

Monte Carlo Methods

Free Energy Monte Carlo Methods

Dieter W. Heermann

Heidelberg University

December 16, 2020

1. Introduction and Overview
2. Umbrella Sampling
3. Multicanonical Sampling and other “extended ensemble” methods
4. Wang-Landau Sampling
5. Transition Path Sampling

- Consider the canonical ensemble average of an observable $A(\vec{x})$, where \vec{x} stands symbolically for the “microstate” coordinate in the configurational part of the phase space of the system, are given by

$$\langle A(\vec{x}) \rangle_T = \frac{1}{Z} \int d\vec{x} \exp[-\mathcal{H}(\vec{x}/k_B T)] A(\vec{x}) , \quad (1)$$

where $\mathcal{H}(\vec{x})$ is the Hamiltonian of the considered many-particle system, and with $k_B =$ Boltzmann’s constant, $T =$ absolute temperature, and Z being the partition function which is related to the free energy F ,

$$Z = \int d\vec{x} \exp[-\mathcal{H}(\vec{x})/k_B T] \quad , \quad F = -k_B T \ln Z. \quad (2)$$

- By simple sampling Monte Carlo the integrals over the (very high-dimensional!) space $\int d\vec{x} \dots$ in Eqs.1,2 are replaced by sums over a sample of M randomly chosen points $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_M\}$, but we have seen that simple sampling works only if the number N of the particles (or degrees of freedom, such as Ising or Heisenberg spins) is extremely small.

- We have introduced the Importance Sampling Monte Carlo method, where the points $\{\vec{x}_\nu\}$ no longer are chosen completely at random, but preferentially from the important region of configuration space. Namely, choosing a state \vec{x}_ν with a probability proportional to the Boltzmann factor $\exp(-\mathcal{H}(\vec{x}_\nu)/k_B T)$, the average in Eq.1 is simply replaced by an arithmetic average over the M states generated:

$$\langle A(\vec{x}) \rangle_T \approx \overline{A(\vec{x})} = \frac{1}{M} \sum_{i=1}^M A(\vec{x}_i) \quad . \quad (3)$$

- We have seen that the Importance Sampling Monte Carlo method is very powerful, it allows a very large number of useful applications. However, when we compare Eqs.1 and 3, we see that some important information has been lost: Eq.3 has as a normalizing denominator no longer the partition function Z {Eq.2}, but rather simply the total number M of generated configurations.
- In fact, doing Importance Sampling Monte Carlo the knowledge of Z and hence F , as well as of the entropy S (remember $F = E - T S$ where the internal energy E is accessible as a thermal average of the Hamiltonian, of course, $E = \langle \mathcal{H}(\vec{x}) \rangle_T$) has been lost.

- When we study (in the thermodynamic limit, $N \rightarrow \infty$) the free energy per spin F/N of the Potts ferromagnet as a function of temperature (in zero field H), we observe that for temperatures T less than the transition temperature T_c the free energy of the (q -fold degenerate) ordered phase is lower than that of the (metastable) continuation of the free energy branch of the (nondegenerate) disordered phase, while for $T > T_c$ the disordered branch has the lower free energy.
- At T_c , the two branches $F_{\text{ord}}(T)$, $F_{\text{dis}}(T)$ are precisely equal, but the two branches meet there under different angles. Since we know from elementary thermodynamics that ($\beta \equiv 1/k_B T$)

$$E = (\partial F / \partial \beta)_H = -k_B T^2 \left(\frac{\partial F}{\partial T} \right)_H, \quad (4)$$

the slopes $\partial F_{\text{ord}} / \partial T|_T$, $\partial F_{\text{dis}} / \partial T|_T$ simply are related to the energies E_- , E_+ , $E_+ - E_-$ being the latent heat at the first order transition.

- So one would like to know the free energies of the ordered and disordered phases $F_{\text{ord}}(T)$, $F_{\text{dis}}(T)$ explicitly, since then the transition simply could be located from the condition $F_{\text{ord}}(T_c) = F_{\text{dis}}(T_c)$. The “poor man’s recipe” to achieve this has been based on a method called “thermodynamic integration” (6.1; 6.2). From Eq.4 we readily recognize that free energy differences can be computed from

$$\Delta F \equiv F_2 - F_1 = \int_{\beta_1}^{\beta_2} E(\beta) d\beta = \int_{\beta_1}^{\beta_2} \langle \mathcal{H} \rangle_{\beta} d\beta \quad . \quad (5)$$

- In order to carry out the integral in Eq.5 numerically with sufficient accuracy, one needs to compute $\langle \mathcal{H} \rangle_{\beta}$ at a large number of inverse temperatures β intermediate between β_1 and β_2 (and this need may be cumbersome: that’s why we call this approach a “poor man’s recipe”).
- In order to get an absolute free energy, one needs to know the free energy F_1 of a reference state. For the Potts ferromagnet in the disordered phase, a convenient reference state is the completely random state at $\beta_1 = 0$, where $E(0) = 0$ and $S(\beta = 0) = Nk_B \ln q$ is trivially known.
- For the ordered phase, a similarly simple reference state would be $\beta_2 \rightarrow \infty$, where the system is fully ordered, E coincides with the ground state energy, and $S = k_B \ln q$ (without a factor N : we just have a q -fold degenerate ground state).

- In practice it suffices to choose a large but finite β_2 as a reference state, for which the deviation of $E(\beta_2)$ from the ground state energy is negligibly small.
- Although we have called this thermodynamic integration approach a “poor man’s recipe”, one should not be misled: there are cases where this simple-minded approach actually is the method of choice, superior in accuracy to both finite size scaling approaches and the sampling techniques that we shall describe in the present chapter.
- Even though thermodynamic integration can (at least in favorable cases) yield information on the free energies of the two phases that coexist at the transition, it does not yield any information on the barrier between these phases in configuration space. This free energy barrier is responsible for the hysteresis that very often is observed in simulations and also in experiments!) near first-order transitions.
- In many contexts, it is of great interest to gain information on a suitable path in a (suitably coarse-grained) “free energy landscape” that connects the free energy minima that correspond to the coexisting phases with each other.
- When one discusses $\mathcal{H}(\vec{x})$ as a function of the configuration space coordinate \vec{x} one may speak of a “potential energy landscape”, but clearly the considered space is extremely high-dimensional, and the landscape in this space is presumably extremely “rugged”. As always in statistical physics, one is more interested in a reduced description, referred to as “coarse-graining” above.

