A RANDOM ENERGY MODEL FOR SIZE DEPENDENCE: RECURRENCE VS. TRANSIENCE*

Christof Külske[1 2]

Courant Institute
251 Mercer Street
New York, NY 10012-1185, USA

Abstract: We investigate the size dependence of disordered mean field models having an infinite number of Gibbs measures in the framework of a simplified 'random energy model for size dependence'. We introduce two versions (involving either independent random walks or branching processes), that can be seen as generalizations of Derrida's random energy model. Our model is shown to exhibit a recurrence/transience transition for Gibbs measures (meaning: a.s. existence/nonexistence of subsequences of volumes converging to a given random infinite volume state), depending on the growthrate of a function M_N describing the 'number of extremal Gibbs states that can be observed in a finite volume N'. We investigate the model in detail in the 'critical regime' $M_N \sim (\log N)^p$ (with critical point $p=1$). We obtain the a.s. large volume asymptotics of the relative weights for finding a particular state and we compute the set of a.s. cluster points of the corresponding occupation times (corresponding to the 'empirical metastate'). In the course of the proof we obtain Laws of the Iterated Logarithm for the pair $X_N^\mu, Y_N^\nu := \sup_{\mu=1,2,\dots,M_N;\mu\neq\nu} X_N^\mu$ where $(X_N^\mu)_{\mu=1,\dots,M_N;N=1,2,\dots}$ is either a branching random walk or a collection of random walks, independent over μ.

Key Words: Disordered Systems, Size Dependence, Metastates, Random Walks, Branching Processes, Law of Iterated Logarithm, Large Deviations

I. Introduction

In a preceding paper [K2] some simple examples of 'chaotic' size dependence of disordered mean field spin models having a finite number of pure phases were given. For these the set

---

* Work supported by the DFG

1 e-mail: kuelske@wias-berlin.de, kuelske@cims.nyu.edu

2 After Sept. 97: WIAS, Mohrenstrasse 39, D-10117 Berlin, Germany

of cluster points of finite volume Gibbs measures and the limiting behavior of their metastates was described. Chaotic size dependence of the Gibbs measures can arise in spin systems in the nonuniqueness regime if the boundary conditions do not preselect one of the pure phases. The empirical metastate is the average $N^{-1}\sum_{n=1}^{N}\delta_{\mu_{\Lambda_n}(\eta)}$ of random Gibbs measures $\mu_{\Lambda_n}(\eta)$, along a subsequence of volumes Λ_n. Here η denotes a frozen realization of the underlying random variables describing the disorder. The interesting concept of metastates was recently introduced by Newman and Stein to study chaotic size dependence for disordered lattice spin systems having multiple Gibbs states. A main motivation for this is the study of spin glass models. (See [NS1]-[NS4],[N],[AW]).

Our present paper is motivated by the study of systems having an *infinite* number of pure phases. Of course a rigorous analysis of realistic models is difficult. Our more modest aim is here to propose and study in detail a simple *random energy model for size dependence* (REMSD). It is a natural ansatz, analogous in spirit to Derrida's random energy model (REM) for the energy levels in a spinglass ([D]). To connect to concrete lattice systems we will however sketch below two examples whose size dependence should be well approximated by the REMSD. An interesting phenomenon we can exhibit explicitly in our model is the transition between *recurrent* and *transient* size dependence. By transient size dependence we mean that, for any given infinite volume Gibbs measure there are only finitely many volumes Λ_n for which the system is 'close' to it. While one might naively expect that in a most disordered system recurrent size dependence should occur we will see that our model makes a case for transience occurring in a 'more natural way.' (It occurs in our model in a 'large regime of parameters', to be explained below.) [1]

To introduce our model let us imagine that, for large volumes Λ_N, the state of the system $\mu_{\Lambda_N}(\eta)$ we are dealing with will be close to a mixture

$$\mu_{\Lambda_N}(\eta) \approx \sum_{\nu} p_N^{\nu}(\eta)\mu_{\infty}^{\nu}(\eta) \tag{1.1}$$

of random infinite volume Gibbs measures $\mu_{\infty}^{\nu}(\eta)$. Physically speaking the weight $p_N^{\nu}(\eta)$ of the ν'th phase should be thought of as the normalized exponential of the finite volume free energy in the state ν. In the mean field examples of [K2] their form could already be guessed from the ground state energy differences. Recall as the simplest example for this the Curie Weiss Random Field Ising model whose Hamiltonian is $H_N = -\frac{1}{2N}\sum_{1\le i,j\le N}\sigma_i\sigma_j - \epsilon\sum_{1\le i\le N}\eta_i\sigma_i$. Here $\sigma_i\in\{-1,1\}$ are Ising variables and $\eta_i=\pm 1$ (with equal probability) are the frozen random fields at site i, independent over i. Then the energy difference of the two ground states $\sigma_i\equiv 1$ and $\sigma_i\equiv -1$

[1] *We remark that in a real spin system these issues can depend crucially on the choice of the topology, e.g. whether only local observables or, say, overlaps involving all spins in a volume whose size goes to infinity are considered.*

is $H_N(\sigma_i \equiv 1) - H_N(\sigma_i \equiv -1) = -2\epsilon \sum_{1\le i\le N} \eta_i$. In fact, also for large inverse temperatures β and small ϵ there are distinct plus resp. minus infinite volume Gibbs states and the weights of finding them in the finite volume are given by $p_N^{\pm}(\eta) = \exp\left(\pm const\,(\beta) \sum_{1\le i\le N} \eta_i\right) / Norm.$.

Now, for a given system having an infinite number of extremal Gibbs states the collection of free energies $(X_N^{\nu}(\eta))_{\nu,N}$ of the ν'th state in the volume N will form a complicated stochastic process w.r.t. the volume label N (which plays the role of the 'time'). The analysis then breaks into the two parts 1) to rigorously justify the approximation (1.1) in some sense and 2) study the behavior of the resulting random sequences of measures $p_N^{\nu}(\eta)$.

To get a more concrete idea of possible behavior in this complicated situation it seems therefor natural as a first step to look at an Ansatz for the form of the weights. Then we will be left with the probabilistic study of their behavior.

To make this ansatz, let us assume that in the volume labeled by N the system is in a superposition of only the 'first' M_N states $\nu = 1, \dots, M_N$. M_N is thus the maximal number of states that can be 'seen' by a system of size N. It will be treated as a parameter of our model. Next we will make the ansatz that the distribution of $(X_N^{\nu})_{\nu=1,2,\dots,M_N;N}$ has a simple form. We will consider two variations for this and make the following

DEFINITION:

Independent walk REMSD: *Let $(M_N)_{N=1,2,\dots}$ be an increasing sequence of natural numbers. Let $(X_N^{\nu})_{\nu\in I\!N, N\in I\!N}$ be Gaussian random walks in N, independent over ν, i.e.*

$$X_N^{\nu} = \sum_{i=1}^{N} \eta_i^{\nu} \tag{1.2}$$

where $(\eta_i^{\nu})_{\nu\in I\!N, i\in I\!N}$ are i.i.d. Normal variables. Define, for $\nu = 1, \dots, M_N$, the weights

$$q_N^{\nu} := \frac{e^{\lambda X_N^{\nu}}}{\sum_{\mu=1,\dots,M_N} e^{\lambda X_N^{\mu}}} \tag{1.3}$$

where $\lambda > 0$ is a constant. So, for each fixed N, $q_N := (q_N^{\nu})_{\nu=1,2,\dots,M_N}$ is a random probability distribution supported on $\{1, 2, \dots, M_N\} \subset I\!N$.

Branching REMSD: *$(X_t^{\mu})_{\mu=1,\dots,M_t}$ is a branching Brownian motion with (nonconstant) branching rate $\gamma(t)$ starting with one particle at the origin at time $t = 0$. Thus, each of the M_t particles at time t at the positions $X_t^1, \dots, X_t^{M_t}$ moves independently according to a Brownian motion. A particle gives birth to a new particle at the same spatial position in the time interval $[t, t+dt]$ with probability $\gamma(t)dt$. The particles are labeled by the indices $\nu \in I\!N$ according to* *their order of temporal appearance. Again, for integer times $t = N$, weights q_N^ν are defined by the formula (1.3) where M_N is the (random) number of particles at time $t = N$.*

Before we comment on the relations of this model to other known models in statistical mechanics, let us exhibit two entirely different types of lattice systems whose size dependence is expected to be well-described by the independent walk REMSD, as we will heuristically explain now. Our first example may seem a little artificial but we believe that it provides a useful illustration. It is the

(i) Disordered Solid-on-solid (SOS) interface model in dimensions $d+1 \geq 3+1$, at large inverse temperatures β, small disorder, in the presence of a volume dependent ceiling M_N

Let us define finite volume Gibbs measures with periodic boundary conditions on a torus $T_N = \{-N, -N+1, \dots, N\}^d$ by the formula

$$\mu_N^{per}(\eta)[(h_x)_{x\in\Lambda_N}] = \frac{1}{Norm.} \exp\left(\beta \sum_{\substack{x,y\in T_N \\ x,y \text{near.neighb.}}} |h_x - h_y| + \beta\epsilon \sum_{x\in T_N, h\in \mathbb{Z}} \eta_{x,h} 1_{h_x=h}\right) 1_{|h_z|\leq M_N \ \forall z\in T_N} \tag{1.4}$$

Here $(h_x)_{x\in T_N} \in \mathbb{Z}^{T_N}$ is an integer valued height configuration on the torus. The $\eta_{x,h}$ are Normal random variables, i.i.d. over x and h, and $\epsilon > 0$ is assumed to be sufficiently small.

It was proved in [BoK],[K1] that under the above conditions on dimension, and smallness of temperature and disorder, the Solid-on-solid model has random infinite volume Gibbs measures μ_∞^h, for all $h \in \mathbb{Z}$ (for almost all η). μ_∞^h describes essentially flat interfaces at height h. (A μ_∞^h is obtained as infinite volume limit of finite volume Gibbs measures $\mu_{\Lambda_N}^h$, the latter being the finite volume Gibbs measure in a box Λ_N obtained by imposing a boundary condition $h_x \equiv h$ on the boundary of Λ_N.) [2]

Now, we expect the finite volume Gibbs measures with periodic boundary conditions to be approximately in a superposition of the μ_∞^h, with $|h| \leq M_N$. Note that also states describing interfaces composed of plateaus in different heights are supposed to exist in high enough dimensions (generalizing Dobrushin states in the Ising model to the SOS-model.) However, these would carry an energetic disadvantage of the order L^{d-1} that should suppress them. So we argue that

$$\mu_N(\eta)[(h_x)_{x\in\Lambda_N}] \approx \sum_{h,|h|\leq M_N} p_N^h(\eta)\mu_\infty^h(\eta) \tag{1.5}$$

[2] Note that, in view of the SOS model being a simplified model for interface states in the Ising model, the above model corresponds in fact to an Ising model with antiperiodic boundary condition in one dimension and periodic boundary conditions in the other ones.

with weights $p_N^h(\eta)$ whose form is approximately given by the energetic contributions of the flat interfaces $h_x \equiv h$, i.e.

$$p_N^h(\eta) \approx\approx \frac{\exp\left(\beta \sum_{x\in\Lambda_N, h\in\mathbb{Z}} \eta_{x,h}\right)}{\sum_{|h'|\leq M_N} \exp\left(\beta \sum_{x\in\Lambda_N, h\in\mathbb{Z}} \eta_{x,h}\right)} \tag{1.6}$$

This form obviously resembles (1.3) (taken along a suitable subsequence of volumes.) It is however a strong simplification and of course a precise analysis would be needed to rigorously justify (some modification of) (1.6). [3]

Of course the same size dependence problem could (much easier) be studied in mean field versions of the above and similar layered lattice models.

Our second example which has more 'flavor' of a spinglass is the

(ii) Hopfield model with small external field h in the region of perfect memory

Let σ_i, $i = 1, \dots, N$ be Ising spins as above. Let $\xi = (\xi_i^\mu)_{i\in\mathbb{N}, \mu=1,\dots,M_N}$ denote i.i.d. Bernoulli variables taking the values $1, -1$ with probability $\frac{1}{2}$. In the neural network interpretation $\xi^\nu = (\xi_i^\nu)_{i\in\mathbb{N}}$ are the *patterns* the neural net with neurons $(\sigma_i)_{i=1,\dots,N}$ is supposed to learn ([Ho]). For $\beta > 0$ define the finite volume Gibbs measures

$$\begin{aligned}\mu_N(\xi)[(\sigma_i)_{i=1,\dots,N}] &:= \frac{1}{Norm.} \exp\left(-\beta H_N(\xi)[(\sigma_i)_{i=1,\dots,N}]\right)\\ H_N(\xi)[(\sigma_i)_{i=1,\dots,N}] &= -\frac{1}{2N} \sum_{1\leq i,j\leq N} \sum_{1\leq\nu\leq M_N} \xi_i^\nu \xi_j^\nu \sigma_i\sigma_j - \epsilon \sum_{i=1}^N \sigma_i\end{aligned} \tag{1.7}$$

A rigorous study of the model based on much of the techniques in [BG] for the standard Hopfield model will be published in [GK]. Let us however sketch some of its properties. Note that this model is equivalent to a model where the ϵ-term in the energy function is replaced by $\epsilon \sum_{i=1}^N \eta_i\sigma_i$. Here η_i are again *random* fields taking the values $1, -1$ with probability $\frac{1}{2}$, independent over i and independent of the patterns. This equivalence is clear from the trivial gauge transformation $\sigma_i \mapsto \eta_i\sigma_i$ since the distribution of $(\eta_i\xi_i^\mu)_{i\in\mathbb{N},\mu=1,\dots,M_N}$ is the same as that of $(\xi_i^\mu)_{i\in\mathbb{N},\mu=1,\dots,M_N}$. So, alternatively the above model can be thought of as a Hopfield model with 'randomly biased' neurons (which might seem a little more natural). We also see in this way that the ϵ-term in the

[3] Indeed, let us denote by $E_{\Lambda_N}^h$ the groundstate energies for flat boundary condition at height h. Then the $E_{\Lambda_N}^h$ for nearby heights h will be dependent random variables due to local fluctuations of the interfaces. In fact, the corresponding minimizing interfaces will glue together at some small fraction of sites in Λ_N. However, we can expect independence to hold in good approximation for large distances in the height.

energy function plays a role analogous to the random field term in the Curie Weiss Random field Ising model.

The analysis in [GK] shows that, β large enough, ϵ small enough, and M_N growing sublinearly in N (sufficiently slowly) the system is in a region of perfect memory and the form of the weights is given to leading order by a nice expression. More precisely, a system with N spins is in good approximation in a mixture of $2M_N$ 'pure' states $\mu_\infty^{s\nu}(\xi)$, $s = \pm 1$ whose typical spin configuration looks like $s\xi^\nu$ up to small fluctuations. We have

$$\begin{aligned} \mu_N(\xi) &\approx \sum_{\nu=1,\dots,M_N, s=\pm 1} p_N^{s\nu}(\xi)\mu_\infty^{s\nu}(\xi) \\ p_N^{s\nu}(\xi) &= \frac{\exp\left(Const\,(\epsilon,\beta)s\sum_{i=1}^N \xi^\nu\right)}{Norm.} \end{aligned} \tag{1.8}$$

Note that, up to the s-dependence and the Bernoulli nature of the random variables, the weights take the form of the independent walk REMSD. Here, in contrast to the standard Hopfield model, the magnetic field $\epsilon > 0$ breaks the spinflip symmetry.

The form of (1.8) can again heuristically be understood by considering energy differences of the configurations $s\xi^\nu$. This gives, to *leading order*,

$$H_N(\xi)[s\xi^\nu] - H_N(\xi)[s'\xi^{\nu'}] \approx -\epsilon \sum_{i=1}^N \left(s\xi^\nu - s'\xi^{\nu'}\right) \tag{1.9}$$

There are however interesting differences whether the model is looked at on the level of overlaps that involve all spins in the volume or *local* observables. This point is discussed thoroughly in [GK] where also a construction of the conditioned *and* the empirical metastates in the *local topology* will be given. For a comprehensive review of the standard Hopfield model we refer the reader to [BG2], [T], and the very recent [BG3]. In the last reference also the conditioned one-pattern metastate in the standard Hopfield model was constructed in the regime of parameters where the 'replica symmetric solution' holds.

