METASTABILITY IN STOCHASTIC DYNAMICS OF DISORDERED MEAN-FIELD MODELS

Anton Bovier[1], Michael Eckhoff[2], Véronique Gayrard[3], Markus Klein[4]

Abstract: We study a class of Markov chains that describe reversible stochastic dynamics of a large class of disordered mean field models at low temperatures. Our main purpose is to give a precise relation between the metastable time scales in the problem to the properties of the rate functions of the corresponding Gibbs measures. We derive the analog of the Wentzell-Freidlin theory in this case, showing that any transition can be decomposed, with probability exponentially close to one, into a deterministic sequence of "admissible transitions". For these admissible transitions we give upper and lower bounds on the expected transition times that differ only by a factor $\sqrt{N}$, where N denotes the volume of the system. The distribution rescaled transition times are shown to converge to the exponential distribution. We exemplify our results in the context of the random field Curie-Weiss model.

Keywords: Metastability, stochastic dynamics, Markov chains, Wentzell-Freidlin theory, disordered systems, mean field models, random field Curie-Weiss model.

AMS Subject Classification: 82C44, 60K35

[1]Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, D-10117 Berlin, Germany. e-mail: bovier@wias-berlin.de

[2]Institut für Mathematik, Universität Potsdam, Am Neuen Palais 10, D-14469 Potsdam, Germany. e-mail: eckhoff@rz.uni-potsdam.de

[3]Centre de Physique Théorique, CNRS, Luminy, Case 907, F-13288 Marseille, Cedex 9, France. email: Veronique.Gayrard@cpt.univ-mrs.fr

[4]Institut für Mathematik, Universität Potsdam, Am Neuen Palais 10, D-14469 Potsdam, Germany. e-mail: mklein@felix.math.uni-potsdam.de

Table of contents

1. Introduction . . . 3
  1.1. General introduction . . . 3
  1.2. General set-up . . . 5
  1.3. Outline of the general strategy . . . 12
2. Precise estimates on transition probabilities . . . 15
3. Laplace transforms of the transition times in the elementary situation . . . 21
4. Valleys, trees, and graphs . . . 26
  4.1. Valleystructure and its tree-representation . . . 26
  4.2. Construction of the transition process . . . 28
5. Transition times of admissible transitions . . . 33
  5.1. Expected times of admissible transitions . . . 33
  5.2. Laplace transforms of transition times of admissible transitions . . . 41
  5.3. The distribution of transition times . . . 47
6. Miscallaneous consequences for the process . . . 49
7. Mean field models and mean field dynamics . . . 53
8. The random field Curie Weiss model . . . 56
References . . . 62

1. Introduction

1.1 General introduction

This paper is devoted to developing a systematic approach to the analysis of the long time behaviour of the dynamics of certain mean field spin systems, where by dynamics we understand of course a stochastic dynamics of Glauber type. For the purposes of this paper, we will always choose this as reversible with respect to the Gibbs measure of the model. By long time behaviour we mean that we are interested in time scales on which the phenomena of "meta-stability" occur, i.e. time scales that increase with the volume of the system exponentially fast. Our primary motivation comes from the study of *disordered* spin systems, and most particularly the so called Hopfield model [Ho,BG1], although in the present paper we only illustrate our results in a much simpler setting, that of the random field Curie-Weiss (RFCWM) model (see e.g. [K1]). Our chief objective is to be able to control in a precise manner the effect of the randomness on the metastable phenomena.

On a heuristic level, metastable phenomena in mean field models are well understood. The main idea is to consider the dynamics induced on the order parameters by the Glauber dynamics on the spin space, i.e. the macroscopic variables that characterize the model. A first issue that arises here, and that we will discuss at length below, is that this induced dynamics is in general not Markovian. However, one may always define a new Markovian dynamics that "mimics" the old one and that is reversible with respect to the measures induced on order parameters by the Gibbs measures. This dynamics on the order parameters is essentially a random walk in a landscape given by the "rate function" associated to the distribution of the order parameters. The accepted picture of the resulting motion is that this walk will spend most of its time in the "most profound valleys" of the rate function and stay in a given valley for an exponentially long time of order $\exp(N\Delta F)$ where ΔF is the difference between the minimal value of the rate function in the valley and its value at the lowest "saddle point" over which the process may exit the valley. An excellent survey on this type of processes is given in van Kampen's textbook [vK], although most of the results presented there, and in particular all those related to the long time behaviour, concern the one-dimensional case. Rather surprisingly, one finds very few papers in the literature that really treat this problem with any degree of mathematical rigour. One exception is the classical paper by Cassandro, Galves, Olivieri, and Vares [CGOV] (see also [Va] for a broader review on metastability) who consider (amongst others) the case of the Curie-Weiss model in which there is only a single order parameter and thus the resulting dynamics is that of a one-dimensional random walk. More recently, a particular version of the RFCWM that leads to a two-dimensional problem was treated by Mathieu and Picco [MP]. However, there is an abundant literature on two types of related problems. One of these concerns Markov chains with finite state space and exponentially small transition probabilities. They are treated in the work of Freidlin and Wentzell (but see below for a discussion) and have since then been investigated intensely (for a small selection of recent references see [OS1,OS2,CC,GT]. In the context of stochastic dynamics of spin systems, they occur if finite systems are considered in the limit of zero temperature.[5] A second class of problems, that is in a sense closer to our situation, and that

[5] Let us mention, however, that there has been condiderable work done on the dynamics of spin systems on infinite lattices; see inparticular the recent paper by Schonmann and Shlosman [SS] on the metastable behaviour in the two-dimensional Ising model in infinite volume, and references therein.

can be obtained from it formally by passing to the limit of continuous space and time, is that of "small random perturbations of dynamical systems" i.e. a stochastic differential equation of the form

$$dx^\epsilon(t) = b(x^\epsilon(t))dt + \sqrt{\epsilon}a(x^\epsilon(t))dW(t) \tag{1.1}$$

where $x^\epsilon(t) \in \mathbb{R}^d$, and $W(t)$ is a d-dimensional Wiener process, and in the case of a reversible dynamics the drift term $b(x^\epsilon(t), \epsilon)$ is given by $b(x, \epsilon) = \nabla F_\epsilon(x)$, $F_e(x)$ being the rate function.

The basic reference on the problem (1.1) is the seminal book by Wentzell and Freidlin [FW] which discusses this problem (as well as a number of related ones) in great detail. Many further references can be found in the forthcoming second edition of this book. One of the important aspects of this work is that is devises a scheme that allows to control the long-time dynamics of the problem through an associated Markov chain with finite state space and exponentially small transition probabilities. The basic input here are large deviation estimates on the short time behaviour of the associated processes. This treatment has inspired a lot of consecutive works which it is impossible to summarize to any degree of completeness. For our purposes, an important development is a refinement of the estimates which in [FW] are given only to the leading exponential order in ϵ to a full asymptotic expansion. Relevant references are [Ki1-4,Az,FJ]. The work of [FJ] in particular is very interesting in that it develops full asymptotic expansions to all orders for certain exit probabilities using purely analytic techniques based on WKB methods. Very similar results are obtained in [Az] using refined large deviation techniques. To our knowledge all the refined treatments that have appeared in the literature treat only specific "local" questions, and there seems to be no coherent treatment of the global problem in a complicated (multi-valley) situation that takes into account sub-leading terms.

The problems we will study require essentially to redo the work of Freidlin and Wentzell in the setting of our Markov chains. Moreover, for the problems we are interested in, it will be important to have a more precise control, beyond the leading exponential asymptotics, for the global problem, if we want to be able to exhibit the influence of the residual randomness. The point is that in many disordered mean field models very precise estimates of the large deviation properties of the Gibbs measures are available. Typically, the rate function is deterministic to leading order (although *not* equal to the rate function of the averaged system[6]!) while the next order corrections (typically, but not always, of order $N^{-1/2}$) are random. To capture this effect, some degree of precision in the estimates is thus needed. On the other hand, we will not really need a full asymptotic expansion[7] of our quantities, and we will put more effort on the control of the "global" behaviour than on the overly precise treatment of "local" problems. A main difference is of course that we do not have a stochastic differential equation but a Markov chain on a discrete state space[8]. Therefore one may draw intuition from the proofs given in the continuous case without being able to use any result proved in that context directly. Finally, our goal is to give a treatment that is as simple and

[6] It is important to keep in mind that the main effect of the disorder manifests itself in a deterministic modification of the rate function. This effect is somewhat reminiscent to the phenomenon of homogenization.

[7] We believe that it is possible to obtain such an expansion for the global problem. However, this will require a much more elaborate analysis which we postpone to future publications.

[8] The state space is even finite for any N, but its size increases with N, which renders this fact rather useless.

transparent as possible. This is the main reason to concentrate on the reversible case, and one of our strategies is to use reversibility to as large an extent as possible. This allows to replace refined large deviation estimates by simple reversibility arguments. Large deviation estimates are then only used in a less delicate situations. In the same spirit, we will take advantage of the discrete nature of the problem whenever this is possible (just to compensate for all the disadvantages we encounter elsewhere). This will surprise the reader familiar with the continuous case, but we hope she will be convinced at the end that this was a pleasant surprise.

Let us say a final word concerning our preoccupations with the dependence on dimensionality. One of our ultimate goals is to be able to treat, e.g., the Hopfield model in the case where the number of order parameters grows with the volume of the system. On the level of the mean field dynamics, this requires us to be able to treat a system where the dimension of the space grows with the large parameter. Although we will not consider this situation in this first paper, we will achieve a precise control of the dimension dependence of sub-leading corrections.

1.2. The general set-up.

We will now describe the general class of Markov chains we will consider. Their relation to disordered spin systems will be explained in Section 7 and a specific example will be discussed in Section 8. Section 7 can be read now, if desired; on the other hand, the bulk of the paper can also be read without reference to this motivation.

We consider canonical Markov chains on a state space Γ_N where Γ_N is the intersection of some lattice[9] (of spacing $O(1/N)$) in $\mathbb{R}^d$ with some connected $\Lambda \subset \mathbb{R}^d$ which is either open or the closure of an open set. To avoid some irrelevant issues, we will assume that Λ is either $\mathbb{R}^d$ or a bounded and convex subset of $\mathbb{R}^d$. Γ_N is assumed to have spacing of order $1/N$, i.e. the cardinality of the state space is of order N^d. Moreover, we identify Γ_N with a graph with finite (d-dependent) coordination number respecting the Euclidean structure in the sense that a vertex $x \in \Gamma_N$ is connected only to vertices at Euclidean distances less than c/N from x. The main example the reader should have in mind is $\Gamma_N = \mathbb{Z}^d/N \cap \Lambda$, with edges only between nearest neighbors. We denote the set of edges of Γ_N by $E(\Gamma_N)$.

Let $\mathbb{Q}_N$ be a probability measure on $(\Gamma_N, \mathcal{B}(\Gamma_N))$. We will set, for $x \in \Gamma_N$,

$$F_N(x) \equiv -\frac{1}{N} \ln \mathbb{Q}_N(x) \tag{1.2}$$

We will assume the following properties of $F_N(x)$.

Assumptions:

R1 $F \equiv \lim_{N\uparrow\infty} F_N$ exists and is a smooth function $\Lambda \to \mathbb{R}$; the convergence is uniform in compact subsets of $\mathbb{R}^d$.

R2 F_N can be represented as $F_N = F_{N,0} + \frac{1}{N} F_{N,1}$ where $F_{N,0}$ is twice Lipshitz, i.e. $|F_{N,0}(x) - F_{N,0}(y)| \leq C\|x-y\|$ and for any generator of the lattice, k, $N|F_{N,0}(x) - F_{N,0}(x+k/N) -$

[9] The requirement that Γ_N is a lattice is made for convenience and can be weakened considerably, if desired. What is needed are some homogeneity and rather minimal isotropy assumptions.

$(F_{N,0}(y) - F_{N,0}(y+k/N)| \leq C\|x-y\|$, with C uniform on compact subsets of the interior of Λ. $F_{N,1}$ is only required to be Lipshitz, i.e. $|F_{N,1}(x) - F_{N,1}(y)| \leq C\|x-y\|$.

For the purposes of the present paper we will make a number of assumptions concerning the functions F_N which we will consider as "generic". An important assumption concerns the structure of the set of minima of the functions F_N. We will assume that the set $\mathcal{M}_N \subset \Gamma_N$, of *local* minima of F_N is finite and of constant cardinality for all N large enough, and that the sets $\mathcal{M}_N$ converge, as N tends to infinity, to the set $\mathcal{M}$ of local minima of the function F[10].

Another set of points that will be important is the set, $\mathcal{E}_N$, of "essential" saddle points (i.e. the lowest saddle points one has to cross to go from one minimum to another) of F_N. A precise definition of essential saddle points will be given in Section 4. By the assumptions on $\mathcal{M}_N$ this set is also finite.

G1 We will assume that there exists $\alpha > 0$ such that $\min_{x\neq y \in \mathcal{M}_N \cup \mathcal{E}_N} |F_N(x) - F_N(y)| = K_N \geq N^{\alpha-1}$.

G2 We assume that at each minimum the eigenvalues of the Hessian of F are strictly positive and at each essential saddle there is one strictly negative eigenvalue while all others are strictly positive.

G3 All minima and saddles are well in the interior of Λ, i.e. there exists a $\delta > 0$ such that for any $x \in \mathcal{M}_N \cup \mathcal{E}_N$, $\mathrm{dist}(x, \Lambda^c) > \delta$.

Remark: We make the rather strong assumptions above in order to be able to formulate very general theorems that do not depend on specific properties of the model. They can certainly be relaxed. The regularity conditions **R2** are necessary only for the application of certain large deviation results in Section 4 and are otherwise not needed.

We recall that in our main applications, $\mathbb{Q}_N$ will be random measures, but we will forget this fact for the time being and think of $\mathbb{Q}_N$ as some particular realization.

We can now construct a Markov chain $X_N(t)$ with state space given by the set of vertices of Γ_N and time parameter set either $\mathbb{N}$ or $\mathbb{R}_+$. For this we first define for any x, y such that $(x,y) \in E(\Gamma_N)$ transition rates

$$p_N(x,y) \equiv \sqrt{\frac{\mathbb{Q}_N(y)}{\mathbb{Q}_N(x)}} f_N(x,y) \tag{1.3}$$

for some non-negative, symmetric function f_N. We will assume that f_N does not introduce too much anisotropy. This can be expressed by demanding that

R3 There exists $c > 0$ such that if $(x,y) \in E(\Gamma_N)$, and $\mathrm{dist}(x, \Lambda^c) > \delta/2$, (where δ is the same as in assumption **G3**) $p_N(x,y) \geq c$.

[10]This assumption can easily be relaxed somewhat. For example, it would be no problem if the function F_N is degenerate on a small set of points in the very close (order $N^{-1/2}$) neighborhood of a minimum. One then would just choose one of them to represent this cluster. Other situations, e.g. when the function F has local minima on large sets and would lead to new effects would require special treatments.

Moreover, for applications of large deviation results we need stronger regularity properties analogous to **R2**.

R4 $\ln f_N(x,y)$ as a function of any of its arguments is uniformly Lipshitz on compact subsets of the interior of Λ.

For the case of discrete time, i.e. $t \in \mathbb{N}$, we then define the transition matrix

$$P_N(x,y) \equiv \begin{cases} p_N(x,y), & \text{if} \quad (x,y) \in E(\Gamma_N) \\ 1 - \sum_{z\in\Gamma_N:(x,z)\in E(\Gamma_N)} p_N(x,z), & \text{if} \quad x = y \\ 0, & \text{else} \end{cases} \tag{1.4}$$

choosing f such that $\sup_{x\in\Gamma_N} \sum_{z\in\Gamma_N:(x,z)\in E(\Gamma_N)} p_N(x,y) \leq 1$.

Similarly, in the continuous time case, we can use the rates to define the generator

$$A_N(x,y) \equiv \begin{cases} p_N(x,y), & \text{if} \quad (x,y) \in E(\Gamma_N) \\ -\sum_{z\in\Gamma_N:(x,z)\in E(\Gamma_N)} p_N(x,y), & \text{if} \quad x = y \\ 0, & \text{else} \end{cases} \tag{1.5}$$

Our basic approach to the analysis of these Markov chains is to observe the process when it is visiting the positions of the minima of the function F_N, i.e. the points of the set $\mathcal{M}_N$, and to record the elapsed time. The ideology behind this is that we suspect the process to show the following typical behaviour: starting at any given point, it will rather quickly (i.e. in some time of order N^k) visit a nearby minimum, and then visit this same minimum at similar time interval an exponentially large number of times without visiting any other minimum between successive returns. Then, at some random moment it will go, quickly again, to some other minimum which will then be visited regularly a large number of times, and so on. Moreover, between successive visits of a minimum the process will typically not only avoid visits at other minima, but will actually stay very close to the given minimum. Thus, recording the visits at the minima will be sufficient information on the behaviour of the process. These expectations will be shown to be justified (see in particular Section 7). Incidentally, we mention that the "quick" processes of transitions can be analysed in detail using large deviation methods [WF1-4]. In [BG2] a large deviation principle is proven for a class of Markov chains including those considered here that shows that the "paths" of such quick processes concentrate asymptotically near the classical trajectories of some (relativistic) Hamiltonian system. More precisely, the transitions between minima can be identified as instanton solutions of the corresponding Hamiltonian system.

Let us mention that the strategy to record visits at single points is specific to the discrete state space. In the diffusion setting, visits at single points do not happen with sufficient probability to contain pertinent information on the process. Indeed, the crucial fact we use is that in the discrete case it is excessively difficult for the process to stay for a time of order N^k (we will discuss the values of k later) in the vicinity of a minimum without visiting it[11]

[11] The reader may wonder at this point why the minima are so special compared e.g. with their neighboring points. In fact they are not, and nothing would change if we chose some other point close to the minimum rather than the exact minimum. But of course the minima themselves are the optimal choice, and also the most natural ones.

which in the continuum is not the case. For this reason Freidlin and Wentzell record visits not at single points but at certain neighborhoods of minima and critical points which has the disadvantage that such visits do not exactly allow a splitting of the process and this introduces some error terms in estimates which in our setting can easily be avoided. This is the main advantage we draw from working in a discrete space.

The informal discussion above will be made precise in the sequel. We place ourselves in the discrete time setting throughout this paper, but everything can be transferred to the continuous time setup with mild modifications, if desired. Let us first introduce some notation. We will use the symbol $\mathbb{P}$ for the law of our Markov chain, omitting the explicit mention of the index N, and denote by X_t the coordinate variables. We will write τ_x^y for the first time the process conditioned to starting at y hits the point x, i.e. we write

$$\mathbb{P}\left[\tau_x^y = t\right] \equiv \mathbb{P}\left[X_t = x, \forall_{0<s<t} X_s \neq x | X_0 = y\right] \tag{1.6}$$

for $t > 0$. In the case $x = y$, we will insist that τ_y^y is the time of the first visit to y *after* $t = 0$, i.e. $\mathbb{P}[\tau_y^y = 0] = 0$. This notation may look unusual at first sight, but we are convinced that the reader will come to appreciate its convenience.

One of the most useful basic identities which follows directly from the strong Markov property and the fact that τ_x^y is a stopping time is the following:

Lemma 1.1: *Let x, y, z be arbitrary points in Γ_N. Then*

$$\begin{aligned}\mathbb{P}\left[\tau_x^y = t\right] = &\ \mathbb{P}\left[\tau_x^y = t, \tau_x^y < \tau_z^y\right] \\ &+ \sum_{0<s<t} \mathbb{P}\left[\tau_z^y = s, \tau_z^y < \tau_x^y\right] \mathbb{P}\left[\tau_x^z = t - s\right]\end{aligned} \tag{1.7}$$

Proof: Just note that the process either arrives at x before visiting z, or it visits z a first time before x. $\diamondsuit$

A simple consequence is the following basic renewal equation.

Lemma 1.2: *Let $x, y \in \Gamma_N$. Then*

$$\mathbb{P}\left[\tau_x^y = t\right] = \sum_{n=0}^{\infty} \sum_{\substack{t_1, \dots t_{n+1} \\ \sum_i t_i = t}} \prod_{i=1}^{n} \mathbb{P}\left[\tau_y^y = t_i, \tau_y^y < \tau_x^y\right] \mathbb{P}\left[\tau_x^y = t_{n+1}, \tau_x^y < \tau_y^y\right] \tag{1.8}$$

The fundamental importance in the decomposition of Lemma 1.2 lies in the fact that objects like the last factor in (1.8) are “reversible”, i.e. they can be compared to their time-reversed counterpart. To formulate a general principle, let us define the time-reversed chain corresponding to a transition from y to x via $X_t^r \equiv X_{\tau_x^y - t}$. For an event $\mathcal{A}$ that is measurable with respect to the sigma algebra $\mathcal{F}\left(X_s, 0 \leq s \leq \tau_x^y\right)$ we then define the time reversed event $\mathcal{A}^r$ as the event that takes place for the chain X_t if and only if the event $\mathcal{A}$ takes place for the chain X_t^r. This allows us to formulate the next lemma:

Lemma 1.3: *Let $x, y \in \Gamma_N$, and let $\mathcal{A}$ be any event measurable with respect to the sigma algebra $\mathcal{F}(X_s, 0 \leq s \leq \tau_x^y)$. Let $\mathcal{A}^r$ denote the time reversion of the event $\mathcal{A}$. Then*

$$\mathbb{Q}_N(y)\mathbb{P}\left[\mathcal{A}, \tau_x^y < \tau_y^y\right] = \mathbb{Q}_N(x)\mathbb{P}\left[\mathcal{A}^r, \tau_y^x < \tau_x^x\right] \tag{1.9}$$

For example, we have

$$\mathbb{Q}_N(y)\mathbb{P}\left[\tau_x^y = t, \tau_x^y < \tau_y^y\right] = \mathbb{Q}_N(x)\mathbb{P}\left[\tau_y^x = t, \tau_y^x < \tau_x^x\right] \tag{1.10}$$

Of course the power of Lemma 1.3 comes to bear when x and y are such that the ratio between $\mathbb{Q}_N(x)$ and $\mathbb{Q}_N(y)$ is very large or very small.

Formulas like (1.8) invite the use of Laplace transforms. Let us first generalize the notion of stopping times to arrival times in sets. I.e. for any set $I \subset \Gamma_N$ we will set τ_I^x to be the time of the first visit of the process, starting at x, to the set I. With this notion we define the corresponding Laplace transforms

$$G_{x,I}^y(u) \equiv \sum_{t \geq 0} e^{ut}\mathbb{P}\left[\tau_x^y = t, \tau_x^y \leq \tau_I^y\right] \equiv \mathbb{E}\left[e^{u\tau_x^y} 1\!\!1_{\tau_x^y \leq \tau_I^y}\right] \tag{1.11}$$

(We want to include the possibility that I contains x and/or y for later convenience). Note that in particular

$$G_{x,I}^y(0) = \mathbb{P}\left[\tau_x^y \leq \tau_I^y\right] \tag{1.12}$$

and

$$\frac{d}{du}G_{x,I}^y(u = 0) \equiv \dot{G}_{x,I}^y(0) = \mathbb{E}\left[\tau_x^y 1\!\!1_{\tau_x^y \leq \tau_I^y}\right] \tag{1.13}$$

The expected time of reaching x from y conditioned on the event not to visit I in the meantime is expressed in terms of these functions as

$$\frac{\dot{G}_{x,I}^y(0)}{G_{x,I}^y(0)} = \mathbb{E}\left[\tau_x^y | \tau_x^y \leq \tau_I^y\right] \tag{1.14}$$

It is important to keep in mind that the Laplace transforms defined in (1.11) are Laplace transforms of the distributions of positive random variables. Thus, all these functions exist and are analytic at least for all $u \in \mathbb{C}$ with $Re(u) \leq 0$. Moreover, if $G_{x,I}^y(u_0)$ is finite for some $u_0 \in \mathbb{R}_+$, then it is analytic in the half-space $Re(u) \leq u_0$. As we will see later, all the functions introduced in (1.11) will exist for some $u_0 > 0$.

An important consequence of Lemma 1.3 is

Lemma 1.4: *Assume that I is any subset of Γ_N containing x and y. Then*

$$\mathbb{Q}_N(y)G_{x,I}^y(u) = \mathbb{Q}_N(x)G_{y,I}^x(u) \tag{1.15}$$

Proof: Immediate from Lemma 1.3. $\diamondsuit$

Lemma 1.4 implies in particular that the Laplace transforms of the *conditional* times are invariant under reversal, i.e.

$$\frac{G^y_{x,I}(u)}{G^y_{x,I}(0)} = \frac{G^x_{y,I}(u)}{G^x_{y,I}(0)} \tag{1.16}$$

and in particular

$$\mathbb{E}[\tau^y_x | \tau^y_x \leq \tau^y_I] = \mathbb{E}\left[\tau^x_y | \tau^x_y \leq \tau^x_I\right] \tag{1.17}$$

(1.16) expresses the well-known but remarkable fact that in a reversible process the conditional times to reach a point x from y without return to y are equal to those to reach y from x without return to x.

