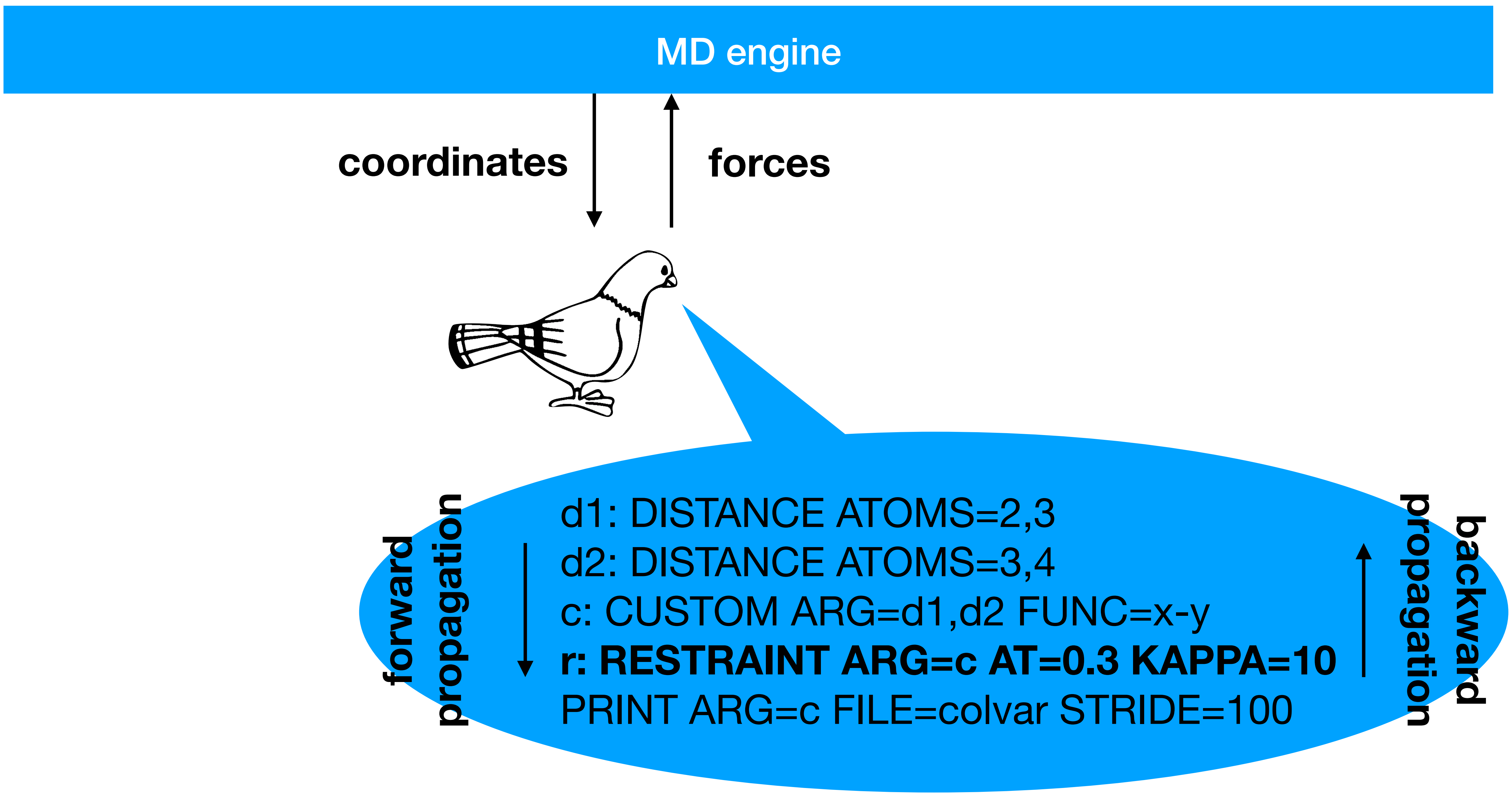
Hub et al, JCTC (2010)
Estimating error using bootstrap

Tan, Gallicchio, Lapelosa, and Levy JCP (2012)
Show that WHAM can be used in a binless fashion

Using PLUMED for on-the-fly analysis



Using PLUMED to bias a simulation



Using PLUMED for WHAM analysis

PLUMED driver on concatenated trajectory

+

REWEIGHT_WHAM action
(integrated with PLUMED)

Or

An external python tool, wham.py
(allow more flexibility for bootstrapping)

Note: both approaches are bin-less, and allow arbitrary bias potentials to be used in all replicas (i.e., not just a series of restraints on the same variable)

Using PLUMED for WHAM analysis

PLUMED driver on concatenated trajectory

+

REWEIGHT_WHAM action
(integrated with PLUMED)

Or

An external python tool, wham.py
(allow more flexibility for bootstrapping)

Note: both approaches are bin-less, and allow arbitrary bias potentials to be used in all replicas (i.e., not just a series of restraints on the same variable)

Instructions

Class	Topic	Lecture I	Lecture II	Instructor
21.1	PLUMED syntax and analysis	January 18, 2021	January 25, 2021	M. Bonomi
21.2	Statistical errors in MD	February 1, 2021	February 8, 2021	G. Tribello
21.3	Umbrella sampling	February 15, 2021	February 22, 2021	G. Bussi
21.4	Metadynamics	March 1, 2021	March 8, 2021	M. Bonomi
21.5	Replica exchange methods	March 15, 2021	March 22, 2021	G. Bussi
21.6	Dimensionality reduction	April 12, 2021	April 19, 2021	G. Tribello
21.7	Performance optimization	April 26, 2021	May 3, 2021	M. Bonomi
21.8	Poster session	May 10, 2021		

1. Go to www.plumed.org
2. Click on the Masterclass tab
3. Click on the Topic of class 21.3
4. 1 week to complete the exercises
5. Questions/discussions on Slack channel masterclass-21-3
6. Lecture I and II available on YouTube