The Spectral Gap of the REM under Metropolis Dynamics

L. R. G. Fontes
Universidade de São Paulo

M. Isopi
Universitá di Roma 1

Y. Kohayakawa
Universidade de São Paulo

P. Picco
CPT, CNRS, Marseille

April 25, 1997

Abstract

We derive upper and lower bounds for the spectral gap of the Random Energy Model under Metropolis dynamics which are sharp in exponential order. They are based on the variational characterization of the gap. For the lower bound, a Poincaré inequality derived by Diaconis and Stroock is used. The scaled asymptotic expression is a linear function of the temperature. The corresponding function for a global version of the dynamics exhibits phase transition instead.

We also study the dependence of lower order terms on the volume. In the global dynamics, we observe a phase transition. For the local dynamics, the expressions we have, which are possibly not sharp, do not change their order of dependence on the volume as the temperature changes.

Mathematics Subject Classification 1991: 60K35, 82B44.
Key words and phrases: Disordered systems, spin glasses, random energy model, Metropolis dynamics, Glauber dynamics, convergence to equilibrium, spectral gap, dynamical phase transition

1 Introduction

The Random Energy Model (REM)([1], [2]) is a disordered hamiltonian spin system designed as a caricature of the Sherrington-Kirkpatrick (SK) Spin-Glass Model ([8]). Both models are mean-field ones. While in the SK model one has gaussian pair interactions only, in the REM there are gaussian multibody interactions. The hamiltonian for the SK model is (for σ a given configuration of spins + or -1 in a volume Λ)

$$H_{SK}(\sigma) = -|\Lambda|^{-1/2} \sum_{i,j\in\Lambda} J_{ij}\sigma_i\sigma_j, \tag{1.1}$$

where the sum is over all pairs of distinct sites in Λ and $\{J_{ij},\, i,j\}$ is a family of i.i.d. standard gaussians, whereas that for the REM is

$$H_{REM}(\sigma) = -\frac{|\Lambda|^{1/2}}{2^{|\Lambda|/2}} \sum_{\alpha\subset\Lambda} J_\alpha\sigma_\alpha, \tag{1.2}$$

where the sum is over all the 2^N subsets α of Λ, $\{J_\alpha,\, \alpha\}$ is a family of i.i.d. standard gaussians defined on a common probability space $(\mathcal{E}, \Sigma, I\!P)$ and $\sigma_\alpha = \prod_{i\in\alpha}\sigma_i$ $(\sigma_\emptyset = 1)$.

We observe that for the REM the hamiltonians of all configurations form a family of i.i.d. gaussians with mean zero and variance $|\Lambda|$. A proof of this elementary fact is provided in the appendix A. That could be taken as an alternative description of this model, indeed it is the usual one. We will adopt it in the next section.

The equilibrium statistical mechanics of the REM has been much studied. In a non rigorous way in ([1], [2]) and in a rigorous way in ([3], [4], [7]). We quote some of these rigorous results that will be important for understanding some aspects of the dynamics. Given $\beta \geq 0$, the inverse temperature, taking $\Lambda = \{1, \dots, N\}$, let us denote by

$$Z_N \equiv Z_N(\beta) = \sum_\sigma e^{-\beta H_{REM}(\sigma)} \tag{1.3}$$

the finite volume partition function and by

$$F_N(\beta) = \frac{1}{N}\log Z_N(\beta) \tag{1.4}$$

the finite volume free energy. It was proved in [4] that for all $\beta \geq 0$, the limit $\lim_{N\to\infty} F_N(\beta) = F(\beta)$ exists $I\!P$-almost surely and in $L^p(\mathcal{E}, I\!P)$ for $1 \leq p < \infty$. $F(\beta)$ is a non random function which is differentiable in β but the second derivative has a jump at $\beta_c = \sqrt{2\log 2}$. In fact, $F(\beta)$ is equal to $\beta^2/2 + \beta_c^2/2$ for $\beta < \beta_c$ and $\beta_c\beta$ for $\beta \geq \beta_c$, as expected from the results of [1]. This is called in the physics literature a third order phase transition.

The point is, depending if we are in a high temperature regime ($\beta < \beta_c$) or in a low temperature one ($\beta \geq \beta_c$), not only that the free energy changes from a quadratic function of β to a linear one but that the difference between the finite volume free energy and its infinite volume limit is exponential in the high temperature case and this occurs almost surely, whereas in the low temperature regime this difference behaves as $C(\omega, \beta, N)\frac{\log N}{N}$ for some random function $C(\omega, \beta, N)$. This function converges in $I\!P$-probability to a non-random limit but does not converge $I\!\!P$-almost surely and an interval where $C(\omega,\beta,N)$ fluctuates $I\!\!P$-almost surely is identified. More precisely, it was proved in [4] that if $0\le\beta<\beta_c$ then

$$F(\beta)-e^{-\lambda(\beta)N}\le F_N(\beta)\le F(\beta)+e^{-\lambda(\beta)N} \tag{1.5}$$

$I\!\!P$-almost surely and in $L^1(\mathcal{E},I\!\!P)$ for some $\lambda(\beta)>0$ if $\beta<\beta_c$. If $\beta\ge\beta_c$ then the rate of convergence is more subtle, depending if we want a $I\!\!P$-almost sure result or one in $I\!\!P$-probablity. Namely, calling

$$\mathcal{V}_N(\beta)=\frac{\beta_c}{\beta}\frac{\log\left[Z_Ne^{-NF(\beta)}\right]}{\log N}, \tag{1.6}$$

this quantity has a behavior which is radically different in $I\!\!P$-probability and $I\!\!P$ almost surely. It was proven in [4] that $I\!\!P$-almost surely

$$\limsup_{N\to\infty}\mathcal{V}_N(\beta)\le\frac{1}{2} \tag{1.7}$$

and also $I\!\!P$-almost surely

$$\liminf_{N\to\infty}\mathcal{V}_N(\beta)\ge-\frac{1}{2}. \tag{1.8}$$

These two results do not imply that $\mathcal{V}_N(\beta)$ does not converge $I\!\!P$ almost surely, this question was solved later, namely, it was proven in [7] that we have $I\!\!P$-almost surely

$$\limsup_{N\to\infty}\mathcal{V}_N(\beta)=\frac{1}{2} \tag{1.9}$$

and also $I\!\!P$-almost surely

$$\liminf_{N\to\infty}\mathcal{V}_N(\beta)=-\frac{1}{2}. \tag{1.10}$$

The result in $I\!\!P$-probability (and in $L^1(\Omega,I\!\!P)$) is simpler. In [7] it is proved that

$$\lim_{N\to\infty}\mathcal{V}_N(\beta)=-\frac{1}{2} \tag{1.11}$$

as it was expected from [2].

In this work we consider a dynamical version of the model, namely, the REM undergoing a Glauber dynamics (Metropolis). That is, we are considering a dynamics in random environment. We study the speed of convergence to equilibrium, the exponential rate of which is given by the *spectral gap* (or just *gap*) of the dynamics, which is the difference between the first and second eigenvalues of the transition probability matrix of the continuous time Markov chain defining it (see [5] Proposition 3, also [13]). As in other dynamics of spin systems, this gap goes to zero when the volume goes to infinity and one of the main questions in the study of the approach to equilibrium is the exact rate at which it does so. It is natural to consider the inverse of the gap instead of the gap, the former quantity being linked to the relaxation time. In non random mean field models at low temperature, the logarithm of the inverse of the gap grows like the volume. At high temperature, it is $o(N)$ and grows at least as the logarithm of the volume. In short range random systems, it is expected that the inverse of the spectral gap grows in a slower way at low temperature. This is a very active line of research. Therefore, it is important to clarify at least the case of one of the standard random mean field models, where we know very well the statics.

To get bounds for the inverse of the gap, we use a variational characterization and a bound derived by Diaconis and Stroock based on that ([5]). A nice percolation problem in the hypercube with $|\Lambda|$ dimensions comes into play. We prove that for the REM the logarithm of the inverse of the spectral gap grows like the volume. We give its exact asymptotic behavior by dividing it by the volume and proving that this normalised quantity converges $I\!P$-almost surely *for all* β to the linear non random function $\beta_c\beta$ (which is also the free energy of the REM at low temperature). We give $I\!P$ almost sure upper and lower bounds for the finite volume error of approximation of this quantity to its limit, in the very same spirit of (1.7) and (1.8). The magnitude of these bounds are of order $\sqrt{\frac{\log N}{N}}$, which suggests that the actual error may be bigger than that for the free energy (which is of order $\frac{\log N}{N}$ for $\beta \geq \beta_c$ and exponentially decaying in N for $\beta < \beta_c$, as pointed above). We conjecture that this is indeed the case and that the constant of proportionality is random. Also, we do not expect to see any phase transition in the behavior of this constant.

Our first theorem (Theorem 1 at the end of Section 3) implies that there is no dynamical phase transition (interpreted as a change of behavior in the limiting scaled gap as a function of the temperature) for this model. Note that in the Curie-Weiss model, there is a dynamical phase transition for the dynamics induced on the magnetizations, in the sense that, at low temperature, the logarithm of the spectral gap is proportional to the volume times the difference between the canonical free energy computed at magnetization zero and the canonical free energy computed at its minimum, whereas at high temperature it is $o(N)$.