A special rôle will be played by the Laplace transforms for which the exclusion set are all the minima. We will denote these by $g^y_x(u) \equiv G^y_{x,\mathcal{M}_N}(u)$. Indeed, we think of the events $\{\tau^y_x \leq \tau^y_{\mathcal{M}_N}\}$, for $x, y \in \mathcal{M}_N$, as elementary transitions and decompose any process going from one minimum to another into such elementary transitions. This gives:

Lemma 1.5: *Let $x, y \in \mathcal{M}_N$. denote by ω an arbitrary sequence $\omega = \omega_0, \omega_1, \omega_2, \omega_3, \dots, \omega_{|\omega|}$ of elements $\omega_i \in \mathcal{M}_N$. Then we have*

$$G^y_x(u) = \sum_{\omega : x \to y} p(\omega) \prod_{i=1}^{|\omega|} \frac{g^{\omega_{i-1}}_{\omega_i}(u)}{g^{\omega_{i-1}}_{\omega_i}(0)} \tag{1.18}$$

where

$$p(\omega) \equiv \prod_{i=1}^{|\omega|} \mathbb{P}\left[\tau^{\omega_{i-1}}_{\omega_i} \leq \tau^{\omega_{i-1}}_{\mathcal{M}_N}\right] \tag{1.19}$$

and $\omega : y \to x$ indicates that the sum is over such walks for which $\omega_0 = y$ and $\omega_{|\omega|} = x$, and $\omega_i \neq x$ for all $0 < i < |\omega|$.

Lemma 1.5 can be thought of as a random walk representation of our process as observed on the minima only. As we will show soon, the quantities $\frac{g^{\omega_{i-1}}_{\omega_i}(u)}{g^{\omega_{i-1}}_{\omega_i}(0)}$ are rather harmless, i.e. they do not explode in a small neighborhood of zero, and e.g. their derivative at zero is at most polynomially large in N. On the other hand, we will also see that the "transition probabilities" $\mathbb{P}\left[\tau^{\omega_{i-1}}_{\omega_i} < \tau^{\omega_{i-1}}_{\mathcal{M}_N}\right]$ are all exponentially small provided that $\omega_{i-1} \neq \omega_i$. This means that a typical walk will contain enormously long "boring" chains of repeated returns to the same point. It is instructive to observe that these repeated returns to a given minimum can be re-summed, to obtain a representation in terms of walks that do not contain zero steps:

Lemma 1.6: *Let $x, y \in \mathcal{M}_N$. denote by $\tilde{\omega}$ a sequence $\tilde{\omega} = \omega_0, \omega_1, \omega_2, \omega_3, \dots, \omega_{|\omega|}$ of elements $\omega_i \in \mathcal{M}_N$ such that for all i, $\omega_i \neq \omega_{i+1}$. Then we have*

$$G^y_x(u) = \sum_{\tilde{\omega} : x \to y} \tilde{p}(\tilde{\omega}) \prod_{i=1}^{|\omega|} \frac{1 - g^{\omega_{i-1}}_{\omega_{i-1}}(0)}{1 - g^{\omega_{i-1}}_{\omega_{i-1}}(u)} \frac{g^{\omega_{i-1}}_{\omega_i}(u)}{g^{\omega_{i-1}}_{\omega_i}(0)} \tag{1.20}$$

where

$$\tilde{p}(\tilde{\omega}) \equiv \prod_{i=1}^{|\omega|} \frac{\mathbb{P}\left[\tau_{\omega_i}^{\omega_{i-1}} \leq \tau_{\mathcal{M}_N}^{\omega_{i-1}}\right]}{\mathbb{P}\left[\tau_{\mathcal{M}_N \backslash \omega_{i-1}}^{\omega_{i-1}} < \tau_{\omega_{i-1}}^{\omega_{i-1}}\right]} \tag{1.21}$$

The reason for writing Lemma 1.6 in the above form is that it entails as a corollary the following expression for the expected transition time:

$$\mathbb{E}\tau_x^y = \sum_{\tilde{\omega}:x\to y} \tilde{p}(\tilde{\omega}) \sum_{i=1}^{|\omega|} \left(\frac{\dot{g}_{\omega_{i-1}}^{\omega_{i-1}}(0)}{1-g_{\omega_{i-1}}^{\omega_{i-1}}(0)} + \frac{\dot{g}_{\omega_i}^{\omega_{i-1}}(0)}{g_{\omega_i}^{\omega_{i-1}}(0)} \right) \tag{1.22}$$

Note that $\tilde{p}(\tilde{\omega})$ has indeed a natural interpretation as the probability of the sequence of steps $\tilde{\omega}$, while each term in the sum is the expected time such a step takes. Moreover, this time consists of two pieces: the first is a waiting time which in fact arises from the re-summation of the many returns before a transition takes place while the second is the time of the actual transition, once it really happens. Note that the first term is enormous since the denominator, $1-g_{\omega_{i-1}}^{\omega_{i-1}}(0) = \mathbb{P}\left[\tau_{\mathcal{M}_N \backslash \omega_{i-1}}^{\omega_{i-1}} < \tau_{\omega_{i-1}}^{\omega_{i-1}}\right]$, is, as we will see, exponentially small.

Remark: Lemma 1.6 does provide a representation of the process on the minima in terms of an embedded Markov chain with exponentially small transition probabilities. Moreover, we expect that for N large, the waiting times will be almost exponentially distributed (but with very different rates!), while transitions happen essentially instantaneously on the scale of even the fastest waiting time. This is the analogue of the controlling Markov processes constructed in Freidlin and Wentzell (see in particular Chap. 6.2 of [FW]).

In the case where $\mathcal{M}_N$ consists of only two points, Lemma 1.6 already provides the full solution to the problem since the only walk left is the single step (y, x).

Corollary 1.7: *Assume that $\mathcal{M}_N = \{x, y\}$. Then*

$$G_x^y(u) = \frac{g_x^y(u)}{1-g_y^y(u)} \tag{1.23}$$

and

$$\mathbb{E}\tau_x^y = \frac{\dot{g}_y^y(0)}{1-g_y^y(0)} + \frac{\dot{g}_x^y(0)}{g_x^y(0)} \tag{1.24}$$

Proof: Just use that in this particular setting, $1-g_y^y(0) = g_x^y(0)$. $\diamondsuit$

Remark: (1.24) can be written in the maybe more instructive form

$$\mathbb{E}\tau_x^y = \frac{1-g_x^y(0)}{g_x^y(0)} \frac{\dot{g}_y^y(0)}{g_y^y(0)} + \frac{\dot{g}_x^y(0)}{g_x^y(0)} \tag{1.25}$$

As we will see, all the ratios of the type $\dot{g}(0)/g(0)$ represent the expected times of a transition conditioned on the event that this transition happens and should be thought of as "small". On the other hand, the probability $g_x^y(0)$ will be shown to be exponentially small so that the first factor in the first term in (1.25) is extremely large. Thus, to get a precise estimate on the expected transition time in this case, it suffices to compute precisely the two quantities $\dot{g}_y^y(0)$ and $g_x^y(0)$ only (the second term in (1.25) being negligible in comparison). One might be tempted to think that in the general case the random walk representation given through Lemma 1.6 would similarly lead to a reduction to the problem to that of computing the corresponding quantities at and between all minima. This however is not so. The reason is that the walks $\tilde{\omega}$ still can perform more complicated multiple loops and these loops will introduce new and more singular terms when appear explicitly in (1.20) and (1.22). This renders this representation much less useful than it appears at first sight. On the other hand, the structure of the representation of Corollary 1.7 will be rather universal. Indeed, it is easy to see that with our notations we have the following

Lemma 1.8: *Let $I \subset \Gamma_N$. Then for all $y \notin I \cup x$,*

$$G_{x,I}^y(u) = \frac{G_{x,\{I\cup y\}}^y(u)}{1 - G_{y,\{I\cup x\}}^y(u)} \tag{1.26}$$

holds for all u for which the left-hand side exists.

Proof: Separating paths that reach x from y without return to y from those that do return, and splitting the latter at the first return time, using the strong Markov inequality, we get that

$$G_{x,I}^y(u) = G_{x,\{I\cup y\}}^y(u) + G_{y,\{I\cup x\}}^y(u) G_{x,I}^y(u) \tag{1.27}$$

By construction, if $G_{x,I}^y(u)$ is finite, the second summand being less than the left-hand side, we have that $+G_{y,\{I\cup x\}}^y(u) < 1$ and so (1.26) follows.$\diamondsuit$

Lemma 1.8 will be one of our crucial tools. In particular, since it relates functions with exclusion sets I to functions with larger exclusion sets, it suggests control over the Laplace transforms via induction over the size of the exclusion sets.

A special case of particular importance is obtained by setting $u = 0$ in Lemma 1.8. This give the

Corollary 1.9: *Let $I \subset \Gamma_N$. Then for all $y \notin I \cup x$,*

$$\mathbb{P}\left[\tau_x^y < \tau_I^y\right] = \frac{\mathbb{P}\left[\tau_x^y < \tau_{I\cup y}^y\right]}{\mathbb{P}\left[\tau_{I\cup x}^y < \tau_y^y\right]} \tag{1.28}$$

1.3. Outline of the general strategy.

As indicated above, an important tool in our analysis will be the use of induction over the size of exclusion sets by the help of Lemmata 1.1 and 1.8. One of the basic inputs for this will be a priori estimates on the quantities $g_y^x(u)$. These will be based on the representation of these functions as solutions of certain Dirichlet problems associated to the operator $(1-e^u P_N)$ with Dirichlet boundary conditions in set containing $\mathcal{M}_N$.

The crucial point here is to have Dirichlet boundary conditions at all the minima of F_N and at y. Without these boundary conditions, the stochastic matrix P is symmetric in the space $\ell_2(\Gamma_N, \mathbb{Q}_N)$ and has a maximal eigenvalue 1 with corresponding (right) eigenvector 1; since this eigenvector does not satisfy the Dirichlet boundary conditions at the minima, the spectrum of the Dirichlet operator lies strictly below 1, so that for sufficiently small values of u, $1 - e^u P$ is invertible. It is essential to know by how much the Dirichlet conditions push the spectrum down. It turns out that Dirichlet boundary conditions at all the minima push the spectrum by an amount of at least CN^{-d-1} below one, and this will allow us not only to construct the solution but to get very good control on its behaviour. If, on the other hand, not all the minima had received Dirichlet conditions, we must expect that the spectrum is only pushed down by an exponentially small amount, and we will have to devise different techniques to deal with these quantities.

As a matter of fact, while the spectral properties discussed above follow from our estimates, we will not use these to derive them. The point is that what we really need are pointwise estimates on our functions, rather than ℓ_2 estimates, and we will actually use more probabilistic techniques to prove ℓ_∞ estimates as key inputs. The main result, proven in Section 3, will be the following theorem:

Theorem 1.10: *There exists a constant $c > 0$ such that for all $x \in \Gamma_N$, $y \in \mathcal{M}_N$ the functions $g_y^x(u)$ are analytic in the half-plane $Re(u) < cN^{-d-3/2}$. Moreover, for such u, for any non-negative integer k there exists a constant C_k such that*

$$\left| \frac{d^k}{du^k} g_y^x(u) \right| \leq C_k N^{k(d+3/2)+d/2} e^{N[F_N(x) - F_N(z^*(x,y))]} \tag{1.29}$$

where

$$z^*(y,x) \equiv \arg \inf_{c:c(0)=y, c(1)=x} \left(\sup_{t\in[0,1]} [F_N(c(t))] \right) \tag{1.30}$$

where the infimum is over all paths $c : [0,1] \to \Gamma_N$ going from y to x.[12] *Γ_N with jumps along the edges of Γ_N only.*

These estimates are not overly sharp, and there are no corresponding lower bounds. Therefore, our strategy will be to use these estimates only to control sub-leading expressions and to use different methods to control the leading quantities which will be seen to be certain of the expected return times, like $\frac{\dot g_x^x(0)}{g_x^x(0)}$ and the transition probabilities $\mathbb{P}\left[\tau_y^x < \tau_x^x\right]$. The latter quantities will be estimated up to a multiplicative error of order $N^{1/2}$ in Section 2. In fact we will prove there the following theorem:

Theorem 1.11: *With the notation of Theorem 1.10 there exists finite positive constants c, C such that if $x \neq y \in \mathcal{M}_N$, then*

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \leq cN^{\frac{d-1}{2}} e^{-N[F_N(z^*(y,x)) - F_N(y)]} \tag{1.31}$$

and

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \geq CN^{\frac{d-2}{2}} e^{-N[F_N(z^*(y,x)) - F_N(y)]} \tag{1.32}$$

---

[12] Note that here we think of a path as a discontinuous (càdlàg) function that stay at a site in Γ_N for some time interval δt and then jumps to a neighboring site along an edge of Γ_N. This parametrization will however be of no importance and allows just some convenient notation. If $z^*(x,y) \notin \{x,y\}$, we will call $z^*(x,y)$ the essential saddle between x and y.

The estimates for the return times require some more preparation and will be stated only in Section 5, but let us mention that the main idea in getting sharp estimates for them is the use of the ergodic theorem .

Equipped with these inputs we will, in Section 4, proceed to the analysis of general transition processes. We will introduce a natural tree structure on the set of minima and show that any transition between two minima can be uniquely decomposed into a sequence of so-called "admissible transitions" in such a way that with probability rapidly tending to one (as $N \uparrow \infty$), the process will consist of this precise sequence of transitions. This will require large deviation estimates in path space that are special cases of more general results that have recently been proven in [BG2].

In Section 5 we will investigate the transition times of admissible transitions. In the first sub-section we will prove sharp bounds on the expected times of such admissible transitions with upper and lower bounds differing only be a factor of $N^{1/2}$. This will be based on more general upper bounds on expected times of general types of transitions that will be proven by induction. In the second sub-section we show that the rescaled transition times converge (along subsequences) to exponentially distributed random variables. This result again is based on an inductive proof establishing control on the rather complicated analytic structure of the Laplace transforms of a general class of transition times. In Section 6 we use these results to derive some consequences: We show that during an admissible transition, at any given time, the process is close to the starting point with probability close to 1, that it converges exponentially to equilibrium, etc. Section 7 motivates the connection between our Markov chains and Glauber dynamics of disordered mean field models, and in Section 8 we discuss a specific example, the random field Curie-Weiss model.

Notation: We have made an effort to use a notation that is at the same time concise and unambiguous. This has required some compromise and it may be useful to outline our policy here. First, all objects associated with our Markov chains depend on N. We make this evident in some cases by a subscript N. However, we have omitted this subscript in other cases, in particular when there is already a number of other indices that are more important (as in $G^x_y(u)$), or in ever recurring objects like $\mathbb{P}$ and $\mathbb{E}$, and which sometimes will have to be distinguished from the laws of modified Markov chains by other subscript. Constants c, C, k etc. will always be understood to depend on the details of the Markov chain, but to be independent of N for N large. There will appear constants $K_N > 0$ that will depend on N in a way depending on the details of the chain, but such that for some $\alpha > 0$, $N^{1-\alpha} K_N \uparrow \infty$ (this can be seen as a requirement on the chain). Specific letters are reserved for a particular meaning only locally in the text.

Acknowledgements: A.B. would like to thank Enzo Olivieri for an inspiring discussion on reversible dynamics that has laid the foundation of this work. M. K. thanks Johannes Sjöstrand and B. Helffer for helpful discussions. Finally, V.G. and A.B. we thank the Weierstrass-Institute, Berlin, and the Centre de Physique Théorique, Marseille, for hospitality and financial support that has made this collaboration possible.

2. Precise estimates on transition probabilities

In this section we will prove the upper and lower bounds of Theorem 1.11 for the quantities

$$G_x^y(0) \equiv \mathbb{P}\left[\tau_x^y < \tau_y^y\right], \quad x, y \in \mathcal{M}_N \tag{2.1}$$

Proof of Theorem 1.11.: The proof of the upper bound (1.31) is very easy. We construct a 'hyper-surface'[13] $\mathcal{S}_N \subset \Gamma_N$ separating x and y such that

i) $z^*(y,x) \in \mathcal{S}_N$.

ii) $\forall z \in \mathcal{S}_N$, $F_N(z) \geq F_N(z^*(y,x))$.

Path splitting allows then the simple upper bounds:

$$\begin{aligned}\mathbb{P}\left[\tau_x^y < \tau_y^y\right] &\leq \sum_{z\in\mathcal{S}_N} \mathbb{P}\left[\tau_z^y < \tau_y^y, \tau_z^y \leq \tau_{\mathcal{S}_N}^y\right] \mathbb{P}\left[\tau_x^z < \tau_y^z\right] \\ &\leq \sum_{z\in\mathcal{S}_N} \mathbb{P}\left[\tau_z^y < \tau_y^y, \tau_z^y \leq \tau_{\mathcal{S}_N}^y\right]\end{aligned} \tag{2.2}$$

Using reversibility, we have furthermore that

$$\mathbb{P}\left[\tau_z^y < \tau_y^y, \tau_z^y \leq \tau_{\mathcal{S}_N}^y\right] = e^{-N[F_N(z)-F_N(y)]}\mathbb{P}\left[\tau_y^z < \tau_z^z, \tau_y^z < \tau_{\mathcal{S}_N}^z\right] \leq e^{-N[F_N(z)-F_N(y)]} \tag{2.3}$$

Since we assume that we have a generic situation with no more than a finite number of equivalent escape saddles and that, by **G2**, F_N is quadratic at these saddle points, then a straightforward computation shows that (2.3) implies (1.31).

The main task of this section will be to establish the corresponding lower bound (1.32). The main idea of the proof of the lower bound is to reduce the problem to a sum of essentially one-dimensional ones which can be solved explicitly. The key observation is the following monotonicity property of the transition probabilities.

Lemma 2.1: *Let $\Delta \subset \Gamma_N$ be a subgraph of Γ_N and let $\widetilde{\mathbb{P}}_\Delta$ denote the law of the Markov chain with transition rates*

$$\widetilde{p}_\Delta(x',x'') = \begin{cases} p_N(x',x''), & \text{if} \quad x' \neq x'' \text{, and } (x',x'') \in E(\Delta) \\ 1 - \sum_{y':(x',y')\in E(\Delta)} p_N(x',y'), & \text{if} \quad x' = x'' \\ 0, & \text{else} \end{cases} \tag{2.4}$$

Assume that $y, x \in \Delta$. Then

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \geq \widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \tag{2.5}$$

[13] We actually require no analytic properties for the set $\mathcal{S}_N$ and the term hyper-surface should not be taken very seriously.

Proof: This lemma is an almost immediate consequence of the following variational representation of the transition probability that can be found in the book by Liggett ([Li], p. 99, Theorem 6.1):

Theorem 2.2: [Li] *Let $\mathcal{H}_x^y$ denote the space of functions*

$$\mathcal{H}_x^y \equiv \{h : \Gamma_N \to [0,1] : h(y) = 0, h(x) = 1\} \tag{2.6}$$

and define the Dirichlet form

$$\Phi_N(h) \equiv \sum_{x',x'' \in \Gamma_N} \mathbb{Q}_N(x') p_N(x', x'')[h(x') - h(x'')]^2 \tag{2.7}$$

Then

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] = \frac{1}{2\mathbb{Q}_N(y)} \inf_{h \in \mathcal{H}_x^y} \Phi_N(h) \tag{2.8}$$

Proof: See Ligget [Li], Chapter II.6. Note that the set R in Liggett's book in will be $\Gamma_N \backslash \{x\}$, and our $\mathcal{H}_x^y$ would be $H_{\Gamma_N \backslash \{x\}}$ in his notation.$\diamondsuit$

To prove Lemma 2.1 from here, just note that for any $h \in \mathcal{H}_x^y$

$$\begin{aligned} \Phi_N(h) &\geq \sum_{(x',x'') \in \Delta} \mathbb{Q}_N(x') p_N(x', x'')[h(x') - h(x'')]^2 \\ &= \mathbb{Q}_N(\Delta) \sum_{(x',x'') \in \Delta} \widetilde{\mathbb{Q}}_\Delta(x') \widetilde{p}_\Delta(x', x'')[h(x') - h(x'')]^2 \end{aligned} \tag{2.9}$$

where $\widetilde{\mathbb{Q}}_\Delta(x) \equiv \mathbb{Q}_N(x)/\mathbb{Q}_N(\Delta)$. This implies immediately that

$$\inf_{h \in \mathcal{H}_x^y} \Phi_N(h) \geq \mathbb{Q}_N(\Delta) \inf_{h \in \mathcal{H}_x^y} \Phi_\Delta(h) = \mathbb{Q}_N(\Delta) \inf_{h \in \mathcal{H}_x^y(\Delta)} \Phi_\Delta(h) \tag{2.10}$$

where $\mathcal{H}_x^y(\Delta) \equiv \{h : \Delta \to [0,1] : h(y) = 0, h(x) = 1\}$. Thus, using Theorem 2.2 for the process $\widetilde{P}_\Delta$, we see that

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] = \frac{1}{2\mathbb{Q}_N(y)} \inf_{h \in \mathcal{H}_x^y} \Phi_N(h) \geq \frac{1}{2\widetilde{\mathbb{Q}}_\Delta(y)} \inf_{h \in \mathcal{H}_x^y(\Delta)} \Phi_\Delta(h) = \widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \tag{2.11}$$

which proves the lemma. $\diamondsuit$

To make use of this lemma, we will choose Δ in a special way. Note that the simplest choice would be to choose Δ as one single path connecting y and x over the saddle point $z^*(y,x)$ in an optimal way. However, such a choice would produce a bound of the form $CN^{-1/2} \exp\left(-NF_N(z^*(y,x)) - F_N(y)\right)$ which differs from the upper bound by a factor $N^{d/2}$. I seems clear that in order to improve this bound we must choose Δ in such a way that it still provides "many" paths connecting y and x. To do this we proceed as follows. Let E be any number s.t. $F_N(z^*(y,x)) > E > \max\left(F_N(y), F_N(x)\right)$ (e.g. choose $E = F_N(z^*(y,x)) -$

$\frac{1}{2}\left(F_N(z^*(y,x)) - \max\left(F_N(y), F_N(x)\right)\right)$. Denote by D_y, D_x the connected components of the level set $\{x' \in \Gamma_N : F_N(x') \leq E\}$ that contain the points y, resp. x.

Note that of course we cannot, due to the discrete nature of the set Γ_N, achieve that the function F_N is constant on the actual discrete boundary of the sets D_y, D_x. The discrete boundary ∂D of any set $D \subset \Gamma_N$, will be defined as

$$\partial D \equiv \{x \in D | \exists y \in \Gamma_N \backslash D\,, \text{s.t.}\, (x,y) \in \Gamma_N\} \tag{2.12}$$

We have, however, that

$$\sup_{x' \in \partial D_y, x'' \in \partial D_x} |F_N(x') - F_N(x'')| \leq CN^{-1} \tag{2.13}$$

Next we choose a family of paths $\gamma_z : [0,1] \to \Gamma_N$, indexed by $z \in B \subset \mathcal{S}_N$ with the following properties:

i) $\gamma_z(0) \in \partial D_y$, $\gamma_z(1) \in \partial D_x$

ii) For $z \neq z'$, γ_z and $\gamma_{z'}$ are disjoint (i.e. they do not have common sites or common edges.

iii) F_N restricted to γ_z attains its maximum at z.

Of course we will choose the set $B \subset \mathcal{S}_N$ to be a small relative neighborhood in $\mathcal{S}_N$ of the saddle $z^*(y,x)$. In fact it will turn out to be enough to take B a disc of diameter $CN^{-1/2}$ so that its cardinality is bounded by $|B| \leq CN^{(d-1)/2}$.

For such a collection, we will set

$$\Delta \equiv D_x \cup D_y \cup \bigcup_{z \in B} V(\gamma_z) \tag{2.14}$$

where $V(\gamma_z)$ denotes the graph composed of the vertices that γ_z visits and the edges along which it jumps; the unions are to be understoof in the sense of the union of the corresponding subgraphs of Γ_N.

