

Edwards-Anderson & Random Field Models:  
Dynamical Correlation Length  
Continuous Time Monte Carlo  
Parallel Tempering

Camille ARON

Supervised by Leticia CUGLIANDOLO & Marco PICCO  
Laboratoire de Physique Théorique et Hautes Énergies  
Université Pierre et Marie Curie

May 17, 2007

## **Abstract**

Coming soon...

## Spin Glasses & Disorder

Spin glasses are magnetic systems in which the interactions between the magnetic moments are “in conflict” with each other, due to some frozen-in structural disorder. Thus no conventional long-range order (of ferromagnetic or antiferromagnetic type) can be established. Nevertheless these systems exhibit a “freezing transition” to a state with a new kind of order in which the spins are aligned in random directions.

Aging means that older systems relax in a slower manner than younger ones. One defines the age of a system as the time spent in the phase under study. The aging properties will be studied by monitoring the time evolution of correlation functions.

Experimental samples blabla...

# Chapter 1

## Two models under observation

Two varieties of disorder are encountered in spin models: randomness in the strength of the bonds and randomness in the strength of an externally applied magnetic field. We study these two types of models by means of their most emblematic representants: the former being the Edwards-Anderson model and the latter being the Random-Field model. We will focus on two three-dimensional Ising models. Spins are bimodal Ising variables:  $s_i = \pm 1, i = 1, \dots, N = L^3$ .  $L$  being the lattice linear size.

### 1.1 Edwards-Anderson (EA)

The Edwards-Anderson spin-glass model is defined by its hamiltonian

$$(1.1) \quad H^{\text{EA}} = - \sum_{\langle i,j \rangle} J_{ij} s_i s_j .$$

The interaction strength  $J_{ij}$  act on nearest neighbours on a cubbic three-dimensional lattice and are independant identically distributed random variables. In this study, we will adopt two different probability distribution laws for thoses  $J_{ij}$ 's: a gaussian distribution with zero mean and unit variance ( $\overline{J_{ij}} = 0, J^2 \triangleq \overline{J_{ij}^2} = 1$ ), or a bimodal distribution ( $J_{ij} = \pm 1$ ). We will see that physical quantities do not strongly depend on the very choice of this distribution.

Simulating an instantaneous quench from infinite temperature is realised by choosing a random initial condition:  $s_i(t = 0) = \pm 1$  with probability one half. If at high temperatures, the systems behaves like a paramagnetic material in thermodynamic equilibrium, under a critical temperature  $T_g \simeq 0.92$ , the system experiences a breakdown of time translation invariance: two-time observables<sup>1</sup>  $\mathcal{O}(t_w, t)$  cannot be written as function of the time difference  $\mathcal{O}(t_w - t)$ . This aging property of this so-called glassy phase does not show any tendency to stop within the accessible time-window: equilibrium is far from being reached. One can tell that the relaxation time becomes greater than the typical observation time. In this study, we will focus on the physic of this glassy phase.

#### 1.1.1 Computation

A good compromise to observe the dynamic of this glassy phase within a reasonable computational time is to quench the system down to a temperature of  $T \simeq 0.60$ . In this work,

---

<sup>1</sup> $t_w$  stands for "waiting-time".

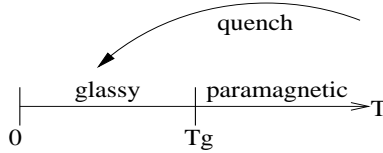


Figure 1.1: The 3D Edwards-Anderson phase diagram exhibits a transition from a paramagnetic to a glassy phase at  $T_g \simeq 0.92$ .

most results are obtained using lattices of linear size  $L = 60$  (representing  $N = 216\,000$  spins) at temperature  $T = 0.60$ . Periodic boundary conditions are used in order to reduce finite size effects. The dynamic is reproduced by a standard Monte Carlo algorithm, using Metropolis *et al.* prescription. With these parameters, it takes around a full day to compute  $10^6$  Monte Carlo sweeps<sup>2</sup> on a 3 GHz processor.

### 1.1.2 Observables

#### Global correlation and order parameter

The two-time self-correlation, commonly defined as

$$(1.2) \quad C(t, t_w) \triangleq \frac{1}{N} \sum_{i=1}^N s_i(t) s_i(t_w),$$

quantifies how two spin configurations of the same system, the one taken at  $t_w$  and the other one at  $t \geq t_w$ , are close to each other. In the large  $N$  limit, this quantity is self-averaging with respect to noise and disorder induced fluctuations. In figure 1.1.2 we clearly see this two-time dependence of the correlation for  $T = 0.60 < T_g$ , whereas for  $T = 2.0$  the correlation decays exponentially with  $t - t_w$  regardless of  $t_w$ . In spin glasses no consensus as to which is the origin of aging has been reached. Still, the qualitative behavior of correlations is rather close to the one in domain growth. Figure 1.1.2 shows that if comparing the configurations at  $t_w$  and at a later time  $\tau + t_w$ , one finds a clear separation of time-scales depending on the relative value of  $\tau$  with respect to  $t_w$ . For  $\tau \ll t_w$ , the correlation function decays regardless of  $t_w$ : it can be interpreted as the effect of thermal fluctuations within motionless domains. For  $\tau \gg t_w$  the decay to zero is interpreted as the result of domain-wall motion. Of course the concept of domain is here a lot less intuitive than with coarsening systems and has still to be precised. In the limit  $t \geq t_w \gg t_0$ , with  $t_0$  some microscopic time scale, the correlation can be written as a sum of two terms representing these two regimes:

$$(1.3) \quad C(t, t_w) = C_{\text{thermal}}(t - t_w) + C_{\text{aging}}(t, t_w)$$

with the limit conditions

$$(1.4) \quad \begin{aligned} C_{\text{thermal}}(0) &= 1 - q_{\text{EA}}, \quad \lim_{t_w \rightarrow t^-} C_{\text{aging}}(t, t_w) = q_{\text{EA}}, \\ \lim_{t \rightarrow t_w \rightarrow \infty} C_{\text{thermal}}(t - t_w) &= 0, \quad \lim_{t \gg t_w \rightarrow \infty} C_{\text{thermal}}(t, t_w) = 0. \end{aligned}$$

This defines  $q_{\text{EA}}$ , the order parameter, which we see is not trivial to compute.

<sup>2</sup>1 MC sweep  $\equiv N$  spins were tried to be flipped.

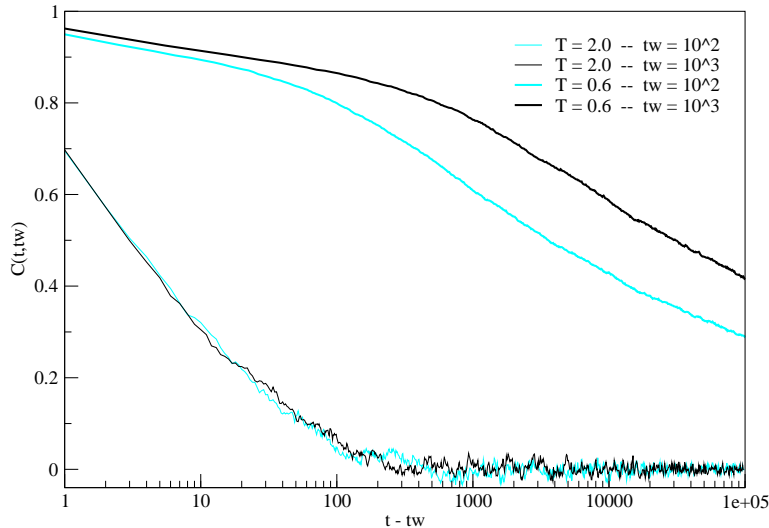


Figure 1.2: The two-time self-correlation  $C(t, t_w)$  is a function of  $t - t_w$  above  $T_g$ , whereas in the glassy phase, the time translation invariance is lost.  $t - t_w \sim t_w$  separates two time-scales associated with thermal and aging regimes.

