Computational Statistical Physics

Part I: Statistical Physics and Phase Transitions

Marina Marinkovic

March 16, 2022

ETH Zürich
Institute for Theoretical Physics
HIT G 41.5
Wolfgang-Pauli-Strasse 27
8093 Zürich

402-0812-00L
FS 2022

The Kasteleyn and Fortuin Theorem

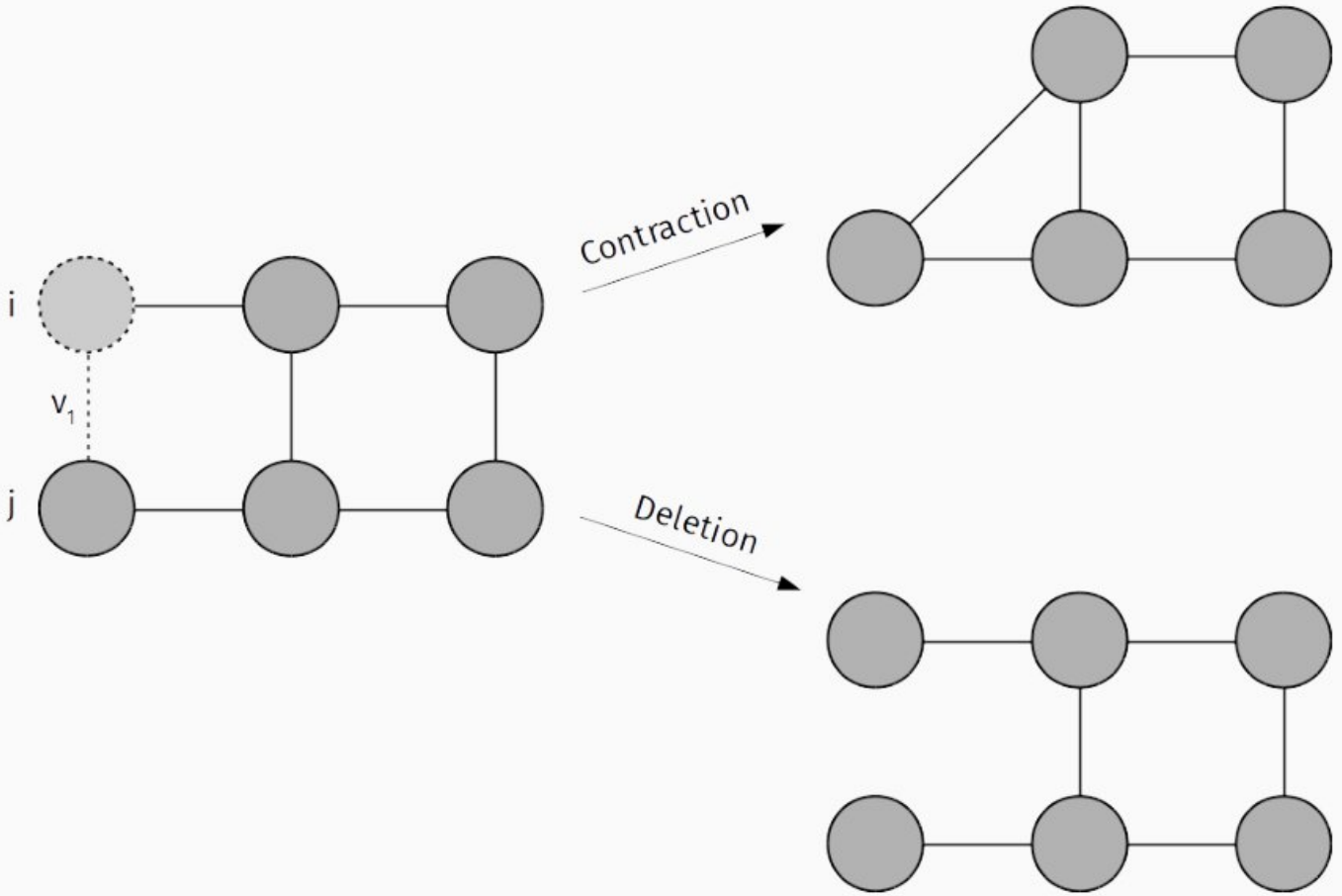

Figure 1: Contraction and deletion on a graph.

