Reweighting

Suppose that a MC simulation generates samples $X_t \sim \pi_\beta(X_t) = \frac{1}{Z} e^{-\beta E(x)}$ from a canonical distribution at temperature $T = \frac{1}{\beta}$.

Using these we can estimate expectation values $\langle A \rangle_\beta = \int A(x)\,\pi_\beta(x)\,dx \approx$

$$\approx \bar{A}_\beta = \frac{1}{N}\sum_t A(x_t)$$

But we can also use the samples to estimate expectation values at different β':

$$\langle A \rangle_{\beta'} = \int A(x)\,\pi_{\beta'}(x)\,dx = \int A(x) \underbrace{\frac{\pi_{\beta'}(x)}{\pi_\beta(x)}}_{\hat{W}(x)} \pi_\beta(x)\,dx =$$

$$= \langle A \hat{W} \rangle_\beta \approx$$

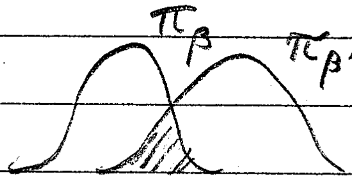
$$\approx \bar{A}_{\beta'} = \frac{1}{N}\sum_t A(x_t)\,\hat{W}(x_t) \approx$$

$$\approx \frac{\sum_t A(x_t)\,W(x_t)}{\sum_t W(x_t)}$$

The last self-normalized version has the advantage that we don't need to know the normalization constants $Z_\beta, Z_{\beta'}$.

Thus, we can extrapolate the simulation at β to nearby β'.

The extrapolation only works in a rather small region around β, since the samples from π_β must contribute significantly to $\pi_{\beta'}$



↑ the overlap needs to be large.

If not the fluctuations in the weight $w(x) = e^{-(\beta'-\beta)E(x)}$ beomes too large.

Mathematically we must require that $\pi_{\beta'}(\cdot)$ is absolutely continuous with respect to $\pi_\beta(\cdot)$.

<u>Single histogram reweighting</u>

In practice one may collect a histogram of the energies encountered in the simulation,

$$H(E) = \sum_{t=1}^{N} \delta_{E,E(x_t)} \quad \text{and} \quad A(E) = \sum_{t=1}^{N} A(x_t)\,\delta_{E,E(x_t)}$$

$$\text{Then} \quad \langle A \rangle_{\beta'} \approx \bar{A}_{\beta'} = \frac{\sum_E A(E)\, e^{-(\beta'-\beta)E}}{\sum_E e^{-(\beta'-\beta)E}\, H(E)}$$

WHAM

To extrapolate over a wider range one can use the multiple histogram method, a.k.a. WHAM, Weigthed Histogram Analysis Method (see also MBAR, which is a binless method working with the time series instead of histograms.)

Recall that $\pi_\beta(x) = \frac{e^{-\beta E(x)}}{Z}$,

$$\pi_\beta(E) = \int \delta(E - E(x))\pi_\beta(x)\,dx = \Omega(E)\frac{e^{-\beta E}}{Z}$$

Density of states

From a single histogram $H_{\beta_i}(E)$ collected at β_i we can estimate

$$\Omega_i(E) = Z_i e^{\beta_i E}\frac{H_{\beta_i}(E)}{N_i} = e^{\beta_i(E-F_i)}\frac{H_{\beta_i}(E)}{N_i}$$

Since $\Omega(E)$ does not depend on β we can combine the estimates Ω_i taken at different β_i into

$$\bar{\Omega}(E) = \sum_i \alpha_i(E)\,\Omega_i(E)$$

Choose $\alpha_i(E)$ to minimize the variance of $\bar{\Omega}(E)$ under the constraint $\sum \alpha_i(E) = 1$.

Natural to assume $\mathrm{Var}\, H_i(E) = \underbrace{\langle H_i(E)\rangle}$ (Poisson distr)

$$\mathrm{Var}\,\bar{\Omega} = \sum_i \alpha_i^2(E)\left(\frac{e^{\beta_i(E-F_i)}}{N_i}\right)^2 N_i e^{\beta_i(F_i-E)}$$

$$\mathrm{Min}\ \mathrm{Var}\,\bar{\Omega} - \lambda\sum_i \alpha_i \Rightarrow \alpha_i(E) = \frac{N_i e^{\beta_i(F_i-E)}}{\sum_i N_i e^{\beta_i(F_i-E)}}$$

$\Rightarrow$ the WHAM eqs

$$\bar{\Omega}(E) = \frac{\sum_i H_i(E)}{\sum_i N_i e^{\beta_i(F_i - E)}}$$

The unknown free energies F_i may be solved for selfconsistently

$$e^{-\beta F} = \sum_E \bar{\Omega}(E) e^{-\beta E} = \sum_E \frac{\sum_i H_i(E) e^{-\beta E}}{\sum_i N_i e^{\beta_i(F_i - E)}}$$

The temperatures β_i should be chosen so that the neighboring distributions overlap