### Correlation length

As in usual ordering processes, one would like to identify a correlation length and determine its temperature and time-dependence. A two-time correlation length,  $\xi(t, t_w)$ , can be extracted from the exponential decay of two-site two-time correlation

$$(1.5) \quad C_4(r; t, t_w) \triangleq \frac{1}{N} \sum_{i,j} [s_i(t)s_i(t_w)s_j(t)s_j(t_w) - C^2(t, t_w)]_{|\vec{r}_i - \vec{r}_j|=r}.$$

Figure 1.1.2 gives an example of a typical  $C_4$  function: exponential fitting is very good (at least at this temperature –  $T = 0.60$  – far below  $T_g$ ). The results of this analysis are shown in figure 1.1.2 where we plot  $\xi(t, t_w)$  as a function of  $t - t_w$  for three values of  $t_w$  given in the caption. A better representation of the same data is given in figure 1.1.2 where we display  $\xi$  as a function of  $1 - C$ , evaluated at the same times. The curves are now monotonic in both  $1 - C$  and  $t_w$ . Note that the values of the correlation length obtained are extremely short (no more than a few lattice spacings). It makes it reasonable to speak of thermodynamic limit for lattices of size  $L = 60$ .

### Local correlations

The random nature of the interactions naturally introduces a different dynamic from site to site. This local dynamics can be described by two-time spin-spin correlations, which instead of being spatially averaged over the whole bulk, are only averaged over a coarse-

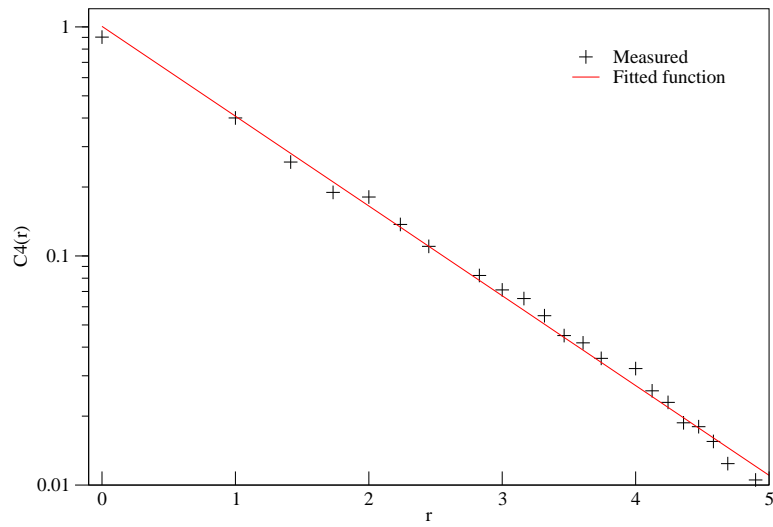


Figure 1.3: The correlation length  $\xi$  is extracted from the exponential behavior of  $C_4(r)$ . Here we show an example of  $C_4(r; t, t_w)$  for  $t_w = 10^4$  and  $t = 10^5$ .

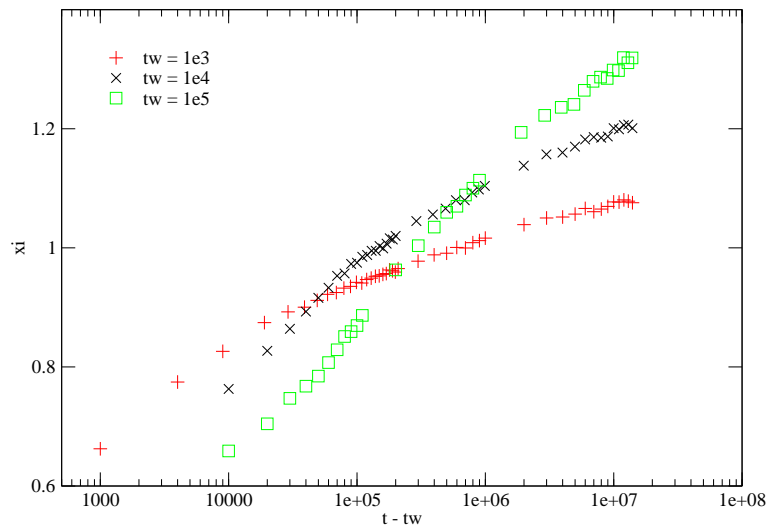


Figure 1.4: The correlation length  $\xi$  is an ever-growing function of  $t - t_w$ .

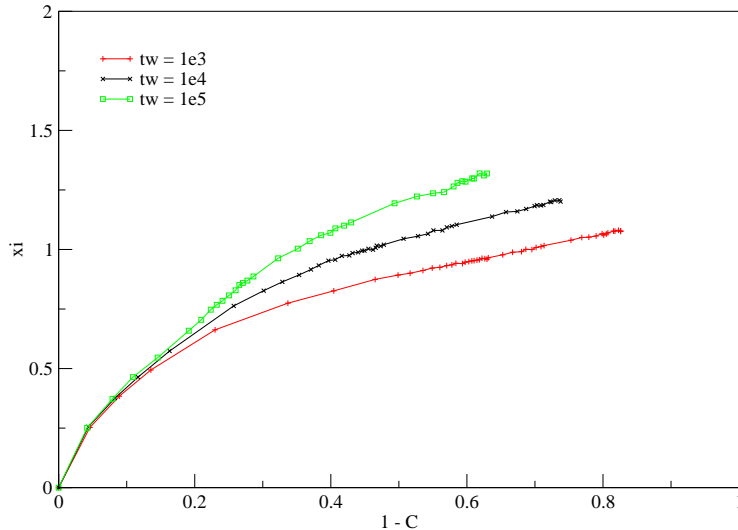


Figure 1.5: The correlation length  $\xi$  is a monotonic function of both  $1 - C(t, t_w)$  and  $t_w$ .

graining cell with volume  $V_r = l^3$  centered at some site  $r$  :

$$(1.6) \quad C_r(t, t_w) = \frac{1}{V_r} \sum_{i \in V_r} s_i(t) s_i(t_w)$$

The values of such local correlations vary spatially, and can be interpreted as representing spatially heterogeneous ages of the dynamical evolution. We study the probability distribution function (pdf) of  $C_r$ :  $\rho(C_r; t, t_w, l, L)$ . At fixed temperature, the pdf  $\rho(C_r; t, t_w, l, L)$  depends on four parameters, two times  $t, t_w$  and two lengths  $l, L$ . The dependence on  $t$  and  $t_w$  can be replaced by a dependence on  $C(t, t_w)$  and  $\xi(t, t_w)$ , the former being the global correlation and the latter the correlation length. Indeed, we show that  $C(t, t_w)$  is a monotonic function on the two times (c.f. figure 1.1.2) and  $\xi$  is a growing function of  $1 - C$  (c.f. figure 1.1.2), thus allowing the inversion  $(t, t_w) \rightarrow (C, \xi)$ . Let us make the scaling assumption that the pdfs depend on the coarse-graining length  $l$  and the system linear size  $L$  through the ratios  $l/\xi$  and  $\xi/L$ . In the end, the pdfs characterizing the heterogeneous aging of the system can be written as

$$(1.7) \quad \rho(C_r; C(t, t_w), l/\xi(t, t_w), \xi(t, t_w)/L).$$

We numerically test this proposal assuming that the thermodynamic limit applies so that the last scaling ratio disappears. Holding  $C$  and the ratio  $l/\xi$  constant gives a full collapse of datas as shown in figure 1.1.2.