Let us now comment on the properties of the REMSD, its relations with other models and discuss our results. Let us denote

$$q_N^* := \max_{\nu=1,\dots,M_N} q_N^\nu = \frac{e^{\lambda X_N^*}}{\sum_{\mu=1,\dots,M_N} e^{\lambda X_N^\mu}} \tag{1.10}$$

with

$$X_N^* := X_N^{\nu_N} := \max_{\nu=1,\dots,M_N} X_N^\nu \tag{1.11}$$

where ν_N is the maximizing index at 'time' N. We will look at first at the distribution of q_N for fixed N, ignoring the dependence structure for different N.

We note that, for the special choice $M_N = 2^N$ and independent walks, q_N is precisely the Boltzmann distribution in Derrida's random energy model (REM). There our parameter λ is just called the inverse temperature. The REMSD seems thus a natural extension of the REM also from a purely probabilistic point of view.[4] Now, however, the main emphasis lies in the study of its *path properties* along the sequence $N = 1, 2, \ldots$.

The standard REM ([D], [DeE]) has been studied in great detail. It is well known that it has a phase transition at $\lambda = \lambda_c = \sqrt{2\log 2}$. The limit of what plays the role of the 'free energy' in the spinglass interpretation of the REM is, almost surely,

$$\begin{aligned}\lim_{N\uparrow\infty} \frac{1}{\lambda N} \sum_{\mu=1}^{2^N} \exp\left(\lambda\sqrt{N}G^\mu\right) &= \frac{\lambda}{2} + \frac{\lambda_c^2}{2\lambda} \quad \text{for} \quad 0 \le \lambda \le \lambda_c \\ &= \lambda_c \quad \text{for} \quad \lambda \ge \lambda_c\end{aligned} \tag{1.12}$$

for a sequence of i.i.d. Normal Gaussians $G^1, G^2, \ldots$ (see [OP]). So, for $\lambda \ge \lambda_c$ the behavior of the system is dominated by the maximum of the Boltzmann weights, the free energy equals to leading order the dominating energy level (it 'freezes'). For $\lambda \le \lambda_c$ the free energy equals the annealed free energy on the volume order. For a thorough analysis of the fluctuations of the free energy and the Poisson limit of the energy levels of the REM (and the GREM a generalization of it, leading to Ruelle's probability cascades) see [GMP].

In [DeSp] a branching process version for the REM with *constant* branching rate has been introduced. There the free energy is related to the (known) behavior of traveling waves solutions of the KPP equation. The qualitative behavior of this system then turns out to be the same as in the independent case. Actually, our branching REMSD generalizes this model to nonconstant branching rate.

Now, as should be clear from our examples, we will however be mostly interested in functions M_N that grow slower (subexponentially) with N. This then corresponds to studying Derrida's REM at a size dependent inverse temperature $\lambda_N^{REM} = \sqrt{\log 2}\sqrt{\frac{N}{\log M_N}} \uparrow \infty$. The system is thus driven to the small temperature region, where we expect the distribution q_N to be well supported on the maximizer ν_N, so that $q_N^{\nu_N} \approx 1$.

To see *how* well, our first result is about the precise large deviation asymptotics. For this point we restrict our attention to the independent walk REMSD. We have

Theorem 1: *Assume that only* $\frac{N}{\log M_N} \to \infty$ *with* $N \uparrow \infty$ *(possibly arbitrarily slowly) and* $\frac{\log M_N}{\log M_{N+1}} \to 1$.

[4] Related 'time dependent' abstract random energy type models have also very recently been studied in [BS] to understand the physicist's 'cavity method' to analyze the Sherrington Kirkpatrick spin-glass.

Then, for all $\frac{1}{2} < x < 1$,

$$\lim_{N\uparrow\infty} \frac{-\log I\!P\left[q_N^* \leq x\right]}{\log \sqrt{\frac{N}{\log M_N}}} = 1 \tag{1.13}$$

Note that the result is independent of x. Under no further restriction on the growth of M_N it depends however strongly on the special form of the Gaussian distribution. Our proof (that will be given in Chapter II) involves conditioning on the maximum X_N^* and is therefore not an immediate consequence of the results in [GMP]. (Probabilistically it is in fact a conditional large deviation result for heavy tailed random variables.)

Let us thus for the sake of the discussion assume that $\frac{N}{\log M_N} \to \infty$ with $N \uparrow \infty$. Having seen that, with large probability, the measure q_N is a perturbation of the random Dirac measure δ_{ν_N}, we are now looking at the path behavior w.r.t. the time N. We make the following natural

Definition 1: *We say that the REMSD is* **recurrent** *if, for any index* $\nu \in I\!N$,

$$\limsup_{N\uparrow\infty} q_N^\nu = 1 \tag{1.14}$$

We say that the REMSD is **transient** *if, for any* $\nu \in I\!N$,

$$\limsup_{N\uparrow\infty} q_N^\nu = 0 \tag{1.15}$$

We will state our theorems for the independent *and* the branching REMSD at the same time. Then, our first main result reads

Theorem 2: *Assume that* $\limsup_{N\uparrow\infty} \frac{\log M_N}{\log\log N} < 1$ *for the independent walk REMSD or* $\limsup_{t\uparrow\infty} \frac{\int_0^t \gamma(t')dt'}{\log\log t} < 1$ *for the branching REMSD. Then the model is recurrent.*

Assume that $\liminf_{N\uparrow\infty} \frac{\log M_N}{\log\log N} > 1$ *or* $\limsup_{t\uparrow\infty} \frac{\int_0^t \gamma(t')dt'}{\log\log t} > 1$. *Then the model is transient.*

To understand the condition for the branching case note that $\int_0^t \gamma(t')dt' = \log I\!EM_t$. We thus see that already a very mild growthrate of the function M_N makes the model transient. We will also call the first regime the 'supercritical' and the second the 'subcritical' regime of growth of M_N. Next, we investigate the models more precisely around the regime of the critical growthrate $M_N \sim \log N$.

We obtain the following precise a.s. statement (proved in Chapter III) that holds for the independent *and* the branching REMSD without change.

Theorem 2': *Assume that* $\lim_{N\uparrow\infty} \frac{\log M_N}{\log\log N} = p$ *for independent walks (or* $\lim_{t\uparrow\infty} \frac{\int_0^t \gamma(t')dt'}{\log\log t} = p$ *for branching walks) with* $p > 0$.

(i) For $p > 1$ *(transience regime) we have, a.s.,*

$$\limsup_{N\uparrow\infty} \frac{\log q_N^\mu}{l_N} = -\lambda(\sqrt{p} - 1) \tag{1.16}$$

and

$$\liminf_{N\uparrow\infty} \frac{\log q_N^\mu}{l_N} = -\lambda(\sqrt{p} + 1) \tag{1.17}$$

(ii) For $p < 1$ *(recurrence regime) we have, a.s.,*

$$\liminf_{N\uparrow\infty} \frac{\log (1 - q_N^\mu)}{l_N} = -\lambda(1 - \sqrt{p}) \tag{1.18}$$

and

$$\liminf_{N\uparrow\infty} \frac{\log q_N^\mu}{l_N} = -\lambda\sqrt{2(p+1)} \tag{1.19}$$

Note not only the jump expressed in (1.16) vs. (1.18) but also the continuous (but nondifferentiable) transition for the $\liminf$, on the right hand sides of (1.17) and (1.19) as a function of p; this might recall a *second order phase transition.*

The statements of Theorem 2' rely on laws of the iterated logarithm for the difference of X_N^ν and $Y_N^\nu := \sup_{\substack{\mu=1,2,\dots,M_N \\ \mu\neq\nu}} X_N^\mu$ that we prove for growthrates of M_N as in the hypothesis of Theorem 2'. They are interesting in their own right and therefor stated as Theorem 4 in Chapter III. For the properties of the limiting distributions of the maxima of i.i.d. random variables we refer the reader to the classical references [LLR], [Be], for branching processes see e.g. [Du], [AN], [AH])

Now, Theorem 2,2' is a statement about the recurrence v.s. transience of the sequence given by the maximizers $(\nu_N)_{N=1,2,\dots}$. To understand better its behavior it is now interesting to investigate also the *empirical distribution* of ν_N. Obviously it is the analogue of the empirical metastate construction. (Care is however in order to relate the results back to spin models when using *local* observables, see [GK]). So, let us define the empirical measure $n_N = (n_N^\nu)_{\nu\in\mathbb{N}}$ given by

$$n_N^\nu := \frac{1}{N}\#\{n \in \{1,2,\dots,N\}, \nu_n = \nu\} \tag{1.20}$$

Thus $(n_N)_{N=1,2,\dots}$ is a sequence of random measures, with the n_N's supported on the growing sets $\{1,2,\dots,M_N\}$.

Our final results deal specifically with the asymptotic almost sure behavior of the fraction of times n_N^ν in one fixed given state ν. Note that of course $n_N^\nu \to 0$ stochastically, since for fixed $\epsilon > 0$,

$$\mathbb{E}[n_N^\nu] = \frac{1}{N} \sum_{1 \le n \le N} \frac{1}{M_n} \to 0 \tag{1.21}$$

for any function $M_N \uparrow \infty$.

Now, we have trivially that $n_N^\nu \to 0$ a.s. in the supercritical regime. In the subcritical regime, however the behavior is more interesting. It turns out that the set of cluster points of n_N^ν is a nontrivial interval that can be computed explicitly. We have

Theorem 3: *Assume that* $\lim_{N\uparrow\infty} \frac{\log M_N}{\log\log N} = p$ *for independent walks (or* $\lim_{t\uparrow\infty} \frac{\int_0^t \gamma(t')dt'}{\log\log t} = p$ *for branching walks), with* $0 < p < 1$. *Then, a.s., we have for the set of cluster points*

$$\mathcal{CP}\{n_N^\nu, N = 1, 2, \ldots\} = \left[0, 1 - e^{-4\left(\frac{1}{p}-1\right)}\right] \tag{1.22}$$

The proof of Theorem 3 (given in Chapter IV) uses large deviation techniques for the sample paths of (X_N^ν, Y_N^ν). As will be clear from the proofs (that we tried to give in a detailed manner) the results of Theorems 2 - 4 should in fact apply to a large class of distributions for the increments. All that is needed are suitable upper and lower Gaussian bounds of 'moderate deviations'.

Acknowledgments:

The author thanks C.Newman for discussions about metastates and helpful comments, A.Bovier and V.Gayrard for discussions about related questions in the Standard Hopfield Model that provided a motivation for this work, and P.Picco for pointing out [DeSp]. He thanks the Courant Institute for a very pleasant stay during which this work was done. It was supported by the DFG.

II. Large deviation estimates

In this chapter (which is technically independent of the rest of the paper) we prove Theorem 1. It is convenient to express the volume label N as a function of the number of states and change notations slightly to avoid $\sqrt{N}$-factors. Assume that G_i, $i = 1, 2, \ldots$, are i.i.d. Standard Normals and denote $G_n^* = \max_{k=1,\ldots,n} G_k$. Let $\lambda_n \uparrow \infty$ denote some function of n and introduce, for $k = 1, \ldots, n$, the weights

$$r_n^k = \frac{e^{\lambda_n G_k}}{\sum_{i=1,\ldots,n} e^{\lambda_n G_i}}, \qquad r_n^* = \max_{k=1,\ldots,n} r_n^k \tag{2.1}$$

Obviously, Theorem 1 follows from the statement that for $\frac{\lambda_n}{\sqrt{\log n}} \to \infty$ with $n \uparrow \infty$, for all $\frac{1}{2} < x < 1$,

$$\lim_{n\uparrow\infty} \frac{\log I\!P\left[r_n^* \leq x\right]}{\log \frac{\lambda_n}{\sqrt{\log n}}} = -1 \tag{2.2}$$

Proof of (2.2): We denote by $I\!P^u$ the conditional distribution of the $n-1$ variables among $G_1, \dots, G_n$ that do not realize the maximum, under the condition that $G_n^* = u$. (By a little abuse of notation we write for these variables $G_1, \dots, G_{n-1}$.) It is easy to verify that

$$I\!P^u\left[G_i \leq x_i, \forall i = 1, \dots, n-1\right] = \prod_{i=1}^{n-1} \frac{\Phi(x_i)}{\Phi(u)} \tag{2.3}$$

for $x_i \leq u$, for all $i = 1, \dots, n$. $\Phi(x) = I\!P[G_1 \leq x]$ denotes the distribution function of the Normal distribution. We may then express the distribution function of r^* by

$$\begin{aligned} I\!P\left[r_n^* \leq x\right] &= \int I\!P^{G_n^*}(du) I\!P^u\left[\frac{e^{\lambda_n u}}{e^{\lambda_n u} + \sum_{i=1,\dots,n-1} e^{\lambda_n G_i}} \leq x\right] \\ &= \int I\!P^{G_n^*}(du) I\!P^u\left[\sum_{i=1,\dots,n-1} Y_i \geq y\right] \end{aligned} \tag{2.4}$$

where we have put $y = 1/x - 1$ and the variables $Y_i := e^{\lambda_n(G_i - u)}$, taking values in $[0,1]$, are i.i.d. with u-dependent distribution

$$I\!P^u\left[Y_i \leq z\right] = \frac{\Phi\left(\frac{\log z}{\lambda} + u\right)}{\Phi(u)} \tag{2.5}$$

In the sequel we will not make the λ_n and u dependence explicit in our notation.

For the estimation of (2.4) we will use that the outer integral has relevant contributions only for u on the n-dependent scale $u \approx x/a_n + b_n$ where

$$\begin{aligned} a_n &= \sqrt{2\log n} \\ b_n &= \sqrt{2\log n} - \frac{\log\log n + \log 4\pi}{2\sqrt{2\log n}} \end{aligned} \tag{2.6}$$

In fact, it is a standard result in elementary extreme value theory (see [LLR]) that

$$I\!P[a_n(G_n^* - b_n) \leq x] \to e^{-e^{-x}} \tag{2.7}$$

Thus, to prove (2.2), we will 1) peel off the tails of the outer integral and 2) control the conditional distributions $I\!P^{u_n}$, for some $u_n = x_n/a_n + b_n$, where x_n does not depend much on n. We will see for this second part that, in the case $0 < y < 1$, it is the *maximum* of the Y_i which is responsible for the large deviations of the Y_i-sum under $I\!P^u$ rather than the collective behavior of the sum.

Hence, for our estimates exponential inequalities (Cramer theory) are not useful. Rather, when viewed e.g. on the scale of the $I\!P^{u_n}$ median of Y_i (which goes to zero with n), we are dealing with sums of heavy tailed variables (see [Nag]); however with n-dependent distribution.