- The most familiar concept of coarse-graining uses the idea of an “order parameter” as introduced by Landau (6.1).
- In simple cases, like Ising ferromagnets or fluids that undergo a vapor-liquid phase transition, this order parameter is a scalar quantity (the magnetization $\langle m \rangle_{T,H}$ per spin in the case of a ferromagnet, or the density $\rho = \langle N \rangle_T / V$ in the case of a fluid occupying a volume V).
- Rather than discussing the full canonic or grandcanonic probability distribution

$$p_{\text{can}}(\vec{x}) = (1/Z) \exp[-\mathcal{H}(\vec{x})/k_B T]$$

or

$$p_{g,c}(\vec{x}) = (1/Y) \exp[(\mu N - \mathcal{H}(\vec{x}))/k_B T] \quad , \quad (6)$$

where $Y(T, V, \mu) = \sum_N \exp(\mu N/k_B T) Z_N(T, V)$ is the grand-canonical partition function and μ the chemical potential, one focusses attention to a reduced description, e.g. in terms of the distribution $p(\rho)$ of the density only,

$$p(\rho) = (1/Y) \sum_N \int d\vec{x} \exp[(\mu N - \mathcal{H}(\vec{x}))/k_B T] \delta(\rho - N/V) \quad . \quad (7)$$

Note that $p(\rho)$ can be interpreted in terms of a free energy $f(\rho)$ per unit volume as

$$p(\rho) = (1/Y) \exp[-Vf(\rho)/k_B T] \quad (8)$$

Fig.1 presents a plot of $f(\rho)$ vs. ρ for a Lennard-Jones model of a fluid at a temperature $T = 0.68 T_c$ and choosing the chemical potential μ such that in the bulk (volume $V = L^3 \rightarrow \infty$) coexistence between vapor and liquid occurs, $\mu = \mu_{\text{coex}}(T)$ (6.7; 6.8).

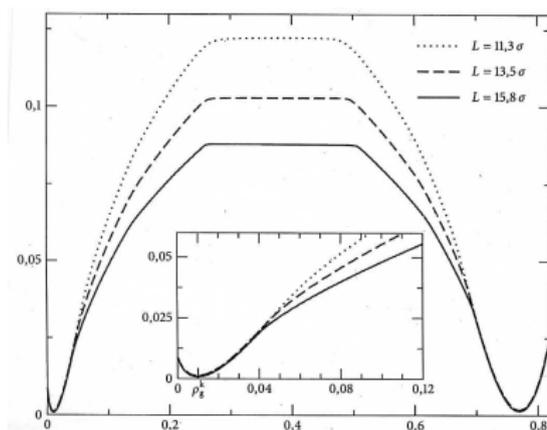


Figure: Free energy per unit volume $f(\rho)$ plotted vs. density ρ , for a Lennard-Jones model of a fluid (the potential being truncated at $r_c = 2.2^{1/6}\sigma$, with the Lennard-Jones parameter σ being the unit of length, and shifted to zero there). The data refer to three choices of cubic $L \times L \times L$ simulation boxes, with periodic boundary conditions, a temperature $T = 0.68T_c$, and the chemical potential being chosen at two-phase coexistence, $\mu = \mu_{\text{coex}}$, so the two minima are equally deep. Insert shows a blow-up of the region near the bulk gas density at coexistence, ρ_g .

- These data have been obtained by the technique of “successive umbrella sampling” (6.9), that will be described in the following section.

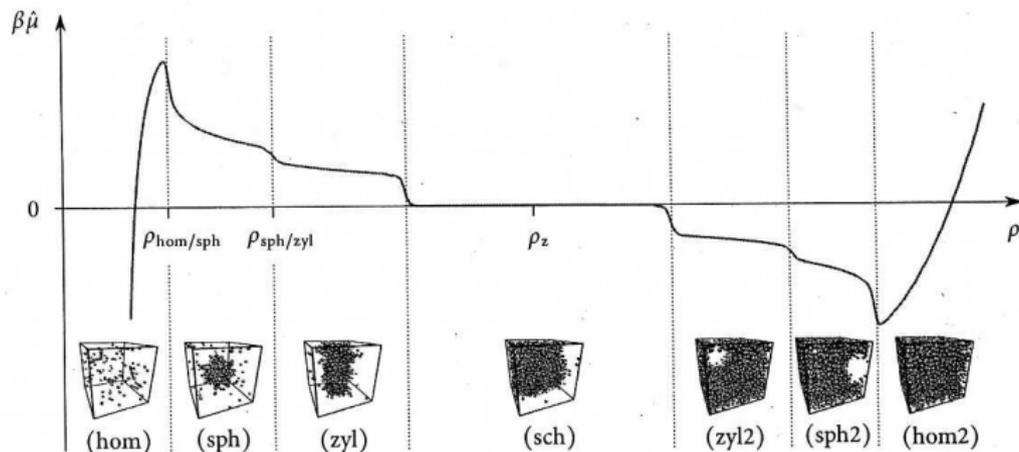


Figure: Plot of the reduced chemical potential difference $\beta\hat{\mu}$ vs. ρ , for the system with $L = 15.8\sigma$ from Fig. 6.1. Typical snapshot pictures of system configurations illustrate the evolution of the system from homogeneous vapor to homogeneous liquid via a sequence of inhomogeneous states when one crosses the free energy barrier of Fig. 1 when ρ is increased. For further explanations cf. text.

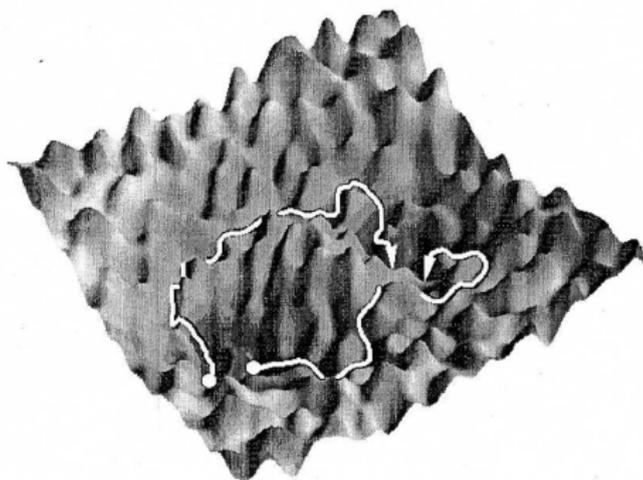


Figure: Schematic picture of a “rugged free energy landscape”, plotted as a function of two order parameter coordinates ψ_x, ψ_y . The two white lines indicate two possible trajectories from the region of one deep minimum to another one. From Dellago et al. (6.28)

- Such “rugged free energy landscapes” (6.29) occur in such diverse systems as domain patterns in random ferromagnets, diluted magnets with competing interactions (“spin glasses” (6.30), undercooled fluids near the glass transition (6.31), and polymers (in particular proteins) under bad solvent conditions (6.32; 6.33; 6.34; 6.35; 6.36; 6.37; 6.38; 6.39; 6.40).
- In all these systems the step analogous to the step from Eqs.6 to Eq. 8 for the simple fluid, where one focuses on a reduced description in terms of a simple “order parameter”, is not at all straightforward.
- In models for spin glasses, this problem has been circumvented by studying two “real replica”, copies of the same system with a specific choice of random exchange interactions, and defining an order parameter q as the (normalized) projection of the spin configurations in the replicas (1), (2) onto each other, $q = (1/N) \sum_{i=1}^N S_i^{(1)} S_i^{(2)}$ (in the case of Ising spins, $S_i^{(\alpha)} = \pm 1$) (6.30; 6.31).
- The sampling of the distribution $P(q)$, analogous to the distribution $P(\rho)$ in Eqs. (7), (8), nevertheless is a challenging computational problem, since in addition to the thermal averaging (with the Boltzmann factor, cf. Eq. (7)) there is the need to carry out an average $[\dots]_{\text{av}}$ over the quenched disorder: i.e., the distribution $P_{\{J_{ij}\}}(q)$ obtained for one particular choice $\{J_{ij}\}$ of the random bonds needs to be averaged over (typically) several hundred realizations of such random bond configurations (6.30; 6.41; 6.42; 6.43; 6.44; 6.45; 6.46).