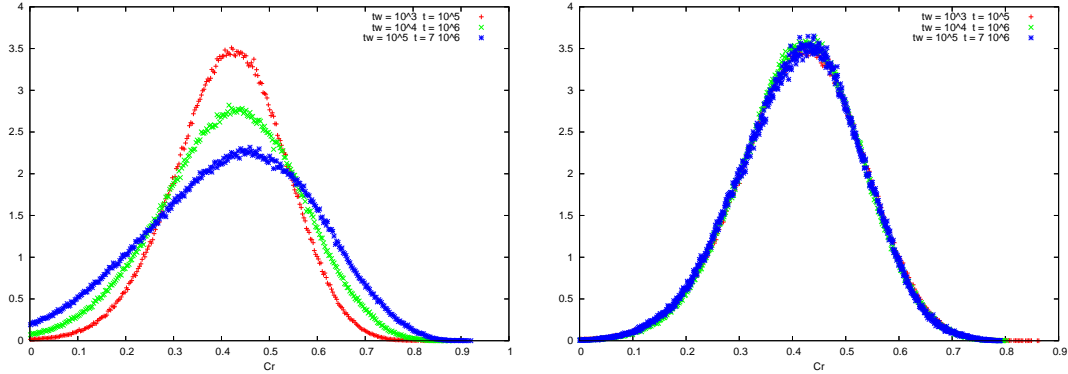


Figure 1.6: Pdf of local correlation  $C_r$ . (left)  $C_r$  is coarse-grained on boxes of linear size  $l = 9$  at  $C = 0.6$  fixed. The curves correspond to three waiting-times given in the key. (right)  $C_r$  is coarse-grained, at  $C = 0.6$  fixed, on boxes with length  $l$  chosen to keep  $l/\xi$  constant. The curves correspond to the same  $t_w$ 's with respectively  $l = 9, 11, 13$ . The collapse is improved with respect to the left figure.

## 1.2 Random Field (RF)

The Random-Field model is defined by its hamiltonian

$$(1.8) \quad H^{\text{RF}} = -J \sum_{\langle i,j \rangle} s_i s_j - \sum_i h_i s_i.$$

Here again,  $J > 0$  is the strength of the short range ferromagnetic interactions between spins,  $h_i$  represents a local random magnetic field on site  $i$ . We will adopt two different distribution functions for those independent identically distributed random variables: gaussian ( $\overline{h_i} = 0$ ,  $\overline{h_i^2} = 1$ ), or bimodal ( $h_i = \pm 1$ ). In the limit  $J \gg 1$ , the model reduces to the well known Ising model with a phase transition from paramagnetic to ferromagnetic state occurring at  $T_c/J = 4.515$ . In the limit  $J \ll 1$ , the random magnetic field destroys all possible long range order. Fine tuning the temperature  $T$  and the spin-spin coupling  $J$ , we can exhibit a spin glass domain. It's aging regime can be observed during a relatively long time if the lattice is large enough ( $L > 200$ ). After that period, the system finally adopts a ferromagnetic order. For a bimodal distribution of  $h_i$  and taking  $T = 1$  and  $J = 1$ , figure 1.2 shows that the global correlation function  $C(t, t_w)$  is the one of an aging system, very similar to the one in EA model (c.f. figure 1.1.2). Here again, we can distinguish two regimes : the thermal one for  $t - t_w < t_w$  where correlation only decays because of thermal fluctuations, and the aging regime for  $t - t_w > t_w$  due to some domain-wall motion.

Using the exponential decay of the two-time two-site correlation function  $C_4(r; t, t_w)$  defined in 1.1.2, we can extract a two-time correlation length  $\xi(t, t_w)$ . Results of this analysis are shown in figure 1.2 where we plot  $\xi(t, t_w)$  versus  $1 - C(t, t_w)$ . The behaviour of  $\xi$  is very similar to the one obtained with EA model (c.f. figure 1.1.2), it is monotonic in both  $1 - C$  and  $t_w$ . Nevertheless, in this coarsening system, the values of the correlation length are much higher than with EA model.

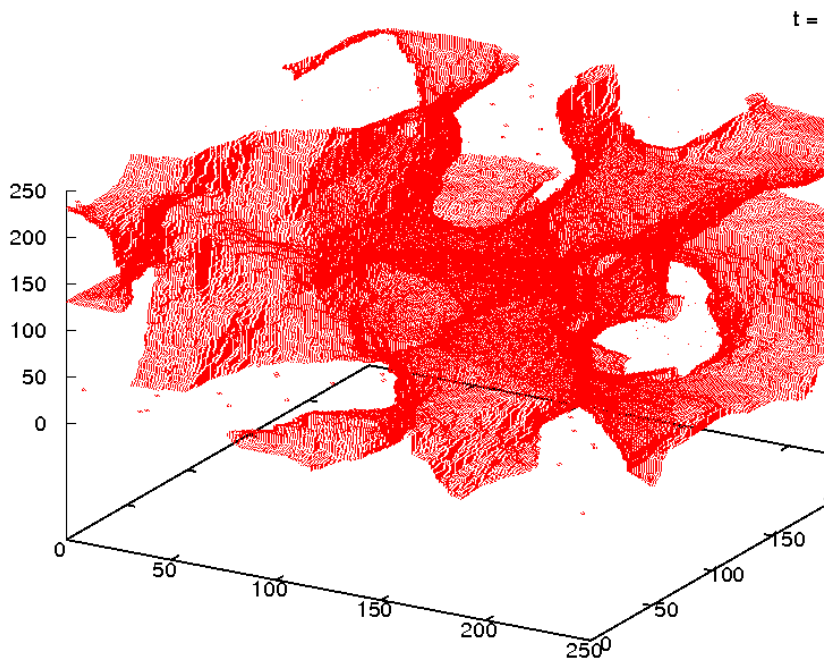
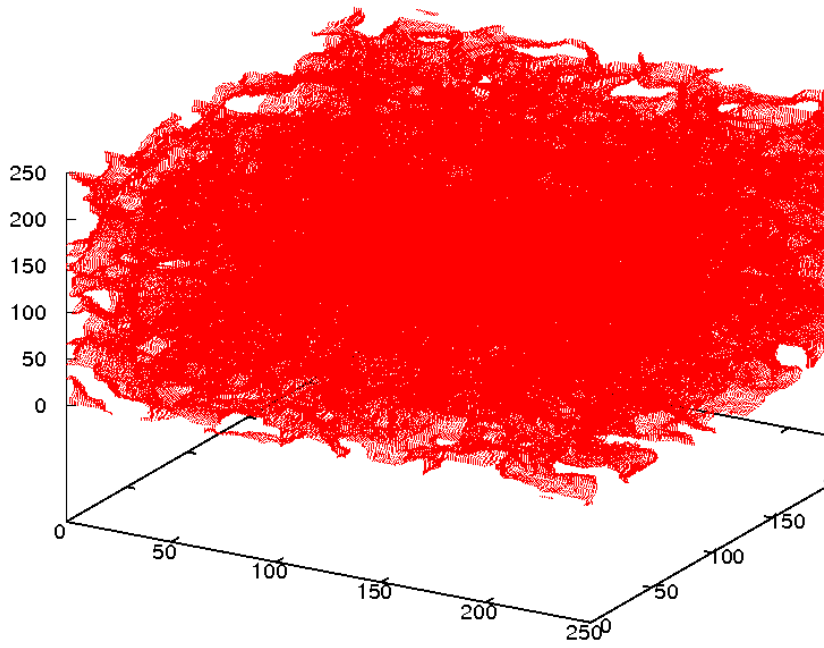


Figure 1.7: (up) Image of the domain-wall structure in a RF lattice of size  $L = 250$ , at temperature  $T = 1$  and at  $t = 10^3$  ( $J = 1$ , bimodal  $h_i$ ). (down) Image of the same system at  $t = 4.85 \cdot 10^6$ .

