Microcanonical Approach to the Simulation of First-Order Phase Transitions.

V. Martin-Mayor\textsuperscript{1,2}

\textsuperscript{1}Departamento de Física Teórica I, Facultad de Ciencias Físicas, Universidad Complutense, 28040 Madrid, Spain.
\textsuperscript{2}Instituto de Biocomputación y Física de Sistemas Complejos, (BIFI, Spain).

(Dated: March 23, 2022)

A generalization of the microcanonical ensemble suggests a simple strategy for the simulation of first order phase transitions. At variance with flat-histogram methods, there is no iterative parameters optimization, nor long waits for tunneling between the ordered and the disordered phases.

We test the method in the standard benchmark: the $Q$-states Potts model ($Q=10$ in 2 dimensions and $Q=4$ in 3 dimensions), where we develop a cluster algorithm. We obtain accurate results for systems with $10^6$ spins, outperforming flat-histogram methods that handle up to $10^4$ spins.

PACS numbers: 64.60.Cn, 75.40.Mg, 05.50.+q.

Phase transitions are ubiquitous (formation of quark-gluon plasmas, vaporization/crystalization of ordinary liquids, Cosmic Inflation, etc.). Most of them are of (Ehrenfest) first order \cite{Matsubara}. Monte Carlo simulations \cite{Metropolis} are crucial for their investigation, but difficulties arise for large system linear size, $L$ (or space dimension, $D$). The intrinsic problem is that, at a first order phase transition, two (or more) phase coexist. The simulated system tunnels between pure phases by building an interface of size $L$. The free-energy cost of such a mixed configuration is $\Sigma L^{D-1}$ (if: surface tension), the interface is built with probability $\exp[-\Sigma L^{D-1}]$ and the natural time scale for the simulation grows with $L$ as $\exp[\Sigma L^{D-1}]$. This disaster is called exponential critical slowing down (ECSD).

No cure is known for ECSD in canonical simulations (cluster methods \cite{Dinero, Swendsen} do not help), which motivated the invention of the multicanonical ensemble \cite{Glauber, Swendsen2}. The multicanonical probability for the energy density is constant, at least in the energy gap $e^o < e < e^d$ ($e^o$ and $e^d$: energy densities of the coexisting low-temperature ordered and high-temperature disordered phase), hence the name flat-histogram methods \cite{Swendsen, Glauber, Swendsen2}. The canonical probability minimum in the energy gap ($x \exp[-\Sigma L^{D-1}]$) is filled by means of an iterative parameter optimization.

In flat-histogram methods the system performs an energy random walk in the energy gap. The elementary step being of order $L^{-D}$ (a single spin-flip), one naively expects a tunneling time from $e^o$ to $e^d$ of order $L^{2D}$ spin-flips. But the (one-dimensional) energy random walk is not Markovian, and these methods suffer ECSD \cite{Swendsen2}. In fact, for the standard benchmark (the $Q=10$ Potts model \cite{Swendsen} in $D=2$), the barrier of $10^4$ spins was reached in 1992 \cite{Swendsen2}, while the largest simulated system (to our knowledge) had $4 \times 10^4$ spins \cite{Swendsen2}.

ECSD in flat histogram simulations is probably understood \cite{Swendsen2}: on its way from $e^d$ to $e^o$, the system undergoes several (four in $D=2$) “transitions”. First comes the condensation transition \cite{Swendsen2, Pikovsky}, at a distance of order $L^{-(D+1)}$ from $e^d$, where a macroscopic droplet of the ordered phase is nucleated. Decreasing $e$, the droplet grows to the point that, for periodic boundary conditions, it reduces its surface energy by becoming a strip \cite{Swendsen2}, see Fig. 2 (in $D=3$, the droplet becomes a cylinder, then a slab \cite{Swendsen}). At lower $e$ the strip becomes a droplet of disordered phase. Finally, at the condensation transition close to $e^o$, we encounter the homogeneous ordered phase.

