6 Reweighting

The configurations generated in a Monte Carlo simulation contain a huge amount of information, from which we usually distill a couple of numbers.

It would be a shame to waste all that information. **Reweighting** is a method which allows us to "expand" the results from the original simulation, performed at inverse temperature β_0, say, to any other β sufficiently close to the simulation point without performing any additional simulations.

The simplest form of the reweighting is based on the fact that the canonical probability of a configuration ϕ at inverse temperature β, $p_\beta(\phi)$, can be easily related to the distribution at other temperature β':

$$p_{\beta'}(\phi) \propto e^{-\beta' E_\phi} = C e^{-(\beta'-\beta)E_\phi} p_\beta(\phi)\,,$$

where C is a proportionality constant (which depends on β and β', and will remain undetermined). Thus, the expectation value of an operator $O(\phi)$ at temperature β' can be written as

$$\langle O\rangle_{\beta'} \equiv \frac{1}{Z_{\beta'}} \int d\phi\, O(\phi)\, p_{\beta'}(\phi) = \frac{C}{Z_{\beta'}} \int d\phi\, O(\phi) e^{-(\beta'-\beta)E_\phi} p_\beta(\phi) = \frac{Z_\beta}{Z_{\beta'}} C \, \langle O e^{-(\beta'-\beta)E}\rangle_\beta\,,$$

where in the last step the expectation value is evaluated at temperature β. In order to get the ratio of the partition functions, we can set $O = 1$, which implies

$$\frac{Z_{\beta'}}{Z_\beta} = C\, \langle e^{-(\beta'-\beta)E}\rangle_\beta\,.$$

Thus, finally we obtain the desired result:

$$\boxed{\langle O\rangle_{\beta'} = \frac{\langle O e^{-(\beta'-\beta)E}\rangle_\beta}{\langle e^{-(\beta'-\beta)E}\rangle_\beta}\,.}$$

This implies that the expectation value of any observable at any inverse temperature β' can be obtained in terms of expectation values evaluated at β.

This appears to indicate that it should be possible to do a simulation at one value of β, and use the above formula to obtain results at any other temperature. In reality the situation is not so simple, due to the finite statistics in realistic simulations. This will be discussed below.

6.1 Reweighting in Monte Carlo simulations

Let us consider the case that we perform a simulation of some system at coupling β. The Monte Carlo simulation gives us a series of configurations $\phi_1, \phi_2 \ldots \phi_N$, and measurements of some observable $O_i = O(\phi_i)$.

Now the standard estimate of the expectation value is:

$$\langle O \rangle_\beta \equiv \frac{1}{Z} \sum_{\{\phi\}} O(\phi) \, \exp\left[-\beta E_\phi\right] \approx \frac{1}{N} \sum_{i=1}^{N} O_i$$

where the first sum goes over full configuration space, and the second over the Monte Carlo configurations/measurements.

Thus, the reweighting formula becomes

$$\boxed{\langle O \rangle_{\beta'} = \frac{\sum_i O_i \, e^{-(\beta'-\beta)E_i}}{\sum_i e^{-(\beta'-\beta)E_i}} = \frac{\langle O \, e^{-(\beta'-\beta)E} \rangle_\beta}{\langle e^{-(\beta'-\beta)E} \rangle_\beta}}$$

Note: $\sum_i$ goes over measurements, and E_i, O_i must be measured from the same configuration.

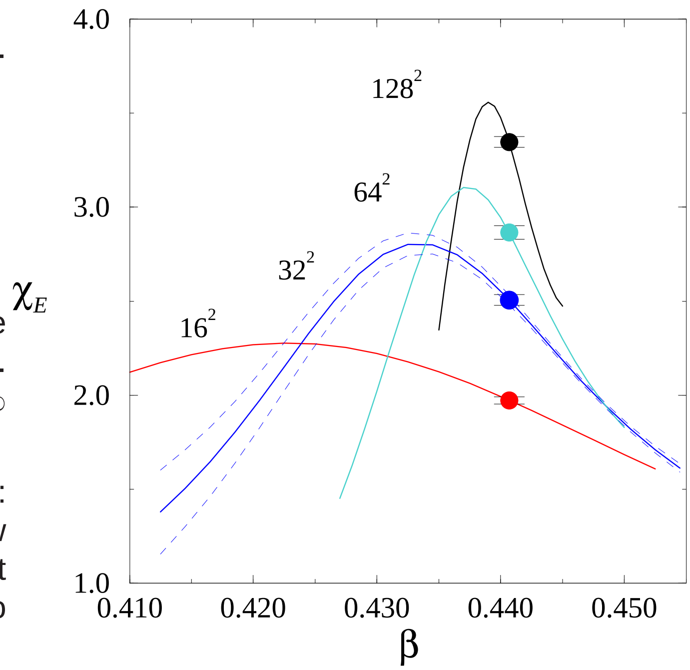
In practice: perform a Monte Carlo simulation at β, and during the simulation write down all measurements E_i and various desired operators O_i in a file. After the run, use the equation above to calculate $\langle O \rangle_\beta$.

