The parabolic Anderson model on Galton-Watson trees

Wang, D.

The Parabolic Anderson Model on Galton-Watson Trees

Daoyi Wang

The Parabolic Anderson Model on Galton-Watson Trees

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 28 mei 2024
klokke 13:45 uur

door

Daoyi Wang
geboren te Beijing, China
in 1996

Promotor:

Prof. dr. W.Th.F. den Hollander

Promotiecommissie:

Prof. dr. ir. G.L.A. Derks
Prof. dr. M.R.H. Mandjes
Prof. dr. W. König (Weierstrass Institut & Technical University Berlin)
Dr. R. Soares dos Santos (Federal University of Minas Gerais, Belo Horizonte, Brazil)
Prof. dr. P. Sousi (University of Cambridge, United Kingdom)

The research in this thesis was supported by the Netherlands Organisation for Scientific Research through NWO Gravitation Grant NETWORKS-024.002.003.

Contents

CHAPTER 1
Introduction

§1.1 Introduction

§1.1.1 Definitions, intermittency and problems

The parabolic Anderson model (abbreviated PAM) is the Cauchy problem for the heat equation with random potential. It is given by the parabolic differential equation

$$\begin{array}{lcll} \frac{\partial}{\partial t}u(x,t) & = & \kappa\Delta u(x,t) + \xi(x)u(x,t), & t>0,\, x\in\mathscr{X}, \\ u(x,0) & = & f(x), & x\in\mathscr{X}, \end{array} \tag{1.1}$$

where $\kappa > 0$ is the diffusion constant, $\mathscr{X}$ an ambient space, Δ the Laplace operator acting on functions on $\mathscr{X}$, ξ the random potential, and f the initial condition. The ambient space $\mathscr{X}$ can be taken to be continuous, and/or the potential taken to be time dependent. However, we shall restrict to discrete spaces (graphs) and a static potential in this thesis. See [34], [18], [20] for more background on $\mathbb{R}^d$, and [10], [17] for works on time-dependent potentials.

Under mild conditions on the initial condition and potential, the unique non-negative solution to (1.1) admits the the well-known *Feynman-Kac* representation

$$u(x,t) = \mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} f(X_t)\right], \tag{1.2}$$

where $X = (X_t)_{t\geq 0}$ is the Markov process with generator $\kappa\Delta$, i.e. a continuous-time random walk with jump rate κ to each neighbour. In addition, $\mathbb{P}_x$ denotes probability with respect to X given that $X_0 = x$. Since the equations

$$\frac{\partial}{\partial t}v(x,t) = \kappa\Delta v(x,t), \qquad \frac{\partial}{\partial t}w(x,t) = \xi(x)w(x,t),$$

have solutions

$$v(x,t) = \mathbb{E}_x(f(X_t)), \qquad w(x,t) = \mathrm{e}^{\int_0^t \xi(x)\mathrm{d}s} f(x),$$

the solution in (1.2) is heuristically plausible.

One interpretation of the PAM concerns the following non-interacting particle system. At time t and site x, particles are killed with rate $\xi^-(x)$ or are split into two with rate $\xi^+(x)$, where $\xi(x) = \xi^+(x) - \xi^-(x)$. At the same time, each particle jumps independently with $\kappa\Delta$ as generator. The solution $u(x,t)$ is the average number of particles (or *mass*) at site x and time t, with initial condition $f(x)$ that is integer valued. The parabolic problem is a classical model of a system evolving in an inhomogeneous random medium and has wide ranging applications. By considering the particles above as organisms, the PAM has an application in population dynamics. The PAM is also related to other physical problems, including the Burger's equation, and the advection-convection equation for a temperature field. See [10] for more details and more applications.

Another contributing factor to the popularity of PAM is that the model exhibits *intermittency*, rather than being spatially homogeneous. This is related to Anderson localisation observed in the associated Anderson Hamiltonian $\kappa\Delta + \xi$ studied in [1], where the eigenvectors of the operator are concentrated around a point and decay exponentially. Heuristically, intermittency is the related phenomenon where the majority of the total mass

$$U(t) = \sum_{x\in\mathscr{X}} u(x,t)$$

is concentrated in specific regions of $\mathscr{X}$ as $t \to \infty$ known as *intermittent islands*. See [29][Section 2.2.4] for further details on how Anderson localisation relates to intermittency. The natural questions regarding intermittency are therefore:

(a) What is the asymptotics of $U(t)$ as $t \to \infty$?

(b) Where are the intermittent island(s) situated in $\mathscr{X}$? How does ξ affect the intermittent islands?

(c) What do ξ and $u(t,x)$ look like on the intermittent islands?

The above intuition is too imprecise to provide a rigorous definition of intermittency. In the literature, intermittency is defined using the moments of the total mass

$$m_p(t) = \frac{1}{t}\log\langle U(t)^p\rangle, \quad p \in \mathbb{N}, \tag{1.3}$$

where $\langle\cdot\rangle$ denotes expectation with respect to the potential ξ. Here $m_p(t)$ is known as the p^{th} *Lyapunov exponent*, and the model is said to be intermittent if

$$m_1 < \frac{m_2}{2} < \cdots < \frac{m_p}{p} < \cdots \tag{1.4}$$

as $t \to \infty$. We refer to [29][Section 1.4] for a more complete picture as to why this definition captures the phenomenon described before. Loosely speaking, (1.4) says that moments grow faster than previous moments by an exponential factor. This can only happen if large amounts of mass concentrate on small regions of $\mathscr{X}$, so that the main contribution to the moments comes from regions where the solution takes large values. In other words, the overwhelming part of the mass has to be concentrated in these regions for large t. This means that the definition in terms of moments is consistent with the heuristics given before.

Understanding the asymptotics of $m_p(t)$ is often the first step in the analysis of the PAM, as it quantifies the intermittency phenomenon and resolves the first question above. There is a distinction between the *quenched* total mass and the *annealed* total mass, i.e. the total mass taken almost surely with respect to the potential or averaged over the potential. In (1.3) and (1.4) the total mass was taken annealed, although the quenched asymptotics of $U(t)$ is also relevant. For the other questions regarding the intermittent islands, further analysis is required. The analysis in this case is done through a *characteristic variational formula* that optimises the large deviation probability of realisations of the potential and the principal eigenvalue of the Anderson

Hamiltonian on the intermittent islands. The variational formula also serves as the second order correction term in the asymptotics of both the quenched and the annealed total mass. As will become clear later, this is because potential profiles that are 'close' to the minimiser of the variational formula contribute most to the total mass. There is again a distinction between the annealed and the quenched variational formulae despite their similarities. In essence they approach the problem from different starting points and hence differ in their interpretation and are also formulated slightly differently.

§1.1.2 The parabolic Anderson model on a lattice

The study of the PAM on discrete spaces originated on the lattice $\mathbb{Z}^d$, $d \geq 1$, through the seminal works [24] and [25] by Gärtner and Molchanov. In particular, they showed that the (non-negative) solution to (1.1) exists and is unique under mild conditions, namely non-negative bounded initial conditions and that the potential does not percolate from below. In addition, the solution has the simple representation given by (1.2).

Henceforward, we shall assume that $f(x) = \delta_0(x)$, the potential ξ to be independent and identically distributed, and $\kappa = 1$. In this case, the PAM in (1.1) reads

$$\begin{array}{rcll} \frac{\partial}{\partial t} u(x,t) & = & \Delta u(x,t) + \xi(x) u(x,t), & t > 0,\, x \in \mathbb{Z}^d, \\ u(x,0) & = & \delta_0(x), & x \in \mathbb{Z}^d, \end{array} \tag{1.5}$$

where the Laplacian is given by

$$(\Delta f)(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |y-x|=1}} [f(y) - f(x)]. \tag{1.6}$$

The solution is then given by

$$u(x,t) = \mathbb{E}_0 \left[\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \mathbb{1}\{X_t = x\} \right], \tag{1.7}$$

where we have used time reversal. Taking a point mass at 0 is the standard choice of initial condition, as it allows for the total mass $U(t)$ to be expressed as

$$U(t) = \sum_{x \in \mathbb{Z}^d} u(x,t) = \mathbb{E}_0 \left[\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \right]. \tag{1.8}$$

The result for arbitrary κ can be obtained from those for $\kappa = 1$, by scaling time.

Since [24] and [25], the PAM on $\mathbb{Z}^d$ has been extensively studied and is now well understood. We give a brief summary of the main results under the above assumptions. We borrow from [29].

Suppose that the cumulant generating function of the potential

$$H(t) = \log\langle e^{t\xi(0)}\rangle$$

belongs to the *de Haan* class of regular functions, i.e. that

$$\lim_{t\to\infty} \frac{H(ct) - cH(t)}{g(t)} = \hat{H}(t) \not\equiv 0 \tag{1.9}$$

for all $c \neq 1$, and that $\lim_{t\to\infty} g(t)/t$ exists. It is known from [35] that there are 4 qualitatively different regimes (known as *universality classes*) for the behaviour of the total mass, as well as the size and number of intermittent islands. Precisely which regime the asymptotics belong to depends on the the limit $\hat{H}$, which is determined by the upper tail of the distribution of $\xi(0)$. The 4 classes are:

(a) *Double exponential.* This is the regime that was originally studied in [25]. Potentials in this class have upper tails approximately equal to

$$\mathsf{P}(\xi(0) > u) = e^{-e^{u/\varrho}}, \qquad u \in \mathbb{R}, \tag{1.10}$$

for a parameter $\varrho \in (0,\infty)$. The unique feature of this class is that the intermittent islands do not shrink nor grow with time, while also being non-trivial. In this class, the moments of the total mass satisfy, as $t \to \infty$,

$$\langle U(t)^p\rangle = \exp[H(pt) - pt\chi(\varrho) + o(t)], \tag{1.11}$$

where $\chi(\varrho)$ is the annealed characteristic variational formula given by

$$\chi(\varrho) := \inf_{p\in\mathcal{P}(\mathbb{Z}^d)} [I(p) + \varrho J(p)], \qquad \varrho \in (0,\infty), \tag{1.12}$$

with

$$I(p) = \sum_{x,y\in\mathbb{Z}^d : x\sim y} \left(\sqrt{p(x)} - \sqrt{p(y)}\right)^2, \qquad J(p) = -\sum_{x\in\mathbb{Z}^d} p(x)\log p(x).$$

In this regime, $H(t) = \varrho t\log(\varrho t) - \varrho t + o(t)$ as $t \to \infty$. Furthermore, the variational formula was studied in [23] and it is known that the minimiser decomposes into the d-fold tensor product of the minimiser for the case $d = 1$. The minimiser of the one-dimensional problem is unimodal (the maximum can be taken at 0 due to translation invariance), and decreasing with the distance to the maximum. Furthermore, it is known that the minimiser is unique up to translations for ϱ sufficiently large.

For the quenched case, [25] showed that as, $t \to \infty$,

$$\frac{1}{t}\log U(t) = \frac{H(d\log t)}{d\log t} - \chi_\star(\varrho) + o(1), \tag{1.13}$$

where $\chi_\star(\varrho)$ is the quenched variational formula. In this case, the annealed and the quenched variational formulae are equal and given by (1.12). However, we emphasise that the two have different interpretations despite them being equal in this case.

(b) *Single point.* This class comprises potentials with heavier tails and corresponds to the double exponential class with $\varrho = \infty$. This class was also analysed in [25], as a limiting case for the double exponential class. Examples of distributions in this class include the normal distribution, the Weibull distribution with

$$\mathsf{P}(\xi(0) > u) = \mathrm{e}^{-Cu^\alpha}, \qquad \alpha > 1,$$

where the condition $\alpha > 1$ ensures all exponential moments are finite. In this class, the results are consistent with the double exponential class with $\varrho = \infty$. The moments of the total mass satisfy, as $t \to \infty$,

$$\langle U(t)^p \rangle = \exp[H(pt) - pt\chi(\infty) + o(t)].$$

The variational formula is also the same as (1.12) at $\varrho = \infty$, in which case the minimiser is taken at a single point and the value is equal to $2d$. Therefore the intermittent island is a single site.

(c) *Bounded Potential.* The classical example is

$$\mathsf{P}(\xi(0) = 0) = p, \qquad \mathsf{P}(\xi(0) = -\infty) = 1 - p,$$

for $p \in [0, 1]$ which corresponds to Bernoulli traps. In this case the total mass can be interpreted as the survival probability of the random walk. In this class, the diameter of the intermittent islands diverges to ∞ at least as fast as some power of t as $t \to \infty$. Furthermore, it is known from [32] that the minimiser of the variational formula exists, is unique and is compactly supported on a ball. The quenched and annealed asymptotics for the total mass are complicated and involve the diameter of the intermittent islands. See [7] for details.

(d) *Almost bounded.* This class lies between the bounded and the double exponential classes in terms of the thickness of the tails. An example is obtained after ϱ is replaced by a regular function $\varrho(u)$ that tends to 0 as $u \to \infty$. Here, the diameter of the intermittent islands diverges to ∞, but slower than any power of t as $t \to \infty$. See [21] for further details.

§1.1.3 The parabolic Anderson model on random graphs

On a general graph $G = (V, E)$, the PAM is defined by

$$\begin{aligned} \tfrac{\partial}{\partial t} u(x,t) &= \Delta u(x,t) + \xi(x) u(x,t), \quad && t > 0,\, x \in V, \\ u(x,0) &= \delta_{\mathcal{O}}(x), && x \in V. \end{aligned} \tag{1.14}$$

The Laplacian is given by

$$(\Delta f)(x) = \sum_{y \sim x} [f(y) - f(x)], \tag{1.15}$$

where the sum runs over all neighbours of x, i.e. $y \sim x$ means that $\{x, y\} \in E$. Note that this is precisely equal to the graph Laplacian matrix on G given by $D - A$, where D is the degree matrix and A is the adjacency matrix. The solution is given by

$$u(x,t) = \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\,\mathbb{1}\{X_t = x\}\right], \tag{1.16}$$

where the random walk has jump rate 1 along the edges E, and $\mathcal{O} \in V$. In the case of a rooted graph, $\mathcal{O}$ is chosen to be the root.

There has been little work in the area of deterministic graphs. The most impactful works in this direction are [16] and [5]. The former concerned the PAM on a complete graph of size N with an i.i.d. exponentially distributed potential. All moments are eventually infinite as $t \to \infty$ under this potential, hence intermittency was quantified by showing that

$$\frac{u(t,y)}{U(t)} \to 1 \quad a.s. \tag{1.17}$$

as $t \to \infty$ and $N \to \infty$ simultaneously (in any relation), where y is the site with the largest value of ξ. This is to be expected, as the exponential distribution has even thicker tails than the single point class above. Furthermore, it was shown that there is a phase transition depending on whether $t/\log N \to 0$ or not, i.e. whether time grows faster than the logarithm of the size of the graph or not. In the case of the hypercube $\{0,1\}^n$, the potential was taken to be i.i.d. normally distributed with mean 0 and variance n. In this case, the PAM exhibits a similar behaviour as on the complete graph: as $t, n \to \infty$, most of the mass is concentrated on one site in the same sense as (1.17). This site is again where the potential is maximal. Moreover, there is again a phase transition, except that the critical time scale is $n \log n$. The PAM on the hypercube has biological applications, as it can model the occurrence of mutants in a large population. The hypercube can be interpreted as the set of all possible gene sequences of 2 alleles (the gene pool), and the potential can be seen as a random fitness landscape.

For random graphs, even less is known. The study of the PAM on random graphs was initiated in [12] on the Galton-Watson tree $(G, V, \mathcal{O})$ with offspring distribution bounded below by a constant $d \geq 2$, and with expectation ϑ. Consequently, the Galton-Watson trees considered are infinite with probability 1. For the double exponential class of potentials, the almost sure asymptotics of the total mass was shown to be

$$\frac{1}{t}\log U(t) = \varrho \log\left(\frac{\varrho t \vartheta}{\log\log t}\right) - \varrho - \chi(\varrho) + o(1), \tag{1.18}$$

almost surely with respect to the potential and the random tree. The argument roughly follows the ideas from the lattice. The total mass is estimated by considering random walk paths that very quickly reach an intermittent island with high values of the potential and subsequently remain there. The probabilistic cost to travel to the island and the maximum of the potential on the island were calculated and optimised over. The first term in (1.18) corresponds to the maximal value of the potential in the island, which is obtained through extreme-value analysis. Since neighbourhoods grow exponentially on the Galton-Watson tree, the first term is larger than (1.13) on $\mathbb{Z}^d$, where neighbourhoods grow polynomially, because there are more sites for the potential to take larger values. It was also shown that the intermittent islands are equal to an infinite regular tree with minimal degree when ϱ is sufficiently large. This was done by analysing the variational formula

$$\chi(\varrho) := \inf_{p \in \mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0, \infty), \tag{1.19}$$

which is the same as (1.12) but now on V, and showing that it is non-increasing under successive 'trimming' of excess branches.

More recently, in [2], the PAM with i.i.d. Pareto distributed potentials was analysed on a critical Galton-Watson tree conditioned on survival. The authors followed the arguments presented in [28] for the lattice and obtained that the asymptotics of the total mass satisfies

$$\lim_{\lambda \to \infty} \liminf_{t \to \infty} \mathsf{P}(\log U(t) \in [\lambda^{-1} t a(t), \lambda t a(t)]) = 1,$$

where λ is a parameter coming from the offspring distribution, and $a(t)$ is the order of the maximal value of the potential in a growing ball that was explicitly identified. Furthermore, it was shown that the solution localises on 2 points almost surely, and on a single point with high probability. Formally, there exist processes V_t and W_t on the vertices such that as $t \to \infty$,

$$\frac{u(t, V_t) + u(t, W_t)}{U(t)} \to 1$$

almost surely with respect to the tree and the potential, and

$$\frac{u(t, V_t)}{U(t)} \to 1$$

in probability with respect to the tree and the potential.

The main difficulty of moving to random graphs is that they complicate the analysis by adding another layer of randomness. Furthermore, many properties of the lattice, including degree regularity and translation invariance that previous works relied on, are no longer present. Another difficulty is the exponentially growing nature of the Galton-Watson tree, which means that any subtree has a boundary comparable to its volume in size. Consequently, bounds that rely on reflection techniques that fold $\mathbb{Z}^d$ into a box of an appropriate size at the cost of a negligible boundary term are no longer applicable. Even though the Galton-Watson tree is not regular, the degrees are i.i.d. so that there is still some regularity. Hence, the Galton-Watson tree is the natural first step towards extending the results beyond the lattice. Furthermore, random graphs that are locally tree-like may share the the same behaviour as the Galton-Watson tree. An example is the *configuration model*, as shown in [12].

§1.1.4 Overview of the results

This thesis aims to contribute to the PAM on random trees. In this section we give an overview of the main results. In all cases, we work under the i.i.d. double exponential class of potentials. We also assume that the minimal offspring is at least 2, so that we always have an infinite tree.

Chapter 2

In Chapter 2 we derive the annealed asymptotics of the total mass on a regular tree with degree $d+1$.

Theorem 1.1.1. *[Growth rate of the total mass] For any $d \geq 4$,*

$$\frac{1}{t}\log\langle U(t)\rangle = \varrho\log(\varrho t) - \varrho - \chi(\varrho) + o(1), \qquad t \to \infty.$$

This shows that the asymptotics on the regular tree exactly matches with the lattice given by (1.11). We try to mimic the argument from [25] for the lattice, which involves finding matching upper lower and upper bounds. For the lattice, the lower bound is obtained by approximating the random walk on the infinite lattice by the random walk on box of length R with zero boundary conditions. The upper bound is obtained by periodising the random walk onto a torus of length R. The bounds coincide in the limit $R \to \infty$, which gives the result. The upper bound heavily relies on translation invariance in the periodisation procedure, which is not present on the regular tree. To overcome this, we devise a *novel procedure* to periodise the random walk onto a subtree analogous to the torus for the lattice. This procedure relies on new ideas as well as delicate computations.

Subsequently, we analysed the variational formula on the regular tree. In [12] it was already shown that the minimiser is attained by the regular tree with minimal degree. The theorem below collects its key properties.

Theorem 1.1.2. *[Properties of the variational formula] For any $d \geq 2$ the following hold:*
(a) The infimum in (1.12) *may be restricted to the set*

$$\mathcal{P}^{\downarrow}_{\mathcal{O}}(V) = \{p \in \mathcal{P}(V)\colon \operatorname{argmax} p = \mathcal{O},\ p \text{ is non-increasing along backbones}. \tag{1.20}$$

(b) For every $\varrho \in (0,\infty)$, the infimum in (1.19) *restricted to $\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$ is attained, every minimiser $\bar p$ is such that $\bar p > 0$ on V, and $\partial S_R = \sum_{\partial B_R(\mathcal{O})} \bar p(x)$, $R \in \mathbb{N}_0$, satisfies*

$$\sum_{R\in\mathbb{N}_0} \partial S_R \log(R+1) \leq \frac{d+1}{\varrho},$$

where $B_R(\mathcal{O})$ is the ball of radius R centred at $\mathcal{O}$.
(c) The function $\varrho \mapsto \chi(\varrho)$ is strictly increasing and globally Lipschitz continuous on $(0,\infty)$, with

$$\lim_{\varrho\downarrow 0}\chi(\varrho) = d-1, \qquad \lim_{\varrho\to\infty}\chi(\varrho) = d+1.$$

Chapter 3

We would like to have a result analogous to Theorem 1.1.1 on the Galton-Watson tree. Chapter 3 makes another step in this direction by extending the results of Chapter 2 to a *periodic* Galton-Watson tree, i.e. a Galton-Watson tree up to distance R from the root, which is repeated every R generations. This is a close approximation of the Galton-Watson tree for R large. On this tree, we again derive the asymptotics of the annealed total mass. We require the following assumption on the offspring distribution D.

Assumption 1.A. [Offspring distribution]
There exist $4 \leq d^- \leq d^+ < \infty$ such that $\mathrm{supp}(D) \subseteq [d^-, d^+]$. ■

Theorem 1.1.3. *Subject to Assumption 1.A, as $t \to \infty$,*

$$\frac{1}{t} \log\langle U(t)\rangle = \varrho \log(\varrho t) - \varrho - \chi_{\mathcal{GW}}(\varrho) + o(1)$$

almost surely with respect to the tree.

The proof relies on the novel techniques developed in the previous chapter, but all the technicalities and fine details needed to be carefully adapted. Obtaining the result requires careful navigation with the degrees, which are now random.

Chapter 4

Chapter 4 concerns the aspymptotics of the total mass, but in the quenched setting. This chapter follows the framework that was set up in [12], which required the offspring distribution to be bounded. This goal of this chapter is to identify the weakest condition on the degree distribution under which the arguments can be pushed through. We require the following assumption on the offspring distribution.

Assumption 1.B. [Super-double-exponential tails] There exists a function $f\colon (0,\infty) \to (0,\infty)$ satisfying $\lim_{s\to\infty} f(s) = 0$, $\lim_{s\to\infty} f'(s) = 0$ and $\lim_{s\to\infty} f(s)\log s = \infty$ such that

$$\limsup_{s\to\infty} \mathrm{e}^{-s} \log \mathcal{P}(D > s^{f(s)}) < -2\vartheta.$$

■

The assumption is slightly weaker than $\mathcal{E}\left(\mathrm{e}^{\mathrm{e}^{\mathrm{e}^{aD}}}\right) < \infty$ for some $a < \infty$.

Theorem 1.1.4. *Subject to Assumption 1.B, as $t \to \infty$,*

$$\frac{1}{t} \log U(t) = \varrho \log\left(\frac{\varrho t \vartheta}{\log\log t}\right) - \varrho - \chi(\varrho) + o(1).$$

The asymptotics are the same as in [12], but we have relaxed the assumption to offspring distributions with unbounded support. Obtaining the result requires a careful analysis of the structural properties of the Galton-Watson tree and control on the occurrence of large degrees uniformly in large subtrees. Vertices with large degrees are problematic and Assumption 1.B is the weakest condition needed to control them.

As mentioned above, the existence and uniqueness of the Feynman-Kac solution to PAM on the lattice was proven in [24]. The existence and uniqueness on trees should hold as well, although was never explicitly proven. In this chapter we follow the arguments of [24] and rigorously prove the existence and uniqueness on Galton-Watson trees provided, the offspring distribution has all exponential moments.

Assumption 1.C. [Exponential tails]
$\mathcal{E}\left[\mathrm{e}^{aD}\right] < \infty$ for all $a \in (0, \infty)$. ■

Theorem 1.1.5. *Subject to Assumption 1.C,* (1.14) *has a unique non-negative solution for all most all realisations of the potential and the tree. This solution admits the Feynman-Kac representation in* (1.16)*.*

Chapter 5

Chapter 5 again considers the PAM on a Galton-Watson tree, but with the normalised Laplacian

$$(\Delta f)(x) = \frac{1}{\deg(x)} \sum_{y \sim x} [f(y) - f(x)],$$

instead of (1.15). The aim of this chapter is to investigate the effect of the normalisation of the Laplacian. In the case of the lattice or regular graphs, the choice does not matter as the normalisation is uniform across all the vertices and is therefore simply a constant. This constant can be absorbed into the diffusion constant κ and all results can be easily inferred. In [5] and [16], the degrees approach infinity with the size of the graph and hence the normalised Laplacian was the only viable choice. In the case of inhomogoenous graphs, which random graphs are, the choice of Laplacian plays a role. We again consider the quenched asymptotics of the total mass.

Assumption 1.D. [Exponential tails]
$\mathcal{E}\left[\mathrm{e}^{aD}\right] < \infty$ for all $a \in (0, \infty)$. ■

Theorem 1.1.6. *[Total mass asymptotics] Subject to Assumption* 1.D*,*

$$\frac{1}{t} \log U(t) = \varrho \log\left(\frac{\varrho t \vartheta}{\log\log t}\right) - \varrho - \widetilde{\chi}(\varrho) + o(1), \quad t \to \infty, \tag{1.21}$$

almost surely with respect to the graph and the potential.

In this case, the variational formula is given by

$$\widetilde{\chi}(\varrho) = \inf_{p \in \mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0, \infty), \tag{1.22}$$

with

$$I_E(p) = \sum_{\{x,y\} \in E} \left(\sqrt{\frac{p(x)}{\deg(x)}} - \sqrt{\frac{p(y)}{\deg(y)}}\right)^2, \qquad J_V(p) = -\sum_{x \in V} p(x) \log p(x).$$

Note that I_E is different compared with (1.19). Hence, the leading-order terms in (1.21) and (1.1.4) are the same, however the normalised Laplacian results in a different the second-order term. Furthermore, comparing Assumption 1.A with Assumption 1.D we see that the asymptotics now holds under vastly milder conditions on the offspring distribution. We again follow the framework of [12]. The main challenge is to investigate the spectral properties of the Anderson Hamiltonian $\Delta+\xi$, which are now different.

It was found in [12] that the minimiser of the variational formula on the Galton-Watson tree is equal to the variational formula on the regular tree minimal degrees $\chi_{\mathcal{T}_d}$. We show that, with the normalised Laplacian and the different I_E-function this is again true, and under weaker conditions on ϱ.

Theorem 1.1.7. *[Identification of the minimiser] If* $\varrho \geq \frac{1}{d_{\min}\log(d_{\min}+1)}$, *then* $\widetilde{\chi}(\varrho) = \chi_{\mathcal{T}_d}(\varrho)$.

§1.1.5 Open problems

We end this introduction by stating a few open problems. The ones that are within reach include:

- Extend the techniques from Chapters 2 and 3 to the Galton-Watson tree.
- For the variational formula on the regular tree analysed in Chapter 2, determine whether the minimiser $\bar{p}$ is unique modulo translation, $\bar{p}(x)$ satisfies $\lim_{|x|\to\infty} |x|^{-1}\log\bar{p}(x) = -\infty$, $\bar{p}$ is radially symmetric, and whether $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$ is analytic on $(0,\infty)$.
- Identify the minimiser of the variational formula for all ϱ, and show that it is unique.
- Calculate the asymptotics of $\langle U(t)^p\rangle$ on the Galton-Watson tree and see whether this equal to
$$\exp\{H(pt) - pt\chi(\varrho) + o(t)\},$$
which is the same form as (1.11) on the lattice.

As stated before, very little is known about the PAM on general graphs. More challenging open problems include:

- Can we prove all our results under a different class of potential?
- What can we say about the PAM on other random graph models?

PART I

PARABOLIC ANDERSON MODEL: ANNEALED

CHAPTER 2

Annealed parabolic Anderson model on a regular tree

This chapter is based on the following paper:
F. den Hollander and D. Wang. Annealed parabolic Anderson model on a regular tree. *Markov Process. Related Fields*, 30:105–147, 2024.

Abstract

We study the total mass of the solution to the parabolic Anderson model on a regular tree with an i.i.d. random potential whose marginal distribution is double-exponential. In earlier work we identified two terms in the asymptotic expansion for large time of the total mass under the *quenched law*, i.e., conditional on the realisation of the random potential. In the present paper we do the same for the *annealed law*, i.e., averaged over the random potential. It turns out that the annealed expansion differs from the quenched expansion. The derivation of the annealed expansion is based on a *new approach* to control the local times of the random walk appearing in the Feynman-Kac formula for the total mass. In particular, we condition on the backbone to infinity of the random walk, truncate and periodise the infinite tree relative to the backbone to obtain a random walk on a finite subtree with a specific boundary condition, employ the large deviation principle for the empirical distribution of *Markov renewal processes* on finite graphs, and afterwards let the truncation level tend to infinity to obtain an asymptotically sharp asymptotic expansion.

§2.1 Introduction and main results

Section 2.1.1 provides background and motivation, Section 2.1.2 lists notations, definitions and assumptions, Section 2.1.3 states the main theorems, while Section 2.1.4 places these theorems in their proper context.

§2.1.1 Background and motivation

The *parabolic Anderson model* (PAM) is the Cauchy problem

$$\partial_t u(x,t) = \Delta_{\mathscr{X}} u(x,t) + \xi(x)u(x,t), \qquad t > 0,\, x \in \mathscr{X}, \tag{2.1}$$

where t is time, $\mathscr{X}$ is an ambient space, $\Delta_{\mathscr{X}}$ is a Laplace operator acting on functions on $\mathscr{X}$, and ξ is a random potential on $\mathscr{X}$. Most of the literature considers the setting where $\mathscr{X}$ is either $\mathbb{Z}^d$ or $\mathbb{R}^d$ with $d \geq 1$, starting with the foundational papers [24], [25], [23] and further developed through a long series of follow-up papers (see the monograph [29] and the survey paper [3] for an overview). More recently, other choices for $\mathscr{X}$ have been considered as well:

(I) *Deterministic graphs* (the complete graph [16], the hypercube [5]).

(II) *Random graphs* (the Galton-Watson tree [12], [13], the configuration model [12]).

Much remains open for the latter class.

The main target for the PAM is a description of *intermittency*: for large t the solution $u(\cdot,t)$ of (2.1) concentrates on well-separated regions in $\mathscr{X}$, called *intermittent islands*. Much of the literature focusses on a detailed description of the size, shape and location of these islands, and on the profiles of the potential $\xi(\cdot)$ and the solution $u(\cdot,t)$ on them. A special role is played by the case where ξ is an i.i.d. random potential with a *double-exponential* marginal distribution

$$\mathsf{P}(\xi(0) > u) = \mathrm{e}^{-\mathrm{e}^{u/\varrho}}, \qquad u \in \mathbb{R}, \tag{2.2}$$

where $\varrho \in (0,\infty)$ is a parameter. This distribution turns out to be critical, in the sense that the intermittent islands neither grow nor shrink with time, and represents a class of its own.

In the present paper we consider the case where $\mathscr{X}$ is an *unrooted regular tree* $\mathcal{T}$. Our focus will be on the asymptotics as $t \to \infty$ of the total mass

$$U(t) = \sum_{x\in\mathcal{T}} u(x,t).$$

In earlier work [12], [13] we were concerned with the case where $\mathscr{X}$ is a *rooted Galton-Watson tree* in the *quenched* setting, i.e., almost surely with respect to the random tree and the random potential. This work was restricted to the case where the random potential is given by (2.2) and the offspring distribution of the Galton-Watson tree has support in $\mathbb{N}\backslash\{1\}$ with a sufficiently thin tail. In the present paper our focus will be on the *annealed* setting, i.e., averaged over the random potential. We derive two terms in the asymptotic expansion as $t \to \infty$ of the average total mass

$$\langle U(t) \rangle = \sum_{x \in \mathcal{T}} \langle u(x,t) \rangle,$$

where $\langle \cdot \rangle$ denotes expectation with respect to the law of the random potential. It turns out that the annealed expansion *differs* from the quenched expansion, even though the same variational formula plays a central role in the two second terms.

The derivation in the annealed setting forces us to follow a different route than in the quenched setting, based on various approximations of $\mathcal{T}$ that are more delicate than the standard approximation of $\mathbb{Z}^d$ (see [11, Chapter VIII]). This is the reason why we consider regular trees rather than Galton-Watson trees, to which we hope to return later. A key tool in the analysis is the large deviation principle for the empirical distribution of *Markov renewal processes* on finite graphs derived in [31], which is recalled in Appendix A.1.

§2.1.2 The PAM on a graph

Notations and definitions

Let $G = (V, E)$ be a *simple connected undirected* graph, either finite or countably infinite, with an arbitrarily designated vertex $\mathcal{O}$. Let Δ_G be the Laplacian on G, i.e.,

$$(\Delta_G f)(x) = \sum_{\substack{y \in V: \\ \{x,y\} \in E}} [f(y) - f(x)], \qquad x \in V,\ f\colon\ V \to \mathbb{R},$$

which acts along the edges of G. Let $\xi := (\xi(x))_{x \in V}$ be a random potential attached to the vertices of G, taking values in $\mathbb{R}$. Our object of interest is the non-negative solution of the Cauchy problem with localised initial condition,

$$\begin{array}{rcll} \partial_t u(x,t) & = & (\Delta_G u)(x,t) + \xi(x) u(x,t), & x \in V,\ t > 0, \\ u(x,0) & = & \delta_{\mathcal{O}}(x), & x \in V. \end{array} \tag{2.3}$$

The quantity $u(x,t)$ can be interpreted as the amount of mass at time t at site x when initially there is unit mass at $\mathcal{O}$. The total mass at time t is $U(t) = \sum_{x \in V} u(x,t)$. The total mass is given by the *Feynman-Kac formula*

$$U(t) = \mathbb{E}_{\mathcal{O}} \left(\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \right), \tag{2.4}$$

where $X = (X_t)_{t \geq 0}$ is the continuous-time random walk on the vertices V with jump rate 1 along the edges E, and $\mathbb{P}_{\mathcal{O}}$ denotes the law of X given $X_0 = \mathcal{O}$. Let $\langle \cdot \rangle$ denote expectation with respect to ξ. The quantity of interest in this paper is the average total mass at time t:

$$\langle U(t) \rangle = \left\langle \mathbb{E}_{\mathcal{O}} \left(\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \right) \right\rangle. \tag{2.5}$$

Assumption on the potential

Throughout the paper we assume that the random potential $\xi = (\xi(x))_{x\in V}$ consists of i.i.d. random variables with a marginal distribution whose cumulant generating function

$$H(u) = \log \left\langle \mathrm{e}^{u\xi(\mathcal{O})} \right\rangle \tag{2.6}$$

satisfies the following:

Assumption 2.A. [Asymptotic double-exponential potential]
There exists a $\varrho \in (0,\infty)$ such that

$$\lim_{u\to\infty} uH''(u) = \varrho. \tag{2.7}$$

■

Remark 2.1.1. [Double-exponential potential] A special case of (2.7) is when $\xi(\mathcal{O})$ has the double-exponential distribution in (2.2), in which case

$$H(u) = \log \Gamma(\varrho u + 1)$$

with Γ the gamma function. ♠

By Stirling's approximation, (2.7) implies

$$H(u) = \varrho u \log(\varrho u) - \varrho u + o(u), \qquad u \to \infty. \tag{2.8}$$

Assumption 2.A is more than enough to guarantee existence and uniqueness of the non-negative solution to (2.3) on any discrete graph with at most exponential growth (as can be inferred from the proof in [24], [25] for the case $G = \mathbb{Z}^d$). Since ξ is assumed to be i.i.d., we have from (2.5) that

$$\langle U(t)\rangle = \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V} H(\ell_t(x))\right]\right), \tag{2.9}$$

where

$$\ell_t(x) = \int_0^t \mathbb{1}\{X_s = x\}\,\mathrm{d}s, \qquad x \in V,\ t \geq 0,$$

is the local time of X at vertex x up to time t.

Variational formula

The following *characteristic variational formula* is important for the description of the asymptotics of $\langle U(t)\rangle$. Denote by $\mathcal{P}(V)$ the set of probability measures on V. For $p \in \mathcal{P}(V)$, define

$$I_E(p) = \sum_{\{x,y\}\in E} \left(\sqrt{p(x)} - \sqrt{p(y)}\right)^2, \qquad J_V(p) = -\sum_{x\in V} p(x)\log p(x), \tag{2.10}$$

and set

$$\chi_G(\varrho) = \inf_{p\in\mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0,\infty). \tag{2.11}$$

The first term in (2.11) is the quadratic form associated with the Laplacian, which is the large deviation rate function for the *empirical distribution*

$$L_t = \frac{1}{t}\int_0^t \delta_{X_s}\,\mathrm{d}s = \frac{1}{t}\sum_{x\in V} \ell_t(x)\delta_x \in \mathcal{P}(V) \tag{2.12}$$

(see e.g. [11, Section IV]). The second term in (2.11) captures the second order asymptotics of $\sum_{x\in V} H(tp(x))$ as $t\to\infty$ via (2.8) (see e.g. [11, Section VIII]).

Reformulation

The following lemma pulls the leading order term out of the expansion and shows that the second order term is controlled by the large deviation principle for the empirical distribution.

Lemma 2.1.2. *[Key object for the expansion] If $G=(V,E)$ is finite, then*

$$\langle U(t)\rangle = \mathrm{e}^{H(t)+o(t)}\,\mathbb{E}_{\mathcal{O}}\left(\mathrm{e}^{-\varrho t J_V(L_t)}\right), \qquad t\to\infty.$$

where J_V is the functional in (2.10) *and L_t is the empirical distribution in* (2.12)*.*

Proof. Because $\sum_{x\in V}\ell_t(x) = t$, we can rewrite (2.9) as

$$\begin{aligned}\langle U(t)\rangle &= \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V} H(\ell_t(x))\right]\right)\\ &= \mathrm{e}^{H(t)}\,\mathbb{E}_{\mathcal{O}}\left(\exp\left\{t\sum_{x\in V}\frac{1}{t}\left[H(\tfrac{\ell_t(x)}{t}t) - \tfrac{\ell_t(x)}{t}H(t)\right]\right\}\right).\end{aligned}$$

Assumption 2.A implies that H has the following scaling property (see [23]):

$$\lim_{t\to\infty}\frac{1}{t}[H(ct)-cH(t)] = \varrho c\log c \quad \text{uniformly in } c\in[0,1].$$

Hence the claim follows. □

§2.1.3 The PAM on an unrooted regular tree: annealed total mass for large times and key variational formula

In this section we specialise to the case where $G=\mathcal{T}=(E,V)$, an unrooted regular tree of degree $d+1$ with $d\geq 2$ (see Fig. 2.1). The main theorem of our paper is the following expansion.

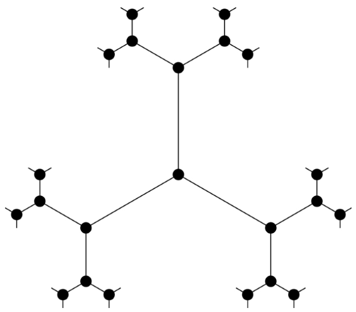

Figure 2.1: An unrooted tree with degree 3 (= offspring size 2).

Theorem 2.1.3. *[Growth rate of the total mass] For any $d \geq 4$, subject to Assumption 2.A,*

$$\frac{1}{t}\log\langle U(t)\rangle = \varrho\log(\varrho t) - \varrho - \chi_{\mathcal{T}}(\varrho) + o(1), \qquad t \to \infty, \tag{2.13}$$

where $\chi_{\mathcal{T}}(\varrho)$ is the variational formula in (2.11) with $G = \mathcal{T}$.

The proof of Theorem 2.1.3 is given in Sections 2.2–2.4 and makes use of technical computations collected in Appendices A.2–A.3.

The main properties of the key quantity

$$\chi_{\mathcal{T}}(\varrho) = \inf_{p\in\mathcal{P}(V)}[I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0,\infty), \tag{2.14}$$

are collected in the following theorem (see Fig. 2.2).

Theorem 2.1.4. *[Properties of the variational formula] For any $d \geq 2$ the following hold:*
(a) The infimum in (2.14) may be restricted to the set

$$\mathcal{P}^{\downarrow}_{\mathcal{O}}(V) = \big\{p \in \mathcal{P}(V)\colon \operatorname{argmax} p = \mathcal{O},\ p \text{ is non-increasing along paths to infinity}\big\}. \tag{2.15}$$

(b) For every $\varrho \in (0,\infty)$, the infimum in (2.14) restricted to $\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$ is attained, every minimiser $\bar{p}$ is such that $\bar{p} > 0$ on V, and $\partial S_R = \sum_{\partial B_R(\mathcal{O})}\bar{p}(x)$, $R \in \mathbb{N}_0$, satisfies

$$\sum_{R\in\mathbb{N}_0}\partial S_R\log(R+1) \leq \frac{d+1}{\varrho},$$

where $B_R(\mathcal{O})$ is the ball of radius R centred at $\mathcal{O}$.
(c) The function $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$ is strictly increasing and globally Lipschitz continuous on $(0,\infty)$, with

$$\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) = d-1, \qquad \lim_{\varrho\to\infty}\chi_{\mathcal{T}}(\varrho) = d+1.$$

The proof of Theorem 2.1.4 is given in Section 2.5 (see Fig. 2.2).

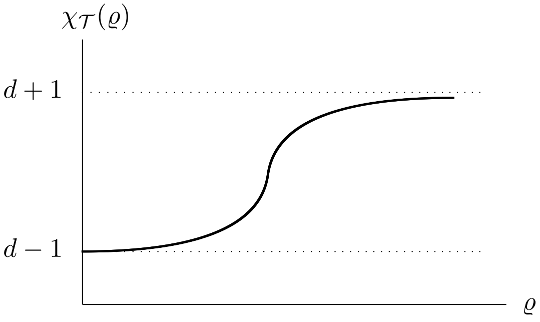

Figure 2.2: Qualitative plot of $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$.

§2.1.4 Discussion

1. Theorem 2.1.3 identifies the scaling of the total mass up to and including terms that are exponential in t. The first two terms in the right-hand side of (2.13) are the same as those of $\frac{1}{t}H(t)$ (recall (2.8)). The third term is a correction that comes from the cost for X in the Feynman-Kac formula in (2.4) to create an *optimal local time profile* somewhere in $\mathcal{T}$, which is captured by the minimiser(s) of the variational formula in (2.14).

2. For the quenched model on a *rooted Galton-Watson tree* $\mathcal{GW}$ we found in [12], [13] that

$$\frac{1}{t}\log U(t) = \varrho \log\left(\frac{\varrho t \vartheta}{\log\log t}\right) - \varrho - \chi(\varrho) + o(1), \qquad t \to \infty, \qquad \mathsf{P} \times \mathfrak{P}\text{-a.s.}, \tag{2.16}$$

where P is the law of the potential, $\mathfrak{P}$ is the law of $\mathcal{GW}$, ϑ is the logarithm of the mean of the offspring distribution, and

$$\chi_{\mathcal{T}}(\varrho) = \inf_{\mathcal{S} \subset \mathcal{GW}} \chi_{\mathcal{S}}(\varrho) \tag{2.17}$$

with $\chi_{\mathcal{S}}(\varrho)$ given by (2.11) and the infimum running over all subtrees of $\mathcal{GW}$. This result was shown to be valid as soon as the offspring distribution has support in $\mathbb{N}\backslash\{1\}$ (i.e., all degrees are at least 3) and has a sufficiently thin tail. The extra terms in (2.16) come from the cost for X in the Feynman-Kac formula in (2.4) to travel in a time of order $o(t)$ to an optimal finite subtree with an optimal profile of the potential, referred to as *intermittent islands*, located at a distance of order $\varrho t/\log\log t$ from $\mathcal{O}$, and to subsequently spend most of its time on that subtree. In this cost the parameter ϑ appears, which is absent in (2.13). It was shown in [12] that if $\varrho \geq 1/\log(d_{\min}+1)$, with $d_{\min}$ the minimum of the support of the offspring distribution, then the infimum in (2.17) is attained at the unrooted regular tree with degree $d_{\min}+1$, i.e., the *minimal unrooted regular tree contained in* $\mathcal{GW}$, for which $\vartheta = \log d_{\min}$. Possibly the bound on ϱ is redundant.

3. In view of Lemma 2.1.2 and the fact that Assumption 2.A implies (2.8), we see that the proof of Theorem 2.1.3 amounts to showing that, on $\mathcal{T} = (V, E)$,

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}_{\mathcal{O}} \left(\mathrm{e}^{-\varrho t J_V(L_t)} \right) = -\chi_{\mathcal{T}}(\varrho).$$

We achieve this by deriving asymptotically matching upper and lower bounds. These bounds are obtained by truncating $\mathcal{T}$ outside a ball of radius R, to obtain a finite tree $\mathcal{T}_R$, deriving the $t \to \infty$ asymptotics for finite R, and letting $R \to \infty$ afterwards. For the lower bound we can use the standard truncation technique based on *killing* X when it exits $\mathcal{T}_R$ and applying the large deviation principle for the empirical distribution of *Markov processes* on finite graphs derived in [15]. For the upper bound, however, we cannot use the standard truncation technique based on *periodisation* of X beyond radius R, because $\mathcal{T}$ is an expander graph (see [29, Chapter IV] for a list of known techniques on $\mathbb{Z}^d$ and $\mathbb{R}^d$). Instead, we follow a route in which $\mathcal{T}$ is approximated in successive stages by a version of $\mathcal{T}_R$ with a *specific boundary condition*, based on monitoring X relative to its backbone to infinity. This route allows us to use the large deviation principle for the empirical distribution of *Markov renewal processes* on finite graphs derived in [31], but we need the condition $d \geq 4$ to *control* the specific boundary condition in the limit as $R \to \infty$ (see Remark 2.4.1 for more details). The reason why the approximation of $\mathcal{T}$ by finite subtrees is successful is precisely because in the parabolic Anderson model the *total mass tends to concentrate on intermittent islands*.

4. Theorem 2.1.4 shows that, modulo translations, the optimal strategy for L_t as $t \to \infty$ is to be close to a minimiser of the variational formula in (2.14) restricted to $\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$. Any minimiser is centred at $\mathcal{O}$, strictly positive everywhere, non-increasing in the distance to $\mathcal{O}$, and rapidly tending to zero. The following questions remain open:

(1) Is the minimiser $\bar{p}$ unique modulo translation?

(2) Does $\bar{p}(x)$ satisfy $\lim_{|x|\to\infty} |x|^{-1} \log \bar{p}(x) = -\infty$, with $|x|$ the distance between x and $\mathcal{O}$?

(3) Is $\bar{p}$ radially symmetric?

(4) Is $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$ analytic on $(0, \infty)$?

We expect the answer to be yes for (1) and (2), and to be no for (3) and (4).