The Kasteleyn and Fortuin Theorem

The partition function is the sum over all the possible configurations weighted by the Boltzmann factor and thus given by

$$Z = \sum_X e^{-\beta E(X)} \overset{(??)}{=} \sum_X e^{-\beta J \sum_\nu \epsilon_\nu} = \sum_X \prod_\nu e^{-\beta J \epsilon_\nu}. \quad (1)$$

The Kasteleyn and Fortuin Theorem

We now consider a graph where bond ν_1 connects two nodes i and j with states σ_i and σ_j, respectively. If we would delete bond ν_1, the partition function is

$$Z_D = \sum_X \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu}. \tag{2}$$

The Kasteleyn and Fortuin Theorem

We can thus rewrite Eq. (1) as

$$Z = \sum_X e^{-\beta J \epsilon_{\nu_1}} \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu} = \sum_{X:\sigma_i = \sigma_j} \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu} + e^{-\beta J} \sum_{X:\sigma_i \neq \sigma_j} \prod_{\nu \neq \nu_1} e^{-}$$

where the first part is the partition function of the contracted graph Z_C and the second part is given by the identity

$$\sum_{X:\sigma_i \neq \sigma_j} \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu} = \sum_X \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu} - \sum_{X:\sigma_i = \sigma_j} \prod_{\nu \neq \nu_1} e^{-\beta J \epsilon_\nu} = Z_D - Z_C. \tag{3}$$

The Kasteleyn and Fortuin Theorem

Summarizing the latter results, we find

$$Z = Z_C + e^{-\beta J}\left(Z_D - Z_C\right) = pZ_C + (1-p)Z_D, \qquad (4)$$

where $p = 1 - e^{-\beta J}$. To be more precise, we expressed the partition function Z as the contracted and deleted partition functions at bond ν_1. We apply the latter procedure to another bond ν_2 and find

$$Z = p^2 Z_{C_{\nu_1},C_{\nu_2}} + p(1-p)Z_{C_{\nu_1},D_{\nu_2}} + (1-p)^2 Z_{D_{\nu_1},D_{\nu_2}}. \qquad (5)$$

The Kasteleyn and Fortuin Theorem

After applying these operations to every bond, the graph is reduced to a set of separated points corresponding to clusters of nodes which are connected and in the same state out of q states. The partition function reduces to

$$Z = \sum_{\substack{\text{configurations of} \\ \text{bond percolation}}} q^{\#\text{ of clusters}} p^c (1-p)^d = \left\langle q^{\#\text{ of clusters}} \right\rangle_{\mathsf{b}}, \quad (6)$$

where c and d are the numbers of contracted and deleted bonds respectively. In the limit of $q \to 1$, one obtains the partition function of bond percolation[1].

[1] In bond percolation [Broadbent, Hammersley (1957)], an edge of a graph is occupied with probability p and vacant with probability $1-p$.

Coniglio-Klein clusters

The probability of a given cluster C to be in a certain state σ_0 is independent of the state itself, i.e.,

$$p(C,\sigma_0) = p^{c_C}(1-p)^{d_C} \sum_{\substack{\text{bond percolation} \\ \text{without cluster C}}} q^{\#\text{ of clusters}} p^c (1-p)^d. \quad (7)$$

Coniglio-Klein clusters

This implies that flipping this particular cluster has no effect on the partition function (and therefore the energy) so that it is possible to accept the flip with probability one. This can be seen by looking at the detailed balance condition of the system

$$p(C,\sigma_1)W\left[(C,\sigma_1)\to(C,\sigma_2)\right] = p(C,\sigma_2)W\left[(C,\sigma_2)\to(C,\sigma_1)\right] \tag{8}$$

and using $p(C,\sigma_1) = p(C,\sigma_2)$.

Coniglio-Klein clusters

We then obtain for acceptance probabilities

$$W[(C,\sigma_2) \to (C,\sigma_1)] = \frac{p(C,\sigma_2)}{p(C,\sigma_1)+p(C,\sigma_2)} = \frac{1}{2} \quad \text{Glauber dyn. (9)}$$

$$W[(C,\sigma_2) \to (C,\sigma_1)] = \min\left[1, \frac{p(C,\sigma_2)}{p(C,\sigma_1)}\right] = 1 \quad \text{Metropolis (10)}$$

Based on these insights, we introduce cluster algorithms which are much faster than single-spin flip algorithms and less prone to the problem of critical slowing down.

Swendsen-Wang algorithm

Swendsen-Wang algorithm

- Occupy the bonds with probability $p = 1 - e^{-\beta J}$ if sites are in the same state.
- Identify the clusters with the Hoshen-Kopelman algorithm.
- Flip the clusters with probability 0.5 for Ising or always choose a new state for $q > 2$.
- Repeat the procedure.