In the final section of this paper, we consider a *global* Metropolis dynamics for the REM for which the scaled gap behaves asymptotically as a function of the temperature which does undergo a (third order) phase transition. We give also $I\!P$-almost sure bounds for the rate at which the scaled gap converges to its limit and we see here also a phase transition on the rate.

As regards other disordered dynamical models, Cassandro et al. have studied a random walk with random traps, with a different approach, using coupling techniques, to get the order of the speed of convergence to equilibrium [6].

Dynamics of spin glasses were also studied (non-rigorously) by Sompolinsky and Zippelius ([9]) which considered a *soft spin* approach so that they could work with quantities varying continuously. This has been made rigorous by Ben Arous and Guionnet ([12]). But their study is for time scale shorter than the one we are dealing with. Some recent work ([10] and [11]) shows that non exponential relaxation is possible in infinite volume for models with short range interaction. This is of course not in contradiction with our result for a (large) finite sample. We plan to investigate infinite volume behaviour and the relationship with the short range models results in a forthcoming paper.

The rest of the paper is organized as follows. In the next section we describe the model in detail. In Section 3 we derive upper and lower bounds for the inverse of the gap of the Metropolis dynamics. In Section 4 we consider a modification of this dynamics. Appendices are devoted to auxiliary results.

2 The Model

Throughout we consider the Random Energy Model (REM) as a non-equilibrium system undergoing Metropolis dynamics in finite volume. We want to study the behavior of the gap between the first and second eigenvalues of the infinitesimal generator of the corresponding continuous time Markov process (or of the probability transition matrix; the gap is the same) as the volume goes to infinity.

Let Λ be a nonempty set with $|\Lambda| = N$ and let Ω denote $\{-1, 1\}^{\Lambda}$. Let $\mathcal{H} = \{H(\sigma), \sigma \in \Omega\}$ be an independent family of Gaussian random variables with common mean zero and common variance N defined on a probability space $(\mathcal{E}, \Sigma, I\!\!P)$. $\mathcal{H}$ plays the role of the random hamiltonian or energy function. We consider a continuous-time Markov chain with state space Ω with transition probabilities that are reversible with respect to the Gibbs measure μ_N on Ω, which is obtained from $\mathcal{H}$ and the inverse temperature parameter β in the usual way, that is

$$\mu_N(\sigma) = \frac{1}{Z_N} \exp\{-\beta H(\sigma)\}, \quad \sigma \in \Omega.$$

More specifically, we consider Metropolis type transition probabilities, given by

$$\begin{aligned} P(\sigma, \sigma') &= \exp\{-\beta(H(\sigma') - H(\sigma))^+\}/N, \quad \text{if} \quad ||\sigma' - \sigma|| = 1, \\ &= 0, \quad \text{if} \quad ||\sigma' - \sigma|| > 1, \\ &= 1 - \sum_{\sigma'' \neq \sigma} P(\sigma, \sigma''), \quad \text{if} \quad \sigma' = \sigma, \end{aligned} \tag{2.1}$$

where $a^+ = \max\{a, 0\}$ and $||x|| = \frac{1}{2} \sum_{i=1}^{N} |x_i|$.

Note that these transition probabilities are random variables defined on $(\mathcal{E}, \Sigma, I\!\!P)$. That is, we have an inhomogenous random walk on the hypercube Ω in the random environment defined by the transition probability valued random variables $P(\sigma, \sigma')$. We are interested in properties of this dynamics that are true for almost all realisations of the random hamiltonian $\mathcal{H}$.

3 Upper and Lower Bounds

We recall here some basic facts about Markov chains. Let $P(\cdot, \cdot)$ be the transition probability for an irreducible Markov chain in a finite state space S which is reversible with respect to a measure μ on S. That is, $\mu(x)P(x, y) = \mu(y)P(y, x)$ for all $x, y \in S$.

Let ϕ be a real valued function on S. Let us define

$$Var(\phi) = \frac{1}{2} \sum_{\sigma, \sigma'} (\phi(\sigma) - \phi(\sigma'))^2 \mu(\sigma)\mu(\sigma'), \tag{3.1}$$

$$\mathcal{E}(\phi, \phi) = \frac{1}{2} \sum_{\sigma, \sigma'} (\phi(\sigma) - \phi(\sigma'))^2 Q(\sigma, \sigma'), \tag{3.2}$$

where $Q(\sigma, \sigma') = \mu(\sigma)P(\sigma, \sigma')$.

$Var(\phi)$ is the variance of ϕ and $\mathcal{E}(\phi, \phi)$ is the Dirichlet form of the Markov semi-group associated to $P(\cdot, \cdot)$.

Since $P(\cdot,\cdot)$ has largest eigenvalue 1 and the constant functions are the only eigenvectors with eigenvalue 1, using the minimax characterization of eigenvalues, if we define

$$\tau(\phi) = \frac{Var(\phi)}{\mathcal{E}(\phi,\phi)}, \tag{3.3}$$

then the inverse of the gap between the first and second eigenvalues of $P(\cdot,\cdot)$ is given by ([5])

$$\tau = \sup_{\phi} \tau(\phi), \tag{3.4}$$

where the sup is taken over non constant ϕ's.

We will use this characterization of the gap to derive bounds for it in the case of the dynamics given by (2.1). (Later on, in the final section, we will consider a *global* modification of it.) In this context, $S = \Omega$ and $\mu = \mu_N$.

We have also the following bound, given in [5], for the distance in variation.

$$\|P_t(x,\cdot) - \mu_N(\cdot)\|_{var} \leq \sqrt{\frac{1-\mu_N(x)}{4\mu_N(x)}} e^{-t/\tau} \tag{3.5}$$

Here $P_t(x,y) = e^{-t}\sum_{n=0}^{\infty}\frac{t^n}{n!}P^n(x,y)$ is the transition kernel. (There is a similar lower bound for the maximum in x of the left hand side. See [13].)

In the next two subsections, we derive upper and lower bounds for τ.

3.1 Lower Bound

As usual in this kind of variational problem, we take a trial function to get a lower bound. We choose for ϕ the indicator of a spin configuration, that is, we define ϕ_σ by $\phi_\sigma(\sigma') = \delta(\sigma,\sigma')$, where $\delta(\cdot,\cdot)$ is the Kronecker delta. We have

$$\tau(\phi_\sigma) = \frac{\exp\{-\beta H(\sigma)\}}{Z_N \sum_{\langle\sigma,\sigma'\rangle} Q(\sigma,\sigma')}(1-\mu_N(\sigma)), \tag{3.6}$$

where the sum is over the N nearest neighbors of σ, denoted σ'.

For the Metropolis dynamics, $Q(\sigma,\sigma') = (NZ_N e^{\beta(H(\sigma)\vee H(\sigma'))})^{-1}$ and thus (3.6) equals

$$\frac{N\exp\{-\beta H(\sigma)\}}{\sum_{\langle\sigma,\sigma'\rangle}\exp\{-\beta(H(\sigma)\vee H(\sigma'))\}}(1-\mu_N(\sigma)). \tag{3.7}$$

Let $\underline{\sigma}$ be the (unique) spin configuration σ for which $H(\sigma)$ is minimal. Then we get the bound

$$\tau \geq \max_{\sigma} \tau(\phi_\sigma) \geq \tau(\phi_{\underline{\sigma}}) = \frac{N\exp\{-\beta H(\underline{\sigma})\}}{\sum_{\langle\underline{\sigma},\sigma'\rangle}\exp\{-\beta H(\sigma')\}}(1-\mu_N(\underline{\sigma})). \tag{3.8}$$

Proposition 3.1 *There exists a positive constant c such that, for all β, with $I\!P$-probability 1, for all but a finite number of indices N we have*

$$\frac{1}{N}\log\tau \geq \beta_c\beta - c\beta\sqrt{\frac{\log N}{N}}, \tag{3.9}$$

where $\beta_c = \sqrt{2\log 2}$. *In particular*

$$\liminf_{N\to\infty} \frac{1}{N}\log\tau \geq \beta_c\beta \tag{3.10}$$

$I\!P$-almost surely.

Proof

Let us first consider the term $1-\mu_N(\underline{\sigma})$. Let $D_N = H(\underline{\sigma}') - H(\underline{\sigma})$ where $\underline{\sigma}'$ is the location of the second minimum of H. Then

$$(1-\mu_N(\underline{\sigma}))^{-1} = 1 + \frac{e^{-\beta H(\underline{\sigma})}}{\sum_{\sigma\neq\underline{\sigma}} e^{-\beta H(\sigma)}} \leq 1 + \frac{e^{-\beta H(\underline{\sigma})}}{e^{-\beta H(\underline{\sigma}')}} \leq 2e^{\beta D_N}. \tag{3.11}$$

Now, given $\epsilon > 0$, one has

$$I\!P(D_N > N\epsilon) = \int_{-\infty}^{\infty} \frac{\bar{F}(x+N\epsilon)}{\bar{F}(x)}\, dF'_N(x),$$

where F is the gaussian with mean 0 and variance N distribution function, $\bar{F} = 1-F$ and F'_N is the distribution function of $-H(\underline{\sigma}')$. Here we use the elementary fact that for a sequence $X_1,\dots,X_n$ of i.i.d. continuous random variables, if $Y_1,\dots,Y_n$ are its increasing order statistics, then

$$I\!P(Y_n > y | Y_{n-1} = x) = \frac{1-F_{X_1}(y)}{1-F_{X_1}(x)}$$

for all $x < y$, where F_{X_1} is the distribution function of X_1.

