

Hylleraas School 2025

4 – enhanced sampling

Molecular dynamics – Sampling technique

$$\langle \mathcal{O} \rangle = \frac{1}{Q} \int \cdots \int \mathcal{O}(\{\mathbf{r}, \mathbf{p}\}) e^{-\frac{E(\{\mathbf{r}, \mathbf{p}\})}{kT}} d\Gamma = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \mathcal{O}(t) dt$$

$$Q = \iint e^{-E(\mathbf{r}, \mathbf{p})/kT} d\Gamma \quad A = -kT \ln Q$$

Discussion time!!!

...What is the free energy?????

Find the most concise definition for it

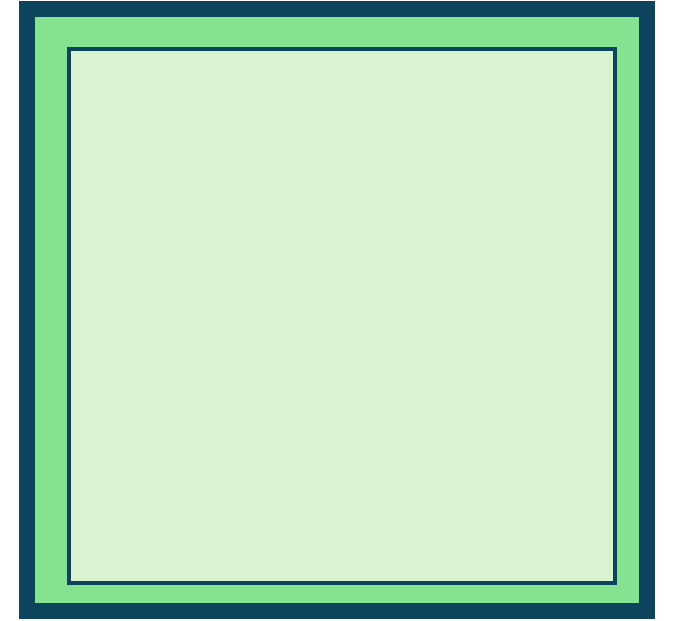
The free energy is nothing else than the

Entropy of the universe,

defined on the variables of a subsystem of interest.

Microcanonical conditions

- Constant N, V, E (close, isolated system)
- $S = S(N, V, E)$ **good** thermodynamic potential

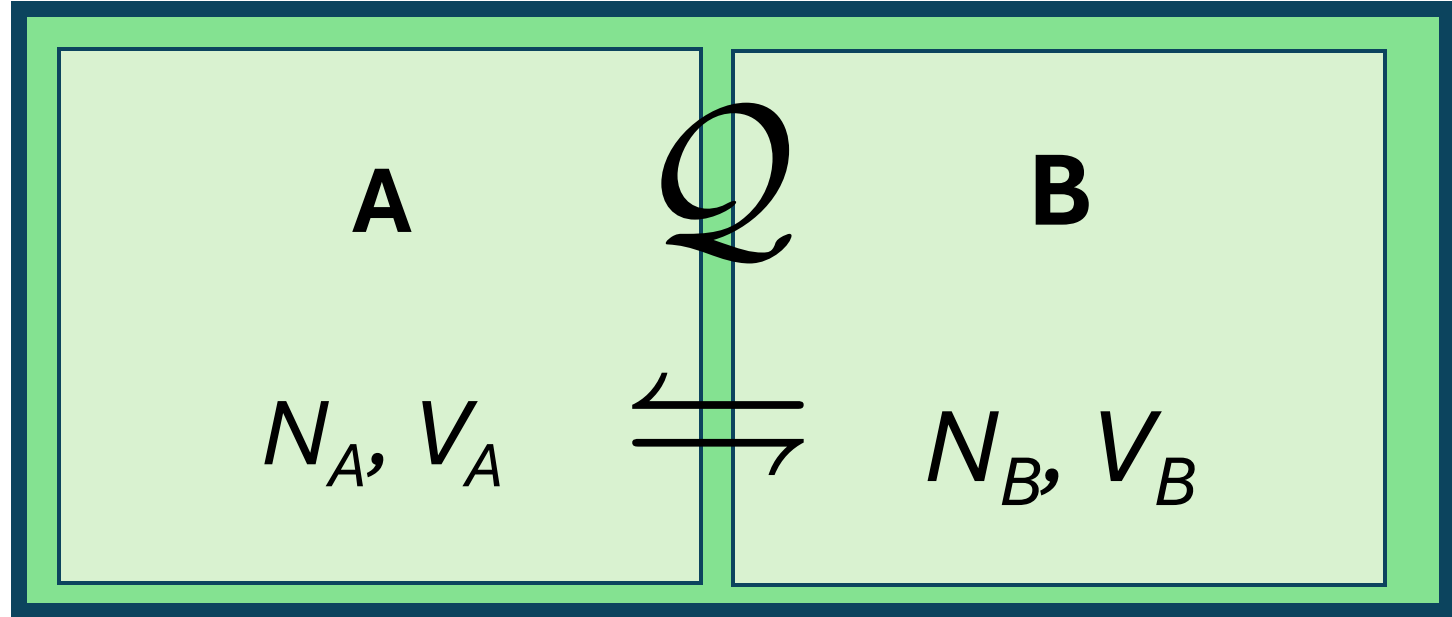


$$\Delta S > 0$$

Condition of equilibrium: ***S is at its maximum***

Microcanonical conditions - subsystems

- $E = E_A + E_B$
- E constant



$$\frac{\partial S_A}{\partial E_A}(\overline{E_A}) = \frac{\partial S_B}{\partial E_B}(\overline{E_B}) \equiv \frac{1}{T}$$

- Natural variables: $N, V, \mathbf{T}$

→ S is **not** a good thermodynamic potential

Thermal bath

- $E = E_A + E_B$

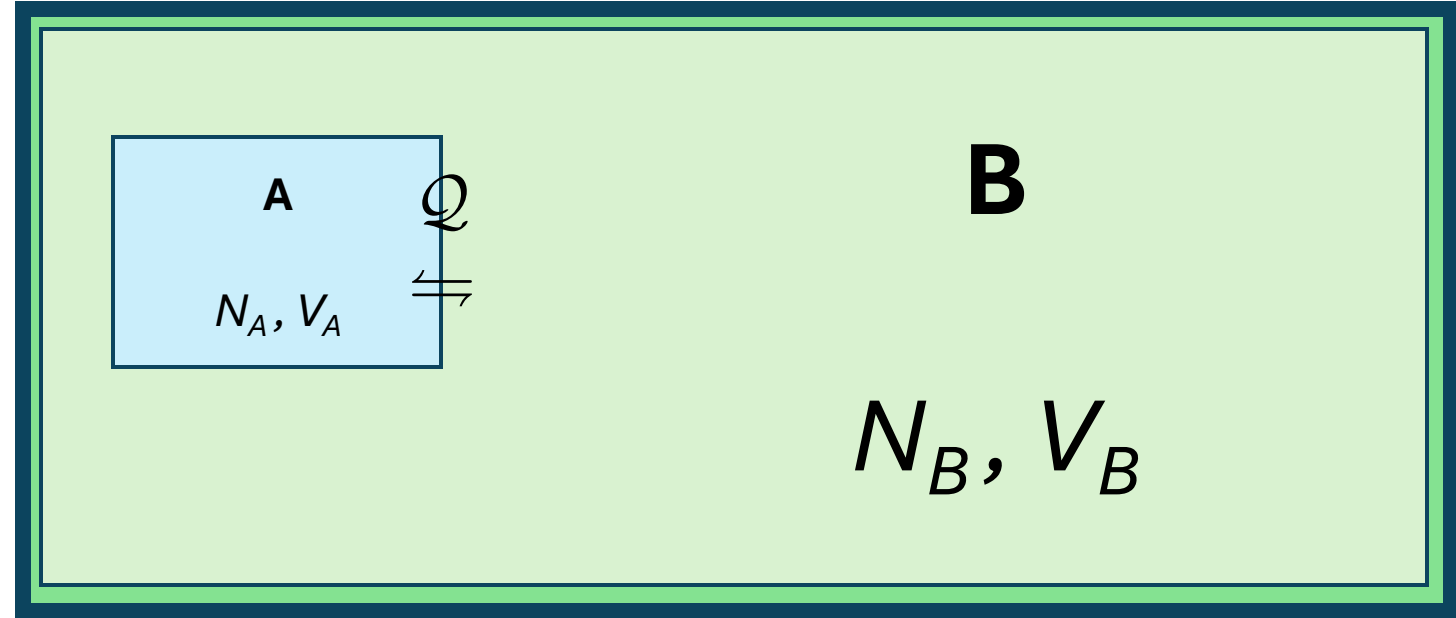
$$\overline{E}_A \ll \overline{E}_B \sim E$$

$$S(E) = S_A(\overline{E}_A) + S_B(\overline{E}_B)$$

$$dS = dS_A(\overline{E}_A) + dS_B(\overline{E}_B) = 0$$

$$dS = dS_A(\overline{E}_A) + \frac{\partial S_B(\overline{E}_B)}{\partial E_B} dE_B = 0$$

$$dS_A(E_A) - \frac{\partial S_B(E_B)}{\partial E_B} dE_A > 0$$



$$dE_B = -dE_A$$

$$dS_A(E_A) - \frac{dE_A}{T_B} > 0$$

$$dS_A(\overline{E}_A) - \frac{dE_A}{T_B} = 0$$

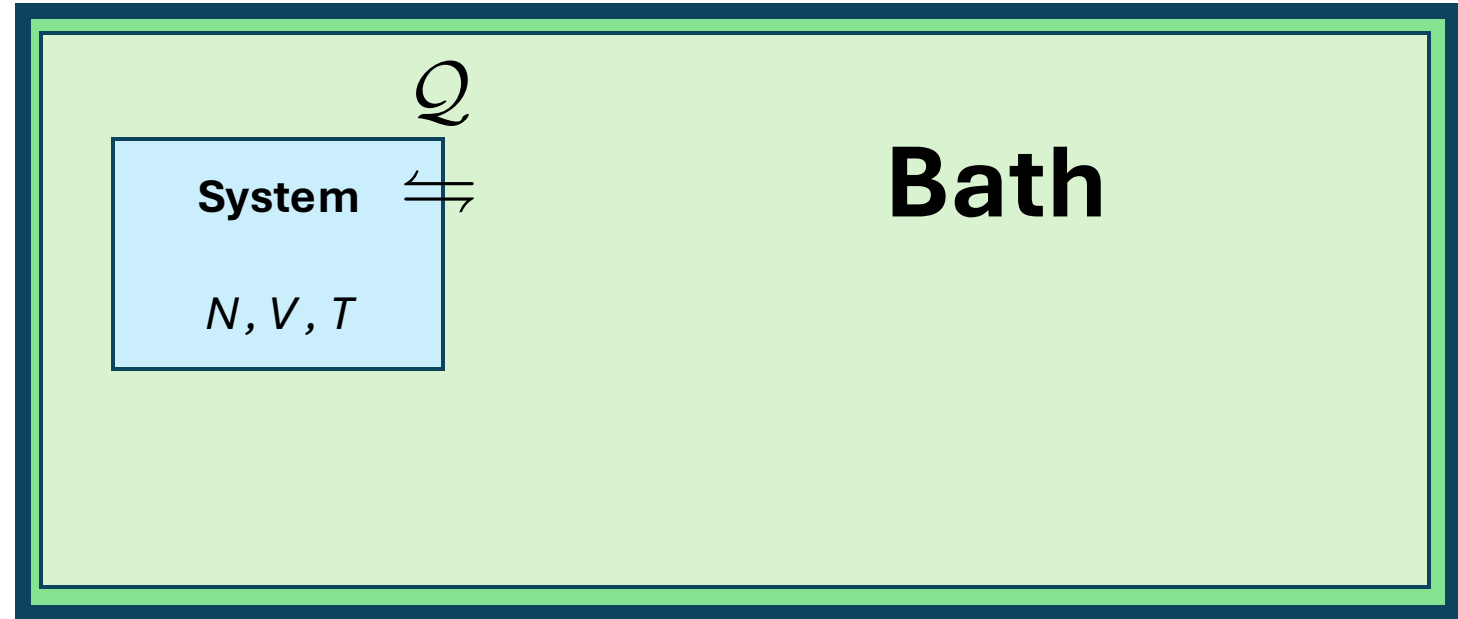
Canonical Conditions: N, V, T

Out of equilibrium:

$$dS(E) - \frac{dE}{T} > 0$$

At equilibrium:

$$dS(E) - \frac{dE}{T} = 0$$



New function: **Helmholtz Free energy**

$$F = F(N, V, T) = U - TS$$

$$dF = dU - TdS \quad dF = 0$$

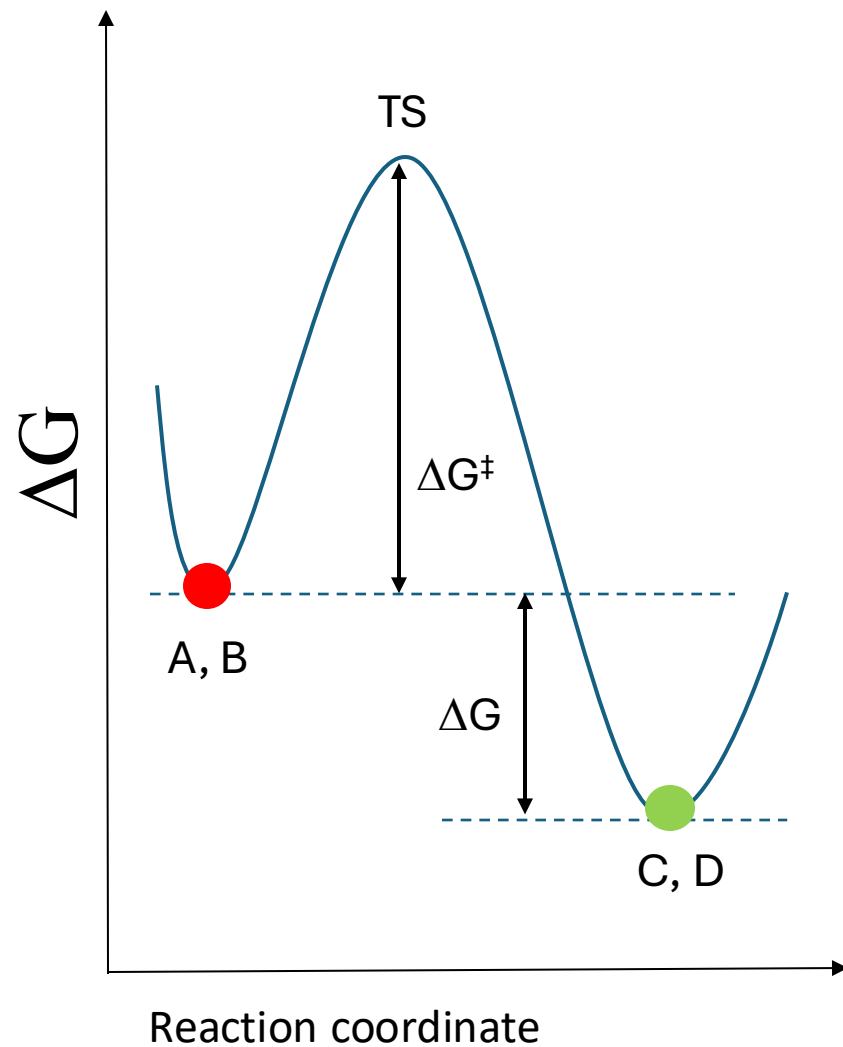
Condition of equilibrium: ***F (or A...) is at its minimum***

U: letter indicating the total energy (E) in thermodynamics

Back to sampling...

$$\langle \mathcal{O} \rangle = \frac{1}{Q} \int \cdots \int \mathcal{O}(\{\mathbf{r}, \mathbf{p}\}) e^{-\frac{E(\{\mathbf{r}, \mathbf{p}\})}{kT}} d\Gamma = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \mathcal{O}(t) dt$$

$$Q = \iint e^{-E(\mathbf{r}, \mathbf{p})/kT} d\Gamma \quad A = -kT \ln Q$$



- The energy is a well-defined property for any microscopic state:

$$H = H(\{\mathbf{r}, \mathbf{p}\})$$

- The free energy (like the entropy) is well defined for an ensemble of microscopic states:

$$A = -kT \ln Q$$

$$Q = \iint e^{-E(\mathbf{r}, \mathbf{p})/kT} d\Gamma$$

Discussion time!!

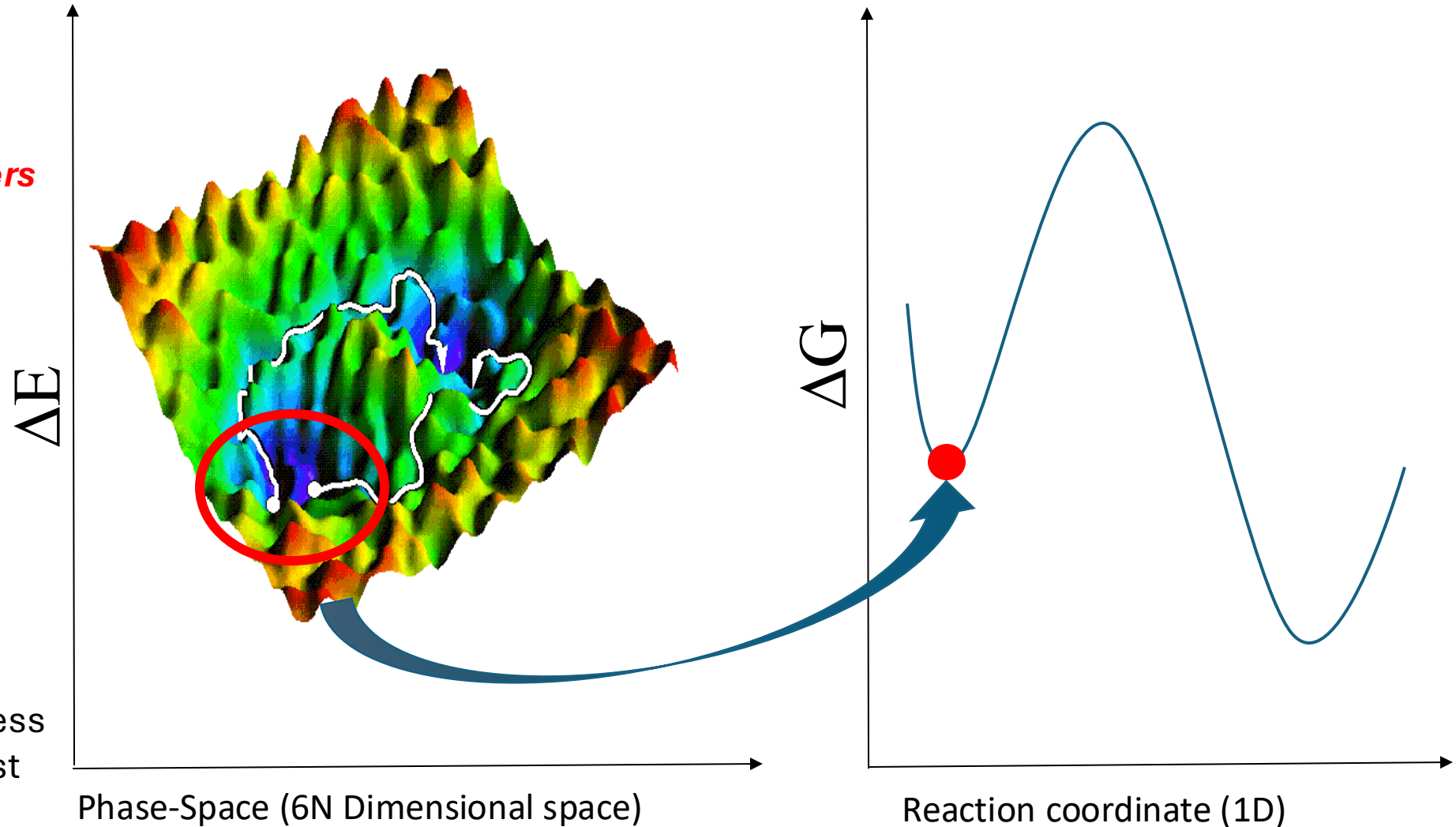
- Is there anything like a reactant state structure?
- Is there anything like a transition state structure?
- Is a geometry optimised structure representative of what?

Contraction from PES to FES

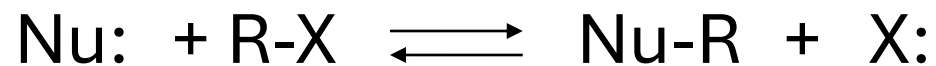
A free energy map is a **contraction** of the phase-space into a subspace of **order parameters** that characterize the system of interest.

This means that one point on a free-energy plot corresponds to **several** microscopic states.