Here we present a method to simulate first order transitions without iterative parameter optimization nor energy random walk. We extend the configuration space as in Hybrid Monte Carlo \cite{Duane}: to our $N$ variables, $\sigma_i$ (named spins here, but they could be atomic positions) we add $N$ real momenta, $p_i$. The microcanonical ensemble for the $\{\sigma_i, p_i\}$ offers two advantages. First, microcanonical simulations \cite{Martin-Mayor} are feasible at any value of $e$ within the gap. Second, we obtain Fluctuation-Dissipation Eqs. \cite{Langer, Evertsz} where the (inverse) temperature $\beta$, a function of $e$ and the spins, plays a role dual to that of $e$ in the canonical ensemble. The $e$ dependence of the mean value $\langle \beta \rangle_e$, interpolated from a grid as it is almost constant over the gap, characterizes the transition. We test the method in the $Q$-states Potts model, for which we develop a cluster algorithm. We handle systems with $10^6$ spins for $Q = 10$ in $D = 2$ and for $Q = 4$ in $D = 3$ (where multibondic simulations handle $N = 10^4$ \cite{Swendsen2}).

Let $U$ be the spin Hamiltonian. Our total energy is

$$E = \sum_{i=1}^{N} \frac{p_i^2}{2} + U, \quad (e \equiv E/N, \quad u \equiv U/N). \quad (1)$$

In the canonical ensemble, the $\{p_i\}$ are a trivial gaussian bath decoupled from the spins. Note that, at inverse temperature $\beta$, one has $\langle e \rangle_\beta = \langle u \rangle_\beta + 1/(2\beta)$.

Microcanonically, the entropy density, $s(e, N)$, is given by $\sum_{\{\sigma_i\}}$ (summation over spin configurations)

$$\exp[NS(e, N)] = \int_{-\infty}^{\infty} \prod_{i=1}^{N} dp_i \sum_{\{\sigma_i\}} \delta(N e - E), \quad (2)$$

or, integrating out the $\{p_i\}$ using Dirac’s delta function,

$$\exp[NS(e, N)] = \frac{(2\pi N)^{N/2}}{\sqrt{\pi} T(N/2)} \sum_{\{\sigma_i\}} (e-u)^{N/2-1} \theta(e-u). \quad (3)$$

In the canonical ensemble, the $\{p_i\}$ are a trivial gaussian bath decoupled from the spins. Note that, at inverse temperature $\beta$, one has $\langle e \rangle_\beta = \langle u \rangle_\beta + 1/(2\beta)$.
The Heaviside step function, \( \theta(e-u) \), enforces \( e > u \). The microcanonical average at fixed \( e \) of a generic function of \( e \) and the spins, \( O(e, \{ \sigma_i \}) \), is (see Eq. (3) and (15))
\[
\langle O \rangle_e = \frac{\sum_{\{\sigma_i\}} O(e, \{\sigma_i\}) (e-u)^{N-1}\theta(e-u)}{\sum_{\{\sigma_i\}} (e-u)^{N-1}\theta(e-u)}.
\]

(4)

The Metropolis simulation of Eq. (4), is straightforward. Calculating \( ds/d\epsilon \) from Eq. (4) we learn that
\[
\frac{ds(e, N)}{de} = \langle \dot{\beta}(e; \{\sigma_i\}) \rangle_e , \tag{5}
\]
\[
\dot{\beta}(e; \{\sigma_i\}) = \frac{N-2}{2N(e-u)} . \tag{6}
\]

Fluctuation-Dissipation follows by derivating Eq. (4):
\[
\frac{d\langle O \rangle_e}{de} = \left\langle \frac{\partial O}{\partial e} \right\rangle_e + N \left[ \langle O \dot{\beta} \rangle_e - \langle O \rangle_e \langle \dot{\beta} \rangle_e \right] . \tag{7}
\]

As in the canonical case \( \dot{c} \), an integral version of (7) allows to extrapolate \( \langle O \rangle_e \) from simulations at \( e \geq e' \):
\[
\langle O \rangle_{e'} = \left\langle \frac{O(e'; \{\sigma_i\})}{\theta(e' - u)} \left[ \frac{e' - u}{e - u} \right] ^{N-1} \right\rangle_e . \tag{8}
\]

For \( e < e' \), configurations with \( e < u < e' \), suppressed by a factor \( (e'-u)^{N/2-1} \), are ignored in (3). Since we are limited in practice to \( |e-e'| \leq \sqrt{(u^2)/d(u)} = (u/\epsilon)/|d(u)/de| \sim N^{-1/2} \), the restriction \( e \geq e' \) can be dropped, as it is numerically negligible.