We break the integral above into an integral over $x<0$ and another over $x>0$. The former is bounded from above by

$$I\!P(H(\underline{\sigma}') > 0) = (2^N+1)2^{-2^N},$$

where the equality follows from elementary computations. To bound the latter integral, we observe that, since x is positive, $\bar{F}(x+N\epsilon) \leq e^{-\epsilon^2 N/2}\bar{F}(x)$ by a simple linear change of variables. Thus $e^{-\epsilon^2 N/2}$ is an upper bound for this integral. Therefore, given $\gamma > 0$, choosing $\epsilon = \sqrt{\frac{2\log N}{N}(1+\gamma)}$, we get

$$I\!P\left(\frac{D_N}{N} \geq \sqrt{\frac{2\log N}{N}(1+\gamma)}\right) \leq (2^N+1)2^{-2^N} + \frac{1}{N^{1+\gamma}}. \tag{3.12}$$

Using the first Borel-Cantelli lemma, we then get that for any $\gamma > 0$, with $I\!P$-probability 1, for all but a finite number of indices N

$$\frac{1}{N}\log(1-\mu_N(\underline{\sigma})) \geq -\frac{\log 2}{N} - \beta\sqrt{\frac{2\log N}{N}(1+\gamma)}. \tag{3.13}$$

To bound from below the term $\exp\{-\beta H(\underline{\sigma})\}$, we use the easily checked fact that for all $\epsilon > 0$

$$I\!P\left(H(\underline{\sigma}) \geq -\beta_c N + \frac{1}{2\beta_c}(1+\epsilon)\log N\right) \leq e^{-cN^\epsilon} \tag{3.14}$$

for some positive constant c. Thus, using again the Borel-Cantelli lemma, we get that for any $\epsilon > 0$, with $I\!P$-probability 1, for all but a finite number of indices N

$$\frac{1}{N} \log \exp\{-\beta H(\underline{\sigma})\} \geq \beta\beta_c - \frac{\beta}{2\beta_c}(1+\epsilon)\frac{\log N}{N}. \tag{3.15}$$

Note that the corrections are of smaller order than the ones in 3.13. It remains to consider the denominator in (3.7). We first bound from above the sum by N times the maximum of the summands. Since $\underline{\sigma}$ is the configuration where the infimum is reached, these summands are not independent. However, the maximum is stochastically dominated by the maximum of N independent summands. A proof of this fact can be found in the appendix.

Now if m_N is the minimum of N independent standard gaussian random variables then if $c > 1$

$$\begin{aligned} I\!P\left(\bigcup_{N=N_o}^{\infty}\left\{m_N \leq -\sqrt{(1+\epsilon)2\log N}\right\}\right) &\leq \sum_{k=k_0}^{\infty} I\!P\left(\bigcup_{N=c^k}^{c^{k+1}}\left\{m_N \leq -\sqrt{(1+\epsilon)2\log c^k}\right\}\right) \\ &\leq \sum_{k=k_0}^{\infty} e^{-\epsilon k \log c} < \infty. \end{aligned} \tag{3.16}$$

Therefore we get that with $I\!P$-probability 1 for all but a finite of indices N

$$m_N \geq -\sqrt{2(1+\epsilon)\log N}. \tag{3.17}$$

We used now the classical fact that if two families $(X_i)_{i \in I\!N}$ and $(Y_i)_{i\in I\!N}$ of real random variables are such that for any $i \in I\!N$ X_i is stochastically dominated by Y_i, then we can construct a common probability space such that, $I\!P$-almost surely, $X_i \leq Y_i$ for all $i \in I\!N$. Therefore, we get for any $\epsilon > 0$, with $I\!P$-probability 1 for all but a finite of indices N

$$\frac{1}{N}\log(\sum_{\langle\sigma,\sigma'\rangle} \exp\{-\beta(H(\sigma)\vee H(\sigma'))\})^{-1} \geq -\beta\sqrt{2(1+\epsilon)\frac{\log N}{N}}. \tag{3.18}$$

Collecting (3.13), (3.15) and (3.18), we get (3.9). □

Remark 3.1 *The bound (3.18) is optimal, in the sense that it can be proved, by restricting the sum to the σ' which realises the maximum, that with $I\!P$-probability 1, infinitely often (in N) we have*

$$\frac{1}{N}\log(\sum_{\langle\sigma,\sigma'\rangle} \exp\{-\beta(H(\sigma)\vee H(\sigma'))\})^{-1} \leq -\beta\sqrt{2(1-\epsilon)\frac{\log N}{N}}. \tag{3.19}$$

Note that it can be also proved, but it is rather long, that

$$\frac{1}{N}\log(1-\mu_N(\underline{\sigma})) \geq -\frac{\log 2}{N} - c\beta\frac{\log N}{N}(1+\gamma) \tag{3.20}$$

for all large enough N. Since the proof we gave here is really shorter and, on the other hand, the upper bound for the correction to the upper bound for τ we will get in the next chapter is of order $\sqrt{\frac{\log N}{N}}$ and we have no proof of the optimality for that part, we prefer not to argue (3.20) in detail.

3.2 Upper Bound

The upper bound is based on the *Poincaré inequality* derived in [5]. It is given in terms of the *canonical paths* of Jerrum and Sinclair [13], [14].

Let $\mathcal{C}_N$ denote the hypercube in N dimensions obtained from Ω by adding nearest neighbor bonds (the ones over which the transition probabilities are positive) between the points of Ω, that we will call with a little abuse sites, *i.e.* a site is just a spin configuration. We will call σ_j the j-coordinate of the site σ. Let Γ be a complete set of directed paths in $\mathcal{C}_N$, that is, Γ is a set of directed paths in $\mathcal{C}_N$ such that every two distinct sites in $\mathcal{C}_N$ are ends of a directed path in Γ. The paths are self avoiding *i.e.* for a given path a bond is visited just once.

¿From Proposition $1'$ in [5] one has the bound

$$\tau \leq \max_b Q(b)^{-1} \sum_{\gamma_{\eta,\eta'} \ni b} |\gamma_{\eta,\eta'}| \mu_N(\eta)\mu_N(\eta'), \tag{3.21}$$

where the max is over the nearest neighbor bonds $b = \langle \sigma, \sigma' \rangle$ of $\mathcal{C}_N$, $Q(b) = Q(\sigma, \sigma')$ and the summation is over all paths in Γ (indexed by their endpoints η and η') which pass through b.

Writing the bound more explicitly, we have

$$\tau \leq \frac{N}{Z_N} \max_{b=\langle\sigma,\sigma'\rangle} e^{\beta(H(\sigma)\vee H(\sigma'))} \sum_{\gamma_{\eta,\eta'} \ni b} |\gamma_{\eta,\eta'}| e^{-\beta(H(\eta)+H(\eta'))}. \tag{3.22}$$

The rest of this subsection is devoted to exhibit a choice of Γ such that one has control on (3.22) enough to obtain an upper bound for it of the same logarithmic order as the lower bound of last subsection, with corrections of the same order.

One term to control is $e^{\beta(H(\sigma)\vee H(\sigma'))}$, which is too big when a path passes through values of H close to the maximum. The strategy is to avoid those *bad* points as often as possible. This cannot be done always since there must be paths that visit sites with high positive values of H. This is because the set of paths is *complete*. On a heuristic level, one wants to take advantage of the fact that the Gibbs measure of such high positive values of the energy is very small. To take this into account in a rigorous way, we define the set of paths in such a way that if these high positive values are visited, this will occur only at the ends of the path and therefore their contribution are depressed by the factor $e^{-\beta(H(\eta)+H(\eta'))}$ appearing in (3.22).

It is not obvious that this can be done and, if it can, then the set of paths we get is certainly dependent on the realizations of $\mathcal{H}$, which in turn puts a problem on the subsequent control of the sum in (3.22).

Fortunately, the bad points are sufficiently rare that the first control is obtained with a set of paths with randomness limited so that the second control becomes possible.

3.2.1 A choice of Γ

To choose Γ, we start by considering a family of sets of paths Γ_i, $i = 1, \ldots, N$, as follows.

