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 constant factor. The distributions of the 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@math.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@math.uni-potsdam.de

Table of contents

1. Introduction 3
1.1. General introduction 3
1.2. General set-up 4
1.3. Outline of the general strategy 10
2. Precise estimates on transition probabilities 12
3. Laplace transforms of the transition times in the elementary situation 18
4. Valleys, trees, and graphs 23
4.1. Valley-structure and its tree-representation 23
4.2. Construction of the transition process 24
5. Transition times of admissible transitions 28
5.1. Expected times of admissible transitions 29
5.2. Laplace transforms of transition times of admissible transitions 37
5.3. The distribution of transition times 41
6. Miscellaneous consequences for the process 43
7. An application to disordered mean field models 47
7.1 Generalities 47
7.2. The random field Curie Weiss model 48
7.3. The landscape given by F. The deterministic picture 50
7.4. Fluctuations 52
7.5. Relation to a one-dimensional problem 52
References 53

1. Introduction

1.1 General introduction

This paper is devoted to developing a systematic approach to obtain *sharp* estimates on the long time behaviour of certain classes of Markov chains. Our main motivation for this investigation arose from the study of the dynamics of mean field spin systems, where by dynamics we understand of course stochastic dynamics of Glauber type. More specifically, we are interested in the dynamics 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 so-called *metastable* phenomena. While our methods and results are rather more general, this motivation has to a considerable extent influenced specific assumptions we are making, and we would like to ask the reader to keep this fact in mind when following us through the paper. For the purposes of this paper, we will always consider reversible Markov chains. From the point of view of many applications, this is rather natural, but we must of course admit that this is at present a serious limitation that we do, however, hope to overcome in subsequent work.

On a heuristic level, metastable phenomena in mean field models and in stochastic dynamics in general 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, important progress has been made in a paper by Mathieu and Picco [MP] that treated a particular case of the random field Curie-Weiss model, leading to a two-dimensional double well problem. 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 [Sc,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 rather close to the situation we are considering, and that 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

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

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

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 $b(x, \epsilon) = \nabla F_\epsilon(x)$, $F_\epsilon(x)$ 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 second edition of this book. One of the important aspects of this work is that it 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. An important point to keep in mind is that the work of Wentzell and Freidlin in as far as it gives a complete description of the long-time behaviour of the process gives only results for the exponential asymptotics of the quantities considered; that is, the quantities that can be computed, like transition probabilities, transition times, etc., are controlled only up to *multiplicative errors* of the order $\exp\left(o(1)\epsilon^{-1}\right)$. *For many practical applications, and in particular also for the applications in disordered systems we have in mind, such a precision is insufficient since one is interested in effects on a refined scale that would be masked completely by the errors.*

Not surprisingly, great efforts have been made to obtain sharper estimates than those of [WF] subsequently, using both probabilistic [Ki1-4,Az] and analytic [FJ] techniques, to name only a few references. However, when studying these results, one finds that all of them concern purely local phenomena, i.e. they concentrate on the computation of *exit times* and *exit probabilities* from neighborhoods of critical points that are carefully stipulated *not to contain any other critical point* of the drift field. In particular, none of these results concern the situation when the process passes from one stable fixpoint to another through the neighborhood of an unstable critical point (saddle point). It has to be stressed that (contrary to popular believe) this is not just a minor detail, but a) the correct computation of transition times (and *spectra!!*) requires to take the effects of saddle point into account and b) doing this leads to considerable technical difficulties that cannot be solved by simple extensions of the same methods. The reason for these difficulties are most easily understood from the analytic point of view. The computation of all the relevant quantities leads to the solution of certain partial differential (or, in the discrete case, difference) equations. The classical methods used to obtain the results in terms of asymptotic expansions in the parameter ϵ are so-called semi-classical or WKB methods. These methods work in their standard form precisely under the conditions mentioned above. The problem of a transition through a saddle point corresponds to a much more difficult situation where straightforward WKB expansions fail. In the analogous quantum-mechanical problem, such a situation corresponds to what is called "tunneling through a non-resonant well" which has been treated in a reasonably general setting only in the third paper of a series of papers on tunneling problems by Helffer and Sjösstrand [HS] using most sophisticated and complicated tools from micro-local analysis. To our knowledge, these methods have never been applied to the corresponding probabilistic problems[6]. One should also mention that the methods of [HS] apply only in the self-adjoint (= reversible) case and require a considerable amount of regularity assumptions, as well as additional geometric hypotheses for the potentials. Their extension to the discrete setting we have in mind, while probably possible, would and will require substantial additional efforts.

It is for this reason that in the present paper we develop an altogether new approach to the problem that will give results that are almost, but not quite, as sharp as could be hoped for from an extension of microlocal techniques, but that have the advantage of being much easier to obtain and that will have a much wider range of applicability.

1.2. The general set-up.

[6] Even in the much simpler one-dimensional case, full asymptotic expansions have been obtained only on the heuristic level [KMST]. They are rigorously justified in [EK] in a specific case.

We will now describe the general class of Markov chains we will consider. The conditions we will impose are motivated by the application to the dynamics of disordered mean field models we have in mind. This relation will be explained in Section 7 and a specific example will be discussed. 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 should also mention that the class of chains we introduce here can be naturally interpreted without this context as a discretized version of diffusion processes in a potential, as have been considered e.g. in [vK,Ku,KMST].

We consider canonical Markov chains on a state space Γ_N where Γ_N is the intersection of some lattice[7] (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[8]:

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 intΛ.

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, denoting by k any generator of the lattice, we have

$$N|F_{N,0}(x) - F_{N,0}(x+k/N) - (F_{N,0}(y) - F_{N,0}(y+k/N)| \leq C\|x-y\| \tag{1.3}$$

, 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.[9]

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). Formally, we define the

[7] 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.

[8] The assumptions made here are sufficient, but far from necessary, to allow the methods we explain here to work. We prefer in this paper to work in a concrete class of examples rather than to formulate abstract conditions. In a forthcoming publications [BEGK] we will deal with related questions in a more abstract setting.

[9] It is possible to generalize the notion of the set of local minima considerably. In particular, in the presence of very many local minima, one may under suitable conditions reduce the set $\mathcal{M}_N$ to a subset representing particularly profound "valleys". We will elaborate on this point in [BEGK].

essential saddle, $z^*(x,y)$, between two minima $x, y \in \mathcal{M}_N$ as

$$z^*(y,x) \equiv \arg \inf_{\gamma:\gamma(0)=y,\gamma(1)=x} \left(\sup_{t\in[0,1]} \left[F_N(\gamma(t)) \right] \right) \tag{1.4}$$

where the infimum is over all paths $\gamma : [0,1] \to \Gamma_N$ going from y to x[10] with jumps along the edges of Γ_N only[11].

A point z is called an essential saddle point if there exist minima $x, y \in \mathcal{M}_N$ such that $z^*(x,y) = z$. The set of all essential saddle point will be denoted by $\mathcal{E}_N$. Our assumptions on the N dependence of the set $\mathcal{M}_N$ apply in the same way to $\mathcal{E}_N$.

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. The basic methods can also be applied in different situations, with ensuing modifications of the results.

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

$$p_N(x,y) \equiv \sqrt{\frac{\mathbb{Q}_N(y)}{\mathbb{Q}_N(x)}} f_N(x,y) \tag{1.5}$$

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**) $p_N(x,y) \geq c$.

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

$$P_N(x,y) \equiv \begin{cases} p_N(x,y), & \text{if} \quad (x,y) \in E(\Gamma_N) \\ 1 - \sum_{z\in\Gamma_N:(x,z)\in E(\Gamma_N)} p_N(x,z), & \text{if} \quad x = y \\ 0, & \text{else} \end{cases} \tag{1.6}$$

[10] 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.

[11] We will extend the definition (1.4) also to general points $x, y \in \Gamma_N$. In that case it may happen that $z^*(x,y)$ is not a saddle point, but one of the endpoints x or y itself.

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

$$A_N(x,y)\equiv\begin{cases} p_N(x,y), & \text{if} \quad (x,y)\in E(\Gamma_N)\\ -\sum_{z\in\Gamma_N:(x,z)\in E(\Gamma_N)} p_N(x,y), & \text{if} \quad x=y\\ 0, & \text{else}\end{cases} \tag{1.7}$$

It is clear that the continuous time process can be studied by passing to the embedded discrete time process where time becomes the number of jumps. We will therefore only consider discrete time chains in this paper.

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 intervals 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 positive probability. 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 e^{Na_N} within a valley $V(x)$ without visiting x 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. We do, however, expect that our methods can be applied to continuous processes by using suitable discretizations, but this is left to further research.

The informal discussion above will be made precise in the sequel. 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. For $y\in\Gamma_N$ and $I\subset\Gamma_N$, we will write τ_I^y for the first time the process conditioned to starting at y hits the set I, i.e. we write

$$\mathbb{P}\left[\tau_I^y=t\right]\equiv\mathbb{P}\left[X_t\in Ix,\forall_{0<s<t}X_s\notin I|X_0=y\right] \tag{1.8}$$

for $t>0$. In the case $I=\{y\}$, 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.

In the remainder of this subsection we state some elementary facts from the theory of Markov chains that will be used repeatedly in the paper and that we therefore choose to restate for the benefit

of the non-expert. No proofs will be given. Readers familiar with Markov processes should only glance through this part to get acquainted with our notation.

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*

$$\begin{aligned}\mathbb{P}\left[\tau_x^y = t\right] &= \mathbb{P}\left[\tau_x^y = t, \tau_x^y < \tau_z^y\right] \\ &\quad + \sum_{0<s<t} \mathbb{P}\left[\tau_z^y = s, \tau_z^y < \tau_x^y\right] \mathbb{P}\left[\tau_x^z = t - s\right]\end{aligned} \tag{1.9}$$

A simple consequence is the following basic renewal equation.

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

$$\mathbb{P}\left[\tau_x^y = t\right] = \sum_{n=0}^{\infty} \sum_{\substack{t_1,\dots t_{n+1} \\ \sum_i t_i = t}} \prod_{i=1}^{n} \mathbb{P}\left[\tau_y^y = t_i, \tau_y^y < \tau_x^y\right] \mathbb{P}\left[\tau_x^y = t_{n+1}, \tau_x^y < \tau_y^y\right] \tag{1.10}$$

The fundamental importance in the decomposition of Lemma 1.2 lies in the fact that objects like the last factor in (1.10) 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 $\mathcal{F}\left(X_s, 0 \le s \le \tau_x^y\right)$ 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 $\mathcal{F}\left(X_s, 0 \le s \le \tau_x^y\right)$. Let $\mathcal{A}^r$ denote the time reversion of the event $\mathcal{A}$. Then*

$$\mathbb{Q}_N(y)\mathbb{P}\left[\mathcal{A}, \tau_x^y < \tau_y^y\right] = \mathbb{Q}_N(x)\mathbb{P}\left[\mathcal{A}^r, \tau_y^x < \tau_x^x\right] \tag{1.11}$$

For example, we have

$$\mathbb{Q}_N(y)\mathbb{P}\left[\tau_x^y = t, \tau_x^y < \tau_y^y\right] = \mathbb{Q}_N(x)\mathbb{P}\left[\tau_y^x = t, \tau_y^x < \tau_x^x\right] \tag{1.12}$$

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 (1.10) 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

$$G_{x,I}^y(u) \equiv \sum_{t \ge 0} e^{ut}\mathbb{P}\left[\tau_x^y = t, \tau_x^y \le \tau_I^y\right] \equiv \mathbb{E}\left[e^{u\tau_x^y} \mathbb{I}_{\{\tau_x^y \le \tau_I^y\}}\right], \quad u \in \mathbb{C} \tag{1.13}$$

(We want to include the possibility that I contains x and/or y for later convenience). As we will see it is important to understand what the domains of these functions are. Since the Laplace transforms

defined in (1.13) are Laplace transforms of the distributions of positive random variables, all these functions exist and are analytic at least for all $u \in \mathbb{C}$ with $Re(u) \leq 0$. Moreover, if $G^y_{x,I}(u_0)$ is finite for some $u_0 \in \mathbb{R}_+$, then it is analytic in the half-space $Re(u) \leq u_0$.[12] As we will see later, each of the functions introduced in (1.13) will actually exist and be finite for some $u_0 > 0$ (depending on N).

Note that in particular

$$G^y_{x,I}(0) = \mathbb{P}\left[\tau^y_x \leq \tau^y_I\right] \tag{1.14}$$

and

$$\frac{d}{du} G^y_{x,I}(u=0) \equiv \dot{G}^y_{x,I}(0) = \mathbb{E}\left[\tau^y_x \mathbb{I}_{\{\tau^y_x \leq \tau^y_I\}}\right] \tag{1.15}$$

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

$$\frac{\dot{G}^y_{x,I}(0)}{G^y_{x,I}(0)} = \mathbb{E}\left[\tau^y_x | \tau^y_x \leq \tau^y_I\right] \tag{1.16}$$

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*

$$\mathbb{Q}_N(y) G^y_{x,I}(u) = \mathbb{Q}_N(x) G^x_{y,I}(u) \tag{1.17}$$

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

$$\frac{G^y_{x,I}(u)}{G^y_{x,I}(0)} = \frac{G^x_{y,I}(u)}{G^x_{y,I}(0)} \tag{1.18}$$

and in particular

$$\mathbb{E}\left[\tau^y_x | \tau^y_x \leq \tau^y_I\right] = \mathbb{E}\left[\tau^x_y | \tau^x_y \leq \tau^x_I\right] \tag{1.19}$$

(1.18) 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.

A crucial rôle will be played by the following *renewal equation*:

Lemma 1.5: *Let $I \subset \Gamma_N$. Then for all $y \notin I \cup x$,*

$$G^y_{x,I}(u) = \frac{G^y_{x,\{I \cup y\}}(u)}{1 - G^y_{y,\{I \cup x\}}(u)} \tag{1.20}$$

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

[12] It is, however, possible to identify the Laplace transform with globally defined meromorphic functions; these are canonically obtained as unique solutions of certain boundary value problems and their singularities are determined by the singularities of the relvant self-adjoint operator in a finite-dimensional space. See Sect. 3.

