Multicanonical methods

Kari Rummukainen, Nordita
(Late) Summer School on Monte Carlo Methods

Normal Monte Carlo algorithms sample configurations with the Boltzmann weight $p \propto \exp(-\beta E)$. Sometimes this is not desirable.

> **Example: if a system has a first order phase transitions between two states the system is preferably in either of the states (at the transition temperature). The tunnelling between the states can be strongly suppressed. If the goal is to study the tunnelling process, the Monte Carlo simulation should enhance the probability of those configurations!**

Multicanonical methods tailor the MC probability in order to enhance configurations in desired phase space domains.

Related methods:

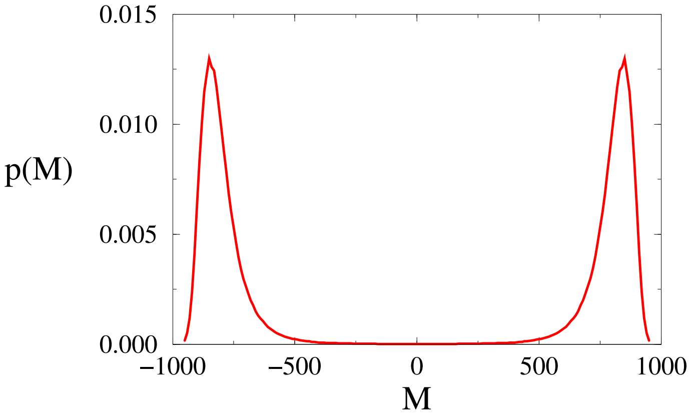
- **Multicanonical** [Berg, Neuhaus, Phys.Lett.B 267 (1991)]
  Umbrella sampling [Torrie, Valleau, J.Comp.Phys. 23 (1977)]
- **Simulated tempering** [Marinari, Parisi, Europhys.Lett. 19 (1992)]
  Method of expanded ensembles [Lyubartsev et al, J.Chem.Phys. 96 (1992)]
- **Parallel tempering** [Hukushima, Nemoto 96]
  [Geyer 91]
- **Parallel multicanonical method**

Recent review: B. Berg, cond-mat/9909236

Multicanonical simulations and 1st order phase transitions

- **Consider again Ising model, this time at $T < T_c$: now it has a 1st order phase transition between states with positive and negative magnetization.**

  Magnetization distribution $p(M)$, volume 32^2, $\beta = 1/T = 0.453$, $h = 0$:



- **Canonical probability**

$$p_{\text{can}}(M) \propto \sum_{\{s\}} \delta(M, M_s) e^{-\beta E_s}$$

  is strongly suppressed between the peaks

  $\rightarrow$ canonical MC simulations hardly ever sample those configurations.

- **Multicanonical method:**

  Instead of using the Bolzmann weight $p \propto \exp[-\beta E]$ in Monte Carlo simulations, use a modified probability

$$p \propto \exp[-\beta E + W(M)] \, ,$$

  where $W(M)$ is a weight function, inserted 'by hand', which is carfully tuned to enhance the probability between the peaks.

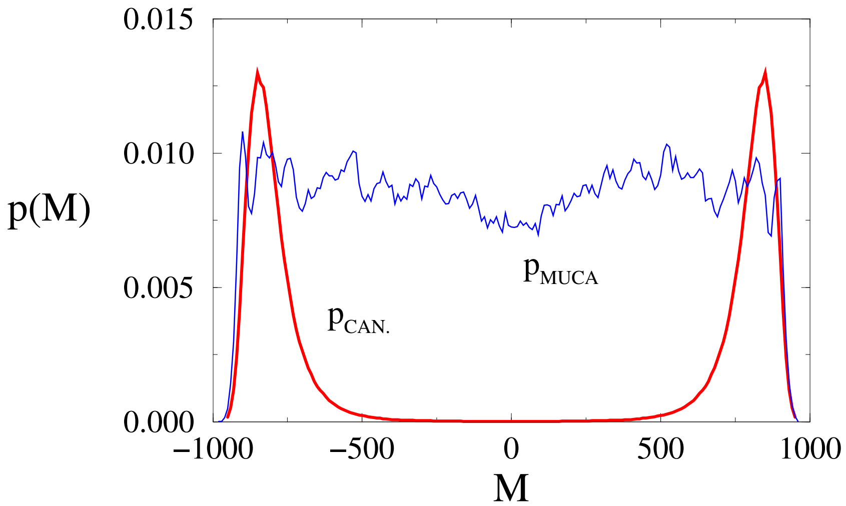
- **Multicanonical probability is produced by multicanonical simulations:**

$$p_{\rm muca}(M) \propto \sum_{\{s\}} \delta(M, M_s) e^{-\beta E_s + W(M_s)} \propto p_{\rm can}(M) e^{W(M)}$$