- Given a rugged free energy landscape as schematically shown in Fig. 3, it is a nontrivial question which path the system will take in order to reach equilibrium. In order to describe the kinetics of such thermally activated processes in complex free energy landscapes, a Monte Carlo sampling over an ensemble of trajectories of such processes is desirable, and this can be achieved by the “transition path sampling”-technique (6.28; 6.77; 6.78; 6.79; 6.80; 6.81; 6.82; 6.83) and its variants to sample rare events (“string method”, “metadynamics”, etc.) (6.84; 6.85; 6.86; 6.87; 6.88; 6.89; 6.90). As Bolhuis et al. (6.80) put it, transition path sampling can be viewed as “throwing ropes over mountain passes in the dark”.
- Of course, the sampling of trajectories of processes requiring thermal activation such as crossing barriers in nucleation events via Monte Carlo methods has its sound theoretical basis in the dynamic interpretation of Monte Carlo sampling (6.91). In fact, the first Monte Carlo simulation studies of nucleation processes in the nearest-neighbor kinetic Ising model were performed decades ago (6.92; 6.93).
- However, this naive (i.e., unbiased) sampling of trajectories over free energy barriers is efficient only if these barriers are not higher than a few $k_B T$. Via transition path sampling one creates trajectories that are biased such that higher free energy barriers (of the order of 20 to 100 $k_B T$) are crossed, and thus very valuable insight into the kinetics of nucleation processes (e.g. (6.94)) can be gained.

- We now return to the problem of calculating free energy differences between two states 1,2, but we now assume these systems have different Hamiltonians $\mathcal{H}_1(\vec{x})$, $\mathcal{H}_2(\vec{x})$. From Eq. (2) we can immediately derive that

$$\Delta F = F_2 - F_1 = -k_B T \ln(Z_2/Z_1) = -k_B T \ln \left(\frac{\int d\vec{x} \exp[-\beta \mathcal{H}_2(\vec{x})]}{\int d\vec{x} \exp[-\beta \mathcal{H}_1(\vec{x})]} \right) . \quad (9)$$

- We first use Eq. 9 to justify the “overlapping distribution method” (6.96).
- Suppose that we carry out an importance sampling study of system 2. For every configuration space point \vec{x} that is generated in the resulting Markov chain we can compute also the potential energy $\mathcal{H}_1(\vec{x})$ of system 1 (we focus here on the case that both systems have exactly the same volume V and same particle number N).
- We then can obtain the probability of $\Delta E = \mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x})$ as

$$p_2(\Delta E) = (1/Z_2) \int d\vec{x} \exp[-\beta \mathcal{H}_2(\vec{x})] \delta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}) - \Delta E] . \quad (10)$$

- Due to the delta function, Eq. (10) can be rewritten as

$$\begin{aligned} p_2(\Delta E) &= (1/Z_2) \exp(-\beta\Delta E) \int d\vec{x} \exp[-\beta\mathcal{H}_1(\vec{x})] \delta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}) - \Delta E] \\ &= (Z_1/Z_2) \exp(-\beta\Delta E) p_1(\Delta E) \end{aligned} \quad (11)$$

- In full analogy to Eq. 10, $p_1(\Delta E)$ is the probability density of finding a potential energy density ΔE between systems 1 and 2, when one samples the states of the system 1.
- Using Eq. 9, we obtain

$$\ln p_2(\Delta E) = \beta(\Delta F - \Delta E) + \ln p_1(\Delta E) \quad . \quad (12)$$

- From Eq. (12) we see that the two functions $\ln p_2(\Delta E)$, $\ln p_1(\Delta E) - \beta\Delta E$ should be identical, apart from a constant offset $\beta\Delta F$.
- Thus, if there exists a range of ΔE where the two distributions $\ln p_2(\Delta E)$, $\ln p_1(\Delta E) - \beta\Delta E$ have sufficient overlap, one can obtain $\beta\Delta F$ from a best fit.
- Of course, in practice this method can work only if the two systems differ only very little (e.g., two Ising systems at slightly different values of the magnetic field H).

- Another consequence of Eq. (11) is seen when one integrates this equation over ΔE from $-\infty$ to $+\infty$. Since p_2 is normalized to unity, this yields

$$1 = \exp(\beta\Delta F) \int_{-\infty}^{+\infty} d\Delta E \exp(-\beta\Delta E) p_1(\Delta E) = \exp(\beta\Delta F) \langle \exp(-\beta\Delta E) \rangle_1$$

i.e.

$$\exp(-\beta\Delta F) = \langle \exp\{-\beta(\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}))\} \rangle_1 \quad (13)$$

- Here $\langle \dots \rangle_1$ means that averages are taken in a sampling with system 1. Again Eq. (13) is practically useful if the two systems 1,2 differ only very little.
- The idea of “umbrella sampling” (6.47; 6.48) now is to enhance the overlap between the distributions $p_1(\Delta E) \exp(-\beta\Delta E)$, $p_2(\Delta E)$ by carrying out a biased sampling with a suitable chosen weight function $\Pi(\vec{x})$.

- Biased sampling is derived by rewriting Eq. (1) as follows

$$\begin{aligned} \langle A(\vec{x}) \rangle_0 &= \int d\vec{x} \Pi(\vec{x}) \exp[-\beta \mathcal{H}_0(\vec{x})] A(\vec{x}) / \Pi(\vec{x}) \int d\vec{x} \Pi(\vec{x}) \exp[-\beta \mathcal{H}_0(\vec{x})] / \Pi(\vec{x}) \\ &= \langle A(\vec{x}) / \Pi(\vec{x}) \rangle_{0, \Pi} / \langle 1 / \Pi(\vec{x}) \rangle_{0, \Pi} \quad , \end{aligned} \quad (14)$$

where the index Π at $\langle \dots \rangle_{0, \Pi}$ is a reminder that the weight contains an additional factor $\Pi(\vec{x})$ in the statistical weight for the states \vec{x} , in addition to the Boltzmann factor.

- We use this to rewrite Eq. (13) as follows

$$\begin{aligned} \langle \exp\{-\beta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x})]\} \rangle_1 &= \frac{\int d\vec{x} \Pi(\vec{x}) \exp[-\beta \mathcal{H}_2(\vec{x})] / \Pi(\vec{x})}{\int d\vec{x} \Pi(\vec{x}) \exp[-\beta \mathcal{H}_1(\vec{x})] / \Pi(\vec{x})} \\ &= \langle \exp[-\beta \mathcal{H}_2(\vec{x})] / \Pi(\vec{x}) \rangle_{\Pi} / \langle \exp[-\beta \mathcal{H}_1(\vec{x})] / \Pi(\vec{x}) \rangle_{\Pi} \quad . \end{aligned} \quad (15)$$

- Here Eq. (14) is used for the special case $\mathcal{H}_0(\vec{x}) = 0$.
- In order to be able to sample both the numerator and the denominator with acceptable accuracy, the “bridging distribution” $\Pi(\vec{x})$ must have sufficient overlap with both the important regions of configuration space of systems 1 and 2.
- This bridging property of $\Pi(\vec{x})$ is responsible for the name umbrella sampling.

