CE 530 Molecular Simulation

Lecture 18
Free-energy calculations

David A. Kofke
Department of Chemical Engineering
SUNY Buffalo
kofke@eng.buffalo.edu

Free-Energy Calculations

- Uses of free energy
  - *Phase equilibria*
  - *Reaction equilibria*
  - *Solvation*
  - *Stability*
  - *Kinetics*
- Calculation methods
  - *Free-energy perturbation*
  - *Thermodynamic integration*
  - *Parameter-hopping*
  - *Histogram interpolation*

Ensemble Averages

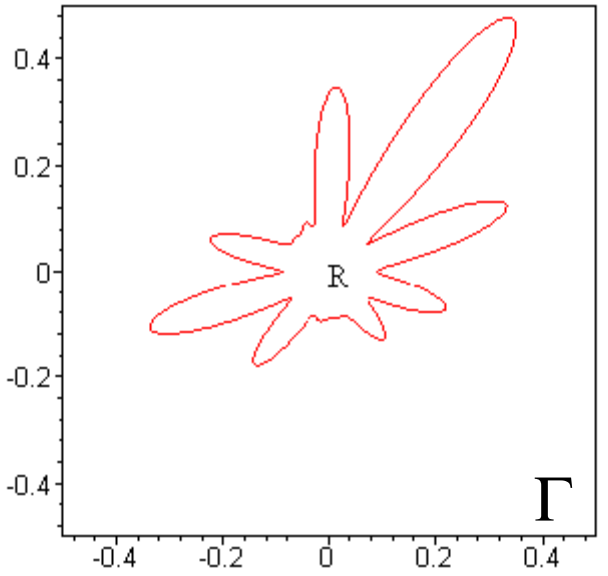
- Simple ensemble averages are of the form

$$\langle M \rangle = \int d\Gamma \pi(\Gamma) M(\Gamma)$$

- To evaluate:
  - *sample points in phase space with probability* $\pi(\Gamma)$
  - *at each point, evaluate* $M(\Gamma)$
  - *simple average of all values gives* $\langle M \rangle$
- Previous example
  - *mean square distance from origin in region R*

$$\langle r^2 \rangle = \int d\Gamma \frac{s(\Gamma)}{\int d\Gamma s(\Gamma)} r^2(\Gamma) \qquad s = \begin{cases} 1 & \text{inside R} \\ 0 & \text{outside R} \end{cases}$$

  - *sample only points in R, average* r^2
- Principle applies to both MD and MC

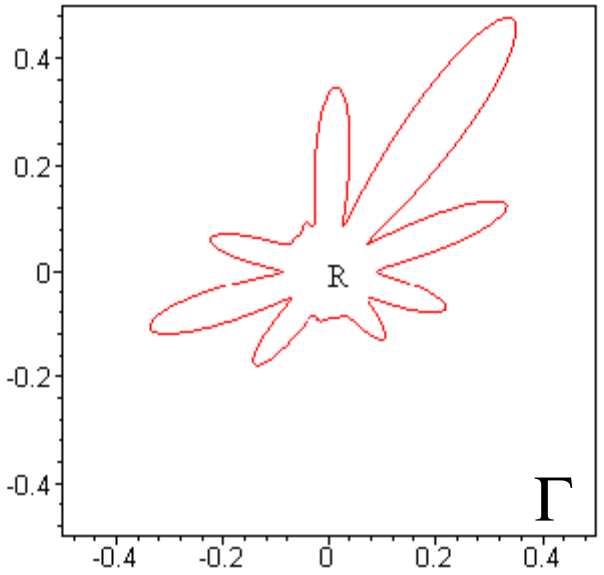


Ensemble Volumes

- Entropy and free energy relate to the size of the ensemble
  - *e.g.*, $S = k \ln\Omega(E,V,N)$ Ω = number of states of given E,V,N
- No effective way to measure the size of the ensemble
  - *no phase-space function that gives size of R while sampling only R*

    imagine being place repeatedly at random points on an island

    what could you measure at each point to determine the size of the island?



- Volume of ensemble is numerically unwieldy
  - *e.g.* for 100 hard spheres

    $r = 0.1$, $\Omega = 5 \times 10^{133}$

    $r = 0.5$, $\Omega = 3 \times 10^{7}$

    $r = 0.9$, $\Omega = 5 \times 10^{-142}$
- Shape of important region is very complex
  - *cannot apply methods that exploit some simple geometric picture*