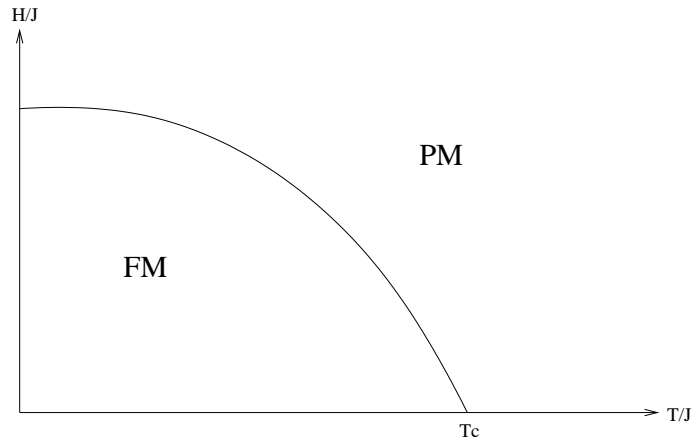


Figure 1.8: The Random Field model phase diagram shows that the critical temperature between the paramagnetic (PM) and the ferromagnetic (FM) order is a decreasing function of the disorder strength  $H/J$ .

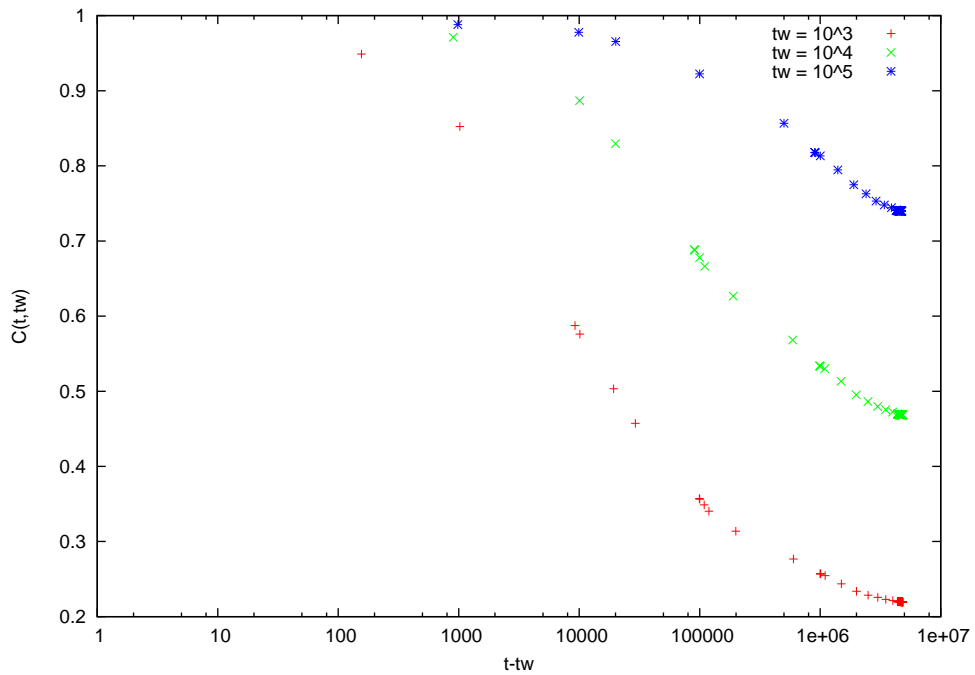


Figure 1.9:  $C(t, tw)$  given for  $t_w = 10^3, 10^4, 10^5$  was obtained with a lattice of size  $L = 250$  quenched to  $T = 1$ , with  $J = 1$  and a bimodal distribution of  $h_i$ . This two-time behaviour is the one of a glassy phase.

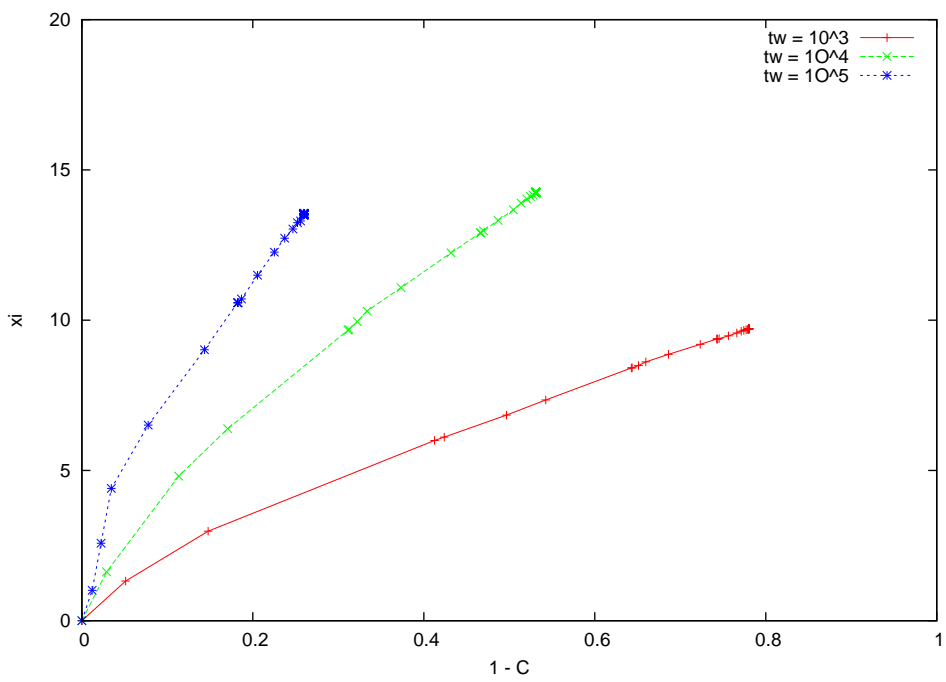


Figure 1.10: The correlation length  $\xi$  is a monotonic function of both  $1 - C(t, t_w)$  and  $t_w$ .

## Chapter 2

# Continuous Time Monte Carlo

### 2.1 The trouble with Metropolis

The standard Monte Carlo algorithm generates a markovian sequence of configurations as follows :

- a spin is selected with uniform probability amongst all spins ;
- the probability  $P$  of flipping this spin is computed according to a thermodynamically reasonable formula: if  $\mathcal{C}$  and  $\mathcal{C}'$  are two configurations separated by one spin flip with energy cost  $\Delta E$ , then  $P$  must satisfy the detailed balance equation:  $P(\mathcal{C} \rightarrow \mathcal{C}')/P(\mathcal{C}' \rightarrow \mathcal{C}) = \exp(-\Delta E/T)$  ;
- a random number  $R$  is drawn over the interval  $[0; 1]$  ;
- the reversal is performed if  $R < P$ , else it is rejected.

The number of attempts can then be used as a measure of time. The Metropolis *et al.* solution for  $P$  is given by

- $P = 1$ , if  $\Delta E \leq 0$  ;
- $P = \exp(-\Delta E/T)$ , otherwise.

The rejection rate thus becomes quite high if a typical spin flip costs, energetically speaking,  $\Delta E \gg T$ . A huge number of trials is then required to generate new configurations. For instance, the Edwards-Anderson model sees its rejection rate grow above 95% in the glassy phase (c.f. figure 2.1).

### 2.2 Algorithms

The continuous time Monte Carlo algorithm is a rejection free method: it generates a new configuration with every choice of spin. Each step of the method involves computing the time to leave the current state (the “waiting-time”  $\tau_i$  which is a random variable) rather than attempting moves that will be rejected. It can also be seen as a method accounting for the *a priori* probability of reversal before, rather than after, choosing the site to change. In that sense, it is a re-organization of the standard Monte Carlo algorithm.