- In practice, it is often advantageous to work not only with a single bridging distribution $\Pi(\vec{x})$, but with many distributions intermediate between $\frac{1}{Z_1} \exp[-\beta\mathcal{H}_1(\vec{x})]$ and $\frac{1}{Z_2} \exp[-\beta\mathcal{H}_2(\vec{x})]$.
- Suppose there exists a parameter $\Delta\psi$ that distinguishes between \mathcal{H}_2 and \mathcal{H}_1 .
- Then it often is useful to divide the interval $\Delta\psi$ in n steps $\Delta\psi/n$.
- The optimum choice of n , however, in the general case is a subtle matter (6.95).
- Another important drawback is that the function $\Pi(\vec{x})$ is not known a priori, it must be guessed, using the knowledge on the Boltzmann weights of the systems 1 and 2.
- However, there are situations where the implementation of the method is rather straightforward.
- One such scheme is the technique of Virnau and Müller (6.9) termed “successive umbrella sampling”, to which we turn now.
- This method has the advantage that by extrapolation of results from one “window” to the next (adjacent) window one can estimate the weight function rather simply, and one also can obtain rather reliable error estimates for this procedure.
- This algorithm has been used for the study of liquid-vapor transitions (6.25; 6.27) and of liquid-liquid phase separation (6.26; 6.97).

- For simplicity, we follow (6.9) using the language appropriate for a liquid-vapor transition, and simply take the number of particles N as the order parameter.
- We are interested in the distribution $P(N)$ for a given choice of variables V, T and μ in the grandcanonical ensemble, at temperatures T distinctly lower than the critical temperature.
- Then $P(N)$ has two peaks, one for particle numbers close to $N_v = V\rho_v$ and another close to $N_\ell = V\rho_\ell$, where ρ_v, ρ_ℓ are the densities of coexisting vapor and liquid in the thermodynamics limit, provided μ is close to μ_{coex} , and in between N_v, N_ℓ there occurs a region with a deep minimum {precisely at μ_{coex} we know that $P(N) \propto \exp[-Vf(\rho)/k_B T]$, with $f(\rho)$ shown in Fig. 1, for example}.
- However, to sample the distribution efficiently it is better to sample the re-weighted distribution $P_{\text{sim}}(N)$

$$P_{\text{sim}}(N) = P(N) \exp[-w(N)] \quad (16)$$

where in the optimum case $w(N) = \ln P(N)$, because then $P_{\text{sim}}(N)$ would be perfectly flat.

- We note that this is the same idea as used in multicanonical sampling (see Sec. 6.3.) and Wang-Landau sampling (see Sec. 6.4), of course.
- However, $P(N)$ initially is completely unknown.

- The strategy (6.9) now is to divide the region of all particle numbers from $N = 0$ to some $N = N_{\max}$ into m overlapping windows of width ω , and investigate one small window after the other.
- A histogram $H_K(N)$ monitors how often each state N is visited in the window $[k\omega, (k+1)\omega]$,
- We denote the values of the k 'th histogram at its left and right boundary by $H_{k\ell} \equiv H_k(k\omega)$, $H_{kr} \equiv H_k[(k+1)\omega]$, and define the ratio r_k as

$$r_k = H_{kr}/H_{k\ell} \quad . \quad (17)$$

- After a predetermined number of Monte Carlo steps per window, the (unnormalized) probability distribution is estimated recursively

$$P(N)/P(0) = (H_{0r}/H_{0\ell})(H_{1r}/H_{1\ell}) \cdots (H_k(N)/H_{k\ell}) = \prod_{i=1}^{k-1} r_i (H_k(N)/H_{k\ell}) \quad (18)$$

with $N \in [k\omega, (k+1)\omega]$. Probability ratios in this equation correspond to free energy differences.

- Care is needed at the boundaries of a window to fulfill detailed balance: if a move attempts to leave the window, it is rejected, and the previous state is counted once more for the histogram.

- The number of Monte Carlo steps (count of insertion or deletion trials) is not increased, however, as these moves neither contribute to the statistics.
- As one samples one window after the other, a weighted simulation amounts to replace $H_k(N)$ by $H_k(N) \exp[(w_k(N))]$, and the weight function $w_k(N)$ is estimated by extrapolation: After $P(N)$ has been estimated, $w(N)$ is extrapolated quadratically into the next window.
- The first window is usually unweighted.

- Let us consider the problem of thermally driven first order transitions, as they occur for instance in the case of the well-know q -state Potts model (6.98).
- The probability distribution $P_L(E)$ of the internal energy in a simulation using lattices of size $L \times L$ ($d = 2$) or $L \times L \times L$ ($d=3$ dimensions) in the transition region has two rather sharp peaks, separated by a deep minimum in between.
- When the peaks have equal weight (6.99; 6.100) (“equal weight rule”) the transition from the ordered phase (stable for $T < T_c$) to the disordered phase (stable for $T > T_c$) occurs, giving rise to a rather sharp peak in the specific heat, that develops into a delta function (representing the latent heat of the transition) as $L \rightarrow \infty$.
- It is very difficult to obtain accurate data on this specific heat peak by naive Monte Carlo sampling, because for large L the transitions between the two peaks of $P(E)$ occur very rarely, but it is necessary to sample many such transitions to estimate the relative weights for these peaks.
- Using a single histogram reweighting (6.101) and multiple histogram reweighting (6.102) the accuracy can be significantly improved, since Monte Carlo data yielding a distribution $P_L(E)$ for one inverse temperature β can be reweighted to neighboring temperatures in an interval $\Delta\beta \propto 1/\sqrt{N}$ (where $N = L^d$ here).
- Nevertheless, if the sampled distribution has a double-peak structure, the problem of estimating correctly the relative weights of the peaks remains.

- The multicanonical Monte Carlo method (6.49; 6.50; 6.51; 6.52; 6.53; 6.54) addresses this problem noting that

$$P_L(E) = g_L(E) \exp(-\beta E) \quad (19)$$

and suggesting to choose a reweighting with the weight function

$$w_{\text{muca}}(E) \propto 1/g_L(E) \quad \text{for} \quad E_{\min} \leq E \leq E_{\max} \quad , \quad (20)$$

where the interval $[E_{\min}, E_{\max}]$ is the range of energies of interest for the considered transition.

- Of course, if $g_L(E)$ were known, the problem would be solved, the specific heat could be calculated from $g_L(E)$ analytically, no Monte Carlo sampling would be required.
- But $g_L(E)$ is not known, one can only try to find a “working approximation” (6.103) of the weight function $w_{\text{muca}}(E)$.
- After such a function has been constructed, one performs an actual Monte Carlo simulation with this function, to obtain a final estimate of $g_L(E)$.
- Then one can use Eq. (19) to obtain averages at the desired values β of interest.

- The drawback of multicanonical Monte Carlo is that finding a suitable approximate weight function is not straightforward (note that the problem is essentially the same as in standard umbrella sampling: this is no surprise at all, since multicanonical Monte Carlo can be viewed as a reformulation of umbrella sampling (6.64; 6.95; 6.103)).
- For systems with rugged free energy landscapes, finite size scaling methods for estimating weight factors $w_{\text{muca}}(E)$ often do not work; e.g., for biomolecules one particular size is only of interest, or for spin glasses, the weights change too much when the system size is increased.
- If one does not want to rely on “ad hoc per hand estimates” (6.103), one must try to use a systematic recursive method (6.103; 6.104; 6.105).