- **If we want that $p_{\rm muca}(M)$ is flat in the region between the peaks we should choose $W(M) = -\log p_{\rm can}(M)$. However, we do not a priori know $p_{\rm can}(M)$. Thus, one has to do with approximations (later).**
- **Multicanonical simulation gives us $p_{\rm muca}(M)$. We obtain the canonical distribution (which is, after all, what we're after!) by reweighting:**

$$p_{\rm can}(M) \propto p_{\rm muca}(M) e^{-W(M)}$$

- **The picture below shows $p_{\rm muca}(M)$, which is $\sim$ constant. The wigglines is due to numerical noise (statistics) and the 'incorrectness' of $W(M)$. $p_{\rm can}$ has been obtained from $p_{\rm muca}$ by reweighting.**



- **How accurately should W be determined? Remember that the simulation is formally correct with arbitrary $W(M)$. Good W can just save a lot of cpu-time. We should only demand that $p_{\rm muca}(M)$ does not become too small in any area of interest.**

  $\mapsto$ **If we allow $p_{\rm muca}$ to vary $\sim 50\%$, W can vary by $W \pm \log 2$.**

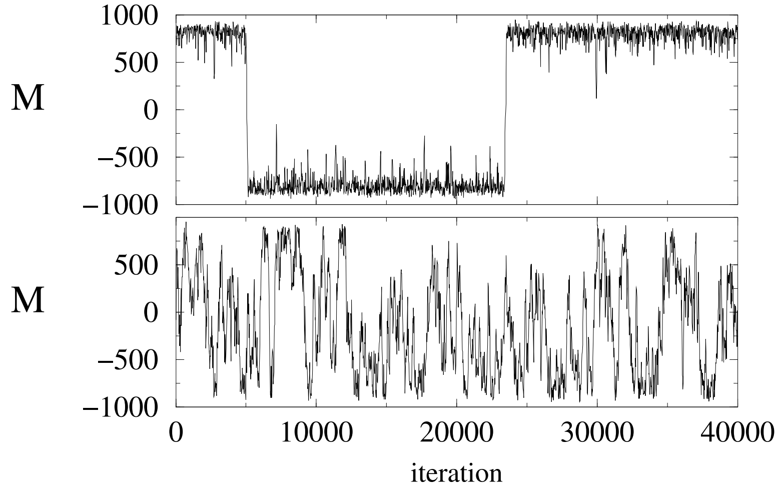
- **The parametrization of $W(M)$ is not unique. Perhaps the most common is a piecewise linear one: choose a discrete small set $\{m_i\}$ of magnetizations, and parametrize $w_i = W(m_i)$, and interpolate linearly between m_i's:**

$$W(M) = w_{i+1}\frac{M - m_i}{m_{i+1} - m_i} + w_i\frac{m_{i+1} - M}{m_{i+1} - m_i} \qquad \text{when } m_i \le M \le m_{i+1}\,.$$

