UNIVERSITÀ DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Matematica Pura ed Applicata

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MATEMATICHE

INDIRIZZO MATEMATICA

CICLO XXI

TWO PROBLEMS CONCERNING INTERACTING SYSTEMS:

1. **ON THE PURITY OF THE FREE BOUNDARY CONDITION POTTS MEASURE ON GALTON-WATSON TREES**
2. **UNIFORM PROPAGATION OF CHAOS AND FLUCTUATION THEOREMS IN SOME SPIN-FLIP MODELS**

Direttore della Scuola: Ch.mo Prof. Paolo dai Pra

Supervisori: Ch.mo Prof. Paolo dai Pra
Prof. Dr. Christof Külske

Dottorando: Marco Formentin

30 giugno 2009

Abstract: A rigorous approach to Statistical Physics issues often produces interesting mathematical questions. This Ph.D. thesis is composed of two different parts. One does not intersect the other, but both research topics lie at the interface between Probability Theory and Statistical Mechanics.

- In the first part we deal with reconstruction of a tree-indexed Markov chain on Galton-Watson trees, improving previous bound by Mossel and Peres, both for symmetric and strongly asymmetric chains. Moreover, we give some numerical estimates to compare our bound with those of other authors. We provide a sufficient condition of the form $\mathbb{Q}(d)c(M) < 1$ for the non-reconstructability of tree-indexed q-state Markov chains obtained by broadcasting a signal from the root with a given transition matrix M. Here $c(M)$ is a constant depending on the transition matrix M and $\mathbb{Q}(d)$ is the expected number of offspring on the Galton-Watson tree. This result is equivalent to proving the extremality of the free boundary condition Gibbs measure within the corresponding Gibbs-simplex. When considering the Potts model case we take this point of view too. Our theorem holds for possibly non-reversible M. In the case of the symmetric Ising model the method produces the correct reconstruction threshold, in the case of the (strongly) asymmetric Ising model where the Kesten-Stigum bound is known to be not sharp the method provides improved numerical bounds.

- In the second part of the thesis we give sharp estimates for time uniform propagation of chaos in some specials mean field spin-flip models exhibiting phase transition. The first model is the dynamical Curie-Weiss model, that can be considered as the most basic mean field model. The second example is a model proposed recently in the context of credit risk in Finance; it describes the time evolution of finantial indicators for a network of interacting firms. Although we have chosen to deal with two specific models, the method we use appear to be rather general, and should work for other classes of models. A substantial limitation of our results is that they are limited to the subcritical case or, in Statistical Mechanical terms, to the high temperature regime.

Sommario: Un approccio rigoroso a questioni di Fisica Statistica spesso produce interessanti problemi matematici. Questa tesi di dottorato è composta da due parti. La prima non interseca la seconda, ma entrambe stanno sul confine tra Teoria della Probabilità e Meccanica Statistica.

- La prima parte tratta il problema della ricostruzione per catene di Markov su alberi di tipo Galton-Watson. Miglioriamo i risultati precedentemente ottenuti da Mossel e Peres, sia per catene simmetriche che fortemente asimmetriche. Dimostriamo una condizione sufficiente della forma $\mathbb{Q}(d)c(M) < 1$ per la non ricostruzione di catene di Markov a q-stati sull'albero. Qui $c(M)$ è una costante che dipende dalla matrice di transizione M e $\mathbb{Q}(d)$ è la media del numero di figli per vertice nell'albero di Galton-Watson. Questo risultato è equivalente alla purezza della misura libera di Gibbs. Quando consideriamo il caso del modello di Potts assumiamo anche questo punto di vista. Il teorema è valido anche per catene non reversibili. Nel caso del modello di Ising il nostro risultato produce la correta soglia di ricostruzione, nel caso di catene (fortemente) asimmetriche dove si sa che il bound di Kesten-Stigum non è esatto il metodo usato dà risultati numerici migliori.

- Nella seconda parte diamo delle stime uniformi nel tempo per la propagazione del caos in alcuni modelli di spin con interazione a campo medio che presentano transizione di fase. Il primo è il modello dinamico di Curie-Weiss, che può essere considerato come il più semplice esempio di sistema con interazione a campo medio. Il secondo è un modello recentemente impiegato per spiegare i meccanismi del rischio di credito; esso descrive l'evoluzione temporale di indicatori finaziari per un gruppo di aziende interagenti quotate sul mercato. Anche se abbiamo trattato modelli specifici, crediamo che il metodo funzioni piuttosto in generale e che sia applicabile anche ad altre classi di modelli. Una limitazione sostanziale dei nostri risultati è che valgono solo nel caso sottocritico, che corrisponde, nel linguaggio della Meccanica Statistica, al regime di alta temperatura.

Acknowledgments

There are few people to whom I am indebted for various reasons: Paolo Dai Pra, Christof Külske, Ida Germana Minelli and Francesca Collet. No word in this thesis would have been possible without their help. I am sincerely grateful to each of them, and I hope they already know. In particular, I want to express my gratitude to Christof. He is a teacher, a co-worker and a friend. To meet him was a gift. Fortune has treated me kindly.

Contents

Part I

On the Purity of the free boundary condition Potts measure on Galton-Watson trees

Introduction to Part I

This first part of the thesis focuses on the purity transition for the Potts model on random and deterministic trees.

Interacting stochastic processes on trees and lattices often differ in a fundamental way: where a lattice model has a single transition point (a critical value for a parameter of the model) the corresponding model on a tree might possess multiple transition points. Such phenomena happen more generally for non-amenable graphs (where surface terms are no smaller than volume terms), trees being major examples [9]. A main example of an interacting model is the usual ferromagnetic Ising model [1]. Here the interesting property which gives rise to a new transition is the purity (sometimes called extremality) of the free boundary condition Gibbs measure.

On a tree the open boundary state will still be extremal in a temperature interval strictly below the ferromagnetic transition temperature. It ceases to be extremal at even lower temperatures.

Ferromagnetic order on a tree is characterized by the fact that a plus-boundary condition at the leaves of a finite tree of depth N persists to have influence to the origin when N tends to infinity. For the tree it now happens in a range of temperatures, that even though an all plus-boundary condition will be felt at the origin, a *typical boundary condition* chosen from the free boundary condition measure itself will not be felt at the origin for a range of temperatures below the ferromagnetic transition. The latter implies the extremality of the free boundary condition state.

In the following we write $\theta = \tanh \beta$ where β is the inverse temperature of the

Ising (or Potts) model and denote by d the number of children on a regular rooted tree. Then the ferromagnetic transition temperature is given by $d\theta = 1$, and the transition temperature where the free boundary condition state ceases to be extremal is given by $d\theta^2 = 1$.

A proof of the latter fact is contained in [5]. A beautiful alternate proof of the extremality for $d\theta^2 \leq 1$ for regular trees was given by Ioffe [6]. The method used therin was elegant but very much dependent on the two-valuedness of the Ising spin variable. This was exploited for the control of conditional probabilities in terms of projections to products of spins. Some care is necessesary to treat the marginal case where equality holds in the condition. Indeed, one needs to control quadratic terms in a recursion; this is difficult for a general tree where the degrees are not fixed. A second paper [7] proves an analogue of the condition for general trees with arbitrary degrees leaves this case open. Finally, for a general tree which does not possess any symmetries, [14] give a sharp criterion for extremality in terms of capacities. It remains an open problem to determine the extremal measures and the weights in the extreme decomposition of the open boundary condition state for $d\theta^2 > 1$.

Let us remark that the problem of extremality of the open boundary condition state is equivalent to the so-called Reconstruction Problem: an issue about noisy information flow on trees.

Reconstruction is a topic where people coming from Probability, Statistical Mechanics, Biology and Computer Science can give contributions: recently it have been of interest in spin glasses [15] and computational biology [13]. The Reconstruction Problem can be stated as follow: we send a signal (a plus or a minus) from the origin to the boundary, making a prescribed error (that is related to the temperature of the Ising model) at every edge of the tree. In this way one obtains a Markov chain indexed by the tree. The reconstruction problem on a tree is called to be solvable, if the measure, obtained on the boundary at distance N by sending an initial $+$, keeps a finite variational distance to the measure obtained by sending a $-$, as n tends to infinity.

Nonsolvability of reconstruction is equivalent to the extremality of the open boundary condition state [12, 10]. This is to say that there can be no transport of information along the tree between root and boundary, for typical signals. This equivalence makes clear what purity transition is about.

In the present part we aim at an explicit sufficient condition ensuring the extremality of the free b.c. state for the Potts model, generalizing the Ising condition $d\theta^2 \leq 1$. We consider the free boundary condition Gibbs measure of the Potts model on a random tree. We provide an explicit temperature interval below the ferromagnetic transition temperature for which this measure is extremal, improving older bounds of Mossel and Peres [11].

Consider an infinite random tree, without leaves, rooted at 0. Call d_i the number of children at each vertex i and let be $\mathbb{Q}$ their distribution. The same and independent at each vertex. The symbol $\mathbb{Q}$ stands also for the mean.
In this situation our main result, formulated for a random tree, is the following. Write

$$P = \{(p_i)_{i=1,\dots,q}, p_i \geq 0 \; \forall i, \; \sum_{i=1}^{q} p_i = 1\} \tag{1}$$

for the simplex of Potts probability vectors.

Theorem 0.0.1 *The free boundary condition Gibbs measure* $\mathbb{P}$ *is extremal, for* $\mathbb{Q}$*-a.e. tree* T *when the condition* $\mathbb{Q}(d_0)\frac{2\theta}{q-(q-2)\theta}\bar{c}(\beta,q) < 1$ *is satisfied. Here,*

$$\bar{c}(\beta,q) := \sup_{p\in P} \frac{\sum_{i=1}^{q}(qp_i - 1)\log(1+(e^{2\beta}-1)p_i)}{\sum_{i=1}^{q}(qp_i-1)\log qp_i}. \tag{2}$$

Let us comment briefly. It is known that on a regular tree of degree d, there is reconstruction beyond the Kesten-Stigum bound ($d\left(\frac{2\theta}{q-(q-2)\theta}\right)^2 > 1$) proven to be sharp, for every degree d, only when $q = 2$ and for d sufficiently large and $q = 3$ [21]. Our bound for non-reconstruction improves the one which has been previously given in [11] and it holds for every number of offspring. We recover the bound in [11] from our when we use the estimate $\bar{c}(\beta,q) \leq \theta$. This estimate we see indeed numerically. Moreover, numerically $\bar{c}(\beta,q)$ seems to decrease monotonically in q at fixed β.

In literature, there are other thresholds and important works: in particular those of Sly [21] and Montanari-Mezard [15] along with the conjecture for deterministic trees therein. The conjecture states that the Kesten-Stigum bound is sharp only when $q \leq 4$. The paper by Sly partially proves this conjecture; he proves that the Kesten-Stigum bound is sharp also when $q = 3$ with the degree d large enough and that it can not be sharp for $q > 5$. Thus, when d is small the problem of finding a rigorous sharp bound is still open for every $q \geq 3$. In this case (d small) our bounds seem to the best rigorous thresholds as of today as pointed out also in [22]. We don't think they can be sharp, however, they differ only few percent form the Kesten-Stigum bound when $q \leq 4$, and from the numerically determined thresholds of Montanari and Mezard for $q > 5$. We are going to describe these aspects in more details in the conclusions.

This thesis, where we present two different proof of Theorem 0.0.1, is organized as follows.

The first chapter is introductory to the the problem. We review the Ising Model on a regular deterministic tree (i.e. the degree of a vertex is fixed) and its ferromagnetic phase transition; then we pass to define purity transition describing also its equivalence to the solvability/non-solvability transition for the Reconstruction Problem, for a $+/-$ signal sent from the root to the boundary of the tree. This problem is equivalent to the extremality of the free boundary Gibbs measure for the Ising model. Here we give also a proof for the threshold of purity transition in the Ising model case; this is a simplification of that due to Peres and Pemantle[14].

In the second chapter we switch to the purity transition for the Potts model where the possible signal sent from the origin can assume q values (i.e. $1, 2, \ldots, q$) instead of only two ($+$ or $-$) as in the former case. We give the first proof of our main result (i.e. Theorem 0.0.1) stated for Galton-Watson trees. A second alternative proof is given in the third chapter. This second proof is more general and it can be applied beyond the Potts model and, as an application, we treat the case of an asymmetric binary channels. For strongly asymmetric channels the bound we derive, improves the known threshold for this model [11]. When the asymmetry is small there exists a tight bound [8] that, until the present, we are not able to recover with our method.

The last chapter is for comments and comparison with thresholds for purity transition obtained by others authors. We propose also some conjectures.

The results of this part of the thesis are in:

- M. Formentin, C. Külske, *On the Purity of the free boundary condition Potts measure on random trees*, to be published in Stochastic Processes and their Applications (2009), available at arXiv:0810.0677, (2008);

- M. Formentin, C. Külske, *A symmetric entropy bound on the non-reconstruction regime of Markov chains on Galton-Watson trees*, preprint, available at arXiv:0903.2962, (2009).

Chapter 1

Ising model on trees: purity and reconstruction

1.1 Introduction

This is an introductory chapter. We review the purity/non-purity phase transition and its equivalence with the non-solvability/solvability transition in the Reconstruction Problem on regular trees [9, 11]. We describe this equivalence to make intuitive the meaning of purity transition: to say that the Free Gibbs Measure is pure is to say that there is no transport of information along the tree, form the boundary to the root.

In this chapter we deal mainly with the Ising model on regular trees for which we prove a bound for the purity transition. The proof is a simplification of [14]. The same method provides a proof of the ferromagnetic transition too.

1.2 Purity for the Ising model and reconstruction for binary channels

In the next, the situation is as follows: we have a tree T, where we have chosen a special point, the root, that we denote by 0 (see Figure 1.1). Recall that a graph is a set of the vertices called V, connected with edges, G, and that a tree is a graph without

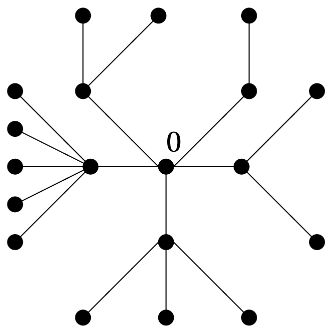


Figure 1.1: A tree rooted at 0.

loop. Thus, on a tree there is a unique chain of edges γ from one vertex v to another vertex ω. This induces a natural notion of distance as the number of edges in γ. We write $\omega > v$ if ω has a greater distance than v from the root. The set of vertices at distance N from the root is the level N of the tree. We indicates with T^N the sub-tree with just N levels and with ∂T^N the level N: i.e. the boundary of the sub-tree. Moreover, T_v^N is the sub-tree of T^N rooted at v.

The *Ising model* is obtained by putting at every vertex v of T^N a random variable $\eta(v) \in \{+1,-1\}$ (also called spin), and assigning to every configuration $\eta \in \{+1,-1\}^{\#(V)}$ the probability:

$$\mathbb{P}_\beta^\beta = \frac{1}{Z} \prod_{v \to \omega, \omega > v} e^{\beta \eta(v) \eta(w)} \tag{1.1}$$

where Z is a normalization factor.

The product runs over all the vertices; $v \to \omega$, $\omega > v$ means all the couples (v,ω) where v is at distance one from ω and ω is a child of v meaning that it has a greater distance to the origin. We can look at the parameter β as the strength of the interactions between vertices at distance one. Coming the Ising model from Physics, β is often interpreted as the inverse of the temperature.

In (1.1) no boundary condition is specified, thus this is called the free Gibbs measure.

When dealing with the ferromagnetic transition we are interested in $\mathbb{P}_\beta^{N,+}$, where + means that we set to +1 all the spins in ∂T^N. More precisely, we are interested in the limit:

$$\lim_{N\to\infty} \mathbb{P}_\beta^{N,+}\left(\eta(0)=+1\right). \tag{1.2}$$

We want to know if (1.2) is greater than $\frac{1}{2}$ for a certain range of the parameter β. Or, in other words, if for some values of β the plus-boundary condition persists to have influence at the root even when its distance from the root grows and N tends to infinity. In this chapter we deal with regular tree of degree d. To us, the degree of a vertex is the number of its children and a tree will be said to be regular if all the vertices have the same degree.

For a regular tree where the degree d is the same for every vertex one has the following:

Theorem 1.2.1 *The inequality*

$$\lim_{N\to\infty} \mathbb{P}_\beta^{N,+}\left(\eta(0)=+1\right) > \frac{1}{2}, \tag{1.3}$$

holds only if $d\theta > 1$*, where* $\theta = \tanh\beta$.

For a proof see [1, 14] or later in this chapter.

So, ferromagnetic order on a tree is characterized by the fact that a plus-boundary condition at the leaves of a finite tree of depth N persists to have influence to the origin when N tends to infinity. From an heuristics point of view we can say that, if β is sufficiently large, there is a transport of information from the boundary of the tree to the origin even when N goes to infinity. Actually, there is a way to make precise this information theoretic interpretation of the Ising model: i.e. to show that non-solvabilty of the reconstruction problem is equivalent to the extremality of the free Gibbs measure.

The *Reconstruction Problem on symmetric binary channels* can be stated as follows. We send a signal from the root to the boundary, making a prescribed error at every edge of the tree. In this way one obtains a Markov chain indexed by the tree. That is, on the tree T we construct the following Markov process [11, 12]: to each edge e we associate a random variable $X(e)$ with

$$\mathbb{P}(X(e)=1)=\epsilon=1-\mathbb{P}(X(e)=-1). \tag{1.4}$$

All the variables $X(e)$ are independent. The value of the spin $\eta(v)$ at the vertex v will be:

$$\eta(v)=\eta(0)\prod_{e\in\gamma}X(e), \tag{1.5}$$

where γ is the unique path going from the root 0 to the vertex v. While the initial value of $\eta(0)$ is chosen at random uniformly.
Suppose you know the values of the spins at distance N from the origin of the tree. What can you say about the spin at the root of the tree? Which is the probability of guessing the original value of the spin at 0 knowing that the configuration at the boundary ∂T^N is ξ, when N goes to infinity? These questions define the Reconstruction Problem. More formally

Definition 1.2.2 *The Reconstruction Problem is said to be solvable when*

$$\lim_{N\to\infty}\frac{1}{2}\sum_{\xi\in\eta_{\partial T^N}}\left|\mathbb{P}^M(\partial T^N=\xi|\eta(0)=1)-\mathbb{P}^M(\partial T^N=\xi|\eta(0)=-1)\right|>0. \tag{1.6}$$

To us, the quantity of interest is[1]

$$\Delta_N(T,\epsilon)=\mathbb{E}\left(\left|\mathbb{P}^M(\eta(0)=1|\partial T^N=\xi)-\mathbb{P}^M(\eta(0)=-1|\partial T^N=\xi)\right|\right). \tag{1.7}$$

Lemma 1.2.3 *The Reconstruction Problem is solvable only if*

$$\lim_{N\to\infty}\Delta_N(T,\epsilon)>0. \tag{1.8}$$

Proof: Because of the symmetry of the model $\mathbb{P}^M(\eta(0)=1)=\mathbb{P}^M(\eta(0)=-1)=\frac{1}{2}$ and moreover, the Bayes' formula gives:

$$\mathbb{P}^M(\partial T^N=\xi|\eta(0)=1)=\frac{\mathbb{P}^M(\eta(0)=1|\partial T^N=\xi)}{\mathbb{P}^M(\eta(0)=1)}\mathbb{P}^M(\partial T^N=\xi). \tag{1.9}$$

Substituting in (1.6) the result follows. □

[1] Here M stands for Markov. This is the probability coming from the Markov process constructed before. For N finite this is different from the Gibbs measure on the tree (see few lines below). In the sequel we often drop the index M.

$\Delta_N(T,\epsilon)$, intuitively, can be regarded as the difference between the probabilities of a correct and incorrect reconstruction knowing the configuration at the boundary of the tree. One wants to investigate if there exists a critical value ϵ_c such that:

$$\lim_{N\to\infty} \Delta_N(T,\epsilon) = 0 \text{ if } \epsilon \geq \epsilon_c. \tag{1.10}$$

For $\epsilon \geq \epsilon_c$ the problem is non-solvable and solvable otherwise.

Now, it happens that if you choose

$$\frac{\epsilon}{1-\epsilon} = \exp(-2\beta). \tag{1.11}$$

on an infinite tree, the law of the Markov process defined before is the limit of the Gibbs measure

$$\mathbb{P}^N_\beta(\eta) = \frac{1}{Z} \prod_{v\to\omega,\omega>v} e^{\beta\eta(v)\eta(\omega)}, \tag{1.12}$$

as the size of the tree grows to infinity. Then, $\lim_{N\to\infty} \Delta_N(T,\epsilon) = 0$ means that

$$\lim_{N\to\infty} \mathbb{E}\left|\mathbb{P}^\beta_N(\eta(0) = 1|\partial T^N = \xi) - \mathbb{P}^\beta_N(\eta(0) = -1|\partial T^N = \xi)\right| = 0. \tag{1.13}$$

Equation (1.13) is equivalent to the definition of purity for the limiting Gibbs measure. *In this way one states that non-solvability of reconstruction is equivalent to the purity of the free Gibbs measure* [10, 12], which is to say that there can be no transport of information along the tree between root and boundary, for typical signals.

In the sequel we write $\mathbb{P}^{N,\xi}_\beta(\cdot)$ for $\mathbb{P}^N_\beta(\eta(0) = \cdot|\partial T^N = \xi)$ and $\mathbb{P}^{N,\xi}_{\beta,v}(\cdot)$ when we are considering the sub-tree of T^N rooted at v. Some times we drop the dependence from β and N when it is clear.

Remark: Notice that here $\mathbb{P}^{N,\xi}_\beta(\cdot)$ is a random variable with respect to the free Gibbs measure on the boundary condition $\partial T^N = \xi$.

Let us say here that (1.13) is equivalent to

$$\lim_{N\to\infty} \mathbb{P}\left(\xi : \left|\mathbb{P}^{N,\xi}_\beta(i) - \frac{1}{2}\right| > \epsilon\right) = 0, \tag{1.14}$$

for $i = +, -$, and all $\epsilon > 0$. We will use often this characterization later. The proof of the equivalence is immediate using that $|\mathbb{P}_\beta^{N,\xi}(+) - \mathbb{P}_\beta^{N,\xi}(-)|$ is always positive but in $\mathbb{P}_\beta^{N,\xi}(+) = \mathbb{P}_\beta^{N,\xi}(-) = \frac{1}{2}$.

For a regular tree of degree d, it holds the following theorem [1, 5, 6, 14].

Theorem 1.2.4 *The limiting Gibbs measure on a regular tree of degree d is pure, i.e.*

$$\lim_{N\to\infty} \mathbb{E}\left|\mathbb{P}_\beta^{N,\xi}(\eta(0) = +) - \mathbb{P}_\beta^{N,\xi}(\eta(0) = -)\right| = 0 \tag{1.15}$$

if $d\theta^2 < 1$ *and* $\theta = \tanh(\beta)$.

The Theorem is true with equality too: that is the Kesten-Stigum bound is sharp for $q = 2$. Here we give this version to have a proof with all the main ideas, but simpler. A proof of the latter fact is contained in [5]. A beautiful alternative proof of the extremality for $d\theta^2 < 1$ for regular trees was given by Ioffe [6]. The method used there in was elegant, but very much dependent on the nature of the Ising model's spin variable: i.e. on the fact that it can assume only two values. The proof we give could be generalized to the Potts model [17] to obtain the same bound of [11]. This is a simplified version of [14].

1.3 Proof of Theorem 1.2.4

The method of the proof is by controlling the recursions from the outside to the inside of a tree of the log-likelihood ratios:

$$X_v^N := \log\left(\frac{\mathbb{P}_v^N(\eta(v) = +)}{\mathbb{P}_v^N(\eta(v) = -)}\right). \tag{1.16}$$

We have to prove that:

$$\lim_{N\to\infty} \mathbb{E}(|X_0^N|) = 0, \tag{1.17}$$

because this condition is equivalent to (1.15). Following [14] we prove (1.17) with the help of another quantity. We define:

$$Q_v^{N,+}(X_v^N) = \int X_v^N \, dQ_v^{N,+}(\xi) \tag{1.18}$$

where $Q_v^{N,+}$ is the shorthand for the probability for the boundary configuration ξ of T^N knowing that $\eta(v) = +1$:

$$Q_v^{N,+}(\xi) = \mathbb{P}_\beta^N \left(\eta : \eta_{\partial T_v^N = \xi} | \eta(v) = +1\right). \tag{1.19}$$

In the same way one defines:

$$Q_v^{N,-}(\xi) = \mathbb{P}_\beta^N \left(\eta : \eta_{\partial T_v^N = \xi} | \eta(v) = -1\right). \tag{1.20}$$

We need the following Lemmas [11]:

Lemma 1.3.1 *For X_v^N one has the iteration:*

$$X_v^N = \sum_{\omega : v \to \omega} g(X_w^N), \tag{1.21}$$

with

$$g(x) = \log \left[\frac{\cosh\left(\frac{x}{2}\right) + \theta \sinh\left(\frac{x}{2}\right)}{\cosh\left(\frac{x}{2}\right) - \theta \sinh\left(\frac{x}{2}\right)} \right]. \tag{1.22}$$

Proof: To derive (1.21) write:

$$\mathbb{P}_v^{N,\xi}(\eta_v) = Z_v^{-1} \prod_{\substack{v \to \omega, \omega > v \\ \partial T_v^N = \xi}} \exp\left(\beta \eta(v) \eta(\omega)\right) \tag{1.23}$$

for the probability of the state η_v with boundary condition fixed equal to ξ and η_v is the restriction of η to the sub-tree T_v^N. Writing d_v for te degree of the vertex v, we obtain:

$$\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)=Z_v^{-1}\sum_{\eta_v:\eta(v)=+}\exp(\beta\eta(\omega))\prod_{\omega\to y,y>\omega}\exp\left(\beta\eta(y)\eta(\omega)\right)$$

$$=Z_v^{-1}\sum_{\eta_v:\eta(v)=+}\left(\exp(\beta\eta(\omega_1))\prod_{\omega_1\to y,y>\omega_i}\exp\left(\beta\eta(y)\eta(\omega_1)\right)\right)\times\dots$$
$$\times\left(\exp(\beta\eta(\omega_{d_v}))\prod_{\omega_{d_v}\to y,y>\omega_{d_v}}\exp\left(\beta\eta(y)\eta(\omega_{d_v})\right)\right)$$

$$=Z_v^{-1}\sum_{\eta_{\omega_1},\dots,\eta_{\omega_{d_v}}}\left(\exp(\beta\eta(\omega_1))\prod_{\omega_1\to y,y>\omega_1}\exp\left(\beta\eta(y)\eta(\omega_1)\right)\right)\times\dots$$
$$\times\left(\exp(\beta\eta(\omega_{d_v}))\prod_{\omega_{d_v}\to y,y>\omega_{d_v}}\exp\left(\beta\eta(y)\eta(\omega_{d_v})\right)\right)$$

$$=Z_v^{-1}\prod_{i=1}^{d_v}\left(\sum_{\eta_{\omega_i}}\exp(\beta\eta(\omega_i))\prod_{\omega_i\to y,y>\omega_i}\exp\left(\beta\eta(y)\eta(\omega_i)\right)\right)$$

$$=Z_v^{-1}\prod_{i=1}^{d_v}\left(\sum_{\eta_{\omega_i}}\exp(\beta\eta(\omega_i))Z_{\omega_i}\exp\left(\beta\eta(\omega_i)\right)\mathbb{P}_\omega^{N,\xi}\left(\eta_{\omega_i}\right)\right)$$

$$=Z_v^{-1}\prod_{i=1}^{d_v}\sum_{j=+,-}\left(\sum_{\eta_{\omega_i}:\eta(\omega_i)=j}Z_{\omega_i}\exp(\beta j)\mathbb{P}_\omega^{N,\xi}\left(\eta_{\omega_i}\right)\right)$$

$$=Z_v^{-1}\prod_{i=1}^{d_v}Z_{\omega_i}\sum_{j=+,-}\exp(\beta j)\mathbb{P}_\omega^{N,\xi}\left(\eta\omega_i=j\right).\quad(1.24)$$

In the same way:

$$\mathbb{P}_v^{N,\xi}\left(\eta(v)=-\right)=Z_v^{-1}\prod_{i=1}^{d_v}Z_{\omega_i}\sum_{j=+,-}\exp(-\beta j)\mathbb{P}_\omega^{N,\xi}\left(\eta(\omega_i)=j\right).\quad(1.25)$$

Take the ratio

$$\frac{\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)}{\mathbb{P}_v^{N,\xi}\left(\eta(v)=-\right)}=\prod_{i=1}^{d_v}\frac{\sum_{j=+,-}\exp(\beta j)\mathbb{P}_\omega^{N,\xi}\left(\eta(\omega_i)=j\right)}{\sum_{j=+,-}\exp(-\beta j)\mathbb{P}_\omega^{N,\xi}\left(\eta(\omega_i)=j\right)}$$

$$=\prod_{i=1}^{d_v}\frac{\exp(\beta)\exp\left(X_{\omega_i}^N\right)+\exp(-\beta)}{\exp(-\beta)\exp\left(X_{\omega_i}^N\right)+\exp(\beta)}=\prod_{i=1}^{d_v}\frac{\exp(\beta)\exp\left(\frac{X_{\omega_i}^N}{2}\right)+\exp(-\beta)\exp\left(-\frac{X_{\omega_i}^N}{2}\right)}{\exp(-\beta)\exp\left(\frac{X_{\omega_i}^N}{2}\right)+\exp(\beta)\exp\left(-\frac{X_{\omega_i}^N}{2}\right)},\tag{1.26}$$

and remember that $\exp(x)=\cosh(x)+\sinh(x)$ and $\exp(-x)=\cosh(x)-\sinh(x)$ to write:

$$\frac{\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)}{\mathbb{P}_v^{N,\xi}\left(\eta(v)=-\right)}=\frac{\cosh\left(\frac{X_{\omega_i}^N}{2}\right)\left(e^\beta+e^{-\beta}\right)+\sinh\left(\frac{X_{\omega_i}^N}{2}\right)\left(e^\beta-e^{-\beta}\right)}{\cosh\left(\frac{X_{\omega_i}^N}{2}\right)\left(e^\beta+e^{-\beta}\right)-\sinh\left(\frac{X_{\omega_i}^N}{2}\right)\left(e^\beta-e^{-\beta}\right)}.\tag{1.27}$$