§2.2 Proof of the main theorem: lower bound

In this section we prove the lower bound in Theorem 2.1.3, which is standard and straightforward. In Section 2.2.1 we obtain a lower bound in terms of a variational formula by *killing* the random walk when it exits $\mathcal{T}_R$. In Section 2.2.2 we derive the lower bound of the expansion by letting $R \to \infty$ in the variational formula.

§2.2.1 Killing and lower variational formula

Fix $R \in \mathbb{N}$. Let $\mathcal{T}_R$ be the subtree of $\mathcal{T} = (V, E)$ consisting of all the vertices that are within distance R of the root $\mathcal{O}$ and all the edges connecting them. Put $V_R = V_R(\mathcal{T}_R)$ and $E_R = E(\mathcal{T}_R)$. Let $\tau_R = \inf\{t \geq 0\colon\ X_t \notin V_R\}$ denote the first time that X exits $\mathcal{T}_R$. It follows from (2.9) that

$$\langle U(t)\rangle \geq \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V_R} H(\ell_t(x))\right]\mathbb{1}\{\tau_R > t\}\right).$$

Since $\mathcal{T}_R$ is finite, Lemma 2.1.2 gives

$$\langle U(t)\rangle \geq \mathrm{e}^{H(t)+o(t)}\,\mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{-\varrho t J_V(L_t)}\mathbb{1}\{\tau_R > t\}\right]$$

with J_V the functional defined in (2.10). As shown in [22] (see also [25]), the family of sub-probability distributions $\mathbb{P}_{\mathcal{O}}(L_t \in \cdot, \tau_R > t)$, $t \geq 0$, satisfies the LDP on $\mathcal{P}^R(V) = \{p \in \mathcal{P}(V)\colon \operatorname{supp}(p) \subset V_R\}$ with rate function I_E, with I_E the functional defined in (2.10). This is the *standard* LDP for the empirical distribution of *Markov processes*. Therefore, by Varadhan's Lemma,

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{-\varrho t J_V(L_t)}\mathbb{1}\{\tau_R > t\}\right] = -\chi_R^-(\varrho)$$

with

$$\chi_R^-(\varrho) = \inf_{p\in\mathcal{P}^R(V)}[I_E(p) + \varrho J_V(p)], \tag{2.18}$$

where we use that $p \mapsto J_V(p)$ is bounded and continuous (in the discrete topology) on $\mathcal{P}^R(V)$. Note that

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{P}_{\mathcal{O}}(\tau_R > t) = -\inf_{p\in\mathcal{P}^R(V)} I_E(p) < 0,$$

which is non-zero because any $p \in \mathcal{P}^R(V)$ is non-constant on V. The expression in (2.18) is the same as (2.11) with $G = \mathcal{T}$, except that p is restricted to V_R.

§2.2.2 Limit of the lower variational formula

Clearly, $R \mapsto \chi_R^-(\varrho)$ is non-increasing. To complete the proof of the lower bound in Theorem 2.1.3, it remains is to show the following.

Lemma 2.2.1. $\limsup_{R\to\infty}\chi_R^-(\varrho) \leq \chi_{\mathcal{T}}(\varrho)$.

Proof. Pick any $p \in \mathcal{P}(V)$ such that $I_E(p) < \infty$ and $J_V(p) < \infty$. Let p^R be the projection of p onto V_R, i.e.,

$$p^R(x) = \begin{cases} p(x), & x \in \operatorname{int}(V_R),\\ \sum_{y\geq x} p(y), & x \in \partial V_R,\end{cases}$$

where $y \geq x$ means that y is an element of the progeny of x in $\mathcal{T}$. Since $p^R \in \mathcal{P}^R(V)$, we have from (2.18) that $\chi_R^-(\varrho) \leq I_E(p^R) + \varrho J_V(p^R)$. Trivially, $\lim_{R\to\infty} I_E(p^R) = I_E(p)$ and $\lim_{R\to\infty} J_V(p^R) = J_V(p)$, and so we have $\limsup_{R\to\infty} \chi_R^-(\varrho) \leq I_E(p) + \varrho J_V(p)$. Since this bound holds for arbitrary $p \in \mathcal{P}(V)$, the claim follows from (2.18). □

§2.3 Proof of the main theorem: upper bound

In this section we prove the upper bound in Theorem 2.1.3, which is more laborious and requires a more delicate approach than the standard periodisation argument used on $\mathbb{Z}^d$. In Section 2.3.1 we obtain an upper bound in terms of a variational formula on a version of $\mathcal{T}_R$ with a specific boundary condition. The argument comes in four steps, encapsulated in Lemmas 2.3.1–2.3.6 below:

(I) *Condition on the backbone* of X (Section 2.3.1).

(II) *Project* X onto a concatenation of finite subtrees attached to this backbone that are *rooted* versions of $\mathcal{T}_R$ (Section 2.3.1).

(III) *Periodise* the projected X to obtain a *Markov renewal process* on a single finite subtree and show that the periodisation can be chosen such that the local times at the vertices on the *boundary* of the finite subtree are negligible (Section 2.3.1).

(IV) Use the *large deviation principle* for the empirical distribution of *Markov renewal processes* derived in [31] to obtain a variational formula on a single subtree (Section 2.3.1).

In Section 2.3.2 we derive the upper bound of the expansion by letting $R \to \infty$ in the variational formula.

§2.3.1 Backbone, projection, periodisation and upper variational formula

Backbone

For $r \in \mathbb{N}$, let τ_r be the last time when X visits $\partial B_r(\mathcal{O}) = \{x \in V : d(x, \mathcal{O}) = r\}$, the boundary of the ball of radius r around $\mathcal{O}$. Then the sequence $\mathcal{B} = (X_{\tau_r})_{r\in\mathbb{N}_0}$ forms the backbone of X, running from $\mathcal{O}$ to infinity.

Lemma 2.3.1. *[Condition on a backbone] For every backbone* bb *and every* $t \geq 0$*,*

$$\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{T})} H(\ell_t(x))\right]\right) = \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{T})} H(\ell_t(x))\right] \,\middle|\, \mathcal{B} = \mathrm{bb}\right).$$

Proof. By symmetry, the conditional expectation in the right-hand side does not depend on the choice of bb. Indeed, permutations of the edges away from the root do not affect the law of $\sum_{x\in V(\mathcal{T})} H(\ell_t(x))$. □

Turn the one-sided backbone into a two-sided backbone by adding a second backbone from $\mathcal{O}$ to infinity. By symmetry, the choice of this second backbone is arbitrary, say bb$'$. Redraw $\mathcal{T}$ by representing bb$' \cup$ bb as $\mathbb{Z}$ and representing the rest of $\mathcal{T}$ as a sequence of *rooted trees* $\mathcal{T}^* = (\mathcal{T}_x^*)_{x\in\mathbb{Z}}$ hanging off $\mathbb{Z}$ (see Fig. 2.3). In $\mathcal{T}_x^*$, the root sits at x and has $d-1$ downward edges, while all lower vertices have d downward edges.

Let $\mathcal{T}^{\mathbb{Z}}$ denote the redrawn tree and $X^{\mathbb{Z}} = (X_t^{\mathbb{Z}})_{t\geq 0}$ be the random walk on $\mathcal{T}^{\mathbb{Z}}$. Furthermore let $(\ell_t^{\mathbb{Z}}(x))_{x\in\mathcal{T}^{\mathbb{Z}}}$ denote the local times of $X^{\mathbb{Z}}$ at time t.

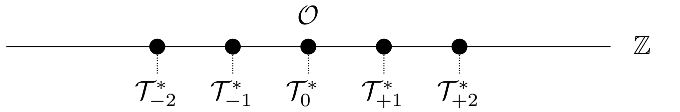

Figure 2.3: Redrawing of $\mathcal{T}$ as $\mathcal{T}^{\mathbb{Z}}$: a two-sided backbone $\mathbb{Z}$ with a sequence $\mathcal{T}^* = (\mathcal{T}_x^*)_{x\in\mathbb{Z}}$ of rooted trees hanging off. The upper index $*$ is used to indicate that the tree is rooted.

Lemma 2.3.2. *[Representation of $\mathcal{T}$ as a backbone with rooted trees] For every* bb *and $t \geq 0$,*

$$\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{T})} H(\ell_t(x))\right] \,\middle|\, \mathcal{B} = \mathrm{bb}\right) = \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{T}^{\mathbb{Z}})} H(\ell_t^{\mathbb{Z}}(x))\right] \,\middle|\, X_\infty^{\mathbb{Z}} = +\infty\right).$$

Proof. Simply redraw $\mathcal{T}$ as $\mathcal{T}^{\mathbb{Z}}$. □

Note that $X^{\mathbb{Z}}$ is a Markov process whose sojourn times have distribution EXP$(d+1)$ and whose steps are drawn uniformly at random from the $d+1$ edges that are incident to each vertex.

Projection

For $R \in \mathbb{N}\backslash\{1\}$, cut $\mathbb{Z}$ into slices of length R, i.e.,

$$\mathbb{Z} = \cup_{k\in\mathbb{Z}}(z + (kR + I)), \qquad I = \{0, 1, \dots, R-1\},$$

where z is to be chosen later. Apply the following two maps to $\mathcal{T}^{\mathbb{Z}}$ (in the order presented):

(i) For each $k \in \mathbb{Z}$, fold $\mathcal{T}_{z+(kR+(R-1))}^*$ onto $\mathcal{T}_{z+(k+1)R}^*$ by folding the $d-1$ edges downwards from the root on top of the edge in $\mathbb{Z}$ connecting $z + (kR + (R-1))$ and $z + (k+1)R$, and putting the d infinite rooted trees hanging off each of these $d-1$ edges on top of the rooted tree $\mathcal{T}_{z+(k+1)R}^*$ hanging off $z + (k+1)R$. Note that each of the d infinite rooted trees is a copy of $\mathcal{T}_{z+(k+1)R}^*$.

(ii) For each $k \in \mathbb{Z}$ and $m \in \{0, 1, \ldots, R-2\}$, cut off all the infinite subtrees in $\mathcal{T}^*_{z+(kR+m)}$ whose roots are at depth $(R-1)-m$. Note that the total number of leaves after the cutting equals

$$(d-1)\sum_{m=0}^{R-2} d^{(R-2)-m} = (d-1)d^{R-2}\,\frac{1-d^{-(R-1)}}{1-d^{-1}} = d^{R-1}-1,$$

which is the same as the total number of leaves of the rooted tree $\mathcal{T}^*_R$ of depth $R-1$ (i.e., with R generations) minus 1 (a fact we will need below).

By doing so we obtain a *concatenation* of finite units

$$\mathcal{U}_R = (\mathcal{U}_R[k])_{k\in\mathbb{Z}}$$

that are rooted trees of depth $R-1$ (see Fig. 2.4). Together with the two maps that turn $\mathcal{T}^{\mathbb{Z}}$ into $\mathcal{U}_R$, we apply two maps to $X^{\mathbb{Z}}$:

(i) All excursions of $X^{\mathbb{Z}}$ in the infinite subtrees that are *folded to the right and on top* are projected accordingly.

(ii) All excursions of $X^{\mathbb{Z}}$ in the infinite subtrees that are *cut off* are replaced by a sojourn of $X^{\mathcal{U}_R}$ in the *tadpoles* that replace these subtrees (see Fig. 2.4)

The resulting path, which we call $X^{\mathcal{U}_R} = (X^{\mathcal{U}_R}_t)_{t\geq 0}$, is a *Markov renewal process* with the following properties:

- The sojourn times in all the vertices that are not tadpoles have distribution $\mathrm{EXP}(d+1)$.
- The sojourn times in all the tadpoles have distribution ψ, defined as the conditional distribution of the *return time* τ of the random walk on the infinite rooted tree $\mathcal{T}^*$ *given* that $\tau < \infty$ (see [30] for a proper definition).
- The transitions into the tadpoles have probability $\frac{d}{d+1}$, the transitions out of the tadpoles have probability 1 (because of the condition $X^{\mathbb{Z}}_\infty = +\infty$).
- The transitions from $z + (kR + (R-1))$ to $z + (k+1)R$ have probability $\frac{d}{d+1}$, while the reverse transitions have probability $\frac{1}{d+1}$.

Write $(\ell^{\mathcal{U}_R}_t(x))_{x\in V_{\mathcal{U}_R}}$ to denote the local times of $X^{\mathcal{U}_R}$ at time t.

Lemma 2.3.3. *[Projection onto a concatenation of finite subtrees] For every $R \in \mathbb{N}\backslash\{1\}$ and $t \geq 0$,*

$$\begin{aligned}
&\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{T}^{\mathbb{Z}})} H(\ell^{\mathbb{Z}}_t(x))\right] \,\middle|\, X^{\mathbb{Z}}_\infty = +\infty\right)\\
&\quad\leq \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{U}_R)} H(\ell^{\mathcal{U}_R}_t(x))\right] \,\middle|\, X^{\mathcal{U}_R}_\infty = +\infty\right).
\end{aligned}$$

Proof. The maps that are applied to turn $X^{\mathbb{Z}}$ into $X^{\mathcal{U}_R}$ are such that local times are *stacked on top of each other*. Since H defined in (2.6) is convex and $H(0) = 0$, we have $H(\ell) + H(\ell') \leq H(\ell + \ell')$ for all $\ell, \ell' \in \mathbb{N}_0$, which implies the inequality. □

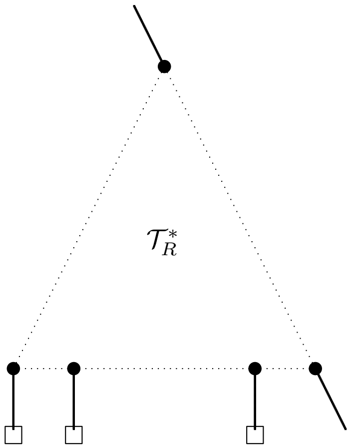

Figure 2.4: A unit in $\mathcal{U}_R$. Inside is a rooted tree $\mathcal{T}_R^*$ of depth $R-1$, of which only the root and the leaves are drawn. Hanging off the leaves at depth $R-1$ from the root are tadpoles, except for the right-most bottom vertex, which has a downward edge that connects to the root of the next unit. The vertices marked by a bullet form the boundary of $\mathcal{U}_R$, the vertices marked by a square box form the tadpoles of $\mathcal{U}_R$.

Periodisation

Our next observation is that the condition $\{X_\infty^{\mathcal{U}_R} = +\infty\}$ is *redundant*.

Lemma 2.3.4. *[Condition redundant] For every $R \in \mathbb{N}\backslash\{1\}$ and $t \geq 0$,*

$$\mathbb{E}_\mathcal{O}\left(\exp\left[\sum_{x\in V(\mathcal{U}_R)} H(\ell_t^{\mathcal{U}_R}(x))\right] \,\middle|\, X_\infty^{\mathcal{U}_R} = +\infty\right) = \mathbb{E}_\mathcal{O}\left(\exp\left[\sum_{x\in V(\mathcal{U}_R)} H(\ell_t^{\mathcal{U}_R}(x))\right]\right).$$

Proof. The event $\{X_\infty^{\mathcal{U}_R} = +\infty\}$ has probability 1 because on the edges connecting the units of $\mathcal{U}_R$ (see Fig. 2.4) there is a drift downwards. To see why, note that $\frac{1}{d+1} < \frac{1}{2} < \frac{d}{d+1}$ because $d \geq 2$, and use that a one-dimensional random walk with drift is transient to the right [33]. □

Since $\mathcal{U}_R$ is periodic, we can *fold* $X^{\mathcal{U}_R}$ onto a single unit $\mathcal{W}_R$, to obtain a Markov renewal process $X^{\mathcal{W}_R}$ on $\mathcal{W}_R$ (see Fig. 2.5) in which the transition from the top vertex to the right-most bottom vertex has probability $\frac{1}{d+1}$, while the reverse transition has probability $\frac{d}{d+1}$. Clearly, the sojourn time distributions are not affected by the folding and therefore remain as above. Write $(\ell_t^{\mathcal{W}_R}(x))_{x\in V(\mathcal{W}_R)}$ to denote the local times of $X^{\mathcal{W}_R}$ at time t.

Lemma 2.3.5. *[Periodisation to a single finite subtree] For every $R \in \mathbb{N}\backslash\{1\}$ and $t \geq 0$,*

$$\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{U}_R)} H(\ell_t^{\mathcal{U}_R}(x))\right]\right) \leq \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{W}_R)} H(\ell_t^{\mathcal{W}_R}(x))\right]\right).$$

Proof. The periodisation again stacks local time on top of each other. □

Before we proceed we make a *crucial observation*, namely, we may still choose the shift $z \in \{0,1,\ldots,R-1\}$ of the cuts of the two-sided backbone $\mathbb{Z}$ (recall Fig. 2.3). We will do so in such a way that the *local time* up to time t spent in the set $\partial_{\mathcal{U}_R}$ defined by

$$\begin{aligned}\partial_{\mathcal{U}_R} &= \text{all vertices at the top or at the bottom of a unit in } \mathcal{U}_R\\ &= \text{all vertices marked by } \bullet \text{ in Fig. 2.4}\end{aligned} \tag{2.19}$$

is at most t/R. After the periodisation these vertices are mapped to the set $\partial_{\mathcal{W}_R}$ defined by

$$\begin{aligned}\partial_{\mathcal{W}_R} &= \text{all vertices at the top or at the bottom of } \mathcal{W}_R\\ &= \text{all vertices marked by } \bullet \text{ in Fig. 2.5.}\end{aligned}$$

Lemma 2.3.6. *[Control on the time spent at the boundary] For every $R \in \mathbb{N}\backslash\{1\}$ and $t \geq 0$,*

$$\begin{aligned}&\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{U}_R)} H(\ell_t^{\mathcal{U}_R}(x))\right]\right)\\ &\qquad \leq \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x\in V(\mathcal{W}_R)} H(\ell_t^{\mathcal{W}_R}(x))\right] 1_{\left\{\frac{1}{t}\sum_{x\in\partial_{\mathcal{W}_R}} \ell_t^{\mathcal{W}_R}(x)\leq 1/R\right\}}\right).\end{aligned}$$

Proof. For different z the sets of vertices making up ∂_R correspond to *disjoint* sets of vertices in $\mathcal{T}^{\mathbb{Z}}$ (see Fig. 2.4). Since $\sum_{x\in\mathcal{T}^{\mathbb{Z}}} \ell_t^{\mathbb{Z}}(x) = t$ for all $t \geq 0$, it follows that there exists a z for which $\sum_{x\in\partial_R} \ell_t^{\mathbb{Z}}(x) \leq t/R$. Therefore the upper bound in Lemma 2.3.3 can be strengthened to the one that is claimed. □

Upper variational formula

Lemmas 2.3.1–2.3.6 provide us with an upper bound for the average total mass (recall ((2.9)) on the *infinite* tree $\mathcal{T}$ in terms of the same quantity on the *finite* tree-like unit $\mathcal{W}_R$ with a *specific boundary condition*. Along the way we have paid a price: the sojourn times in the tadpoles are *no longer* exponentially distributed, and the transition probabilities into and out of the tadpoles and between the top vertex and the right-most bottom vertex are *biased*. We therefore need the large deviation principle for the empirical distribution of Markov renewal processes derived in [31], which we can now apply to the upper bound.

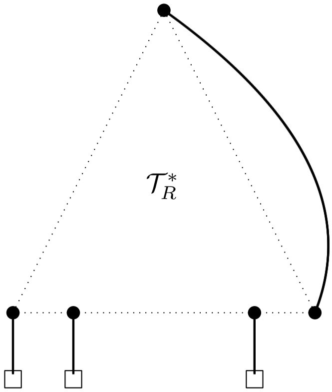

Figure 2.5: A unit $\mathcal{W}_R$ with the top vertex and the right-most bottom vertex connected by an edge.

Since $\mathcal{W}_R$ is finite, Lemma 2.1.2 gives

$$\langle U(t)\rangle \leq \mathrm{e}^{H(t)+o(t)}\,\mathbb{E}_{\mathcal{O}}\left(\mathrm{e}^{-\varrho J_{V(\mathcal{W}_R)}(L_t^{\mathcal{W}_R})}\,1_{\left\{L_t^{\mathcal{W}_R}(\partial_{\mathcal{W}_R})\leq 1/R\right\}}\right)$$

with J_V the functional defined in (2.10). The following lemma controls the expectation in the right-hand side.

Lemma 2.3.7. *[Scaling of the key expectation] For every $R\in\mathbb{N}\backslash\{1\}$,*

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}_{\mathcal{O}}\left(\mathrm{e}^{-\varrho t J_{V(\mathcal{W}_R)}(L_t^{\mathcal{W}_R})}\,1_{\left\{L_t^{\mathcal{W}_R}(\partial_{\mathcal{W}_R})\leq 1/R\right\}}\right) = -\chi_R^+(\varrho),$$

where

$$\chi_R^+(\varrho) = \inf_{\substack{p\in\mathcal{P}(V(\mathcal{W}_R)):\\ p(\partial_{\mathcal{W}_R})\leq 1/R}}\left\{I^\dagger_{E(\mathcal{W}_R)}(p)+\varrho J_{V(\mathcal{W}_R)}(p)\right\}, \tag{2.20}$$

with

$$I^\dagger_{E(\mathcal{W}_R)}(p) = \inf_{\beta\in(0,\infty)}\inf_{q\in\mathcal{P}(V(\mathcal{W}_R))}\left[\widehat{K}(\beta q)+\widetilde{K}(p\mid\beta q)\right], \tag{2.21}$$

where

$$\widehat{K}(\beta q) = \sup_{\widehat{q}\in\mathcal{P}(V(\mathcal{W}_R))}\sum_{x\in V(\mathcal{W}_R)}\beta q(x)\log\left(\frac{\widehat{q}(x)}{\sum_{y\in V(\mathcal{W}_R)}\pi_{x,y}\widehat{q}(y)}\right), \tag{2.22}$$

$$\widetilde{K}(p\mid\beta q) = \sum_{x\in V(\mathcal{W}_R)}\beta q(x)\,(\mathcal{L}\lambda_x)\left(\tfrac{p(x)}{\beta q(x)}\right), \tag{2.23}$$

with

$$(\mathcal{L}\lambda_x)(\alpha) = \sup_{\theta\in\mathbb{R}}[\alpha\theta-\lambda_x(\theta)], \qquad \alpha\in[0,\infty), \tag{2.24}$$

$$\lambda_x(\theta) = \log\int_0^\infty \mathrm{e}^{\theta\tau}\psi_x(\mathrm{d}\tau), \quad \theta\in\mathbb{R}, \tag{2.25}$$

where $\psi_x = \psi$ when x is a tadpole, $\psi_x = \mathrm{EXP}(d+1)$ when x is not a tadpole, and $\pi_{x,y}$ is the transition kernel of the discrete-time Markov chain on $V(\mathcal{W}_R)$ embedded in $X^{\mathcal{W}_R}$.

Proof. Apply the large deviation principle derived in [31], which we recall in Proposition A.1.1 in Appendix A.1. □

The expression in (2.20) is similar to (2.11) with $G = \mathcal{W}_R$, expect that the rate function $I_{E(\mathcal{W}_R)}$ in (2.21) is more involved than the rate function I_E in (2.10).

§2.3.2 Limit of the upper variational formula

The prefactor $\mathrm{e}^{H(t)+o(1)}$ in Lemma 2.1.2 accounts for the terms $\varrho\log(\varrho t) - \varrho$ in the right-hand side of (2.13) (recall 2.8). In view of Lemma 2.3.7, in order to complete the proof of the upper bound in Theorem 2.1.3 it suffices to prove the following lemma.

Lemma 2.3.8. *For any $d \geq 4$, $\liminf_{R\to\infty} \chi_R^+(\varrho) \geq \chi_{\mathcal{T}}(\varrho)$.*

Proof. The proof is given in Section 2.4 and relies on two steps:

- Show that, for $d \geq 4$,

$$I^{\dagger}_{E(\mathcal{W}_R)}(p) \geq I^{+}_{E(\mathcal{W}_R)}(p) + O(1/R) \tag{2.26}$$

  with $I^{+}_{E(\mathcal{W}_R)}$ a rate function similar to the *standard rate function* $I_{E(\mathcal{W}_R)}$ given by (2.10).

- Show that, $d \geq 2$,

$$\widehat{\chi}_R^+(\varrho) = \inf_{\substack{p\in\mathcal{P}(V(\mathcal{W}_R)):\\ p(\partial\mathcal{W}_R)\leq 1/R}} \left\{ I^{+}_{E(\mathcal{W}_R)}(p) + \varrho J_{V(\mathcal{W}_R)}(p) \right\}$$

  satisfies

$$\liminf_{R\to\infty} \widehat{\chi}_R^+(\varrho) \geq \chi_{\mathcal{T}}(\varrho). \tag{2.27}$$

□

§2.4 Analysis of the upper variational formula

In this section we carry out the proof of the claims in Section 2.3.2, namely, we settle (2.26) in Section 2.4.1 and (2.27) in Section 2.4.2. The computations carried out in Appendix A.3 guide us along the way.

§2.4.1 Identification of the rate function for the local times on the truncated tree

To identify the rate function $I^{\dagger}_{E(\mathcal{W}_R)}$ in Lemma 2.3.7, we need to work out the two infima between braces in (2.20). The computation follows the same line of argument as in Appendix A.3, but is more delicate. We will only end up with a lower bound. However, this is sufficient for the upper variational formula.

To simplify the notation we write (recall Fig. 2.5):

$$
\begin{array}{lcl}
(V_R, E_R) & = & \text{vertex and edge set of } \mathcal{W}_R \textit{ without the tadpoles},\\
\mathcal{O} & = & \textit{top} \text{ vertex of } V_R,\\
\star & = & \textit{right-most bottom} \text{ vertex of } V_R,\\
\partial V_R & = & \text{set of vertices at the } \textit{bottom} \text{ of } V_R,\\
\square & = & \text{set of } \textit{tadpoles},\\
\square_x & = & \text{tadpole attached to } x \in \partial V_R \backslash \star.
\end{array}
$$

Note that ∂V_R consists of $\star$ and the vertices to which the tadpoles are attached. Note that $\mathrm{int}(V_R) = V_R \setminus \partial V_R$ *includes* $\mathcal{O}$.

1. Inserting (A.9) in Appendix A.2 into (2.22)–(2.23), we get

$$
I^{\dagger}_{E(\mathcal{W}_R)}(p) = (d+1) \sum_{x \in V_R} p(x) + \inf_{\beta \in (0,\infty)} \inf_{q \in \mathcal{P}(V_R)} \sup_{\widehat{q} \in \mathcal{P}(V_R)} L(\beta, q, \widehat{q} \mid p)
$$

with

$$
L(\beta, q, \widehat{q} \mid p) = -A - B - C - D,
$$

where

$$
\begin{aligned}
A &= \sum_{x \in \mathrm{int}(V_R)} \beta q(x) \left\{ 1 + \log \left(\frac{\sum_{y \sim x} \widehat{q}(y)}{\widehat{q}(x)} \frac{p(x)}{\beta q(x)} \right) \right\},\\
B &= \sum_{x \in \partial V_R \backslash \star} \beta q(x) \left\{ 1 + \log \left(\frac{\widehat{q}(x^{\uparrow}) + d\widehat{q}(\square_x)}{\widehat{q}(x)} \frac{p(x)}{\beta q(x)} \right) \right\},\\
C &= \beta q(\star) \left\{ 1 + \log \left(\frac{\widehat{q}(\star^{\uparrow}) + d\widehat{q}(\mathcal{O})}{\widehat{q}(\star)} \frac{p(\star)}{\beta q(\star)} \right) \right\},\\
D &= \sum_{x \in \square} \beta q(x) \left\{ \log \left(\frac{\widehat{q}(x^{\uparrow})}{\widehat{q}(x)} \right) - (\mathcal{L}\lambda) \left(\frac{p(x)}{\beta q(x)} \right) \right\},
\end{aligned}
$$

with $\mathcal{L}\lambda$ the Legende transform of the cumulant generating function of ψ (recall (2.25)) and $x^{\uparrow}$ the unique vertex to which x is attached upwards. (Recall that $y \sim x$ means that x and y are connected by an edge in E_R.) Note that A, B, C each combine two terms, and that A, B, C, D depend on p. We suppress this dependence because p is fixed.

2. Inserting the parametrisation $\widehat{q} = u/\|u\|_1$ and $q = v/\|v\|_1$ with $u, v \colon V_R \to (0,\infty)$ and putting $\beta q = v$, we may write

$$
I^{\dagger}_{E(\mathcal{W}^R)}(p) = (d+1) \sum_{x \in V_R} p(x) + \inf_{v \colon V_R \to (0,\infty)} \sup_{u \colon V_R \to (0,\infty)} L(u, v) \tag{2.28}
$$

with

$$L(u,v) = -A - B - C - D,$$

where

$$\begin{aligned}
A &= \sum_{x\in \mathrm{int}(V_R)} v(x)\left\{1+\log\left(\frac{\sum_{y\sim x} u(y)}{u(x)}\,\frac{p(x)}{v(x)}\right)\right\},\\
B &= \sum_{x\in \partial V_R\backslash\star} v(x)\left\{1+\log\left(\frac{u(x^{\uparrow})+du(\Box_x)}{u(x)}\,\frac{p(x)}{v(x)}\right)\right\},\\
C &= v(\star)\left\{1+\log\left(\frac{u(\star^{\uparrow})+du(\mathcal{O})}{u(\star)}\,\frac{p(\star)}{v(\star)}\right)\right\},\\
D &= \sum_{x\in\Box} v(x)\left\{\log\left(\frac{u(x^{\uparrow})}{u(x)}\right) - (\mathcal{L}\lambda)\left(\frac{p(x)}{v(x)}\right)\right\}.
\end{aligned} \tag{2.29}$$

Our task is to carry out the supremum over u and the infimum over v in (2.28).

3. First, we compute the infimum over v for fixed u. (Later we will make a judicious choice for u to obtain a lower bound.) Abbreviate

$$\begin{aligned}
A_u(x) &= \frac{\sum_{y\sim x} u(y)}{u(x)}\,p(x), && x\in \mathrm{int}(V_R),\\
B_u(x) &= \frac{u(x^{\uparrow})+du(\Box_x)}{u(x)}\,p(x), && x\in\partial V_R\backslash\star,\\
C_u(\star) &= \frac{u(\star^{\uparrow})+du(\mathcal{O})}{u(\star)}\,p(\star).
\end{aligned} \tag{2.30}$$

- For $z\in V_R$, the first derivatives of L are

$$\begin{aligned}
z\in \mathrm{int}(V_R)\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{A_u(z)}{v(z)}\right),\\
z\in \partial V_R\backslash\star\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{B_u(z)}{v(z)}\right),\\
z=\star\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{C_u(z)}{v(z)}\right),
\end{aligned}$$

while the second derivatives of L equal $1/v(z)>0$. Hence the infimum is uniquely taken at

$$\begin{aligned}
x\in \mathrm{int}(V_R)\colon &\quad \bar v(x) = A_u(x),\\
x\in V_R\backslash\star\colon &\quad \bar v(x) = B_u(x),\\
x=\star\colon &\quad \bar v(x) = C_u(x).
\end{aligned}$$

- For $z\in\Box$, the computation is more delicate. Define (see (A.6) in Appendix A.2)

$$\mu(\alpha) = \alpha(\mathcal{L}\lambda)'(\alpha) - (\mathcal{L}\lambda)(\alpha).$$

The function μ has range $(-\infty,\log\sqrt{d}]$, with the maximal value uniquely taken at $\alpha=\infty$. Therefore there are two cases.

▶ $u(x)/u(x^{\uparrow}) \leq \sqrt{d}$: Abbreviate $\alpha_u(z) = p(z)/v(z)$. For $z \in \square$,

$$\begin{aligned}\frac{\partial L(u,v)}{\partial v(z)} &= \log\left(\frac{u(z)}{u(z^{\uparrow})}\right) + (\mathcal{L}\lambda)\left(\frac{p(z)}{v(z)}\right) - \frac{p(z)}{v(z)}(\mathcal{L}\lambda)'\left(\frac{p(z)}{v(z)}\right)\\ &= \log\left(\frac{u(z)}{u(z^{\uparrow})}\right) - \mu(\alpha_u(z)),\\ \frac{\partial^2 L(u,v)}{v(z)^2} &= \frac{p^2(z)}{v^3(z)}(\mathcal{L}\lambda)''\left(\frac{p(z)}{v(z)}\right) > 0,\end{aligned}$$

where we use that $\mathcal{L}\lambda$, being a Legendre transform, is strictly convex. Hence the infimum is uniquely taken at

$$\bar{v}(x) = \frac{p(x)}{\alpha_u(x)}, \qquad x \in \square,$$

with $\alpha_u(x)$ solving the equation

$$\log\left(\frac{u(x)}{u(x^{\uparrow})}\right) = \mu(\alpha_u(x)), \qquad x \in \square.$$

Since $\mu'(\alpha) = \alpha(\mathcal{L}\lambda)''(\alpha)$ and $\mathcal{L}\lambda$ is strictly convex (see Fig. A.3 in Appendix A.2), μ is strictly increasing and therefore invertible. Consequently,

$$\alpha_u(x) = \mu^{-1}\left(\log\left(\frac{u(x)}{u(x^{\uparrow})}\right)\right), \qquad x \in \square.$$

Putting the above formulas together, we arrive at (recall (2.30))

$$\begin{aligned}L(u) &= \inf_{v\colon V_R\to(0,\infty)} L(u,v)\\ &= -\sum_{x\in\mathrm{int}(V_R)} A_u(x) \quad - \sum_{x\in\partial V_R\setminus\star} B_u(x) \quad - C_u(\star) \quad + \sum_{x\in\square} D_u(x)\end{aligned} \tag{2.31}$$

with (recall (2.29))

$$\begin{aligned}D_u(x) &= -\frac{p(x)}{\alpha_u(x)}\left[\log\left(\frac{u(x^{\uparrow})}{u(x)}\right) - (\mathcal{L}\lambda)(\alpha_u(x))\right]\\ &= \frac{p(x)}{\alpha_u(x)}\left[(\mathcal{L}\lambda)(\alpha_u(x)) - \mu(\alpha_u(x))\right]\\ &= p(x)\,(\mathcal{L}\lambda)'(\alpha_u(x)) = p(x)\left((\mathcal{L}\lambda)'\circ\mu^{-1}\right)\left(\log\left(\frac{u(x)}{u(x^{\uparrow})}\right)\right).\end{aligned}$$

In (A.8) in Appendix A.2 we show that $(\mathcal{L}\lambda)'\circ\mu^{-1} = \lambda^{-1}$. Moreover In (A.14) in Appendix A.2 we show that $(\lambda^{-1}\circ\log) = S$ with

$$S(\beta) = d + 1 - \beta - \frac{d}{\beta}, \qquad \beta \in (0, \sqrt{d}].$$

Since S has domain $(0, \sqrt{d}]$, $D_u(x)$ is only defined when $u(x)/u(x^\uparrow) \leq \sqrt{d}$, in which case

$$D_u(x) = p(x)\, S\left(\frac{u(x)}{u(x^\uparrow)}\right), \qquad x \in \square. \tag{2.32}$$

▶ $u(x)/u(x^\uparrow) > \sqrt{d}$: In this case $\frac{\partial L(u,v)}{\partial v(z)} > 0$, the infimum is uniquely taken at $\bar{v}(x) = 0$, and

$$D_u(x) = p(x)\,(\sqrt{d}-1)^2 = p(x)\, S(\sqrt{d}), \qquad x \in \square,$$

where we use (A.13). Note that the right-hand side does not depend on u.

4. Next, we compute the supremum over u. The first derivatives of L are

$$\begin{aligned}
&z \in \mathrm{int}(V_R)\backslash\mathcal{O}: && \frac{\partial L(u)}{\partial u(z)} = \frac{\sum_{y\sim z} u(y)}{u^2(z)}\, p(z) - \sum_{y\sim z} \frac{1}{u(y)}\, p(y),\\
&z = \mathcal{O}: && \frac{\partial L(u)}{\partial u(\mathcal{O})} = \frac{\sum_{y\sim \mathcal{O}} u(y)}{u(\mathcal{O})^2}\, p(\mathcal{O}) - \sum_{y:y^\uparrow=\mathcal{O}} \frac{1}{u(y)} p(y) - \frac{d}{u(\star)}\, p(\star),\\
&z = \star: && \frac{\partial L(u)}{\partial u(\star)} = -\frac{1}{u(\mathcal{O})}\, p(\mathcal{O}) + \frac{u(\star^\uparrow) + du(\mathcal{O})}{u(\star)^2}\, p(\star),\\
&z \in \partial V_R\backslash\star: && \frac{\partial L(u)}{\partial u(z)} = -\frac{1}{u(z^\uparrow)}\, p(z^\uparrow) + \frac{u(z^\uparrow) + du(\square_z)}{u(z)^2}\, p(z)\\
& && \qquad + \left[\frac{u(\square_z)}{u(z)^2} - \frac{d}{u(\square_z)}\right] p(\square_z) 1_{\left\{\frac{u(z)}{u(z^\uparrow)} \leq \sqrt{d}\right\}},\\
&z \in \square: && \frac{\partial L(u)}{\partial u(z)} = -\frac{d}{u(z^\uparrow)}\, p(z^\uparrow) + \left[-\frac{1}{u(z^\uparrow)} + \frac{du(z^\uparrow)}{u(z)^2}\right] p(z)\, 1_{\left\{\frac{u(z)}{u(z^\uparrow)} \leq \sqrt{d}\right\}}.
\end{aligned} \tag{2.33}$$

The second derivates of L are all < 0. The first line in (2.33) can be rewritten as

$$\sum_{y\sim z} u(y) \left[\frac{p(z)}{u^2(z)} - \frac{p(y)}{u^2(y)}\right],$$

which is zero when

$$\bar{u}(x) = \sqrt{p(x)}, \qquad x \in V_R. \tag{2.34}$$

Given the choice in (2.34), the fifth line in (2.33) is zero when

$$\bar{u}(x) = \sqrt{\frac{dp(x^\uparrow)p(x)}{dp(x^\uparrow) + p(x)}}, \qquad x \in \square. \tag{2.35}$$

Indeed, the derivative is strictly negative when the indicator is 0 and therefore the indicator must be 1. But the latter is guaranteed by (2.34)–(2.35), which imply that

$$\frac{\bar{u}(x)}{\bar{u}(x^\uparrow)} = \sqrt{\frac{dp(x)}{dp(x^\uparrow) + p(x)}} \leq \sqrt{d}, \qquad x \in \square.$$

Given the choice in (2.34)–(2.35), also the fourth line in (2.33) is zero. Thus, only the second and third line in (2.33) are non-zero, but this is harmless because $\mathcal{O},\star$ carry a negligible weight in the limit as $R\to\infty$ because of the constraint $p(\partial V_R \cup \mathcal{O}) \leq 1/R$ in Lemma 2.3.7 (recall (2.19)).

Inserting (2.34)–(2.35) into (2.31) and using (2.30), (2.32), we get the following lower bound:

$$
\begin{aligned}
&\sup_{u\colon V_R\to(0,\infty)} L(u)\\
&\geq -\sum_{x\in\mathrm{int}(V_R)} A_{\bar u}(x) \quad -\sum_{x\in\partial V_R\setminus\star} B_{\bar u}(x) \quad - C_{\bar u}(\star) + \sum_{x\in\square} D_{\bar u}(x)\\
&= -\sum_{x\in\mathrm{int}(V_R)}\sum_{y\sim x}\sqrt{p(y)p(x)} - \sum_{x\in\partial V_R\setminus\star}\sqrt{p(x)}\left(\sqrt{p(x^\uparrow)} + d\sqrt{\frac{dp(x)p(\square_x)}{dp(x)+p(\square_x)}}\right)\\
&\quad -\sqrt{p(\star)}\left(\sqrt{p(\star^\uparrow)} + d\sqrt{p(\mathcal{O})}\right)\\
&\quad + \sum_{x\in\square} p(x)\left(d+1-\sqrt{d}\left[\sqrt{\frac{p(x)}{dp(x^\uparrow)+p(x)}} + \sqrt{\frac{dp(x^\uparrow)+p(x)}{p(x)}}\right]\right).
\end{aligned}
$$

5. Using the relation $(d+1)p(x) = \sum_{y\sim x} p(x)$, $x\in\mathrm{int}(V_R)$, we get from (2.28) that

$$
I^\dagger_{E(\mathcal{W}^R)}(p) \geq K^1_R(p) + K^2_R(p)
$$

with

$$
\begin{aligned}
K^1_R(p) &= \sum_{x\in\mathrm{int}(V_R)}\sum_{y\sim x}\left[p(x) - \sqrt{p(x)p(y)}\right]\\
&= \sum_{\{x,y\}\in\widehat{E}_R}\left(\sqrt{p(x)}-\sqrt{p(y)}\right)^2 + \left[p(\mathcal{O}) - \sqrt{p(\mathcal{O})p(\star)}\right] - \sum_{x\in\partial V_R}\left[p(x) - \sqrt{p(x)p(x^\uparrow)}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
K^2_R(p) = &\sum_{x\in\partial V_R\setminus\star}\left[(d+1)p(x) - \sqrt{p(x)}\left(\sqrt{p(x^\uparrow)} + d\sqrt{\frac{dp(x)p(\square_x)}{dp(x)+p(\square_x)}}\right)\right]\\
&+ (d+1)p(\star) - \sqrt{p(\star)}\left(\sqrt{p(\star^\uparrow)} + d\sqrt{p(\mathcal{O})}\right)\\
&+ \sum_{x\in\square} p(x)\left[d+1-\sqrt{d}\left(\sqrt{\frac{p(x)}{dp(x^\uparrow)+p(x)}} + \sqrt{\frac{dp(x^\uparrow)+p(x)}{p(x)}}\right)\right].
\end{aligned}
$$

The first sum in the right-hand side of $K^1_R(p)$ equals the *standard rate function* $I_{\widehat{E}_R}(p)$ given by (2.10), with

$$
\widehat{E}_R = E_R \setminus \{\mathcal{O},\star\}
$$

the set of edges in the unit $\mathcal{W}_R$ *without the tadpoles* and *without the edge* $\{\mathcal{O},\star\}$ (i.e., $\widehat{E}_R = E(\mathcal{T}_R^*)$; recall Fig. 2.4). Rearranging and simplifying terms, we arrive at

$$I^\dagger_{E(\mathcal{W}^R)}(p) \geq I_{\widehat{E}_R}(p) + K_R^3(p) \tag{2.36}$$

with

$$K_R^3(p) = S_{\partial V_R\setminus\star}(p) + S_{\mathcal{O},\star}(p) + S_{(\partial V_R\setminus\star)\cup\square}(p),$$

where

$$\begin{aligned}
S_{\partial V_R\setminus\star}(p) &= d \sum_{x\in\partial V_R\setminus\star} p(x),\\
S_{\mathcal{O},\star}(p) &= \left(\sqrt{p(\mathcal{O})} - \sqrt{p(\star)}\right)^2 + (d-1)\left[p(\star) - \sqrt{p(\mathcal{O})p(\star)}\,\right],\\
S_{(\partial V_R\setminus\star)\cup\square}(p) &= -\sum_{x\in\partial V_R\setminus\star} p(x)\, d\sqrt{\frac{dp(\square_x)}{dp(x)+p(\square_x)}}\\
&\quad + \sum_{x\in\partial V_R\setminus\star} p(\square_x)\left(d+1-\sqrt{d}\left[\sqrt{\frac{p(\square_x)}{dp(x)+p(\square_x)}} + \sqrt{\frac{dp(x)+p(\square_x)}{p(\square_x)}}\right]\right).
\end{aligned} \tag{2.37}$$

6. Since $\sqrt{p(\mathcal{O})p(\star)} \leq \frac{1}{2}[p(\mathcal{O})+p(\star)]$, the boundary constraint $\sum_{x\in\partial V_R\cup\mathcal{O}} p(x) \leq 1/R$ implies that $S_{\partial V_R\setminus\star}(p)+S_{\mathcal{O},\star}(p) = O(1/R)$. The same constraint implies that the first sum in $S_{(\partial V_R\setminus\star)\cup\square}(p)$ is $O(1/R)$. Hence

$$K_R^3(p) = O(1/R) + \sum_{x\in\partial V_R\setminus\star} p(x)\, F\left(\tfrac{p(\square_x)}{p(x)}\right)$$

with

$$F(w) = w\left(d+1-\sqrt{d}\left[\sqrt{\frac{w}{d+w}} + \sqrt{\frac{d+w}{w}}\right]\right).$$

The map $w\mapsto F(w)$ is continuous on $(0,\infty)$ with

$$F(w) = \begin{cases} -\sqrt{w} + (d+1)w + O(w^{3/2}), & w\downarrow 0,\\ [(d+1)-2\sqrt{d}\,]\,w + O(w^{-1}), & w\to\infty. \end{cases}$$

From this we see that if $d\geq 4$, then there exists a $C\in(1,\infty)$ such that

$$F(w) + C \geq \left(1-\sqrt{w}\right)^2, \qquad w\in[0,\infty). \tag{2.38}$$

Hence we have the lower bound

$$\begin{aligned}
K_R^3(p) &\geq O(1/R) + \sum_{x\in\partial V_R\setminus\star} p(x)\left[-C + \left(1-\sqrt{\tfrac{p(\square_x)}{p(x)}}\right)^2\right]\\
&= O(1/R) + \sum_{x\in\partial V_R\setminus\star} \left(\sqrt{p(x)} - \sqrt{p(\square_x)}\right)^2.
\end{aligned}$$

Via (2.36)–(2.37), it follows that

$$I^{\dagger}_{E(\mathcal{W}^R)}(p) \geq O(1/R) + I_{\widetilde{E}_R}(p), \qquad R \in \mathbb{N}, \tag{2.39}$$

with $I_{\widetilde{E}_R}(p)$ the *standard rate function* given by (2.10), with

$$\widetilde{E}_R = \widehat{E}_R \cup \left[\cup_{x \in \partial V_R \setminus \star} \{x, \square_x\}\right]$$

the set of edges in the unit $\widetilde{\mathcal{W}}_R$ that is obtained from the unit $\mathcal{W}_R$ by *removing the edge* $\{\mathcal{O}, \star\}$ (i.e., $\widetilde{E}_R = E(\widetilde{\mathcal{W}}_R)$; recall Fig. 2.5). This completes the proof of (2.26).