The canonical probability density for \( e \), \( P^L(\beta)(e) \propto \exp[N(s(e, N) - \beta e)] \) follows from \( \langle \dot{\beta} \rangle_e \):
\[
\log P^L(\beta)(e_2) - \log P^L(\beta)(e_1) = N \int_{e_1}^{e_2} \text{d}e \left( \langle \dot{\beta} \rangle_e - \beta \right) . \tag{9}
\]

In the thermodynamically stable region (i.e. \( d\beta \langle \dot{\beta} \rangle_e /de < 0 \)), there is a single root of \( \langle \dot{\beta} \rangle_e = \beta \), at the maximum of \( P^L(\beta) \). But, see Fig. 1, in the energy gap \( \langle \dot{\beta} \rangle_e \) has a maximum and a minimum (\( L \)-dependent spinodals \( \beta \)), and there are several roots of \( \langle \dot{\beta} \rangle_e = \beta \). The rightmost (leftmost) root is \( \langle \dot{\beta} \rangle_e = \langle \dot{\beta} \rangle_{\beta = \beta} \), a local maximum of \( P^L(\beta) \) corresponding to the disordered (ordered) phase. We define \( \beta_e^L(\beta) \) as the second rightmost root of \( \langle \dot{\beta} \rangle_e = \beta \).

At the finite-system (inverse) critical temperature, \( T_c^L \), one has \( \langle \dot{\beta} \rangle_e = \beta_e^L(\beta) \), which is equivalent, Eq. (10) and (20), to Maxwell’s construction:
\[
\langle \dot{\beta} \rangle_e = \beta_e^L(\beta) = \int_{\beta_e^L(\beta)}^{\beta_e^L(\beta)} \text{d} \beta \left( \langle \dot{\beta} \rangle_e - \beta^L \right) . \tag{10}
\]

(For large \( N \), \( \beta_c^L - \beta_c^L \propto 1/N \)). Actually, at fixed \( e \) in the gap, also \( \langle \dot{\beta} \rangle_e \) tends to \( \beta_c^L \) for large \( N \). In the strip phase it converges faster than \( \beta_c^L \), see Table II.

In a cubic box the surface tension is estimated as
\[
\Sigma^L = \frac{N}{2L^{D-1}} \int_{\beta_e^L(\beta)}^{\beta_e^L(\beta)} \text{d} \beta \left( \langle \dot{\beta} \rangle_e - \beta^L \right) . \tag{11}
\]

\( L \rightarrow \infty \) extrapolations \( \Sigma^L \propto 1/L \) are popular.

As for the specific heat, for \( N \rightarrow \infty \) the inverse function of the canonical \( \langle \dot{\beta} \rangle_e \) is the microcanonical \( \langle \dot{\beta} \rangle_e \):
\[
\int_{\beta_e^L(\beta)}^{\beta_e^L(\beta)} \text{d} \beta \left( \langle \dot{\beta} \rangle_e - \beta^L \right) . \tag{12}
\]

For large \( N \), \( \beta_e^L(\beta) \), \( \beta_e^L(\beta) \), \( \beta_e^L(\beta) \), \( C_e(\beta_e^L(\beta)) \), \( C_L(\beta_e^L(\beta)) \) tend to \( e^4 \), \( e^6 \), or the specific heat of the coexisting phases (we lack analytical hints about convergence rates).