Example: 2d Ising, $V = 16^2 \ldots 128^2$. Specific heat (susceptibility of E)

$$C_V = -\frac{1}{V} \frac{\partial \langle E \rangle}{\partial \beta}$$
$$= \frac{1}{V} \langle (E - \langle E \rangle)^2 \rangle$$

Susceptibilities diverges with some critical exponent at the critical point. In this case, $C_V \sim L^{\alpha/\nu}$ when $L \to \infty$ and we are at $\beta = \beta_c$:

Points: simulated values, curves: reweighted data. Dashed lines show the error band for 32^2 (for clarity, not shown for others). We will return to errors below.



6.2 Alternative view: reweighting using histograms

Using histograms (probability distributions from simulations) for reweighting is the "original" reweighting method. This is a very intuitive approach but somewhat restricted, and superseded by the approach in Sec. 6.1.

For concreteness, consider an Ising model simulation at β. Now the energy E has only discrete values (E is integer).

- During a MC run (at β), measure the energy histogram (energy probability distribution) $h_\beta(E)$ just by counting how many times energy value E occurs during the simulation.
- We know that $h_\beta(E) \propto n(E)\exp(-\beta E)$, where $n(E)$ is the number of states (density of states) at energy E and independent of β.
- Thus, we can reweight h: $h_{\beta'}(E) \propto h_\beta(E)\exp[-(\beta' - \beta)E]$.
- Finally, if $O(E)$ is a function of energy, we can calculate

$$\langle O \rangle_{\beta'} = \frac{\sum_E O(E) h_{\beta'}(E)}{\sum_E h_{\beta'}(E)} = \frac{\sum_E O(E) h_\beta(E)\exp[-(\beta' - \beta)E]}{\sum_E h_\beta(E)\exp[-(\beta' - \beta)E]}$$

- **Main weakness:** this works only if $O[\{s\}] = O[E(\{s\})]$ is a function of energy, whereas the equation in page 65 works for arbitrary observable.
- If the energy is not discrete, we need to bin the energy values. This causes binning errors.

Note: histogram method can be obtained directly from the equation on page 65 using $O = \delta_{E,E'}$.)

History of reweighting:

- First proposed in ’59 [Salzburg et al, J.Chem.Phys 30 (1959) 60]
- First used by McDonald and Singer 1967, no success (reweighting range too small?)
- Shown to be very effective by **Ferrenberg and Swendsen** [PRL 61 (1988) 2635]; now the whole thing goes under the name F-S reweighting.
- Multihistogram method: F+S [PRL 63 (1989) 1195; Computers in Physics, Sep/Oct 1988]