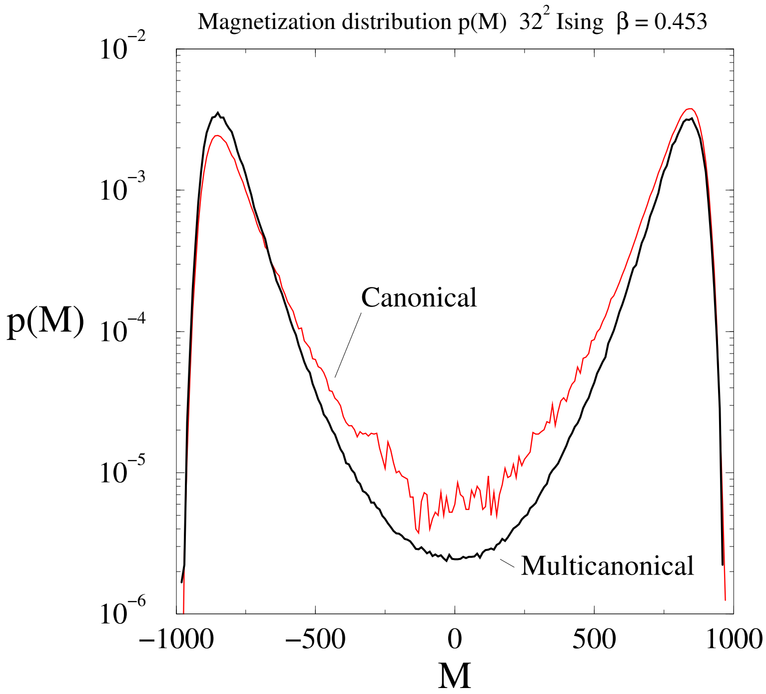
- **In the example above W was a function of M ("multimagnetic" ensemble). In general, W can be a function of any order parameter. Original formulation was in terms of E (useful for temperature-driven transitions).**

Performance comparison: again 32^2 Ising, at $\beta = 0.453$.

• **Monte Carlo time history of M, from canonical and multicanonical simulations:**



• **Canonical distributions $p_{\mathrm{can}}(M)$, obtained from multicanonical and canonical simulations (200000 measurements):**



• **In 1st order phase transitions, the probability suppression between the pure phase peaks increases exponentially. The configurations in the middle are 'mixed': they contain a fraction of both pure phases, with an interface in between. The interface has extra free energy, $f_A = \sigma A$, where σ = surface tension, and A is the area of the interface. Minimally $A = 2L$ on 2-dim. lattice with periodic boundaries.**

$$p_{\min}/p_{\max} \sim e^{-\sigma 2L}$$

Example:
20-state 2-d Potts model:

$$E = \sum_{ij}(1 - \delta_{s_i,s_j})$$

$s_i = 1 \ldots 20$

Temperature-driven 1st order phase transition at $\beta_c = \log(1 + \sqrt{20}) \approx 2.900$

[A. Billoire, T. Neuhaus and B. Berg, 1994]

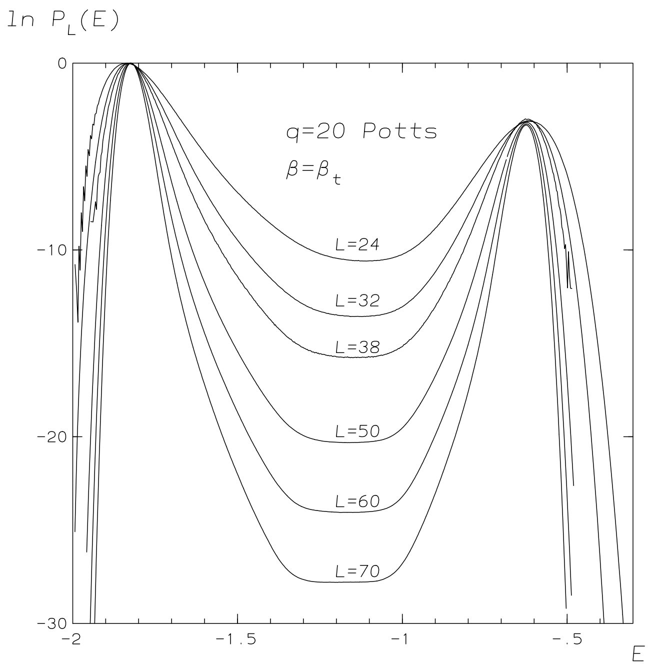
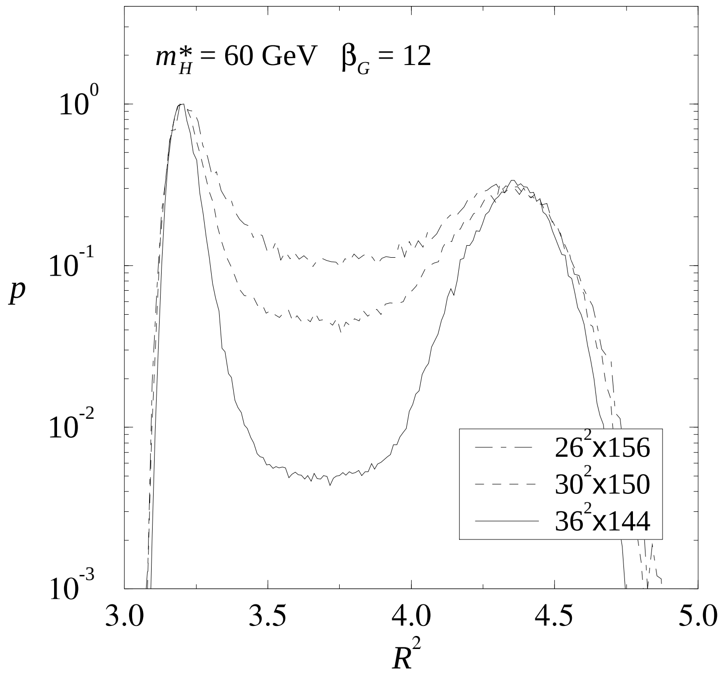