Hoshen-Kopelman algorithm

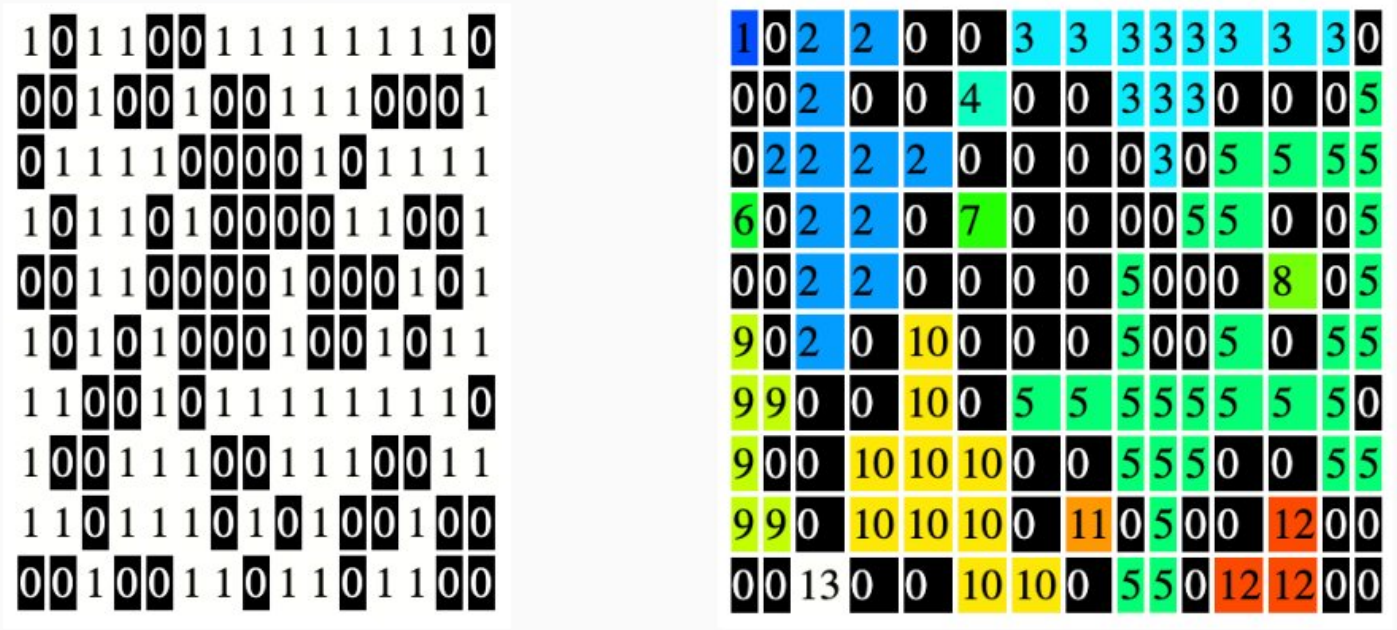

Figure 2: Left: spin configuration, 1's represent spin up; 0's are spin down. Right: the result of applying the HK algorithm to the grid on the left. Credit: `https://www.ocf.berkeley.edu/~fricke/projects/hoshenkopelman/hoshenkopelman.html`

Swendsen-Wang algorithm

Swendsen-Wang algorithm

- Occupy the bonds with probability $p = 1 - e^{-\beta J}$ if sites are in the same state.
- Identify the clusters with the Hoshen-Kopelman algorithm.
- Flip the clusters with probability 0.5 for Ising or always choose a new state for $q > 2$.
- Repeat the procedure.

Wolff algorithm

Wolff algorithm

- Choose a site randomly.
- If the neighboring sites are in the same state, add them to the cluster with probability $p = 1 - e^{-\beta J}$.
- Repeat this for any site on the boundaries of the cluster, until all the bonds of the cluster have been checked exactly once.
- Choose a new state for the cluster.
- Repeat the procedure.

Other Ising-like models

One of the possible generalizations of the Ising model is the so called n-vector model. Unlike the Potts model, it describes spins as vectors with n components. This model has applications in modelling magnetism or the Higgs mechanism. The Hamiltonian resembles the one of the Potts model in the sense that it favors spin alignment

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j + \vec{H} \sum_i \vec{S}_i. \tag{11}$$

with $\vec{S}_i = \left(S_i^1, S_i^2, \ldots, S_i^n\right)$ and $\left|\vec{S}_i\right| = 1$.