A useful (and not completely obvious) stochastic domination property that is implied by the Normal distribution is

Lemma 2: *For all $\lambda > 0$, $u_1 \leq u_2$, $0 \leq z \leq 1$*

$$I\!P^{u_1}\left[Y_i \leq z\right] \leq I\!P^{u_2}\left[Y_i \leq z\right] \tag{2.8}$$

Proof: We have to verify that, for fixed $a > 0$, $\Phi(u-a)/\Phi(u)$ is an increasing function in u. Taking a u-derivative it is equivalent to show that

$$\frac{p(u-a)}{p(u)} - \frac{\Phi(u-a)}{\Phi(u)} \geq 0 \tag{2.9}$$

for all $a \geq 0$ and $u \in I\!R$. Taking logarithms of this inequality we see that this follows from

$$\begin{aligned}&\log p(u-a) - \log p(u) - \log \Phi(u-a) + \log \Phi(u)\\ &= -a\frac{\partial}{\partial u}\Big|_{\tilde u}(\log p(u) - \log \Phi(u)) = a\left(\tilde u + \frac{p(\tilde u)}{\Phi(\tilde u)}\right) \geq 0\end{aligned} \tag{2.10}$$

But this inequality is clear for nonnegative $\tilde u$; for negative $\tilde u$ it is also true, which follows from the well known equality $\Phi(u) \leq p(u)/(-u)$, for *all* $u < 0$.$\diamondsuit$

Using stochastic domination from the lemma follows that the conditional probability under the outer integral in (2.4) is a decreasing function of u. Consequently

$$I\!P\left[q_n^* \leq x\right] \geq I\!P[G_n^* \leq u] I\!P^u\left[\sum_{i=1,\ldots,n-1} Y_i \geq y\right] \tag{2.11}$$

for all u. Similarly

$$I\!P\left[q_n^* \leq x\right] \leq I\!P[G_n^* \leq u] + I\!P^u\left[\sum_{i=1,\ldots,n-1} Y_i \geq y\right] \tag{2.12}$$

The last two inequalities are the starting point for our estimates; we will then have to choose $u = u_n = \frac{x}{a_n} + b_n$ in a useful way. We distinguish two different regimes of growth for λ_n: The case of slow divergence, $\lim_{n\uparrow\infty}\frac{\log\log\lambda_n}{a_n} = 0$ and the case of fast divergence, $\liminf_{n\uparrow\infty}\frac{\log\log\lambda_n}{a_n} > 0$. For easy reference we note the formula for the Gaussian density

$$np\left(\frac{x}{a_n} + b_n\right) = \sqrt{2\log n}\exp\left(-\frac{x^2}{2a_n} - \frac{xb_n}{a_n} - \frac{(\log\log n + \log 4\pi)^2}{16\log n}\right) \tag{2.13}$$

Let us begin with the first case which is more difficult. We treat the lower bound first because it explains the form of the asymptotics. We use (2.11) and choose $u = u_n = b_n$. Then, from $I\!P[G_n^* \leq u]$, $1 - \Phi(x) \sim \frac{p(x)}{x}$ and (2.13) follows that the first factor in (2.11) converges to a constant and can be ignored. Now we estimate

$$\begin{aligned} & I\!P^{u_n}\left[\sum_{i=1,\ldots,n-1} Y_i \geq y\right] \geq I\!P^{u_n}\left[\exists! i_0 \in \{1,\ldots,n-1\} : Y_{i_0} \geq y\right] \\ & = (n-1) I\!P^{u_n}\left[Y_1 \geq y\right] I\!P^{u_n}\left[Y_1 \leq y\right]^{n-2} \end{aligned} \tag{2.14}$$

Therefore we have for $1 > x > \frac{1}{2}$ from (2.14)

$$\begin{aligned} & \liminf_{n\uparrow\infty} \frac{\log I\!P\left[r_n^* \leq x\right]}{\log \frac{\lambda_n}{a_n}} \\ & \geq \liminf_{n\uparrow\infty} \frac{\log\left((n-1) I\!P^{u_n}\left[Y_1 \geq y\right]\right)}{\log \frac{\lambda_n}{a_n}} + \liminf_{n\uparrow\infty} \frac{(n-2)\log\left(1 - I\!P^{u_n}\left[Y_1 \geq y\right]\right)}{\log \frac{\lambda_n}{a_n}} \end{aligned} \tag{2.15}$$

For the large n-asymptotics of the distribution of Y_i under $I\!P^{u_n}$ we write

$$I\!P^{u_n}\left[Y_i \geq y\right] = \frac{1}{\Phi(u_n)}\left(\Phi(u_n) - \Phi\left(u_n - \frac{\log y^{-1}}{\lambda_n}\right)\right) = \frac{p(\tilde{u}_n)}{\Phi(u_n)}\frac{\log y^{-1}}{\lambda_n} \tag{2.16}$$

with $u_n - \frac{\log y^{-1}}{\lambda_n} \leq \tilde{u}_n \leq u_n$. Using that $p(\tilde{u}_n) \geq p(u_n)$ for positive arguments we have from this for the first term in (2.14)

$$\begin{aligned} & \liminf_{n\uparrow\infty} \frac{\log\left((n-1) I\!P^{u_n}\left[Y_1 \geq y\right]\right)}{\log \frac{\lambda_n}{a_n}} \\ & \geq \liminf_{n\uparrow\infty} \frac{1}{\log \frac{\lambda_n}{a_n}}\left(\log \frac{n p(u_n)}{\lambda_n} + \log\log y^{-1} - \log\Phi(u_n)\right) \end{aligned} \tag{2.17}$$

The last two terms vanish under the limit. For the first term we have from (2.13) that

$$\frac{1}{\log \frac{\lambda_n}{a_n}} \log \frac{n p(u_n)}{\lambda_n} = -1 - \frac{1}{\log \frac{\lambda_n}{a_n}} \frac{(\log\log n + \log 4\pi)^2}{16 \log n} \to -1 \tag{2.18}$$

For the second term in (2.15) we define $\tilde{x}_n$ by $\tilde{u}_n = \frac{\tilde{x}_n}{a_n} + b_n$ and note that $\tilde{x}_n \to 0$ due to $\frac{a_n}{\lambda_n} \to 0$, using the upper and lower bounds for $\tilde{u}_n$. In particular we have $p(\tilde{u}_n) \downarrow 0$ and expanding the log we have

$$\begin{aligned} & (n-2)\log\left(1 - I\!P^{u_n}\left[Y_1 \geq y\right]\right) = (n-2)\log\left(1 - \frac{p(\tilde{u}_n)}{\Phi(u_n)}\frac{\log y^{-1}}{\lambda_n}\right) \\ & \geq -const\,(n-2)\frac{p(\tilde{u}_n)}{\Phi(u_n)}\frac{\log y^{-1}}{\lambda_n} = -\frac{a_n}{\lambda_n}\mathcal{O}(1) \end{aligned} \tag{2.19}$$

where the last equality follows from (2.13). This already goes to zero with $n \uparrow \infty$ which finishes the proof of the lower bound.

For the large deviation upper bound we use the quadratic Chebychev inequality

$$\log I\!P^u \left[\sum_{i=1,\dots,n-1} Y_i \geq y \right] \leq \log \left((n-1) I\!E^u (Y_1^2) \right) - \log y^2 \tag{2.20}$$

which gives us from (2.12)

$$\limsup_{n\uparrow\infty} \frac{\log I\!P\left[r_n^* \leq x\right]}{\log \frac{\lambda_n}{a_n}} \leq \max \left\{ \limsup_{n\uparrow\infty} \frac{\log I\!P[G_n^* \leq u_n]}{\log \frac{\lambda_n}{a_n}}, \limsup_{n\uparrow\infty} \frac{\log \left(n I\!E^{u_n}(Y_1^2) \right)}{\log \frac{\lambda_n}{a_n}} \right\} \tag{2.21}$$

In contrast to the lower bound we have to be a bit more careful with the choice of the 'split point' $u_n = \frac{x_n}{a_n} + b_n$. We look for the slowest possible way to let $x_n \downarrow -\infty$ such that the first term is bounded by -1. For $u > 0$ we write

$$\log I\!P[G_n^* \leq u] = n \log \left(1 - I\!P[G_1 \geq u]\right) \leq -n I\!P[G_1 \geq u] \tag{2.22}$$

and use the well known estimate on the Gaussian tails $I\!P[G_1 \geq u] \geq \frac{p(u)}{u} \frac{1}{1+u^{-2}}$, valid for *all* $u > 0$, to get

$$\frac{\log I\!P[G_n^* \leq u_n]}{\log \frac{\lambda_n}{a_n}} \leq -\frac{np(u_n)}{u_n} \frac{1}{1+u_n^{-2}} \frac{1}{\log \frac{\lambda_n}{a_n}} \tag{2.23}$$

Below we will choose x_n s.t. $\lim_{n\uparrow\infty} \frac{x_n}{a_n} = 0$, note that from this follows that $\lim_{n\uparrow\infty} \frac{a_n}{u_n} = 1$. In particular $u_n \uparrow +\infty$ and we have from (2.13)

$$\begin{aligned} &\limsup_{n\uparrow\infty} \frac{\log I\!P[G_n^* \leq u_n]}{\log \frac{\lambda_n}{a_n}} \\ &\leq -\liminf_{n\uparrow\infty} \exp\left(-\frac{x^2}{2a_n} - \frac{xb_n}{a_n} - \frac{(\log\log n + \log 4\pi)^2}{16 \log n} \right) \frac{1}{\log \frac{\lambda_n}{a_n}} \\ &= -\liminf_{n\uparrow\infty} \exp\left(-\frac{x^2}{2a_n} - \frac{xb_n}{a_n} \right) \frac{1}{\log \frac{\lambda_n}{a_n}} \end{aligned} \tag{2.24}$$

Now it is convenient to take x_n to be the largest solution of the quadratic equation such that the r.h.s. becomes identically -1. This amounts to $x_n = -b_n \left(1 - \sqrt{1 - \frac{2a_n}{b_n^2} \log\log \frac{\lambda_n}{a_n}} \right)$.

Using the fact that we are in the region of restricted growth $\frac{\log\log \lambda_n}{a_n} \downarrow 0$, it is now easy to verify a posteriori that in fact $\frac{x_n}{a_n} \downarrow 0$. (This is the point of the proof where some restriction on the growth is needed.) Thus we are done with the estimate of the first term in (2.21).

To deal with the second term in (2.21) we must look at the second moment

$$I\!E^u \left[Y_i^2 \right] = \frac{\Phi(u - 2\lambda)}{\Phi(u)} e^{-2\lambda u + 2\lambda^2} \tag{2.25}$$

Let us compute the asymptotics for $u_n, \lambda_n \uparrow \infty$ under the assumption that $2\lambda_n - u_n \uparrow \infty$. Then we have

$$I\!E^{u_n}\left[Y_i^2\right] \le \frac{p(u_n)}{\Phi(u_n)(2\lambda_n - u_n)} \tag{2.26}$$

We are now ready to estimate

$$\limsup_{n\uparrow\infty} \frac{\log\left(nI\!E^{u_n}(Y_1^2)\right)}{\log\frac{\lambda_n}{a_n}} \le \limsup_{n\uparrow\infty}\left(\frac{\log\frac{p(u_n)}{\lambda_n}}{\log\frac{\lambda_n}{a_n}} - \frac{\log\Phi(u_n)}{\log\frac{\lambda_n}{a_n}} - \frac{\log(2-\frac{u_n}{\lambda_n})}{\log\frac{\lambda_n}{a_n}}\right) \tag{2.27}$$

The second term converges to zero since $u_n \uparrow \infty$. The third term converges to zero since $\frac{u_n}{\lambda_n} = \frac{u_n}{a_n}\frac{a_n}{\lambda_n} \downarrow 0$ since $\frac{u_n}{a_n} \to 1$ with our choice of u_n. For the first term we have

$$\frac{1}{\log\frac{\lambda_n}{a_n}}\log\frac{p(u_n)}{\lambda_n} = \frac{1}{\log\frac{\lambda_n}{a_n}}\left(\log\frac{a_n}{\lambda_n} - \log\log\frac{\lambda_n}{a_n} - \frac{(\log\log n + \log 4\pi)^2}{16\log n}\right) \tag{2.28}$$

where we have used the definition of x_n and (2.13). Obviously this converges to -1 which finishes the proof of the upper bound.

It remains to conclude the analysis with the easier case $\liminf_{n\uparrow\infty}\frac{\log\log\lambda_n}{a_n} > 0$. In particular $\lambda_n \uparrow \infty$ faster than any power of n. Since we have $\frac{\log a_n}{\log\lambda_n} \to 0$ it suffices to consider the equivalent 'rate function' $\log\lambda_n$ in the denominator of (2.2). To get the lower bound we set $u_n = 1 = Const$ and note that

$$\begin{aligned}&\liminf_{n\uparrow\infty}\frac{\log\left((n-1)I\!P^{u_n}\left[Y_1 \ge y\right]\right)}{\log\lambda_n}\\&= \liminf_{n\uparrow\infty}\frac{1}{\log\lambda_n}\left(-\log\lambda_n + \log n + \log p(\tilde{u}_n) + \log\log y^{-1} - \log\Phi(u_n)\right) = -1\end{aligned} \tag{2.29}$$

where we have used that $\frac{\log n}{\log\lambda_n} \to 0$. Secondly, as above, already the numerator of the second term in (2.18) goes to zero since $\frac{n}{\lambda_n} \to 0$. This proves the lower bound.

For the upper bound we determine the split point $u_n \downarrow -\infty$ by a good choice that makes the first limsup in (2.18) smaller that -1. Since

$$\begin{aligned}&\log I\!P[G_n^* \le u] = n\log\left(I\!P[G_1 \le u]\right)\\&\le n\log p(u) = -\frac{nu^2}{2} - n\log\sqrt{2\pi}\end{aligned} \tag{2.30}$$

the choice $u_n = -\sqrt{\frac{2\log\lambda_n}{n}}$ implies that in fact $\limsup\frac{\log I\!P[G_n^* \le u_n]}{\log\lambda_n} \le -1$. Since u_n is negative we have the estimate $I\!E^{u_n}\left[Y_i^2\right] \le \frac{|u_n|(1+u_n^{-2})}{2\lambda_n - u_n}$. Now, since $\frac{u_n}{\lambda_n} \to 0$ we have from that in fact the desired

$$\limsup_{n\uparrow\infty}\frac{\log\left(nI\!E^{u_n}(Y_1^2)\right)}{\log\frac{\lambda_n}{a_n}} = \limsup_{n\uparrow\infty}\frac{\log\left(I\!E^{u_n}(Y_1^2)\right)}{\log\lambda_n} = -1 \tag{2.31}$$

$\diamondsuit$

III. Laws of the Iterated Logarithm

In this chapter we prove Theorems 2,2' and the LIL Theorem 4 (see below). For fixed ν we write

$$Y_N := Y_N^\nu := \sup_{\substack{\mu=1,2,\ldots,M_N \\ \mu\neq\nu}} X_N^\mu \tag{3.1}$$

Bounding the walks in the definition of the weight q_N^ν for $\mu \neq \nu$ by Y_N^ν we get the simple estimate

$$\left(1+(M_N-1)\exp\lambda\left(Y_N^\nu - X_N^\nu\right)\right)^{-1} \leq q_N^\nu \leq \left(1+\exp\lambda\left(Y_N^\nu - X_N^\nu\right)\right)^{-1} \tag{3.2}$$

for fixed $\nu \leq M_N$. In particular, transience follows, if, a.s. $\liminf_{N\uparrow\infty}\left(Y_N^\nu - X_N^\nu\right) = \infty$. Recurrence follows if $\liminf_{N\uparrow\infty}\left(\lambda^{-1}\log M_N + Y_N^\nu - X_N^\nu\right) = -\infty$. It is thus natural to look at the fluctuations of (X_N^ν, Y_N^ν) in more detail which we will do in this chapter.

We treat this issue only in the interesting *transition regime* of growthrates as in the hypothesis of Theorem 2' where we consider p as a parameter. An example for a branching intensity that satisfies the hypothesis of theorem 2' is e.g. $\gamma(t) = \frac{p}{(t+e)\log(t+e)}$. The aim of this chapter is to establish a 'phase transition' w.r.t. the parameter p with critical value $p_c = 1$.

Since we will treat the independent walk case and the branching case in a unified way, some remarks are in order about the latter. From $I\!E\,[M_t] = \exp\left(\int_0^t \gamma(t')dt'\right)$ we see that the hypothesis of Theorem 2' amounts to $\lim_{t\uparrow\infty}\frac{\log I\!E M_t}{\log\log t} = p$. It is then natural to introduce the new time scale $\tau = \tau(t) = \log I\!E M_t$ on which the branching intensity is constant. Then $\tilde{M}_\tau := M_{t(\tau)}$ is just the number of 'particles' resulting from one 'particle' at time 0 from binary fission at constant rate 1 (a.k.a. Yule process). It is well known (and can be easily computed from the Kolmogorov backward equation) that the generating function satisfies

$$I\!E\left[s^{\tilde{M}_\tau}\right] = \frac{se^{-\tau}}{1-(1-e^{-\tau})s} \tag{3.3}$$

for $s < 1$ (see [AN], page 109). A well known result in the elementary theory of branching processes (that follows from the a.s. martingale convergence theorem) is that $e^{-\tau}\tilde{M}_\tau$ converges a.s. to a random variable W ([AN], page 111). (In this case W is exponentially distributed, as can be immediately computed from (3.3).) Writing $\tilde{M}_\tau(\omega) = e^\tau W(\omega) g_\tau(\omega)$ with $\lim_{\tau\uparrow\infty} g_\tau(\omega) = 1$ this implies in particular that, a.s.

$$\lim_{t\uparrow\infty}\frac{\log M_t(\omega)}{p\log\log t} = \lim_{\tau\uparrow\infty}\frac{\log\tilde{M}_\tau(\omega)}{\tau} = 1 + \lim_{\tau\uparrow\infty}\frac{\log W(\omega)g_\tau(\omega)}{\tau} = 1 \tag{3.4}$$

Thus, from the hypothesis on its expectation follows that the number of random walks satisfies the hypothesis of the theorem for the independent case also a.s.. This will be useful for the analysis.

