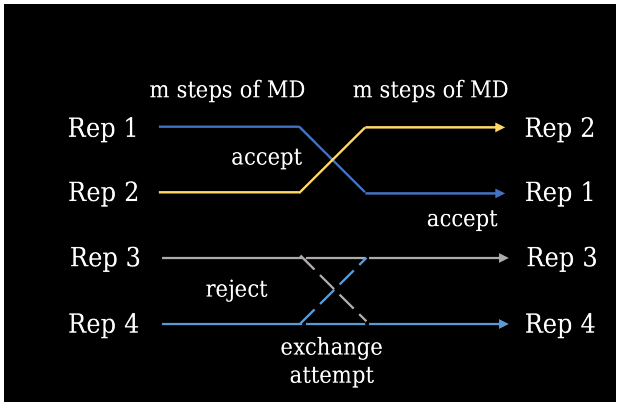
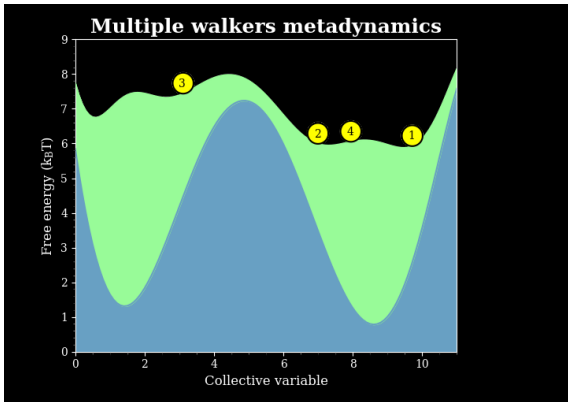
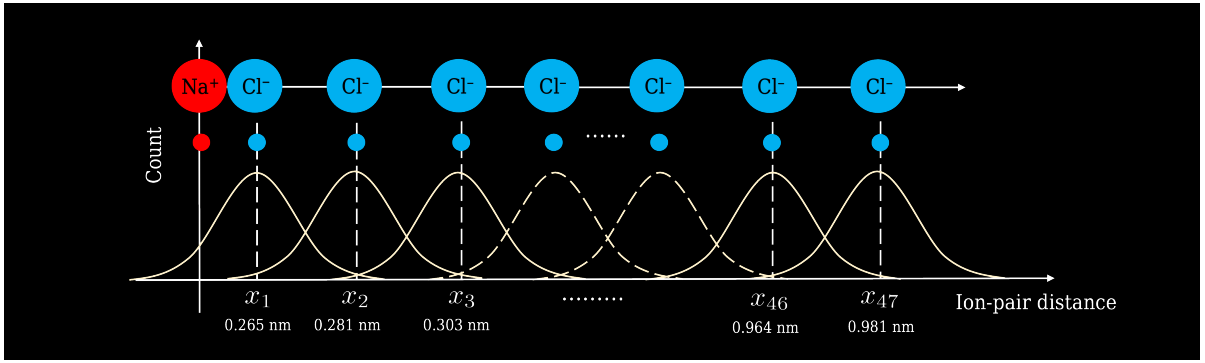
Enhanced sampling molecular dynamics on HPC systems: Ensembles of Simulations