- Writing the weight of state α as

$$w(\alpha) = \exp[-b(E_\alpha)E_\alpha + a(E_\alpha)] \equiv [-S(E_\alpha)] \quad , \quad (21)$$

where S can be interpreted as the associated entropy, and assuming that the energy spectrum is discrete (step size ε), one can derive a recursion (6.54)

$$b_{n+1}(E) = b_n(E) + \hat{g}_n(E)[\ln H_n(E + \varepsilon) - \ln H_n(E)]/\varepsilon \quad (22)$$

$$\hat{g}_n(E) = g_n(E)/[h_n(E) + g_n(E)] \quad , \quad (23)$$

$$g_n(E) = H_n(E + \varepsilon)H_n(E)/[H_n(E + \varepsilon) + H_n(E)] \quad , \quad (24)$$

$$h_{n+1}(E) = h_n(E) + g_n(E), \quad h_0(E) = 0 \quad , \quad (25)$$

$H_n(E)$ being the histogram entry at the n 'th step of the recursion.

- A much more straightforward sampling strategy is to create an extended ensemble, which is known as “parallel tempering” or “replica exchange Monte Carlo” (6.65; 6.66; 6.68).
- One performs a set of m canonical Monte Carlo simulations at different values $\beta_i, \dots i = 0, \dots m - 1$, with $\beta_0 < \beta_1 < \dots < \beta_{m-2} < \beta_{m-1}$.

- For these runs standard Boltzmann weight factors $\exp[-\beta_i \mathcal{H}(\vec{x})]$ are used.
- From time to time one attempts to exchange neighboring β values,

$$\beta_i \rightarrow \beta_{i-1}, \quad \beta_{i-1} \rightarrow \beta_i, \quad i = 1, \dots, n-1 \quad (26)$$

at fixed configuration of the systems, just treating the exchange in Eq. (26) as an additional type of Monte Carlo move, which is subjected to the standard importance sampling acceptance criterion.

- Of course, the set of values $\{\beta_i\}$ has to be chosen such that a reasonably large acceptance rate results.
- Obviously, this requires that the distributions $P_L(E)$ for β_i and β_{i-1} have a strong overlap.
- Thus, the number m of neighboring temperatures must scale like $m \propto \sqrt{N}$ as $N \rightarrow \infty$, but nevertheless the method is useful, particularly for systems such as glasses and spin glasses (6.29; 6.31; 6.41; 6.42; 6.43; 6.44; 6.45; 6.46).

- The Wang-Landau algorithm (6.56; 6.57; 6.58; 6.59; 6.60; 6.61; 6.62; 6.63) is an elegant iteration method to construct directly the energy density of states, $g(E)$.
- The idea is to perform a random walk in energy space with a flat histogram.
- One can also generalize this method to other variables (e.g. sample for an Ising ferromagnet $g(E, M)$ where M is the magnetization), but such generalizations will not be considered here.
- Unlike histogram methods where one extracts estimates of $g(E)$ from probability distributions generated by standard Monte Carlo simulations, temperature plays no role in this algorithm at all!
- Transitions between microstates (e.g., spin flips in an Ising model) are carried out according to probability

$$p(E \rightarrow E') = \min\{g(E)/g(E'), 1\} \quad (27)$$

- If $g(E)$ were known in beforehand, a Markov process based on Eq. (27) would generate a flat histogram.
- Since $g(E)$ is not known in beforehand, an iteration process needs to be performed to construct $g(E)$, starting from a simple initial guess.
- In the absence of any a priori knowledge on $g(E)$, it is natural to choose as an initial density of states simply $g(E) = 1$, for all E .

- For simplicity, we restrict here attention to a case where the energy spectrum is discrete and bounded ($E_{\min} \leq E \leq E_{\max}$), as it is the case for an Ising model or a lattice model for a polymer chain, but we emphasize at the outset that such a restriction is not necessary, and many successful applications of Wang-Landau sampling to off-lattice models (like fluids or polymer chains with Lennard-Jones interactions) have been given (6.61).
- Since the actual $g(E)$ is very different from $g(E) \equiv 1$, it is necessary to have a procedure that leads quickly away from this initial condition.
- This is achieved by replacing $g(E)$ by $g(E)$ times f , where the initial guess of the modification factor f_1 is $f_1 = e^1$, whenever a state with energy E is visited.
- During this random walk in energy space controlled by Eq. (27), one accumulates a histogram $H(E)$; i.e. starting out with $H(E) = 0$ for all E , one replaces $H(E)$ by $H(E) + 1$ whenever E is visited.
- The moves that are carried out to realize Eq. (27) are just the standard moves that one also would use in a standard Monte Carlo simulation that attempts to sample the Boltzmann distribution (e.g., spin flips of an Ising model, random displacements of an effective monomer to a new position in a model for a polymer chain, etc.)
- This sampling process of the histogram is continued, until the histogram $H(E)$ is “reasonably flat”.

- In practice, it has turned out that a useful criterion for this “flatness” is to require that the minimum entry in the histogram is not smaller than 80% of its mean value (6.56; 6.57; 6.58).
- When in the first iteration step this flatness has been achieved, one resets all histogram entries $H(E)$ to zero, and performs a sampling with a modification factor $f_2 = \sqrt{f_1}$ (and in the i 'th step, one chooses $f_i = \sqrt{f_{i-1}}$).
- Otherwise the procedure is identical to the procedure followed in the first step.
- This procedure is iterated, until f_i reaches a minimum value, which in practice is chosen to be of order $f_{\min} \approx \exp(10^{-8})$.
- While the detailed balance principle clearly is violated in the early stages of the iteration, this is no problem in the final run.
- The final density of states $g(E)$ then is used to calculate averages, e.g. the specific heat

$$C(T) = (\langle E^2 \rangle_T - \langle E \rangle_T^2) / (k_B T^2) \quad (28)$$

where moments $\langle E^k \rangle$ are straightforwardly computed from $g(E)$ as

$$\langle E^k \rangle = \sum_E E^k g(E) \exp(-\beta E) / \sum_E g(E) \exp(-\beta E) \quad (29)$$

- Of course, it is possible to use the run with the final choice f_{\min} to obtain more detailed information.
- E.g., in the study of Ising models one would like to estimate moments of the magnetization distribution $\langle m^k \rangle_T$.
- Sampling of the full (two-dimensional) joint density of states $g(E, m)$ would solve this problem, but for large systems this would require a huge computational effort.
- Typically it is sufficient (and much easier) to only take “microcanonical” averages, $\langle |m|^k \rangle_E$, from which then the desired canonical averages follow,

$$\langle |m|^k \rangle_T = \frac{\sum_E \langle |m|^k \rangle_E g(E) \exp(-\beta E)}{\sum_E g(E) \exp(-\beta E)} \quad , k = 1, 2, 3, \dots . \quad (30)$$

- Here we return to the problem already alluded to in Fig. 3. Suppose the system is stable or metastable in the basin near the starting points (white dots) of the two trajectories, i.e. typically it will stay there for a very long time.
- Only very rarely it will follow a path over high saddle points in a complex free energy landscape to make a transition into another deep basin (arrows).
- The aim of transition path sampling then is to gain information on all probable path ways, in order to analyze the transition mechanism. E.g., in the context of chemical reactions one would like to identify the “reaction coordinate” (6.77); in the case of crystal nucleation from the melt, one wants to get information on the crystal, structure size, and shape of the nucleus that forms (6.110); etc.
- This “reaction coordinate” need not be a scalar, but may contain several variables: e.g., in nucleation from an un-mixing binary fluid mixture one expects that both the size of the droplet and its composition change during the growth of the droplet from subcritical to supercritical.
- Since the pathways collected with transition path sampling are trajectories consistent with the Markovian master equation description of Monte Carlo sampling, the full information on the kinetics of the transition can be extracted (6.83).
- Transition path sampling can be adapted to cases where the underlying dynamic simulation method is not Monte Carlo but Molecular Dynamics or Brownian Dynamics (6.83).