Remark 2.4.1. The condition $d \geq 4$ is needed only in (2.38). For $d = 2, 3$ we have $F(w) + C \geq \theta_c(1 - \sqrt{w})^2$ with $\theta_c = d + 1 - 2\sqrt{d} \in (0, 1)$. Consequently, the edges $\{x, \square_x\}$, $x \in \partial V_R \setminus \star$, carry a weight that is smaller than that of the edges in $\mathcal{T}$, which may cause the optimal p to stick to the boundary as $R \to \infty$, in which case we do not have (2.39). ♠

§2.4.2 Limit of the upper variational formula

Note that

$$\widetilde{\mathcal{W}}_R \subseteq \mathcal{T},$$

with $\mathcal{T}$ the infinite tree. Consequently,

$$I_{\widetilde{E}_R}(p) = I_{E(\mathcal{T})}(p) - (d-1) \sum_{x \in \partial V_R \setminus \star} p(x), \qquad \forall\, p \in \mathcal{P}(V(\mathcal{T}))\colon\ \mathrm{supp}(p) = V(\widetilde{\mathcal{W}}_R),$$

where the sum compensates for the contribution coming from the edges in $\mathcal{T}$ that link the vertices in $\partial V_R \setminus \star$ to the vertices one layer deeper in $\mathcal{T}$ that are not tadpoles. Since this sum is $O(1/R)$, we obtain (recall (2.20))

$$\begin{aligned}
\chi^{+}_R(\varrho) &= \inf_{\substack{p \in \mathcal{P}(V(\mathcal{W}_R))\colon \\ p(\partial \mathcal{W}_R) \leq 1/R}} \left\{ I^{\dagger}_{E(\mathcal{W}_R)}(p) + \varrho J_{V(\mathcal{W}_R)}(p) \right\} \\
&\geq O(1/R) + \inf_{\substack{p \in \mathcal{P}(V(\mathcal{T}))\colon \\ \mathrm{supp}(p) = V(\widetilde{\mathcal{W}}_R),\, p(\partial_{\widetilde{\mathcal{W}}_R}) \leq 1/R}} \left\{ I_{E(\mathcal{T})}(p) + \varrho J_{V(\mathcal{T})}(p) \right\} \\
&\geq O(1/R) + \chi_{\mathcal{T}}(\rho),
\end{aligned}$$

where the last inequality follows after dropping the constraint under the infimum and recalling (2.14). This completes the proof of (2.27).

§2.5 Analysis of the variational problem on the infinite regular tree

In this Section we prove Theorem 2.1.4. Section 2.5.1 formulates two theorems that imply Theorem 2.1.4, Section 2.5.2 provides the proof of these theorems. Recall the definition of $\mathcal{P}(V)$, $I_E(p)$ and $J_V(p)$ from (2.10). Set

$$\chi_{\mathcal{T}}(\varrho) = \inf_{p\in\mathcal{P}_{\mathcal{O}}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0,\infty), \tag{2.40}$$

where $\mathcal{P}_{\mathcal{O}}(V) = \{p \in \mathcal{P}(V) \colon \operatorname{argmax} p = \mathcal{O}\}$. Since $\mathcal{P}(V)$, I_E and J_V are invariant under translations, the centering at $\mathcal{O}$ is harmless.

§2.5.1 Two properties

Theorem 2.5.1. *For every $\varrho \in (0,\infty)$ the infimum in (2.40) is attained, and every minimiser $\bar{p}$ is strictly positive, non-increasing along backbones, and such that*

$$\sum_{N\in\mathbb{N}_0} \partial S_R \log(R+1) \le \frac{d+1}{\varrho}, \qquad \partial S_R = \sum_{\partial B_R(\mathcal{O})} \bar{p}(x),$$

where $B_R(\mathcal{O})$ is the ball of radius R around $\mathcal{O}$.

Theorem 2.5.2. *The function $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$ is strictly increasing and globally Lipschitz continuous on $(0,\infty)$, with $\lim_{\varrho\downarrow 0} \chi_{\mathcal{T}}(\varrho) = d-1$ and $\lim_{\varrho\to\infty} \chi_{\mathcal{T}}(\varrho) = d+1$.*

Theorems 2.5.1–2.5.2 settle Theorem 2.1.4. Their proof uses the following two lemmas.

Lemma 2.5.3. *For every $\varrho \in (0,\infty)$, the infimum in (2.40) may be restricted to $p \in \mathcal{P}_{\mathcal{O}}(V)$ such that $J_V(p) \le \frac{d+1}{\varrho}$.*

Proof. Let $\delta_{\mathcal{O}} \in \mathcal{P}_{\mathcal{O}}(V)$ denote the point measure at $\mathcal{O}$. Then, for all $\varrho \in (0,\infty)$,

$$\chi_{\mathcal{T}}(\varrho) \le I_E(\delta_{\mathcal{O}}) + \varrho J_V(\delta_{\mathcal{O}}) = (d+1) + \varrho \times 0 = d+1.$$

Since $I_V \ge 0$, we may restrict the infimum in (2.40) to p with $J_V(p) \le \frac{d+1}{\varrho}$. □

Lemma 2.5.4. *For every $\varrho \in (0,\infty)$, there exists a $c(\varrho) > 0$ such that the infimum in (2.40) may be restricted to $p \in \mathcal{P}_{\mathcal{O}}(V)$ such that $J_V(p) \ge c(\varrho)$.*

Proof. Since $J_V(p) = 0$ if and only if $p = \delta_{\mathcal{O}}$ is a point measure, it suffices to show that $\delta_{\mathcal{O}}$ is not a minimiser of $\chi_{\mathcal{T}}(\varrho)$. To that end, for $y \in V$ compute

$$\frac{\partial}{\partial p(y)} [I_E(p) + \varrho J_V(p)] = 1 - \sum_{z\sim y} \sqrt{\frac{p(z)}{p(y)}} - \varrho \log p(y) - \varrho. \tag{2.41}$$

Because $p(\mathcal{O}) > 0$, it follows that the right-hand side tends to $-\infty$ as $p(y) \downarrow 0$ for every $y \sim \mathcal{O}$. Hence, no $p \in \mathcal{P}_{\mathcal{O}}(V)$ with $p(y) = 0$ for some $y \sim \mathcal{O}$ can be a minimiser of (2.40), or be the weak limit point of a minimising sequence. In particular, $\delta_{\mathcal{O}}$ cannot. □

§2.5.2 Proof of the two properties

Proof of Theorem 2.5.1. First observe that $\mathcal{P}(V)$ and J_V are invariant under permutations, i.e., for any $p \in \mathcal{P}(V)$ and any relabelling π of the vertices in V, we have $\pi p \in \mathcal{P}(V)$ and $J_V(\pi p) = J_V(p)$. The same does not hold for I_E, but we can apply permutations such that $I_E(\pi p) \leq I_E(p)$.

1. Pick any $p \in \mathcal{P}(V)$. Pick any backbone bb $= \{x_0, x_1, \cdots\}$ that runs from $x_0 = \mathcal{O}$ to infinity. Consider a permutation π that *reorders* the vertices in bb such that $\{(\pi p)(x)\}_{x \in \mathrm{bb}}$ becomes *non-increasing*. Together with the reordering, transport all the trees that hang off bb as well. Since πp is non-increasing along bb, while all the edges that do not lie on bb have the same neighbouring values in p and in πp, we have

$$I_E(\pi p) \leq I_E(p). \tag{2.42}$$

Indeed,

$$\tfrac{1}{2}\left[I_E(p) - I_E(\pi p)\right] = \sum_{k \in \mathbb{N}_0} \sqrt{(\pi p)(x_k)(\pi p)(x_{k+1})} - \sum_{k \in \mathbb{N}_0} \sqrt{p(x_k)p(x_{k+1})}, \tag{2.43}$$

where we use that $p(x_0) = (\pi p)(x_0)$ (because $p(x_0) \geq p(x_k)$ for all $k \in \mathbb{N}$) and $\sum_{k \in \mathbb{N}} p(x_k) = \sum_{k \in \mathbb{N}} (\pi p)(x_k)$. The right-hand side of (2.43) is ≥ 0 by the rearrangement inequality for sums of products of two sequences [26, Section 10.2, Theorem 368]. In fact, strict inequality in (2.43) holds unless p is constant along bb. But this is impossible because it would imply that $p(\mathcal{O}) = 0$ and hence $p(x) = 0$ for all $x \in V$. Thus, p and bb being arbitrary, it follows from (2.42) that any minimiser or minimising sequence must be non-increasing in the distance to $\mathcal{O}$. Indeed, if it were not, then there would be a bb along which the reordering would lead to a lower value of $I_E + \varrho J_V$. Hence we may replace (2.40) by

$$\chi_{\mathcal{T}}(\varrho) = \inf_{p \in \mathcal{P}^{\downarrow}_{\mathcal{O}}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0, \infty), \tag{2.44}$$

with $\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$ defined in (2.15).

2. Let $p \in \mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$. Estimate

$$J_V(p) = \sum_{R \in \mathbb{N}_0} \sum_{x \in \partial B_R(\mathcal{O})} [-p(x) \log p(x)] \geq \sum_{R \in \mathbb{N}_0} \sum_{x \in \partial B_R(\mathcal{O})} \Big[-p(x) \log \big(\tfrac{1}{R+1}\big)\Big],$$

where we use that $p(x) \leq \frac{1}{R+1}$ for all $x \in \partial B_R(\mathcal{O})$. Hence

$$J_V(p) \geq \sum_{R \in \mathbb{N}_0} \partial S_R \log(R+1)$$

with $\partial S_R = \sum_{x \in \partial B_R(\mathcal{O})} p(x)$. By Lemma 2.5.3, $J_V(p) \leq \frac{d+1}{\varrho}$, and so

$$\sum_{R \in \mathbb{N}_0} \partial S_R \log(R+1) \leq \frac{d+1}{\varrho}. \tag{2.45}$$

The computation in (2.41) shows that any p for which there exist $z \sim y$ with $p(z) > 0$ and $p(y) = 0$ cannot be minimiser nor a weak limit point of a minimising sequence. Hence all minimisers or weak limit points of minimising sequences are strictly positive everywhere.

3. Take any minimising sequence $(p_n)_{n\in\mathbb{N}}$ of (2.44). By (2.45), $\lim_{R\to\infty}\sum_{x\notin B_R(\mathcal{O})} p_n(x) = 0$ uniformly in $n \in \mathbb{N}$, and so $(p_n)_{n\in\mathbb{N}}$ is tight. By Prokhorov's theorem, tightness is equivalent to $(p_n)_{n\in\mathbb{N}}$ being relatively compact, i.e., there is a subsequence $(p_{n_k})_{k\in\mathbb{N}}$ that converges weakly to a limit $\bar{p} \in \mathcal{P}^{\downarrow}_{\mathcal{O}}(V)$. By Fatou's lemma, we have $\liminf_{k\to\infty} I_E(p_{n_k}) \geq I_E(\bar{p})$ and $\liminf_{k\to\infty} J_V(p_{n_k}) \geq J_V(\bar{p})$. Hence

$$\chi_{\mathcal{T}}(\varrho) = \lim_{k\to\infty} [I_E(p_{n_k}) + \varrho J_V(p_{n_k})] \geq I_E(\bar{p}) + \varrho J_V(\bar{p}).$$

Hence $\bar{p}$ is a minimiser of (2.44). □

Proof of Theorem 2.5.2. The proof uses approximation arguments.

1. We first show that $\varrho \mapsto \chi_{\mathcal{T}}(\varrho)$ is strictly increasing and globally Lipschitz. Pick $\varrho_1 < \varrho_2$. Let $\bar{p}_{\varrho_1}$ be any minimiser of (2.40) at ϱ_1, i.e.,

$$\chi_{\mathcal{T}}(\varrho_1) = I_E(\bar{p}_{\varrho_1}) + \varrho_1 J_V(\bar{p}_{\varrho_1}).$$

Estimate

$$\begin{aligned}&[I_E(\bar{p}_{\varrho_1}) + \varrho_1 J_V(\bar{p}_{\varrho_1})] = [I_E(\bar{p}_{\varrho_1}) + \varrho_2 J_V(\bar{p}_{\varrho_1})] - (\varrho_2 - \varrho_1) J_V(\bar{p}_{\varrho_1})\\ &\geq \chi_{\mathcal{T}}(\varrho_2) - (\varrho_2 - \varrho_1) J_V(\bar{p}_{\varrho_1}) \geq \chi(\varrho_2) - (\varrho_2 - \varrho_1)\tfrac{d+1}{\varrho_1},\end{aligned}$$

where we use Lemma 2.5.3. Therefore

$$\chi_{\mathcal{T}}(\varrho_2) - \chi_{\mathcal{T}}(\varrho_1) \leq (\varrho_2 - \varrho_1)\tfrac{d+1}{\varrho_1}.$$

Similarly, let $\bar{p}_{\varrho_2}$ be any minimiser of (2.40) at ϱ_2, i.e.,

$$\chi_{\mathcal{T}}(\varrho_2) = I_E(\bar{p}_{\varrho_2}) + \varrho_2 J_V(\bar{p}_{\varrho_2}).$$

Estimate

$$\begin{aligned}&[I_E(\bar{p}_{\varrho_2}) + \varrho_2 J_V(\bar{p}_{\varrho_2})] = [I_E(\bar{p}_{\varrho_2}) + \varrho_1 J_V(\bar{p}_{\varrho_2})] + (\varrho_2 - \varrho_1) J_V(\bar{p}_{\varrho_2})\\ &\geq \chi_{\mathcal{T}}(\varrho_1) + (\varrho_2 - \varrho_1) J_V(\bar{p}_{\varrho_2}) \geq \chi_{\mathcal{T}}(\varrho_1) + (\varrho_2 - \varrho_1) c(\varrho_2),\end{aligned}$$

where we use Lemma 2.5.4. Therefore

$$\chi_{\mathcal{T}}(\varrho_2) - \chi_{\mathcal{T}}(\varrho_1) \geq c(\varrho_2)(\varrho_2 - \varrho_1).$$

2. Because $\chi_{\mathcal{T}}(\varrho) \leq d+1$ for all $\varrho \in (0,\infty)$, it follows that $\lim_{\varrho\to\infty}\chi_{\mathcal{T}}(\varrho) \leq d+1$. To obtain the reverse inequality, let $\bar{p}_\varrho$ be any minimiser of (2.44) at ϱ. By Lemma 2.5.3, we may assume that $J_V(\bar{p}_\varrho) \leq \frac{d+1}{\varrho}$. Hence $\lim_{\varrho\to\infty} J_V(\bar{p}_\varrho) = 0$, and consequently $\lim_{\varrho\to\infty}\bar{p}_\varrho = \delta_{\mathcal{O}}$ weakly. Therefore, by Fatou's lemma, $\lim_{\varrho\to\infty}\chi_{\mathcal{T}}(\varrho) = \lim_{\varrho\to\infty}[I_E(\bar{p}) + \varrho J_V(\bar{p})] \geq \liminf_{\varrho\to\infty} I_E(\bar{p}_\varrho) \geq I_E(\delta_{\mathcal{O}}) = d+1$.

3. To prove that $\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) \le d-1$, estimate

$$\chi_{\mathcal{T}}(\varrho) \le \inf_{\substack{p\in\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)\\ \mathrm{supp}(p)\subseteq B_R(\mathcal{O})}} [I_E(p)+\varrho J_V(p)], \qquad R\in\mathbb{N}_0.$$

Because

$$\sup_{\substack{p\in\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)\\ \mathrm{supp}(p)\subseteq B_R(\mathcal{O})}} J_V(p) = J_V(p_R) = \log|B_R(\mathcal{O})|, \qquad R\in\mathbb{N}_0,$$

with

$$p_R(x) = \begin{cases} |B_R(\mathcal{O})|^{-1}, & x\in B_R(\mathcal{O}),\\ 0, & \text{else,}\end{cases}$$

it follows that

$$\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) \le \inf_{\substack{p\in\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)\\ \mathrm{supp}(p)\subseteq B_R(\mathcal{O})}} I_E(p) \le I_E(p_R), \qquad R\in\mathbb{N}_0.$$

Compute (recall (2.10)) ,

$$I_E(p_R) = \frac{|\partial B_{R+1}(\mathcal{O})|}{|B_R(\mathcal{O})|}, \qquad R\in\mathbb{N}_0.$$

Inserting the relations

$$|\partial B_R(\mathcal{O})| = \begin{cases} 1, & R=0,\\ (d+1)d^{R-1}, & R\in\mathbb{N},\end{cases}$$

$$|B_R(\mathcal{O})| = \sum_{R'=0}^{R} |\partial B_{R'}(\mathcal{O})| = 1+\frac{d+1}{d-1}(d^R-1), \quad R\in\mathbb{N}_0,$$

we get

$$I_E(p_R) = (d-1)\,\frac{(d+1)d^R}{(d+1)d^R-2}.$$

Hence $\lim_{R\to\infty} I_E(p_R) = d-1$, and so $\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) \le d-1$.

4. To prove that $\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) \ge d-1$, note that because $J_V\ge 0$ we can estimate

$$\lim_{\varrho\downarrow 0}\chi_{\mathcal{T}}(\varrho) \ge \inf_{p\in\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)} I_E(p).$$

It therefore suffices to show that

$$\inf_{p\in\mathcal{P}^{\downarrow}_{\mathcal{O}}(V)} I_E(p) \ge d-1,$$

i.e., $(p_R)_{R\in\mathbb{N}_0}$ is a minimising sequence of the infimum in the left-hand side. The proof goes as follows. Write (recall (2.10))

$$\begin{aligned} I_E(p) &= \tfrac12 \sum_{\substack{x,y\in V\\ x\sim y}} \left(\sqrt{p(x)}-\sqrt{p(y)}\right)^2\\ &= \tfrac12 \sum_{\substack{x,y\in V\\ x\sim y}} \left[p(x)+p(y)-2\sqrt{p(x)p(y)}\right] = (d+1) - \sum_{\substack{x,y\in V\\ x\sim y}} \sqrt{p(x)p(y)}. \end{aligned}$$

Since $\mathcal{T}$ is a tree, each edge can be labelled by the end-vertex that is farthest from $\mathcal{O}$. Hence the sum in the right-hand side can be written as

$$\sum_{x \in V \setminus \mathcal{O}} 2\sqrt{p(x)p(x^{\downarrow})},$$

where $x^{\downarrow}$ is the unique neighbour of x that is closer to $\mathcal{O}$ than x. Since $2\sqrt{p(x)p(x^{\downarrow})} \leq p(x) + p(x^{\downarrow})$, it follows that

$$\sum_{x \in V \setminus \mathcal{O}} 2\sqrt{p(x)p(x^{\downarrow})} \leq \sum_{x \in V \setminus \mathcal{O}} p(x) + \sum_{x \in V \setminus \mathcal{O}} p(x^{\downarrow}) = [1 - p(\mathcal{O})] + 1.$$

Therefore

$$I_E(p) \geq d - 1 + p(\mathcal{O}),$$

which settles the claim. $\square$

CHAPTER 3

The parabolic Anderson model on a periodic Galton-Watson tree

Abstract

In [14], the annealed total mass of the solution to the parabolic Anderson model on a regular tree with an i.i.d. double-exponential random potential was studied. The first two terms in the asymptotic expansion for large time of the total mass was identified. This chapter extends the analysis to a periodic Galton-Watson tree with large periodicity and is therefore a crucial step towards understanding the annealed total mass on the regular Galton-Watson tree. To do this we need to carefully deal with the non-homogeneity of the periodic Galton-Watson tree.

§3.1 Introduction and main results

In [12], the quenched asymptotic growth rate of the total mass of the PAM on a Galton-Watson tree was indentified. The ultimate goal in this chapter is to obtain the corresponding result in the annealed setting, i.e. the growth rate when averaged over the potential. Chapter 2 made the first step in this direction by considering the regular tree. This chapter considers the Galton-Watson tree with large periodicity, which approximates the full Galton-Watson tree and is therefore another step towards our goal. The periodicity allows us to build on the previous chapter, where the key techniques developed heavily rely on an underlying periodic structure. The main challenge here is to navigate the non-homogeneity of the periodic Galton-Watson tree, which has to be dealt with carefully.

In Section 3.1, the periodic Galton-Watson tree is defined, and the PAM and key quantities are introduced. In Section 3.2 the main theorem is stated. Sections 3.3 and 3.4 concern the proof of the main theorem through a lower, and respectively an upper bound.

§3.1.1 Definition of the periodic Galton-Watson tree

We analyse the PAM on the graph generated by a 'periodic' Galton-Watson tree. The graph is generated by first taking $\mathbb{Z}$, and looking at $[N] = \{0, 1, \cdots, N\} \subset \mathbb{Z}$. From each $x \in [N]$, independently generate a Galton-Watson tree with offspring distribution D, except at the roots $x \in [N]$ where the offspring distribution is $D - 1$. Let the tree rooted at x be denoted by $\mathcal{T}_x$. For $x \in \mathbb{Z} \setminus [N]$, repeat the same Galton-Watson tree that was generated at $x \mod (N+1)$. In other words, the trees $\mathcal{T}_x$ and $\mathcal{T}_y$ are equal if $x = y \mod (N+1)$. This can be equivalently viewed as the infinite concatenation of $[N]$ with all the generated Galton-Watson trees hanging off. Denote the resulting tree by $\mathcal{GW} = (V, E, 0)$ and the probability with respect to $\mathcal{GW}$ by $\mathfrak{P}$. Note that the standard Galton-Watson tree corresponds to $N = \infty$, so that there is no periodicity, or alternatively, each vertex in $\mathbb{Z}$ has an independent Galton-Watson hanging off it, and not just the vertices $[N]$.

§3.1.2 The PAM on a periodic Galton-Watson tree

See (1.14) for the definition of the PAM on a general (locally-finite) graph and other relevant notation. Recall that the annealed total mass is given by the Feynman-Kac representation

$$\langle U(t) \rangle = \left\langle \mathbb{E}_0 \left(\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \right) \right\rangle. \tag{3.1}$$

where $\langle \cdot \rangle$ denotes expectation with respect to the potential $\xi = (\xi(x))_{x \in \mathcal{GW}}$, $X = (X_t)_{t \geq 0}$ a continuous-time random walk on the vertices V with jump rate 1 along the edges E, and $\mathbb{P}_0$ denotes the law of X given $X_0 = 0$.

§3.1.3 Assumptions on the potential

We make the same assumptions as in the previous chapter. Throughout the paper we assume that the random potential ξ consists of i.i.d. random variables with a marginal distribution whose cumulant generating function

$$H(u) = \log \left\langle \mathrm{e}^{u\xi(\mathcal{O})} \right\rangle \tag{3.2}$$

satisfies the following:

Assumption 3.A. [Asymptotic double-exponential potential]
There exists a $\varrho \in (0, \infty)$ such that

$$\lim_{u \to \infty} u H''(u) = \varrho. \tag{3.3}$$

■

It will be useful later on to observe that, since ξ is assumed to be i.i.d., we have from (3.1)-(3.2) that

$$\langle U(t) \rangle = \mathbb{E}_{\mathcal{O}} \left(\exp \left[\sum_{x \in V} H(\ell_t(x)) \right] \right), \tag{3.4}$$

where

$$\ell_t(x) = \int_0^t \mathbb{1}\{X_s = x\} \, \mathrm{d}s, \qquad x \in V, \, t \geq 0,$$

is the local time of X at vertex x up to time t. See Chapter 2.1.2 for further details.

§3.2 Main theorem: annealed total mass for large times

To state the main theorem, we introduce the following *characteristic variational formula*. Denote by $\mathcal{P}(V)$ the set of probability measures on V. For $p \in \mathcal{P}(V)$, define

$$I_E(p) = \sum_{\{x,y\} \in E} \left(\sqrt{p(x)} - \sqrt{p(y)} \right)^2, \qquad J_V(p) = -\sum_{x \in V} p(x) \log p(x), \tag{3.5}$$

and set

$$\chi_{\mathcal{GW}}(\varrho) = \inf_{p \in \mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0, \infty). \tag{3.6}$$

The first term in (3.5) is the quadratic form associated with the Laplacian, which is the large deviation rate function for the *empirical distribution*

$$L_t = \frac{1}{t} \int_0^t \delta_{X_s} \, \mathrm{d}s = \frac{1}{t} \sum_{x \in V} \ell_t(x) \delta_x \in \mathcal{P}(V). \tag{3.7}$$

The following lemma links the Feynman-Kac formula (3.4) to the main theorem, as well as introduces several quantities that are used later on in the proof. The lemma pulls the leading order term out of the expansion and shows that the second order term is controlled by the large deviation principle for the empirical distribution of the normalised local times.

Lemma 3.2.1. *[Key object for the expansion] If $G = (V, E)$ is finite, then*

$$\langle U(t)\rangle = \mathrm{e}^{H(t)+o(t)}\, \mathbb{E}_{\mathcal{O}}\left(\mathrm{e}^{-\varrho t J_V(L_t)}\right), \qquad t \to \infty, \tag{3.8}$$

where J_V is the functional in (3.6) *and L_t is the empirical distribution in* (3.7).

Proof. See Lemma 2.1.2 in Chapter 2 □

For the offspring distribution D, denote its support by $\mathrm{supp}(D)$. For technical reasons we require the following assumption.

Assumption 3.B. [Offspring distribution]
There exist $4 \le d^- \le d^+ < \infty$ such that $\mathrm{supp}(D) \subseteq [d^-, d^+]$. ■

Theorem 3.2.2. *Subject to Assumptions 3.A and 3.B, as $t \to \infty$,*

$$\frac{1}{t}\log\langle U(t)\rangle = \varrho\log(\varrho t) - \varrho - \chi_{\mathcal{GW}}(\varrho) + o(1), \qquad \mathfrak{P}\text{-}a.s. \tag{3.9}$$

The variational formula depends on the realisation of $\mathcal{GW}$ and hence is a random object. However, as will become clear from later analysis, it is actually deterministic and equal to the one of the embedded regular tree with degrees $d^- + 1$. The latter has been studied in Chapter 2.5.

§3.3 Proof of main theorem: lower bound

A lower bound for (3.9) in Theorem 3.2.2 is obtained though a standard and straightforward argument. Let $B_R(0) \subset \mathcal{GW}$ be the ball of radius R around 0 in the graph distance. We consider a random walk that is killed when it leaves $B_R(0)$. This gives

$$\frac{1}{t}\log\langle U(t)\rangle \ge \varrho\log(\varrho t) - \varrho - \chi_R^-(\varrho) + o(1), \qquad t \to \infty,$$

with $\chi_R^-(\varrho)$ the variational formula on $B_R(0)$ with zero boundary condition. It is easily shown that $\chi_R^-(\varrho) \to \chi_{\mathcal{GW}}(\varrho)$ as $R \to \infty$. Hence, letting $R \to \infty$ we get the desired lower bound. See Chapter 2.2 for technical details. The inhomogeneity of the periodic Galton-Watson tree plays no role in the argument, which carries over exactly.

§3.4 Proof of main theorem: upper bound

In this section we prove the upper bound for (3.9) in Theorem 3.2.2. We try to follow the argument used on the regular tree in the previous chapter. Again, the argument is comprised of four steps:

(I) *Condition on the backbone* of X (Section 3.5.1).

(II) *Project* X onto a concatenation of finite subtrees attached to this backbone that have depth R and special *tadpoles* at the bottom. (Section 3.5.2).

(III) *Periodise* the projected X to obtain a *Markov renewal process* on a finite graph (Section 3.5.3).

(IV) Use the *large deviation principle* for the empirical distribution of *Markov renewal processes* derived in [31] to obtain a variational formula on a single subtree (Section 3.5.4).

Finally, in Section 3.6 we derive the upper bound of the expansion by letting $R \to \infty$ in the variational formula.

§3.5 Backbone, projection, periodisation and upper variational formula

§3.5.1 Backbone

Although the intuition is the same as on the regular tree, the backbone has to be defined more carefully due to the inhomogeneity. Since X is transient, it escapes to infinity along a path in $\mathcal{T}_n$ for some $n \in \mathbb{Z}$. We can assume $n \in \mathbb{N}_0$, for if not, we can reflect the labelling in Section 3.1.1 so that this is the case. For $k \in \mathbb{Z}_+$, let Z_k denote all the vertices at distance k from 0 *and* belonging to $\cup_{l \in \mathbb{Z}_+} \mathcal{T}_l$. Similarly for $k \in \mathbb{Z}_-$, Z_k denotes all the vertices at distance k from 0, *and* belonging to $\cup_{l \in \mathbb{Z}_-} \mathcal{T}_l$. For fixed $k \in \mathbb{Z}$, define T_k to be the last time X visits Z_k, i.e.

$$T_k = \sup\{t > 0 : X_t \in Z_k\}.$$

The backbone is the sequence of vertices $(X_{T_k})_{k \geq 0}$ with the convention that $X_{T_0} = 0$, and can be interpreted as the path along which the random walk escapes to infinity. The following lemma shows that we may assume that the backbone can be taken equal to $\mathbb{Z}$ without loss of generality.

Lemma 3.5.1. *For every $t \geq 0$,*

$$\mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x \in V(\mathcal{GW})} H(\ell_t(x))\right]\right) = \mathbb{E}_{\mathcal{O}}\left(\exp\left[\sum_{x \in V(\mathcal{GW})} H(\ell_t(x))\right] \,\middle|\, (X_{T_k})_{k \geq 0} = \mathbb{Z}\right)$$

in distribution.

Proof. We apply the following permutations to $\mathcal{GW}$ that turns the backbone into $\mathbb{Z}$. We do the permutations inductively on k as follows:

- Since X_{T_0} is already in $\mathbb{Z}$ by convention, we do not permute it.

- Suppose that $\mathcal{GW}$ has been permuted such that $X_{T_i} = i \in \mathbb{Z}$ for all $i = 0, \cdots, j-1$. Then X_{T_j} is in the subtree rooted at $X_{T_{j-1}}$. If $X_{T_j} \neq j \in \mathbb{Z}$, then we can swap it and the tree hanging below it with j and the tree hanging below it. In other words, we swap $\mathcal{T}_{X_{T_j}}$ with $\mathcal{T}_j$.

This permutation procedure preserves the edges and the vertices and so the resulting tree is isomorphic to $\mathcal{GW}$. Therefore, conditioning on the backbone being $\mathbb{Z}$ does not affect the distribution of the total mass given in (3.4). □

§3.5.2 Projection

For every vertex that is distance R from the backbone, replace the tree hanging below it by a special *tadpole* vertex. R is chosen such that it is a multiple of N, i.e. $R = nN$ for some $n \in \mathbb{N}$, with N from Section 3.1.1. Denote this truncated version of $\mathcal{GW}$ by $\mathcal{GW}_R$. We apply the following map to X. Whenever X travels farther than distance R from the backbone, its excursion is cut off and replaced by a sojourn time at the corresponding tadpole. The resulting path, which we call $X^R = (X_t^R)_{t\geq 0}$, is a Markov renewal process on $\mathcal{GW}_R$ with the following properties:

- The sojourn times in all the vertices that are not tadpoles have distribution $\mathrm{EXP}(D_x + 1)$.

- The sojourn times in all the tadpole attached to vertex x have distribution ψ_x, defined as the conditional distribution of the *return time* τ_x of the random walk on the Galton-Watson tree rooted at x *given* that $\tau_x < \infty$ (again see [30] for a proper definition).

- The transitions into the tadpoles have probability $\frac{D_x}{D_x+1}$, the transitions out of the tadpoles have probability 1 (because of the conditioning on the backbone). Here D_x denotes the number of offspring of x, which has the same distribution as D (and $D_x + 1$ is the degree of the vertex).

Write $(\ell_t^{\mathcal{GW}_R}(x))_{x\in V_{\mathcal{GW}_R}}$ to denote the local times of X^R at time t.

Lemma 3.5.2. *[Projection onto tadpoles] For every $R \in (N\mathbb{Z}) \cap \mathbb{N}$ and $t \geq 0$,*

$$\mathbb{E}_0\left(\exp\left[\sum_{x\in V(\mathcal{GW})} H(\ell_t(x))\right]\right) \leq \mathbb{E}_0\left(\exp\left[\sum_{x\in V(\mathcal{GW}_R)} H(\ell_t^{\mathcal{GW}_R}(x))\right]\right).$$

Proof. See Lemma 2.3.3. The map stacks local times on top of each other and the inequality follows from the convexity of H defined in (3.2). □

§3.5.3 Periodisation

We cut $\mathcal{GW}_R$ into periodic units of length $R+1$ with the same R as in Section 3.5.2. We do this in the natural way following the inherent periodic structure.By the contruction of the perodic Galton-Waton tree, we can fold all paths of X^R that are in $\mathcal{T}_m$ into paths in $\mathcal{T}_x$ where $m = x \mod (R+1)$ for all $m \in \mathbb{Z}$ since the trees are identical. Finally, paths that go from $R \mod (R+1)$ to $0 \mod (R+1)$ are folded by adding an additional edge between 0 and R. Denote the periodised graph by $\mathcal{GW}_{\pi,R}$ and the random walk on $\mathcal{GW}_{\pi,R}$ by $X^{\pi,R}$, which is obtained by folding X^R. Write $(\ell_t^{\pi,R}(x))_{x\in V_{\mathcal{GW}_R}}$ to denote the local times of $X^{\pi,R}$ at time t.

Lemma 3.5.3. *[Periodisation to a finite graph] For every $R \in (N\mathbb{Z}) \cap \mathbb{N}$ and $t \geq 0$,*

$$\mathbb{E}_0\left(\exp\left[\sum_{x\in V(\mathcal{GW}_R)} H(\ell_t^{GW_R}(x))\right]\right) \leq \mathbb{E}_0\left(\exp\left[\sum_{x\in V(\mathcal{GW}_{\pi,R})} H(\ell_t^{\pi,R}(x))\right]\right).$$

Proof. The periodisation again stacks local times on top of each other. □

Crucial observation. Let ∂V_R be the set of vertices to which a tadpole is attached. Due to shift invariance, we may assume that the total local time spent at $\partial V_R \cup 0 \cup \star$, is at most t/R, without loss of generality. See Lemma 2.3.6 for more details.

§3.5.4 Large deviation rate function

We use the following large deviation principle for Markov renewal processes derived in [31].

To simplify notation, we define

$\star$	=	vertex R on the backbone,
$\square$	=	set of *tadpoles*,
∂V_R	=	set of vertices neighbouring $\square$,
$\text{int}(V_R)$	=	$V_R \setminus (\square \cup \partial V_R)$,
$\square_x$	=	tadpole attached to $x \in \partial V_R \backslash \star$.

1. For $x \notin \square$, $\psi_x = \text{EXP}(D_x + 1)$, and so

$$\lambda_x(\theta) = \begin{cases} \log\left(\frac{D_x+1}{D_x+1-\theta}\right), & \theta < D_x + 1, \\ \infty, & \theta \geq D_x + 1. \end{cases}$$

To find $\lambda_x^*(\alpha)$ we compute

$$\frac{\partial}{\partial\theta}[\alpha\theta - \log(\tfrac{D_x+1}{D_x+1-\theta})] = \alpha - \frac{1}{D_x+1-\theta}, \qquad \frac{\partial^2}{\partial\theta^2}[\alpha\theta - \log(\tfrac{D_x}{D_x+1-\theta})] = -\frac{1}{(D_x+1-\theta)^2} < 0.$$

This gives that the supremum in (A.4) is uniquely taken at

$$\theta^* = D_x + 1 - \tfrac{1}{\alpha}, \qquad \alpha > 0,$$

so that

$$\lambda_x^*(\alpha) = \alpha(D_x + 1) - \log[\alpha(D_x + 1)] - 1, \qquad \alpha > 0. \tag{3.10}$$

2. Inserting (3.10) into (A.1)–(A.3), we get

$$I^{\dagger}_{E_R}(p) = \sum_{x\in V_R}\sum_{y\sim x} p(x) + \inf_{\beta\in(0,\infty)}\inf_{q\in\mathcal{P}(V_R)}\sup_{\hat{q}\in\mathcal{P}(V_R)} L(\beta,q,\hat{q}\mid p),$$

where we recall that $y\sim x$ means that x and y are connected by an edge in $\mathcal{GW}_{\pi,R}$ (denoted by E_R), and

$$L(\beta,q,\widehat{q}\mid p) = -A-B-C-D,$$

where

$$\begin{aligned}
A &= \sum_{x\in\mathrm{int}(V_R)} \beta q(x)\left\{1+\log\left(\frac{\sum_{y\sim x}\widehat{q}(y)}{\widehat{q}(x)}\,\frac{p(x)}{\beta q(x)}\right)\right\},\\
B &= \sum_{x\in\partial V_R\setminus\star} \beta q(x)\left\{1+\log\left(\frac{\widehat{q}(x^{\uparrow})+D_x\widehat{q}(\square_x)}{\widehat{q}(x)}\,\frac{p(x)}{\beta q(x)}\right)\right\},\\
C &= \beta q(\star)\left\{1+\log\left(\frac{\widehat{q}(\star^{\uparrow})+D_\star\widehat{q}(\mathcal{O})}{\widehat{q}(\star)}\,\frac{p(\star)}{\beta q(\star)}\right)\right\},\\
D &= \sum_{x\in\square}\beta q(x)\left\{\log\left(\frac{\widehat{q}(x^{\uparrow})}{\widehat{q}(x)}\right)-(\mathcal{L}\lambda_x)\left(\frac{p(x)}{\beta q(x)}\right)\right\},
\end{aligned}$$

with $\mathcal{L}\lambda_x$ the Legendre transform of the cumulant generating function of ψ_x and $x^{\uparrow}$ the unique vertex to which x is attached upwards. Note that A,B,C each combine two terms, and that A,B,C,D depend on p. We suppress this dependence because p is fixed.

3. Inserting the parametrisation $\widehat{q}=u/\|u\|_1$ and $q=v/\|v\|_1$ with $u,v\colon V_R\to(0,\infty)$ and putting $\beta q = v$, we may write

$$I^{\dagger}_{E_R}(p) = \sum_{x\in V_R}(D_x+1)p(x) + \inf_{v\colon V_R\to(0,\infty)}\sup_{u\colon V_R\to(0,\infty)} L(u,v) \tag{3.11}$$

with

$$L(u,v) = -A-B-C-D,$$

where

$$\begin{aligned}
A &= \sum_{x\in\mathrm{int}(V_R)} v(x)\left\{1+\log\left(\frac{\sum_{y\sim x}u(y)}{u(x)}\,\frac{p(x)}{v(x)}\right)\right\},\\
B &= \sum_{x\in\partial V_R\setminus\star} v(x)\left\{1+\log\left(\frac{u(x^{\uparrow})+D_xu(\square_x)}{u(x)}\,\frac{p(x)}{v(x)}\right)\right\},\\
C &= v(\star)\left\{1+\log\left(\frac{u(\star^{\uparrow})+D_xu(\mathcal{O})}{u(\star)}\,\frac{p(\star)}{v(\star)}\right)\right\},\\
D &= \sum_{x\in\square} v(x)\left\{\log\left(\frac{u(x^{\uparrow})}{u(x)}\right)-(\mathcal{L}\lambda_x)\left(\frac{p(x)}{v(x)}\right)\right\}.
\end{aligned} \tag{3.12}$$

Our task is to carry out the supremum over u and the infimum over v in (3.11).

4. First, we compute the infimum over v for fixed u. (Later we will make a judicious choice for u to obtain a lower bound.) Abbreviate

$$\begin{aligned} A_u(x) &= \frac{\sum_{y\sim x} u(y)}{u(x)}\, p(x), && x \in \mathrm{int}(V_R),\\ B_u(x) &= \frac{u(x^{\uparrow}) + D_x u(\Box_x)}{u(x)}\, p(x), && x \in \partial V_R \backslash \star, \\ C_u(\star) &= \frac{u(\star^{\uparrow}) + D_\star u(\mathcal{O})}{u(\star)}\, p(\star). \end{aligned} \tag{3.13}$$

• For $z \in V_R$, the first derivatives of L are

$$\begin{aligned} z \in \mathrm{int}(V_R)\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{A_u(z)}{v(z)}\right),\\ z \in \partial V_R\backslash\star\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{B_u(z)}{v(z)}\right),\\ z = \star\colon &\quad \frac{\partial L(u,v)}{\partial v(z)} = -\log\left(\frac{C_u(z)}{v(z)}\right), \end{aligned}$$

while the second derivatives of L equal $1/v(z) > 0$. Hence the infimum is uniquely taken at

$$\begin{aligned} x \in \mathrm{int}(V_R)\colon &\quad \bar{v}(x) = A_u(x),\\ x \in V_R\backslash\star\colon &\quad \bar{v}(x) = B_u(x),\\ x = \star\colon &\quad \bar{v}(x) = C_u(x). \end{aligned}$$

• For $z \in \Box$, the computation is more delicate. Define

$$\mu_x(\alpha) = \alpha(\mathcal{L}\lambda_x)'(\alpha) - (\mathcal{L}\lambda_x)(\alpha).$$

The function μ_x has range $(-\infty, \log M_x]$ for some $M_x < \infty$. The maximal value is uniquely taken at $\alpha = \infty$. Therefore there are two cases.

► $u(z)/u(z^{\uparrow}) \leq M_z$: Abbreviate $\alpha_u(z) = p(z)/v(z)$. For $z \in \Box$,

$$\begin{aligned} \frac{\partial L(u,v)}{\partial v(z)} &= \log\left(\frac{u(z)}{u(z^{\uparrow})}\right) + (\mathcal{L}\lambda_z)\left(\frac{p(z)}{v(z)}\right) - \frac{p(z)}{v(z)}(\mathcal{L}\lambda_z)'\left(\frac{p(z)}{v(z)}\right)\\ &= \log\left(\frac{u(z)}{u(z^{\uparrow})}\right) - \mu_z(\alpha_u(z)),\\ \frac{\partial^2 L(u,v)}{v(z)^2} &= \frac{p^2(z)}{v^3(z)}(\mathcal{L}\lambda_z)''\left(\frac{p(z)}{v(z)}\right) > 0, \end{aligned}$$

where we use that $\mathcal{L}\lambda_x$, being a Legendre transform, is strictly convex. Hence the infimum is uniquely taken at

$$\bar{v}(x) = \frac{p(x)}{\alpha_u(x)}, \qquad x \in \Box,$$

with $\alpha_u(x)$ solving the equation

$$\log\left(\frac{u(x)}{u(x^\uparrow)}\right) = \mu_x(\alpha_u(x)), \qquad x \in \square.$$

Since $\mu_x'(\alpha) = \alpha(\mathcal{L}\lambda_x)''(\alpha)$ and $\mathcal{L}\lambda_x$ is strictly convex (see Fig. A.3 in Appendix A.2), μ_x is strictly increasing and therefore invertible. Consequently,

$$\alpha_u(x) = \mu_x^{-1}\left(\log\left(\frac{u(x)}{u(x^\uparrow)}\right)\right), \qquad x \in \square.$$

Putting the above formulas together, we arrive at (recall (3.13))

$$\begin{aligned} L(u) &= \inf_{v\colon V_R \to (0,\infty)} L(u,v) \\ &= -\sum_{x\in \mathrm{int}(V_R)} A_u(x) \quad - \sum_{x\in\partial V_R\setminus\star} B_u(x) \quad - C_u(\star) \quad + \sum_{x\in\square} D_u(x) \end{aligned} \tag{3.14}$$

with (recall (3.12))

$$\begin{aligned} D_u(x) &= -\frac{p(x)}{\alpha_u(x)}\left[\log\left(\frac{u(x^\uparrow)}{u(x)}\right) - (\mathcal{L}\lambda_x)(\alpha_u(x))\right] \\ &= \frac{p(x)}{\alpha_u(x)}\left[(\mathcal{L}\lambda_x)(\alpha_u(x)) - \mu_x(\alpha_u(x))\right] \\ &= p(x)\,(\mathcal{L}\lambda_x)'(\alpha_u(x)) = p(x)\left((\mathcal{L}\lambda_x)'\circ\mu_x^{-1}\right)\left(\log\left(\frac{u(x)}{u(x^\uparrow)}\right)\right). \end{aligned}$$

► $u(x)/u(x^\uparrow) > M_z$: In this case $\frac{\partial L(u,v)}{\partial v(z)} > 0$, the infimum is uniquely taken at $\bar{v}(x) = 0$, and

$$D_u(x) = p(x)\,\theta_{c,x}, \qquad x \in \square,$$

where we use (A.13). Note that the right-hand side does not depend on u.

5. Recall that $(\mathcal{L}\lambda_x)'(\alpha) = \theta_x^*(\alpha)$ and let $\alpha_x = \mu_x^{-1}(\log(\frac{u(x)}{u(x^\uparrow)}))$. For $u(x)/u(x^\uparrow) \in [1, M_x]$,

$$\theta_x^*(\alpha_x) \geq \theta_{\min}^*(\alpha_x),$$

while for $u(x)/u(x^\uparrow) \in (-\infty, 1)$,

$$\theta_x^*(\alpha_x) < \theta^*(\alpha_x).$$

This follows from the fact that the sojourn time at x is stochastically smaller than that on the minimal tree. It also follows that

$$\theta_{c,x} \geq \theta_{c,\min}.$$

We may therefore estimate

$$L(u) \geq \bar{L}(u) := -\sum_{x\in \mathrm{int}(V_R)} A_u(x) \quad - \sum_{x\in\partial V_R\setminus\star} B_u(x) \quad - C_u(\star) \quad + \sum_{x\in\square} \bar{D}_u(x),$$

where

$$\bar{D}_u(x) = \begin{cases} p(x)\left((\mathcal{L}\lambda_{\max})' \circ \mu_{\max}^{-1}\right)\left(\log\left(\frac{u(x)}{u(x^\uparrow)}\right)\right), & u(x)/u(x^\uparrow) \in (-\infty, 1), \\ p(x)\left((\mathcal{L}\lambda_{\min})' \circ \mu_{\min}^{-1}\right)\left(\log\left(\frac{u(x)}{u(x^\uparrow)}\right)\right), & u(x)/u(x^\uparrow) \in [1, \sqrt{d^-}], \\ p(x)\theta_{c,\min}, & u(x)/u(x^\uparrow) > \sqrt{d^-}. \end{cases}$$

To simplify notation, we suppress min from the notation. For $u(x)/u(x^\uparrow) \in [1, d^-]$, it can be shown that $(\mathcal{L}\lambda)' \circ \mu^{-1} = \lambda^{-1}$. Moreover, $(\lambda^{-1} \circ \log) = S$ with

$$S(\beta) = d^- + 1 - \beta - \frac{d^-}{\beta}, \qquad \beta \in (0, \sqrt{d^-}\,]. \tag{3.15}$$

Since S has domain $(0, \sqrt{d^-}\,]$, $D_u(x)$ is only defined when $u(x)/u(x^\uparrow) \le \sqrt{d^-}$, in which case

$$\bar{D}_u(x) = p(x)\, S\left(\frac{u(x)}{u(x^\uparrow)}\right), \qquad x \in \square. \tag{3.16}$$

On the other hand for $u(x)/u(x^\uparrow) > d^-$,

$$\bar{D}_u(x) = p(x)\,(\sqrt{d^-} - 1)^2 = p(x)\, S(\sqrt{d^-}), \qquad x \in \square.$$

For $u(x)/u(x^\uparrow) \in (-\infty, 1)$, (3.15) and (3.16) hold with d^- replaced by d^+.