Notice that $\frac{e^\beta-e^{-\beta}}{e^\beta+e^{-\beta}}=\tanh(\beta)$ and take the log in both sides of (1.27) to get the conclusion. □

Lemma 1.3.2 *The projection of $Q_v^{N,+}(\xi)$ onto the boundary condition of T_ω^N is*

$$\frac{1+\theta}{2}Q_\omega^{N,+}+\frac{1-\theta}{2}Q_\omega^{N,-}.\tag{1.28}$$

Proof: To prove the lemma we have to show that

$$Q_v^{N,+}=\prod_{i=1}^{d_v}\left(\frac{1+\theta}{2}Q_{\omega_i}^{N,+}+\frac{1-\theta}{2}Q_{\omega_i}^{N,-}\right).\tag{1.29}$$

$$\mathbb{P}_v^N\left(\partial T_v^N=\xi|\eta(v)=+\right)=\frac{\mathbb{P}_v^N\left(\partial T_v^N=\xi,\eta(v)=+\right)}{\mathbb{P}_v^N\left(\eta(v)=+\right)}$$

$$=\frac{Z_v^{-1}\prod_{i=1}^{d_v}Z_{\omega_i}\sum_{j=+,-}\exp(\beta j)\mathbb{P}_\omega^N\left(\partial T_\omega^N=\xi,\eta(\omega_i)=+\right)}{Z_v^{-1}\prod_{i=1}^{d_v}Z_{\omega_i}\sum_{j=+,-}\exp(\beta j)\mathbb{P}_\omega^N\left(\eta(\omega_i)=+\right)}$$

$$=\prod_{i=1}^{d_v}\left(\frac{\exp(\beta)\mathbb{P}_\omega^N\left(\partial T_\omega^N=\xi,\eta(\omega_i)=+\right)}{\left(\exp(-\beta)+\exp(+\beta)\right)\mathbb{P}_\omega^N\left(\eta(\omega_i)=+\right)}+\frac{\exp(-\beta)\mathbb{P}_\omega^N\left(\partial T_\omega^N=\xi,\eta(\omega_i)=-\right)}{\left(\exp(-\beta)+\exp(\beta)\right)\mathbb{P}_\omega^N\left(\eta(\omega_i)=-\right)}\right)$$

$$=\prod_{i=1}^{d_v}\left(\frac{1+\theta}{2}Q_{\omega_i}^{N,+}+\frac{1-\theta}{2}Q_{\omega_i}^{N,-}\right),\quad(1.30)$$

where $\theta=\tanh(\beta)$. □

Lemma 1.3.3 *For every function f odd,*

$$\int f\left(X_v^N\right)dQ_v^{N,+}(\xi)=\int f\left(|X_v^N|\right)\tanh\left(\frac{|X_v^N|}{2}\right)d\mathbb{P}_N^\beta(\partial T_v^N\xi).\quad(1.31)$$

Moreover, for the symmetry of the model with respect to the change of $+1$ with -1, one has :

$$\int f\left(X_v^N\right)dQ_v^{N,+}(\xi)=-\int f\left(X_v^N\right)dQ_v^{N,-}(\xi).\quad(1.32)$$

Proof: To prove (1.32) notice that

$$Q_v^{N,+}+Q_v^{N,-}=\mathbb{P}_v^N\left(\partial T_v^N=\xi\right)\quad(1.33)$$

and because of symmetry

$$\int f\left(X_v^N(\xi)\right)\mathbb{P}_v^N\left(\partial T_v^N=\xi\right)=0.\quad(1.34)$$

For (1.31) we have to compute the Radon-Nykodin derivative

$$\frac{dQ_v^{N,+}(\xi)}{d\mathbb{P}_v^N\left(\partial T_v^N=\xi\right)}=\frac{\mathbb{P}_v^N\left(\partial T_v^N=\xi,\eta(v)=+\right)}{\mathbb{P}_v^N\left(\eta(v)=+\right)\mathbb{P}_v^N\left(\partial T_v^N=\xi\right)}=$$

$$\frac{\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)}{\mathbb{P}_v^N\left(\eta(v)=+\right)}=2\frac{\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)}{\mathbb{P}_v^{N,\xi}\left(\eta(v)=+\right)+\mathbb{P}_v^{N,\xi}\left(\eta(v)=-\right)}=2\frac{\exp\left(X_v^N\right)}{1+\exp\left(X_v^N\right)}.\quad(1.35)$$

Moreover,

$$\frac{\exp(x)}{1+\exp(x)} = \frac{\exp(x/2)}{\exp(-x/2)+\exp(x/2)}$$

$$= \frac{1}{2}\frac{\sinh(x/2)+\cosh(x/2)}{\cosh(x/2)} = \frac{1}{2}\left(1+\tanh(x/2)\right) \quad (1.36)$$

and we conclude:

$$\frac{dQ_v^{N,+}(\xi)}{d\mathbb{P}_v^N\left(\partial T_v^N = \xi\right)} = 1+\tanh\left(\frac{X_v^N}{2}\right). \quad (1.37)$$

Thus

$$\int f\left(X_v^N(\xi)\right) dQ_v^{N,+}(\xi) = \int f\left(X_v^N(\xi)\right)\tanh\left(\frac{X_v^N}{2}\right) d\mathbb{P}_v^N\left(\partial T_v^N = \xi\right). \quad (1.38)$$

The integrand is even and we can take the absolute value of $X_v^N(\xi)$. □

Notice that for Lemma 1.3.3 to prove (1.17) one could prove that

$$\lim_{N\to\infty} Q_0^{N,+}(X_0^N) = 0 \quad (1.39)$$

as we do in following, using the Banach-Cacciopoli's fixed point lemma. We use first Lemma 1.3.1 and Lemma 1.3.2 to compute:

$$Q_v^{N,+}(X_v^N) = Q_v^{N,+}\left(\sum_{\omega:v\to\omega} g(X_\omega^N)\right)$$

$$= \sum_{\omega:v\to\omega}\left(\frac{1+\theta}{2}Q_\omega^{N,+} + \frac{1-\theta}{2}Q_\omega^{N,-}\right) g(X_\omega^N)$$

$$= \theta \sum_{\omega:v\to\omega} Q_\omega^{N,+} g(X_\omega^N). \quad (1.40)$$

Suppose you are on a regular tree of degree d. In this case, because of the symmetry, we can say that $Q_\omega^{N,+} g_\omega(X_\omega^N)$ are all equal even if rooted on different ω. So

$$Q_v^{N,+}(X_v^N) = d\theta Q_\omega^{N,+} g(X_\omega^N). \quad (1.41)$$

Now we take the Taylor expansion of $g(x)$. The function $g(x)$ is odd and concave for $x > 0$ thus, this for the Lemma 1.3.1 implies

$$Q_{\omega}^{N,+}(g(X_{\omega}^{N})) < Q_{\omega}^{N,+}(\theta X_{\omega}^{N}) \tag{1.42}$$

and so

$$Q_{v}^{N,+}(X_{v}^{N}) < d\theta^{2} Q_{\omega}^{N,+}(X_{\omega}^{N}). \tag{1.43}$$

Thus, if $d\theta^2 < 1$ the Banach-Cacciopoli's fixed point lemma can be applied and the iteration of (1.43) goes to zero when $N \to 0$. This concludes the proof.

1.4 Proof of Theorem 1.2.1

At this point, the proof of the theorem is quite simple. Consider

$$X_{v}^{N,+} := \log\left(\frac{\mathbb{P}_{\beta,v}^{N,+}\left(\eta(0) = +\right)}{\mathbb{P}_{\beta,v}^{N,+}\left(\eta(0) = -\right)}\right). \tag{1.44}$$

We have to show that $X_{0}^{N,+}$ goes to zero only if $d\theta \leq 1$.

The boundary condition is fixed, thus (1.44) is no more a random variable and to control iteration of numbers is much more simpler.

Lemma 1.3.1 holds also in this case,

$$X_{v}^{N,+} = \sum_{\omega: v \to \omega} g\left(X_{\omega}^{N,+}\right) \tag{1.45}$$

and the addendums of the sum on the right hand side are all equal due to symmetry of the tree. We have:

$$X_{v}^{N,+} = d g\left(X_{\omega}^{N,+}\right). \tag{1.46}$$

The function g is concave with $g(0) = 0$ and the supremum of the first derivative is $g'(0) = \theta$. This implies that (3.35) has a unique fixed point only when $d\theta \leq 1$. Moreover when $d\theta \leq 1$, (3.35) is a contraction (see fig. 1.2, and 1.3). This proves the theorem.

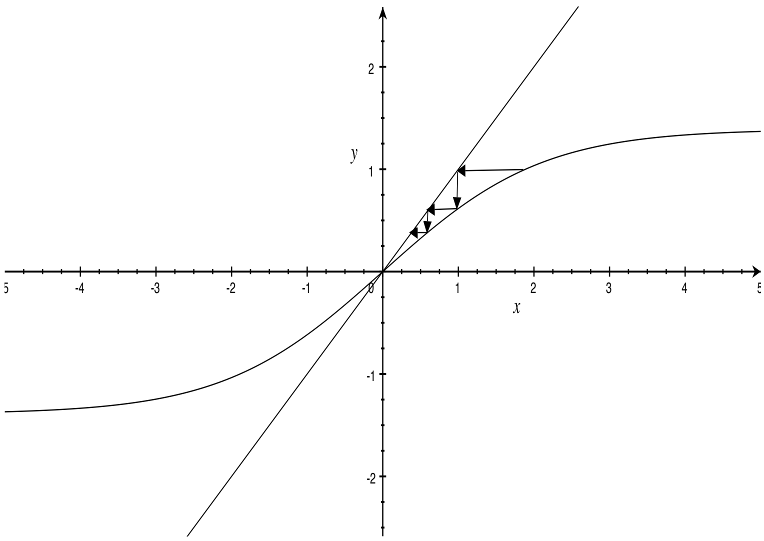


Figure 1.2: For $d\theta \leq 1$ there is a unique attractive stable point.

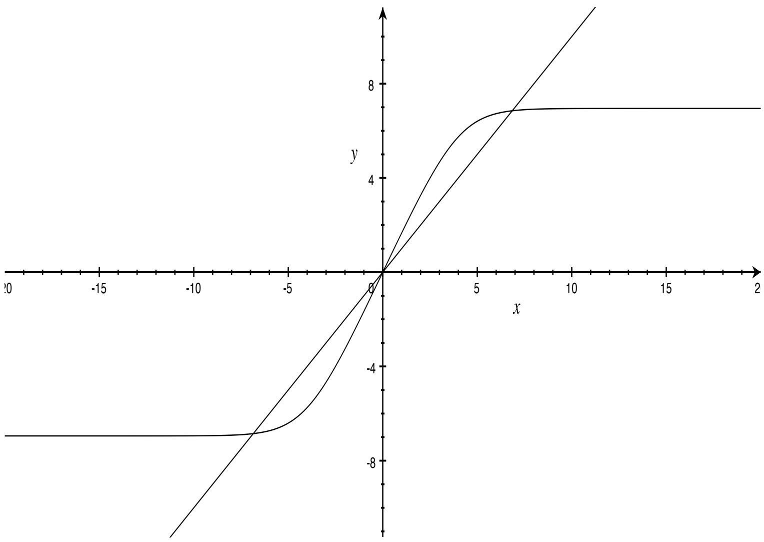


Figure 1.3: There are three fixed points when $d\theta > 1$.

Chapter 2

Purity transition on Galton-Watson trees (I): the Potts model

2.1 Introduction

In this chapter we give the first version of the proof of Theorem 0.0.1. We consider the free boundary condition Gibbs measure of the Potts model on a random tree. We provide an explicit temperature interval below the ferromagnetic transition temperature for which this measure is extremal, improving older bounds of Mossel and Peres. In information theoretic language extremality of the Gibbs measure corresponds to non-reconstructability for symmetric q-ary channels. We assume this point of view in the next chapter. The bounds for the corresponding threshold value of the inverse temperature are optimal for the Ising model and appear to be close to the Kesten-Stigum bounds on d-ary trees up to a factor of 0.0150 in the case $q = 3$ and 0.0365 for $q = 4$, independently of d. See the discussion in the last chapter for details.

Our proof uses an iteration of random boundary entropies from the outside of the tree to the inside, along with a symmetrization argument.

2.2 The Potts model on trees

We denote by T^N a finite tree rooted at 0 of depth N. Then the *free b.c. Potts measure on* T^N is the probability distribution $\mathbb{P}^N$ that assigns to a configuration $\eta_{T^N} = (\eta(v))_{v\in T^N} \in \{1,2,\dots,q\}^{T^N}$ the probability weights

$$\mathbb{P}_N^\beta = \frac{1}{Z}\prod_{v\to\omega} e^{2\beta\delta_{(\eta(v),\eta(\omega))}}. \tag{2.1}$$

Here $\delta_{(\cdot,\cdot)}$ is the Kronecker's delta. The sum is over all edges (v,w) of the tree T^N and Z is the partition function that makes the r.h.s. a probability measure.

The *free b.c. Potts measure on an infinite tree* T is by definition the weak limit $\mathbb{P} = \lim_{N\uparrow\infty}\mathbb{P}_{T^N}$ when T^N is an exhaustion of T. $\mathbb{P}$ is identical to what is called *the symmetric chain on* q*-symbols* in the context of the reconstruction problems in [12]. This chain has one parameter, namely the probability to change the symbol that is transmitted to any of the $q-1$ others, which is given by $\frac{1}{e^{2\beta}+q-1}$. Actually,the extremality for this model
(i.e. $\lim_{N\to\infty}\mathbb{E}|\mathbb{P}_N^\beta(\eta(0)=i)-\mathbb{P}_N^\beta(\eta(0)=j)|=0,\ \forall\ i,j=1,\dots,q$) is equivalent to the solvability of the Reconstruction Problem with a q-ary symmetric channel where, the probability of changing the signal passing from v to ω is

$$M(\eta(v)=i,\eta(\omega)=i) = 1-(q-1)l, \tag{2.2}$$

$$M(\eta(v)=i,\eta(\omega)=j) = l \tag{2.3}$$

with $l = \frac{1}{e^{2\beta}+q-1}$.

2.3 A criterion for extremality on random trees

Consider a random tree T with vertices i and number of children at the site i given by d_i. We choose d_i to be independent random variables with the same distribution $\mathbb{Q}$. We use the symbol $\mathbb{Q}$ also to describe the expected value. As is well known these appear as local approximations of random graphs which has newly emphasized their interest [16]. Our results however are already interesting in the case of regular trees where every vertex i has precisely d children.

Our main result, formulated for a random tree, is the following. Write

$$P = \{(p_i)_{i=1,\dots,q}, p_i \geq 0 \ \forall i, \ \sum_{i=1}^{q} p_i = 1\}$$

for the simplex of Potts probability vectors.

Theorem 2.3.1 *The free boundary condition Gibbs measure* $\mathbb{P}$ *is extremal, for* $\mathbb{Q}$*-a.e. tree* T *when the condition* $\mathbb{Q}(d_0)\frac{2\theta}{q-(q-2)\theta}\bar{c}(\beta,q) < 1$ *is satisfied. Here,*

$$\bar{c}(\beta,q) := \sup_{p\in P} \frac{\sum_{i=1}^{q}(qp_i - 1)\log(1+(e^{2\beta}-1)p_i)}{\sum_{i=1}^{q}(qp_i-1)\log qp_i}. \tag{2.4}$$

Remark : It appears that the supremum over P is achieved at the symmetric point $\frac{1}{q}(1,1,\dots,1)$ *only* in the Ising model $q = 2$. This implies the sharpness of the bound in the Ising case, see also the discussion at the end of the thesis.

Recall that Mossel and Peres proved for d-ary trees the bound

Theorem 2.3.2 *Consider a tree* T *with degree* d*. For*

$$d\frac{2\theta^2}{q-(q-2)\theta} < 1, \tag{2.5}$$

with $\theta = \tanh\beta$*, the free boundary condition Gibbs measure is extremal.*

The same type of result holds for random trees:

Theorem 2.3.3 *For* $\mathbb{Q}(d_0)\frac{2\theta^2}{q-(q-2)\theta} < 1$*, the free boundary condition Gibbs measure* $\mathbb{P}_T$ *is extremal, for* $\mathbb{Q}$*-a.e. tree* T*.*

We recover it from our bounds when we use the estimate $\bar{c}(\beta,q) \leq \theta$. This estimate we see indeed numerically. Moreover, numerically $\bar{c}(\beta,q)$ seems to decrease monotonically in q at fixed β.

Note also the bounds of Martinelli et al. [20] (see Theorem 9.3., Theorem 9.3.' Theorem 9.3") who give a nice criterion for non-reconstruction involving a Dobrushin constant of the corresponding Markov specification which however give worse estimates in the Potts model.

Let us put this result in perspective. For the purpose of the discussion we specialize to the case of the regular tree with d children. Denote by $\mathbb{P}^{N,k}$ the measures on T^N obtained by putting the boundary condition k to all Potts-spins at the outer boundary, and denote by $\mathbb{P}^k$ the corresponding limiting measures on T.

Absence of ferromagnetic order (uniqueness of the Gibbs measure) can be detected by the fact that the distribution of the spin η_0 at the origin under the infinite volume measure $\mathbb{P}^k$ is the equidistribution, independently of the boundary condition k. This condition is easy to obtain by considering a simple one-dimensional recursion of numbers (instead of measures). For more details see Section 3.13. Absence of ferromagnetic order in particular implies purity of the free b.c. state. In the language of the reconstruction problem this means non-solvability and as such the condition is mentioned as Proposition 4 in [12].

Let us compare with opposite results: It is known as the so-called Kesten-Stigum bound [3] that $d\lambda_2(\theta,q)^2 > 1$ implies reconstructability (i.e. non-extremality of the free boundary condition measure). Here $\lambda_2(\theta,q)$ is the second eigenvalue of the transition matrix that produces the free b.c. Potts model by broadcasting from the origin to the boundary; it is decreasing in q at fixed θ, and increasing in θ at fixed q. This is intuitively clear: the bigger the number of states q and the smaller the inverse temperature, the easier it is to forget about the information put at the boundary. Moreover it is proved as Theorem 2 in [12] that when one fixes d and a value of $d\lambda_2(\theta,q) \equiv \lambda > 1$, for q large enough the reconstruction problem is solvable for the corresponding value of θ.

Now, our method of proof is based on controlling recursions for the probability distributions at roots of subtrees from the outside to the inside of a tree. These are recursions on log-likelihood ratios of Potts probability vectors for the root of subtrees, and these ratios are random w.r.t. the boundary condition (which is chosen according to the free b.c. condition measure).

Understanding recursions for probability distributions (needed to investigate the purity of the free b.c. state) is much less straightforward than controlling recursions for real numbers (needed for investigating the existence of ferromagnetic order). We prove convergence to a Dirac-distribution by controlling the boundary relative entropy, generalizing from the approach of [14] for the Ising model. Novelties appear for the Potts model, a key point being proper symmetrization to bring out the constant (4.4), beginning with Lemma 2.4.2.

2.4 Proof of Theorem 2.3.1

To show the triviality of a measure μ on the tail sigma-algebra it suffices to show that, for any fixed cylinder event A we have

$$\lim_{N\uparrow\infty} \mu \left|\mu(A|\mathcal{T}_N) - \mu(A)\right| = 0, \tag{2.6}$$

where $\mathcal{T}_N$ is the sigma-algebra created by the spins that have at least distance N to the origin (see [4] Proposition 7.9).

We denote by T^N the tree rooted at 0 of depth N. The notation T_v^N indicates the sub-tree of T^N rooted at v obtained from "looking to the outside" on the tree T^N. We denote by $\mathbb{P}_v^{N,\xi}$ the correponding Potts-Gibbs measure on T_v^N with boundary condition on ∂T_v^N given by $\xi = (\xi_i)_{i\in\partial T_v^N}$. We denote by $\mathbb{P}_v^N$ the correponding Potts-Gibbs measure on T_v^N with free boundary conditions, as in (6.33).

We are going to show that the distribution of the probabilities to see a value s at the origin, obtained by putting a boundary condition ξ at distance N that is chosen according to the free measure $\mathbb{P}$ itself, converges to the equidistribution in probability. This reads

$$\lim_{N\uparrow\infty} \mathbb{P}\Big(\xi : \Big|\mathbb{P}^{N,\xi}(\eta(0) = s) - \frac{1}{q}\Big| \geq \varepsilon\Big) \to 0. \tag{2.7}$$

This then implies (2.6).

Sometimes we write

$$\pi_v^N = \Big(\mathbb{P}^{N,\xi}(\eta(v) = s)\Big)_{s=1,\dots,q} \tag{2.8}$$

To achieve (3.48) it is more convenient to look at the probability distribution for the spin at the root v obtained with the boundary condition ξ in terms of the "log-likelihood ratios" defined by

$$X_k^j(v;\xi) := \log \frac{\mathbb{P}_v^{N,\xi}(\eta(v) = j)}{\mathbb{P}_v^{N,\xi}(\eta(v) = k)}, \tag{2.9}$$

where $1 \leq j \neq k \leq q$. Ultimately we are interested to show the convergence of these quantities at $v = 0$ to zero, for all pairs j,k, in $\mathbb{P}$-probability, as the depth N of the tree tends to infinity.

We denote the measure at the boundary at distance N from the root on the tree emerging from v, which is obtained by conditioning the spin in the site v to take the value to be j, by

$$Q_v^{N,j}(\xi) := \mathbb{P}_v^N(\eta : \eta_{|\partial T_v^N} = \xi | \eta(v) = j). \tag{2.10}$$

Definition 2.4.1 *Denote the relative entropy of the boundary measures between the states obtained by conditioning the spin at v to be 1 respectively 2, by*

$$m_v^{(N)} = S(Q_v^{N,2}|Q_v^{N,1}) = \int Q_v^{N,2}(d\xi)\log\frac{Q_v^{N,2}(\xi)}{Q_v^{N,1}(\xi)}. \tag{2.11}$$

Here and in the sequel denote by w the children of v, indicated by the symbol $v \to w$.

Lemma 2.4.2 *The boundary relative entropy can be written as an expected value w.r.t. the open boundary condition Gibbs measure $\mathbb{P}$ in the form*

$$S(Q_v^{N,2}|Q_v^{N,1}) = \frac{1}{q-1}\int \mathbb{P}(d\xi)\sum_{i=1}^{q}\varphi\Big(q\mathbb{P}_v^{N,\xi}(\eta(v)=i)\Big), \tag{2.12}$$

with $\varphi(x) = (x-1)\log x$.

Proof: In the first step we express the relative entropy as an expected value

$$S(Q_v^{N,2}|Q_v^{N,1}) = q\int \mathbb{P}(d\xi) g\Big(\mathbb{P}_v^{N,\xi}(\eta(v)=2), \mathbb{P}_v^{N,\xi}(\eta(v)=1)\Big), \tag{2.13}$$

with

$$g(p_2,p_1) = p_2\log\frac{p_2}{p_1}. \tag{2.14}$$

To see this, we use that

$$\frac{dQ_v^{N,2}}{d\mathbb{P}_v^N}(\xi) = q\mathbb{P}_v^{N,\xi}(\eta(v)=2), \tag{2.15}$$

by the definition of the conditional probability and the fact that the marginal of $\mathbb{P}$ at any site is the equidistribution.

In the next step we use the invariance of $\mathbb{P}$ under permutation of the Potts-indices to write

$$S(Q_v^{N,2}|Q_v^{N,1}) = q\int \mathbb{P}(d\xi)(Rg)\Big(\mathbb{P}_v^{N,\xi}(\eta(v)=1),\mathbb{P}_v^{N,\xi}(\eta(v)=2),\dots,\mathbb{P}_v^{N,\xi}(\eta(v)=q)\Big), \tag{2.16}$$

where R is the symmetrization operator acting on functions $f(p_1,\dots,p_q)$ of Potts-probability vectors by

$$(Rf)(p_1,p_1,\dots,p_q) = \frac{1}{q!}\sum_\pi f(p_{\pi(1)},p_{\pi(2)},\dots,p_{\pi(q)}), \tag{2.17}$$

where π runs over the permutations of $\{1,\dots,q\}$.

One verifies that

$$(Rg)(p_1,p_1,\dots,p_q) = \frac{1}{q(q-1)}\sum_{i=1}^{q}(qp_i-1)\log qp_i, \tag{2.18}$$

which proves the lemma. □

2.4.1 Recursions for the boundary entropy for subtrees

Proposition 2.4.3 *The boundary relative entropy $m_v^{(N)}$ at the site v obeys the following linear recursive inequalities in terms of the values at the children w, given by*

$$m_v^{(N)} \le \frac{2\theta}{q-\theta(q-2)}\bar{c}(\beta,q)\sum_{w:v\to w} m_w^{(N)}. \tag{2.19}$$

Remark: Noting that $\frac{Q_v^{N,j}(\xi)}{Q_v^{N,k}(\xi)} = X_k^j(v;\xi)$ we may write

$$m_v^{(N)} = \int Q_v^{N,2}(d\xi)X_1^2(v;\xi). \tag{2.20}$$

Remark: Suppose that we are considering a spherically symmetric tree. This means that the number of offspring depends only on the generation, e.g. $d_v = d_{|v|}$ where $|v|$ is the distance of v to the origin (that is the length of the unique path from the origin to v). Then $m_v^{(N)} = m_{|v|}^{(N)}$ and so

$$m_k^{(N)} \le \frac{2\theta}{q-\theta(q-2)}\bar{c}(\beta,q)d_k m_{k+1}^{(N)}. \tag{2.21}$$

So $\lim_{N\uparrow\infty} m_0^{(N)} = 0$ is implied by $\sum_{k=1}^\infty \log(cd_k) = -\infty$ with $c = \frac{2\theta}{q-\theta(q-2)}\bar{c}(\beta,q)$.

Proof of Theorem 1.1 Taking expectation w.r.t. the random graph we note that $\mathbb{E}m_v^{(N)} = \mathbb{E}m_{|v|}^{(N)}$. Now, using Wald's inequality we have

$$\mathbb{E}m_k^{(N)} \le \frac{2\theta}{q-\theta(q-2)}\bar{c}(\beta,q)\mathbb{E}d_0\,\mathbb{E}(m_{k+1}^{(N)}). \tag{2.22}$$

From this follows that $\lim_{N\uparrow\infty}\mathbb{E}m_0^{(N)} = 0$ using the uniform boundedness in N, $\mathbb{E}m_{N-1}^{(N)} \le C\mathbb{E}(d_0)$. This can be seen from Lemma 2.4.4 a few lines below. □

To prove Proposition 2.4.3 at first a recursion for the log-likelihood ratios $X_k^j(v;\xi)$ has to be derived, for fixed finite tree of depth N from the outside to the inside. This iteration is standard, but we include its derivation for the convenience of the reader. The proof is quite similar to that for the Ising model. In the following we omit the dependence on the fixed boundary condition ξ in the notation.