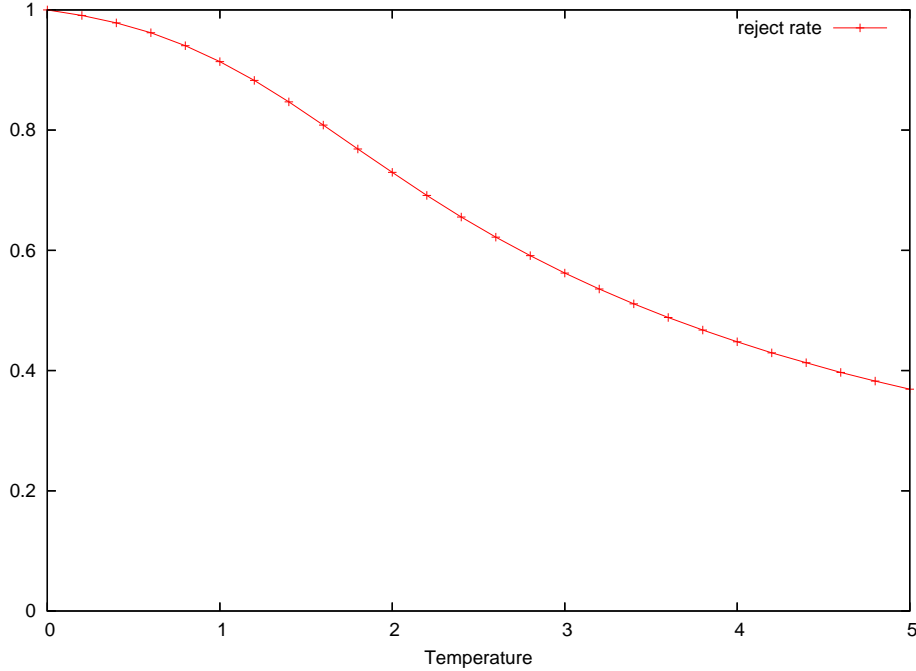


Figure 2.1: With the EA model, Metropolis Monte Carlo algorithm has a rejection rate above 95 % in the glassy phase ( $T < T_g \simeq 0.92$ ).

In the standard MC algorithm, when the probability to reject the change is  $\lambda_i$ , the probability  $p_i(t)$  to remain in the current state for  $t$  time steps decays exponentially as

$$(2.1) \quad p_i(t) = \lambda_i^t = \exp(t \ln \lambda_i).$$

The “waiting-time”  $\tau_i$  has thus to be

$$(2.2) \quad \tau_i = \frac{\ln R}{\ln \lambda_i},$$

where  $R$  is a uniform random number on  $]0; 1]$ .

So each spin  $s_i$ , according to the energy cost of its reversal  $\Delta E_i$ , is being attributed a random “waiting-time”  $\tau_i = -\max[1, \exp(-\Delta E_i/T)] \ln R$  where  $R$  is a uniform random number on the interval  $]0; 1]$ . Afterwards the procedure is :

- spin  $s_{i_0}$  having the shortest flipping time  $\tau_{i_0}$  is flipped ;
- $\tau_{i_0}$  is reactualised, as well as  $\tau_j$ 's of nearest neighbours  $s_j$  of  $s_{i_0}$  ;
- time is updated:  $t \leftarrow t + \tau_{i_0}$ .

The principal drawback of this method stands on the necessity to maintain updated an ordered list of  $N$  spins, sorted in terms of  $\tau_i$ . If the computational time spent to create this ordered list from scratch is spent once and for all, the overhead number of manipulations required to keep it updated after each spin flip, is proportionnal to  $(z + 1)N$ ,  $z$  beeing the number of nearest neighbours<sup>1</sup>. One can see that this leads to a serious slow down of the efficiency of this so-called rejection free scheme.

---

<sup>1</sup>In our 3D case  $z = 6$ .

Nevertheless, adopting a convenient model allows us to get rid of this sorting issue. Let us stick to a bimodal distribution of the  $J_{ij}$ 's (resp. the  $h_i$ 's). Therefore the energy cost to flip a spin can only take a finite number of values :

$$(2.3) \quad \Delta E_i^{\text{EA}} = 2s_i \left\{ \sum_{j \in v(i)} J_{ij} s_j \right\}$$

$$(2.4) \quad \Delta E_i^{\text{RF}} = 2J s_i \left\{ \sum_{j \in v(i)} s_j + h_i \right\}$$

For instance, in the Edwards-Anderson model, it can take seven different values so that spins can be gathered in seven different boxes :

$s_i \sum_{j \in v(i)} J_{ij} s_j$	$\Delta E_i^{\text{RF}}$	Box #
6	12	1
4	8	2
2	4	3
0	0	4
-2	-4	5
-4	-8	6
-6	-12	7

Each spin flip requires the following calculations :

- a box  $c$  is chosen with probability  $N_c P_c / Q$ , where  $N_c$  is the number of spins in box  $c$  and  $Q = \sum_c N_c P_c$  ;
- a spin  $s_{i_0}$  is chosen with uniform probability within the box  $c$  ;
- $s_{i_0}$  is flipped ;
- the flipped spin is moved to its new box according to its new state, as well as its nearest neighbours ;
- the time is updated  $t \leftarrow t - \ln(R)/Q$  where  $R$  is a uniform random number drawn within  $]0; 1]$ .

It is straightforward that this last algorithm reduces to the former one if the number of boxes goes to  $N$ . Using this algorithm, it is not necessary anylonger to maintain updated an ordered list of  $N$  spins: this problem reduces to the one of arranging spins in a few boxes, with no matter of how spins are classified in each box.

## 2.3 Benchmarking

Looking back at this algorithm, we see that the time spent for one Monte Carlo sweep is strictly proportional to  $N$ , the total number of spins. Moreover, it is rejection free: no computational time is spent trying to flip spins that do not have a favorable neighbourhood. The speed of this algorithm is therefore completely independent of the temperature. If in standard Monte Carlo, the physical effect of the temperature is “coded” in the rejection rate making it uneasy to obtain old samples, here it is directly coded on the age of the sample:

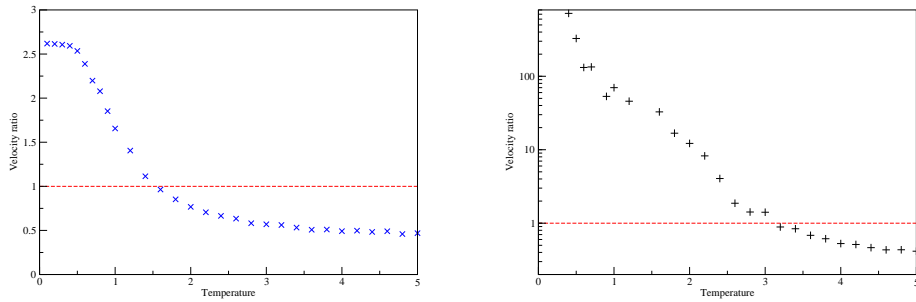


Figure 2.2: (left) EA : The velocity ratio of continuous time algorithm over standard MC is larger than 1 in the glassy phase. (right) RF : it is much larger than 1 for  $T < 3$ . ( $J=1$ ).

a sample that has experienced 1000 spin flips at very low temperature has to be older than its twin brother also having experienced 1000 spin flips but at a higher temperature. We make the following measurements with the EA model: for different temperatures we measure the time (in seconds) needed to obtain a system of a certain age with standard MC and with continuous time algorithm on a 2.6 MHz processor. Of course measured times depend on the very structure of the simulation programs, however a great deal of attention has been paid to each program efficiency (to minimize the number of operations, and the memory access.). In order to get rid of these technical contingencies, we present the results in the ratio of measured time for standard MC and continuous time versus temperature. We repeat this operation for two waiting times ( $t_w = 0$  and  $t_w = 1000$ ). Figure 2.3 shows clearly that in the glassy phase ( $T < T_g \simeq 0.92$ ) it is worthwhile to use continuous time algorithm.