6. Next, we compute the supremum over u for $\bar{L}$. The first derivatives of $\bar{L}$ are

$$\begin{aligned}
z \in \mathrm{int}(V_R)\backslash\mathcal{O}: \quad & \frac{\partial \bar{L}(u)}{\partial u(z)} = \frac{\sum_{y\sim z} u(y)}{u^2(z)}\, p(z) - \sum_{y\sim z} \frac{1}{u(y)}\, p(y), \\
z = \mathcal{O}: \quad & \frac{\partial \bar{L}(u)}{\partial u(\mathcal{O})} = \frac{\sum_{y\sim \mathcal{O}} u(y)}{u(\mathcal{O})^2}\, p(\mathcal{O}) - \sum_{y:y^\uparrow = \mathcal{O}} \frac{1}{u(y)} p(y) - \frac{D_\star}{u(\star)}\, p(\star), \\
z = \star: \quad & \frac{\partial \bar{L}(u)}{\partial u(\star)} = -\frac{1}{u(\mathcal{O})}\, p(\mathcal{O}) + \frac{u(\star^\uparrow) + D_\star u(\mathcal{O})}{u(\star)^2}\, p(\star), \\
z \in \partial V_R\backslash\star: \quad & \frac{\partial \bar{L}(u)}{\partial u(z)} = -\frac{1}{u(z^\uparrow)}\, p(z^\uparrow) + \frac{u(z^\uparrow) + D_z u(\square_z)}{u(z)^2}\, p(z) \\
& \qquad + \left[\frac{u(\square_z)}{u(z)^2} - \frac{d^-}{u(\square_z)}\right] p(\square_z) 1_{\left\{\frac{u(z)}{u(z^\uparrow)} \in [1, \sqrt{d}]\right\}} \\
& \qquad + \left[\frac{u(\square_z)}{u(z)^2} - \frac{d^+}{u(\square_z)}\right] p(\square_z) 1_{\left\{\frac{u(z)}{u(z^\uparrow)} < 1\right\}}, \\
z \in \square: \quad & \frac{\partial \bar{L}(u)}{\partial u(z)} = -\frac{D_{z^\uparrow}}{u(z^\uparrow)}\, p(z^\uparrow) + \left[-\frac{1}{u(z^\uparrow)} + \frac{d^- u(z^\uparrow)}{u(z)^2}\right] p(z)\, 1_{\left\{\frac{u(z)}{u(z^\uparrow)} \in [1, \sqrt{d}]\right\}} \\
& \qquad + \left[-\frac{1}{u(z^\uparrow)} + \frac{d^+ u(z^\uparrow)}{u(z)^2}\right] p(z)\, 1_{\left\{\frac{u(z)}{u(z^\uparrow)} < 1\right\}}.
\end{aligned} \tag{3.17}$$

CHAPTER 3

The second derivatives of L are all < 0. The first line in (3.17) can be rewritten as

$$\sum_{y\sim z} u(y)\left[\frac{p(z)}{u^2(z)} - \frac{p(y)}{u^2(y)}\right],$$

which is zero when

$$\bar{u}(x) = \sqrt{p(x)}, \qquad x \in \mathrm{int}(V_R) \setminus \mathcal{O}. \tag{3.18}$$

Given the choice in (3.18), the fourth line in (3.17) is zero when

$$\bar{u}(x) = \sqrt{\frac{d^- p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}1_{\left\{\frac{u(x)}{u(x^\uparrow)}\in[1,\sqrt{d^-}]\right\}} + \sqrt{\frac{d^+ p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}1_{\left\{\frac{u(x)}{u(x^\uparrow)}<1\right\}}, \qquad x \in \Box. \tag{3.19}$$

Furthermore, the derivative in the fifth line is strictly negative when both indicators are 0 and therefore at least one indicator must be 1. This is guaranteed since the quotient

$$\frac{\bar{u}(x)}{\bar{u}(x^\uparrow)} = \sqrt{\frac{d^- p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}1_{\left\{\frac{u(x)}{u(x^\uparrow)}\in[1,\sqrt{d^-}]\right\}} + \sqrt{\frac{d^+ p(x^\uparrow)p(x)}{D_{x^\uparrow}+p(x)}}1_{\left\{\frac{u(x)}{u(x^\uparrow)}<1\right\}}, \qquad x \in \Box,$$

is bounded from above by $\sqrt{d}$ for all $D_{x^\uparrow}$. In addition, we can rewrite the indicators in (3.19) to get

$$\bar{u}(x) = \sqrt{\frac{d^- p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}1_{\left\{\frac{p(x)}{p(x^\uparrow)}\geq\frac{D_{x^\uparrow}}{d^- -1}\right\}} + \sqrt{\frac{d^+ p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}1_{\left\{\frac{p(x)}{p(x^\uparrow)}<\frac{D_{x^\uparrow}}{d^+ -1}\right\}}, \qquad x \in \Box. \tag{3.20}$$

Given the choice in (3.18)–(3.19), also the fourth line in (3.17) is zero. Thus, only the second and third line in (3.17) are non-zero, but this is harmless because $\mathcal{O}, \star$ carry a negligible weight in the limit as $R \to \infty$,because of the constraint $p(\partial V_R \cup \mathcal{O}) \leq 1/R$ (recall Section 3.5.3). To simplify notation, define

$$\Omega^-_\Box(x) = \sqrt{\frac{d^- p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}}, \qquad \Omega^+_\Box(x) = \sqrt{\frac{d^+ p(x^\uparrow)p(x)}{D_{x^\uparrow}p(x^\uparrow)+p(x)}},$$

and

$$\Omega^-(x) = \sqrt{\frac{d^- p(x)p(\Box_x)}{D_x p(x)+p(\Box_x)}}, \qquad \Omega^+(x) = \sqrt{\frac{d^+ p(x)p(\Box_x)}{D_x p(x)+p(\Box_x)}}.$$

Inserting (3.18)–(3.19) into (3.14) and using (3.13) and (3.16), we get the following lower bound:

$$
\begin{aligned}
&\sup_{u\colon V_R\to(0,\infty)} \bar{L}(u)\\
&\geq -\sum_{x\in \mathrm{int}(V_R)} A_{\bar{u}}(x) \quad -\sum_{x\in\partial V_R\backslash\star} B_{\bar{u}}(x) \quad - C_{\bar{u}}(\star)+\sum_{x\in\square}\bar{D}_{\bar{u}}(x)\\
&= -\sum_{x\in \mathrm{int}(V_R)}\sum_{y\sim x}\sqrt{p(y)p(x)}\\
&\quad -\sum_{x\in\partial V_R\backslash\star}\sqrt{p(x)}\left(\sqrt{p(x^{\uparrow})}+D_x\Omega^{-}(x)1_{\left\{\frac{p(\square_x)}{p(x)}\geq\frac{D_x}{d^{-}-1}\right\}}+D_x\Omega^{+}(x)1_{\left\{\frac{p(\square_x)}{p(x)}<\frac{D_x}{d^{-}-1}\right\}}\right)\\
&\quad -\sqrt{p(\star)}\left(\sqrt{p(\star^{\uparrow})}+D_\star\sqrt{p(\mathcal{O})}\right)\\
&\quad +\sum_{x\in\square}p(x)\left[S^{-}\left(\frac{\Omega^{-}_{\square}(x)}{\sqrt{p(x^{\uparrow})}}\right)1_{\left\{\frac{p(x)}{p(x^{\uparrow})}\geq\frac{D_{x^{\uparrow}}}{d^{-}-1}\right\}}+S^{+}\left(\frac{\Omega^{+}_{\square}(x)}{\sqrt{p(x^{\uparrow})}}\right)1_{\left\{\frac{p(x)}{p(x^{\uparrow})}<\frac{D_{x^{\uparrow}}}{d^{+}-1}\right\}}\right]
\end{aligned}
$$

7. Using the relation $(D_x+1)p(x)=\sum_{y\sim x}p(x)$, $x\in\mathrm{int}(V_R)$, we get from (3.11) that

$$
I^{\dagger}_{E_R}(p)\geq K^1_R(p)+K^2_R(p)
$$

with

$$
\begin{aligned}
K^1_R(p) &= \sum_{x\in\mathrm{int}(V_R)}\sum_{y\sim x}\left[p(x)-\sqrt{p(x)p(y)}\right]\\
&= \sum_{\{x,y\}\in E_R\backslash\{\mathcal{O},\star\}}\left(\sqrt{p(x)}-\sqrt{p(y)}\right)^2+\left[p(\mathcal{O})-\sqrt{p(\mathcal{O})p(\star)}\right]-\sum_{x\in\partial V_R}\left[p(x)-\sqrt{p(x)p(x^{\uparrow})}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
K^2_R(p) = &\sum_{x\in\partial V_R\backslash\star}\Bigg[(D_x+1)p(x)\\
&\qquad -\sqrt{p(x)}\left(\sqrt{p(x^{\uparrow})}+D_x\Omega^{-}(x)1_{\left\{\frac{p(\square_x)}{p(x)}\geq\frac{D_x}{d^{-}-1}\right\}}+D_x\Omega^{+}(x)1_{\left\{\frac{p(\square_x)}{p(x)}<\frac{D_x}{d^{-}-1}\right\}}\right)\Bigg]\\
&\qquad +(D_\star+1)p(\star)-\sqrt{p(\star)}\left(\sqrt{p(\star^{\uparrow})}+D_\star\sqrt{p(\mathcal{O})}\right)\\
&+\sum_{x\in\square}p(x)\left[S^{-}\left(\frac{\Omega^{-}_{\square}(x)}{\sqrt{p(x^{\uparrow})}}\right)1_{\left\{\frac{p(x)}{p(x^{\uparrow})}\geq\frac{D_{x^{\uparrow}}}{d^{-}-1}\right\}}+S^{+}\left(\frac{\Omega^{+}_{\square}(x)}{\sqrt{p(x^{\uparrow})}}\right)1_{\left\{\frac{p(x)}{p(x^{\uparrow})}<\frac{D_{x^{\uparrow}}}{d^{+}-1}\right\}}\right].
\end{aligned}
$$

The first sum in the right-hand side of $K^1_R(p)$ equals the *standard rate function* $I_{\widehat{E}_R}(p)$ without $\{\mathcal{O},\star\}$ and the tadpoles. Rearranging and simplifying terms, we arrive at

$$
I^{\dagger}_{E_R}(p)\geq I_{\widehat{E}_R}(p)+K^3_R(p) \tag{3.21}
$$

with

$$
K^3_R(p)=S_{\partial V_R\backslash\star}(p)+S_{\mathcal{O},\star}(p)+S_{(\partial V_R\backslash\star)\cup\square}(p),
$$

where

$$S_{\partial V_R\setminus\star}(p) = \sum_{x\in\partial V_R\setminus\star} D_x p(x),$$

$$S_{\mathcal{O},\star}(p) = \left(\sqrt{p(\mathcal{O})}-\sqrt{p(\star)}\right)^2 + (D_\star-1)\left[p(\star)-\sqrt{p(\mathcal{O})p(\star)}\right],$$

$$\begin{aligned} S_{(\partial V_R\setminus\star)\cup\Box}(p) = &-\sum_{x\in\partial V_R\setminus\star} p(x)\,D_x\left(\frac{\Omega^-(x)}{\sqrt{p(x)}}1_{\left\{\frac{p(\Box_x)}{p(x)}\geq\frac{D_x}{d^--1}\right\}}+\frac{\Omega^+(x)}{\sqrt{p(x)}}1_{\left\{\frac{p(\Box_x)}{p(x)}<\frac{D_x}{d^--1}\right\}}\right)\\ &+\sum_{x\in\Box} p(x)\left[S^-\left(\frac{\Omega_\Box^-(x)}{\sqrt{p(x^\uparrow)}}\right)1_{\left\{\frac{p(x)}{p(x^\uparrow)}\geq\frac{D_{x^\uparrow}}{d^--1}\right\}}+S^+\left(\frac{\Omega_\Box^+(x)}{\sqrt{p(x^\uparrow)}}\right)1_{\left\{\frac{p(x)}{p(x^\uparrow)}<\frac{D_{x^\uparrow}}{d^+-1}\right\}}\right]. \end{aligned} \tag{3.22}$$

8. Since $\sqrt{p(\mathcal{O})p(\star)} \leq \frac{1}{2}[p(\mathcal{O})+p(\star)]$, the boundary constraint $\sum_{x\in\partial V_R\cup\mathcal{O}} p(x)\leq 1/R$ (recall Section 3.5.3) implies that $S_{\partial V_R\setminus\star}(p)+S_{\mathcal{O},\star}(p)=O(1/R)$. The same constraint implies that the first sum in $S_{(\partial V_R\setminus\star)\cup\Box}(p)$ is $O(1/R)$. Hence

$$K_R^3(p)\geq O(1/R)+\sum_{x\in\partial V_R\setminus\star} p(x)\,F\left(\tfrac{p(\Box_x)}{p(x)}\right)$$

with

$$\begin{aligned} F(w) = &\,w\left(d^-+1-\sqrt{d}\left[\sqrt{\frac{w}{d+w}}+\sqrt{\frac{d^++w}{w}}\right]\right)1_{\left\{\frac{p(\Box_x)}{p(x)}\geq\frac{D_x}{d^--1}\right\}}\\ &+w\left(d^++1-\sqrt{d^+}\left[\sqrt{\frac{w}{d+w}}+\sqrt{\frac{d^++w}{w}}\right]\right)1_{\left\{\frac{p(\Box_x)}{p(x)}<\frac{D_x}{d^--1}\right\}}. \end{aligned}$$

The map $w\mapsto F(w)$ is continuous on $(0,\infty)$ with

$$F(w)=\begin{cases} -\frac{\sqrt{d^-}}{\sqrt{d^+}}\sqrt{w}+(d^-+1)w+O(w^{3/2}), & w\downarrow 0,\\ [(d^-+1)-2\sqrt{d^-}\,]\,w+\sqrt{d^-}(d^--d^+)/2+O(w^{-1}), & w\to\infty, \end{cases}$$

on the first indicator, while

$$F(w)=\begin{cases} -\sqrt{w}+(d^++1)w+O(w^{3/2}), & w\downarrow 0,\\ [(d^++1)-2\sqrt{d^+}\,]\,w+\sqrt{d^-}(d^--d^+)/2+O(w^{-1}), & w\to\infty, \end{cases}$$

on the second indicator. From this we see that if $d^+\geq d^-\geq 4$, then there exists a $C\in(1,\infty)$ such that

$$F(w)+C\geq\left(1-\sqrt{w}\right)^2,\qquad w\in[0,\infty).$$

Hence we have the lower bound

$$\begin{aligned} K_R^3(p)&\geq O(1/R)+\sum_{x\in\partial V_R\setminus\star} p(x)\left[-C+\left(1-\sqrt{\tfrac{p(\Box_x)}{p(x)}}\right)^2\right]\\ &=O(1/R)+\sum_{x\in\partial V_R\setminus\star}\left(\sqrt{p(x)}-\sqrt{p(\Box_x)}\right)^2. \end{aligned}$$

Via (3.21)–(3.22), it follows that

$$I^{\dagger}_{E_R} \geq O(1/R) + I_{\widetilde{E}_R}(p), \qquad R \in \mathbb{N},$$

with $I_{\widetilde{E}_R}(p)$ the *standard rate function* and

$$\widetilde{E}_R = \widehat{E}_R \cup \left[\cup_{x \in \partial V_R \setminus \star} \{x, \square_x\} \right]$$

the set of edges obtained by *removing the edge* $\{\mathcal{O}, \star\}$.

§3.6 Limit of the upper variational formula

With the above, we have shown that

$$\frac{1}{t} \log \langle U(t) \rangle \leq \varrho \log(\varrho t) - \varrho - \chi^{+}_{R}(\varrho) + o(1), \qquad t \to \infty,$$

with

$$\chi^{+}_{R}(\varrho) = \inf_{p \in \mathcal{P}(\mathcal{GW}_{\pi,R})} \left\{ I^{\dagger}_{E}(p) + \varrho J_{V(\mathcal{GW}_{\pi,R})}(p) \right\}. \tag{3.23}$$

It only remains to show that $\liminf_{R\to\infty} \chi^{+}_{R}(\varrho) \geq \chi_{\mathcal{GW}}(\varrho)$. Due to shift invariance (recall Section 3.5.3), we may assume that the minimiser in (3.23) satisfies $\sum_{x \in \partial V_R \cup \mathcal{O}} p(x) \leq 1/R$. Next

$$\mathcal{GW}_{\pi,R} \subseteq \mathcal{GW}.$$

Consequently,

$$I_{\widetilde{E}_R}(p) = I_E(p) - \sum_{x \in \partial V_R \setminus \star} (D_x - 1)p(x), \qquad \forall\, p \in \mathcal{P}(\mathcal{GW})\colon\ \mathrm{supp}(p) \subseteq \mathcal{GW}_{\pi,R},$$

where the sum compensates for the contribution coming from the edges in $\mathcal{GW}$ that link the vertices in $\partial V_R \setminus \star$ to the vertices one layer deeper in $\mathcal{GW}$ that are not tadpoles. Since this sum is $O(1/R)$, we obtain

$$\begin{aligned}
\chi^{+}_{R}(\varrho) &= \inf_{p \in \mathcal{P}(V(\mathcal{GW}_{\pi,R}))} \left\{ I^{\dagger}_{E}(p) + \varrho J_{V(\mathcal{GW}_R)}(p) \right\} \\
&\geq O(1/R) + \inf_{p \in \mathcal{P}(V(\mathcal{GW}))\colon\ \mathrm{supp}(p) \subseteq V(\mathcal{GW}_{\pi,R})} \left\{ I_{E(\mathcal{GW})}(p) + \varrho J_{V(\mathcal{GW})}(p) \right\} \\
&\geq O(1/R) + \chi_{\mathcal{GW}}(\varrho),
\end{aligned}$$

where the last inequality follows after dropping the constraint under the infimum.

Appendix of Part I

APPENDIX A

Appendix: Part I

§A.1 Large deviation principle for the local times of Markov renewal processes

The following LDP, which was used in the proof of Lemma 2.3.7, was derived in [31, Proposition 1.2], and generalises the LDP for the empirical distribution of a Markov proceses on a finite state space derived in [15]. See [11, Chapter III] for the definition of the LDP.

Proposition A.1.1. *Let $Y = (Y_t)_{t\geq 0}$ be the Markov renewal process on the finite graph $G = (V, E)$ with transition kernel $(\pi_{x,y})_{\{x,y\}\in E}$ and with sojourn times whose distributions $(\psi_x)_{x\in V}$ have support $(0,\infty)$. For $t > 0$, let L_t^Y denote the empirical distribution of Y at time t (see (2.12)). Then the family $(\mathbb{P}(L_t^Y \in \cdot))_{t>0}$ satisfies the LDP on $\mathcal{P}(V)$ with rate t and with rate function $I_E^\dagger$ given by*

$$I_E^\dagger(p) = \inf_{\beta\in(0,\infty)} \inf_{q\in\mathcal{P}(V)} \left[\widehat{K}(\beta q) + \widetilde{K}(p \mid \beta q)\right] \tag{A.1}$$

with

$$\widehat{K}(\beta q) = \sup_{\widehat{q}\in\mathcal{P}(V)} \sum_{x\in V} \beta q(x) \log\left(\frac{\widehat{q}(x)}{\sum_{y\in V}\pi_{x,y}\widehat{q}(y)}\right), \tag{A.2}$$

$$\widetilde{K}(p \mid \beta q) = \sum_{x\in V} \beta q(x)\,(\mathcal{L}\lambda_x)\left(\tfrac{p(x)}{\beta q(x)}\right), \tag{A.3}$$

where

$$\begin{aligned} (\mathcal{L}\lambda_x)(\alpha) &= \sup_{\theta\in\mathbb{R}}[\alpha\theta - \lambda_x(\theta)], \quad \alpha\in[0,\infty),\\ \lambda_x(\theta) &= \log\int_0^\infty \mathrm{e}^{\theta\tau}\psi_x(\mathrm{d}\tau), \quad \theta\in\mathbb{R}. \end{aligned} \tag{A.4}$$

The rate function $I_E^\dagger$ consist of two parts: $\widehat{K}$ in (A.2) is the rate function of the LDP on $\mathcal{P}(V)$ for the empirical distribution of the discrete-time Markov chain on V with transition kernel $(\pi_{x,y})_{\{x,y\}\in E}$ (see [11, Theorem IV.7]), while $\widetilde{K}$ in (A.3) is the rate function of the LDP on $\mathcal{P}(0,\infty)$ for the empirical mean of the sojourn times, given the empirical distribution of the discrete-time Markov chain. Moreover, λ_x is the cumulant generating function associated with ψ_x, and $\mathcal{L}\lambda_x$ is the Legendre transform of λ_x, playing the role of the Cramèr rate function for the empirical mean of the i.i.d.

sojourn times at x. The parameter β plays the role of the ratio between the continuous time scale and the discrete time scale.

§A.2 Sojourn times: cumulant generating functions and Legendre tranforms

In Appendix A.2.1 we recall general properties of cumulant generating functions and Legendre transforms, in Appendices A.2.2 and A.2.3 we identify both for the two sojourn time distributions arising in Lemma 2.3.7, respectively.

§A.2.1 General observations

Let λ be the cumulant generating function of a non-degenerate sojourn time distribution ϕ, and $\mathcal{L}\lambda$ be the Legendre transform of λ (recall (2.25)). Both λ and $\mathcal{L}\lambda$ are strictly convex, are analytic in the interior of their domain, and achieve a unique zero at $\theta = 0$, respectively, $\alpha = \alpha_c$ with $\alpha_c = \int_0^\infty \tau\phi(\mathrm{d}\tau)$. Furthermore, λ diverges at some $\theta_c \in (0,\infty]$ and has slope α_c at $\theta = 0$. Moreover, if the slope of λ diverges at θ_c, then $\mathcal{L}\lambda$ is finite on $(0,\infty)$.

The supremum in the Legendre transform defining $(\mathcal{L}\lambda)(\alpha)$ is uniquely taken at $\theta = \theta(\alpha)$ solving the equation

$$\lambda'(\theta(\alpha)) = \alpha, \quad \alpha > 0.$$

The tangent of λ with slope α at $\theta(\alpha)$ intersects the vertical axis at $(-\mathcal{L}\lambda)(\alpha)$, i.e., putting

$$\mu(\alpha) = \lambda(\theta(\alpha)) \tag{A.5}$$

we have

$$\mu(\alpha) = \alpha(\mathcal{L}\lambda)'(\alpha) - (\mathcal{L}\lambda)(\alpha). \tag{A.6}$$

(See Fig. A.1.) Note that by differentiating (A.6) we get

$$\mu'(\alpha) = \alpha(\mathcal{L}\lambda)''(\alpha),$$

which shows that $\alpha \mapsto \mu(\alpha)$ is strictly increasing and hence invertible, with inverse function μ^{-1}. Note that by differentiating the relation $(\mathcal{L}\lambda)(\alpha) = \alpha\theta(\alpha) - \lambda(\theta(\alpha))$ we get

$$(\mathcal{L}\lambda)'(\alpha) = \theta(\alpha). \tag{A.7}$$

A further relation that is useful reads

$$(\mathcal{L}\lambda)' \circ \mu^{-1} = \lambda^{-1}, \tag{A.8}$$

which follows because $\mu = \lambda \circ \theta$ by (A.5) and $(\mathcal{L}\lambda)' = \theta$ by (A.7).

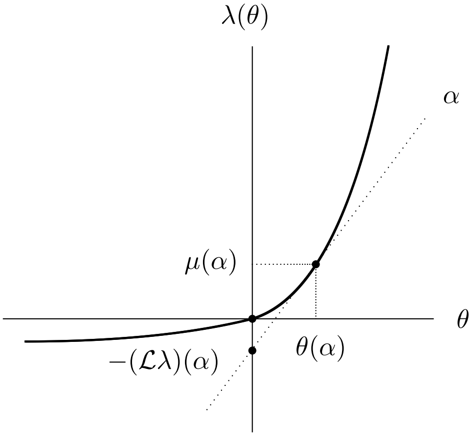

Figure A.1: Picture exhibiting the link between $\lambda(\theta)$, $(\mathcal{L}\lambda)(\alpha)$, $\theta(\alpha)$, $\mu(\alpha)$. The dotted line is the tangent of λ with slope α, crossing the horizontal axis at $-(\mathcal{L}\lambda)(\alpha)$, and touching λ at the point $(\theta(\alpha), \mu(\alpha))$. All are analytic on the interior of their domain.

§A.2.2 Exponential sojourn time

If $\phi = \mathrm{EXP}(d+1)$, then the cumulant generating function $\lambda(\theta) = \log \int_0^\infty \mathrm{e}^{\theta\tau}\psi(\mathrm{d}\tau)$ is given by

$$\lambda(\theta) = \begin{cases} \log\left(\frac{d+1}{d+1-\theta}\right), & \theta < d+1, \\ \infty, & \theta \geq d+1. \end{cases}$$

To find $(\mathcal{L}\lambda)(\alpha)$, we compute

$$\frac{\partial}{\partial\theta}[\alpha\theta - \log(\tfrac{d+1}{d+1-\theta})] = \alpha - \frac{1}{d+1-\theta}, \qquad \frac{\partial^2}{\partial\theta^2}[\alpha\theta - \log(\tfrac{d}{d+1-\theta})] = -\frac{1}{(d+1-\theta)^2} < 0.$$

Hence the supremum in (2.24) is uniquely taken at

$$\theta(\alpha) = d + 1 - \tfrac{1}{\alpha}, \qquad \alpha > 0,$$

so that

$$(\mathcal{L}\lambda)(\alpha) = \alpha(d+1) - 1 - \log[\alpha(d+1)], \qquad \alpha > 0. \tag{A.9}$$

Thus, λ and $\mathcal{L}\lambda$ have the shape in Fig. A.2, with $\theta_c = d+1$ and $\alpha_c = \frac{1}{d+1}$, and with $\lim_{\theta\uparrow\theta_c} \lambda(\theta) = \infty$ and $\lim_{\theta\uparrow\theta_c} \lambda'(\theta) = \infty$.
Note that μ has domain $(0, \infty)$ and range $\mathbb{R}$.

§A.2.3 Non-exponential sojourn time

For $\phi = \psi$ the computations are more involved. Let $\mathcal{T}^* = (E, V)$ be the infinite rooted regular tree of degree $d+1$. Write $\mathcal{O}$ for the root. Let $X = (X_n)_{n\in\mathbb{N}_0}$ be the discrete-time simple random walk on $\mathcal{T}^* = (E, V)$ starting from $\mathcal{O}$. Write $\tau_{\mathcal{O}}$ to denote the

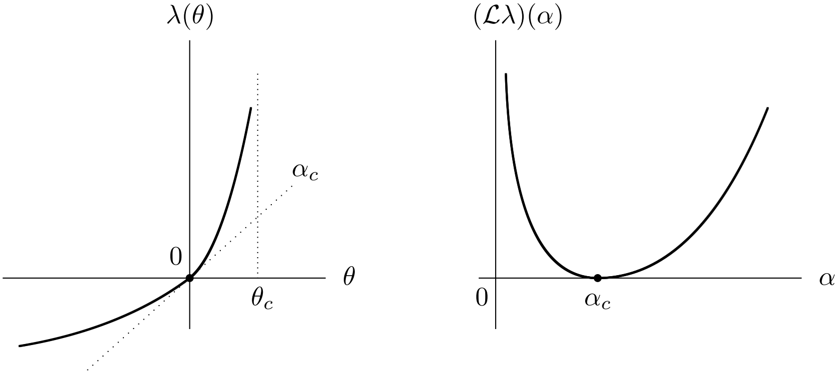

Figure A.2: Picture of $\theta \mapsto \lambda(\theta)$ (left) and $\alpha \mapsto (\mathcal{L}\lambda)(\alpha)$ (right) for $\phi = \mathrm{EXP}(d+1)$.

time of the *first return* of X to $\mathcal{O}$. Define $r = \mathbb{P}_{\mathcal{O}}(\tau_{\mathcal{O}} < \infty)$. It is easy to compute r by projecting X on $\mathbb{N}_0$: r is the return probability to the origin of the random walk on $\mathbb{N}_0$ that jumps to the right with probability $p = \frac{d}{d+1}$ and to the left with probability $q = \frac{1}{d+1}$, which equals $\frac{p}{q}$ (see [33, Section 8]). Thus, $r = \frac{1}{d}$.

For $y \in \mathcal{T}^*$, define $h_y = \mathbb{P}_y(\tau_{\mathcal{O}} < \infty)$. Then h_y can be explicitly calculated, namely,

$$h_y = \begin{cases} d^{-|y|}, & y \in \mathcal{T}^* \setminus \{\mathcal{O}\}, \\ 1, & y = \mathcal{O}. \end{cases}$$

Note that h is a harmonic function on $\mathcal{T}^* \setminus \mathcal{O}$, i.e., $h_y = \sum_{z \in \mathcal{T}^*} \widehat{\pi}_{y,z} h_z$, $y \in \mathcal{T}^* \setminus \mathcal{O}$. We can therefore consider the Doob-transform of X, which is the random walk with transition probabilities away from the root given by

$$\check{\sigma}_{y,z} = \begin{cases} \frac{d}{d+1}, & z = y^{\uparrow}, \\ \frac{1}{d}\frac{1}{d+1}, & z \neq y^{\uparrow}, \{y,z\} \in E, \\ 0, & \text{else}, \end{cases} \qquad y \in \mathcal{T}^* \setminus \{\mathcal{O}\},$$

and transition probabilities from the root are given by

$$\check{\sigma}_{\mathcal{O},z} = \begin{cases} \frac{1}{d}, & \{\mathcal{O}, z\} \in E, \\ 0, & \text{else}. \end{cases}$$

Thus, the Doob-transform reverses the upward and the downward drift of X.

Recall from Lemma 2.3.7 that ψ is the distribution of $\tau_{\mathcal{O}}$ *conditional* on $\{\tau_{\mathcal{O}} < \infty\}$ and on X leaving $\mathcal{O}$ at time 0.

Lemma A.2.1. *Let* $\lambda(\theta) = \log \int_0^\infty \mathrm{e}^{\theta\tau} \psi(\mathrm{d}\tau)$. *Then*

$$\mathrm{e}^{\lambda(\theta)} = \begin{cases} \frac{d+1-\theta}{2} \left[1 - \sqrt{1 - \frac{4d}{(d+1-\theta)^2}}\,\right], & \theta \in (-\infty, \theta_c], \\ \infty, & \textit{else}, \end{cases} \tag{A.10}$$

with $\theta_c = (\sqrt{d}-1)^2$*. The range of* $\exp\circ\lambda$ *is* $(0,\sqrt{d}\,]$*, with the maximal value is uniquely taken at* $\theta = \theta_c$*.*

Proof. To compute the moment-generating function of $\tau_{\mathcal{O}}$, we consider the Doob-transform of X and its projection onto $\mathbb{N}_0$. Let $p_{2k} = P(\tau_{\mathcal{O}} = 2k)$. It is well-known that (see [33, Section 8])

$$G^{p,q}(s) = \mathbb{E}(s^{\tau_{\mathcal{O}}} \mid \tau_{\mathcal{O}} < \infty) = \sum_{k\in\mathbb{N}} s^{2k} p_{2k} = \frac{1}{2p}\left[1-\sqrt{1-4pqs^2}\right], \qquad |s| \le 1. \tag{A.11}$$

Therefore we have

$$\begin{aligned} e^{\lambda(\theta)} &= \mathbb{E}(e^{\theta\tau_{\mathcal{O}}}) = \sum_{k\in\mathbb{N}} p_{2k}\left[\mathbb{E}\left(e^{\theta\,\mathrm{EXP}(d+1)}\right)\right]^{2k-1} \\ &= \sum_{k\in\mathbb{N}} p_{2k}\left(\frac{d+1}{d+1-\theta}\right)^{2k-1} = \left(\frac{d+1-\theta}{d+1}\right) G^{p,q}(s) \end{aligned} \tag{A.12}$$

with

$$p = \tfrac{1}{d+1}, \qquad q = \tfrac{d}{d+1}, \qquad s = \frac{d+1}{d+1-\theta}.$$

Inserting (A.11) into (A.12), we get the formula for $\lambda(\theta)$. From the term in the square root we see that $\lambda(\theta)$ is finite if and only if $\theta \le \theta_c = d+1-2\sqrt{d} = (\sqrt{d}-1)^2$. □

There is no easy closed form expression for $(\mathcal{L}\lambda)(\alpha)$, but it is easily checked that λ and $\mathcal{L}\lambda$ have the shape in Fig. A.3, with $\theta_c = (\sqrt{d}-1)^2$ and $\alpha_c = \int_0^\infty \tau\psi(\mathrm{d}\tau) < \infty$, and with $\lambda(\theta_c) = \log\sqrt{d} < \infty$ and $\lambda'(\theta_c) = \infty$, i.e., there is a *cusp* at the threshold θ_c, implying that $\mathcal{L}\lambda$ is finite on $(0,\infty)$. It follows from (A.7) that

$$\lim_{\alpha\to\infty} \frac{1}{\alpha}(\mathcal{L}\lambda)(\alpha) = \lim_{\alpha\to\infty} \theta(\alpha) = \theta_c. \tag{A.13}$$

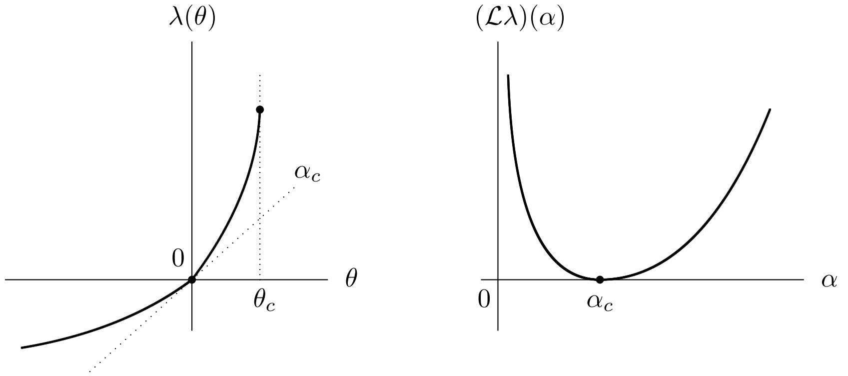

Figure A.3: Picture of $\theta \mapsto \lambda(\theta)$ *(left) and* $\alpha \mapsto (\mathcal{L}\lambda)(\alpha)$ *(right) for* $\phi = \psi$*.*

Lemma A.2.2. *The function* $\lambda^{-1} \circ \log = (\exp\circ\lambda)^{-1}$ *is given by*

$$(\exp\circ\lambda)^{-1}(\beta) = d + 1 - \beta - \frac{d}{\beta}, \qquad \beta \in (0, \sqrt{d}]. \tag{A.14}$$

The range of $(\exp\circ\lambda)^{-1}$ *is* $(-\infty, \theta_c]$, *with the maximal value* θ_c *uniquely taken at* $\beta = \sqrt{d}$.

Proof. We need to invert $\exp\circ\lambda$ in (A.10). Abbreviate $\chi = \frac{d+1-\theta}{2}$. Then

$$\beta = \chi\left[1 - \sqrt{1 - \frac{d}{\chi^2}}\right] \quad\Longrightarrow\quad \chi = \frac{\beta^2 + d}{2\beta} \quad\Longrightarrow\quad \theta = d + 1 - \frac{\beta^2 + d}{\beta}.$$

□

Note that $(\sqrt{d}, \infty)$ is not part of the domain of $(\exp\circ\lambda)^{-1}$, even though the right-hand side of (A.14) still makes sense (as a second branch). Note that μ has domain $(0, \infty)$ and range $(-\infty, \sqrt{d}]$ (see Fig. A.1).

§A.3 Large deviation estimate for the local time away from the backbone

In this appendix we derive a large deviation principle for the *total local times at successive depths* of the random walk on $\mathcal{T}^{\mathbb{Z}}$ (see Fig. 2.3). This large deviation principle is not actually needed, but serves as a warm up for the more elaborate computations in Section 2.4.

For $k \in \mathbb{N}_0$, let V_k be the set of vertices in $\mathcal{T}^{\mathbb{Z}}$ that are at distance k from the backbone (see Fig. 2.3). For $R \in \mathbb{N}$, define

$$\begin{aligned} \ell_t^R(k) &= \textstyle\sum_{x \in V_k} \ell_t^{\mathbb{Z}}(x), & k &= 0, 1, \ldots, R, \\ \ell_t^R &= \textstyle\sum_{k > R} \sum_{x \in V_k} \ell_t^{\mathbb{Z}}(x), & k &= R + 1, \end{aligned}$$

and

$$L_t^R = \frac{1}{t}\Big((\ell_t(k))_{k=0}^R, \ell_t^R \Big).$$

Abbreviate $V_R^* = \{0, 1, \ldots, R, R+1\}$,

Lemma A.3.1. *For every* $R \in \mathbb{N}$, $(L_t^R)_{t \geq 0}$ *satisfies the large deviation principle on* $\mathcal{P}(V_R^*)$ *with rate* t *and with rate function* $I_R^\dagger$ *given by*

$$\begin{aligned} I_R^\dagger(p) = {} & \big[\sqrt{(d-1)p(0)} - \sqrt{dp(1)}\,\big]^2 + \sum_{k=1}^{R-1} \big[\sqrt{p(k)} - \sqrt{dp(k+1)}\,\big]^2 \\ & + \big[\sqrt{p(R) + p(R+1)} - \sqrt{dp(R+1)}\,\big]^2. \end{aligned} \tag{A.15}$$

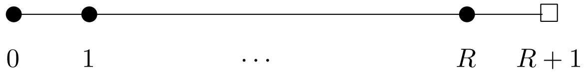

Figure A.4: Depths $k = 0, 1, \dots, R$ and $k > R$.

Proof. By monitoring the random walk on the tree in Fig. 2.3 and projecting its depth on the vertices $0, 1, \dots, R$, respectively, $R+1$, we can apply the LDP in Proposition A.1.1 (see Fig. A.4).

1. The sojourn times have distribution $\mathrm{EXP}(d+1)$ at vertices $k = 0, 1, \dots, R$ and distribution ψ at vertex $k = R+1$. The transition probabilities are

$$\begin{aligned}
&\pi_{0,0} = \tfrac{2}{d+1}, \qquad \pi_{0,1} = \tfrac{d-1}{d+1},\\
&\pi_{k,k+1} = \tfrac{1}{d+1}, \quad \pi_{k,k-1} = \tfrac{d}{d+1}, \qquad k = 1, \dots, R,\\
&\pi_{R+1,R} = 1.
\end{aligned}$$

Proposition A.1.1 therefore yields that $(L_t^R)_{t \geq 0}$ satisfies the LDP on on $\mathcal{P}(V_R^*)$ with rate t and with rate function $I_R^\dagger$ given by

$$I_R^\dagger(p) = (d+1) \sum_{k=0}^{R} p(k) + \inf_{v\colon V_R^* \to (0,\infty)} \sup_{u\colon V_R^* \to (0,\infty)} L(u,v) \tag{A.16}$$

with

$$L(u,v) = -A - B - C, \tag{A.17}$$

where

$$\begin{aligned}
A &= \sum_{k=1}^{R} v(x) \left\{ 1 + \log\left(\frac{du(k-1) + u(k+1)}{u(k)} \, \frac{p(k)}{v(k)} \right) \right\},\\
B &= v(0) \left\{ 1 + \log\left(\frac{2u(0) + (d-1)u(1)}{u(0)} \, \frac{p(0)}{v(0)} \right) \right\},\\
C &= v(R+1) \left\{ \log\left(\frac{u(R)}{u(R+1)} \right) - (\mathcal{L}\lambda) \left(\frac{p(R+1)}{v(R+1)} \right) \right\}.
\end{aligned}$$

Here we use (A.9) to compute A and B, and for C we recall that $\mathcal{L}\lambda$ is the Legendre transform of the cumulant generation function λ of ψ computed in Lemma A.10.

2. We compute the infimum of $L(u,v)$ over v for fixed u.

- For $k = 1, \dots, R$,

$$\begin{aligned}
\frac{\partial A}{\partial v(k)} &= \log\left(\frac{du(k-1) + u(k+1)}{u(k)} \, \frac{p(k)}{v(k)} \right),\\
\Longrightarrow \bar{v}_u(k) &= p(k) \, \frac{du(k-1) + u(k+1)}{u(k)}.
\end{aligned}$$

The second derivative is $1/v(k) > 0$.

• For $k = 0$,

$$\frac{\partial B}{\partial v(0)} = \log\left(\frac{2u(0) + (d-1)u(1)}{u(0)} \frac{p(0)}{v(0)}\right),$$
$$\Longrightarrow \bar{v}_u(0) = p(0)\,\frac{2u(0) + (d-1)u(1)}{u(0)}.$$

The second derivative is $1/v(0) > 0$.

• For $k = R + 1$, the computation is more delicate. Define (recall (A.6) in Appendix A.2)

$$\mu(\alpha) = \alpha(\mathcal{L}\lambda)'(\alpha) - (\mathcal{L}\lambda)(\alpha).$$

The function μ has range $(-\infty, \log\sqrt{d}]$, with the maximal value uniquely taken at $\alpha = \infty$. Therefore there are two cases.

► $u(R+1)/u(R) \leq \sqrt{d}$. Compute

$$\frac{\partial C}{\partial v(R+1)} = \mu\left(\frac{p(R+1)}{v(R+1)}\right) - \log\left(\frac{u(R+1)}{u(R)}\right),$$
$$\Longrightarrow \bar{v}(R+1) = \frac{p(R+1)}{\alpha_u(R+1)}$$

with $\alpha_u(R+1)$ solving the equation

$$\log\left(\frac{u(R+1)}{u(R)}\right) = \mu\big(\alpha_u(R+1)\big).$$

Since $\mu'(\alpha) = \alpha(\mathcal{L}\lambda)''(\alpha)$ and $\mathcal{L}\lambda$ is strictly convex (see Fig. A.3 in Appendix A.2), μ is strictly increasing and therefore invertible. Consequently,

$$\alpha_u(R+1) = \mu^{-1}\left(\log\left(\frac{u(R+1)}{u(R)}\right)\right). \tag{A.18}$$

Putting (A.17)–(A.18) together, we get

$$L(u) = \inf_{v\colon V_R^* \to (0,\infty)} L(u,v) = -\sum_{k=1}^{R} A_u(k) - B_u + C_u \tag{A.19}$$

with

$$A_u(k) = \frac{du(k-1) + u(k+1)}{u(k)}\,p(k), \qquad k = 1, \ldots, R,$$
$$B_u = \frac{2u(0) + (d-1)u(1)}{u(0)}\,p(0),$$

and

$$\begin{aligned}
C_u &= \frac{p(R+1)}{\alpha_u(R+1)}\left[(\mathcal{L}\lambda)\big(\alpha_u(R+1)\big) - \log\left(\frac{u(R+1)}{u(R)}\right)\right]\\
&= \frac{p(R+1)}{\alpha_u(R+1)}\big[(\mathcal{L}\lambda)\big(\alpha_u(R+1)\big) - \mu\big(\alpha_u(R+1)\big)\big]\\
&= p(R+1)\,(\mathcal{L}\lambda)'\big(\alpha_u(R+1)\big)\\
&= p(R+1)\,((\mathcal{L}\lambda)'\circ\mu^{-1})\left(\log\left(\frac{u(R+1)}{u(R)}\right)\right).
\end{aligned}$$

In (A.8) in Appendix A.2 we showed that $(\mathcal{L}\lambda)'\circ\mu^{-1} = \lambda^{-1}$. Moreover, in (A.14) in Appendix A.2 we showed that $(\lambda^{-1}\circ\log) = S$ with

$$S(\beta) = d+1-\beta-\frac{d}{\beta}, \qquad \beta\in(0,\sqrt{d}\,]. \tag{A.20}$$

Since S has domain $(0,\sqrt{d}\,]$, $C_u(R+1)$ is only defined when $u(R+1)/u(R)\le\sqrt{d}$, in which case

$$C_u = p(R+1)\,S\left(\frac{u(R+1)}{u(R)}\right). \tag{A.21}$$

► $u(R+1)/u(R)\le\sqrt{d}$. In this case $\frac{\partial C}{\partial v(R+1)}>0$, the infimum is taken at $\bar{v}(R+1)=0$, and hence (recall (A.13))

$$C_u = p(R+1)\,(\sqrt{d}-1)^2 = p(R+1)\,S(\sqrt{d}). \tag{A.22}$$

Note that the right-hand side does not depend on u. The expressions in (A.21)–(A.22) can be summarised as

$$C_u = p(R+1)\,S\left(\sqrt{d}\wedge\frac{u(R+1)}{u(R)}\right).$$

3. Next we compute the supremum over u of

$$L(u) = L(u,\bar{v}_u) = -A_u - B_u + C_u. \tag{A.23}$$

with $A_u=\sum_{k=1}^R A_u(k)$. We only write down the derivatives that are non-zero.

• For $k=2,\dots,R-1$,

$$-\frac{\partial A_u}{\partial u(k)} = -p(k+1)\,\frac{d}{u(k+1)} - p(k-1)\,\frac{1}{u(k-1)} + p(k)\,\frac{du(k-1)+u(k+1)}{u(k)^2}.$$

• For $k=1$,

$$\begin{aligned}
-\frac{\partial A_u}{\partial u(1)} &= -p(2)\,\frac{d}{u(2)} + p(1)\,\frac{du(0)+u(2)}{u(1)^2},\\
-\frac{\partial B_u}{\partial u(1)} &= -p(0)\,\frac{d-1}{u(0)}.
\end{aligned}$$

- For $k = R$,

$$- \frac{\partial A_u}{\partial u(R)} = -p(R-1)\, \frac{1}{u(R-1)} + p(R)\, \frac{du(R-1) + u(R+1)}{u(R)^2},$$
$$\frac{\partial C_u}{\partial u(R)} = p(R+1)\, \left[\frac{u(R+1)}{u(R)^2} - \frac{d}{u(R+1)} \right] 1_{\left\{ \frac{u(R+1)}{u(R)} \leq \sqrt{d} \right\}}.$$

- For $k = 0$,

$$- \frac{\partial A_u}{\partial u(0)} = -p(1)\, \frac{d}{u(1)},$$
$$- \frac{\partial B_u}{\partial u(0)} = p(0)\, \frac{(d-1)u(1)}{u(0)^2}.$$

- For $k = R+1$,

$$- \frac{\partial A_u}{\partial u(R+1)} = -p(R)\, \frac{1}{u(R)},$$
$$\frac{\partial C_u}{\partial u(R+1)} = p(R+1)\, \left[-\frac{1}{u(R)} + \frac{du(R)}{u(R+1)^2} \right] 1_{\left\{ \frac{u(R+1)}{u(R)} \leq \sqrt{d} \right\}}.$$

All the first derivatives of $A_u + B_u + C_u$ are zero when we choose

$$\begin{aligned} \bar{u}(0) &= \sqrt{(d-1)p(0)}, \qquad \bar{u}(k) = \sqrt{d^k p(k)}, \quad k = 1, \dots, R, \\ \bar{u}(R+1) &= \sqrt{d^{R+1}\, \frac{p(R)p(R+1)}{p(R) + p(R+1)}}. \end{aligned} \tag{A.24}$$

All the second derivatives are strictly negative, and so $\bar{u}$ is the unique maximiser.

4. Inserting (A.24) into (A.19), we get

$$\begin{aligned} L(\bar{u}) = L(\bar{u}, \bar{v}_{\bar{u}}) &= - \sum_{k=2}^{R-1} A_{\bar{u}}(k) - \left[A_{\bar{u}}(1) + B_{\bar{u}} \right] - A_{\bar{u}}(R) + C_{\bar{u}} \\ &= - \sum_{k=2}^{R-1} \sqrt{dp(k)}\, \left[\sqrt{p(k-1)} + \sqrt{p(k+1)} \right] \\ &\quad - \left[2\sqrt{d(d-1)p(0)p(1)} + 2p(0) + \sqrt{dp(1)p(2)} \right] \\ &\quad - \left[\sqrt{dp(R-1)p(R)} + \sqrt{\frac{p(R)}{p(R) + p(R+1)}}\, \sqrt{dp(R)p(R+1)} \right] \\ &\quad + p(R+1)\, S\left(\sqrt{\frac{dp(R+1)}{p(R) + p(R+1)}} \right). \end{aligned}$$

Recalling (A.16), (A.20) and (A.23), and rearranging terms, we find the expression in (A.15). □

Note that $I_R^\dagger$ has a unique zero at p given by

$$p(0) = \tfrac{1}{2}, \qquad p(k) = \tfrac{1}{2}(d-1)d^{-k}, \quad k = 1, \dots, R, \qquad p(R+1) = \tfrac{1}{2}d^{-R}.$$

This shows that the fraction of the local time typically spent a distance k away from the backbone decays exponentially fast in k.