Prof. Michael Shirts
Dr. Wei-Tse Hsu
CU Boulder



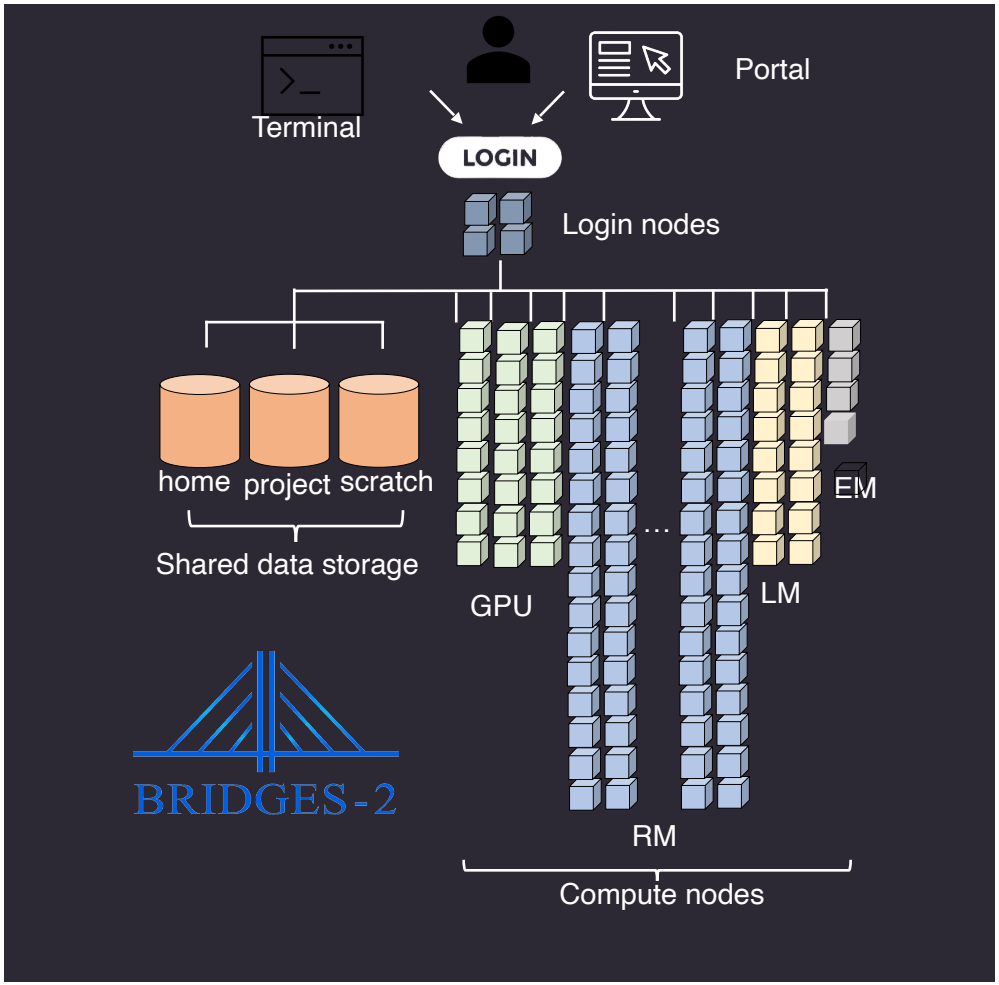




i-CoSE

High-performance computing (HPC) systems are powerful for complex and resource-intensive tasks

- Interactive session: `interact -N 1 --ntasks-per-node=128 -p RM -p time 00:30:00`
- SLURM job submission script