## Chapter 3

# Tempering

We study the possibility to reproduce the aging dynamics of glassy systems by use of the Parallel Tempering algorithm. This algorithm uses several copies of the same physical system, at different temperatures. Each copy freely evolves during a certain period of time with standard Monte Carlo dynamics, until copies are periodically offered the possibility to exchange their temperature.

The basic ideas behind this algorithm relies on the reality that at low temperatures, in spin glasses and disordered systems, a lot of local energy minima appear in the phase space, separated by barriers. In order to properly thermalize the system, all these regions have to be visited. But the characteristic time needed by the system to escape from a local minimum increases very rapidly upon lowering the temperature. We therefore need to help the system escaping local minima, avoiding in this way long relaxations. Introducing additional possibilities of warming and cooling through the exchanges, the system will explore more easily the complicated phase space. The formalism makes use of an extended canonical ensemble made of  $M$  different replicas of the sample we want to simulate, which is described by a Hamiltonian  $H(X)$ . With  $X$  we denote, for simplicity, all the microscopic spin variables. To the  $m$ -th replica we associate an inverse temperature  $\beta_m$ . The state of the extended system is specified by the set of  $\mathbf{X} \triangleq X_1, \dots, X_M$ . With  $\beta \triangleq \beta_1, \dots, \beta_M$  we indicate the set of temperatures, and for simplicity we assume  $\beta_m \beta_{m+1}$ . The partition function is given by

$$(3.1) \quad \mathcal{Z}(\beta) = \prod_{m=1}^M e^{-\beta_m H(X_m)} = \prod_{m=1}^M Z(\beta_m)$$

where  $Z(\beta_m)$  refers to the original system. Once the set  $\beta$  is specified, the probability distribution of  $\mathbf{X}$  is

$$(3.2) \quad \mathcal{P}(\mathbf{X}, \beta) = \prod_{m=1}^M P(X_m, \beta_m)$$
$$P(X_m, \beta_m) = \frac{e^{-\beta_m H(X)}}{Z(\beta)}$$

Introducing a transition matrix  $W(X_m, \beta_m \rightleftharpoons X_n, \beta_n)$ , which defines the probability of exchanging temperatures between two replicas with configurations  $X_m$  and  $X_n$  and temperatures  $\beta_m$  and  $\beta_n$ , the detailed balance is satisfied if

$$(3.3) \quad \mathcal{P}(\dots; X_m, \beta_m; \dots; X_n, \beta_n; \dots) W(X_m, \beta_m \rightleftharpoons X_n, \beta_n) = \\ \mathcal{P}(\dots; X_n, \beta_n; \dots; X_m, \beta_m; \dots) W(X_n, \beta_n \rightleftharpoons X_m, \beta_m).$$

Combining last equations we obtain the condition

$$(3.4) \quad \frac{W(X_m, \beta_m \rightleftharpoons X_n, \beta_n)}{W(X_n, \beta_n \rightleftharpoons X_m, \beta_m)} = e^{-\Delta} \text{ with } \Delta \triangleq (\beta_n - \beta_m)[H(X_m) - H(X_n)].$$

### 3.1 Algorithm

We will only focus on this “exchange” part of this algorithm. Let us consider two replicas of the same physical system (same realisation of disorder), at two different temperatures  $T_i$  and  $T_{i+1} = T_i + \delta T_i$  ( $\delta T_i > 0$ ). We note  $\Delta \triangleq (\beta_{i+1} E_i + \beta_i E_{i+1}) - (\beta_i E_i + \beta_{i+1} E_{i+1}) = \delta \beta_i (E_{i+1} - E_i)$  with  $\delta \beta_i = -\frac{\delta T_i}{T_i^2}$ . The rules for the exchange process are an extension of the usual Metropolis criterion :

- swap the temperatures of the two replicas if  $\Delta < 0$  ;
- swap the temperatures with probability  $P_{\text{swap}} = \exp(-\Delta)$  otherwise.

In order to measure the traditional observables and compare them with standard Monte Carlo dynamics, we first use the Parallel Tempering method until  $t_w$ , we then extract the replica being at  $T = 0.60$  at time  $t_w$ . We finally use standard Monte Carlo Metropolis algorithm on this particular system, and make the measurements on it.

### 3.2 Implementation

At this stage, this algorithm still has a lot of parameters to be fixed. How many replicas are necessary ? What should be the temperature range ? Which are the pairs of replicas that should be offered a temperature swap ? How often these swaps should be attempted ?

#### 3.2.1 Temperature sampling

To avoid a high rejection rate on the exchange process, the mean probability of swaps needs to be not too close to zero :

$$(3.5) \quad \langle P_{\text{swap}} \rangle = \langle \exp(-\Delta) \rangle = O(1).$$

It is clear that we can't say anything *a priori* on the relative values of  $E_i$  and  $E_{i+1}$ . Since

$$(3.6) \quad \langle \exp(-\Delta) \rangle \simeq \exp(-\langle \Delta \rangle) + \frac{1}{2}(\langle \Delta^2 \rangle - \langle \Delta \rangle^2) + \dots,$$

keeping the first moment (the other ones are difficult to estimate but are *a posteriori* negligible), we have :

$$(3.7) \quad \langle \Delta \rangle = \delta \beta_i (\langle E_{i+1} \rangle - \langle E_i \rangle),$$

$$(3.8) \quad \langle \Delta \rangle \simeq -\frac{\delta T_i}{T_i^2} \left. \frac{d \langle E \rangle}{dT} \right|_i.$$

We are thus able to estimate the temperature sampling by

$$(3.9) \quad \delta T_i = T_i \sqrt{\frac{\ln(P_{\text{swap}})}{-\left. \frac{d \langle E \rangle}{dT} \right|_i}}.$$

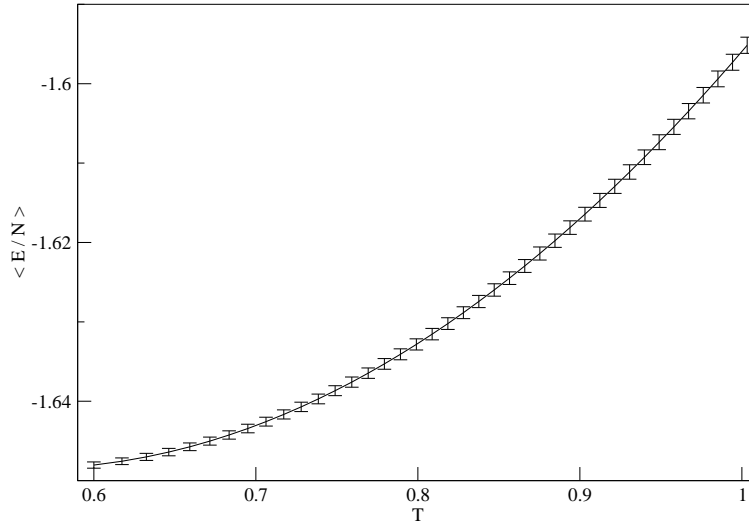


Figure 3.1: Mean energy per spin for the 39 replicas (3D EA, gaussian  $J_{ij}$ 's) sampling the temperature window  $T \in [0.60; 1.0]$ . Error bars indicates their measured energy variance. A reasonable rejection rate is guaranteed if those variances fill the mean energy gap between two adjacent replicas. We notice that the energy depends weakly on the temperature.