Lemma 2.3: With Δ defined above we have, for any $\delta > 0$, and for N large enough

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \geq \left(1 - CN^{d/2} e^{-N[F_N(z^*(y,x)) - E]}\right) \sum_{z \in B} \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)} \widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \tag{2.15}$$

Proof: All paths on Δ contributing to the event $\left\{\tau_x^y < \tau_y^y\right\}$ must now pass along one of the paths γ_z. Using the strong Markov property, we split the paths at the first arrival point in D_x which gives the equality

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] = \sum_{z \in B} \widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(1)}^y \leq \tau_{D_x \cup y}^y\right] \widetilde{\mathbb{P}}_\Delta\left[\tau_x^{\gamma_z(1)} < \tau_y^{\gamma_z(1)}\right] \tag{2.16}$$

By reversibility,

$$
\begin{aligned}
\widetilde{\mathbb{P}}_\Delta\left[\tau^y_{\gamma_z(1)} \le \tau^y_{D_x\cup y}\right] &= \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)}\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_y \le \tau^{\gamma_z(1)}_{D_x}\right] \\
&= \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)}\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{\gamma_z(0)} \le \tau^{\gamma_z(1)}_{D_x}\right]\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(0)}_y \le \tau^{\gamma_z(0)}_{D_x}\right]
\end{aligned}
\tag{2.17}
$$

where in the last line we used that the path going from $\gamma_z(1)$ to y without further visits to D_x must follow γ_z. Note further that we have the equality

$$
\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{\gamma_z(0)} \le \tau^{\gamma_z(1)}_{D_x}\right] = \widetilde{\mathbb{P}}_{\gamma_z}\left[\tau^{\gamma_z(1)}_{\gamma_z(0)} \le \tau^{\gamma_z(1)}_{\gamma_z(1)}\right]
\tag{2.18}
$$

where the right hand side is a purely one-dimensional object. We will now show that the probabilities $\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_y\right]$ and $\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(0)}_y \le \tau^{\gamma_z(0)}_{D_x}\right]$ are exponentially close to 1. To see this, write

$$
\begin{aligned}
1-\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_y\right] &= \widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_y < \tau^{\gamma_z(1)}_x\right] \\
&= \frac{\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_y < \tau^{\gamma_z(1)}_{x\cup\gamma_z(1)}\right]}{1-\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{\gamma_z(1)} < \tau^{\gamma_z(1)}_{x\cup y}\right]}
\end{aligned}
\tag{2.19}
$$

The latter equality arises from the by now usual decomposition of the path into multiple returns to $\gamma_z(1)$ and by summing the resulting geometric series. Now by reversibility, the denominator in (2.19) satisfies the bound

$$
\begin{aligned}
\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_y < \tau^{\gamma_z(1)}_{x\cup\gamma_z(1)}\right] &\le \sum_{x'\in B}\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{x'} < \tau^{\gamma_z(1)}_{x\cup\gamma_z(1)}\right] \\
&\le |B|\frac{\mathbb{Q}_\Delta(z^*(y,x))}{\mathbb{Q}_\Delta(\gamma_z(1))} = |B|\frac{\mathbb{Q}_N(z^*(y,x))}{\mathbb{Q}_N(\gamma_z(1))}
\end{aligned}
\tag{2.20}
$$

On the other hand,

$$
\begin{aligned}
1-\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{\gamma_z(1)} < \tau^{\gamma_z(1)}_{x\cup y}\right] &= \widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_{x\cup y} < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right] \\
&\ge \widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right] \\
&\ge \widetilde{\mathbb{P}}_\gamma\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right]
\end{aligned}
\tag{2.21}
$$

where γ is a a one dimensional path going from $\gamma_z(1)$ to x. We will show later that

$$
\widetilde{\mathbb{P}}_\gamma\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right] \ge CN^{-1/2}
\tag{2.22}
$$

Thus we get that

$$
\widetilde{\mathbb{P}}_\Delta\left[\tau^{\gamma_z(1)}_x < \tau^{\gamma_z(1)}_y\right] \ge 1 - CN^{d/2}e^{-N[F_N(z^*(y,x))-E]}
\tag{2.23}
$$

By the same procedure, we get also that

$$\widetilde{\mathbb{P}}_{\Delta}\left[\tau_{y}^{\gamma_z(0)} \leq \tau_{D_x}^{\gamma_z(0)}\right] \geq 1 - CN^{d/2}e^{-N[F_N(z^*(y,x))-E]} \tag{2.24}$$

Putting all these estimates together, we arrive at the affirmation of the lemma.◊

We are left to prove the lower bounds for the purely one-dimensional problems whose treatment is explained for instance in [vK]. In fact, we will show that

Proposition 2.4: Let γ_z be a one dimensional path such that F_N attains its maximum on γ_z at z. Then there is a constant $0 < C < \infty$ such that

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \geq CN^{-1/2}e^{-N[F_N(z)-F_N(\gamma_z(1))]} \tag{2.25}$$

Proof: Let $K \equiv |\gamma_z|$ denote the number of edges in the path γ_z. Let us fix the notation $\omega_0, \omega_1, \dots, \omega_K$, for the ordered sites of the path γ_z, with $\gamma_z(1) = \omega_0, \gamma_z(0) = \omega_{|\gamma_z|}$.

For any site ω_n we introduce the probabilities to jump to the right, resp. the left

$$p(n) = p_N(\omega_n, \omega_{n+1}), \quad q(n) = p_N(\omega_n, \omega_{n-1}) \tag{2.26}$$

We will first show that

Lemma 2.5: *With the notation introduced above,*

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] = \left[\sum_{n=1}^{K} \frac{\mathbb{Q}_N(\omega_0)}{\mathbb{Q}_N(\omega_n)} \frac{1}{p(n)}\right]^{-1} \tag{2.27}$$

Proof: Let us denote by $r(n)$ the solution of the boundary value problem

$$\begin{aligned} &r(n)(p(n)+q(n)) = p(n)r(n+1) + q(n)r(n-1), \quad \text{for } 0 < n < K \\ &r(0) = 0, \quad r(K) = 1 \end{aligned} \tag{2.28}$$

Obviously we have that

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] = p(0)r(1) \tag{2.29}$$

(2.28) has the following well know unique solution

$$r(n) = \frac{\sum_{k=1}^{n} \prod_{\ell=k}^{K-1} \frac{p(\ell)}{q(\ell)}}{\sum_{k=1}^{K} \prod_{\ell=k}^{K-1} \frac{p(\ell)}{q(\ell)}} \tag{2.30}$$

hence,

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] = \frac{p(0)\prod_{\ell=1}^{K-1} \frac{p(\ell)}{q(\ell)}}{\sum_{k=1}^{K} \prod_{\ell=k}^{K-1} \frac{p(\ell)}{q(\ell)}} = \frac{p(0)}{\sum_{k=1}^{K} \prod_{\ell=1}^{k-1} \frac{q(\ell)}{p(\ell)}} \tag{2.31}$$

Now reversibility reads $\mathbb{Q}_N(\omega_\ell)p(\ell) = \mathbb{Q}_N(\omega_{\ell+1})q(\ell+1)$, and this allows to simplify

$$\prod_{\ell=1}^{k-1} \frac{p(\ell)}{q(\ell)} = \frac{q(k)\mathbb{Q}_N(\omega_k)}{q(1)\mathbb{Q}_N(\omega_1)} \tag{2.32}$$

and finally

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau^{\gamma_z(1)}_{\gamma_z(0)} < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right] = \frac{1}{\mathbb{Q}_N(\omega_0)\sum_{k=1}^K \left[q(k)\mathbb{Q}_N(\omega_k)\right]^{-1}} \tag{2.33}$$

which is the assertion of the lemma. $\diamondsuit$

We are left to estimate the sum $\mathbb{Q}_N(\omega_0)\sum_{k=1}^K \frac{1}{q(k)\mathbb{Q}_N(\omega_k)}$ uniformly in K. Since $q(k) \geq c > 0$ for all $1 \leq k \leq K$, for an upper bound on this sum it is enough to consider

$$\mathbb{Q}_N(\omega_0)\sum_{k=1}^K \frac{1}{\mathbb{Q}_N(\omega_k)} = \frac{\mathbb{Q}_N(\omega_0)}{\mathbb{Q}_N(z)}\sum_{k=1}^K e^{-N[F_N(z)-F_N(\omega_k)]} \tag{2.34}$$

Now in the neighborhood of z, we can certainly bound

$$F_N(z) - F_N(\omega_k) \geq c\left(\frac{k}{N}\right)^2 \tag{2.35}$$

while elsewhere $F_N(z) - F_N(\omega_k) > \epsilon > 0$ (of course nothing changes if the paths have to pass over finitely many saddle points of equal height), and from this it follows immediately by elementary estimates that uniformly in K

$$\sum_{k=1}^K e^{-N[F_N(z)-F_N(\omega_k)]} \leq CN^{1/2} \tag{2.36}$$

which in turn concludes the proof of Proposition 2.4.[14] $\diamondsuit\diamondsuit$

Combining Proposition 2.4 with Lemma 2.3, we get that

$$\begin{aligned}\widetilde{\mathbb{P}}_\Delta\left[\tau^y_x < \tau^y_y\right] &\geq \left(1 - CN^{d/2}e^{-N[F_N(z^*(y,x))-E]}\right)\sum_{z\in B}\frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)}\frac{\mathbb{Q}_N(z)}{\mathbb{Q}_N(\gamma_z(1))}CN^{-1/2} \\ &= e^{-N[F_N(z^*(y,x))-F_N(y)]}\left(1 - CN^{d/2}e^{-N[F_N(z^*(y,x))-E]}\right)CN^{-1/2}\sum_{z\in B}e^{-N[F_N(z)-F_N(z^*(y,x))]}\end{aligned} \tag{2.37}$$

By our assumptions $F_N(z) - F_N(z^*(y,x))$ restricted to the surface $\mathcal{S}_N$ is bounded from above by a quadratic function in a small neighborhood of $z^*(y,x)$ and so, if B is chosen to be such a

[14] Of course we could easily be more precise and identify the constant in (2.36) to leading order with the second derivative of $F(z)$ in the direction of γ (see e.g. [vK] where this computation is given in the case of the continuum setting, and [KMST] where a formal asymptotic expansion is derived in the discrete case), but this would not really help us as we do not have the corresponding upper bound.

neighborhood, the lower bound claimed in Theorem 1.11 follows immediately by a standard gaussian approximation of the last sum. $\diamond\diamond$

3. Laplace transforms of transition times in the elementary situation

In this section we shall prove Theorem 1.10, which is our basic estimate for the Laplace transforms of elementary transition times. We shall need the sharp estimates on the transition probabilities which we obtained in the previous section based on Lemma 2.2. Combined with reversibility they lead to an estimate on the hitting time $\tau^x_{\mathcal{M}_N}$. This is the basic analytic result needed to estimate the Laplace transforms, using their usual representation as solutions of an appropriate boundary value problem. Let us recall the notation

$$G^x_{y,\Sigma}(u) = \mathbb{E}\left[e^{u\tau^x_y}\mathbb{1}_{\tau^x_y \le \tau^x_\Sigma}\right], \quad g^x_y(u) = G^x_{y,\mathcal{M}_N}(u)$$

In this section Σ will always denote a proper nonempty subset of Γ_N that contains $\mathcal{M}_N$. Moreover, we will assume that y is not in the interior of Σ, i.e. it is not impossible that y is reached before $\Sigma\backslash y$ from x, since otherwise $G^x_{y,\Sigma}(u) = 0$ trivially.

To prove Theorem 1.10, it is enough to show that

$$g^x_y(u) \le C_0 N^{d/2} e^{N[F_N(z^*(x,y))-F_N(x)]} \tag{3.1}$$

for real and positive $u \le cN^{-d-3/2}$. Note that $z^*(x,y)$ is defined in (1.30) in such a way that $z^*(x,y)$ equals to x if y can be reached from x without passing a point at which F_N is larger than $F_N(x)$. Analyticity then follows since $g^x_y(u)$ is a Laplace transform of the distribution of a positive random variable, and the estimates for $k \ge 1$ follow using Cauchy's inequality.

In the sequel we will fix $y \in \Sigma$ and $\mathcal{M}_N \subset \Sigma \subset \Gamma_N$. It will be useful to define the function

$$v_u(x) = \begin{cases} G^x_{y,\Sigma}(u) & \text{for} \quad x \notin \Sigma \\ 1 \quad \text{for} \quad x = y \\ 0 \quad \text{for} \quad x \in \Sigma\backslash y. \end{cases} \tag{3.2}$$

As explained in the introduction, $v_u(x)$ is analytic near $u = 0$ (so far without any control in N on the region of analyticity).

Similarly, we define the function

$$w_0(x) = \begin{cases} \mathbb{E}[\tau^x_{\mathcal{M}_N}] & \text{for} \quad x \notin \mathcal{M}_N \\ 0 \quad \text{for} \quad x \in \mathcal{M}_N \end{cases} \tag{3.3}$$

Observe that as a consequence of Lemma 1.1 of the introduction we get (for any $x, y \in \Gamma_N, \Sigma \subset \Gamma_N$) that

$$G^x_{y,\Sigma}(u) = e^u P_N(x,y) + e^u \sum_{z \notin \Sigma} P_N(x,z) G^z_{y,\Sigma}(u). \tag{3.4}$$

Using this identity one readily deduces that v_u is the unique solution of the boundary value problem

$$(1-e^u P_N)v_u(x)=0 \quad (x\notin\Sigma), \qquad v_u(y)=1, \quad v_u(x)=0 \quad (x\in\Sigma\backslash y). \tag{3.5}$$

and, in the same way, w_u is the unique solution of

$$(1-P_N)w_0(x)=1 \quad (x\notin\mathcal{M}_N), \qquad w_0(x)=0 \quad (x\in\mathcal{M}_N). \tag{3.6}$$

We shall use these auxiliary functions to prove the crucial

Lemma 3.1: *There is a constant $C\in\mathbb{R}$ such that for all $y\in\Gamma_N$ and all N large enough*

$$T_N := \max_{y\in\Gamma_N}\mathbb{E}[\tau^y_{\mathcal{M}_N}] \leq CN^{d+1} \tag{3.7}$$

Proof: In view of the Kolmogorov forward equations it suffices to consider the case $y\notin\mathcal{M}_N$. We set $\Sigma=\mathcal{M}_N\cup y$, where $y\notin\mathcal{M}_N$. Then $v_0(x)$ defined in (3.2) solves the Dirichlet problem

$$\begin{aligned}&(1-P_N)v_0(x)=0 \quad (x\notin\Sigma)\\ &v_0(y)=1, \qquad v_0(x)=0 \quad (x\in\mathcal{M}_N)\end{aligned} \tag{3.8}$$

Moreover, (3.4) with $u=0$ and $x=y$S reads (since $G^y_{y,\Sigma}(0)=\mathbb{P}[\tau^y_y\leq\tau^y_\Sigma]$)

$$1-\mathbb{P}\left[\tau^y_\Sigma<\tau^y_y\right]=\sum_{z\in\Gamma_N}p_N(y,z)v_0(z) \tag{3.9}$$

which can be written as

$$(1-P_N)v_0(y)=\mathbb{P}[\tau^y_{\mathcal{M}_N}<\tau^y_y] \tag{3.10}$$

We shall use $v_0(x)$ as a fundamental solution for $1-P_N$ and, using the symmetry of P_N in $\ell^2(\Gamma_N,\mathbb{Q}_N)$, we get

$$\begin{aligned}\mathbb{Q}_N(y)\mathbb{P}[\tau^y_{\mathcal{M}_N}<\tau^y_y]\mathbb{E}[\tau^y_{\mathcal{M}_N}] &= \langle(1-P_N)v_0,w_0\rangle_{\mathbb{Q}}\\ &=\langle v_0,(1-P_N)w_0\rangle_{\mathbb{Q}}\\ &=\mathbb{Q}_N(y)+\sum_{x\notin\Sigma}\mathbb{Q}_N(x)\mathbb{P}[\tau^x_y<\tau^x_{\mathcal{M}_N}],\end{aligned} \tag{3.11}$$

where in the last step we have used equation (3.6) and the fact that $y\notin\mathcal{M}_N$. This gives the crucial formula for the expected hitting time in terms of Boltzmann factors and transition probabilities, namely

$$\mathbb{E}[\tau^y_{\mathcal{M}_N}]=\sum_{x\notin\Sigma}\frac{\mathbb{Q}_N(x)}{\mathbb{Q}_N(y)}\frac{\mathbb{P}[\tau^x_y<\tau^x_{\mathcal{M}_N}]}{\mathbb{P}[\tau^y_{\mathcal{M}_N}<\tau^y_y]}+\frac{1}{\mathbb{P}[\tau^y_{\mathcal{M}_N}<\tau^y_y]}. \tag{3.12}$$

We remark that in this sum only those values of x with $F_N(x) \leq F_N(y)$ contribute. To estimate the probabilities in equation (3.12) we choose, given the starting point $y \notin \mathcal{M}_N$, an appropriate minimum $z \in \mathcal{M}_N$ near y such that there is a path $\gamma : y \to z$ (of moderate cardinality) so that F_N attains its maximum on γ at y (note that such a z exists trivially always). Then the variational principle in equation (2.11) (with γ as the subgraph Δ) gives

$$\mathbb{P}[\tau^y_{\mathcal{M}_N} < \tau^y_y] \geq \mathbb{P}[\tau^y_z < \tau^y_y] \geq \widetilde{\mathbb{P}}_\gamma[\tau^y_z < \tau^y_y], \tag{3.13}$$

where the first inequality is a trivial consequence of $z \in \mathcal{M}_N$. But then Proposition 2.5 can be applied to get the lower bound

$$\mathbb{P}[\tau^y_{\mathcal{M}_N} < \tau^y_y] \geq CN^{-1/2} \tag{3.14}$$

for some constant C.

To estimate the other probability in (3.12) we use Corollary 1.9 to write, for $x \notin \Sigma$,

$$\mathbb{P}[\tau^x_y < \tau^x_{\mathcal{M}_N}] = \frac{\mathbb{P}[\tau^x_y < \tau^x_{\mathcal{M}_N \cup x}]}{\mathbb{P}[\tau^x_\Sigma < \tau^x_x]} \tag{3.15}$$

Since $\mathcal{M}_N \subset \Sigma$, we obtain from (3.14) that for $x \notin \Sigma$,

$$\mathbb{P}[\tau^x_\Sigma < \tau^x_x] \geq \mathbb{P}[\tau^x_{\mathcal{M}_N} < \tau^x_x] \geq CN^{-1/2} \tag{3.16}$$

Reversibility then gives the upper bound

$$\mathbb{P}[\tau^x_y < \tau^x_{\mathcal{M}_N \cup x}] = \frac{\mathbb{Q}_N(y)}{\mathbb{Q}_N(x)} \mathbb{P}[\tau^y_x < \tau^y_{\mathcal{M}_N \cup y}] \leq \min\left(1, \frac{\mathbb{Q}_N(y)}{\mathbb{Q}_N(x)}\right). \tag{3.17}$$

Thus, inserting (3.16) and (3.17) into (3.15) we obtain from the representation (3.12) that

$$\mathbb{E}[\tau^y_{\mathcal{M}_N}] \leq CN(1 + \sum_{x \notin \Sigma} 1) \leq CN^{d+1}. \tag{3.18}$$

for some constant C. This proves the lemma. $\diamondsuit$

Next we need an estimate on the Laplace transform $G^x_{y,\Sigma}(u)$. This will be obtained from an integral representation of our auxiliary function $v_u(x)$, choosing u smaller than the estimate on the inverse of the maximal expected time T_N obtained in Lemma 3.1. More precisely, we shall prove

Lemma 3.2: *Assume that $\mathcal{M}_N \subset \Sigma \subset \Gamma_N$. Then there is a constant $c > 0$ such that for all $u \leq cN^{-d-1}$ and all $x, y \in \Gamma_N$,*

$$G^x_{y,\Sigma}(u) \leq 2 \tag{3.19}$$

Furthermore, there are constants $b, c > 0$ such that for all $u \leq cN^{-d-3/2}$ and $y \in \Gamma_N \backslash \mathcal{M}_N$,

$$1 - G^y_{y,\Sigma}(u) \geq bN^{-1/2} \tag{3.20}$$

Proof: As mentioned in the beginning of this chapter we can assume without loss of generality that $y \in \partial\Sigma$. Then it follows from equation (3.5) that the function $w_u(x) := v_u(x) - v_0(x)$ solves the Dirichlet problem

$$\begin{aligned} &(1-P_N)w_u(x) = (1-P_N)v_u(x) = (1-e^{-u})v_u(x), \quad (x \notin \Sigma) \\ &w_u(x) = 0 \quad (x \in \Sigma) \end{aligned} \tag{3.21}$$

The relation between resolvent and semi-group gives the following representation for $x \notin \Sigma$

$$w_u(x) = \mathbb{E}\left[\sum_{t=0}^{\tau_\Sigma^x - 1} f(X_t)\right], \quad f(x) := (1-e^{-u})v_u(x) \tag{3.22}$$

that in turn yields the integral equation

$$v_u(x) = \mathbb{P}[\tau_y^x = \tau_\Sigma^x] + (1-e^{-u})\mathbb{E}\left[\sum_{t=0}^{\tau_\Sigma^x - 1} v_u(X_t)\right]. \tag{3.23}$$

for the function v_u. We can now use our a priori bounds from Lemma 3.1 on the expectation of the stopping time τ_Σ^x to extract an upper bound for the sup-norm of this function. Namely, setting $M(u) := \sup_{x\notin\Sigma} v_u(x)$ we obtain the estimate

$$M(u) \le 1 + |1-e^{-u}| \max_{x\in\Gamma_N} \mathbb{E}[\tau_\Sigma^x] M(u) \le 1 + \frac{1}{3}M(u), \tag{3.24}$$

where we have used that $|u| < cN^{-d-1}$ with c sufficiently small. This gives for $x \notin \Sigma$,

$$G_{y,\Sigma}^x(u) \le 3/2 \tag{3.25}$$

The estimate of the Laplace transform $G_{y,\Sigma}^x(u)$ is trivial for negative u or for $x \in \Sigma\backslash\partial\Sigma$. In the case $x \in \partial\Sigma$, (3.19) follows from (3.4), using (3.25).

To prove the estimate (3.20) on the Laplace transform $G_{y,\Sigma}^y$ of the recurrence time to the boundary point $y \in \partial\Sigma$, (in particular $y \in \Sigma\backslash\mathcal{M}_N$ under our assumptions), observe that for any $\delta > 0$, there exists $c >$ such that for $|u| < cN^{-d-3/2}$, using Lemma 3.2 to estimate $\mathbb{E}[\tau_\Sigma^x] \le \mathbb{E}[\tau_{\mathcal{M}_N}^x]$ from above, it follows that

$$\mathbb{E}\left[\sum_{t=0}^{\tau_\Sigma^x - 1} (1-e^{-u})\right] \le \delta N^{-1/2} \tag{3.26}$$

Inserting this estimate and the a priori bound (3.19) into (3.23) together with the a priori bound gives then that

$$G_{y,\Sigma}^x(u) \le \mathbb{P}[\tau_y^x = \tau_\Sigma^x] + 2\delta N^{-1/2}, \tag{3.27}$$

Inserting (3.27) into (3.4), which represents $G^y_{y,\Sigma}(u)$ via $G^x_{y,\Sigma}(u)$ for $x \notin \Sigma$, it follows that modulo $\delta N^{-1/2}$ one has, for $|u| < cN^{-d-3/2}$,

$$\begin{aligned} 1 - G^y_{y,\Sigma}(u) &\geq 1 - e^u P_N(y,y) - e^u \sum_{x \notin \Sigma} p_N(y,x)\mathbb{P}[\tau^x_y = \tau^x_\Sigma] \\ &= 1 - \mathbb{P}[\tau^y_\Sigma = \tau^y_y] \\ &= \mathbb{P}[\tau^y_\Sigma < \tau^y_y]. \end{aligned} \tag{3.28}$$

Since $\mathcal{M}_N \subset \Sigma$ and $y \in \Sigma \backslash \mathcal{M}_N$ one obtains from (3.16) that

$$1 - G^y_{y,\Sigma}(u) \geq \mathbb{P}[\tau^y_{\mathcal{M}_N} < \tau^y_y] - 2\delta N^{-1/2} \geq bN^{-1/2} \tag{3.29}$$

for some $b > 0$, choosing δ sufficiently small in equation (3.26). This proves Lemma 3.3. $\diamondsuit$

We are now ready to give the

Proof of Theorem 1.10: Note that when $F_N(x) = F_N(z^*(x,y))$, Lemma 3.3 already provides the desired (actually a sharper) estimate. It remains to consider the case $z^*(x,y) \neq x$.

Here we can, as in the proof of Theorem 1.11 in Section 2, construct a discrete separating hyper-surface $\mathcal{S}_N$ containing the minimal saddle $z^*(x,y)$ and separating y and x. Since the process starting at x must hit $\mathcal{S}_N$ before hitting y, path splitting at $\mathcal{S}_N$ gives

$$g^x_y(u) = \sum_{z \in \mathcal{S}_N} G^x_{z,\Omega}(u) g^z_y(u), \qquad \Omega = \mathcal{M}_N \cup \mathcal{S}_N. \tag{3.30}$$

We treat the cases $x \in \mathcal{M}_N$ and $x \notin \mathcal{M}_N$ separately. In the latter case we need an additional renewal argument, while in the former all loops are suppressed since the process is killed upon arrival at $x \in \mathcal{M}_N$. For $x \notin \mathcal{M}_N$ the renewal equation (1.26) reads

$$G^x_{z,\Omega}(u) = (1 - G^x_{x,\Omega}(u))^{-1} G^x_{z,\Omega\cup x}(u) \tag{3.31}$$

By Lemma 3.2 and reversibility we have

$$G^x_{z,\Omega\cup x}(u) = \frac{\mathbb{Q}_N(z)}{\mathbb{Q}_N(x)} G^z_{x,\Omega}(u) \leq 2\frac{\mathbb{Q}_N(z)}{\mathbb{Q}_N(x)}, \tag{3.32}$$

using Lemma 3.3. Combining (3.32) and (3.20) of Lemma 3.2 we get from the renewal equation (3.31)

$$G^x_{z,\Omega}(u) \leq CN^{1/2} \frac{\mathbb{Q}_N(z)}{\mathbb{Q}_N(x)}, \quad (z \in \mathcal{S}_N, u \leq cN^{-d-3/2}) \tag{3.33}$$

for $c > 0$ sufficiently small.

If $x \in \mathcal{M}_N$, we directly apply the reversibility argument to $G^x_{z,\Omega}(u)$ (without renewal) and obtain a sharper estimate, i.e. (3.33) with $N^{1/2}$ deleted on the right hand side.

Inserting (3.33) into (3.30) and using (3.19) to estimate the Laplace transform $g_y^z(u) = G^z_{y,\mathcal{M}_N}(u)$ we finally get, for $u \leq cN^{-d-3/2}$,

$$g_y^x(u) \leq CN^{1/2}\mathbb{Q}_N(x)^{-1} \sum_{z\in\mathcal{S}_N} \mathbb{Q}_N(z) = \mathcal{O}(N^{d/2})e^{-N(F_N(z^*(x,y))-F_N(x))}, \tag{3.34}$$

where the last equality is obtained by a standard gaussian approximation as (2.37). All estimates on the derivatives $k \geq 1$ now follow from Cauchy's inequality and the obvious extension of our estimates to complex values of u. This completes the proof of Theorem 1.10.$\diamondsuit\diamondsuit$

4. Valleys, trees and graphs

In this chapter we provide the setup for the inductive treatment of the global problem. Although this description is not particularly original, and is essentially equivalent to the approach of Freidlin and Wentzell [WF], we give a self-contained exposition of our version that we find particularly suitable for the specific problem at hand. To keep the description as simple as possible, we make the assumption that F_N is "generic" in the sense that no accidental symmetries or other "unusual" structures occur. This will be made more precise below. For the case of a random system, this appears a natural assumption.