Reference Systems

- All free-energy methods are based on calculation of free-energy *differences*
- Example
  - *volume of R can be measured as a fraction of the total volume*

$$\frac{\Omega_R}{\Omega_\Gamma} = \langle s(\Gamma) \rangle_\Gamma$$

sample the reference system

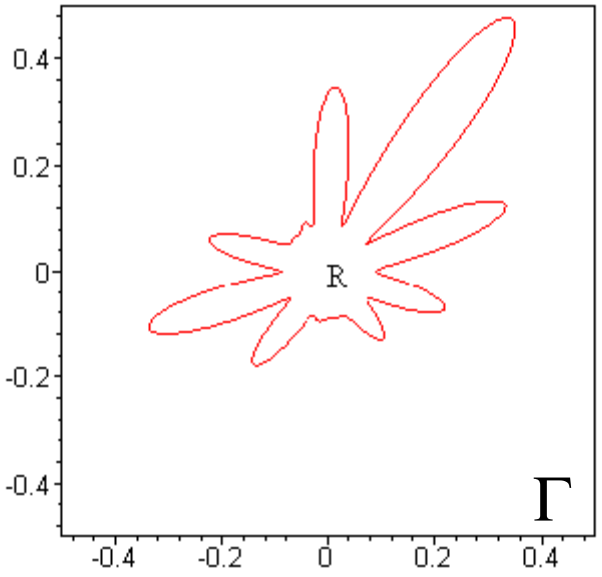
keep an average of the fraction of time occupying target system

  - *what we get is the difference*

$$S_R - S_\Gamma = k \ln(\Omega_R / \Omega_\Gamma)$$

- Usefulness of free-energy difference
  - *it may be the quantity of interest anyway*
  - *if reference is simple, its absolute free energy can be evaluated analytically*

e.g., ideal gas, harmonic crystal



Hard Sphere Chemical Potential

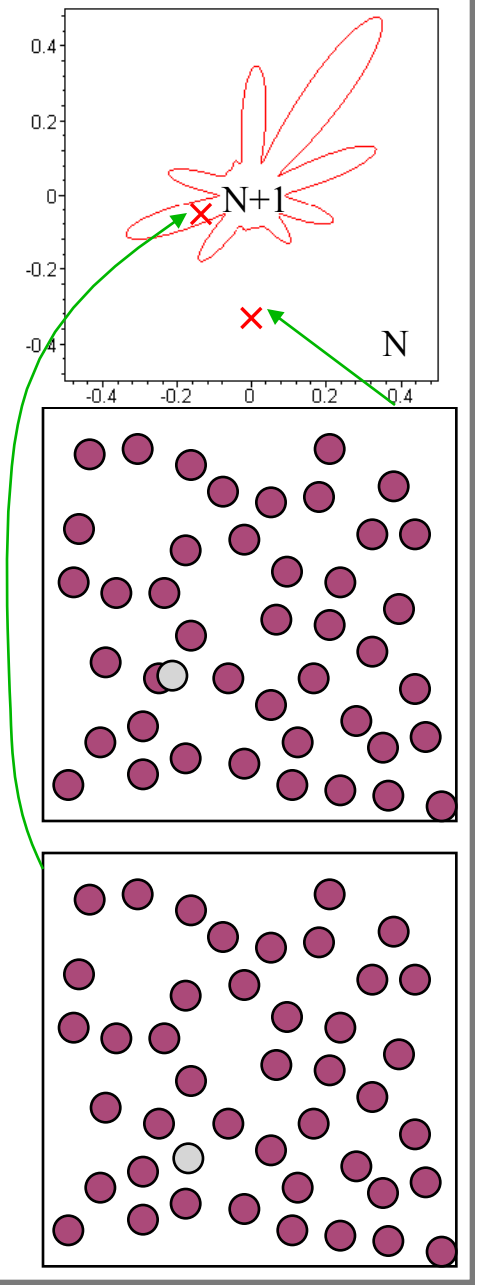
- Chemical potential is an entropy difference

$$\beta\mu = -\left(\frac{\partial S/k}{\partial N}\right)_{U,V} \approx -\left[S(U,V,N+1) - S(U,V,N)\right]$$

- For hard spheres, the energy is zero or infinity
  - *any change in N that does not cause overlap will be change at constant U*
- To get entropy difference
  - *simulate a system of N+1 spheres, one non-interacting "ghost"*
  - *occasionally see if the ghost sphere overlaps another*
  - *record the fraction of the time it does not overlap*

$$e^{\Delta S/k} = e^{-\beta\mu} = \frac{V}{N\Lambda^3}\left(\frac{\Omega_{N+1}}{\frac{V}{N\Lambda^3}\Omega_N}\right) = \frac{V}{N\Lambda^3}\left\langle f_{\text{non-overlap}}\right\rangle$$

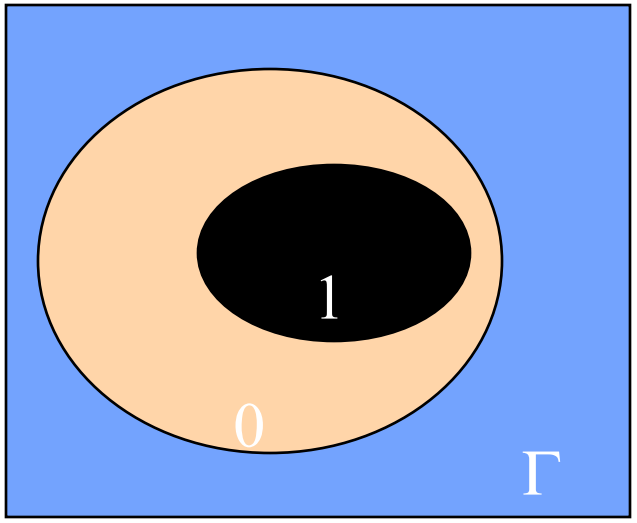
- Here is an applet demonstrating this calculation