The energy  $E$  being extensive, it is noticeable that the number of replicas should be proportional to  $L^{3/2}$ . One can easily get  $\langle E \rangle (T)$  by numerically measuring the energy of a small system at different temperatures.

With Edwards-Anderson model,  $\langle E \rangle (T)$  can be modeled by a parabol (see figure 3.2.1). We obtain the following phenomenological law for Edwards-Anderson model :

$$(3.10) \quad \delta T_i = 1.36 \sqrt{-\ln(P_{\text{swap}})} L^{-3/2} f(T_i)$$

with  $f(T_i) \triangleq \frac{T_i}{\sqrt{T_i - 0.557}}$  which is bounded by 1.5 and 2.9 for a temperature window  $T \in [0.6; 1.0]$ , so the dependence on temperature is quite smooth. Satisfying results are obtained choosing  $\delta T \simeq 0.1$ .

With the random-field model, the temperature dependence of the energy is much stronger (c.f. figure 3.2.1) so that  $\delta T$  becomes drastically small ( $\delta T \simeq 0.001$ ). A too large number of replicas is needed, making it impossible to compute within a reasonable time on an ordinary computer. We conclude to the impossibility of using Parallel Tempering with this particular model. Note that using bigger lattices allows us to have a longer aging regime, and consequently a smoother temperature dependence of the energy (at least during this aging regime). However  $\delta T$  still remains too small since it is proportional to  $L^{-3/2}$ .

From now on, we will only focus on the Edwards-Anderson model, with a gaussian distribution of  $J_{ij}$ 's.

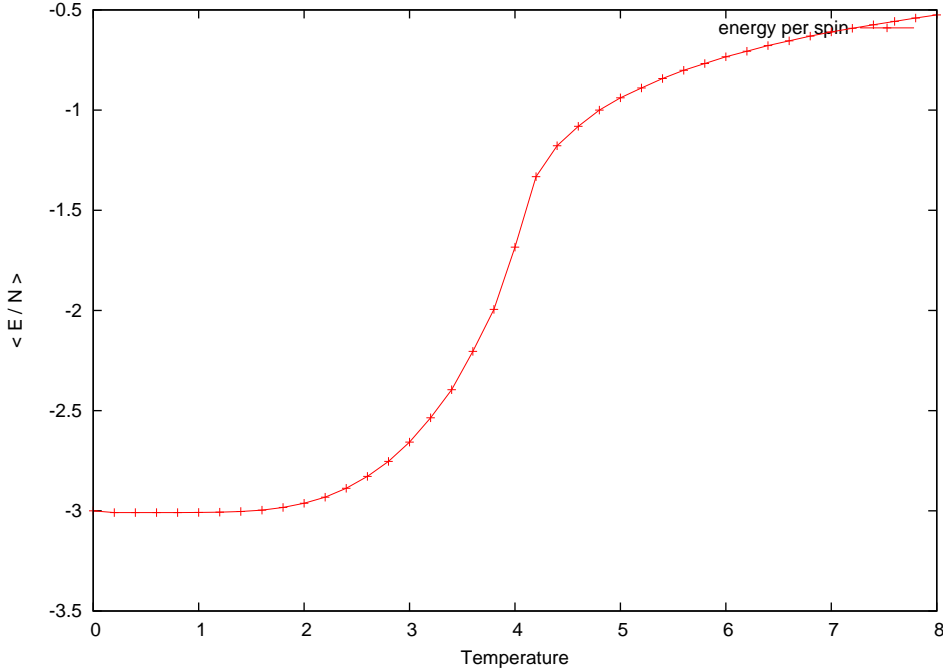


Figure 3.2: Mean energy per spin in the RF model ( $J = 1$ ). We shall notice that energy strongly depends on temperature.

### 3.2.2 Temperature range

We test the influence of the temperature range, all other parameters being kept constant. Several temperature windows are tried:  $[0.5; 1.0]$ ,  $[0.5; 1.1]$ ,  $[0.6; 1.0]$ ,  $[0.6; 1.1]$ ,  $[0.6; 0.9]$ . Clearly all windows exceeding  $T_g \simeq 0.92$  give similar results: a faster aging of the system. We choose the less time-consuming temperature range, that is to say the smallest:  $[0.6; 1.0]$ .

### 3.2.3 Tempering rate

We test different rates of attempts to exchange replica temperature: one every Monte Carlo sweep, one every ten, one every hundred. Reducing this rate clearly slows down the dynamics of replicas on the temperature axis. Nevertheless it does not have a strong influence on the correlation functions. Finally, we decide that a unitary ratio is the most natural choice.

### 3.2.4 Update Rules

Which are the pairs of replicas that should be offered a temperature swap ? We test two methods illustrated in figure 3.2.4:

It turns out that both rules gives similar results. Monitoring the temperature of a given replica (c.f. figure 3.2.4), we see as expected that it explores the whole temperature range. Since it is a spin glass, we clearly note a slowing down of this particular dynamic in time (time scale of figure 3.2.4 is logarithmic).

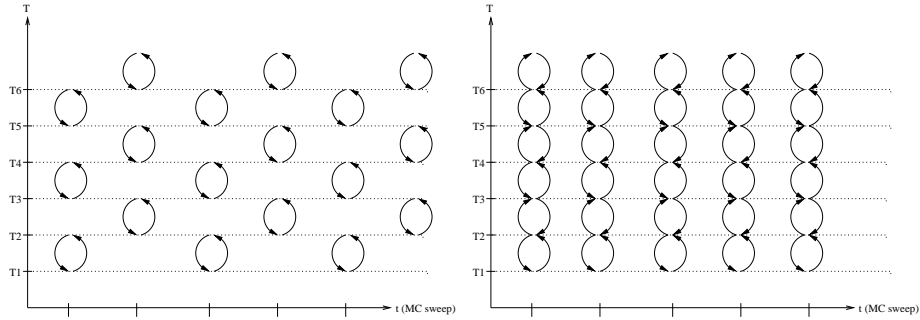


Figure 3.3: Representation of the exchange schemes for the parallel tempering method. (left) alternating scheme used to attempt to exchange temperatures between different systems. (right) a system can travel up to far temperatures in only one step.

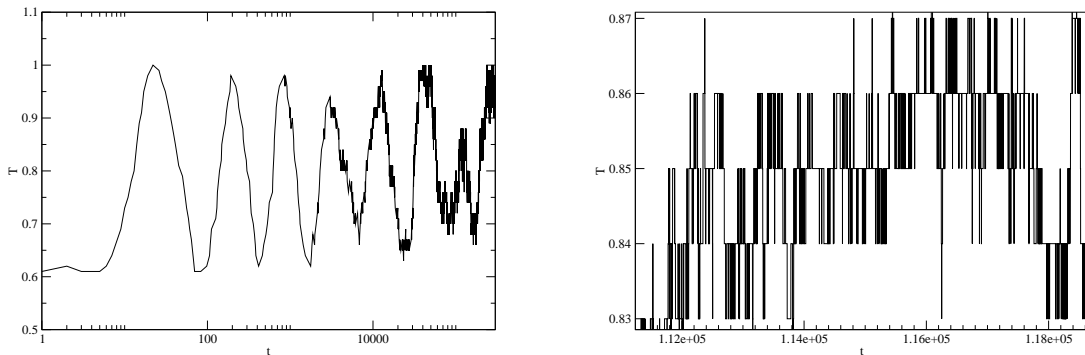


Figure 3.4: (left) The 'trajectory' of a particular system along temperature axis shows that replicas explore the whole temperature range. (right) Zoom between  $t = 1.1210^5$  and  $t = 1.1810^5$ .