PART II

PARABOLIC ANDERSON MODEL: QUENCHED

CHAPTER 4

Parabolic Anderson model on a Galton-Watson tree revisited

This chapter is based on the following paper:
F. den Hollander and D. Wang. The parabolic Anderson model on a Galton-Watson tree revisited. *J. Stat. Phys.*, 189(1):Paper 8, 2022.

Abstract

In [12] a detailed analysis was given of the large-time asymptotics of the total mass of the solution to the parabolic Anderson model on a supercritical Galton-Watson random tree with an i.i.d. random potential whose marginal distribution is double-exponential. Under the assumption that the degree distribution has bounded support, two terms in the asymptotic expansion were identified under the quenched law, i.e., conditional on the realisation of the random tree and the random potential. The second term contains a variational formula indicating that the solution concentrates on a subtree with minimal degree according to a computable profile. The present paper extends the analysis to degree distributions with unbounded support. We identify the weakest condition on the tail of the degree distribution under which the arguments in [12] can be pushed through. To do so we need to control the occurrence of large degrees uniformly in large subtrees of the Galton-Watson tree.

§4.1 Introduction and main results

Section 4.1.1 provides a brief introduction to the parabolic Anderson model. Section 4.1.2 introduces basic notation and key assumptions. Section 4.1.3 states the main theorem and gives an outline of the remainder of the paper.

§4.1.1 The PAM and intermittency

The *parabolic Anderson model* (PAM) is the Cauchy problem

$$\partial_t u(x,t) = \Delta_{\mathscr{X}} u(x,t) + \xi(x)u(x,t), \qquad t > 0,\, x \in \mathscr{X},$$

where $\mathscr{X}$ is an ambient space, $\Delta_{\mathscr{X}}$ is a Laplace operator acting on functions on $\mathscr{X}$, and ξ is a random potential on $\mathscr{X}$. Most of the literature considers the setting where $\mathscr{X}$ is either $\mathbb{Z}^d$ or $\mathbb{R}^d$ with $d \geq 1$ (for mathematical surveys we refer the reader to [3], [29]). More recently, other choices for $\mathscr{X}$ have been considered as well: the complete graph [16], the hypercube [5], Galton-Watson trees [12], and random graphs with prescribed degrees [12].

The main target for the PAM is a description of *intermittency*: for large t the solution $u(\cdot,t)$ of (2.1) concentrates on well-separated regions in $\mathscr{X}$, called *intermittent islands*. Much of the literature has focussed on a detailed description of the size, shape and location of these islands, and the profiles of the potential $\xi(\cdot)$ and the solution $u(\cdot,t)$ on them. A special role is played by the case where ξ is an i.i.d. random potential with a *double-exponential* marginal distribution

$$\mathsf{P}(\xi(0) > u) = \mathrm{e}^{-\mathrm{e}^{u/\varrho}}, \qquad u \in \mathbb{R},$$

where $\varrho \in (0,\infty)$ is a parameter. This distribution turns out to be critical, in the sense that the intermittent islands neither grow nor shrink with time, and therefore represents a class of its own.

The analysis of intermittency typically starts with a computation of the large-time asymptotics of the total mass, encapsulated in what are called *Lyapunov exponents*. There is an important distinction between the *annealed* setting (i.e., averaged over the random potential) and the *quenched* setting (i.e., almost surely with respect to the random potential). Often both types of Lyapunov exponents admit explicit descriptions in terms of *characteristic variational formulas* that contain information about where and how the mass concentrates in $\mathscr{X}$. These variational formulas contain a *spatial part* (identifying where the concentration on islands takes place) and a *profile part* (identifying what the size and shape of both the potential and the solution are on the islands).

In the present paper we focus on the case where $\mathscr{X}$ is a Galton-Watson tree, in the quenched setting (i.e., almost surely with respect to the random tree and the random potential). In [12] the large-time asymptotics of the total mass was derived under the assumption that the degree distribution has bounded support. The goal of the present paper is to relax this assumption to unbounded degree distributions. In particular, we identify the *weakest condition on the tail of the degree distribution* under which the arguments in [12] can be pushed through. To do so we need to control the occurrence of large degrees *uniformly in large subtrees* of the Galton-Watson tree.

§4.1.2 The PAM on a graph

We begin with some basic definitions and notations (and refer the reader to [3], [29] for more background).

Let $G = (V, E)$ be a *simple connected undirected* graph, either finite or countably infinite. Let Δ_G be the Laplacian on G, i.e.,

$$(\Delta_G f)(x) := \sum_{\substack{y \in V: \\ \{x,y\} \in E}} [f(y) - f(x)], \qquad x \in V,\ f\colon V \to \mathbb{R}. \tag{4.1}$$

Our object of interest is the non-negative solution of the Cauchy problem with localised initial condition,

$$\begin{array}{rcll} \partial_t u(x,t) & = & (\Delta_G u)(x,t) + \xi(x) u(x,t), & x \in V,\ t > 0, \\ u(x,0) & = & \delta_{\mathcal{O}}(x), & x \in V, \end{array} \tag{4.2}$$

where $\mathcal{O} \in V$ is referred to as the *root* of G. We say that G is *rooted at* $\mathcal{O}$ and call $G = (V, E, \mathcal{O})$ a *rooted graph*. The quantity $u(x,t)$ can be interpreted as the amount of mass present at time t at site x when initially there is unit mass at $\mathcal{O}$.

Criteria for existence and uniqueness of the non-negative solution to (4.2) are well known (see [24], [25] for the case $G = \mathbb{Z}^d$), and the solution is given by the *Feynman-Kac formula*

$$u(x,t) = \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s}\, \mathbb{1}\{X_t = x\} \right], \tag{4.3}$$

where $X = (X_t)_{t \geq 0}$ is the continuous-time random walk on the vertices V with jump rate 1 along the edges E, and $\mathbb{P}_{\mathcal{O}}$ denotes the law of X given $X_0 = \mathcal{O}$. We are interested in the *total mass* of the solution,

$$U(t) := \sum_{x \in V} u(x,t) = \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \right]. \tag{4.4}$$

Often we suppress the dependence on G, ξ from the notation. Note that, by time reversal and the linearity of (4.2), $U(t) = \hat{u}(0,t)$ with $\hat{u}$ the solution of (4.2) with a different initial condition, namely, $\hat{u}(x,0) = 1$ for all $x \in V$.

As in [12], throughout the paper we assume that the random potential $\xi = (\xi(x))_{x \in V}$ consists of i.i.d. random variables with marginal distribution satisfying:

Assumption 4.A. [Asymptotic double-exponential potential]
For some $\varrho \in (0, \infty)$,

$$\mathsf{P}(\xi(0) \geq 0) = 1, \qquad \mathsf{P}(\xi(0) > u) = \mathrm{e}^{-\mathrm{e}^{u/\varrho}} \text{ for } u \text{ large enough.} \tag{4.5}$$

■

Chapter 4

The restrictions in (4.5) are helpful to avoid certain technicalities that require no new ideas. In particular, (4.5) is enough to guarantee existence and uniqueness of the non-negative solution to (4.2) on any graph whose largest degrees grow modestly with the size of the graph (as can be inferred from the proof in [25] for the case $G = \mathbb{Z}^d$; see Section 4.6 for more details). All our results remain valid under milder restrictions (e.g. [25, Assumption (F)] plus an integrability condition on the lower tail of $\xi(0)$).

The following *characteristic variational formula* is important for the description of the asymptotics of $U(t)$ when ξ has a double-exponential tail. Denote by $\mathcal{P}(V)$ the set of probability measures on V. For $p \in \mathcal{P}(V)$, define

$$I_E(p) := \sum_{\{x,y\}\in E} \left(\sqrt{p(x)} - \sqrt{p(y)}\right)^2, \qquad J_V(p) := -\sum_{x\in V} p(x)\log p(x),$$

and set

$$\chi_G(\varrho) := \inf_{p\in\mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0,\infty). \tag{4.6}$$

The first term in (4.6) is the quadratic form associated with the Laplacian, describing the solution $u(\cdot,t)$ in the intermittent islands, while the second term in (4.6) is the Legendre transform of the rate function for the potential, describing the highest peaks of $\xi(\cdot)$ in the intermittent islands.

§4.1.3 The PAM on a Galton-Watson tree

Chapter 4

Let D be a random variable taking values in $\mathbb{N}$. Start with a root vertex $\mathcal{O}$, and attach edges from $\mathcal{O}$ to D first-generation vertices. Proceed recursively: after having attached the n-th generation of vertices, attach to each one of them independently a number of vertices having the same distribution as D, and declare the union of these vertices to be the $(n+1)$-th generation of vertices. Denote by $\mathcal{GW} = (V,E)$ the graph thus obtained and by $\mathfrak{P}$ its probability law and $\mathfrak{E}$ the expectation. Write $\mathcal{P}$ and $\mathcal{E}$ to denote probability and expectation for D, and $\mathrm{supp}(D)$ to denote the support of $\mathcal{P}$. The law of D can be viewed as the offspring distribution of $\mathcal{GW}$, and the law of $D+1$ the degree distribution of $\mathcal{GW}$.

Throughout the paper, we assume that the degree distribution satisfies:

Assumption 4.B. [Exponential tails]
(1) $d_{\min} := \min \mathrm{supp}(D) \geq 2$ and $\mathcal{E}[D] \in (2,\infty)$.
(2) $\mathcal{E}\left[\mathrm{e}^{aD}\right] < \infty$ for all $a \in (0,\infty)$. ■

Under this assumption, $\mathcal{GW}$ is $\mathfrak{P}$-a.s. an infinite tree. Moreover,

$$\lim_{r\to\infty} \frac{\log|B_r(\mathcal{O})|}{r} = \log \mathcal{E}[D] =: \vartheta \in (0,\infty) \qquad \mathfrak{P} - a.s., \tag{4.7}$$

where $B_r(\mathcal{O}) \subset V$ is the ball of radius r around $\mathcal{O}$ in the graph distance (see e.g. [30, pp. 134–135]). Note that this ball depends on $\mathcal{GW}$ and therefore is random. Furthermore, under this assumption on the tail of D, the solution to the PAM exists and is unique.

Theorem 4.1.1. *Let $G = \mathcal{GW}$. Subject to Assumption 4.B(2),* (4.1) *has a unique non-negative solution* $(\mathsf{P} \times \mathfrak{P})$ *almost surely. This solution admits the Feynman-Kac representation* (4.3).

For our main result we need an assumption that is much stronger than Assumption 4.B(2).

Assumption 4.C. [Super-double-exponential tails] There exists a function $f\colon (0,\infty) \to (0,\infty)$ satisfying $\lim_{s\to\infty} f(s) = 0$, $\lim_{s\to\infty} f'(s) = 0$ and $\lim_{s\to\infty} f(s)\log s = \infty$ such that

$$\limsup_{s\to\infty} \mathrm{e}^{-s} \log \mathcal{P}(D > s^{f(s)}) < -2\vartheta. \tag{4.8}$$

■

To state our main result, we define the constant

$$\widetilde{\chi}(\varrho) := \inf\big\{\chi_T(\varrho)\colon\ T \text{ is an infinite tree with degrees in } \operatorname{supp}(D)\big\}, \tag{4.9}$$

with $\chi_G(\varrho)$ defined in (4.6), and abbreviate

$$\mathfrak{r}_t = \frac{\varrho t}{\log\log t}. \tag{4.10}$$

Theorem 4.1.2. *[Quenched Lyapunov exponent] Subject to Assumptions* 4.A–4.C*,*

$$\frac{1}{t}\log U(t) = \varrho\log(\vartheta\mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) + o(1), \quad t\to\infty, \qquad (\mathsf{P}\times\mathfrak{P})\text{-a.s.} \tag{4.11}$$

With Theorem 4.1.2 we have completed our task to relax the main result in [12] to degree distributions with unbounded support. The extension comes at the price of having to assume a tail that decays faster than double-exponential as shown in (4.8). This property is needed to control the occurrence of large degrees *uniformly in large subtrees* of $\mathcal{GW}$. No doubt Assumption 4.C is stronger than is needed, but to go beyond would require a major overhaul of the methods developed in [12], which remains a challenge.

In (4.2) the initial mass is located at the root. The asymptotics in (4.11) is robust against different choices. A heuristic explanation where the terms in (4.11) come from was given in [12, Section 1.5]. The asymptotics of $U(t)$ is controlled by random walk paths in the Feynman-Kac formula in (4.4) that run within time $\mathfrak{r}_t/\varrho\log\mathfrak{r}_t$ to an *intermittent island* at distance r_t from $\mathcal{O}$, and afterwards stay near that island for the rest of the time. The intermittent island turns out to consist of a subtree with degree $d_{\min}$ where the potential has a height $\varrho\log(\vartheta\mathfrak{r}_t)$ and a shape that is the solution of a variational formula restricted to that subtree. The first and third term in (4.11) are the contribution of the path after it has reached the island, the second term is the cost for reaching the island.

For $d \in \mathbb{N}\setminus\{1\}$, let $\mathcal{T}_d$ be the infinite homogeneous tree in which every node has downward degree d. It was shown in [12] that if $\varrho \geq 1/\log(d_{\min}+1)$, then

$$\widetilde{\chi}(\varrho) = \chi_{\mathcal{T}_{d_{\min}}}(\varrho). \tag{4.12}$$

CHAPTER 4

Presumably $\mathcal{T}_{d_{\min}}$ is the *unique* minimizer of (4.9), but proving so would require more work.

Outline. The remainder of the chapter is organised as follows. Section 4.2 collects some structural properties of Galton-Watson trees. Section 4.3 contains several preparatory lemmas, which identify the maximum size of the islands where the potential is suitably high, estimate the contribution to the total mass in (4.4) by the random walk until it exits a subset of $\mathcal{GW}$, bound the principal eigenvalue associated with the islands, and estimate the number of locations where the potential is intermediate. Section 4.4 uses these preparatory lemmas to find the contribution to the Feynman-Kac formula in (4.4) coming from various sets of paths. Section 4.5 uses these contributions to prove Theorem 4.1.2. In Section 4.6 we prove Theorem 4.1.1.

Assumptions 4.A–4.B are needed throughout the paper. Only in Sections 4.4–4.5 do we need Assumption 4.C.

§4.2 Structural properties of the Galton-Watson tree

In the section we collect a few structural properties of $\mathcal{GW}$ that play an important role throughout the paper. None of these properties were needed in [12]. Section 4.2.1 looks at volumes, Section 4.2.2 at degrees, Section 4.2.3 at tree animals.

§4.2.1 Volumes

Let Z_k be the number of offspring in generation k, i.e.,

$$Z_k = |\{x \in V\colon\ d(x,\mathcal{O}) = k\}|, \tag{4.13}$$

where $d(x,\mathcal{O})$ is the distance from $\mathcal{O}$ to x. Let $\mu = \mathcal{E}[D]$. Then there exists a random variable $W \in (0,\infty)$ such that

$$W_k := \mathrm{e}^{-k\vartheta} Z_k = \mu^{-k} Z_k \to W \qquad \mathfrak{P}\text{-a.s. as } k \to \infty. \tag{4.14}$$

It is shown in [4, Theorem 5] that

$$\exists\, C < \infty, c > 0\colon \quad \mathfrak{P}(|W_k - W| \geq \varepsilon) \leq C\mathrm{e}^{-c\,\varepsilon^{2/3}\mu^{k/3}} \qquad \forall\, \varepsilon > 0,\ k \in \mathbb{N}. \tag{4.15}$$

In addition, it is shown in [6, Theorems 2–3] that if D is bounded, then

$$-\log \mathfrak{P}(W \geq x) \;=\; x^{\gamma^+/(\gamma^+-1)}\,[L^+(x) + o(1)], \qquad x \to \infty, \tag{4.16}$$

$$-\log \mathfrak{P}(W \leq x) \;=\; x^{-\gamma^-/(1-\gamma^-)}\,[L^-(x) + o(1)], \qquad x \downarrow 0, \tag{4.17}$$

where $\gamma^+ \in (1,\infty)$ and $\gamma^- \in (0,1)$ are the unique solutions of the equations

$$\mu^{\gamma^+} = d_{\max}, \qquad \mu^{\gamma^-} = d_{\min}, \tag{4.18}$$

with $L^+, L^-: (0,\infty) \to (0,\infty)$ real-analytic functions that are multiplicatively periodic with period μ^{γ^+-1}, respectively, $\mu^{1-\gamma^-}$. Note that Assumption 4.B(1) guarantees that $\gamma^- \neq 1$.

The tail behaviour in (4.16) requires that $d_{\max} < \infty$. In our setting we have $d_{\max} = \infty$, which corresponds to $\gamma^+ = \infty$, and so we expect exponential tail behaviour. The following lemma provides a rough bound.

Lemma 4.2.1. *[Exponential tail for generation sizes] If there exists an $a > 0$ such that $\mathcal{E}[\mathrm{e}^{aD}] < \infty$, then there exists an $a_* > 0$ such that $\mathfrak{E}[\mathrm{e}^{a_* W}] < \infty$.*

Proof. First note that if there exists an $a > 0$ such that $\mathcal{E}[\mathrm{e}^{aD}] < \infty$, then there exist $b > 0$ large and $c > 0$ small such that

$$\varphi(a) := \mathcal{E}[\mathrm{e}^{aD}] \leq \mathrm{e}^{\mu a + ba^2} \qquad \forall\, 0 < a < c. \tag{4.19}$$

Hence

$$\mathfrak{E}[\mathrm{e}^{aZ_{n+1}}] = \mathfrak{E}[\varphi(a)^{Z_n}] \leq \mathfrak{E}[\mathrm{e}^{(\mu a + ba^2)Z_n}] \tag{4.20}$$

and consequently, because $\mu > 1$,

$$\mathfrak{E}\left[\mathrm{e}^{aW_{n+1}}\right] \leq \mathfrak{E}\left[\mathrm{e}^{(a+ba^2\mu^{-(n+2)})W_n}\right] \leq \mathfrak{E}\left[\mathrm{e}^{a\exp(bc\mu^{-(n+2)})W_n}\right]. \tag{4.21}$$

Put $a_n := c\exp(-bc\sum_{k=0}^{n-1}\mu^{-(k+2)})$, which satisfies $0 < a_n \leq c$. From the last inequality in (4.21) it follows that

$$\mathfrak{E}\left[\mathrm{e}^{a_{n+1}W_{n+1}}\right] \leq \mathfrak{E}\left[\mathrm{e}^{a_n W_n}\right]. \tag{4.22}$$

Since $n \mapsto a_n$ is decreasing with $\lim_{n\to\infty} a_n = a_* > 0$, Fatou's lemma gives

$$\mathfrak{E}\left[\mathrm{e}^{a_* W}\right] \leq \mathfrak{E}\left[\mathrm{e}^{a_0 W_0}\right]. \tag{4.23}$$

Because $\mathcal{E}[\mathrm{e}^{a_0 W_0}] = \mathrm{e}^{a_0} < \infty$, we get the claim. □

The following lemma says that $\mathfrak{P}$-a.s. a ball of radius R_r centred anywhere in $B_r(\mathcal{O})$ has volume $\mathrm{e}^{\vartheta R_r + o(R_r)}$ as $r \to \infty$, provided R_r is large compared to $\log r$.

Lemma 4.2.2. *[Volumes of large balls] Subject to Assumption 4.B(1), if there exists an $a > 0$ such that $\mathcal{E}[\mathrm{e}^{aD}] < \infty$, then for any R_r satisfying $\lim_{r\to\infty} R_r/\log r = \infty$,*

$$\liminf_{r\to\infty} \frac{1}{R_r} \log\Big(\inf_{x\in B_r(\mathcal{O})} |B_{R_r}(x)|\Big) = \limsup_{r\to\infty} \frac{1}{R_r} \log\Big(\sup_{x\in B_r(\mathcal{O})} |B_{R_r}(x)|\Big) = \vartheta \qquad \mathfrak{P}\text{-a.s.} \tag{4.24}$$

Proof. For $y \in \mathcal{GW}$ that lies k generations below $\mathcal{O}$, let $y[-i]$, $0 \leq i \leq k$ be the vertex that lies i generations above y. Define the *lower ball* of radius r around y as

$$B_r^{\downarrow}(y) := \{x \in V\colon\ \exists\, 0 \leq i \leq r \text{ with } x[-i] = y\}. \tag{4.25}$$

Note that $B_r^{\downarrow}(\mathcal{O}) = B_r(\mathcal{O})$.

We first prove the claim for lower balls. Afterwards we use a sandwich argument to get the claim for balls.

Let $\mathcal{Z}_k$ denote the vertices in the k-th generation. To get the upper bound, pick $\delta > 0$ and estimate

$$\begin{aligned}
\mathfrak{P}\Big(\sup_{x\in B_r(\mathcal{O})} |B^{\downarrow}_{R_r}(x)| \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big) &\leq \sum_{k=0}^{r} \mathfrak{P}\Big(\sup_{x\in\mathcal{Z}_k} |B^{\downarrow}_{R_r}(x)| \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big) \\
&= \sum_{k=0}^{r}\sum_{l\in\mathbb{N}} \mathfrak{P}\Big(\sup_{x\in\mathcal{Z}_k} |B^{\downarrow}_{R_r}(x)| \geq \mathrm{e}^{(1+\delta)\vartheta R_r} \;\Big|\; Z_k = l\Big)\mathfrak{P}(Z_k = l) \\
&\leq \sum_{k=0}^{r}\sum_{l\in\mathbb{N}} l\, \mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big)\mathfrak{P}(Z_k = l) \\
&= \mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big)\sum_{k=0}^{r}\mathfrak{E}(Z_k).
\end{aligned} \tag{4.26}$$

By (4.7), $\sum_{k=0}^r \mathfrak{E}(Z_k) = \frac{\mathrm{e}^{\vartheta(r+1)}-1}{\mathrm{e}^{\vartheta}-1} = O(\mathrm{e}^{\vartheta r})$, and so in order to be able to apply the Borel-Cantelli lemma, it suffices to show that the probability in the last line decays faster than exponentially in r for any $\delta > 0$. To that end, estimate

$$\begin{aligned}
\mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big) &= \mathfrak{P}\Big(\sum_{k=0}^{R_r} Z_k \geq \mathrm{e}^{(1+\delta)\vartheta R_r}\Big) \\
&= \mathfrak{P}\Big(\sum_{k=0}^{R_r} W_k \geq \mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \leq \sum_{k=0}^{R_r}\mathfrak{P}\Big(W_k \geq \frac{1}{R_r+1}\mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \\
&= \sum_{k=0}^{R_r}\mathfrak{P}\Big(W + (W_k - W) \geq \frac{1}{R_r+1}\mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \\
&\leq \sum_{k=0}^{R_r}\mathfrak{P}\Big(W \geq \frac{1}{2(R_r+1)}\mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \\
&\qquad + \sum_{k=0}^{R_r}\mathfrak{P}\Big(|W_k - W| \geq \frac{1}{2(R_r+1)}\mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \\
&\leq \mathfrak{E}[\mathrm{e}^{a_* W}]\sum_{k=0}^{R_r}\exp\Big(-a_*\frac{1}{2(R_r+1)}\mathrm{e}^{\delta\vartheta R_r}\mathrm{e}^{\vartheta(R_r-k)}\Big) \\
&\qquad + \sum_{k=0}^{R_r} C\exp\Big(-c\Big[\frac{1}{2(R_r+1)}\,\mathrm{e}^{\delta\vartheta R_r}\,\mathrm{e}^{\vartheta(R_r-k)}\Big]^{2/3}(\mathrm{e}^{\vartheta})^{k/3}\Big) \\
&\leq \mathfrak{E}[\mathrm{e}^{a_* W}](R_r+1)\exp\Big(-a_*\frac{1}{2(R_r+1)}\mathrm{e}^{\delta\vartheta R_r}\Big) \\
&\qquad + C(R_r+1)\exp\Big(-c\Big[\frac{1}{2(R_r+1)}\,\mathrm{e}^{\delta\vartheta R_r}\Big]^{2/3}\Big),
\end{aligned} \tag{4.27}$$

where we use (4.15) with $\mu = \mathrm{e}^{\vartheta}$. This produces the desired estimate.

To get the lower bound, pick $0 < \delta < 1$ and estimate

$$\begin{aligned}
\mathfrak{P}\Big(\inf_{x \in B_r(\mathcal{O})} |B^{\downarrow}_{R_r}(x)| \leq \mathrm{e}^{(1-\delta)\vartheta R_r}\Big) &\leq \sum_{k=0}^{r} \mathfrak{P}\Big(\inf_{x \in \mathcal{Z}_k} |B^{\downarrow}_{R_r}(x)| \leq \mathrm{e}^{(1-\delta)\vartheta R_r}\Big) \\
&= \sum_{k=0}^{r} \sum_{l \in \mathbb{N}} \mathfrak{P}\Big(\inf_{x \in \mathcal{Z}_k} |B^{\downarrow}_{R_r}(x)| \leq \mathrm{e}^{(1-\delta)\vartheta R_r} \;\Big|\; Z_k = l\Big) \mathfrak{P}(Z_k = l) \\
&\leq \sum_{k=0}^{r} \sum_{l \in \mathbb{N}} l\, \mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \leq \mathrm{e}^{(1-\delta)\vartheta R_r}\Big) \mathfrak{P}(Z_k = l) \\
&= \mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \leq \mathrm{e}^{(1-\delta)\vartheta R_r}\Big) \sum_{k=0}^{r} \mathfrak{E}(Z_k).
\end{aligned} \tag{4.28}$$

It again suffices to show that the probability in the last line decays faster than exponentially in r for any $\delta > 0$. To that end, estimate

$$\begin{aligned}
\mathfrak{P}\Big(|B^{\downarrow}_{R_r}(\mathcal{O})| \leq \mathrm{e}^{(1-\delta)\vartheta R_r}\Big) &= \mathfrak{P}\Big(\mathrm{e}^{-\vartheta R_r} \sum_{k=0}^{R_r} Z_k \leq \mathrm{e}^{-\delta\vartheta R_r}\Big) \\
&\leq \mathfrak{P}\Big(W_{R_r} \leq \mathrm{e}^{-\delta\vartheta R_r}\Big) \leq \mathfrak{P}(W \leq 2\,\mathrm{e}^{-\delta\vartheta R_r}) + \mathfrak{P}(W - W_{R_r} \geq \mathrm{e}^{-\delta\vartheta R_r}) \\
&\leq \exp\Big(-c^{-}(2\mathrm{e}^{\delta\vartheta R_r})^{\frac{\gamma^-}{1-\gamma^-}}[1+o(1)]\Big) + C\exp\Big(-c\,[\mathrm{e}^{-\frac{2}{3}\delta\vartheta}(\mathrm{e}^{\vartheta})^{\frac{1}{3}}]^{R_r}\Big),
\end{aligned} \tag{4.29}$$

where we use (4.17), (4.15) with $\mu = \mathrm{e}^{\vartheta}$, and put $c^- := \inf L^- \in (0,\infty)$. For δ small enough this produces the desired estimate. This completes the proof of (4.24) for lower balls.

To get the claim for balls, we observe that

$$B^{\downarrow}_r(x) \subseteq B_r(x) \subseteq \bigcup_{k=0}^{r} B^{\downarrow}_r(x[-k]), \tag{4.30}$$

and therefore

$$|B^{\downarrow}_r(x)| \leq |B_r(x)| \leq \sum_{k=0}^{r} |B^{\downarrow}_r(x[-k])|. \tag{4.31}$$

It follows from (4.31) that

$$\inf_{x \in B_r(\mathcal{O})} |B^{\downarrow}_r(x)| \leq \inf_{x \in B_r(\mathcal{O})} |B_r(x)| \leq \sup_{x \in B_r(\mathcal{O})} |B_r(x)| \leq (r+1) \sup_{x \in B_r(\mathcal{O})} |B^{\downarrow}_r(x)|. \tag{4.32}$$

Hence we get (4.24). □

§4.2.2 Degrees

Write D_x to denote the degree of vertex x. The following lemma implies that, $\mathfrak{P}$-a.s. and for $r \to \infty$, D_x is bounded by a vanishing power of $\log r$ for all $x \in B_{2r}(\mathcal{O})$.

CHAPTER 4

Lemma 4.2.3. *[Maximal degree in a ball around the root]*
(a) Subject to Assumption 4.B(2)*, for every* $\delta > 0$*,*

$$\sum_{r\in\mathbb{N}} \mathfrak{P}\big(\exists\, x \in B_{2r}(\mathcal{O})\colon\ D_x > \delta r\big) < \infty. \tag{4.33}$$

(b) Subject to Assumption 4.C*, there exists a function* $\delta_r\colon (0,\infty) \to (0,\infty)$ *satisfying* $\lim_{r\to\infty} \delta_r = 0$ *and* $\lim_{r\to\infty} r\frac{\mathrm{d}}{\mathrm{d}r}\delta_r = 0$ *such that*

$$\sum_{r\in\mathbb{N}} \mathfrak{P}\big(\exists\, x \in B_{2r}(\mathcal{O})\colon\ D_x > (\log r)^{\delta_r}\big) < \infty. \tag{4.34}$$

Proof. (a) Estimate

$$\begin{aligned}
\mathfrak{P}\big(\exists\, x \in B^{\downarrow}_{2r}(\mathcal{O})\colon\ D_x > \delta r\big) &\le \sum_{k=0}^{2r} \mathfrak{P}\big(\exists\, x \in \mathcal{Z}_k\colon\ D_x > \delta r\big)\\
&= \sum_{k=0}^{2r}\sum_{l\in\mathbb{N}} \mathfrak{P}\big(\exists\, x \in \mathcal{Z}_k\colon\ D_x > \delta r \mid Z_k = l\big)\,\mathfrak{P}(Z_k = l)\\
&\le \mathcal{P}(D > \delta r)\sum_{k=0}^{2r}\sum_{l\in\mathbb{N}} l\,\mathfrak{P}\big(Z_k = l\big) = \mathcal{P}(D > \delta r)\sum_{k=0}^{2r}\mathfrak{E}(Z_k).
\end{aligned} \tag{4.35}$$

Since $\sum_{k=0}^{2r}\mathfrak{E}(Z_k) = \frac{\mathrm{e}^{(2r+1)\vartheta}-1}{\mathrm{e}^{\vartheta}-1} = O(\mathrm{e}^{2r\vartheta})$, it suffices to show that $\mathcal{P}(D > \delta r) = O(\mathrm{e}^{-cr})$ for some $c > 2\vartheta$. Since $\mathcal{P}(D > \delta r) \le \mathrm{e}^{-a\delta r}\mathcal{E}(\mathrm{e}^{aD})$, the latter is immediate from Assumption 4.B(2) when we choose $a > 2\vartheta/\delta$.
(b) The only change is that in the last line $\mathcal{P}(D > \delta r)$ must be replaced by $\mathcal{P}(D > (\log r)^{\delta_r})$. To see that the latter is $O(\mathrm{e}^{-cr})$ for some $c > 2\vartheta$, we use the tail condition in (4.8) with $\delta_r = f(s)$ and $s = \log r$. □

Chapter 4

§4.2.3 Tree animals

For $n \in \mathbb{N}_0$ and $x \in B_r(\mathcal{O})$, let

$$\mathcal{A}_n(x) = \{\Lambda \subset B_n(x)\colon\ \Lambda \text{ is connected}, \Lambda \ni x, |\Lambda| = n+1\} \tag{4.36}$$

be the set of *tree animals* of size $n+1$ that contain x. Put $a_n(x) = |\mathcal{A}_n(x)|$.

Lemma 4.2.4. *[Number of tree animals] Subject to Assumption* 4.B(2)*,* $\mathfrak{P}$*-a.s. there exists an* $r_0 \in \mathbb{N}$ *such that* $a_n(x) \le r^n$ *for all* $r \ge r_0$*,* $x \in B_r(\mathcal{O})$ *and* $0 \le n \le r$*.*

Proof. For $n \in \mathbb{N}_0$ and $x \in B^{\downarrow}_r(\mathcal{O})$, let

$$\mathcal{A}^{\downarrow}_n(x) = \{\Lambda \subset B^{\downarrow}_n(x)\colon\ \Lambda \text{ is connected}, \Lambda \ni x, |\Lambda| = n+1\} \tag{4.37}$$

be the set of *lower tree animals* of size $n+1$ that contain x. Put $a^{\downarrow}_n(x) = |\mathcal{A}^{\downarrow}_n(x)|$.

We first prove the claim for lower tree animals. Afterwards we use a sandwich argument to get the claim for tree animals.

Fix $\delta > 0$. By Lemma 4.2.3(a) and the Borel-Cantelli lemma, $\mathfrak{P}$-a.s. there exists an $r_0 = r_0(\delta) \in \mathbb{N}$ such that $D_x \leq \delta r$ for all $x \in B^{\downarrow}_{2r}(\mathcal{O})$. Any lower tree animal of size $n+1$ containing a vertex in $B^{\downarrow}_r(\mathcal{O})$ is contained in $B^{\downarrow}_{r+n}(\mathcal{O})$. Any lower tree animal of size $n+1$ can be created by adding a vertex to the outer boundary of a lower tree animal of size n. This leads to the recursive inequality

$$a^{\downarrow}_n(x) \leq (\delta r) a^{\downarrow}_{n-1}(x) \qquad \forall\, x \in B^{\downarrow}_r(\mathcal{O}), \quad 1 \leq n \leq r. \tag{4.38}$$

Since $a^{\downarrow}_0(x) = 1$, it follows that

$$a^{\downarrow}_n(x) \leq (\delta r)^n \qquad \forall\, x \in B^{\downarrow}_r(\mathcal{O}),\ 0 \leq n \leq r. \tag{4.39}$$

Pick $\delta \leq 1$ to get the claim for lower tree animals.

To get the claim for tree animals, pick $\delta \leq \frac{1}{(n+1)}$ and note that $a_n(x) \leq \sum_{k=0}^{n} a^{\downarrow}_n(x[-k])$ (compare with (4.31)), and so $a_n(x) \leq r^n$ for all $x \in B_r(\mathcal{O})$ and all $0 \leq n \leq r$. □

§4.3 Preliminaries

In this section we extend the lemmas in [12, Section 2]. Section 4.3.1 identifies the maximum size of the islands where the potential is suitably high. Section 4.3.2 estimates the contribution to the total mass in (4.4) by the random walk until it exits a subset of $\mathcal{GW}$. Section 4.3.3 gives a bound on the principal eigenvalue associated with the islands. Section 4.3.5 estimates the number of locations where the potential is intermediate.

Abbreviate $L_r = L_r(\mathcal{GW}) = |B_r(\mathcal{O})|$ and put

$$S_r := (\log r)^{\alpha}, \qquad \alpha \in (0,1). \tag{4.40}$$

§4.3.1 Maximum size of the islands

For every $r \in \mathbb{N}$ there is a unique a_r such that

$$\mathsf{P}(\xi(0) > a_r) = \frac{1}{r}. \tag{4.41}$$

By Assumption 4.A, for r large enough

$$a_r = \varrho \log\log r. \tag{4.42}$$

For $r \in \mathbb{N}$ and $A > 0$, let

$$\Pi_{r,A} = \Pi_{r,A}(\xi) := \{z \in B_r(\mathcal{O})\colon\ \xi(z) > a_{L_r} - 2A\} \tag{4.43}$$

be the set of vertices in $B_r(\mathcal{O})$ where the potential is close to maximal,

$$D_{r,A} = D_{r,A}(\xi) := \{z \in B_r(\mathcal{O})\colon\ \mathrm{dist}(z, \Pi_{r,A}) \leq S_r\} \tag{4.44}$$

be the S_r-neighbourhood of $\Pi_{r,A}$, and $\mathfrak{C}_{r,A}$ be the set of connected components of $D_{r,A}$ in $\mathcal{GW}$, which we think of as *islands*. For $M_A \in \mathbb{N}$, define the event

$$\mathcal{B}_{r,A} := \big\{\exists \mathcal{C} \in \mathfrak{C}_{r,A} \colon \ |\mathcal{C} \cap \Pi_{r,A}| > M_A\big\}. \tag{4.45}$$

Note that $\Pi_{r,A}, D_{r,A}, \mathcal{B}_{r,A}$ depend on $\mathcal{GW}$ and therefore are random.

Lemma 4.3.1. *[Maximum size of the islands] Subject to Assumptions* 4.A–4.B*, for every $A > 0$ there exists an $M_A \in \mathbb{N}$ such that*

$$\sum_{r\in\mathbb{N}} \mathsf{P}(\mathcal{B}_{r,A}) < \infty \qquad \mathfrak{P} - a.s. \tag{4.46}$$

Proof. We follow [8, Lemma 6.6]. By Assumption 4.A, for every $x \in V$ and r large enough,

$$\mathsf{P}(x \in \Pi_{r,A}) = \mathsf{P}(\xi(x) > a_{L_r} - 2A) = L_r^{-c_A} \tag{4.47}$$

with $c_A = e^{-2A/\varrho}$. By Lemma 4.2.2, $\mathfrak{P}$-a.s. for every $y \in B_r(\mathcal{O})$ and r large enough,

$$|B_{S_r}(y)| \le |B_{o(r)}(\mathcal{O})| = L_{o(r)} = L_r^{o(1)}, \tag{4.48}$$

where we use that $S_r = o(\log r) = o(r)$, and hence for every $m \in \mathbb{N}$,

$$\mathsf{P}(|B_{S_r}(y) \cap \Pi_{r,A}| \ge m) \le \binom{|B_{S_r}(y)|}{m} L_r^{-c_A m} \le (|B_{S_r}(y)| L_r^{-c_A})^m \le L_r^{-c_A m[1+o(1)]}. \tag{4.49}$$

Consequently, $\mathfrak{P}$-a.s.

$$\begin{aligned} \mathsf{P}(\exists \mathcal{C} \in \mathfrak{C}_{r,A} \colon \ |\mathcal{C} \cap \Pi_{r,A}| \ge m) &\le \mathsf{P}(\exists y \in B_r(\mathcal{O}) \colon \ |B_{S_r}(y) \cap \Pi_{r,A}| \ge m) \\ &\le |B_r(\mathcal{O})| L_r = L_r^{(1-c_A m)[1+o(1)]}. \end{aligned} \tag{4.50}$$

By choosing $m > 1/c_A$, we see that the above probability becomes summable in r, and so we have proved the claim with $M_A = \lceil 1/c_A \rceil$. ☐

Lemma 4.3.1 implies that $(\mathsf{P} \times \mathfrak{P})$-a.s. $\mathcal{B}_{r,A}$ does not occur eventually as $r \to \infty$. Note that $\mathfrak{P}$-a.s. on the event $[\mathcal{B}_{r,A}]^c$,

$$\forall \mathcal{C} \in \mathfrak{C}_{r,A} \colon \ |\mathcal{C} \cap \Pi_{r,A}| \le M_A, \ \mathrm{diam}_{\mathcal{GW}}(\mathcal{C}) \le 2 M_A S_r, \ |\mathcal{C}| \le \mathrm{e}^{2\vartheta M_A S_r}, \tag{4.51}$$

where the last inequality follows from Lemma 4.2.2.

§4.3.2 Mass up to an exit time

Lemma 4.3.2. *[Mass up to an exit time] Subject to Assumption* 4.B(2)*, $\mathfrak{P}$-a.s. for any $\delta > 0$, $r \ge r_0$, $y \in \Lambda \subset B_r(\mathcal{O})$, $\xi \in [0,\infty)^V$ and $\gamma > \lambda_\Lambda = \lambda_\Lambda(\xi, \mathcal{GW})$,*

$$\mathbb{E}_y\left[\mathrm{e}^{\int_0^{\tau_{\Lambda^c}} (\xi(X_s)-\gamma)\,\mathrm{d}s}\right] \le 1 + \frac{(\delta r)\,|\Lambda|}{\gamma - \lambda_\Lambda}. \tag{4.52}$$

Chapter 4

Proof. We follow the proof of [25, Lemma 2.18] and [19, Lemma 4.2]. Define

$$u(x) := \mathbb{E}_x \left[\mathrm{e}^{\int_0^{\tau_{\Lambda^c}} (\xi(X_s)-\gamma)\,\mathrm{d}s} \right]. \tag{4.53}$$

This is the solution to the boundary value problem

$$\begin{aligned} (\Delta + \xi - \gamma)u = 0 \quad &\text{on } \Lambda \\ u = 1 \quad &\text{on } \Lambda^{\mathrm{c}}. \end{aligned} \tag{4.54}$$

Via the substitution $u =: 1 + v$, this turns into

$$\begin{aligned} (\Delta + \xi - \gamma)v = \gamma - \xi \quad &\text{on } \Lambda \\ v = 0 \qquad\quad &\text{on } \Lambda^{\mathrm{c}}. \end{aligned} \tag{4.55}$$

It is readily checked that for $\gamma > \lambda_\Lambda$ the solution exists and is given by

$$v = \mathcal{R}_\gamma(\xi - \gamma), \tag{4.56}$$

where $\mathcal{R}_\gamma$ denotes the resolvent of $\Delta + \xi$ in $\ell^2(\Lambda)$ with Dirichlet boundary condition. Hence

$$v(x) \le (\delta r)\,(\mathcal{R}_\gamma \mathbb{1})(x) \le (\delta r)\,\langle \mathcal{R}_\gamma \mathbb{1}, \mathbb{1}\rangle_\Lambda \le \frac{(\delta r)\,|\Lambda|}{\gamma - \lambda_\Lambda}, \quad x \in \Lambda, \tag{4.57}$$

where $\mathbb{1}$ denotes the constant function equal to 1, and $\langle \cdot, \cdot \rangle_\Lambda$ denotes the inner product in $\ell^2(\Lambda)$. To get the first inequality, we combine Lemma 4.2.3(a) with the lower bound in (B.2) from Lemma B.1.1, to get $\xi - \gamma \le \lambda_\Lambda + \delta r - \gamma \le \delta r$ on Λ. The positivity of the resolvent gives

$$0 \le [\mathcal{R}_\gamma(\delta r - (\xi - \gamma))](x) = (\delta r)\,[\mathcal{R}_\gamma \mathbb{1}](x) - [\mathcal{R}_\gamma(\xi - \gamma)](x). \tag{4.58}$$

To get the second inequality, we write

$$(\delta r)\,(\mathcal{R}_\gamma \mathbb{1})(x) \le (\delta r) \sum_{x\in\Lambda} (\mathcal{R}_\gamma \mathbb{1})(x) = (\delta r) \sum_{x\in\Lambda} (\mathcal{R}_\gamma \mathbb{1})(x)\mathbb{1}(x) = (\delta r)\,\langle \mathcal{R}_\gamma \mathbb{1}, \mathbb{1}\rangle_\Lambda. \tag{4.59}$$

To get the third inequality, we use the Fourier expansion of the resolvent with respect to the orthonormal basis of eigenfunctions of $\Delta + \xi$ in $\ell^2(\Lambda)$. □

§4.3.3 Principal eigenvalue of the islands

The following lemma provides a spectral bound.

Lemma 4.3.3. *[Principal eigenvalues of the islands] Subject to Assumptions* 4.A *and* 4.B(2), *for any* $\varepsilon > 0$, $(\mathsf{P} \times \mathfrak{P})$*-a.s. eventually as* $r \to \infty$,

$$\textit{all } \mathcal{C} \in \mathfrak{C}_{r,A} \textit{ satisfy: } \lambda_{\mathcal{C}}(\xi; \mathcal{GW}) \le a_{L_r} - \widehat{\chi}_{\mathcal{C}}(\mathcal{GW}) + \varepsilon. \tag{4.60}$$

Proof. We follow the proof of [12, Lemma 2.6]. For $\varepsilon > 0$ and $A > 0$, define the event

$$\bar{\mathcal{B}}_{r,A} := \left\{ {\text{there exists a connected subset } \Lambda\subset V \text{ with } \Lambda\cap B_r(\mathcal{O})\neq\emptyset, \atop |\Lambda|\le \mathrm{e}^{2\vartheta M_A S_r},\ \lambda_\Lambda(\xi;\mathcal{GW})>a_{L_r}-\widehat{\chi}_\Lambda(\mathcal{GW})+\varepsilon} \right\} \tag{4.61}$$

with M_A as in Lemma 4.3.1. Note that, by (4.A), $\mathrm{e}^{\xi(x)/\varrho}$ is stochastically dominated by $Z \vee N$, where Z is an Exp(1) random variable and $N > 0$ is a constant. Thus, for any $\Lambda \subset V$, using [12, Eq. (2.17)], putting $\gamma = \sqrt{\mathrm{e}^{\varepsilon/\varrho}} > 1$ and applying Markov's inequality, we may estimate

$$\begin{aligned} &\mathsf{P}\left(\lambda_\Lambda(\xi;\mathcal{GW}) > a_{L_r} - \widehat{\chi}_\Lambda(\mathcal{GW}) + \varepsilon\right) \le \mathsf{P}\left(\mathcal{L}_\Lambda(\xi - a_{L_r} - \varepsilon) > 1\right) \\ &= \mathsf{P}\left(\gamma^{-1}\mathcal{L}_\Lambda(\xi) > \gamma \log L_r\right) \le \mathrm{e}^{-\gamma\log L_r}\mathrm{E}[\mathrm{e}^{\gamma^{-1}\mathcal{L}_\Lambda(\xi)}] \le \mathrm{e}^{-\gamma\log L_r} K_\gamma^{|\Lambda|} \end{aligned} \tag{4.62}$$

with $K_\gamma = \mathrm{E}[\mathrm{e}^{\gamma^{-1}(Z\vee N)}] \in (1,\infty)$. Next, by Lemma 4.2.4, for any $x \in B_r(\mathcal{O})$ and $1 \le n \le r$, the number of connected subsets $\Lambda \subset V$ with $x \in \Lambda$ and $|\Lambda| = n + 1$ is $\mathfrak{P}$-a.s. at most $(n+1)r^n \le \mathrm{e}^{2n\log r}$ for $r \ge r_0$. Noting that $\mathrm{e}^{S_r} \le r$, we use a union bound and that by Lemma 4.2.2 $\log L_r = \vartheta r + o(r)$ as $r \to \infty$ $\mathfrak{P}$-a.s., to estimate for r large enough,

$$\begin{aligned} \mathsf{P}(\bar{\mathcal{B}}_{r,A}) &\le \mathrm{e}^{-(\gamma-1)\log L_r} \sum_{n=1}^{\lfloor \mathrm{e}^{2\vartheta M_A S_r}\rfloor} \mathrm{e}^{2n\log r} K_\gamma^n \\ &\le \mathrm{e}^{2\vartheta M_A S_r} \exp\left\{-\vartheta(\gamma-1)r + o(r) + (2\log r + \log K_\gamma)\,\mathrm{e}^{2\vartheta M_A S_r}\right\} \\ &= r^{o(1)} \exp\left\{-\vartheta(\gamma-1)r + o(r) + (\log r)\, r^{o(1)}\right\} \le \mathrm{e}^{-\frac{1}{2}\vartheta(\gamma-1)r}. \end{aligned} \tag{4.63}$$

Via the Borel-Cantelli lemma this implies that $(\mathsf{P}\times\mathfrak{P})$-a.s. $\bar{\mathcal{B}}_{r,A}$ does not occur eventually as $r \to \infty$. The proof is completed by invoking Lemma 4.3.1. ☐

Corollary 4.3.4. *[Uniform bound on principal eigenvalue of the islands] Subject to Assumptions* 4.A–4.B*, for ϑ as in* (4.7)*, and any $\varepsilon > 0$, $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $r \to \infty$,*

$$\max_{\mathcal{C}\in\mathfrak{C}_{r,A}} \lambda^1_{\mathcal{C}}(\xi;G) \le a_{L_r} - \widetilde{\chi}(\varrho) + \varepsilon. \tag{4.64}$$

Proof. See [12, Corollary 2.8]. The proof carries over verbatim because the degrees play no role. ☐

§4.3.4 Maximum of the potential

The next lemma shows that a_{L_r} is the leading order of the maximum of ξ in $B_r(\mathcal{O})$.