Other Ising-like models

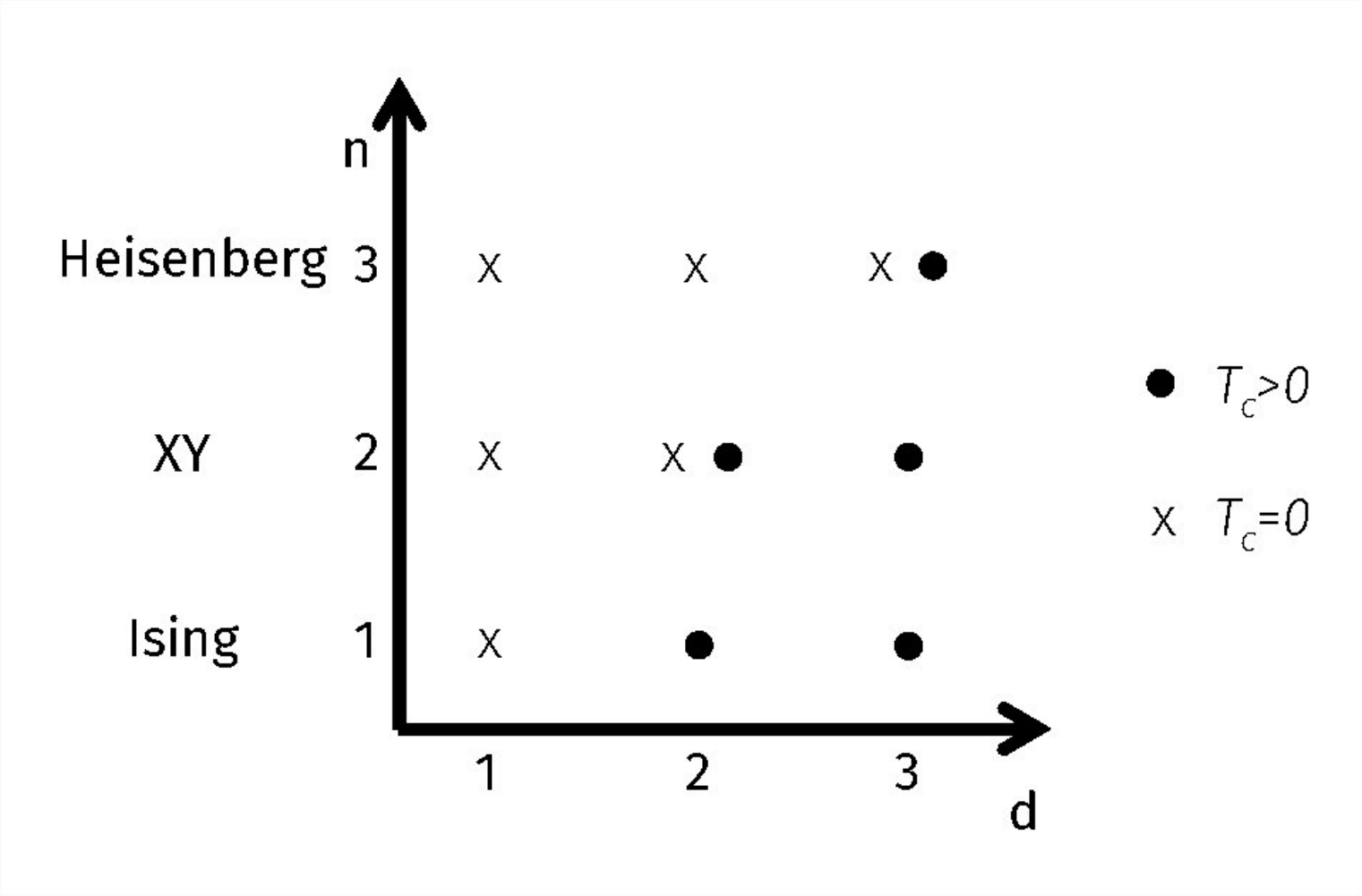

Figure 3: The dependence of the critical temperature on the number of vector components n.

Other Ising-like models

For Monte Carlo simulations with vector-valued spins we have to adapt our simulation methods. The classical strategy is to flip spins by modifying the spin locally trough adding a small $\Delta\vec{S}$ such that $\vec{S}_i' = \vec{S}_i + \Delta\vec{S}$ and $\Delta\vec{S} \perp \vec{S}_i$. The classical Metropolis algorithm can then be used in the same fashion as in the Ising model.