Order parameters are **global variables** that signal the progress of any transformation of interest



An example: Nucleophilic substitution

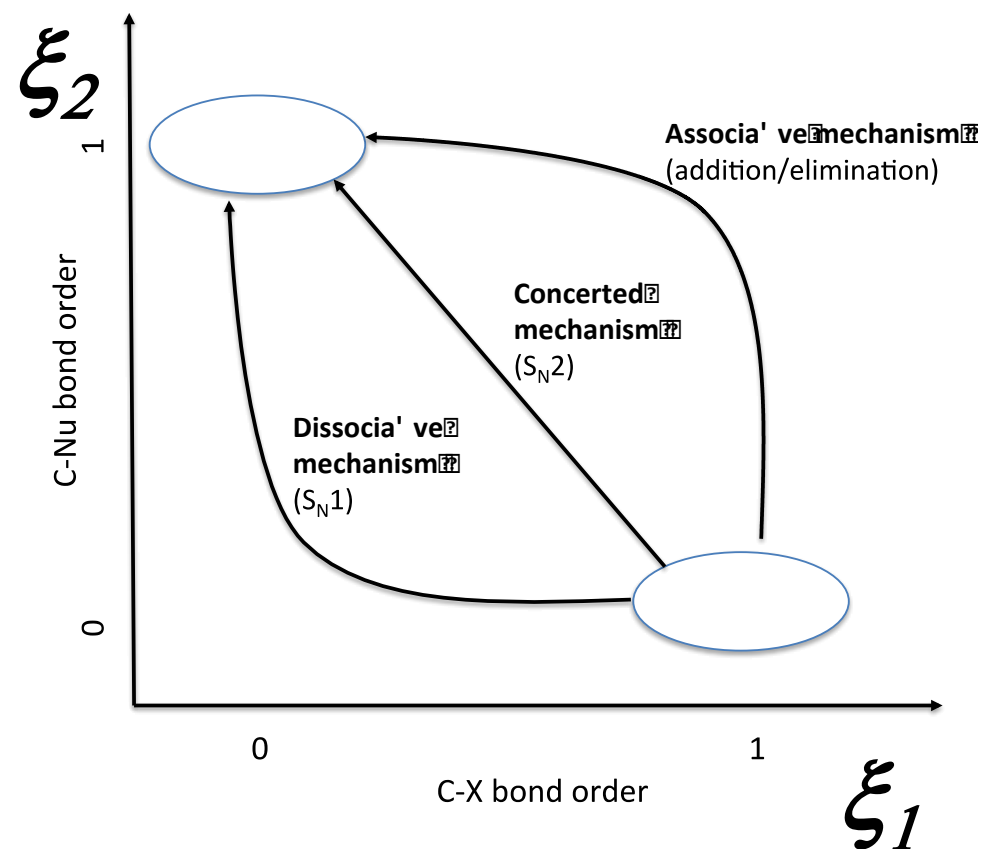


$$\xi = \xi(\{\mathbf{r}, \mathbf{p}\})$$

- **Global** variable (in principle it depends on **all** DoF)
- **Slow-varying** variable (It signals the large-scale transformation in the chemical space, not some fast motion in the conformational space)

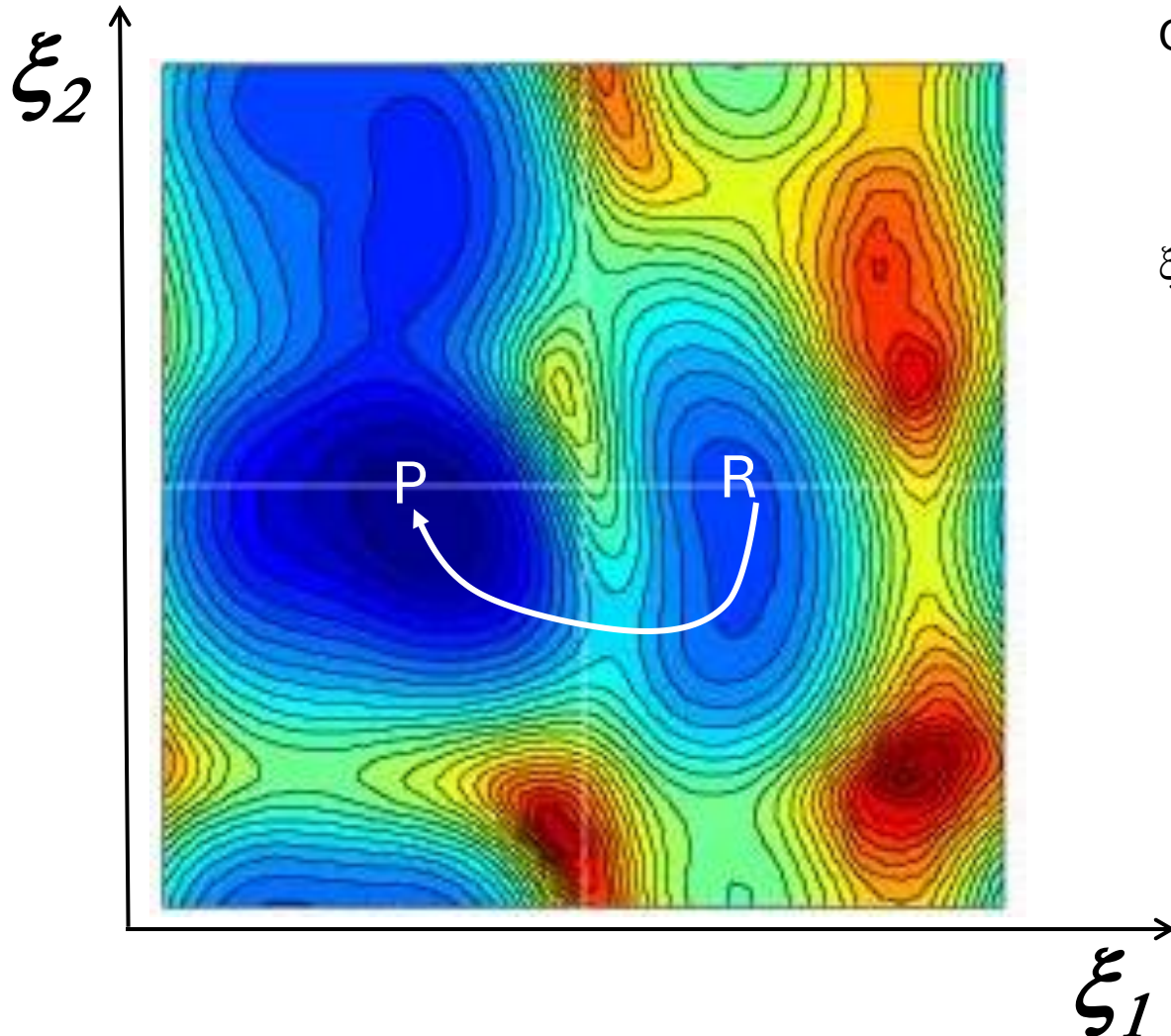
C-X distance: **A BAD** ORDER PARAMETER!!!

C-X bond order: **A GOOD** ORDER PARAMETER



A free energy map is obtained by properly weighing states

$$Q = \iint e^{-E(\mathbf{r}, \mathbf{p})/kT} d\Gamma$$



Canonical partition function: $A = -kT \ln Q$

ξ are defined by the microstates: $\xi = \xi(\{\mathbf{r}\}, \{\mathbf{p}\})$

The partition function conditional to the order parameters is:

$$Q(\xi_1, \xi_2) = \iint \delta(\xi_1(\mathbf{r}, \mathbf{p}) - \xi_1) \delta(\xi_2(\mathbf{r}, \mathbf{p}) - \xi_2) e^{-\beta E(\mathbf{r}, \mathbf{p})} d\Gamma$$

$$A(\xi_1, \xi_2) = -kT \ln Q(\xi_1, \xi_2)$$

Reaction coordinate:

$\lambda = \lambda(\xi_1, \xi_2)$ Variational path connecting two minima of interest over the free energy surface.

A free energy map is obtained by properly weighing states

$$Q = \iint e^{-E(\mathbf{r}, \mathbf{p})/kT} d\Gamma$$

Canonical partition function:

$$Q$$

ξ are defined

$$\xi = \xi(\{\mathbf{r}\}, \{\mathbf{p}\})$$

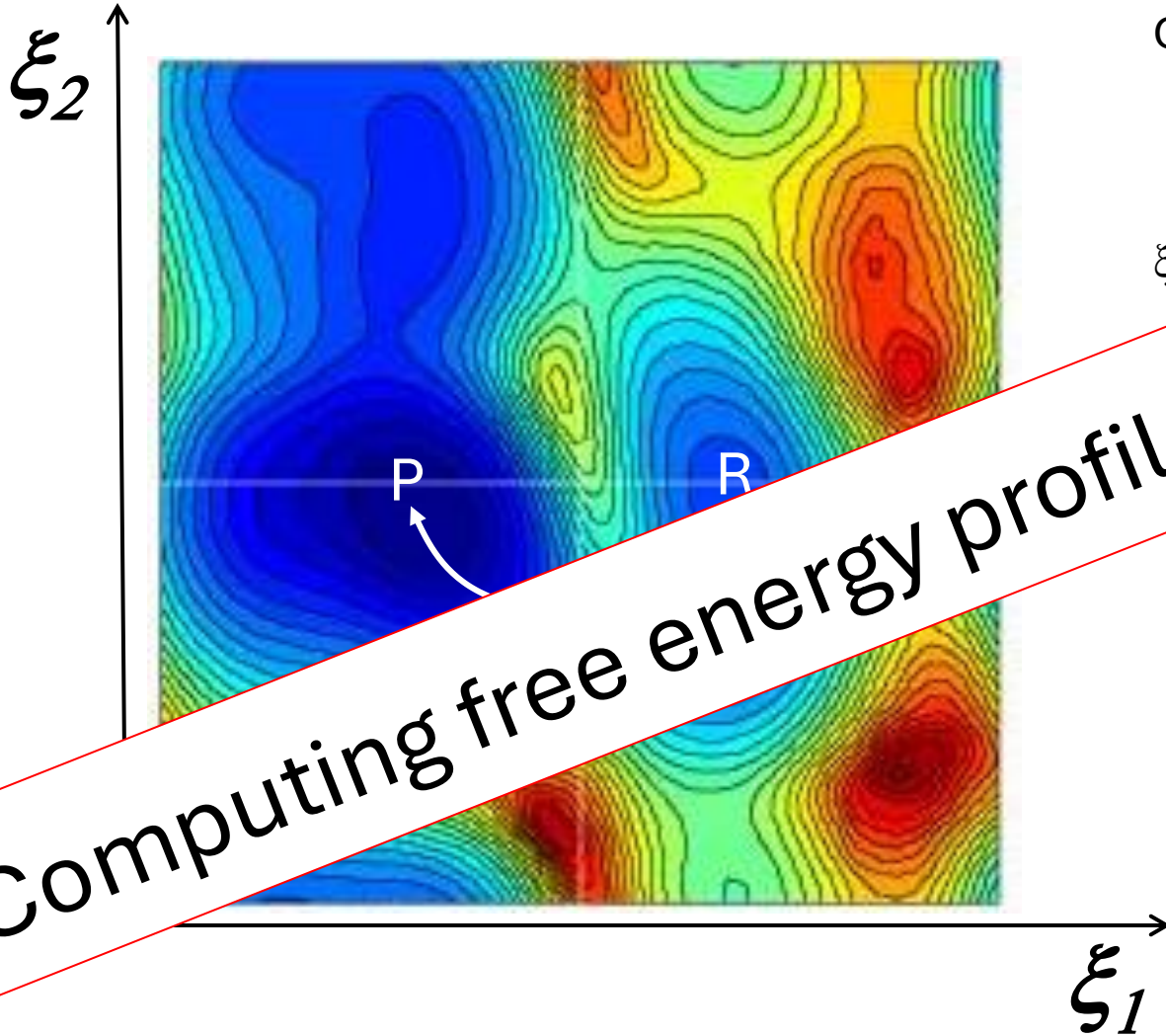
function conditional to the order parameters is:

$$Q(\xi_1, \xi_2) = \iint \delta(\xi_1(\mathbf{r}, \mathbf{p}) - \xi_1) \delta(\xi_2(\mathbf{r}, \mathbf{p}) - \xi_2) e^{-\beta E(\mathbf{r}, \mathbf{p})} d\Gamma$$

$$A(\xi_1, \xi_2) = -kT \ln Q(\xi_1, \xi_2)$$

Reaction coordinate:

$\lambda = \lambda(\xi_1, \xi_2)$ Variational path connecting two minima of interest over the free energy surface.