To state our Laws of the Iterated Logarithm let us introduce the scale of the classical LIL

$$l_N := \sqrt{2N \log\log N} \tag{3.5}$$

Then the result of this chapter is

Theorem 4: *Assume that* $\lim_{N\uparrow\infty} \frac{\log M_N}{\log\log N} = p$ *for independent walks, or* $\lim_{t\uparrow\infty} \frac{\int_0^t \gamma(t')dt'}{\log\log t} = p$ *for branching walks, with* $p > 0$*. Then we have, a.s., for any fixed* ν*,*

(i) $\limsup_{N\uparrow\infty} \frac{X_N^*}{l_N} = \sqrt{p+1}$

(ii) $\liminf_{N\uparrow\infty} \frac{X_N^*}{l_N} = \sqrt{p}$

(iii) $\limsup_{N\uparrow\infty} \frac{X_N^\nu - Y_N^\nu}{l_N} = 1 - \sqrt{p}$

(iv) $\limsup_{N\uparrow\infty} \frac{Y_N^\nu - X_N^\nu}{l_N} = \sqrt{2(p+1)}$ *for* $0 \le p \le 1$

$\limsup_{N\uparrow\infty} \frac{Y_N^\nu - X_N^\nu}{l_N} = \sqrt{p} + 1$ *for* $p \ge 1$

The first two statements are given here for the sake of completeness, but they will also be needed in the course of the proof of (iii), (iv). Note that the r.h.s. of (iv) changes in a continuous but nondifferentiable way when p passes through its critical value $p = 1$. This may seen to be somewhat reminiscent of a *second order phase transition.* Note also that the trivial case of two random walks, $M \equiv 2$, is formally contained in (iv). In view of the hypothesis (a), for (i) and (ii) it makes no difference to put Y_N^ν or X_N^* (but of course for (iii) and (iv)).

The recurrence/transience transition stated in Theorem 2 is now a direct consequence of (iii) with (3.1). For $p > 1$, in fact one obtains from (iii) direct the quantitative estimate of Theorem 2' (1.16); (1.17) follows from (iv). In a similar way we get in the case $p < 1$ the estimates (1.18) from (iii) and (1.19) from (iv).

Corollary: For $p = 0$ we have, a.s., $\limsup_{N\uparrow\infty} \frac{X_N^*}{\sqrt{2N \log\log N}} = 1$.

Indeed, we have, by (i) and the trivial fact that the limsup is increasing as a function of the function $N \mapsto M_N$, the inequality $\limsup_{N\uparrow\infty} \frac{X_N^*}{l_N} \le \sqrt{p+1}$ for any (arbitrarily small) $p > 0$. This shows that $\limsup_{N\uparrow\infty} \frac{X_N^*}{l_N} \le 1$. The lower bound $\limsup_{N\uparrow\infty} \frac{X_N^*}{l_N} \ge \limsup_{N\uparrow\infty} \frac{X_N^1}{l_N} = 1$ holds trivially from the usual one dimensional LIL.

Proof: We consider first the case of independent walks. Let us start with the proof of (i) and prove the upper bound $\limsup_{N\uparrow\infty} \frac{X_N^*}{l_N} \le \sqrt{p+1}$. This explains also the skeleton of the proof used in the other cases.

We follow as close as possible the usual arguments that are used in the case of one simple random walk, $M \equiv 1$ (as presented in a more general setting e.g. in [LT] p.199.ff) As in the simple case, we will consider the process X_N^* at exponentially spaced sampling points ρ^k with $\rho > 1$. The information at these points will turn out to be essentially enough to describe the paths. The occurrence of $\sqrt{p+1}$ will then be explained by the computation of the fixed time distributions $X_{N_k}^*$.

We will show that for each $\delta > 0$ there exists a subsequence N_k s.t.

$$\sum_{k=1}^{\infty} I\!P\left[\max_{N_{k-1}<N\leq N_k} \frac{X_N^*}{\sqrt{p+1}l_N} \geq 1+\delta\right] < \infty \tag{3.6}$$

The claim follows then from Borel Cantelli. Now, to control the behavior of the paths between the sampling points, we estimate

$$\begin{aligned}
& I\!P\left[\max_{N_{k-1}<N\leq N_k} \frac{X_N^*}{\sqrt{p+1}l_N} \geq 1+\delta\right] \leq I\!P\left[\max_{N_{k-1}<N\leq N_k} X_N^* \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right] \\
&= I\!P\left[\max_{N_{k-1}<N\leq N_k} \max_{1\leq\mu\leq M_N} X_N^\mu \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right] \\
&\leq I\!P\left[\max_{1\leq\mu\leq M_{N_k}} \max_{N_{k-1}<N\leq N_k} X_N^\mu \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right] \\
&= 1-\left(1-I\!P\left[\max_{N_{k-1}<N\leq N_k} X_N^1 \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right]\right)^{M_{N_k}}
\end{aligned} \tag{3.7}$$

using the independence over μ in the last line. Now let us, for $\rho > 1$, denote by N_k the smallest N s.t. $N > \rho^k$. Then $\frac{N_{k+1}}{N_k} \sim \rho$. Note that for all $\delta > 0$ we can choose $\epsilon > 0$ (small enough) and $\rho > 1$ (small enough) s.t. $(1+\delta)l_{N_{k-1}} \geq (1+\epsilon)l_{N_k}$. Fix these ϵ and ρ. Then we have

$$I\!P\left[\max_{N_{k-1}\leq N\leq N_k} X_N^1 \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right] \leq I\!P\left[\max_{N_{k-1}\leq N\leq N_k} X_N^1 \geq (1+\epsilon)\sqrt{p+1}l_{N_k}\right] \tag{3.8}$$

Using the well known estimate $I\!P\left[\max_{1\leq N\leq L} X_N^1 \geq x\right] \leq e^{-\frac{x^2}{2L}}$ we get from this

$$\begin{aligned}
& I\!P\left[\max_{N_{k-1}<N\leq N_k} X_N^1 \geq (1+\delta)\sqrt{p+1}l_{N_{k-1}}\right] \\
& \leq (\log N_k)^{-(1+\epsilon)^2(p+1)} \leq Const\, k^{-(1+\epsilon)^2(p+1)}
\end{aligned} \tag{3.9}$$

But now we note that the hypothesis on M_N gives us that $M_{N_k} \leq k^{p(1+\epsilon')}$, for any $\epsilon' > 0$ (for k sufficiently large). Thus the r.h.s. of (3.7) is bounded by $Const\, k^{-(1+\epsilon)^2(p+1)} \times k^{p(1+\epsilon')}$ (for k sufficiently large). But since this expression is summable over k, for any $\epsilon > 0$ (if $\epsilon' > 0$ is sufficiently small), also the series (3.6) converges for any $\delta > 0$ and we are done for the proof of the upper bound of (i).

Let us now comment on the modifications that are needed to take care of the branching case. We regard from now on the branching times and the family structure as fixed, i.e. what particle gives offspring to what child at given times. Since we are dealing with a.s. results we put ourselves on the full measure set where we have for the number of walks $\lim_{N\uparrow\infty}\frac{\log M_{t=N}}{\log\log N}=p$. Let us denote by the symbols $I\!\!P'$, $I\!\!E'$ the probability resp. average *conditioned* on a particular branching history in the above set; so the expectation is really only over the Brownian paths. We will have to show that the branching does not produce enough dependence between the walks to alter the behavior of the independent case in a significant way.

The following stochastic inequality between the maxima for the independent and the branching case makes things a bit easier.

Lemma III.1: *Let U_t^μ denote a binary 'branching process' of random walks or diffusions with fixed branching history in the above sense. Assume the processes $(U_t^\mu)_{t\geq 0}$ have the same distribution for each μ, and denote by V_t^μ fresh versions of them, independent over μ. Then*

$$I\!\!P\left[\max_{\mu=1,\dots,M_t} U_t^\mu\leq u\right]\geq I\!\!P\left[\max_{\mu=1,\dots,M_t} V_t^\mu\leq u\right]=I\!\!P\left[V_t^1\leq u\right]^{M_t} \tag{3.10}$$

Proof: Induction over the number of branches M. We assume that the claim holds at any fixed time t' if $M_{t'}\leq M$. So, consider any time t where $M_t=M+1$. Denote by t_1 the first splitting time in the family tree and by A and B the set of indices of the trees coming from the two offspring particles. Then $|A|,|B|\leq M$ and $|A|+|B|=M+1$ and we have

$$\begin{aligned}
&I\!\!P\left[\max_{\mu=1,\dots,M+1} U_t^\mu\leq u\right]=I\!\!E_{U_{t_1}^1}\, I\!\!P\left[\max_{\mu\in A} U_t^\mu\leq u\middle|U_{t_1}^1\right] I\!\!P\left[\max_{\mu\in B} U_t^\mu\leq u\middle|U_{t_1}^1\right]\\
&\geq I\!\!E_{U_{t_1}^1}\, I\!\!P\left[U_t^1\leq u\middle|U_{t_1}^1\right]^{|A|} I\!\!P\left[U_t^1\leq u\middle|U_{t_1}^1\right]^{|B|}\\
&=I\!\!E_{V_{t_1}^1}\, I\!\!P\left[V_t^1\leq u\middle|V_{t_1}^1\right]^{M+1}\geq I\!\!P\left[V_t^1\leq u\right]^{M+1}
\end{aligned} \tag{3.11}$$

Here we have used the induction hypothesis in the first inequality and Jensen inequality in the last one. $\diamondsuit$

Since, as we have seen, the maximum of independent walks gives a stochastic upper bound on the maximum of branching walks, (i) and (ii) follows for the independent *and* the branching case at the same time, if we prove: the *upper bounds* for the equalities for lim sup and lim inf for the *independent case* and *lower bounds* for lim sup and lim inf for the *branching case*.

So, let us continue with the proof of the lower bound on the lim sup (i) for the branching model. As indicated above, it suffices to consider a fixed configuration of branching times in the full measure set where (3.4) holds. We will show that for all $\epsilon>0$ and $\rho>1$ large enough

$$I\!\!P'\left[\sup_{1\leq\mu\leq M_{N_{k-1}}}\left(X_{N_k}^\mu-X_{N_{k-1}}^\mu\right)\geq(1-\epsilon)\sqrt{p+1}\,l_{N_k-N_{k-1}}\quad\text{for inf. many}\quad k\right]=1 \tag{3.12}$$

To see that this is enough we write

$$\sup_{1\leq\mu\leq M_{N_k}} X^{\mu}_{N_k} \geq X^{\mu_k}_{N_k} = \left(X^{\mu_k}_{N_k} - X^{\mu_k}_{N_{k-1}} \right) + X^{\mu_k}_{N_{k-1}} \tag{3.13}$$

where μ_k denotes the maximizer among the indices $\mu \in \{1,\dots,M_{N_{k-1}}\}$ of the increments $X^{\mu}_{N_k} - X^{\mu}_{N_{k-1}} =: W_{N_k - N_{k-1}}$.

Note that all processes X^{μ}_N involved in the search of this maximum are born before the time N_{k-1}, thus their increments are just independent Gaussians (as they would be trivially in the independent walk case). Note also that the random variable μ_k is independent of the 'past history' up to the time N_{k-1} and the $X^{\mu_k}_{N_{k-1}}$ have the same distribution as $X^1_{N_{k-1}}$. Since $I\!P\left[X^1_{N_{k-1}} \leq -(1+\tilde{\epsilon})l_{N_{k-1}}\right]$ is summable over k (which is used to prove the classical LIL) we conclude that the last term in (3.13) is bounded below by $-(1+\tilde{\epsilon})l_{N_{k-1}}$ (except for finitely many exceptions).

So we have on a full measure set for infinitely many k that (say)

$$\sup_{1\leq\mu\leq M_{N_k}} X^{\mu}_{N_k} \geq (1-\epsilon)\sqrt{p+1}\, l_{N_k - N_{k-1}} - (1+\tilde{\epsilon})l_{N_{k-1}} \tag{3.14}$$

For large k the r.h.s. behaves like

$$\left((1-\epsilon)\sqrt{p+1}\sqrt{1-\frac{1}{\rho}} - \frac{1+\tilde{\epsilon}}{\sqrt{\rho}} \right) l_{N_k} \tag{3.15}$$

Thus, if we have (3.14) for all large $\rho > 1$ and small $\epsilon > 0$ the first bracket can be made arbitrarily close to one and from (3.12) we see that this proves that $\limsup_{N\uparrow\infty} \frac{X^*_N}{\sqrt{p+1}\,l_N} \geq 1$.

Now, to show (3.12) we use the independence over μ to write

$$\begin{aligned} &I\!P\Big[\sup_{1\leq\mu\leq M_{N_{k-1}}} W^{\mu}_{N_k-N_{k-1}} \geq (1-\epsilon)\sqrt{p+1}\, l_{N_k-N_{k-1}} \Big] \\ &= 1 - \left(1 - I\!P\left[W_{N_k-N_{k-1}} \geq (1-\epsilon)\sqrt{p+1}\, l_{N_k-N_{k-1}} \right] \right)^{M_{N_{k-1}}} \end{aligned} \tag{3.16}$$

To get a lower bound on the occurring probability in the last line we use the simple lower bound on the tails of the Normal distribution stating that, $1-\Phi(x) \geq \exp\left(-\frac{x^2(1+\epsilon')}{2}\right)$, for any arbitrarily small $\epsilon' > 0$, for sufficiently large x. This gives us

$$\begin{aligned} &I\!P\left[W_{N_k-N_{k-1}} \geq (1-\epsilon)^2\sqrt{p+1}\, l_{N_k-N_{k-1}} \right] \geq \left(\log\left(N_k - N_{k-1}\right)\right)^{-(1+\epsilon')(1-\epsilon)^2(p+1)} \\ &\geq const\, k^{-(1+\epsilon')(1-\epsilon)^2(p+1)} \end{aligned} \tag{3.17}$$

for sufficiently large k, with some ϵ, ϵ' -independent constant. We note that the growthrate of M_N is such that $M_{N_{k-1}} \geq k^{p(1-\epsilon'')}$ for any arbitrarily small $\epsilon'' > 0$, for any fixed ρ, for large k. From this it is easy to see that the r.h.s. of (3.16) is bounded from below by $Const\, k^{-(1+\epsilon')(1-\epsilon)^2(p+1)} \times k^{p(1-\epsilon'')}$. But, for any small $\epsilon > 0$ we can find $\epsilon', \epsilon'' > 0$ s.t. this expression is bigger than $Const\, k^{-1}$. Hence it is nonsummable over k. Which, by the second half of the Borel Cantelli Lemma shows (3.12), as desired.

The proof of (ii) parallels of course the preceding; however, there are some differences that need to be taken care of so we need to run through the skeleton of the proof of (i). Let us start with the lower bound on the liminf for the branching case (which is analogous to the upper bound for the limsup). Again we condition on a fixed branching history s.t. (3.4) holds.