4.1. The valley structure and its tree-representation

The first important concept will be that of the set of essential saddle points.

Definition 4.1: *We call a point $z \in \Gamma_N$ an essential saddle point, if the connected (according to the graph structure on Γ_N) component of the level set $\Lambda_z \equiv \{x \in \Gamma_N : F_N(x) \leq F_N(z)\}$ that contains z falls into two*[15] *disconnected components when z is removed from it.*

These two components are called "valleys" and denoted by $V^{\pm}(z)$, with the understanding that

$$\inf_{x\in V^+(z)} F_N(x) < \inf_{x\in V^-(z)} F_N(x) \tag{4.1}$$

holds. We denote by $\mathcal{E}_N$ the set of all essential saddle points.

With any valley we associated two characteristics: its "height",

$$h(V^i(z)) \equiv F_N(z) \tag{4.2}$$

and its "depth"

$$d(V^i(z)) \equiv F_N(z) - \inf_{x\in V^i(z)} F_N(x) \tag{4.3}$$

The essential topological structure of the landscape F_N is encoded in a tree structure that we now define on the set $\mathcal{M}_N \cup \mathcal{E}_N$. To construct this, we define, for any essential saddle

[15] Under the assumption **G1** no points will exist where this connected component would fall into several components.

$z \in \mathcal{E}_N$, the two points

$$z_z^{\pm} = \begin{cases} \arg\max_{z_i \in \mathcal{E}_N \cap V^{\pm}(z)} F_N(z_i), & \text{if} \quad \mathcal{E}_N \cap V^{\pm}(z) \neq \emptyset \\ \mathcal{M}_N \cap V^{\pm}(z), & \text{else} \end{cases} \tag{4.4}$$

(note that necessarily the set $\mathcal{M}_N \cap V^{\pm}(z)$ consists of a single point if $\mathcal{E}_N \cap V^{\pm}(z) = \emptyset$). Now draw a link from any essential saddle to the two points $z_z^{\pm}$. This produces a connected tree, $\mathcal{T}_N$, with vertex set $\mathcal{E}_N \cup \mathcal{M}_N$ having the property that all the vertices with coordination number 1 (endpoints) correspond to local minima, while all other vertices are essential saddle points. An alternative equivalent way to construct this tree is by starting from below: Form each local minimum, draw a link to the lowest essential saddle connecting it to other minima. Then from each saddle point that was reached before, draw a line to the lowest saddle above it that connects it to further minima. Continue until exhaustion. We see that under our assumption of non-degeneracy, both procedures give a unique answer. (But note that in a random system the answer can depend on the value of N!)

The tree $\mathcal{T}_N$ induces a natural hierarchical distance between two points in $\mathcal{E}_N \cup \mathcal{M}_N$, given by the length of the shortest path on $\mathcal{T}_N$ needed to join them. We will also call the "level" of a vertex its distance to the root, z_0.

The properties of the long-time behaviour of the process will be mainly read-off from the structure of the tree $\mathcal{T}_N$ and the values of F_N on the vertices of $\mathcal{T}_N$. However, this information will not be quite sufficient. In fact, we will see that the information encoded in the tree contains all information on the time-scales of "exits" from valleys; what is still missing is how the process descends into a neighboring valley after such an exit. It turns out that all we need to know in addition is which minimum the process visits *first* after crossing a saddle point. This point deserves some discussion. First, we note that the techniques we have employed so far in this paper are insufficient to answer such a question. Second, it is clear that without further assumptions, there will not be a deterministic answer to this question; that is, in general it is possible that the process has the option to visit various minima first with certain probabilities. If this situation occurs, one should compute these probabilities; this appears, however, an exceedingly difficult task that is beyond the scope of the present paper. We will therefore restrict our attention to the situation where F_N is such that there is always one minimum that is visited first with overwhelming probability. To analyse this problem, we need to discuss an issue that we have so far avoided, that of sample path large deviations for the (relatively) short time behaviour of our processes. A detailed treatment of this problem is given in [BG2] and, as this issues concerns the present paper only marginally, we will refer the interested reader to that paper and keep the discussion here to a minimum. What we will need here is that for "short" times, i.e. for times $t = TN$, $T < \infty$, the process starting at any point x_0 at time 0 will remain (arbitrarily) close (on the macroscopic scale) to certain deterministic trajectories $x(t, x_0)$ with probability exponentially close to one[16]. These trajectories are solutions of certain differential equations involving the function F. In the continuum approximation they are just the gradient flow of F, i.e. $\frac{d}{dt}x(t) = -\nabla F(x(t)), \quad x(0) = x_0$, and while the equations are more complicated in

[16] Convergence of this type of processes to deterministic trajectories was first proved on the level of the law of large numbers by Kurtz [Ku].

the discrete case they are essentially of similar nature. In particular, all critical points are always fixpoints. We will assume that the probability to reach and δ-neighborhood of the boundary of Λ in finite time T will be exponentially small for all fixed T. We will assume further that at each essential saddle the deterministic paths starting in a neighborhood of z lead into uniquely specified minima within the two valleys connected through z. As we will see, these paths will determine the behaviour of the process.

We will incorporate these information in our graphical representation by decorating the tree by adding two yellow[17] arrows pointing from each essential saddle to the minima in each of the branches of the tree emanating from it into which the deterministic paths lead. (These branches are essentially obtained by following the gradient flow from the saddle into the next minimum on both sides.) We denote the tree decorated with the yellow arrows by $\widetilde{\mathcal{T}}_N$.

4.2. Construction of the transition process

We are in principle interested in questions like "how long does the process take to get from one minimum to another?". This question is more subtle than one might think. A related question, that should precede the previous one, is actually "how does the process get from one minimum to another one?", and we will first make this question precise and provide an answer.

We recall that in (1.18) we have given a representation of the process going from y to x in terms of a random walk on the minima. As we pointed out there, this representation was not extremely useful. We will now show that it is possible to give another decomposition of the process that is much more useful.

Let us consider the event $\mathcal{F}(x,y) \equiv \{\tau_x^y < \infty\}$ with $x, y \in \mathcal{M}_N$. Of course this event has probability one. We now describe an algorithm that will allow to decompose this event, up to a set of exponentially small measure, into a sequence of "elementary" transitions of the form

$$\mathcal{F}(x_i, z_i, x_{i+1}) \equiv \left\{\tau_{x_{i+1}}^{x_i} \leq \tau_{\mathcal{T}_{z_i,x_i}^c \cap \mathcal{M}_N}^{x_i}\right\} \tag{4.5}$$

where $x_i, x_{i+1} \in \mathcal{M}_N$, z_i is the first common ancestor of x_i and x_{i+1} in the tree $\mathcal{T}_N$, and $\mathcal{T}_{z_i,x_i}$ is the branch of $\mathcal{T}_N$ emanating from z_i that contains both x_i and $\mathcal{T}_{z_i,x_i}^c \equiv \mathcal{T}_N \backslash \mathcal{T}_{z_i,x_i}$. We will write $\mathcal{T}_z$ for the union of all branches emanating from z. The motivation for this definition is contained in the following

Proposition 4.2: *Let $x, y \in \mathcal{T}_{z,y} \cap \mathcal{M}_N$, and $\bar{y} \in \mathcal{T}_{z,x}^c \cap \mathcal{M}_N$. Then there is a constant $C < \infty$ such that*

$$\mathbb{P}\left[\tau_{\bar{y}}^x < \tau_y^x\right] \leq \inf_{z' \in \mathcal{T}_{z,y} \backslash z} C\sqrt{N} e^{-N[F_N(z) - F_N(z')]} \tag{4.6}$$

Remark: Note that by construction we have $F_N(z) - F_N(z') > 0$ for all lower saddles in the branch $\mathcal{T}_{z,y}$. Thus the proposition asserts that with enormous probability, the process starting from any minimum in a given valley visits all other minima in that same valley before visiting any minimum outside of this valley. As a matter of fact, the same also holds

[17] This color was used in the original drawing on a blackboard in the office of V. G. in the CPT, Marseille, and is retained here for historical reasons.

for general points. Thus what the proposition says is that up to the first exit from a valley, the process restricted to this valley behaves like an ergodic one.

Proof: We use as usual path-splitting at the consecutive returns of the process to x. This yields

$$\begin{aligned} \mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{y}^{x}\right] &= \frac{\mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{x}^{x}, \tau_{x}^{x} < \tau_{y}^{x}\right] + \mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{y}^{x}, \tau_{y}^{x} < \tau_{x}^{x}\right]}{\mathbb{P}\left[\tau_{y}^{x} < \tau_{x}^{x}, \tau_{y}^{x} < \tau_{\bar{y}}^{x}\right] + \mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{x}^{x}, \tau_{\bar{y}}^{x} < \tau_{y}^{x}\right]} \\ &= \frac{\mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{x}^{x}\right] - \mathbb{P}\left[\tau_{y}^{x} < \tau_{\bar{y}}^{x} < \tau_{x}^{x}\right]}{\mathbb{P}\left[\tau_{y}^{x} < \tau_{x}^{x}\right] + \mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{x}^{x} < \tau_{y}^{x}\right]} \\ &\leq \frac{\mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{x}^{x}\right]}{\mathbb{P}\left[\tau_{y}^{x} < \tau_{x}^{x}\right]} \end{aligned} \tag{4.7}$$

Now using the upper and lower bounds from Theorem 1.11, we get

$$\mathbb{P}\left[\tau_{\bar{y}}^{x} < \tau_{y}^{x}\right] \leq C\sqrt{N} e^{-N[F_N(z) - F_N(z')]} \tag{4.8}$$

where z' is the lowest saddle connecting x and y. (4.8) yields the proposition.$\diamondsuit$

Proposition 4.2 implies in particular that the process will visit the lowest minimum in a given valley before exiting from it, with enormous probability. This holds true on any level of the hierarchy of valleys. These visits at the lowest minima thus serve as a convenient breakpoint to organize any transition into elementary steps that start at a lowest minimum of a given valley and exit just into the next hierarchy. This leads to the following definition.

Definition 4.3: *A transition $\mathcal{F}(x,z,y)$ is called admissible, if*

i) x is the deepest minimum in the branch $\mathcal{T}_{z,x}$, i.e. $F_N(x) = \inf_{x' \in \mathcal{T}_{z,x}} F_N(x)$.

ii) z and y are connected by a yellow arrow in $\widetilde{\mathcal{T}}_N$.

Remark: We already understand why an admissible transition should start at deepest minimum: if it would not, we would know that the process would first go there, and we could decompose it into a first transition to this lowest minimum, and then an admissible transition to y. What we do not see yet, is where the condition on the endpoint (the yellow arrow) comes from. The point here is that upon exiting the branch $\mathcal{T}_{z,x}$, the process has to arrive somewhere in the other branch emanating from z. We will show later that with exponentially large probability this is the first minimum which the deterministic path staring from z leads to.

Proposition 4.4: *If $\mathcal{F}(x,z,y)$ is an admissible transition, then there exists $K_N > 0$, satisfying $N^{1-\alpha} K_N \uparrow \infty$ such that*

$$\mathbb{P}\left[\mathcal{F}(x,z,y)\right] \geq 1 - e^{-NK_N} \tag{4.9}$$

Remark: To proof this proposition, we will use the large deviation estimates that require the stronger regularity assumptions $R2, R4$, as well as the structural assumptions discussed in the beginning of this section. These are to some extent technical and clearly not necessary. Alternatively, one can replace these by the assumption that Proposition 4.4 holds, i.e. for any $z \in \mathcal{E}$ and $x \in \mathcal{T}_{z,x}$ there is a unique $y \in \mathcal{T}_{z,x}^c \cup \mathcal{M}_N$ such that (4.9) holds.

Proof: The proof is based on the fact that the process will, with probability one, hit the set $\mathcal{T}_{z,x}^c$ eventually. Thus, if we show that given x and z, for all $\tilde y \in \mathcal{T}_{z,x}^c$ with $\tilde y \neq y$, $\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{\mathcal{T}_{z,x}^c \setminus \tilde y}^x\right]$ is exponentially small, the proposition follows. To simplify the notation, let us set $I = \mathcal{T}_{z,x}^c \cap \mathcal{M}_N$. Note that the case $\tilde y \notin \mathcal{T}_z$ is already covered by Proposition 4.2, so we assume that $\tilde y \in \mathcal{T}_z$. Using Corollary 1.9

$$\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{I\setminus \tilde y}^x\right] = \frac{\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{I\setminus \tilde y \cup x}^x\right]}{\mathbb{P}\left[\tau_I^x < \tau_x^x\right]} \tag{4.10}$$

By reversibility,

$$\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{I\setminus \tilde y \cup x}^x\right] = e^{N[F_N(x)-F_N(\tilde y)]}\mathbb{P}\left[\tau_x^{\tilde y} < \tau_I^{\tilde y}\right] \tag{4.11}$$

Now construct the separating hyper-surface $\mathcal{S}_N$ passing through z as in the proof of Theorem 1.11. Then

$$\mathbb{P}\left[\tau_x^{\tilde y} < \tau_I^{\tilde y}\right] = \sum_{z'\in\mathcal{S}_N} \mathbb{P}\left[\tau_{z'}^{\tilde y} \leq \tau_{I\cup\mathcal{S}_N}^{\tilde y}\right]\mathbb{P}\left[\tau_x^{z'} < \tau_I^{z'}\right] \tag{4.12}$$

Putting all things together, and using reversibility once more, we see that

$$\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{I\setminus \tilde y}^x\right] = \frac{1}{\mathbb{P}[\tau_I^x < \tau_x^x]}\sum_{z'\in\mathcal{S}_N} e^{-N[F_N(z')-F_N(x)]}\mathbb{P}\left[\tau_{\tilde y}^{z'} \leq \tau_{I\cup\mathcal{S}_N}^{z'}\right]\mathbb{P}\left[\tau_x^{z'} < \tau_I^{z'}\right] \tag{4.13}$$

Using that $\mathbb{P}[\tau_I^x < \tau_x^x] \geq \mathbb{P}[\tau_{y'}^x < \tau_x^x]$ for any $y' \in I$, together with the lower bound of Theorem 1.11 and the trivial bound $\mathbb{P}\left[\tau_x^{z'} < \tau_I^{z'}\right] \leq 1$, we see that

$$\begin{aligned}\mathbb{P}\left[\tau_{\tilde y}^x < \tau_{I\setminus \tilde y}^x\right] &\leq C^{-1}N^{-(d-2)/2}\sum_{z'\in\mathcal{S}_N} e^{-N[F_N(z')-F_N(z)]}\mathbb{P}\left[\tau_{\tilde y}^{z'} \leq \tau_{I\cup\mathcal{S}_N}^{z'}\right]\\ &\leq C^{-1}N^{-(d-2)/2}\sum_{\substack{z'\in\mathcal{S}_N\\ F_N(z')-F_N(z)\leq K_N}} e^{-N[F_N(z')-F_N(z)]}\mathbb{P}\left[\tau_{\tilde y}^{z'} \leq \tau_{I\cup\mathcal{S}_N}^{z'}\right]\\ &+ N^{d/2+1}e^{-NK_N}\end{aligned} \tag{4.14}$$

Under our assumptions the condition $F_N(z') - F_N(z) \leq K_N$ implies that $|z'-z| \leq C'\sqrt{K_N}$, i.e. all depends on the term $\mathbb{P}\left[\tau_{\tilde y}^{z'} \leq \tau_{I\cup\mathcal{S}_N}^{z'}\right]$ for z' very close to the saddle point z. Now, heuristically, we must expect that with large probability the process will first arrive at the minimum that is reached from z' by following the 'gradient' of F_N.

Let us now show that this is the case. Let us first remark that using the same arguments as in the proof of Proposition 4.2, it is clear that the probability that the process will hit the set where $F_N(x') > F_N(z^*(x,\tilde y)) + \delta'$, $\delta' > 0$, before reaching $\tilde y$ is of order $\exp(-\delta' N)$ so that this possibility is negligible. Denote the complement of this set by $L_{\delta'}$. Now consider the ball D_δ of radius δ centered at z, where δ should be large enough such that the intersection of $L_{\delta'}$ with $\mathcal{S}_N$ is well contained in the interior of D_δ. The set $L_{\delta'} \cap D_\delta$ is then separated by $\mathcal{S}_N$ into two parts, and we call C_δ the part that is on the side of I. According to the previous discussion, if the process is to reach I, it has to pass through the surface $\Sigma \equiv \partial C_\delta \cap \partial D_\delta$. Finally, let R_δ denote the ball of radius δ centered at y. Note first that

$$\begin{aligned} &\mathbb{P}\left[\tau_{\tilde y}^{z'} < \tau_{I\cup\mathcal{S}_N\cup L_{\delta'}^c}^{z'}\right] \\ &\leq \mathbb{P}\left[\tau_{\tilde y}^{z'} < \tau_{I\cup\mathcal{S}_N\cup L_{\delta'}^c}^{z'}, \tau_{\tilde y}^{z'} < \tau_{R_\delta}^{z'}\right] + \mathbb{P}\left[\tau_{\tilde y}^{z'} < \tau_{I\cup\mathcal{S}_N\cup L_{\delta'}^c}^{z'}, \tau_{\tilde y}^{z'} > \tau_{R_\delta}^{z'}\right] \\ &\leq \mathbb{P}\left[\tau_{\tilde y}^{z'} < \tau_{R_\delta\cup\mathcal{S}_N\cup L_{\delta'}^c}^{z'}\right] + \sum_{x''\in R_\delta} \mathbb{P}\left[\tau_{\tilde y}^{x''} < \tau_{y}^{x''}\right] \end{aligned} \tag{4.15}$$

The second term is exponentially small by standard reversibility arguments. It remains to control the first.

$$\begin{aligned} &\mathbb{P}\left[\tau_{\tilde y}^{z'} < \tau_{R_\delta\cup\mathcal{S}_N\cup L_{\delta'}^c}^{z'}\right] \\ &= \sum_{x'\in\Sigma} \mathbb{P}\left[\tau_{x'}^{z'} \leq \tau_{\Sigma}^{z'}\right] \mathbb{P}\left[\tau_{\tilde y}^{x'} < \tau_{R_\delta}^{x'}\right] \\ &\leq |\Sigma| \sup_{x'\in\Sigma} \mathbb{P}\left[\tau_{\tilde y}^{x'} < \tau_{R_\delta}^{x'}\right] \end{aligned} \tag{4.16}$$

Now under the assumptions on F, for all $x' \in \Sigma$, the deterministic paths $x(t,x')$ reach R_δ in finite time T (i.e. in a microscopic time TN) without getting close to $\tilde y$. Therefore, for some $\rho > 0$

$$\mathbb{P}\left[\tau_{\tilde y}^{x'} < \tau_{R_\delta}^{x'}\right] \leq \mathbb{P}\left[\sup_{t\in[0,NT]} |X_t - x(t,x')| > \rho \big| X_0 = x'\right] \tag{4.17}$$

But the large deviation theorem of [BG2] implies that there exists $\epsilon \equiv \epsilon(\rho,T) > 0$, such that

$$\limsup_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}\left[\sup_{t\in[0,NT]} |X_t - x(t,x')| > \rho \big| X_0 = x'\right] \leq -\epsilon(\rho,T) \tag{4.18}$$

so that, e.g., for all large enough N,

$$\mathbb{P}\left[\sup_{t\in[0,NT]} |X_t - x(t,x')| > \rho \big| X_0 = x'\right] \leq e^{-N\epsilon(\rho,T)/2} \tag{4.19}$$

Tt then suffices to observe that

$$\sum_{y'\in I} \mathbb{P}\left[\tau_{y'}^{x} < \tau_{I\setminus y'}^{x}\right] = \mathbb{P}\left[\tau_I^x < \infty\right] = 1 \tag{4.20}$$

and so, since $\mathbb{P}\left[\tau_{\tilde y}^{x} < \tau_{I\setminus\tilde y}^{x}\right] \leq \exp(-NK_N)$, for all $\tilde y \neq y$, $\mathbb{P}\left[\tau_{y}^{x} < \tau_{I\setminus y}^{x}\right] > 1-\exp(-NK_N)$.$\diamondsuit$

Note that the above argument also shows that if $\mathcal{F}(x,z,y)$ is admissible, and $y' \in I$, then

$$\mathbb{P}\left[\tau_{y'}^{x} \leq \tau_{I\cup x}^{x}\right] \leq \mathbb{P}\left[\tau_{y}^{x} \leq \tau_{I\cup x}^{x}\right] e^{-NK_N} \tag{4.21}$$

Theorem 4.5: *Let $x, y \in \mathcal{M}_N$. Then there is a unique sequence of admissible events $\mathcal{F}(x_i, z_i, x_{i+1})$, $i = 1, \dots, k$, such that*[18]

$$\{\tau_y^x < \infty\} \supset \{\tau_y^x = \sum_{i=1}^{k} \tau_{x_{i+1}}^{x_i} < \infty\} \cap \bigcap_{i=1}^{k} \mathcal{F}(x_i, z_i, x_{i+1}) \tag{4.22}$$

and such that the sequences are free of cycles, i.e. the points $x_i, i = 1, \dots k+1$ are all distinct. Moreover, there is a strictly positive constant K_N, such that

$$\mathbb{P}\left[\{\tau_y^x = \sum_{i=1}^{k} \tau_{x_{i+1}}^{x_i} < \infty\} \cap \bigcap_{i=1}^{k} \mathcal{F}(x_i, z_i, x_{i+1})\right] \geq 1 - e^{-NK_N} \tag{4.23}$$

Proof: There is a simple algorithm that allows to construct the sequence of admissible transitions. Let z be the first common ancestor of x and y in $\mathcal{T}_N$. First we notice that we will 'never' (that is to say with exponentially small probability) visit a minimum that is not contained in the two branches emanating from z before visiting all of $\mathcal{T}_z$. Given this restriction, starting from x, we make the maximal admissible transition, i.e. one traverses the highest possible saddle for which the starting point is a lowest minimum of its branch. This leads to some point x_2, from which we continue as before, with the restriction that the first common ancestor of x_2 and y now determines the maximal allowed transition. This process is continued until an admissible transition reaches y. It is clear that this algorithm determines a sequence of admissible transitions. We have to show that this is the only one containing no loops.

Note first that the condition that no transition leaves the branches of the youngest common ancestor follows since Proposition 4.4 ensures that the target point is reached before exit from this valley with probability close to one. It is easy to see that we should always choose the maximal admissible transition. Suppose we start in some point that is the deepest minimum in some valley that does not contain the target point, and we perform an admissible transition that does not exit from this valley. Then we must return to this point at least once more before reaching the target which means that our sequence of admissible transitions contains a loop. Therefore, at each step the choice of the next admissible transition is uniquely determined.

Finally, from Proposition 4.4 the estimate (4.23) follows immediately. $\diamondsuit$

Remark: We see that the same type of reasoning would also allow us to deal with degenerate situations where e.g. integral curves of the gradient bifurcate and transitions to several points y may have non-vanishing probabilities. The picture of the deterministic sequence of admissible transitions should then be replaced by a (cycle free) random process of admissible transitions. The precise computation of the corresponding probabilities would however require more refined estimates than those presented here (except if this can be done by using exact symmetries).

[18] We hope the notation used here is self-explanatory: E.g. $\{\tau_y^x < \infty\}$ stands for $\cup_{t<\infty}\{X_0 = x, X_1 \neq y, \dots, X_{t-1} \neq y, X_t = y\}$.

Remark: Theorem 4.5 asserts that for fixed large N a transition occurs along an essentially deterministic sequence of admissible transitions. When dealing with the dynamics of system with quenched disorder, this deterministic (with respect to the Markov chain) sequence will however depend on the realization of the quenched disorder, *and on the volume* N. In a typical situation, this will give rise to a manifestation of dynamical "chaotic size dependence" (in the spirit of Newman and Stein (see e.g. [NS] for an overview).

In the sequel we will always be interested in computing the times (expected or distribution) of transitions conditioned on the canonical chain of admissible transitions constructed in Theorem 4.5. We mention that in general, these do not coincide with the unconditional transition times. Namely, in general, there can occur unlikely excursions (into deeper valleys) that take extremely long times so that they dominate e.g. the expected transition times. Physically, this is clearly not the most interesting quantity.

5. Transition times of admissible transitions

From the discussion above it is clear that the most basic quantities we need to control to describe the long time behaviour of our processes are the times associated with an admissible transition. Note that an admissible transition $\mathcal{F}(x,z,y)$ can also be considered as a first exit from the valley associated with the saddle z and the minimum x. We proceed in three steps, considering first the expectations the expectations of these times, then the Laplace transforms, and finally the probability distributions themselves.

5.1. Expected times of admissible transitions.

A first main result is the following theorem.

Theorem 5.1: *Let $\mathcal{F}(x,z,y)$ be an admissible transition, and assume that x is a generic quadratic minimum. Then there exist finite positive constants c, C such that*

$$\begin{aligned} \mathbb{E}\tau_y^x \mathbb{1}_{\mathcal{F}(x,z,y)} &\leq CNe^{N[F_N(z)-F_N(x)]}(1+e^{-NK_N}) \\ \mathbb{E}\tau_y^x \mathbb{1}_{\mathcal{F}(x,z,y)} &\geq cN^{1/2}e^{N[F_N(z)-F_N(x)]}(1-e^{-NK_N}) \end{aligned} \tag{5.1}$$

where K_N satisfies $N^{1-\alpha}K_N \uparrow \infty$ for some $\alpha > 0$.