What is a “state” in chemical space?

A geometry optimised structure of a system of interest is **NOT** a reactant (or product, or whatever) state in any chemical space

What is the probability of finding a molecule in its minimum potential energy if it is at a temperature $T > 0$?

ZERO

In a continuum of states, we can only define a probability density:

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta H(\mathbf{r}, \mathbf{p})}}{Q}$$

... then, why do we bother so much about it?

Minimum Energy Structure $\neq$ Reactant (or product) state

... then, why do we bother so much about it?

Because sampling is complicated...

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta H(\mathbf{r}, \mathbf{p})}}{Q}$$

We do not know the probability function unless all the states are sampled (Q is unknown)

But we know the relative probability density for two states:

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = e^{-\beta \Delta E}$$

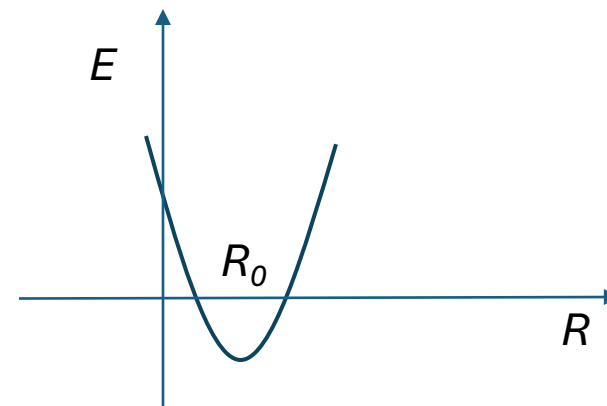
requires sampling by MD or Monte Carlo, which is computationally expensive,
and, actually, not sufficient, as we will see later

The Harmonic approximation

- Let's assume that we have identified a model geometry that can be unequivocally assigned to the reactant state.
- By optimising its geometry, we can determine the probability of all the states that are produced by its harmonic deformations.

$$E(R) \approx E(R_0) + E'(R_0)\Delta R + \frac{1}{2}E''(R_0)\Delta R^2 \rightarrow \Delta E \approx \frac{1}{2}\kappa\Delta R^2$$

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = e^{-\beta\Delta E}$$



The Harmonic approximation (2)

$$Q_{vib}^{QM} = Q_{vib}^{CM}$$

(See lesson 1)

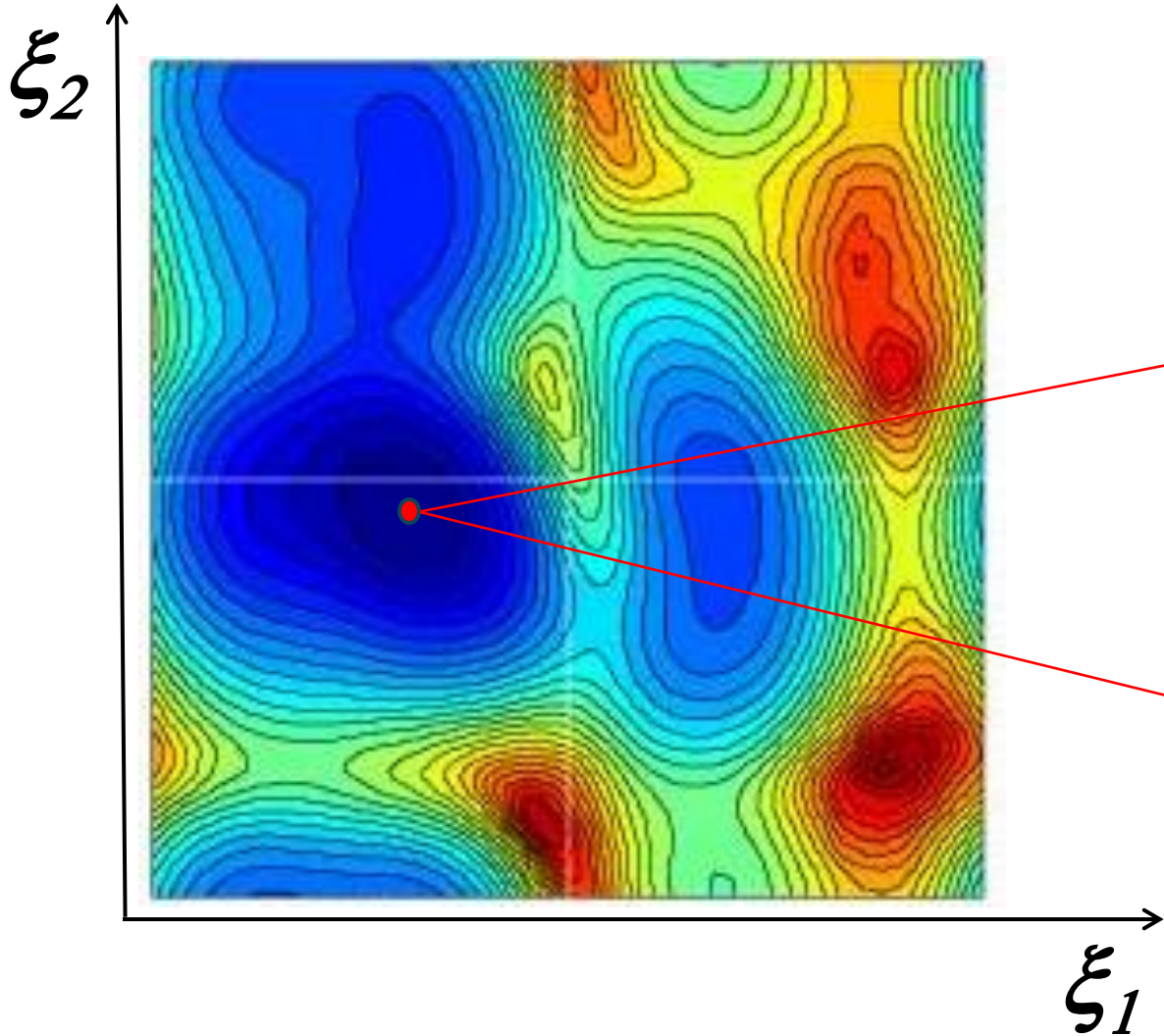
- **One very useful finding:** statistically averaged quantities over quantum mechanical vibrations are the same as predicted by classical mechanics!

Harmonic approximation for the entropy

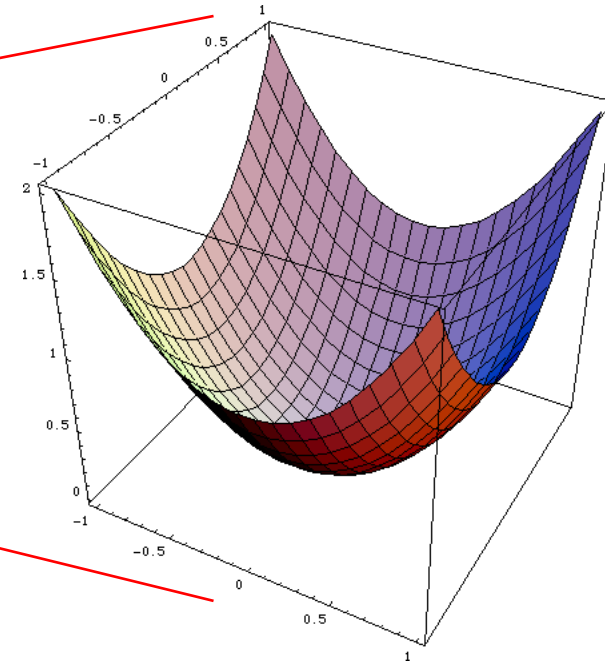
$$S = -\frac{\partial A}{\partial T} = k \ln Q + \frac{kT}{Q} \frac{\partial Q}{\partial T} \quad \frac{\partial Q}{\partial T} = \frac{k}{\hbar\omega}$$

$$S = k \ln Q + k = k \ln \frac{kTe}{\hbar\omega}$$

The Harmonic approximation (3)



In practice, you assume that all microscopic states mapped to one point of your free-energy landscape are summarised by a parabola of states on your PES



The Harmonic approximation (4)

... *Does it make sense?*



- Small, gas-phase systems
- Low-temperature systems
- Rigid moieties
- One order parameter
- The reaction coordinate is associated to *local* order parameters



- Large, condensed-phase systems
- High temperature
- Flexible moieties involved in the reaction (one torsional angle is sufficient...)
- Many order parameters
- The reaction coordinate is associated to *global* order parameters