We now specialize to the Potts model \( D \). The spins \( \sigma_i = 0, 1, \ldots, Q - 1 \), live in the \( N = L^D \) nodes of a (hyper)cubic lattice of side \( L \) with periodic boundary conditions, and interaction \( <ij> : \) lattice nearest-neighbors:
\[
U = -\sum_{<ij>} \delta_{\sigma_i, \sigma_j} . \tag{13}
\]

A cluster method is feasible. Let \( \kappa \) be a tunable parameter and \( w(e, u, \kappa) = (e - u)^{N/2-1} \exp[\kappa N u] \theta(e - u) \). Our weight is \( w(e, u, \kappa) \exp[-\kappa U] \), see (14), or, introducing bond occupation variables, \( n_{ij} = 0, 1 \), and \( p \equiv 1 - \exp[\kappa] \),
\[
w(e, u, \kappa) \prod_{<i,j>} \left[ (1-p)\delta_{n_{ij}, 0} + p \delta_{n_{ij}, 1}\delta_{\sigma_i, \sigma_j} \right] . \tag{14}
\]
which is the canonical statistical weight at $\beta = \kappa$ [24], but for the $\{n_{ij}\}$ independent factor $w(e, u, \kappa)$. Hence, clusters are traced in the standard way, but we accept a single-cluster flip [4] with Metropolis probability
\[
p(e, \kappa) = \min\{1, w(e, u_{\text{final}}, \kappa)/w(e, u_{\text{initial}}, \kappa)\}.
\]
Eqs. (5-8) suggest that $\kappa = \langle \hat{\beta} \rangle_e$ maximizes $p(e, \kappa)$ (a short Metropolis run provides a first $\kappa$ estimate). We obtain $\langle p(e, \kappa) \rangle_e > 0.99$ for $e \leq e_d$, and still $\langle p(e, \kappa) \rangle_e=e' > 0.78$.

We simulated the $(Q = 10, D = 2)$ Potts model [22], for $L = 32, 64, 128, 256, 512$ and 1024, sampling $\langle \hat{\beta} \rangle_e$ at 30 points evenly distributed in $-1.41666 \leq e \leq -0.45$. For $L = 512$, we made 15 extra simulations to resolve the narrow spinodal peaks (26 extra points for $L = 1024$). Our Elementary Monte Carlo Step (EMCS) was: max$\{10N/(Nc\langle p(e, \kappa) \rangle_e)\}$ cluster-flip attempts ($N$: number of spins in the traced cluster; it is of order one at $e_d$ and of order $N$ at $e_o$). So, every EMCS we flip at least $N$ spins. For each $e$, we performed $2 \times 10^6$ EMCS, dropping the first 10% for thermalization. A similar computation was carried out for the $(Q = 4, D = 3)$ Potts model [16] (for details see Table I and [25]).

Our $\langle \hat{\beta} \rangle_e$ in $D = 2$ is shown in Fig. I. Data reweighing [8] was used only to reconstruct the narrow spinodal peaks. To find the roots of $\langle \hat{\beta} \rangle_e = \beta$, or to calculate the integrals in Eqs. (10-11), we interpolated $\langle \hat{\beta} \rangle_e$ using a cubic spline [25]. For $L \geq 512$ the strip-droplet transitions produce two “jumps” in $\langle \hat{\beta} \rangle_e$, causing oscillations in the interpolation (Gibbs phenomenon), cured by either of two interpolation schemes, see Fig. I.

We obtain $\beta^L_e$, $\Sigma^L_e$, $e^L_d(\beta^L_e)$, $e^L_1(\beta^L_e)$, $C_L(e^L_d(\beta^L_e))$ and $C_L(e^L_1(\beta^L_e))$ from the interpolation of $\langle \hat{\beta} \rangle_e$, and of $d(\langle \hat{\beta} \rangle_e)/de$, see [7]. Statistical errors are Jack-Knife’s [26] (the $i$-th block is obtained interpolating the $i$-th Jack-Knife blocks for $\langle \hat{\beta} \rangle_e$). There are also interpolation and integration errors. Fortunately, errors of order $\epsilon$ in $e^L_d(\beta^L_e)$ or $e^L_1(\beta^L_e)$ yield errors of order $\epsilon^2$ in $\beta^L_e$: the main error in $\beta^L_e$ is the quadrature error for $\langle \hat{\beta} \rangle_e$ divided by the latent heat. On the other hand, $e^L_1(\beta^L_e)$ is near to the droplet-strip transition, and an error on it does have an impact on $\Sigma^L_e$.