Lemma 4.3.5. *[Maximum of the potential] Subject to Assumptions* 4.A–4.B*, for any $\vartheta > 0$, $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $r \to \infty$,*

$$\left| \max_{x\in B_r(\mathcal{O})} \xi(x) - a_{L_r} \right| \le \frac{2\varrho\log r}{\vartheta r}. \tag{4.65}$$

Proof. See [12, Lemma 2.4]. The proof carries over verbatim and uses Lemma 4.2.2. ☐

Chapter 4

§4.3.5 Number of intermediate peaks of the potential

We recall the following Chernoff bound for a binomial random variable with parameters n and p (see e.g. [9, Lemma 5.9]):

$$P\left(\mathrm{Bin}(n,p) \geq u\right) \leq \mathrm{e}^{-u[\log(\frac{u}{np})-1]}, \qquad u > 0. \tag{4.66}$$

Lemma 4.3.6. *[Number of intermediate peaks of the potential] Subject to Assumptions* 4.A *and* 4.B(2)*, for any* $\beta \in (0,1)$ *and* $\varepsilon \in (0, \frac{1}{2}\beta)$ *the following holds. For a self-avoiding path* π *in* $\mathcal{GW}$*, set*

$$N_\pi = N_\pi(\xi) := |\{z \in \mathrm{supp}(\pi)\colon\ \xi(z) > (1-\varepsilon)a_{L_r}\}|. \tag{4.67}$$

Define the event

$$\mathcal{B}_r := \left\{ {\textit{there exists a self-avoiding path } \pi \textit{ in } \mathcal{GW} \textit{ with} \atop \mathrm{supp}(\pi)\cap B_r \neq \emptyset,\, |\mathrm{supp}(\pi)| \geq (\log L_r)^\beta \textit{ and } N_\pi > \frac{|\mathrm{supp}(\pi)|}{(\log L_r)^\varepsilon}} \right\}. \tag{4.68}$$

Then

$$\sum_{r\in\mathbb{N}_0} \mathsf{P}(\mathcal{B}_r) < \infty \qquad \mathfrak{P} - a.s. \tag{4.69}$$

Proof. We follow the proof of [12, Lemma 2.9]. Fix $\beta \in (0,1)$ and $\varepsilon \in (0, \frac{1}{2}\beta)$. (4.5) implies

$$p_r := \mathsf{P}(\xi(0) > (1-\varepsilon)a_{L_r}) = \exp\left\{-(\log L_r)^{1-\varepsilon}\right\}. \tag{4.70}$$

Fix $x \in B_r(\mathcal{O})$ and $k \in \mathbb{N}$. The number of self-avoiding paths π in $B_r(\mathcal{O})$ with $|\mathrm{supp}(\pi)| = k$ and $\pi_0 = x$ is at most $\mathrm{e}^{k\log r}$ by Lemma 4.2.4 for r sufficiently large. For such a π, the random variable N_π has a $\mathrm{Bin}(k, p_r)$-distribution. Using (4.66), we obtain

$$\begin{aligned}\mathsf{P}\Big(&\exists \text{ self-avoiding } \pi \text{ with } |\mathrm{supp}(\pi)| = k, \pi_0 = x \text{ and } N_\pi > k/(\log L_r)^\varepsilon\Big) \\ &\leq \exp\Big\{ -k\Big((\log L_r)^{1-2\varepsilon} - \log r - \frac{1+\varepsilon\log\log L_r}{(\log L_r)^\varepsilon}\Big)\Big\}.\end{aligned} \tag{4.71}$$

By the definition of ε, together with the fact that $L_r > r$ and $x \mapsto (\log\log x)/(\log x)^\varepsilon$ is eventually decreasing, the expression in parentheses above is at least $\frac{1}{2}(\log L_r)^{1-2\varepsilon}$. Summing over $k \geq (\log L_r)^\beta$ and $x \in B_r(\mathcal{O})$, we get $\mathfrak{P} - a.s.$

$$\mathsf{P}\left(\mathcal{B}_r\right) \leq 2L_r \exp\Big\{-\tfrac{1}{2}(\log L_r)^{1+\beta-2\varepsilon}\Big\} \leq c_1 \exp\Big\{-c_2(\log L_r)^{1+\delta}\Big\} \tag{4.72}$$

for some $c_1, c_2, \delta > 0$. Since $L_r > r$, (4.72) is summable in r. □

Lemma 4.3.6 implies that $(\mathsf{P} \times \mathfrak{P})$-a.s. for r large enough, all self-avoiding paths π in $\mathcal{GW}$ with $\mathrm{supp}(\pi) \cap B_r \neq \emptyset$ and $|\mathrm{supp}(\pi)| \geq (\log L_r)^\beta$ satisfy $N_\pi \leq \frac{|\mathrm{supp}(\pi)|}{(\log L_r)^\varepsilon}$.

Lemma 4.3.7. *[Number of high exceedances of the potential] Subject to Assumptions* 4.A *and* 4.B(2)*, for any $A > 0$ there is a $C \geq 1$ such that, for all $\delta \in (0,1)$, the following holds. For a self-avoiding path π in $\mathcal{GW}$, let*

$$N_\pi := |\{x \in \operatorname{supp}(\pi)\colon\ \xi(x) > a_{L_r} - 2A\}|. \tag{4.73}$$

Define the event

$$\mathcal{B}_r := \left\{ {\textit{there exists a self-avoiding path } \pi \textit{ in } G \textit{ with} \atop \operatorname{supp}(\pi)\cap B_r \neq \emptyset,\, |\operatorname{supp}(\pi)| \geq C(\log L_r)^\delta \textit{ and } N_\pi > \frac{|\operatorname{supp}(\pi)|}{(\log L_r)^\delta}} \right\}. \tag{4.74}$$

Then $\sum_{r\in\mathbb{N}_0} \sup_{G\in\mathfrak{G}_r} \mathsf{P}(\mathcal{B}_r) < \infty$. In particular, $(\mathsf{P}\times\mathfrak{P})$-a.s. for r large enough, all self-avoiding paths π in $\mathcal{GW}$ with $\operatorname{supp}(\pi)\cap B_r \neq \emptyset$ and $|\operatorname{supp}(\pi)| \geq C(\log L_r)^\delta$ satisfy

$$N_\pi = |\{x \in \operatorname{supp}(\pi)\colon\ \xi(x) > a_{L_r} - 2A\}| \leq \frac{|\operatorname{supp}(\pi)|}{(\log L_r)^\delta}. \tag{4.75}$$

Proof. Proceed as for Lemma 4.3.6, noting that this time

$$p_r := \mathsf{P}\left(\xi(0) > a_{L_r} - 2A\right) = L_r^{-\epsilon} \tag{4.76}$$

where $\epsilon = \mathrm{e}^{-2A/\varrho}$, and taking $C > 2/\epsilon$. □

§4.4 Path expansions

In this section we extend [12, Section 3]. Section 4.4.1 proves three lemmas that concern the contribution to the total mass in (4.4) coming from various sets of paths. Section 4.4.2 proves a key proposition that controls the entropy associated with a key set of paths. The proof is based on the three lemmas in Section 4.4.1.

We need various sets of nearest-neighbour paths in $\mathcal{GW} = (V, E, \mathcal{O})$, defined in [12]. For $\ell \in \mathbb{N}_0$ and subsets $\Lambda, \Lambda' \subset V$, put

$$\begin{aligned} \mathscr{P}_\ell(\Lambda,\Lambda') &:= \left\{ (\pi_0,\dots,\pi_\ell) \in V^{\ell+1}\colon \quad \begin{matrix} \pi_0 \in \Lambda, \pi_\ell \in \Lambda', \\ \{\pi_i,\pi_{i-1}\} \in E\ \forall\, 1 \leq i \leq \ell \end{matrix} \right\}, \\ \mathscr{P}(\Lambda,\Lambda') &:= \bigcup_{\ell\in\mathbb{N}_0} \mathscr{P}_\ell(\Lambda,\Lambda'), \end{aligned} \tag{4.77}$$

and set

$$\mathscr{P}_\ell := \mathscr{P}_\ell(V,V), \qquad \mathscr{P} := \mathscr{P}(V,V). \tag{4.78}$$

When Λ or Λ' consists of a single point, write x instead of $\{x\}$. For $\pi \in \mathscr{P}_\ell$, set $|\pi| := \ell$. Write $\operatorname{supp}(\pi) := \{\pi_0,\dots,\pi_{|\pi|}\}$ to denote the set of points visited by π.

Let $X = (X_t)_{t\geq 0}$ be the continuous-time random walk on G that jumps from $x \in V$ to any neighbour $y \sim x$ at rate 1. Denote by $(T_k)_{k\in\mathbb{N}_0}$ the sequence of jump times (with $T_0 := 0$). For $\ell \in \mathbb{N}_0$, let

$$\pi^\ell(X) := (X_0,\dots,X_{T_\ell}) \tag{4.79}$$

be the path in $\mathscr{P}_\ell$ consisting of the first ℓ steps of X. For $t \geq 0$, let

$$\pi(X_{[0,t]}) = \pi^{\ell_t}(X), \qquad \text{with } \ell_t \in \mathbb{N}_0 \text{ satisfying } T_{\ell_t} \leq t < T_{\ell_t+1}, \tag{4.80}$$

denote the path in $\mathscr{P}$ consisting of all the steps taken by X between times 0 and t.

Recall the definitions from Section 4.3.1. For $\pi \in \mathscr{P}$ and $A > 0$, define

$$\lambda_{r,A}(\pi) := \sup\left\{\lambda^1_{\mathcal{C}}(\xi; G)\colon\ \mathcal{C} \in \mathfrak{C}_{r,A},\ \mathrm{supp}(\pi) \cap \mathcal{C} \cap \Pi_{r,A} \neq \emptyset\right\}, \tag{4.81}$$

with the convention $\sup\emptyset = -\infty$. This is the largest principal eigenvalue among the components of $\mathfrak{C}_{r,A}$ in $\mathcal{GW}$ that have a point of high exceedance visited by the path π.

Lemma 4.4.1. *[Mass up to an exit time] Subject to Assumption* 4.C*, $\mathfrak{P}$-a.s. for any $r \geq r_0$, $y \in \Lambda \subset B_r(\mathcal{O})$, $\xi \in [0,\infty)^V$ and $\gamma > \lambda_\Lambda = \lambda_\Lambda(\xi, \mathcal{GW})$,*

$$\mathbb{E}_y\left[\mathrm{e}^{\int_0^{\tau_{\Lambda^c}}(\xi(X_s)-\gamma)\,\mathrm{d}s}\right] \leq 1 + \frac{(\log r)^{\delta r}\,|\Lambda|}{\gamma - \lambda_\Lambda}. \tag{4.82}$$

Proof. The proof is identical to that of Lemma 4.3.2, with δr replaced by $(\log r)^{\delta r}$ (recall Lemma 4.2.3). □

§4.4.1 Mass of the solution along excursions

Lemma 4.4.2. *[Path evaluation] For $\ell \in \mathbb{N}_0$, $\pi \in \mathscr{P}_\ell$ and $\gamma > \max_{0 \leq i < |\pi|}\{\xi(\pi_i) - D_{\pi_i}\}$,*

$$\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^{T_\ell}(\xi(X_s)-\gamma)\,\mathrm{d}s}\ \middle|\ \pi^\ell(X) = \pi\right] = \prod_{i=0}^{\ell-1}\frac{D_{\pi_i}}{\gamma - [\xi(\pi_i) - D_{\pi_i}]}. \tag{4.83}$$

Proof. The proof is identical to that of [12, Lemma 3.2]. The left-hand side of (4.83) can be evaluated by using the fact that T_ℓ is the sum of ℓ independent $\mathrm{Exp}(\deg(\pi_i))$ random variables that are independent of $\pi^\ell(X)$. The condition on γ ensures that all ℓ integrals are finite. □

For a path $\pi \in \mathscr{P}$ and $\varepsilon \in (0,1)$, we write

$$M^{r,\varepsilon}_\pi := \left|\left\{0 \leq i < |\pi|\colon\ \xi(\pi_i) \leq (1-\varepsilon)a_{L_r}\right\}\right|, \tag{4.84}$$

with the interpretation that $M^{r,\varepsilon}_\pi = 0$ if $|\pi| = 0$.

Lemma 4.4.3. *[Mass of excursions] Subject to Assumptions* 4.A–4.C*, for every $A, \varepsilon > 0$, $(\mathsf{P} \times \mathfrak{P})$-a.s. there exists an $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$, all $\gamma > a_{L_r} - A$ and all $\pi \in \mathscr{P}(B_r(\mathcal{O}), B_r(\mathcal{O}))$ satisfying $\pi_i \notin \Pi_{r,A}$ for all $0 \leq i < \ell := |\pi|$,*

$$\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^{T_\ell}(\xi(X_s)-\gamma)\,\mathrm{d}s}\ \middle|\ \pi^\ell(X) = \pi\right] \leq q^\ell_{r,A}\,\mathrm{e}^{M^{r,\varepsilon}_\pi \log[(\log r)^{\delta r}/a_{L_r,A,\varepsilon}q_{r,A}]}, \tag{4.85}$$

Chapter 4

where

$$a_{L_r,A,\varepsilon} := \varepsilon a_{L_r} - A, \qquad q_{r,A} := \left(1 + \frac{A}{(\log r)^{\delta_r}}\right)^{-1}. \tag{4.86}$$

Note that $\pi_\ell \in \Pi_{r,A}$ *is allowed.*

Proof. The proof is identical to that of [12, Lemma 3.3], with $d_{\max}$ replaced by $(\log r)^{\delta_r}$ (recall Lemma 4.2.3). □

We follow [12, Definition 3.4] and [9, Section 6.2]. Note that the distance between $\Pi_{r,A}$ and $D^{\mathrm{c}}_{r,A}$ in $\mathcal{GW}$ is at least $S_r = (\log L_r)^\alpha$ (recall (4.43)–(4.44)).

Definition 4.4.4. [Concatenation of paths] (a) When π and π' are two paths in $\mathscr{P}$ with $\pi_{|\pi|} = \pi'_0$, we define their *concatenation* as

$$\pi \circ \pi' := (\pi_0, \dots, \pi_{|\pi|}, \pi'_1, \dots, \pi'_{|\pi'|}) \in \mathscr{P}. \tag{4.87}$$

Note that $|\pi \circ \pi'| = |\pi| + |\pi'|$.

(b) When $\pi_{|\pi|} \neq \pi'_0$, we can still define the *shifted concatenation* of π and π' as $\pi \circ \hat{\pi}'$, where $\hat{\pi}' := (\pi_{|\pi|}, \pi_{|\pi|} + \pi'_1 - \pi'_0, \dots, \pi_{|\pi|} + \pi'_{|\pi'|} - \pi'_0)$. The shifted concatenation of multiple paths is defined inductively via associativity. ■

Now, if a path $\pi \in \mathscr{P}$ intersects $\Pi_{r,A}$, then it can be decomposed into an initial path, a sequence of excursions between $\Pi_{r,A}$ and $D^{\mathrm{c}}_{r,A}$, and a terminal path. More precisely, there exists $m_\pi \in \mathbb{N}$ such that

$$\pi = \check{\pi}^1 \circ \hat{\pi}^1 \circ \dots \circ \check{\pi}^{m_\pi} \circ \hat{\pi}^{m_\pi} \circ \bar{\pi}, \tag{4.88}$$

where the paths in (4.88) satisfy

$$\begin{array}{llll} \check{\pi}^1 \in \mathscr{P}(V, \Pi_{r,A}) & \text{with} & \check{\pi}^1_i \notin \Pi_{r,A}, & 0 \le i < |\check{\pi}^1|, \\ \hat{\pi}^k \in \mathscr{P}(\Pi_{r,A}, D^{\mathrm{c}}_{r,A}) & \text{with} & \hat{\pi}^k_i \in D_{r,A}, & 0 \le i < |\hat{\pi}^k|,\ 1 \le k \le m_\pi - 1, \\ \check{\pi}^k \in \mathscr{P}(D^{\mathrm{c}}_{r,A}, \Pi_{r,A}) & \text{with} & \check{\pi}^k_i \notin \Pi_{r,A}, & 0 \le i < |\check{\pi}^k|,\ 2 \le k \le m_\pi, \\ \hat{\pi}^{m_\pi} \in \mathscr{P}(\Pi_{r,A}, V) & \text{with} & \hat{\pi}^{m_\pi}_i \in D_{r,A}, & 0 \le i < |\hat{\pi}^{m_\pi}|, \end{array} \tag{4.89}$$

while

$$\begin{array}{ll} \bar{\pi} \in \mathscr{P}(D^{\mathrm{c}}_{r,A}, V) \text{ and } \bar{\pi}_i \notin \Pi_{r,A}\ \forall i \ge 0 & \text{if } \hat{\pi}^{m_\pi} \in \mathscr{P}(\Pi_{r,A}, D^{\mathrm{c}}_{r,A}), \\ \bar{\pi}_0 \in D_{r,A}, |\bar{\pi}| = 0 & \text{otherwise.} \end{array} \tag{4.90}$$

Note that the decomposition in (4.88)–(4.90) is unique, and that the paths $\check{\pi}^1$, $\hat{\pi}^{m_\pi}$ and $\bar{\pi}$ can have zero length. If π is contained in $B_r(\mathcal{O})$, then so are all the paths in the decomposition.

Whenever $\mathrm{supp}(\pi) \cap \Pi_{r,A} \neq \emptyset$ and $\varepsilon > 0$, we define

$$s_\pi := \sum_{i=1}^{m_\pi} |\check{\pi}^i| + |\bar{\pi}|, \qquad k^{r,\varepsilon}_\pi := \sum_{i=1}^{m_\pi} M^{r,\varepsilon}_{\check{\pi}^i} + M^{r,\varepsilon}_{\bar{\pi}} \tag{4.91}$$

Chapter 4

to be the total time spent in exterior excursions, respectively, on moderately low points of the potential visited by exterior excursions (without their last point).

In case $\mathrm{supp}(\pi) \cap \Pi_{r,A} = \emptyset$, we set $m_\pi := 0$, $s_\pi := |\pi|$ and $k_\pi^{r,\varepsilon} := M_\pi^{r,\varepsilon}$. Recall from (4.81) that, in this case, $\lambda_{r,A}(\pi) = -\infty$.

We say that $\pi, \pi' \in \mathscr{P}$ are *equivalent*, written $\pi' \sim \pi$, if $m_\pi = m_{\pi'}$, $\check{\pi}'^i = \check{\pi}^i$ for all $i = 1, \dots, m_\pi$, and $\bar{\pi}' = \bar{\pi}$. If $\pi' \sim \pi$, then $s_{\pi'}$, $k_{\pi'}^{r,\varepsilon}$ and $\lambda_{r,A}(\pi')$ are all equal to the counterparts for π.

To state our key lemma, we define, for $m, s \in \mathbb{N}_0$,

$$\mathscr{P}^{(m,s)} = \{\pi \in \mathscr{P} \colon\ m_\pi = m, s_\pi = s\}, \tag{4.92}$$

and denote by

$$C_{r,A} := \max\{|\mathcal{C}| \colon\ \mathcal{C} \in \mathfrak{C}_{r,A}\} \tag{4.93}$$

the maximal size of the islands in $\mathfrak{C}_{r,A}$.

Lemma 4.4.5. *[Mass of an equivalence class] Subject to Assumptions* 4.A *and* 4.C*, for every $A, \varepsilon > 0$, $(\mathsf{P} \times \mathfrak{P})$-a.s. there exists an $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$, all $m, s \in \mathbb{N}_0$, all $\pi \in \mathscr{P}^{(m,s)}$ with* $\mathrm{supp}(\pi) \subset B_r(\mathcal{O})$*, all $\gamma > \lambda_{r,A}(\pi) \vee (a_{L_r} - A)$ and all $t \geq 0$,*

$$\begin{aligned}&\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^t (\xi(X_u)-\gamma)\,\mathrm{d}u}\,\mathbb{1}_{\{\pi(X_{[0,t]})\sim\pi\}}\right]\\ &\leq \left(C_{r,A}^{1/2}\right)^{\mathbb{1}_{\{m>0\}}}\left(1+\frac{(\log r)^{\delta_r}\,C_{r,A}}{\gamma-\lambda_{r,A}(\pi)}\right)^m\left(\frac{q_{r,A}}{d_{\min}}\right)^s \mathrm{e}^{k_\pi^{r,\varepsilon}\log[(\log r)^{\delta_r}/a_{L_r,A,\varepsilon}q_{r,A}]}.\end{aligned} \tag{4.94}$$

Proof. The proof is identical to that of [12, Lemma 3.5], with $d_{\max}$ is replaced by $(\log r)^{\delta_r}$ (recall Lemma 4.2.3). □

CHAPTER 4

§4.4.2 Key proposition

The main result of this section is the following proposition.

Proposition 4.4.6. *[Entropy reduction] Let $\alpha \in (0,1)$ and $\kappa \in (\alpha, 1)$. Subject to Assumption* 4.C*, there exists an $A_0(r)$ such that, for all $A \geq A_0(r)$, with $\mathfrak{P}$-probability tending to one as $r \to \infty$, the following statement is true. For each $x \in B_r(\mathcal{O})$, each $\mathcal{N} \subset \mathscr{P}(x, B_r(\mathcal{O}))$ satisfying* $\mathrm{supp}(\pi) \subset B_r(\mathcal{O})$ *and* $\max_{1\leq\ell\leq|\pi|}\mathrm{dist}_G(\pi_\ell, x) \geq (\log L_r)^\kappa$ *for all $\pi \in \mathcal{N}$, and each assignment $\pi \mapsto (\gamma_\pi, z_\pi) \in \mathbb{R} \times V$ satisfying*

$$\gamma_\pi \geq \left(\lambda_{r,A}(\pi) + \mathrm{e}^{-S_r}\right) \vee (a_{L_r} - A) \qquad \forall\, \pi \in \mathcal{N} \tag{4.95}$$

and

$$z_\pi \in \mathrm{supp}(\pi) \cup \bigcup_{\substack{\mathcal{C}\in\mathfrak{C}_{r,A}\colon\\ \mathrm{supp}(\pi)\cap\mathcal{C}\cap\Pi_{r,A}\neq\emptyset}} \mathcal{C} \qquad \forall\, \pi \in \mathcal{N}, \tag{4.96}$$

the following inequality holds for all $t \geq 0$:

$$\log\mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\,\mathbb{1}_{\{\pi(X_{[0,t]})\in\mathcal{N}\}}\right] \leq \sup_{\pi\in\mathcal{N}}\left\{t\gamma_\pi + \mathrm{dist}_G(x, z_\pi)\log[(\log r)^{\delta_r}/a_{L_r,A,\varepsilon}q_{r,A}]\right\}. \tag{4.97}$$

Proof. The proof is based on [12, Section 3.4]. First fix $c_0 > 2$ and define

$$A_0(r) = (\log r)^{\delta_r} \left(\mathrm{e}^{4c_0(\log r)^{1-\alpha}} - 1 \right). \tag{4.98}$$

Fix $A \geq A_0(r)$, $\beta \in (0, \alpha)$ and $\varepsilon \in (0, \frac{1}{2}\beta)$ as in Lemma 4.3.6. Let $r_0 \in \mathbb{N}$ be as given in Lemma 4.4.5, and take $r \geq r_0$ so large that the conclusions of Lemmas 4.2.3, 4.3.1, 4.3.3 and 4.3.6 hold, i.e., assume that the events $\mathcal{B}_r$ and $\mathcal{B}_{r,A}$ in these lemmas do not occur. Fix $x \in B_r(\mathcal{O})$. Recall the definitions of $C_{r,A}$ and $\mathscr{P}^{(m,s)}$. Note that the relation $\sim$ is an equivalence relation in $\mathscr{P}^{(m,s)}$, and define

$$\widetilde{\mathscr{P}}^{(m,s)}_x := \big\{\text{equivalence classes of the paths in } \mathscr{P}(x, V) \cap \mathscr{P}^{(m,s)}\big\}. \tag{4.99}$$

The following bounded on the cardinality of this set is needed.

Lemma 4.4.7. *[Bound equivalence classes] Subject to Assumption* 4.C, $\mathfrak{P}$*-a.s.,*$|\widetilde{\mathscr{P}}^{(m,s)}_x| \leq (2C_{r,A})^m (\log r)^{\delta_r (m+s)}$ *for all* $m, s \in \mathbb{N}_0$.

Proof. We can copy the proof of [12, Lemma 3.6], replacing $d_{\max}$ by $(\log r)^{\delta_r}$.

The estimate is clear when $m = 0$. To prove that it holds for $m \geq 1$, write $\partial\Lambda := \{z \notin \Lambda\colon \operatorname{dist}_G(z, \Lambda) = 1\}$ for $\Lambda \subset V$. Then $|\partial\mathcal{C} \cup \mathcal{C}| \leq ((\log r)^{\delta_r} + 1)|\mathcal{C}| \leq 2(\log r)^{\delta_r} C_{r,A}$ by Lemma 4.2.3. Define the map $\Phi\colon \widetilde{\mathscr{P}}^{(m,s)}_x \to \mathscr{P}_s(x, V) \times \{1, \ldots, 2(\log r)^{\delta_r} C_{r,A}\}^m$ as follows. For each $\Lambda \subset V$ with $1 \leq |\Lambda| \leq 2(\log r)^{\delta_r} C_{r,A}$, fix an injection $f_\Lambda\colon \Lambda \to \{1, \ldots, 2(\log r)^{\delta_r} C_{r,A}\}$. Given a path $\pi \in \mathscr{P}^{(m,s)} \cap \mathscr{P}(x, V)$, decompose π, and denote by $\widetilde{\pi} \in \mathscr{P}_s(x, V)$ the shifted concatenation of $\check{\pi}^1, \ldots, \check{\pi}^m, \bar{\pi}$. Note that, for $2 \leq k \leq m$, the point $\check{\pi}^k_0$ lies in $\partial\mathcal{C}_k$ for some $\mathcal{C}_k \in \mathfrak{C}_{r,A}$, while $\bar{\pi}_0 \in \partial\overline{\mathcal{C}} \cup \overline{\mathcal{C}}$ for some $\overline{\mathcal{C}} \in \mathfrak{C}_{r,A}$. Thus, it is possible to set

$$\Phi(\pi) := \big(\widetilde{\pi}, f_{\partial\mathcal{C}_2}(\check{\pi}^2_0), \ldots, f_{\partial\mathcal{C}_m}(\check{\pi}^m_0), f_{\partial\bar{\mathcal{C}} \cup \bar{\mathcal{C}}}(\bar{\pi}_0)\big). \tag{4.100}$$

It is readily checked that $\Phi(\pi)$ depends only on the equivalence class of π and, when restricted to equivalence classes, Φ is injective. Hence the claim follows. □

Now take $\mathcal{N} \subset \mathscr{P}(x, V)$ as in the statement, and set

$$\widetilde{\mathcal{N}}^{(m,s)} := \big\{\text{equivalence classes of paths in } \mathcal{N} \cap \mathscr{P}^{(m,s)}\big\} \subset \widetilde{\mathscr{P}}^{(m,s)}_x. \tag{4.101}$$

For each $\mathcal{M} \in \widetilde{\mathcal{N}}^{(m,s)}$, choose a representative $\pi_{\mathcal{M}} \in \mathcal{M}$, and use Lemma 4.4.7 to write

$$\begin{aligned} \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u} \mathbb{1}_{\{\pi(X_{[0,t]}) \in \mathcal{N}\}} \right] &= \sum_{m,s \in \mathbb{N}_0} \sum_{\mathcal{M} \in \widetilde{\mathcal{N}}^{(m,s)}} \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u} \mathbb{1}_{\{\pi(X_{[0,t]}) \sim \pi_{\mathcal{M}}\}} \right] \\ &\leq \sum_{m,s \in \mathbb{N}_0} (2(\log r)^{\delta_r} C_{r,A})^m ((\log r)^{\delta_r})^s \sup_{\pi \in \mathcal{N}^{(m,s)}} \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u} \mathbb{1}_{\{\pi(X_{[0,t]}) \sim \pi\}} \right] \end{aligned} \tag{4.102}$$

with the convention $\sup \emptyset = 0$. For fixed $\pi \in \mathcal{N}^{(m,s)}$, by (4.95), apply (4.94) and Lemma 4.3.1 to obtain, for all r large enough and with $c_0 > 2$,

$$\begin{aligned}(2(c\log r)^{\delta_r})^m (\log r)^{\delta_r s}\, \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u} \mathbb{1}_{\{\pi(X_{[0,t]})\sim \pi\}} \right] \\ \le \mathrm{e}^{t\gamma_\pi} \mathrm{e}^{c_0 m \log r} [q_{r,A}(\log r)^{\delta_r}]^s \, \mathrm{e}^{k_\pi^{r,\varepsilon} \log[(\log r)^{\delta_r}/a_{L_r,A,\varepsilon} q_{r,A}]}.\end{aligned} \tag{4.103}$$

We next claim that, for r large enough and $\pi \in \mathcal{N}^{(m,s)}$,

$$s \ge [(m-1) \vee 1]\, S_r. \tag{4.104}$$

Indeed, when $m \ge 2$, $|\operatorname{supp}(\breve{\pi}^i)| \ge S_r$ for all $2 \le i \le m$. When $m = 0$, $|\operatorname{supp}(\pi)| \ge \max_{1\le \ell \le |\pi|} |\pi_\ell - x| \ge (\log L_r)^\kappa \gg S_r$ by assumption. When $m = 1$, the latter assumption and Lemma 4.3.1 together imply that $\operatorname{supp}(\pi) \cap D^{\mathrm{c}}_{r,A} \neq \emptyset$, and so either $|\operatorname{supp}(\breve{\pi}^1)| \ge S_r$ or $|\operatorname{supp}(\bar{\pi})| \ge S_r$. Thus, (4.104) holds by the definition of S_r and s.

Note that $q_{r,A}^{S_r} < \mathrm{e}^{-4c_0 \log r}$, so

$$\begin{aligned}&\sum_{m\ge 0} \sum_{s \ge [(m-1)\vee 1] S_r} \mathrm{e}^{c_0 m \log r} [q_{r,A}(\log r)^{\delta_r}]^s \\ &= \frac{[q_{r,A}(\log r)^{\delta_r}]^{S_r} + \mathrm{e}^{c_0 \log r}[q_{r,A}(\log r)^{\delta_r}]^{S_r} + \sum_{m\ge 2} \mathrm{e}^{m c_0 \log r}[q_{r,A}(\log r)^{\delta_r}]^{(m-1)S_r}}{1 - q_{r,A}(\log r)^{\delta_r}} \\ &\le \frac{3\mathrm{e}^{-c_0 \log r}}{1 - q_{r,A}(\log r)^{\delta_r}} < 1\end{aligned} \tag{4.105}$$

for r large enough. Inserting this back into (4.102), we obtain

$$\log \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\pi(X_{0,t}) \in \mathcal{N}\}} \right] \le \sup_{\pi \in \mathcal{N}} \left\{ t\gamma_\pi + k_\pi^{r,\varepsilon} \log[(\log r)^{\delta_r}/a_{L_r,A,\varepsilon} q_{r,A}] \right\}. \tag{4.106}$$

Thus the proof will be finished once we show that, for some $\varepsilon' > 0$ and whp, respectively, a.s. eventually as $r \to \infty$,

$$k_\pi^{r,\varepsilon} \ge \operatorname{dist}_G(x, z_\pi)(1 - 2(\log L_r)^{-\varepsilon'}) \qquad \forall\, \pi \in \mathcal{N}. \tag{4.107}$$

We can copy the argument at the end of [12, Section 3.4]. For each $\pi \in \mathcal{N}$ define an auxiliary path $\pi_\star$ as follows. First note that by using our assumptions we can find points $z', z'' \in \operatorname{supp}(\pi)$ (not necessarily distinct) such that

$$\operatorname{dist}_G(x, z') \ge (\log L_r)^\kappa, \qquad \operatorname{dist}_G(z'', z_\pi) \le 2 M_A S_r, \tag{4.108}$$

where the latter holds by (4.51). Write $\{z_1, z_2\} = \{z', z''\}$ with z_1, z_2 ordered according to their hitting times by π, i.e., $\inf\{\ell \colon \pi_\ell = z_1\} \le \inf\{\ell \colon \pi_\ell = z_2\}$. Define π_e as the concatenation of the loop erasure of π between x and z_1 and the loop erasure of π between z_1 and z_2. Since π_e is the concatenation of two self-avoiding paths, it visits each point at most twice. Finally, define $\pi_\star \sim \pi_e$ by replacing the excursions of π_e from $\Pi_{r,A}$ to $D^{\mathrm{c}}_{r,A}$ by direct paths between the corresponding endpoints, i.e., replace each $\hat{\pi}^i_e$ by $|\hat{\pi}^i_e| = \ell_i$, $(\hat{\pi}^i_e)_0 = x_i \in \Pi_{r,A}$, and $(\hat{\pi}^i_e)_{\ell_i} = y_i \in D^{\mathrm{c}}_{r,A}$ by a shortest-distance path $\widetilde{\pi}^i_\star$ with the same endpoints and $|\widetilde{\pi}^i_\star| = \mathrm{dist}_G(x_i, y_i)$. Since $\pi_\star$ visits each $x \in \Pi_{r,A}$ at most 2 times,

$$k^{r,\varepsilon}_{\pi} \geq k^{r,\varepsilon}_{\pi_\star} \geq M^{r,\varepsilon}_{\pi_\star} - 2|\,\mathrm{supp}(\pi_\star) \cap \Pi_{r,A}|(S_r + 1) \geq M^{r,\varepsilon}_{\pi_\star} - 4|\,\mathrm{supp}(\pi_\star) \cap \Pi_{r,A}|S_r. \tag{4.109}$$

Note that $M^{r,\varepsilon}_{\pi_\star} \geq |\{x \in \mathrm{supp}(\pi_\star) \colon\ \xi(x) \leq (1-\varepsilon)a_{L_r}\}| - 1$ and, by (4.108), $|\,\mathrm{supp}(\pi_\star)| \geq \mathrm{dist}_G(x, z') \geq (\log L_r)^\kappa \gg (\log L_r)^{\alpha+2\varepsilon'}$ for some $0 < \varepsilon' < \varepsilon$. Applying Lemmas 4.3.6–4.3.7 and using (4.40) and $L_r > r$, we obtain, for r large enough,

$$k^{r,\varepsilon}_{\pi} \geq |\,\mathrm{supp}(\pi_\star)| \left(1 - \frac{2}{(\log L_r)^\varepsilon} - \frac{4S_r}{(\log L_r)^{\alpha+2\varepsilon'}}\right) \geq |\,\mathrm{supp}(\pi_\star)| \left(1 - \frac{1}{(\log L_r)^{\varepsilon'}}\right). \tag{4.110}$$

On the other hand, since $|\,\mathrm{supp}(\pi_\star)| \geq (\log L_r)^\kappa$, by (4.108) we have

$$\begin{aligned} |\mathrm{supp}(\pi_\star)| &= \big(|\mathrm{supp}(\pi_\star)| + 2M_A S_r\big) - 2M_A S_r \\ &= \big(|\mathrm{supp}(\pi_\star)| + 2M_A S_r\big)\left(1 - \frac{2M_A S_r}{|\mathrm{supp}(\pi_\star)| + 2M_A S_r}\right) \\ &\geq (\mathrm{dist}_G(x, z'') + 2M_A S_r)\left(1 - \frac{2M_A S_r}{(\log L_r)^\kappa}\right) \\ &\geq \mathrm{dist}_G(x, z_\pi)\left(1 - \frac{1}{(\log L_r)^{\varepsilon'}}\right), \end{aligned} \tag{4.111}$$

where the first inequality uses that the distance between two points on $\pi_\star$ is less than the total length of $\pi_\star$. Now (4.107) follows from (4.110)–(4.111). □

§4.5 Proof of the main theorem

Define

$$U^*(t) := \mathrm{e}^{t[\varrho \log(\vartheta \mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho)]}, \tag{4.112}$$

where we recall (4.10). To prove Theorem 4.1.2 we show that

$$\frac{1}{t}\log U(t) - \frac{1}{t}\log U^*(t) = o(1), \quad t \to \infty, \qquad (\mathsf{P} \times \mathfrak{P})\text{-a.s.} \tag{4.113}$$

The proof proceeds via upper and lower bound, proved in Sections 4.5.1 and 4.5.2, respectively. Throughout this section, Assumptions 4.A, 4.B(1) and 4.C are in force.

§4.5.1 Upper bound

We follow [12, Section 4.2]. The proof of the upper bound in (4.113) relies on two lemmas showing that paths staying inside a ball of radius $\lceil t^\gamma \rceil$ for some $\gamma \in (0,1)$ or leaving a ball of radius $t \log t$ have a negligible contribution to (4.4), the total mass of the solution.

Lemma 4.5.1. *[No long paths] For any $\ell_t \geq t \log t$,*

$$\lim_{t\to\infty} \frac{1}{U^*(t)} \mathbb{E}_{\mathcal{O}} \left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\ell_t}]^{\mathrm{c}}}<t\}} \right] = 0 \quad (\mathsf{P} \times \mathfrak{P}) - a.s. \tag{4.114}$$

Proof. We follow [12, Lemma 4.2]. For $r \geq \ell_t$, let

$$\mathcal{B}_r := \left\{ \max_{x\in B_r(\mathcal{O})} \xi(x) \geq a_{L_r} + 2\varrho \right\}. \tag{4.115}$$

Since $\lim_{t\to\infty} \ell_t = \infty$, Lemma 4.3.5 gives that P-a.s.

$$\bigcup_{r\geq\ell_t} \mathcal{B}_r \text{ does not occur eventually as } t \to \infty. \tag{4.116}$$

Therefore we can work on the event $\bigcap_{r\geq\ell_t} [\mathcal{B}_r]^{\mathrm{c}}$. On this event, we write

$$\begin{aligned} \mathbb{E}_{\mathcal{O}} \left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\ell_t}]^{\mathrm{c}}}<t\}} \right] &= \sum_{r\geq\ell_t} \mathbb{E}_{\mathcal{O}} \left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\sup_{s\in[0,t]} |X_s|=r\}} \right] \\ &\leq \mathrm{e}^{2\varrho t} \sum_{r\geq\ell_t} \mathrm{e}^{\varrho t \log r + \log(\delta_r \log\log r)} \,\mathbb{P}_{\mathcal{O}} \left(J_t \geq r \right), \end{aligned} \tag{4.117}$$

where J_t is the number of jumps of X up to time t, and we use that $|B_r(\mathcal{O})| \leq (\log r)^{\delta_r r}$. Next, J_t is stochastically dominated by a Poisson random variable with parameter $t(\log r)^{\delta_r}$. Hence

$$\mathbb{P}_{\mathcal{O}} \left(J_t \geq r \right) \leq \frac{[\mathrm{e}t \,(\log r)^{\delta_r}]^r}{r^r} \leq \exp\left\{ -r \log \left(\frac{r}{\mathrm{e}t \,(\log r)^{\delta_r}} \right) \right\} \tag{4.118}$$

for large r. Using that $\ell_t \geq t \log t$, we can easily check that, for $r \geq \ell_t$ and t large enough,

$$\varrho t \log r - r \log \left(\frac{r}{\mathrm{e}t \,(\log r)^{\delta_r}} \right) < -3r, \qquad r \geq \ell_t. \tag{4.119}$$

Thus (4.117) is at most

$$\mathrm{e}^{2\varrho t} \sum_{r\geq\ell_t} \mathrm{e}^{-3r+\log(\delta_r \log\log r)} \leq \mathrm{e}^{2\varrho t} \sum_{r\geq\ell_t} \mathrm{e}^{-2r} \leq 2\,\mathrm{e}^{2\varrho t}\,\mathrm{e}^{-2\ell_t} \leq \mathrm{e}^{-\ell_t}. \tag{4.120}$$

Since $\lim_{t\to\infty} \ell_t = \infty$ and $\lim_{t\to\infty} U^*(t) = \infty$, this settles the claim. □

Lemma 4.5.2. *[No short paths] For any $\gamma \in (0, 1)$,*

$$\lim_{t\to\infty} \frac{1}{U^*(t)} \mathbb{E}_{\mathcal{O}} \left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\lceil t^\gamma\rceil}]^{\mathrm{c}}}>t\}} \right] = 0 \quad (\mathsf{P} \times \mathfrak{P}) - a.s. \tag{4.121}$$

Proof. We follow [12, Lemma 4.3]. By Lemma 4.3.5 with $r = \lceil t^\gamma \rceil$, we may assume that

$$\max_{x\in B_{\lceil t^\gamma\rceil}} \xi(x) \leq \varrho \log\log L_{\lceil t^\gamma\rceil} + \frac{2\varrho \log\lceil t^\gamma\rceil}{\vartheta\lceil t^\gamma\rceil} \leq \gamma\varrho \log t + O(1), \quad t \to \infty, \tag{4.122}$$

where the second inequality uses that $\log L_{\lceil t^\gamma\rceil} \sim \log |B_{\lceil t^\gamma\rceil}(\mathcal{O})| \sim \vartheta \lceil t^\gamma\rceil$. Hence

$$\frac{1}{U^*(t)}\,\mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\tau_{[B_{\lceil t^\gamma\rceil}]^{\mathrm{c}}}>t\}}\right] \le \frac{1}{U^*(t)}\,\mathrm{e}^{\gamma\varrho t\log t+O(1)} \le \mathrm{e}^{(1-\gamma)\varrho t\log t + C\log\log\log t}, \quad t\to\infty, \tag{4.123}$$

for any constant $C>1$. □

The proof of the upper bound in (4.113) also relies on a third lemma estimating the contribution of paths leaving a ball of radius $\lceil t^\gamma\rceil$ for some $\gamma\in(0,1)$ but staying inside a ball of radius $t\log t$. We slice the annulus between these two balls into layers, and derive an estimate for paths that reach a given layer but do not reach the next layer. To that end, fix $\gamma\in(\alpha,1)$ with α as in (4.40), and let

$$K_t := \lceil t^{1-\gamma}\log t\rceil, \qquad r_t^{(k)} := k\lceil t^\gamma\rceil, \quad 1\le k\le K_t, \qquad \ell_t := K_t\lceil t^\gamma\rceil \ge t\log t. \tag{4.124}$$

For $1\le k\le K_t$, define (recall (4.77))

$$\mathcal{N}_t^k := \left\{\pi\in\mathscr{P}(\mathcal{O},V)\colon\ \mathrm{supp}(\pi)\subset B_{r_t^{k+1}}(\mathcal{O}),\, \mathrm{supp}(\pi)\cap B^{\mathrm{c}}_{r_t^k}(\mathcal{O})\neq\emptyset\right\} \tag{4.125}$$

and set

$$U^k(t) := \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\pi_{[0,t]}(X)\in\mathcal{N}_t^k\}}\right]. \tag{4.126}$$

Lemma 4.5.3. *[Upper bound on $U^k(t)$] For any $\varepsilon>0$, $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $t\to\infty$,*

$$\sup_{1\le k\le K_t}\frac{1}{t}\log U_t^k \le \frac{1}{t}\log U^*(t)+\varepsilon. \tag{4.127}$$

Proof. We follow [12, Lemma 4.4] Fix $k\in\{1,\dots,K_t\}$. For $\pi\in\mathcal{N}_t^k$, let

$$\gamma_\pi := \lambda_{r_t^{k+1},A}(\pi)+\mathrm{e}^{-S_{\lceil t^\gamma\rceil}}, \qquad z_\pi\in\mathrm{supp}(\pi),\, |z_\pi|>r_t^k, \tag{4.128}$$

chosen such that (4.95)–(4.96) are satisfied. By Proposition 4.4.6 and (4.86), $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $t\to\infty$,

$$\frac{1}{t}\log U_t^k \le \gamma_\pi - \frac{|z_\pi|}{t}\Big(\log[\varepsilon\varrho\log(\vartheta r_t^{(k+1)})] - \delta_r\log[\log(r_t^{(k+1)})]+o(1)\Big). \tag{4.129}$$

Using Corollary 4.3.4 and $\log L_r\sim\vartheta r$, we bound

$$\gamma_\pi \le \varrho\log(\vartheta r_t^{(k+1)}) - \widetilde{\chi}(\varrho)+\tfrac{1}{2}\varepsilon+o(1). \tag{4.130}$$

Moreover, $|z_\pi|>r_t^{k+1}-\lceil t^\gamma\rceil$ and

$$\begin{aligned}&\frac{\lceil t^\gamma\rceil}{t}\Big(\log[\varepsilon\varrho\log(\vartheta r_t^{(k+1)})] - \delta_r\log[\log(r_t^{(k+1)})]\Big)\\ &\quad\le \frac{1}{t^{1-\gamma}}\log\log(2t\log t) = o(1).\end{aligned} \tag{4.131}$$

Hence

$$\gamma_\pi \le F_t(r_t^{(k+1)}) - \widetilde{\chi}(\varrho) + \tfrac{1}{2}\varepsilon + o(1) \tag{4.132}$$

with

$$F_t(r) := \varrho \log(\vartheta r) - \frac{r}{t}\big[\log(\varepsilon\varrho\log(\vartheta r)) - \delta_r \log(\log r)\big], \qquad r > 0. \tag{4.133}$$

The function F_t is maximized at any point r_t satisfying

$$\varrho t = r_t \left[\log(\varepsilon\varrho\log(\vartheta r_t)) - (\delta_r + r_t \tfrac{\mathrm{d}}{\mathrm{d}r}\delta_r)\log\log r_t + \frac{1}{\log(\vartheta r_t)} - \frac{\delta_r}{\log r_t}\right]. \tag{4.134}$$

In particular, $r_t = \mathfrak{r}_t[1+o(1)]$, which implies that

$$\sup_{r>0} F_t(r) \le \varrho\log(\vartheta\mathfrak{r}_t) - \varrho + o(1), \qquad t \to \infty. \tag{4.135}$$

Inserting (4.135) into (4.132), we obtain $\frac{1}{t}\log U_t^k < \varrho\log(\vartheta\mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) + \varepsilon$, which is the desired upper bound because $\varepsilon > 0$ is arbitrary. □

Proof of the upper bound in (4.113). To avoid repetition, all statements hold $(\mathfrak{P}\times\mathsf{P})$-a.s. eventually as $t \to \infty$. Set

$$U^0(t) := \mathbb{E}_\mathcal{O}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\tau_{[B_{\lceil t\gamma\rceil}]^c} > t\}}\right], \quad U^\infty(t) := \mathbb{E}_\mathcal{O}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\tau_{[B_{\lceil t\log t\rceil}]^c} \le t\}}\right]. \tag{4.136}$$

Then

$$U(t) \le U^0(t) + U^\infty(t) + K_t \max_{1\le k\le K_t} U^k(t). \tag{4.137}$$

From Lemmas 4.5.1–4.5.3 and the fact that $K_t = o(t)$, we get

$$\limsup_{t\to\infty}\left\{\frac{1}{t}\log U(t) - \frac{1}{t}\log U^*(t)\right\} \le \varepsilon. \tag{4.138}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the upper bound in (4.11). □

§4.5.2 Lower bound

We first introduce an alternative representation for χ in (2.11) in terms of a 'dual' variational formula. Fix $\varrho \in (0,\infty)$ and a graph $G = (V,E)$. The functional

$$\mathcal{L}(q;G) := \sum_{x\in V} \mathrm{e}^{q(x)/\varrho} \in [0,\infty], \qquad q\colon V \to [-\infty,\infty), \tag{4.139}$$

plays the role of a large deviation rate function for the potential ξ in V (compare with (2.7)). For $\Lambda \subset V$, define

$$\widehat{\chi}_\Lambda(G) := -\sup_{\substack{q\colon V\to[-\infty,\infty),\\ \mathcal{L}(q;G)\le 1}} \lambda_\Lambda(q;G) \in [0,\infty). \tag{4.140}$$

The condition $\mathcal{L}(q;G) \le 1$ under the supremum ensures that the potentials q have a fair probability under the i.i.d. double-exponential distribution. Write $\widehat{\chi}(G) = \widehat{\chi}_V(G)$.