Lemma 1.5 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.

Lemma 1.5 has two important consequences that are obtained by setting $u = 0$ in (1.20) and by taking the derivative of (1.20) with respect to u and evaluating the result at $u = 0$:

Corollary 1.6: *Let $I \subset \Gamma_N$. Then for all $y \notin I \cup x$,*

$$\mathbb{P}\left[\tau_x^y < \tau_I^y\right] = \frac{\mathbb{P}\left[\tau_x^y < \tau_{I\cup y}^y\right]}{\mathbb{P}\left[\tau_{I\cup x}^y < \tau_y^y\right]} \tag{1.21}$$

and

$$\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_I^x\right] = \mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_{I\cup x}^x\right] + \frac{\mathbb{E}\left[\tau_x^x | \tau_x^x < \tau_{I\cup y}^x\right]}{\mathbb{P}\left[\tau_{I\cup y}^x < \tau_x^x\right]} \mathbb{P}\left[\tau_x^x < \tau_{I\cup y}^x\right] \tag{1.22}$$

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 1.5. 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 sets 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 1.7: *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*

$$\left|\frac{d^k}{du^k} g_y^x(u)\right| \leq C_k N^{k(d+3/2)+d/2} e^{N[F_N(x) - F_N(z^*(x,y))]} \tag{1.23}$$

where $z^*(y,x)$ is defined in (1.4).

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 expected return times and the transition probabilities $\mathbb{P}\left[\tau_y^x < \tau_x^x\right]$. The latter quantities will be estimated in Section 2 where we will prove the following theorem:

Theorem 1.8: *With the notation of Theorem 1.7 there exists finite positive constants c, C such that if $x \neq y \in \mathcal{M}_N$, then*

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \leq cN^{\frac{d-2}{2}} e^{-N[F_N(z^*(y,x))-F_N(y)]} \tag{1.24}$$

and

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \geq CN^{\frac{d-2}{2}} e^{-N[F_N(z^*(y,x))-F_N(y)]} \tag{1.25}$$

As will be seen in the proofs, the values of the constants c, C can be estimated rather well, but of course their values depend on the specific situation. This rather central result is based on a judicious use of well-known monotonicity properties of the probabilities considered that follow from a variational representation that is at the heart of the so called electric resistor network representation of reversible Markov chains. In this context the basic ingredient of the proof of Theorem 1.7 is known as Rayleigh's short-cut method. See the monograph by Doyle and Snell [DS] for an exposition of the relation between reversible Markov chains and electric resistor networks. We will comment on this analogy (that may provide helpful intuition) in the course of the proof without actually formally using it.

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 constant. 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 $G_y^x(u)$), 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. The final draft of the paper has benefited from comments by and discussions with G. Ben Arous, E. Bolthausen, Francesca Nardi, Enzo Olivieri, Elisabetta Scoppola, and various anonymous referee's. Finally, V.G. and A.B. 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 prove Theorem 1.8. A key ingredient is the following variational representation of the probabilities

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

that can be found in Ligget's book ([Li], p. 99, Theorem 6.1)[13]

Theorem 2.1 [Li] *Let $\mathcal{H}_x^y$ denote the space of functions*

$$\mathcal{H}_x^y \equiv \{h : \Gamma_N \to [0,1] : h(y) = 0, h(x) = 1\} \tag{2.2}$$

and define the Dirichlet form

$$\Phi_N(h) \equiv \sum_{x', x'' \in \Gamma_N} \mathbb{Q}_N(x') p_N(x', x'')[h(x') - h(x'')]^2 \tag{2.3}$$

Then

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] = \frac{1}{2\mathbb{Q}_N(y)} \inf_{h \in \mathcal{H}_x^y} \Phi_N(h) \tag{2.4}$$

Proof: See Ligget [Li], Chapter II.6. Note that the set R in Liggett's book (page 98) will be $\Gamma_N \backslash \{x\}$, and our $\mathcal{H}_x^y$ would be $H_{\Gamma_N \backslash \{x\}}$ in his notation.$\Diamond$

Proof of Theorem 1.8.: The proof of the upper bound (1.24) is very easy. We will just construct a suitable trial-function $h \in \mathcal{H}_y^x$ and bound the infimum in (2.4) by the value of the Dirichlet form Φ_N at this function.

To this end we construct a 'hyper-surface'[14] $\mathcal{S}_N \subset \Gamma_N$ separating x and y such that

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

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

$\mathcal{S}_N$ splits Γ_N into two components Γ_x and Γ_y which contain x and y, respectively. Let $\chi \in \mathcal{C}_0^\infty([0,\infty),[0,1])$ with $\chi(s) = 1$ for $s \in [0,1)$ and $\chi(s) = 0$ for $s > 2$. Then put

$$h_N(x') = \begin{cases} \chi(N^{1/2}\,\mathrm{dist}(x', S_N)), & \text{for } x' \in \Gamma_x \\ 1, & \text{for } x' \in \Gamma_y \end{cases} \tag{2.5}$$

[13] Let us note that this representation is also the basis of the electric network analogy as can be found e.g. in [DS]. The quantities $\mathbb{Q}_N(y)\mathbb{P}[\tau_x^y < \tau_y^y]$ play in fact the rôle of effective conductances of the electric network between points x and y. In fact this entire chapter could be easily reformulated in this language; however, we prefer to remain with the more probabilistic terminology and only comment on the electric analogies in footnotes.

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

Clearly, h_N is constant outside a layer $\mathcal{L}_N$ of width $N^{-1/2}$ around $\mathcal{S}_N$. Since the transition matrix $p_N(x',x'')$ vanishes for $|x'-x''| \geq CN^{-1}$, for some constant $C < \infty$ and since by construction

$$|h_N(x') - h_N(x'')| \leq cN^{1/2}|x'-x''| \tag{2.6}$$

we obtain

$$\begin{aligned} \mathbb{P}\left[\tau_y^x < \tau_x^x\right] &\leq \frac{1}{2\mathbb{Q}_N(y)}\Phi_N(h_N) \\ &\leq const.N^{-1}\sum_{x',x''\in\mathcal{L}_N}\frac{\mathbb{Q}_N(x')}{\mathbb{Q}_N(x)}p_N(x',x'') \\ &\leq \frac{C}{N}\frac{\mathbb{Q}_N(z^*(x,y))}{\mathbb{Q}_N(x)}\sum_{x'\in\mathcal{L}_N}\frac{\mathbb{Q}_N(x')}{\mathbb{Q}_N(z^*(x,y))} \end{aligned} \tag{2.7}$$

Since F_N is assumed to have a quadratic saddle point at $z^*(x,y)$, using a standard Gaussian approximation the final sum is readily seen to be $O(N^{d/2})$ which gives the upper bound (1.24).

The main task of this section will be to establish the corresponding lower bound (1.25). 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[15].

Lemma 2.2 *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*

$$\widetilde{p}_\Delta(x',x'') = \begin{cases} p_N(x',x''), & \text{if} \quad x' \neq x'', \text{and } (x',x'') \in E(\Delta) \\ 1 - \sum_{y':(x',y')\in E(\Delta)} p_N(x',y'), & \text{if} \quad x' = x'' \\ 0, & \text{else} \end{cases} \tag{2.8}$$

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

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] \geq \widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \tag{2.9}$$

Proof: To prove Lemma 2.2 from here, just note that for any $h \in \mathcal{H}_x^y$

$$\begin{aligned} \Phi_N(h) &\geq \sum_{(x',x'')\in\Delta}\mathbb{Q}_N(x')p_N(x',x'')[h(x')-h(x'')]^2 \\ &= \mathbb{Q}_N(\Delta)\sum_{(x',x'')\in\Delta}\widetilde{\mathbb{Q}}_\Delta(x')\widetilde{p}_\Delta(x',x'')[h(x')-h(x'')]^2 \end{aligned} \tag{2.10}$$

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

$$\inf_{h\in\mathcal{H}_x^y}\Phi_N(h) \geq \mathbb{Q}_N(\Delta)\inf_{h\in\mathcal{H}_x^y}\Phi_\Delta(h) = \mathbb{Q}_N(\Delta)\inf_{h\in\mathcal{H}_x^y(\Delta)}\Phi_\Delta(h) \tag{2.11}$$

where $\mathcal{H}_x^y(\Delta) \equiv \{h : \Delta \to [0,1] : h(y) = 0, h(x) = 1\}$. Thus, using Theorem 2.1 for the process $\widetilde{P}_\Delta$, we see that

$$\mathbb{P}\left[\tau_x^y < \tau_y^y\right] = \frac{1}{2\mathbb{Q}_N(y)}\inf_{h\in\mathcal{H}_x^y}\Phi_N(h) \geq \frac{1}{2\widetilde{\mathbb{Q}}_\Delta(y)}\inf_{h\in\mathcal{H}_x^y(\Delta)}\Phi_\Delta(h) = \widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \tag{2.12}$$

[15] Lemma 2.2 is most easily understood in the electrical network interpretation: the effective conductance can only be decreased by removal of elementary conductors. In this context the Lemma below is well-known as Rayleigh's monotonicity law.

which proves the lemma. $\diamondsuit$

To make use of this lemma, we will choose Δ in a special way[16]. 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 $N^{d/2}$. It 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

$$\partial D \equiv \{x \in D | \exists y \in \Gamma_N \backslash D \,, \text{s.t. } (x,y) \in E(\Gamma_N)\} \tag{2.13}$$

We have, however, that

$$\sup_{x' \in \partial D_y, x'' \in \partial D_x} |F_N(x') - F_N(x'')| \leq CN^{-1} \tag{2.14}$$

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.

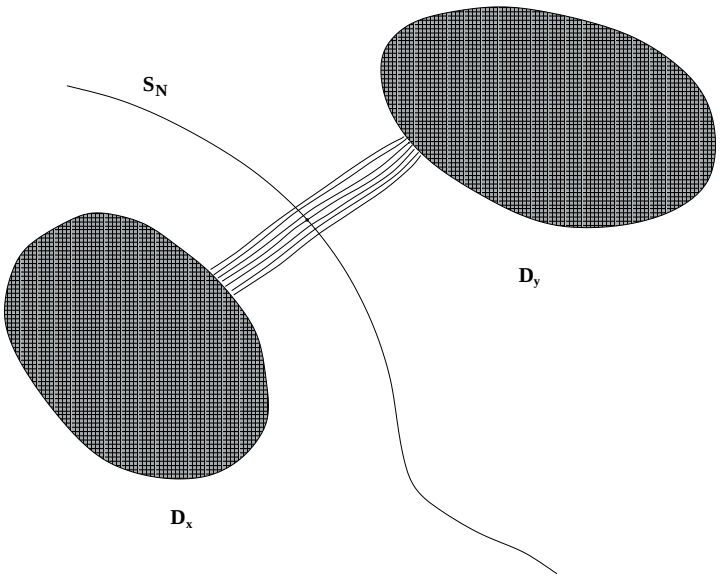
Of course we will choose the set $B \subset \mathcal{S}_N$ to be a small relative neighborhood in $\mathcal{S}_N$ of the saddle $z^*(y,x)$. In fact it will turn out to be enough to take B a disc of diameter $CN^{-1/2}$ so that its cardinality is bounded by $|B| \leq C^{d/2}N^{(d-1)/2}/\Gamma((d+1)/2)$.

For such a collection, we will set

$$\Delta \equiv D_x \cup D_y \cup \bigcup_{z \in B} V(\gamma_z) \tag{2.15}$$

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 understood in the sense of the union of the corresponding subgraphs of Γ_N.

[16] The idea of simplifying the network by removing links is of course old and known in the electrical language as Rayleigh's short-cut method. See e.g. [DS] for many examples. It is remarkable how well this method works for our purposes.



Schematic representation of the graph D

Lemma 2.3 With Δ defined above we have

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] \geq \left(1 - \frac{C^d N^{d/2}}{\Gamma((d+1)/2)} e^{-N[F_N(z^*(y,x))-E]}\right) \sum_{z\in B} \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)} \widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \tag{2.16}$$

Remark: Note that the only condition on d (in relation to N) for the validity of Theorem 1.8 comes from the demand that the term in the first bracket in (2.16) be positive. One verifies easily that this is true whenever $d \leq \alpha N$, with α sufficiently small depending on the properties of F.