FIG. 8

Multicanonical simulations are not sensitive to this suppression. However, they have to 'random walk' through the transition region. The region widens $\propto V$, but the 'step size' (change in one update cycle) increases only as $\propto V^{1/2}$:

- $\tau \sim e^{-\sigma 2L}$ **in canonical simulations (supercritical slowing down)**
- $\tau \sim V^2 = L^{2d}$ **in multicanonical simulations**

SU(2)-Higgs model (Standard Model phase transition at $T \sim 100$ GeV.



[Kajantie, Laine, K.R, Shaposhnikov, 1995]

Strong 1st order transition if Higgs particle mass $\lesssim 60$ GeV.

How to calculate W?

If we know p_{can}, we know W – but p_{can} is the quantity we are trying to determine! To determine W:

1. Use canonical simulations to obtain rough p_{can}
2. Finite-size scaling
3. Iterate: $W^1 \to W^2 \to \ldots$
4. Recursive computation of W: automatic iterative process.

- **Let us divide our order parameter M into bins, so that the system is in bin m when $M_m \le M < M_{m+1}$. The iteration is based on the relative probabilities $p(m+1)$ and $p(m)$ that the system is in bins m and $m+1$ during a Monte Carlo run.**
- **For presentational simplicity, I parametrize $W(M) = w(m)$, when $M_m \le M < M_{m+1}$. Generalization to the piecewise linear W is straightforward.**

The iteration proceeds as follows:

a During a (short) Monte Carlo run (number i), measure

$$n_i(m) = \#\ \text{hits in bin}\ m \qquad h_i(m) = n_i(m) e^{-w(m)}\,.$$

Thus, $n_i(m)$ is multicanonical histogram, $h_i(m)$ canonical.

b After k runs, we obtain the estimate $k+1$ for W:

$$w_{k+1}(m+1) - w_{k+1}(m) = \sum_{i\le k} g_i(m) \log \frac{h_i(m+1)}{h_i(m)} \Big/ \sum_{i \le k} g_i(m)\,.$$

Here $g_i(m) = n_i(m+1) + n_i(m)$ is the statistical importance of the run i at bins m and $m+1$.