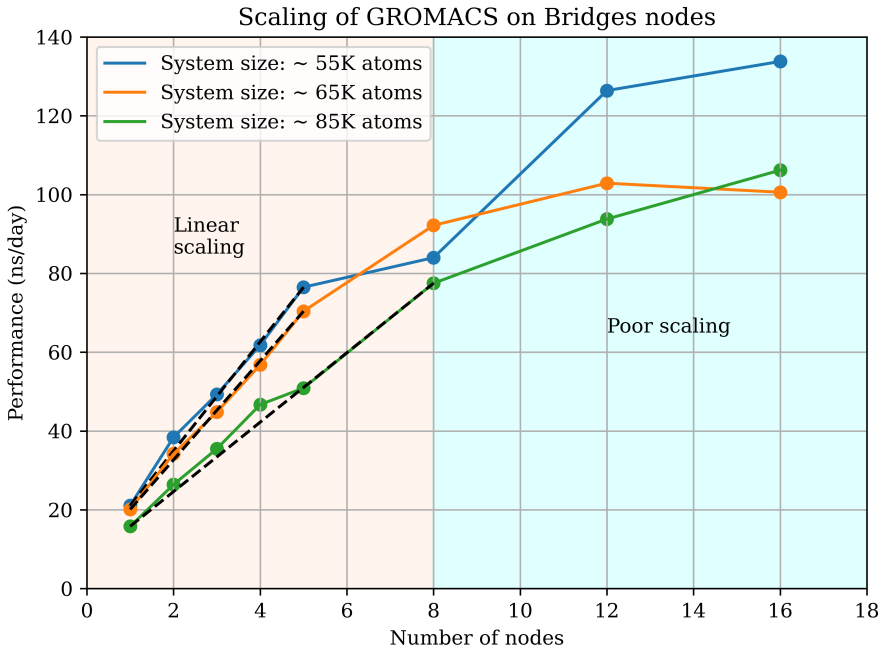



```
#!/bin/sh
#SBATCH --job-name MD_simulation
#SBATCH -N 1
#SBATCH -p RM
#SBATCH -t 2:00:00
#SBATCH --ntasks-per-node=64

module load gromacs
module load openmpi/3.1.6-gcc10.2.0

mpirun -np 16 gmx_mpi mdrun -deffnm md
```

It's important to make the best use of compute resources when running parallel simulations

- Parallelization can be performed across
  - Multiple cores in a single node using multiple threads (e.g. OpenMP thread)
  - Multiple nodes using multiple MPI processes by OpenMPI (launched by mpirun or mpiexec)

Communications take time!