***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

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_x^y < \tau_y^y\right] = \sum_{z\in B} \widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(1)}^y \leq \tau_{D_x\cup y}^y\right] \widetilde{\mathbb{P}}_\Delta\left[\tau_x^{\gamma_z(1)} < \tau_y^{\gamma_z(1)}\right] \tag{2.17}$$

By reversibility,

$$\begin{aligned} \widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(1)}^y \leq \tau_{D_x\cup y}^y\right] &= \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)} \widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(1)} \leq \tau_{D_x}^{\gamma_z(1)}\right] \\ &= \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)} \widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} \leq \tau_{D_x}^{\gamma_z(1)}\right] \widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(0)} \leq \tau_{D_x}^{\gamma_z(0)}\right] \end{aligned} \tag{2.18}$$

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

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} \leq \tau_{D_x}^{\gamma_z(1)}\right] = \widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} \leq \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \tag{2.19}$$

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

$$1 - \widetilde{\mathbb{P}}_\Delta\left[\tau_x^{\gamma_z(1)} < \tau_y^{\gamma_z(1)}\right] = \widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(1)} < \tau_x^{\gamma_z(1)}\right] = \frac{\widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(1)} < \tau_{x\cup\gamma_z(1)}^{\gamma_z(1)}\right]}{1 - \widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(1)}^{\gamma_z(1)} < \tau_{x\cup y}^{\gamma_z(1)}\right]} \tag{2.20}$$

where the second equality follows from Corollary 1.6. Now by reversibility, the numerator in (2.20) satisfies the bound

$$\begin{aligned}\widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(1)} < \tau_{x\cup\gamma_z(1)}^{\gamma_z(1)}\right] &\leq \sum_{x'\in B}\widetilde{\mathbb{P}}_\Delta\left[\tau_{x'}^{\gamma_z(1)} < \tau_{x\cup\gamma_z(1)}^{\gamma_z(1)}\right] \\ &\leq |B|\frac{\mathbb{Q}_\Delta(z^*(y,x))}{\mathbb{Q}_\Delta(\gamma_z(1))} = |B|\frac{\mathbb{Q}_N(z^*(y,x))}{\mathbb{Q}_N(\gamma_z(1))}\end{aligned} \tag{2.21}$$

On the other hand,

$$\begin{aligned}1-\widetilde{\mathbb{P}}_\Delta\left[\tau_{\gamma_z(1)}^{\gamma_z(1)} < \tau_{x\cup y}^{\gamma_z(1)}\right] &= \widetilde{\mathbb{P}}_\Delta\left[\tau_{x\cup y}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \\ &\geq \widetilde{\mathbb{P}}_\Delta\left[\tau_x^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \\ &\geq \widetilde{\mathbb{P}}_\gamma\left[\tau_x^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right]\end{aligned} \tag{2.22}$$

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

$$\widetilde{\mathbb{P}}_\gamma\left[\tau_x^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \geq CN^{-1/2} \tag{2.23}$$

Thus we get that

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_x^{\gamma_z(1)} < \tau_y^{\gamma_z(1)}\right] \geq 1 - \frac{C^d N^{d/2}}{\Gamma((d+1)/2)} e^{-N[F_N(z^*(y,x))-E]} \tag{2.24}$$

By the same procedure, we get also that

$$\widetilde{\mathbb{P}}_\Delta\left[\tau_y^{\gamma_z(0)} \leq \tau_{D_x}^{\gamma_z(0)}\right] \geq 1 - \frac{C^d N^{d/2}}{\Gamma((d+1)/2)} e^{-N[F_N(z^*(y,x))-E]} \tag{2.25}$$

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

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau_{\gamma_z(0)}^{\gamma_z(1)} < \tau_{\gamma_z(1)}^{\gamma_z(1)}\right] \geq CN^{-1/2} e^{-N[F_N(z)-F_N(\gamma_z(1))]} \tag{2.26}$$

Proof: Let $K \equiv |\gamma_z|$ denote the number of edges in the path γ_z. Let us fix the notation $\omega_0, \omega_1, \ldots, \omega_K$, 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

$$p(n) = p_N(\omega_n, \omega_{n+1}), \quad q(n) = p_N(\omega_n, \omega_{n-1}) \tag{2.27}$$

The following lemma is a well-known fact[17]:

[17] In the electrical language, it states that the resistance of a chain of resistors is the sum of the resistances[DS].

Lemma 2.5 *With the notation introduced above,*

$$\widetilde{\mathbb{P}}_{\gamma_z}\left[\tau^{\gamma_z(1)}_{\gamma_z(0)} < \tau^{\gamma_z(1)}_{\gamma_z(1)}\right] = \left[\sum_{n=1}^{K} \frac{\mathbb{Q}_N(\omega_0)}{\mathbb{Q}_N(\omega_n)} \frac{1}{q(n)}\right]^{-1} \tag{2.28}$$

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

$$\mathbb{Q}_N(\omega_0) \sum_{k=1}^{K} \frac{1}{\mathbb{Q}_N(\omega_k)} = \frac{\mathbb{Q}_N(\omega_0)}{\mathbb{Q}_N(z)} \sum_{k=1}^{K} e^{-N[F_N(z)-F_N(\omega_k)]} \tag{2.29}$$

Now in the neighborhood of z, we can certainly bound

$$F_N(z) - F_N(\omega_k) \geq c\left(\frac{k}{N}\right)^2 \tag{2.30}$$

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

$$\sum_{k=1}^{K} e^{-N[F_N(z)-F_N(\omega_k)]} \leq CN^{1/2} \tag{2.31}$$

which in turn concludes the proof of Proposition 2.4.[18] ◊◊

Combining Proposition 2.4 with Lemma 2.3, we get that

$$\begin{aligned}
\widetilde{\mathbb{P}}_{\Delta}\left[\tau^{y}_{x} < \tau^{y}_{y}\right] &\geq \left(1 - \frac{C^d N^{d/2}}{\Gamma((d+1)/2)} e^{-N[F_N(z^*(y,x))-E]}\right) \sum_{z\in B} \frac{\mathbb{Q}_N(\gamma_z(1))}{\mathbb{Q}_N(y)} \frac{\mathbb{Q}_N(z)}{\mathbb{Q}_N(\gamma_z(1))} CN^{-1/2} \\
&= e^{-N[F_N(z^*(y,x))-F_N(y)]} \left(1 - \frac{C^d N^{d/2}}{\Gamma((d+1)/2)} e^{-N[F_N(z^*(y,x))-E]}\right) \\
&\quad \times CN^{-1/2} \sum_{z\in B} e^{-N[F_N(z)-F_N(z^*(y,x))]}
\end{aligned} \tag{2.32}$$

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 neighborhood, the lower bound claimed in Theorem 1.8 follows immediately by a standard Gaussian approximation of the last sum. ◊◊

[18] Of course we could easily be more precise and identify the constant in (2.31) to leading order with the second derivative of $F(z)$ in the direction of γ (see e.g. [vK,KMST,EK]).

3. Laplace transforms of transition times in the elementary situation

In this section we shall prove Theorem 1.7, 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

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

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 1.7, 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 (1.4) 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

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

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

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

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[19]

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

[19] More precisely, one shows that for all $u \in \mathbb{C}$ for which v_u (as defined in (3.2)) exists, it coincides with the unique solution of (3.5). Moreover, since $G^x_{y,\Sigma}(u)$ is the Laplace transform of the distribution of a positive random variable, it exists for all u such that $\lim_{u' \uparrow Re(u)} G^x_{y,\Sigma}(u') < \infty$. But this implies immediately that v_u coincides with the solution of (3.5) considered as a meromorphic function of u for all u to the left of the first pole of this function. This justifies to identify v_u with the solution of (3.5) in the sequel.

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 N large enough*

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

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 $x=y$ reads (since $G^y_{y,\Sigma}(0)=\mathbb{P}[\tau^y_y\le\tau^y_\Sigma]$)

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

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 the invariant measure $\mathbb{Q}_N$ 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 $\mathbb{Q}_N(x)\ge\mathbb{Q}_N(y)$ can give a large contribution. 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 (2.12) (with γ as the subgraph Δ) gives

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

where the first inequality is a trivial consequence of $z\in\mathcal{M}_N$. But then Proposition 2.4 can be applied to get the lower bound

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

for some constant C.

To estimate the other probability in (3.12) we use Corollary 1.6 to write, for $x \notin \Sigma$,

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

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

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

Reversibility then gives the upper bound

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

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[20]

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

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

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

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 $\mathbb{E}[\tau_\Sigma^x] \leq \mathbb{E}[\tau_{\mathcal{M}_N}^x]$ from above, it follows that

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

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

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

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

$$\begin{aligned} 1 - G_{y,\Sigma}^y(u) &\geq 1 - e^u P_N(y,y) - e^u \sum_{x\notin\Sigma} p_N(y,x)\mathbb{P}[\tau_y^x = \tau_\Sigma^x] \\ &= 1 - \mathbb{P}[\tau_\Sigma^y = \tau_y^y] \\ &= \mathbb{P}[\tau_\Sigma^y < \tau_y^y]. \end{aligned} \tag{3.28}$$

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

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

for some $b > 0$, choosing δ sufficiently small in equation (3.26). This proves Lemma 3.2. $\diamondsuit$

We are now ready to give the

Proof of Theorem 1.7: Note that when $F_N(x) = F_N(z^*(x,y))$, Lemma 3.2 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 1.8 in Section 2, construct a discrete separating hypersurface $\mathcal{S}_N$ 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

$$g_y^x(u) = \sum_{z\in\mathcal{S}_N} G_{z,\Omega}^x(u) g_y^z(u), \qquad \Omega = \mathcal{M}_N \cup \mathcal{S}_N. \tag{3.30}$$

[20] Recall that $M(u)$ is a meromorphic function with poles only on the real axis. In particular $M(u)$ is finite except at isolated points.

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 (1.20) 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 3.2. 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 $g^z_y(u) = G^z_{y,\mathcal{M}_N}(u)$ we finally get, for $u \leq cN^{-d-3/2}$,

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

where the last equality is obtained by a standard Gaussian approximation as (2.32). 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 1.7.$\diamondsuit\diamondsuit$

4. Valleys, trees, and graphs

In this chapter we provide the setup for the inductive treatment of the global problem. Our purpose here is not to reinvent the cycle decomposition of Freidlin and Wentzell [WF], but rather to show that with the sharp estimates on the transition probabilities obtained in Section 2, we can also resolve the fine 'finite N' effects on the behaviour of the process. To keep the description as simple as possible, we make use of the assumption made in the introduction that F_N is "generic" in the sense that no accidental symmetries or other "unusual" structures occur. For the case of a random system, this appears natural.

4.1. The valley structure and its tree-representation

We recall from Section 1 the definition (1.4) of essential saddle points. Under our general assumptions (G1), any essential saddle has the property that 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 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 $z \in \mathcal{E}_N$, 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[21]. 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 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 a 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 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.

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

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

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 x_i, and $\mathcal{T}^c_{z_i,x_i} \equiv \mathcal{T}_N \backslash \mathcal{T}_{z_i,x_i}$. 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

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

Proposition 4.1: *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*

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

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 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 Corollary 1.6 with I consisting of a single point. This gives

$$\mathbb{P}\left[\tau_{\bar{y}}^x < \tau_y^x\right] = \frac{\mathbb{P}\left[\tau_{\bar{y}}^x < \tau_{x \cup \bar{y}}^x\right]}{\mathbb{P}\left[\tau_{y \cup \bar{y}}^x < \tau_x^x\right]} \leq \frac{\mathbb{P}\left[\tau_{\bar{y}}^x < \tau_x^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_x^x\right]} \tag{4.7}$$

Using the upper and lower bounds from Theorem 1.8 for the numerator and denominator, resp., we get

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

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

Proposition 4.1 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 4.2 *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 $\widetilde{\mathcal{T}}_N$.

A transition is called metastable, if it is admissible and if the lowest minimum in the branch $\mathcal{T}_{z,y}$ is lower than $F_N(x)$.

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 starting from z leads to. Metatable transition are also known as "exits from a cycle" in the language of [WF].

Proposition 4.3: *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*

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

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 4.3 holds, i.e. for any $z \in \mathcal{E}$ and $x \in \mathcal{T}_{z,x}$ there is a unique $y \in \mathcal{T}^c_{z,x} \cup \mathcal{M}_N$ such that (4.9) holds.

Proof: The proof is based on the fact that the process will, with probability one, hit the set $\mathcal{T}^c_{z,x}$ eventually. Thus, if we show that given x and z, for all $\tilde{y} \in \mathcal{T}^c_{z,x}$ with $\tilde{y} \neq y$, $\mathbb{P}\left[\tau^x_{\tilde{y}} < \tau^x_{\mathcal{T}^c_{z,x}\setminus\tilde{y}}\right]$ is exponentially small, the proposition follows. To simplify the notation, let us set $I = \mathcal{T}^c_{z,x} \cap \mathcal{M}_N$. Note that the case $\tilde{y} \notin \mathcal{T}_z$ is already covered by Proposition 4.1, so we assume that $\tilde{y} \in \mathcal{T}_z$. Using Corollary 1.6

$$\mathbb{P}\left[\tau^x_{\tilde{y}} < \tau^x_{I\setminus\tilde{y}}\right] = \frac{\mathbb{P}\left[\tau^x_{\tilde{y}} < \tau^x_{I\setminus\tilde{y}\cup x}\right]}{\mathbb{P}\left[\tau^x_I < \tau^x_x\right]} \tag{4.10}$$

By reversibility,

$$\mathbb{P}\left[\tau^x_{\tilde{y}} < \tau^x_{I\setminus\bar{y}\cup x}\right] = e^{N[F_N(x)-F_N(\bar{y})]}\mathbb{P}\left[\tau^{\bar{y}}_x < \tau^{\bar{y}}_I\right] \tag{4.11}$$

Now construct the separating hyper-surface $\mathcal{S}_N$ passing through z as in the proof of Theorem 1.8. Then

$$\mathbb{P}\left[\tau^{\bar{y}}_x < \tau^{\bar{y}}_I\right] = \sum_{z'\in\mathcal{S}_N} \mathbb{P}\left[\tau^{\bar{y}}_{z'} \leq \tau^{\bar{y}}_{I\cup\mathcal{S}_N}\right]\mathbb{P}\left[\tau^{z'}_x < \tau^{z'}_I\right] \tag{4.12}$$

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

$$\mathbb{P}\left[\tau^x_{\tilde{y}} < \tau^x_{I\setminus\bar{y}}\right] = \frac{1}{\mathbb{P}[\tau^x_I < \tau^x_x]}\sum_{z'\in\mathcal{S}_N} e^{-N[F_N(z')-F_N(x)]}\mathbb{P}\left[\tau^{z'}_{\tilde{y}} \leq \tau^{z'}_{I\cup\mathcal{S}_N}\right]\mathbb{P}\left[\tau^{z'}_x < \tau^{z'}_I\right] \tag{4.13}$$

Using that $\mathbb{P}[\tau^x_I < \tau^x_x] \geq \mathbb{P}[\tau^x_{y'} < \tau^x_x]$ for any $y' \in I$, together with the lower bound of Theorem 1.8 and the trivial bound $\mathbb{P}\left[\tau^{z'}_x < \tau^{z'}_I\right] \leq 1$, we see that

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

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 $\mathbb{P}\left[\tau^{z'}_{\tilde{y}} \leq \tau^{z'}_{I\cup\mathcal{S}_N}\right]$ 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 4.1, 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 $\rho > 0$

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

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

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

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

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

It 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 $\tilde{y} \neq y$, $\mathbb{P}\left[\tau_y^x < \tau_{I\setminus y}^x\right] > 1 - \exp(-NK_N)$.♢

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 4.4 *Let $x, y \in \mathcal{M}_N$. Then there is a unique sequence of admissible transitions $\mathcal{F}(x_i, z_i, x_{i+1})$, $i = 1, \ldots, k$, such that*[22]

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

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

and such that the sequence is free of cycles, i.e. the points $x_i, i = 1, \ldots, k+1$ *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 4.3 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 4.3 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).

