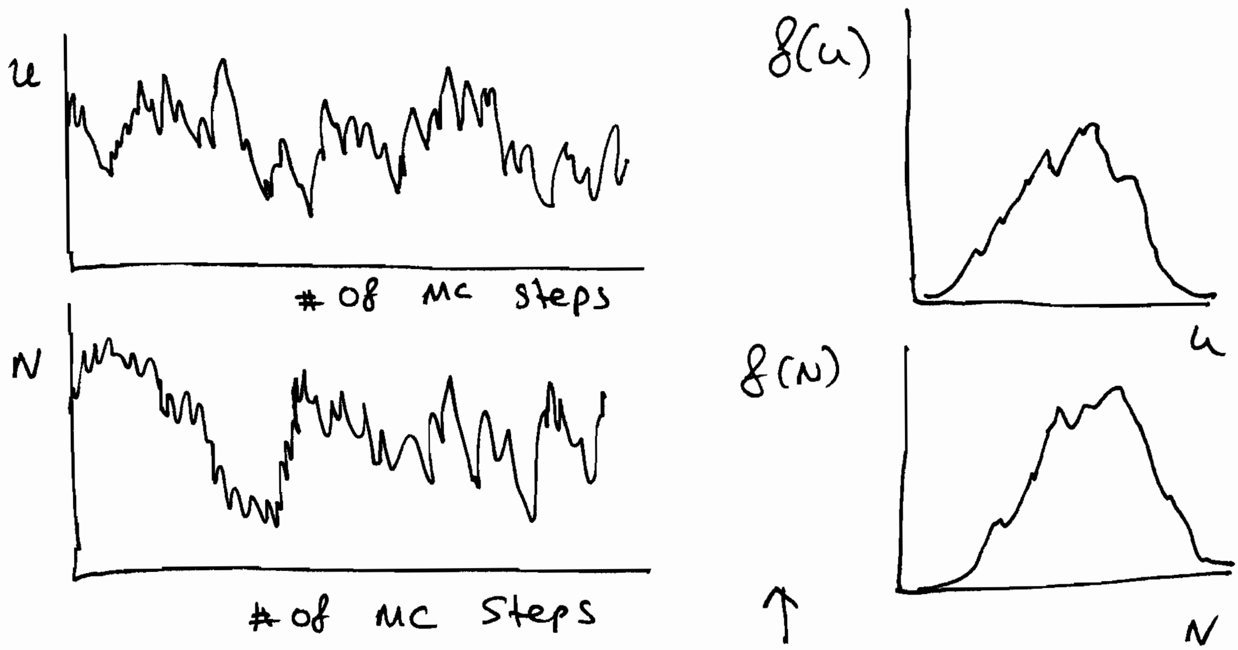
Histograms - Single

Even though we have stated that "statistical" properties such as A, G (or μ_i), S cannot be readily obtained from simulations, fluctuations provide useful quantities related to differences in free energy.

E.g. NVT or μVT simulations (at equilibrium)



↑ Count of how many times given u, N was observed

$$f(u) = \frac{\Omega(N,V,u)\cdot \exp(-\beta u)}{Q(N,V,T)} \Rightarrow$$

$$\ln f(u) + \beta u = \ln \Omega(N,V,u) + C$$

$\boxed{k \ln \Omega = S \ !}$

unknown constant related to Q

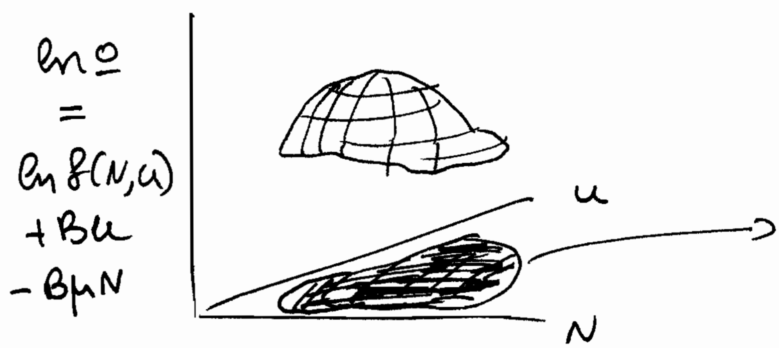
or – for μVT, keeping track of both $U + N$ histogram:

$$f(N,U) = \frac{\Omega(N,V,U)\exp(-\beta U + \beta\mu N)}{\Xi(\mu,V,T)} \Rightarrow$$

$$\ln f(N,U) + \beta U - \beta\mu N = \ln\Omega(N,V,U) + \underbrace{c}_{\text{unknown constant}}$$

So, a given simulation "illuminates" a section of the entropy surface for a system, from which all other quantities can be obtained.

e.g. $A = U - TS \qquad G = U - TS + PV$



region of U, N visited in simulation (depends on μ, T, V)

"Nearby" states can be quickly predicted from the original run by "reweighting" of $f(N,U)$

$$f_{expected}(N,U;\mu',T',V) = \frac{\Omega(N,V,U)\cdot\exp(-\beta'U + \beta'\mu'N)}{\Xi'(\mu',V,T')}$$

$$\Rightarrow f_{expected}(N,U;\mu',T',V) = f(N,U;\mu,V,T)\cdot$$

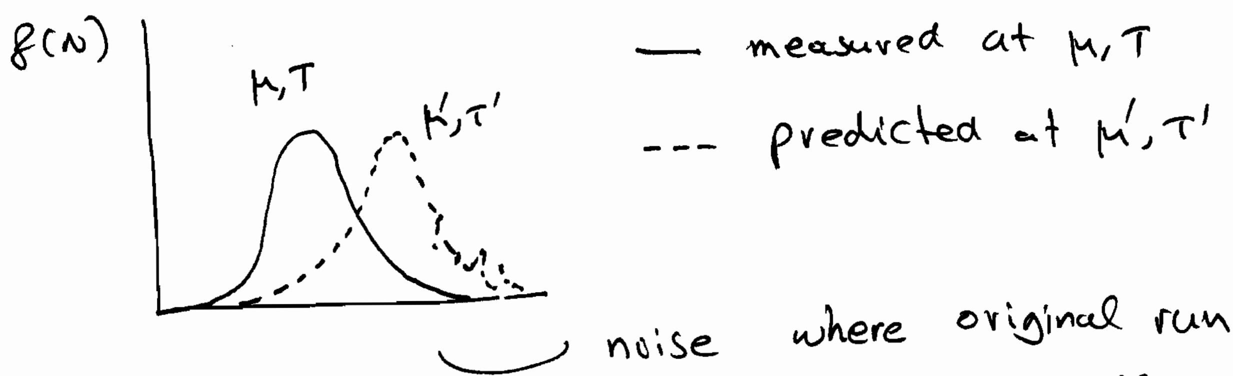
$$\cdot\exp\left(-\beta'U + \beta U + (\beta'\mu' - \beta\mu)N\right)\cdot\underbrace{\frac{\Xi(\mu,V,T)}{\Xi'(\mu',V,T')}}_{\text{constant}}$$

Even though the ratio $\Xi(\mu, V, T)/\Xi(\mu', V, T')$ is not known, it is constant for the two states; the new probability distribution $f_{expected}(N, u; \mu', T', V)$ can be obtained from the normalization condition,

$$\sum_{\text{all } N} \sum_{\text{all } U} \tilde{f}_{expected}(N, u) = 1$$

↳ normalized f

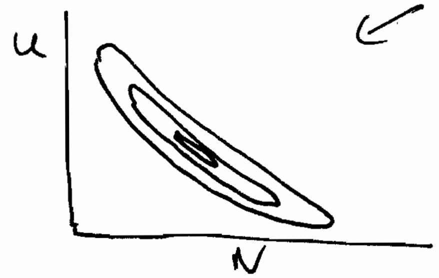
In 1-dimensional projection



Clearly, this approach will fail if μ', T' correspond to very different states than μ, T

For smaller systems, the relative range of N, u visited is greater than for large systems.