Lemma 2.4.4 *For all indices* $1 \le j,k \le q$ *we have*

$$X_k^j(v) = \sum_{\omega: v\to w}\log\frac{\sum_{i\neq k,j}\exp[X_k^i(w)] + 1 + \exp(2\beta)\exp[X_k^j(w)]}{\sum_{i\neq k,j}\exp[X_k^i(w)] + \exp(2\beta) + \exp[X_k^j(w)]}. \tag{2.23}$$

Proof: Note that the Potts-measure $\mathbb{P}_v^{N,\xi}$ is proportional to the weight

$$W(\eta) = \prod_{x\to y, x\ge v}\exp[2\beta\delta_{\eta(x),\eta(y)}], \tag{2.24}$$

where the product is taken over the neighboring vertices coming after v looking from the root of the tree. The normalization factor will be Z_v^{-1}.
We want to rewrite $X_k^j(v)$ as a function of $X_k^j(w)$ where w are the children of v. The key observation is that

$$W(\eta_v) = \prod_{w:v\to w} W(\eta_w)\exp[2\beta\delta_{\eta(v),\eta(w)}], \tag{2.25}$$

where we have written η_v for the restriction of η to the sub-tree T_v^N. Now,

$$\mathbb{P}_v^{N,\xi}(\eta(v)=j) = Z_v^{-1}\prod_{w:v\to w}\sum_{\eta_w} W(\eta_w)\exp[2\beta\delta_{j,\eta(w)}]$$

$$= Z_v^{-1}\prod_{w:v\to w} Z_w\sum_{i=1}^{q} Z_w^{-1}\exp[2\beta\delta_{j,i}]\sum_{\eta_w:\eta(w)=i} W(\eta_w)$$

$$= Z_v^{-1}\prod_{w:v\to w} Z_w\sum_{i=1}^{q}\exp[2\beta\delta_{j,i}]\,\mathbb{P}_w^{N,\xi}(\eta(w)=i). \tag{2.26}$$

The same computation can be done for $\mathbb{P}_v^{N,\xi}(\eta(v)=k)$ to obtain:

$$\mathbb{P}_v^{N,\xi}(\eta(v)=k) = Z_v^{-1}\prod_{w:v\to w} Z_w \sum_{i=1}^{q}\exp[2\beta\delta_{k,i}]\mathbb{P}_w^{N,\xi}(\eta(w)=i). \tag{2.27}$$

Now consider the ratio and then divide everything by $\mathbb{P}_w^{N,\xi}(\eta(w)=k)$:

$$\frac{\mathbb{P}_v^{N,\xi}(\eta(v)=j)}{\mathbb{P}_v^{N,\xi}(\eta(v)=k)} = \prod_{w:v\to w}\frac{\sum_{i=1}^{q}\exp[2\beta\delta_{j,i}]\mathbb{P}_w^{N,\xi}(\eta(w)=i)}{\sum_{i=1}^{q}\exp[2\beta\delta_{k,i}]\mathbb{P}_w^{N,\xi}(\eta(w)=i)} =$$

$$= \prod_{w:v\to w}\frac{\sum_{i\neq k,j}\frac{\mathbb{P}_w^{N,\xi}(\eta(w)=i)}{\mathbb{P}_w^{N,\xi}(\eta(w)=k)} + 1 + \exp(2\beta)\frac{\mathbb{P}_w^{N,\xi}(\eta(w)=j)}{\mathbb{P}_w^{N,\xi}(\eta(w)=k)}}{\sum_{i\neq k,j}\frac{\mathbb{P}_w^{N,\xi}(\eta(w)=i)}{\mathbb{P}_w^{N,\xi}(\eta(w)=k)} + \exp(2\beta) + \frac{\mathbb{P}_w^{N,\xi}(\eta(w)=j)}{\mathbb{P}_w^{N,\xi}(\eta(w)=k)}}, \tag{2.28}$$

which proves the result. □

2.4.2 Controlling the recursion relation for the boundary entropy

Lemma 2.4.5

$$X_i^j(v) = \sum_{\omega:v\to w}\left[u\Big(\mathbb{P}_v^{N,\xi}(\eta(v)=j)\Big) - u\Big(\mathbb{P}_v^{N,\xi}(\eta(v)=i)\Big)\right], \tag{2.29}$$

where

$$u(p_1) = \log(1+p_1(e^{2\beta}-1)). \tag{2.30}$$

Proof: Remember the recursion given in Lemma 2.4.4. Now re-express the X's by the p-variables and use the fact that they form a probability vector. □

Using this we may derive the following equality on the iteration of the boundary entropy.

Lemma 2.4.6

$$Q_v^{N,2}X_1^2(v) = \frac{2\theta}{q-(q-2)\theta}\sum_{\omega:v\to\omega} Q_\omega^{N,2}\left[u\Big(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=2)\Big) - u\Big(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=1)\Big)\right]. \tag{2.31}$$

Proof: As the second piece of information next to Lemma 2.4.5 which is needed to understand the iteration for the boundary relative entropy $m_v^{(N)}$ we must see how the boundary measure $Q_v^{N,j}(d\xi)$, obtained by conditioning at v, relates to the boundary measures obtained by conditioning at the children, denoted by w.

For the Potts model we have

$$Q_v^{N,j} = \prod_{v\to\omega}\left[\frac{\exp(2\beta)}{(q-1)+\exp(2\beta)}Q_\omega^{N,j} + \frac{1}{(q-1)+\exp(2\beta)}\sum_{i\neq j}Q_\omega^{N,i}\right]$$

$$= \prod_{v\to\omega}\left[\frac{1+\theta}{q-(q-2)\theta}Q_\omega^{N,j} + \frac{1-\theta}{q-(q-2)\theta}\sum_{i\neq j}Q_\omega^{N,i}\right]. \quad (2.32)$$

Let us make this computation explicit.

$$Q_v^{N,j}(\xi) = \mathbb{P}_v^N\left(\eta : \eta_{|\partial T_v^N} = \xi | \eta(v) = j\right) = \frac{\mathbb{P}\left(\eta : \eta_{|\partial T_v^N} = \xi, \eta(v) = j\right)}{\mathbb{P}_v^N}$$

$$= \frac{Z_v^{-1}\prod_{\omega:v\to\omega} Z_\omega \sum_{i=1}^q \exp(2\beta\delta_{i,j})\mathbb{P}_w^{N,\xi}\left(\eta(\omega)=i\right)}{Z_v^{-1}\prod_{\omega:v\to\omega} Z_\omega \sum_{i=1}^q \exp(2\beta\delta_{i,j})\mathbb{P}_\omega^{N}\left(\eta(\omega)=i\right)}$$

$$= \prod_{\omega:v\to\omega}\frac{\sum_{i=1}^q \exp(2\beta\delta_{i,j})\mathbb{P}_\omega^{N,\xi}\left(\eta(\omega)=i\right)}{(q-1)\mathbb{P}_\omega^N\left(\eta(\omega)=i\right) + \exp(2\beta)\mathbb{P}_\omega^N\left(\eta(\omega)=i\right)}$$

$$= \prod_{\omega:v\to\omega}\sum_{i=1}^q \frac{\exp\left(2\beta\delta_{i,j}\right)}{(q-1)+\exp\left(2\beta\right)}\mathbb{P}_\omega^N\left(\eta : \eta_{|\partial T_\omega^N} = \xi | \eta(\omega) = i\right)$$

$$= \prod_{\omega:v\to\omega}\left[\frac{\exp(2\beta)}{(q-1)+\exp(2\beta)}Q_\omega^{N,j}(\xi) + \frac{1}{(q-1)+\exp(2\beta)}\sum_{i\neq j}Q_\omega^{N,i}(\xi)\right]. \quad (2.33)$$

Thus, from (3.43), to control the iteration we must look at the terms

$$\left[\frac{1+\theta}{q-(q-2)\theta}Q_\omega^{N,2} + \frac{1-\theta}{q-(q-2)\theta}Q_\omega^{N,1} + \frac{1-\theta}{q-(q-2)\theta}\sum_{i\geq 3}Q_\omega^{N,i}\right]$$
$$\left[u\left(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=2)\right) - u\left(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=1)\right)\right]. \quad (2.34)$$

We first note that, by symmetry under the measure $Q_\omega^{N,i}$, for $i = 3,\ldots,q$, the corresponding terms in the sum vanish. Now we use the permutation symmetry of the Potts indices to see the proof. □

Next we use the following representation.

Lemma 2.4.7

$$Q_v^{N,2} X_1^2(v) = \frac{2\theta}{q-(q-2)\theta} \sum_{\omega: v\to\omega} \int \mathbb{P}(d\xi) h\Big(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=2), \mathbb{P}_\omega^{N,\xi}(\eta(\omega)=1)\Big), \tag{2.35}$$

with

$$h(p_2,p_1) = q p_2 (u(p_2) - u(p_1)). \tag{2.36}$$

Proof: This follows as in the Proof of Lemma 2.4.2 by plugging in the Radon-Nikodym derivative of $Q_w^{N,2}$ w.r.t. the open b.c. measure.

□

With these preparations we can now finish the proof of the main proposition.

Proof of Proposition 2.4.3: Recalling the definition of the symmetrization operator (2.17) we obtain

$$Q_v^{N,2} X_1^2(v) = \frac{2\theta}{q-(q-2)\theta} \sum_{w: v\to w} \int \mathbb{P}(d\xi)(Rh)\Big(\mathbb{P}_w^{N,\xi}(\eta(w)=1), \dots, \mathbb{P}_w^{N,\xi}(\eta(w)=q)\Big), \tag{2.37}$$

where

$$(Rh)(p_1,\dots,p_q) = \frac{1}{q-1}\sum_{i=1}^{q}(q p_i - 1) u(p_i). \tag{2.38}$$

From here follows that

$$Q_v^{N,2} X_1^2(v) = \frac{2\theta}{q-(q-2)\theta} \sum_{\omega: v\to w} \int \mathbb{P}(d\xi) H\Big(\mathbb{P}_\omega^{N,\xi}(\eta(\omega)=1), \dots, \mathbb{P}_\omega^{N,\xi}(\eta(\omega)=q)\Big), \tag{2.39}$$

where

$$H(p_1,\dots,p_q) = \frac{1}{q-1}\sum_{i=1}^{q}(q p_i - 1)\tilde{u}(p_i), \tag{2.40}$$

with

$$\tilde{u}(p_1) = \log\frac{1+p_1(e^{2\beta}-1)}{1+\frac{1}{q}(e^{2\beta}-1)}. \tag{2.41}$$

From (2.39) we have the linear recursion relation

$$m^N(v) = Q_v^{N,2} X_1^2(v)$$

$$\leq \frac{2\theta}{q-(q-2)\theta}\bar{c}(\beta,q)\sum_{\omega: v\to\omega}\int \mathbb{P}(d\xi) Rg\Big(\mathbb{P}_\omega^{N,\xi}(\eta(w)=1),\dots,\mathbb{P}_w^{N,\xi}(\eta(\omega)=q)\Big)$$

$$\leq \frac{2\theta}{q-(q-2)\theta}\bar{c}(\beta,q)\sum_{\omega: v\to w} m^N(\omega) \quad (2.42)$$

and from here the result of the proposition follows. □

2.5 The ferromagnetic ordering

Let us discuss the threshold value for the ferromagnetic ordering (where the infinite volume states with uniform boundary conditions cease to be different).

Observe that for a boundary condition ξ that is all q we have that $X_k^j(v) = 0$ for all $1 \leq i, j \leq q-1$, and further that $X_i^q(v) = X_1^q(v)$ for all $i = 1,\dots,q-1$. So the iteration runs on the one-dimensional quantity $X_1^q(v)$ and reads

$$\begin{aligned} X_1^q(v) &= \sum_{\omega: v\to w} \log\frac{q-1+\exp(2\beta)\exp[X_1^q(w)]}{q-2+\exp(2\beta)+\exp[X_1^q(w)]} \\ &=: \sum_{\omega: v\to w} \psi(X_1^q(w)). \end{aligned} \quad (2.43)$$

For a regular tree with d children we have

$$X_1^q(k) = d\psi(X_1^q(k+1)). \quad (2.44)$$

We have to distinguish now the cases of $q = 2$ and $q \geq 3$. For $q = 2$ we see by computation of the second derivative that the function ψ is concave. This means that the critical value β for which a positive solution X ceases to exist is given by $1 = d\psi'(0)$.

The derivative at $X = 0$ (which we state now for general q) reads

$$\frac{\partial}{\partial X}\psi(X)\big|_{X=0} = \frac{e^{2\beta}-1}{e^{2\beta}+q-1} = \frac{2\theta}{q-(q-2)\theta}. \quad (2.45)$$

Hence, the critical value in the Ising case is given by $d \tanh \beta = 1$, for a regular tree where every vertex has d children.

We note that this quantity equals λ_2, the second eigenvalue of the transition matrix associated to the model.

Let us now turn to the Potts model with $q \geq 3$. A computation shows that $\psi"(0) > 0$ for $\beta > 0$ and $q \geq 3$, and hence the function ψ is *not* concave. This reflects the fact that the transition at the critical point where a positive solution ceases is a first order transition, where the nonzero solution is bounded away from zero.

For a regular tree with d children we can derive the transition value $\beta(q, d)$ as follows: We must have $1 = d\psi'(X^*)$, meaning that the function ψ touches the line X with the same slope. This equation translates into $\frac{1}{d} = \frac{ax}{q-1+ax} - \frac{x}{q-2+a+x}$ in the variables $a = e^{2\beta}$, $x = \exp[X^*]$. The fixed point equation itself reads $x^{\frac{1}{d}} = \frac{q-1+ax}{q-2+a+x}$.

From these two equations the critical values can be derived numerically for any d, q. We note moreover that, for the special case of a binary tree $d = 2$, the fixed point equation is cubic in the variable $y := x^{\frac{1}{2}}$. The fixed point equation is equivalent to $y(q-2+a+y^2) - ((q-1)+ay^2) = 0$. We already know one root, it is $y = 1$, so we can produce a quadratic equation by polynomial division. Writing $y = 1+u$ we get the solutions $u = \frac{1}{2}(-3+a-\sqrt{5-2a+a^2-4q})$ and $u = \frac{1}{2}(-3+a+\sqrt{5-2a+a^2-4q})$. The solution ceases to exist when the argument of the squareroot becomes negative which results in a critical value $a = 1+2\sqrt{q-1}$, or $\beta(d=2, q) = \frac{1}{2}\log(1+2\sqrt{q-1})$. We note the numerical values $\beta(d=2, q=3) = 0.671227$, $\beta(d=2, q=4) = 0.748034$.

The same type of reasoning can be used for $d = 3$ where the fixed point equation requires the solution of a fourth order equation in $z = x^{\frac{1}{3}}$, which can be reduced to a third order equation dividing out the root $z = 1$.

Chapter 3

Purity transition on Galton-Watson trees (II): entropy is Lyapunov

3.1 Introduction

In this chapter we give an alternative proof for the purity threshold for the Potts model. Actually, here, the set up is more general, the proof turn out to be simpler and the result of the previous chapter is obtained as a special case. The method presented here permits to treat asymmetric channels too, with very good bounds for strongly asymmetric chains.
We look to the problem from the point of view of Markov chains indexed by the tree. We give a criterion of the form $\mathbb{Q}(d_0)c(M) < 1$ for the non-reconstructability of tree-indexed q-state Markov chains obtained by broadcasting a signal from the root with a given transition matrix M. Here $c(M)$ is an explicit constant defined in terms of a $q-1$-dimensional variational problem over symmetric entropies, and $\mathbb{Q}(d_0)$ is the expected number of offspring on the Galton-Watson tree.

Our theorem holds for possibly non-reversible M and its proof is based on a general "Recursion Formula" for expectations of a symmetrized relative entropy function, which invites their use as a Lyapunov function. In the case of the Potts model, the present theorem reproduces earlier results, with a simplified proof.

3.2 Purity transition for q-state Markov chain on trees

As usual, we consider an infinite rooted tree T having no leaves. For $v, \omega \in T$ we write $v \to \omega$, if ω is the child of v, and we denote by $|v|$ the distance of a vertex v to the root. We write T^N for the subtree of all vertices with distance $\leq N$ to the root.

To each vertex v there is associated a (spin-) variable $\eta(v)$ taking values in a finite space which, without loss of generality, will be denoted by $\{1,2,\dots,q\}$. Our model will be defined in terms of the stochastic matrix with non-zero entries

$$M = (M(v,\omega))_{1\leq v,\omega\leq q}\,. \tag{3.1}$$

By the Perron-Frobenius theorem there is a unique single-site measure $\alpha = (\alpha(j))_{j=1,\dots,q}$ which is invariant under the application of the transition matrix M, meaning that $\sum_{i=1}^{q}\alpha(i)M(i,j) = \alpha(j)$.

The object our study is the corresponding *tree-indexed Markov chain in equilibrium.* This is the probability distribution $\mathbb{P}$ on $\{1,\dots,q\}^T$ whose restrictions $\mathbb{P}_{T^N}$ to the state spaces of finite trees $\{1,\dots,q\}^{T^N}$ are given by

$$\mathbb{P}_{T^N}(\eta_{T^N}) = \alpha(\eta(0)) \prod_{\substack{v,\omega:\\ v\to\omega}} M(\eta(v),\eta(\omega))\,. \tag{3.2}$$

The notion equilibrium refers to the fact that all single-site marginals are given by the invariant measure α.

Now, our present aim is to provide a general criterion, depending on the model only in a local (finite-dimensional) way, which implies the extremality of $\mathbb{P}$, and which works also in regimes of non-uniqueness.

To formulate our result we need the following notation.

We write for the simplex of length-q probability vectors

$$P = \{(p(i))_{i=1,\dots,q}, p(i)\geq 0\ \forall\, i,\ \sum_{i=1}^{q} p(i) = 1\} \tag{3.3}$$

and we denote the relative entropy between probability vectors $p, \alpha \in P$ by $S(p|\alpha) = \sum_{i=1}^{q} p(i)\log\frac{p(i)}{\alpha(i)}$. We introduce the *symmetrized entropy* between p and α and write

$$L(p) = S(p|\alpha) + S(\alpha|p) = (p-\alpha)\log\frac{dp}{d\alpha}\,. \tag{3.4}$$

While the symmetrized entropy is not a metric (since the triangle inequality fails) it serves us as a “distance” to the invariant measure α.

Let us define the constant, depending solely on the transition matrix M, in terms of the following supremum over probability vectors

$$c(M) = \sup_{p \in P} \frac{L(pM^{\text{rev}})}{L(p)}, \tag{3.5}$$

where $M^{\text{rev}}(i,j) = \frac{\alpha(j)M(j,i)}{\alpha(i)}$ is the transition matrix of the reversed chain. Note that numerator and denominator vanish when we take for p the invariant distribution α. Consider a Galton-Watson tree with i.i.d. offspring distribution concentrated on $\{1,2,\dots\}$ and denote the corresponding expected number of offspring by $\mathbb{Q}(d_0)$.

Here is our main result.

Theorem 3.2.1 *If $\mathbb{Q}(d_0)c(M) < 1$ then the tree-indexed Markov chain $\mathbb{P}$ on the Galton-Watson tree T is extremal for $\mathbb{Q}$-almost every tree T. Equivalently, in information theoretic language, there is no reconstruction.*

Remark : The computation of the constant $c(M)$ for a given transition matrix M is a simple numerical task. Note that fast mixing of the Markov chain corresponds to small $c(M)$. See the next chapter for numerical estimates of $c(M)$ in the Potts model case.

3.3 Applications: two special models

Here we give two applications of the general Theorem 3.2.1, computing the constant $c(M)$ for two special models; namely the Potts model and asymmetric binary channels. For the Potts case we recover the threshold of the previous chapter, while for asymmetric binary channels we see that the entropy method permits to improve the bound appearing in [11].

3.3.1 Potts model

The Potts model with q states at inverse temperature β is defined by the transition matrix

$$M_\beta = \frac{1}{e^{2\beta}+q-1}\begin{pmatrix} e^{2\beta} & 1 & 1 & \dots & 1 \\ 1 & e^{2\beta} & 1 & \dots & 1 \\ 1 & 1 & e^{2\beta} & \dots & 1 \\ & & \dots & & \end{pmatrix}. \tag{3.6}$$

This Markov chain is reversible for the equidistribution. In the case $q = 2$, the Ising model, one computes $c(M_\beta) = (\tanh\beta)^2$ which yields the correct reconstruction threshold.

Theorem 3.2.1 is a generalization of the main result given in our paper [18] for the specific case of the Potts model. To see this connection we rewrite

$$c(M_\beta) = \frac{e^{2\beta}-1}{e^{2\beta}+q-1}\bar{c}(\beta,q) \tag{3.7}$$

and note that the main theorem of the previous chapter was formulated in terms of the quantity

$$\bar{c}(\beta,q) = \sup_{p\in P}\frac{\sum_{i=1}^{q}(qp_i-1)\log(1+(e^{2\beta}-1)p_i)}{\sum_{i=1}^{q}(qp_i-1)\log qp_i}. \tag{3.8}$$

In fact, since for the Potts model

$$M^{\mathrm{rev}}(i,j) = M_\beta(j,i) = M_\beta(i,j) \tag{3.9}$$

and

$$pM^{\mathrm{rev}}(i) = \frac{\left(e^{2\beta}-1\right)p(i)+1}{e^{2\beta}+q-1} \tag{3.10}$$

we have

$$\begin{aligned}
L\left(pM^{\mathrm{rev}}\right) &= \sum_{i=1}^{q}\left(pM^{\mathrm{rev}}(i)-\alpha(i)\right)\log qpM^{\mathrm{rev}}(i)\\
&= \sum_{i=1}^{q}\left(\frac{(e^{2\beta}-1)p(i)+1}{e^{2\beta}+q-1}-\frac{1}{q}\right)\log q\frac{(e^{2\beta}-1)p(i)+1}{e^{2\beta}+q-1}\\
&= \frac{e^{2\beta}-1}{e^{2\beta}+q-1}\sum_{i=1}^{q}\left(p(i)-\frac{1}{q}\right)\log(1+(e^{2\beta}-1)p(i)).
\end{aligned} \tag{3.11}$$

Now, dividing by the entropy $L(p)$ we recover the constant $\bar{c}(\beta,q)$.

3.3.2 Asymmetric binary channels

Consider the following transition matrix for a Markov chain on a tree:

$$M = \begin{pmatrix} 1-\delta_1 & \delta_1 \\ 1-\delta_2 & \delta_2 \end{pmatrix} \quad \text{with} \quad d_1, \delta_2 \in (0,1). \tag{3.12}$$

The chain is not symmetric when $1-\delta_1 \neq \delta_2$.
We write

$$\pi_v^N = \pi_v^{N,\xi} = \left(\mathbb{P}^{N,\xi}(\eta(v) = s)\right)_{s=1,\dots,q}. \tag{3.13}$$

Call $(\alpha(+), \alpha(-))$ the invariant distribution then to prove non reconstruction one has to show that:

$$\lim_{N\uparrow\infty} \mathbb{P}\left(\xi : \left|\pi^{N,\xi}(s) - \alpha(s)\right| \geq \varepsilon\right) \to 0, \tag{3.14}$$

for $s = +, -$, for all $\varepsilon > 0$. This is the same of

$$\lim_{N\to\infty} \mathbb{E}\left(\left|\alpha(+)\pi^{N,\xi}(-) - \alpha(-)\pi^{N,\xi}(+)\right|\right) = 0, \tag{3.15}$$

that is equivalent to non solvability. Infact

Lemma 3.3.1

$$\lim_{N\to\infty} \frac{1}{2} \sum_{\xi\in\eta_{\partial T^N}} \left|\mathbb{P}(\partial T^N = \xi|\eta(0) = 1) - \mathbb{P}(\partial T^N = \xi|\eta(0) = -1)\right| = 0 \tag{3.16}$$

if and only if

$$\lim_{N\to\infty} \mathbb{E}\left(\left|\alpha(+)\pi^{N,\xi}(-) - \alpha(-)\pi^{N,\xi}(+)\right|\right) = 0. \tag{3.17}$$

Proof: Notice that

$$\mathbb{P}(\partial T^N = \xi|\eta(0) = \cdot) = \frac{\pi^{N,\xi}(\cdot)}{\alpha(\cdot)} \mathbb{P}(\xi \in \partial T^N), \tag{3.18}$$

and substitute in the formula for the total variation distance to obtain:

$$\frac{1}{2} \sum_{\xi\in\eta_{\partial T^N}} \left|\mathbb{P}(\partial T^N = \xi|\eta(0) = 1) - \mathbb{P}(\partial T^N = \xi|\eta(0) = -1)\right|$$

$$= \frac{1}{2\alpha(+)\alpha(-)} \mathbb{E}\left(\left|\alpha(+)\pi^{N,\xi}(-) - \alpha(-)\pi^{N,\xi}(+)\right|\right). \tag{3.19}$$

□

Let us focus on regular trees. Mossel and Peres in [11] prove the following:

Theorem 3.3.2 *On a regular tree of degree d the Reconstruction Problem defined by the matrix M* (6.33) *is unsolvable when*

$$d\frac{(\delta_2-\delta_1)^2}{\min\{\delta_1+\delta_2, 2-\delta_1-\delta_2\}} \le 1. \tag{3.20}$$

It is known that there is reconstruction when $d(\delta_2-\delta_1)^2 > 1$, that, being $\lambda_2 = \delta_2-\delta_1$ is the Kesten–Stigum bound . When $\delta_1+\delta_2 = 1$, the matrix M is symmetric and the Kesten–Stigum bound is sharp. Recently, Borgs, Chayes, Mossel and Roch in [8], have shown with an elegant proof that the Kesten–Stigum threshold is tight for roughly symmetric binary channels; i.e. when $|1-(\delta_1+\delta_2)| < \delta$, for some δ small. Even if the threshold we give is very near to Kesten–Stigum bound when the chain has a small asymmetry, by now, we are not able to recover this sharp estimate with our method. However, the entropy method of Theorem 3.2.1 improve (3.20) for the values of δ_1 and δ_2 giving a strongly asymmetric chain.

A computation gives:

$$\alpha(+) = \frac{1-\delta_2}{1-(\delta_2-\delta_1)}, \ \ \alpha(-) = \frac{\delta_1}{1-(\delta_2-\delta_1)}, \tag{3.21}$$

and

$$L(p) = \left(p - \frac{1-\delta_2}{1-(\delta_2-\delta_1)}\right)\log\left(\frac{p}{1-p}\frac{\delta_1}{1-\delta_2}\right), \tag{3.22}$$

$$L(pM^{\text{rev}}) = (\delta_2-\delta_1)\left(p - \frac{1-\delta_2}{1-(\delta_2-\delta_1)}\right)\log\left(\frac{(1-\delta_2)+p(\delta_2-\delta_1)}{\delta_2-p(\delta_2-\delta_1)}\frac{\delta_1}{1-\delta_2}\right). \tag{3.23}$$

Thus:

$$c(M) = \sup_p \frac{(\delta_2-\delta_1)\log\left(\frac{(1-\delta_2)+p(\delta_2-\delta_1)}{\delta_2-p(\delta_2-\delta_1)}\frac{\delta_1}{1-\delta_2}\right)}{\log\left(\frac{p}{1-p}\frac{\delta_1}{1-\delta_2}\right)}. \tag{3.24}$$

It is quite simple to compute numerically the constant $c(M)$; the numerical outputs and the comparisons with (3.20) and the Kesten-Stigum bound are in tables 3.1 and 3.2. For the couples of values of $(\delta_1.\delta_2)$ we checked the Kesten-Stigum upper bound on the non-reconstruction thresholds for asymmetric chains are very near to ours. More than those coming from [11] when they are not equal.

$\delta_1 = 0.3$	KS Kesten-Stigum	FK Formentin-Külske	MP Mossel-Peres
$\delta_2 = 0.1$	0.04	0.0579	0.1
$\delta_2 = 0.2$	0.01	0.0125	0.02
$\delta_2 = 0.4$	0.01	0.0107	0.143
$\delta_2 = 0.5$	0.04	0.0413	0.05
$\delta_2 = 0.6$	0.09	0.0907	0.1
$\delta_2 = 0.7$	0.16	0.16	0.16
$\delta_2 = 0.8$	0.25	0.2525	0.28
$\delta_2 = 0.9$	0.36	0.3787	0.45

Table 3.1: For $\delta_1 = 0.3$, the Kesten-Stigum upper bound on the non-reconstruction thresholds for asymmetric chains are very near to ours. More than those coming from [11] when they are not equal.

3.4 Proof of Theorem 3.2.1: entropy is Lyapunov

We denote by T^N the tree rooted at 0 of depth N. The notation T^N_v indicates the sub-tree of T^N rooted at v obtained from "looking to the outside" on the tree T^N. We denote by $\mathbb{P}^N_v$ the measure on T^N_v with free boundary conditions, or, equivalently the Markov chain obtained from broadcasting on the subtree with the root v with the same transition kernel, starting in α. We denote by $\mathbb{P}^{N,\xi}_v$ the correponding measure on T^N_v with boundary condition on ∂T^N_v given by $\xi = (\xi_i)_{i\in\partial T^N_v}$. Obviously it is obtained by conditioning the free boundary condition measure $\mathbb{P}^{N,\xi}_v$ to take the value ξ on the boundary.

To control a recursion for these quantities along the tree we find it useful to make explicit the following notion.

Definition 3.4.1 *We call a real-valued function $\mathscr{L}$ on P a linear stochastic Lyapunov function with center p^* if there is a constant c such that*

- $\mathscr{L}(p) \geq 0\ \forall p \in P$ *with equality if and only if* $p = p^*$*;*
- $\mathbb{E}\mathscr{L}(\pi^N_v) \leq c\sum_{\omega: v\mapsto\omega} \mathbb{E}\mathscr{L}(\pi^N_\omega)$.

$\delta_1 = 0.7$	KS Kesten-Stigum	FK Formentin-Külske	MP Mossel-Peres
$\delta_2 = 0.1$	0.36	0.3787	0.45
$\delta_2 = 0.2$	0.25	0.2525	0.28
$\delta_2 = 0.3$	0.16	0.16	0.16
$\delta_2 = 0.4$	0.09	0.0907	0.1
$\delta_2 = 0.5$	0.04	0.0413	0.05
$\delta_2 = 0.6$	0.01	0.0107	0.0143
$\delta_2 = 0.8$	0.01	0.0125	0.02
$\delta_2 = 0.9$	0.04	0.0579	0.1

Table 3.2: For $\delta_1 = 0.7$, the Kesten-Stigum upper bound on the non-reconstruction thresholds for asymmetric chains together with ours and those coming from [11].

Proposition 3.4.2 *Consider a tree-indexed Markov chain $\mathbb{P}$, with transition kernel $M(i,j)$ and invariant measure $\alpha(i)$.*

Then the function

$$L(p) = S(p|\alpha) + S(\alpha|p) = (p-\alpha)\log\frac{dp}{d\alpha} \tag{3.25}$$

is a linear stochastic Lyapunov function with center α w.r.t. the measure $\mathbb{P}$ for the constant (3.5).

Proposition 3.4.2 immediately follows from the following invariance property of the recursion which is the main result of our paper.

Proposition 3.4.3 *Recursion Formula for expected symmetrized entropy.*

$$\int \mathbb{P}(d\xi)L(\pi_v^{N,\xi}) = \sum_{\omega: v\to\omega}\int \mathbb{P}(d\xi)L(\pi_\omega^{N,\xi}M^{rev}). \tag{3.26}$$

Warning: Pointwise, that is for fixed boundary condition, things fail and one has

$$L(\pi_v^{N,\xi}) \neq \sum_{\omega: v\to\omega} L(\pi_\omega^{N,\xi}M^{\mathrm{rev}}) \tag{3.27}$$

in general. In this sense the proposition should be seen as an invariance property which limits the possible behavior of the recursion.