Free-Energy Perturbation

- Widom method is an example of a free-energy perturbation (FEP) technique
- FEP gives free-energy difference between two systems
  - *labeled 0, 1*
- Working equation $e^{-\beta(A_1 - A_0)} = \frac{Q_1}{Q_0} = \frac{\int d\Gamma e^{-\beta U_1}}{\int d\Gamma e^{-\beta U_0}}$

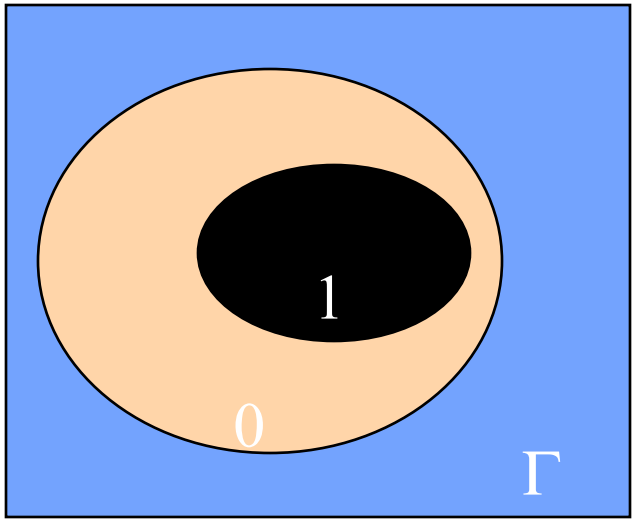
Free-energy difference is a ratio of partition functions



=

Free-Energy Perturbation

- Widom method is an example of a free-energy perturbation (FEP) technique
- FEP gives free-energy difference between two systems
  - *labeled 0, 1*
- Working equation $e^{-\beta(A_1-A_0)} = \frac{Q_1}{Q_0} = \frac{\int d\Gamma e^{-\beta U_1}}{\int d\Gamma e^{-\beta U_0}}$



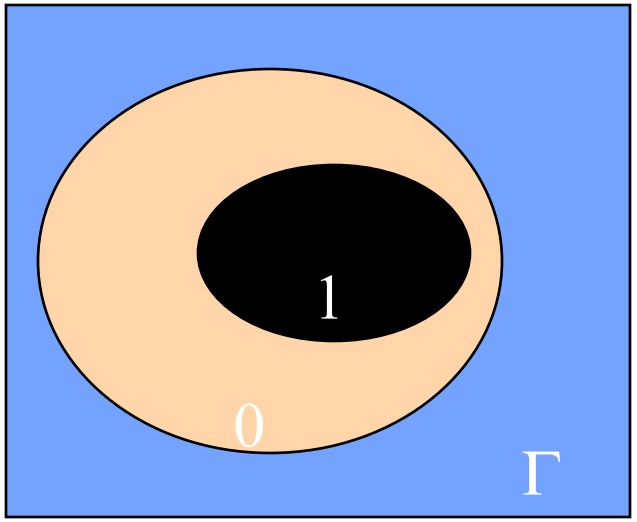
$$= \frac{\int d\Gamma e^{-\beta(U_1-U_0)} e^{-\beta U_0}}{\int d\Gamma e^{-\beta U_0}}$$

Add and subtract reference-system energy

$$=$$

Free-Energy Perturbation

- Widom method is an example of a free-energy perturbation (FEP) technique
- FEP gives free-energy difference between two systems
  - *labeled 0, 1*
- Working equation $e^{-\beta(A_1-A_0)} = \frac{Q_1}{Q_0} = \frac{\int d\Gamma e^{-\beta U_1}}{\int d\Gamma e^{-\beta U_0}}$



$$= \frac{\int d\Gamma e^{-\beta(U_1-U_0)} e^{-\beta U_0}}{\int d\Gamma e^{-\beta U_0}}$$

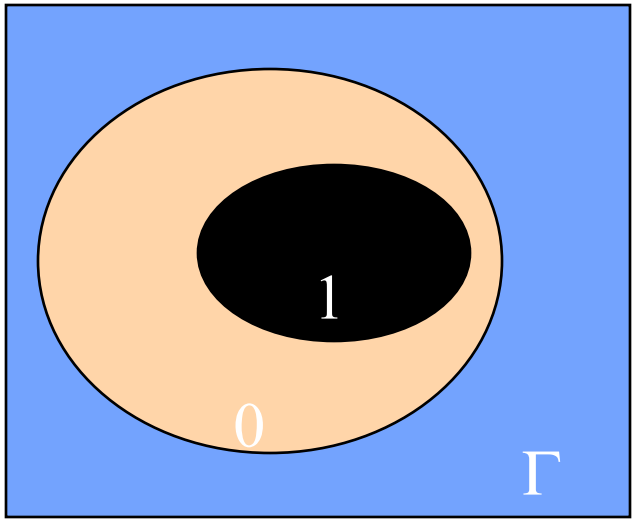
$$= \int d\Gamma e^{-\beta(U_1-U_0)} \pi_0(\Gamma)$$