Practical Issues: How are histograms stored?

(a) 

← typical projection

– sparse matrix

need N_{min}, N_{max}
U_{min}, U_{max}

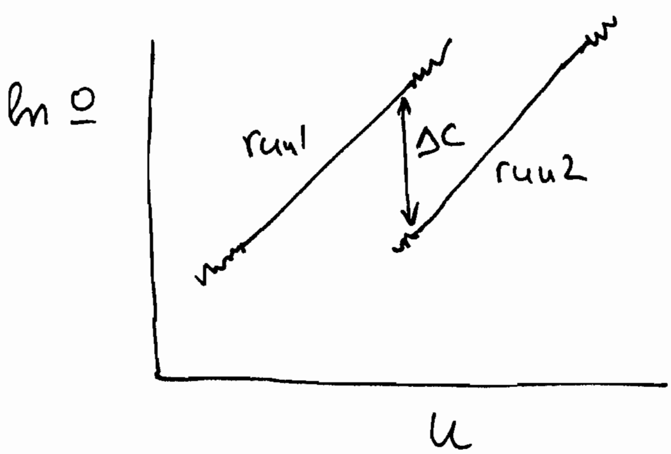
(b) Periodically sample N, U, print to file

→ can grow to large size

→ works for multicomponent systems: $N_1, N_2, \ldots N_n, U$.

Multiple Histograms

The real power of the method comes when multiple histograms at different conditions are combined to obtain a global free energy (or entropy) surface.



run1 @ T_1 gives

$$\ln \Omega_1(N,V,U) = \ln f_1(U) + \beta U + C_1$$

run 2 @ T_2:

$$\ln \Omega_2(N,V,U) = \ln f_2(U) + \beta U + C_2$$

But at same N, U → Ω is unique!

Need to determine ΔC to bring runs into agreement

$$\Delta C = \ln \frac{Q_1(N,V,T_1)}{Q_2(N,V,T_2')}$$ (for $f(U)$ histograms)

$$\Delta C = \ln \frac{\Xi_1(N,V,T_1)}{\Xi_2(\mu_2,V,T_2)}$$

for $f(U,N)$ histograms

$$= \Delta(PV/kT)$$

since $k \ln \Xi = S - \frac{U}{T} + \frac{\mu N}{T} = \frac{TS - U + G}{T} = \frac{PV}{T}$

The "shifts" ΔC are related to free energy differences / partition function ratios.

Ferrenberg + Swendsen, Phys. Rev. Lett. 63, 1195 (1989) derived a method to combine multiple histograms so as to minimize deviations between observed and predicted frequencies:

* guess C_i
* Obtain (unnormalized) frequencies $f_i(N,U)$ at μ_i, T_i
* Compute $$\wp(N,U;\mu_i,T_i) = \frac{\sum_{i=1}^{R} f_i(N,U)\, e^{-\beta_i U+\beta_i\mu_i N}}{\sum_{i=1}^{R} k_i\, e^{-\beta_i U+\beta_i\mu_i N - C_i}} \quad (1)$$

where R: # of runs

k_i # of observations $= \sum_{N,U} f_i(N,U)$ for run i

* Calculate new "weights" C_i from $$\exp(C_i) = \sum_U \sum_N \wp(N,U;\mu_i,T_i) \quad (2)$$

Iterating (1) + (2) leads to a self-consistent set of weights (free energies)

If sufficient data exist to low pressures for which $\frac{pV}{kT} = \langle N\rangle$, the absolute value of the pressure can be calculated in addition to the free energy (chemical potential) as a function of density ρ.