We will show that for each $\delta > 0$ there exists an exponential subsequence $N_k \sim \rho^k$ as above (with $\rho > 1$ sufficiently small) s.t.

$$\sum_{k=1}^{\infty} I\!P' \left[\min_{N_{k-1} < N \leq N_k} \frac{X_N^*}{\sqrt{p} l_N} \leq 1 - 2\delta \right] < \infty \tag{3.18}$$

This time we split of the maximum of the branching processes in question in three parts. The main part will be a maximum over independent processes that can be estimated as before while the remainder terms will be shown to be irrelevant. We fix $0 < \eta < 1$ that will be chosen sufficiently small later. We write

$$\begin{aligned} X_N^* &\geq \max_{1 \leq \nu \leq M_{\eta N_{k-1}}} \left(X^\nu_{\eta N_{k-1}} + \left(X^\nu_{N_{k-1}} - X^\nu_{\eta N_{k-1}} \right) + \left(X^\nu_N - X^\nu_{N_{k-1}} \right) \right) \\ &\geq \max_{1 \leq \nu \leq M_{\eta N_{k-1}}} \left(X^\nu_{N_{k-1}} - X^\nu_{\eta N_{k-1}} \right) + \left(X^{\nu^\eta_{k-1}}_{\eta N_{k-1}} + X^{\nu^\eta_{k-1}}_N - X^{\nu^\eta_{k-1}}_{N_{k-1}} \right) \\ &=: \tilde{U}_{k-1} - \tilde{X}_{N,k-1} \end{aligned} \tag{3.19}$$

Here ν^η_{k-1} is a maximizer of the maximum in the second line. $\tilde{X}_{N,k-1}$ is a walk of length $N + \eta N_{k-1} - N_{k-1}$ that is independent of the first term in the last line. In this way we get, for $\delta > 0$,

$$\begin{aligned} &I\!P' \left[\min_{N_{k-1} < N \leq N_k} \frac{X_N^*}{\sqrt{p} l_N} \leq 1 - 2\delta \right] \\ &\leq I\!P' \left[\tilde{U}_{k-1} \leq (1-\delta) \sqrt{p} l_{N_k} \right] + I\!P' \left[\max_{N_{k-1} < N \leq N_k} \tilde{X}_{N,k-1} \geq \delta \sqrt{p} l_{N_k} \right] \end{aligned} \tag{3.20}$$

Let us focus on the second term first which gives the corrections inside the block to the main term. Using the Gaussian upper estimate we get as in (3.9)

$$I\!P' \left[\max_{N_{k-1} < N \leq N_k} \tilde{X}_{N,k-1} \geq \delta \sqrt{p} l_{N_k} \right] \leq \exp\left(- \frac{\delta^2 p N_k \log\log N_k}{2\left(N_k - (1-\eta) N_{k-1}\right)} \right) \tag{3.21}$$

Now, the argument in the last exponential behaves as $-\frac{\delta^2 \rho}{2(\rho - 1 + \eta)} p \log\log N_k$. Thus, for fixed $\delta > 0$, we may choose $\rho > 1$ sufficiently small and $\eta > 0$ sufficiently small to achieve that for sufficiently large k the r.h.s. of (3.21) is bounded above by (say) $Const\,(\rho) \frac{1}{k^2}$. This is summable.

To prove the summability of the first term on the r.h.s. of (3.20) we choose $\epsilon < \delta$. By suitable further shrinkage of $\rho > 1$ and $\eta > 0$ we can obtain that

$$I\!P' \left[\tilde{U}_{k-1} \leq (1-\delta)\sqrt{p} l_{N_k} \right] \leq I\!P' \left[\tilde{U}_{k-1} \leq (1-\epsilon)\sqrt{p} l_{(1-\eta)N_{k-1}} \right] \tag{3.22}$$

Let us write for short $M \equiv M_{\eta N_k}$. Using the Gaussian estimate from below on the tails of the distribution function we have, with arbitrary small $\epsilon' > 0$, for large enough N that

$$I\!P \left[W_N \geq (1-\epsilon)\sqrt{p} l_N \right] \geq (\log N)^{-(1+\epsilon')(1-\epsilon)^2 p} \tag{3.23}$$

where W_N is a dummy walk and hence

$$\begin{aligned} & I\!P'[\tilde{U}_{k-1} \leq (1-\epsilon)\sqrt{p} l_{(1-\eta)N_{k-1}}] \\ & = \exp\left(M \log\left(1 - I\!P[W_{(1-\eta)N_{k-1}} \geq (1-\epsilon)\sqrt{p} l_{(1-\eta)N_{k-1}}] \right) \right) \\ & \leq \exp\left(-M I\!P[W_{(1-\eta)N_{k-1}} \geq (1-\epsilon)\sqrt{p} l_{(1-\eta)N_{k-1}}] \right) \\ & \leq \exp\left(-Const\, M \times (\log((1-\eta)N_{k-1}))^{-(1+\epsilon')(1-\epsilon)^2 p} \right) \end{aligned} \tag{3.24}$$

We notice that, for any fixed $\eta > 0$ the lower bound $M_{\eta N_k} \geq k^{p(1-\epsilon'')}$ holds, for k sufficiently large. This gives us that, with arbitrarily small $\epsilon', \epsilon'' > 0$, for large k,

$$I\!P' \left[\tilde{U}_{k-1} \leq (1-\delta)\sqrt{p} l_{N_k} \right] \leq \exp\left(-Const\, k^{p(1-\epsilon'')} \times k^{-(1+\epsilon')(1-\epsilon)^2 p} \right) \tag{3.25}$$

For $\epsilon', \epsilon'' > 0$ small enough this expression is bounded from above by $\exp(-Const\, k^{\alpha})$ with a positive exponent α. Since this is summable over k we are done with the proof of the lower bound for the liminf.

To complete the proof of (ii) it remains to show the upper bound for independent walks. The proof follows the lower bound for the limsup. We show that for all $\epsilon > 0$ and $\rho > 1$ large enough

$$I\!P\Big[\sup_{1 \leq \mu \leq M_{N_k}} \left(X^{\mu}_{N_k} - X^{\mu}_{N_{k-1}} \right) \leq (1+\epsilon)\sqrt{p} l_{N_k - N_{k-1}} \quad \text{for inf. many} \quad k \Big] = 1 \tag{3.26}$$

To see that this is enough we write

$$\sup_{1 \leq \mu \leq M_{N_k}} X^{\mu}_{N_k} \leq \sup_{1 \leq \mu \leq M_{N_k}} \left(X^{\mu}_{N_k} - X^{\mu}_{N_{k-1}} \right) + \sup_{1 \leq \mu \leq M_{N_k}} X^{\mu}_{N_{k-1}} \tag{3.27}$$

Then we use (i) to bound the last sup. With the same type of argument as before we then see that assuming (3.26) for all large $\rho > 1$ and small $\epsilon > 0$ proves that $\liminf_{N \uparrow \infty} \frac{X^*_N}{\sqrt{p} l_N} \leq 1$.

To show (3.26) we write

$$
\begin{aligned}
&I\!P\Big[\sup_{1\le\mu\le M_{N_k}}\left(X^{\mu}_{N_k}-X^{\mu}_{N_{k-1}}\right)\le(1+\epsilon)\sqrt{p}\,l_{N_k-N_{k-1}}\Big]\\
&=\left(1-I\!P\left[W_{N_k-N_{k-1}}\ge(1+\epsilon)\sqrt{p}\,l_{N_k-N_{k-1}}\right]\right)^{M_{N_k}}
\end{aligned}
\tag{3.28}
$$

with a dummy walk W_N. We have the from Gaussian bound

$$
\begin{aligned}
&I\!P\left[X_{N_k-N_{k-1}}\ge(1+\epsilon)\sqrt{p}\,l_{N_k-N_{k-1}}\right]\\
&\le\left(\log\left(N_k-N_{k-1}\right)\right)^{-(1+\epsilon)^2p}\le Const\,k^{-(1+\epsilon)^2p}
\end{aligned}
\tag{3.29}
$$

for sufficiently large k. For large k we have $M_{N_k}\le k^{p(1+\epsilon'')}$ for arbitrarily small ϵ''. From this follows that the r.h.s. of (3.28) goes to 1 with $k\uparrow\infty$ for any $\epsilon>0$. This in particular shows the nonsummability and thus (3.26).

We turn to the proof of (iii). Note that the *upper bound* for both the independent walks and the branching walks simply follows from

$$
\limsup_{N\uparrow\infty}\frac{X^{\nu}_N-Y^{\nu}_N}{l_N}\le\limsup_{N\uparrow\infty}\frac{X^{\nu}_N}{l_N}-\liminf_{N\uparrow\infty}\frac{Y^{\nu}_N}{l_N}=1-\sqrt{p}
\tag{3.30}
$$

It remains to prove the lower bound. Let us assume at first that $0<p<1$ so that the lim sup will be positive. Following the standard pattern of our proof of (i) and (ii) we show that for all $\epsilon>0$ and $\rho>1$ large enough we have for the events

$$
A_k:=\left\{\left(X^{\nu}_{N_k}-X^{\nu}_{N_{k-1}}\right)-\sup_{1\le\mu\le M_{N_k},\mu\ne\nu}\left(X^{\mu}_{N_k}-X^{\mu}_{N_{k-1}}\right)\ge(1-\epsilon)(1-\sqrt{p})l_{N_k-N_{k-1}}\right\}
\tag{3.31}
$$

that

$$
I\!P\Big[A_k\quad\text{for inf. many}\quad k\Big]=1
\tag{3.32}
$$

To check that this is enough also in this case we write

$$
\begin{aligned}
X^{\nu}_{N_k}-\sup_{1\le\mu\le M_{N_k},\mu\ne\nu}X^{\mu}_{N_k}\ge&\left(X^{\nu}_{N_k}-X^{\nu}_{N_{k-1}}\right)-\sup_{1\le\mu\le M_{N_k},\mu\ne\nu}\left(X^{\mu}_{N_k}-X^{\mu}_{N_{k-1}}\right)\\
&+X^{\nu}_{N_{k-1}}-\sup_{1\le\mu\le M_{N_k},\mu\ne\nu}X^{\mu}_{N_{k-1}}
\end{aligned}
\tag{3.33}
$$

Then we use again the upper bounds (i) on the last sup (and the classic LIL on the first term in the last line). So we see in the same fashion as before that (3.32) for all large $\rho>1$ and small $\epsilon>0$ implies the desired $\limsup_{N\uparrow\infty}\frac{X^{\nu}_N-Y^{\nu}_N}{(1-\sqrt{p})\ l_N}\ge1$.

To show (3.32) we need of course an estimate on the joint behavior of X^{ν}_N and Y^{ν}_N at fixed time N. We have

Lemma III.2: *Let X^μ_N be either independent or branching Brownian motions with fixed branching scheme, $M_N \uparrow \infty$, and W_N denote an independent walk. Then, for fixed $y > -1$ and any (arbitrarily small) fixed $\tilde\epsilon > 0$, for N sufficiently large the following estimate holds*

$$\mathbb{P}\left[W_N - \sup_{1\le\mu\le M_N} X^\mu_N \ge y\sqrt{2N\log M_N}\right] \ge M_N^{-\tilde\epsilon-(1+y)^2} \tag{3.34}$$

Remark: Note that the bound can be seen as a generalization of the obvious bound for independent walks with $y = 0$; here the l.h.s. equals $\frac{1}{M_N+1}$, due to symmetry.

Proof: Using the stochastic bound of Lemma III.1 we need only consider the independent walk case, since W_N and $\sup_{1\le\mu\le M_N} X^\mu_N$ are independent. Set $d = y\sqrt{2N\log M_N}$. We introduce a 'split point' $z > 0$ to write

$$\mathbb{P}\left[W_N - \sup_{1\le\mu\le M_N} X^\mu_N \ge d\right] \ge \mathbb{P}\left[\sup_{1\le\mu\le M_N} X^\mu_N \le z\right] \mathbb{P}\left[W_N \ge d+z\right] \tag{3.35}$$

Setting now $z = \sqrt{2N\log M_N}$ we have for the first term with the usual Gaussian upper bounds

$$\begin{aligned}\mathbb{P}\left[\sup_{1\le\mu\le M_N} X^\mu_N \le z\right] &= \left(1-\mathbb{P}\left[X^\mu_N \le z\right]\right)^{M_N}\\ \ge \left(1-\exp\left(-\frac{z^2}{2N}\right)\right)^{M_N} &= \left(1-\frac{1}{M_N}\right)^{M_N} \to \frac{1}{e} > 0\end{aligned} \tag{3.36}$$

For the second term in (3.35) we use the Gaussian lower bound which gives us, for any $\epsilon' > 0$, if N is sufficiently large, that

$$\mathbb{P}\left[W_N \ge d+z\right] = \mathbb{P}\left[W_N \ge (1+y)\sqrt{2N\log M_N}\right] \ge \exp\left(-(1+\epsilon')(1+y)^2\log M_N\right) \tag{3.37}$$

From this the lemma follows.$\diamondsuit$

We will first consider the independent case and discuss the changes for the branching case later. To apply the Lemma and show (3.32) we express the r.h.s. in the occurring inequality in terms of M_{N_k}. In fact, our hypothesis on the growth of M_N implies that for all $\bar\epsilon > 0$, for any $\rho > 1$ we have

$$l_{N_k-N_{k-1}} \le (1+\bar\epsilon)\frac{1}{\sqrt{p}}\sqrt{2(N_k-N_{k-1})\log M_{N_k}} \tag{3.38}$$

for k large enough. So we have

$$\begin{aligned}&\mathbb{P}\Big[A_k\Big]\\ &\ge \mathbb{P}\Big[X^\nu_{N_k-N_{k-1}} - \sup_{\substack{1\le\mu\le M_{N_k}\\ \mu\ne\nu}} X^\mu_{N_k-N_{k-1}} \ge (1-\epsilon)(1+\bar\epsilon)(\frac{1}{\sqrt{p}}-1)\sqrt{2(N_k-N_{k-1})\log M_{N_k}}\Big]\end{aligned} \tag{3.39}$$

Lemma III.2 estimates the last line from below to give

$$\mathbb{P}\left[A_k\right] \geq k^{-p(1+\epsilon'')\left(\tilde{\epsilon}+\left(1+(1-\epsilon)(1+\bar{\epsilon})(\frac{1}{\sqrt{p}}-1)\right)^2\right)} \tag{3.40}$$

using $M_{N_k} \leq k^{p(1+\epsilon'')}$ for k large enough. Now, for small $\epsilon > 0$ and $\tilde{\epsilon} = \bar{\epsilon} = \epsilon'' = 0$, the exponent in the last line equals $-1 + 2p(\frac{1}{\sqrt{p}} - 1)\epsilon + \mathcal{O}(\epsilon^2)$ which is *bigger* than -1. Now, allowing for sufficiently small nonzero $\tilde{\epsilon}, \bar{\epsilon}, \epsilon''$ does not destroy this property. Hence the r.h.s. of (3.40) is nonsummable, as we wanted to show.

In the case $p > 1$ where the lim inf will be negative it is easy to check that the above arguments work in an analogous way. In fact, now we need to show

$$\begin{aligned} \mathbb{P}\Big[\left(X^\nu_{N_k} - X^\nu_{N_{k-1}}\right) - \sup_{\substack{1\leq\mu\leq M_{N_k}\\ \mu\neq\nu}} \left(X^\mu_{N_k} - X^\mu_{N_{k-1}}\right) \\ \geq -(1+\epsilon)(\sqrt{p}-1) l_{N_k - N_{k-1}} \quad \text{for inf. many} \quad k\Big] = 1 \end{aligned} \tag{3.41}$$

We need a bound of the form (3.38), but $\leq$ has to be replaced by $\geq$ and $\bar{\epsilon}$ by $-\bar{\epsilon}$. Note that when it comes to the application of Lemma 10 the resulting value of $y \approx \frac{1}{\sqrt{p}} - 1$ is still bigger than -1 which justifies its validity. Finally these arguments give the bound

$$\begin{aligned} &\mathbb{P}\Big[\left(X^\nu_{N_k} - X^\nu_{N_{k-1}}\right) - \sup_{\substack{1\leq\mu\leq M_{N_k}\\ \mu\neq\nu}} \left(X^\mu_{N_k} - X^\mu_{N_{k-1}}\right) \geq -(1+\epsilon)(\sqrt{p}-1) l_{N_k - N_{k-1}}\Big] \\ &\geq k^{-p(1+\epsilon'')\left(\tilde{\epsilon}+\left(1+(1+\epsilon)(1-\bar{\epsilon})(\frac{1}{\sqrt{p}}-1)\right)^2\right)} \end{aligned} \tag{3.42}$$

In this case, for small $\epsilon > 0$ and $\tilde{\epsilon} = \bar{\epsilon} = \epsilon'' = 0$, the exponent in the last line equals $-1 + 2p(1 - \frac{1}{\sqrt{p}})\epsilon + \mathcal{O}(\epsilon^2)$ which is bigger than -1. This finishes the proof of (iii) for the independent case.