Histogram methods

Histogram methods

For computing the thermal average

$$\langle Q \rangle = \frac{1}{Z_T} \sum_X Q\left(X\right) e^{-\frac{E(X)}{k_B T}}. \tag{12}$$

we need to sample different configurations at different temperatures. Another possibility would be to determine an average at a certain temperature T_0 and extrapolate to another temperature T. In the case of a canonical ensemble, an extrapolation can be achieved by reweighting the histogram of energies $p_{T_0}(E)$ with the Boltzmann factor $e^{\frac{E}{T}-\frac{E}{T_0}}$.

Histogram methods

Such histogram methods have first been described by Salzburg et al. in 1959. We now reformulate the computation of the thermal average of a quantity Q and of the partition function as a sum over all possible energies instead of over all possible configurations and find

$$Q\left(T_0\right) = \frac{1}{Z_{T_0}} \sum_E Q\left(E\right) p_{T_0}\left(E\right) \quad \text{with} \quad Z_{T_0} = \sum_E p_{T_0}\left(E\right), \tag{13}$$

where $p_{T_0}\left(E\right) = g\left(E\right) e^{-\frac{E}{k_B T_0}}$ with $g\left(E\right)$ defining the *degeneracy of states*, i.e., the number of states with energy E.

Histogram methods

This takes into account the fact that multiple configurations can have the same energy. The goal is to compute the quantity Q at another temperature T

$$Q\left(T\right)=\frac{1}{Z_T}\sum_E Q\left(E\right)p_T\left(E\right). \tag{14}$$

The degeneracy of states contains all the information needed. Using the definition of $g\left(E\right)$ yields

$$p_T\left(E\right)=g\left(E\right)e^{-\frac{E}{k_BT}}=p_{T_0}\left(E\right)\exp\left[-\frac{E}{k_BT}+\frac{E}{k_BT_0}\right] \tag{15}$$

and with $f_{T_0,T}\left(E\right)=\exp\left[-\frac{E}{k_BT}+\frac{E}{k_BT_0}\right]$ we finally obtain

$$Q\left(T\right)=\frac{\sum_E Q\left(E\right)p_{T_0}\left(E\right)f_{T_0,T}\left(E\right)}{\sum_E p_{T_0}\left(E\right)f_{T_0,T}\left(E\right)}. \tag{16}$$

Histogram methods

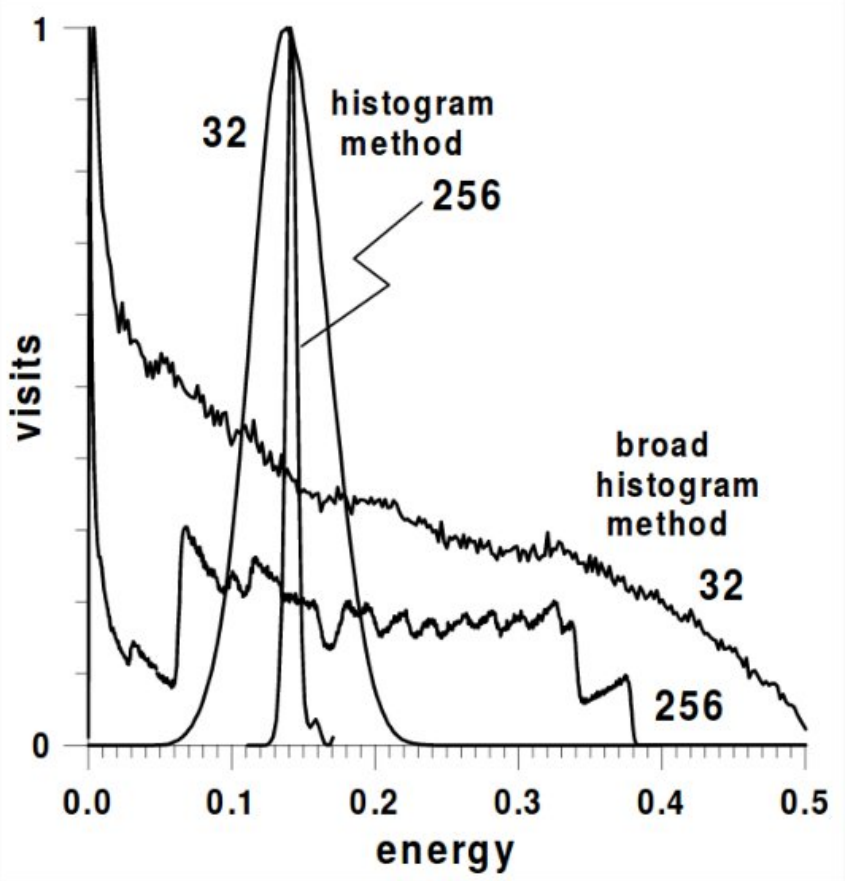

Figure 4: An example of the histogram and the broad histogram method for different system sizes. The figure is taken from [Oliveria et al. (1996)]