$$=$$

Identify reference-system probability distribution

Free-Energy Perturbation

- Widom method is an example of a free-energy perturbation (FEP) technique
- FEP gives free-energy difference between two systems
  - *labeled 0, 1*
- Working equation $e^{-\beta(A_1-A_0)} = \frac{Q_1}{Q_0} = \frac{\int d\Gamma e^{-\beta U_1}}{\int d\Gamma e^{-\beta U_0}}$



$$= \frac{\int d\Gamma e^{-\beta(U_1-U_0)} e^{-\beta U_0}}{\int d\Gamma e^{-\beta U_0}}$$

$$= \int d\Gamma e^{-\beta(U_1-U_0)} \pi_0(\Gamma)$$

$$= \left\langle e^{-\beta(U_1-U_0)} \right\rangle_0$$

Write as reference-system ensemble average

- Sample the region important to 0 system, measure properties of 1 system

Chemical potential

- For chemical potential, U_1 - U_0 is the energy of turning on the ghost particle
  - *call this u_t, the "test-particle" energy*

$$e^{-\beta(A_1-A_0)} = e^{-\beta\mu}$$

$$= \frac{V}{N\Lambda^3}\left\langle e^{-\beta u_t}\right\rangle_0$$

  - *test-particle position may be selected at random in simulation volume*
  - *for hard spheres, $e^{-\beta u_t}$ is 0 for overlap, 1 otherwise*

    then (as before) average is the fraction of configurations with no overlap

- This is known as Widom's insertion method

Deletion Method

- The FEP formula may be used also with the roles of the reference and target system reversed

Original: 0 → 1

$$e^{-\beta(A_1-A_0)} = \left\langle e^{-\beta(U_1-U_0)} \right\rangle_0$$

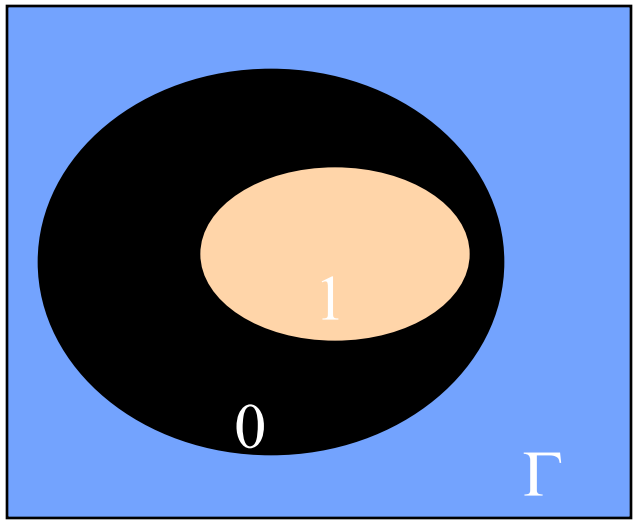
Modified: 1 → 0

$$e^{+\beta(A_1-A_0)} = \left\langle e^{+\beta(U_1-U_0)} \right\rangle_1$$

  - *sample the 1 system, evaluate properties of 0 system*

- Consider application to hard spheres

$$e^{+\beta\mu} = \frac{N}{V}\left\langle e^{+\beta u_t} \right\rangle_1$$



  - *$e^{\beta u_t}$ is infinity for overlap, 0 otherwise*
  - *but overlaps are never sampled*
  - *true average is product of 0 × ∞*
    technically, formula is correct
  - *in practice simulation average is always zero*
    method is flawed in application
    many times the flaw with deletion is not as obvious as this

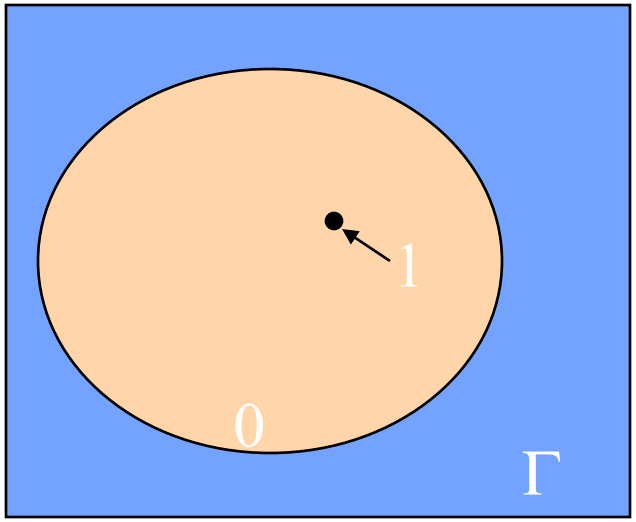
Other Types of Perturbation

- Many types of free-energy differences can be computed
- Thermodynamic state
  - *temperature, density, mixture composition*
- Hamiltonian
  - *for a single molecule or for entire system*
  - *e.g., evaluate free energy difference for hard spheres with and without electrostatic dipole moment*
- Configuration
  - *distance/orientation between two solutes*
  - *e.g, protein and ligand*
- Order parameter identifying phases
  - *order parameter is a quantity that can be used to identify the thermodynamic phase a system is in*
  - *e.g, crystal structure, orientational order, magnetization*