Remark: In dimension $d=1$ the upper bound captures the true behaviour (see e.g. [vK] where the expected transition time in $d=1$ is computed in the continuous case. Note that the extra factor N in our estimates is just a trivial scaling factor between the microscopic discrete time and the appropriate macroscopic time scale). We expect that the upper bound has the correct behaviour in all dimensions.

Before proving the theorem, we will prove some more crude but more general estimates. For this we introduce some notation. Let $I \subset \mathcal{M}_N$. We define

$$d_I(x,y) \equiv \inf_{x' \in I \cup y} [F_N(z^*(x,x')) - F_N(x)] \tag{5.2}$$

to be the effective depth of a valley associated with the minimum x with exclusion at the set I. Here $z^*(x,y)$ is the lowest saddle connecting x and y, as defined in (1.30). Note that equivalently we have that

$$z^*(x,y) \equiv \arg \inf_{\substack{z\in\mathcal{E}_N \\ x\in\mathcal{T}_{z,x}, y\in\mathcal{T}_{z,y}}} F_N(z) \tag{5.3}$$

With these notations we will show the following

Lemma 5.2: *Let $I \subset \mathcal{M}_N$. There exist $C < \infty$ and $\kappa < \infty$ such that for any $x, y \in \mathcal{M}_N$, we have that for all N large enough,*

$$\begin{aligned} &\mathbb{E}\left[\tau_y^x | \tau_y^x \leq \tau_I^x\right] \leq CN^\kappa \\ &+ CN^\kappa \sup_{x'\in\mathcal{M}_N\backslash\{I\cup y\}} \min\left(e^{-N[F_N(z^*(x,x'))-F_N(z^*(x,y))]}, 1\right) e^{Nd_I(x',y)} \end{aligned} \tag{5.4}$$

If the set $\mathcal{M}_N\backslash\{I\cup y\}$ is empty, we use the convention that the sup takes the value one. Moreover, we set $z^*(x,x) \equiv x$.

Remark: In order to understand (5.4), it is helpful to realise that

$$\min\left(e^{-N[F_N(z^*(x,x'))-F_N(z^*(x,y))]}, 1\right) \approx \mathbb{P}\left[\tau_{x'}^x < \tau_y^x | \tau_y^x \leq \tau_I^x\right] \tag{5.5}$$

is essentially the probability that the process visits x' on its way to y. Also, $e^{Nd_I(x',y)}$ should be thought of as the time it takes to reach I or y from x'. Note that formula (5.4) implies that the process conditioned on avoiding a set I will leave a valley in a time that corresponds to reaching the point in I "closest" to it. This may at the first glance appear surprising.

Proof: Note first that if $I = \mathcal{M}_N$ the estimate follows from the estimates of Theorems 1.10 and 1.11. starting point to prove the lemma by downward induction over the size of the set I. Actually, the structure of the induction is a bit more complicated. We have to distinguish the cases where the starting point x is contained in the exclusion set I and where it is not. We will then proceed in two steps:

(i) Show that if (5.4) holds for all $J \subset \mathcal{M}_N$ with cardinality $|J| = k$ and all $x, y \in \mathcal{M}_N$, and if (5.4) holds for all J of cardinality $|J| = k-1$ for all $y \in \mathcal{M}_N$ and $x \notin J \cup y$, then (5.4) holds for all I with cardinality $|I| = k-1$ and all $x, y \in \mathcal{M}_N$.

(ii) Show that if (5.4) holds for all J with cardinality $|J| = k$ and all $x, y \in \mathcal{M}_N$, then (5.4) holds for all J of cardinality $|J| = k-1$ for all $y \in \mathcal{M}_N$ and $x \notin J \cup y$.

If we can establish both steps, we can conclude that since (5.4) holds for $I = \mathcal{M}_N$ and all $x, y \in \mathcal{M}_N$, it holds for all $I \subset \mathcal{M}_N$.

We now proof both assertions. Note that C and κ will denote in the course of the proof generic finite numerical constants. We will not keep track of the changes of their values in the course of the induction.

Case (i): $(x \in J \cup y)$. We can assume without loss that $y \notin J$.

$$
\begin{aligned}
&\mathbb{E}\left[\tau_y^x \mathbb{1}_{\tau_y^x < \tau_J^x}\right] \\
&= \mathbb{E}\left[\tau_y^x \mathbb{1}_{\tau_y^x \leq \tau_{\mathcal{M}_N}^x}\right] + \sum_{x' \in \mathcal{M}_N \backslash J \backslash \{x,y\}} \mathbb{E}\left[(\tau_{x'}^x + \tau_y^{x'}) \mathbb{1}_{\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x} \mathbb{1}_{\tau_y^{x'} < \tau_J^{x'}}\right] \\
&= \mathbb{E}\left[\tau_y^x \mathbb{1}_{\tau_y^x \leq \tau_{\mathcal{M}_N}^x}\right] + \sum_{x' \in \mathcal{M}_N \backslash J \backslash \{x,y\}} \Bigg(\mathbb{E}\left[\tau_{x'}^x \mathbb{1}_{\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x}\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right] \\
&+ \mathbb{P}\left[\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x\right] \mathbb{E}\left[\tau_y^{x'} \mathbb{1}_{\tau_y^{x'} < \tau_J^{x'}}\right]\Bigg)
\end{aligned}
\tag{5.6}
$$

The first summand in (5.6), divided by the probability $\mathbb{P}[\tau_y^x < \tau_J^x]$, produces a term of the order of the first terms in (5.4), by Theorems 1.10 and 1.11. The second term in (5.6), again divided by the same probability, is at most of the order N^κ. Thus the only dangerous term is the last one. We see that we have to bound, for any of the x' occuring in (5.6), a term of the form

$$
\frac{\mathbb{P}\left[\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x\right] \mathbb{E}\left[\tau_y^{x'} \mathbb{1}_{\tau_y^{x'} < \tau_J^{x'}}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} = \frac{\mathbb{P}\left[\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \mathbb{E}\left[\tau_y^{x'} | \tau_y^{x'} < \tau_J^{x'}\right] \tag{5.7}
$$

The conditional expectation on the right hand side satisfies (5.4) by the inductive hypothesis. Now we have that

$$
\begin{aligned}
&\mathbb{P}\left[\tau_y^x < \tau_J^x\right] \geq C N^{(d-2)/2} e^{-N[F_N(z^*(x,y)) - F_N(x)]} \\
&\mathbb{P}\left[\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x\right] \leq C N^{(d-1)/2} e^{-N[F_N(z^*(x,x')) - F_N(x)]}
\end{aligned}
\tag{5.8}
$$

and

$$
\mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right] \leq \min\left(C\sqrt{N} \inf_{x'' \in J} e^{-N[F_N(z^*(x',y)) - F_N(z^*(x',x''))]}, 1\right) \tag{5.9}
$$

This last inequality is proven using first that $\mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right] \leq \inf_{x'' \in J} \mathbb{P}\left[\tau_y^{x'} < \tau_{x''}^{x'}\right]$ Now Proposition 4.2 and Theorem 1.11 give (5.9). Therefore we get

$$
\begin{aligned}
&\frac{\mathbb{P}\left[\tau_{x'}^x \leq \tau_{\mathcal{M}_N}^x\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \mathbb{E}\left[\tau_y^{x'} | \tau_y^{x'} < \tau_J^{x'}\right] \\
&\leq C^2 N^\kappa e^{N F_N(z^*(x,y))} \min\left(e^{-N F_N(z^*(x',y))}, e^{-N F_N(z^*(x,x'))}\right) \\
&\times \left(\sup_{x'' \in \mathcal{M}_N \backslash J \backslash y} \min\left(e^{N[F_N(z^*(x',y)) - F_N(z^*(x',x''))]}, 1\right) e^{N d_J(x'',y)} + C N^\kappa\right)
\end{aligned}
\tag{5.10}
$$

We will bound each term in the supremum over x'' and treat all the four cases corresponding to the different possibilities in the minimum separately. A crucial observation is the following general statement that is tied to the tree structure of the set of saddle points: Whenever

x, x', x'' are minima, $F_N(z^*(x,x')) < F_N(z^*(x,x''))$ implies that $z^*(x',x'') = z^*(x,x'')$. This observation will be used repeatedly in what follows.

a) $F_N(z^*(x',y)) \geq F_N(z^*(x,x'))$, $F_N(z^*(x',y)) \leq F_N(z^*(x',x''))$. Then

$$\begin{aligned} & e^{NF_N(z^*(x,y))} \min\left(e^{-NF_N(z^*(x',y))}, e^{-NF_N(z^*(x,x'))}\right) \\ & \times \min\left(e^{N[F_N(z^*(x',y))-F_N(z^*(x',x''))]}, 1\right) = e^{N[F_N(z^*(x,y))-F_N(z^*(x',x''))]} \end{aligned} \tag{5.11}$$

But in this case we must have $z^*(x,x'') = z^*(x'x'')$, and $F_N(z^*(x,y)) < F_N(z^*(x,x''))$ so that the left hand side of (5.11) in fact equals

$$e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]} = \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]}, 1\right) \tag{5.12}$$

b) $F_N(z^*(x',y)) \geq F_N(z^*(x,x'))$, $F_N(z^*(x',y)) > F_N(z^*(x',x''))$.
Here the left hand side of (5.11) equals

$$e^{N[F_N(z^*(x,y))-F_N(z^*(x',y))]} = 1 \tag{5.13}$$

since we must have that $z^*(x',y) = z^*(y,x)$. Moreover, it is also clear that $F_N(z^*(x,y)) > F_N(z^*(x,x''))$, so that (5.13) is also equal to
$\min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]}, 1\right)$.

c) $F_N(z^*(x',y)) < F_N(z^*(x,x'))$, $F_N(z^*(x',y)) < F_N(z^*(x',x''))$.
Here we get

$$e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))+F_N(z^*(x',y))-F_N(z^*(x',x''))]} \tag{5.14}$$

There are two sub-cases to consider:

c.1) $F_N(z^*(x,x')) > F_N(z^*(x',x''))$.
In that case we must have $z^*(x,x') = z^*(x,x'') = z^*(x,y)$. Thus (5.14) is

$$\leq e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]} = 1 = \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]}, 1\right) \tag{5.15}$$

c.2) $F_N(z^*(x,x')) < F_N(z^*(x',x''))$.
Then $z^*(x',x'') = z^*(x,x'')$ and $F_N(z^*(x,y)) \leq F_N(z^*(x,x''))$, so that (5.14) satisfies

$$\leq e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]} = \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]}, 1\right) \tag{5.16}$$

d) $F_N(z^*(x',y)) < F_N(z^*(x,x'))$, $F_N(z^*(x',y)) \geq F_N(z^*(x',x''))$.
Here we obtain

$$\begin{aligned} & e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]} = e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]} = \\ & \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x''))]}, 1\right) \end{aligned} \tag{5.17}$$

since $z^*(x,x') = z^*(x,x'') = z^*(x,y)$.

We have seen that in all cases we get the claimed bound.

To complete the proof we need to turn to the case

Case (ii): $x \notin J \cup y$**.** Using the by now familiar renewal argument, we get

$$\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_J^x\right] = \frac{\mathbb{E}\left[\tau_y^x \mathbb{1}_{\tau_y^x < \tau_{J\cup x}^x}\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} + \frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x < \tau_{J\cup y}^x}\right] \mathbb{P}\left[\tau_y^x < \tau_{J\cup x}^x\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]^2} \tag{5.18}$$

Since $\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right] \geq \mathbb{P}\left[\tau_y^x < \tau_{J\cup x}^x\right]$, and $J \cup x$ has cardinality k, the first term in (5.18) satisfies the desired bound by the induction hypothesis. For the second term we have

$$\frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x < \tau_{J\cup y}^x}\right] \mathbb{P}\left[\tau_y^x < \tau_{J\cup x}^x\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]^2} \leq \frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x < \tau_{J\cup y}^x}\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} \tag{5.19}$$

Since $\mathbb{P}\left[\tau_x^x < \tau_{J\cup y}^x\right] \geq 1 - e^{-\mathcal{O}(N)}$ we can replace the numerator by the corresponding conditional expectation. Since $J \cup y$ has cardinality k, we can use the induction hypothesis for this term. On the other hand, by Theorem 1.11,

$$\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right] \geq C N^{(d-2)/2} \sup_{x'' \in J\cup y} e^{-N[F_N(z^*(x,x'')) - F_N(x)]} \tag{5.20}$$

This gives

$$\begin{aligned} \frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x < \tau_{J\cup y}^x}\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} \leq C^2 N^\kappa \Bigg(& \inf_{x'' \in J\cup y} e^{N[F_N(z^*(x,x'')) - F_N(x)]} + \\ & + \sup_{x' \in \mathcal{M}_N \setminus \{J\cup y\cup x\}} \inf_{x'' \in J\cup y} e^{N[F_N(z^*(x,x'')) - F_N(z^*(x,x')) + d_{J\cup y}(x',x)]} \Bigg) \end{aligned} \tag{5.21}$$

We want to show that this is equal to the left hand side of (5.4). Note first that the first term in the bracket corresponds to the choice $x' = x$ in the sup in (5.4). To see that the remaining terms are o.k., we just have to observe that by definition

$$\begin{aligned} d_{J\cup y}(x',x) &\equiv \inf_{x'' \in J\cup y\cup x} \left[F_N(z^*(x',x'')) - F_N(x')\right] \\ &\leq \inf_{x'' \in J\cup y} \left[F_N(z^*(x',x'')) - F_N(x')\right] \\ &= d_J(x',y) \end{aligned} \tag{5.22}$$

Let x^* b any point in $J\cup y\cup x$ that realizes the minimum of $F_N(z^*(x',x''))$ over $x' \in J\cup y\cup x$.

a) $x^* \neq x$, $F_N(z^*(x',x^*)) < F_N(z^*(x',x))$.
In this case $d_{J\cup y}(x',x) = d_J(x',y)$. Moreover, we have that $z^*(x,x^*) = z^*(x',x)$, so that

$$\begin{aligned} &\inf_{x''\in J\cup y} e^{N[F_N(z^*(x,x''))-F_N(z^*(x,x'))+d_{J\cup y}(x',x)]} \\ &\leq \min\left(e^{N[F_N(z^*(x,x^*))-F_N(z^*(x,x'))]}, e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]}\right) e^{Nd_J(x',y)} \\ &= \min\left(1, e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]}\right) e^{Nd_J(x',y)} \end{aligned} \tag{5.23}$$

b) $x^* = x$.
Set $x^{**} \equiv \arg\inf_{x''\in J\cup y} F_N(z^*(x',x''))$.
Then $d_{J\cup y}(x',x) - d_J(x',y) = F_N(z^*(x',x)) - F_N(z^*(x',x^{**}))$. Now again $F_N(z^*(x,x')) \leq F_N(z^*(x',x^{**}))$ implies that
$F_N(z^*(x,x^{**})) \leq F_N(z^*(x',x^{**}))$. Therefore,

$$\begin{aligned} &\inf_{x''\in J\cup y} e^{N[F_N(z^*(x,x''))-F_N(z^*(x,x'))+d_{J\cup y}(x',x)]} \\ &\leq \min\left(e^{N[F_N(z^*(x,x^{**}))-F_N(z^*(x',x^{**}))]}, e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]}\right) e^{Nd_J(x',y)} \\ &\leq \min\left(1, e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]}\right) e^{Nd_J(x',y)} \end{aligned} \tag{5.24}$$

Thus in all cases we obtain the claimed estimate. Putting everything together we have finally proven Lemma 5.2.◇

Proof of Theorem 5.1: Lemma 5.2 can now be used to prove the theorem. For this, let $\mathcal{F}(x,z,y)$ be an admissible transition and fix $I = \mathcal{T}^c_{z,x}\backslash y \cap \mathcal{M}_N$. We rewrite (5.18) in the form

$$\begin{aligned} \mathbb{E}\left[\tau^x_y|\tau^x_y < \tau^x_I\right] &= \mathbb{E}\left[\tau^x_y|\tau^x_y < \tau^x_{I\cup x}\right] \frac{\mathbb{P}[\tau^x_y < \tau^x_{I\cup x}]}{\mathbb{P}[\tau^x_{I\cup y} < \tau^x_x]} \\ &+ \frac{\mathbb{E}\left[\tau^x_x|\tau^x_x < \tau^x_{I\cup y}\right]}{\mathbb{P}\left[\tau^x_{I\cup y} < \tau^x_x\right]} \frac{\mathbb{P}[\tau^x_x < \tau^x_{I\cup y}]\mathbb{P}[\tau^x_y < \tau^x_{I\cup x}]}{\mathbb{P}[\tau^x_{I\cup y} < \tau^x_x]} \end{aligned} \tag{5.25}$$

Corollary 1.9 gives

$$\frac{\mathbb{P}\left[\tau^x_y < \tau^x_{I\backslash y\cup x}\right]}{\mathbb{P}\left[\tau^x_{y\cup I} < \tau^x_x\right]} = \mathbb{P}\left[\tau^x_y < \tau^x_I\right] \tag{5.26}$$

Thus

$$\begin{aligned} \mathbb{E}\left[\tau^x_y|\tau^x_y < \tau^x_I\right] &= \mathbb{E}\left[\tau^x_y|\tau^x_y < \tau^x_{I\cup x}\right]\mathbb{P}\left[\tau^x_y < \tau^x_I\right] \\ &+ \frac{\mathbb{E}\left[\tau^x_x|\tau^x_x < \tau^x_{I\cup y}\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_x\right]} \frac{\mathbb{P}\left[\tau^x_y < \tau^x_x\right]}{\mathbb{P}\left[\tau^x_{I\cup y} < \tau^x_x\right]} \\ &\times \mathbb{P}\left[\tau^x_y < \tau^x_I\right]\mathbb{P}\left[\tau^x_x < \tau^x_{I\cup y}\right] \end{aligned} \tag{5.27}$$

But

$$\mathbb{P}\left[\tau^x_y < \tau^x_x\right] = \mathbb{P}\left[\tau^x_y < \tau^x_{I\cup x}\right] + \mathbb{P}\left[\tau^x_I < \tau^x_y < \tau^x_x\right] \tag{5.28}$$

where the last term satisfies, by virtue of (4.21),

$$\begin{aligned}\mathbb{P}\left[\tau_I^x < \tau_y^x < \tau_x^x\right] &= \sum_{y'\in I} \mathbb{P}\left[\tau_{y'}^x \leq \tau_{I\cup x\cup y}^x\right] \mathbb{P}\left[\tau_y^{y'} < \tau_x^{y'}\right] \\ &\leq \sum_{y'\in I} \mathbb{P}\left[\tau_{y'}^x \leq \tau_{I\cup x\cup y}^x\right] \\ &\leq \mathbb{P}\left[\tau_y^x < \tau_{I\cup x}^x\right] |I| e^{-NK_N}\end{aligned} \tag{5.29}$$

for some $K_N > 0$. This implies that

$$\begin{aligned}\frac{\mathbb{P}\left[\tau_y^x < \tau_x^x\right]}{\mathbb{P}\left[\tau_{I\cup y}^x < \tau_x^x\right]} &= \frac{\mathbb{P}\left[\tau_y^x < \tau_x^x\right] \mathbb{P}\left[\tau_y^x < \tau_I^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_{I\cup x}^x\right]} \\ &= \mathbb{P}\left[\tau_y^x < \tau_I^x\right] + |I| e^{-NK_N} = 1 + \mathcal{O}\left(e^{-NK_N}\right)\end{aligned} \tag{5.30}$$

Putting everything together we arrive at

$$\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_I^x\right] = \frac{\mathbb{E}\left[\tau_x^x | \tau_x^x < \tau_{I\cup y}^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_x^x\right]} \left(1 + \mathcal{O}(e^{-NK_N})\right) + \mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_{I\cup x}^x\right] \left(1 - \mathcal{O}(e^{-NK_N})\right) \tag{5.31}$$

Now we have already precise bounds on the denominator of the first term (seeSection 2), and using the upper bound from Lemma 5.2 we see that, under the assumptions of the theorem, the second term is by a factor $N^\kappa \exp(-K_N N)$ smaller than the first. It remains to estimate precisely the numerator in the first term. The essential idea here is to use the ergodic theorem. It may be useful to explain this first in a simpler situation where there is only a single minimum present and consider the quantity $g_y^y(u)$. Let $D \subset \Gamma_N$ be the local valley associated to y, that is the connected component of the level set of the saddle point that connects y to the rest of the world. The basic idea is to show that the expected recurrence time at y (without visits at other points of $\mathcal{M}_N$) is up to exponentially small errors equal to the same time of another Markov chain $\widetilde{X}_D(t)$ with state space D with transition rates $\widetilde{p}_D(x,z)$ defined as in (2.4) and whose invariant measure, $\widetilde{\mathbb{Q}}_D$, is easily seen to be just $\mathbb{Q}_N$ conditioned on D, i.e. $\widetilde{\mathbb{Q}}_D(x) \equiv \mathbb{Q}_N(x)/\mathbb{Q}_N(D)$ for any $x \in D$. Then, by the ergodic theorem, we have that

$$\widetilde{\mathbb{E}}_D \tau_y^y = \frac{1}{\widetilde{\mathbb{Q}}_D(y)} = \frac{\mathbb{Q}_N(D)}{\mathbb{Q}_N(y)} \tag{5.32}$$

This quantity can be estimated very precisely via sharp large deviation estimates. It will typically exhibits a behaviour of the form $CN^{d/2}$.

To arrive at this comparison, we simply divide the paths in our process into those reaching the boundary of D and those who don't, i.e. we write

$$g_y^y(u) = \mathbb{E}\left[e^{u\tau_y^y} 1\!\!1_{\tau_y^y \leq \tau_{\mathcal{M}_N}^y} 1\!\!1_{\tau_{\partial D}^y < \tau_y^y}\right] + \mathbb{E}\left[e^{u\tau_y^y} 1\!\!1_{\tau_{\partial D}^y > \tau_y^y}\right] \tag{5.33}$$

Let us denote by D^+ and D^- the two sets obtained by adding and removing, respectively, one layer of points to, resp. from, D. Note that on the event $\{\tau_{\partial D}^y > \tau_y^y\}$ the processes $X(t)$ and $\widetilde{X}_{D^+}(t)$ have the same law until time τ_y^y, so that

$$\mathbb{E}\left[e^{u\tau_y^y} 1\!\!1_{\tau_{\partial D}^y > \tau_y^y}\right] = \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y} 1\!\!1_{\tau_{\partial D}^y > \tau_y^y}\right] \leq \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right] \tag{5.34}$$

We will show that this is the dominant term in (5.33), the first summand on the right being exponentially small. Indeed

$$\begin{aligned}\mathbb{E}\left[e^{u\tau_y^y}\mathbb{1}_{\tau_y^y\leq\tau_{\mathcal{M}_N}^y}\mathbb{1}_{\tau_{\partial D}^y<\tau_y^y}\right] &= \sum_{z\in\partial D}\mathbb{E}\left[e^{u(\tau_z^y+\tau_y^z)}\mathbb{1}_{\tau_z^y\leq\tau_{y\cup\partial D}^y}\mathbb{1}_{\tau_y^z\leq\tau_{\mathcal{M}_N}^z}\right]\\ &= \sum_{z\in\partial D}\mathbb{E}\left[e^{u\tau_z^y}\mathbb{1}_{\tau_z^y\leq\tau_{y\cup\partial D}^y}\right]\mathbb{E}\left[e^{u\tau_y^z}\mathbb{1}_{\tau_y^z\leq\tau_{\mathcal{M}_N}^z}\right]\end{aligned} \tag{5.35}$$

Using Theorem 1.10, for small enough u, the first factor is bounded by $const.N^\kappa e^{-N[F_N(z)-F_N(y)]}$, while the second is bounded by $const.N^\kappa$. This gives the desired upper bound

$$g_y^y(u)\leq\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right]+CN^\kappa e^{-N[F_N(z^*)-F_N(y)]} \tag{5.36}$$

where z^* denotes the lowest saddle point in ∂D. To get the corresponding lower bound, just note that

$$\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\mathbb{1}_{\tau_{\partial D}^y>\tau_y^y}\right]=\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right]-\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\mathbb{1}_{\tau_{\partial D}^y\leq\tau_y^y}\right] \tag{5.37}$$

But the last term in (5.37) can be treated precisely as in (5.35), so that we arrive at

$$g_y^y(u)\geq\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right]-CN^\kappa e^{-N[F_N(z^*)-F_N(y)]} \tag{5.38}$$

Differenting and using reversibility and the upper bounds from Theorem 1.10, as well as the obvious lower bound

$$g_y^y(0)=1-\mathbb{P}\left[\tau_{\mathcal{M}}^y<\tau_y^y\right]\geq 1-\sum_{x\in\mathcal{M}_N\backslash y}g_x^y(0)\geq 1-|\mathcal{M}_N|N^{d-1}e^{-N[F_N(z^*)-F_N(y)]} \tag{5.39}$$

gives then that

$$\frac{\dot g_y^y(0)}{g_y^y(0)}=\frac{\mathbb{Q}_N(D^+)}{\mathbb{Q}_N(y)}+\mathcal{O}(N^\kappa)e^{-N[F_N(z^*)-F_N(y)]} \tag{5.40}$$

The same ideas can now be carried over to the estimation of the return time in an admissible situation, using the estimates from Lemma 5.2.