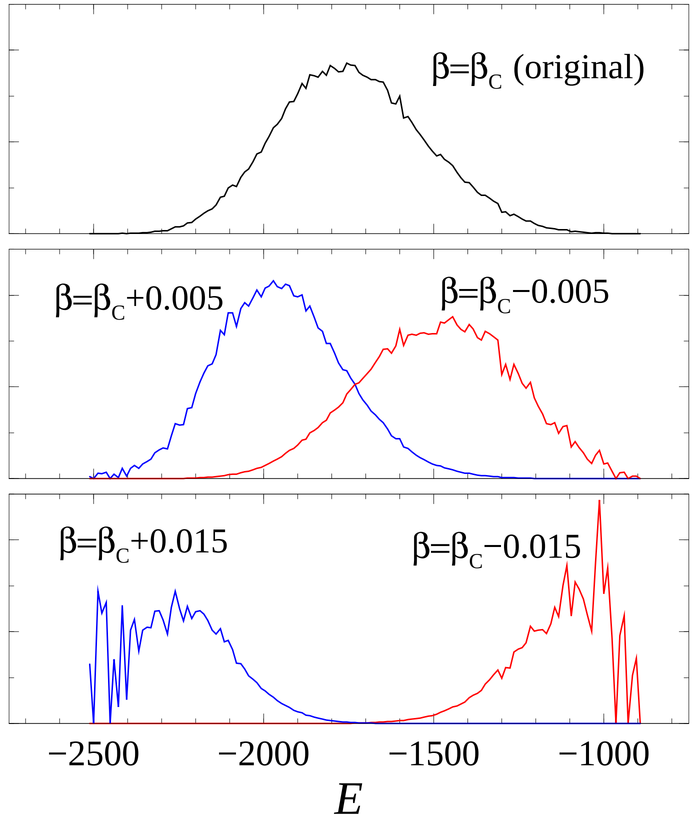
Distribution functions:
Histograms of energy can be obtained using $O_i = \delta(E, E_i)$:

$$h_\beta(E) = \frac{\sum_i \delta(E, E_i) e^{-\delta\beta E_i}}{\sum_i e^{-\delta\beta E_i}},$$

where $\delta\beta = \beta' - \beta$.

Histograms of 64^2 Ising model, with original simulation at $\beta = \beta_c \approx 0.44$.

The histograms become "exponentially" worse when the reweighting distance increases. This is due to the limited statistics in the tails of the original distribution. This is what restricts the range of β where one can reweight.



6.3 How far one can reweight?

When we reweight, the expectation value of the histogram shifts sideways. The *bulk* and *expectation value* of the shifted histogram should be well within the original histogram, in order not to lose statistical significance and keeping the errors down to a manageable level (see the histograms on page 67).

We can estimate the allowed reweighting range as follows:

Let us again consider Ising model. If we perform the simulation at β_0, we should only go to β where

$$|\delta\langle E\rangle| = |\langle E\rangle_\beta - \langle E\rangle_{\beta_0}| \lesssim \langle (E - \langle E\rangle_{\beta_0})^2\rangle_{\beta_0}^{1/2} = [V\, C_E(\beta_0)]^{1/2},$$

where C_V is the specific heat ($\sim$ susceptibility of energy). Because

$$V\, C_V = -\left(\frac{\partial\langle E\rangle}{\partial\beta}\right)_{\beta=\beta_0} = \langle (E - \langle E\rangle)^2\rangle,$$

we can Taylor expand

$$\langle E\rangle_\beta = \langle E\rangle_{\beta_0} - (\beta - \beta_0) V\, C_V$$

to obtain

$$|\delta\beta| = |\beta - \beta_0| \lesssim [V\,C_V]^{-1/2}.$$

This actually should be valid for *any* observable, not only E, if we substitute the corresponding susceptibility.
Thus, we see that valid reweighting range is $\lesssim$ natural 'fluctuation' range in the simulation, in order not to lose statistical signficance.

- Simulation at a non-critical point: now C_V is a finite number (with regular β-dependence). Thus, in this case $\Delta\beta \propto 1/\sqrt{V}$. This is a rule of thumb in reweighting: the allowed range is *reduced* as $V^{-1/2}$ as volume increases (at constant β).
- At a critical point (as in our Ising model example): $\Delta\beta \propto 1/V^x$, where x is some critical exponent. In our case above, the specific heat behaves as $C_V \sim V^{\alpha/(d\nu)}$, and we obtain $x = (1 + \alpha/(d\nu))/2$, where $d = 2$ is the dimensionality.

PRAGMATIC VIEW: check that the 'mass' of the reweighted histogram does not shift too far away into tails of the original histogram, where there is insufficient amount of data! Naturally, more statistics $\rightarrow$ slightly larger range.

The scaling laws are bad:

- Reweighting range $\propto 1/\sqrt{V}$.
- If we insist on increasing range by a factor of n, $\delta\beta \rightarrow n\delta\beta$, and if we assume that $\delta\langle E\rangle_{n\delta\beta} = n\delta\langle E\rangle_{\delta\beta}$, then we have to increase statistics by a factor $\sim \exp[n^2 - 1]$ to achieve comparable accuracy (tail of a Gaussian)!