For $i \in \{1, \ldots, N\}$, η and $\eta' \in \Omega$ fixed such that $\eta_i \neq \eta'_i$, let $\gamma^i_{\eta,\eta'}$ be the path from η to η' obtained by going left to right cyclically from η to η' successively flipping the disagreeing spins, starting at the coordinate i. Let

$$\Gamma_i = \{\gamma^i_{\eta,\eta'},\ \eta, \eta' \in \Omega\}. \tag{3.23}$$

Given η, η' and $\gamma_{\eta,\eta'}$, let $\overline{\gamma_{\eta,\eta'}}$ be the set of points visited by the path $\gamma_{\eta,\eta'}$ and $\gamma^o_{\eta,\eta'} = \overline{\gamma_{\eta,\eta'}} \setminus \{\eta, \eta'\}$ the set of interior point of the path $\gamma_{\eta,\eta'}$. We say that a family of paths $\gamma_1, \ldots, \gamma_n$ is *interior-disjoint* if $\gamma^o_i \cap \gamma^o_j = \emptyset$ for all $1 \le i \ne j \le n$. We will need the following result, which is easy to check but fundamental.

Lemma 3.1 *Given η and η' in Ω at distance n (that is, $||\eta - \eta'|| = n$), there exist n interior-disjoint paths in $\{\gamma^i_{\eta,\eta'},\ i = 1, \ldots, N\}$.*

Sketch of Proof

Indeed, if $i_1, \ldots, i_n$ are the coordinates where η and η' disagree, then one easily checks that $\gamma^{i_1}_{\eta,\eta'}, \ldots, \gamma^{i_n}_{\eta,\eta'}$ are interior-disjoint. Notice all such paths have $n-2$ interior points. $\square$

The set Γ will be chosen depending on a positive parameter c_e to be chosen later. It will be formed by indicating for each pair (η, η') the path connecting the respective sites.

We will distinguish between *good* and *bad* sites of Ω. Good sites are those σ for which $H(\sigma) \le \sqrt{2(1+c_e)N \log N}$; otherwise, they are bad. We say that a path γ is good if all its interior points γ^o are good, and that a set of paths is good if all its elements are good.

We construct the random set of paths Γ. For $||\eta - \eta'|| \ge \frac{N}{\log N}$, if there is a good path in $\{\gamma^i_{\eta,\eta'},\ i = 1, \ldots, N\}$, choose the first such for Γ; otherwise, choose $\gamma^1_{\eta,\eta'}$.

For $||\eta - \eta'|| < \frac{N}{\log N}$, if there exists a good site η'' in Ω such that $||\eta - \eta''|| \ge \frac{N}{\log N}$, $||\eta' - \eta''|| \ge \frac{N}{\log N}$ and there are good paths, one in $\{\gamma^i_{\eta,\eta''},\ i = 1, \ldots, N\}$ and another in $\{\gamma^i_{\eta',\eta''},\ i = 1, \ldots, N\}$, such that the union of these two good paths is a self avoiding path, select this union as the path connecting η and η' in Γ (notice that this is a good path since η'' is good); otherwise, select $\gamma^1_{\eta,\eta'}$. Notice that all the paths constructed in this way have length smaller than N, so we have the bound

$$\tau \le \frac{N^2}{Z_N} \max_{b=\langle\sigma,\sigma'\rangle} e^{\beta(H(\sigma)\vee H(\sigma'))} \sum_{\gamma_{\eta,\eta'}\ni b} e^{-\beta(H(\eta)+H(\eta'))}. \tag{3.24}$$

The next result controls the term $e^{\beta(H(\sigma)\vee H(\sigma'))}$.

Proposition 3.2 *With $I\!P$-probability 1, for all but a finite number of indices N, the set of paths Γ previously constructed is good.*

Proof

We will argue that for an arbitrary pair (η, η') in Ω, the probability not to find a good path connecting them as prescribed in the construction above is not bigger that $e^{-c(e)N}$ for some constant $c(e)$ that depends on c_e which can be chosen as big as we need. Since the number of such pairs does not exceed 4^N, choosing $c(e) > \log 4$, the result follows from the first Borel-Cantelli Lemma. We assume that N is large enough to keep only the exponential factor in gaussian estimates.

For pairs of sites more than distance $\frac{N}{\log N}$ apart, using the previous lemma, there are at least $\frac{N}{\log N}$ disjoint paths of length at most N connecting them. The probability for a given site to be bad is no bigger than $\exp\{-(1+c_e)\log N\}$. Thus the probability of a given path among the disjoint ones to be not good is at most $N\exp\{-(1+c_e)\log N\} = \exp\{-c_e \log N\}$, since all the paths constructed have length smaller than N. We conclude that the probability that *all* the $\frac{N}{\log N}$ disjoint paths constructed in the previous lemma are not good is at most $\exp\{-c_e N\}$.

For (η,η') less than distance $\frac{N}{\log N}$ apart, let $D(\eta,\eta')$ be the coordinates where η and η' disagree. Given $\tilde\eta$ that coincides with η on $D(\eta,\eta')$ and has $\frac{N}{\log N}$ discrepancies with η, let $\gamma_{\tilde\eta,\tilde\eta'}$ be the path starting at the site $\tilde\eta$, constructed by flipping the coordinates in $D(\eta,\eta')$ in increasing order. The probability that the set of visited points by this path is not good is at most $\exp\{-c_e\log N\}$. Since, for any $\epsilon>0$, there are at most $\exp cN^{1-\epsilon}$ many such $\tilde\eta$s and, as it is easy to check, all the paths $\gamma_{\tilde\eta,\tilde\eta'}$ obtained by varying $\tilde\eta$ are disjoint, we get that the probability that *all* the disjoints paths $\gamma_{\tilde\eta,\tilde\eta'}$ are not good is at most $\exp\{-c_e\log N\exp\{cN^{1-\epsilon}\}\}$ for some positive constant c. Therefore, for N large enough, we can find at least one such good site $\tilde\eta$, say η" and the corresponding path, say $\gamma_{\eta",\eta"'}$ with all its visited points good. By construction, $\eta"'$ coincides with η' on $D(\eta,\eta')$ and is at distance $\frac{N}{\log N}$ apart from it. Therefore we are in the very same hypothesis as before and we can find good paths $\gamma_{\eta"',\eta'}$ and $\gamma_{\eta,\eta"}$ for N large enough. Now glueing the three good paths $\gamma_{\eta,\eta"}$, $\gamma_{\eta",\eta"'}$ and $\gamma_{\eta"',\eta'}$, we get a good path $\gamma_{\eta,\eta'}$ and by construction this path is self avoiding.

All cases have now been covered and the result is proven. □

We will now suppose that N is large enough, how large is an almost surely finite random variable by the previous result, so that we can take Γ good. Notice that, in this case, a bad site can appear only at the ends of any path of Γ. So, if b contains a bad site, say σ, and a path γ of Γ contains b, then σ is an end of γ and summing over all such paths is equivalent to summing over all sites of Ω but σ.

So, if b contains a bad site, the term inside the max sign in (3.24) can be bounded above by Z_N.

Let G be the collection of bonds of $\mathcal{C}_N$ that contain no bad site. By the last paragraph,

$$\frac{\tau}{N^2}\le 1\vee Z_N^{-1}\max_{b\in G}e^{\beta(H(\sigma)\vee H(\sigma'))}\sum_{\gamma_{\eta,\eta'}\ni b}e^{-\beta(H(\eta)+H(\eta'))}. \tag{3.25}$$

Using the fact that $b\in G$, we get

$$\frac{\tau}{N^2}\le 1\vee\left[e^{\beta\sqrt{2(1+c_e)N\log N}}\tau_1\right] \tag{3.26}$$

where

$$\tau_1\equiv Z_N^{-1}\max_{b\in G}\sum_{\gamma_{\eta,\eta'}\ni b}e^{-\beta(H(\eta)+H(\eta'))}. \tag{3.27}$$

3.2.2 Estimates for τ_1

We write $\tau_1=\tau_1^1+\tau_1^2$, where in τ_1^1 the sum in (3.27) is over paths connecting sites at distance $N/\log N$ or more apart and in τ_1^2 the sum is over paths connecting sites at distance less than $N/\log N$ apart.

We consider first τ_1^1. The sum can be estimated by

$$Z_N^{-1}\sum_{i=1}^{N}\sum_{\gamma_{\eta,\eta'}\ni b}^{(i)}e^{-\beta(H(\eta)+H(\eta'))}, \tag{3.28}$$

where the sum $\sum^{(i)}$ is over all paths of Γ_i. Note that we have replaced here a random set of paths by non random ones. This follows from the fact that the subset of paths of Γ connecting sites which are more than distance $O(N/\log N)$ apart is contained in $\cup_{i=1}^N \Gamma_i$. We have then the estimate

$$\tau_1^1 \leq N Z_N^{-1} \max_i \max_b \sum_{\gamma_{\eta,\eta'} \ni b}^{(i)} e^{-\beta(H(\eta)+H(\eta'))}. \tag{3.29}$$

The random variables

$$Z_N(i) \equiv \max_b \sum_{\gamma_{\eta,\eta'} \ni b}^{(i)} e^{-\beta(H(\eta)+H(\eta'))},$$

$i = 1, \dots, N$, have the same distribution. It will suffice to consider the first one. We will need some further notation.