Proposition 4.5.4. *[Alternative representations for χ] For any graph $G = (V, E)$ and any $\Lambda \subset V$,*

$$\widehat{\chi}_\Lambda(\varrho; G) \geq \widehat{\chi}_V(\varrho; G) = \widehat{\chi}_G(\varrho) = \chi_G(\varrho). \tag{4.141}$$

Proof. See [12, Section A.1] □

For the lower bound we follow [12, Section 4.1]. Fix $\varepsilon > 0$. By the definition of $\widetilde{\chi}$, there exists an infinite rooted tree $T = (V', E', \mathcal{Y})$ with degrees in $\operatorname{supp}(D_g)$ such that $\chi_T(\varrho) < \widetilde{\chi}(\varrho) + \frac{1}{4}\varepsilon$. Let $Q_r = B_r^T(\mathcal{Y})$ be the ball of radius r around $\mathcal{Y}$ in T. By Proposition 4.5.4 and (4.140), there exist a radius $R \in \mathbb{N}$ and a potential profile $q\colon B_R^T \to \mathbb{R}$ with $\mathcal{L}_{Q_R}(q; \varrho) < 1$ (in particular, $q \leq 0$) such that

$$\lambda_{Q_R}(q; T) \geq -\widehat{\chi}_{Q_R}(\varrho; T) - \tfrac{1}{2}\varepsilon > -\widetilde{\chi}(\varrho) - \varepsilon. \tag{4.142}$$

For $\ell \in \mathbb{N}$, let $B_\ell = B_\ell(\mathcal{O})$ denote the ball of radius ℓ around $\mathcal{O}$ in $\mathcal{GW}$. We will show next that, $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $\ell \to \infty$, B_ℓ contains a copy of the ball Q_R where the potentail ξ is bounded from below by $\varrho \log\log |B_\ell| + q$.

Proposition 4.5.5. *[Balls with high exceedances] $(\mathfrak{P} \times \mathsf{P})$-almost surely eventually as $\ell \to \infty$, there exists a vertex $z \in B_\ell$ with $B_{R+1}(z) \subset B_\ell$ and an isomorphism $\varphi : B_{R+1}(z) \to Q_{R+1}$ such that $\xi \geq \varrho \log\log |B_\ell| + q \circ \varphi$ in $B_R(z)$. In particular,*

$$\lambda_{B_R(z)}(\xi; \mathcal{GW}) > \varrho \log\log |B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon. \tag{4.143}$$

Any such z necessarily satisfies $|z| \geq c\ell$ $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $\ell \to \infty$ for some constant $c = c(\varrho, \vartheta, \widetilde{\chi}(\varrho), \varepsilon) > 0$.

Proof. See [12, Proposition 4.1]. The proof carries over verbatim because the degrees play no role. □

Proof of the lower bound in (4.11). Let z be as in Proposition 4.5.5. Write τ_z for the hitting time of z by the random walk X. For $s \in (0, t)$, we estimate

$$\begin{aligned} U(t) &\geq \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^t \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{\tau_z \leq s\}}\, \mathbb{1}_{\{X_u \in B_R(z)\,\forall u \in [\tau_z, t]\}}\Big] \\ &= \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^{\tau_z} \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{\tau_z \leq s\}}\, \mathbb{E}_z\Big[\mathrm{e}^{\int_0^v \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{X_u \in B_R(z)\,\forall u \in [0,v]\}}\Big]\Big|_{v = t - \tau_z}\Big], \end{aligned} \tag{4.144}$$

where we use the strong Markov property at time τ_z. We first bound the last term in the integrand in (4.144). Since $\xi \geq \varrho \log\log |B_\ell| + q$ in $B_R(z)$,

$$\begin{aligned} \mathbb{E}_z\Big[\mathrm{e}^{\int_0^v \xi(X_u)\,\mathrm{d}u} \mathbb{1}_{\{X_u \in B_R(z)\,\forall u \in [0,v]\}}\Big] &\geq \mathrm{e}^{v\varrho \log\log |B_\ell|} \mathbb{E}_{\mathcal{Y}}\Big[\mathrm{e}^{\int_0^v q(X_u)\,\mathrm{d}u} \mathbb{1}_{\{X_u \in Q_R\,\forall u \in [0,v]\}}\Big] \\ &\geq \mathrm{e}^{v\varrho \log\log |B_\ell|} \mathrm{e}^{v\lambda_{Q_R}(q;T)} \phi^1_{Q_R}(\mathcal{Y})^2 \\ &> \exp\big\{v\,(\varrho \log\log |B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon)\,\big\} \end{aligned} \tag{4.145}$$

for large v, where we used that $B_{R+1}(z)$ is isomorphic to Q_{R+1} for the indicators in the first inequality, and applied Lemma B.1.2 and (4.142) to obtain the second and third inequalities, respectively. On the other hand, since $\xi \geq 0$,

$$\mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^{\tau_z} \xi(X_u)\,\mathrm{d}u}\mathbb{1}\{\tau_z \leq s\}\Big] \geq \mathbb{P}_{\mathcal{O}}(\tau_z \leq s), \tag{4.146}$$

and we can bound the latter probability from below by the probability that the random walk runs along a shortest path from the root $\mathcal{O}$ to z within a time at most s. Such a path $(y_i)_{i=0}^{|z|}$ has $y_0 = \mathcal{O}$, $y_{|z|} = z$, $y_i \sim y_{i-1}$ for $i = 1, \dots, |z|$, has at each step from y_i precisely $\deg(y_i)$ choices for the next step with equal probability, and the step is carried out after an exponential time E_i with parameter $\deg(y_i)$. This gives

$$\mathsf{P}_{\mathcal{O}}(\tau_z \leq s) \geq \Big(\prod_{i=1}^{|z|} \frac{1}{\deg(y_i)}\Big) P\Big(\sum_{i=1}^{|z|} E_i \leq s\Big) \geq ((\log|z|)^{\delta_\ell})^{-|z|}\mathrm{Poi}_{d_{\min}s}([|z|,\infty)), \tag{4.147}$$

where Poi_γ is the Poisson distribution with parameter γ, and P is the generic symbol for probability. Summarising, we obtain

$$\begin{aligned} U(t) &\geq ((\log|z|)^{\delta_l})^{-|z|}\mathrm{e}^{-d_{\min}s}\frac{(d_{\min}s)^{|z|}}{|z|!}\mathrm{e}^{(t-s)\left[\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon\right]} \\ &\geq \exp\left\{-d_{\min}s + (t-s)\left[\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon\right] - |z|\log\left(\frac{(\log|z|)^{\delta_\ell}}{d_{\min}}\frac{|z|}{s}\right)\right\} \\ &\geq \exp\left\{-d_{\min}s + (t-s)\left[\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon\right] - \ell\log\left(\frac{(\log\ell)^{\delta_\ell}}{d_{\min}}\frac{\ell}{s}\right)\right\}, \end{aligned} \tag{4.148}$$

where in the last inequality we use that $s \leq |z|$ and $\ell \geq |z|$. Further assuming that $\ell = o(t)$, we see that the optimum over s is obtained at

$$s = \frac{\ell}{d_{\min} + \varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon} = o(t). \tag{4.149}$$

Note that, by Proposition 4.5.5, this s indeed satisfies $s \leq |z|$. Applying (4.7) we get, after a straightforward computation, $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $t \to \infty$,

$$\frac{1}{t}\log U(t) \geq \varrho\log\log|B_\ell| - \frac{\ell}{t}\log\log\ell - \frac{\ell}{t}\delta_\ell\log\log\ell - \widetilde{\chi}(\varrho) - \varepsilon + O\left(\frac{\ell}{t}\right). \tag{4.150}$$

Inserting $\log|B_\ell| \sim \vartheta\ell$, we get

$$\frac{1}{t}\log U(t) \geq F_\ell - \widetilde{\chi}(\varrho) - \varepsilon + o(1) + O\left(\frac{\ell}{t}\right) \tag{4.151}$$

with

$$F_\ell = \varrho\log(\vartheta\ell) - \frac{\ell}{t}\log\log\ell - \frac{\ell}{t}\delta_\ell\log\log\ell. \tag{4.152}$$

The optimal ℓ for F_ℓ satisfies

$$\varrho t = \ell\big[1 + (\delta_\ell + \ell\tfrac{\mathrm{d}}{\mathrm{d}\ell}\delta_\ell)\big]\log\log\ell + \frac{\ell\delta_\ell}{\log\ell} + \frac{\ell}{\log\ell}, \tag{4.153}$$

i.e., $\ell = \mathfrak{r}_t[1+o(1)]$. For this choice we obtain

$$\frac{1}{t}\log U(t) \geq \varrho\log(\vartheta\mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) - \varepsilon + o(1). \tag{4.154}$$

Hence $(\mathfrak{P}\times \mathsf{P})$-a.s.

$$\liminf_{t\to\infty}\left\{\frac{1}{t}\log U(t) - \frac{1}{t}\log U^*(t)\right\} \geq -\varepsilon. \tag{4.155}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the lower bound in (4.11). □

Remark: It is clear from (4.134) and (4.153) that, in order to get the correct asymptotics, it is crucial that both δ_r and $r\frac{\mathrm{d}}{\mathrm{d}r}\delta_r$ tend to zero as $r\to\infty$. This is why Assumption 4.C is the weakest condition on the tail of the degree distribution under which the arguments in [12] can be pushed through.

§4.6 Existence and uniqueness of the Feynman-Kac formula

We follow the argument in [24, Section 2], where existence and uniqueness of the Feynman-Kac formula in (4.3) was shown for $G=\mathbb{Z}^d$.

Theorem 4.6.1. *Subject to Assumptions 4.A and 4.B,* (4.2) *has a unique nonnegative solution* $(P\times\mathfrak{P})$*-almost surely. This solution admits the Feynman-Kac representation in* (4.3)*.*

We note that, due to the exponential growth of the Galton-Watson tree, the condition on the potential needed here is stronger than the one required in [24] on $\mathbb{Z}^d$.

The proof of Theorem 4.6.1 requires several preparatory results. Lemmas 4.6.2 and 4.6.3 below show the existence and uniqueness, respectively, of the Feynman-Kac solution for a deterministic potential. Lemma 4.6.4 extends this to a random potential.

Consider the problem

$$\begin{array}{rcll} \partial_t u(x,t) &=& (\Delta_G u)(x,t) + q(x)u(x,t), & x\in V,\ t>0,\\ u(x,0) &=& \delta_{\mathcal{O}}(x), & x\in V, \end{array} \tag{4.156}$$

where q is a deterministic potential that is bounded from below. Without loss of generality, we may assume that q is nonnegative.

Define

$$v(x,t) = \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t q(X_s)\mathrm{d}s}\,\mathbb{1}\{X_t = x\}\right]. \tag{4.157}$$

Lemma 4.6.2. *[Existence]* (4.156) *admits at least one nonnegative solution if and only if*

$$v(x,t) < \infty \qquad \forall\ (t,x) \in \mathbb{R}_+ \times V. \tag{4.158}$$

If (4.158) *is fulfilled, then v is the minimal nonnegative solution of* (4.156).

Proof. See [12, Lemma 2.2]. The proof relies on restricting the Feynman-Kac functional in (4.158) to cubes of length $2N$ around the origin and letting $N \to \infty$. On the tree we restrict to balls of radius R around the root and let $R \to \infty$. The arguments carry over with this change. □

Lemma 4.6.3. *[Uniqueness] If q is bounded from below, then* (4.156) *admits at most one nonnegative solution $\mathfrak{P}$-almost surely.*

Proof. It suffices to show that (4.156) with initial condition $u(x,0) = 0$, $x \in V$, only has the 0 solution. We follow the proof of [24, Lemma 2.3]. For $R \in \mathbb{N}$, define Γ_R to be the set of paths

$$\gamma : \mathcal{O} = x_0 \to x_1 \to \cdots \to x_n \tag{4.159}$$

consisting of neighbouring vertices in V such that $x_0, \cdots, x_{n-1} \in B_R(\mathcal{O})$ and $x_n \in Z_{R+1}$. Furthermore, define

$$\gamma_+ = \{x \in \gamma : q(x) > 0\}. \tag{4.160}$$

Let τ_R be the first time when the random walk hits Z_{R+1}, and let v be a solution of (4.156). The Feynman-Kac representation of v reads

$$v(t,0) = \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^{\tau_R} q(X_s)\mathrm{d}s} v\Big(t-\tau_R, X(\tau_R)\Big) \mathbb{1}\{\tau_R \le t\}\right]. \tag{4.161}$$

We are done once we show that

$$v(T,0) \ge \left(\frac{T}{t}\right)^{R-1} \mathrm{e}^{-(\log R)(T-t)} v(t,0) \tag{4.162}$$

for all $0 < t \le T$ and all $R \in \mathbb{N}$. Indeed, in that case the right-hand side tends to infinity as $R \to \infty$, and therefore so does the left-hand side, which leaves $v(t,0) = 0$ for all $t \in (0,T]$ as the only possible solution.

To prove (4.162), fix an arbitrary path $\gamma \in \Gamma_R$. The contribution of the random walk moving along the path γ equals $\chi_\gamma(t)$ with

$$\chi_\gamma(t) = \left(\prod_{i=0}^{n-1} \frac{1}{D_i}\right) \mathbb{E}_{\mathcal{O}}\left(\exp\left\{\sum_{i=0}^{n-1} q_i\sigma_i\right\} v\left(t - \sum_{i=0}^{n-1}\sigma_i, x_n\right) \mathbb{1}\left\{\sum_{i=0}^{n-1}\sigma_i \le t\right\}\right), \tag{4.163}$$

where $q_i = q(x_i)$ and the σ_i are the successive waiting times of the random walk, which are independent and exponentially distributed with parameter D_i. Letting m be such

Chapter 4

that $q(x_m) = \min_{x\in\gamma} q(x)$, we can rewrite (4.163) as

$$\begin{aligned}\chi_\gamma(T) &= \int\cdots\int_{\sum_{i=0}^{n-1} s_i \le T} \sum_{i=0}^{n-1} \mathrm{d}s_i \exp\left\{\sum_{i=0}^{n-1} s_i(q_m - D_i)\right\} v\left(T - \sum_{i=0}^{n-1} s_i, x_n\right) \exp\left\{\sum_{i=0}^{n-1} s_i(q_i - q_m)\right\} \\ &\ge \int\cdots\int_{\sum_{i=0}^{n-1} s_i \le T} \sum_{i=0}^{n-1} \mathrm{d}s_i \exp\left\{\sum_{i=0}^{n-1} s_i(q_m - \log R)\right\} v\left(T - \sum_{i=0}^{n-1} s_i, x_n\right) \exp\left\{\sum_{i=0}^{n-1} s_i(q_i - q_m)\right\},\end{aligned} \tag{4.164}$$

where the inequality uses Lemma 4.2.3(a), which shows that the maximal degree is $o(R)$ as $R \to \infty$. After some straightforward manipulations and making a change of integration variables (for full details see [24, (2.14)–(2.15)]), we arrive at (4.162). □

Lemma 4.6.4. *For each $t > 0$,*

$$\mathbb{P}_\mathcal{O}\left(\max_{s\in[0,t]} |X_s| = R\right) \le \mathrm{e}^{-[1+o(1)]R\log R}, \qquad R \to \infty, \quad \mathfrak{P} - a.s., \tag{4.165}$$

where, for $x \in V$, $|x| = \mathrm{dist}(\mathcal{O}, x)$ denotes the distance between x and the root $\mathcal{O}$.

Proof. For fixed R, let $(\tilde{X}_t)_{t\ge0}$ be the random walk on the regular tree with offspring D_R such that $D_x \le D_R$ for all $x \in B_R(\mathcal{O})$. From Lemma 4.2.3, $D_R = o(R)$. Define $(N(t))_{t\ge0}$ to be the Poisson process with rate D_R, associated with the jumps of $\tilde{X}$. We estimate

$$\mathbb{P}_\mathcal{O}\left(\max_{s\in[0,t]} |X_s| = R\right) \le \mathbb{P}_\mathcal{O}\left(\max_{s\in[0,t]} |\tilde{X}_s| = R\right) \le \mathbb{P}_\mathcal{O}(N(t) = R).$$

Since

$$\mathbb{P}_\mathcal{O}(N(t) = R) = \frac{(D_R)^R}{R!}\mathrm{e}^{-tD_R},$$

(4.165) follows from Stirling's formula. □

Proof of Theorem 4.6.1. We follow the proof of [24, Theorem 2.1 a)]. We need to check that the expression in (4.3) is finite for arbitrary $(t,x) \in \mathbb{R}_+ \times V$. To that end we estimate

$$\mathbb{E}_\mathcal{O}\left[\mathrm{e}^{\int_0^t q(X_s)\mathrm{d}s} \mathbb{1}\{X_t = x\}\right] \le \sum_{R\in\mathbb{N}} \mathbb{P}_\mathcal{O}\left(\max_{s\in[0,t]} |X_s| = R\right) \exp\left\{t \max_{y\in B_R(\mathcal{O})} \xi(y)\right\}.$$

We know from Lemma 4.3.5 that $\max_{y\in B_R(\mathcal{O})} \xi(y) \sim \varrho\log(\theta R)$ as $R \to \infty$, $(P \times \mathfrak{P})$-a.s. (recall (4.7) and (4.42)). Applying Lemma 4.6.4, we see that the sum on the right-hand side is indeed finite. □

Chapter 4

CHAPTER 5

The parabolic Anderson model on a Galton-Watson tree with normalised Laplacian

This chapter is based on the following paper:
D. Wang. The parabolic Anderson model on a Galton-Watson tree with normalised Laplacian, 2023. Preprint, arXiv:2310.05602.

Abstract

In [12], the asymptotics of the total mass of the solution to the parabolic Anderson model was studied on an almost surely infinite Galton-Watson tree with an i.i.d. potential having a double-exponential distribution. The second-order contribution to this asymptotics was identified in terms of a variational formula that gives information about the local structure of the region where the solution is concentrated.

The present paper extends this work to the degree-normalised Laplacian. The normalisation causes the Laplacian to be non-symmetric, which leads to different spectral properties. We find that the leading order asymptotics of the total mass remains the same, while the second-order correction coming from the variational formula is different. We also find that the optimiser of the variational formula is again an infinite tree with minimal degrees. Both of these results are shown to hold under much milder conditions than for the regular Laplacian.

§5.1 Introduction and main results

§5.1.1 The PAM and intermittency

The *parabolic Anderson model* (PAM) on a graph $G = (V, E)$, is the Cauchy problem for the heat equation with a random potential:

$$\begin{aligned} \partial_t u(x,t) &= (\Delta u)(x,t) + \xi(x)u(x,t), \quad && x \in V,\, t > 0, \\ u(x,0) &= \delta_{\mathcal{O}}(x), && x \in V, \end{aligned} \tag{5.1}$$

where $\mathcal{O}$ is a vertex in V, $\{\xi(x)\}_{x\in V}$ is the random potential defined on V, and Δ is the normalised discrete Laplacian, defined by

$$(\Delta f)(x) := \frac{1}{\deg(x)} \sum_{\substack{y\in V:\\ \{x,y\}\in E}} [f(y) - f(x)], \qquad x \in V,\, f\colon V \to \mathbb{R}. \tag{5.2}$$

Most of the literature has focused on $\mathbb{Z}^d$ and we refer the reader to [29] for a comprehensive study. Other choices include the complete graph [16], the hypercube [5], and more recently, the regular tree [14] and the Galton-Watson tree [12], [13] and [2]. The PAM can also be studied on continuous spaces, the most extensively studied being $\mathbb{R}^d$. On such spaces, the Laplacian and the potential are defined analogously and we again refer the reader to [29] for more background.

The PAM may be interpreted as a system of particles such that particles are killed with rate $\xi^-(x)$ or are split into two with rate $\xi^+(x)$ at every vertex x. At the same time, each particle jumps independently with Δ as generator. The solution $u(x,t)$ can be interpreted as the expected number of particles or *mass* present at vertex x at time t when the initial condition at time 0 is $\delta_{\mathcal{O}}(x)$. See [24, Section 1.2] for further details.

The solution is known to exhibit a phenomenon known as *intermittency*, meaning that the solution concentrates on small regions of the graph known as *intermittent islands*. This is well studied on $\mathbb{Z}^d$, where it is known that the sizes of the island(s) depend on the distribution of the tail of the potential, and can be separated into four classes (see [29, Section 3.4]). Throughout the paper, we work in the *double-exponential* class where the potential $\xi = \{\xi(x)\}_{x\in V}$ consists of i.i.d. random variables satisfying

Assumption 5.A. [Asymptotic double-exponential potential]
For some $\varrho \in (0, \infty)$,

$$\mathsf{P}(\xi(0) \geq 0) = 1, \qquad \mathsf{P}(\xi(0) > u) = \mathrm{e}^{-\mathrm{e}^{u/\varrho}} \text{ for } u \text{ large enough.} \tag{5.3}$$

■

The main feature of this choice is that the intermittent islands are not single vertices, whilst also having the property that their sizes do not change with time - a critical fact in our analysis.

§5.1.2 The PAM on a Galton-Watson tree

The present paper analyses the PAM on the graph generated by the Galton-Watson process. The graph is generated by taking a root $\mathcal{O}$ and attaching D vertices (known as offspring), where D is a random variable. Each offspring has $D-1$ offspring attached to it, where D is an identically distributed but independent copy of D. This is repeated forever, or until the process dies out. Let $\mathcal{GW} = (V, E, \mathcal{O})$ be the resulting graph and let $\mathfrak{P}$ and $\mathfrak{E}$ denote probability and expectation with respect to $\mathcal{GW}$. Similarly, let $\mathcal{P}$ and $\mathcal{E}$ denote probability and expectation with respect to D.

Assumption 5.B. [Exponential tails]
(1) $d_{\min} := \min \operatorname{supp}(D) \geq 2$ and $\mathcal{E}[D] \in (2, \infty)$.
(2) $\mathcal{E}\left[e^{aD}\right] < \infty$ for all $a \in (0, \infty)$. ■

Under this assumption, $\mathcal{GW}$ is $\mathfrak{P}$-a.s. an infinite tree. Moreover,

$$\lim_{r\to\infty} \frac{\log |B_r(\mathcal{O})|}{r} = \log \mathcal{E}[D] =: \vartheta \in (0, \infty) \qquad \mathfrak{P} - a.s., \tag{5.4}$$

where $B_r(\mathcal{O}) \subset V$ is the ball of radius r around $\mathcal{O}$ in the graph distance (see e.g. [30, pp. 134–135]). Note that this ball depends on $\mathcal{GW}$ and is therefore random.

The proper choice of Laplacian depends on the setting. In the case of the complete graph and the hypercube, when the limit of the number of vertices going to infinity is taken, only the normalised Laplacian gives a meaningful limit. In the case of regular graphs such as $\mathbb{Z}^d$ and the regular tree, normalising the Laplacian by the degree simply amounts to rescaling time, and both the techniques and the results can be easily inferred accordingly. We will focus on the Galton-Watson tree, which is not only inhomogeneous but is also random, and hence the choice of Laplacian does play a role. The present paper considers the normalised Laplacian, in contrast to [12], [13], and [2], and investigate how this choice affects the results and methods used in [12].

Under Assumptions 5.A and 5.B, the criteria for existence and uniqueness of a non-negative solution of (5.1) are met (see [24] and [13, Appendix C]) and is given by the well-known Feynman-Kac representation. With the choice of initial condition in (5.1) and Laplacian in (5.2), this amounts to

$$u(x,t) = \mathbb{E}_{\mathcal{O}}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{X_t = x\}\right], \tag{5.5}$$

where $X = (X_t)_{t\geq 0}$ is the continuous-time random walk on the vertices V with jump rate 1 on each vertex (or equivalently with jump rate equal to the inverse of the degree along the edges E), and $\mathbb{P}_{\mathcal{O}}$ denotes the law of X given $X_0 = \mathcal{O}$. The quantity we will be interested in is the *total mass*, given by

$$U(t) := \sum_{x\in V} u(x,t) = \mathbb{E}_{\mathcal{O}}\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\right], \tag{5.6}$$

in particular, its asymptotics as $t \to \infty$.

An important distinction is made between the *quenched* and the *annealed* total mass, i.e. the total mass taken almost surely with respect to or averaged over the sources of randomness, respectively. Since this paper aims to follow the framework of [12], we also consider the quenched setting for both the graph and potential. We refer to [14] for corresponding results for the annealed total mass on a regular tree. The annealed setting for the Galton-Watson tree (averaged over just the potential or over both the graph and potential) remains open.

§5.1.3 Main results and discussion

To state our results, we first introduce some quantities of interest pertaining to the *characteristic variational formula* associated with Assumption 5.A. The latter describes the shape and profile of the solution in the intermittent islands and captures the second-order asymptotics of the total mass. We refer to [29] for more details on the variational formula and its relationship with the PAM.

Denote the set of probability measures on V by $\mathcal{P}(V)$. For $p \in \mathcal{P}(V)$, define

$$I_E(p) := \sum_{\{x,y\}\in E} \left(\sqrt{\frac{p(x)}{\deg(x)}} - \sqrt{\frac{p(y)}{\deg(y)}} \right)^2, \qquad J_V(p) := -\sum_{x\in V} p(x)\log p(x), \tag{5.7}$$

and set

$$\chi_G(\varrho) := \inf_{p\in\mathcal{P}(V)} [I_E(p) + \varrho J_V(p)], \qquad \varrho \in (0,\infty). \tag{5.8}$$

The first term arises from the Laplacian and coincides with the large deviation rate function for the empirical distribution of the random walk in (5.5) (see [11, Exercise IV.24], while the second term comes from the choice of the double-exponential potential. Furthermore, define the constant

$$\widetilde{\chi}(\varrho) := \inf\big\{\chi_T(\varrho) \colon\ T \text{ is an infinite tree with degrees in } \operatorname{supp}(D)\big\}, \tag{5.9}$$

with $\chi_G(\varrho)$ defined in (5.8), and abbreviate

$$\mathfrak{r}_t = \frac{\varrho t}{\log\log t}. \tag{5.10}$$

Theorem 5.1.1. *[Total mass asymptotics] Subject to Assumptions* 5.A–5.B*,*

$$\frac{1}{t}\log U(t) = \varrho\log(\vartheta\mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) + o(1), \quad t\to\infty, \qquad (\mathsf{P}\times\mathfrak{P})\text{-}a.s. \tag{5.11}$$

The proof of Theorem 5.1.1 is given in Section 5.4. For a heuristic explanation on how the terms in (5.11) arise and how they relate to the asymptotics of the total mass, we refer the reader to [12, Section 1.5].

For $d \geq 2$, let $\mathcal{T}_d$ denote the *infinite homogeneous tree* with degree equal to d at every vertex.

Theorem 5.1.2. *[Identification of the minimiser] If $\varrho \geq \frac{1}{d_{\min}\log(d_{\min}+1)}$, then $\widetilde{\chi}(\varrho) = \chi_{\mathcal{T}_{d_{\min}}}(\varrho)$.*

The proof of Theorem 5.1.2 is given in Section 5.5.2.

Comparing our results to those obtained in [12] and [13], we see that the choice of Laplacian indeed has an effect, albeit in a subtle way. The leading-order terms in (5.11) remain unchanged, while the second-order (variational formula) term stemming from (5.8) is different, due to I_E in (5.7) being normalised by the degrees. This normalisation was not present in [12] and [13]. In addition, normalising the Laplacian results in a 'slow down' of the random walk in (5.5) compared with the analogous formula in [12]. As will be shown later on, this leads to simplifications in several key lemmas and leads to Theorem 5.1.1 holding under the milder tail condition in Assumption 5.B(2).

The different Laplacian and I_E function have surprisingly minimal effects on Theorem 5.1.2: the optimal tree is still $\mathcal{T}_{d_{\min}}$, exactly as was found in [12]. The main difference is that our result holds for a greater range of ϱ values compared to [12], which required the sharper restriction $\varrho \geq 1/\log(d_{\min}+1)$. We believe that the minimal tree is the minimiser for all ϱ and that it is also the unique minimiser, however this remains open. It is also worth noting that the object $\widetilde{\chi}(\varrho)$ is well understood. The case $d_{\min} = 2$ corresponds to the the variational problem on $\mathbb{Z}$ and has been studied in [23]. For $d_{\min} > 2$ we refer the reader to [14], where the variational formula

$$\bar{\chi}_{d_{\min}}(\varrho) = \inf_{p \in \mathcal{P}(V)} [\bar{I}_E(p) + \varrho J_V(p)], \quad \bar{I}_E(p) = \sum_{\{x,y\} \in E} \left(\sqrt{p(x)} - \sqrt{p(y)}\right)^2,$$

was studied. Clearly,

$$\widetilde{\chi}(\varrho) = \tfrac{1}{d_{\min}} \bar{\chi}_{d_{\min}}(d_{\min}\varrho).$$

Outline. The remainder of the paper is dedicated to the proof of Theorems 5.1.1 and 5.1.2, and follows the framework developed in [12]. Section 2.1 is novel and deals with the spectral estimates of the Anderson Hamiltonian $\Delta + \xi$, which are different due to Δ no longer being symmetric with respect to the usual inner product. Sections 2.2 and 2.3 collect the necessary results regarding the Galton-Watson tree and the potential from [12] and [13]. All of these results carry over directly since the Laplacian and random walk play no role. Section 3 follows the *path expansion* technique from [12] and adapts the results to the random walk in (5.5). Section 4 is dedicated to the proof of Theorem 5.1.1, and follows [12]. Section 5 deals with the analysis of the variational formula (5.9) including the proof of Theorem 5.1.2, which applies the gluing argument from [12].

§5.2 Preliminaries

In this section we collect results that are needed later. Section 5.2.1 investigates how the normalisation affects the spectral properties of the Laplacian. Section 5.2.2 collects two vital facts about the Galton-Watson tree. Section 5.2.3 collects results regarding the potential.

§5.2.1 Related Spectral Problems

Recall the Rayleigh-Ritz formula for the principal eigenvalue $\lambda_\Lambda(q;G)$,

$$\lambda_\Lambda(q;G) = \sup\big\{\langle(\Delta+q)\phi,\phi\rangle\colon\ \phi\in\mathbb{R}^V,\ \operatorname{supp}\phi\subset\Lambda,\ \|\phi\|=1\big\}. \tag{5.12}$$

As alluded to in Section 5.1.3, Δ (and therefore also $\Delta+q$) is not symmetric with respect to the usual ℓ^2 inner product, but is symmetric with respect to the degree-weighted inner product

$$\langle\phi,\psi\rangle := \sum_{x\in\Lambda}\deg(x)\phi(x)\psi(x), \tag{5.13}$$

in the sense that $\langle\Delta\phi,\psi\rangle = \langle\phi,\Delta\psi\rangle$ (see [27, Section 2] for further details). Henceforward, all inner products will be with respect to (5.13).

We introduce an alternative representation for χ in (5.8) in terms of a 'dual' variational formula. Fix $\varrho\in(0,\infty)$ and a graph $G=(V,E)$. The functional

$$\mathcal{L}(q;G) := \sum_{x\in V}\mathrm{e}^{q(x)/\varrho}\in[0,\infty], \qquad q\colon\ V\to[-\infty,\infty), \tag{5.14}$$

plays the role of a large deviation rate function for the potential ξ in V (compare with (5.3)). For $\Lambda\subset V$, define

$$\widehat{\chi}_\Lambda(G) := -\sup_{\substack{q\colon V\to[-\infty,\infty),\\ \mathcal{L}(q;G)\le 1}}\lambda_\Lambda(q;G)\in[0,\infty), \tag{5.15}$$

where $\lambda_\Lambda(q;G)$ is the principal eigenvalue of the Anderson Hamiltonian $\Delta+q$ on the set Λ with zero boundary condition. The condition $\mathcal{L}(q;G)\le 1$ under the supremum ensures that the potentials q have a fair probability under the i.i.d. double-exponential distribution.

CHAPTER 5

Proposition 5.2.1. *[Alternative representations for χ] For any graph $G=(V,E)$ and any $\Lambda\subset V$,*

$$\widehat{\chi}_\Lambda(\varrho;G)\ge\widehat{\chi}_V(\varrho;G)=\chi_G(\varrho). \tag{5.16}$$

Proposition 5.2.1 is not essential in the proof of Theorem 5.1.1, but is stated here to provide additional context to some of the results below. The proof is given in Section 5.5.1.

Lemma 5.2.2. *[Spectral bounds]*

(1) *For any $\Gamma\subset\Lambda\subset V$,*

$$\max_{z\in\Gamma}q(z)-1\le\lambda_\Gamma(q;G)\le\lambda_\Lambda(q;G)\le\max_{z\in\Lambda}q(z). \tag{5.17}$$

(2) *The eigenfunction corresponding to $\lambda_\Lambda(q;G)$ can be taken to be non-negative.*

(3) *If q is real-valued and $\Gamma \subsetneq \Lambda$ is finite and connected in G, then the second inequality in* (5.17) *is strict and the eigenfunction corresponding to $\lambda_\Lambda(q;G)$ is strictly positive.*

Proof. Write

$$\langle(\Delta+q)\phi,\phi\rangle = -\sum_{\{x,y\}\in E_\Lambda}[\phi(x)-\phi(y)]^2 + \sum_{x\in\Lambda}\deg(x)q(x)\phi(x)^2. \tag{5.18}$$

The upper bound in (5.17) follows from the estimate

$$\langle(\Delta+q)\phi,\phi\rangle \le \sum_{x\in\Lambda}\deg(x)q(x)\phi(x)^2 \le \max_{z\in\Lambda} q(z)\sum_{x\in\Lambda}\deg(x)\phi(x)^2 = \max_{z\in\Lambda} q(z).$$

To get the lower bound in (5.17), we use the fact that λ_Λ is non-decreasing in q. Let $\bar{z} = \arg\max q(z)$. Replacing $q(z)$ by $-\infty$ for every $z \neq \bar{z}$ and taking the test function $\bar{\phi} = \frac{1}{\sqrt{\deg(\bar{z})}}\delta_{\bar{z}}$, we get that

$$\begin{aligned}\lambda_\Lambda(q;G) &\ge -\sum_{\substack{x,y\in\Lambda:\\ \{x,y\}\in E_\Lambda}}\left[\bar{\phi}(x)-\bar{\phi}(y)\right]^2 + \sum_{x\in\Lambda}\deg(x)q(x)\bar{\phi}(x)^2\\ &= -\sum_{\substack{y\in\Lambda:\\ \{\bar{z},y\}\in E_\Lambda}}\frac{1}{\deg(\bar{z})} + q(\bar{z}) = -1 + \max_{z\in\Lambda} q(z),\end{aligned} \tag{5.19}$$

which settles the claim in (1). The claims in (2) and (3) are standard. □

Inside $\mathcal{GW}$, fix a finite connected subset $\Lambda \subset V$, and let H_Λ denote the Anderson Hamiltonian in Λ with zero Dirichlet boundary conditions on $\Lambda^c = V\backslash\Lambda$ (i.e. the restriction of the operator $H_G = \Delta + \xi$ to the class of functions supported on Λ). For $y \in \Lambda$, let u_Λ^y be the solution of

$$\begin{array}{rcll}\partial_t u(x,t) &=& (H_\Lambda u)(x,t), & x\in\Lambda,\ t>0,\\ u(x,0) &=& \delta_y(x), & x\in\Lambda,\end{array} \tag{5.20}$$

and set $U_\Lambda^y(t) := \sum_{x\in\Lambda} u_\Lambda^y(x,t)$. Let τ_{Λ^c} be the hitting time of Λ^c and

$$u_\Lambda^y(x,t) = \mathbb{E}_y\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\}\mathbb{1}\{\tau_{\Lambda^c} > t, X_t = x\}\right], \tag{5.21}$$

the Feynman-Kac solution to (5.1) with Dirichlet boundary conditions on Λ^c. Then $u_\Lambda^y(x,t)$ also admits the spectral representation

$$u_\Lambda^y(x,t) = \sum_{k=1}^{|\Lambda|} \mathrm{e}^{t\lambda_\Lambda^k}\phi_\Lambda^k(y)\phi_\Lambda^k(x), \tag{5.22}$$

where $\lambda_\Lambda^1 \ge \lambda_\Lambda^2 \ge \cdots \ge \lambda_\Lambda^{|\Lambda|}$ and $\phi_\Lambda^1, \phi_\Lambda^2, \ldots, \phi_\Lambda^{|\Lambda|}$ are, respectively the eigenvalues and the corresponding orthonormal eigenfunctions of $\Delta + \xi$ restricted to Λ. These two representations may be exploited to obtain bounds for one in terms of the other, as shown by the following lemma.

Lemma 5.2.3. *[Bounds on the solution] For any $y \in \Lambda$ and any $t > 0$,*

$$\mathrm{e}^{t\lambda_\Lambda^1}\phi_\Lambda^1(y)^2 \le \mathbb{E}_y\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\tau_{\Lambda^c}>t, X_t=y\}}\right] \le \mathbb{E}_y\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\mathbb{1}_{\{\tau_{\Lambda^c}>t\}}\right] \tag{5.23}$$

Proof. The first inequality follows from a suitable application of Parseval's identity. The second inequality is elementary. □

Lemma 5.2.4. *[Mass up to an exit time] For any $y \in \Lambda$, $\xi \in [0,\infty)^V$ and $\gamma > \lambda_\Lambda = \lambda_\Lambda(\xi, \mathcal{GW})$,*

$$\mathbb{E}_y\left[\mathrm{e}^{\int_0^{\tau_{\Lambda^c}}(\xi(X_s)-\gamma)\,\mathrm{d}s}\right] \le 1 + \frac{|\Lambda|}{\gamma - \lambda_\Lambda}. \tag{5.24}$$

Proof. We follow the proof of [13, Lemma 3.2] and [25, Lemma 2.18]. Define

$$u(x) := \mathbb{E}_x\left[\mathrm{e}^{\int_0^{\tau_{\Lambda^c}}(\xi(X_s)-\gamma)\,\mathrm{d}s}\right]. \tag{5.25}$$

This is the solution to the boundary value problem

$$\begin{aligned}(\Delta + \xi - \gamma)u &= 0 \quad \text{on } \Lambda,\\ u &= 1 \quad \text{on } \Lambda^{\mathrm{c}}.\end{aligned} \tag{5.26}$$

Via the substitution $u =: 1 + v$, this turns into

$$\begin{aligned}(\Delta + \xi - \gamma)v &= \gamma - \xi \quad \text{on } \Lambda,\\ v &= 0 \qquad\quad \text{on } \Lambda^{\mathrm{c}}.\end{aligned} \tag{5.27}$$

It is readily checked that for $\gamma > \lambda_\Lambda$ the solution exists and is given by

$$v = \mathcal{R}_\gamma(\xi - \gamma), \tag{5.28}$$

where $\mathcal{R}_\gamma$ denotes the resolvent of $\Delta + \xi$. Hence

$$v(x) \le (\mathcal{R}_\gamma \mathbb{1})(x) \le \langle \mathcal{R}_\gamma \mathbb{1}, \mathbb{1}\rangle \le \frac{|\Lambda|}{\gamma - \lambda_\Lambda}, \quad x \in \Lambda, \tag{5.29}$$

where $\mathbb{1}$ denotes the constant function equal to 1, and $\langle\cdot,\cdot\rangle$ denotes the weighted inner product. To get the first inequality, we apply the lower bound in (5.17) from Lemma 5.2.2, to get $\xi - \gamma \le \lambda_\Lambda + 1 - \gamma \le 1$ on Λ. The positivity of the resolvent gives

$$0 \le [\mathcal{R}_\gamma(1 - (\xi - \gamma))](x) = [\mathcal{R}_\gamma \mathbb{1}](x) - [\mathcal{R}_\gamma(\xi - \gamma)](x). \tag{5.30}$$

To get the second inequality, we write

$$(\mathcal{R}_\gamma \mathbb{1})(x) \le \sum_{x\in\Lambda}(\mathcal{R}_\gamma \mathbb{1})(x) = \sum_{x\in\Lambda}(\mathcal{R}_\gamma \mathbb{1})(x)\mathbb{1}(x) \le \sum_{x\in\Lambda}(\mathcal{R}_\gamma \mathbb{1})(x)\mathbb{1}(x)\deg(x) = \langle \mathcal{R}_\gamma \mathbb{1}, \mathbb{1}\rangle. \tag{5.31}$$

To get the third inequality, we use the Fourier expansion of the resolvent with respect to the orthonormal basis of eigenfunctions of $\Delta + \xi$ in $\langle\cdot,\cdot\rangle$. □

§5.2.2 Structural properties of the Galton-Watson tree

All of the results below can be lifted directly from [13] since the normalisation of the Laplacian plays no role for the properties of the Galton-Watson tree. The results are included for the sake of completeness.

Lemma 5.2.5. *[Maximal degree in a ball around the root]*
(a) Subject to Assumption 5.B(2)*, for every* $\delta > 0$*,*

$$\sum_{r\in\mathbb{N}} \mathfrak{P}\big(\exists\, x \in B_{2r}(\mathcal{O})\colon\ \deg(x) \geq \delta r\big) < \infty. \tag{5.32}$$

Proof. See [13, Lemma 2.3]. □

Lemma 5.2.5 shows that $\mathfrak{P}$-almost surely, as $r \to \infty$ all degrees in a ball of radius r are eventually less than δr for any $\delta > 0$.

Lemma 5.2.6. *[Volumes of large balls] If there exists an* $a > 0$ *such that* $\mathcal{E}[\mathrm{e}^{aD}] < \infty$*, then for any* R_r *satisfying* $\lim_{r\to\infty} R_r/\log r = \infty$*,*

$$\liminf_{r\to\infty} \frac{1}{R_r} \log\Big(\inf_{x\in B_r(\mathcal{O})} |B_{R_r}(x)|\Big) = \limsup_{r\to\infty} \frac{1}{R_r} \log\Big(\sup_{x\in B_r(\mathcal{O})} |B_{R_r}(x)|\Big) = \vartheta \qquad \mathfrak{P}\text{-}a.s. \tag{5.33}$$

Proof. See [13, Lemma 2.2]. □

Lemmma 5.2.6 gives that $\mathfrak{P}$-almost surely, any ball of radius r centred within distance r to the root also has volume $\mathrm{e}^{r\vartheta+o(1)}$ as $r \to \infty$.

§5.2.3 Estimates on the potential

All of the results below are lifted directly from [13], since the normalisation of the Laplacian plays no role in the properties of the potential. The results are included for the sake of completeness. Abbreviate $L_r = |B_r(\mathcal{O})|$ and put

$$S_r := (\log r)^\alpha, \qquad \alpha \in (0,1). \tag{5.34}$$

For every $r \in \mathbb{N}$ there is a unique a_r such that

$$\mathsf{P}(\xi(0) > a_r) = \frac{1}{r}. \tag{5.35}$$

By Assumption 5.A, for r large enough

$$a_r = \varrho \log\log r. \tag{5.36}$$

For $r \in \mathbb{N}$ and $A > 0$, let

$$\Pi_{r,A} = \Pi_{r,A}(\xi) := \{z \in B_r(\mathcal{O})\colon\ \xi(z) > a_{L_r} - 2A\} \tag{5.37}$$

CHAPTER 5

be the set of vertices in $B_r(\mathcal{O})$ where the potential is close to maximal,

$$D_{r,A} = D_{r,A}(\xi) := \{z \in B_r(\mathcal{O})\colon \operatorname{dist}(z, \Pi_{r,A}) \le S_r\} \tag{5.38}$$

be the S_r-neighbourhood of $\Pi_{r,A}$, and $\mathfrak{C}_{r,A}$ be the set of connected components of $D_{r,A}$ in $\mathcal{GW}$, which we think of as *islands*. For $M_A \in \mathbb{N}$, define the event

$$\mathcal{B}_{r,A} := \big\{\exists \mathcal{C} \in \mathfrak{C}_{r,A}\colon\ |\mathcal{C} \cap \Pi_{r,A}| > M_A\big\}. \tag{5.39}$$

Note that $\Pi_{r,A}, D_{r,A}, \mathcal{B}_{r,A}$ depend on $\mathcal{GW}$ and therefore are random.

Lemma 5.2.7. *[Maximum size of the islands] Subject to Assumptions* 5.A–5.B*, for every $A > 0$ there exists an $M_A \in \mathbb{N}$ such that*

$$\sum_{r\in\mathbb{N}} \mathsf{P}(\mathcal{B}_{r,A}) < \infty \qquad \mathfrak{P} - a.s. \tag{5.40}$$

Proof. See [13, Lemma 3.1] and [8, Lemma 6.6]. □

Lemma 5.2.7 implies that $(\mathsf{P} \times \mathfrak{P})$-a.s. $\mathcal{B}_{r,A}$ does not occur eventually as $r \to \infty$. Note that $\mathfrak{P}$-a.s. on the event $[\mathcal{B}_{r,A}]^c$,

$$\forall \mathcal{C} \in \mathfrak{C}_{r,A}\colon\ |\mathcal{C} \cap \Pi_{r,A}| \le M_A,\ \operatorname{diam}_{\mathcal{GW}}(\mathcal{C}) \le 2M_A S_r,\ |\mathcal{C}| \le \mathrm{e}^{2\vartheta M_A S_r}, \tag{5.41}$$

where the last inequality follows from Lemma 5.2.6.