Proposition 5.3: *Let $\mathcal{F}(x,z,y)$ be an admissible transition. Let D denote the level set of the saddle z. Then there are finite constants C,k and $K_N>0$ such that*

$$\mathbb{E}\left[\tau_x^x|\tau_x^x<\tau_{I\cup y}^x\right]=\frac{\mathbb{Q}_N(D^+)}{\mathbb{Q}_N(x)}+\mathcal{O}\left(e^{-NK_N}\right) \tag{5.41}$$

where $I=\mathcal{T}_{z,x}^c\backslash y\cap\mathcal{M}_N$.

Proof: Basically, the proof goes as outlined above. With D defined as the level set of the saddle z, we can decompose

$$\mathbb{E}\left[\tau_x^x\mathbb{1}_{\tau_x^x<\tau_{I\cup y}^x}\right]=\mathbb{E}\left[\tau_x^x\mathbb{1}_{\tau_x^x<\tau_{I\cup y}^x}\mathbb{1}_{\tau_x^x<\tau_{\partial D}^x}\right]+\mathbb{E}\left[\tau_x^x\mathbb{1}_{\tau_x^x<\tau_{I\cup y}^x}\mathbb{1}_{\tau_x^x>\tau_{\partial D}^x}\right] \tag{5.42}$$

The first summand gives precisely

$$\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x<\tau_{I\cup y}^x} \mathbb{1}_{\tau_x^x<\tau_{\partial D}^x}\right] = \mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x<\tau_{\partial D}^x}\right] \tag{5.43}$$

so that from this term alone we would get the same estimate as in (5.40). We have to show that the second term does not give a relevant contribution. Note that as in (5.35) we can split paths at the first visits to ∂D. This gives

$$\begin{aligned}\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x<\tau_{I\cup y}^x} \mathbb{1}_{\tau_x^x>\tau_{\partial D}^x}\right] = \sum_{z'\in\partial D} \mathbb{E}\left[\tau_{z'}^x \mathbb{1}_{\tau_{z'}^x\le\tau_{\partial D}^x} \mathbb{1}_{\tau_{z'}^x<\tau_x^x}\right] \mathbb{P}\left[\tau_x^{z'} < \tau_{I\cup y}^{z'}\right]\\ + \mathbb{P}\left[\tau_{z'}^x \le \tau_{\partial D}^x, \tau_{z'}^x < \tau_x^x\right] \mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}<\tau_{I\cup y}^{z'}}\right]\end{aligned} \tag{5.44}$$

Now $\mathbb{P}\left[\tau_{z'}^x \le \tau_{\partial D}^x, \tau_{z'}^x < \tau_x^x\right]$ is bounded by $e^{-N[F_N(z)-F_N(x)]}$, and by reversibility $\mathbb{E}\left[\tau_{z'}^x \mathbb{1}_{\tau_{z'}^x\le\tau_{\partial D}^x} \mathbb{1}_{\tau_{z'}^x<\tau_x^x}\right] \le e^{-N[F_N(z)-F_N(x)]}\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}<\tau_{\partial D}^{z'}}\right]$, so all we have to show is that the two quantities $\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}<\tau_{\partial D}^{z'}}\right]$ and $\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}<\tau_{I\cup y}^{z'}}\right]$ (which are more or less the same) are not too large. But this follows from our previous bounds by splitting the process going from z' to x at its first visit to a point in $\mathcal{T}_{z,x} \cap \mathcal{M}_N$, e.g.

$$\begin{aligned}&\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}<\tau_{I\cup y}^{z'}}\right] = \mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\tau_x^{z'}\le\tau_{\mathcal{M}_N}^{z'}}\right]\\ &+ \sum_{x'\in\mathcal{M}_N\setminus I}\left(\mathbb{E}\left[\tau_{x'}^{z'} \mathbb{1}_{\tau_{x'}^{z'}\le\tau_{\mathcal{M}_N}^{z'}}\right] \mathbb{P}\left[\tau_x^{x'} < \tau_{I\cup y}^{x'}\right] + \mathbb{P}\left[\tau_{x'}^{z'} \le \tau_{\mathcal{M}_N}^{z'}\right] \mathbb{E}\left[\tau_x^{x'} \mathbb{1}_{\tau_x^{x'}<\tau_{I\cup y}^{x'}}\right]\right)\end{aligned} \tag{5.45}$$

Lemma 5.2 and Theorem 1.10 can now be used on the expectations in (5.45), and this implies the desired result.$\diamondsuit$

Now if (as we assume) x is a quadratic minimum of F_N, $\frac{\mathbb{Q}_N(D^+)}{\mathbb{Q}_N(x)} = CN^{d/2}$, and using this together with Theorem 4.1 we get the estimates of Theorem 5.1.$\diamondsuit\diamondsuit$

Remark: The reader will have observed that we could also prove lower bounds for more general transitions, complementing Lemma 5.2. But the point is that these would depend in a complicated way on the global specifics of the function F_N, contrary to the situation of admissible transitions for which we get the very simple estimates of Theorem 5.1. The beauty of the construction lies in some sense in the fact that the general "worst case" upper bounds of Lemma 5.2 suffice to obtain the precise estimates of the theorem.

5.2 Laplace transforms of transition times of admissible transitions

Theorem 5.1 gives precise estimates on the expected transition times for an elementary transition. We will now show that as expected, the distribution of these transition times is asymptotically exponential. This will be done by controlling the Laplace transforms for small arguments.

Theorem 5.4: *Let $\mathcal{F}(x,z,y)$ be an admissible transition. Set $\bar\tau_y^x \equiv \mathbb{E}\left[\tau_y^x|\mathcal{F}(x,z,y)\right]$. Then*

$$\mathbb{E}\left[e^{v\tau_z^x/\bar\tau_y^x}|\mathcal{F}(x,z,y)\right] = \frac{1}{1-v} + e^{-NK_N} f(v) \tag{5.46}$$

where for any $\delta > 0$, for N large enough, f is bounded and analytic in the domain $|Re(v)| < 1-\delta, |Im(v)| < \exp(NK_N)$.

Proof: The main ingredient of the proof lies in controlling the analytic structure of the Laplace transforms. The procedure will be similar to that in the proof of Lemma 5.2, that is we consider the entire family of functions $G_{y,I}^x(u)$ and establish the corresponding domains by induction, starting with the case $I = \mathcal{M}_N$ where the analytic estimates of Theorem 1.10 hold. It will be convenient to use functions where the argument u has been properly rescaled. The naive expectation might be that the Laplace transform will exist for values of u up to the inverse of the corresponding expected transition time. However, this is not so. The point is that Laplace transforms are much more sensitive to "deep valleys" than the expected times for which such valleys contribute less if they are unlikely to be visited. However the Laplace transform will only partly benefit from this, but simply explode at a value corresponding to the deepest valley that is at all allowed to be visited.

We introduce some more notation for convenience. Set

$$t_I(x,y) \equiv e^{Nd_I(x,y)} \tag{5.47}$$

and

$$T_I(y) \equiv \sup_{x\in\mathcal{M}_N\backslash\{I\cup y\}} t_I(x,y) \tag{5.48}$$

With the notation of Section 1, we define

$$\widehat{G}_{y,I}^x(v) \equiv \frac{G_{y,I}^x\left(v/T_I(y)\right)}{\mathbb{P}\left[\tau_y^x \leq \tau_I^x\right]} \tag{5.49}$$

The following key lemma gives us control over how this happens. It is the analogue of Lemma 5.2.

Lemma 5.5: *Let $I \subset \mathcal{M}_N$, and let $x,y \in \mathcal{M}_N$. Then $G_{y,I}^x(u)$ can be represented in the form*

$$\begin{aligned} &\widehat{G}_{y,I}^x(v) = a_{x,y,I}^0(v/T_I(y)) \\ &+ \sum_{x'\in\mathcal{M}_N\backslash\{I\cup y\}} \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]},1\right) a_{x,y,I}^{x'}\left(v\frac{t_I(x',y)}{T_I(y)}\right) \end{aligned} \tag{5.50}$$

where $a_{x,y,I}^0$ and $a_{x,y,I}^{x'}$, for any $x' \in \mathcal{M}_N\backslash\{I\cup y\}$ are complex functions that have the properties (for finite constants C,κ):

(i) They are bounded by CN^κ and analytic in the domain $|Re(u)| < CN^{-\kappa}$, $|Im(u)| < N$,

(ii) They are real positive for real v.

Proof: An important corollary of analyticity are corresponding bounds on the derivatives. Namely, by a standard application of the Cauchy integral formula it follows that for any function a which is bounded and analytic in the domain $|Re(v)| < CN^{-\kappa}$, $|Im(v)| < N$, for $|Re(v)| < \frac{C}{2}N^{-\kappa}$ we have,

$$\left|\frac{d^n}{dv^n}a(v)\right| \leq n!C\frac{2^n N^{n\kappa}}{C^n} \sup_{v:Re(v)<CN^{-\kappa}} a(v) \tag{5.51}$$

This will be used repeatedly in the sequel.

We observe that for $I = \mathcal{M}_N$, the representation (5.50) holds trivially due to Theorem 1.10. This provides the starting point for our induction like in Lemma 5.2. Again we assume that the Lemma holds for all I of cardinality greater than or equal to ℓ, and we consider sets $J \subset \mathcal{M}_N$ of cardinality $\ell - 1$. As before, we first show that the case $x \in J$ reduces easily to the case $x \notin J$. Without loss of generality we assume $y \notin J$. Namely, in the former case,

$$\begin{aligned}\widehat{G}^x_{y,J}(v) = &\frac{g^x_y(v/T_J(y))}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]}\\ &+ \sum_{x''\in\mathcal{M}_N\backslash\{J\cup y\}} \frac{g^x_{x''}(v/T_J(y))\mathbb{P}\left[\tau^{x''}_y < \tau^{x''}_J\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]}\widehat{G}^{x''}_{y,J}(v)\end{aligned}$$

Inserting the induction hypothesis for the $\widehat{G}^{x'}_{y,J}(v)$, we get

$$\begin{aligned}\widehat{G}^x_{y,J}(v) = &\frac{g^x_y(v/T_J(y))}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]}\\ &+ \sum_{x'\in\mathcal{M}_N\backslash\{J\cup y\}} \sum_{x''\in\mathcal{M}_N\backslash\{J\cup y\}} \frac{g^x_{x''}(v/T_J(y))}{g^x_{x''}(0)} \frac{\mathbb{P}\left[\tau^x_{x''} \leq \tau^x_{\mathcal{M}_N}\right]\mathbb{P}\left[\tau^{x''}_y < \tau^{x''}_J\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]}\\ &\times \min\left(e^{N[F_N(z^*(x'',y))-F_N(z^*(x'',x'))]}, 1\right) a^{x'}_{x'',y,J}\left(v\frac{t_J(x',y)}{T_J(y)}\right)\end{aligned} \tag{5.52}$$

Remember that in the proof of Lemma 5.2 we have established that

$$\begin{aligned}&\frac{\mathbb{P}\left[\tau^x_{x''} \leq \tau^x_{\mathcal{M}_N}\right]\mathbb{P}\left[\tau^{x''}_y < \tau^{x''}_J\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]} \min\left(e^{N[F_N(z^*(x'',y))-F_N(z^*(x'',x'))]}, 1\right)\\ &\leq \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x,x'))]}, 1\right)\end{aligned} \tag{5.53}$$

which shows that (5.52) provides the claimed representation.

The more subtle part of the proof concerns the case $x \notin J$. Here we use of course that

$$G^x_{y,J}(u) = \frac{G^x_{y,J\cup x}(u)}{1 - G^x_{x,J\cup y}(u)} \tag{5.54}$$

which implies

$$
\begin{aligned}
\widehat{G}^x_{y,J}(v) &= \frac{\mathbb{P}\left[\tau^x_y < \tau^x_{J\cup x} | \tau^x_y < \tau^x_J\right] \widehat{G}^x_{y,J\cup x}\left(v\frac{T_{J\cup x}(y)}{T_J(y)}\right)}{1-\mathbb{P}\left[\tau^x_x < \tau^x_{J\cup y}\right] \widehat{G}^x_{x,J\cup y}\left(v\frac{T_{J\cup y}(x)}{T_J(y)}\right)} \\
&= \frac{\mathbb{P}\left[\tau^x_y < \tau^x_{J\cup x} | \tau^x_y < \tau^x_J\right] \widehat{G}^x_{y,J\cup x}\left(v\frac{T_{J\cup x}(y)}{T_J(y)}\right)}{\mathbb{P}\left[\tau^x_x > \tau^x_{J\cup y}\right] - \mathbb{P}\left[\tau^x_x < \tau^x_{J\cup y}\right] v\frac{T_{J\cup y}(x)}{T_J(y)} \int_0^1 d\theta \widehat{G}'^x_{x,J\cup y}\left(\theta v\frac{T_{J\cup y}(x)}{T_J(y)}\right)}
\end{aligned}
\tag{5.55}
$$

The numerator again poses no problem since it permits to obtain the desired representation by the inductive hypothesis. Potential danger comes from the denominator. But using the induction hypothesis, we see that

$$
\begin{aligned}
&\frac{T_{J\cup y}(x)}{T_J(y)} \widehat{G}'^x_{x,J\cup y}\left(\theta v \frac{T_{J\cup y}(x)}{T_J(y)}\right) \\
&= \frac{T_{J\cup y}(x)}{T_J(y)} \Bigg[\sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}} \min\left(e^{N[F_N(z^*(x,x))-F_N(z^*(x,x'))]}, 1\right) \\
&\times a'^{x'}_{x,x,J\cup y}\left(\theta v \frac{t_{J\cup y}(x',x)}{T_J(y)}\right) \frac{t_{J\cup y}(x',x)}{T_{J\cup y}(x)} + a^0_{x,x,J\cup y}\left(\frac{\theta v}{T_J(y)}\right) \frac{1}{T_{J\cup y}(x)} \Bigg] \\
&= \sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}} e^{N[F_N(x)-F_N(z^*(x,x'))]} \frac{t_{J\cup y}(x',x)}{T_J(y)} a'^{x'}_{x,x,J\cup y}\left(\theta v \frac{t_{J\cup y}(x',x)}{T_J(y)}\right) \\
&+ a^0_{x,x,J\cup y}\left(\frac{\theta v}{T_J(y)}\right) \frac{1}{T_J(y)}
\end{aligned}
\tag{5.56}
$$

All we need is to bound the modulus of this expression from above. This gives

$$
\begin{aligned}
&\left|\frac{T_{J\cup y}(x)}{T_J(y)} \widehat{G}'^x_{x,J\cup y}\right| \\
&\leq \sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}} e^{N[F_N(x)-F_N(z^*(x,x'))]} \frac{t_{J\cup y}(x',x)}{T_J(y)} CN^\kappa + \frac{cN^\kappa}{T_J(y)}
\end{aligned}
\tag{5.57}
$$

Now in the second part of the proof of Lemma 5.2 we have shown that

$$
\begin{aligned}
&e^{N[F_N(x)-F_N(z^*(x,x'))]} \frac{t_{J\cup y}(x',x)}{t_J(x',y)} \\
&\leq \mathbb{P}\left[\tau^x_{J\cup y} < \tau^x_x\right] \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x',y))]}, 1\right)
\end{aligned}
\tag{5.58}
$$

Thus we deduce from (5.57) that

$$
\begin{aligned}
&\left|\frac{T_{J\cup y}(x)}{T_J(y)} \widehat{G}'^x_{x,J\cup y}\right| \leq +\frac{cN^\kappa}{T_J(y)} \\
&+ CN^\kappa \mathbb{P}\left[\tau^x_{J\cup y} < \tau^x_x\right] \sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}} \frac{t_J(x',y)}{T_J(y)} \min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x',y))]}, 1\right)
\end{aligned}
\tag{5.59}
$$

Since $T_J(y) \geq t_J(x',y)$, we see that in any case the numerator in (5.55) will not vanish for $v < C^{-1}N^{-\kappa}$. Now we have to distinguish two cases. If there exists an x' such that $T_J(y) = t_J(x',y)$ and $\min\left(e^{N[F_N(z^*(x,y))-F_N(z^*(x',y))]},1\right) = 1$, then the considerations above show that

$$\frac{\mathbb{P}\left[\tau_y^x < \tau_{J\cup x}^x | \tau_y^x < \tau_J^x\right]}{1-\mathbb{P}\left[\tau_x^x < \tau_{J\cup y}^x\right]\widehat{G}_{x,J\cup y}^x\left(v\frac{T_{J\cup y}(x)}{T_J(y)}\right)} \tag{5.60}$$

is a function of v analytic in the strip $Re(v) < C^{-1}N^{-\kappa}$. Since there exists an x' with the above properties, in the representation of (5.50) such a term weighted with a factor one is allowed to appear, so we are some.

Otherwise, the denominator is far from vanishing and we can simply expand in geometric series. Since $t_{J\cup y}(x',x) \leq t_J(x',y)$, one sees that

$$\int_0^1 d\theta v \frac{t_J(x',y)}{T_J(y)} a'^{x'}_{x,x,J\cup y}\left(\theta v\frac{t_{J\cup y}(x',x)}{T_J(y)}\right) \equiv \tilde{a}^{x'}_{x,x,J\cup y}\left(\theta v\frac{t_J(x',y)}{T_J(y)}\right) \tag{5.61}$$

is an analytic function in the required domain and bounded by N^κ. Note that this requirement is the reason why we must restrict ourselves to bounded imaginary parts of v. Using that sums and products preserve the structure required in (5.50), we see that again we obtain the desired representation for our function. This concludes the proof of the lemma.$\diamondsuit$

We are now ready to prove Theorem 5.4. For this we have to improve the previous analysis in the case where x is the deepest minimum in the allowed set $\mathcal{M}_N\backslash I$. In that case $T_I(y)$ is strictly larger than any of the terms $t_{I\cup x}(x',y)$ and $t_{I\cup y}(x',x)$ and it will pay to use Taylor expansions to second order. Also, we will be more precise in the rescaling of the variables u and define

$$\widetilde{G}_{y,I}^x(v) \equiv \widehat{G}_{y,I}^x(vT_I(y)/\bar{T}) \tag{5.62}$$

with

$$\bar{T} \equiv \frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\tau_x^x<\tau_{y,I}^x}\right]}{\mathbb{P}\left[\tau_{I\cup y}^x < \tau_x^x\right]} \tag{5.63}$$

Note that we use $\bar{T}$ instead of $\bar{\tau}$ in the proof because this will simplify the following formulas. But in the present situation the two quantities are essentially equal, namely

$$\left|\frac{\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_I^x\right]}{\bar{T}} - 1\right| \leq e^{-O(N)} \tag{5.64}$$

This follows easily from equation (5.18) together with estimates on the expected transition times and probabilities (which shows that the first term in (5.18) is negligible in our situation), and the fact that by definition of an admissible transition, $P[\tau_y^x < \tau_{I\cup x}^x]/P[\tau_{I\cup y}^x < \tau_x^x]$ is exponentially close to one (use (5.26)!). Thus in the final results they can be interchanged without harm.

Then

$$\widetilde{G}^x_{y,I}(v)=\frac{\mathbb{P}\left[\tau^x_y<\tau^x_{I\cup x}|\tau^x_y<\tau^x_I\right]\widehat{G}^x_{y,I\cup x}\left(v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1-\mathbb{P}\left[\tau^x_x<\tau^x_{I\cup y}\right]\widehat{G}^x_{x,I\cup y}\left(v\frac{T_{I\cup y}(x)}{\bar{T}}\right)} \tag{5.65}$$

Using the analyticity properties established in the preceding lemma, we now proceed to a more careful computation, using second order Taylor expansions. This yields

$$\begin{aligned}&\widetilde{G}^x_{y,I}(v)\\ &=\frac{\mathbb{P}[\tau^x_y<\tau^x_{I\cup x}|\tau^x_y<\tau^x_I]\left(1+\frac{v}{\bar{T}}\mathbb{E}[\tau^x_y|\tau^x_y<\tau^x_{I\cup x}]+\frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)\right)}{\mathbb{P}[\tau^x_{y\cup I}<\tau^x_x]-\frac{v}{\bar{T}}\mathbb{E}[\tau^x_x\mathbb{1}_{\tau^x_x<\tau^x_{I\cup y}}]-\frac{(vT_{I\cup y}(x))^2}{2\bar{T}^2}\mathbb{P}[\tau^x_x<\tau^x_{I\cup y}]\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)}\end{aligned} \tag{5.66}$$

for some $0\leq\tilde{\theta}\leq 1$. We use (5.26) to get

$$\begin{aligned}\widetilde{G}^x_{y,I}(v)&=\frac{1+\frac{v}{\bar{T}}\mathbb{E}\left[\tau^x_y|\tau^x_y<\tau^x_{I\cup x}\right]+\frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1-\frac{v}{\bar{T}}\frac{\mathbb{E}[\tau^x_x\mathbb{1}_{\tau^x_x<\tau^x_{I\cup y}}]}{\mathbb{P}\left[\tau^x_{y\cup I}<\tau^x_x\right]}-\frac{(vT_{I\cup y}(x))^2\mathbb{P}[\tau^x_x<\tau^x_{I\cup y}]}{2\bar{T}^2\mathbb{P}[\tau^x_{I\cup y}<\tau^x_x]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)}\\ &=\frac{1+\frac{v}{\bar{T}}\mathbb{E}\left[\tau^x_y|\tau^x_y<\tau^x_{I\cup x}\right]+\frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1-v-\frac{(vT_{I\cup y}(x))^2}{2\bar{T}\mathbb{E}\left[\tau^x_x|\tau^x_x<\tau^x_{I\cup y}\right]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)}\end{aligned} \tag{5.67}$$

The term we must be most concerned with is the second order term in the denominator. Here we must use the full analyticity properties proven in Lemma 5.5. This gives, after computations analogous to those leading to (5.57) and using the obvious lower bound on $\bar{T}$

$$\begin{aligned}&\left|\frac{(vT_{I\cup y}(x))^2}{2\bar{T}\mathbb{E}\left[\tau^x_x|\tau^x_x<\tau^x_{I\cup y}\right]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)\right|\\ &\leq C^2N^{2k}\frac{|v|^2}{2\mathbb{E}\left[\tau^x_x|\tau^x_x<\tau^x_{I\cup y}\right]^2}\\ &\times\sum_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}}e^{-N[F_N(z^*(x',x))-F_N(x)+F_N(z^*(x,y))-F_N(x)]}\left(t_{I\cup y}(x',x)\right)^2\end{aligned} \tag{5.68}$$

Now according to our hypothesis that x is the lowest minimum in I, it follows that $t_{I\cup y}(x',x)\leq e^{N[F_N(z^*(x,x'))-F_N(x)]}$ so that (5.68) is finally bounded by

$$\begin{aligned}&C^2N^{2k}C'N^{-d}\frac{|v|^2}{2}|I|\sup_{x'\in\mathcal{M}_N\setminus\{I\cup y\cup x\}}e^{-N[F_N(z^*(x,y)-F_N(z^*(x',x))]}\\ &=C^2N^{2k}C'N^{-d}\frac{|v|^2}{2}|I|e^{-N\delta}\end{aligned} \tag{5.69}$$

where δ is strictly positive.

All the other terms in (5.67) except the leading ones are even smaller. Note moreover that both $T_{I\cup y}(x)$ and $T_{I\cup x}(y)$ are exponentially small compared to $\bar{T}$, so that all these error terms as functions of v are analytic if $|Re(v)| < 1$, and $|Im(v)| < \exp(NK_N/4)$. This allows to write $\widetilde{G}^x_{y,I}(v)$ in the form

$$\widetilde{G}^x_{y,I}(v) = \frac{1}{1-v} + e^{-NK_N/2}\frac{e_1(v)}{1-v} + \frac{e_2(v)}{(1-v)(1-v-e^{-NK_N/2}e_3(v))} \tag{5.70}$$

where all e_i are analytic and uniformly (in N) bounded in the domain $|Re(v)| < 1$, $|Im(v)| < e^{NK_N/4}$. This concludes the proof of the theorem.$\diamondsuit\diamondsuit$

5.3. The distribution of transition times.

From Theorem 5.4 one obtains of course some information on the distribution function.

Corollary 5.6: *Under the same assumptions as in Theorem 5.4, we have:*

(i) For any $\delta > 0$, for sufficiently large N,

$$\mathbb{P}\left[\tau^x_y > t\bar{T}|\mathcal{F}(x,z,y)\right] \leq e^{-(1-\delta)t}/\delta \tag{5.71}$$

(ii) Assume that N_i is a sequence of volumes tending to infinity such that for all i, $\mathcal{F}(x,z,y)$ is an admissible transition. Then, for any $t \geq 0$,

$$\lim_{i\uparrow\infty}\mathbb{P}_{N_i}\left[\tau^x_y > t\bar{\tau}_{N_i}|\mathcal{F}(x,z,y)\right] = e^{-t} \tag{5.72}$$

Proof: (i) is an immediate consequence of the the Laplace transform is bounded for real positive v with $v < 1$ and the exponential Chebyshev inequality. (ii) is a standard consequence of the fact that the Laplace transform converges pointwise for any purely imaginary v to that of the exponential distribution, and is analytic in a neighborhood of zero.$\diamondsuit$

With a little more work we can also complement the upper bound (5.71) by a lower bound on the distribution of the survival time in a valley.