Given a bond $b = \langle \sigma, \sigma' \rangle$, let $\ell = \ell(b)$ be the coordinate where there is a discrepancy. Given a coordinate j and a site σ, define the collections of sites

$$\Omega_j^-(\sigma) = \{\eta \in \Omega : \eta(i) = \sigma(i),\ i = j, \dots, N\} \tag{3.30}$$

$$\Omega_j^+(\sigma) = \{\eta \in \Omega : \eta(i) = \sigma(i),\ i = 1, \dots, j\}. \tag{3.31}$$

Now

$$\sum_{\gamma_{\eta,\eta'} \ni b}^{(1)} e^{-\beta(H(\eta)+H(\eta'))} = \sum_{\eta \in \Omega_\ell^-(\sigma)} e^{-\beta H(\eta)} \sum_{\eta \in \Omega_\ell^+(\sigma')} e^{-\beta H(\eta)}, \tag{3.32}$$

since the paths in Γ_1 through b connect sites in $\Omega_\ell^-(\sigma)$ to sites in $\Omega_\ell^+(\sigma')$.

Thus

$$\begin{aligned} &\max_b \sum_{\gamma_{\eta,\eta'} \ni b}^{(1)} e^{-\beta(H(\eta)+H(\eta'))} \\ &= \max_j \max_{b=\langle\sigma,\sigma'\rangle : \ell(\sigma,\sigma')=j} \sum_{\eta\in\Omega_{j-1}} e^{-\beta H(\eta,\sigma_j,\dots,\sigma_N)} \sum_{\eta'\in\Omega_{N-j}} e^{-\beta H(\sigma'_1,\dots,\sigma'_j,\eta')} \end{aligned} \tag{3.33}$$

where $\Omega_i = \{1, -1\}^i$. Therefore calling

$$Z_{j-1}(\xi, \zeta') = \sum_{\eta \in \Omega_{j-1}} e^{-\beta H(\eta,\xi,\zeta')} \tag{3.34}$$

and

$$Z_{N-j}(-\xi, \zeta) = \sum_{\eta' \in \Omega_{N-j}} e^{-\beta H(\zeta,-\xi,\eta')}, \tag{3.35}$$

the right hand side of (3.33) equals

$$\max_j \max_{\xi=\pm 1} \max_{\zeta' \in \Omega_{N-j}} Z_{j-1}(\xi, \zeta') \max_{\zeta \in \Omega_{j-1}} Z_{N-j}(-\xi, \zeta). \tag{3.36}$$

We now estimate $Z_{j-1}(\xi, \zeta')$ and $Z_{N-j}(-\xi, \zeta)$ in the same way Olivieri and Picco estimated the REM partition function (see Section III in [4] through to equation (III.17)), rather using the exponential Markov inequality than the simple one in order to be able to control the max signs in (3.36) that involve an exponential number of terms. However, this section is self contained.

Let j and $\zeta' \in \Omega_{N-j}$ be fixed and take $\xi = 1$ (for definiteness). Let M be a positive integer to be chosen later and make a partition of the real line with the intervals

$$\begin{aligned} \Delta_0 &= \left(-\infty, \beta_c \frac{N}{M}\right], \\ \Delta_k &= \left(\beta_c \frac{kN}{M}, \beta_c \frac{N(k+1)}{M}\right], \quad \text{if } 1 \le k \le M, \\ \Delta_{M+1} &= (\beta_c(1+1/M)N, +\infty). \end{aligned}$$

Now write

$$Z_{j-1}(\xi = 1, \zeta') = \sum_{\sigma \in \Omega_{j-1}} \sum_{k=0}^{M+1} \mathbb{I}_{\Delta_k}(-H(\sigma, 1, \zeta')) e^{-\beta H(\sigma,1,\zeta')}$$

$$\le 2^{j-1} e^{\beta_c \beta N/M} + \sum_{k=1}^{M} N_k e^{\beta_c \beta (k+1) N/M} + N^* e^{-\beta H(\underline{\sigma})}, \tag{3.37}$$

where

$$N_k = N_k(\zeta', i = 1, \xi = +1) = \sum_{\eta \in \Omega_{j-1}} \mathbb{I}_{\Delta_k}(-H(\eta, 1, \zeta'))$$

for $k = 0, \ldots, M$ and

$$N^* = \sum_{\sigma \in \Omega} \mathbf{1}_{\Delta_{M+1}}(-H(\sigma)).$$

Let us suppose now that $j - 1 = \alpha N$ and focus on the middle sum in (3.37). Let p_k denote $I\!P(-H(\sigma) \in \Delta_k)$. We then have

$$\beta_c \frac{\sqrt{N}}{M} 2^{-\frac{(k+1)^2}{M^2} N} < p_k < \beta_c \frac{\sqrt{N}}{M} 2^{-\frac{k^2}{M^2} N}, \tag{3.38}$$

for all $k = 1, \ldots, M$ and N large enough.

Optimizing in the exponential Markov inequality, we get

$$I\!P(N_k > \rho_k E(N_k)) \le e^{-\lambda_k 2^{\alpha N}}, \tag{3.39}$$

where

$$\rho_k = 2^{(\frac{(k+1)^2+1}{M^2} - \alpha)^+ N + 2} \tag{3.40}$$

$$\begin{aligned} \lambda_k &= \rho_k p_k \log \frac{\rho_k(1-p_k)}{1-\rho_k p_k} - \log\left[1 - p_k + \frac{\rho_k p_k (1-p_k)}{1-\rho_k p_k}\right], \\ &\quad \text{if} \quad \rho_k p_k < 1, \\ &= \infty, \quad \text{otherwise.} \end{aligned} \tag{3.41}$$

In any case, we have $\lambda_k \ge c \rho_k p_k$ for some positive constant c and thus,

$$I\!P(N_k > \rho_k E(N_k)) \le \exp\{-c \rho_k p_k 2^{\alpha N}\}. \tag{3.42}$$

Also

$$\rho_k p_k 2^{\alpha N} \ge 2^{N/M^2}.$$

Let the reader be reminded that $N_k = N_k(\zeta', j, i = 1, \xi = 1)$ and that the previous estimates are done for a fixed configuration $\zeta' \in \Omega_{N-j}$ with $j \in \{1, \ldots, N\}$, $i \in \{1, \ldots, N\}$ and $\xi = \pm 1$. Recalling (3.36), we have to make this estimates uniformly with respect to all those possible values. Since there are not more than $2N^2 2^N$ random variables that come into account, we need to have a probability estimate in (3.42) that compensates this factor. This suggests that we choose M in such a way that $2^{N/M^2} \geq c_u N$ for a positive constant c_u that will be chosen later. That is, we take

$$M = M(N) = \sqrt{\frac{N \log 2}{\log c_u N}} \tag{3.43}$$

and we get, for all $\epsilon > 0$,

$$\begin{aligned} &I\!P \left[\max_i \max_j \max_{\zeta' \in \Omega_{N-j}} \max_{1 \leq k \leq M} N_k(\zeta', i, \xi) > \rho_k E(N_k) \right] \\ &\leq\ 2N^2 M(N) 2^N e^{-c_u N} \leq e^{-\epsilon N} \end{aligned} \tag{3.44}$$

choosing $c_u > \log 2 + 2\epsilon$.

We conclude that the middle sum in (3.37) is bounded above by constant times

$$\frac{\sqrt{N}}{M} \sum_{k=1}^{M} \rho_k 2^{\alpha N} 2^{-\frac{k^2}{M^2} N} e^{\beta_c \beta (k+1) N/M} \tag{3.45}$$

$$= \frac{\sqrt{N}}{M} \sum_{k=1}^{\sqrt{\alpha M^2 - 1}} e^{(\alpha - k^2/M^2) \log 2 + \beta_c \beta (k+1) N/M} + \frac{\sqrt{N}}{M} \sum_{k=\sqrt{\alpha M^2 - 1}}^{M} e^{\beta_c \beta (k+1) N/M} 2^{N/M^2} \tag{3.46}$$

$$\leq \sqrt{N} \left(\sup_{x \in [0,1]} e^{\alpha G_{\beta/\sqrt{\alpha}}(x) N} e^{\beta_c \beta N/M} + e^{\beta_c \beta (1+1/M) N} 2^{N/M^2} \right), \tag{3.47}$$

where $G_\beta(x) = \beta_c \beta x + \frac{\beta_c^2}{2}(1 - x^2)$, except for an event of probability not bigger than $\exp\{-\epsilon N\}$.

The sup term in (3.47) equals $\exp\{\alpha F(\beta/\sqrt{\alpha}) N\}$, where

$$\begin{aligned} F(x) &= (x^2 + \beta_c^2)/2, \quad \text{if } 0 \leq x \leq \beta_c, & (3.48) \\ &= \beta_c x, \quad \text{if } x \geq \beta_c & (3.49) \end{aligned}$$

and $\alpha F(\beta/\sqrt{\alpha})|_{\alpha=0} \equiv 0$.