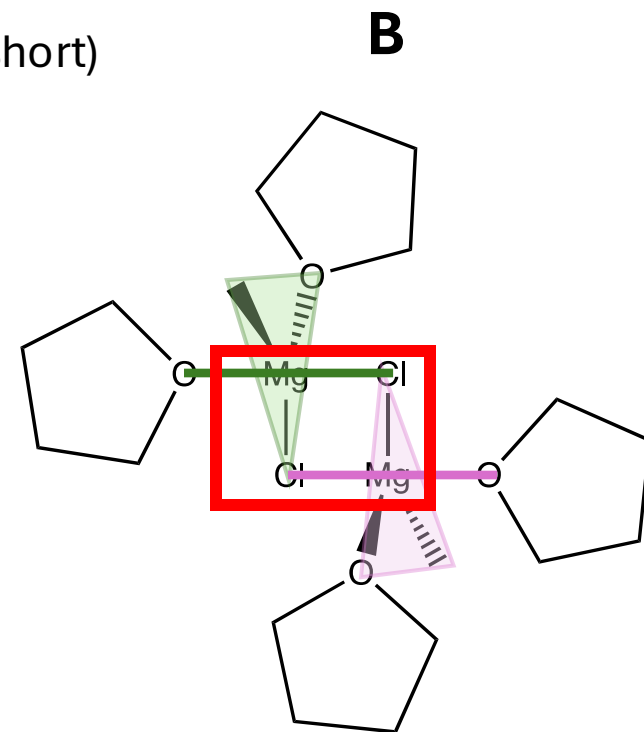
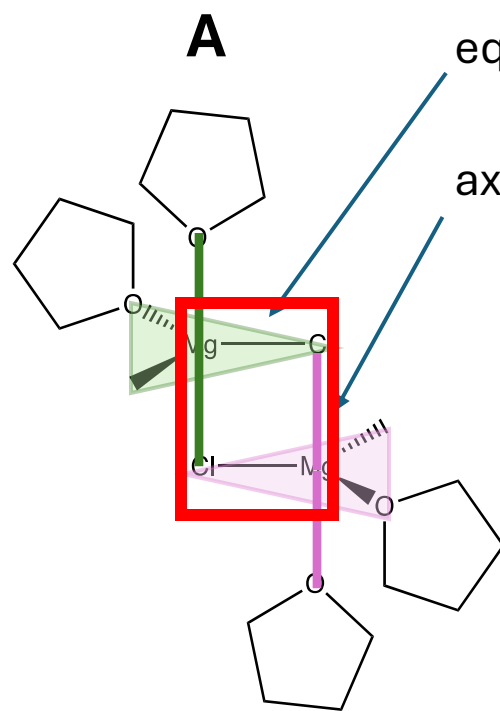
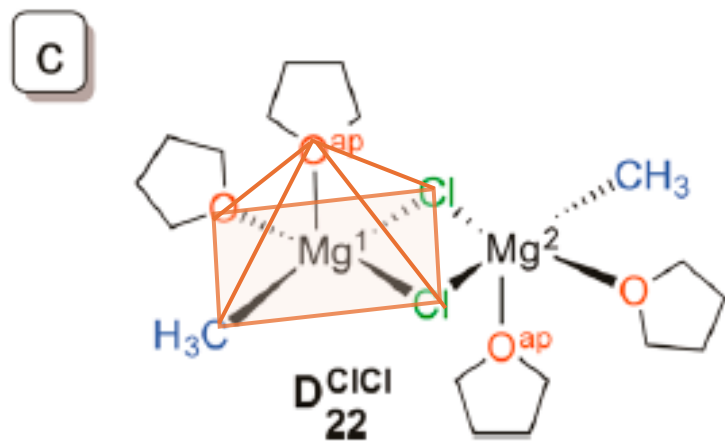
And what about entropy??

$$\Delta G = \Delta H - T\Delta S$$

A minimum (or a maximum, or a saddle point...) of the free energy is NOT a minimum of the total energy!!!!

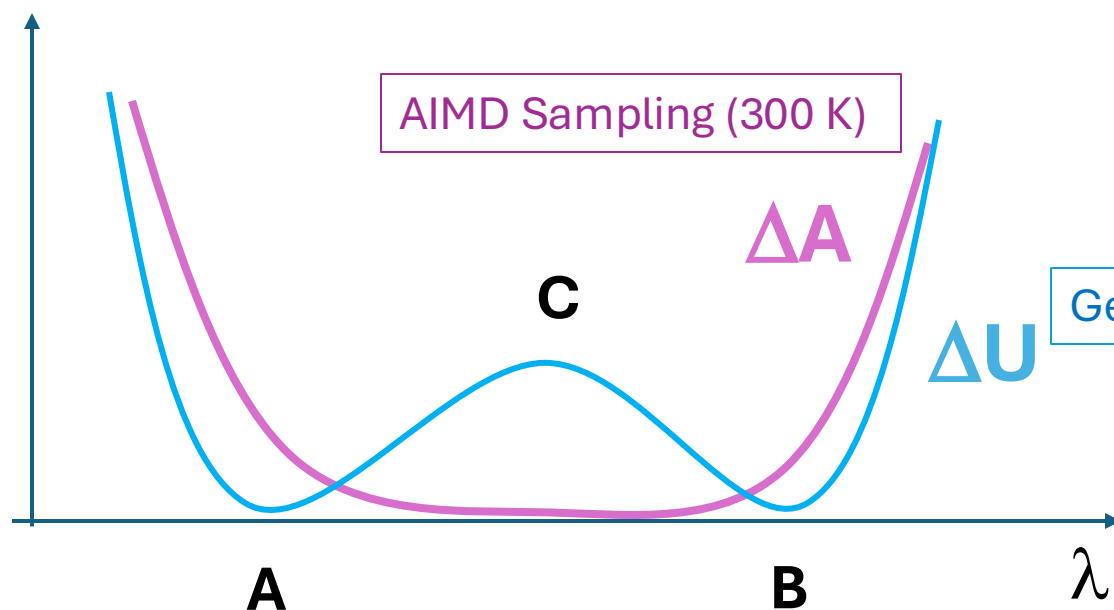
You may deal with systems where the free energy features
do not correspond
to those of the potential energy landscape

solvation structure of the CH_3MgCl dimer



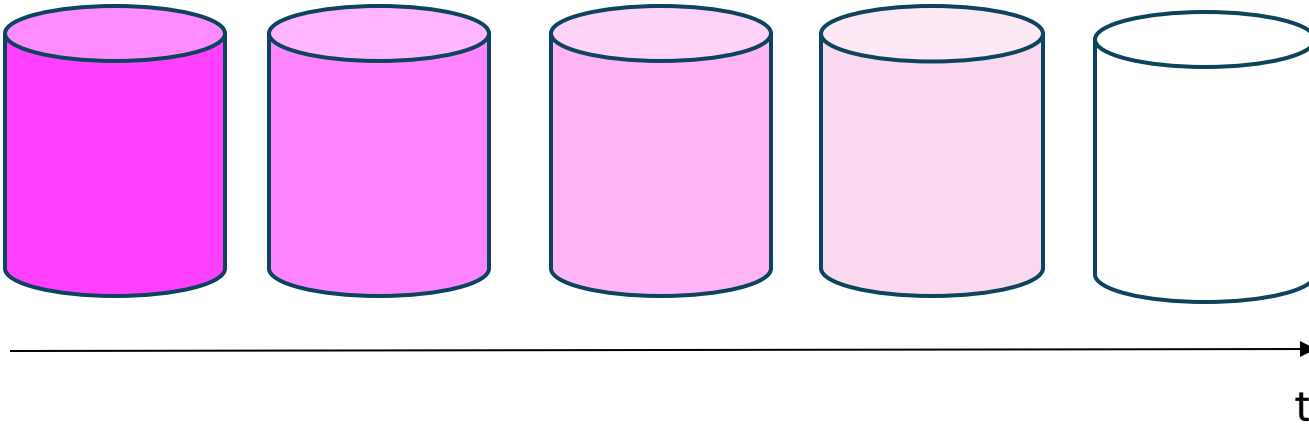
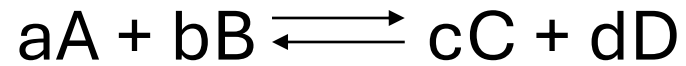
AIMD Sampling (300 K)

Geometry optimized



Sampling at operando conditions (300K, expl. solvent):
0 K TS $\rightarrow$ entropy favoured, becomes the GS!!

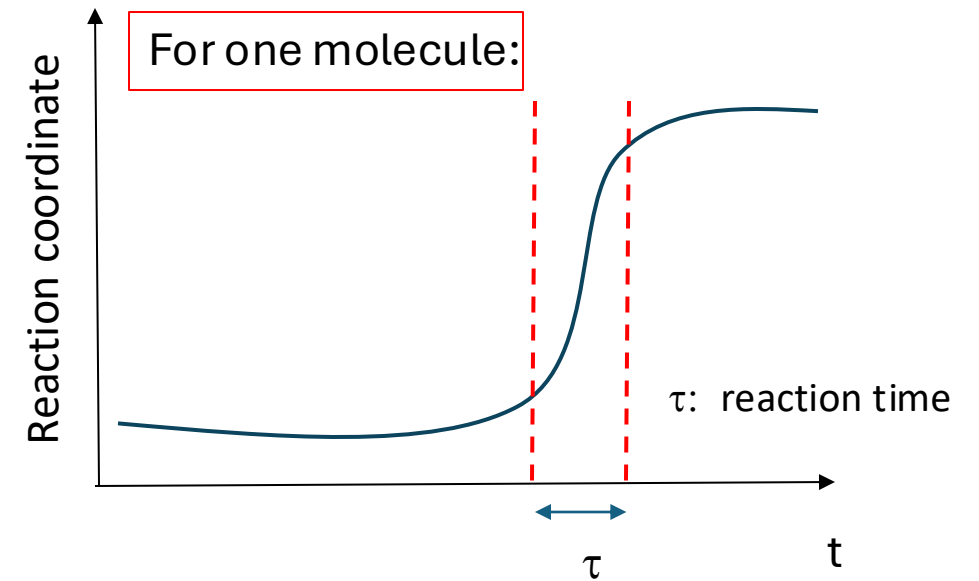
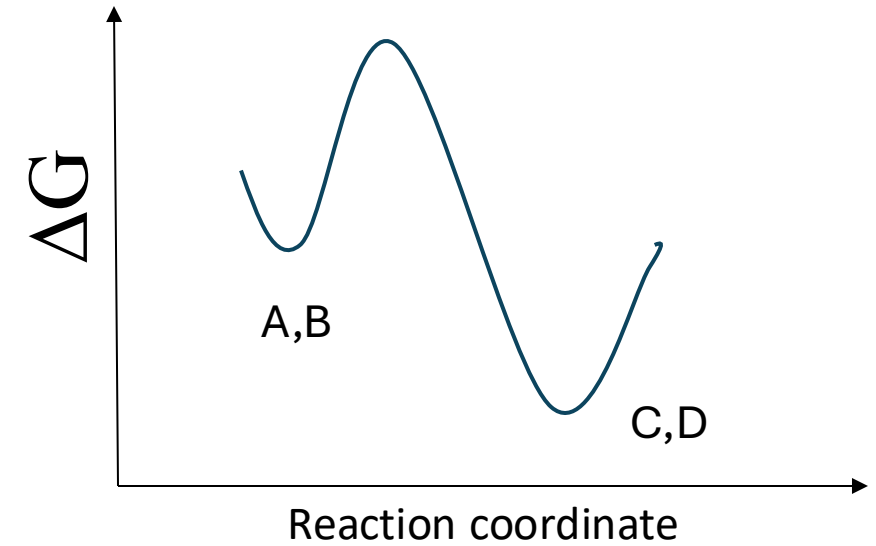
Activated processes