Remark: Theorem 4.4 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 4.4. 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 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, for N large enough,*

$$cNe^{N[F_N(z)-F_N(x)]} \leq \mathbb{E}\left[\tau_y^x | \mathcal{F}(x,z,y)\right] \leq CNe^{N[F_N(z)-F_N(x)]} \tag{5.1}$$

Remark: The result is easy to understand heuristically on the basis of representation (1.22). By the definition of an admissible transition, the denominator in the second term is essentially equal to $\mathbb{P}[\tau_y^x < \tau_x^x]$ and can be estimated by Theorem 1.8. The probability multiplying this term is virtually one, and it is not very hard to see that by the ergodic theorem, the conditional expected return time in the numerator is equal to $\mathbb{Q}_N(D)/\mathbb{Q}_N(x)$, D being the valley containing x. The reader will check that this gives the behaviour claimed in the theorem. The main work we will have to do is thus to show that the contribution of the first term in (1.22) does not alter this result.

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. Recall that $z^*(x,y)$ denotes the lowest saddle connecting x and y, as defined in (1.4). Note that Theorem 1.8 implies that

$$CN^{(d-2)/2}e^{-Nd_I(x,y)} \leq \mathbb{P}\left[\tau_{I\cup y}^x \leq \tau_x^x\right] \leq c(|I|+1)N^{(d-2)/2}e^{-Nd_I(x,y)} \tag{5.3}$$

With these notations we will show the following

Lemma 5.2 *Let $I \subset \mathcal{M}_N$, and $\mathbb{P}\left[\tau_y^x \leq \tau_I^x\right] > 0$ (this can of course only fail in a one-dimensional situation). There exist $C < \infty$ such that for any $x, y \in \mathcal{M}_N$, we have that for all N large enough,*

$$\mathbb{E}\left[\tau_y^x | \tau_y^x \leq \tau_I^x\right] \leq CN^{d+3} + CN^{d+3} \sup_{x' \in \mathcal{M}_N \setminus \{I \cup y\}} \left(\mathbb{P}\left[\tau_{x'}^x < \tau_y^x | \tau_y^x \leq \tau_I^x\right] e^{Nd_I(x',y)}\right) \tag{5.4}$$

If the set $\mathcal{M}_N \setminus \{I \cup y\}$ is empty, we use the convention that the sup takes the value one.

Remark: Note that

$$\mathbb{P}\left[\tau_{x'}^x < \tau_y^x | \tau_y^x < \tau_I^x\right] = \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_y^x < \tau_I^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_I^x\right]} = \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_{y\cup I}^x\right]\mathbb{P}\left[\tau_y^{x'} < \tau_I^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_I^x\right]} \leq \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_y^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_I^x\right]} \tag{5.5}$$

and thus, using the same arguments as in the proof of Proposition 4.1, whenever $\mathbb{P}\left[\tau_y^x < \tau_I^x\right]$ is close to one, we have the more explicit bound

$$\mathbb{P}\left[\tau_{x'}^x < \tau_y^x | \tau_y^x < \tau_I^x\right] \leq C \min\left(e^{-N[F_N(z^*(x,x'))-F_N(z^*(x,y))]}, 1\right) \tag{5.6}$$

This will be important in the application of this lemma to the proof of Theorem 5.1.

Proof: The starting point of the proof of Lemma 5.2 is the observation that it holds for $I = \mathcal{M}_N$. We formulate this as a distinct lemma.

Lemma 5.3 *Assume that* $\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] > 0$. *Then there exists a constant* $C < \infty$ *such that for all* $x, y \in \mathcal{M}_N$, *and all* N *large enough,*

$$\mathbb{E}\left[\tau_y^x | \tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] \leq CN^{d+3} \tag{5.7}$$

Proof: The first important observation is then that $\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right]$ cannot be too small, i.e.

Lemma 5.4 *For any* $x, y \in \mathcal{M}_N$, *there exists* $L < \infty$, *such that for all* N *large enough*

$$\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] \geq e^{-NL} \tag{5.8}$$

Proof: Clearly

$$\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] \geq \mathbb{P}\left[X_{[TN]} = y, \forall 0 < t < [NT], X_t \notin \mathcal{M}_N | X_0 = x\right] \tag{5.9}$$

So all we have to show is that the finite-time probability in (5.9) is larger than $\exp(-LN)$ for some constant $L < \infty$. But this is obvious by just fixing a trajectory consisting of $[NT]$ steps and leading from x to y without visiting the set $\mathcal{M}_N$ on the way (making sure that T is chosen large enough to allow such a trajectory) and observing that the probability that the process is doing just this is at least c^{TN}, with c is the constant from assumption (R3). $\diamondsuit$

Next we will use the fact that in Lemma 3.2 we have shown that $g_y^x(u) \leq 2$ for $u < N^{-d-1}$. Now for all $T < \infty$,

$$\begin{aligned}
\mathbb{E}\left[\tau_y^x \mathbb{1}_{\{\tau_y^x \leq \tau_{\mathcal{M}_N}^x\}}\right] &\leq T\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] + \sum_{i=T}^{\infty} \mathbb{E}\left[\tau_y^x \mathbb{1}_{\{i \leq \tau_y^x \leq i+1\}} \mathbb{1}_{\{\tau_y^x \leq \tau_{\mathcal{M}_N}^x\}}\right] \\
&\leq T\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] + \sum_{i=T}^{\infty} (i+1) \inf_{v \geq 0} e^{-vi} \mathbb{E}\left[e^{v\tau_y^x} \mathbb{1}_{\{\tau_y^x \leq \tau_{\mathcal{M}_N}^x\}}\right] \\
&\leq T\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] + \sum_{i=T}^{\infty} (i+1) 2e^{-cN^{-(d+1)}i} \\
&\leq T\mathbb{P}\left[\tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] + Ce^{-cN^{-(d+1)}T} N^{d+1}(T + N^{d+1})
\end{aligned} \tag{5.10}$$

Now choose $T = N^{d+3}$. Then

$$\mathbb{E}\left[\tau_y^x | \tau_y^x \leq \tau_{\mathcal{M}_N}^x\right] \leq N^{d+3} + CN^{2d+4} \exp(-cN^2 + LN) \tag{5.11}$$

which for N large enough is bounded by CN^{d+3} for some constant C, as desired. $\diamondsuit$

This gives us a 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 when the starting point x is contained in the exclusion set I and when 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 prove both assertions. Note that C will denote in the course of the proof a generic finite numerical constant. We will not keep track of the changes of its value in the course of the induction. We will set $\kappa = d + 3$.

Step (i): We need only to consider sets J of cardinality $k-1$ with $x \in J \cup y$. We can assume without loss that $y \notin J$.

$$
\begin{aligned}
&\mathbb{E}\left[\tau_y^x \mathbb{I}_{\{\tau_y^x < \tau_J^x\}}\right] \\
&= \mathbb{E}\left[\tau_y^x \mathbb{I}_{\{\tau_y^x \le \tau_{\mathcal{M}_N}^x\}}\right] + \sum_{x' \in \mathcal{M}_N \setminus J \setminus \{x,y\}} \mathbb{E}\left[(\tau_{x'}^x + \tau_y^{x'}) \mathbb{I}_{\{\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\}} \mathbb{I}_{\{\tau_y^{x'} < \tau_J^{x'}\}}\right] \\
&= \mathbb{E}\left[\tau_y^x \mathbb{I}_{\{\tau_y^x \le \tau_{\mathcal{M}_N}^x\}}\right] + \sum_{x' \in \mathcal{M}_N \setminus J \setminus \{x,y\}} \Bigg(\mathbb{E}\left[\tau_{x'}^x \mathbb{I}_{\{\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\}}\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right] \\
&+ \mathbb{P}\left[\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\right] \mathbb{E}\left[\tau_y^{x'} \mathbb{I}_{\{\tau_y^{x'} < \tau_J^{x'}\}}\right] \Bigg)
\end{aligned}
\tag{5.12}
$$

Dividing by $\mathbb{P}\left[\tau_y^x < \tau_J^x\right]$, we get from (5.12) that

$$
\begin{aligned}
\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_J^x\right] &\le \mathbb{E}\left[\tau_y^x | \tau_y^x \le \tau_{\mathcal{M}_N}^x\right] \\
&+ \sum_{x' \in \mathcal{M}_N \setminus J \setminus \{x,y\}} \Bigg(\frac{\mathbb{E}\left[\tau_{x'}^x | \tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right]} \frac{\mathbb{P}\left[\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \\
&+ \mathbb{E}\left[\tau_y^{x'} | \tau_y^{x'} < \tau_J^{x'}\right] \frac{\mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right] \mathbb{P}\left[\tau_{x'}^x < \tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \Bigg)
\end{aligned}
\tag{5.13}
$$

The first summand in the last line in (5.13) produces exactly the first term in (5.4). For the second, observe that

$$
\begin{aligned}
\frac{\mathbb{P}\left[\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} &\le \frac{\mathbb{P}\left[\tau_{x'}^x \le \tau_{J \cup y}^x\right] \mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \\
&= \mathbb{P}\left[\tau_y^x < \tau_y^x | \tau_y^x < \tau_J^x\right]
\end{aligned}
\tag{5.14}
$$

which makes the entire term smaller than CN^κ. For the last term we may use the induction hypothesis

for the conditional expectation time appearing in it to get

$$\begin{aligned}
&\mathbb{E}\left[\tau_y^{x'}|\tau_y^{x'}<\tau_J^{x'}\right]\frac{\mathbb{P}\left[\tau_y^{x'}<\tau_J^{x'}\right]\mathbb{P}\left[\tau_{x'}^x<\tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_y^x<\tau_J^x\right]}\\
&\leq CN^\kappa+CN^\kappa\sup_{x''\in\mathcal{M}_N\setminus\{J\cup y\}}\frac{\mathbb{P}\left[\tau_{x''}^{x'}<\tau_y^{x'}<\tau_J^{x'}\right]\mathbb{P}\left[\tau_{x'}^x<\tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_y^x<\tau_J^x\right]}e^{Nd_J(x'',y)}\\
&=CN^\kappa+CN^\kappa\sup_{x''\in\mathcal{M}_N\setminus\{J\cup y\}}\frac{\mathbb{P}\left[\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}{\mathbb{P}\left[\tau_y^x<\tau_J^x\right]}e^{Nd_J(x'',y)}\frac{\mathbb{P}\left[\tau_{x''}^{x'}<\tau_y^{x'}<\tau_J^{x'}\right]\ P\left[\tau_{x'}^x<\tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}
\end{aligned}\tag{5.15}$$

But the last factor satisfies

$$\begin{aligned}
\frac{\mathbb{P}\left[\tau_{x''}^{x'}<\tau_y^{x'}<\tau_J^{x'}\right]\ P\left[\tau_{x'}^x<\tau_{\mathcal{M}_N}^x\right]}{\mathbb{P}\left[\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}&\leq\frac{\mathbb{P}\left[\tau_{x''}^{x'}<\tau_y^{x'}<\tau_J^{x'}\right]\ P\left[\tau_{x'}^x<\tau_{J\cup y\cup x''}^x\right]}{\mathbb{P}\left[\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}\\
&\leq\frac{\mathbb{P}\left[\tau_{x'}^x<\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}{\mathbb{P}\left[\tau_{x''}^x<\tau_y^x<\tau_J^x\right]}\leq 1
\end{aligned}\tag{5.16}$$

Thus (5.15) actually gives a term of the desired form. This proves the first inductive step.

To complete the proof we need to turn to

Step (ii): Here we must consider J such that $x\notin J\cup y$.

We will first consider the sub-case when x is such that

$$d_J(x,y)=\sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\}}d_J(x',y)\tag{5.17}$$

Note that in this situation

$$\begin{aligned}
\sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\}}\mathbb{P}\left[\tau_{x'}^x<\tau_y^x|\tau_y^x<\tau_J^x\right]e^{Nd_J(x',y)}&\geq\frac{\mathbb{P}\left[\tau_x^x<\tau_y^x<\tau_J^x\right]}{\mathbb{P}\left[\tau_y^x<\tau_J^x\right]}e^{Nd_J(x,y)}\\
&=\mathbb{P}\left[\tau_x^x<\tau_{y\cup J}^x\right]e^{Nd_J(x,y)}
\end{aligned}\tag{5.18}$$

But

$$\mathbb{P}\left[\tau_x^x<\tau_{y\cup J}^x\right]=1-\mathbb{P}\left[\tau_{J\cup y}^x<\tau_x^x\right]\geq 1-CN^{(d-2)/2}e^{-Nd_J(x,y)}\tag{5.19}$$

so that in this case it will be enough to prove that

$$\mathbb{E}\left[\tau_y^x|\tau_y^x<\tau_J^x\right]\leq CN^\kappa e^{Nd_J(x,y)}\tag{5.20}$$

Recall from Corollary 1.6 that

$$\mathbb{E}\left[\tau_y^x|\tau_y^x<\tau_J^x\right]=\mathbb{E}\left[\tau_y^x|\tau_y^x<\tau_{J\cup x}^x\right]+\frac{\mathbb{E}\left[\tau_x^x|\tau_x^x<\tau_{J\cup y}^x\right]}{\mathbb{P}\left[\tau_{J\cup y}^x<\tau_x^x\right]}\mathbb{P}\left[\tau_x^x<\tau_{J\cup y}^x\right]\tag{5.21}$$