So, uniformly in i, j, ξ, ζ' we have

$$\max_{\zeta' \in \Omega_{N-j}} Z_{j-1}(\zeta', \xi) \leq cM e^{\alpha F(\beta/\sqrt{\alpha}) N} e^{\beta_c \beta N/M} + M e^{\beta_c \beta (1+1/M) N} 2^{N/M^2} + N^* e^{-\beta H(\underline{\sigma})} \tag{3.50}$$

with $I\!P$-probability not smaller than $1 - \exp\{-\epsilon N\}$, where we have absorbed the factor $2^{\alpha N}$ in (3.37) by changing the constant c, since $2^{\alpha N} \leq e^{\alpha F(\beta/\sqrt{\alpha}) N}$, as it can be easily checked.

Notice that $I\!P$-almost surely, $N^* = 0$ for all but a finite number of indices N, since

$$P(-H(\sigma) > (1 + 1/M) N \beta_c) \leq 2^{-(1+1/M)^2 N}.$$

After a similar reasoning with

$$Z_{N-j}(-\xi, \zeta) = \max_{\zeta \in \Omega_{j-1}} \sum_{\eta' \in \Omega_{N-j}} e^{-\beta H(\zeta, -1, \eta')},$$

we conclude that $I\!P$-almost surely for all but a finite number of indices N

$$\tau_1^1 \le N Z_N^{-1} \max_{0\le\alpha\le 1} \psi(\alpha,N,M)\psi(1-\alpha,N,M), \tag{3.51}$$

where

$$\begin{aligned}\psi(\alpha,N,M) &= cMe^{\alpha F(\beta/\sqrt{\alpha})N} e^{\beta_c\beta N/M} + Me^{\beta_c\beta(1+1/M)N}2^{N/M^2} \\ &\le e^{c\sqrt{N\log N}} e^{\left[\alpha F(\beta/\sqrt{\alpha})\vee\beta_c\beta\right]N}\end{aligned}$$

and c is a positive constant, not necessarily the same everytime it appears.

Collecting, we get that

$$\frac{1}{N}\log\tau_1^1 \le \max_{0\le\alpha\le1} \Psi(\alpha,\beta) - F(\beta) + c\beta\sqrt{\frac{\log N}{N}} + c'\frac{\log N}{N}. \tag{3.52}$$

where

$$\Psi(\alpha,\beta) = [\alpha F(\frac{\beta}{\sqrt{\alpha}}) + (1-\alpha)F(\frac{\beta}{\sqrt{1-\alpha}})] \vee [\alpha F(\frac{\beta}{\sqrt{\alpha}}) + \beta_c\beta] \vee 2\beta_c\beta$$

and c' is a positive constant.

Now one checks that $\max_{0\le\alpha\le1}\Psi(\alpha,\beta) - F(\beta) \le \beta_c\beta$ for all β.

Therefore, we get that $I\!P$-almost surely for all but a finite number of indices N

$$\frac{1}{N}\log\tau_1^1 \le \beta\beta_c + c\beta\sqrt{\frac{\log N}{N}} + c'\frac{\log N}{N} \tag{3.53}$$

for some constants c, c' and all β.

We consider now the term τ_1^2. We estimate it in two different ways, one for β close to 0, another for β away from 0.

The first bound follows from the fact that the sum is over a set of paths connecting sites in a hypercube of dimension at most $(N/\log N)$ around b, so we have

$$\tau_1^2 \le Z_N^{-1} e^{cN/\log N} e^{-2\beta H(\underline{\sigma})} \tag{3.54}$$

Unrestricting the paths to go through b, we get the other bound

$$\tau_1^2 \le Z_N^{-1} \max_{\eta\in\Omega} \sum_{\eta':||\eta'-\eta||<\frac{N}{\log N}} e^{-\beta H(\eta')}. \tag{3.55}$$

Notice that (3.55) is very similar to (3.34), with a sum over less than $2^{\alpha N}$ (independent) terms, with an arbitrary $\alpha > 0$ (for N large enough). We can thus proceed to estimate (3.55) in the same way as we did (3.34) to get the bound (3.50) from which the right hand side of (3.53) follows as an upper bound for the log of (3.55) divided by N when $\beta > \beta_0$, where β_0 is a suitable positive fixed number close to 0 (the bound is in fact true for $\beta > 0$, but then N large enough will depend on β).

For β close to 0, we use (3.54) to get

$$\frac{1}{N}\log\tau_1^2 \le 2\beta_c\beta - F(\beta) + c/\log N \tag{3.56}$$

which is negative for N large enough. This finishes the estimation of τ_1^2. The bound (3.53) is seen to carry to τ by (3.26).

Together with the previous results, this proves the following.

Proposition 3.3 *For all β, there exist finite constant c and c' such that, with $I\!P$-probability 1, for all but a finite number of indices N, we have*

$$\frac{1}{N}\log\tau \le \beta\beta_c + c\sqrt{\frac{\log N}{N}} + c'\frac{\log N}{N}. \tag{3.57}$$

In particular,

$$\limsup_{N\to\infty}\frac{1}{N}\log\tau \le \beta_c\beta. \tag{3.58}$$

Remark 3.2 *The argument leading to inequality (3.53) proves a weak form of the inequality*

$$\sum_{\eta\in\Omega_{j-1}} e^{-\beta H(\eta,\xi,\zeta')} \sum_{\eta'\in\Omega_{N-j}} e^{-\beta H(\zeta,-\xi,\eta')} \le e^{-\beta H(\underline{\sigma})} Z_N \tag{3.59}$$

for all $1\le j\le N$, $\zeta\in\Omega_{j-1}$, $\xi=\pm1$ and $\zeta'\in\Omega_{N-j}$. Namely, it is true almost surely after taking logs, dividing by N and passing to the limit as $N\to\infty$. One might wonder whether a stronger, deterministic form of this inequality is true. It would simplify the long probabilistic argument of the last subsubsection and make extensions to similar models easier. But this is unfortunately not the case, as the following example shows. Let N be an odd number, $j=(N+1)/2$, $\xi=-1$ and define

$$\begin{aligned} x(\eta,-1,-1,\ldots,-1) &= 1 \quad \textit{for all } \eta\in\Omega_{j-1},\\ x(-1,\ldots,-1,1,\eta') &= 1 \quad \textit{for all } \eta'\in\Omega_{N-j},\\ x(\sigma) &= 0 \quad \textit{in all other cases.} \end{aligned}$$

Then it is clear that the left hand side in (3.59) (with the Boltzmann factors replaced by $x(\cdot)$) equals 2^{N-1} while the right hand side equals $2^{(N+1)/2}$.

Combining the bounds of Propositions 3.1 and 3.3, we have the following result.

Theorem 1 *For all β, there exist finite constants c, c' and c'' such that, with $I\!P$-probability 1, for all but a finite number of indices N, we have*

$$\beta\beta_c - c\beta\sqrt{\frac{\log N}{N}} \le \frac{1}{N}\log\tau \le \beta\beta_c + c'\beta\sqrt{\frac{\log N}{N}} + c''\frac{\log N}{N}. \tag{3.60}$$

In particular,

$$T(\beta) := \lim_{N\to\infty}\frac{1}{N}\log\tau \tag{3.61}$$

exists with probability 1 and equals $\beta_c\beta$.

Remark 3.3 *Although the REM as an equilibrium model exhibits a (third order) phase transition at $\beta=\beta_c$, this leaves no trace in the dynamical model (as regards T), neither at β_c nor elsewhere. We might then say that there is no dynamical phase transition for the REM under Metropolis dynamics. We did not work out the details, but believe this is so also for other (local) Glauber dynamics. In an appendix to the paper, we consider a modification of Metropolis so that it is a global dynamics for which (3.61) has a third order phase transition at β_c.*

Remark 3.4 *An explanation for the result of Theorem 1 is as follows. Since the energies of the REM are independent, the minimum one is surrounded by order 1 energies that can have fluctuations of order* $\sqrt{N \log N}$. *Therefore, the time to exit the ground state under a local dynamics is of the order of* $\exp \beta_c \beta N$ *and the fluctuations are of order* $\exp \sqrt{N \log N}$. *This should be the main contribution to* τ *to leading order. In the high temperature regime, the dominant contribution to the free energy comes from energies that are larger than the ground state, but the entropy comes into play to give a smaller free energy. However, with the local dynamics, the stochastic process is trapped into the ground states, that do not contribute to the statics, for a time which is bigger than all the exit times of the other states.*

By using equations (3.5) and (1.7), it is not difficult to check the following result.