Macroscopically: slow, continuous process

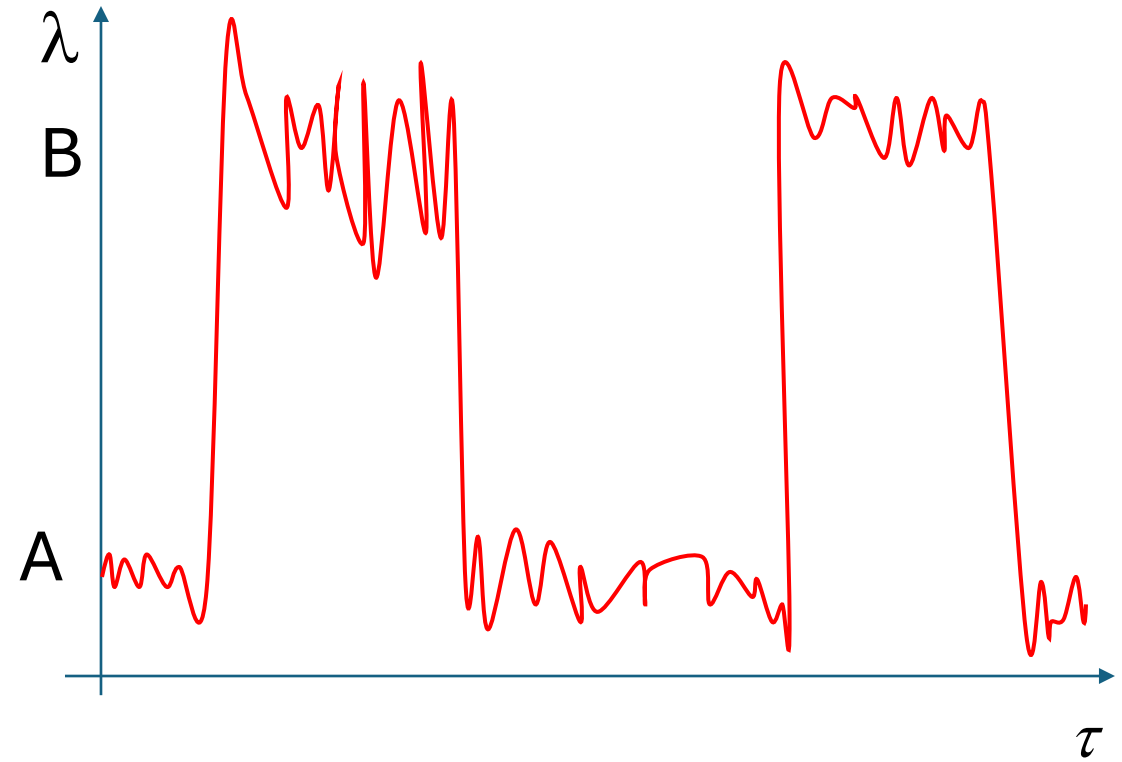
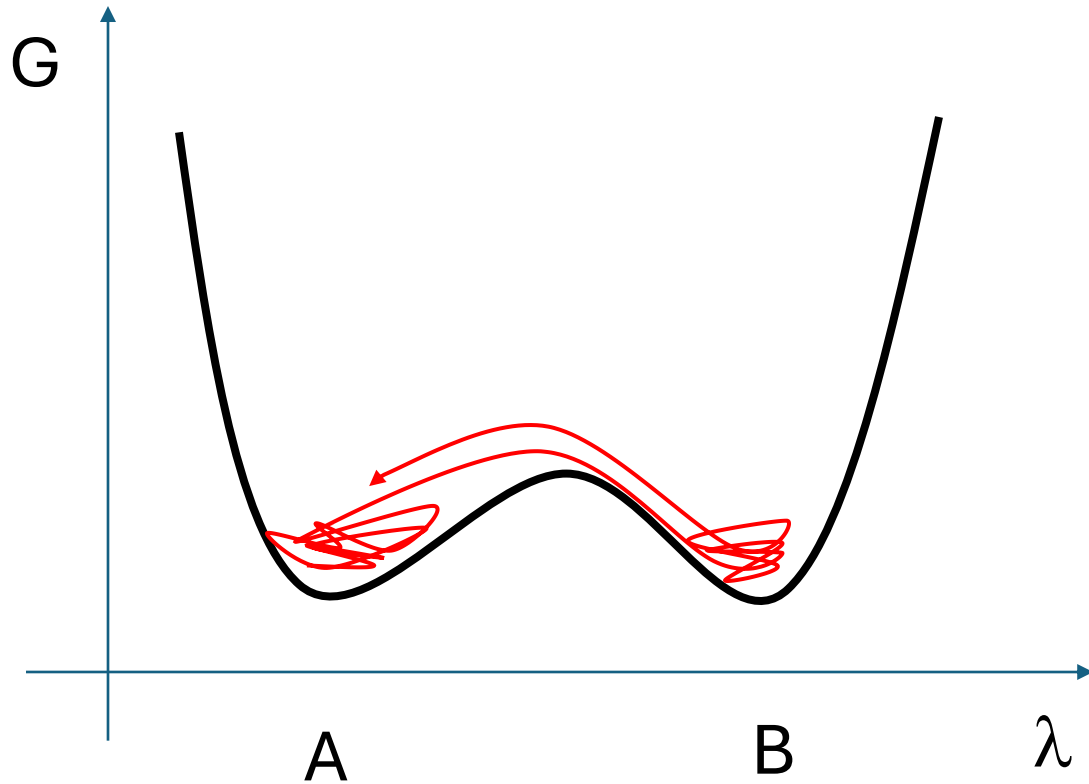
Microscopically: fast, but unlikely event

→ the slow steady macroscopic process is the statistical outcome of several unlikely microscopic events



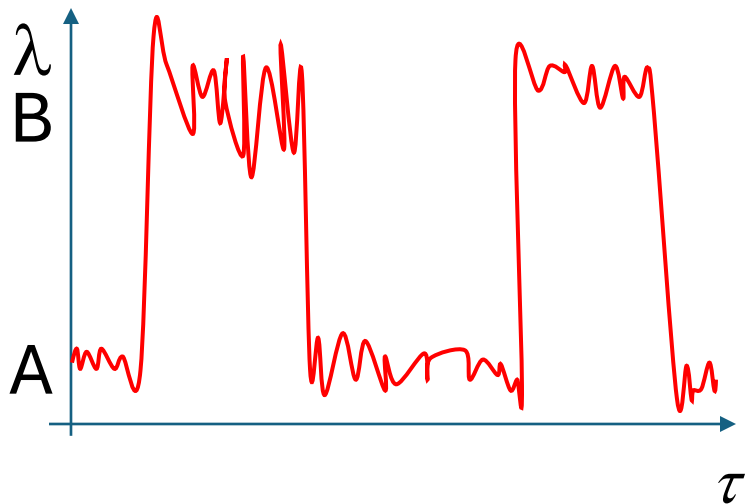
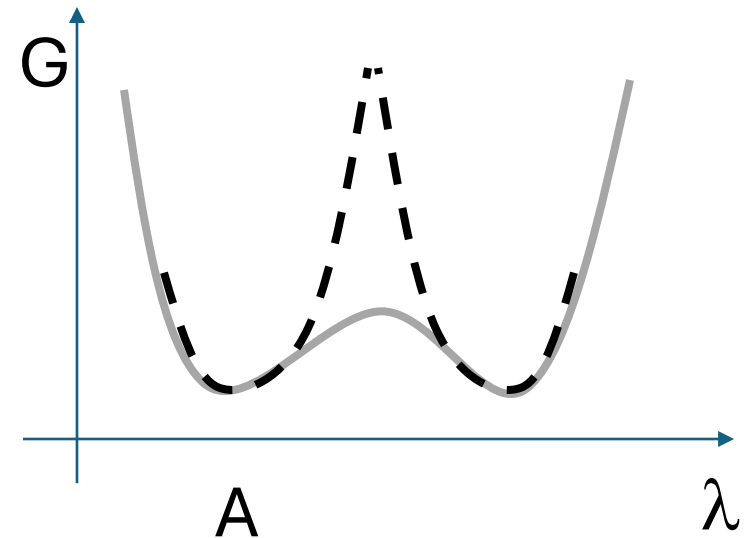
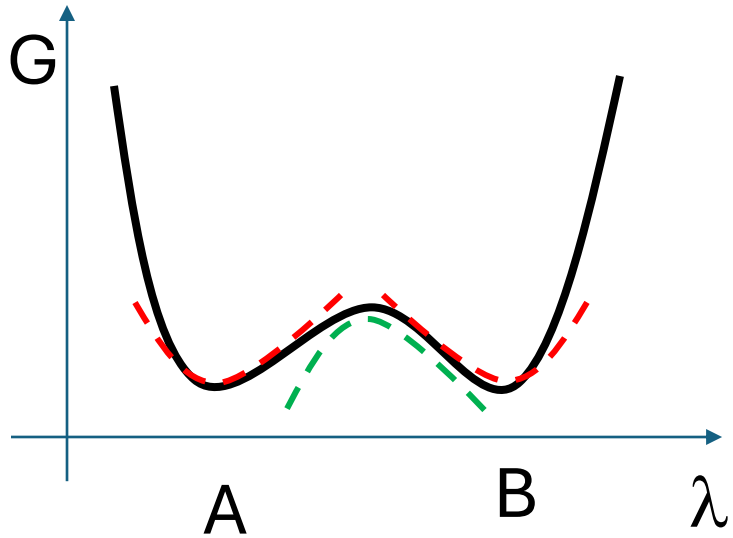
Sampling is not enough!!!

$$\Delta G = -kT \ln \mathcal{P}(\xi_1, \xi_2) \quad \textbf{Boltzmann Inversion}$$



In the best case (activation energies in the order of $k_B T$), direct sampling produces oscillatory probability distributions

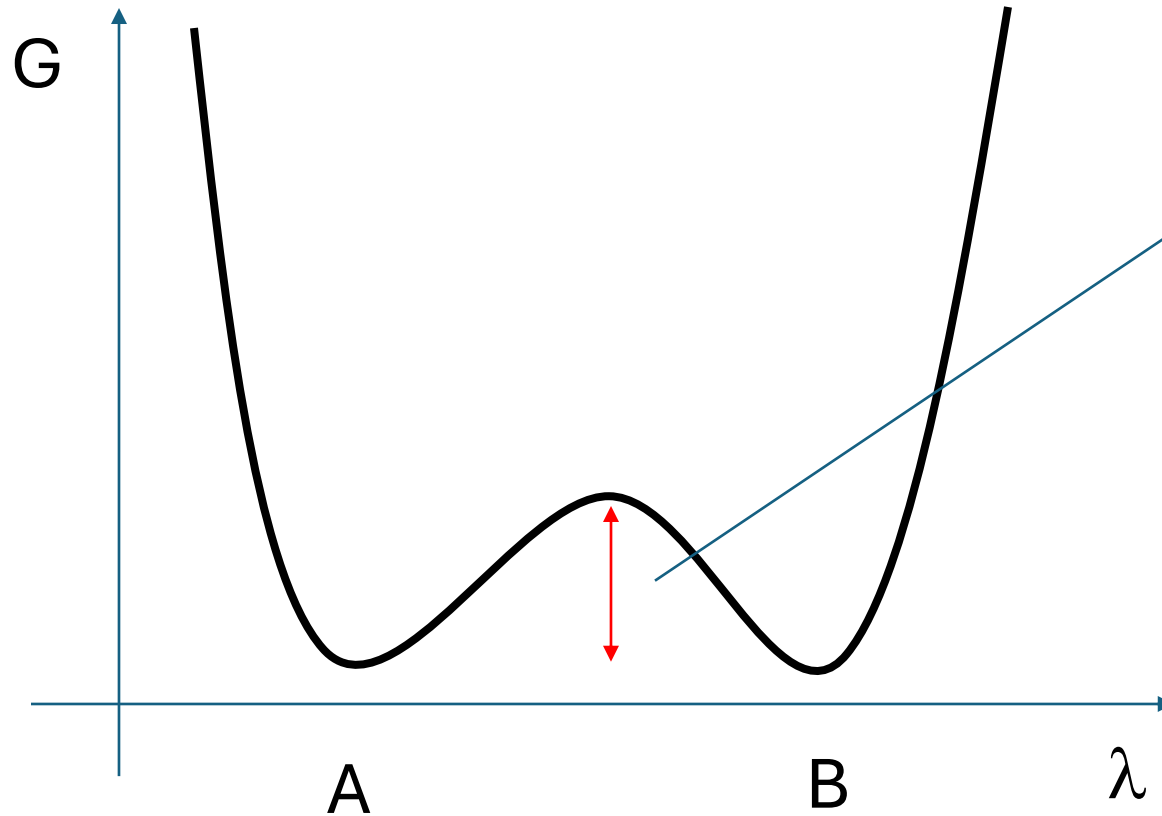
Sampling is not enough!!!