In Table II are our results for $(D = 2, Q = 10)$ and the known large $L$ limits. A fit for $\epsilon$ in $\beta^L_e - \beta^L_e = c/L^D$ [21] is unacceptable for $L \geq 32$ ($\chi^2/d.o.f. = 14.32/4$), but good for $L \geq 64$ ($\chi^2/d.o.f. = 1.77/3$); our accuracy allows to detect subleading corrections. A fit $e^L_1(\beta^L_e) - e^L_0 = b/L^D$ works only for $L \geq 256$ ($\chi^2/d.o.f. = 1.90/2$; for $e^L_1(\beta^L_e)$ we get $\chi^2/d.o.f. = 1.41/2$). However, $\beta^{\text{strip}, L}$ (see caption to Table II) is compatible with $\beta^{\infty}_e$ for $L \geq 256$. Then, the simplest strategy to get $\beta^{\infty}_e$ and the latent heat is: (1) for $L$ large enough to display a strip phase, locate it with short runs, (2) get $\beta^{\text{strip}, L}$ accurately, and (3) find the leftmost(rightmost) root for $\langle \hat{\beta} \rangle_e = \beta^{\text{strip}, L}$.

As for $\Sigma^L_e$, the inequality $\Sigma^L_e \leq 0.0473505$ [27] (equality under the hypothesis of complete wetting) was violated by $1/L$ extrapolations performed with $L \leq 100$ [3]. The reader may check (Table II) that our data for $L \leq 256$ extrapolate above 0.0473505, but drop below for $L \geq 512$. Indeed, the consistency of our results for $\beta^L_e$ imply that the integration error for $\langle \hat{\beta} \rangle_e$ is (at most) $2 \times 10^{-6}$ for $L = 1024$. Hence, the integration error for $\Sigma^L_e$ is at most $10^{-3}$. Adding it to the difference between the linear and the step-like interpolation, Fig. I, we obtain $\Sigma^L_e = 0.043(2)$, which is slightly below 0.0473505.

As for $(Q = 4, D = 3)$, see Table II, $\beta^{\text{strip}, L}$ has converged (within accuracy) for $L \geq 64$. Hence, our preferred estimate is $\beta^{\infty}_e = 0.6286206(10)$, that may be compared with Janke and Kapler’s $\beta^{\infty}_e = 0.628632(2)$ [16]. Accordingly, we find $e^L_o(\beta^{\text{strip}, L}) = -1.10537(4)$, $e^L_d(\beta^{\text{strip}, L}) = -0.52291(2)$, $C_L(e^L_o(\beta^{\text{strip}, L})) = 35.4(9)$, and $C_L(e^L_d(\beta^{\text{strip}, L})) = 4.24(18)$. The reader will note that $\beta^{L=128}_e$ is far too high (for instance, from the $\chi^2$ for $D=4$ extrapolation of $\beta^L_e = \beta^{\infty}_e + c/L^D$). Therefore, the integration error is $\sim 4 \times 10^{-6}$ (larger than the statistical one), which provides a bound for the error in the surface tension: $\Sigma^L=128 = 0.0118(4)$. This is compatible with $\Sigma^L=64$, and provides a reasonable $\Sigma^L$.

We propose a microcanonical strategy for the Monte Carlo simulation of first-order phase transitions. The method is demonstrated in the standard benchmarks: the $Q = 10, D = 2$ Potts model (where we compare with exact results), and the $Q = 4, D = 3$ Potts model. For both, we obtain accurate results in systems with more than $10^6$ spins (preexisting methods handle $10^4$ spins). Envisaged applications include first-order transitions with quenched disorder [16, 28], colloid crystallization [29], peptide aggregation [30] and the condensation transition [11].