For the branching case we argue as follows. Note first of all that the events A_k in (3.31) are independent over k, also in this case. Let us, in a more explicit way then in our earlier proceeding, denote by $\tilde{\mathbb{P}}$ the average over the branching times and family tree (and by $\mathbb{P}'$ the conditional average given the branching structure over the walks). So we have $\mathbb{P} = \tilde{\mathbb{P}} \otimes \mathbb{P}'$. We denote by B^ν_k the event that the particle ν has no offspring for $t \in [N_{k-1}, N_k]$. We may write

$$\mathbb{P}\Big[A_k\Big] \geq \tilde{\mathbb{E}}\left[1_{B_k} \mathbb{P}'\left[A_k\right]\right] \tag{3.43}$$

Now, for any fixed branching history in B_k we can apply Lemma III.2; in fact the increments $X^\nu_{N_k} - X^\nu_{N_{k-1}}$ and the process $\sup_{1\leq\mu\leq M_{N_k}, \mu\neq\nu}\left(X^\mu_{N_k} - X^\mu_{N_{k-1}}\right)$ are now independent w.r.t. $\mathbb{P}'$. So we get again the lower bounds (3.40) resp (3.42), uniformly for all splitting times in B_k. It remains thus to estimate $\tilde{\mathbb{P}}[B_k]$. We have

$$\begin{aligned} \tilde{\mathbb{P}}\Big[B_k\Big] &= \exp\left(-\int_{N_{k-1}}^{N_k} \gamma(t)dt\right) \\ &= \exp\left(-p(1+o(1))\log\log N_k + p(1+o(1))\log\log N_{k-1}\right) = k^{o(1)} \end{aligned} \tag{3.44}$$

But this does not change the nonsummability of the estimate resulting from (3.43) with (3.40) resp. (3.42). This concludes the proof of Theorem 4.$\diamondsuit$

Finally we come to the proof of (iv). By now the structure of the proof should be clear, so that we can be a bit briefer; we will however take time to explain the emergence of the different form of the p-dependence on the r.h.s of (iv).

Let us first discuss the case $p \geq 1$. Note that in this case the lower bound follows from (ii) and the classical LIL. Due to Lemma III.1 we need only prove the upper bound for independent walks. To do so, we need to show upper bounds on the probabilities of the events

$$D_k := \left\{ \max_{N_{k-1} \leq N \leq N_k} \left(\sup_{1 \leq \mu \leq M_{N_k}, \mu \neq \nu} X_N^\mu - X_N^\nu \right) \geq (1+2\delta) C(p) l_{N_{k-1}} \right\} \tag{3.45}$$

to show their summability, where $C(p) = 1 + \sqrt{p}$. We have

$$\begin{aligned} D_k \subset & \left\{ \sup_{1 \leq \mu \leq M_{N_k}, \mu \neq \nu} X_{N_{k-1}}^\mu - X_{N_{k-1}}^\nu \geq (1+\delta) C(p) l_{N_{k-1}} \right\} \\ \cup & \left\{ \max_{N_{k-1} \leq N \leq N_k} \left(\sup_{1 \leq \mu \leq M_{N_k}, \mu \neq \nu} \left(X_N^\mu - X_{N_{k-1}}^\mu \right) - \left(X_N^\nu - X_{N_{k-1}}^\nu \right) \right) \geq \delta C(p) l_{N_{k-1}} \right\} =: D_k^1 \cup D_k^2 \end{aligned} \tag{3.46}$$

Let us denote by $I\!E^\nu$ the expectation over the ν'th walk and by $I\!E^{\nu^c}$ the expectation over the walks for $\mu \neq \nu$. Then we have

$$\begin{aligned} & I\!P\left[D_k^2\right] = 1 \\ & \quad - I\!E^\nu \left(1 - I\!P^{\nu^c} \left[\max_{N_{k-1} < N \leq N_k} \left(\left(X_N^\mu - X_{N_{k-1}}^\mu \right) - \left(X_N^\nu - X_{N_{k-1}}^\nu \right) \right) \geq \delta C(p) l_{N_{k-1}} \right] \right)^{M_{N_k}} \\ & \leq 1 - \left(1 - I\!P \left[\max_{N_{k-1} < N \leq N_k} \left(\left(X_N^\mu - X_{N_{k-1}}^\mu \right) - \left(X_N^\nu - X_{N_{k-1}}^\nu \right) \right) \geq \delta C(p) l_{N_{k-1}} \right] \right)^{M_{N_k}} \\ & \leq 1 - \left(1 - \exp\left(- \frac{\delta^2 C(p)^2 N_{k-1} \log\log N_{k-1}}{N_k - N_{k-1}} \right) \right)^{M_{N_k}} \end{aligned} \tag{3.47}$$

In the first equation we have used the independent over ν, in the next inequality we have used the Jensen inequality and in the last inequality we have used the Gaussian maximal inequality on the walk $X_N^\mu - X_N^\nu$ to get an estimate like (3.21). Using now the same arguments as after (3.21) we see that we can bound from above the exponential in the last line of (3.47) by k^{-q} for any desired large positive q, by choosing ρ sufficiently close to one. This makes (3.47) summable. Thus it remains to treat the event D_k^1.

Before doing so, let us comment on the case $0 < p < 1$. There we need to prove upper bounds in the independent walk case on the probabilities D_k^1 with $C(p) = \sqrt{2(p+1)}$. (The estimate of the event D_k^2 is of course completely identical.) In this case we also need lower bounds in the branching case of the probabilities of the events

$$E_k := \left\{ \sup_{1\leq\mu\leq M_{N_k},\mu\neq\nu} \left(X_{N_k}^{\mu} - X_{N_{k-1}}^{\mu}\right) - \left(X_{N_k}^{\nu} - X_{N_{k-1}}^{\nu}\right) \geq (1-\delta)\sqrt{2(p+1)} l_{N_k - N_{k-1}} \right\} \quad (3.48)$$

The same argument as for the proof of (iii) (see (3.43)) however show, that the occurring variables in (3.46) can be treated as independent over μ.

The key estimate to treat the D_k^1's and E_k's is the following analogue of Lemma III.2:

Lemma III.3: *Let $G, G_1, G_2, \dots, G_M$ be i.i.d. Normal variables. Write $a_M = \sqrt{2\log M}$, as in Chapter II. Then, for any $\tilde{\epsilon} > 0$ we have, for $M \geq M_0(\epsilon)$ sufficiently large, that*

$$\begin{aligned} I\!P\left[\max_{i=1,\dots,M} G_i \geq G + y a_M\right] &\leq M^{\tilde{\epsilon}-(y-1)^2} \quad, \geq M^{-\tilde{\epsilon}-(y-1)^2} \quad \text{for} \quad 1 < y < 2 \\ &\leq M^{\tilde{\epsilon}-\frac{y^2}{2}+1} \quad, \geq M^{-\tilde{\epsilon}-\frac{y^2}{2}+1} \quad \text{for} \quad y > 2 \end{aligned} \quad (3.49)$$

Proof: We write

$$\begin{aligned} &I\!P\left[\max_{i=1,\dots,M} G_i \geq G + y a_M\right] = \int I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z - y a_M) dz \\ &= \int_{-\infty}^{(1+\epsilon)a_M} I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z - y a_M) dz + \int_{(1+\epsilon)a_M}^{\infty} I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z - y a_M) dz \end{aligned} \quad (3.50)$$

For the first term we have the upper bound

$$\int_{-\infty}^{(1+\epsilon)a_M} p(z - y a_M) dz \leq e^{-\frac{(y-1-\epsilon)^2 a_M^2}{2}} = M^{-(y-1-\epsilon)^2} \quad (3.51)$$

and the lower bound

$$\begin{aligned} &\int_{-\infty}^{(1-\epsilon)a_M} I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z - y a_M) dz \geq const \int_{-\infty}^{(1-\epsilon)a_M} p(z - y a_M) dz \\ &\geq e^{-\frac{(1+\epsilon')(y-1+\epsilon)^2 a_M^2}{2}} = M^{-(1+\epsilon')(y-1+\epsilon)^2} \end{aligned} \quad (3.52)$$

or any $y > 1$ where ϵ is small enough, for large M. For the second term in (3.50) we use

Lemma III.4: *Let $\epsilon, \bar{\epsilon} > 0$. Then there exists $M_0 = M_0(\epsilon, \bar{\epsilon})$ s.t. for all $M \geq M_0(\epsilon)$ for all $z \geq (1+\epsilon)a_M$*

$$I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] \geq M e^{-\frac{z^2(1+\bar{\epsilon})}{2}}, \quad \leq Const\, M e^{-\frac{z^2}{2}} \quad (3.53)$$

Proof: We write, with arbitrarily small $\epsilon' > 0$,

$$I\!P[\max_{i=1,\dots,M} G_i \geq z] = 1-(1-I\!P[G\geq z])^M \geq 1-\left(1-e^{-\frac{(1+\epsilon')z^2}{2}}\right)^M, \quad \leq 1-\left(1-e^{-\frac{z^2}{2}}\right)^M \tag{3.54}$$

Since $M \times \sup_{z\geq(1+\epsilon)a_M} \exp\left(-\frac{(1+\epsilon')z^2}{2}\right) = M\times M^{-(1+\epsilon')(1+\epsilon)^2} \to 0$ the bounds follow. $\diamondsuit$

Applying those bounds of Lemma III.4 directly gives after a little Gaussian computation

$$\begin{aligned}
&\int_{(1+\epsilon)a_M}^{\infty} I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z-ya_M)dz \\
&\qquad \geq M^{1-\frac{1+\bar\epsilon}{2+\bar\epsilon}y^2} \times \frac{1}{\sqrt{2+\bar\epsilon}} I\!P\left[G \geq \left(1+\epsilon-\frac{y}{2+\bar\epsilon}\right)\frac{a_M}{\sqrt{2+\bar\epsilon}}\right], \\
&\qquad \leq Const\, M^{1-\frac{1}{2}y^2} \times \frac{1}{\sqrt{2}} I\!P\left[G \geq \left(1+\epsilon-\frac{y}{2}\right)\frac{a_M}{\sqrt{2}}\right]
\end{aligned} \tag{3.55}$$

Now we have to distinguish the cases $y > 2$ and $1 < y < 2$. Let at first $y > 2$. Then the Gaussian probability converges to 1 and we have, for some small $\epsilon_1 > 0$,

$$\int_{(1+\epsilon)a_M}^{\infty} I\!P\left[\max_{i=1,\dots,M} G_i \geq z\right] p(z-ya_M)dz \geq M^{1-\frac{y^2}{2}-\epsilon_1}, \quad \leq Const\, M^{1-\frac{y^2}{2}} \tag{3.56}$$

But since $-(y-1)^2 \leq -\frac{y^2}{2}+1$ iff $(y-2)^2 \geq 0$ and the upper bound in (3.56) is valid for any $y > 1$, Lemma III.3 follows from (3.56) and (3.51) and (3.52). $\diamondsuit$

Now the estimation of the desired events is done as always (with the usual play of ϵ-corrections to ensure summability resp. nonsummability over k.) We spare the reader the details this time. Let us only explain here, how the functional form of the r.h.s. of (iv) arises. We have $M_{N_k} \sim k^p$, $a_{M_{N_k}} \sim \sqrt{2p\log k}$, $l_{N_k} \sim \sqrt{N_k}\frac{a_{M_{N_k}}}{\sqrt{p}}$. Thus, we apply Lemma III.3 with $y = \frac{C(p)}{\sqrt{p}}$.

To (formally) produce the harmonic series in k we must have $p(y-1)^2 = 1$ for $1 < y < 2$ resp. $p(y^2/2-1) = 1$ for $y > 2$. Solving for $C(p)$ yields the correct r.h.s. of (iv). $\diamondsuit$

IV. The a.s. Behavior of the Empirical Distribution

In this chapter we give the proof of Theorem 3, using large deviation arguments for sample paths. The proof consists of two parts. Let us put

$$x_p = e^{-4\left(\frac{1}{p}-1\right)} \tag{4.1}$$

In Part I we show that $\limsup_{N\uparrow\infty} n_N^\nu \leq 1-x_p$. a.s. As we will see, this result follows from the lower bound for Y_N^ν from Theorem 4 (ii), Strassen's theorem applied to the walk X_N^ν, and the solution of a variational problem (see Lemmas IV.1,IV.2). In Part II we show the more subtle statement that each $x \in [0, 1-x_p]$ is a cluster point of the sequence n_N^ν. This result will need a refined upper bound on the probability of exceedences of *segments of the path* Y_N^ν above a curve of the form $N \mapsto (\sqrt{p}-\epsilon)l_N$. These will be used to prove that, infinitely often, portions of the paths of X_N^ν and Y_N^ν become jointly close to fixed curves, in such a way that a prescribed value of n_N^ν can be approximated. In the course of the proof we replace the sums by continuous time expressions, so we will need some approximation and regularization arguments.

We start with Part I and consider the independent walk case and the branching case at the same time. From the LIL Theorem 4 (ii) lower bound applied to Y_N^ν follows that, for all (arbitrarily small) $\epsilon > 0$, a.s.,

$$\begin{aligned} &\limsup_{N\uparrow\infty} \frac{1}{N}\#\{n : 1 \le n \le N \big| X_n^\nu > Y_n^\nu\} \\ &\le \limsup_{N\uparrow\infty} \frac{1}{N}\#\{n : 1 \le n \le N \big| X_n^\nu > (\sqrt{p}-\epsilon)l_n\} \end{aligned} \tag{4.2}$$

So we are left with a onedimensional problem on which we apply the classical Strassen theorem (see [DS] p.21, Theorem 1.4.1). To do so, we find it convenient to represent the walk X_n^ν as sampling of a Brownian path $X_\cdot : t \mapsto X_t$, defined for *all* $t \ge 0$. (For which we now drop the superscript for simplicity.) Define, following the notations of [DS], the space of Cameron-Martin functions $H^1 := \{\phi,\ \phi(x) = \int_0^x \dot\phi(t)dt,\ \dot\phi \in L^1[0,1]\}$, with energy norm given by $\|\phi\|_{H^1}^2 := \int_0^1 \dot\phi(x)^2 dx$ and denote $K := \{\phi \in H^1, \|\phi\|_{H^1}^2 \le 1\}$. Write $\|\phi\|_\infty = \sup_{0\le t\le 1}|\phi(t)|$.

Then, Strassen's theorem[5] tells us that

(i) $\limsup_{N\uparrow\infty} \left\| \frac{X_{\cdot N}}{l_N} - K \right\|_\infty = 0$

(ii) For all $\phi \in K$ there exists a subsequence N_s s.t. $\limsup_{s\uparrow\infty} \left\| \frac{X_{\cdot N_s}}{l_{N_s}} - \phi \right\|_\infty = 0$

To use this on the n-sum in the definition of n_N^ν we introduce the continuous time functional given by

$$I_p(\phi) = \text{Leb}\,(x \in [0,1], \phi(x) \ge \sqrt{px}) \tag{4.3}$$

for $\phi \in \mathcal{C}[0,1]$, where Leb denotes the Lebesgue measure. We have then

Lemma IV.1: *For all $\epsilon > 0$, a.s.,*

$$\limsup_{N\uparrow\infty} \frac{1}{N}\#\{n : 1 \le n \le N \big| X_n > \sqrt{p}l_n\} \le \limsup_{N\uparrow\infty} I_{p-\epsilon}\left(\frac{X_{\cdot N}}{l_N}\right) \tag{4.4}$$

[5] see also [DS], page 22 (1.4.5); note that the results for compact interval [0,1] are trivially implied by the ones on the real line with weighted norms.

whose proof is deferred to the end of the chapter. So, with Strassen statement (i), it remains essentially to optimize I_p over functions in K. The maximizer will turn out to be one of the functions $\phi_{p,b}$, that are given, for $0 \le b \le 1$, by the piecewise definition

$$\begin{aligned} \phi_{p,b}(x) &= \sqrt{\frac{p}{b}}x, \quad \text{for} \quad 0 \le x \le b \\ &= \sqrt{px}, \quad \text{for} \quad b \le x \le 1 \end{aligned} \tag{4.5}$$

Note that we have

$$\|\phi_{p,b}\|_{H^1}^2 = \int_0^b \frac{p}{b}dx + \frac{p}{4}\int_b^1 \frac{1}{x}dx = p\left(1 - \frac{\log b}{4}\right) \tag{4.6}$$

Solving the equation $\|\phi_{p,b}\|_{H^1}^2 = 1$ for b gives then $b = x_p$ which explains the form of x_p. Thus $\phi_{p,b} \in K$ for $b \ge x_p$. It is now natural to guess that the maximum of I_p with energy $\|\phi_{p,b}\|_{H^1}^2 \le 1$ is actually attained for ϕ_{p,x_p}. Indeed, we have

Lemma IV.2: *For any $0 < p < 1$ we have*

$$\limsup_{N\uparrow\infty} I_p\left(\frac{X_{\cdot N}}{l_N}\right) = \sup_{\phi\in K} I_p(\phi) = I_p(\phi_{p,x_p}) = 1 - x_p \tag{4.7}$$

The right equality is clear. We prove the other equalities by regularization of the functional I_p (since it is noncontinuous w.r.t. sup-norm) and application of Strassen's theorem. To show that ϕ_{p,x_p} is in fact the minimizer we will use an approximation by piecewise linear functions. The details are given at the end of this chapter.