$$\Delta G = -kT \ln \mathcal{P}(\xi_1, \xi_2)$$

Poor convergence in Transition State regions

Violations of ergodicity



$$\Delta G^\ddagger \gg k_B T$$

$$\mathcal{P}(\xi_1, \xi_2) = e^{-\beta \Delta G(\xi_1, \xi_2)}$$

Probability to cross the barrier $\rightarrow 0 !!$

The system remains confined in the starting configuration

Enhanced sampling methods

Route to the computation of thermodynamically equilibrated properties

- Perturbation of the original system by an external potential
- Determination of the free energy by response to the perturbation

$H = H_0 + W(\xi_1, \xi_2)$ $\longrightarrow$ Collect some quantity of interest (depends on the method)

$\longrightarrow$ Derive the free energy from that



The bias is built to force the system to explore and sample appropriately the chemical space of interest

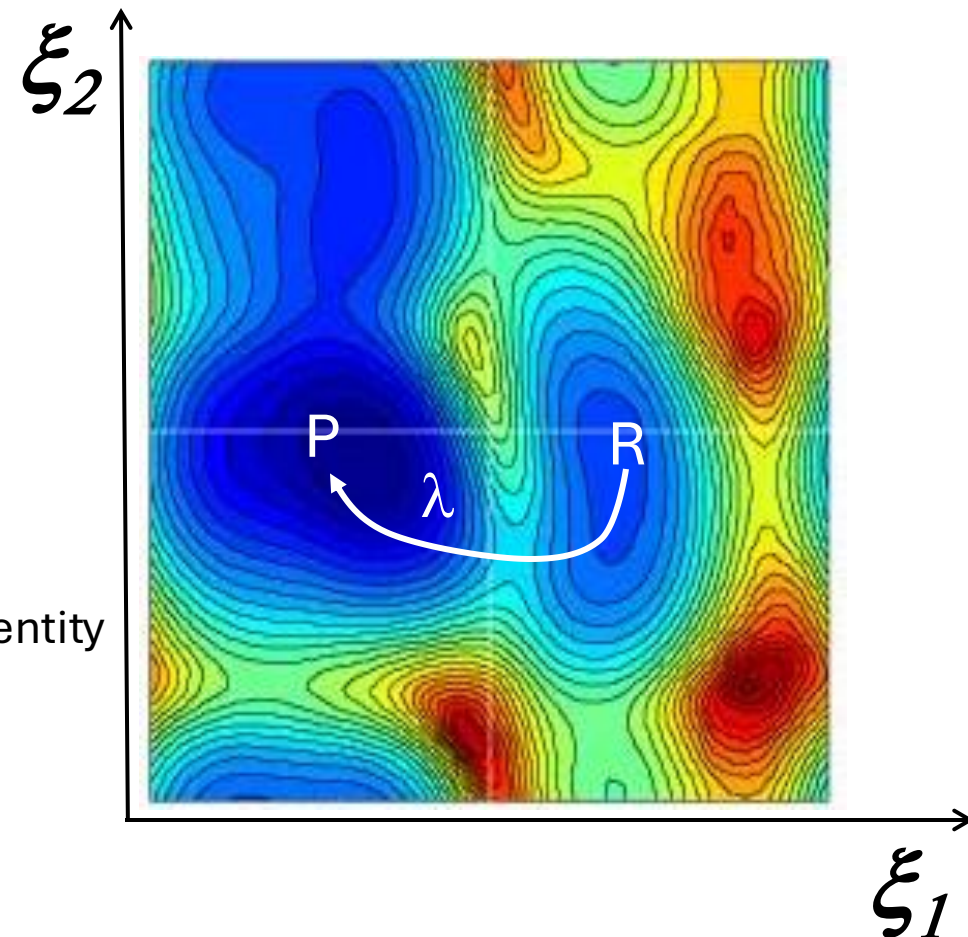
- Sampling is limited by the diffusional properties of the variables of interest and the dimensionality of the problem.
- All methods are affected by the necessary trade-off between sampling accuracy and sampling effectiveness.

Enhanced sampling methods

Route to the computation of thermodynamically equilibrated properties

Different methods for different situations

- **R, P** are known, interested only in $\Delta G = G(P) - G(R)$.
→ Thermodynamic cycles, free energy perturbation
- **R, P, λ** are known, interested in the complete $G(\lambda)$ profile.
→ Thermodynamic integration, umbrella sampling, EVB, Jarzynski identity
- ξ_s are known (eventually **R, P**), λ unknown (or multiple mechanisms).
→ metadynamics, elastic path sampling, conformational flooding
- ξ_s are unknown (or too many).
→ Bias-exchange metadynamics, Replica-exchange (parallel tempering), Markov chain models



Example: Umbrella sampling

$$W(\lambda) = \text{restraint}$$

$$H = H_0 + W(\lambda(\{\mathbf{r}\}, \{\mathbf{p}\})) = H_0 + \frac{1}{2}k(\lambda - \lambda_0)^2$$

$$\mathcal{P}_0(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta H_0}}{Q_0}$$

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) = \frac{e^{-\beta H}}{Q}$$

↑
"soft" bias on the reaction coordinate
(Typically harmonic potential)

← The bias changes the probability of a microscopic state!!

$$\mathcal{P}(\lambda) = \iint \mathcal{P}(\mathbf{r}, \mathbf{p}) \delta(\lambda(\mathbf{r}, \mathbf{p}) - \lambda) d\Gamma = \frac{1}{Q} \iint e^{-\beta W(\lambda)} e^{-\beta H_0(\mathbf{r}, \mathbf{p})} \delta(\lambda(\mathbf{r}, \mathbf{p}) - \lambda) d\Gamma = \frac{Q_0}{Q} e^{-\beta W(\lambda)} \mathcal{P}_0(\lambda)$$

Example: Umbrella sampling

$$\mathcal{P}(\lambda) = \iint \mathcal{P}(\mathbf{r}, \mathbf{p}) \delta(\lambda(\mathbf{r}, \mathbf{p}) - \lambda) d\Gamma = \frac{1}{Q} \iint e^{-\beta W(\lambda)} e^{-\beta H_0(\mathbf{r}, \mathbf{p})} \delta(\lambda(\mathbf{r}, \mathbf{p}) - \lambda) d\Gamma = \frac{Q_0}{Q} e^{-\beta W(\lambda)} \mathcal{P}_0(\lambda)$$

$$\mathcal{P}_0(\lambda) = c \mathcal{P}(\lambda) e^{\beta W(\lambda)}$$

$$G(\lambda) = -kT \ln \mathcal{P}_0(\lambda) = -kT \ln \mathcal{P}(\lambda) - W(\lambda) + \text{constant}$$

Good method for transformations vs. “soft” order parameters

- Non-covalent binding
- Diffusion / a(d)sorption / percolation

$$W(\lambda) = \frac{1}{2} k (\lambda - \lambda_0)^2$$

$$\mathcal{P}_W(\lambda) = e^{-\beta W(\lambda)} = e^{-\frac{\beta k}{2} (\lambda - \lambda_0)^2} \quad \longleftarrow \text{Gaussian!!}$$

$$G(\lambda) = -kT \ln \mathcal{P}_0(\lambda) = -kT \ln \frac{\mathcal{P}(\lambda)}{\mathcal{P}_W(\lambda)}$$

*Free energy obtained by
WHAM
against a Normal distribution*

In practice:

- Add bias on reaction coordinate:

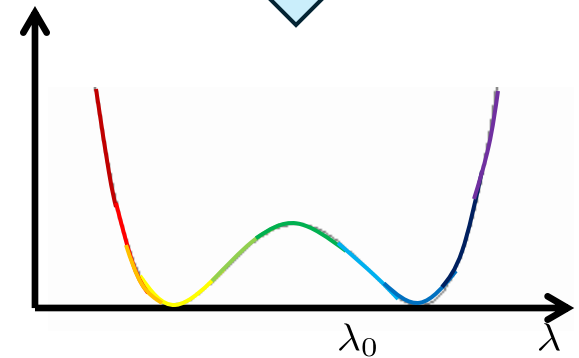
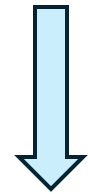
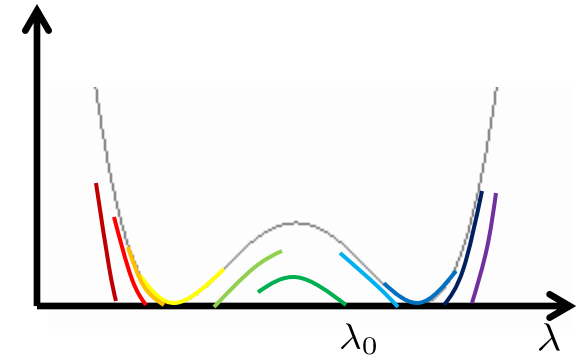
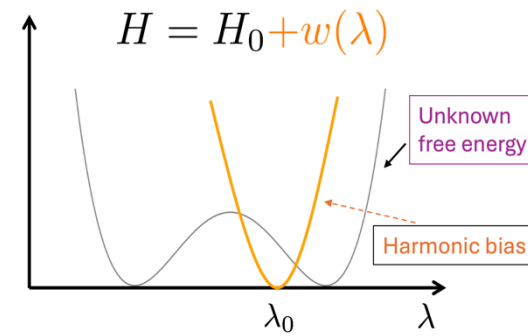
$$W(\lambda) = \frac{1}{2}k(\lambda - \lambda_0)^2$$