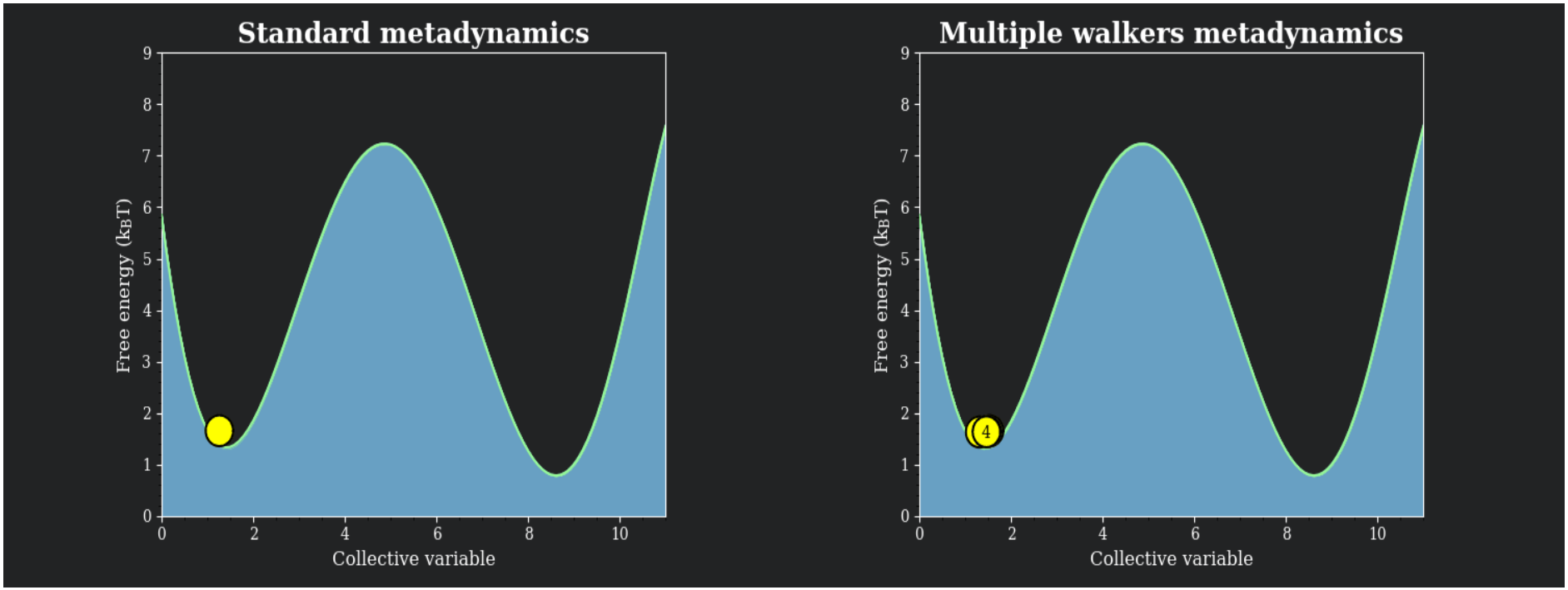


Hands-on exercise 1

Running uncoupled MD simulations in parallel

Metadynamics converge can be sped up by running multiple walkers in the same energy surface

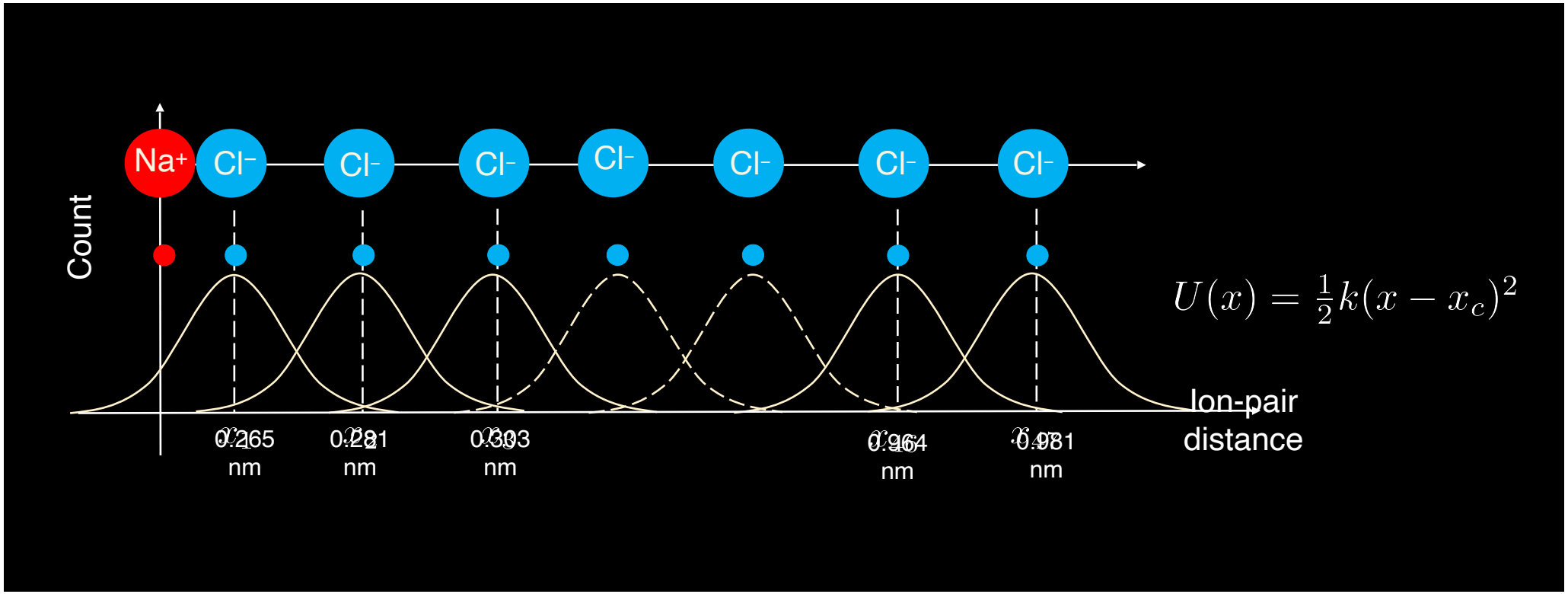
- In multiple walkers metadynamics, the deposited bias is shared between replicas so each replica is aware of the whole history of the bias.