Proposition 5.7: *Let $\mathcal{F}(x,z,y)$ be an admissible transition, and set $I = \mathcal{M}_N\backslash\mathcal{T}_{z,x}$. Let $h(N)$ be any sequence tending to zero as N tends to infinity. Then, for some $\kappa < \infty$, for any $0 < \alpha_t < 1$ we have that*

$$\mathbb{P}\left[\tau^x_I > t\right] \geq \begin{cases} e^{-t/(\bar{T}(1-\alpha_t))}\alpha_t^2\left(1-h(N)\right) & \text{if } t > h(N)\frac{CN^{d/2}}{N^\kappa + T_I(x)} \\ 1-o(1) & \text{if } t \leq h(N)\frac{CN^{d/2}}{N^\kappa+T_I(x)} \end{cases} \tag{5.73}$$

Proof: The proof of this lower bound consists essentially in guessing the strategy the process will follow in order to realize the event in question which will be to return a specific number of times to x without visiting the set I. For, obviously,

$$
\begin{aligned}
\mathbb{P}\left[\tau_I^x > t\right] &\geq \sum_{\substack{s_1,\dots,s_n\geq 1\\ s_1+\dots+s_n>t}} \mathbb{P}\left[\forall_{i=1}^n X_{s_i} = x, \forall_{s\leq s_1+\dots+s_n} X_s \notin I\right] \\
&= \sum_{\substack{s_1,\dots,s_n\geq 1\\ s_1+\dots+s_n>t}} \prod_{i=1}^n \mathbb{P}\left[\tau_x^x = s_i \leq \tau_I^x\right] \\
&= \left(\mathbb{P}\left[\tau_x^x \leq \tau_I^x\right]\right)^n \sum_{\substack{s_1,\dots,s_n\geq 1\\ s_1+\dots+s_n>t}} \prod_{i=1}^n \mathbb{P}\left[\tau_x^x = s_i | \tau_x^x \leq \tau_I^x\right]
\end{aligned}
\tag{5.74}
$$

We introduce the family of independent, identically distributed variables Y_i taking values in the positive integers such that for $\mathbb{P}\left[Y_i = s\right] = \mathbb{P}\left[\tau_x^x = s | \tau_x^x \leq \tau_I^x\right]$. Then (5.74) can be written as

$$
\mathbb{P}\left[\tau_I^x > t\right] \geq \left(\mathbb{P}\left[\tau_x^x \leq \tau_I^x\right]\right)^n \mathbb{P}\left[\sum_{i=1}^n Y_i > t\right]
\tag{5.75}
$$

We have good control on the first factor in (5.75). We need a lower bound on the second probability. The simplest way to proceed is to use the inequality, going back to Paley and Zygmund, that asserts that for any random variable X with finite expectation, and any $\alpha > 0$, $\mathbb{P}[X > (1-\alpha)\mathbb{E}X] \geq \alpha^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}X^2}$. We will use this with $X = \sum_i Y_i$. Thus

$$
\begin{aligned}
\mathbb{P}\left[\sum_{i=1}^n Y_i > t\right] &= \mathbb{P}\left[\sum_{i=1}^n Y_i > \frac{t}{n\mathbb{E}Y_1} n\mathbb{E}Y_1\right] \\
&\geq \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{n^2(\mathbb{E}Y_1)^2}{n(n-1)(\mathbb{E}Y_1)^2 + n\mathbb{E}Y_1^2} \\
&= \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{1}{1 + \frac{1}{n}\left(\frac{\mathbb{E}Y_1^2}{(\mathbb{E}Y_1)^2} - 1\right)}
\end{aligned}
\tag{5.76}
$$

Now using Lemma 5.5 one verifies easily that

$$
\mathbb{E}Y_1^2 \leq CN^\kappa + T_I(x)e^{-NK_N}
\tag{5.77}
$$

So that (5.76) gives

$$
\mathbb{P}\left[\sum_{i=1}^n Y_i > t\right] \geq \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{1}{1 + \frac{1}{n}\left(CN^\kappa + T_I(x)e^{-NK_N}\right)}
\tag{5.78}
$$

Thus the second factor is essentially equal to one if $n \gg \max(N^\kappa, T_I)$. We now choose n as the integer part of $n(t)$ where

$$
n(t) = \min\left(\frac{t}{\mathbb{E}Y_1(1-\alpha_t)}, \frac{1}{h(N)(CN^\kappa + T_I)}\right)
\tag{5.79}
$$

where $0 < \alpha_t < 1$ is such that $n(t) \in \mathbb{N}$. This yields

$$\begin{aligned} \mathbb{P}[\tau_I^x > t] &\geq (1 - \mathbb{P}[\tau_I^x \leq \tau_x^x])^{n(t)} \alpha_t^2 \frac{1}{1 + \frac{1}{n(t)}(CN^\kappa + T_I(x)e^{-NK_N})} \\ &\geq e^{-t\frac{\mathbb{P}[\tau_I^x \leq \tau_x^x]}{\mathbb{E}Y_1(1-\alpha_t)} - t\mathcal{O}\left(\mathbb{P}[\tau_I^x \leq \tau_x^x]^2\right)} \alpha_t^2 (1 - h(N)) \end{aligned} \tag{5.80}$$

if t is such that the $n(t)$ is given by the second term in the minimum in (5.79). This yields the first case in (5.78). If t is smaller than that, one sees easily that $n(t)\mathbb{P}[\tau_I^x \leq \tau_x^x]$ tends to zero uniformly in t, as $N \uparrow \infty$ (in fact exponentially fast!) This implies the second case. $\diamondsuit$

6. Miscellaneous consequences for the processes

In this section we collect some soft consequences of the preceding analysis. We begin by substantiating the claim made in Section 1 that the process spends most of the time in the immediate vicinity of the minima. We formulate this in the following form.

Proposition 6.1: *There exists finite positive constants C, k such that for any $\rho > 0$, $x \in \mathcal{M}_N$ and $t > 0$,*

$$\mathbb{P}\left[|X_t - x| > \rho \middle| \tau^x_{\mathcal{M}_N \backslash x} > t, X_0 = x\right] \leq CN^k \inf_{\rho' < \rho} \sup_{y \in \Gamma_N : \rho' \leq |x-y| \leq 3\rho'/2} e^{-N[F_N(y) - F_N(x)]} \tag{6.1}$$

Proof: We start to decompose the event $\{\tau^x_{\mathcal{M}_N \backslash x} > t\}$ as follows:

$$\{\tau^x_{\mathcal{M}_N \backslash x} > t\} = \bigcup_{0<s<t} \{\tau^x_{\mathcal{M}_N \backslash x} > s\} \cap \{X_s = x, \forall_{s<s'<t} X_{s'} \notin \mathcal{M}_N\} \tag{6.2}$$

Then

$$\begin{aligned} &\mathbb{P}\left[|X_t - x| > \rho \middle| \tau^x_{\mathcal{M}_N \backslash x} > t, X_0 = x\right] \\ &= \sum_{0<s<t} \frac{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > s, X_s = x\right]}{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > t\right]} \mathbb{P}\left[|X_{t-s} - x| > \rho, \tau^x_{\mathcal{M}_N} > t - s | X_0 = x\right] \\ &\leq \sum_{0<s<t} \frac{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > s, X_s = x\right]}{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > s, X_s = x\right] \mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > t - s\right]} \mathbb{P}\left[|X_{t-s} - x| > \rho, \tau^x_{\mathcal{M}_N} > t - s | X_0 = x\right] \\ &\inf_{\rho' < \rho} \leq \sum_{y : \rho' \leq |x-y| \leq 3\rho'/2} \sum_{0<s<t} \frac{\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{M}_N}, \tau^x_{\mathcal{M}_N} > t - s\right]}{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > t - s\right]} \\ &\leq \inf_{\rho' < \rho} \sum_{y : \rho' \leq |x-y| \leq 3\rho'/2} \sum_{0<s<t} \frac{\min\left(\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{M}_N}\right], \mathbb{P}\left[\tau^x_{\mathcal{M}_N} > t - s\right]\right)}{\mathbb{P}\left[\tau^x_{\mathcal{M}_N \backslash x} > t - s\right]} \end{aligned} \tag{6.3}$$

Now

$$\mathbb{P}\left[\tau_y^x < \tau_{\mathcal{M}_N}^x\right] \le e^{-N[F_N(y)-F_N(x)]} \tag{6.4}$$

while

$$\mathbb{P}\left[\tau_{\mathcal{M}_N}^x > t-s\right] = \sum_{y\in\mathcal{M}_N} \mathbb{P}\left[\tau_y^x > t-s|\tau_y^x < \tau_{\mathcal{M}_N\backslash y}^x\right] \mathbb{P}\left[\tau_y^x < \tau_{\mathcal{M}_N\backslash y}^x\right] \tag{6.5}$$

Note that by Theorem 1.10 and the exponential Markov inequality, for some $\kappa < \infty$

$$\mathbb{P}\left[\tau_y^x > t-s|\tau_y^x < \tau_{\mathcal{M}_N\backslash y}^x\right] \le CN^\kappa e^{-(t-s)cN^{-\kappa}} \tag{6.6}$$

while the factors $\mathbb{P}\left[\tau_y^x < \tau_{\mathcal{M}_N\backslash y}^x\right]$ are exponentially small except if $y = x$. The denominator is bounded below by Proposition 5.7. Since it decays with $t-s$ at an exponentially smaller rate than the numerator (see (6.6)), and is close to one for times up to the order $\exp(NC)$ (for some C), it is completely irrelevant.

Thus we see that the second term in the minimum takes over for $t-s > N^{\kappa+1}[F_N(y) - F_N(x)]$, a number small compared to the inverse of the first term in the minimum. Thus using that, for $a, b \ll 1$,

$$\sum_{t=0}^{\infty} \min\left(e^{-at}, b\right) \le \frac{b|\ln b|}{a} + \frac{b}{1-e^{-a}} \approx \frac{b}{a}(|\ln b| + 1) \tag{6.7}$$

the result follows immediately. $\diamondsuit$

Based on this result, we will now show that during an admissible transition the process also stays mostly close to its starting point, i.e. the lowest minimum of the valley concerned. The following proposition makes this precise.

Proposition 6.2: *Let $\mathcal{F}(x, z, x')$ be an admissible transition. Then there exists finite positive constants C, k and K_N s.t. $\lim N^{1-\alpha} K_N \uparrow \infty$, for some $\alpha > 0$, such that for any t and $\rho > 0$,*

$$\begin{aligned} &\mathbb{P}\left[|X_t - x| > \rho|\tau_{\mathcal{T}_{z,x}^c}^x > t, X_0 = x\right] \\ &\le CN^\kappa \inf_{\rho'<\rho} \sup_{y\in\Gamma_N:\rho'\le|x-y|\le 3\rho'/2} e^{-N[F_N(y)-F_N(x)]} + Ce^{-NK_N} \end{aligned} \tag{6.8}$$

Proof: The proof of this proposition is in principle similar to that of Proposition 6.1. We begin by applying the same decomposition as before to get

$$
\begin{aligned}
&\mathbb{P}\left[|X_t - x| > \rho \middle| \tau^x_{\mathcal{T}^c_{z,x}} > t, X_0 = x\right] \\
&= \sum_{0<s<t} \frac{\mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}} > s, X_s = x\right]}{\mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}} > t\right]} \mathbb{P}\left[|X_{t-s} - x| > \rho, \tau^x_{\mathcal{T}^c_{z,x}\cup x} > t - s \middle| X_0 = x\right] \\
&\inf_{\rho'<\rho} \leq \sum_{y:\rho'\leq|x-y|\leq 3\rho'/2} \sum_{0<s<t} \frac{\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x}, \tau^x_{\mathcal{T}^c_{z,x}\cup x} > t - s\right]}{\mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}} > t - s\right]} \\
&\leq \inf_{\rho'<\rho} \sum_{y:\rho'\leq|x-y|\leq 3\rho'/2} \sum_{0<s<t} \frac{\min\left(\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x}\right], \mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}\cup x} > t - s\right]\right)}{\mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}} > t - s\right]}
\end{aligned}
\tag{6.9}
$$

As in the proof of Proposition 6.1, the denominator is bounded by Proposition 5.7 and is seen to be insignificant. We concentrate on the estimates of the numerator. Again we have the obvious bound

$$
\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x}\right] \leq e^{-N[F_N(y)-F_N(x)]}
\tag{6.10}
$$

but to deal with the second probability in the minimum will be a little more complicated. Note first that as in (6.5) we can write

$$
\begin{aligned}
\mathbb{P}\left[\tau^x_{\mathcal{T}^c_{z,x}\cup x} > t - s\right] &= \mathbb{P}\left[\tau^x_x > t - s \middle| \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right] \mathbb{P}\left[\tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right] \\
&+ \sum_{y\in\mathcal{T}^c_{z,x}} \mathbb{P}\left[\tau^x_y > t - s \middle| \tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right] \mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right]
\end{aligned}
\tag{6.11}
$$

Now the terms in the second sum are all harmless, since by the estimates of Lemma 5.10 and the geometry of our setting

$$
\mathbb{P}\left[\tau^x_y > t - s \middle| \tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right] \mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right] \leq e^{-(t-s)/T_{\mathcal{T}^c_{z,x}\cup x}(y)} e^{-N[F_N(z)-F_N(x)]}
\tag{6.12}
$$

with $T_{\mathcal{T}^c_{z,x}\cup x}(y)$ much smaller than $e^{N[F_N(z)-F_N(x)]}$. The remaining term is potentially dangerous. To deal with this efficiently, we need to classify the trajectories according to the deepest minimum they have visited before returning to x. In the present situation the relevant effective depth of a minimum $y \in \mathcal{T}_{z,x}$ is (recall (5.2))

$$
d(y) \equiv d_{\mathcal{T}^c_{z,x}}(y, x) = F_N(z^*(y, x)) - F_N(y)
\tag{6.13}
$$

We will enumerate the minima in $\mathcal{T}_{z,x}$ according to increasing depth by $x = y_0, \ldots, y_k$ (we assume for simplicity that no degeneracies occur). We set $L(y) \equiv \{y' \in \mathcal{T}_{z,x} : d(y') \geq d(y)\}$. Then the family of disjoint events $\{\tau^x_x \geq \tau^x_{y_i}\} \cap \{\tau^x_x < \tau^x_{L(y_{i+1})}\}$ can serve as a partition of unity, i.e. we have that

$$\mathbb{P}\left[\tau_x^x > t-s|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right] = \sum_{i=0}^k \mathbb{P}\left[\tau_x^x > t-s, \tau_x^x \geq \tau_{y_i}^x, \tau_x^x < \tau_{L(y_{i+1})}^x|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right]$$
$$\leq \sum_{i=0}^k \min\left(\mathbb{P}\left[\tau_x^x \geq \tau_{y_i}^x|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right], \mathbb{P}\left[\tau_x^x > t-s, \tau_x^x < \tau_{L(y_{i+1})}^x|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right]\right) \tag{6.14}$$
$$\leq \sum_{i=0}^k \min\left(\mathbb{P}\left[\tau_x^x \geq \tau_{y_i}^x|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right], \mathbb{P}\left[\tau_x^x > t-s|\tau_x^x < \tau_{L(y_{i+1})}^x, \tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right]\right)$$

Now again

$$\mathbb{P}\left[\tau_x^x > t-s|\tau_x^x < \tau_{L(y_{i+1})}^x, \tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right] \leq e^{-(t-s)e^{-Nd(y_i)}} \tag{6.15}$$

while

$$\mathbb{P}\left[\tau_x^x \geq \tau_{y_i}^x|\tau_x^x < \tau_{\mathcal{T}_{z,x}^c}^x\right] \leq e^{-N[F_N(z^*(y_i,x))-F_N(x)]} \tag{6.16}$$

which is much smaller than $e^{-Nd(y_i)}$ (except in the case $i=0$ where we are back in the situation of Proposition 6.1). Combining all these estimates and using again (6.7) yields the claim of the proposition. $\diamondsuit$

Remark: Note that Proposition 6.2 again exhibits the special rôle played by admissible transitions. It justifies the idea that the behaviour of the process during an admissible transition can be described, on the time scale of the expected transition time[19], as waiting in the immediate vicinity of the starting minimum for an exponential time until jumping quasi-instantaneously to the destination point. This idea can also be expressed by passing to a measure valued description (as was done in [GOSV]) which will exhibit that the empirical measure of the process on any time scale small compared to the expected transition time but long compared to the next-smallest transition time within the admissible transition, is close to the Dirac mass at the minimum; since this, in turn, is asymptotically the invariant measure of the process conditioned to stay in the valley associated to the admissible transition, it can thus justly be seen as a metastable state associated with this time scale. The corresponding measure-valued process is than close to a jump process on the Dirac measures centered at these points. These results can be derived easily from the preceding Propositions, and we will not go into the details.

Let us also mention that from the preceding results and Theorem 5.9 one can easily extract statements concerning "exponential convergence to equilibrium". E.g., one has the following.

Corollary 6.3: *Let $N_k \uparrow \infty$ be a subsequence such that for all k the topological structure of the tree from Section 5 is the same and such that along the subsequence, F_{N_k} is generic. Let m_0 denote the lowest minimum of F_{N_k}. Let $f \in C(\Lambda, \mathbb{R})$ be any continuous function on the state space. Consider the process starting in some point $x \in \Gamma_N$. Then there is*

[19] A finer resolution will of course exhibit rare and rapid excursions to other minima during the time of the admissible transition, and we have all the tools to investigate these interior cycles.

a unique minimum $m(x)$ of F_{N_k}, converging to a minium $m(x)$ of F, such that, setting $\bar\tau^x(k) \equiv \mathbb{E}\left[\tau_{m_0}^{m_{N_k}(x)}\right]$

$$\lim_{k\uparrow\infty} \mathbb{E} f(X_{t/\bar\tau^x(k)}) = e^{-t} f(m(x)) + (1-e^{-t}) f(m_0) \tag{6.17}$$

(where on the right hand side $m(x), m_0$ denote the corresponding minima of the limiting function F). The point $m(x)$ is the lowest minimum of the deepest valley visited by the process in the canonical decomposition of the transition x, m_0 given in Theorem 4.5.

We leave the proof of the corollary to the reader. In a way such statements that involve convergence on a single time-scale are rather poor reflections of the complex structure of the behaviour of the process that is encoded in the description given in Section 4.

Relation to spectral theory.

Contrary to much of the work on the dynamics of spin systems we have not used the notion of "spectral gap" in this paper, and in fact the analysis of spectra has been limited in general to the rather auxiliary estimates in Section 2. Of course these approaches are closely related and our results could be re-interpreted in terms of spectral theory.

Most evidently, the estimate given in Theorem 5.1 can also be seen as precise estimates on the largest eigenvalue of the Dirichlet operator associated with the admissible transition $\mathcal{F}(x,z,y)$. Moreover, these Dirichlet eigenvalues are closely related to the low-lying spectrum of the stochastic matrix P_N. Sharp estimates on this relation require however some work, and we will not pursue this analysis in this paper but relegate it to forthcoming work in which the relation between the metastability problem and associated quantum mechanical tunneling problem will be further elucidated.

7. Mean field models and mean field dynamics

Our main motivation is to study the properties of stochastic dynamics for a class of models called "generalized random mean field models" that were introduced in [BG1]. We recall that such models require the following ingredients:

(i) A single spin space $\mathcal{S}$ that we will always take to be a subset of some linear space, equipped with some a priori probability measure q.

(ii) A state space $\mathcal{S}^N$ whose elements we denote by σ and call *spin configurations*, equipped with the product measure $\prod_i q(d\sigma_i)$.

(iii) The dual space $(\mathcal{S}^N)^{*M}$ of linear maps $\xi^T_{N,M} : \mathcal{S}^N \to \mathbb{R}^M$.

(iv) A mean field potential which is some real valued function $E_M : \mathbb{R}^M \to \mathbb{R}$.

(v) An abstract probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and measurable maps $\xi^T : \Omega \to (\mathcal{S}^{\mathbb{N}})^{*\mathbb{N}}$. Note that if Π_N is the canonical projection $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}^N$, then $\xi^T_{M,N}[\omega] \equiv \Pi_M \xi^T[\omega] \circ \Pi_N^{-1}$ are random elements of $(\mathcal{S}^N)^{*M}$.

(vi) The random order parameter

$$m_{N,M}[\omega](\sigma) \equiv \frac{1}{N}\xi^T_{M,N}[\omega]\sigma \in \mathbb{R}^M \tag{7.1}$$

(vii) A random Hamiltonian

$$H_{N,M}[\omega](\sigma) \equiv -NE_M\left(m_{N,M}[\omega](\sigma)\right) \tag{7.2}$$

In [BG1] the equilibrium properties of such models were studied in the case where $M = M(N)$ grows with N. Our aim in the long run is to be able to study dynamics in this situation, but in the present paper we restrict us to the case of fixed $M = d$. Also, we will only consider the case where $\mathcal{S}$ is a finite set.

Typical dynamics studied for such models are Glauber dynamics, i.e. (random) Markov chains $\sigma(t)$, defined on the configuration space $\mathcal{S}^N$ that are reversible with respect to the (random) Gibbs measures

$$\mu_{\beta,N}(\sigma)[\omega] \equiv \frac{e^{-\beta H_N[\omega](\sigma)}\prod_{i=1}^N q(\sigma_i)}{Z_{\beta,N}[\omega]} \tag{7.3}$$

and in which the transition rates are non-zero only if the final configuration can be obtained from the initial one by changing the value of one spin only. To simplify notation we will henceforth drop the reference to the random parameter ω.

As always the final goal will be to understand the macroscopic dynamics, i.e. the behaviour of $m_N(\sigma(t))$ as a function of time. It would be very convenient in this situation if $m_N(\sigma(t))$ were itself a Markov chain with state space $\mathbb{R}^d$. Such a Markov chain would be reversible with respect to the measure induced by the Gibbs measure on $\mathbb{R}^d$ through the map $\frac{1}{N}\xi^T$, and this measure has nice large deviation properties. Unfortunately, $m_N(\sigma(t))$ is almost never a Markov chain. A notable exception is the (non-random) Curie-Weiss model (see the next section). There are special situations in which it is possible to introduce a larger number of macroscopic order parameters in such a way that the corresponding induced process will be Markovian; in general this will not be possible. However, there is a canonical construction of a new Markov process on $\mathbb{R}^d$ that can be expected to be a good approximation to the induced process. This construction and the following results are all adapted from Ligget [Li], Section II.6.

Let $r_N(\sigma,\sigma')$ be transition rates of a Glauber dynamics reversible with respect to the measure $\mu_{\beta,N}$, i.e. for $\sigma \neq \sigma'$, $p_N(\sigma,\sigma') = \sqrt{\frac{\mu_N(\sigma')}{\mu_N(\sigma)}}g_N(\sigma,\sigma')$ where $g_N(\sigma,\sigma') = g_N(\sigma',\sigma)$. We denote by $\mathcal{R}_N$ the law of this Markov chain and by $\sigma(t)$ the coordinate variables. Define the induced measure

$$\mathbb{Q}_{\beta,N} = \mu_{\beta,N} \circ m^{-1}_{N,d} \tag{7.4}$$

and the new transition rates for a Markov chain with state space the $\Gamma_N = m_{N,d}(\mathcal{S}^N)$ (we drop the indices of $m_{N,d}$ in the sequel) by

$$p_N(x,y) \equiv \frac{1}{\mathbb{Q}_{\beta,N}(x)}\sum_{\sigma:m(\sigma)=x}\sum_{\sigma':m(\sigma')=y}\mu_{\beta_N}(\sigma)r_N(\sigma,\sigma') \tag{7.5}$$

Theorem 7.1: *Let $\mathbb{P}_N$ be the law of the Markov chain $x(t)$ with state space Γ_N and transitions rates $p_N(x,y)$ given by (7.5). Then $\mathcal{Q}_{\beta,N}$ is the unique reversible invariant measure for the chain $x(t)$. Moreover, for any $\sigma \in \mathcal{S}_N$ and $D \subset \mathcal{S}_N$, one has*

$$\mu_{\beta,N}(\sigma)\mathcal{R}_N\left[\tau_D^\sigma \leq \tau_\sigma^\sigma\right] \leq \mathbb{Q}_{\beta,N}(m(\sigma))\mathbb{P}_N\left[\tau_{m(D)}^{m(\sigma)} \leq \tau_{m(\sigma)}^{m(\sigma)}\right] \tag{7.6}$$

Finally, the image process $m(\sigma(t))$ is Markovian and has law $\mathbb{P}_N$ if for all σ, σ'' such that $m(\sigma) = m(\sigma'')$, $r(\sigma,\cdot) = r(\sigma'',\cdot)$. If the initial measure π_0 is such that for all σ, $\pi_0(\sigma) > 0$, then this condition is also necessary.

Remark: Notice that by the ergodic theorem, we can rewrite (7.6) in the less disturbing form

$$\frac{\mathbb{E}\tau_\sigma^\sigma}{\mathcal{R}_N\left[\tau_D^\sigma \leq \tau_\sigma^\sigma\right]} \geq \frac{\mathbb{E}\tau_{m(\sigma)}^{m(\sigma)}}{\mathbb{P}\left[\tau_{m(D)}^{m(\sigma)} \leq \tau_{m(\sigma)}^{m(\sigma)}\right]} \tag{7.7}$$

from which we see that the theorem really implies an ineqality for the arrival times in D.

Proof: Note that we can write

$$p_N(x,y) = \sqrt{\frac{\mathbb{Q}_{\beta,N}(y)}{\mathbb{Q}_{\beta,N}(x)}} \sum_{\sigma:m(\sigma)=x} \sum_{\sigma':m(\sigma')=y} \frac{\sqrt{\mu_{\beta_N}(\sigma)\mu_{\beta,N}(\sigma')}}{\sqrt{\mathbb{Q}_{\beta,N}(x)\mathbb{Q}_{\beta,N}(y)}} g_N(\sigma,\sigma') \tag{7.8}$$

which makes the reversibility of the new chain obvious. Note also that if $r_N(\sigma,\sigma')$ is constant on the sets $m^{-1}(x)$, then

$$p_N(x,y) = \sum_{\sigma':m(\sigma')=y} r_N(\sigma,\sigma') = \mathcal{R}_N\left[m(\sigma(t+1)) = y | \sigma(t) = \sigma\right] \tag{7.9}$$

which is only a function of $m(\sigma)$. From this one sees easily that in this case, the law of $m(\sigma(t))$ is $\mathbb{P}$. The proof of the converse statement is a little more involved and can be found in [BR].