By the induction hypothesis, the first term in (5.21) satisfies the bound

$$\mathbb{E}\left[\tau_y^x|\tau_y^x<\tau_{J\cup x}^x\right]\leq CN^\kappa\sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}}\mathbb{P}\left[\tau_{x'}^x<\tau_y^x|\tau_y^x<\tau_{J\cup x}^x\right]e^{Nd_{J\cup x}(x',y)}\leq CN^\kappa e^{Nd_J(x,y)}\tag{5.22}$$

as desired (note that $d_{J\cup x}(x',y) \le d_J(x',y)$ by definition and $d_J(x',y) \le d_J(x,y)$ by assumption (5.17)). For the second term, we use again the induction hypothesis to get that

$$\begin{aligned}
&\frac{\mathbb{E}\left[\tau_x^x | \tau_x^x < \tau_{J\cup y}^x\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} \mathbb{P}\left[\tau_x^x < \tau_{J\cup y}^x\right] \\
&\le CN^\kappa \sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_x^x < \tau_{J\cup y}^x\right]}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} e^{Nd_{J\cup y}(x',x)} + \frac{CN^\kappa}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} \\
&\le C^{-1}N^{-(d-2)/2} e^{Nd_J(x,y)} CN^\kappa \sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}} \left(\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right] e^{Nd_{J\cup y}(x',x)}\right) + \frac{CN^\kappa}{\mathbb{P}\left[\tau_{J\cup y}^x < \tau_x^x\right]} \\
&\le C^{-1}ce^{Nd_J(x,y)} CN^\kappa \sup_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}} \left(e^{-N\left(F_N(z^*(x,x'))-F_N(x)\right)} e^{Nd_{J\cup y}(x',x)}\right) \\
&+ N^{-(d-2)/2} c^{-1} CN^\kappa e^{Nd_J(x,y)}
\end{aligned} \tag{5.23}$$

It remains to show that

$$\begin{aligned}
&e^{-N\left(F_N(z^*(x,x'))-F_N(x)\right)} e^{Nd_{J\cup y}(x',x)} \\
&= e^{-N\left(F_N(z^*(x,x'))-F_N(x)\right)} e^{\inf_{x''\in J\cup y\cup x} N\left(F_N(z^*(x'',x'))-F(x')\right)}
\end{aligned} \tag{5.24}$$

is bounded by one. We consider two cases:

(i) Assume that

$$\inf_{x''\in J\cup y\cup x} F_N(z^*(x'',x')) - F(x') = F_N(z^*(x,x')) - F_N(x') \tag{5.25}$$

Define $z^*(x,A)$ by $F_N(z^*(x,A)) = \inf_{x'\in A} F_N(z^*(x,x'))$. Then in this case, x is 'closer' to x' then to $J\cup y$, so that

$$z^*(x, J\cup y) = z^*(x', J\cup y) \tag{5.26}$$

Thus by assumption (5.17)

$$\inf_{x''\in J\cup y} F_N(z^*(x'',x')) - F(x') = F_N(z^*(x,J\cup y)) - F_N(x') < F_N(z^*(x,J\cup y)) - F(x) \tag{5.27}$$

This implies that $F_N(x) < F_N(x')$, and since

$$e^{-N\left(F_N(z^*(x,x'))-F_N(x)\right)} e^{\inf_{x''\in J\cup y\cup x} N\left(F_N(z^*(x'',x'))-F(x')\right)} \le e^{-N\left(F_N(x')-F_N(x)\right)} \tag{5.28}$$

the sup is bounded by 1 as desired.

(ii) We are left with the case

$$\inf_{x''\in J\cup y\cup x} F_N(z^*(x'',x')) - F(x') < F_N(z^*(x,x')) - F_N(x') \tag{5.29}$$

Here x' is closer to $J\cup y$ than to x, and so

$$z^*(x, J\cup y) = z^*(x, x') \tag{5.30}$$

But then, again by (5.17)

$$F_N(z^*(x,x') - F_N(x) = F_N(z^*(x,J\cup y)) - F_N(x) \ge \inf_{x''\in J\cup y\cup x} F_N(z^*(x',x'')) - F_N(x') \tag{5.31}$$

which has the desired implication. This covers the case when (5.17) holds.

Let us now turn to the case when (5.17) does not hold. Let $x^* \in \mathcal{M}_N$ be such that

$$d_J(x^*, y) = \sup_{x' \in \mathcal{M}_N \setminus \{J \cup y\}} d_J(x', y) \tag{5.32}$$

By assumption $x^* \neq x$. We can write

$$\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_J^x\right] = \frac{\mathbb{E}\left[\tau_y^x \mathbb{I}_{\{\tau_y^x < \tau_J^x\}} \mathbb{I}_{\{\tau_{x^*}^x > \tau_y^x\}}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} + \frac{\mathbb{E}\left[\tau_y^x \mathbb{I}_{\{\tau_y^x < \tau_J^x\}} \mathbb{I}_{\{\tau_{x^*}^x < \tau_y^x\}}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \equiv (I) + (II) \tag{5.33}$$

Now

$$(I) = \mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_{J \cup x^*}^x\right] \frac{\mathbb{P}\left[\tau_y^x < \tau_{J \cup x^*}^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \tag{5.34}$$

and using the induction hypothesis,

$$\begin{aligned} (I) &\leq CN^\kappa + CN^\kappa \sup_{x' \in \mathcal{M}_N \setminus \{J \cup y \cup x^*\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_y^x < \tau_{J \cup x^*}^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} e^{N d_{J \cup x^*}(x', y)} \\ &\leq CN^\kappa + CN^\kappa \sup_{x' \in \mathcal{M}_N \setminus \{J \cup y\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_y^x < \tau_J^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} e^{N d_J(x', y)} \end{aligned} \tag{5.35}$$

as desired. On the other hand,

$$\begin{aligned} (II) =& \frac{\mathbb{E}\left[\tau_{x^*}^x \mathbb{I}_{\{\tau_{x^*}^x < \tau_{J \cup y}^x\}}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \mathbb{P}\left[\tau_y^{x^*} < \tau_J^{x^*}\right] + \frac{\mathbb{E}\left[\tau_y^{x^*} \mathbb{I}_{\{\tau_y^{x^*} < \tau_J^{x^*}\}}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \mathbb{P}\left[\tau_{x^*}^x < \tau_{J \cup y}^x\right] \\ \equiv& (IIa) + (IIb) \end{aligned} \tag{5.36}$$

To treat (IIa) we can use the induction hypothesis to get

$$\begin{aligned} (IIa) &= \mathbb{E}\left[\tau_{x^*}^x | \tau_{x^*}^x < \tau_{J \cup y}^x\right] \mathbb{P}\left[\tau_{x^*}^x < \tau_y^x | \tau_y^x < \tau_J^x\right] \\ &\leq CN^\kappa + CN^\kappa \sup_{x' \in \mathcal{M}_N \setminus \{J \cup x^* \cup y\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_{x^*}^x < \tau_{J \cup y}^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \mathbb{P}\left[\tau_y^{x^*} < \tau_J^{x^*}\right] e^{N d_{J \cup y}(x', x^*)} \\ &= CN^\kappa + CN^\kappa \sup_{x' \in \mathcal{M}_N \setminus \{J \cup x^* \cup y\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_{x^*}^x < \tau_y^x < \tau_J^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} e^{N d_{J \cup y}(x', x^*)} \\ &\leq CN^\kappa + CN^\kappa \sup_{x' \in \mathcal{M}_N \setminus \{J \cup x^* \cup y\}} \frac{\mathbb{P}\left[\tau_{x'}^x < \tau_y^x < \tau_J^x\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} e^{N d_J(x', y)} \end{aligned} \tag{5.37}$$

since $d_{J \cup y}(x', x^*) = \inf_{x'' \in J \cup y \cup x^*} [F_N(z^*(x', x'')) - F_N(x')] \leq d_J(x', y)$, and finally

$$\begin{aligned} (IIb) &= \mathbb{E}\left[\tau_y^{x^*} | \tau_y^{x^*} < \tau_J^{x^*}\right] \frac{\mathbb{P}\left[\tau_{x^*}^x < \tau_{J \cup y}^x\right] \mathbb{P}\left[\tau_y^{x^*} < \tau_J^{x^*}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \\ &= \mathbb{E}\left[\tau_y^{x^*} | \tau_y^{x^*} < \tau_J^{x^*}\right] \mathbb{P}\left[\tau_{x^*}^x < \tau_y^x | \tau_y^x < \tau_J^x\right] \\ &\leq CN^\kappa \mathbb{P}\left[\tau_{x^*}^x < \tau_y^x | \tau_y^x < \tau_J^x\right] e^{N d_J(x^*, y)} \end{aligned} \tag{5.38}$$

where the last line is obtained by using that the conditional expectation in the one-but-last line is of the form considered just before. Putting all together we see that the lemma is proven.♢

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 $I = \mathcal{T}^c_{z,x} \setminus y \cap \mathcal{M}_N$.

We have already seen in (5.19) that $\mathbb{P}\left[\tau^x_x < \tau^x_{I\cup y}\right]$ differs from 1 only by an exponentially small term. Moreover, using (4.21), we see that in the case of an admissible transition,

$$\begin{aligned}\mathbb{P}\left[\tau^x_y < \tau^x_x\right] \leq & \mathbb{P}\left[\tau^x_{I\cup y} < \tau^x_x\right] \leq \mathbb{P}\left[\tau^x_y < \tau^x_x\right] + \sum_{y'\in I} \mathbb{P}\left[\tau^x_{y'} < \tau^x_{x\cup I}\right] \\ & \leq \mathbb{P}\left[\tau^x_y < \tau^x_x\right](1 + |I| e^{-NK_N})\end{aligned} \tag{5.39}$$

Therefore (5.21) implies in this case that

$$\mathbb{E}\left[\tau^x_y | \tau^x_y < \tau^x_I\right] = \frac{\mathbb{E}\left[\tau^x_x | \tau^x_x < \tau^x_{I\cup y}\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_x\right]}\left(1 + \mathcal{O}(e^{-NK_N})\right) + \mathbb{E}\left[\tau^x_y | \tau^x_y < \tau^x_{I\cup x}\right] \tag{5.40}$$

Now we have already precise bounds on the denominator of the first term (see Section 2), and using the upper bound from Lemma 5.2 (taking into account that we are now in the situation of the remark following that lemma!) 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 (2.8) 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

$$\widetilde{\mathbb{E}}_D \tau^y_y = \frac{1}{\widetilde{\mathbb{Q}}_D(y)} = \frac{\mathbb{Q}_N(D)}{\mathbb{Q}_N(y)} \tag{5.41}$$

This quantity can be estimated very precisely via sharp large deviation estimates. It will typically exhibit 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

$$g^y_y(u) = \mathbb{E}\left[e^{u\tau^y_y} \mathbb{1}_{\{\tau^y_y \leq \tau^y_{\mathcal{M}_N}\}} \mathbb{1}_{\{\tau^y_{\partial D} < \tau^y_y\}}\right] + \mathbb{E}\left[e^{u\tau^y_y} \mathbb{1}_{\{\tau^y_{\partial D} > \tau^y_y\}}\right] \tag{5.42}$$

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 $\{\tau^y_{\partial D} > \tau^y_y\}$ the processes $X(t)$ and $\widetilde{X}_{D^+}(t)$ have the same law until time τ^y_y, so that

$$\mathbb{E}\left[e^{u\tau^y_y} \mathbb{1}_{\{\tau^y_{\partial D} > \tau^y_y\}}\right] = \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau^y_y} \mathbb{1}_{\{\tau^y_{\partial D} > \tau^y_y\}}\right] \leq \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau^y_y}\right] \tag{5.43}$$

We will show that this is the dominant term in (5.42), the first summand on the right being exponentially small. Indeed

$$\begin{aligned}\mathbb{E}\left[e^{u\tau^y_y} \mathbb{1}_{\{\tau^y_y \leq \tau^y_{\mathcal{M}_N}\}} \mathbb{1}_{\{\tau^y_{\partial D} < \tau^y_y\}}\right] &= \sum_{z\in\partial D} \mathbb{E}\left[e^{u(\tau^y_z + \tau^z_y)} \mathbb{1}_{\{\tau^y_z \leq \tau^y_{y\cup\partial D}\}} \mathbb{1}_{\{\tau^z_y \leq \tau^z_{\mathcal{M}_N}\}}\right] \\ &= \sum_{z\in\partial D} \mathbb{E}\left[e^{u\tau^y_z} \mathbb{1}_{\{\tau^y_z \leq \tau^y_{y\cup\partial D}\}}\right] \mathbb{E}\left[e^{u\tau^z_y} \mathbb{1}_{\{\tau^z_y \leq \tau^z_{\mathcal{M}_N}\}}\right]\end{aligned} \tag{5.44}$$

Using Theorem 1.7, 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

$$g_y^y(u) \leq \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right] + CN^{\kappa}e^{-N[F_N(z^*)-F_N(y)]} \tag{5.45}$$

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

$$\widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\mathbb{1}_{\{\tau_{\partial D}^y > \tau_y^y\}}\right] = \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right] - \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\mathbb{1}_{\{\tau_{\partial D}^y \leq \tau_y^y\}}\right] \tag{5.46}$$

But the last term in (5.46) can be treated precisely as in (5.44), so that we arrive at

$$g_y^y(u) \geq \widetilde{\mathbb{E}}_{D^+}\left[e^{u\tau_y^y}\right] - CN^{\kappa}e^{-N[F_N(z^*)-F_N(y)]} \tag{5.47}$$

Differentiating and using reversibility and the upper bounds from Theorem 1.7, as well as the obvious lower bound

$$g_y^y(0) = 1 - \mathbb{P}\left[\tau_{\mathcal{M}_N}^y < \tau_y^y\right] \geq 1 - \sum_{x\in\mathcal{M}_N\setminus y} g_x^y(0) \geq 1 - |\mathcal{M}_N|N^{d-1}e^{-N[F_N(z^*)-F_N(y)]} \tag{5.48}$$

gives in the same way that

$$\frac{\dot{g}_y^y(0)}{g_y^y(0)} = \frac{\mathbb{Q}_N(D^+)}{\mathbb{Q}_N(y)} + \mathcal{O}(N^{\kappa})e^{-N[F_N(z^*)-F_N(y)]} \tag{5.49}$$