Theorem 2 *There exists a positive constant c such that for all* $\epsilon > 0$, *all* $\beta > 0$ *and all* $t' > 0$, *if*

$$t_N(t') = e^{N\beta\beta_c + c\sqrt{N\log N}} \left[\frac{N}{2}(F(\beta) + \beta\beta_c) + \frac{\beta}{2\beta_c}(\log N + \epsilon) + t' \right] \tag{3.62}$$

then, with $I\!P$*-probability 1, for all but a finite number of indices N we have*

$$sup_\sigma \| P_{t_N(t')}(\sigma, \cdot) - \mu_N(\cdot) \|_{var} \leq e^{-t'}. \tag{3.63}$$

Remark 3.5 *For the proof of this theorem, we need only an upper bound for the square root term in (3.5), so a lower bound on* $\mu(\sigma)$ *is needed. Since we want a result which is uniform with respect to the initial conditions, we cannot exclude a priori to start with a configuration which corresponds to the spin configuration that makes* $H(\sigma)$ *maximal. Even if the Gibbs measure of such configuration is very small. Also, we cannot expect to have any cancellation of the error terms that are of order* $\log N$. *This is because a part of the errors terms are coming from the fluctuations of a Gibbs factor computed on a maxima and the other part is coming from the fluctuations of the partition function, the latter coming from the fluctuations of the minimum.*

4 A Global Dynamics

In this chapter, we consider a *global* Metropolis dynamics, where the system can make all possible jumps, not only nearest neighbor ones as in the *local* dynamics. When scaled as for the local dynamics of the previous sections, τ is shown to exhibit a phase transition as a function of β (see Theorem 3 below).

Consider the following continuous-time Markov chain in Ω.

$$\begin{aligned} P(\sigma, \sigma') &= 2^{-N} \exp\{-\beta(H(\sigma') - H(\sigma))^+\}, \quad \text{if} \quad \sigma' \neq \sigma, \\ &= 1 - \sum_{\sigma'' \neq \sigma} P(\sigma, \sigma''), \quad \text{if} \quad \sigma' = \sigma, \end{aligned} \tag{4.1}$$

This is an irreducible, reversible with respect to μ_N chain and thus we can apply the variational characterization of the gap (3.4). We use the same trial function for τ as in Subsection 3.1 to obtain

$$\tau \geq \tau(\phi_{\underline{\sigma}}) = \frac{2^N e^{-\beta H(\underline{\sigma})}}{Z_N}. \tag{4.2}$$

The upper bound is much simpler to derive than the corresponding local one, either directly from the variational characterization or by using the Poincaré inequality (3.21). We choose the latter, using the set of paths

$$\Gamma = \{(\sigma, \sigma') : \sigma, \sigma' \in \Omega, \, \sigma \neq \sigma'\}$$

constituted of all bonds in Ω (notice that the global transition probabilities are nonzero over these) to get

$$\frac{2^N}{Z_N} \max_{\sigma \neq \sigma'} e^{\beta(H(\sigma) \vee H(\sigma'))} e^{-\beta(H(\sigma)+H(\sigma'))} = \frac{2^N e^{-\beta H(\underline{\sigma})}}{Z_N}. \tag{4.3}$$

We conclude from (4.2) and (4.3) that

$$\tau = \frac{2^N e^{-\beta H(\underline{\sigma})}}{Z_N}. \tag{4.4}$$

Therefore, since $\lim_{N\to\infty} F_N(\beta) = F(\beta)$ $I\!P$-almost surely, we get the following result.

Theorem 3 *$I\!P$-almost surely,*

$$T(\beta) = \lim_{N\to\infty} (\log \tau)/N = \beta_c \beta + \beta_c^2/2 - F(\beta), \tag{4.5}$$

where F is given by (3.48) and (3.49) above.

Remark 4.1 *Since F undergoes a third order phase transition in $\beta = \beta_c$, then so does T.*

Remark 4.2 *The identity (4.4) is valid for the same dynamics for any spin system. Applied to the Ising model, it shows that T inherits its phase transition (whenever that occurs).*

To consider the error terms, we introduce the quantity

$$\mathcal{T}(\beta) \equiv \log \left[\tau e^{-NT(\beta)} \right]. \tag{4.6}$$

The first result is an *almost sure* one.

Theorem 4 *If $\beta \geq \beta_c$, then*

$$\mathcal{T}(\beta) \leq 0 \quad \textit{for all } N \tag{4.7}$$

and, with $I\!P$-probability 1,

$$\liminf_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log \log N} \geq -\frac{\beta}{\beta_c}. \tag{4.8}$$

If $\beta < \beta_c$, then, with $I\!P$-probability 1,

$$\limsup_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = -\liminf_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = \frac{\beta}{2\beta_c}. \tag{4.9}$$

Remark 4.3 *The important fact is that the almost sure finite volume corrections to the free energy, that are of order* $\log N/N$ *in the low temperature regime, are coming precisely from the fluctuations of the ground states, and there is an exact cancellation of these fluctuations when τ is considered. This gives fluctuations that are of order at most* $\log\log N/N$ *for $T(\beta)$. In particular this implies that there is also a phase transitions in the errors terms.*

One could be interested in the error terms in probability. In this case the results are simpler.

Theorem 5 *If $\beta \geq \beta_c$, then*

$$\lim_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = 0 \quad \textit{in } I\!P\textit{-probability.} \tag{4.10}$$

If $\beta < \beta_c$, then

$$\lim_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = -\frac{\beta}{2\beta_c} \quad \textit{in } I\!P\textit{-probability.} \tag{4.11}$$

Proof of Theorem 4

We consider first the case where $\beta < \beta_c$. Using the explicit formula for $F(\beta)$ see (3.48) and (3.49) we get

$$\mathcal{T}(\beta) = -\beta(H(\underline{\sigma}) + N\beta_c) - \log Z_N e^{-NF(\beta)}. \tag{4.12}$$

Using the proposition 5 in ([4]), that is (1.5), we get that with $I\!P$-probability 1, for all but a finite number of indices N the last term is of order at most $Ne^{-\lambda(\beta)N}$. For the first term, using (3.14) we get with $I\!P$-probability 1, for all but a finite number of indices N

$$\mathcal{T}(\beta) \geq -\frac{\beta}{2\beta_c}(1+\epsilon)\log N. \tag{4.13}$$

On the other hand we have also, with $I\!P$-probability 1

$$\liminf_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = -\frac{\beta}{2\beta_c}. \tag{4.14}$$

This is a direct consequence of the proof of the formula (2.7) in [7] (see pages 520-521 there). Moreover we have with $I\!P$-probability 1, for all but a finite number of indices N

$$\mathcal{T}(\beta) \leq \frac{\beta}{2\beta_c}(1+\epsilon)\log N, \tag{4.15}$$

which is a consequence of the following estimate that is easy to check.

$$I\!P\left(-H(\underline{\sigma}) \leq \beta_c N(1 - \frac{\log N}{2\beta_c^2 N} - \frac{\log\log N^{1+\delta} + \log\sqrt{2\pi}}{\beta_c^2 N})\right) \leq \frac{c}{N^{1+\delta}} \tag{4.16}$$

for some positive constant c.

Moreover it is not too difficult to see that we have, with $I\!P$-probability 1

$$\limsup_{N\to\infty} \frac{\mathcal{T}(\beta)}{\log N} = \frac{\beta}{2\beta_c}, \tag{4.17}$$

from which we get (4.9).

For the proof of (4.7), we use again the explicit formula for $F(\beta)$. If $\beta \geq \beta_c$ we have

$$\mathcal{T}(\beta) = -\log Z_N e^{\beta H(\underline{\sigma})}. \tag{4.18}$$

Since $Z_N e^{\beta H(\underline{\sigma})} \geq 1$, we get $\mathcal{T}(\beta) \leq 0$.

The proof of (4.8) is just a little more involved and will make use of results of [4] with some modifications.

We define as in [4] (page 135), the real interval

$$I_1 = \left[N\beta_c \left(1 - \frac{1}{2\beta_c^2} \frac{\log N}{N} + \frac{(1+\epsilon)}{\beta_c^2} \frac{\log\log N}{N} \right), (1+\epsilon) N\beta_c \right]. \tag{4.19}$$

Then we have

$$\sum_\sigma e^{-\beta(H(\sigma)-H(\underline{\sigma}))} \mathbb{I}_{I_1}(-H(\sigma)) \leq \sum_\sigma \mathbb{I}_{I_1}(-H(\sigma)), \tag{4.20}$$

since $H(\sigma) - H(\underline{\sigma}) \geq 0$.