Broad histogram methods

Let N_{up} and N_{down} be the numbers of processes which lead to an increasing and decreasing energy, respectively. Furthermore, we have to keep in mind that the degeneracy of states increases exponentially with energy E since the number of possible configurations increases with energy. To explore all energy regions equally, we find a condition equivalent to the one of detailed balance, i.e.,

$$g\left(E+\Delta E\right) N_{\mathrm{down}}\left(E+\Delta E\right) = g\left(E\right) N_{\mathrm{up}}\left(E\right). \qquad (17)$$

Broad histogram methods

The motion in phase space towards higher energies can then be penalized with a Metropolis-like dynamics:

- Choose a new configuration,
- if the new energy is lower, accept the move,
- if the new energy is higher then accept with probability $\frac{N_{\mathrm{down}}(E+\Delta E)}{N_{\mathrm{up}}(E)}$.

Broad histogram methods

We obtain the function $g(E)$ by taking the logarithm of Eq. (17) and divide by ΔE

$$\log\left[g\left(E+\Delta E\right)\right]-\log\left[g\left(E\right)\right]=-\log\left[N_{up}\left(E\right)\right]-\log\left[N_{down}\left(E+\Delta E\right)\right]. \tag{18}$$

Broad histogram methods

In the limit of small energy differences, we can approximate the latter equation by

$$\frac{\partial \log\left[g\left(E\right)\right]}{\partial E} = \frac{1}{\Delta E} \log\left[\frac{N_{\mathrm{up}}\left(E\right)}{N_{\mathrm{down}}\left(E+\Delta E\right)}\right] \tag{19}$$

which we can numerically integrate to obtain $g\left(E\right)$.

Broad histogram methods

Distributions of N_{up} and N_{down} can be obtained by keeping track of these numbers for each configuration at a certain energy. In addition, we also need to store the values of the quantity $Q(E)$ we wish to compute as a thermal average according to

$$Q(T) = \frac{\sum_E Q(E)\, g(E)\, e^{-\frac{E}{k_B T}}}{\sum_E g(E)\, e^{-\frac{E}{k_B T}}}. \tag{20}$$

Based on a known degeneracy of states $g(E)$, we can now compute quantities at any temperature.

Flat histogram methods

Flat histogram method

- Start with $g(E) = 1$ and set $f = e$.
- Make a Monte Carlo update with $p(E) = 1/g(E)$.
- If the attempt is successful at E: $g(E) \leftarrow f \cdot g(E)$.
- Obtain a histogram of energies $H(E)$.
- If $H(E)$ is flat enough, then $f \leftarrow \sqrt{f}$.
- Stop when $f \leq 10^{-8}$.

Umbrella sampling

The Umbrella sampling technique was developed and proposed in [Torrie and Valleau, J. Comp. Phys., 1979].The aim is to overcome the problem of the lacking ergodicity for certain energy landscapes.

As an example, in the Ising model the system could have difficulties in jumping from a positive to a negative magnetization or vice versa if the system is very large.

Umbrella sampling

The basic idea is to multiply transition probabilities with a function that is large at the free energy barrier and later remove this correction in the averaging step.

$$\widetilde{p}\left(C\right) = \frac{w\left(C\right) e^{-\frac{E(C)}{k_B T}}}{\sum_C w\left(C\right) e^{-\frac{E(C)}{k_B T}}} \quad \text{with} \quad \left\langle A \right\rangle = \frac{\left\langle A/w \right\rangle_w}{\left\langle 1/w \right\rangle_w}. \qquad (21)$$

Summary of the Histogram Methods

Summarizing, some of the most common techniques related to the histogram methods are

- Wang-Landau method [Wang 2001, Zhou 2006]
- Multiple histogram method [Ferrenberg 1989]
- Multicanonical Monte Carlo [Berg 2000]
- Flat Histogram method [Wang 2001]
- Umbrella sampling [Torrie and Valleau, 1979]