  The reweighting range cannot practically be increased by statistics. Much better to perform new simulations with different β.

6.4 Error analysis in reweighting

- Errors analyzed in previous plot using jackknife error analysis Use of **jackknife** or **bootstrap** methods strongly recommended! (Return to that in Sec. 8.)
- Errors increase when reweighting distance increases. Must not do reweighting too far from original simulation point! (how far?)

- Note: using normal error analysis in upstairs and downstairs of

$$\langle O\rangle_\beta = \frac{\langle O\, e^{-(\beta-\beta_c)E}\rangle_{\beta_c}}{\langle e^{-(\beta-\beta_c)E}\rangle_{\beta_c}}$$

  is usually not reliable: the exponential factors make the 'observable' inside $\langle\cdot\rangle$ very skewed. Furthermore, this *overestimates* errors, since the deviations up and down are correlated!

- Calculating the errors using **jackknife** (for details of jackknife, see Sec. 8)

  1. Divide the data (N measurements) into M blocks, length $m = N/M \gg \tau$ (this to make individual blocks statistically independent).
  2. For each m, *delete* block m from the full data, and calculate

     $$\langle O\rangle_\beta^m = \frac{\sum_i O_i\, e^{-(\beta-\beta_c)E_i}}{\sum_i e^{-(\beta-\beta_c)E_i}}$$

     where the sums go over the $N-m$ measurements which do not belong to block m.
  3. Calculate the error through

     $$\delta\langle O\rangle_\beta = \sqrt{\frac{M-1}{M}\sum_m(\langle O\rangle_\beta^m - \langle O\rangle_\beta)^2}$$

     where $\langle O\rangle_\beta$ is either the full dataset reweighted value, or, usually the average of $\langle O\rangle_\beta^m$. The difference is very small, but the average of $\langle O\rangle_\beta^m$ is used for *bias correction*.

6.5 Reweighting with respect to arbitrary parameters

Above we discussed reweighting wrt. the inverse temperature β. However, it can be done using any parameter of the action/energy function.

As an example, let us consider reweighting with respect to **external magnetic field** h in the Ising model

$$Z_{\beta,h} = \sum_s \exp[-\beta E + hM] \quad \textbf{where} \quad M = \sum_i s_i$$

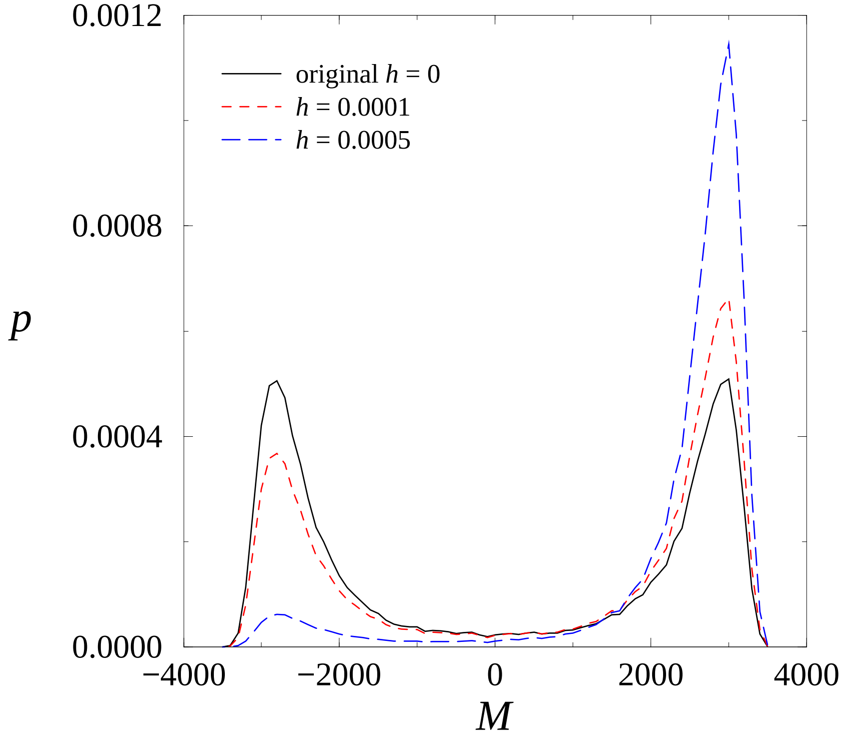
If the original simulation was performed with h_s, reweighting in h is completely analogous to eq. :

$$\langle O\rangle_h = \frac{\sum_i O_i\, e^{(h-h_s)M_i}}{\sum_i e^{(h-h_s)M_i}}$$

Thus, even if originally $h_s = 0$ (no field in the original simulation), we can reweight to finite external field.