- However, also in these cases the weight of the trajectories in the “transition path ensemble ” is obtained by Monte Carlo methods.
- The idea is to carry out a random walk in the space of trajectories rather than in configuration space.
- The basic step generates a new path $\{\bar{x}^{(n)}(\{t\})\}$ from an old one $\{\bar{x}^{(0)}(\{t\})\}$.
- The underlying dynamics of the system that is simulated defines a “path ensemble” $P_{AB}\{\bar{x}(\{t\})\}$ where we denote the initial state as A and the final state as B.
- The initial condition is prepared by placing the system in a heat bath at temperature T , and hence the distribution of initial conditions is just the standard canonical distribution, but some variable is constrained so that the state point falls in the region of one minimum in the free energy landscape (Fig. 3) where all trajectories start.
- In order that all trajectories are compatible with the path ensemble, one introduces a transition probability $p[\{\bar{x}^{(0)}(t)\} \rightarrow \{\bar{x}^{(n)}(t)\}]$ that satisfies the detailed balance condition with $P_{AB}\{\bar{x}(\{t\})\}$,

$$\begin{aligned} P_{AB}\{\bar{x}^{(0)}(\{t\})\}p[\{\bar{x}^{(0)}(t)\} \rightarrow \{\bar{x}^{(n)}(t)\}] &= \\ = P_{AB}\{\bar{x}^{(n)}(\{t\})\}p[\{\bar{x}^{(n)}(t)\} \rightarrow \{\bar{x}^{(0)}(t)\}] & . \end{aligned} \quad (31)$$

- The transition between individual states (at one time t of the Markov process, $\vec{x}^{(0)}(t) \rightarrow \vec{x}^{(n)}(t)$) hence is replaced by a transition between two full trajectories, such as shown in Fig. 3.
- This probability p then can again be written as a product of a “generation probability” of a new path and an “acceptance probability”, and from Eq. (31) one then readily can postulate a Metropolis importance sampling rule.
- Of course, the subtle problem is the generation of new paths with a reasonably high acceptance probability.
- To solve this problem, one mostly relies on the so-called “shooting algorithm” (6.28).
- For this purpose the path is divided into many small time slices.
- From the randomly selected time slice t' , one carries out a move to a new state, $\vec{x}(t') \rightarrow \vec{x}'(t')$ according to the rules of the underlying dynamics.
- From the new state $\vec{x}'(t')$ forward trajectories (that end up in B) and backward trajectories (that end up in A) are generated, so that one can glue one forward and one backward trajectory together to get a new full trajectory going from A to B.
- This new trajectory eventually is accepted (or rejected), depending on the Metropolis test based on Eq. (31).

- Of course, in order to be able to start such a sampling, an initial trajectory must be available.
- Just as in ordinary Monte Carlo sampling, where the initial state does not need to be similar to the states characteristic for the final thermal equilibrium, e.g. one can start the simulation of an Ising ferromagnet always from a perfectly ordered spin configuration irrespective of the temperature of the simulation, one can start transition path sampling from a completely atypical trajectory, and then try to relax the system to converge towards the dominating trajectories.
- Of course, it depends on the model chosen whether such an approach is practically feasible or not (and also the task of finding an initial trajectory is a nontrivial task!).
- A useful concept in transition path sampling is the “committor distribution”. It is defined as the fraction of trajectories started in A and going via a state \vec{r} to reach the state B after a time t (6.83)

$$P_B(\vec{r}, t) \equiv \frac{\int \mathcal{D}\vec{x}(t') P(\{\vec{x}(t')\}) \delta(\vec{r}_0 - \vec{r}) h_B(\vec{x}(t))}{\int \mathcal{D}\vec{x}(t') P(\{\vec{x}(t')\}) \delta(\vec{r}_0 - \vec{r})} \quad (32)$$

Here $h_B(\vec{x}) = 1$ if $x \in B$ and zero else, paths that start at time $t = 0$ at $\vec{r} = \vec{r}_0$ but do not end up at time t in the region B are not included in the numerator, while they are included in the denominator.

- Thus the committor distribution is a statistical measure for “how committed is a given configuration to reach the “product state” B” (6.83). In an practical simulation, only a finite sample of \mathcal{N} trajectories started at $t = 0$ at a point \vec{r} in configuration space is available, and then Eq. (32) can be written simply as

$$p_B(\vec{r}, t) \approx \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} h_B(\vec{x}_i(t)) \quad . \quad (33)$$

Analogously one can define $p_A(\vec{r}, t)$, and a formal definition of the “transition state ensemble” then becomes (6.28; 6.77; 6.78; 6.79; 6.80; 6.81; 6.82; 6.83)

$$p_A(\vec{r}) = p_B(\vec{r}) \quad (34)$$

where we also have invoked a time-independent form of the committor (counting all events that first reach A or first reach B, irrespective how long it has taken).

- While transition path sampling is rather straightforward to implement in the example of nucleation in the kinetic Ising model (Fig. ??), where one has a qualitative insight into the kinetics of the process and the nature of transition states a priori, there are problems where no such knowledge is available, e.g. in the folding process of complex off-lattice models for proteins (6.34).

- In such cases, it is useful to combine then the transition path sampling with complementary sampling strategies to obtain information on the free energy landscape, e.g. the replica exchange method, that has been mentioned in Sec. 6.3 (6.34). Since at the time of writing “for large proteins the computational effort due to both system size and long time scales becomes prohibitive” (6.34), this subject will remain an area of further research.

- [6.1] L. D. Landau, E. M. Lifshitz: *Statistical Physics, 3rd Ed.* (Pergamon Press, Oxford, 1980)
- [6.2] K. Binder: *Z. Phys.* **B45**, 61 (1981)
- [6.3] S. Kämmerer, B. Dünweg, K. Binder, M. D'Onorio De Meo: *Phys. Rev.* **B53**, 2345 (1996)
- [6.4] W. Schweika, D. P. Landau, K. Binder: *Phys. Rev.* **B53**, 8937 (1996)
- [6.5] F. F. Haas, F. Schmid, K. Binder: *Phys. Rev.* **B61**, 15077 (2000)
- [6.6] R. Lipowsky: *J. Appl. Phys.* **55**, 2485 (1984)
- [6.7] M. Schrader: *Diplomarbeit* (Johannes-Gutenberg-Universität Mainz, 2009, unpublished)
- [6.8] M. Schrader, P. Virnau, K. Binder: In preparation
- [6.9] P. Virnau, M. Müller: *J. Chem. Phys.* **120**, 10925 (2004)
- [6.10] L. G. Mac Dowell, P. Virnau, M. Müller, K. Binder: *J. Chem. Phys.* **120**, 5293 (2004)
- [6.11] L. G. Mac Dowell, V. K. Shen, J. R. Errington: *J. Chem. Phys.* **125**, 034705 (2006)
- [6.12] M. Müller, J. J. de Pablo: *In Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol. 1*, ed. by M. Ferrario, G. Ciccotti, K. Binder (Springer, Berlin-Heidelberg-New York 2006) p. 67