Proof of Proposition 3.4.3. We need the *measure on boundary configurations* at distance N from the root on the tree emerging from v which is obtained by conditioning the spin in the site v to take the value to be j, namely

$$Q_v^{N,j}(\xi) := \mathbb{P}_v^N(\eta : \eta_{|\partial T_v^N} = \xi | \eta(v) = j). \tag{3.28}$$

Then the double expected value w.r.t. to the a priori measure α between boundary relative entropies can be written as an expected value w.r.t. $\mathbb{P}$ over boundary conditions w.r.t. to the open b.c. measure of the symmetrized entropy between the distributions at v and α in the following form.

Lemma 3.4.4

$$\int \mathbb{P}(d\xi) \underbrace{L(\pi_v^{N,\xi})}_{\textit{symmetric entropy at } v} = \int \alpha(dx_1) \int \alpha(dx_2) \underbrace{S(Q_v^{N,x_2} | Q_v^{N,x_1})}_{\textit{boundary entropy}}. \tag{3.29}$$

Proof of Lemma 3.4.4: In the first step we express the relative entropy as an expected value

$$S(Q_v^{N,x_2} | Q_v^{N,x_1}) = \int \mathbb{P}(d\xi) \frac{d\pi_v^N}{d\alpha}(x_2) \Big(\log \frac{d\pi_v^N}{d\alpha}(x_2) - \log \frac{d\pi_v^N}{d\alpha}(x_1) \Big). \tag{3.30}$$

Here we have used that, with obvious notations,

$$\frac{dQ_v^{N,x_2}}{d\mathbb{P}_v^N}(\xi) = \frac{\mathbb{P}_v(\eta(v) = x_2, \xi)}{\mathbb{P}_v(\eta(v) = x_2)\mathbb{P}_v(\xi)} = \frac{d\pi_v^N}{d\alpha}(x_2). \tag{3.31}$$

Further we have used that

$$\log \frac{dQ_v^{N,x_2}}{dQ_v^{N,x_1}} = \log \frac{d\pi_v^N}{d\alpha}(x_2) - \log \frac{d\pi_v^N}{d\alpha}(x_1), \tag{3.32}$$

for $x_1, x_2 \in \{1, \dots, q\}$. This gives

$$\int \alpha(dx_1) \int \alpha(dx_2) S(Q_v^{N,x_2} | Q_v^{N,x_1})$$

$$= \int \mathbb{P}(d\xi) \int \alpha(dx_2) \frac{d\pi_v^N}{d\alpha}(x_2) \log \frac{d\pi_v^N}{d\alpha}(x_2)$$
$$- \int \mathbb{P}(d\xi) \int \alpha(dx_1) \underbrace{\int \alpha(dx_2) \frac{d\pi_v^N}{d\alpha}(x_2)}_{1} \log \frac{d\pi_v^N}{d\alpha}(x_1)$$

$$= \int \mathbb{P}(d\xi) S(\pi_v^{N,\xi} | \alpha) + \int \mathbb{P}(d\xi) S(\alpha | \pi_v^{N,\xi}) \quad (3.33)$$

and finishes the proof of Lemma 3.4.4. □

Let us continue with the proof of the Recursion Formula. We need two more ingredients formulated in the next two lemmas. The first gives the recursion of the probability vectors π_v^N in terms of the values π_ω^N of their children w, which is valid for any fixed choice of the boundary condition ξ.

Lemma 3.4.5 *Deterministic recursion.*

$$\pi_v^N(j) = \frac{\alpha(j) \prod_{w: v \to \omega} \sum_i \frac{M(j,i)}{\alpha(i)} \pi_\omega^N(i)}{\sum_k \alpha(k) \prod_{\omega: v \to \omega} \sum_i \frac{M(k,i)}{\alpha(i)} \pi_\omega^N(i)}, \quad (3.34)$$

or, equivalently: for all pairs of values j,k we have

$$\log \frac{d\pi_v^N}{d\alpha}(j) - \log \frac{d\pi_v^N}{d\alpha}(k) = \sum_{\omega: v \to \omega} \log \frac{\sum_i \frac{M(j,i)}{\alpha(i)} \pi_\omega^N(i)}{\sum_i \frac{M(k,i)}{\alpha(i)} \pi_\omega^N(i)}. \quad (3.35)$$

Proof: We have:

$$\mathbb{P}_v^N\big(\eta(v) = j, \partial T_v^N = \xi\big) = \sum_{\eta_v: \eta(v) = j} \alpha(j) \prod_{x \geq v: x \to y} M\big(\eta(x), \eta(y)\big), \quad (3.36)$$

where it is understood that $\eta(v) = \xi(v)$ if $v \in \partial T^N$. Thus, with the same notation:

$$\mathbb{P}_v^{N,\xi}\big(\eta(v) = j\big) = Z_v^{-1} \alpha(j) \sum_{\eta_v: \eta(v) = j} \prod_{x \geq v: x \to y} M\big(\eta(x), \eta(y)\big), \quad (3.37)$$

with

$$Z_v^{-1} = \sum_{k=1}^{q} \alpha(k)\mathbb{P}_v^N\left(\eta(v) = k, \partial T_v^N = \xi\right). \tag{3.38}$$

We want to rewrite π_v^N as a function of π_ω^N where ω are the children of v. The key observation is that:

$$\mathbb{P}_v^N\left(\eta(v) = j, \partial T_v^N = \xi\right) = \alpha(j) \prod_{\omega: v\to\omega} \sum_{i=1}^{q} \frac{M(j,i)}{\alpha(i)} \mathbb{P}_\omega^N\left(\eta(\omega) = i, \partial T_\omega^N = \xi\right). \tag{3.39}$$

Once you have this the proof is simple. In fact:

$$\begin{aligned}\pi_v^N(j) &= Z_v^{-1}\mathbb{P}_v^N\left(\eta(v) = j, \partial T_v^N = \xi\right) \\ &= Z_v^{-1}\alpha(j) \prod_{\omega: v\to\omega} Z_\omega \sum_{i=1}^{q} \frac{M(j,i)}{\alpha(i)} \underbrace{Z_\omega^{-1}\mathbb{P}_\omega^N\left(\eta(\omega) = i, \partial T_\omega^N = \xi\right)}_{=\pi_\omega^N(i)},\end{aligned} \tag{3.40}$$

and

$$Z_v^{-1} = \sum_{k=1}^{q} \alpha(k) \prod_{\omega: v\to\omega} Z_\omega \sum_{i=1}^{q} \frac{M(k,i)}{\alpha(i)} \pi_\omega^N(i). \tag{3.41}$$

To derive (3.39) write:

$$\mathbb{P}_v^N\big(\eta(v)=j,\partial T^N=\xi\big)=\sum_{\eta_v:\eta(v)=j}\alpha(j)\prod_{x\geq v:x\to y}M\big(\eta(x),\eta(y)\big)$$

$$=\alpha(j)\sum_{\eta_v:\eta(v)=j}\prod_{\omega:v\to\omega}\underbrace{M\big(j,\eta(\omega)\big)\prod_{x\geq\omega:x\to y}M\big(\eta(x),\eta(y)\big)}_{:=f(\eta_\omega)}$$

$$=\alpha(j)\sum_{\eta_{\omega_1},\dots,\omega_{d_v}}f(\eta_{\omega_1})\times\ldots\times f(\eta_{\omega_{d_v}})$$

$$=\alpha(j)\left(\sum_{\eta_{\omega_1}}f(\eta_{\omega_1})\right)\times\ldots\times\left(\sum_{\eta_{\omega_{d_v}}}f(\eta_{\omega_{d_v}})\right)$$

$$=\alpha(j)\prod_{\omega:v\to\omega}\left(\sum_{\eta_\omega}f(\eta_\omega)\right)=\alpha(j)\prod_{\omega:v\to\omega}\sum_{i=1}^{q}\sum_{\eta_\omega:\eta(\omega)=i}f(\eta_\omega)$$

$$=\alpha(j)\prod_{\omega:v\to\omega}\sum_{i=1}^{q}\frac{M(j,i)}{\alpha(i)}\underbrace{\sum_{\eta_\omega:\eta(\omega)=i}\alpha(i)\prod_{x\geq\omega:x\to y}M\big(\eta(x),\eta(y)\big)}_{=\mathbb{P}_\omega^N\left(\eta(\omega)=i,\partial T_v^N=\xi\right)}. \tag{3.42}$$

□

We also need to take into account the *forward propagation* of the distribution of boundary conditions from the parents to the children, formulated in the next lemma.

Lemma 3.4.6 *Propagation of the boundary measure.*

$$Q_v^{N,j}=\prod_{\omega:v\to\omega}\sum_i M(j,i)Q_\omega^{N,i}. \tag{3.43}$$

Proof: This statement follows from the previous lemma. By definition:

$$Q_v^{N,j}(\xi) = \frac{\mathbb{P}_v^N\left(\eta(v) = j, \partial T_v^N = \xi\right)}{\mathbb{P}_v^N\left(\eta(v) = j\right)}$$

$$= \frac{\alpha(j)\prod_{\omega:v\to\omega}\sum_{i=1}^q \frac{M(j,i)}{\alpha(i)}\mathbb{P}_\omega^N\left(\eta(\omega) = i, \partial T_\omega^N = \xi\right)}{\alpha(j)}$$

$$= \prod_{\omega:v\to\omega}\sum_{i=1}^q M(j,i)Q_\omega^{N,i}. \quad (3.44)$$

□

Now we are ready to head for the Recursion Formula.

We use (3.32) and the second form of the statement of the deterministic recursion Lemma 3.4.5 to write the boundary entropy in the form

$$S(Q_v^{N,j}|Q_v^{N,k}) = Q_v^{N,j}\sum_{w:v\to w}\log\frac{\sum_i \frac{M(j,i)}{\alpha(i)}\pi_w^N(i)}{\sum_i \frac{M(k,i)}{\alpha(i)}\pi_w^N(i)}. \quad (3.45)$$

Next, substituting the Propagation-of-the-boundary-measure-Lemma 3.4.6 and

(3.31) we write

$$S(Q_v^{N,j}|Q_v^{N,k}) = Q_v^{N,j} \sum_{\omega:v\to\omega} \log \frac{\sum_i \frac{M(j,i)}{\alpha(i)}\pi_\omega^N(i)}{\sum_i \frac{M(k,i)}{\alpha(i)}\pi_\omega^N(i)}$$

$$= \sum_{\omega:v\to\omega}\sum_l M(j,l)Q_\omega^{N,l} \log \frac{\sum_i \frac{M(j,i)}{\alpha(i)}\pi_\omega^N(i)}{\sum_i \frac{M(k,i)}{\alpha(i)}\pi_\omega^N(i)}$$

$$= \sum_{\omega:v\to\omega}\int d\mathbb{P}(\xi)\sum_l M(j,l)\frac{\pi_\omega^N(l)}{\alpha(l)} \log \frac{\sum_i \frac{M(j,i)}{\alpha(i)}\pi_\omega^N(i)}{\underbrace{\sum_i \frac{M(k,i)}{\alpha(i)}\pi_\omega^N(i)}_{\frac{\pi_\omega^N M^{\text{rev}}(k)}{\alpha(k)}}}$$

$$= \sum_{\omega:v\to\omega}\int d\mathbb{P}(\xi)\frac{\pi_\omega^N M^{\text{rev}}(j)}{\alpha(j)} \log \frac{\frac{\pi_\omega^N M^{\text{rev}}(j)}{\alpha(j)}}{\frac{\pi_\omega^N M^{\text{rev}}(k)}{\alpha(k)}}, \quad (3.46)$$

using in the last step the definition of the reversed Markov chain. Finally applying the sum $\sum_{j,k}\alpha(j)\alpha(k)\cdots$ to both sides of (3.46) we get the Recursion Formula. To see this, note that the l.h.s. of (3.46) together with this sum becomes the r.h.s. of the equation in Lemma 3.4.4. For the r.h.s. of (3.46) we note that

$$\sum_{j,k}\alpha(j)\alpha(k)\frac{\pi_\omega^N M^{\text{rev}}(j)}{\alpha(j)} \log \frac{\frac{\pi_\omega^N M^{\text{rev}}(j)}{\alpha(j)}}{\frac{\pi_\omega^N M^{\text{rev}}(k)}{\alpha(k)}} = L(\pi_\omega^N M^{\text{rev}}). \quad (3.47)$$

This finishes the proof of the Recursion Formula Proposition 3.4.3. □

Finally, Theorem 3.2.1 follows from Proposition 3.4.2 with the aid of the Wald equality with respect to the expectation over Galton-Watson trees since the contraction of the recursion and the Lyapunov function properties yield

$$\lim_{N\uparrow\infty}\mathbb{P}\Big(\xi : \Big|\pi^{N,\xi}(s) - \alpha(s)\Big| \geq \varepsilon\Big) \to 0, \quad (3.48)$$

for all s, for all $\varepsilon > 0$, and this implies the extremality of the measure $\mathbb{P}$. This ends the proof of Theorem 3.2.1. □

Chapter 4

Conclusions

4.1 Introduction

In this last chapter we comment on the result. We give some numerical estimates for our bound along with some conjectures. A comparison with other rigorous bound [11, 21], but also with thresholds coming from algorithms [15, 22] is made for different values of q. For example, for $q = 3$ and d small this comparison shows our bound is very good: the best as of today [22].

4.2 Conjectures and comparisons

In this part of the thesis we have proven that the Free Gibbs Measure on a Galtson-Watson tree is pure when

$$\mathbb{Q}(d_0)\frac{2\theta}{q-(q-2)\theta}\bar{c}(\beta,q) < 1 \tag{4.1}$$

where

$$\bar{c}(\beta,q) := \sup_{p\in P}\frac{\sum_{i=1}^{q}(qp_i-1)\log(1+(e^{2\beta}-1)p_i)}{\sum_{i=1}^{q}(qp_i-1)\log qp_i}. \tag{4.2}$$

Let us comment on the constant appearing, and provide the following conjecture. Define

$$\hat{c}(\beta,q) := \sup_{p\in P, p_2=\cdots=p_q}\frac{H(p_1,\ldots,p_q)}{Rg(p_1,\ldots,p_q)}. \tag{4.3}$$

Conjecture 4.2.1 *We believe that* $\hat{c}(\beta, q) = \bar{c}(\beta, q)$.

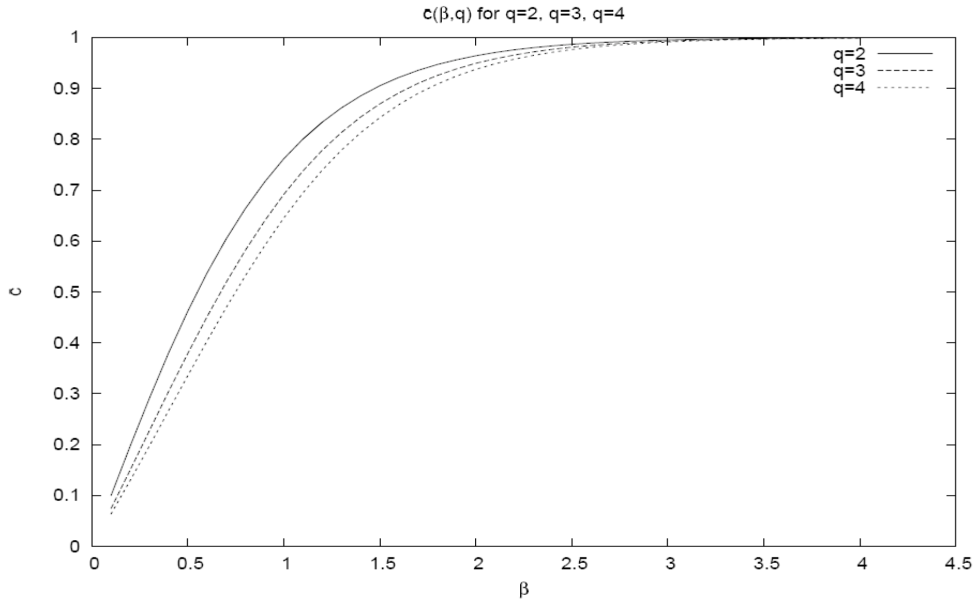


Figure 4.1: Numerical outputs of $\bar{c}$ for different values of q.

We checked this numerically for small values of q. If the previous conjecture is true, the two properties of $\bar{c}(\beta, q)$, namely, monotonicity in q and the bound θ (see figure 4.1) carry over. These two properties are seen as follows.

Lemma 4.2.2

$$\bar{c}(\beta, q) = \sup_{x \in D_q} \bar{\varphi}(q, l_q)(x), \tag{4.4}$$

with the function

$$\bar{\varphi}(q, l_q)(x) = \frac{\log\left(\frac{1+l_q x}{1-l_q(q-1)x}\right)}{\log\left(\frac{1+qx}{1-q(q-1)x}\right)}, \tag{4.5}$$

with parameter $\lambda_q = \frac{e^{2\beta}-1}{1+\frac{1}{q}(e^{2\beta}-1)}$ *on the range* $D_q = \left[-\frac{1}{q}, \frac{1}{q(q-1)}\right]$ *with* $D_{(q-1)} \supset D_q$.

Remark: Notice that $\frac{l_q}{q} = \lambda_2$.

Proof: Change to new coordinates on the simplex of probability vectors $(p_1,\dots,p_q)$ given by

$$x_i = p_i - \frac{1}{q} \text{ for } i = 1,\dots,q-1, \tag{4.6}$$

take $x = x_i$ for $i = 1,\dots,q-1$ and use Conjecture 4.2.1 □

Lemma 4.2.3 *For all $q \geq 3$ we have that*

$$\hat{c}(\beta,q) < \hat{c}(\beta,q-1) \leq \frac{l_2}{2} = \theta. \tag{4.7}$$

Proof: We use that

$$\frac{\partial\bar{\varphi}(q,l_q)(x)}{\partial q} < 0, \tag{4.8}$$

for $x \in D_q$. This gives

$$\bar{\varphi}(q,l_q)(x) < \bar{\varphi}(q-1,l_{q-1})(x) < \bar{\varphi}(2,l_2)(x), \quad x \in D_q. \tag{4.9}$$

□

Remark: Conjecture 4.2.1 (if true) makes it very simple to compute $\hat{c}(\beta,q)$ numerically, for every q. The problem of finding the sup would remain one dimensional even when q grows.

Next, what about the sharpness of the constant? Could it be possible that Theorem 2.3.1 in fact holds with the sharp value $\frac{e^{2\beta}-1}{q-1+e^{2\beta}}$ replacing the constant $\bar{c}(\beta,q)$? In our approach such a conjecture would be based on looking at the Hessian of the function

$$\partial_{x_i,x_j}\varphi_{l_q,q}(x_1,\dots,x_{q-1})\big|_{x_k=0\,\forall k} = 4\lambda_2 1_{i=j} + 2\lambda_2 1_{i\neq j}. \tag{4.10}$$

Indeed, heuristically it should suffice to look at the quadratic approximation around the equidistribution. This results in the rigorous lower bound $\bar{c}(\beta,q) \geq \frac{l_q}{q} = \frac{e^{2\beta}-1}{q-1+e^{2\beta}} = \lambda_2$ which we recognize as the Kesten-Stigum bound. For the Ising model we have equality, which is not true for $q = 3$.

Let us compare with the recent literature. In their paper [15] Montanari and Mezard make the conjecture that the Kesten-Stigum bound is sharp for $q \leq 4$, or more precisely:

Conjecture 4.2.4 *(Mézard and Montanari 2006) Consider the Potts model with q symbols on a d-ary tree and let $\lambda_2 = \frac{e^{2\beta}-1}{e^{2\beta}+q-1} = \frac{2\theta}{q-(q-2)\theta}$, with $\theta = \tanh(\beta)$, then if $q \leq 4$ and $d < d_{max}$, there is reconstruction if and only if $d\lambda_2^2 > 1$.*

This conjecture is based on extensive numerical simulations of the random recursion. Moreover, the restriction on d comes from the limitation on the values of d they can treat numerically and they actually think that $d_{max} = +\infty$.
Let us compare our bound with this conjecture. First, how close are the Kesten-Stigum bounds and our constants? We obtain numerically $\bar{c}(\beta, q) = \frac{e^{2\beta}-1}{q-1+e^{2\beta}}(1+\varepsilon(q))$ with $\varepsilon(3) = 0.0150$ and $\varepsilon(4) = 0.0365$. If we specialize to a binary tree, and take advantage of the possible temperature dependence of ε we obtain $\beta_c := \sup\{\beta, 2\frac{2\theta}{3-\theta}\bar{c}(\beta,3) < 1\} = 1.0434$ for $q = 3$ and $\beta_c := \sup\{\beta, 2\frac{2\theta}{4-2\theta}\bar{c}(\beta,4) < 1\} = 1.1555$ in the case $q = 4$.

After completion of the first draft of our work [18] Sly's preprint [21] appeared where he proves the following.

Theorem 4.2.5 *(Sly 2008) When $q \leq 3$, and $d > d_{min}$, then Kesten-Stigum bound is sharp, while the Kesten-Stigum bound is never sharp when $q \geq 5$.*

His method uses large degrees to justify quadratic expansions by means of Central Limit Theorem (CLT) approximation and makes no statements for small degrees where our estimates apply. So, he proves conjecture 6.33 partially; the case $q = 4$ is critical along his line of reasoning and the condition $d > d_{min}$ is needed to prove a concentration result via CLT.

Just the line of the Sly's proof. He considers the iteration of the following quantity

$$x_N := \mathbb{E}\left(X^+(N) - \frac{1}{q}\right) \tag{4.11}$$

where in our notation

$$X^+(N) := \mathbb{P}\left(\eta(0) = 1 | \partial T^N = \eta^1(N)\right). \tag{4.12}$$

This is a random variable with respect to the boundary configuration conditioned at the origin of the tree: the notation $\eta^1(N)$ means a random configuration on ∂T^N

chosen among the configuration on the tree with $\eta(0) = 1$.
Then, he proves that $x_N \geq 0$ and that the condition

$$\lim_{N\to+\infty} x_N = 0, \tag{4.13}$$

is equivalent to non-reconstruction.
For x_N small he has the following recursion (formula (1.1) on pag.6 of [21]):

$$x_{N+1} = k\lambda_2^2 x_N + (1+o(1))\frac{k(k-1)}{2}\frac{q(q-4)}{q-1}\lambda_2^4 x_N^2. \tag{4.14}$$

From this equation one can see why $q = 4$ is a critical case. When $q \leq 3$, the second term in (4.14) is negative and if, moreover, x_N is sufficiently small when $d\lambda^2 \leq 1$, then $\lim_{N\to\infty} x_N = 0$. The CLT is necessary to show that this concentration holds; the limitation on the number of children ($d > d_{min}$) he can treat comes by this reason (see the explanation subsequent to formula (1.1) on pag.6 of [21]).
So, his proof has two steps: first he shows that the expected distance form the unconditional distribution is below some small constant, and then, that below this small constant that distance goes to zero when N grows. The first step makes the analysis possible only the case of large degree trees.

Thus, for $q \leq 4$ and d small, the problem of finding a sharp bound is still open. Moreover no sharp bound is known when $q \geq 5$.

In view of these considerations and results it is natural for us to conjecture that the Kesten-Stigum bound holds $\mathbb{Q}$-a.s. for small q, or more precisely:

Conjecture 4.2.6 *For $q \leq 4$ there is $\mathbb{Q}$ a.s. reconstruction if and only if*

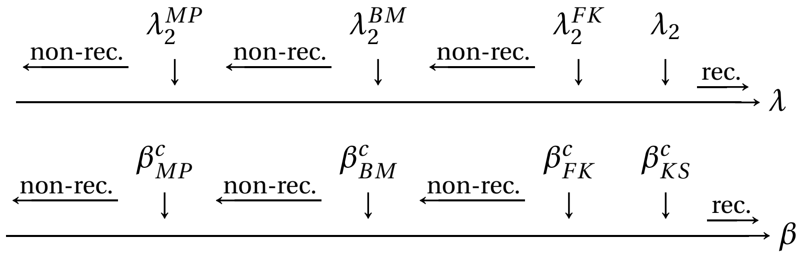
$$\mathbb{Q}(d_0)\left(\frac{2\theta}{q-(q-2)\theta}\right)^2 > 1. \tag{4.15}$$

Let us compare our threshold with others bounds in literature for $q = 3$ and d small. The Kesten-Stigmun bound says that for a tree of degree d, when $d\lambda_2^2 > 1$ reconstruction holds. In table 4.1 we report some of them to compare. As usual, d represent the number of children at each vertex of a regular tree, and $\mathbb{Q}(d)$ its mean when we are considering a random tree.

In the following pictures we illustrate the situation arising from table 4.1:

q=3	KS Kesten-Stigum	MP Mossel-Peres	FK Formentin-Külske	BM Bhatnagar-Maneva
$d=2$	0.7071...	$\frac{2}{3}=0.\bar{6}$	0.7018...	0.69
$d=3$	0.5773...	0.5302...	0.5731...	0.555
$\mathbb{Q}(d)=2.5$	0.6324...	0.5873...	0.6278...	0.61

Table 4.1: The Kesten-Stigum upper bound on the non-reconstruction threshold and the values of λ_2 up to which non-reconstruction has be shown for $q=3$ in the ferromagnetic regime. The table is taken from [22].



What can we say about large values of $q \geq 5$? Montanari and Mezard [15] find in all the finitely many test-cases of $q \geq 5$ and d which they treat by simulations that the Kesten-Stigum bound is *not* sharp. Let us therefore conclude by making a comparison of our and their values in Tables 4.2 and 4.3 in the case $q=5$, showing closeness of our bounds with the simulation values also here.

q=5	ϵ_r	$\beta_r = -0.5\log\left(\frac{\epsilon_r}{(q-1)(1-\epsilon_r)}\right)$	$\lambda_r = 1-\frac{q}{q-1}\epsilon_r$
$d=2$	0.2348	1.2838	0.7065
$d=3$	0.33881	1.0285	0.5765
$d=4$	0.4008	0.8942	0.499
$d=7$	0.4986	0.6955	0.37675
$d=15$	0.5955	0.4998	0.255625

Table 4.2: Simulation results for the reconstruction thresholds by Mezard and Montanari [15]

q=5	β_c	λ_c
$d=2$	1.2425	0.6875
$d=3$	0.98535	0.5526
$d=4$	0.8520	0.473457
$d=7$	0.65465	0.35095
$d=15$	0.4640	0.2342

Table 4.3: Numerical thresholds coming from our bound of Theorem 2.3.1

Finally, as an application of our second method (entropy is Lyapunov) we derive a threshold for non reconstruction in asymmetric channels, that even not tight when the asymmetry is small [8] improves the known bounds for strongly asymmetric channels [11], see tables below. In every case we checked they very near to the Kesten–Stigum bound, that is the threshold for reconstruction.

$\delta_1 = 0.3$	KS Kesten-Stigum	FK Formentin-Külske	MP Mossel-Peres
$\delta_2 = 0.1$	0.04	0.0579	0.1
$\delta_2 = 0.2$	0.01	0.0125	0.02
$\delta_2 = 0.4$	0.01	0.0107	0.0143
$\delta_2 = 0.5$	0.04	0.0413	0.05
$\delta_2 = 0.6$	0.09	0.0907	0.1
$\delta_2 = 0.7$	0.16	0.16	0.16
$\delta_2 = 0.8$	0.25	0.2525	0.28
$\delta_2 = 0.9$	0.36	0.3787	0.45

Table 4.4: For $\delta_1 = 0.3$, the Kesten-Stigum upper bound on the non-reconstruction thresholds for asymmetric chains are very near to ours. More than those coming from [11] when they are not equal.

$\delta_1 = 0.7$	KS Kesten-Stigum	FK Formentin-Külske	MP Mossel-Peres
$\delta_2 = 0.1$	0.36	0.3787	0.45
$\delta_2 = 0.2$	0.25	0.2525	0.28
$\delta_2 = 0.3$	0.16	0.16	0.16
$\delta_2 = 0.4$	0.09	0.0907	0.1
$\delta_2 = 0.5$	0.04	0.0413	0.05
$\delta_2 = 0.6$	0.01	0.0107	0.0143
$\delta_2 = 0.8$	0.01	0.0125	0.02
$\delta_2 = 0.9$	0.04	0.0579	0.1

Table 4.5: For $\delta_1 = 0.7$, the Kesten-Stigum upper bound on the non-reconstruction thresholds for asymmetric chains together with ours and those coming from [11].

Part II

Uniform propagation of chaos and fluctuation theorems in some spin-flip models

Introduction to Part II

In this part we deal with uniform propagation of chaos in some spin-flip models.

Propagation of chaos is a very well known phenomenon in stochastic processes with mean field interaction. The aim of this part of the thesis is to provide sharp estimate on propagation of chaos for some special models. No attempt of reaching some generality will be made, although we believe our methods should apply well beyond the cases considered here.

Somewhat in contrast with our aim of dealing with special models, here we set up a general language to give an overview of the results proved in later sections. By *mean field* stochastic process we mean a family $x^{(N)} = (x^{(N)}(t))_{t\geq 0}$ with the following features:

- $x^{(N)}(t) = (x_1^{(N)}(t), x_2^{(N)}(t), \ldots, x_N^{(N)}(t))$ is a Markov process with N real-valued components;
- Consider the *empirical measure*

  $$\rho_N(t) := \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k^{(N)}(t)}, \tag{4.16}$$

  which is a random probability on $\mathbb{R}$. Then $(\rho_N(t))_{t\geq 0}$ is a measure-valued *Markov* process.

Although this is by no means a *standard* definition of mean field model, it captures the basic features of the specific models we will consider.