### 3.3 Benchmarking

We just saw that the Parallel Tempering method, applied to the Edwards-Anderson model, is insensitive to a lot of parameters. Figure (3.3) gives an example of the two-time correlation function measured for a system prepared until  $t_w$  with standard MC and with Tempering. The correlation function of the system obtained with Tempering is the one of an older system than the one obtained with standard MC. From this point of view, Tempering is a success since it produces older systems for the same waiting time.

In order to quantify the velocity of this algorithm, we try to figure out the relation between the age of a system prepared with standard MC (given by the waiting time  $t_w$ ) and the effective age of a system prepared with Tempering (denoted by  $t_w^{\text{eff}}$ ). One way of doing so would be to prepare a system with Tempering, and then try to find which is value  $t_w^{\text{eff}}$  of the waiting time needed to obtain a system showing the same correlation function with a full standard MC procedure. This can be easily done since the age of a system can be determined with the change of correlation function slope around  $t - t_w \sim t_w$ . Note that these functions can never be superposed exactly entirely, but we are only interested in the aging part of it. From this first study, we guess a power law:  $t_w^{\text{eff}} \propto (t_w)^\alpha$ .

In order to get the exponent  $\alpha$  more precisely, we make the following measurements: for

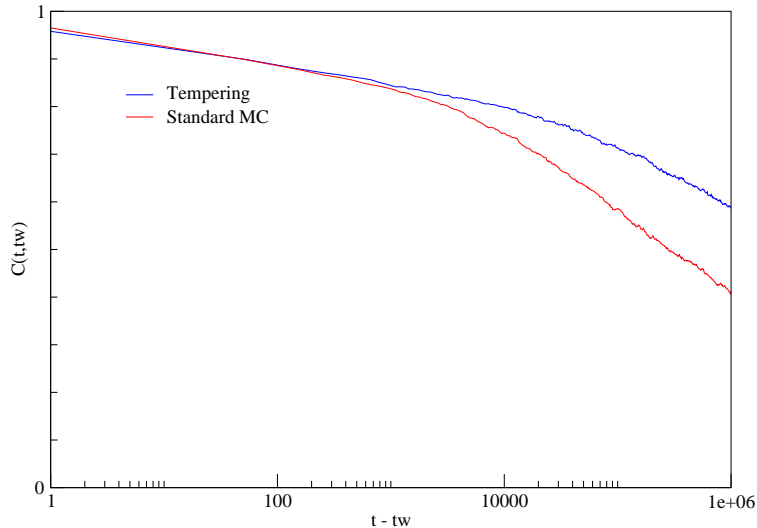


Figure 3.5: Global correlation  $C(t, t_w)$  is measured after the system evolved with standard MC or parallel tempering until  $t_w = 10^4$ . Parallel tempering dynamics give an older sample for the same  $t_w$ .

both methods, and for different waiting times, we measure the time  $\tau^* = t^* - t_w$  needed for the two-time correlation to reach  $C^* = 0.6$  (this particular value of  $C$  was chosen because it is situated in the aging regime). As shown in figure 3.3, we find power laws (fits being really precise)

$$(3.11) \quad \begin{aligned} \tau_{\text{Tempering}}^* &\propto (t_w)^{\alpha_T} \\ \tau_{\text{MC}}^* &\propto (t_w^{\text{eff}})^{\alpha_{\text{MC}}}, \end{aligned}$$

thus obtaining the following power law

$$(3.12) \quad t_w^{\text{eff}} \propto (t_w)^\alpha, \text{ with } \alpha \triangleq \frac{\alpha_T}{\alpha_{\text{MC}}} \simeq 1.27.$$

We repeat this operation for several other temperatures  $T$ , the behaviour of  $\alpha$  with  $T$  is given in the inset of 3.3.

Parallel Tempering is a heavy method to set up: a lot of parameters are to be tuned, and the simulation of a consequent number  $M$  of replicas until  $t_w$  is time-consuming and requires a certain amount of hardware facilities (especially memory). However, the power law between the age of a sample obtained with Tempering and the one obtained with standard MC, indicates that this new method is worthwhile if one wants to prepare very old samples. Moreover, we saw that its efficiency is increasing with the temperature going down.

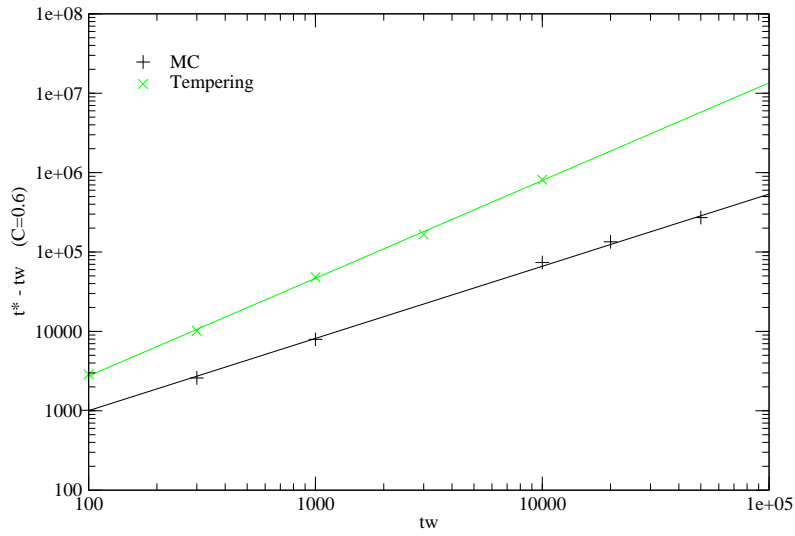


Figure 3.6: A quantitative comparison of Tempering versus standard Monte Carlo is done by measuring times needed to reach  $C = 0.6$  for different  $t_w$ . For a given  $t_w$ , the tempering process gives an older system than standard MC.

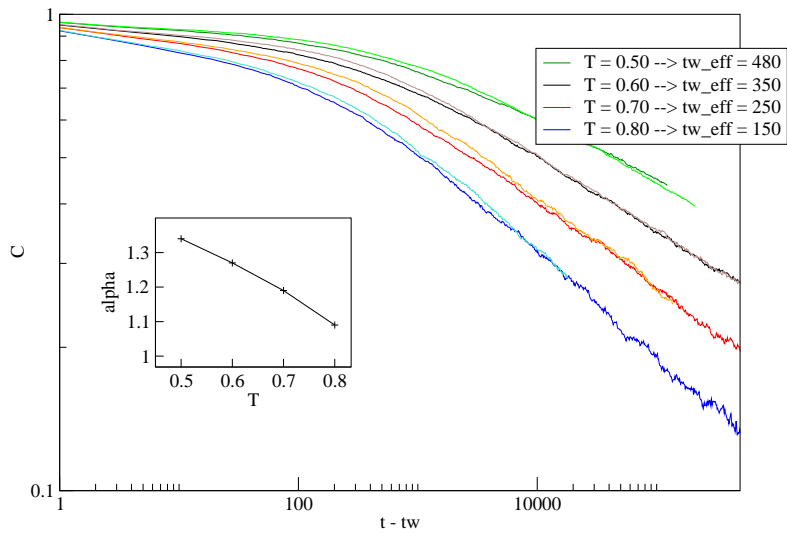


Figure 3.7: We superpose correlation functions obtained using standard MC after different waiting-times  $t_w$  with ones obtained using Tempering after corresponding  $t_w^{\text{eff}}$ .

### 3.4 Conclusion

Nice physical result : scaling of  $\rho(C_r; C, \dots)$  by use of the dynamical correlation length  $\xi$ .

Continuous Time Monte Carlo = very good for Random Field model.

Parallel Tempering = good to get very old samples of Edward-Anderson systems.

## Acknowledgements

M. Picco  
L. Cugliandolo  
A. Sicilia  
B. Diu

# Bibliography

- [1] N. Metropolis, A. W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [2] H. Honenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).