But assuming Lemma IV.1, Lemma IV.2 the proof of Claim I follows immediately from (4.2), by the continuity of the function $p \mapsto x_p$.

Let us come to Part II which is more involved. Again we consider independent and branching case at the same time. Let $b \in \left(e^{-4\left(\frac{1}{p}-1\right)}, 1\right)$. As usual, we use exponentially spaced blocks with endpoints $N_k = [\rho^k]$. We will show that, for all $\eta > 0$, there exists a (large) $\rho > 1$, s.t.

$$I\!P\left[\left|n_{N_k}^{\nu} - (1-b)\right| < \eta \quad \text{for inf. many} \quad k\right] = 1 \tag{4.8}$$

As a basic piece of information we will need an upper bound on the probability of exceedences of the path of Y_N^{ν}. Let now $\epsilon > 0$, $\phi \in H^1$. (Later we will choose $\phi = \phi_{p+\delta,b}$ for small $\delta > 0$.) Write

$$d_N := \sqrt{\log\log N} \tag{4.9}$$

Let us then define the sets

$$\begin{aligned} A_{1,k} &:= \left\{ \sup_{\substack{1\le\mu\le M_{N_k} \\ \mu\ne\nu}} \left(X_N^{\mu} - X_{N_{k-1}}^{\mu}\right) \le (\sqrt{p}+\epsilon)\sqrt{N}d_{N_k} \quad \text{for all} \quad N_{k-1} \le N \le N_k \right\} \\ A_{2,k} &:= \left\{ \left| \frac{X_N^{\nu} - X_{N_{k-1}}^{\nu}}{l_{N_k}} - \phi\left(\frac{N}{N_k}\right) \right| < \epsilon \quad \text{for all} \quad N_{k-1} \le N \le N_k \right\} \end{aligned} \tag{4.10}$$

Note that, in the independent walk case, these events are independent, for different k and different index $1, 2$. For the branching case, we will restrict ourselves to branching histories in the events B_k (see the definition before (3.43)); indeed, for a branching history satisfying B_k, the events $A_{1,k}, A_{2,k}$ are independent w.r.t the measure $I\!P'$ describing the diffusion part. To treat both cases simultaneously it is convenient to denote by B_k the whole probability space in the case of independent walks.

We will show below that, for all $\phi \in H^1$, for all (arbitrarily small) $\epsilon > 0$, for $\rho > 1$ large enough,

$$I\!P\Big[A_{1,k} \cap A_{2,k} \quad \text{occurs for inf. many} \quad k \in I\!N\Big] = 1 \tag{4.11}$$

Our first claim is that this implies (4.8). Indeed, by the the LILs on Y_N^ν, (4.11) implies that, on a full measure set for infinitely many k, for all $N_{k-1} \le N \le N_k$

$$\begin{aligned}
&X_N^\nu - \sup_{\substack{1\le\mu\le M_N\\ \mu\neq\nu}} X_N^\mu \ge X_N^\nu - \sup_{\substack{1\le\mu\le M_{N_k}\\ \mu\neq\nu}} X_N^\mu \\
&\ge X_N^\nu - X_{N_{k-1}}^\nu + X_{N_{k-1}}^\nu - \sup_{\substack{1\le\mu\le M_{N_k}\\ \mu\neq\nu}} \left(X_N^\mu - X_{N_{k-1}}^\mu\right) - \sup_{\substack{1\le\mu\le M_{N_k}\\ \mu\neq\nu}} X_{N_{k-1}}^\mu \\
&\ge \phi\left(\frac{N}{N_k}\right) l_{N_k} - \epsilon l_{N_k} - Const\, l_{N_{k-1}} - (\sqrt{p}+\epsilon)\sqrt{N} d_{N_k} - Const\, l_{N_{k-1}}
\end{aligned} \tag{4.12}$$

In the last line we have used the definition of the events $A_{1,k}$ and $A_{2,k}$ and the LILs we have already proven for the N_{k-1}-terms.

For large (of these infinitely many) k we have from the last inequality that for all $N_{k-1} \le N \le N_k$

$$X_N^\nu - \sup_{\substack{1\le\mu\le M_N\\ \mu\neq\nu}} X_N^\mu \ge \left(\phi\left(\frac{N}{N_k}\right) - \sqrt{p}\sqrt{\frac{N}{N_k}} - \frac{Const'}{\sqrt{\rho}} - Const\,\epsilon\right) l_{N_k} \tag{4.13}$$

For an analogous upper bound we use the LIL-lower bound directly for the sup-term (as in part I) to get for infinitely many k that for all $N_{k-1} \le N \le N_k$

$$\begin{aligned}
X_N^\nu - \sup_{\substack{1\le\mu\le M_N\\ \mu\neq\nu}} X_N^\mu &\le \phi\left(\frac{N}{N_k}\right) l_{N_k} + \epsilon l_{N_k} - (\sqrt{p}-\epsilon) l_N \\
&\le \left(\phi\left(\frac{N}{N_k}\right) - \sqrt{p}\sqrt{\frac{N}{N_k}} + Const\,\epsilon\right) l_{N_k}
\end{aligned} \tag{4.14}$$

We will now choose the function $\phi = \phi_{p+\delta,b}$ in a suitable way that ensures that from the last two estimates follows that $n_{N_k}^\nu$ is close to b. Our choice will yield a slight modification of the equation Leb $\left[x \in [0,1], \phi_{p,b}(x) \ge \sqrt{px}\right] = 1 - b$: It is elementary to verify that, for all $\bar\epsilon > 0$, there exist a (small) $\delta > 0$ and a (small) $\epsilon' > 0$ s.t.

(i) Leb $\left[x \in [0,1], \phi_{p+\delta,b}(x) \ge \sqrt{px} - \epsilon'\right] \le 1 - b + \bar\epsilon$

(ii) $\text{Leb}\left[x \in [0,1], \phi_{p+\delta,b}(x) \geq \sqrt{px} + \epsilon'\right] \geq 1 - b - \bar{\epsilon}$

Note moreover, for any fixed $b \in \left(e^{-4\left(\frac{1}{p}-1\right)}, 1\right)$, that $\delta > 0$ can always be chosen sufficiently small s.t. $\|\phi_{p+\delta,b}\|_{H^1} < 1$.

Now, for given $\eta > 0$, we choose $\phi = \phi_{p+\delta,b}$ such that (i) and (ii) hold with $\bar{\epsilon} = \frac{\eta}{2}$ and some $\delta > 0$ and $\epsilon' > 0$.

Then it follows from (4.14) that, for infinitely many k, for $\epsilon > 0$ sufficiently small and $\rho > 1$ sufficiently large,

$$\begin{aligned} &\#\{N : 1 \leq N \leq N_k \big| X_N^\nu - \sup_{\substack{1 \leq \mu \leq M_{N_k} \\ \mu \neq \nu}} X_N^\mu \geq 0\} \\ &\leq N_{k-1} + \#\left\{N : N_{k-1} < N \leq N_k \big| \phi_{p+\delta,b}\left(\frac{N}{N_k}\right) - \sqrt{p}\sqrt{\frac{N}{N_k}} + Const\,\epsilon \geq 0\right\} \\ &\leq N_{k-1} + N_k \text{Leb}\left[\frac{1}{\rho} \leq x \leq 1 : \phi_{p+\delta,b}\left(x\right) - \sqrt{p}\sqrt{x} + \frac{Const\,(\rho)}{N_k} + Const\,\epsilon \geq 0\right] \\ &\leq N_{k-1} + (1-b)N_k + \frac{\eta}{2}N_k \end{aligned} \tag{4.15}$$

Here the $\frac{Const\,(\rho)}{N_k}$ term controls the continuum approximation and we have chosen $\rho > 1$ large enough and $\epsilon > 0$ small enough s.t. the ρ and ϵ-dependent error terms in are controlled by $\epsilon' > 0$. Similarly we have a lower bound

$$\begin{aligned} &\#\{N : 1 \leq N \leq N_k \big| X_N^\nu - \sup_{\substack{1 \leq \mu \leq M_{N_k} \\ \mu \neq \nu}} X_N^\mu \geq 0\} \\ &\geq \#\left\{N : N_{k-1} < N \leq N_k \big| \phi_{p+\delta,b}\left(\frac{N}{N_k}\right) - \sqrt{p}\sqrt{\frac{N}{N_k}} - \frac{Const\,'}{\sqrt{\rho}} - Const\,\epsilon \geq 0\right\} \\ &\geq -N_{k-1} + (1-b)N_k - \frac{\eta}{2}N_k \end{aligned} \tag{4.16}$$

But from the last two inequalities follows in fact (4.8) (after further enlarging $\rho > 1$, if necessary).

Now, let us turn to the proof of (4.11). Note that, in the independent walk case, the events $A_{1,k}, A_{2,k}$ are independent, for different k and different index $1, 2$. For the branching case, we will restrict ourselves to branching histories in the events B_k (see the definition before (3.43)). Indeed, let us denote in short the branching history (splitting times and structure of family tree) by the random variable τ. Then we have that for all configurations of the family tree $v \in B_k$ that the events $A_{1,k}, A_{2,k}$ are independent w.r.t the conditional measure $I\!P[\cdot|\tau = v]$ describing the diffusion part. The key ingredient we need for this is the following lemma that provides us with information about the Y_N^ν.

Lemma IV.3: *For all $\epsilon > 0$ and all $\rho > 1$ (in the definition of the event $A_{1,k}$) we have $\lim_{k\uparrow\infty} I\!P[A_{1,k}] = 1$ for the independent walk case and $\lim_{k\uparrow\infty} \inf_{v\in B_k} I\!P[A_{1,k}|\tau = v] = 1$ for the branching model.*

Proof: We can assume that we are in the independent case. (Indeed, in the branching case, $v \in B_k$ means that the increments of the Brownian motions are conditionally independent. Trivially, all constants arising are uniform in the branching history v.)

We fix one block between N_{k-1} and N_k. The key idea is to break up this block into R exponentially spaced sub-blocks: We define for $i = 0, 1, \ldots, R$ their endpoints by $L_i := [\rho^{\frac{i}{R}} \rho^{k-1}]$. Later R will have be chosen to be sufficiently large, independent of k. Thus $L_0 = N_{k-1}$ and $L_R = N_k$. We have

$$
\begin{aligned}
I\!P[A^c_{1,k}] &= I\!P\left[\exists N_{k-1} \le N \le N_k : \sup_{\substack{1 \le \mu \le M_{N_k} \\ \mu \ne \nu}} \left(X^\mu_N - X^\mu_{N_{k-1}}\right) \ge (\sqrt{p} + \epsilon)\sqrt{N} d_{N_k}\right] \\
&\le \sum_{i=1}^{R} I\!P\left[\exists L_{i-1} < N \le L_i : \sup_{\substack{1 \le \mu \le M_{N_k} \\ \mu \ne \nu}} \left(X^\mu_N - X^\mu_{N_{k-1}}\right) \ge (\sqrt{p} + \epsilon)\sqrt{N} d_{N_k}\right] \\
&\le \sum_{i=1}^{R} \left(1 - I\!P\left[\sup_{\substack{1 \le \mu \le M_{N_k} \\ \mu \ne \nu}} \sup_{L_{i-1} < N \le L_i} \left(X^\mu_N - X^\mu_{N_{k-1}}\right) \le (\sqrt{p} + \epsilon)\sqrt{L_{i-1}} d_{N_k}\right]\right) \\
&= \sum_{i=1}^{R} \left(1 - \left\{1 - I\!P\left[\sup_{L_{i-1} < N \le L_i} \left(X^\mu_N - X^\mu_{N_{k-1}}\right) \ge (\sqrt{p} + \epsilon)\sqrt{L_{i-1}} d_{N_k}\right]\right\}^{M_{N_k}}\right)
\end{aligned}
\tag{4.17}
$$

Using the maximal inequality as usual we get

$$
I\!P\left[\sup_{L_{i-1} < N \le L_i} \left(X^\mu_N - X^\mu_{N_{k-1}}\right) \ge (\sqrt{p} + \epsilon)\sqrt{L_{i-1}} d_{N_k}\right] \le e^{-\frac{(\sqrt{p}+\epsilon)^2 L_{i-1} d^2_{N_k}}{2L_i}} \tag{4.18}
$$

Due to our definition for the sub-blocks we have, for large k, $\frac{L_{i-1}}{L_i} \sim \rho^{\frac{1}{R}}$. Thus, for any fixed $\rho > 1$, $\epsilon > 0$, by choosing R sufficiently large we can ensure that for all sufficiently large k we have $(\sqrt{p} + \epsilon)\sqrt{\frac{L_{i-1}}{L_i}} \ge \sqrt{p} + \frac{\epsilon}{2}$ for all $i = 1, \ldots, R$.

Recall now that $M_{N_k} \sim Const\, k^p$, $N_k \sim Const\, \rho^k$. With this we obtain from (4.17), (4.18) that, for sufficiently large k,

$$
I\!P[A^c_{1,k}] \le R\left(1 - \left(1 - Const\, k^{-\left(\sqrt{p}+\frac{\epsilon}{2}\right)^2}\right)^{M_{N_k}}\right) \le Const' k^{-\left(\sqrt{p}+\frac{\epsilon}{2}\right)^2} \times k^p \tag{4.19}
$$

which goes to zero with $k \uparrow \infty$ and proves the lemma. $\diamondsuit$

To treat both the independent and branching case at the same time it is convenient to denote by B_k the whole probability space in the independent case. Now, using the independence of the events $A_{1,k} \cap A_{2,k} \cap B_k$ over k, by Borel Cantelli, (4.11) is implied by $\sum_{k=1}^{\infty} I\!P\left[A_{1,k} \cap A_{2,k} \cap B_k\right] =$

∞. Using Lemma IV.3 and recalling $I\!P[B_k] = k^{o(k)}$ from (3.44), the divergence of this series is implied by

$$\sum_{k=1}^{\infty} k^{-o(1)} I\!P[A_{2,k}(\epsilon,\rho)] = \infty \tag{4.20}$$

for all arbitrarily small $\epsilon > 0$ and $\rho > 1$ large enough for any function $o(1)$. This divergence follows as in the proof of Strassen's theorem by a standard estimate using Schilder's theorem, which we however include for the convenience of the reader:

Drop the superscript ν and denote by X_t a Brownian motion. Write

$$\begin{aligned} & I\!P\left[\left|\frac{X_N - X_{N_{k-1}}}{l_{N_k}} - \phi\left(\frac{N}{N_k}\right)\right| < \epsilon \quad \text{for all} \quad N_{k-1} \le N \le N_k\right] \\ & \ge I\!P\left[\left|\frac{X_N}{l_{N_k}} - \phi\left(\frac{N}{N_k}\right)\right| < \frac{\epsilon}{2} \quad \text{for all} \quad N_{k-1} \le N \le N_k\right] - I\!P\left[\left|\frac{X_{N_{k-1}}}{l_{N_k}}\right| \ge \frac{\epsilon}{2}\right] \end{aligned} \tag{4.21}$$

But $I\!P\left[\left|\frac{X_{N_{k-1}}}{l_{N_k}}\right| \ge \frac{\epsilon}{2}\right] \sim k^{-\frac{\epsilon^2\rho^2}{4}}$ for large k. This, for ρ large enough, is summable and thus can be ignored. For the first term we have

$$I\!P\left[\left|\frac{X_N}{l_{N_k}} - \phi\left(\frac{N}{N_k}\right)\right| < \frac{\epsilon}{2} \quad \text{for all} \quad N_{k-1} \le N \le N_k\right] \ge I\!P\left[\sup_{0\le t\le 1}\left|\frac{X_t}{d_{N_k}} - \phi\left(t\right)\right| < \frac{\epsilon}{2}\right] \tag{4.22}$$

We are now employing Schilder's theorem which states that $\frac{X_t}{l}$ with $l \uparrow \infty$ obeys a LDP with rate function $\phi \mapsto \frac{\|\phi\|^2_{H^1}}{2}$ and rate l^2. The large deviation lower bound then gives that, for all $\bar\epsilon_1 > 0$, we have, for l large enough that $I\!P\left[\sup_{0\le t\le 1}\left|\frac{X_t}{l} - \phi\left(t\right)\right| < \frac{\epsilon}{2}\right] \ge \exp\left(-\frac{l^2(1+\bar\epsilon_1)\|\phi\|^2_{H^1}}{2}\right)$. Using that $\|\phi\|_{H^1} < 1$, by making $\bar\epsilon_1 > 0$ sufficiently small we can thus find a small $\bar\epsilon_2 > 0$ such that

$$I\!P\left[\left|\frac{X_N}{l_{N_k}} - \phi\left(\frac{N}{N_k}\right)\right| < \frac{\epsilon}{2} \quad \text{for all} \quad N_{k-1} \le N \le N_k\right] \ge k^{-1+\bar\epsilon_2} \tag{4.23}$$

for k large enough which proves (4.15). $\diamondsuit\diamondsuit$

It remains to give the proofs of the Lemma IV.1 and IV.2 used in Part I of the proof of the theorem.