Suppose we are given a model as above, and assume that the law of $x^{(N)}(0)$ is a product measure with marginal λ, where λ is a fixed probability on $\mathbb{R}$. We say that *propagation of chaos* holds, if

- for every $t > 0$, $\rho_N(t)$ converges in probability, as $N \uparrow \infty$, to a deterministic measure $\rho(t)$;

- let $i_1, i_2, \ldots, i_m$ be fixed indices, and $T > 0$; then the stochastic process

$$\left(x_{i_1}^{(N)}(t), x_{i_2}^{(N)}(t), \ldots, x_{i_m}^{(N)}(t)\right)_{t \in [0,T]} \tag{4.17}$$

  converges in law, as $N \uparrow \infty$, to a process $(\xi_1(t), \ldots, \xi_m(t))_{t\in[0,T]}$ with independent and identically distributed components; moreover $\xi_1(t)$ has law $\rho(t)$.

We then say that a *fluctuation theorem* holds if, for every $h : \mathbb{R} \to \mathbb{R}$ continuous and bounded and $T > 0$, the *fluctuation process*

$$\left(\sqrt{N}\left[\int h d\rho_N(t) - \int h d\rho(t)\right]\right)_{t\in[0,T]} \tag{4.18}$$

converges weakly to a Gaussian process.

Mean field models, that are often given as dependent on various parameters, may exhibit several types of *phase transitions*. In this thesis we only consider phase transitions that can be described in terms of the large time behavior of the limiting distribution $\rho(t)$, that we have introduced above in defining propagation of chaos. We will say that the mean field model is *subcritical* if $\rho(t)$ converges, as $t \to +\infty$, to a probability $\rho(\infty)$ that does not depend on the choice of the initial distribution λ. Otherwise, we say the process is *supercritical*. The examples treated in this thesis are subcritical for certain values of the parameters, and supercritical for others, and therefore we say they exhibit phase transition.

The main object of this part of the thesis is the analysis of the time uniformity in propagation of chaos and fluctuation theorems. More precisely:

- is the convergence (in law) of $(x_{i_1}^{(N)}(t), x_{i_2}^{(N)}(t), \ldots, x_{i_m}^{(N)}(t))$ to $(\xi_1(t), \ldots, \xi_m(t))$ uniform in t?

- Is the convergence of the fluctuations $\left(\sqrt{N}\left[\int h d\rho_N(t) - \int h d\rho(t)\right]\right)$ toward their Gaussian limit uniform in t?

Although it is reasonable to expect that such uniformity holds for subcritical models, only few results in this direction can be found in the literature, concerning interacting diffusions [23, 26].

In this thesis we study time uniformity for two specific models, which exhibit phase transition. The first model is the dynamical Curie-Weiss model, than can be considered as the most basic mean field model. The second example is a model, proposed recently in [24], in the context of credit risk in finance; it describes the time evolution of finantial indicators for a network of interacting firms. The ferromagnetic-type interaction is used to describe the effects of business relationships between firms; it should be noted that, in applications to social sciences, the mean field assumptions may be reasonable in many circumstances. From the more mathematical point of view the model is not reversible, and an explicit expression for its invariant measure is not known. The following considerations further motivate the study of time uniformity in limit theorems.

- Many key financial quantities are based on normal approximations of empirical means (see e.g. [25, 24]), that are obtained as a combination of law of large numbers and central limit theorems. In real financial applications N is given, and it is of the order of few hundreds. Uniform limit theorems rule out the possibility that normal approximations deteriorate in time, becoming unreliable for large times.

- Uniform propagation of chaos and fluctuation theorem provide the law of large numbers and the central limit theorem for the empirical means with respect to the invariant measure of the system, which is not known. Thus, results about the *statics* of the system can be derived by the *dynamics*.

Although we have chosen to deal with two specific models, the method we use appear to be rather general, and should be usable for other classes of models. A substantial limitation of our results is that they are limited to the subcritical case or, in statistical mechanical terms, to the high temperature regime. In the supercritical case, when the limiting flow $\rho(t)$ has multiple limit points as $t \to +\infty$, the dynamics for large but finite N may have a metastable behavior. This means that the empirical measure $\rho_N(t)$ may fluctuate between different limit points, on a large, N-dependent, time scale. Thus, a uniform propagation of chaos as in the definition given above is not possible. Nevertheless, we believe uniform limit theorem could be state also in the supercritical case, concerning for instance the distance between ρ_N and the set of attractors for the evolution of $\rho(t)$. This interesting problem is not treated in this thesis.

The plan of this part is as follows. In the first chapter we introduce the Curie-Weiss model and prove that in the subcritical regime uniform propagation of chaos holds. In the proof are introduced the basic techniques we use also in the sequel.

The method used for the Curie-Weiss model works also for the non-reversible model proposed in [24] to model credit risk in finance. The second chapter is dedicated to this spin-flip system. After having reviewed quickly the model and its financial meaning we show that uniform propagation of chaos property holds also in this case. In Chapters 3 we prove uniform fluctuation theorem for the Curie-Weiss model.

Chapter 5

Propagation of chaos for the Curie-Weiss mean field model

5.1 Introduction

The Curie-Weiss model can be considered as the simplest mean-field model exhibiting a phase transition. In this chapter we prove that the uniform propagation of chaos property holds when the McKean-Vlasov equation for the system has an unique stable equilibrium. Remember that the Mckean-Vlasov equation describes the time evolution of the weak limit of the empirical average $\bar{\eta}$ (see below) of the spins when N grows to infinity.
The key point in the proof is that, in the subcritical regime, the L^2 distance between the empirical average and its weak limit remains uniformly small in time. The latter fact implies uniform propagation of chaos.

5.2 The Curie-Weiss model and uniform propagation of chaos

Let $\eta = (\eta_1, \dots, \eta_N)$ be the N-spin system defined to be the Markov chain with generator acting on functions $f : \{-1,1\}^N \to \mathbb{R}$, given by:

$$L_N f(\eta) = \sum_{i=1}^{N} c^{\omega}\left(i,\eta\right) \nabla_i^{\eta} f(\eta) \tag{5.1}$$

where

$$c^{\omega}\left(i,\eta\right) = \exp\left(-\beta\eta_i(\bar{\eta}+\omega)\right) \text{ with } \bar{\eta} = \frac{1}{N}\sum_{i=1}^{N}\eta_i \tag{5.2}$$

and

$$\nabla_i^{\eta} f(\eta) = f(\eta^i) - f(\eta). \tag{5.3}$$

Here $\eta^i = (\eta_1, \dots, -\eta_i, \dots, \eta_N)$ if $\eta = (\eta_1, \dots, \eta_i, \dots, \eta_N)$: η^i is the state obtained fron η by flipping the i-th spin η_i.

This N-spin system is known as the Curie-Weiss model with external magnetic field $h = \beta\omega$.

Let us introduce here some notations we use in the sequel. Sometimes we write $\eta_i(t)$ when we want to make explicit the time dependence of the spin values. Moreover, we indicate with $\mathbb{E}^t(\cdot)$ and $\mathbb{P}^t(\cdot)$ the probabilities and means computed at time t.

Consider the process $\bar{\eta} = \frac{1}{N}\sum_{i=1}^{N}\eta_i$. It is markovian with a deterministic weak limit, that we call m_t.

Proposition 5.2.1 *Assume that $\eta_i(0)$ for $i = 1, \dots, N$ is an i.i.d. sequence of random variables with $\mathbb{E}^0(\eta_i) = m(0)$. The process $m_t \in [-1,1]$ satisfying the ordinary differential equation:*

$$\frac{d}{dt} m_t = 2\sinh\left(\beta(m_t+\omega)\right) - 2m_t \cosh\left(\beta(m_t+\omega)\right) . \tag{5.4}$$

is the weak limit of $\bar{\eta}$.

From now on we refer to (5.4) as the McKean-Vlasov equation.

Proof of Proposition 5.2.1: We use the following theorem [32]:

Theorem 5.2.2 *Let $Y_N(t)$ a sequence of Markov processes with values in $\mathscr{Y}_N$ and infinitesimal generator L_N, defined on $D(L_N)$. Moreover, let L on $D(L)$ be the infinitesimal generator of another Markov process $Y(t)$ with values in $\mathscr{Y} \supseteq \mathscr{Y}_N$. Call $\mathscr{C}$ a core for L and suppose that every function in $\mathscr{C}$ is an element of $D(L_N)$ when restricted to $\mathscr{Y}_N$.*

If the condition

$$\lim_{N\to\infty} \sup_{\alpha\in\mathscr{Y}_N} |L_N f(\alpha) - Lf(\alpha)| = 0 \tag{5.5}$$

holds for every $f \in \mathscr{C}$ and $Y_N(0)$ converges in distribution to $Y(0)$, then the sequence $Y_N(t)$ converges to $Y(t)$ in distribution.

The Weak Law of Large Numbers assure that $m(0)$ is the weak limit of $\bar{\eta}$ once $\mathbb{E}^0(\eta_i) = m(0)$.
Then, to find the limit generator L, write

$$L_N f(\bar{\eta}) = \sum_{i=1}^{N} e^{-\beta\eta_i(\bar{\eta}+\omega)} \nabla_i^{\eta} f(\bar{\eta}) = \sum_{i=1}^{N} e^{-\beta\eta_i(\bar{\eta}+\omega)} \left(f\left(\bar{\eta} - \frac{2\eta_i}{N}\right) - f(\bar{\eta})\right)$$

$$= \sum_{i=1}^{N} \left[\sinh(\beta(\bar{\eta}+\omega)) - \eta_i \cosh(\beta(\bar{\eta}+\omega))\right] \left(f\left(\bar{\eta} - \frac{2\eta_i}{N}\right) - f(\bar{\eta})\right)$$

$$\begin{aligned} &= \frac{N}{2}(1+\bar{\eta}) \left[\sinh(\beta(\bar{\eta}+\omega)) - \cosh(\beta(\bar{\eta}+\omega))\right] \left(f\left(\bar{\eta} - \frac{2}{N}\right) - f(\bar{\eta})\right) \\ &\quad + \frac{N}{2}(1-\bar{\eta}) \left[\sinh(\beta(\bar{\eta}+\omega)) + \cosh(\beta(\bar{\eta}+\omega))\right] \left(f\left(\bar{\eta} + \frac{2}{N}\right) - f(\bar{\eta})\right). \end{aligned} \tag{5.6}$$

This implies that the process $m_t^N = \bar{\eta}(t)$ is Markovian with generator H_N given by

$$\begin{aligned} H_N f(m) &= \frac{N}{2}(1+m) \left[\sinh(\beta(m+\omega)) - \cosh(\beta(m+\omega))\right] \left(f\left(m - \frac{2}{N}\right) - f(m)\right) \\ &\quad + \frac{N}{2}(1-m) \left[\sinh(\beta(m+\omega)) + \cosh(\beta(m+\omega))\right] \left(f\left(m + \frac{2}{N}\right) - f(m)\right) \end{aligned} \tag{5.7}$$

Now, for $f \in \mathscr{C} = \mathscr{C}^1$-functions with compact support,

$$H_N f(m) = Hf(m) + O\left(\frac{1}{N}\right),$$

where

$$Hf(m) = \left(2\sinh\left(\beta(m+\omega)\right) - 2m\cosh\left(\beta(m+\omega)\right)\right)\frac{\partial}{\partial x} f(m)$$

and $O\left(\frac{1}{N}\right)$ goes to zero uniformly in m. Moreover H is the generator of a process with the deterministic evolution (5.4). Thus, by Theorem 5.2.2, the conclusion follows.

□

The number and the stability of the equilibrium points of (5.4) depend on the values of β and ω.

- When $\beta < 1$ it has only one linearly stable equilibrium for every values of ω.
- If $\beta > 1$ there esists a curve

$$\omega(\beta) = \sqrt{\frac{\beta-1}{\beta}} - \frac{1}{\beta}\arctan\left(\sqrt{\frac{\beta-1}{\beta}}\right) \tag{5.8}$$

  such that if $|\omega| < \omega(\beta)$ there are three fixed points: two stable and one instable. For $|\omega| = \omega(\beta)$ two of these points coincide; while, for $|\omega| > \omega(\beta)$, there is an unique linearly stable equilibrium. Remember that in dimension 1 uniqueness together with linear stability means global attractivity.

From now on we restrict ourselves to the region of parameters where (5.4) has only one stable solution. In view of the Proposition 5.2.1 it is natural to compare the Curie-Weiss model with the N-spin system $\sigma = (\sigma_1, \ldots, \sigma_N)$ with generator:

$$G_N f(\sigma) = \sum_{i=1}^{N} a^{\omega}(i,t)\nabla_i^{\sigma} f(\sigma), \tag{5.9}$$

where,

$$a^{\omega}(i,t) = \exp\left(-\beta\sigma_i(m_t+\omega)\right). \tag{5.10}$$

Remark: If $\sigma_i(0)$, $i = 1, \ldots, N$, are independent random variables, they remains independent for every $t > 0$.

Remark: If $\mathbb{E}^0(\sigma_1) = (m_0)$, then $\mathbb{E}^t(\sigma_1) = m_t$. Actually, $\mathbb{E}^t(\sigma_1)$ obeys to the following homogeneous linear differential equation:

$$\frac{d}{dt}\mathbb{E}^t[\sigma_1] = \mathbb{E}^t[G_N(\sigma_1)] = \mathbb{E}^t\left[-2\sigma_1 a^{\omega}(1,t)\right]$$

$$= \mathbb{E}^t\left[2\sinh\big(\beta(m_t+\omega)\big) - 2\sigma_1\cosh\big(\beta(m_t+\omega)\big)\right]$$

$$= 2\sinh\big(\beta(m_t+\omega)\big) - 2\mathbb{E}^t(\sigma_1)\cosh\big(\beta(m_t+\omega)\big). \quad (5.11)$$

Thus,

$$\frac{d}{dt}\mathbb{E}^t(\sigma_1) = 2\sinh\big(\beta(m_t+\omega)\big) - 2\mathbb{E}^t(\sigma_1)\cosh\big(\beta(m_t+\omega)\big). \quad (5.12)$$

If m_t is a solution of (5.4) is a solution of (5.12), too. Uniqueness implies $\mathbb{E}^t(\sigma_1) = m_t$ when $\mathbb{E}^0(\sigma_1) = m_0$.

Now we are ready for the statement of the main theorem of this chapter. Let us introduce the following notation. For any quantity $g(N) \in \mathbb{R}^+$, we write:

$$g(N) \le O\left(\frac{1}{N^{\alpha}}\right), \quad (5.13)$$

with $\alpha > 0$, if there exists a constant $C > 0$ such that $g(N) \le \frac{C}{N^{\alpha}}$.

Theorem 5.2.3 *In the region of the parameters* (β,ω) *where the McKean-Vlasov equation has an unique stable fixed point there exists a probability space where both the processes with generators* (5.1) *and* (5.9) *can be realized. Moreover, if*

$$\mathbb{P}^0\left(\eta_i \neq \sigma_i\right) \le O\left(\sqrt{\frac{1}{N}}\right), \ \ \mathbb{E}^0\left[(\bar{\eta} - m_0)^2\right] \le O\left(\sqrt{\frac{1}{N}}\right) \quad (5.14)$$

then

$$\sup_{t\in[0,\infty)} \mathbb{P}^t\left(\eta_i \neq \sigma_i\right) \le O\left(\sqrt{\frac{1}{N}}\right). \quad (5.15)$$

In order words, we show that in the region of the parameters (β,ω) where (5.4) has an unique linearly stable equilibrium it is possible to obtain for $\mathbb{P}^t(\eta_j \neq \sigma_j)$ a time uniform upper bound: i.e. the bound depends on N, but not on t. It means that the systems are close for every t. When it happens we say that there is *uniform propagation of chaos.*

Remark: Clearly condition (5.14) is satisfied if $\{\eta_i(0) : i = 1, \ldots N\}$ are i.i.d., and we set, for $i = 1, \ldots, N$, $\sigma_i(0) = \eta_i(0)$.

In order to prove the theorem, we put the two processes defined before in the same probability space via Basic Coupling (see few lines below) and then we try to estimate the quantity $\mathbb{P}^t(\eta_j \neq \sigma_j)$ using the Gronwall's Lemma that we state here for completeness.

Lemma 5.2.4 *Let $F(t)$ be a differentiable function in $[0, +\infty)$, such that:*

$$\frac{d}{dt}F(t) \leq a(t) + b(t)F(t), \tag{5.16}$$

then

$$F(t) \leq \exp\left(\int_0^t b(r)dr\right)\left[\int_0^t a(r)\exp\left(-\int_0^r b(u)du\right)dr + F(0)\right]. \tag{5.17}$$

Remark: If $a(t) = a$ and $b(t) = b$, formula (5.17) reads

$$F(t) \leq \frac{a}{b}e^{bt} - \frac{a}{b} + F(0)e^{bt}. \tag{5.18}$$

Actually, we will use this simplified form of the Gronwall's lemma.

Proof of Lemma 5.2.4: Equation (5.16) is equivalent to

$$\left(\frac{d}{dt}F(t) - b(t)F(t)\right)\exp\left(-\int_0^t b(r)dr\right) \leq a(t)\exp\left(-\int_0^t b(r)dr\right) \tag{5.19}$$

that's the same of

$$\frac{d}{dt}\left[F(t)\exp\left(-\int_0^t b(r)dr\right)\right] \leq a(t)\exp\left(-\int_0^t b(r)dr\right). \tag{5.20}$$

Integrating, the conclusion follows. □

As we said, the other technique we use is "coupling". Coupling is a powerful method to compare two stochastic processes by realizing them in the same probability space without changing their marginals. In this section we use a kind of coupling for spin systems called Basic Coupling as described in [27]. This coupling for η and σ has the aim of making them flip together with the largest possible rate, but each process has to flip at the correct rate to maintain its marginal unchanged. In [27] the Basic Coupling is unformally described in the following way:

- if $\eta_i(t) \neq \sigma_i(t)$ the processes flip the spin at i independently with their own rates;

- while, when $\eta_i(t) = \sigma_i(t)$ they flip together with the largest possible rate according with the requirement of keeping the rate of each single process unchanged.

The best way to write the generator of the coupled process is probably the following, because in this way it is clear which is the rate of each possible transition:

$$\Omega f(\eta,\sigma) = \sum_{i=1,\eta_i\neq\sigma_i}^{N} a^{\omega}(i,t)\,\nabla_i^{\sigma} f(\eta,\sigma) + \sum_{i=1,\eta_i\neq\sigma_i}^{N} c^{\omega}\left(i,\eta\right)\nabla_i^{\eta} f(\eta,\sigma) +$$

$$+\sum_{i:\sigma_i=\eta_i} \min\left\{a^{\omega}(i,t), c^{\omega}\left(i,\eta\right)\right\}\left(f\left(\sigma^i,\eta^i\right) - f\left(\sigma,\eta\right)\right)$$

$$+\sum_{i:\sigma_i=\eta_i} \left(a^{\omega}(i,t) - \min\left\{a^{\omega}(i,t), c^{\omega}\left(i,\eta\right)\right\}\right)\nabla_i^{\sigma} f(\eta,\sigma)$$

$$+\sum_{i:\sigma_i=\eta_i} \left(c^{\omega}\left(i,\eta\right) - \min\left\{a^{\omega}(i,t), c^{\omega}\left(i,\eta\right)\right\}\right)\nabla_i^{\eta} f(\eta,\sigma)$$

$$= \sum_{i=1}^{N} a^{\omega}(i,t)\,\nabla_i^{\sigma} f(\eta,\sigma) + \sum_{i=1}^{N} c^{\omega}\left(i,\eta\right)\nabla_i^{\eta} f(\eta,\sigma)$$

$$+\sum_{i:\sigma_i=\eta_i} \min\left\{a^{\omega}(i,t), c^{\omega}\left(i,\eta\right)\right\}\left(f\left(\sigma^i,\eta^i\right) - f\left(\sigma^i,\eta\right) - f\left(\sigma,\eta^i\right) + f\left(\sigma,\eta\right)\right). \quad (5.21)$$

5.3 Proof of Theorem 5.2.3

The first step is to compute the derivative $\frac{d}{dt}\mathbb{P}^t(\eta_j \neq \sigma_j)$. Notice that

$$\mathbb{P}^t(\eta_j \neq \sigma_j) = \frac{1}{4}\mathbb{E}^t\left[\left(\eta_j - \sigma_j\right)^2\right], \tag{5.22}$$

so [27]

$$\frac{d}{dt}\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) = \frac{d}{dt}\left\{\frac{1}{4}\mathbb{E}^t\left[\left(\eta_j - \sigma_j\right)^2\right]\right\} = \frac{1}{4}\mathbb{E}^t\left[\Omega\left(\eta_j - \sigma_j\right)^2\right]. \tag{5.23}$$

Now, with $f = \left(\eta_j - \sigma_j\right)^2$, we have:

$$\nabla_i^\sigma f = \nabla_i^\eta f = 4\sigma_i\eta_i\delta_{ij} \tag{5.24}$$

and

$$f\left(\sigma^i, \eta^i\right) - f\left(\sigma^i, \eta\right) - f\left(\sigma, \eta^i\right) + f\left(\sigma, \eta\right) = 8\eta_i\sigma_i\delta_{ij} \tag{5.25}$$

where δ_{ij} is the Kronecker's delta.
Replacing in the expression of the generator, Ω, of the coupling we get:

$$\begin{aligned}\frac{d}{dt}\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) &= \frac{1}{4}\mathbb{E}^t\left[4\left(c^\omega\left(j,\eta\right) + a^\omega\left(j,t\right)\right)\sigma_j\eta_j - 8\min\left\{c^\omega\left(j,\eta\right), a^\omega\left(j,t\right)\right\}\chi_{(\sigma_j=\eta_j)}\right] = \\ &= \mathbb{E}^t\Big[-\left(c^\omega\left(j,\eta\right) + a^\omega\left(j,t\right)\right)\chi_{(\sigma_j\neq\eta_j)} + \\ &+ \left(c^\omega\left(j,\eta\right) + a^\omega\left(j,t\right)\right)\chi_{(\sigma_j=\eta_j)} - 2\min\left\{c^\omega\left(j,\eta\right), a^\omega\left(j,t\right)\right\}\chi_{(\sigma_j=\eta_j)}\Big] = \\ &= -\mathbb{E}^t\left[\left(c^\omega\left(j,\eta\right) + a^\omega\left(j,t\right)\right)\chi_{(\sigma_j\neq\eta_j)}\right] + \mathbb{E}^t\left[\left|c^\omega\left(j,\eta\right) - a^\omega\left(j,t\right)\right|\chi_{(\sigma_j=\eta_j)}\right].\end{aligned} \tag{5.26}$$

The quantity $\left(c^\omega\left(j,\eta\right) + a^\omega\left(j,t\right)\right)$ is bounded from below by a costant $M > 0$ and we can write:

$$\frac{d}{dt}\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) \leq -M\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) + \mathbb{E}^t\left[\left|\exp\left(-\beta\sigma_j(m_t + \omega)\right) - \exp\left(-\beta\sigma_j(\bar{\eta} + \omega)\right)\right|\right]. \tag{5.27}$$

At this point, we observe that

$$|\exp\left(-\beta\sigma_j(m_t+\omega)\right)-\exp\left(-\beta\sigma_j(\bar{\eta}+\omega)\right)|\leq D|\left(\bar{\eta}-m_t\right)| \tag{5.28}$$

with $D>0$; in fact

$$|\exp\left(-\beta\sigma_j(m_t+\omega)\right)-\exp\left(-\beta\sigma_j(\bar{\eta}+\omega)\right)|= \tag{5.29}$$

$$\exp\left(-\beta\sigma_j(m_t+\omega)\right)|1-\exp(-\beta\sigma_j(\bar{\eta}-m_t))|\leq D|\bar{\eta}-m_t| \tag{5.30}$$

because $|1-\exp(-\beta x)|\leq B|x|$, for some B finite and positive, on every compact subset of $\mathbb{R}$. Moreover, $|\left(-\sigma_j(\bar{\eta}-m_t)\right)|\leq 2$.

Finally, we have:

$$\frac{d}{dt}\mathbb{P}^t\left(\eta_j\neq\sigma_j\right)\leq -M\mathbb{P}^t\left(\eta_j\neq\sigma_j\right)+D\mathbb{E}^t\left[|\bar{\eta}-m_t|\right]\leq$$

$$\leq -M\mathbb{P}^t\left(\eta_j\neq\sigma_j\right)+D\sqrt{\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]}. \tag{5.31}$$

Now we need to find an upper bound for $\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]$. If $\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]\leq O\left(\frac{1}{N^\alpha}\right)$ $\forall\ t\in[0,\infty)$ with $\alpha>0$ we can obtain a time uniform bound for $\mathbb{P}^t\left(\eta_j\neq\sigma_j\right)$. Infact, using (5.18) we would have:

$$\mathbb{P}^t\left(\eta_j\neq\sigma_j\right)\leq -\frac{D}{MN^{\alpha/2}}\exp\left(-Mt\right)+\frac{D}{MN^{\alpha/2}}+\mathbb{P}^0\left(\eta_j\neq\sigma_j\right)\exp\left(-Mt\right). \tag{5.32}$$

Actually, we have $\alpha=1$.

Lemma 5.3.1 *In the region of parameters* (β,ω), *where* (5.4) *has an unique stable fixed point, the inequality*

$$\mathbb{E}^0\left[\left(\overline{\eta}-m_0\right)^2\right]\leq O\left(\frac{1}{N}\right) \tag{5.33}$$

implies,

$$\sup_{t\in[0,+\infty)}\mathbb{E}^t\left[\left(\overline{\eta}-m_t\right)^2\right]\leq O\left(\frac{1}{N}\right). \tag{5.34}$$

Proof: We try again with the Gronwall's Lemma.

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]=\mathbb{E}^t\left[L_N\left(\bar{\eta}-m_t\right)^2\right]+\mathbb{E}^t\left[\frac{\partial}{\partial t}\left(\bar{\eta}-m_t\right)^2\right]. \tag{5.35}$$

With $f=(\bar{\eta}-m_t)^2$, we get:

$$\nabla_i^{\eta} f=(\bar{\eta}-\frac{2\eta_i}{N}-m_t)^2-(\bar{\eta}-m_t)^2=-\frac{4}{N}\eta_i(\bar{\eta}-m_t)+\frac{4}{N^2} \tag{5.36}$$

and

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]=\mathbb{E}^t\left\{-\frac{4}{N}\sum_{i=1}^N c^{\omega}\left(i,\eta\right)\eta_i\left(\bar{\eta}-m_t\right)+\sum_{i=1}^N c^{\omega}\left(i,\eta\right)\frac{4}{N^2}\right\}$$

$$-2\left(\frac{d}{dt}m_t\right)\mathbb{E}^t\left(\bar{\eta}-m_t\right). \tag{5.37}$$

Now, we notice that $c^{\omega}\left(i,\eta\right)=-\eta_i\sinh\left(\beta(\bar{\eta}+\omega)\right)+\cosh\left(\beta(\bar{\eta}+\omega)\right)$. Remember also that $\frac{d}{dt}m_t=2\sinh\left(\beta(m_t+\omega)\right)-2m_t\cosh\left(\beta(m_t+\omega)\right)$ and substituting it in the last equation for the derivative of the mean, it yields:

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]=\left[-4\sinh\left(\beta(m_t+\omega)\right)+4m_t\cosh\left(\beta(m_t+\omega)\right)\right]\mathbb{E}^t\left(\bar{\eta}-m_t\right)+$$

$$+\mathbb{E}^t\left\{\sum_{i=1}^N-\frac{4}{N}\left[-\eta_i\sinh\left(\beta(\bar{\eta}+\omega)\right)+\cosh\left(\beta(\bar{\eta}+\omega)\right)\right]\eta_i\left(\bar{\eta}-m_t\right)+\sum_{i=1}^N c^{\omega}\left(i,\eta\right)\frac{4}{N^2}\right\}\le$$

$$\le\mathbb{E}^t\left\{4\left[\sinh\left(\beta(\bar{\eta}+\omega)\right)-\bar{\eta}\cosh\left(\beta(\bar{\eta}+\omega)\right)-\right.\right.$$

$$\left.\left.-\sinh\left(\beta(m_t+\omega)\right)+m_t\cosh\left(\beta(m_t+\omega)\right)\right]\left(\bar{\eta}-m_t\right)\right\}+4\exp\left(\beta\right)\frac{1}{N}. \tag{5.38}$$

If we set $\varphi_{\beta}^{\omega}\left(x\right)=4\sinh\left(\beta(x+\omega)\right)-4x\cosh\left(\beta(x+\omega)\right)$ we can write:

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\eta}-m_t\right)^2\right]\le\mathbb{E}^t\left\{\left[\varphi_{\beta}^{\omega}\left(\bar{\eta}\right)-\varphi_{\beta}^{\omega}\left(m_t\right)\right]\left(\bar{\eta}-m_t\right)\right\}+4\exp\left(\beta\right)\frac{1}{N}. \tag{5.39}$$

5.3.1 Case $\omega = 0$

The simplest case is when $\omega = 0$. In fact, if $\beta < 1$ the derivative of $\varphi_\beta^0(x)$ with respect to x is always negative and because of this, the following inequality holds:

$$\frac{\varphi_\beta^0(\bar{\eta}) - \varphi_\beta^0(m_t)}{\bar{\eta} - m_t} \leq -K, \tag{5.40}$$

where $-K = \max_{x\in[-1,1]} \frac{\partial \varphi_\beta^0}{\partial x}(x) < 0$.
It implies that,

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\eta} - m_t\right)^2\right] \leq -K\mathbb{E}^t\left[\left(\bar{\eta} - m_t\right)^2\right] + 4\exp\left(\beta\right)\frac{1}{N}. \tag{5.41}$$

Now we can apply the Gronwall's Lemma to obtain:

$$\mathbb{E}^t\left[\left(\bar{\eta} - m_t\right)^2\right] \leq -\frac{4\exp\left(\beta\right)}{KN}\exp\left(-Kt\right) + \frac{4\exp\left(\beta\right)}{KN} + \mathbb{E}^0\left[\left(\bar{\eta} - m_0\right)^2\right] \tag{5.42}$$

In this way, we have obtained an uniform upper bound for $\mathbb{E}^t\left[\left(\bar{\eta} - m_t\right)^2\right]$:

$$\sup_{t\in[0,\infty)} \mathbb{E}^t\left[\left(\bar{\eta} - m_t\right)^2\right] \leq O\left(\frac{1}{N}\right). \tag{5.43}$$

5.3.2 Case $\omega \neq 0$

If $\omega \neq 0$ the derivative of $\varphi_\beta^\omega(x)$ fails to be negative at some points $x \in [-1,1]$. However, if $|\omega| > \omega(\beta)$, the McKean-Vlasov equation

$$\frac{d}{dt}m_t = 2\sinh\left(\beta(m_t + \omega)\right) - 2m_t\cosh\left(\beta(m_t + \omega)\right) = \frac{1}{2}\varphi_\beta^\omega(m_t) \tag{5.44}$$

has an unique, globally attractive, linearly stable fixed point that we call $m_*(\omega)$. The equilibrium point is the unique solution of the equation

$$\varphi_\beta^\omega(m_*(\omega)) = 0 \tag{5.45}$$

which is equivalent to

$$m_*(\omega) = \tanh[\beta(m_*(\omega) + \omega)]. \tag{5.46}$$

To fix the idea we can think of $\omega > 0$; in this case $0 < m_*(\omega) < 1$.
We compute:

$$\frac{\partial}{\partial x}\varphi_\beta^\omega(x) = (\beta-1)\cosh(\beta(x+\omega)) - \beta x \sinh(\beta(x+\omega)) =$$
$$= \cosh(\beta(x+\omega))[\beta - 1 - \beta x \tanh(\beta x + h)] \quad (5.47)$$

so,

$$\frac{\partial}{\partial x}\varphi_\beta^\omega(m_*(\omega)) = \cosh(\beta(m_*(\omega)+\omega))[\beta - 1 - \beta(m_*(\omega))^2] < 0, \quad (5.48)$$

and it remains strictly negative in a neighbourhood V of $m_*(\omega)$ of the form $V = [m_*(\omega)-\xi, 1]$, $\xi > 0$, because of the continuity of the derivative of $\varphi_\beta^\omega(x)$ and because $\beta x \tanh(\beta x + h)$ is an increasing function of x in $[m_*(\omega), 1]$. We recall that to be globally attractive, since we are in 1-dimension, means that $\varphi_\beta^\omega(x) < 0$ when $x \in (m_*(\omega), 1]$ and $\varphi_\beta^\omega(x) > 0$ for $x \in [-1, m_*(\omega))$. So, one can choose ξ such that $\max_{x\in V}\varphi_\beta^\omega(x) \le \inf_{[-1,1]\setminus V}\varphi_\beta^\omega(x)$ and that $\frac{\partial}{\partial x}\varphi_\beta^\omega(m_*(\omega)-\xi)$ is still negative.
With this choice of ξ, if $m_t \in V$, we have:

$$\frac{\varphi_\beta^\omega(\bar\eta) - \varphi_\beta^\omega(m_t)}{\bar\eta - m_t} \le -K \text{ if } m_t \in V. \quad (5.49)$$

Here $-K = \sup_{m_t \in V, \bar\eta\in[-1,1]} \frac{\varphi_\beta^\omega(\bar\eta)-\varphi_\beta^\omega(m_t)}{\bar\eta - m_t} < 0$.