Now is it easy to check that

$$2^N I\!E(\mathbb{I}_{I_1}(-H(\sigma))) \leq \frac{1}{\beta_c} \left(\frac{1}{\log N} \right)^{1+\epsilon}. \tag{4.21}$$

Using now the following nice, but not known as it deserves, inequality, which is a simple consequence of the Markov inequality,

$$I\!P(S_n \geq r) \leq e\,(np(n))^r, \tag{4.22}$$

where $S_n = \sum_{i=1}^N x_i$ and $(x_i)_{i=1,\ldots,N}$ is a family of independent identicaly distributed random variables with values in $0,1$ and $I\!P(x_1 = 1) = p(n)$, we get

$$I\!P\left(\sum_\sigma \mathbb{I}_{I_1}(-H(\sigma)) \geq \frac{\log N}{\log\log N} \right) \leq \frac{1}{N^{1+\epsilon}}. \tag{4.23}$$

We consider now the interval

$$I_2 = (-\infty, N\beta_c(1 - \frac{1}{2\beta_c^2} \frac{\log N}{N} + \frac{(1+\epsilon)}{\beta_c^2} \frac{\log\log N}{N})]. \tag{4.24}$$

Using Lemma 7 in [4], we have, with $I\!P$-probability 1, for all large enough N

$$\sum_\sigma \mathbb{I}_{I_2}(-H(\sigma)) e^{-\beta H(\sigma)} \leq \frac{(1+\epsilon)}{\beta_c} \log\log N e^{N\beta\beta_c(1-\frac{\log N}{2\beta_c^2 N})}. \tag{4.25}$$

Therefore, collecting on the one hand (4.16) together with (4.25) and on the other hand (4.23) together with (4.20), using the first Borel-Cantelli Lemma, we get that, with $I\!P$-probability 1, for all large enough N

$$Z_N e^{\beta H(\underline{\sigma})} \leq \frac{\log N}{\log\log N} + \frac{(1+\epsilon)}{\beta_c} \log\log N e^{\frac{\beta}{\beta_c} \log\log N^{1+\delta}}, \tag{4.26}$$

from which we get immediately

$$\mathcal{T}(\beta) \geq -\frac{\beta}{\beta_c} \log\log N - \log\log\log N \tag{4.27}$$

and this concludes the proof of the Theorem 4. The proof of Theorem 5 is immediate.

We leave as an exercise to the reader to state and prove the corresponding of Theorem 2 for this dynamics.

A Microscopic Representation

We establish here the microscopic representation for the REM hamiltonian mentioned in the introduction. This representation already appears in Derrida's papers without proof. We give an argument here for completeness.

Proposition A.1 *Let $\mathcal{H} = \{H(\sigma), \sigma \in \Omega\}$ be defined by Equation (1.2), where $\{J_\alpha,\, \alpha \subset \Lambda\}$ is a family of i.i.d. standard gaussian random variables. Then $\mathcal{H}$ is a family of i.i.d. gaussian random variables with mean 0 and variance N.*

Proof

The gaussianness, correct marginal mean and variance are clear. It suffices thus to establish independence. It is enough to show that the matrix

$$\{\sigma_\alpha;\, \sigma \in \Omega, \alpha \subset \Lambda\} \tag{A.1}$$

is orthogonal, that is,

$$\sum_{\alpha \subset \Lambda} \sigma_\alpha \sigma'_\alpha = 0 \tag{A.2}$$

for all distinct $\sigma, \sigma' \in \Omega$.

Given $\sigma, \sigma' \in \Omega$ with $\sigma \neq \sigma'$, let Δ denote the (nonempty) set where σ and σ' disagree, that is

$$\Delta = \{i \in \Lambda : \sigma_i \neq \sigma'_i\}. \tag{A.3}$$

We have

$$\sum_{\alpha \subset \Lambda} \sigma_\alpha \sigma'_\alpha = \sum_{\alpha \subset \Lambda} (-1)^{|\alpha \cap \Delta|}. \tag{A.4}$$

The last sum can be rewritten as

$$\sum_{k=0}^{M} \sum_{\alpha \subset \Lambda : |\alpha \cap \Delta| = k} (-1)^k, \tag{A.5}$$

where $M = |\Delta|$. It then equals

$$\sum_{k=0}^{M} g_\Delta(k)(-1)^k, \tag{A.6}$$

where $g_\Delta(k)$ is the number of distinct subsets α of Λ intersecting Δ at exactly k points. There are k choices out of M in Δ, which yields $\binom{M}{k}$ possibilities for $\alpha \cap \Delta$, and total freedom in $\Lambda \backslash \Delta$, which yields 2^{N-M} possibilities for $\alpha \backslash \Delta$. Thus

$$g_\Delta(k) = \binom{M}{k} 2^{N-M}. \tag{A.7}$$

We finally have

$$\sum_{\alpha \subset \Lambda} \sigma_\alpha \sigma'_\alpha = 2^{N-M} \sum_{k=0}^{M} \binom{M}{k} (-1)^k = 2^{N-M}(1-1)^M = 0, \tag{A.8}$$

since $M > 0$. The result is proven. $\square$

B A Domination Lemma

It is enough for the purposes of supporting the argument in the proof of Proposition 3.1 to consider the following setup. Let $X_1, \ldots, X_{n+1}$ be a sequence of continuous i.i.d. random variables with distribution function F and let M_n denote their maximum (which is almost surely unique) and let M denote the corresponding index (so that $M_n = X_M$). Define $Y_i = X_i$ if $i < M$ and $Y_i = X_{i+1}$ otherwise.

Lemma B.1

$$P(Y_i < y_i,\, i = 1, \ldots, n) \geq \prod_{i=1}^{n} P(X_1 < y_i). \tag{B.1}$$

It follows immediately that the maximum of any subset of $\{Y_1, \ldots, Y_n\}$ is stochastically dominated by the maximum of the same number of independent X's.

Proof

$$
\begin{aligned}
& P(Y_i < y_i,\, i = 1, \ldots, n) && \text{(B.2)}\\
= & \sum_{j=1}^{n+1} P(Y_i < y_i,\, i = 1, \ldots, n,\, M = j) && \text{(B.3)}\\
= & \sum_{j=1}^{n+1} P(X_i < y_i, X_i < X_j,\, i < j,\, X_i < y_{i-1}, X_i < X_j,\, i > j) && \text{(B.4)}\\
= & \sum_{j=1}^{n+1} \int P(X_i < y_i, X_i < x,\, i < j,\, X_i < y_{i-1}, X_i < x,\, i > j)\, dF(x) && \text{(B.5)}\\
= & (n+1) \int \prod_{i=1}^{n} P(X_1 < y_i \wedge x)\, dF(x) && \text{(B.6)}\\
\geq & \prod_{i=1}^{n} P(X_1 < y_i) \int (n+1) F^n(x)\, dF(x) && \text{(B.7)}\\
= & \prod_{i=1}^{n} P(X_1 < y_i), && \text{(B.8)}
\end{aligned}
$$

where the inequality follows from

$$F(x \wedge y) \geq F(x)F(y), \tag{B.9}$$

for all x, y. □

Acknowledgements.

L. F. and M. I. thank M. Cassandro, A. Galves and A. Pellegrinotti for stimulating discussions. L. F. further thanks R. Férnandez, P. Ferrari, E. Jordão Neves, E. Kira and V. Sidoravicius for valuable comments, suggestions and discussions. M. I. was partially supported by CNPq grant no. 453865/93-3. L. F. work was partially supported by Fapesp grants nos. 94/4303-5 and

95/8932-0 and is part of Fapesp Theme Grant no. 95/0790-1. He acknowledges support and hospitality when visiting the University of Rome ”La Sapienza” and the CPT, Luminy. Part of this work was written while he was visiting the Courant Institute (under CNPq grant no. 200125/96-6), where he found warm hospitality. M. I. acknowledges hospitality in São Paulo. Y. K. was partially supported by FAPESP (Proc. 96/04505–2) and by CNPq (Proc. 300334/93–1 and ProTeM-CC-II Project ProComb).

References

[1] Derrida B., Random Energy Model: Limit of a Family of Disordered Models, *Phys. Rev. Lett.* **45**, 79-82 (1980)

[2] Derrida B., Random Energy Model: An Exactly Solvable Model of Disordered Systems, *Phys. Rev. B* **24**, 2613-2626 (1981)

[3] Eisele Th., On a Third Order Phase Transition, *Comm. Math. Phys.* **90**, 125-159 (1983)

[4] Olivieri E., Picco P., On the Existence of Thermodynamics for the Random Energy Model, *Comm. Math. Phys.* **96**, 125-144 (1984)

[5] Diaconis P., Stroock D., Geometric Bounds for Eigenvalues of Markov Chains, *Ann. Appl. Probab.* **1**, 36-61 (1991)

[6] Cassandro M., Galves J.A. and Picco P., Dynamical phase transition in disordered systems: the study of a random walk model, *Ann. Inst. H. Poincaré (Physique Théorique)* **55**, 689-706 (1991)

[7] Galves J.A., Martinez S. and Picco P., Fluctuations in Derrida's Random Energy and Generalized Random Energy Models, *J. Stat. Phys.* **54** 515-529 (1989)

[8] Mezard M., Parisi G. and Virasoro M., Spin Glass theory and beyond, *World Scientific* (1980)

[9] Sompolinsky H. and Zippelius A., Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses, *Phys. Rev. B* **25** 6860-6875 (1982)

[10] Zegarlinski B., Strong Decay to Equilibrium in one Dimensional Random Spin Systems, *J. Stat. Phys.* **77** 717-732 (1994)

[11] Guionnet A. and Zegarlinski B., Decay to Equilibrium in Random Spin Systems on a Lattice, *Comm. Math. Phys.* **181** 703-732 (1996)

[12] Ben Arous G. and Guionnet A., Large Deviations for Langevin Spin Glass Dynamics *Probab. Theory Relat. Fields* **102**, 455-509 (1995)

[13] Sinclair A.J., Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.* **1** 351–370 (1992)

[14] Jerrum, M., Sinclair, A., Approximate counting, uniform generation and rapidly mixing Markov chains, *Inform. and Comput.* **82**, 93–133 (1989)