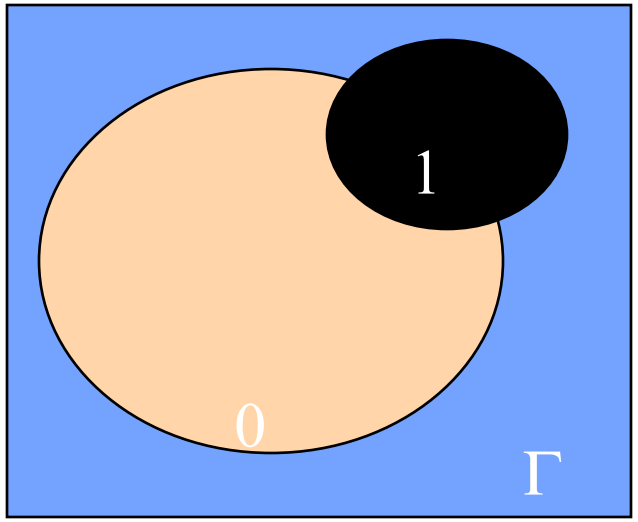
General Numerical Problems

- Sampling problems limit range of FEP calculations
- Target system configurations must be encountered when sampling reference system
- Two types of problem arise

target-system space very small



target system outside of reference



- *first situation is more common*

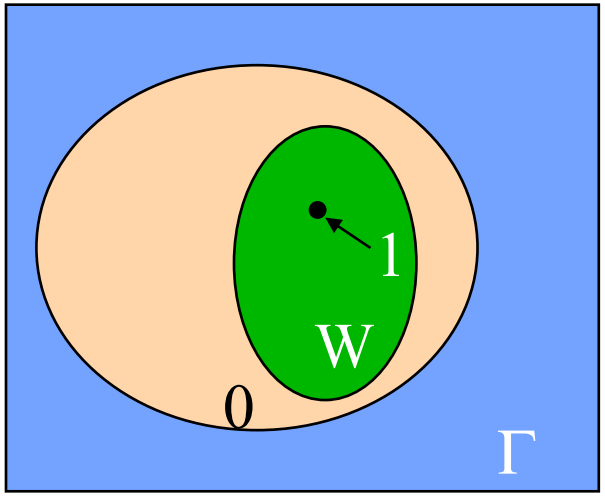
  although deletion FEP provides an avoidable example of the latter

Staging Methods

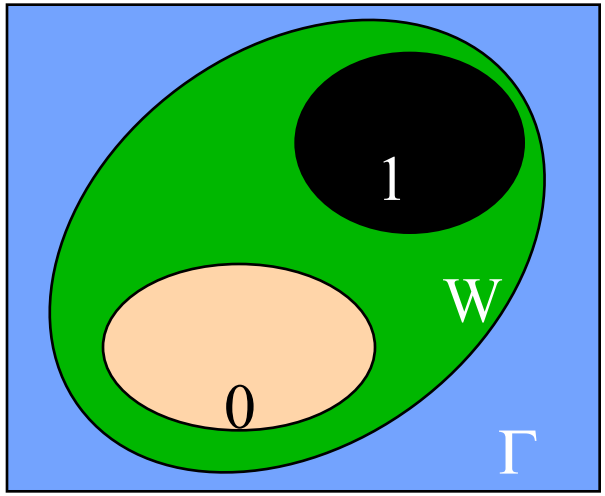
- Multistage FEP can be used to remedy the sampling problem
  - *define a potential U_w intermediate between 0 and 1 systems*
  - *evaluate total free-energy difference as* $A_1 - A_0 = (A_1 - A_w) + (A_w - A_0)$
- Each stage may be sampled in either direction
  - *yielding four staging schemes*
  - *choose to avoid deletion calculation*

$0 \leftarrow W \rightarrow 1$	Umbrella sampling
$0 \rightarrow W \leftarrow 1$	Bennett's method
$0 \leftarrow W \leftarrow 1$	Staged deletion
$0 \rightarrow W \rightarrow 1$	Staged insertion