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 5.5: *Let $\mathcal{F}(x,z,y)$ be an admissible transition. Let D denote the level set of the saddle z. Then for $K_N > 0$ satisfying $N^{1-\alpha}K_N \uparrow \infty$, for some $\alpha > 0$,*

$$\mathbb{E}\left[\tau_x^x | \tau_x^x < \tau_{I\cup y}^x\right] = \frac{\mathbb{Q}_N(D^+)}{\mathbb{Q}_N(x)} + \mathcal{O}\left(e^{-NK_N}\right) \tag{5.50}$$

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

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

$$\mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{I\cup y}^x\}}\right] = \mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{I\cup y}^x\}}\mathbb{1}_{\{\tau_x^x < \tau_{\partial D}^x\}}\right] + \mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{I\cup y}^x\}}\mathbb{1}_{\{\tau_x^x > \tau_{\partial D}^x\}}\right] \tag{5.51}$$

The first summand gives precisely

$$\mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{I\cup y}^x\}}\mathbb{1}_{\{\tau_x^x < \tau_{\partial D}^x\}}\right] = \mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{\partial D}^x\}}\right] \tag{5.52}$$

so that from this term alone we would get the same estimate as in (5.49). We have to show that the second term does not give a relevant contribution. Note that as in (5.44) we can split paths at the first visits to ∂D. This gives

$$\begin{aligned}\mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau_{I\cup y}^x\}}\mathbb{1}_{\{\tau_x^x > \tau_{\partial D}^x\}}\right] = \sum_{z'\in\partial D} &\mathbb{E}\left[\tau_{z'}^x \mathbb{1}_{\{\tau_{z'}^x \leq \tau_{\partial D}^x\}}\mathbb{1}_{\{\tau_{z'}^x < \tau_x^x\}}\right]\mathbb{P}\left[\tau_x^{z'} < \tau_{I\cup y}^{z'}\right] \\ &+ \mathbb{P}\left[\tau_{z'}^x \leq \tau_{\partial D}^x, \tau_{z'}^x < \tau_x^x\right]\mathbb{E}\left[\tau_x^{z'}\mathbb{1}_{\{\tau_x^{z'} < \tau_{I\cup y}^{z'}\}}\right]\end{aligned} \tag{5.53}$$

Now $\mathbb{P}\left[\tau_{z'}^x \leq \tau_{\partial D}^x, \tau_{z'}^x < \tau_x^x\right]$ is bounded by $e^{-N[F_N(z)-F_N(x)]}$, and by reversibility $\mathbb{E}\left[\tau_{z'}^x \mathbb{1}_{\{\tau_{z'}^x \leq \tau_{\partial D}^x\}} \mathbb{1}_{\{\tau_{z'}^x < \tau_x^x\}}\right] \leq e^{-N[F_N(z)-F_N(x)]}\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\{\tau_x^{z'} < \tau_{\partial D}^{z'}\}}\right]$, 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.

$$\begin{aligned}
&\mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\{\tau_x^{z'} < \tau_{I\cup y}^{z'}\}}\right] = \mathbb{E}\left[\tau_x^{z'} \mathbb{1}_{\{\tau_x^{z'} \leq \tau_{\mathcal{M}_N}^{z'}\}}\right] \\
&+ \sum_{x' \in \mathcal{M}_N \setminus I} \left(\mathbb{E}\left[\tau_{x'}^{z'} \mathbb{1}_{\{\tau_x^{z'} \leq \tau_{\mathcal{M}_N}^{z'}\}}\right] \mathbb{P}\left[\tau_x^{x'} < \tau_{I\cup y}^{x'}\right] + \mathbb{P}\left[\tau_{x'}^{z'} \leq \tau_{\mathcal{M}_N}^{z'}\right] \mathbb{E}\left[\tau_x^{x'} \mathbb{1}_{\{\tau_x^{x'} < \tau_{I\cup y}^{x'}\}}\right]\right)
\end{aligned} \tag{5.54}$$

Lemma 5.2 and Theorem 1.7 can now be used on the expectations in (5.54), 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 1.8 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 5.6 *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*

$$\mathbb{E}\left[e^{v\tau_y^x/\bar{\tau}_y^x} | \mathcal{F}(x,z,y)\right] = \frac{1}{1-v} + e^{-NK_N} f(v) \tag{5.55}$$

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

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 1.7 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

$$t_I(x,y) \equiv e^{Nd_I(x,y)} \tag{5.56}$$

and

$$T_I(y) \equiv \sup_{x \in \mathcal{M}_N \setminus \{I \cup y\}} t_I(x,y) \tag{5.57}$$

With the notation of Section 1, we define

$$\widehat{G}^x_{y,I}(v) \equiv \frac{G^x_{y,I}\left(v/T_I(y)\right)}{\mathbb{P}\left[\tau^x_y \leq \tau^x_I\right]} \tag{5.58}$$

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

Lemma 5.7 *Let $I \subset \mathcal{M}_N$, and let $x, y \in \mathcal{M}_N$. Then $\widehat{G}^x_{y,I}(v)$ can be represented in the form*

$$\begin{aligned} \widehat{G}^x_{y,I}(v) &= a^0_{x,y,I}(v/T_I(y)) \\ &+ \sum_{x' \in \mathcal{M}_N \setminus \{I \cup y\}} \mathbb{P}\left[\tau^x_{x'} < \tau^x_y | \tau^x_y \leq \tau^x_I\right] a^{x'}_{x,y,I}\left(v \frac{t_I(x',y)}{T_I(y)}\right) \end{aligned} \tag{5.59}$$

where $a^0_{x,y,I}$ and $a^{x'}_{x,y,I}$, for any $x' \in \mathcal{M}_N \setminus \{I \cup y\}$ are complex functions that have the properties (for a finite constant C, and $\kappa = d+3$):

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

(ii) They are real and 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 $|Re(v)| < CN^{-\kappa}$, for $|Re(v)| < \frac{C}{2}N^{-\kappa}$ we have,

$$\left|\frac{d^n}{dv^n} a(v)\right| \leq n! C \frac{2^n N^{n\kappa}}{C^n} \sup_{v : Re(v) < CN^{-\kappa}} a(v) \tag{5.60}$$

This will be used repeatedly in the sequel.

The first step in the proof is to show that (5.59) holds if $I = \mathcal{M}_N$. The proof of this fact is completely analogous to that of Lemma 5.3 and we will skip the details. Let us just mention that we will get the boundedness of $\widehat{G}^x_{y,\mathcal{M}_N}(v)$ in a domain $Re(u) < cN^{-d-3}$ (while analyticity was established in the larger domain $Re(u) < cN^{-d-1}$ in Section 3). 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,

$$\widehat{G}^x_{y,J}(v) = \frac{g^x_y(v/T_J(y))}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]} + \sum_{x' \in \mathcal{M}_N \setminus \{J \cup y\}} \frac{g^x_{x'}(v/T_J(y)) \mathbb{P}\left[\tau^{x'}_y < \tau^{x'}_J\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]} \widehat{G}^{x'}_{y,J}(v) \tag{5.61}$$

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

$$\begin{aligned} \widehat{G}^x_{y,J}(v) &= \frac{g^x_y(v/T_J(y))}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]} + \sum_{x'' \in \mathcal{M}_N \setminus \{J \cup y\}} \frac{g^x_{x'}(v/T_J(y)) \mathbb{P}\left[\tau^{x'}_y < \tau^{x'}_J\right] \mathbb{P}\left[\tau^x_{x'} \leq \tau^x_{\mathcal{M}_N}\right]}{g^x_{x'}(0) \mathbb{P}\left[\tau^x_y < \tau^x_J\right]} a^0_{x',y,J}(v/T_J(y)) \\ &+ \sum_{x' \in \mathcal{M}_N \setminus \{J \cup y\}} \sum_{x'' \in \mathcal{M}_N \setminus \{J \cup y\}} \frac{g^x_{x'}(v/T_J(y))}{g^x_{x'}(0)} \frac{\mathbb{P}\left[\tau^x_{x'} \leq \tau^x_{\mathcal{M}_N}\right] \mathbb{P}\left[\tau^{x'}_y < \tau^{x'}_J\right]}{\mathbb{P}\left[\tau^x_y < \tau^x_J\right]} \\ &\times \mathbb{P}\left[\tau^{x'}_{x''} < \tau^{x'}_y | \tau^{x'}_y < \tau^{x'}_J\right] a^{x''}_{x',y,J}\left(v \frac{t_J(x'',y)}{T_J(y)}\right) \end{aligned} \tag{5.62}$$

Remember that in the proof of Lemma 5.2 ((5.15)-(5.16)) we have established that

$$\frac{\mathbb{P}\left[\tau_{x'}^x \le \tau_{\mathcal{M}_N}^x\right]\mathbb{P}\left[\tau_y^{x'} < \tau_J^{x'}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]}\mathbb{P}\left[\tau_{x''}^{x'} < \tau_y^{x'} | \tau_y^{x'} < \tau_J^{x'}\right] \le \mathbb{P}\left[\tau_{x''}^x < \tau_y^x | \tau_y^x < \tau_J^x\right] \tag{5.63}$$

which, since the analytic properties of $a_{x',y,J}^{x''}$ and $a_{x,y,J}^{x''}$ are the same, shows that (5.62) provides the claimed representation.

The more subtle part of the proof concerns the case $x \notin J$. We proceed again as in the proof of Lemma 5.2 and consider first the case when $T_J(y) = d_J(x,y)$. Note again that in this case, the representation in the lemma reduces to

$$\widehat{G}_{y,J}^x(v) = a_{x,y,J}^x(v) \tag{5.64}$$

since all the other terms in the sum (5.59) are smaller and more regular.

We use of course that

$$G_{y,J}^x(u) = \frac{G_{y,J\cup x}^x(u)}{1 - G_{x,J\cup y}^x(u)} \tag{5.65}$$

which implies

$$\begin{aligned}\widehat{G}_{y,J}^x(v) &= \frac{\mathbb{P}\left[\tau_y^x < \tau_{J\cup x}^x | \tau_y^x < \tau_J^x\right]\widehat{G}_{y,J\cup x}^x\left(v\frac{T_{J\cup x}(y)}{T_J(y)}\right)}{1 - \mathbb{P}\left[\tau_x^x < \tau_{J\cup y}^x\right]\widehat{G}_{x,J\cup y}^x\left(v\frac{T_{J\cup y}(x)}{T_J(y)}\right)} \\ &= \frac{\widehat{G}_{y,J\cup x}^x\left(v\frac{T_{J\cup x}(y)}{T_J(y)}\right)}{1 - \frac{\mathbb{P}\left[\tau_x^x < \tau_{J\cup y}^x\right]}{\mathbb{P}\left[\tau_x^x > \tau_{J\cup y}^x\right]}v\frac{T_{J\cup y}(x)}{T_J(y)}\int_0^1 d\theta \widehat{G}'^x_{x,J\cup y}\left(\theta v\frac{T_{J\cup y}(x)}{T_J(y)}\right)}\end{aligned} \tag{5.66}$$

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 (similar to (5.22))

$$\begin{aligned}&\frac{T_{J\cup y}(x)}{T_J(y)}\widehat{G}'^x_{x,J\cup y}\left(\theta v\frac{T_{J\cup y}(x)}{T_J(y)}\right) \\ &= \frac{T_{J\cup y}(x)}{T_J(y)}\Bigg[\sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}}\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right] \\ &\times a'^{x'}_{x,x,J\cup y}\left(\theta v\frac{t_{J\cup y}(x',x)}{T_J(y)}\right)\frac{t_{J\cup y}(x',x)}{T_{J\cup y(x)}} + a^0_{x,x,J\cup y}\left(\frac{\theta v}{T_J(y)}\right)\frac{1}{T_{J\cup y}(x)}\Bigg] \\ &= \sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}}\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right]\frac{t_{J\cup y}(x',x)}{T_J(y)}a'^{x'}_{x,x,J\cup y}\left(\theta v\frac{t_{J\cup y}(x',x)}{T_J(y)}\right) \\ &+ a^0_{x,x,J\cup y}\left(\frac{\theta v}{T_J(y)}\right)\frac{1}{T_J(y)}\end{aligned} \tag{5.67}$$

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

$$\begin{aligned}&\left|\frac{T_{J\cup y}(x)}{T_J(y)}\widehat{G}'^x_{x,J\cup y}\left(\theta v\frac{T_{J\cup y}(x)}{T_J(y)}\right)\right| \\ &\le \sum_{x'\in\mathcal{M}_N\setminus\{J\cup x\cup y\}}\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right]\frac{t_{J\cup y}(x',x)}{T_J(y)}CN^\kappa + \frac{cN^\kappa}{T_J(y)}\end{aligned} \tag{5.68}$$

In the second part of the proof of Lemma 5.2 ((5.23) we have shown that

$$\mathbb{P}\left[\tau_{x'}^x < \tau_x^x\right] t_{J\cup y}(x', x) \leq CN^{(d-2)/2} \tag{5.69}$$

Thus we deduce from (5.68) that

$$\left| \frac{T_{J\cup y}(x)}{T_J(y)} \widehat{G'}^x_{x,J\cup y}\left(\theta v \frac{T_{J\cup y}(x)}{T_J(y)}\right)\right| \leq \frac{CN^\kappa}{T_J(y)} \tag{5.70}$$

We see thus that the numerator in (5.66) will not vanish for $v < C^{-1}N^{-\kappa}$. Thus $\widehat{G}^x_{y,J}(v)$ is bounded and analytic in the strip $|Re(v)| < C^{-1}N^{-\kappa}$.