In fact, if $\bar\eta$ and m_t belong both to V the inequality holds because here the derivative is strictly negative, while if $\bar\eta \notin V$ the difference $\varphi_\beta^\omega(\bar\eta) - \varphi_\beta^\omega(m_t)$ is positive and $\bar\eta - m_t$ is negative.
The equilibrium $m_*(\omega)$ is globally attractive, so m_t belongs to V for $t \ge T$ for some T finite and we can estimate $\mathbb{E}^t\left[(\bar\eta - m_t)^2\right]$ fot $t \ge T$. The same computation used to derive (5.43), now gives, for $t = T + s$, $s \ge 0$,

$$\frac{d}{dt}\mathbb{E}^t\left[(\bar\eta - m_t)^2\right] = \frac{d}{ds}\mathbb{E}^{s+T}\left[(\bar\eta - m_{s+T})^2\right] \le$$
$$\le -K\mathbb{E}^{s+T}\left[(\bar\eta - m_{s+T})^2\right] + \frac{4\exp(\beta)}{N}. \quad (5.50)$$

And, by the Gronwall's Lemma:

$$\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] \leq O\left(\frac{1}{N}\right) + \mathbb{E}^T\left[(\bar{\eta}-m_T)^2\right] \tag{5.51}$$

With the same technique one can see that for $0 \leq t \leq T$

$$\frac{d}{dt}\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] \leq C\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] + \frac{4\exp(\beta)}{N}, \tag{5.52}$$

where $C = \max_{x\in[-1,1]} \frac{\partial}{\partial x}\varphi_\beta^\omega(x) > 0$. Hence, we have:

$$\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] \leq \frac{4\exp(\beta)}{CN}\left[\exp(Ct)-1\right] + \mathbb{E}^0\left[(\bar{\eta}-m_0)^2\right], \tag{5.53}$$

that, with the right initial conditions (see (5.43)) implies

$$\mathbb{E}^T\left[(\bar{\eta}-m_T)^2\right] \leq O\left(\frac{1}{N}\right). \tag{5.54}$$

So, using again the Gronwall's Lemma, we obtain:

$$\sup_{t\in[0,\infty)} \mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] \leq O\left(\frac{1}{N}\right). \tag{5.55}$$

also when $\omega > 0$.

The same steps bring to the same results also when $\omega < 0$. We conclude that a time uniform upper bound for $\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right]$ holds for every ω when $|\omega| > \omega(\beta)$ (see (5.43) and (5.55)).

□

Turning back to the proof of the main theorem, in (5.39) we were left with:

$$\frac{d}{dt}\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) \leq -M\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) + D\sqrt{\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right]} \tag{5.56}$$

Now, we know that $\mathbb{E}^t\left[(\bar{\eta}-m_t)^2\right] \leq O\left(\frac{1}{N}\right)$ whatever ω is if $\beta < 1$, and under the condition (**??**) if $\beta \geq 1$. Here we have uniform propagation of chaos. In fact

$$\frac{d}{dt}\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) \leq -M\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) + O\left(\sqrt{\frac{1}{N}}\right). \tag{5.57}$$

Using again the Gronwall's Lemma we have:

$$\mathbb{P}^t\left(\eta_j \neq \sigma_j\right) \leq O\left(\sqrt{\frac{1}{N}}\right), \ \forall t \in [0, \infty). \tag{5.58}$$

Notice that the inequality holds with the right initial conditions. The quantity η, σ, m_0 have to be near at the beginning in order to have $\mathbb{E}^0\left[\left(\bar{\eta} - m_0\right)^2\right] \leq O\left(\frac{1}{N}\right)$ and $\mathbb{P}^0\left(\eta_j \neq \sigma_j\right) \leq O\left(\sqrt{\frac{1}{N}}\right)$. This concludes the proof of Theorem 5.2.3.

Chapter 6

Propagation of chaos in a model for large portfolio losses

6.1 Introduction

In this chapter we consider uniform propagation of chaos for a mean-field interacting particle system modeling the propagation of financial distress among a network of firms linked by financial relationships [24]. The model depends on two real positive parameters and it exhibits a phase transition. We show that in the uniqueness regime (i.e. the associated McKean-Vlasov equations have an unique stable solution) uniform propagation of chaos holds. The argument used for the Curie-Weiss model can be adapted to this case, even though the model is non reversible and a one dimensional parameter is not sufficient anymore.

6.2 Description of the model

We briefly describe the model appearing in [24]. Consider N sites that we look at as a network of firms linked by some financial relationship; for instance, they could belong to the same sector of the market. The model in [24] is applied to describe the propagation of financial distress (or default) in a network of firms, and the related credit risk of a financial institution that, for example, lent money to the firms. Because of the relationships between the firms of the portfolio, default may be con-

tagious and there might be clustering of default.
The financial state of each firm is represented by a couple of spin variables $(\sigma_i, \omega_i) \in \{-1,+1\}^2$. The first one , σ_i, could be interpreted as a rating indicator, and a negative value means that there is a higher probability of not being able to pay back obligations. The second spin variable, ω_i, amplifies or reduces the effect of changes at the level of the σ_i indicators; it represents how a firm is able to face a crises and a positive value could be interpreted as a positive reaction to financial distress. The variable ω_i is a more fundamental indicator than σ_i and in [24] it is supposed to be not directly observable from the market; one could think of ω_i as a liquidity indicator. The success in reacting to a crisis depends also on the global situation of the market that in [24] is represented by the empirical avarage $\bar{\sigma} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i$.
The dynamics of the contagion can be schematized as follows:

ω_i	$\rightarrow$	σ_i	$\rightarrow$	$\bar{\sigma}$	$\rightarrow$	ω_i
health of the firm		rating indicator		situation of the market		health of the firm

The system flips with rates:

$$\begin{array}{lllll} \omega_i & \rightarrow & -\omega_i & \text{at rate} & \exp(-\gamma\omega_i\bar{\sigma}) \\ \sigma_i & \rightarrow & -\sigma_i & \text{at rate} & \exp(-\beta\sigma_i\omega_i) \end{array}$$

Thus, the system evolves in time according to the following infinitesimal generator:

$$L_N f(\sigma,\omega) = \sum_{i=1}^{N} \exp(-\beta\sigma_i\omega_i)\nabla_i^{\sigma} f(\sigma,\omega) + \sum_{i=1}^{N} \exp(-\gamma\omega_i\bar{\sigma})\nabla_i^{\omega} f(\sigma,\omega). \tag{6.1}$$

Notice that the rate of the transition $\omega_i \rightarrow -\omega_i$ depends on the empirical average $\bar{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i$ meaning that the situation of the network influences the variable ω_i, so, the financial distress somewhere in the network may increase the default probability of the partners. In the same way the flip $\sigma_i \rightarrow -\sigma_i$ depends on the value of ω_i.

Let us switch to the Mckean-Vlasov equations for the system just described. On the compact interval $[0, T]$ when N goes to infinity the time evolution of $\bar{\sigma}$ becomes deterministic [24]. More precisely define the quantities:

$$\bar{\omega} = \frac{1}{N}\sum_{i=1}^{N}\omega_i, \;\; \bar{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i, \;\; \text{and} \;\; \overline{\sigma\omega} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i\omega_i. \tag{6.2}$$

The evolution of this triplet is markovian and their weak limit when N goes to infinity has a deterministic dynamics. Let's call m_t^{ω}, m_t^{ω} and ${}_t^{\omega\sigma}$ this limit, the following differential equations are satisfied (McKean-Vlasov equations):

$$\begin{cases} \frac{d}{dt} m_t^{\sigma} = 2\sinh(\beta) m_t^{\omega} - 2\cosh(\beta) m_t^{\sigma}; \\ \frac{d}{dt} m_t^{\omega} = 2\sinh\left(\gamma m_t^{\sigma}\right) - 2\cosh\left(\beta m_t^{\sigma}\right) m_t^{\omega}; \\ \frac{d}{dt} m_t^{\omega\sigma} = 2\sinh(\beta) + 2 m_t^{\sigma}\sinh(\gamma m_t^{\sigma}) - 2\left(\cosh(\beta) - \cosh(\gamma m_t^{\sigma})\right). \end{cases} \tag{6.3}$$

To prove this, like in the previous chapter, we use Theorem 5.2.2. A long but straightforward computation yields (see [24] for details)

$$\begin{aligned} L_N f(\bar\sigma,\bar\omega,\overline{\omega\sigma}) = \frac{N}{4} \sum_{j,k\in\{-1,1\}} & [j(\bar\sigma + k\bar\omega + jk\overline{\omega\sigma} + 1] \\ & \times \left\{ e^{-\beta jk} \left[f\left(\bar\sigma - \frac{2}{N} j, \bar\omega, \overline{\omega\sigma} - \frac{2}{N} jk\right)) - f(\bar\sigma,\bar\omega,\overline{\omega\sigma}) \right] \right. \\ & \left. + e^{-\gamma\bar\sigma k} \left[f\left(\bar\sigma, \bar\omega - \frac{2}{N} k, \overline{\omega\sigma} - \frac{2}{N} jk\right)) - f(\bar\sigma,\bar\sigma,\overline{\omega\sigma}) \right] \right\}. \end{aligned} \tag{6.4}$$

This implies that the process $(\bar\sigma(t),\bar\omega(t),\overline{\omega\sigma}(t))$ is markovian with generator

$$\begin{aligned} K_N f(\xi,\eta,\theta) = \frac{N}{4} \sum_{j,k\in\{-1,1\}} & [j(\xi + k\eta + jk\theta + 1] \\ & \times \left\{ e^{-\beta jk} \left[f\left(\xi - \frac{2}{N} j, \eta, \theta - \frac{2}{N} jk\right)) - f(\xi,\eta,\theta) \right] \right. \\ & \left. + e^{-\gamma\xi k} \left[f\left(\xi, \eta - \frac{2}{N} k, \theta - \frac{2}{N} jk\right)) - f(\xi,\eta,\theta) \right] \right\}. \end{aligned} \tag{6.5}$$

Since, for functions f of class $\mathscr{C}^1$ with compact support,

$$K_N f(\xi,\eta,\theta) = K f(\xi,\eta,\theta) + O\left(\frac{1}{N}\right)$$

uniformly in (ξ,η,θ), with

$$\begin{aligned} K f(\xi,\eta,\theta) = {} & 2\left(\sinh(\gamma\xi) - \eta\cosh(\gamma\xi)\right) \frac{\partial}{\partial x} f(\xi,\eta,\theta) + 2\left(\eta\sinh(\beta) - \xi\cosh(\beta)\right) \frac{\partial}{\partial y} f(\xi,\eta,\theta) \\ & + 2\left(\sinh(\beta) - \theta\cosh(\beta) + \xi\sinh(\gamma\xi) - \theta\cosh(\gamma\xi)\right) \frac{\partial}{\partial z} f(\xi,\eta,\theta) \end{aligned} \tag{6.6}$$

as in Proposition 5.2.1 we conclude that the process $(\bar\sigma(t),\bar\omega(t),\overline{\omega\sigma}(t))$ converges weakly to a solution of (6.3)

The numbers and the stability properties of solutions of (6.3) depends on the parameters γ and β. Note that $m_t^{\sigma\omega}$ does not appears in the first and in the second equation, so system (6.3) is essentially two-dimensional. Moreover, the equilibria $m_t^{\sigma\omega}$ are completely determined by those of m_t^σ. So we reduce to the study of equilibria and relative stability for the first two equations. The situation is the following [24]:

- If $\gamma \leq \frac{1}{\tanh(\beta)}$, then (6.3) has $(0,0)$ as unique equilibrium solution, which is globally asymptotically stable, i.e. for every initial condition (m_0^σ, m_0^ω), we have:

$$\lim_{t\to\infty}(m_t^\sigma, m_t^\omega) = (0,0). \tag{6.7}$$

  When $\gamma < \frac{1}{\tanh(\beta)}$, the equilibrium is also linearly stable.

- For $\gamma > \frac{1}{\tanh(\beta)}$, the point $(0,0)$ is still an equilibrium but it is unstable. Two new equilibria (m_*^σ, m_*^ω) and $(-m_*^\sigma, -m_*^\omega)$ arise. They are lineraly stable with open basin of attraction Γ^+ and Γ^-. Moreover, $\Gamma^+ \cup \Gamma^- = [-1,1]^2 \setminus \Gamma$, where Γ is an invariant curve for (6.3) containing $(0,0)$.

From now on, we limit our study to the case $\gamma < \frac{1}{\tanh(\beta)}$.

6.3 Uniform propagation of chaos

As in the previous chapter, looking for a uniform propagation of chaos property, it's natural to compare the system with generator L_N with the following one where the dynamics of each spin is independent and m_t^σ appears in place of $\bar{\sigma}$ in the transition rates.
The spin variables are $(x_i, y_i) \in \{-1,+1\}^2$ and the infinitesimal generator is given by:

$$G_N f(x,y) = \sum_{i=1}^N \exp\left(-\beta x_i y_i\right)\nabla_i^x f(x,y) + \sum_{i=1}^N \exp\left(-\gamma y_i m_t^\sigma\right)\nabla_i^y f(x,y). \tag{6.8}$$

We prove the following

Theorem 6.3.1 *When $\gamma < \frac{1}{\tanh(\beta)}$ there exists a probability space where both the systems with infinitesimal generators* (6.1) *and* (6.8) *can be realized. Moreover, if*

$$\mathbb{P}^0(\sigma_i \neq x_i) \leq O\left(\sqrt{\frac{1}{\sqrt{N}}}\right) \ \text{and}\ \mathbb{P}^0\left(\omega_i \neq y_i\right) \leq O\left(\sqrt{\frac{1}{\sqrt{N}}}\right), \tag{6.9}$$

and

$$\mathbb{E}^0\left[\left(\bar{\sigma}-m_0^{\sigma}\right)^2\right]\le O\left(\frac{1}{N}\right),\ \ \mathbb{E}^0\left[\left(\bar{\omega}-m_0^{\omega}\right)^2\right]\le O\left(\frac{1}{N}\right), \tag{6.10}$$

then

$$\mathbb{P}^t\left(\sigma_i\neq x_i\right)+\mathbb{P}^t\left(\omega_i\neq y_i\right)\le O\left(\frac{1}{\sqrt{N}}\right). \tag{6.11}$$

Remark: Clearly conditions (6.9) and (6.10) ares satisfied if $\{(\sigma_i(0),\omega_i(0) : i = 1,\dots N\}$ are i.i.d., and we set, for $i = 1,\dots,N$, $\sigma_i(0) = x_i(0)$, $\omega_i(0) = y_i(0)$.

The strategy of the proof is the same as in the the Curie-Weiss model's case. We use a coupling to realize the two systems in the same probability space, then we use twice the Gronwall's Lemma to bound the distance between the systems with a decreasing function of N, independent of t.

6.4 Proof of Theorem 6.3.1

The first step is to construct a suitable coupling to make the systems living in the same probability space. We use Basic Coupling. Infinitesimal generators (6.1) and (6.8) are composed by two pieces corresponding to gradients with respect to different variables. We couple the dynamics of the spin σ_i with x_i and that of ω_i with y_i. The coupling and the infinitesimal generator will be

$$\Omega f(\sigma,\omega,x,y)=\Omega_1 f(\sigma,\omega,x,y)+\Omega_2 f(\sigma,\omega,x,y) \tag{6.12}$$

where

$$\Omega_1 f(\sigma,\omega,x,y)=\sum_{i=1}^N \exp\left(-\beta\sigma_i\omega_i\right)\nabla_i^{\sigma} f(\sigma,\omega,x,y)+\sum_{i=1}^N \exp\left(-\beta x_i y_i\right)\nabla_i^{x} f(\sigma,\omega,x,y)$$

$$+\sum_{i=1}^N \min\left\{\exp\left(-\beta\sigma_i\omega_i\right),\exp\left(-\beta x_i y_i\right)\right\}\Big(f(\sigma^i,\omega,x^i,y)-f(\sigma^i,\omega,x,y)$$
$$-f(\sigma,\omega,x^i,y)+f(\sigma,\omega,x,y)\Big), \tag{6.13}$$

and

$$\Omega_2 f(\sigma,\omega,x,y) = \sum_{i=1}^{N} \exp(-\gamma\omega_i\bar{\sigma})\nabla_i^{\omega} f(\sigma,\omega,x,y) + \sum_{i=1}^{N} \exp(-\gamma y_i m_t^{\sigma})\nabla_i^{y} f(\sigma,\omega,x,y)$$

$$+ \sum_{i=1}^{N} \min\left\{\exp(-\gamma\omega_i\bar{\sigma}), \exp(-\gamma y_i m_t^{\sigma})\right\} \Big(f(\sigma,\omega^i,x,y^i) - f(\sigma,\omega^i,x,y)$$
$$- f(\sigma,\omega,y^i,x) + f(\sigma,\omega,y,x)\Big). \quad (6.14)$$

We want to give an uniform bound for the probability

$$\mathbb{P}^t(x_i \neq \sigma_i) + \mathbb{P}^t\left(\omega_i \neq y_i\right). \quad (6.15)$$

To do this, following the method used for the Curie-Weiss model, we consider functions analogous to (5.22) counting sites where spins are different. In the present case, to make computations simpler, we write them in the form :

$$f_1(\sigma,x) = \frac{1}{2N}\sum_{i=1}^{N}(1-\sigma_i x_i) \quad (6.16)$$

and

$$f_2(\omega,y) = \frac{1}{2N}\sum_{i=1}^{N}(1-\omega_i y_i). \quad (6.17)$$

Notice that $\mathbb{P}^t(x_i \neq \sigma_i) = \mathbb{E}^t\left[f_1(\sigma,x)\right]$ and that $\mathbb{P}^t\left(\omega_i \neq y_i\right) = \mathbb{E}^t\left[f_2(\omega,y)\right]$. Thus, to prove Theorem 6.3.1 it is sufficient to show a uniform bound for $\mathbb{E}^t\left[f_1(\sigma,x) + \lambda f_2(\omega,y)\right]$, where the constant λ is positive and has to be suitably chosen.

The second step is to use the Gronwall's Lemma. So we have to compute:

$$\frac{d}{dt}\mathbb{E}^t\left[f\left(\sigma,\omega,x,y\right)\right] = \mathbb{E}^t\left[\Omega f\left(\sigma,\omega,x,y\right)\right], \quad (6.18)$$

with $f\left(\sigma,\omega,x,y\right) = f_1(\sigma,x) + \lambda f_2\left(\omega,y\right)$. We have:

$$\Omega f\left(\sigma,\omega,x,y\right) = \Omega_1 f_1(\sigma,x) + \lambda\Omega_2 f_2\left(\omega,y\right). \quad (6.19)$$

We split this computation in two pieces. For $\Omega_1 f_1(\sigma,x)$ we have

$$\nabla_i^{\sigma} f_1(\sigma,x) = \frac{1}{2N}\sum_{j}^{N}\nabla_i^{\sigma}(1-\sigma_j x_j) = \frac{1}{N}\sigma_i x_i; \quad (6.20)$$

$$\nabla_i^x f_1(\sigma, x) = \frac{1}{N}\sigma_i x_i; \tag{6.21}$$

$$f_1(\sigma^i, x^i) - f_1(\sigma^i, x) - f_1(\sigma, x^i) + f_1(\sigma, x) = -\frac{2}{N}\sigma_i x_i. \tag{6.22}$$

We obtain:

$$\begin{aligned}
\Omega_1 f_1(\sigma, x) &= \sum_{i=1}^{N} \exp(-\beta\sigma_i\omega_i)\frac{1}{N}\sigma_i x_i + \sum_{i=1}^{N} \exp(-\beta x_i y_i)\frac{1}{N}\sigma_i x_i \\
&\quad -2\sum_{i=1,\sigma_i=x_i}^{N} \min\left\{\exp(-\beta\sigma_i\omega_i), \exp(-\beta x_i y_i)\right\}\frac{1}{N}\sigma_i x_i \\
&= \sum_{i=1,\sigma_i\neq x_i}^{N} \big(\exp(-\beta\sigma_i\omega_i) + \exp(-\beta x_i y_i)\big)\left(\frac{-1}{N}\right) \\
&\quad + \frac{1}{N}\sum_{i=1,\sigma_i=x_i}^{N} \big(\exp(-\beta\sigma_i\omega_i) + \exp(-\beta x_i y_i)\big) - 2\min\left\{\exp(-\beta\sigma_i\omega_i), \exp(-\beta x_i y_i)\right\} \\
&= -\frac{2}{N}\sum_{i=1,\sigma_i\neq x_i}^{N} \cosh(\beta) + \frac{2}{N}\sum_{i=1,\sigma_i=x_i}^{N} \sinh(\beta)|\omega_i - y_i| \\
&\leq -2\cosh(\beta) f_1(\sigma, x) + 2\sinh(\beta) f_2(\omega, y).
\end{aligned} \tag{6.23}$$

In the same way, for $\Omega_2 f_2(\omega, y)$ we obtain

$$\begin{aligned}
\Omega_2 f_2(\omega, y) &= -\sum_{i=1,\omega_i\neq y_i} \frac{1}{N}\big(\exp(-\gamma\omega_i\bar{\sigma}) + \exp(-\gamma y_i m_t^{\sigma})\big) \\
&\quad + \frac{1}{N}\sum_{i=1,\omega_i=y_i} |\exp(-\gamma\omega_i\bar{\sigma}) - \exp(-\gamma y_i m_t^{\sigma})| \\
&\leq -2\exp(-\gamma) f_2(\omega, y) + K\sum_{i=1,\omega_i=y_i} \frac{1}{N}|\bar{\sigma} - m_t^{\sigma}| \\
&\leq -2\exp(-\gamma) f_2(\omega, y) + \frac{K}{N}\sum_{i=1}^{N} |\bar{\sigma} - m_t^{\sigma}|.
\end{aligned} \tag{6.24}$$

Taking the mean it reads:

$$\frac{d}{dt}\mathbb{E}^t\left(f_1(\sigma,x)+\lambda f_2(\omega,y)\right)=\mathbb{E}^t\left[\Omega\left(f_1(\sigma,x)+\lambda f_2(\omega,y)\right)\right]$$

$$=\mathbb{E}^t\left[\Omega_1\left(f_1(\sigma,x)\right)+\lambda\Omega_2\left(f_2(\omega,y)\right)\right]$$

$$=\mathbb{E}^t\left[-2\cosh(\beta)f_1(\sigma,x)-2\lambda\left(\exp(-\gamma)-\frac{\sinh(\beta)}{\lambda}\right)f_2(\omega,y)+\frac{K}{N}\sum_{i=1}^{N}|\bar{\sigma}-m_t^{\sigma}|\right] \quad (6.25)$$

Notice that for λ large enough, there exists a positive constant C such that

$$2\cosh(\beta)>C \quad (6.26)$$

and

$$2\left(\exp(-\gamma)-\frac{\sinh(\beta)}{\lambda}\right)>C, \quad (6.27)$$

thus, the following inequality holds:

$$\frac{d}{dt}\mathbb{E}^t\left(f_1(\sigma,x)+\lambda f_2(\omega,y)\right)\leq -C\mathbb{E}^t\left(f_1(\sigma,x)+\lambda f_2(\omega,y)\right)+\mathbb{E}^t\left(\frac{K}{N}\sum_{i=1}^{N}|\bar{\sigma}-m_t^{\sigma}|\right). \quad (6.28)$$

Looking at (6.28) we see that to prove uniform propagation of chaos we have to estimate $\mathbb{E}^t\left(\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right)$; if $\mathbb{E}^t\left(\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right)\leq O\left(\frac{1}{N}\right)$, $\forall t\in[0,\infty)$ we can conclude using (5.18). It turns out to be simpler to find a bound for $\mathbb{E}^t\left(\left(\bar{\sigma}-m_t^{\sigma}\right)^2+\left(\bar{\omega}-m_t^{\omega}\right)^2\right)$.