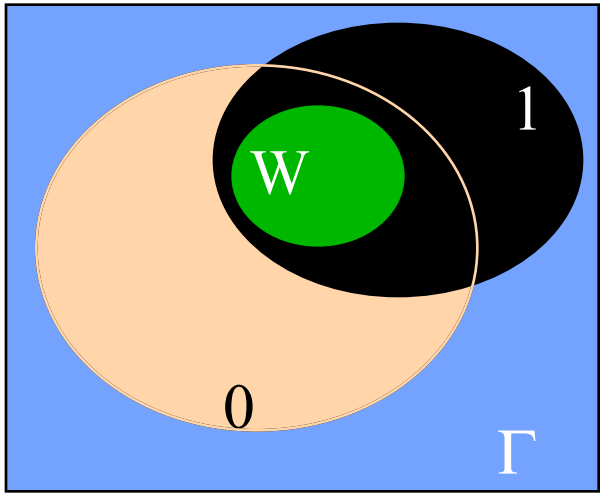
Use staged insertion



Use umbrella sampling



Use Bennett's method



Example of Staging Method

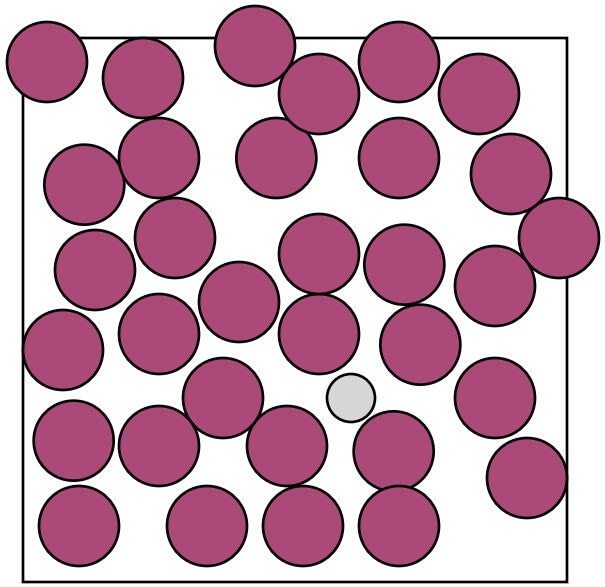
- Hard-sphere chemical potential
- Use small-diameter sphere as intermediate

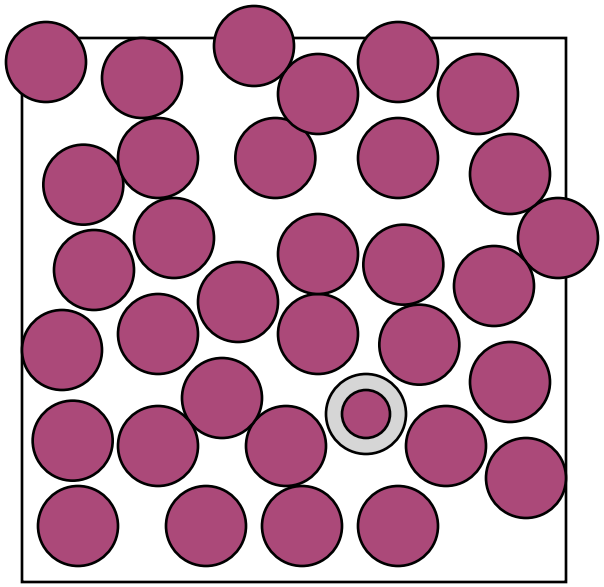
$$e^{-\beta\Delta(A_w - A_N)} = \left\langle e^{-\beta u(\sigma_t)} \right\rangle_N$$

In first stage, measure fraction of time random insertion of small sphere finds no overlap

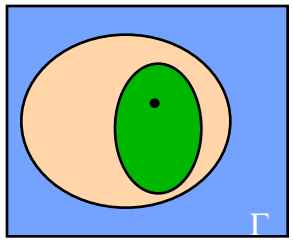
$$e^{-\beta\Delta(A_{N+1} - A_w)} = \left\langle e^{-\beta(u(\sigma) - u(\sigma_t))} \right\rangle_w$$

In second stage, small sphere moves around with others. Measure fraction of time no overlap is found when it is grown to full-size sphere

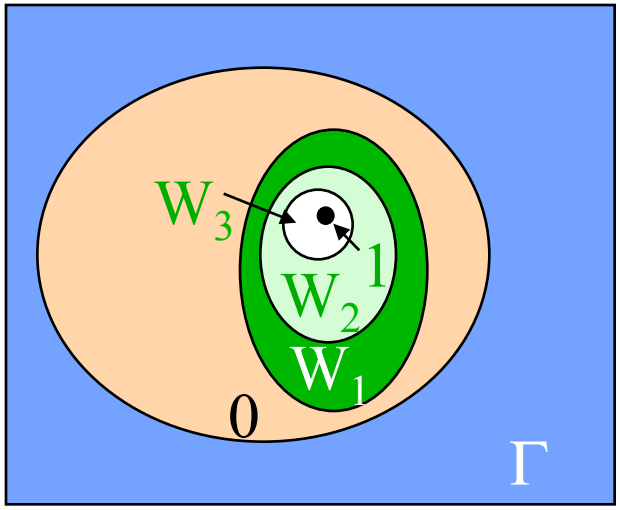




$$e^{-\beta\Delta(A_{N+1} - A_N)} = \left\langle e^{-\beta u(\sigma_t)} \right\rangle_N \left\langle e^{-\beta(u(\sigma) - u(\sigma_t))} \right\rangle_w$$