We thank for discussions L. A. Fernandez (who also helped with figures and C code), L. G. Macdowell, W. Janke, G. Parisi and P. Verrocchio, as well as BIFI and the RTN3 collaboration for computer time. We were partly supported by BSCH—UCM and by MEC (Spain) through contracts BFM2003-08532, FIS2004-05073.
\[
\begin{array}{cccccccc}
L^D & \beta_c^L & \Sigma^L & -e_c^L(\beta_c^L) & -e_c^L(\beta_c^L) & -c_L(e_c^L(\beta_c^L)) & -c_L(e_c^L(\beta_c^L)) & \beta^{\text{strip},L} \\
32^2 & 1.423082(17) & 0.05174(9) & 1.3318(2) & 0.5736(3), & 5.13(13) & 3.99(7) & 1.42028(7) \\
64^2 & 1.425287(9) & 0.05024(11) & 1.3220(2) & 0.5999(2), & 6.44(17) & 5.78(19) & 1.42479(4) \\
128^2 & 1.425859(7) & 0.049225(14) & 1.31676(16) & 0.61164(16), & 7.4(3) & 7.8(3) & 1.42592(2) \\
256^2 & 1.426021(5) & 0.0488(2) & 1.31478(8) & 0.61578(8), & 8.0(3) & 8.7(4) & 1.42604(2) \\
512^{(A)} & 1.426051(4) & 0.0473(3) & 1.31399(6) & 0.61710(4), & 8.6(4) & 9.1(4) & 1.426048(4) \\
512^{(B)} & 1.426048(4) & 0.0467(4) & 1.31399(6) & 0.61708(5), & 8.6(4) & 9.1(4) & 1.426048(4) \\
1024^{(A)} & 1.4260599(19) & 0.0430(3) & 1.31373(5) & 0.61748(3), & 9.7(5) & 8.7(4) & 1.426066(9) \\
1024^{(B)} & 1.4260600(18) & 0.0424(2) & 1.31373(5) & 0.61748(3), & 9.7(5) & 8.7(4) & 1.426066(9) \\
\infty^2 & 1.426062438(9) & \infty^2 & 0.6286133(7) & 0.005591(10) & 5.13(13) & 3.99(7) & 1.426062438(9) \\
8^4 & 0.6276394(7) & 0.005591(10) & 1.1553(7) & 0.51412(12), & 23.0(5) & 3.856(16) & 0.626254(4) \\
16^3 & 0.6284403(8) & 0.0079596(6) & 1.189(4) & 0.51818(5), & 30.1(8) & 3.620(13) & 0.626687(15) \\
32^3 & 0.6285957(10) & 0.00824(6) & 1.10751(15) & 0.522606(16), & 34.2(9) & 4.019(17) & 0.627889(6) \\
64^3 & 0.6286133(7) & 0.011557(6) & 1.10542(8) & 0.522831(8), & 33.2(9) & 4.11(2) & 0.628621(3) \\
128^{(A)} & 0.6286237(5) & 0.011778(7) & 1.10548(3) & 0.522932(2), & 35.4(9) & 4.25(17) & 0.6286206(10) \\
128^{(B)} & 0.6286239(5) & 0.011674(9) & 1.10549(2) & 0.522932(2), & 35.4(9) & 4.25(17) & 0.6286206(10) \\
\end{array}
\]

Table I: System size dependent estimates of the quantities characterizing the first order transition, as obtained for the \( Q = 10, D = 2 \) Potts model (top) and \( Q = 4, D = 3 \) (bottom). Errors are Jack-Knife's. Also shown is \( \beta^{\text{strip},L} = \langle \beta \rangle_{\text{strip}} = -0.764443 \) (for \( D = 2 \)) or \( \beta^{\text{strip},L} = \langle \beta \rangle_{D = 3} \). The results with superscript \( A \) obtained with the linear(step-like) interpolation scheme, see Fig. 1.