👨‍💻 **Hands-on exercise 2**

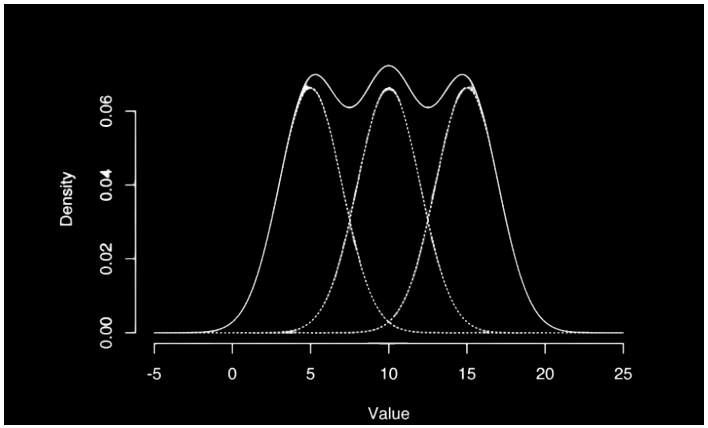
Umbrella sampling runs simulations at different values of some CV

- Pulling simulation: Fix Na^+ and pull Cl^- away to generate a range of initial locations
- Production runs: Fix the whole ion-pair and perform the simulations under harmonic restraint.
- A reasonable force constant should ensure overlap between states.



We can use MBAR to recover the free energy profile along the CV of interest from umbrella sampling

- Possible methods to reconstruct the free energy: WHAM, TI, BAR, MBAR and even more!
- Protocol of using MBAR
  - Decorrelate the time series of the collective variables.
  - Reweight from the mixture distribution of all the biases
    - Compute the free energy of each biased simulation
    - Approximate the free energy surface