Lemma 6.4.1 *In the region of parameters* (γ,β) *where* $\gamma<\frac{1}{\tanh(\beta)}$ *the conditions*

$$\mathbb{E}^0\left[\left(\bar{\sigma}-m_0^{\sigma}\right)^2\right]\leq O\left(\frac{1}{N}\right), \quad (6.29)$$

$$\mathbb{E}^0\left[\left(\bar{\omega}-m_0^{\omega}\right)^2\right]\leq O\left(\frac{1}{N}\right), \quad (6.30)$$

imply

$$\mathbb{E}^t\left[\left(\bar{\sigma}-m_t^{\sigma}\right)^2+\left(\bar{\omega}-m_t^{\omega}\right)^2\right]\leq O\left(\frac{1}{N}\right). \quad (6.31)$$

Proof: To use the Gronwall's Lemma we compute the derivative:

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right]+\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\omega}-m_t^{\omega}\right)^2\right]. \tag{6.32}$$

The first summand in the left hand side of (6.32) is

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right]=\mathbb{E}^t\left[L_N\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right]+\mathbb{E}^t\left[\frac{\partial}{\partial t}\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right]$$

$$=\mathbb{E}^t\left[\sum_{i=1}^{N}\exp\left(-\beta\sigma_i\omega_i\right)\left(\left(\bar{\sigma}-\frac{2}{N}\sigma_i-m_t^{\sigma}\right)^2-\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right)\right]-\left(\frac{d}{dt}m_t^{\sigma}\right)\mathbb{E}^t\left(\sigma-m_t^{\sigma}\right)$$

$$=\mathbb{E}^t\left[\sum_{i=1}^{N}\exp\left(-\beta\sigma_i\omega_i\right)\left(\frac{4}{N^2}-\frac{4}{N}\sigma_i\left(\bar{\sigma}-m_t^{\sigma}\right)\right)\right]-\left(\frac{d}{dt}m_t^{\sigma}\right)\mathbb{E}^t\left(\sigma-m_t^{\sigma}\right)$$

$$=\mathbb{E}^t\left[\sum_{i=1}^{N}\left((\sigma_i\omega_i)\sinh(\beta)+\cosh(\beta)\right)\left(\frac{4}{N^2}-\frac{4}{N}\sigma_i\left(\bar{\sigma}-m_t^{\sigma}\right)\right)\right]-\left(\frac{d}{dt}m_t^{\sigma}\right)\mathbb{E}^t\left(\sigma-m_t^{\sigma}\right)$$

$$=O\left(\frac{1}{N}\right)+\mathbb{E}^t\left[\left(4\bar{\omega}\sinh(\beta)-4\bar{\sigma}\cosh(\beta)\right)\left(\sigma-m_t^{\sigma}\right)\right]-\left(\frac{d}{dt}m_t^{\sigma}\right)\mathbb{E}^t\left(\sigma-m_t^{\sigma}\right). \tag{6.33}$$

The same computation for the second piece gives:

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\omega}-m_t^{\omega}\right)^2\right]$$

$$=O\left(\frac{1}{N}\right)+\mathbb{E}^t\left[\left(4\sinh(\gamma\bar{\sigma})-4\bar{\omega}\cosh(\gamma\bar{\sigma})\right)\left(\bar{\omega}-m_t^{\omega}\right)\right]-\left(\frac{d}{dt}m_t^{\omega}\right)\mathbb{E}^t\left(\omega-m_t^{\omega}\right). \tag{6.34}$$

Now putting (6.33) and (6.34) together and using (6.3) we obtain:

$$\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\sigma}-m_t^{\sigma}\right)^2\right]+\frac{d}{dt}\mathbb{E}^t\left[\left(\bar{\omega}-m_t^{\omega}\right)^2\right]$$

$$=O\left(\frac{1}{N}\right)+\mathbb{E}^t\left[\left(4\sinh(\beta)\left(\bar{\omega}-m_t^{\omega}\right)-4\cosh(\beta)\left(\bar{\sigma}-m_t^{\sigma}\right)\right)\left(\bar{\sigma}-m_t^{\sigma}\right)\right.$$
$$\left.+\left(4\sinh(\gamma\bar{\sigma})-4\bar{\omega}\cosh(\gamma\bar{\sigma})-4\sinh(\gamma m_t^{\sigma})+4m_t^{\omega}\cosh(\gamma m_t^{\sigma})\right)\left(\bar{\omega}-m_t^{\omega}\right)\right]. \tag{6.35}$$

At this point we use the same line of reasoning of the previous chapter when we was dealing with the Curie-Weiss model in a magnetic field. Since $(0,0)$ is a globally attractive equilibrium when $\gamma \leq \frac{1}{\tanh(\beta)}$, for t large, let's say $t > T$, $(m_t^\sigma, m_t^\omega) \in V$, where V is a neighborhood of $(0,0)$. Thus, it is sufficient to prove that (6.31) holds true for $(m_t^\sigma, m_t^\omega) \in V$. Moreover, (6.31) holds when $t \leq T$; the computations are the same as in the proof of Theorem 5.2.3 to obtain (5.55) and we skip the details here. To begin with, we consider the case with $m_t^\omega = m_t^\sigma = 0$. In this limit case the right hand side of (6.35) becomes:

$$\left(4\sinh(\beta)\bar{\omega} - 4\cosh(\beta)\bar{\sigma}\right)\bar{\sigma} + \left(4\sinh(\gamma\bar{\sigma}) - 4\bar{\omega}\cosh(\gamma\bar{\sigma})\right)\bar{\omega}$$

$$\leq 4\sinh(\beta)\bar{\omega}\bar{\sigma} - 4\cosh(\beta)\bar{\sigma}^2 + 4\sinh(\gamma\bar{\sigma})\bar{\omega} - 4\bar{\omega}^2. \quad (6.36)$$

In the region of the parameters where $\gamma < \frac{1}{\tanh(\beta)}$ the McKean-Vlasov equations for the system have an unique, globally, asymptotically stable solution that is also linearly stable (see [24]). The last expression in (6.36) is the driving term for $\frac{d}{dt}\left(\left(m_t^\sigma\right)^2 + \left(m_t^\omega\right)^2\right)$ and the results for the McKean-Vlasov equations imply that (6.36) is always negative but in $(0,0)$. Moreover, $(\bar{\omega}, \bar{\sigma}) \in [-1,1]^2$ and the Taylor's expansion of (6.36) in $(0,0)$ is a negative quadratic form. Thus, there exist a positive constant, let's say A, such that

$$4\sinh(\beta)\bar{\omega}\bar{\sigma} - 4\cosh(\beta)\bar{\sigma}^2 + 4\sinh(\gamma\bar{\sigma})\bar{\omega} - 4\bar{\omega}^2 \leq -A\left(\bar{\sigma}^2 + \bar{\omega}^2\right). \quad (6.37)$$

Because of continuity this holds also for (6.35) when $\left(m_t^\omega, m_t^\sigma\right)$ is in a neighborhood of $(0,0)$ and it happens when t is large enough. This is the same as the Curie-Weiss model in presence of a magnetic field, thus the same line of reasoning leads to the desired result. □

In view of this last Lemma, (6.28) becomes:

$$\frac{d}{dt}\mathbb{E}^t\left(f_1(\sigma,x) + \lambda f_2(\omega,y)\right) \leq -C\mathbb{E}^t\left(f_1(\sigma,x) + \lambda f_2(\omega,y)\right) + O\left(\sqrt{\frac{1}{N}}\right), \quad (6.38)$$

and the Gronwall's Lemma can be applied, concluding the proof of Theorem 6.3.1.

Chapter 7

Uniform fluctuation theorem for the Curie-Weiss Model

7.1 Uniform fluctuation theorem

In this chapter we deal with uniformity in time of the fluctuation theorem for the Curie-Weiss model in the subcritical regime: i.e. $\beta < 1$.

We resume the Curie-Weiss model. Recall that, if $\eta = (\eta_1, \dots, \eta_N)$ is a configuration of the N spins, the Curie-Weiss model is the spin-flip system with generator:

$$L_N f(\eta) = \sum_{i=1}^{N} \exp(-\beta \eta_i \bar{\eta}) \nabla_i^{\eta} f(\eta). \tag{7.1}$$

Throughout this chapter we assume that $\eta_i(0)$, $1 = 1, \dots, N$ are i.i.d. random variables with $\mathbb{E}^0 \left(\eta_1(0)\right) = 0$. The object of interest is

$$X^N(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \eta_i = \sqrt{N} \bar{\eta}, \tag{7.2}$$

that sometimes we refer to as the empirical fluctuation process[1]. The dynamics of

[1] Usually the fluctuation has the form

$$X^N(t) = \sqrt{N}(\bar{\eta} - m_t).$$

Our assumption on the initial condition ($\eta_i(0)$, $i = 1, \dots, N$ i.i.d. with $\mathbb{E}^0 \left(\eta_1(0)\right) = 0$) implies $m_t = 0$, $\forall\, t \geq 0$. In this way, the fluctuation process has the form (7.2). Actually, the arguments we use depend

$X^N(t)$ are markovian with infinitesimal generator given by:

$$\mathscr{L}_N f(x) = c(x,+)\nabla^+ f(x) + c(x,-)\nabla^- f(x), \tag{7.3}$$

where:

$$c(x,+) = \frac{N}{2}\left(1 - \frac{x}{\sqrt{N}}\right) e^{(\beta \frac{x}{\sqrt{N}})}, \tag{7.4}$$

$$c(x,-) = \frac{N}{2}\left(1 + \frac{x}{\sqrt{N}}\right) e^{(-\beta \frac{x}{\sqrt{N}})}, \tag{7.5}$$

and

$$\nabla^+ f(x) = f\left(x + \frac{2}{\sqrt{N}}\right) - f(x), \tag{7.6}$$

$$\nabla^- f(x) = f\left(x - \frac{2}{\sqrt{N}}\right) - f(x). \tag{7.7}$$

Moreover, the empirical fluctuations have a weak limit that is an Ornstein-Ulembeck type diffusion equation. See few lines below for a proof. Here, we call $x(t)$ the limit process of $X^N(t)$; the diffusion equation is:

$$\begin{aligned} dx(t) &= -2(1-\beta)x\,dt + dW(t) \\ x(0) &\sim N(0,1) \end{aligned} \tag{7.8}$$

where $W(t)$ is a standard Brownian motion.

We are ready to state the main result of this chapter:

Theorem 7.1.1 *In the subcritical regime, i.e. $\beta < 1$, if h is a continuous bounded function, then, when $\left|\mathbb{E}^0\left[h(X^N)\right] - \mathbb{E}^0\left[h(x)\right]\right| \le O\left(\frac{1}{N^\gamma}\right)$, with $\gamma > 0$,*

$$\lim_{N\to+\infty} \sup_{t\in[0,+\infty)} \left|\mathbb{E}^t\left[h(X^N)\right] - \mathbb{E}^t\left[h(x)\right]\right| = 0. \tag{7.9}$$

Before starting with the proof we try to explain its strategy.

Step 1: We consider the process Y^N, with $Y^N(0) = X^N(0)$ and infinitesimal generator $\bar{\mathscr{L}}_N$ obtained from $\mathscr{L}_N$ by linearizing the transition rates:

$$\bar{\mathscr{L}}_N f(y) = d(y,+)\nabla^+ f(y) + d(y,-)\nabla^- f(y), \tag{7.10}$$

on this hypothesis only on one point. See Remark after Proposition 7.2.7. We believe the proof can be generalized to $\mathbb{E}^0\left(\eta_1(0)\right) = m_0$, $m_0 \in [-1,1]$.

where

$$d(y,+) = \frac{N}{2}\left(1-(1-\beta)\frac{y}{\sqrt{N}}\right)\chi_{\{y<\sqrt{N}\}}, \tag{7.11}$$

and

$$d(y,-) = \frac{N}{2}\left(1+(1-\beta)\frac{y}{\sqrt{N}}\right)\chi_{\{y>\sqrt{N}\}}, \tag{7.12}$$

are the linearization of $c(x,+)$ and $c(x,-)$ around $x=0$. We prove the following:

Proposition 7.1.2 *In the same hypothesis of Theorem 7.1.1,*

$$\lim_{N\to\infty}\sup_{t\in[0,+\infty)}\left|\mathbb{E}^t\left[h(Y^N)\right]-\mathbb{E}^t\left[h(x)\right]\right|=0. \tag{7.13}$$

This first step is the difficult one.

Step 2: In the second part of the proof we prove:

Proposition 7.1.3 *In the same hypothesis of Theorem 7.1.1,*

$$\lim_{N\to\infty}\sup_{t\in[0,+\infty)}\left|\mathbb{E}^t\left[h(Y^N)\right]-\mathbb{E}^t\left[h(X^N)\right]\right|=0. \tag{7.14}$$

The second step follows showing that the L^1 distance between X^N and Y^N is uniformly small in time.
It is clear that (7.13) plus (7.14) implies (7.9).

7.2 Proof of Theorem 7.1.1: step1

7.2.1 Ornstein-Ulembeck equation for the empirical fluctuations

We derive an Ornstein-Ulembeck type equation for the empirical fluctuations of the Curie-Weiss model: i.e. $X^N(t)=\frac{1}{\sqrt{N}}\sum_{i=1}^N\eta_i$. The resulting diffusion equation will be our limit process.

Proposition 7.2.1 *The process $x(t)$ is the weak limit of $X^N(t)$, and it obeys to the diffusion equation:*

$$dx(t) = -2(1-\beta)x(t)dt + dW(t). \tag{7.15}$$

Proof: Consider the infinitesimal generator of the process X^N:

$$\mathscr{L}_N f(x) \quad = \quad \frac{N}{2}\left(1-\frac{x}{\sqrt{N}}\right)e^{\beta\frac{x}{\sqrt{N}}}\nabla^+ f(x) \quad + \quad \frac{N}{2}\left(1+\frac{x}{\sqrt{N}}\right)e^{-\beta\frac{x}{\sqrt{N}}}\nabla^- f(x) \quad (7.16)$$

Here we make a Taylor expansion of the generator to find its limit as N grows to infinity. The limit is the generator of a diffusion. The process described by the diffusion equations is the weak limit of X^N. Actually, we are using Theorem 5.5 [32]. For the Taylor expansion we have, for $f \in \mathscr{C} = \mathscr{C}^2$-functions with compact support:

$$\begin{aligned}\mathscr{L}_N f(x) = N\left(1-\frac{x}{\sqrt{N}}(1-\beta)\right)&\left(\frac{1}{\sqrt{N}}\frac{\partial f}{\partial x}(x)+\frac{1}{N}\frac{\partial^2 f}{\partial x^2}(x)\right)\\ +N\left(1+\frac{x}{\sqrt{N}}(1-\beta)\right)&\left(-\frac{1}{\sqrt{N}}\frac{\partial f}{\partial x}(x)+\frac{1}{N}\frac{\partial^2 f}{\partial x^2}(x)\right)+o(1)\end{aligned}$$

$$= -2(1-\beta)x\frac{\partial f}{\partial x}(x)+2\frac{\partial^2 f}{\partial x^2}(x)+o(1), \quad (7.17)$$

where the terms "$o(1)$" converge to zero uniformly in x. We obtain the following limit:

$$Lf(x) := \lim_{N\to\infty} L_N f(x) = -2(1-\beta)x\frac{\partial f}{\partial x}(x)+2\frac{\partial^2 f}{\partial x^2}(x). \quad (7.18)$$

Because of Theorem 5.5, this is sufficient to prove that the weak limit of X^N is the process $x(t)$ obeying the following linear diffusion equation:

$$dx = -2(1-\beta)x(t)dt + 2dW(t), \quad (7.19)$$

where $W(t)$, is a standard Brownian motion. The fact that $X^N(0)$ converges in distribution to a $N(0,1)$ comes from our assumptions on the initial condition and the standard Cenrtal Limit Theorem. □

In the sequel we need the ordinary differential equation for the characteristic function $\varphi(u,t) = \mathbb{E}\left(e^{iux_t}\right)$ for x_t.

Corollary 7.2.2 *Set* $\varphi(u,t) = \mathbb{E}^t\left(e^{iux_t}\right)$ *then*

$$\frac{d}{dt}\varphi(u,t) = -2(1-\beta)\frac{\partial}{\partial u}\varphi(u,t) - u^2\varphi(u,t) = A\varphi(u,t). \quad (7.20)$$

Proof: Now, $\frac{d}{dt}\varphi(u,t) = \frac{d}{dt}\mathbb{E}(e^{iux_t}) = \mathbb{E}(\frac{d}{dt}e^{iux_t})$. Using the Ito's formula one gets:

$$d(e^{iux_t}) = iue^{iux_t}(-2(1-\beta))x_t dt - u^2 e^{iux_t} dt + d\mathscr{M}, \tag{7.21}$$

with $\mathscr{M}$ martingale. We take the mean, and divide formally for dt. Moreover, notice that $\mathbb{E}(iue^{iux_t}x_t)$ is equal to $\frac{\partial}{\partial u}\varphi(u,t)$. Thus:

$$\frac{d}{dt}\varphi(u,t) = -2(1-\beta)\frac{\partial}{\partial u}\varphi(u,t) - u^2\varphi(u,t) = \mathrm{A}\varphi(u,t). \tag{7.22}$$

□

7.2.2 Uniform distance for distribution functions

In this section we do the first step toward the completion of the proof of the uniform fluctuation theorem. Before we continue, we need the definition of *distribution function.*

Definition 7.2.3 *Consider the process Y_t^N defined in (7.10). The function*

$$F_{Y_t^N}(x) = \mathbb{P}^t\left(Y_t^N \le x\right) \tag{7.23}$$

is called the distribution function for the process Y_t^N. In the same way, $F_{x_t}(x)$ is the distribution function for the diffusion $x(t)$.

The aim of this section is to prove that:

Theorem 7.2.4 *The following uniform bound holds:*

$$\sup_{t\in[0,+\infty)} \sup_x \left|F_{Y_t^N}(x) - F_{x_t}(x)\right| \le O\left(\frac{1}{N^{1/12}}\right). \tag{7.24}$$

Set $\varphi(u,t) = \mathbb{E}^t\left(e^{iux_t}\right)$ and $\varphi_N(u,t) = \mathbb{E}^t\left(e^{iuY_t^N}\right)$, then the key point is to prove the following

Proposition 7.2.5 *Define:*

$$D_N(t) := \int_{-\infty}^{+\infty}\left(\frac{\varphi(u,t)-\varphi_N(u,t)}{u}\right)^2 du. \tag{7.25}$$

Then

$$\sup_{t\in[0,+\infty)} D_N(t) \le O\left(\frac{1}{N^{1/4}}\right). \tag{7.26}$$

Once we get this result, we use the Essen's Inequality [31], that we recall here.

Theorem 7.2.6 *Let $F(x)$ e $G(x)$ be distribution functions with characteristic functions $f(u)$ and $g(u)$. Moreover, let $G(x)$ have a finite derivative $G'(x)$ for every x. Then for every $R > 0$*

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \leq \frac{2}{\pi} \int_0^R \left| \frac{f(u) - g(u)}{u} \right| du + \frac{24}{\pi R} \sup_x \left| G'(x) \right|. \tag{7.27}$$

This, together with Proposition 7.2.5 allows us to conclude. Indeed, in our case, (7.27) gives:

$$\sup_{x\in\mathbb{R}} \left| F_{Y_t^N} - F_{x_t}(x) \right| \leq \frac{2}{\pi} \int_0^R \left| \frac{\varphi(u,t) - \varphi_N(u,t)}{u} \right| du + \frac{24}{\pi R} \sup_x \left| F'_{x_t}(x) \right|$$

$$\underbrace{\leq}_{\text{Holder's Inequality}} \frac{\sqrt{R}}{2\pi} (D_N(t))^{\frac{1}{2}} + \frac{24}{\pi R} \sup_x \left| F'_{x_t}(x) \right| \underbrace{\leq}_{\text{Proposition 7.2.5}} \frac{\sqrt{R}}{2\pi} O\left(\frac{1}{N^{\frac{\gamma}{2}}} \right) + \frac{24}{\pi R} \sup_x \left| F'_{x_t}(x) \right|. \tag{7.28}$$

Moreover, our assumption on the initial condition (i.e. $\eta_i(0)$, $i = 1, \ldots, N$ i.i.d. with $\mathbb{E}(\eta_1(0)) = 0$) implies

$$F'_{x_t}(x) = \frac{1}{\sqrt{2\pi \mathbb{E}^t(x^2(t))}} e^{-\frac{x}{2\mathbb{E}^t(x^2(t))}}. \tag{7.29}$$

From (7.15), using Ito's formula one can derive the following ordinary differential equation

$$\frac{d}{dt} \mathbb{E}^t(x^2(t)) = -4(1-\beta)\mathbb{E}^t(x^2(t)) + 4, \tag{7.30}$$

with initial condition

$$\mathbb{E}^0(x^2(0)) = 1. \tag{7.31}$$

The differential equation (7.30) has the solution:

$$\mathbb{E}^t(x^2(t)) = \frac{1 + (1-\beta)}{1-\beta} e^{-4(1-\beta)t} - \frac{1}{1-\beta}. \tag{7.32}$$

This finally prove that

$$\sup_{t\in[0,+\infty)} \sup_x \left| F'_{x_t}(x) \right| < +\infty. \tag{7.33}$$

Now, choosing $R = N^{1/12}$, with $\alpha < \gamma$, we prove Theorem 7.2.4, since the r.h.s. of equation (7.28) does not depend on t.

To prove Proposition 7.2.5 we need to show the probability for Y_t^N to stay at the boundary is exponentially small in N uniformly in time when N grows. More precisely we prove:

Proposition 7.2.7 *If $\eta_i(0)$, $i = 1,\dots,N$ are i.i.d. random variables with $\mathbb{E}^0\left(\eta_1(0)\right)$, then*

$$\sup_{t\in(0,+\infty]} \mathbb{P}^t\left(\left|Y_t^N\right| = \sqrt{N}\right) \le a^N \tag{7.34}$$

with $a < 1$.

Proof: We need some lemmas. Set $S_t^N = \sqrt{N}Y_t^N$. We consider S_t^N to make simpler the computations involved in the proof of the lemmas, but it should be clear that the conclusion holds for Y_t^N, too.

We have $S^N \in \{-N,\dots,N\}$ and the possible jumps are

$$\begin{array}{lclll} S^N & \to & S^N + 2 & \text{at rate} & \frac{N}{2}(1-\frac{\rho S^N}{N})\chi_{\{S^N\neq N\}} \\ S^N & \to & S^N - 2 & \text{at rate} & \frac{N}{2}(1+\frac{\rho S^N}{N})\chi_{\{S^N\neq -N\}} \end{array}$$

with $\rho = 1-\beta \in (0,1)$. We call L_ρ the infinitesimal generator of this process.

The lemmas we need are the following.

Lemma 7.2.8 *The reversible measure μ_N^ρ, for the process S^N, is*

$$\mu_N^\rho(m) = \frac{1}{Z_N}\frac{1}{\Gamma\left(\frac{N}{2\rho}-\frac{m}{2}+1\right)\Gamma\left(\frac{N}{2\rho}+\frac{m}{2}+1\right)}, \tag{7.35}$$

where Γ is the Euler's function and Z_N is a normalization factor.

Proof: The measure $\mu_N^\rho(m)$ satisfies the detailed balance condition:

$$\left(1-\rho\frac{m}{N}\right)\frac{1}{\Gamma\left(\frac{N}{2\rho}-\frac{m}{2}+1\right)\Gamma\left(\frac{N}{2\rho}+\frac{m}{2}+1\right)}$$

$$= \left(1+\rho\frac{m+2}{N}\right)\frac{1}{\Gamma\left(\frac{N}{2\rho}-\frac{m+2}{2}+1\right)\Gamma\left(\frac{N}{2\rho}+\frac{m+2}{2}+1\right)}. \tag{7.36}$$

To see this, it suffices to use the identity $\Gamma(z+1) = z\Gamma(z)$. □

Lemma 7.2.9 *By Stirling's formula, when N is large:*

$$\mu_N^\rho(N) \leq Ca^N \tag{7.37}$$

with

$$a = \left[\left(\frac{\rho^2}{1-\rho^2}\right)^{1/\rho}\left(\frac{1-\rho}{1+\rho}\right)\right]^{\frac{1}{2}}. \tag{7.38}$$

Proof: For N large, use the Stirling's approximation for the Gamma function $\Gamma(z) \simeq (\sqrt{2z\pi}\frac{z^z}{e^z})$, with $\Re(z) > 0$:

$$\frac{1}{\Gamma\left(\frac{N}{2\rho} - \frac{1}{2N} + 1\right)}\frac{1}{\Gamma\left(\frac{N}{2\rho} + \frac{1}{2N} + 1\right)}$$

$$\simeq \frac{1}{2\pi}\frac{1}{\sqrt{\frac{N}{2}\frac{1-\rho}{\rho}}}\frac{1}{\sqrt{\frac{N}{2}\frac{1+\rho}{\rho}}}\frac{\exp\left(\frac{N}{2}\frac{1-\rho}{\rho}\right)}{\left(\frac{N}{2}\frac{1-\rho}{\rho}\right)^{\frac{N}{2}\frac{1-\rho}{\rho}}}\frac{\exp\left(\frac{N}{2}\frac{1+\rho}{\rho}\right)}{\left(\frac{N}{2}\frac{1+\rho}{\rho}\right)^{\frac{N}{2}\frac{1+\rho}{\rho}}}$$

$$= \frac{1}{\pi}\frac{1}{N}\sqrt{\frac{\rho^2}{1-\rho^2}}\frac{\exp\left(\frac{N}{\rho}\right)}{\left(\frac{N}{2}\right)^{\frac{N}{\rho}}}\left(\frac{\rho^2}{1-\rho^2}\right)^{\frac{N}{2\rho}}\underbrace{\left(\frac{\rho}{1-\rho}\right)^{-\frac{N}{2}}\left(\frac{\rho}{1+\rho}\right)^{\frac{N}{2}}}_{=\left(\frac{1-\rho}{1+\rho}\right)^{\frac{N}{2}}}$$

$$= \frac{1}{\pi}\frac{\exp\left(\frac{N}{\rho}\right)}{\left(\frac{N}{2}\right)^{\frac{N}{\rho}+1}}\left(\frac{\rho^2}{1-\rho^2}\right)^{\frac{N}{2\rho}+\frac{1}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{N}{2}}. \tag{7.39}$$

Since, $\lim_{N\to\infty}\frac{\exp\left(\frac{N}{\rho}\right)}{\left(\frac{N}{2}\right)^{\frac{2N}{\rho}+1}} = 0$, for N large,

$$\frac{1}{\Gamma\left(\frac{N}{2\rho} - \frac{N}{2} + 1\right)}\frac{1}{\Gamma\left(\frac{N}{2\rho} + \frac{N}{2} + 1\right)} \leq C\underbrace{\left[\left(\frac{\rho^2}{1-\rho^2}\right)^{\frac{1}{\rho}}\left(\frac{1-\rho}{1+\rho}\right)\right]^{\frac{N}{2}}}_{=a(\rho)^N}, \tag{7.40}$$

where

$$a(\rho) = \left[\left(\frac{\rho^2}{1-\rho^2} \right)^{\frac{1}{\rho}} \left(\frac{1-\rho}{1+\rho} \right) \right]^{1/2}.$$

It remains to show that $a(\rho) < 1$ for $\rho \in (0,1)$. By direct computations we get

$$\rho \log a(\rho) = \log \rho - \frac{1}{2}[(1-\rho)\log(1-\rho) + (1+\rho)\log(1+\rho)].$$

By strict convexity of the function $x \mapsto x \log x$, we have

$$\frac{1}{2}[(1-\rho)\log(1-\rho) + (1+\rho)\log(1+\rho)] > 0;$$

thus $\rho \log a(\rho) < 1$ which implies $a(\rho) < 1$. □

Lemma 7.2.10 *(Stochastic Domination) Let S_i^N $i = 1,2$ be the processes with infinitesimal generators:*

$$L_{\rho_i} f(s) = \frac{N}{2}(1 - \frac{\rho_i s}{N})\chi_{\{s \neq N\}} \nabla^+ f(s) + \frac{N}{2}(1 + \frac{\rho_i s}{N})\chi_{\{s \neq -N\}} \nabla^- f(s), \quad (7.41)$$

with $\nabla^{\pm} f(s) = f(s \pm 2) - f(s)$. Assume that $\rho_1 \leq \rho_2$ then, there exists a coupling such that ,if $|S_2^N(0)| \leq |S_1^N(0)|$, then $|S_2^N(t)| \leq |S_1^N(t)|$.

Proof: We show that it is possible to construct a coupling in such a way that the inequality $|S_2^N(0)| \leq |S_1^N(0)|$ is preserved by the dynamics (i.e. $|S_2^N(t)| \leq |S_1^N(t)|$, $\forall t > 0$).

We do the following coupling:

$$\Omega f(s_1,s_2) = (\Omega_1 f(s_1,s_2))\chi_{\{s_1 s_2 > 0 \vee s_1 = s_2 = 0\}} + (\Omega_2 f(s_1,s_2))\chi_{\{s_1 s_2 < 0\}} + (\Omega_3 f(s_1,s_2))\chi_{\{s_2 = 0, s_1 \neq 0\}}. \quad (7.42)$$

Let us explain the terms appearing in (7.42). We have:

$$\Omega_3 f = L_{\rho_1} f + L_{\rho_2} f, \quad (7.43)$$

where L_{ρ_i} is meant to act only on the variable s_i.

$$\Omega_1 f(s_1,s_2) = c_1(s_1,+)\nabla^{s_1,+}f(s_1,s_2) + c_2(s_2,+)\nabla^{s_2,+}f(s_1,s_2)$$
$$+\min\left(c_1(s_1,+),c_2(s_2,+)\right)\left(\nabla^{s_1,s_2,+,+}f(s_1,s_2) - \nabla^{s_1,+}f(s_1,s_2) - \nabla^{s_2,+}f(s_1,s_2)\right)$$
$$+\,c_1(s_1,+)\nabla^{s_1,+}f(s_1,s_2) + c_2(s_2,+)\nabla^{s_2,+}f(s_1,s_2)$$
$$+\min\left(c_1(s_1,-),c_2(s_2,-)\right)\left(\nabla^{s_1,s_2,-,-}f(s_1,s_2) - \nabla^{s_1,-}f(s_1,s_2) - \nabla^{s_2,-}f(s_1,s_2)\right), \quad (7.44)$$

$$\Omega_2 f(s_1,s_2) = c_1(s_1,+)\nabla^{s_1,+}f(s_1,s_2) + c_2(s_2,-)\nabla^{s_2,-}f(s_1,s_2)$$
$$+\min\left(c_1(s_1,+),c_2(s_2,-)\right)\left(\nabla^{s_1,s_2,+,-}f(s_1,s_2) - \nabla^{s_1,+}f(s_1,s_2) - \nabla^{s_2,-}f(s_1,s_2)\right)$$
$$+\,c_1(s_1,-)\nabla^{s_1,-}f(s_1,s_2) + c_2(s_2,+)\nabla^{s_2,+}f(s_1,s_2)$$
$$+\min\left(c_1(s_1,-),c_2(s_2,+)\right)\left(\nabla^{s_1,s_2,-,+}f(s_1,s_2) - \nabla^{s_1,-}f(s_1,s_2) - \nabla^{s_2,+}f(s_1,s_2)\right), \quad (7.45)$$

where
$$c_i(s,\pm) = \frac{N}{2}(1\mp\rho_i s)\chi_{\{s\neq\pm N\}}$$
and
$$\nabla^{s_1,s_2,\pm,\pm}f(s_1,s_2) = f(s_1\pm 2, s_2\pm 2) - f(s_1,s_2).$$

The important thing to notice is that when $|S_1^N| = |S_2^N|$ a jump of the process S_2^N toward a greater modulo, obligates S_1^N to do the same. Thus, $\left|S_1^N(t)\right| \geq \left|S_2^N(t)\right|$ for every $t>0$, provided that we start from $\left|S_1^N(0)\right| \geq \left|S_2^N(0)\right|$. □

This implies that $\mu_N^{\rho_2} \lesssim \mu_N^{\rho_1}$, where the pseudo-order $\lesssim$ means:
$$\int f(|s|)\mu_N^{\rho_2}(ds) \leq \int f(|s|)\mu_N^{\rho_1}(ds)$$
for any f increasing. Note that the initial condition corresponding to the symmetric spins is μ_N^1. Thus, for every $t\geq 0$, we have $\mu_N^1 e^{tL_\rho} \lesssim \mu_N^\rho$. This implies that
$$\mathbb{P}\left(|Y_t^N| = \sqrt{N}\right) = \int \chi_{\{|s|\geq N\}}\mu_N^1 e^{tL_\rho}(ds) \leq \int \chi_{\{|s|\geq N\}}\mu_N^\rho(ds) = \mu_N^\rho(\{|s|=N\}),$$
and so it decays exponentially in N, uniformly in t. □

Remark: Our hypothesis on the initial condition for $\eta_i(0)$ $i=1,\ldots,N$ is fundamental for the proof of Proposition 7.2.7, but this is the only point in our line of reasoning. We believe that this assumption can be relaxed to a non-symmetric one: $\eta_i(0)$ $i=1,\ldots,N$ i.i.d. with $\mathbb{E}^0\left(\eta_1(0)\right)\neq 0$.