Reweighting can be done simultaneously in β and h. (How? And why the original histogram reweighting is not practical in this case?)

Magnetization histogram $p(M)$, size 64^2, in simulation $h = 0$, and $\beta = \beta_c$



Let us consider now general reweighting wrt. arbitrary parameters g^a, which couple to action/energy function $S(\vec{g}; \{\phi\})$. The partition function is

$$Z_{\vec{g}} = \int [d\phi] \exp[-S(\vec{g}; \{\phi\})]$$

Simulation is performed with $g^a = g_0^a$, and we measure $S_i(\vec{g}_0) = S(\vec{g}_0; \{\phi\}_i)$ and $S_i(\vec{g}) = S(\vec{g}; \{\phi\}_i)$ (note that configs are generated with $S(\vec{g}_0; \{\phi\})$). Reweighting observable O:

$$\langle O \rangle_{\vec{g}} = \frac{\sum_i O_i \exp[-(S_i(\vec{g}) - S_i(\vec{g}_0))]}{\sum_i \exp[-(S_i(\vec{g}) - S_i(\vec{g}_0))]}$$

Often the action factorizes to form $S = \sum_a g^a S^a$. In this case it is sufficient to measure the “pieces” of action S^a, and we can reweight wrt. any component g^a.
$g^1 = -\beta$, $S^1 = E$ recovers standard Ising reweighting.

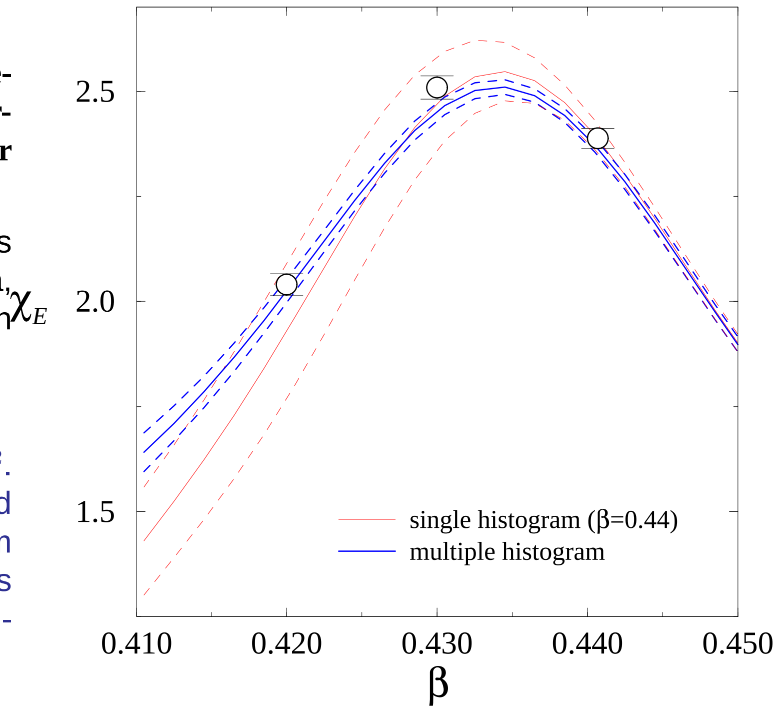
---

7 Multiple histogram reweighting

NOTE: this section is extra material, but included here for reference. This is still very useful for practitioners of the art!

Multiple histogram method joins together several sets of data, run at different β-values, in an optimized way.

Example:
Susceptibility C_V, volume 32^2. The data from 3 runs is joined together with multiple histogram reweighting (blue curve). Errors are much smaller than with single histogram reweighting.



Multiple histogram reweighting is a powerful method – should be part of standard toolbox for everybody who can join points from different simulations. However, somewhat cumbersome to use.