In the general case, we proceed as in the proof of Lemma 5.2 and decompose all paths into those avoiding the deepest minimum x^* and those visiting it. A simple computation yields then

$$\begin{aligned}\widehat{G}^x_{y,J}(v) =& \widehat{G}^x_{y,J\cup x^*}\left(v\frac{T_{J\cup x^*}(y)}{T_J(y)}\right)\frac{\mathbb{P}\left[\tau_y^x < \tau^x_{J\cup x^*}\right]}{\mathbb{P}\left[\tau_y^x < \tau_J^x\right]} \\ &+ \mathbb{P}\left[\tau^x_{x^*} < \tau_y^x | \tau_y^x < \tau_J^x\right] \widehat{G}^x_{x^*,J\cup y}\left(v\frac{T_{J\cup y}(x^*)}{T_J(y)}\right)\widehat{G}^{x^*}_{y,J}(v)\end{aligned} \tag{5.71}$$

Clearly the first term is by the inductive assumption at least as good as desired, and since we have just shown that the last factor in the second term is bounded and analytic in $|Re(v)| < CN^{-\kappa}$, using the inductive assumption for $\widehat{G}^x_{x^*,J\cup y}\left(v\frac{T_{J\cup y}(x^*)}{T_J(y)}\right)$, one sees easily that the desired representation holds. This concludes the proof of the lemma.$\diamondsuit$

We are now ready to prove Theorem 5.6. 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

$$\widetilde{G}^x_{y,I}(v) \equiv \widehat{G}^x_{y,I}(vT_I(y)/\bar{T}) \tag{5.72}$$

with

$$\bar{T} \equiv \frac{\mathbb{E}\left[\tau_x^x \mathbb{1}_{\{\tau_x^x < \tau^x_{y\cup I}\}}\right]}{\mathbb{P}\left[\tau^x_{I\cup y} < \tau_x^x\right]} \tag{5.73}$$

Note that we use $\bar{T}$ instead of $\bar{\tau}_y^x$ in the proof because this will simplify the following formulas. But recall from the proof of Theorem 5.1 that

$$\left|\frac{\mathbb{E}\left[\tau_y^x | \tau_y^x < \tau_I^x\right]}{\bar{T}} - 1\right| \leq e^{-NK_N} \tag{5.74}$$

Thus in the final results they can be interchanged without harm. Then

$$\widetilde{G}^x_{y,I}(v) = \frac{\mathbb{P}\left[\tau_y^x < \tau^x_{I\cup x} | \tau_y^x < \tau_I^x\right] \widehat{G}^x_{y,I\cup x}\left(v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1 - \mathbb{P}\left[\tau_x^x < \tau^x_{I\cup y}\right]\widehat{G}^x_{x,I\cup y}\left(v\frac{T_{I\cup y}(x)}{\bar{T}}\right)} \tag{5.75}$$

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

$$\begin{aligned}
&\widetilde{G}^x_{y,I}(v) \\
&= \frac{\mathbb{P}[\tau^x_y < \tau^x_{I\cup x} | \tau^x_y < \tau^x_I] \left(1 + \frac{v}{\bar{T}}\mathbb{E}[\tau^x_y | \tau^x_y < \tau^x_{I\cup x}] + \frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)\right)}{\mathbb{P}[\tau^x_{y\cup I} < \tau^x_x] - \frac{v}{\bar{T}}\mathbb{E}[\tau^x_x \mathbb{1}_{\{\tau^x_x < \tau^x_{I\cup y}\}}] - \frac{(vT_{I\cup y}(x))^2}{2\bar{T}^2}\mathbb{P}[\tau^x_x < \tau^x_{I\cup y}]\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)}
\end{aligned} \tag{5.76}$$

for some $0 \leq \tilde{\theta} \leq 1$. We use Corollary 1.6 to get

$$\begin{aligned}
\widetilde{G}^x_{y,I}(v) &= \frac{1 + \frac{v}{\bar{T}}\mathbb{E}\left[\tau^x_y | \tau^x_y < \tau^x_{I\cup x}\right] + \frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1 - \frac{v}{\bar{T}}\frac{\mathbb{E}[\tau^x_x \mathbb{1}_{\tau^x_x < \tau^x_{I\cup y}}]}{\mathbb{P}\left[\tau^x_{y\cup I} < \tau^x_x\right]} - \frac{(vT_{I\cup y}(x))^2\mathbb{P}[\tau^x_x < \tau^x_{I\cup y}]}{2\bar{T}^2\mathbb{P}[\tau^x_{I\cup y} < \tau^x_x]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)} \\
&= \frac{1 + \frac{v}{\bar{T}}\mathbb{E}\left[\tau^x_y | \tau^x_y < \tau^x_{I\cup x}\right] + \frac{(vT_{I\cup x}(y))^2}{2\bar{T}^2}\widehat{G}''^x_{y,I\cup x}\left(\tilde{\theta}v\frac{T_{I\cup x}(y)}{\bar{T}}\right)}{1 - v - \frac{(vT_{I\cup y}(x))^2}{2\bar{T}\mathbb{E}\left[\tau^x_x | \tau^x_x < \tau^x_{I\cup y}\right]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)}
\end{aligned} \tag{5.77}$$

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 5.7. This gives, after computations analogous to those leading to (5.68) and using the obvious lower bound on $\bar{T}$

$$\begin{aligned}
&\left|\frac{(vT_{I\cup y}(x))^2}{2\bar{T}\mathbb{E}\left[\tau^x_x | \tau^x_x < \tau^x_{I\cup y}\right]}\widehat{G}''^x_{x,I\cup y}\left(\tilde{\theta}v\frac{T_{I\cup y}(x)}{\bar{T}}\right)\right| \leq C^2N^{2k}\frac{|v|^2}{2\mathbb{E}\left[\tau^x_x | \tau^x_x < \tau^x_{I\cup y}\right]^2} \\
&\times \sum_{x'\in\mathcal{M}_N\setminus\{J\cup y\cup x\}} e^{-N[F_N(z^*(x',x))-F_N(x)+F_N(z^*(x,y))-F_N(x)]}\left(t_{I\cup y}(x',x)\right)^2
\end{aligned} \tag{5.78}$$

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 (5.78) is finally bounded by

$$\begin{aligned}
&C^2N^{2k}C'N^{-d}\frac{|v|^2}{2}|I| \sup_{x'\in\mathcal{M}_N\setminus\{I\cup y\cup x\}} e^{-N[F_N(z^*(x,y)-F_N(z^*(x',x))]} \\
&= C^2N^{2k}C'N^{-d}\frac{|v|^2}{2}|I|e^{-N\delta}
\end{aligned} \tag{5.79}$$

where δ is strictly positive.

All the other terms in (5.77) 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 $|Re(v)| < 1$. This allows to write $\widetilde{G}^x_{y,I}(v)$ in the form

$$\widetilde{G}^x_{y,I}(v) = \frac{1}{1-v} + e^{-NK_N/2}\frac{e_1(v)}{1-v} + \frac{e_2(v)}{(1-v)(1-v-e^{-NK_N/2}e_3(v))} \tag{5.80}$$

where all e_i are analytic and uniformly (in N) bounded in the domain $|Re(v)| < 1$. This concludes the proof of the theorem.$\diamondsuit\diamondsuit$

5.3. The distribution of transition times.

From Theorem 5.6 one obtains of course some information on the distribution function.

Corollary 5.8 *Under the same assumptions as in Theorem 5.6, we have:*

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

$$\mathbb{P}\left[\tau_y^x > t\bar{T}|\mathcal{F}(x,z,y)\right] \leq e^{-(1-\delta)t}/\delta \tag{5.81}$$

(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 $t \geq 0$,

$$\lim_{i\uparrow\infty} \mathbb{P}_{N_i}\left[\tau_y^x > t\bar{\tau}_{N_i}|\mathcal{F}(x,z,y)\right] = e^{-t} \tag{5.82}$$

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.$\Diamond$

With a little more work we can also complement the upper bound (5.81) by a lower bound on the distribution of the survival time in a valley.

Proposition 5.9: *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*

$$\mathbb{P}\left[\tau_I^x > t\right] \geq \begin{cases} e^{-t/(\bar{T}(1-\alpha_t))}\alpha_t^2\left(1-h(N)\right) & \text{if } t > h(N)\frac{CN^{d/2}}{N^\kappa + T_I(x)} \\ 1 - o(1) & \text{if } t \leq h(N)\frac{CN^{d/2}}{N^\kappa + T_I(x)} \end{cases} \tag{5.83}$$

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,

$$\begin{aligned} \mathbb{P}\left[\tau_I^x > t\right] &\geq \sum_{\substack{s_1,\ldots,s_n \geq 1 \\ s_1+\cdots+s_n > t}} \mathbb{P}\left[\forall_{i=1}^n X_{s_i} = x, \forall_{s \leq s_1+\cdots+s_n} X_s \notin I\right] \\ &= \sum_{\substack{s_1,\ldots,s_n \geq 1 \\ s_1+\cdots+s_n > t}} \prod_{i=1}^n \mathbb{P}\left[\tau_x^x = s_i \leq \tau_I^x\right] \\ &= \left(\mathbb{P}\left[\tau_x^x \leq \tau_I^x\right]\right)^n \sum_{\substack{s_1,\ldots,s_n \geq 1 \\ s_1+\cdots+s_n > t}} \prod_{i=1}^n \mathbb{P}\left[\tau_x^x = s_i | \tau_x^x \leq \tau_I^x\right] \end{aligned} \tag{5.84}$$

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 (5.84) can be written as

$$\mathbb{P}\left[\tau_I^x > t\right] \geq \left(\mathbb{P}\left[\tau_x^x \leq \tau_I^x\right]\right)^n \mathbb{P}\left[\sum_{i=1}^n Y_i > t\right] \tag{5.85}$$

We have good control on the first factor in (5.85). 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

$$\begin{aligned}\mathbb{P}\left[\sum_{i=1}^n Y_i > t\right] &= \mathbb{P}\left[\sum_{i=1}^n Y_i > \frac{t}{n\mathbb{E}Y_1} n\mathbb{E}Y_1\right] \\ &\geq \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{n^2(\mathbb{E}Y_1)^2}{n(n-1)(\mathbb{E}Y_1)^2 + n\mathbb{E}Y_1^2} \\ &= \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{1}{1 + \frac{1}{n}\left(\frac{\mathbb{E}Y_1^2}{(\mathbb{E}Y_1)^2} - 1\right)}\end{aligned} \tag{5.86}$$

Now using Lemma 5.7 one verifies easily that

$$\mathbb{E}Y_1^2 \leq CN^\kappa + T_I(x)e^{-NK_N} \tag{5.87}$$

So that (5.86) gives

$$\mathbb{P}\left[\sum_{i=1}^n Y_i > t\right] \geq \left(1 - \frac{t}{n\mathbb{E}Y_1}\right)^2 \frac{1}{1 + \frac{1}{n}\left(CN^\kappa + T_I(x)e^{-NK_N}\right)} \tag{5.88}$$

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

$$n(t) = \min\left(\frac{t}{\mathbb{E}Y_1(1-\alpha_t)}, \frac{1}{h(N)(CN^\kappa + T_I)}\right) \tag{5.89}$$

This yields

$$\begin{aligned}\mathbb{P}\left[\tau_I^x > t\right] &\geq \left(1 - \mathbb{P}\left[\tau_I^x \leq \tau_x^x\right]\right)^{n(t)} \alpha_t^2 \frac{1}{1 + \frac{1}{n(t)}\left(CN^\kappa + T_I(x)e^{-NK_N}\right)} \\ &\geq e^{-t\frac{\mathbb{P}[\tau_I^x \leq \tau_x^x]}{\mathbb{E}Y_1(1-\alpha_t)} - t\mathcal{O}\left(\mathbb{P}[\tau_I^x \leq \tau_x^x]^2\right)} \alpha_t^2 \left(1 - h(N)\right)\end{aligned} \tag{5.90}$$

if t is such that the $n(t)$ is given by the second term in the minimum in (5.89). This yields the first case in (5.88). 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 $t > 0$,*

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

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

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

Then

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

Now

$$\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{M}_N}\right] \leq e^{-N[F_N(y) - F_N(x)]} \qquad (6.4)$$

while

$$\mathbb{P}\left[\tau^x_{\mathcal{M}_N} > t - s\right] = \sum_{y \in \mathcal{M}_N} \mathbb{P}\left[\tau^x_y > t - s | \tau^x_y < \tau^x_{\mathcal{M}_N \setminus y}\right] \mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{M}_N \setminus y}\right] \qquad (6.5)$$

Note that by Theorem 1.7 and the exponential Markov inequality, for some $\kappa < \infty$

$$\mathbb{P}\left[\tau^x_y > t - s | \tau^x_y < \tau^x_{\mathcal{M}_N \setminus y}\right] \leq C N^\kappa e^{-(t-s) c N^{-\kappa}} \qquad (6.6)$$

while the factors $\mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{M}_N \setminus y}\right]$ are exponentially small except if $y = x$. The denominator is bounded below by Proposition 5.9. 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 $a, b \ll 1$,

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

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 $\rho > 0$,*

$$\begin{aligned}
&\mathbb{P}\left[|X_t - x| > \rho | \tau^x_{\mathcal{T}^c_{z,x}} > t, X_0 = x\right] \\
&\leq C N^\kappa \inf_{\rho' < \rho} \sup_{y \in \Gamma_N : \rho' \leq |x-y| \leq 3\rho'/2} e^{-N[F_N(y) - F_N(x)]} + C e^{-N K_N}
\end{aligned} \qquad (6.8)$$

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

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

As in the proof of Proposition 6.1, the denominator is bounded by Proposition 5.9 and is seen to be insignificant. We concentrate on the estimates of the numerator. Again we have the obvious bound

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

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

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

Now the terms in the second sum are all harmless, since by the estimates of Lemma 5.7 and the geometry of our setting, using the exponential Chebyshev inequality as if Corollary 5.8, for any $\delta > 0$,

$$\mathbb{P}\left[\tau^x_y > t - s | \tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right] \mathbb{P}\left[\tau^x_y < \tau^x_{\mathcal{T}^c_{z,x}\cup x\setminus y}\right] \leq e^{-(t-s)(1-\delta)/T_{\mathcal{T}^c_{z,x}\cup x}(y)} e^{-N[F_N(z)-F_N(x)]} \qquad (6.12)$$

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))

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

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

$$\begin{aligned}
&\mathbb{P}\left[\tau^x_x > t - s | \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right] = \sum_{i=0}^{k} \mathbb{P}\left[\tau^x_x > t - s, \tau^x_x \geq \tau^x_{y_i}, \tau^x_x < \tau^x_{L(y_{i+1})} | \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right] \\
&\leq \sum_{i=0}^{k} \min\left(\mathbb{P}\left[\tau^x_x \geq \tau^x_{y_i} | \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right], \mathbb{P}\left[\tau^x_x > t - s, \tau^x_x < \tau^x_{L(y_{i+1})} | \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right]\right) \qquad (6.14) \\
&\leq \sum_{i=0}^{k} \min\left(\mathbb{P}\left[\tau^x_x \geq \tau^x_{y_i} | \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right], \mathbb{P}\left[\tau^x_x > t - s | \tau^x_x < \tau^x_{L(y_{i+1})}, \tau^x_x < \tau^x_{\mathcal{T}^c_{z,x}}\right]\right)
\end{aligned}$$

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[23], 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 [CGOV]) 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 Corollary 5.8 (ii) 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 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 4.4.