Finally, the inequality (7.8) is proven in [Li] (Theorem 6.10). However, the proof given there is rather cumbersome, and meant to illustrate coupling techniques, while the result follows in a much simpler way from Theorem 6.1 in the same book. It may be worthwhile to outline the argument. Theorem 6.1 in [Li] states that

$$\mu_{\beta,N}(\sigma)\mathcal{R}_N\left[\tau_D^\sigma < \tau_\sigma^\sigma\right] = \frac{1}{2}\inf_{h\in\mathcal{H}_D^\sigma} \Phi_N(h) \tag{7.10}$$

where $\mathcal{H}_D^\sigma$ is the set of functions

$$\mathcal{H}_D^\sigma \equiv \left\{h : \mathcal{S}^N \to [0,1] : h(\sigma) = 0, \forall_{\sigma'\in D} h(\sigma') = 1\right\} \tag{7.11}$$

and Φ_N is the Dirichlet form associated to the chain $\mathcal{R}_N$,

$$\Phi_N(h) \equiv \sum_{\sigma,s'\in\Sigma^N} \mu_{\beta,N}(\sigma)\mathcal{R}_N(\sigma,\sigma')\left[h(\sigma) - h(\sigma')\right]^2 \tag{7.12}$$

Now we clearly majorize the infimum by restricting it to functions that are constant on the level sets of the map m, that is if we define the set

$$\widetilde{\mathcal{H}}^{x}_{\widetilde{D}} \equiv \left\{ \tilde{h} : \Gamma_N \to [0,1] : \tilde{h}(x) = 0, \forall_{y \in \widetilde{D}} \tilde{h}(y) = 1 \right\} \tag{7.13}$$

we have that

$$\inf_{h \in \mathcal{H}_\sigma} \Phi_N(h) \leq \inf_{\tilde{h} \in \mathcal{H}^{m(s)}_{m(D)}} \Phi_N(\tilde{h} \circ m) \tag{7.14}$$

But

$$\begin{aligned} \Phi_N(\tilde{h} \circ m) &= \sum_{x,x' \in \Gamma_N} \left[\tilde{h}(x) - \tilde{h}(x') \right]^2 \sum_{\sigma : m(\sigma) = (x), \sigma' : m(\sigma') = x'} \mu_{\beta,N}(\sigma) \mathcal{R}_N(\sigma, \sigma') \\ &= \sum_{x,x' \in \Gamma_N} \left[\tilde{h}(x) - \tilde{h}(x') \right]^2 \mathcal{Q}_{\beta,N}(x) p_N(x,x') \equiv \widetilde{\Phi}_N(\tilde{h}) \end{aligned} \tag{7.15}$$

where $\widetilde{\Phi}_N$ is the Dirichlet form of the chain $\mathbb{P}$. Using the analog of (7.10) for this new chain we arrive at the inequality (7.6).$\diamondsuit$

We certainly expect that in many situations the Markov chain $x(t)$ under the law $\mathbb{P}_N$ has essentially the same long-time behaviour than the non-Markovian image process $m(\sigma(t))$. However, we have no general results and there are clearly situations imaginable in which this would not be true. In the next section we will apply our general results to a specific model where this issue in particular can be studied nicely.

8. The random field Curie-Weiss model

The simplest example of disordered mean field models is the random field Curie-Weiss model. Here $\mathcal{S} = \{-1, 1\}$, q is the uniform distribution on this set. Its Hamiltonian is

$$H_N[\omega](\sigma) \equiv -N \frac{\left(M^1_N(\sigma)\right)^2}{2} - \sum_{i=1}^{N} \theta_i[\omega]\sigma_i \tag{8.1}$$

where

$$M_N(\sigma) \equiv \frac{1}{N} \sum_{i=1}^{N} \sigma_i \tag{8.2}$$

is called the *magnetization*. Here θ_i, $i \in \mathbb{N}$ are i.i.d. random variables. The dynamics of this model has been studied before: dai Pra and den Hollander studied the short-time dynamics using large deviation results and obtained the analog of the McKeane-Vlasov equations [dPdH]. Matthieu and Picco [MP1] considered convergence to equilibrium in a particularly simple case where the random field takes only the two values $\pm\epsilon$ (with further restrictions on the parameters that exclude the presence of more than two minima).

In this section we take up this simple model in the more general situation where the random field is allowed to take values in an arbitrary finite set. The main idea here is that in this case we are, as we will see, in the position to construct an image of the Glauber dynamic in a finite dimensional space that is Markovian, while it will be possible to compare this to the Markovian dynamics defined on the single parameter M_N in the manner described in the previous section.

We consider the Hamiltonian (8.1) where θ_i take values in the set

$$\mathcal{H} \equiv \{h_1, \dots, h_{K-1}, h_K\} \tag{8.3}$$

Each realization of the random field $\{\theta[\omega]\}_{i\in\mathbb{N}}$ induces a random partition of the set $\Lambda \equiv \{1, \dots, N\}$ into subsets

$$\Lambda_k[\omega] \equiv \{i \in \Lambda : \theta_i[\omega] = h_k\} \tag{8.4}$$

We may introduce k order parameters

$$m_k[\omega](\sigma) \equiv \frac{1}{N} \sum_{i\in\Lambda_k[\omega]} \sigma_i \tag{8.5}$$

We denote by $\underline{m}[\omega]$ the K-dimensional vector $(m_1[\omega], \dots, m_K[\omega])$. Note that these take values in the set

$$\Gamma_N[\omega] \equiv \times_{k=1}^{K} \left\{-\rho_{N,k}[\omega], -\rho_{N,k}[\omega] + \tfrac{2}{N}, \dots, \rho_{N,k}[\omega] - \tfrac{2}{N}, \rho_{N,k}[\omega]\right\} \tag{8.6}$$

where

$$\rho_{N,k}[\omega] \equiv \frac{|\Lambda_k[\omega]|}{N} \tag{8.7}$$

Note that the random variables $\rho_{N,k}$ concentrate exponentially (in N) around their mean values $\mathbb{E}_h \rho_{N,k} = \mathcal{P}[\theta_i = h_k] \equiv p_k$. Obviously $m^1[\omega](\sigma) = \sum_{k=1}^{K} m_k[\omega](\sigma)$ and $m^2[\omega](\sigma) = \sum_{\ell=1}^{k} h_k m_k(\sigma)$, so that the Hamiltonian can be written as a function of the variables $\underline{m}[\omega](\sigma)$, via

$$H_N[\omega](\sigma) = -NE(\underline{m}[\omega](\sigma)) \tag{8.8}$$

where $E : \mathbb{R}^K \to \mathbb{R}$ is the deterministic function

$$E(x) \equiv \frac{1}{2}\left(\sum_{k=1}^{K} x_k\right)^2 + \sum_{k=1}^{K} h_k x_k \tag{8.9}$$

The point is now that the image of the Glauber dynamics under the family of functions m_ℓ is again Markovian. This follows easily by verifying the criterion given in Theorem 7.1.

On the other hand, it is easy to compute the equilibrium distribution of the variables $\underline{m}[\omega]$. Obviously,

$$\mu_{\beta,N}[\omega](\underline{m}[\omega](\sigma) = x) \equiv \mathbb{Q}_{\beta,N}[\omega](x) = \frac{1}{Z_N[\omega]} e^{\beta N E(x)} \prod_{k=1}^{K} 2^{-N\rho_{N,k}[\omega]} \binom{N\rho_{N,k}[\omega]}{N(1+x_k)/2} \tag{8.10}$$

where $Z_N[\omega]$ is the normalizing partition function. Stirling's formula yields the well know asymptotic expansion for the binomial coefficients

$$2^{-N\rho_{N,k}[\omega]}\binom{N\rho_{N,k}[\omega]}{N(1+x_k)/2} = e^{-N\rho_{N,k}[\omega][I(x_k/\rho_{N,k}[\omega])+J_N(x_k,\rho_{N,k}[\omega])]} \tag{8.11}$$

where

$$I(x) \equiv \frac{1+x}{2}\ln(1+x) + \frac{1-x}{2}\ln(1-x) \tag{8.12}$$

is the usual Cramèr entropy and

$$J_N(x,\rho) = -\frac{1}{\rho N}\ln\left(\frac{1-(x/\rho)^2}{4} + \frac{2x/\rho}{1-(x/\rho)^2}\right) + O\left(\tfrac{1}{(\rho N)^2}\right) + \frac{1}{N^2}C(\rho N) \tag{8.13}$$

with $C(\rho N)$ a constant independent of x_k (and thus irrelevant) that satisfies $C(\rho N) = O(\ln(\rho N))$. Thus

$$F_N[\omega](x) \equiv -\frac{1}{\beta N}\ln \mathbb{Q}_{\beta,N}[\omega](x) = F_{0,N}[\omega](x) + F_{1,N}[\omega](x) + C_N \tag{8.14}$$

with

$$F_{0,N}(x) = -E(x) + \frac{1}{\beta}\sum_{k=1}^{K}\rho_{N,k}I(x_k/\rho_{N,k}) \tag{8.15}$$

$C_N = \beta^{-1}\sum_{k=1}^{K}\rho_{N,k}C(\rho_{N,k}N)$ is constant and of order $\ln N$, and $F_{1,N}$ of order $1/N$, uniformly on compact subsets of $\Gamma \equiv \times_{k=1}^{K}(-p_k,p_k)$. Moreover, $F_N(x)$ converges almost surely to the deterministic function

$$F_0(x) = -E(x) + \frac{1}{\beta}\sum_{k=1}^{K}p_k I(x_k/p_k) \tag{8.16}$$

uniformly on compact subsets of Γ. The dominant contribution to the finite volume corrections thus comes from the fluctuations part of the function $F_{0,N}$, $F_{0,N}(x) - F_0(x)$. One easily verifies that all conditions imposed on the functions F_N in Section 1 are verified in this example.

The landscape given by F. The deterministic picture.

To see how the landscape of the function F_N looks like, we begin by studying the deterministic limiting function F_0. Let us first look at the critical points. They are solutions of the equation $\nabla F_0(x) = 0$, which reads explicitly

$$0 = \frac{\partial}{\partial x_k}F_0(x) = -\sum_{\ell=0}^{K}x_\ell - h_k x_k + \frac{1}{\beta}p_k I'(x_k/p_k),\ \ k = 1,\ldots,K \tag{8.17}$$

or equivalently

$$\begin{aligned} x_k &= p_k \tanh(\beta(m+h_k)), \quad k=1,\dots,K \\ m &= \sum_{k=1}^K x_k \end{aligned} \tag{8.18}$$

These equations have a particularly pleasant structure. Their solutions are generated by solutions of the transcendental equation

$$m = \sum_{k=1}^K p_k \tanh(\beta(m+h_k)) = \mathbb{E}_h \tanh\beta(m+h) \tag{8.19}$$

Thus if $m^{(1)},\dots,m^{(r)}$ are the solutions of (8.19), then the full set of solutions of the equations (8.17) is given by the vectors $x^{(1)},\dots,x^{(r)}$ defined by

$$x_k^{(\ell)} \equiv p_k \tanh\beta(m^{(\ell)}+h_k) \tag{8.20}$$

Next we analyze the structure of the critical points. Using that $I''(x) = \frac{1}{1-x^2}$, we see that

$$\frac{\partial^2}{\partial x_k \partial x_{k'}} F_0(x) = -1 + \frac{\delta_{k,k'}}{\beta p_k \left(1 - x_k^2/p_k^2\right)} \tag{8.21}$$

Thus at a critical point $x^{(\ell)}$,

$$\frac{\partial^2}{\partial x_k \partial x_{k'}} F_0\left(x^{(\ell)}\right) = -1 + \delta_{k,k'}\lambda_k(m^{(\ell)}) \tag{8.22}$$

where

$$\lambda_k(m) \equiv \frac{1}{\beta p_k (1-\tanh^2(\beta(m+h_k)))} \tag{8.23}$$

Lemma 8.1: *The Hessian of F_0 at $(x^{(\ell)})$ has at most one negative eigenvalue. A negative eigenvalue exists if and only if*

$$\beta\mathbb{E}_h\left(1-\tanh^2(\beta(x^{(\ell)}+h))\right) > 1 \tag{8.24}$$

Proof: Consider any matrix of the form $A_{kk'} = -1 + \delta_{k,k'}\lambda_k$ with $\lambda_k \geq 0$. To see this, let $\{\zeta_1,\dots,\zeta_L\}$ denote the set of distinct values that are taken by $\lambda_1,\dots,\lambda_K$. Put $k_\ell = \{k : \lambda_k = \zeta_\ell\}$ and denote by $|\kappa_\ell|$ the cardinalities of these sets. Now the eigenvalue equations read

$$-\left(\sum_{k=1}^K u_k\right) + (\lambda_k - \gamma)u_k = 0 \tag{8.25}$$

Let ζ_ℓ be such that $|\kappa_\ell| > 1$, if such a ζ_ℓ exists. Then we will construct $|\kappa_\ell|-1$ orthogonal solutions to (8.25) with eigenvalue $\gamma = \zeta_\ell$. Namely, we set $u_k = 0$ for all $k \notin \kappa_\ell$. The remaining components must satisfy $\sum_{k\in\kappa_\ell} u_k = 0$. But obviously, this equation has $|\kappa_\ell| - 1$ orthonormal solutions. Doing this for every ζ_ℓ, we construct altogether $K - L$ eigenvectors corresponding to the eigenvalues ζ_ℓ. Note that for all these solutions, $\sum_k u_k = 0$. We are left with finding the remaining L eigenfunctions. Now take $\gamma \notin \{\zeta_1, \dots, \zeta_L\}$. Then (8.25) can be rewritten as

$$u_k = \frac{\sum_{k=1}^K u_k}{\lambda_k - \gamma} \tag{8.26}$$

Summing equation (8.26) over k, we get

$$\sum_{k=1}^K u_k = \sum_{k=1}^K u_k \sum_{k=1}^K \frac{1}{\lambda_k - \gamma} \tag{8.27}$$

Since we have already exhausted the solutions with $\sum_{k=1}^K u_k = 0$, we get for the remaining ones the condition

$$1 = \sum_{k=1}^K \frac{1}{\lambda_k - \gamma} = \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\lambda_k - \gamma} \tag{8.28}$$

Inspecting the right-hand side of (8.28) one sees immediately that this equation has precisely L solutions γ_i that satisfy

$$\gamma_1 < \zeta_1 < \gamma_2 < \zeta_2 < \gamma_3 < \cdots < \gamma_L < \zeta_L \tag{8.29}$$

of which at most γ_1 can be negative. Moreover, a negative solution γ implies that

$$1 = \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\lambda_k - \gamma} < \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\lambda_k} = \sum_{k=1}^K \frac{1}{\lambda_k} \tag{8.30}$$

which upon inserting the specific form of λ_κ yields (8.24). On the other hand, if $\sum_{k=1}^K \frac{1}{\lambda_k} > 1$, then by monotonicity there exists a negative solution to (8.28). This proves the lemma.$\diamondsuit$

The following general features are now easily verified due to the fact that the analysis of the critical points is reduced to equations of one single variable. The following facts hold:

(i) For any distribution of the field, there exists β_c such that: If $\beta < \beta_c$, there exists a single critical point and F_0 is strictly convex. If $\beta > \beta_c$, there exist at least 3 critical points, the first and the last of which (according to the value of m) are local minima, and each minimum is followed by a saddle with one negative eigenvalue, and vice versa, with possibly intermediate saddles with one zero eigenvalue interspersed.

(ii) Assume $\beta > \beta_c$. Then each pair of consecutive critical points of F_0 can be joined by a unique integral curve of the the vector field $\nabla F_0(x)$.

The exact picture of the landscape depends of course on the particular distribution of the magnetic field chosen. In particular, the exact number of critical points, and in particular of minima, depends on the distribution (and on the temperature). The reader is invited to use e.g. mathematica and produce diverse pictures for her favorite choices. We see that a major effect of the disorder enters into the form of the deterministic function $F_0(x)$. Only a secondary rôle is played by the remnant disorder whose effect will be most notable in symmetric situations where it can break symmetries present on the level of F_0.

Fluctuations

In the present simple situation it turns out that the fluctuations of the function $F_{0,N}$ can also be controlled in a precise way. We will show the following result.

Proposition 8.3: *Let $g_k, k = 1, \dots, K$ be a family of independent Gaussian random variables with mean zero and variance $p_k(1-p_k)$. Then the function $\sqrt{N}[F_N(x)-F_0(x)]$ converges in distribution, uniformly on compact subsets of Γ to the random function*

$$\frac{1}{\beta}\sum_{k=1}^{K} g_k \left(\frac{x_k}{p_k^2} I'(x_k/p_k) - I(x_k/p_k)\right) \tag{8.31}$$

Proof: Since $F_N - F_{0,N}$ converges to zero uniformly, it is enough to consider

$$\begin{aligned} F_{0,N}(x) - F_0(x) &= \frac{1}{\beta}\sum_k (\rho_{N,k} I(x_k/\rho_{N,k}) - p_k I(x_k/p_k)) \\ &= \frac{1}{\beta}\sum_k \Big((\rho_{N,k} - p_k) I(x_k/p_k) + p_k (I(x_k/\rho_{N,k}) - I(x_k/p_k)) \\ &\quad + (\rho_{N,k} - p_k)(I(x_k/\rho_{N,k}) - I(x_k/p_k))\Big) \end{aligned} \tag{8.32}$$

Now in the interior of Γ we may develop

$$I(x_k/\rho_{N,k}) - I(x_k/p_k) = (\rho_{N,k} - p_k)\frac{1}{p_k^2} I'(x_k/p_k) + O((\rho_{N,k} - p_k)^2) \tag{8.33}$$

Now the $\rho_{N,k}$ are actually sums of independent Bernoulli random variables with mean p_k, namely $\rho_{N,k} = \frac{1}{N}\sum_{i=1}^{N} \delta_{h_k,\theta_i}$. Thus, by the exponential Chebyshev inequality,

$$\mathcal{P}\left[|\rho_{N,k} - p_k| > \epsilon\right] \le 2\exp\left(-N I_{p_k}(\epsilon)\right) \tag{8.34}$$

where $I_p(\epsilon) \ge 0$ is a strictly convex function that takes its minimum value 0 at $\epsilon = 0$. Thus with probability tending to one rapidly, we have that e.g. $(\rho_{N,k} - p_k)^2 \le N^{-3/4}$ which allows us to neglect all second order remainders. Finally, by the central limit theorem the family of random variables $\sqrt{N}(\rho_{N,k} - p_k)$ converges to a family of independent Gaussian random variables with variances $p_k(1 - p_k)$. This yields the proposition.$\diamondsuit$

Relation to a one-dimensional problem.

We note that the structure of the landscape in this case is quasi one-dimensional. This is no coincidence. In fact, it is governed by the rate function of the total magnetization, $-\frac{1}{\beta N}\mu_{\beta,N}(m^1(\sigma) = m)$ which to leading orders is computed, using standard techniques, as

$$G_{0,N}(m) = -\frac{m^2}{2} + \sup_{t\in\mathbb{R}}\left(mt - \frac{1}{\beta}\sum_{k=1}^{K} \rho_{N,k} \ln\cosh\beta(h_k + t)\right) \tag{8.35}$$

The most important facts for us are collected in the following Lemma.

Lemma 8.2: *The functions $G_{0,N}$ and $F_{0,N}$ are related in the following ways.*

(i) For any $m \in [-1,1]$,

$$G_{0,N}(m) = \inf_{x \in \mathbb{R}^K : \sum_k x_k = m} F_{0,N}(x) \tag{8.36}$$

(ii) If x^* is a critical point of $F_{0,N}$, then $m^* \equiv \sum_k x_k^*$ is a critical point of $G_{0,N}$.

(iii) If m^* is a critical point of $G_{0,N}$, then $x^*(m^*)$, with components $x_k^*(m) \equiv \rho_{N,k} \tanh \beta(m^* + h_k)$ is a critical point of $F_{0,N}$.

(iv) At any critical point m^*, $G_{0,N}(m^*) = F_{0,N}(x^*(m^*))$.

The prove of this lemma is based on elementary analysis and will be left to the reader.

The point we want to make here is that while the dynamics induced by the Glauber dynamics on the total magnetization is not Markovian, if we define a Markov chain $m(t)$ that is reversible with respect to the distribution of the magnetization in the spirit of Section 8 and compare its behaviour to that of the Markov chain $\underline{m}(t) = \underline{m}(\sigma(t))$, the preceding result assures that their long-time dynamics are identical since all that matters are the precise values of the respective free-energies at its critical points, and these coincide according to the preceding lemma (up to terms of order $1/N$, and the asymptotics given in (8.12), (8.13), up to (K-dependent) constants). In other words, the two dynamics, when observed on the set of minima of their respective free energies, are identical on the level of our precision.

References

[Az] R. Azencott, "Petites perturbations aléatoires des systèmes dynamiques: développements asymptotiques", Bull. Sc. Math. **109**, 253-308 (1985).

[BG1] A. Bovier and V. Gayrard, "Hopfield models as generalized random mean field models", in "Mathematical aspects of spin glasses and neural networks", A. Bovier and P. Picco (eds.), Progress in Probability **41**, (Birkhäuser, Boston 1998).

[BG2] A. Bovier and V. Gayrard, "Sample path large deviations for a class of Markov chains with slowly varying transition probabilities", WIAS-preprint (1998).

[BR] C.J. Burke and M. Rosenblatt, "A Markovian function of a Markov Chain", Ann. Math. Statist. **29**, 1112-1122, (1958).

[CC] O. Catoni and R. Cerf, "The exit path of a Markov chain with rare transitions", ESAIM Probab. Statist. **1**, 95-144 (1995/97).

[CGOV] M. Cassandro, A. Galves, E. Olivieri, and M.E. Varés, "Metastable behaviour of stochastic dynamics: a pathwise approach, J. Stat. Phys. **35**, 603-634 (1988).

[FJ] W.H. Fleming and M.R. James, "Asymptotic series and exit time probabilities", Ann.

Probab. **20**, 1369-1384 (1992).

[FMP] R.L. Fontes, P. Mathieu and P. Picco, "On the averaged dynamics of the random field Curie-Weiss model", private communication.

[FW] M.I. Freidlin and A.D. Wentzell, "Random perturbations of dynamical systems", Springer, Berlin-Heidelberg-New York, 1984.

[GT] I. Gaudron and A. Trouvé, "Fluctuations of empirical means at low temperature for finite Markov chains with rare transitions in the general case". Probab. Theor. Rel. Fields **111**, 215-251 (1998).

[Ho] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA **79**, 2554-2558 (1982).

[vK] N.G. van Kampen, "Stochastic processes in physics and chemistry", North-Holland, Amsterdam, 1981 (reprinted in 1990).

[KMST] C. Knessl, B.J. Matkowsky, Z. Schuss, and C. Tier, "An asymptotic theory of large deviations for Markov jump processes", SIAM J. Appl. Math. **46**, 1006-1028 (1985).

[Ki1] Y. Kifer, "On the principle eignevalue of a singular perturbation problem with hyperbolic limit points and circles", J. Diff. Eq. **37**, 108-139 (1980).

[Ki1] Y. Kifer, "The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point", Israel J. Math. **40**, 74-94 (1981).

[Ki3] Y. Kifer, "Random perturbations of dynamical systems", Progress in Probability and Statistics 16, Birkhäuser, Boston-Basel, 1988.

[Ki4] Y. Kifer, "A discrete time version of the Wentzell-Freidlin theory", Ann. Probab. **18**, 1676-1692 (1990).

[Ku] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes", J. Appl. Probab. **7**, 49–58 (1970).

[Li] T.M. Liggett, "Interacting particle systems", Springer, Berlin, 1985.

[MP] P. Mathieu and P. Picco, "Metastability and convergence to equilibrium for the random field Curie-Weiss model", J. Stat. Phys. **91**, 679-732 (1998).

[OS1] E. Olivieri, E. Scoppola, "Markov chains with exponentially small transition probabilities: First exit problem from a general domain -I. The reversible case" J. Stat. Phys. **79**, 613 (1995).

[OS2] E. Olivieri, E. Scoppola, " Markov chains with exponentially small transition probabilities: First exit problem from a general domain -II. The general case". J. Stat. Phys. **84**, 987-1041 (1996).

[SS] R. Schonmann and S. Shlosman, "Wulff droplets and the metastable relaxation of kinetic Ising models", Commun. Math. Phys. **194** 389-462 (1998).

[Va] M.E. Vares, "Large deviations and metastability", in *Disordered systems*, 1-62, Travaux en Cours Vol.**53**, Hermann, Paris, 1996.

[W1] A.D. Wentzell, "Rough limit theorems on large deviations for Markov stochastic processes I.", Theor. Probab. Appl. **21**, 227-242 (1976).

[W2] A.D. Wentzell, "Rough limit theorems on large deviations for Markov stochastic processes II.", Theor. Probab. Appl. **21**, 499-512 (1976).

[W3] A.D. Wentzell, "Rough limit theorems on large deviations for Markov stochastic processes III.", Theor. Probab. Appl. **24**, 675-692 (1979).

[W4] A.D. Wentzell, "Rough limit theorems on large deviations for Markov stochastic processes IV.", Theor. Probab. Appl. **27**, 215-234 (1982).