Proof of Lemma IV.1: For arbitrary $\epsilon_1, \epsilon_2 > 0$ we write

$$\begin{aligned} & \frac{1}{N}\#\{n : 1 \le n \le N \big| X_n > \sqrt{p} l_n\} \le \epsilon_1 + \frac{1}{N}\#\{n : \epsilon_1 N \le n \le N \big| \frac{X_n}{l_N} > \sqrt{p}\frac{l_n}{l_N}\} \\ & \le \epsilon_1 + \frac{1}{N}\#\{n : \epsilon_1 N \le n \le N \big| \frac{X_n}{l_N} > (\sqrt{p} - \epsilon_2)\sqrt{\frac{n}{N}}\} \end{aligned} \tag{4.24}$$

where the last line holds for large enough N. Next we have to use the 'continuity' of the paths of X_N. Recall to this end that for $\phi \in H^1$ we have from the Schwartz inequality, for all $a \le b$, the Hoelder estimate $|\phi(b)-\phi(a)| = \left|\int_a^b \dot\phi(x)\times 1 dx\right| \leq \sqrt{b-a}$. From Strassen (i) follows then that, for any $\epsilon_3>0$, for N larger than some $N(\epsilon_3)$ we have $\left|\frac{X_{tN}}{l_N} - \frac{X_{t'N}}{l_N}\right| \leq \epsilon_3 + \sqrt{t-t'}$ for all $0\leq t\leq t'\leq 1$. In particular we have for all $\epsilon_3>0$, for N sufficiently large, that

$$\begin{aligned} &\frac{X_n}{l_N} > (\sqrt{p}-\epsilon_2)\sqrt{\frac{n}{N}} \quad \text{implies} \\ &\frac{X_{tN}}{l_N} > (\sqrt{p}-\epsilon_2)\sqrt{t} - \left(\epsilon_3 + \frac{1}{\sqrt{N}}\right) \geq \left(\sqrt{p}-\epsilon_2 - \left(\epsilon_3+\frac{1}{\sqrt{N}}\right)\frac{1}{\sqrt{\epsilon_1}}\right)\sqrt{t} \end{aligned} \tag{4.25}$$

$\forall t\in[\frac{n-1}{N},\frac{n}{N}]$ on all intervals with $n\geq \epsilon_1 N$. For fixed ϵ_1, we choose ϵ_3 small enough and replace the sum by an integral. Thus we get from (4.24) that, for any $\epsilon>0$,

$$\limsup_{N\uparrow\infty}\frac{1}{N}\#\{n: 1\leq n\leq N \big| X_n > \sqrt{p} l_n\} \leq \epsilon_1 + \limsup_{N\uparrow\infty}\int_{\epsilon_1}^1 1_{\frac{X_{tN}}{l_N}>(\sqrt{p}-\epsilon)\sqrt{t}} dt \tag{4.26}$$

But since ϵ_1 was arbitrary the lemma follows.$\diamondsuit$

Proof of Lemma IV.2: To bypass the noncontinuity of I_p we define a regularization I_{p_1,p_2}: Take an increasing continuous function $I\!R \ni x\mapsto \gamma(x)\geq 0$ s.t. $\gamma(x)=0$ for $x\leq 0$ and $\gamma(x)=1$ for $x\geq 1$ and put, for $p_1<p_2$,

$$I_{p_1,p_2}(\phi) = \int_0^1 \gamma\left(\frac{1}{\sqrt{p_2}-\sqrt{p_1}}\left(\frac{\phi(x)}{\sqrt{x}}-\sqrt{p_1}\right)\right)dx \tag{4.27}$$

Then we have

$$I_{p_2}(\phi)\leq I_{p_1,p_2}(\phi)\leq I_{p_1}(\phi) \tag{4.28}$$

for all measurable ϕ on $[0,1]$ and, by dominated convergence, $\phi\to I_{p_1,p_2}(\phi)$ in continuous in sup-norm.

Thus we have, for any $\epsilon>0$, by Strassen (i),

$$\limsup_{N\uparrow\infty} I_p\left(\frac{X_{\cdot N}}{l_N}\right) \leq \limsup_{N\uparrow\infty} I_{p-\epsilon,p}\left(\frac{X_{\cdot N}}{l_N}\right) \leq \sup_{\phi\in K} I_{p-\epsilon,p}(\phi) \tag{4.29}$$

Let us denote by K_L the space of functions in K consisting of finitely many linear pieces. Then, for all $\phi\in K$ for all $\epsilon>0$ there exists a $\tilde\phi\in K_L$ s.t. $\|\phi-\tilde\phi\|_\infty\leq\epsilon$. (One may see this explicitly using the Hoelder continuity of the functions in K.) Thus, we may further write

$$\sup_{\phi\in K} I_{p-\epsilon,p}(\phi) = \sup_{\tilde\phi\in K_L} I_{p-\epsilon,p}\left(\tilde\phi\right) \leq \sup_{\tilde\phi\in K_L} I_{p-\epsilon}\left(\tilde\phi\right) \tag{4.30}$$

We will show now

$$\sup_{\tilde\phi\in K_L} I_p\left(\tilde\phi\right)\leq 1-x_p \tag{4.31}$$

Then (4.29)-(4.31) gives us the upper bound $\limsup_{N\uparrow\infty} I_p\left(\frac{X_{\cdot N}}{l_N}\right) \leq 1 - x_{p-\epsilon}$ The analogous lower bound follows just from

$$\begin{aligned}&\limsup_{N\uparrow\infty} I_p\left(\frac{X_{\cdot N}}{l_N}\right) \geq \limsup_{N\uparrow\infty} I_{p,p+\epsilon}\left(\frac{X_{\cdot N}}{l_N}\right) = \sup_{\phi\in K} I_{p,p+\epsilon}(\phi)\\&\geq \sup_{\phi\in K} I_{p+\epsilon}(\phi) \geq I_{p+\epsilon}\left(\phi_{p+\epsilon,x_{p+\epsilon}}\right) = 1 - x_{p+\epsilon}\end{aligned} \tag{4.32}$$

The last two inequalities, by continuity of the function $p \mapsto x_p$ then prove $\limsup_{N\uparrow\infty} I_p\left(\frac{X_{\cdot N}}{l_N}\right) = 1 - x_p$.

To show (4.31) is sufficient to realize that for all $\tilde{\phi} \in K_L$ the function $\phi_{p,1-I_p(\tilde{\phi})}$ has clearly $I_p\left(\tilde{\phi}\right) = I_p\left(\tilde{\phi}_{p,1-I_p(\tilde{\phi})}\right)$ but $\|\phi_{p,1-I_p(\tilde{\phi})}\|_{H^1} \leq \|\tilde{\phi}\|_{H^1}$. To prove the latter statement we give three steps of modifications of the function $\tilde{\phi}$ that respect the functional I_p but lower the energy norm:

Each piecewise linear $\tilde{\phi}$ has only finitely many intersection points with the curve $\sqrt{px}$. In the first step we replace those parts of $\tilde{\phi}$ that lie above the curve $y = \sqrt{px}$ between two intersection points by parts that lie *on* the curve. In the second step we replace the parts of $\tilde{\phi}$ that are below $\sqrt{px}$ between two intersection points by straight lines that join these points. Let us write $\bar{\phi}$ for the resulting curve. Then we have $I_p(\tilde{\phi}) = I_p(\bar{\phi})$, but $\|\bar{\phi}\|_{H^1} \leq \|\tilde{\phi}\|_{H^1}$.

Indeed, to see this formally consider (for the first step) two neighboring intersection points $a < b$, i.e. $\tilde{\phi}(a) = \sqrt{pa}$, $\tilde{\phi}(b) = \sqrt{pb}$ and write $\eta(x) = \tilde{\phi}(x) - \sqrt{px} \geq 0$ for $a \leq x \leq b$. Then a partial integration and convexity give

$$\int_a^b \dot{\tilde{\phi}}(x)^2 dx - \int_a^b \left(\dot{\sqrt{px}}\right)^2 dx = \int_a^b \dot{\eta}(x)^2 dx + 2\int_a^b \eta(x)\left(-\ddot{\sqrt{px}}\right) dx \geq 0 \tag{4.33}$$

For the second step note that Schwartz on the integral $\int_a^b \dot{\tilde{\phi}}(x) \times 1 dx$ gives

$$\int_a^b \dot{\tilde{\phi}}(x)^2 dx - \left(\frac{\tilde{\phi}(b) - \tilde{\phi}(a)}{b-a}\right)^2 (b-a) \geq 0 \tag{4.34}$$

For the final step of the construction, denote by I_l all the intervals where $\bar{\phi}$ equals a straight line. Write $x_0 := \inf_{x>0}\{x, \bar{\phi}(x) = \sqrt{px}\} > 0$. (If such a point does not exist there is nothing to show.) Notice that $\int_0^{x_0} \dot{\bar{\phi}}(x)^2 dx = p$ is independent of x_0. Denote by J_m all the intervals where $\bar{\phi}$ equals the parabol. Notice that $\sum_m \int_{J_m} \dot{\bar{\phi}}(x)^2 dx \geq \int_{1-\sum_m |J_m|}^1 \left(\dot{\sqrt{px}}\right)^2 dx$ since the derivative of the squareroot is decreasing. But with this we have

$$\begin{aligned}\|\bar{\phi}\|_{H^1}^2 &= \sum_l \int_{I_l} \dot{\bar{\phi}}(x)^2 dx + \sum_m \int_{J_m} \dot{\bar{\phi}}(x)^2 dx\\&\geq \int_0^{x_0} \dot{\bar{\phi}}(x)^2 dx + \int_{1-\sum_m |J_m|}^1 \left(\dot{\sqrt{px}}\right)^2 dx\\&= \|\phi_{p,\sum_m |J_m|}\|_{H^1}^2 = \|\phi_{p,1-I_p(\tilde{\phi})}\|_{H^1}^2\end{aligned} \tag{4.35}$$

which proves (4.31).

The same arguments also prove $\sup_{\phi\in K} I_p(\phi) = 1 - x_p$.$\diamondsuit$

References

[AN] K.B.Athreya, P.Ney, Branching Processes, Springer, New York, Heidelberg, Berlin (1972)

[AH] S.Asmussen, H.Hering, Branching Processes, Birkhäuser, Boston, Basel, Stuttgart (1983)

[Be] S.M.Berman, Sojourns and extremes of stochastic processes, Wadsworth and Brooks/Cole, Pacific Grove, Calif. (1992)

[AW] M.Aizenman, J.Wehr, Rounding Effects of Quenched Randomness on First-Order Phase Transitions, Comm.Math.Phys **130**, 489-528 (1990)

[BG1] A.Bovier, V.Gayrard, The retrieval phase of the Hopfield model: A rigorous analysis of the overlap distribution, Probab. Theory Related Fields **107** (1997), no. 1, 61–98

[BG2] A.Bovier, V.Gayrard, Hopfield models as a generalized random mean field model, WIAS preprint 253, Berlin (1996), in 'Mathematics of spin glasses and neural networks', A.Bovier and P.Picco, eds., 'Progress in Probability', Birkhäuser (1997)

[BGP] A.Bovier, V.Gayrard, P.Picco, Gibbs states of the Hopfield model in the regime of perfect memory, Prob.Theor.Rel.Fields **100**,329-363, (1994)

Gibbs states of the Hopfield model with extensively many patterns, J.Stat.Phys. **79**, 395-414 (1995)

[BG3] A.Bovier, V.Gayrard, Metastates in the Hopfield model in the replica symmetric regime, WIAS preprint, Berlin (1997)

[BoK] A.Bovier, C.Külske, A rigorous renormalization group method for interfaces in random media, Rev.Math.Phys. **6** No.3, 413-496 (1994)

[BS] E.Bolthausen, A.-S.Sznitman, On Ruelle's Probability Cascades and an abstract cavity method, preprint, Zürich (1997)

[De] B.Derrida, Phys.Rev.Lett.**45**, 79-82 (1980), Phys.Rev.B **24**, 2613-2626 (1981) J.Phys.Lett. **46**, L401, L407 (1985)

[DeE] B.Derrida and E.Gardner, J.Phys.C, **19**, 5783 (1986)

[DeSp] B.Derrida, H.Spohn, Polymers on Disordered Trees, Spin Glasses, and Traveling Waves, J.Stat.Phys **51**, 817-840 (1988)

[DS] J.-D.Deuschel, D.W.Stroock, Large Deviations, Pure and Applied Mathematics, vol. 137, Academic Press, Boston (1989)

[Du] R.Durett, Maxima of Branching Random Walks vs. independent Random Walks, Stoch. Proc. Appl. **9** (1979) 117-135

[DZ] A.Dembo, O.Zeitouni, Large Deviations Techniques, Jones and Bartlett, Boston, London (1993)

[El] R.S.Ellis, Entropy, Large Deviations, and Statistical Mechanics, Grundlehren der mathematischen Wissenschaften, vol.271, Springer, New York

[Geo] H.O. Georgii, Gibbs measures and phase transitions, Studies in mathematics, vol. 9 (de Gruyter, Berlin, New York, 1988)

[GK] V.Gayrard, C.Külske, The empirical Metastate in the Hopfield model with external magnetic field, in preparation

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

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

[K1] Ch. Külske, Ph.D. Thesis, Ruhr-Universität Bochum (1993)

[K2] Ch. Külske, Metastates in Disordered Mean Field Models: Random Field and Hopfield Models, J.Stat.Phys **88** Issue 5/6 (1997)

[K3] Ch. Külske, Limiting Behavior of Random Gibbs Measures: Metastates in some disordered mean field models, to appear in 'Mathematics of spin glasses and neural networks', A.Bovier and P.Picco, eds., 'Progress in Probability', Birkhäuser (1997)

[LLR] M.R.Leadbetter, G.Lindgren, H.Rootzen, Extremes and related properties of random sequences and processes, Springer Series in Statistics, Springer Verlag, New York, Heidelberg, Berlin (1983)

[LT] M.Ledoux, M.Talagrand, Probability in Banach Spaces, Springer Verlag (1991)

[Nag] S.V.Nagaev, Large deviations of sums of independent random variables, Ann.Prob. **7**, No. 5, 745-789 (1979)

[N] C.M.Newman, Topics in Disordered Systems, to appear in 'ETH Lecture Notes Series', Birkhäuser (1997)

[NS1] C.M.Newman, D.L.Stein, Chaotic Size Dependence in Spin Glasses, in Cellular Automata and Cooperative Systems, Boccara, Goles, Martinez, Picco (Eds.), Nato ASI Series C Vol. 396, Kluwer, Dordrecht (1993)

[NS2] C.M.Newman, D.L.Stein, Non-Mean Field Behavior in realistic Spin glasses, Phys.Rev.Lett. **76**, No 3, 515 (1996)

[NS3] C.M.Newman, D.L.Stein, Spatial Inhomogeneity and thermodynamic chaos, Phys.Rev.Lett. **76**, No 25, 4821 (1996)

[NS4] C.M.Newman, D.L.Stein, Thermodynamic chaos and the structure of short range spin glasses, in 'Mathematics of spin glasses and neural networks', A.Bovier and P.Picco, eds.,

[OP] E.Olivieri, P.Picco, On the existence of thermodynamics for the random energy model, Commun.Math.Phys **96**, 125-144 (1984)

[Ru] D.Ruelle, A mathematical reformulation of Derrida's REM and GREM, Comm. Math.Phys **108**, 225-239 (1987)

[T] M.Talagrand, Rigorous results for the Hopfield model with many patterns, preprint 1996, to appear in Prob.Theor.Rel.Fields