- Perform MD, collect $\mathcal{P}(\lambda)$:

- Compare $\mathcal{P}(\lambda)$: to $\mathcal{P}_W(\lambda)$

- Deduce $G(\lambda)$ from discrepancy

- Move λ_0 , repeat



Example: Thermodynamic integration

$W(\lambda)$ = holonomic constraint

$$H(\lambda_0) = H_0 + \varphi(\lambda(\{\mathbf{r}\}, \{\mathbf{p}\}) - \lambda_0)$$

↑
Constraint force on the reaction coordinate

$$\frac{\partial}{\partial \lambda} A(\lambda_0) = -kT \frac{1}{Q(\lambda_0)} \iint -\beta \varphi e^{-\beta H} d\Gamma_{\lambda_0} = \langle \varphi \rangle_{\lambda_0}$$

← Expectation value of the constraint at a fixed value of λ

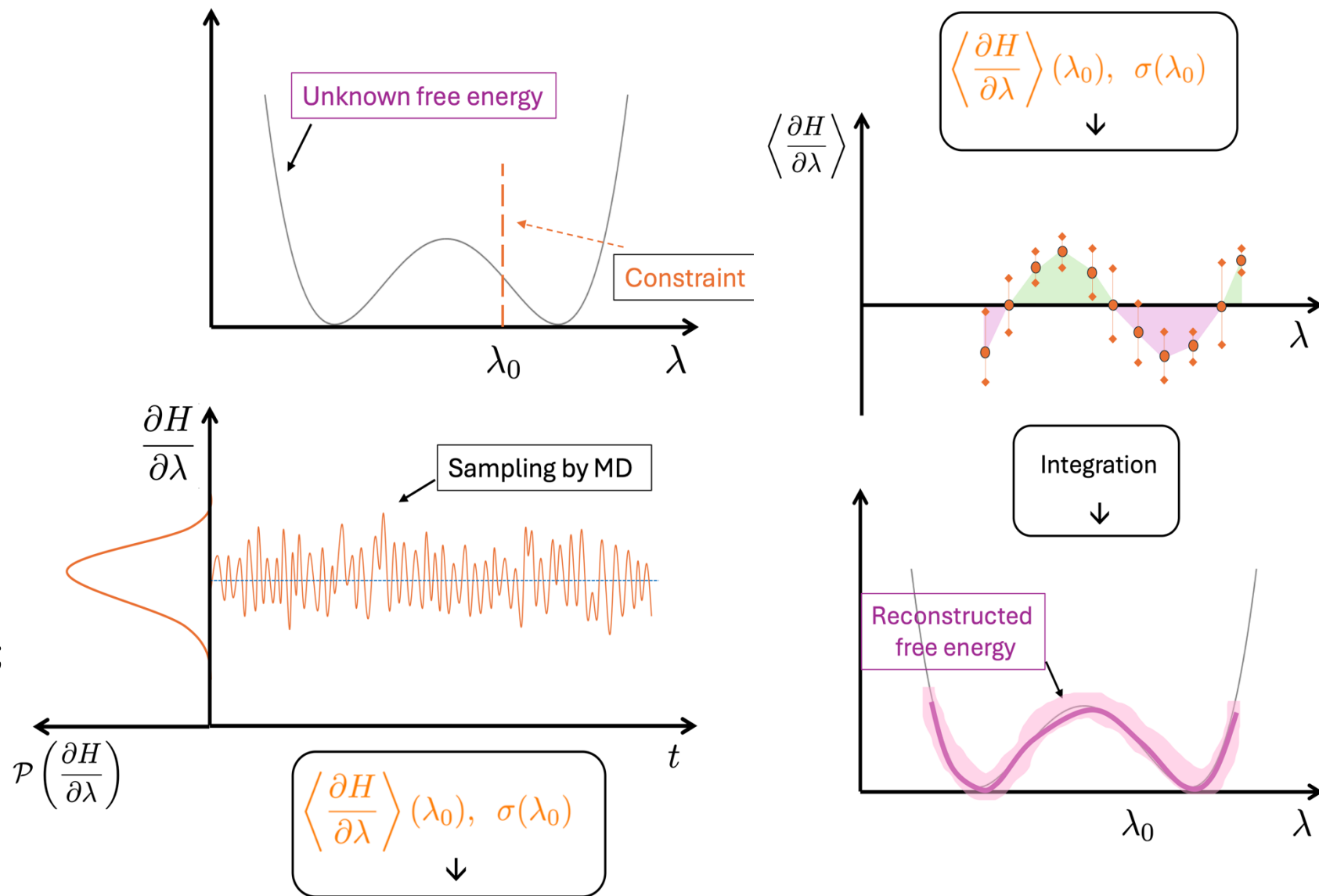
$$\Delta G = \int_{\lambda(R)}^{\lambda(P)} \langle \varphi \rangle_{\lambda} d\lambda$$

Good method for transformations against “hard” order parameters

- Bond breaking / forming
- Alchemical transformations

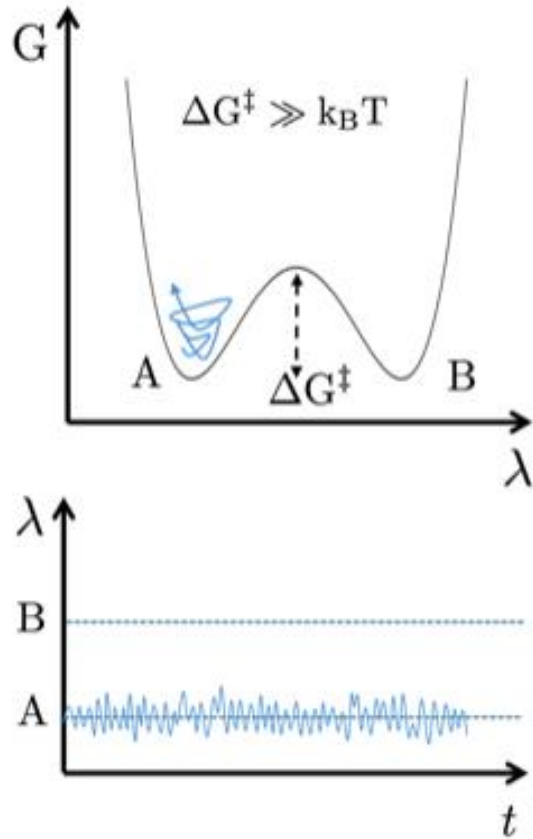
In practice:

- Perform constraint MD;
- Collect bias force;
- Move λ_0 , repeat;
- Integrate average bias force;



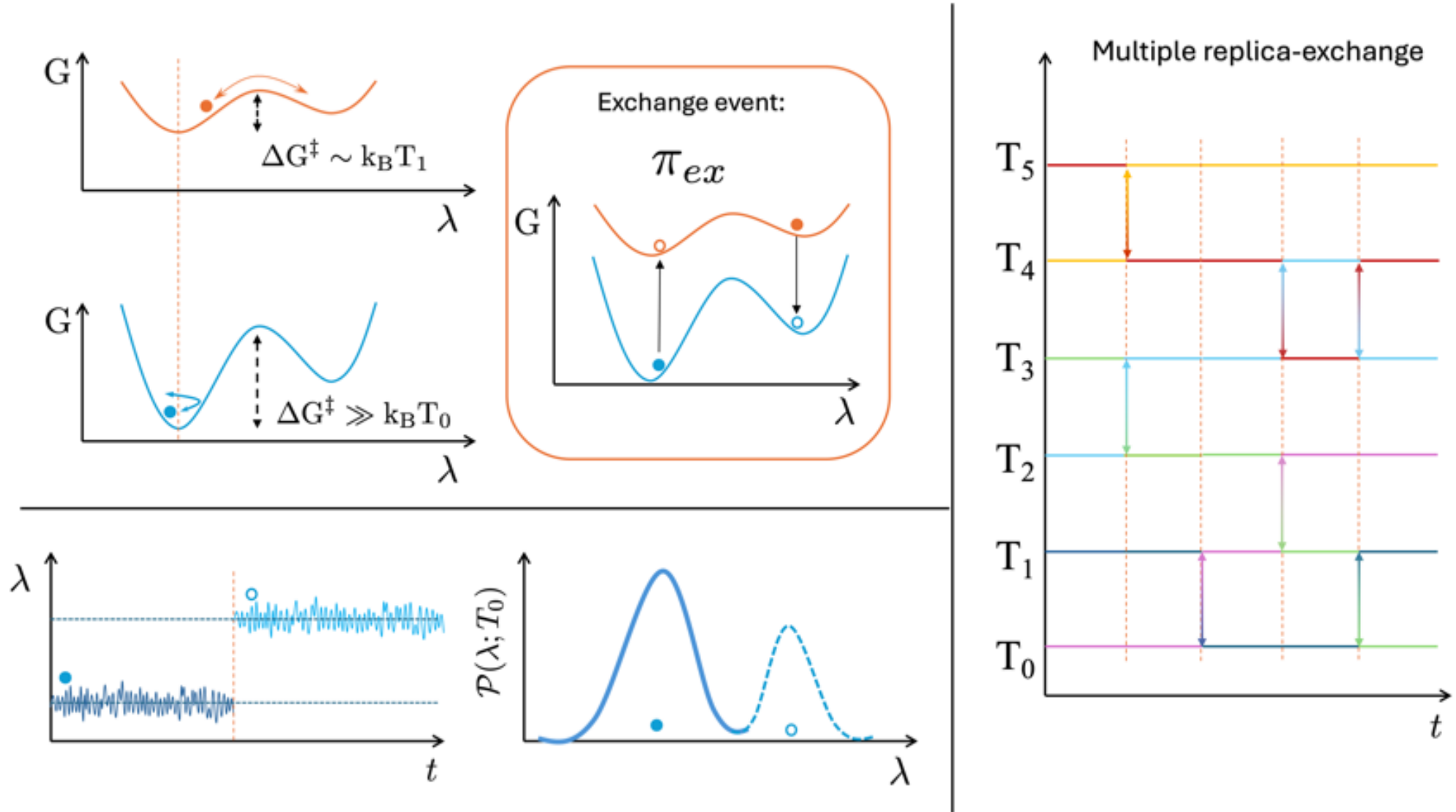
Multiple variable methods: Metadynamics

Conventional MD



<https://www.youtube.com/watch?feature=oembed&v=c4bybwrw6Xg>

Many order parameters: replica exchange



<https://www.youtube.com/watch?feature=oembed&v=3URG5M0BXUY>

Issues with enhanced sampling methods

- Definition of the order parameters / reaction coordinate
- Choice of the method
- Numerical convergence criteria
- Computational costs

Possible advices:

- Learn a few, get familiar with them
- Build personal know-how, be “conservative” with it
- Follow the literature for updates, improvements, innovative approaches
- Wait before jumping on the newest trendy method of the season! Consider learning time, and if it is worth it