- Perform R Monte Carlo simulations, at couplings β_i, with length N_i.
- Measure the energy distributions $p_i(E) = H_i(E)/N_i$, and the autocorrelation time τ_i.
- True distribution is given by

$$p_i(E) = n(E)\, e^{-\beta_i E + f_i},$$

where $n(E)$ is the density of states (does not depend on β), and f_i is (dimensionless) free energy: $f_i = -\log Z_{\beta_i}$.

↦ Each of the simulations gives us an estimate of $n(E) = p_i(E)e^{\beta_i E - f_i}$ (we don't know f_i, so that the normalization is unknown). Since the MC runs were performed at different β_i, each of the run yields a reliable estimate of $n(E)$ only at limited range of E.

- **Optimization:** we obtain an improved estimate for $n(E)$ by combining all runs together:

$$n(E) = \sum_{i=1}^{R} r_i(E) p_i(E) e^{\beta_i E - f_i}\,, \qquad \text{where} \quad \sum_{i=1}^{R} r_i(E) = 1 \text{ for all } E.$$

Note that the relative weights $r_i(E)$ are independent at each E: $\rightarrow$ optimization.

- $r_i(E)$ are determined by minimizing the (error)^2 in $n(E)$ (now follows somewhat technical derivation):
  - What is the uncertainty in histogram values? Assuming that $H_i(E)$ is Poisson distributed around the 'true' value $\bar{H}_i(E)$, we obtain
  $$\delta^2 H_i(E) = g_i \bar{H}_i(E) = g_i N_i n(E) e^{-\beta_i E - f}\,.$$
  Here $g_i = 1 + 2\tau_i$ takes into account the autocorrelations in run i.
  - Thus, (error)^2 in $n(E)$
  $$\delta^2 n(E) = \sum_i r_i^2(E) \frac{\delta^2 H_i(E)}{N_i^2} e^{2(\beta_i E - f_i)} = \sum_i r_i^2(E) \frac{g_i n(E)}{N_i} e^{\beta_i E - f_i}$$
  - Minimize $\delta^2 n(E)$ wrt. $r_i(E)$ with condition $C \equiv \sum_i r_i(E) = 1$. Use Lagrange multipliers (try it):
  $$\frac{\partial}{\partial r_i(E)} [\delta^2 n(E) + \lambda C] = 0 \quad \mapsto \quad r_i(E) = \frac{N_i g_i^{-1} e^{-\beta_i E + f_i}}{\sum_{j=1}^R N_j g_j^{-1} e^{-\beta_j E + f_j}}$$

- Thus, the optimized expression for $n(E)$ is
$$n(E) = \frac{\sum_{i=1}^R g_i^{-1} H_i(E)}{\sum_{j=1}^R N_j g_j^{-1} e^{-\beta_j E + f_j}}$$
- The coefficients f_i are then determined by solving
$$e^{-f_i} = \sum_E n(E) e^{-\beta_i E}$$
To solve this equation use some iterative method (Newton-Raphson). f_i's are determined up to an additional constant.
- Observable expectation values:
$$\langle O \rangle_\beta = \frac{\sum_E O(E) n(E) e^{-\beta E}}{\sum_E n(E) e^{-\beta E}}$$

As with the single histogram method, this can be formulated without resorting to $H_i(E)$. Let E_i^a be the energy measurement number a ($a = 1 \ldots N_i$) from run number i. The expression for the free energy f_β becomes

$$e^{-f_\beta} = \sum_{i=1}^{R} \sum_{a=1}^{N_i} \frac{g_i^{-1} e^{-\beta E_i^a}}{\sum_{j=1}^R N_j g_j^{-1} e^{-\beta_j E_i^a + f_j}}\,.$$

f_i's are then solved by setting $e^{-f_i} = e^{-f_{\beta_i}}$, in analogy to the second eq. in the box above. The expectation value of O at reweighted β:

$$\langle O \rangle_\beta = \sum_{i=1}^{R} \sum_{a=1}^{N_i} \frac{O_i^a g_i^{-1} e^{-\beta E_i^a - f_\beta}}{\sum_{j=1}^{R} N_j g_j^{-1} e^{-\beta_j E_i^a + f_j}}$$