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 this notion has only been mentioned on a heuristic level. But of course these approaches are closely related and our results could be re-interpreted in terms of spectral theory. In a forthcoming publication [BEGK] we will indeed show a very precise relation betwen the transition time studied here and the low-lying spectrum of the operators $1 - P_N$.

[23] 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.

7. An application to disordered mean field models

7.1. Generalities

Our main motivation is to study the properties of stochastic dynamics for a class of random mean field models (see [BG1] for a more extensive discussion). These are models where the Hamiltonian is a function of so called random order parameters

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

Here $\sigma \in \mathcal{S}^N$, where $\mathcal{S}$ is the single spin space, e.g. $\mathcal{S} \equiv \{-1, +1\}$, and $\xi_{M,N}$ is a random $M \times N$ matrix.

The random Hamiltonian of such a model is then

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

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.

Let $r_N(\sigma, \sigma')$ be transition rates of a Glauber dynamics reversible with respect to the measure $\mu_{\beta,N}$, i.e. for $\sigma \neq \sigma'$, $r_N(\sigma, \sigma') = \sqrt{\frac{\mu_N(\sigma')}{\mu_N(\sigma)}}g_N(\sigma, \sigma')$ 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

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

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

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

The follwoing result is a simple adaptation of classical results of Burke and Rosenblatt [BR] from 1958:

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 $x(t)$. Then, 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: 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 the following example we will apply our general results to a specific model where this issue in particular can be studied nicely.

7.2. The random field Curie Weiss model

There are basically two classical models whose Glauber dynamics can be analyzed *completely* with our methods: The Hopfield model with finitely many patterns, and the random field Curie Weiss model with discrete random field. Since the latter is much simpler to present, we chose to give a full account of it as a pedagogical example.

Here $\mathcal{S} = \{-1, 1\}$, q is the uniform distribution on this set. Its Hamiltonian is

$$H_N[\omega](\sigma) \equiv -N\frac{\left(M_N^1(\sigma)\right)^2}{2} - \sum_{i=1}^{N} \theta_i[\omega]\sigma_i \tag{7.6}$$

where

$$M_N(\sigma) \equiv \frac{1}{N}\sum_{i=1}^{N} \sigma_i \tag{7.7}$$

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, in their important paper [MP] considered convergence to equilibrium in a particularly simple case where the random field takes only the two values $\pm\epsilon$, and with further conditions on the parameters β and h that guarantee the existence of only two local minima. Note that even in this case, the spectral approach they used gave only the leading exponential asymptotics for transition times.

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 (7.6) where θ_i take values in the set

$$\mathcal{H} \equiv \{h_1, \ldots, h_{K-1}, h_K\} \tag{7.8}$$

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

$$\Lambda_k[\omega] \equiv \{i \in \Lambda : \theta_i[\omega] = h_k\} \tag{7.9}$$

We may introduce k order parameters

$$m_k[\omega](\sigma) \equiv \frac{1}{N}\sum_{i\in\Lambda_k[\omega]} \sigma_i \tag{7.10}$$

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

$$\Gamma_N[\omega] \equiv \times_{k=1}^K \left\{-\rho_{N,k}[\omega], -\rho_{N,k}[\omega] + \tfrac{2}{N}, \ldots, \rho_{N,k}[\omega] - \tfrac{2}{N}, \rho_{N,k}[\omega]\right\} \tag{7.11}$$

where

$$\rho_{N,k}[\omega] \equiv \frac{|\Lambda_k[\omega]|}{N} \tag{7.12}$$

Note that the random variables $\rho_{N,k}$ concentrate exponentially (in N) around their mean values $\mathbb{E}_h \rho_{N,k} = \Pr[\theta_i = h_k] \equiv p_k$. 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

$$H_N[\omega](\sigma) = -NE(\underline{m}[\omega](\sigma)) \tag{7.13}$$

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

$$E(x) \equiv \frac{1}{2}\left(\sum_{k=1}^K x_k\right)^2 + \sum_{k=1}^K h_k x_k \tag{7.14}$$

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,

$$\mu_{\beta,N}[\omega](\underline{m}[\omega](\sigma) = x) \equiv \mathbb{Q}_{\beta,N}[\omega](x) = \frac{1}{Z_N[\omega]} e^{\beta N E(x)} \prod_{k=1}^K 2^{-N\rho_{N,k}[\omega]} \binom{N\rho_{N,k}[\omega]}{N(\rho_{N,k}[\omega] + x_k)/2} \tag{7.15}$$

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

$$2^{-N\rho_{N,k}[\omega]} \binom{N\rho_{N,k}[\omega]}{N(\rho_{N,k}[\omega] + x_k)/2} = e^{-N\rho_{N,k}[\omega][I(x_k/\rho_{N,k}[\omega]) + J_N(x_k, \rho_{N,k}[\omega])]} \tag{7.16}$$

where

$$I(x) \equiv \frac{1+x}{2}\ln(1+x) + \frac{1-x}{2}\ln(1-x) \tag{7.17}$$

is the usual Cramèr entropy and

$$J_N(x,\rho) = -\frac{1}{\rho N}\ln\left(\frac{1-(x/\rho)^2}{4} + \frac{2x/\rho}{1-(x/\rho)^2}\right) + O\left(\tfrac{1}{(\rho N)^2}\right) + \frac{1}{N}C(\rho N) \tag{7.18}$$

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

$$F_N[\omega](x) \equiv -\frac{1}{\beta N}\ln \mathbb{Q}_{\beta,N}[\omega](x) = F_{0,N}[\omega](x) + F_{1,N}[\omega](x) + C_N \tag{7.19}$$

with

$$F_{0,N}(x) = -E(x) + \frac{1}{\beta}\sum_{k=1}^K \rho_{N,k} I(x_k/\rho_{N,k}) \tag{7.20}$$

$C_N = \beta^{-1}\sum_{k=1}^K \rho_{N,k} C(\rho_{N,k} N)$ is constant and of order $\ln N/N$, 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

$$F_0(x) = -E(x) + \frac{1}{\beta}\sum_{k=1}^K p_k I(x_k/p_k) \tag{7.21}$$

uniformly on compact subsets of Γ. The dominant contribution to the finite volume corrections thus comes from the fluctuation 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.

7.3 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

$$0 = \frac{\partial}{\partial x_k} F_0(x) = -\sum_{\ell=0}^{K} x_\ell - h_k x_k + \frac{1}{\beta} p_k I'(x_k/p_k), \quad k = 1, \ldots, K \tag{7.22}$$

or equivalently

$$\begin{aligned} x_k &= p_k \tanh(\beta(m + h_k)), \quad k = 1, \ldots, K \\ m &= \sum_{k=1}^{K} x_k \end{aligned} \tag{7.23}$$

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

$$m = \sum_{k=1}^{K} p_k \tanh(\beta(m + h_k)) = \mathbb{E}_h \tanh \beta(m + h) \tag{7.24}$$

Thus if $m^{(1)}, \ldots, m^{(r)}$ are the solutions of (7.24), then the full set of solutions of the equations (7.22) is given by the vectors $x^{(1)}, \ldots, x^{(r)}$ defined by

$$x_k^{(\ell)} \equiv p_k \tanh \beta(m^{(\ell)} + h_k) \tag{7.25}$$

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

$$\frac{\partial^2}{\partial x_k \partial x_{k'}} F_0(x) = -1 + \frac{\delta_{k,k'}}{\beta p_k \, (1 - x_k^2/p_k^2)} \tag{7.26}$$

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

$$\frac{\partial^2}{\partial x_k \partial x_{k'}} F_0\left(x^{(\ell)}\right) = -1 + \delta_{k,k'} \lambda_k(m^{(\ell)}) \tag{7.27}$$

where

$$\lambda_k(m) \equiv \frac{1}{\beta p_k (1 - \tanh^2(\beta(m + h_k)))} \tag{7.28}$$

Lemma 7.2: *The Hessian of F_0 at $(x^{(\ell)})$ has at most one negative eigenvalue. A negative eigenvalue exists if and only if*

$$\beta \mathbb{E}_h \left(1 - \tanh^2(\beta(x^{(\ell)} + h))\right) > 1 \tag{7.29}$$

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

$$-\left(\sum_{k=1}^{K} u_k\right) + (\lambda_k - \gamma) u_k = 0 \tag{7.30}$$

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 (7.30) 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 $\gamma \notin \{\zeta_1, \ldots, \zeta_L\}$. Then (7.30) can be rewritten as

$$u_k = \frac{\sum_{k=1}^K u_k}{\lambda_k - \gamma} \tag{7.31}$$

Summing equation (7.31) over k, we get

$$\sum_{k=1}^K u_k = \sum_{k=1}^K u_k \sum_{k=1}^K \frac{1}{\lambda_k - \gamma} \tag{7.32}$$

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

$$1 = \sum_{k=1}^K \frac{1}{\lambda_k - \gamma} = \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\zeta_k - \gamma} \tag{7.33}$$

Inspecting the right-hand side of (7.33) one sees immediately that this equation has precisely L solutions γ_i that satisfy

$$\gamma_1 < \zeta_1 < \gamma_2 < \zeta_2 < \gamma_3 < \cdots < \gamma_L < \zeta_L \tag{7.34}$$

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

$$1 = \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\lambda_k - \gamma} < \sum_{\ell=1}^L \frac{|\kappa_\ell|}{\lambda_k} = \sum_{k=1}^K \frac{1}{\lambda_k} \tag{7.35}$$

which upon inserting the specific form of λ_κ yields (7.29). On the other hand, if $\sum_{k=1}^K \frac{1}{\lambda_k} > 1$, then by monotonicity there exists a negative solution to (7.33). 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.

7.4 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 7.3: *Let $g_k, k = 1, \ldots, K$ 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*

$$\frac{1}{\beta}\sum_{k=1}^{K} g_k \left(\frac{x_k}{p_k^2} I'(x_k/p_k) - I(x_k/p_k) \right) \tag{7.36}$$

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

$$\begin{aligned} F_{0,N}(x) - F_0(x) &= \frac{1}{\beta}\sum_k (\rho_{N,k} I(x_k/\rho_{N,k}) - p_k I(x_k/p_k)) \\ &= \frac{1}{\beta}\sum_k \Big((\rho_{N,k} - p_k) I(x_k/p_k) + p_k (I(x_k/\rho_{N,k}) - I(x_k/p_k)) \\ &\quad + (\rho_{N,k} - p_k)(I(x_k/\rho_{N,k}) - I(x_k/p_k)) \Big) \end{aligned} \tag{7.37}$$

Now in the interior of Γ we may develop

$$I(x_k/\rho_{N,k}) - I(x_k/p_k) = (\rho_{N,k} - p_k)\frac{1}{p_k^2} I'(x_k/p_k) + O((\rho_{N,k} - p_k)^2) \tag{7.38}$$

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,

$$\Pr\left[|\rho_{N,k} - p_k| > \epsilon\right] \leq 2\exp\left(-N I_{p_k}(\epsilon)\right) \tag{7.39}$$

where $I_p(\epsilon) \geq 0$ 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. $(\rho_{N,k} - p_k)^2 \leq N^{-3/4}$ 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. $\diamondsuit$

7.5 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 order is computed, using standard techniques, as

$$G_{0,N}(m) = -\frac{m^2}{2} + \sup_{t\in\mathbb{R}} \left(mt - \frac{1}{\beta}\sum_{k=1}^{K} \rho_{N,k} \ln\cosh\beta(h_k + t) \right) \tag{7.40}$$

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

Lemma 7.3: *The functions $G_{0,N}$ and $F_{0,N}$ are related in the following ways.*

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

$$G_{0,N}(m) = \inf_{x \in \mathbb{R}^K : \sum_k x_k = m} F_{0,N}(x) \tag{7.41}$$

(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 7 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 (7.17), (7.18), 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).

[BEGK] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, "Metastability and low lying spectra in reversible Markov chains", in preparation.

[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 related to disordered mean field models", WIAS-preprint 487 (1999).

[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).

[DS] P.G. Doyle and J.L. Snell, "Random walks and electrical networks", Carus Mathematical Monographs, 22. Mathematical Association of America, Washington, DC, 1984

[EK] M. Eckhoff and M. Klein, "Long time behavior of one-dimensional stochastic dynamics", SFB288 - preprint.

[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).

[dPdH] P. Dai Pra and F. den Hollander, "McKean-Vlasov limit for interacting random processes in random media", J. Statist. Phys. **84**, 735-772 (1996).

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

[HS] B. Helffer and J. Sjöstrand, "Multiple wells in the semiclassical limit. III. Interaction through nonresonant wells", Math. Nachr. **124**, 263-313 (1985).

[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).

[K] Ch. Külske, "Metastates in disordered mean-field models: random field and Hopfield models", J. Statist. Phys. **88**, 1257-1293 (1997).

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

[Ki2] 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).

[NS] C.M. Newman, D.L. Stein, "Thermodynamic chaos and the structure of short-range spin glasses", in *Mathematical aspects of spin glasses and neural networks*, A. Bovier and P. Picco (eds.), Progress in Probability 41, Birkhäuser, Boston-Basel-Berlin, 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).

[Sc] E. Scoppola, "Renormalization group for Markov chains and application to metastability", J. Statist. Phys. **73**, 83-121 (1993).

[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).