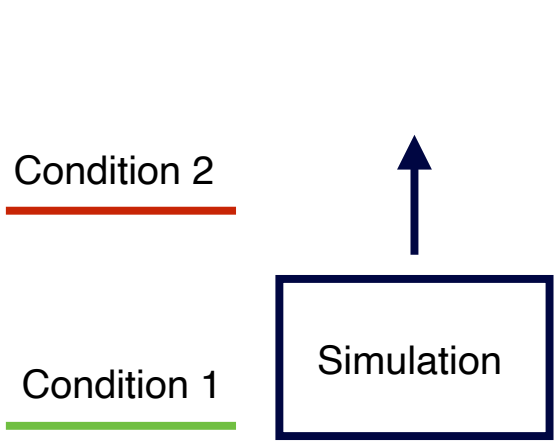


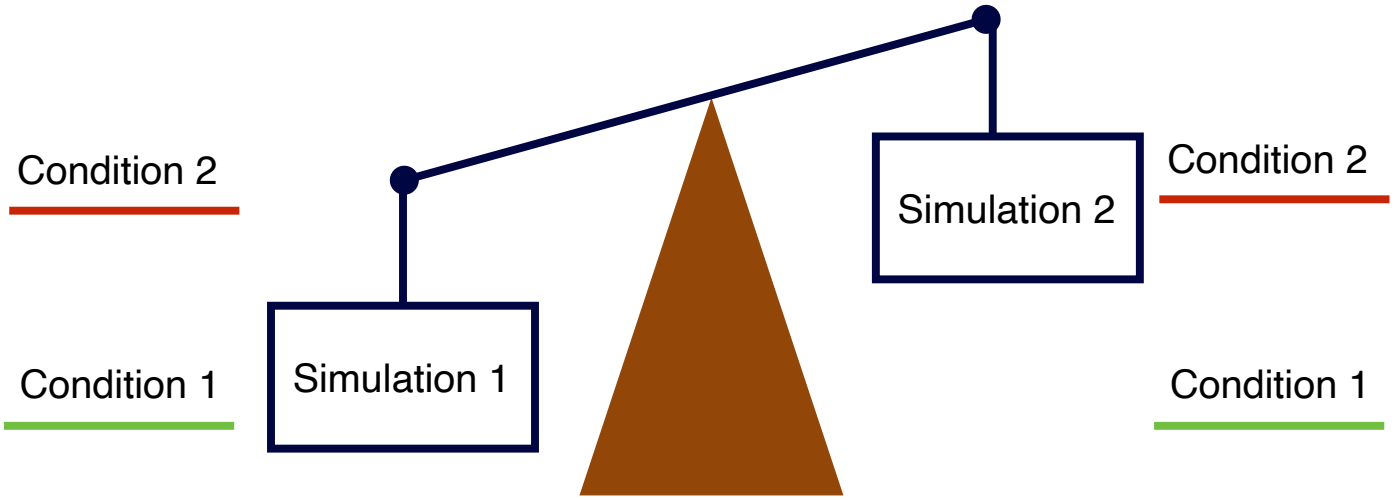
👨‍💻 **Hands-on exercise 3**

Perform and analyze umbrella sampling

Concept of Replica Exchange

- We have an ensemble of simulations, each a different condition (temperature, bias potential
- Rather than changing conditions independently, we allow them to change in a coupled way.





$$\frac{p_1}{p_2} = \frac{p_{A=1}(x_A)p_{B=2}(x_B)}{p_{A=2}(x_A)P_{p=1}(x_B)}$$

- If we were to switch the configurations, what is the total change in probability?

Considerations of Replica Exchange

- Advantages: Don't have to consider the free energies

$$\frac{p_1}{p_2} = \frac{p_{A=1}(x_A)p_{B=2}(x_B)}{p_{A=2}(x_A)P_{p=1}(x_B)}$$

$$= \frac{e^{\cancel{\beta_1 F_1} - \beta_1 U_1(x_A)} e^{\cancel{\beta_2 F_2} - \beta_2 U_2(x_B)}}{e^{\cancel{\beta_1 F_1} - \beta_1 U_1(x_B)} e^{\cancel{\beta_2 F_2} - \beta_2 U_2(x_A)}}$$

$$= \frac{e^{-\beta_1 U_1(x_A)} e^{-\beta_2 U_2(x_B)}}{e^{-\beta_1 U_1(x_B)} e^{-\beta_2 U_2(x_A)}}$$

- Disadvantage:
  - Simulations need to run in lockstep
  - They need to exchange frequently to be useful
  - It gets complicated to exchange when there there are many, including multiple dimensions

Temperature replica exchange periodically exchange simulation coordinates at different temperatures

- The sampling alternates between the configurational and temperature space.
- At higher temperature, it is easier to cross the free energy barrier.
- Useful for estimating free energy or heat capacity as a function of temperature

Transition probability:

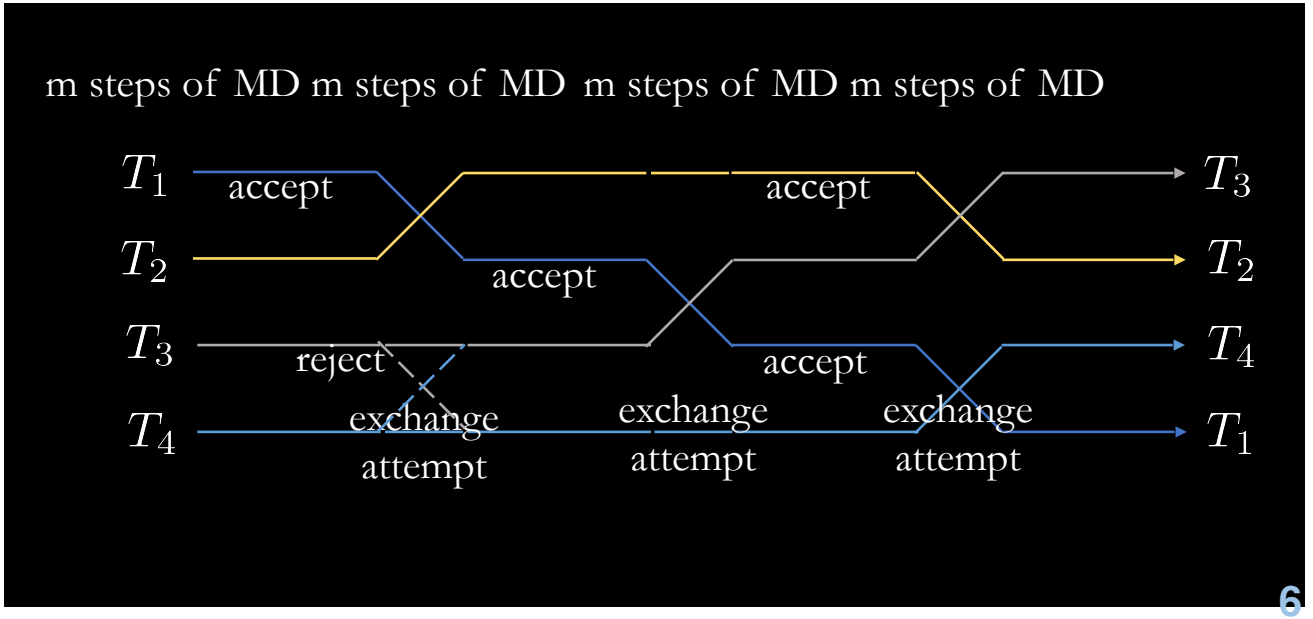
If $\Delta > 1$: Move

If $\Delta < 1$: move with prob Δ

$$\Delta = \frac{p_{T_1}(x_A)p_{T_2}(x_B)}{p_{T_1}(x_B)p_{T_2}(x_A)}$$

$$= \frac{e^{-\beta_1 U(x_A)} e^{-\beta_2 U(x_B)}}{e^{-\beta_2 U(x_A)} e^{-\beta_1 U(x_B)}}$$