7.2.3 Proof of Proposition 7.2.5

First, we have to know the differential equation for $\varphi_N(u,t) = \mathbb{E}\left(e^{iuY_t^N}\right)$. Remember that $\frac{d}{dt}\varphi_N(u,t) = \mathbb{E}\left(L_\rho e^{iuY_t^N}\right)$. We compute:

$$\begin{aligned}
L_\rho e^{itY_t^N} = \frac{N}{2}\left(1-\rho\frac{Y_t^N}{\sqrt{N}}\right)\chi_{(Y_t^N<\sqrt{N})}e^{iuY_t^N}\left(e^{\frac{2iu}{\sqrt{N}}}-1\right) \\
+\frac{N}{2}\left(1+\rho\frac{Y_t^N}{\sqrt{N}}\right)\chi_{(Y_t^N>-\sqrt{N})}e^{iuY_t^N}\left(e^{\frac{-2iu}{\sqrt{N}}}-1\right) \\
=\frac{N}{2}e^{iuY_t^N}\left(e^{\frac{2iu}{\sqrt{N}}}+e^{\frac{-2iu}{\sqrt{N}}}-2\right)+\rho e^{iuY_t^N}\left(e^{\frac{2iu}{\sqrt{N}}}-e^{\frac{-2iu}{\sqrt{N}}}\right) \\
-\frac{N}{2}\left(1-\rho\frac{Y_t^N}{\sqrt{N}}\right)e^{iuY_t^N}\left(e^{\frac{2iu}{\sqrt{N}}}-1\right)\chi_{\{Y_t^N=\sqrt{N}\}} \\
-\frac{N}{2}\left(1+\rho\frac{Y_t^N}{\sqrt{N}}\right)e^{iuY_t^N}\left(e^{-\frac{2iu}{\sqrt{N}}}-1\right)\chi_{\{Y_t^N=-\sqrt{N}\}}. \quad (7.46)
\end{aligned}$$

Set:

$$\begin{aligned}
R(u) = -\frac{N}{2}\left(1-\rho\frac{Y_t^N}{\sqrt{N}}\right)e^{iuY_t^N}\left(e^{\frac{2iu}{\sqrt{N}}}-1\right)\chi_{\{Y_t^N=\sqrt{N}\}} \\
-\frac{N}{2}\left(1+\rho\frac{Y_t^N}{\sqrt{N}}\right)e^{iuY_t^N}\left(e^{-\frac{2iu}{\sqrt{N}}}-1\right)\chi_{\{Y_t^N=-\sqrt{N}\}}. \quad (7.47)
\end{aligned}$$

Thus we have:

$$\begin{aligned}
\frac{\partial}{\partial t}\varphi_N(u,t) = \frac{N}{2}\left(e^{i\frac{2u}{\sqrt{N}}}+e^{-i\frac{2u}{\sqrt{N}}}-2\right)\varphi_N(u,t) \\
-\rho\frac{\sqrt{N}}{2i}\left(e^{i\frac{2u}{\sqrt{N}}}+e^{-i\frac{2u}{\sqrt{N}}}\right)\frac{\partial}{\partial u}\varphi_N(u,t)+\bar{R}(u) \\
= N\left(\cos\left(\frac{2u}{\sqrt{N}}\right)-1\right)\varphi_N(u,t) \\
-\rho\sqrt{N}\sin\left(\frac{2u}{\sqrt{N}}\right)\frac{\partial}{\partial u}\varphi_N(u,t)+\bar{R}(u) =: A_N\varphi_N(u,t)+\bar{R}(u), \quad (7.48)
\end{aligned}$$

where $\bar{R}(u) = \mathbb{E}[R(u)]$.

We try to bound the following distance between $\varphi(u,t)$ and $\varphi_N(u,t)$. Define

$$D_N(t) := \int_{-\infty}^{+\infty} \left(\frac{\varphi(u,t) - \varphi_N(u,t)}{u} \right)^2 du \tag{7.49}$$

We divide this in two pieces:

$$D_N(t) = \int_{-N^{1/4}}^{N^{1/4}} \left(\frac{\varphi(u,t) - \varphi_N(u,t)}{u} \right)^2 du + \int_{\{|u|>N^{1/4}\}} \left(\frac{\varphi(u,t) - \varphi_N(u,t)}{u} \right)^2 du \tag{7.50}$$

Since $|\varphi(u,t)| < 1$ the second term in (7.50) is easy to control and it is $O\left(\frac{1}{N^{\frac{1}{4}}}\right)$. For the first one, we use again the Gronwall's Lemma. We set

$$\bar{D}_N(t) := \int_{-N^{1/4}}^{N^{1/4}} \left(\frac{\varphi(u,t) - \varphi_N(u,t)}{u} \right)^2 du. \tag{7.51}$$

We have

$$\begin{aligned} \frac{d}{dt}\bar{D}_N(t) = \int_{-N^{1/4}}^{N^{1/4}} \big(A(\varphi(u,t) - \varphi_N(u,t)) \\ +(A_N - A)\varphi(u,t) + \bar{R}(u) \big) \frac{\big(\varphi(u,t) - \varphi_N(u,t)\big)}{u^2} 2du = \end{aligned} \tag{7.52}$$

$$\begin{aligned} &\int_{-N^{1/4}}^{N^{1/4}} \left[-u^2((\varphi(u,t) - \varphi_N(u,t)) - 2\rho \frac{\partial}{\partial u}(\varphi(u,t) - \varphi_N(u,t)) \right] \frac{\big(\varphi(u,t) - \varphi_N(u,t)\big)}{u^2} 2du \\ &+ \int_{-N^{1/4}}^{N^{1/4}} \Big[(N(\cos(2u/\sqrt{N}) - 1) + u^2)\varphi(u,t) \\ &-(\rho\sqrt{N}\sin(2u/\sqrt{N}) - 2\rho u)\frac{\partial}{\partial u}\varphi(u,t) \Big] \frac{\big(\varphi(u,t) - \varphi_N(u,t)\big)}{u^2} 2du \\ &+ \int_{-N^{1/4}}^{N^{1/4}} \bar{R}(u) \frac{\big(\varphi(u,t) - \varphi_N(u,t)\big)}{u^2} 2du. \end{aligned} \tag{7.53}$$

We integrate the first term by parts. We obtain:

$$
\begin{aligned}
\frac{d}{dt}\bar{D}_N(t) = -2\int_{-N^{1/4}}^{N^{1/4}} \big(\varphi(u,t)-\varphi_N(u,t)\big)^2\, du + \left[\big(\varphi(u,t)-\varphi_N(u,t)\big)^2 \frac{-2\rho}{u}\right]_{-N^{1/4}}^{N^{1/4}} \\
-2\rho\int_{-N^{1/4}}^{N^{1/4}} \left(\frac{\varphi(u,t)-\varphi_N(u,t)}{u}\right)^2 du + \int_{-N^{1/4}}^{N^{1/4}} \Big[(N(\cos(2u/\sqrt{N})-1)+u^2)\varphi(u,t) \\
-(\rho\sqrt{N}\sin(2u/\sqrt{N})-2\rho u)\frac{\partial}{\partial u}\varphi(u,t)\Big] \frac{\big(\varphi(u,t)-\varphi_N(u,t)\big)}{u^2} 2du \\
+\int_{-N^{1/4}}^{N^{1/4}} \bar{R}(u)\frac{\big(\varphi(u,t)-\varphi_N(u,t)\big)}{u^2} 2du. \quad (7.54)
\end{aligned}
$$

The first two terms coming from integration by parts are negative, while the third one is $-2\rho\bar{D}_N(t)$, so

$$
\begin{aligned}
\frac{d}{dt}\bar{D}_N(t) \le -2\rho\bar{D}_N(t) \\
+\int_{-N^{1/4}}^{N^{1/4}} \Big|(N(\cos(2u/\sqrt{N})-1)+u^2)\varphi(u,t) \\
-(\rho\sqrt{N}\sin(2u/\sqrt{N})-2\rho u)\frac{\partial}{\partial u}\varphi(u,t)\Big| \frac{\big|\varphi(u,t)-\varphi_N(u,t)\big|}{u^2} 2du \\
+\int_{-N^{1/4}}^{N^{1/4}} |\bar{R}(u)|\frac{\big|\varphi(u,t)-\varphi_N(u,t)\big|}{u^2} 2du. \quad (7.55)
\end{aligned}
$$

Now in oder to apply the Gronwall's Lemma we have to show that:

$$
\begin{aligned}
\int_{-N^{1/4}}^{N^{1/4}} \Big|(N(\cos(2u/\sqrt{N})-1)+u^2)\varphi(u,t) \\
-(\rho\sqrt{N}\sin(2u/\sqrt{N})-2\rho u)\frac{\partial}{\partial u}\varphi(u,t)\Big| \frac{\big|\varphi(u,t)-\varphi_N(u,t)\big|}{u^2} 2du \le O\left(\frac{1}{N^{1/4}}\right), \quad (7.56)
\end{aligned}
$$

and

$$
\int_{-N^{1/4}}^{N^{1/4}} |\bar{R}(u)|\frac{\big|\varphi(u,t)-\varphi_N(u,t)\big|}{u^2} 2du \le O\left(\frac{1}{N^{1/4}}\right). \quad (7.57)
$$

To control these terms, note that, for some $B,\ C>0$,

$$
\cos(2u/\sqrt{N})-1+\left(\frac{u}{\sqrt{N}}\right)^2 \le B\left(\frac{u}{\sqrt{N}}\right)^4 \quad (7.58)
$$

$$
\sin(2u/\sqrt{N})-\frac{u}{\sqrt{N}} \le C\left(\frac{|u|}{\sqrt{N}}\right)^3 \quad (7.59)
$$

Hence, (7.56) is less than:

$$K\int_{-N^{1/4}}^{N^{1/4}}\frac{1}{u^2}\left[2NB\frac{u^4}{N^2}+2\rho\sqrt{N}\frac{|u|^3}{\sqrt{N}N}\mathbb{E}\left(\left|\frac{\partial}{\partial u}\varphi(u,t)\right|\right)\right]du. \tag{7.60}$$

It is easy to see that the last expression is $O\left(\frac{1}{N^{\frac{1}{4}}}\right)$. One has only to notice that $\mathbb{E}\left(\left|\frac{\partial}{\partial u}\varphi(u,t)\right|\right)\leq \sqrt{\mathbb{E}\left[\left(x_t^2\right)^2\right]}<M$ when $\beta<1$ (see (5.12)).

The next problem is to control (7.57). Since we have observed that $\mathbb{P}(Y_t^N=\pm\sqrt{N})$ is exponentially small in N uniformly in t, it follows from the definition of $R(u)$ that $\frac{\bar{R}(u)}{|u|}$ is exponentially small in N, uniformly in u and t. Moreover

$$\frac{\left|\varphi(u,t)-\varphi_N(u,t)\right|}{|u|}\leq \mathbb{E}(|x_t-Y_t^N|)\leq C\sqrt{N}.$$

Putting all together we have that (7.57) is exponentially small in N uniformly in t.

This completes the proof of Theorem 7.2.5.

7.2.4 Proof of Proposition 7.1.2

The proof of the Proposition 7.1.2 follows from the next Proposition:

Proposition 7.2.11 *If, for some* $\gamma>0$*,*

$$\sup_t\left|F_{Y_t^N}(x)-F_{x_t}(x)\right|\leq O\left(\frac{1}{N^\gamma}\right) \tag{7.61}$$

then

$$\lim_{N\to\infty}\sup_t\mathbb{E}^t\left(\left|h(Y_t^N(x))-h(x_t(x))\right|\right)=0 \tag{7.62}$$

Proof: Let be r and s such that $\mathbb{P}(X_t\notin(r,s])=1-F_{x_t}(s)+F_{x_t}(r)\leq\epsilon$. We can assume that r and s do not depend on t because X_t is an Ornstein-Ulembeck process. Notice that:

$$\begin{aligned}\mathbb{P}(Y_t^N\notin(r,s])&=1-F_{Y_t^N}(s)+F_{Y_t^N}(r)\\&=1-(F_{Y_t^N}(s)-F_{x_t}(s))+(F_{Y_t^N}(r)-F_{x_t}(r))-F_{x_t}(s)+F_{x_t}(r)\leq O\left(\frac{1}{N^\gamma}\right)+\epsilon.\end{aligned} \tag{7.63}$$

The r.h.s. does not depend on t, thus

$$\sup_t \mathbb{P}(x_t \notin (r,s]) \leq O\left(\frac{1}{N^\gamma}\right) + \epsilon. \tag{7.64}$$

Now, since $[r,s]$ is compact h is uniformly continuous on $[r,s]$ and there exists a partition $r = r_1, \ldots, r_k = s$, such that $|h(x) - h(r_j)| \leq \epsilon$ when $x \in [r_{j-i}, r_i]$. We let k grow as N^α, with $\alpha < \gamma$, to make ϵ small.

Moreover, define

$$g(x) = \sum_{i=1}^{k} h(r_i)\chi_{(r_{i-1},r_i]}, \tag{7.65}$$

and write:

$$\left|\mathbb{E}^t\left(h(Y_t^N) - h(x_t)\right)\right| \leq \left|\mathbb{E}^t\left(h(Y_t^N) - g(Y_t^N)\right)\right| + \\ \left|\mathbb{E}^t\left(g(Y_t^N) - g(x_t)\right)\right| + \left|\mathbb{E}^t\left(g(x_t) - h(x_t)\right)\right|. \tag{7.66}$$

For the first term:

$$\left|\mathbb{E}^t\left(h(Y_t^N) - g(Y_t^N)\right)\right| = \left|\mathbb{E}^t\left(h(Y_t^N) - g(Y_t^N)\right)\chi_{Y_t^N \in (r,s]}\right| + \left|\mathbb{E}^t\left(h(Y_t^N) - g(Y_t^N)\right)\chi_{Y_t^N \notin (r,s]}\right|$$

$$\leq \epsilon + 2\|h\|_\infty \mathbb{P}^t\left(Y_t^N \notin (r,s]\right) \leq \epsilon\,(1 + 2\|h\|_\infty) + O\left(\frac{1}{N^\gamma}\right) \tag{7.67}$$

Since the r.h.s. does not depend on t we conclude:

$$\sup_t \left|\mathbb{E}^t\left(h(Y_t^N) - g(Y_t^N)\right)\right| \leq \epsilon\,(1 + 2\|h\|_\infty) + O\left(\frac{1}{N^\gamma}\right). \tag{7.68}$$

The same line of reasoning for the third term produce:

$$\sup_t \left|\mathbb{E}^t\left(g(x_t) - h(x_t)\right)\right| \leq \epsilon\,(1 + 2\|h\|_\infty). \tag{7.69}$$

For the second term of (7.66) we compute:

$$\sup_t \left|\mathbb{E}^t\left(g(Y_t^N) - g(x_t)\right)\right|$$

$$= \sup_t \sum_{i=1}^{k} h(r_i)[\underbrace{(F_{Y_t^N}(r_i) - F_{x_t}(r_i))}_{\leq O\left(\frac{1}{N^\gamma}\right)} - \underbrace{(F_{Y_t^N}(r_{i-1}) - F_{x_t}(r_{i-1}))}_{\leq O\left(\frac{1}{N^\gamma}\right)}] \leq 2N^\alpha \|h\|_\infty O\left(\frac{1}{N^\gamma}\right). \tag{7.70}$$

So, for N sufficiently large we have:

$$\sup_t \mathbb{E}^t\left(\left|h(Y_t^N(x)) - h(x_t(x))\right|\right) \leq \|h\|_\infty O\left(\frac{1}{N^{\gamma-\alpha}}\right) + \epsilon(1 + \|h\|_\infty), \tag{7.71}$$

and, since ϵ is arbitrary, this completes the proof. □

7.3 Proof of Theorem 7.1.1: step 2

7.3.1 The linearization of the Curie-Weiss model

The next step is to show that the Curie-Weiss system remains uniformly close in time to its linearization. More precisely, we compare the former one with the spin-flip system where we linearize the transition rates. For our aims the quantity of interest is the empirical fluctuations so we investigate the time evolution of $X^N = \frac{1}{\sqrt{N}}\sum_{i=1}\eta_i$ under the Curie-Weiss dynamics and its "linearization" Y^N. We recall that X^N and Y^N are markovian, with generators respectively $\mathscr{L}$ and $\bar{\mathscr{L}}$ defined in (7.3) and (7.10).

We prove the following

Proposition 7.3.1 *When $\beta < 1$ there is a probability space where both the processes X^N and Y^N can be realized and*

$$\sup_{t\in[0,+\infty)} \mathbb{E}^t\left(|X^N - Y^N|\right) < O\left(\frac{1}{\sqrt{N}}\right). \tag{7.72}$$

Proof: The technique is always the same: first, we use coupling to realize the two processes in the same probability space and next, the Gronwall's Lemma helps us to bound $\mathbb{E}^t\left(|X^N - Y^N|\right)$.
The infinitesimal generator of the coupling is

$$\Omega f(x,y) = \Omega^+ f(x,y) + \Omega^- f(x,y) \tag{7.73}$$

with

$$\Omega^+ f(x,y) = c(x,+)\nabla^{x,+} f(x,y) + d(y,+)\nabla^{y,+} f(x,y)$$

$$+ \min\{c(x,+), d(y,+)\}\left(\nabla^{x,y,+,+} f(x,y) - \nabla^{x,+} f(x,y) - \nabla^{y,+} f(x,y)\right) \tag{7.74}$$

$$\Omega^- f(x,y) = c(x,-)\nabla^{x,-} f(x,y) + d(y,-)\nabla^{y,-} f(x,y)$$

$$+\min\{c(x,-),d(y,-)\}\left(\nabla^{x,y,-,-} f(x,y) - \nabla^{x,-} f(x,y) - \nabla^{y,-} f(x,y)\right). \quad (7.75)$$

with

$$\nabla^{x,y,\pm,\pm} f(x,y) = f\left(x \pm \frac{2}{\sqrt{N}}, xy \pm \frac{2}{\sqrt{N}}\right) - f(x,y).$$

Remember that $\frac{d}{dt}\mathbb{E}^t(|X^N - Y^N|) = \mathbb{E}^t(\Omega|X^N - Y^N|)$. Take $f(x,y) = |x-y|$ then, we observe that:

1. $\nabla^{x,+} f(x,y) = \nabla^{y,-} f(x,y) = \frac{2}{\sqrt{N}}\text{sign}(x-y)\chi_{\{x\neq y\}} + \frac{2}{\sqrt{N}}\chi_{\{x=y\}}$,
2. $\nabla^{x,-} f(x,y) = \nabla^{y,+} f(x,y) = -\frac{2}{\sqrt{N}}\text{sign}(x-y)\chi_{\{x\neq y\}} + \frac{2}{\sqrt{N}}\chi_{\{x=y\}}$;

and, so one obtains:

$$\left(\Omega|X^N - Y^N|\right)\chi_{\{X^N\neq Y^N\}} = \Bigg\{\frac{2}{\sqrt{N}}\text{sign}(X^N - Y^N)[c(X^N,+) - d(Y^N,+)]$$

$$-\frac{2}{\sqrt{N}}\text{sign}(X^N - Y^N)[c(X^N,-) - d(Y^N,-)]\Bigg\}\chi_{\{X^N\neq Y^N\}} \quad (7.76)$$

The transition rate $d(Y^N,\cdot)$ is the linearization of $c(X^N,\cdot)$. This implies:

$$c\left(X^N,+\right) - d\left(Y^N,+\right) = d\left(X^N,+\right) + O\left[\left(X^N\right)^2\right] - d\left(Y^N,+\right)$$

$$= \frac{\sqrt{N}}{2}(1-\beta)\left(-X^N + Y^N\right) + O\left[\left(X^N\right)^2\right], \quad (7.77)$$

and that, in the same way,

$$c\left(X^N,-\right) - d\left(Y^N,-\right) = \frac{\sqrt{N}}{2}(1-\beta)\left(X^N - Y^N\right) + O\left[\left(X^N\right)^2\right]. \quad (7.78)$$

For $\left(\Omega f(X^N,Y^N)\right)\chi_{\{X^N\neq Y^N\}}$ the previous computations gives:

$$\left(\Omega f(X^N,Y^N)\right)\chi_{\{X^N\neq Y^N\}} = \left\{-2(1-\beta)|X^N - Y^N| + \frac{O\left[\left(X^N\right)^2\right]}{\sqrt{N}}\right\}\chi_{\{X^N\neq Y^N\}}. \quad (7.79)$$

A similar computation for $\left(\Omega f(X^N,Y^N)\right)\chi_{\{X^N=Y^N\}}$ gives:

$$\left(\Omega f(X^N,Y^N)\right)\chi_{\{X^N=Y^N\}} = \frac{O\left[\left(X^N\right)^2\right]}{\sqrt{N}}\chi_{\{X^N=Y^N\}}. \tag{7.80}$$

For $\beta < 1$ we know that $\mathbb{E}^t\left[\left(X^N\right)^2\right]$ is $O(1)$, so we conclude that

$$\frac{d}{dt}\mathbb{E}^t\left(|X^N - Y^N|\right) \leq -2(1-\beta)\mathbb{E}^t\left(|X^N - Y^N|\right) + O\left(\frac{1}{\sqrt{N}}\right). \tag{7.81}$$

We are in the good situation to apply the Gronwall's Lemma. This completes the proof of the theorem. □

7.3.2 Proof of Proposition 7.1.3

Proposition 7.1.3 directly follows from the proposition below.

Proposition 7.3.2 *If, for some* $\gamma > 0$,

$$\sup_t \mathbb{E}\left(\left|Y_t^N - X_t^N\right|\right) \leq O\left(\frac{1}{N^\gamma}\right), \tag{7.82}$$

then

$$\lim_{N\to\infty}\sup_t \mathbb{E}\left(\left|h(Y_t^N) - h(X_t^N)\right|\right) = 0. \tag{7.83}$$

Proof: From (7.63) of Proposition 7.2.11 we know that there exist r and s such that, $\sup_t \mathbb{P}^t(Y_N \notin (r,s]) \leq \epsilon$ for N sufficiently large. Choose δ such that $|h(x)-h(y)| \leq \epsilon$ when $|x-y| \leq \delta$, $\forall x,y \in (r,s]$, then

$$\begin{aligned}
\mathbb{E}^t\left(\left|h(Y_t^N) - h(X_t^N)\right|\right) &= \mathbb{E}^t\left(\left|h(Y_t^N) - h(X_t^N)\right| : \left|Y_t^N - X_t\right| \leq \delta,\ Y_t^N \notin [r,s]\right) \\
&\quad + \mathbb{E}^t\left(\left|h(Y_t^N) - h(X_t^N)\right| : \left|Y_t^N - X_t\right| \leq \delta,\ Y_t^N \in [r,s]\right) \\
&\quad + \mathbb{E}^t\left(\left|h(Y_t^N) - h(X_t^N)\right| : \left|Y_t^N - X_t\right| > \delta\right) \\
&\leq \epsilon + 2\epsilon\|h\|_\infty + 2\|h\|_\infty \frac{\mathbb{E}\left(\left|Y_t^N - X_t^N\right|\right)}{\delta} \\
&\leq \epsilon + 2\epsilon\|h\|_\infty + 2\|h\|_\infty \frac{\sup_t \mathbb{E}\left(\left|Y_t^N - X_t^N\right|\right)}{\delta}.
\end{aligned} \tag{7.84}$$

This completes the proof of the Proposition 7.1.3 and of the Theorem 7.1.1. □

Bibliography

7.4 Bibliogarphy: Part I

[1] R. J. Baxter, *Exactly solved models in statistical mechanics*, London, Academic Press, 1982.

[2] H. Kesten, B. P. Stigum, *A limit theorem for multidimensional Galton-Watson processes*, Ann. Math. Stat. no. 37, 1211–1223, 1966.

[3] H. Kesten, B. P. Stigum, *Additional limit theorems for undecomposable multidimensional Galton-Watson processes*, Ann. Math. Stat. no. 37, 1463–1481, 1966.

[4] H. O. Georgii, *Gibbs measures and phase transitions,* de Gruyter Studies in Mathematics **9**, Walter de Gruyter & Co., Berlin, 1988.

[5] P. M. Bleher, J. Ruiz, V.A. Zagrebnov, *On the purity of limiting Gibbs state for the Ising model on the Bethe lattice*, J. Stat. Phys. **79**, 473–482, 1995.

[6] D. Ioffe, *A note on the extremality of the disordered state for the Ising model on the Bethe lattice*, Lett. Math. Phys **37**, 137–143, 1996.

[7] D. Ioffe, *Extremality of the disordered state for the Ising model on general trees,* Trees (Versailles), 3–14, Progr. Probab. **40**, Birkhäuser, Basel, 1996.

[8] C. Borgs, J.T. Chayes, E. Mossel, S.Roch, *The Kesten–Stigum reconstruction bound is tight for roughly symmetric binary channels*, 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS), 518-530, 2006.

[9] R. Lyons, *Phase transitions on nonamenable graphs. In: Probabilistic techniques in equilibrium and nonequilibrium statistical physics.* J. Math. Phys. **41** no. 3, 1099–1126, 2000.

[10] S. Janson, E. Mossel, *Robust reconstruction on trees is determined by the second eigenvalue.* Ann. Probab. **32** no. 3B, 2630–2649, 2004.

[11] E. Mossel, Y.Peres, *Information flow on trees,* The Annals of Applied Probability **13** No. 3, 817–844, 2003.

[12] E. Mossel, *Reconstruction on trees: Beating the second eigenvalue,* Ann. Appl. Prob. **11**, 285-300, 2001.

[13] E. Mossel, *Phase transitions in phylogeny,* Trans. Amer. Math. Soc., **356** No.6, 2379-2404, 2004.

[14] R. Pemantle, Y. Peres, *The critical Ising model on trees, concave recursions and nonlinear capacity,* preprint, arXiv:math/0503137v2 [math.PR], 2006.

[15] M. Mezard, A. Montanari, *Reconstruction on trees and spin glass transition,* J. Stat. Phys. **124** no. 6, 1317–1350, 2006.

[16] A. Gerschelfeld, A. Montanari, *Reconstruction Problem for models on random graphs,* 48th Annual Symposium on Foundations of Computer Science, Providence, RI, 2007.

[17] M. Formentin, C. Külske, unpublished, 2008.

[18] M. Formentin, C. Külske, *On the Purity of the free boundary condition Potts measure on random trees,* to be published in Stochastic Processes and their Applications, available at arXiv:0810.0677, 2009.

[19] M. Formentin, C. Külske, *A symmetric entropy bound on the non-reconstruction regime of Markov chains on Galton-Watson trees,* preprint, available at arXiv:0903.2962 , 2009.

[20] F. Martinelli, A. Sinclair, D. Weitz, *Fast mixing for independents sets, colorings and other model on trees,* Random Structures Algorithms, **31** no. 2, 134–172, 2007.

[21] A. Sly, *Reconstruction of symmetric Potts models*, preprint, arXiv:0811.1208 [math.PR], 2008.

[22] N. Bhatnagar, E. Maneva, *A computational method for bounding the probability of reconstruction on trees*, preprint, 2009.

7.5 Bibliography: Part II

[23] P. Cattiaux, A. Guillin, F. Malrieu, *Probabilistic approach for granular media equations in the non-uniformly convex case*, Prob. Th. rel. Fields, **140**, n.1-2, 2008.

[24] P. Dai Pra, W. J. Runggaldier, E. Sartori, and M. Tolotti, *Large Portfolio Losses: A dynamic contagion model, The Annals of Applied Probability*, 19, no. 1, 347-394, 2009.

[25] K. Giesecke, S. Weber, *Credit contagion and aggregate losses, Journal of Economic Dynamics and Control*, **30**, Issue 5, 741-767, 2006.

[26] F. Malrieu, *Convergence to equilibrium for granular media equations and their Euler schemes*, Ann. Appl. Probab., **13**, no. 2, 540Ð560, 2003.

[27] T. Liggett, *Interacting Particle Systems*, New York, Springer, 1985.

[28] E. Sartori, *Some aspects of spin systems with mean–field interaction*, Ph.D. thesis, Scuola di Dottorato in Matematica Computazionale, Padova, 2007.

[29] M. Tolotti, *The impact of contagion on large portfolios. Modeling aspects*, Ph.D. thesis, Scuola Normale Superiore, Pisa, 2008.

[30] F. Collet, *The impact of disorder in the critical dynamics of mean–field models*, Ph.D. thesis, Scuola di Dottorato in Matematica Computazionale, Padova, 2009.

[31] A.N. Shirayaev, *Probability*, second edition, Graduate Text in Mathematics, Springer, 1986.

[32] S. N. Ethier, T. G. Kurtz, *Markov processes, Characterization and Convergence, John Wiley and Sons, New York, 1986.*