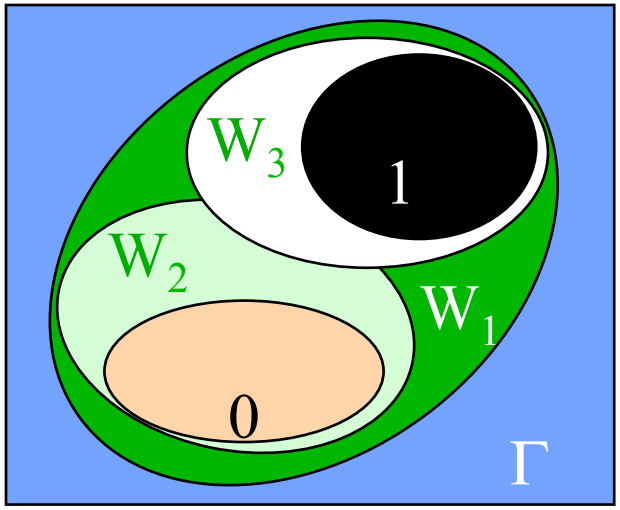


Multiple Stages



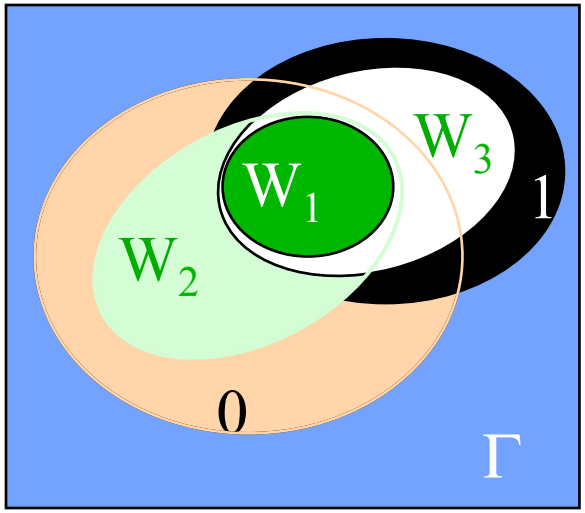
Multistage insertion

$0 \to W_1 \to W_2 \to W_3 \to 1$



Multistage umbrella sampling

$0 \leftarrow W_2 \leftarrow W_1 \to W_3 \to 1$



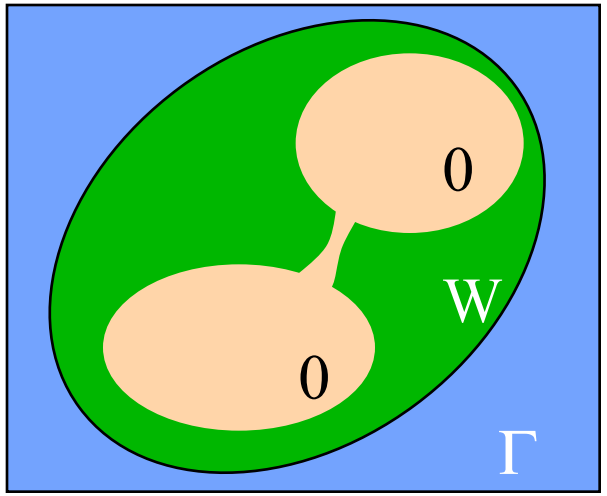
Multistage Bennett's method

$0 \to W_2 \to W_1 \leftarrow W_3 \leftarrow 1$

Non-Boltzmann Sampling

- The FEP methods are an instance of a more general technique that aims to improve sampling
- Unlike biasing methods, improvement entails a change in the limiting distribution
- Apply a formula to recover the correct average

$$\begin{aligned}\langle M \rangle_0 &= \frac{1}{Q_0}\int d\Gamma M(\Gamma) e^{-\beta U_0} \\ &= \frac{Q_W}{Q_0}\frac{1}{Q_W}\int d\Gamma M(\Gamma) e^{-\beta(U_0-U_W)} e^{-\beta U_W} \\ &= \frac{\left\langle M e^{-\beta(U_0-U_W)} \right\rangle_W}{\left\langle e^{-\beta(U_0-U_W)} \right\rangle_W}\end{aligned}$$



Thermodynamic Integration 1.

- Thermodynamics gives formulas for variation of free energy with state

$$d(\beta A) = U d\beta - \beta P dV + \beta \mu dN$$

$$\left(\frac{\partial \beta A}{\partial \beta}\right)_{V,N} = U \qquad \left(\frac{\partial \beta A}{\partial V}\right)_{T,N} = -\beta P$$

- These can be integrated to obtain a free-energy difference
  - *derivatives can be measured as normal ensemble averages*

$$\beta A(V_2) = \beta A(V_1) - \int_{V_1}^{V_2} P(V) dV$$

  - *this is usually how free energies are "measured" experimentally*

Thermodynamic Integration 2.

- TI can be extended to follow uncommon (or unphysical) integration paths
  - *much like FEP, can be applied for any type of free-energy change*
- Formalism
  - *Let λ be a parameter describing the path*
  - *the potential energy is a function of λ*
  - *ensemble formula for the derivative*

$$\frac{\partial \beta A}{\partial \lambda} = -\frac{\partial \ln Q}{\partial \lambda} = -\frac{1}{Q}\frac{\partial}{\partial \lambda}\left[\frac{1}{\Lambda^{3N} N!}\int dr^N e^{-\beta U(r^N;\lambda)}\right]$$

$$= +\frac{1}{Q}\left[\frac{1}{\Lambda^{3N} N!}\int dr^N e^{-\beta U(r^N;\lambda)}\frac{\partial}{\partial \lambda}\beta U(r^N;\lambda)\right]$$

$$= \left\langle \frac{\partial \beta U}{\partial \lambda} \right\rangle$$

  - *then*

$$\beta A(\lambda_2) = \beta A(\lambda_1) + \int_{\lambda_1}^{\lambda_2} \left\langle \frac{\partial \beta U}{\partial \lambda} \right\rangle d\lambda$$