- [6.13] H. Furukawa, K. Binder: Phys. Rev. A **26**, 556 (1982)
- [6.14] B. A. Berg, U. Hansmann, T. Neuhaus: Z. Phys. B**90**, 229 (1993)
- [6.15] J. E. Hunter, W. P. Reinhardt: J. Chem. Phys. **103**, 8627 (1995)
- [6.16] K. Binder, M. H. Kalos: J. Stat. Phys. **22**, 363 (1980)
- [6.17] M. Biskup, L. Chayes, R. Kotecky: Europhys. Lett. **60**, 21 (2002)
- [6.18] K. Binder, Physica A **319**, 99 (2003)
- [6.19] E. A. Carlen, M. C. Carvalho, R. Esposito, J. L. Lebowitz, R. Marra: Molec. Phys. **103**, 3141 (2005)
- [6.20] P. Virnau, M. Müller, L. G. Mac Dowell, K. Binder: New J. Phys. **6**, 7 (2004)
[<http://stacks.iop.org/1367-2630/6/7>]
- [6.21] K. Binder: Phys. Rev. A**25**, 1699 (1982)
- [6.22] K. Binder, M. Müller, W. Oed: J. Chem. Soc., Faraday Trans. **91**, 2369 (1995)
- [6.23] M. Müller, M. Schick: J. Chem. Phys. **105**, 8885 (1996)
- [6.24] M. Müller, L. G. Mac Dowell: Macromolecules **33**, 3902 (2000)
- [6.25] P. Virnau, M. Müller, L. G. Mac Dowell, K. Binder: J. Chem. Phys. **121**, 2169 (2004)
- [6.26] R. L. C. Vink, J. Horbach, K. Binder: Phys. Rev. E**71**, 011401 (2005)
- [6.27] B. M. Mognetti, L. Yelash, P. Virnau, W. Paul, K. Binder, M. Müller, L. G. Mac Dowell: J. Chem. Phys. **128**, 104501 (2008)

- [6.28] C. Dellago, P. G. Bolhuis, P. L. Geissler: *Adv. Chem. Phys.* **123**, 1 (2002)
- [6.29] W. Janke (ed.): *Rugged Free Energy Landscapes: Common Computational Approaches to Spin Glasses, Structural Glasses and Biological Macromolecules* Lect. Notes in Physics 736 (Springer, Berlin-Heidelberg-New York, 2008)
- [6.30] K. Binder, A. P. Young: *Rev. Mod. Phys.* **58**, 801 (1986)
- [6.31] K. Binder, W. Kob: *Glassy Materials and Disordered Solids. An Introduction to Their Statistical Mechanics* (Singapore, World Scientific, 2005)
- [6.32] U. H. E. Hansmann, Y. Okamoto: In *Annual Reviews in Computational Physics VI* edited by D. Stauffer (World Scientific, Singapore, 1999), p. 129
- [6.33] Y. Sugita, Y. Okamoto, In *Lecture Notes in Computational Science and Engineering* edited by T. Schlick, H. H. Gan (Springer, Berlin-Heidelberg-New York, 2002) p. 304
- [6.34] P. G. Bolhuis: In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, edited by M. Ferrario, G. Ciccotti, K. Binder (Springer Berlin-Heidelberg-New York, 2006) p. 393
- [6.35] M. Bachmann, W. Janke: in (6.29), p. 203
- [6.36] A. Irbäck: in (6.29), p. 269
- [6.37] U. H. E. Hansmann: in (6.29), p. 293
- [6.38] Y. Sugita, A. Mitsutake, Y. Okamoto: in (6.29), p. 369.

- [6.39] J. Luettmmer-Strathmann, F. Rampf, W. Paul, K. Binder: *J. Chem. Phys.* **128**, 064903 (2008)
- [6.40] K. Binder, W. Paul, T. Strauch, F. Rampf, V. Ivanov, J. Luettmmer-Strathmann: *J. Phys.: Cond. Matter* **20**, 494215 (2008)
- [6.41] A. P. Young (ed.): *Spin Glasses and Random Fields* (World Scientific, Singapore, 1998)
- [6.42] A. P. Young: in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology, Vol 2, Lect. Notes in Phys. 704*, edited by M. Ferrario, G. Ciccotti and K. Binder (Springer, Berlin-Heidelberg-New York, 2006) p. 31.
- [6.43] A. Billoire: in (6.29), p.11
- [6.44] K. Binder, C. Brangian, W. Kob: in (6.29), p. 47
- [6.45] T. Aspelmeier, A. Billoire, E. Marinari, M. A. Moore: *J. Phys. A: Math. Theor.* **41** 324008 (2008)
- [6.46] A. P. Young,: *J. Phys. A: Math. Theor.* **41**, 324016 (2008)
- [6.47] G. M. Torrie, J. P. Valleau: *Chem. Phys. Lett.* **28**, 578 (1974)
- [6.48] G. M. Torrie, J. P. Valleau: *J. Comp. Phys.* **23**, 187 (1977)
- [6.49] B. A. Berg, T. Neuhaus: *Phys. Lett. B***267**, 249 (1991)
- [6.50] B. A. Berg, T. Neuhaus: *Phys. Rev. Lett* **68**, 9 (1992)

- [6.51] B. A. Berg: *Int. J. Mod. Phys. C* **4**, 249 (1993)
- [6.52] W. Janke: *Physica A* **254**, 164 (1998)
- [6.53] B. A. Berg: *Comp. Phys. Commun.* **153**, 397 (2003)
- [6.54] B. A. Berg: *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore, 2004)
- [6.55] M. Bachmann, W. Janke: *J. Chem. Phys.* **120**, 6779 (2004)
- [6.56] F. Wang, D. P. Landau: *Phys. Rev. Lett.* **86**, 2050 (2001)
- [6.57] F. Wang, D. P. Landau: *Phys. Rev.* **E64**, 056101 (2001)
- [6.58] D. P. Landau, F. Wang: *Comput. Phys. Commun.* **147**, 674 (2002)
- [6.59] B. J. Schulz, K. Binder, M. Müller, D. P. Landau: *Phys. Rev.* **E67**, 067102 (2003)
- [6.60] S. Trebst, D. A. Huse, M. Troyer: *Phys. Rev.* **E70**, 046701 (2004)
- [6.61] D. P. Landau, K. Binder: *A Guide to Monte Carlo Simulation in Statistical Physics 3rd ed.* (Cambridge Univ. Press, Cambridge, 2009)
- [6.62] C. Zhou, R. N. Bhatt: *Phys. Rev.* **E72**, 025701 (2005)
- [6.63] D. P. Landau: in (6.29), p. 353
- [6.64] N. B. Wilding: In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol 1, Lect. Notes in Phys. 703*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin-Heidelberg-New York, 2006) p. 39