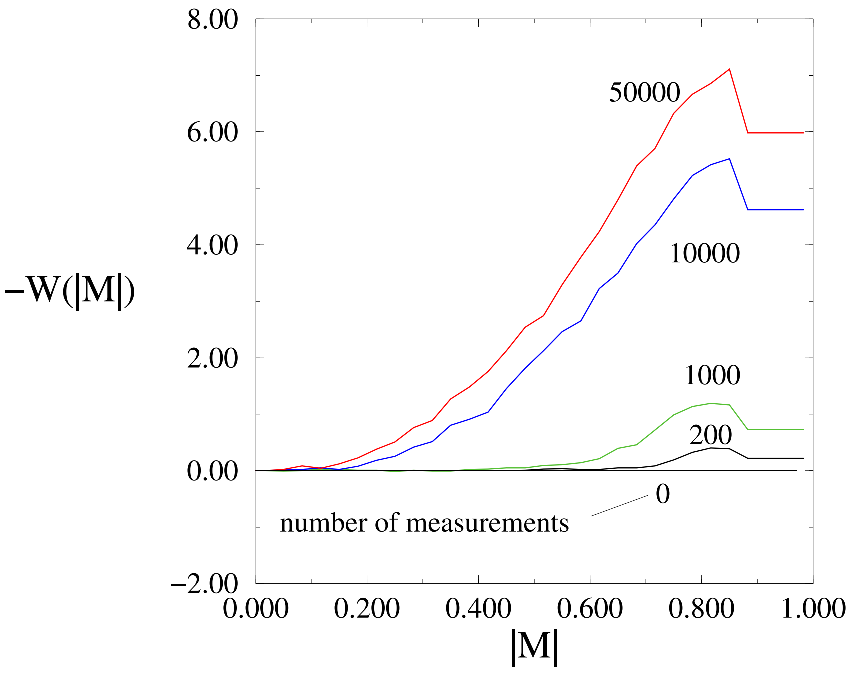
c The convergence can be accelerated by overcorrection of W: let us add to $w_{k+1}(m)$, as obtained above, a term

$$-C \log \sum_{i \le k} n_i(m)\,,$$

C a suitable constant. This disfavors bins where the system has visited often before, and pushes it towards unexplored phase space. The final W is still given by the original expression.

- The steps a,b,c are iterated until W has converged.
- Initially $w_1(m)$ is set to the initial guess of W, and $n_1(m)$ proportional to its 'quality'. Note that we must have $n_1(m) > 0$ everywhere. The larger $n_1(m)$ is, the more damped is the evolution of W.
- Note that the recursion uses all of the available information to the final W, not only the last simulation.

Evolution of the weight function (32^2 Ising)



Simulated tempering

- **Idea: upgrade β to dynamical variable; that is, allow it to change during a simulation.**
- **Canonical probability:** $p_\beta(E) = n(E)e^{-\beta E + f_\beta}$
- **Let now β_i be a set of β-values chosen so that $p_{\beta_i}(E) \equiv p_i(E)$ and $p_{i+1}(E)$ overlap.**

1. **Perform a (small) number of standard MC simulation steps at fixed $\beta = \beta_i$.**
2. **Propose a change $\beta_i \to \beta_j$. This is accepted with probability**

$$p(\beta_i \to \beta_j) = \min\left(1\,,\ \exp[-(\beta_i - \beta_j)E + f_i - f_j]\right)$$

- **Iterate steps 1 and 2.**
- **The update in 2. satisfies detailed balance. However, we do not a priori know the free energies f_i! These have to be estimated, as with the multicanonical algorithm.**
- **Note: f_i's are same parameters which appear in multiple histogram reweighting! So, we can perform simulations at fixed β_i's and use multiple histogram method to get f_i's.**
- **Advantage of simulated tempering over running with fixed temperatures: in simulated tempering, the system goes up and down in temperature - thus, it becomes periodically very disordered. This avoids it getting 'stuck' in one configuration at low temperatures.**
- **Same result can be achieved with multicanonical simulation with $W(E)$.**
- **Very suitable for spin glass simulations, polymers etc.**

Parallel tempering

- **Modification of simulated tempering idea:**
- **Again let β_i's, $i = 1 \ldots N$, be a set of β-values chosen so that $p_{\beta_i}(E) \equiv p_i(E)$ and $p_{i+1}(E)$ overlap.**

1. **Perform a (small) number of simulation steps simultaneously on N systems, with $\beta = \beta_1, \beta_2 \ldots \beta_N$ respectively.**
2. **Propose a swap of beta-values $\beta_i \leftrightarrow \beta_j$. This is accepted with probability**

$$p(\beta_i \leftrightarrow \beta_j) = \min\left(1\,,\, \exp[-(\beta_i - \beta_j)(E_i - E_j)]\,\right)$$

where E_i is the energy of system number i.

- **Iterate steps 1 and 2.**
- **Does essentially the same trick as parallel tempering - applicable to similar problems.**
- **Advantages: no f_i's needed, thus, no tuning! Also, method is inherently parallel.**