$$= e^{(\beta_2-\beta_1)U(x_A)-(\beta_1-\beta_2)U(x_B)}$$



Hamiltonian replica exchange

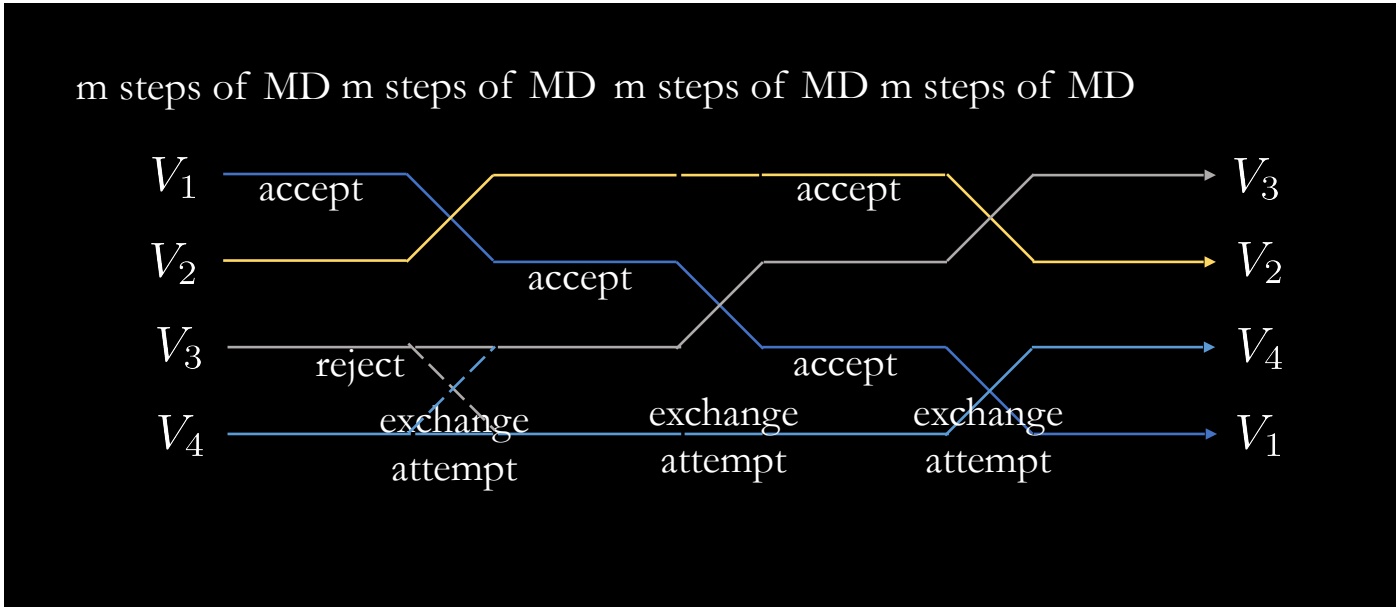
- In Hamiltonian replica exchange, our different conditions are different potentials
- In our case, our different Hamiltonians are different localization biases

If $\Delta > 1$: Move
If $\Delta < 1$: move with prob Δ

$$\Delta = \frac{p_{T_1}(x_A)p_{T_2}(x_B)}{p_{T_1}(x_B)p_{T_2}(x_A)}$$

$$= \frac{e^{-\beta U_1(x_A)}e^{-\beta U_2(x_B)}}{e^{-\beta U_2(x_A)}e^{-\beta U_1(x_B)}}$$

$$= e^{\beta[(U_1(x_B)-U_1(x_A))+(U_2(x_A)-U_2(x_B))]}$$



Hands-on exercise 4

Perform and analyze REUS

Alchemical intermediate states are useful for calculating solvation, binding or mutation ΔG

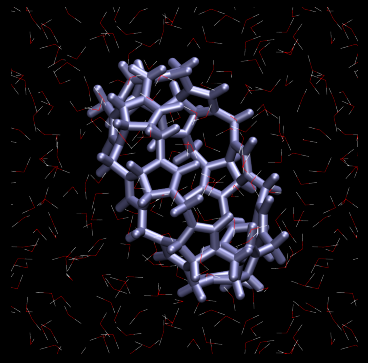
- An alchemical intermediate state is defined by a vector of coupling parameters.

```
; lambda-states     = 1     2     3     4     5     6     7     8     9
coul-lambdas        = 0.00  0.25  0.50  0.75  1.00  1.00  1.00  1.00  1.00
vdw-lambdas         = 0.00  0.00  0.00  0.00  0.00  0.25  0.50  0.75  1.00
```

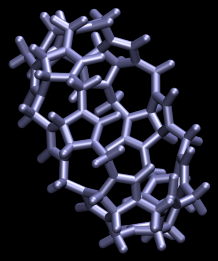
$\vec{\lambda}_i = (\lambda_{vdW,i}, \lambda_{coul,i})$ $\quad \lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4 \quad \lambda_5 \quad \lambda_6 \quad \lambda_7 \quad \lambda_8 \quad \lambda_9$

$\vec{\lambda_1} = (0.0, 0.0)$

Fully coupled (interactions on)



ΔG_{solv}



$\vec{\lambda_9} = (1.0, 1.0)$

Fully decoupled (interactions off)

Alchemical replica exchange generalizes the ensemble with the alchemical direction

- Hamiltonian replica exchange (HREX) alternates the sampling between the coordinate and space of Hamiltonans

Hands-on exercise 5