- [6.65] A. P. Lyubartsev, A. A. Matsinovski, S. V. Shevkunov, P. N. Vorontsov-Velyaminov: *J. Chem. Phys.* **96**, 1776 (1992)
- [6.66] E. Marinari, G. Parisi: *Europhys. Lett.* **19**, 451 (1992)
- [6.67] J. Lee: *Phys. Rev. Lett.* **71**, 211 (1993); *Erratum ibid.* **71**, 2353 (1993)
- [6.68] K. Hukushima, H. Takayama and K. Nemoto: *Int. Mod. Phys.* **C3**, 337 (1996)
- [6.69] P. M. C. de Oliveira, T. J. P. Penna, H. J. Herrmann: *Braz. J. Phys.* **26**, 677 (1996)
- [6.70] P. M. C. de Oliveira, T. J. P. Penna, H. J. Herrmann: *Eur. Phys. J. B* **1**, 205 (1998)
- [6.71] J. S. Wang, L. W. Lee: *Comput. Phys. Commun.* **127**, 131 (2000)
- [6.72] Y. Iba: *Int. J. Mod. Phys. C* **12**, 623 (2001)
- [6.73] Q. Yan, F. Faller, J. J. de Pablo: *J. Chem. Phys.* **116**, 8745 (2002)
- [6.74] J. S. Wang, R. H. Swendsen: *J. Stat. Phys.* **106**, 245 (2002)
- [6.75] I. Carmesin, K. Kremer: *Macromolecules* **21**, 2819 (1988)
- [6.76] H. P. Deutsch, K. Binder: *J. Chem. Phys.* **94**, 2294 (1991)
- [6.77] C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler: *J. Chem. Phys.* **108**, 1964 (1998)
- [6.78] C. Dellago, P. G. Bolhuis, D. Chandler: *J. Chem. Phys.* **108**, 9236 (1998)
- [6.79] P. G. Bolhuis, C. Dellago, D. Chandler: *Faraday Discuss.* **110**, 421 (1998)

- [6.80] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler: *Ann. Rev. Phys. Chem.* **53**, 291 (2002)
- [6.81] C. Dellago, D. Chandler: In *Bridging Time Scales: Molecular Simulation for the Next Decade* edited by P. Nielaba, M. Mareschal, and G. Ciccotti (Springer, Berlin-Heidelberg-New York, 2002)
- [6.82] C. Dellago: In *Handbook of Materials Modeling*, edited by S. Yip (Springer, Berlin-Heidelberg-New York, 2005) p. 1585
- [6.83] C. Dellago, P. G. Bolhuis, P. L. Geissler: In *Computer Simulations in Condensed Matter. From Materials to Chemical Biology, Vol 1.* edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin-Heidelberg-New York, 2006), p. 349
- [6.84] W. E., W. Ren, E. Vanden-Eijnden: *Phys. Rev.* **B66**, 05301 (2002)
- [6.85] A. Laio. M. Parrinello: *Proc. Natl. Acad. Sci. USA* **99**, 12562 (2002)
- [6.86] R. Elber, A. Ghosh, A. Gardenas, H. Stern: *Adv. Chem. Phys.* **126**, 93 (2004)
- [6.87] C. Micheletti, A. Laio, M. Parrinello: *Phys. Rev. Lett.* **92**, 170601 (2004)
- [6.88] E. Weinan, W. Ren, E. Vanden-Eijnden: *J. Phys. Chem.* **B109**, 6688 (2005)
- [6.89] G. Bussi. A. Laio, M. Parrinello: *Phys. Rev. Lett.* **96**, 090601 (2006)
- [6.90] A. Laio, M. Parrinello: In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol. 1*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer, Berlin-Heidelberg-New York, 2006) p. 315

- [6.91] H. Müller-Krumbhaar, K. Binder: J. Stat. Phys. **8**, 1(1973)
- [6.92] K. Binder, H. Müller-Krumbhaar: Phys. Rev. **B9**, 2328 (1974)
- [6.93] K. Binder, D. Stauffer: Advanc. Phys. **25**, 343 (1976)
- [6.94] A. C. Pan, D. Chandler: J. Phys. Chem. **B108**, 19681 (2004)
- [6.95] D. Frenkel, B. Smit: *Understanding Molecular Simulation: From Algorithms to Applications*. 2nded. (Academic Press, San Diego, 2002)
- [6.96] C. H. Bennett: J. Comp. Phys. **22**, 245 (1976)
- [6.97] B. M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, L. G. MacDowell: J. Chem. Phys. (2009, in press)
- [6.98] F.Y. Wu: Rev. Mod. Phys. **54**, 235 (1982)
- [6.99] K. Binder, D. P. Landau: Phys. Rev. **B30**, 1477 (1984)
- [6.100] C. Borgs, R. Kotecky: J. Stat. Phys. **61**, 79 (1990)
- [6.101] A. M. Ferrenberg, R. H. Swendsen: Phys. Rev. Lett. **61**, 2635 (1988)
- [6.102] A. M. Ferrenberg, R. H. Swendsen: Phys. Rev. Lett. **63**, 1195 (1989)
- [6.103] B. A. Berg: In (6.29), p. 317
- [6.104] B. A. Berg: J. Stat Phys. **82**, 323 (1996)
- [6.105] B. A. Berg, A. Billoire, W. Janke: Phys. Rev. **E61**, 12143 (2000)
- [6.106] B. J. Schulz, K. Binder, M. Müller: Phys. Rev. **E71**, 046705 (2005)

- [6.107] A. M. Ferrenberg, D. P. Landau, K. Binder: Phys. Rev. **E58**, 3353 (1998)
- [6.108] K. Vollmayr, J. D. Reger, M. Scheucher, K. Binder: Z. Phys. B: (Condens. Matter) **91**, 113 (1993)
- [6.109] T. Neuhaus, S. Hager: J. Stat.Phys. **113**, 47 (2003)
- [6.110] D. Moroni, P. R. ten Wolde, P. G. Bolhuis: Phys. Rev. Lett. **94**, 235703 (2005)
- [6.111] A. C. Zettelmoyer: *Nucleation* (Marcel Dekker, New York, 1969)
- [6.112] Z. W. Salsburg, J. D. Jacobson, W. S. Fickett, W. W. Wood: J. Chem. Phys. **30**, 65 (1959)
- [6.113] S.-H. Tsai, F. Wang, D. P. Landau; H.-Q. Lin: Phys. Rev. **E75**, 061108 (2007)
- [6.114] Z. Chen, Y. Gu: J. Bioinformatics and Comput. Biol. **4**, 317 (2006)
- [6.115] E. Luijten, H. W. J. Blöte: Int. J. Mod. Phys. **C6**, 359 (1995)
- [6.116] J. Lui, E. Luijten: Phys. Rev. **E71**, 066701 (2005)
- [6.117] Y. C. Kim, M. E. Fisher: Phys. Rev. **E68**, 041506 (2003)
- [6.118] J. C. Grossman: J. Chem. Phys. **117**, 1434 (2002)
- [6.119] M. Nekovee, W. M. C. Foulkes, R. J. Needs: Math. Comput. Simulat. **62**, 463 (2003)
- [6.120] W. Schattke, R. Bahnsen, R. Redmer: Progr. Surf. Sci. **72**, 87 (2003)
- [6.121] S.-I. Lui, J. Chem. Phys. **121**, 10365 (2004)