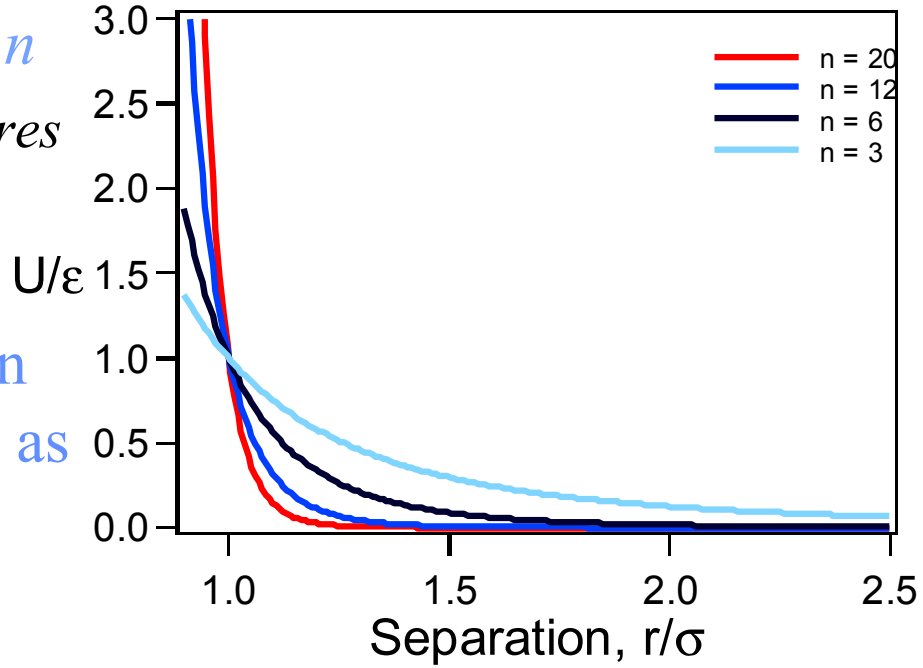
Thermodynamic Integration Example

- The soft-sphere pair potential is given by

$$u(r) = \varepsilon\left(\frac{\sigma}{r}\right)^n$$

Exhibits simplifying behavior because $\varepsilon\sigma^n$ is the only potential parameter

- Softness and range varies with *n*
  - *large n limit leads to hard spheres*
  - *small n leads to Coulombic behavior*



- Thermodynamic integration can be used to measure free energy as a function of softness $s = 1/n$
  - *Integrand is*

$$\frac{\partial \beta A}{\partial s} = -\frac{\beta\varepsilon}{s^2}\langle u(r)\ln(\sigma/r)\rangle$$

Parameter Hopping. Theory

- View free-energy parameter λ as another dimension in phase space

$$E = E(\mathbf{p}^N, \mathbf{r}^N, \lambda)$$

- Partition function

$$Q = \sum_{\lambda} \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda)}$$

$$= \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda_0)} + \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda_1)}$$

$$= Q_0 + Q_1$$

- Monte Carlo trials include changes in λ
- Probability that system has $\lambda = \lambda_0$ or $\lambda = \lambda_1$

$$\pi(\lambda_0) = \frac{1}{Q_0 + Q_1} \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda_0)} = \frac{Q_0}{Q_0 + Q_1}$$

$$\pi(\lambda_1) = \frac{1}{Q_0 + Q_1} \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda_1)} = \frac{Q_1}{Q_0 + Q_1}$$

Parameter Hopping. Implementation

- Monte Carlo simulation in which l-change trials are attempted
- Accept trials as usual, with probability $\min[1, e^{-\beta\Delta U}]$
- Record fractions f_0, f_1 of configurations spent in $\lambda = \lambda_0$ and $\lambda = \lambda_1$
- Free energy is given by ratio

$$e^{-\beta(A_1 - A_0)} = \frac{Q_1}{Q_0} = \frac{Q_1/(Q_0 + Q_1)}{Q_0/(Q_0 + Q_1)} = \frac{f_1}{f_0}$$

- In practice, system may spend almost no time in one of the values
  - *Can apply weighting function* $w(\lambda)$ *to encourage it to sample both*
  - *Accept trials with probability* $\min[1, (w^n/w^o)\, e^{-\beta\Delta U}]$
  - *Free energy is*

$$e^{-\beta(A_1 - A_0)} = \frac{w_0 f_1}{w_1 f_0}$$

  - *Good choice for w has* $f_1 = f_0$
- Multivalue extension is particularly effective
  - *l takes on a continuum of values*

Summary

- Free energy calculations are needed to model the most interesting physical behaviors
  - *All useful methods are based on computing free-energy difference*
- Four general approaches
  - *Free-energy perturbation*
  - *Thermodynamic integration*
  - *Parameter hopping*
  - *Distribution-function methods*
- FEP is asymmetric
  - *Deletion method is awful*
- Four approaches to basic multistaging
  - *Umbrella sampling, Bennett's method, staged insertion/deletion*
- Non-Boltzmann methods improve sampling