Lemma 5.2.8. *[Maximum of the potential] Subject to Assumptions* 5.A–5.B*, for any $\vartheta > 0$, $(\mathsf{P} \times \mathfrak{P})$-a.s. eventually as $r \to \infty$,*

$$\left|\max_{x\in B_r(\mathcal{O})} \xi(x) - a_{L_r}\right| \le \frac{2\varrho \log r}{\vartheta r}. \tag{5.42}$$

Proof. See [12, Lemma 2.4]. The proof carries over verbatim and uses Lemma 5.2.6. □

Lemma 5.2.9. *[Number of intermediate peaks of the potential] Subject to Assumptions* 5.A *and* 5.B(2)*, for any $\beta \in (0,1)$ and $\varepsilon \in (0, \frac{1}{2}\beta)$ the following holds. For a self-avoiding path π in $\mathcal{GW}$, set*

$$N_\pi = N_\pi(\xi) := |\{z \in \operatorname{supp}(\pi)\colon\ \xi(z) > (1-\varepsilon)a_{L_r}\}|. \tag{5.43}$$

Define the event

$$\mathcal{B}_r := \left\{ \substack{\text{there exists a self-avoiding path } \pi \text{ in } \mathcal{GW} \text{ with} \\ \operatorname{supp}(\pi)\cap B_r(\mathcal{O})\neq\emptyset,\, |\operatorname{supp}(\pi)|\ge(\log L_r)^\beta \text{ and } N_\pi > \frac{|\operatorname{supp}(\pi)|}{(\log L_r)^\varepsilon}} \right\}. \tag{5.44}$$

Then

$$\sum_{r\in\mathbb{N}_0} \mathsf{P}(\mathcal{B}_r) < \infty \qquad \mathfrak{P} - a.s. \tag{5.45}$$

Proof. See [13, Lemma 3.6]. □

CHAPTER 5

Lemma 5.2.9 implies that $(\mathsf{P}\times\mathfrak{P})$-a.s. for r large enough, all self-avoiding paths π in $\mathcal{GW}$ with $\mathrm{supp}(\pi)\cap B_r(\mathcal{O})\neq\emptyset$ and $|\mathrm{supp}(\pi)|\geq(\log L_r)^\beta$ satisfy $N_\pi\leq\frac{|\mathrm{supp}(\pi)|}{(\log L_r)^\varepsilon}$.

Lemma 5.2.10. *[Number of high exceedances of the potential] Subject to Assumptions* 5.A *and* 5.B(2)*, for any $A>0$ there is a $C\geq 1$ such that, for all $\delta\in(0,1)$, the following holds. For a self-avoiding path π in $\mathcal{GW}$, let*

$$N_\pi := |\{x\in\mathrm{supp}(\pi)\colon\ \xi(x)>a_{L_r}-2A\}|. \tag{5.46}$$

Define the event

$$\mathcal{B}_r := \left\{ {\text{there exists a self-avoiding path } \pi \text{ in } G \text{ with} \atop \mathrm{supp}(\pi)\cap B_r(\mathcal{O})\neq\emptyset,\, |\mathrm{supp}(\pi)|\geq C(\log L_r)^\delta \text{ and } N_\pi>\frac{|\mathrm{supp}(\pi)|}{(\log L_r)^\delta}} \right\}. \tag{5.47}$$

Then $\sum_{r\in\mathbb{N}_0}\sup_{G\in\mathfrak{G}_r}\mathsf{P}(\mathcal{B}_r)<\infty$. In particular, $(\mathsf{P}\times\mathfrak{P})$-a.s. for r large enough, all self-avoiding paths π in $\mathcal{GW}$ with $\mathrm{supp}(\pi)\cap B_r(\mathcal{O})\neq\emptyset$ and $|\mathrm{supp}(\pi)|\geq C(\log L_r)^\delta$ satisfy

$$N_\pi = |\{x\in\mathrm{supp}(\pi)\colon\ \xi(x)>a_{L_r}-2A\}|\leq\frac{|\mathrm{supp}(\pi)|}{(\log L_r)^\delta}. \tag{5.48}$$

Proof. See [13, Lemma 3.7]. □

Lemma 5.2.11. *[Principal eigenvalues of the islands] Subject to Assumptions* 5.A *and* 5.B(2)*, for any $\varepsilon>0$, $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $r\to\infty$,*

$$\text{all } \mathcal{C}\in\mathfrak{C}_{r,A} \text{ satisfy: } \lambda_{\mathcal{C}}(\xi;\mathcal{GW})\leq a_{L_r}-\widehat{\chi}_{\mathcal{C}}(\mathcal{GW})+\varepsilon. \tag{5.49}$$

Proof. See [13, Lemma 3.3]. □

Corollary 5.2.12. *[Uniform bound on principal eigenvalue of the islands] Subject to Assumptions* 5.A–5.B*, for ϑ as in* (5.4)*, and any $\varepsilon>0$, $(\mathsf{P}\times\mathfrak{P})$-a.s. eventually as $r\to\infty$,*

$$\max_{\mathcal{C}\in\mathfrak{C}_{r,A}}\lambda^1_{\mathcal{C}}(\xi;G)\leq a_{L_r}-\widetilde{\chi}(\varrho)+\varepsilon. \tag{5.50}$$

Proof. See [12, Corollary 2.8]. □

§5.3 Path expansions

In this section we adapt [12, Section 3] to fit with the random walk generated by the normalised Laplacian. Section 5.3.1 proves three lemmas that concern the contribution to the total mass in (5.6) coming from various sets of paths. Section 5.3.2 proves a key proposition that controls the entropy associated with a key set of paths. The proof of which is based on the three lemmas in Section 5.3.1.

We need various sets of nearest-neighbour paths in $\mathcal{GW}=(V,E,\mathcal{O})$, defined in [12]. For $\ell\in\mathbb{N}_0$ and subsets $\Lambda,\Lambda'\subset V$, put

$$\begin{aligned}\mathscr{P}_\ell(\Lambda,\Lambda') &:= \left\{(\pi_0,\dots,\pi_\ell)\in V^{\ell+1}\colon \quad {\pi_0\in\Lambda,\pi_\ell\in\Lambda', \atop \{\pi_i,\pi_{i-1}\}\in E\ \forall\, 1\leq i\leq\ell}\right\},\\ \mathscr{P}(\Lambda,\Lambda') &:= \bigcup_{\ell\in\mathbb{N}_0}\mathscr{P}_\ell(\Lambda,\Lambda'),\end{aligned} \tag{5.51}$$

and set

$$\mathscr{P}_\ell := \mathscr{P}_\ell(V,V), \qquad \mathscr{P} := \mathscr{P}(V,V). \tag{5.52}$$

When Λ or Λ' consists of a single point, write x instead of $\{x\}$. For $\pi \in \mathscr{P}_\ell$, set $|\pi| := \ell$. Write $\mathrm{supp}(\pi) := \{\pi_0, \dots, \pi_{|\pi|}\}$ to denote the set of points visited by π.

Let $X = (X_t)_{t\geq 0}$ be the continuous-time random walk on G that jumps from $x \in V$ to any neighbour $y \sim x$ at rate 1. Denote by $(T_k)_{k\in\mathbb{N}_0}$ the sequence of jump times (with $T_0 := 0$). For $\ell \in \mathbb{N}_0$, let

$$\pi^\ell(X) := (X_0, \dots, X_{T_\ell}) \tag{5.53}$$

be the path in $\mathscr{P}_\ell$ consisting of the first ℓ steps of X. For $t \geq 0$, let

$$\pi(X_{[0,t]}) = \pi^{\ell_t}(X), \qquad \text{with } \ell_t \in \mathbb{N}_0 \text{ satisfying } T_{\ell_t} \leq t < T_{\ell_t+1}, \tag{5.54}$$

denote the path in $\mathscr{P}$ consisting of all the steps taken by X between times 0 and t.

Recall the definitions from Section 5.2.3. For $\pi \in \mathscr{P}$ and $A > 0$, define

$$\lambda_{r,A}(\pi) := \sup\big\{\lambda^1_{\mathcal{C}}(\xi; G)\colon\ \mathcal{C} \in \mathfrak{C}_{r,A},\ \mathrm{supp}(\pi) \cap \mathcal{C} \cap \Pi_{r,A} \neq \emptyset\big\}, \tag{5.55}$$

with the convention $\sup\emptyset = -\infty$. This is the largest principal eigenvalue among the components of $\mathfrak{C}_{r,A}$ in $\mathcal{GW}$ that have a point of high exceedance visited by the path π.

§5.3.1 Mass of the solution along excursions

Lemma 5.3.1. *[Path evaluation] For $\ell \in \mathbb{N}_0$, $\pi \in \mathscr{P}_\ell$ and $\gamma > \max_{0\leq i<|\pi|}\{\xi(\pi_i) - 1\}$,*

$$\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^{T_\ell}(\xi(X_s)-\gamma)\,\mathrm{d}s} \;\Big|\; \pi^\ell(X) = \pi\right] = \prod_{i=0}^{\ell-1} \frac{1}{\gamma - [\xi(\pi_i) - 1]}. \tag{5.56}$$

Proof. The proof is identical to that of [12, Lemma 3.2], except that the random walk now jumps with rate 1. □

For a path $\pi \in \mathscr{P}$ and $\varepsilon \in (0,1)$, we write

$$M^{r,\varepsilon}_\pi := \big|\big\{0 \leq i < |\pi|\colon\ \xi(\pi_i) \leq (1-\varepsilon)a_{L_r}\big\}\big|, \tag{5.57}$$

with the interpretation that $M^{r,\varepsilon}_\pi = 0$ if $|\pi| = 0$.

Lemma 5.3.2. *[Mass of excursions] Subject to Assumption* 5.A*, for every $A, \varepsilon > 0$, there exists $c > 0$ and $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$, all $\gamma > a_{L_r} - A$ and all $\pi \in \mathscr{P}(B_r(\mathcal{O}), B_r(\mathcal{O}))$ satisfying $\pi_i \notin \Pi_{r,A}$ for all $0 \leq i < \ell := |\pi|$,*

$$\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^{T_\ell}(\xi(X_s)-\gamma)\,\mathrm{d}s} \;\Big|\; \pi^\ell(X) = \pi\right] \leq q^\ell_{r,A}\,\mathrm{e}^{M^{r,\varepsilon}_\pi(c-\log\log\log L_r)}, \tag{5.58}$$

where

$$q_A := \frac{1}{1+A} \quad \text{and} \quad c = \log[2(q_A\varepsilon\varrho)^{-1}]. \tag{5.59}$$

Note that $\pi_\ell \in \Pi_{r,A}$ is allowed.

Proof. The proof is identical to that of [12, Lemma 3.3], and uses Lemma 5.3.1. □

We follow [12, Definition 3.4] and [9, Section 6.2]. Note that the distance between $\Pi_{r,A}$ and $D^{c}_{r,A}$ in $\mathcal{GW}$ is at least $S_r = (\log L_r)^{\alpha}$ (recall (5.37)–(5.38)).

Definition 5.3.3. [Concatenation of paths] (a) When π and π' are two paths in $\mathscr{P}$ with $\pi_{|\pi|} = \pi'_0$, we define their *concatenation* as

$$\pi \circ \pi' := (\pi_0, \dots, \pi_{|\pi|}, \pi'_1, \dots, \pi'_{|\pi'|}) \in \mathscr{P}. \tag{5.60}$$

Note that $|\pi \circ \pi'| = |\pi| + |\pi'|$.

(b) When $\pi_{|\pi|} \neq \pi'_0$, we can still define the *shifted concatenation* of π and π' as $\pi \circ \hat{\pi}'$, where $\hat{\pi}' := (\pi_{|\pi|}, \pi_{|\pi|} + \pi'_1 - \pi'_0, \dots, \pi_{|\pi|} + \pi'_{|\pi'|} - \pi'_0)$. The shifted concatenation of multiple paths is defined inductively via associativity. ■

Now, if a path $\pi \in \mathscr{P}$ intersects $\Pi_{r,A}$, then it can be decomposed into an initial path, a sequence of excursions between $\Pi_{r,A}$ and $D^{c}_{r,A}$, and a terminal path. More precisely, there exists $m_\pi \in \mathbb{N}$ such that

$$\pi = \check{\pi}^1 \circ \hat{\pi}^1 \circ \cdots \circ \check{\pi}^{m_\pi} \circ \hat{\pi}^{m_\pi} \circ \bar{\pi}, \tag{5.61}$$

where the paths in (5.61) satisfy

$$\begin{array}{llll} \check{\pi}^1 \in \mathscr{P}(V, \Pi_{r,A}) & \text{with} & \check{\pi}^1_i \notin \Pi_{r,A}, & 0 \le i < |\check{\pi}^1|, \\ \hat{\pi}^k \in \mathscr{P}(\Pi_{r,A}, D^{c}_{r,A}) & \text{with} & \hat{\pi}^k_i \in D_{r,A}, & 0 \le i < |\hat{\pi}^k|,\ 1 \le k \le m_\pi - 1, \\ \check{\pi}^k \in \mathscr{P}(D^{c}_{r,A}, \Pi_{r,A}) & \text{with} & \check{\pi}^k_i \notin \Pi_{r,A}, & 0 \le i < |\check{\pi}^k|,\ 2 \le k \le m_\pi, \\ \hat{\pi}^{m_\pi} \in \mathscr{P}(\Pi_{r,A}, V) & \text{with} & \hat{\pi}^{m_\pi}_i \in D_{r,A}, & 0 \le i < |\hat{\pi}^{m_\pi}|, \end{array} \tag{5.62}$$

while

$$\begin{array}{ll} \bar{\pi} \in \mathscr{P}(D^{c}_{r,A}, V) \text{ and } \bar{\pi}_i \notin \Pi_{r,A}\ \forall\, i \ge 0 & \text{if } \hat{\pi}^{m_\pi} \in \mathscr{P}(\Pi_{r,A}, D^{c}_{r,A}), \\ \bar{\pi}_0 \in D_{r,A}, |\bar{\pi}| = 0 & \text{otherwise.} \end{array} \tag{5.63}$$

Note that the decomposition in (5.61)–(5.63) is unique, and that the paths $\check{\pi}^1$, $\hat{\pi}^{m_\pi}$ and $\bar{\pi}$ can have zero length. If π is contained in $B_r(\mathcal{O})$, then so are all the paths in the decomposition.

Whenever $\mathrm{supp}(\pi) \cap \Pi_{r,A} \neq \emptyset$ and $\varepsilon > 0$, we define

$$s_\pi := \sum_{i=1}^{m_\pi} |\check{\pi}^i| + |\bar{\pi}|, \qquad k^{r,\varepsilon}_\pi := \sum_{i=1}^{m_\pi} M^{r,\varepsilon}_{\check{\pi}^i} + M^{r,\varepsilon}_{\bar{\pi}}, \tag{5.64}$$

to be the total time spent in exterior excursions, respectively, on moderately low points of the potential visited by exterior excursions (without their last point).

In case $\mathrm{supp}(\pi) \cap \Pi_{r,A} = \emptyset$, we set $m_\pi := 0$, $s_\pi := |\pi|$ and $k^{r,\varepsilon}_\pi := M^{r,\varepsilon}_\pi$. Recall from (5.55) that, in this case, $\lambda_{r,A}(\pi) = -\infty$.

We say that $\pi, \pi' \in \mathscr{P}$ are *equivalent*, written $\pi' \sim \pi$, if $m_\pi = m_{\pi'}$, $\check{\pi}'^i = \check{\pi}^i$ for all $i = 1, \dots, m_\pi$, and $\bar{\pi}' = \bar{\pi}$. If $\pi' \sim \pi$, then $s_{\pi'}$, $k^{r,\varepsilon}_{\pi'}$ and $\lambda_{r,A}(\pi')$ are all equal to the counterparts for π.

To state our key lemma, we define, for $m, s \in \mathbb{N}_0$,

$$\mathscr{P}^{(m,s)} = \{\pi \in \mathscr{P}\colon\ m_\pi = m, s_\pi = s\}\,, \tag{5.65}$$

and denote by

$$C_{r,A} := \max\{|\mathcal{C}|\colon\ \mathcal{C} \in \mathfrak{C}_{r,A}\} \tag{5.66}$$

the maximal size of the islands in $\mathfrak{C}_{r,A}$.

Lemma 5.3.4. *[Mass of an equivalence class] For every $A, \varepsilon > 0$ there exist $c > 0$ and $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$, all $m, s \in \mathbb{N}_0$, all $\pi \in \mathscr{P}^{(m,s)}$ with $\operatorname{supp}(\pi) \subset B_r(\mathcal{O})$, all $\gamma > \lambda_{r,A}(\pi) \vee (a_{L_r} - A)$ and all $t \geq 0$,*

$$\begin{aligned} &\mathbb{E}_{\pi_0}\left[\mathrm{e}^{\int_0^t (\xi(X_u)-\gamma)\,\mathrm{d}u}\,\mathbb{1}_{\{\pi(X_{[0,t]})\sim\pi\}}\right] \\ &\quad\leq \left(C_{r,A}^{1/2}\right)^{\mathbb{1}_{\{m>0\}}}\left(1+\frac{C_{r,A}}{\gamma-\lambda_{r,A}(\pi)}\right)^m\left(\frac{q_A}{d_{\min}}\right)^s \mathrm{e}^{\left(c-\log^3 L_r\right)k_\pi^{r,\varepsilon}}. \end{aligned} \tag{5.67}$$

Proof. The proof is identical to that of [12, Lemma 3.5], except that the normalised Laplacian gives rise to Lemma 5.2.4 and Lemma 5.3.2, which are used instead. □

§5.3.2 Key proposition

The main result of this section is the following proposition.

Proposition 5.3.5. *[Entropy reduction] Let $\alpha \in (0,1)$ be as in (5.34) and $\kappa \in (\alpha, 1)$. Subject to Assumption 5.B, there exists an A_0 such that, for all $A \geq A_0$, with $\mathfrak{P}$-probability tending to one as $r \to \infty$, the following statement is true. For each $x \in B_r(\mathcal{O})$, each $\mathcal{N} \subset \mathscr{P}(x, B_r(\mathcal{O}))$ satisfying $\operatorname{supp}(\pi) \subset B_r(\mathcal{O})$ and $\max_{1\leq\ell\leq|\pi|}\operatorname{dist}_G(\pi_\ell, x) \geq (\log L_r)^\kappa$ for all $\pi \in \mathcal{N}$, and each assignment $\pi \mapsto (\gamma_\pi, z_\pi) \in \mathbb{R} \times V$ satisfying*

$$\gamma_\pi \geq \left(\lambda_{r,A}(\pi) + \mathrm{e}^{-S_r}\right) \vee (a_{L_r} - A) \qquad \forall\, \pi \in \mathcal{N} \tag{5.68}$$

and

$$z_\pi \in \operatorname{supp}(\pi) \cup \bigcup_{\substack{\mathcal{C}\in\mathfrak{C}_{r,A}\colon \\ \operatorname{supp}(\pi)\cap\mathcal{C}\cap\Pi_{r,A}\neq\emptyset}} \mathcal{C} \qquad \forall\, \pi \in \mathcal{N}, \tag{5.69}$$

the following inequality holds for all $t \geq 0$:

$$\log \mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s}\,\mathbb{1}_{\{\pi(X_{[0,t]})\in\mathcal{N}\}}\right] \leq \sup_{\pi\in\mathcal{N}}\Big\{t\gamma_\pi + \operatorname{dist}_G(x, z_\pi)(c - \log\log\log L_r)\Big\}. \tag{5.70}$$

Proof. The proof is based on [12, Section 3.4]. First fix $c_0 > 2$ and define

$$A_0 = \mathrm{e}^{4c_0} - 1. \tag{5.71}$$

Fix $A \geq A_0$, $\beta \in (0, \alpha)$ and $\varepsilon \in (0, \frac{1}{2}\beta)$ as in Lemma 5.2.9. Let $r_0 \in \mathbb{N}$ be as given in Lemma 5.3.4, and take $r \geq r_0$ so large that the conclusions of Lemmas 5.2.5, 5.2.7, 5.2.11 and 5.2.9 hold, i.e. assume that the events $\mathcal{B}_r$ and $\mathcal{B}_{r,A}$ in these lemmas do not occur. Fix $x \in B_r(\mathcal{O})$. Recall the definitions of $C_{r,A}$ and $\mathscr{P}^{(m,s)}$. Note that the relation $\sim$ is an equivalence relation in $\mathscr{P}^{(m,s)}$, and define

$$\widetilde{\mathscr{P}}_x^{(m,s)} := \left\{\text{equivalence classes of the paths in } \mathscr{P}(x,V) \cap \mathscr{P}^{(m,s)}\right\}. \tag{5.72}$$

The following bounded on the cardinality of this set is needed.

Lemma 5.3.6. *[Bound equivalence classes] Subject to Assumption* 5.B, $\mathfrak{P}$*-a.s.*,$|\widetilde{\mathscr{P}}_x^{(m,s)}| \le (2C_{r,A})^m(\delta r)^{(m+s)}$ *for all* $m, s \in \mathbb{N}_0$.

Proof. We can copy the proof of [12, Lemma 3.6], replacing $d_{\max}$ by δr. □

Now take $\mathcal{N} \subset \mathscr{P}(x,V)$ as in the statement, and set

$$\widetilde{\mathcal{N}}^{(m,s)} := \left\{\text{equivalence classes of paths in } \mathcal{N} \cap \mathscr{P}^{(m,s)}\right\} \subset \widetilde{\mathscr{P}}_x^{(m,s)}. \tag{5.73}$$

For each $\mathcal{M} \in \widetilde{\mathcal{N}}^{(m,s)}$, choose a representative $\pi_{\mathcal{M}} \in \mathcal{M}$, and use Lemma 5.3.6 to write

$$\begin{aligned}\mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u}\mathbb{1}_{\{\pi(X_{[0,t]})\in\mathcal{N}\}}\right] &= \sum_{m,s\in\mathbb{N}_0}\sum_{\mathcal{M}\in\widetilde{\mathcal{N}}^{(m,s)}} \mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u}\mathbb{1}_{\{\pi(X_{[0,t]})\sim\pi_{\mathcal{M}}\}}\right] \\ &\le \sum_{m,s\in\mathbb{N}_0}(2(\delta r)C_{r,A})^m(\delta r)^s \sup_{\pi\in\mathcal{N}^{(m,s)}} \mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u}\mathbb{1}_{\{\pi(X_{[0,t]})\sim\pi\}}\right]\end{aligned} \tag{5.74}$$

with the convention $\sup\emptyset = 0$. For fixed $\pi \in \mathcal{N}^{(m,s)}$, by (5.68), apply (5.67) and Lemma 5.2.7 to obtain, for all r large enough and with $c_0 > 2$,

$$\begin{aligned}&(2(\delta r))^m(\delta r)^s\,\mathbb{E}_x\left[\mathrm{e}^{\int_0^t \xi(X_u)\mathrm{d}u}\mathbb{1}_{\{\pi(X_{[0,t]})\sim\pi\}}\right] \\ &\quad\le \mathrm{e}^{t\gamma_\pi}\mathrm{e}^{c_0 m S_r}[q_A(\delta r)]^s\,\mathrm{e}^{k_\pi^{r,\varepsilon}(c-\log\log\log L_r)}.\end{aligned} \tag{5.75}$$

We next claim that, for r large enough and $\pi \in \mathcal{N}^{(m,s)}$,

$$s \ge [(m-1)\vee 1]\,S_r. \tag{5.76}$$

Indeed, when $m \ge 2$, $|\operatorname{supp}(\check{\pi}^i)| \ge S_r$ for all $2 \le i \le m$. When $m = 0$, $|\operatorname{supp}(\pi)| \ge \max_{1\le\ell\le|\pi|}|\pi_\ell - x| \ge (\log L_r)^\kappa \gg S_r$ by assumption. When $m = 1$, the latter assumption and Lemma 5.2.7 together imply that $\operatorname{supp}(\pi) \cap D_{r,A}^{\mathrm{c}} \ne \emptyset$, and so either $|\operatorname{supp}(\check{\pi}^1)| \ge S_r$ or $|\operatorname{supp}(\bar{\pi})| \ge S_r$. Thus, (5.76) holds by the definition of S_r and s.

Note that $q_A < \mathrm{e}^{-4c_0}$, so

$$\begin{aligned}&\sum_{m\ge0}\sum_{s\ge[(m-1)\vee1]S_r} \mathrm{e}^{c_0 m S_r}[q_A(\delta r)]^s \\ &= \frac{[q_A(\delta r)]^{S_r} + \mathrm{e}^{c_0 S_r}[q_A(\delta r)]^{S_r} + \sum_{m\ge2}\mathrm{e}^{mc_0 S_r}[q_A(\delta r)]^{(m-1)S_r}}{1-q_A\delta r} \\ &\le \frac{3\mathrm{e}^{-c_0\log r}}{1-q_A\delta r} < 1\end{aligned} \tag{5.77}$$

CHAPTER 5

for r large enough. Inserting this back into (5.74), we obtain

$$\log \mathbb{E}_x \left[\mathrm{e}^{\int_0^t \xi(X_s) \mathrm{d}s} \mathbb{1}_{\{\pi(X_{0,t}) \in \mathcal{N}\}} \right] \le \sup_{\pi \in \mathcal{N}} \left\{ t\gamma_\pi + k_\pi^{r,\varepsilon}(c - \log\log\log L_r) \right\}. \tag{5.78}$$

The remainder of the proof is identical to the end of [12, Section 3.4] and is included for completeness.

The proof will be finished once we show that, for some $\varepsilon' > 0$ and whp, respectively, a.s. eventually as $r \to \infty$,

$$k_\pi^{r,\varepsilon} \ge \operatorname{dist}_G(x, z_\pi)(1 - 2(\log L_r)^{-\varepsilon'}) \qquad \forall \pi \in \mathcal{N}. \tag{5.79}$$

For each $\pi \in \mathcal{N}$ define an auxiliary path $\pi_\star$ as follows. First note that by using our assumptions we can find points $z', z'' \in \operatorname{supp}(\pi)$ (not necessarily distinct) such that

$$\operatorname{dist}_G(x, z') \ge (\log L_r)^\kappa, \qquad \operatorname{dist}_G(z'', z_\pi) \le 2M_A S_r, \tag{5.80}$$

where the latter holds by (5.41). Write $\{z_1, z_2\} = \{z', z''\}$ with z_1, z_2 ordered according to their hitting times by π, i.e. $\inf\{\ell \colon \pi_\ell = z_1\} \le \inf\{\ell \colon \pi_\ell = z_2\}$. Define π_e as the concatenation of the loop erasure of π between x and z_1 and the loop erasure of π between z_1 and z_2. Since π_e is the concatenation of two self-avoiding paths, it visits each point at most twice. Finally, define $\pi_\star \sim \pi_e$ by replacing the excursions of π_e from $\Pi_{r,A}$ to $D_{r,A}^{\mathrm{c}}$ by direct paths between the corresponding endpoints, i.e. replace each $\hat{\pi}_e^i$ by $|\hat{\pi}_e^i| = \ell_i$, $(\hat{\pi}_e^i)_0 = x_i \in \Pi_{r,A}$, and $(\hat{\pi}_e^i)_{\ell_i} = y_i \in D_{r,A}^{\mathrm{c}}$ by a shortest-distance path $\widetilde{\pi}_\star^i$ with the same endpoints and $|\widetilde{\pi}_\star^i| = \operatorname{dist}_G(x_i, y_i)$. Since $\pi_\star$ visits each $x \in \Pi_{r,A}$ at most 2 times,

$$k_\pi^{r,\varepsilon} \ge k_{\pi_\star}^{r,\varepsilon} \ge M_{\pi_\star}^{r,\varepsilon} - 2|\operatorname{supp}(\pi_\star) \cap \Pi_{r,A}|(S_r + 1) \ge M_{\pi_\star}^{r,\varepsilon} - 4|\operatorname{supp}(\pi_\star) \cap \Pi_{r,A}| S_r. \tag{5.81}$$

Note that $M_{\pi_\star}^{r,\varepsilon} \ge |\{x \in \operatorname{supp}(\pi_\star) \colon \xi(x) \le (1-\varepsilon) a_{L_r}\}| - 1$ and, by (5.80), $|\operatorname{supp}(\pi_\star)| \ge \operatorname{dist}_G(x, z') \ge (\log L_r)^\kappa \gg (\log L_r)^{\alpha + 2\varepsilon'}$ for some $0 < \varepsilon' < \varepsilon$. Applying Lemmas 5.2.9–5.2.10 and using (5.34) and $L_r > r$, we obtain, for r large enough,

$$k_\pi^{r,\varepsilon} \ge |\operatorname{supp}(\pi_\star)| \left(1 - \frac{2}{(\log L_r)^\varepsilon} - \frac{4S_r}{(\log L_r)^{\alpha+2\varepsilon'}}\right) \ge |\operatorname{supp}(\pi_\star)| \left(1 - \frac{1}{(\log L_r)^{\varepsilon'}}\right). \tag{5.82}$$

On the other hand, since $|\operatorname{supp}(\pi_\star)| \ge (\log L_r)^\kappa$, by (5.80) we have

$$\begin{aligned} |\operatorname{supp}(\pi_\star)| &= \big(|\operatorname{supp}(\pi_\star)| + 2M_A S_r\big) - 2M_A S_r \\ &= \big(|\operatorname{supp}(\pi_\star)| + 2M_A S_r\big)\left(1 - \frac{2M_A S_r}{|\operatorname{supp}(\pi_\star)| + 2M_A S_r}\right) \\ &\ge (\operatorname{dist}_G(x, z'') + 2M_A S_r)\left(1 - \frac{2M_A S_r}{(\log L_r)^\kappa}\right) \\ &\ge \operatorname{dist}_G(x, z_\pi)\left(1 - \frac{1}{(\log L_r)^{\varepsilon'}}\right), \end{aligned} \tag{5.83}$$

where the first inequality uses that the distance between two points on $\pi_\star$ is less than the total length of $\pi_\star$. Now (5.79) follows from (5.82)–(5.83). ☐

§5.4 Proof of the main theorem

Define

$$U^*(t) := \mathrm{e}^{t[\varrho \log(\vartheta \mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho)]}, \tag{5.84}$$

where we recall (5.10). To prove Theorem 5.1.1 we show that

$$\frac{1}{t}\log U(t) - \frac{1}{t}\log U^*(t) = o(1), \quad t \to \infty, \qquad (\mathsf{P} \times \mathfrak{P})\text{-a.s.} \tag{5.85}$$

The proof proceeds via upper and lower bound, proved in Sections 5.4.1 and 5.4.2, respectively.

§5.4.1 Upper bound

We follow [12, Section 4.2]. The proof of the upper bound in (5.85) relies on two lemmas showing that paths staying inside a ball of radius $\lceil t^\gamma \rceil$ for some $\gamma \in (0,1)$ or leaving a ball of radius $t \log t$ have a negligible contribution to (5.6), the total mass of the solution.

Lemma 5.4.1. *[No long paths] For any $\ell_t \geq t \log t$,*

$$\lim_{t\to\infty} \frac{1}{U^*(t)} \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\ell_t}]^{\mathrm{c}}} < t\}}\right] = 0 \quad (\mathsf{P} \times \mathfrak{P}) - a.s. \tag{5.86}$$

Proof. We follow [12, Lemma 4.2]. For $r \geq \ell_t$, let

$$\mathcal{B}_r := \left\{ \max_{x \in B_r(\mathcal{O})} \xi(x) \geq a_{L_r} + 2\varrho \right\}. \tag{5.87}$$

Since $\lim_{t\to\infty} \ell_t = \infty$, Lemma 5.2.8 gives that P-a.s.

$$\bigcup_{r \geq \ell_t} \mathcal{B}_r \text{ does not occur eventually as } t \to \infty. \tag{5.88}$$

Therefore we can work on the event $\bigcap_{r \geq \ell_t} [\mathcal{B}_r]^{\mathrm{c}}$. On this event, we write

$$\begin{aligned}\mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\ell_t}]^{\mathrm{c}}} < t\}}\right] &= \sum_{r \geq \ell_t} \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\sup_{s\in[0,t]} |X_s| = r\}}\right] \\ &\leq \mathrm{e}^{2\varrho t} \sum_{r \geq \ell_t} \mathrm{e}^{\varrho t \log r + \varrho \log\log(\delta r)}\, \mathbb{P}_{\mathcal{O}}\left(J_t \geq r\right),\end{aligned} \tag{5.89}$$

where J_t is the number of jumps of X up to time t, and we use that $|B_r(\mathcal{O})| \leq (\delta r)^r$. Next, J_t is stochastically dominated by a Poisson random variable with parameter t. Hence

$$\mathbb{P}_{\mathcal{O}}\left(J_t \geq r\right) \leq \frac{(\mathrm{e}t)^r}{r^r} \leq \exp\left\{-r \log\left(\frac{r}{\mathrm{e}t}\right)\right\} \tag{5.90}$$

Chapter 5

for large r. Using that $\ell_t \geq t \log t$, we can easily check that, for $r \geq \ell_t$ and t large enough,

$$\varrho t \log r - r \log\left(\frac{r}{\mathrm{e}t}\right) < -3r, \qquad r \geq \ell_t. \tag{5.91}$$

Thus (5.89) is at most

$$\mathrm{e}^{2\varrho t} \sum_{r \geq \ell_t} \mathrm{e}^{-3r+\log\log(\delta r)} \leq \mathrm{e}^{2\varrho t} \sum_{r \geq \ell_t} \mathrm{e}^{-2r} \leq 2\,\mathrm{e}^{2\varrho t}\,\mathrm{e}^{-2\ell_t} \leq \mathrm{e}^{-\ell_t}. \tag{5.92}$$

Since $\lim_{t\to\infty} \ell_t = \infty$ and $\lim_{t\to\infty} U^*(t) = \infty$, this settles the claim. □

Lemma 5.4.2. *[No short paths] For any $\gamma \in (0,1)$,*

$$\lim_{t\to\infty} \frac{1}{U^*(t)} \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\lceil t^\gamma \rceil}]^{\mathrm{c}}} > t\}}\right] = 0 \quad (\mathsf{P} \times \mathfrak{P}) - a.s. \tag{5.93}$$

Proof. We follow [12, Lemma 4.3]. By Lemma 5.2.8 with $r = \lceil t^\gamma \rceil$, we may assume that

$$\max_{x \in B_{\lceil t^\gamma \rceil}} \xi(x) \leq \varrho \log\log L_{\lceil t^\gamma \rceil} + \frac{2\varrho \log\lceil t^\gamma \rceil}{\vartheta \lceil t^\gamma \rceil} \leq \gamma\varrho \log t + O(1), \quad t \to \infty, \tag{5.94}$$

where the second inequality uses that $\log L_{\lceil t^\gamma \rceil} \sim \log |B_{\lceil t^\gamma \rceil}(\mathcal{O})| \sim \vartheta \lceil t^\gamma \rceil$. Hence

$$\frac{1}{U^*(t)} \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\lceil t^\gamma \rceil}]^{\mathrm{c}}} > t\}}\right] \leq \frac{1}{U^*(t)} \mathrm{e}^{\gamma\varrho t \log t + O(1)} \leq \mathrm{e}^{-(1-\gamma)\varrho t \log t + Ct \log\log\log t}, \quad t \to \infty, \tag{5.95}$$

for any constant $C > 1$. □

The proof of the upper bound in (5.85) also relies on a third lemma estimating the contribution of paths leaving a ball of radius $\lceil t^\gamma \rceil$ for some $\gamma \in (0,1)$ but staying inside a ball of radius $t \log t$. We slice to annulus between these two balls into layers, and derive an estimate for paths that reach a given layer but do not reach the next layer. To that end, fix $\gamma \in (\alpha, 1)$ with α as in (5.34), and let

$$K_t := \lceil t^{1-\gamma} \log t \rceil, \qquad r_t^{(k)} := k\lceil t^\gamma \rceil, \quad 1 \leq k \leq K_t, \qquad \ell_t := K_t \lceil t^\gamma \rceil \geq t \log t. \tag{5.96}$$

For $1 \leq k \leq K_t$, define (recall (5.51))

$$\mathcal{N}_t^k := \left\{\pi \in \mathscr{P}(\mathcal{O}, V)\colon\ \mathrm{supp}(\pi) \subset B_{r_t^{k+1}}(\mathcal{O}),\ \mathrm{supp}(\pi) \cap B^{\mathrm{c}}_{r_t^k}(\mathcal{O}) \neq \emptyset\right\} \tag{5.97}$$

and set

$$U^k(t) := \mathbb{E}_{\mathcal{O}}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\pi_{[0,t]}(X) \in \mathcal{N}_t^k\}}\right]. \tag{5.98}$$

Lemma 5.4.3. *[Upper bound on $U^k(t)$] For any $\varepsilon > 0$, $(\mathsf{P} \times \mathfrak{P})$-a.s. eventually as $t \to \infty$,*

$$\sup_{1 \leq k \leq K_t} \frac{1}{t} \log U_t^k \leq \frac{1}{t} \log U^*(t) + \varepsilon. \tag{5.99}$$

Proof. We follow [12, Lemma 4.4]. Fix $k \in \{1, \dots, K_t\}$. For $\pi \in \mathcal{N}_t^k$, let

$$\gamma_\pi := \lambda_{r_t^{k+1},A}(\pi) + \mathrm{e}^{-S_{\lceil t^\gamma \rceil}}, \qquad z_\pi \in \mathrm{supp}(\pi), |z_\pi| > r_t^k, \tag{5.100}$$

chosen such that (5.68)–(5.69) are satisfied. By Proposition 5.3.5 and (5.59), $(\mathsf{P} \times \mathfrak{P})$-a.s. eventually as $t \to \infty$,

$$\frac{1}{t} \log U_t^k \le \gamma_\pi - \frac{|z_\pi|}{t} \Big(\log\log(\vartheta r_t^{(k+1)})] - c + o(1) \Big). \tag{5.101}$$

Using Corollary 5.2.12 and $\log L_r \sim \vartheta r$, we bound

$$\gamma_\pi \le \varrho \log(\vartheta r_t^{(k+1)}) - \widetilde{\chi}(\varrho) + \tfrac{1}{2}\varepsilon + o(1). \tag{5.102}$$

Moreover, $|z_\pi| > r_t^{k+1} - \lceil t^\gamma \rceil$ and

$$\frac{\lceil t^\gamma \rceil}{t} \Big(\log\log(\vartheta r_t^{(k+1)})] - c \Big) \le \frac{1}{t^{1-\gamma}} \log\log(2t \log t) = o(1). \tag{5.103}$$

Hence

$$\gamma_\pi \le F_t(r_t^{(k+1)}) - \widetilde{\chi}(\varrho) + \tfrac{1}{2}\varepsilon + o(1) \tag{5.104}$$

with

$$F_{c,t}(r) := \varrho \log(\vartheta r) - \frac{r}{t} \big[\log\log(\vartheta r) - c \big], \qquad r > 0. \tag{5.105}$$

The function $F_{c,t}$ is maximised at any point $r_{c,t}$ satisfying

$$\varrho t = r_{c,t} \log\log r_{c,t} - c r_{c,t} + \frac{r_{c,t}}{\log r_{c,t}}. \tag{5.106}$$

In particular, $r_t = \mathfrak{r}_t[1 + o(1)]$, which implies that

$$\sup_{r>0} F_t(r) \le \varrho \log(\vartheta \mathfrak{r}_t) - \varrho + o(1), \qquad t \to \infty. \tag{5.107}$$

Inserting (5.107) into (5.104), we obtain $\frac{1}{t} \log U_t^k < \varrho \log(\vartheta \mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) + \varepsilon$, which is the desired upper bound because $\varepsilon > 0$ is arbitrary. □

Proof of the upper bound in (5.85). To avoid repetition, all statements hold $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $t \to \infty$. Set

$$U^0(t) := \mathbb{E}_\mathcal{O}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\lceil t^\gamma \rceil}]^\mathrm{c}} > t\}} \right], \quad U^\infty(t) := \mathbb{E}_\mathcal{O}\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{[B_{\lceil t \log t \rceil}]^\mathrm{c}} \le t\}} \right]. \tag{5.108}$$

Then

$$U(t) \le U^0(t) + U^\infty(t) + K_t \max_{1 \le k \le K_t} U^k(t). \tag{5.109}$$

From Lemmas 5.4.1–5.4.3 and the fact that $K_t = o(t)$, we get

$$\limsup_{t\to\infty} \left\{ \frac{1}{t} \log U(t) - \frac{1}{t} \log U^*(t) \right\} \le \varepsilon. \tag{5.110}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the upper bound in (5.11). □

§5.4.2 Lower bound

We follow [12, Section 4.1]. Fix $\varepsilon > 0$. By the definition of $\widetilde{\chi}$, there exists an infinite rooted tree $T = (V', E', \mathcal{Y})$ with degrees in $\mathrm{supp}(D_g)$ such that $\chi_T(\varrho) < \widetilde{\chi}(\varrho) + \frac{1}{4}\varepsilon$. Let $Q_r = B_r^T(\mathcal{Y})$ be the ball of radius r around $\mathcal{Y}$ in T. By Proposition 5.2.1 and (5.15), there exist a radius $R \in \mathbb{N}$ and a potential profile $q\colon B_R^T \to \mathbb{R}$ with $\mathcal{L}_{Q_R}(q;\varrho) < 1$ (in particular, $q \le 0$) such that

$$\lambda_{Q_R}(q;T) \ge -\widehat{\chi}_{Q_R}(\varrho;T) - \tfrac{1}{2}\varepsilon > -\widetilde{\chi}(\varrho) - \varepsilon. \tag{5.111}$$

For $\ell \in \mathbb{N}$, let $B_\ell = B_\ell(\mathcal{O})$ denote the ball of radius ℓ around $\mathcal{O}$ in $\mathcal{GW}$. We will show next that, $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $\ell \to \infty$, B_ℓ contains a copy of the ball Q_R where the potentail ξ is bounded from below by $\varrho \log\log|B_\ell(\mathcal{O})| + q$.

Proposition 5.4.4. *[Balls with high exceedances]* $(\mathfrak{P} \times \mathsf{P})$*-almost surely eventually as* $\ell \to \infty$*, there exists a vertex* $z \in B_\ell$ *with* $B_{R+1}(z) \subset B_\ell$ *and an isomorphism* $\varphi : B_{R+1}(z) \to Q_{R+1}$ *such that* $\xi \ge \varrho \log\log|B_\ell(\mathcal{O})| + q \circ \varphi$ *in* $B_R(z)$*. In particular,*

$$\lambda_{B_R(z)}(\xi;\mathcal{GW}) > \varrho \log\log|B_\ell(\mathcal{O})| - \widetilde{\chi}(\varrho) - \varepsilon. \tag{5.112}$$

Any such z *necessarily satisfies* $|z| \ge c\ell$ $(\mathfrak{P} \times \mathsf{P})$*-a.s. eventually as* $\ell \to \infty$ *for some constant* $c = c(\varrho, \vartheta, \widetilde{\chi}(\varrho), \varepsilon) > 0$.

Proof. We follow [12, Proposition 4.1]. Only the last step changes as a result of the normalised Laplacian. First note that, as a consequence of the definition of $\mathcal{GW}$, it may be shown straightforwardly that, for some $p = p(T,R) \in (0,1)$ and $\mathfrak{P}$-almost surely eventually as $\ell \to \infty$, there exist $N \in \mathbb{N}$, $N \ge p|B_\ell|$ and distinct $z_1, \ldots, z_N \in B_\ell$ such that $B_{R+1}(z_i) \cap B_{R+1}(z_j) = \emptyset$ for $1 \le i \ne j \le N$ and, for each $1 \le i \le N$, $B_{R+1}(z_i) \subset B_\ell$ and $B_{R+1}(z_i)$ is isomorphic to Q_{R+1}. Now, by (5.3), for each $i \in \{1, \ldots, N\}$,

$$\mathsf{P}\left(\xi \ge \varrho \log\log|B_\ell| + q \text{ in } B_R(z_i)\right) = |B_\ell|^{-\mathcal{L}_{Q_R}(q)}. \tag{5.113}$$

Using additionally that $|B_\ell| \ge \ell$ and $1 - x \le \mathrm{e}^{-x}$, $x \in \mathbb{R}$, we obtain

$$\mathsf{P}(\nexists i \in \{1,\ldots,N\}\colon \xi \ge \varrho\log\log|B_\ell| + q \text{ in } B_R(z_i)) = \left(1 - |B_\ell|^{-\mathcal{L}_{Q_R}(q)}\right)^N \le \mathrm{e}^{-p\ell^{1-\mathcal{L}_{Q_R}(q)}},$$

which is summable in $\ell \in \mathbb{N}$, so the proof of the first statement is completed using the Borel-Cantelli lemma. As for the last statement note that by (5.17) and Lemma 5.2.8

$$\lambda_{B_{c\ell}}(\xi;\mathcal{GW}) \le \max_{x \in B_{c\ell}(\mathcal{O})} \xi(x) < a_{L_{c\ell}} + o(1) < a_{L_\ell} + \varrho \log c\vartheta + o(1) < a_{L_\ell} - \widetilde{\chi}(\varrho) - \varepsilon \tag{5.114}$$

provided $c > 0$ is small enough. □

Lemma 5.4.5. *Let* $z \in \mathcal{GW}$ *and let* $v_z = (v_{z,i})_{i=0}^{|z|}$ *be the shortest path from* $\mathcal{O}$ *to* z*, i.e.* $v_{z,0} = \mathcal{O}$*,* $v_{z,|z|} = z$*, and* $v_{z,i} \sim v_{z,i-1}$ *for* $i = 1, \ldots, |z|$*. Then*

$$\sum_{L \in \mathbb{N}} \mathfrak{P}\left(\bigcup_{z \in Z_L} \left\{\prod_{i=1}^{L} \frac{1}{\deg(v_{z,i})} \le \frac{1}{(\log L)^{\delta_L L}}\right\}\right) < \infty,$$

where δ_L *satisfies* $\lim_{L\to\infty} \delta_L \log\log L = \infty$.

Proof. For $L \in \mathbb{N}$, let $\mathcal{Z}_L$ be the L-th generation of $\mathcal{GW}$ rooted at $\mathcal{O}$. For $z \in \mathcal{Z}_L$, let

$$\mathcal{E}_z = \left\{ \prod_{i=1}^{L} \deg(v_{z_i}) \geq (\log L)^{\delta_L L} \right\}.$$

We want to estimate

$$\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z\right).$$

Pick any $K \in \mathbb{N}$ and estimate

$$\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z\right) \leq \mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z,\, |\mathcal{Z}_L| > K\right) + \mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z,\, |\mathcal{Z}_L| \leq K\right).$$

Estimate

$$\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z,\, |\mathcal{Z}_L| > K\right) \leq \mathfrak{P}\left(|\mathcal{Z}_L| > K\right) \leq \frac{1}{K}\mathfrak{E}(\mathcal{Z}_L) =: \frac{e_L}{K}.$$

Also estimate

$$\begin{aligned}
&\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z,\, |\mathcal{Z}_L| \leq K\right)\\
&= \sum_{\ell=1}^{K} \mathfrak{P}\left(\cup_{k=1}^{\ell}\mathcal{E}_{z_\ell},\, |\mathcal{Z}_L| = \ell\right) \leq \sum_{\ell=1}^{K}\sum_{k=1}^{\ell} \mathfrak{P}\left(\mathcal{E}_{z_\ell},\, |\mathcal{Z}_L| = \ell\right)\\
&\leq \sum_{\ell=1}^{K}\sum_{k=1}^{\ell} \mathfrak{P}\left(\mathcal{E}_{z_1},\, |\mathcal{Z}_L| = \ell\right) \leq K \sum_{\ell=1}^{K} \mathfrak{P}\left(\mathcal{E}_{z_1},\, |\mathcal{Z}_L| = \ell\right)\\
&= K\mathfrak{P}\left(\mathcal{E}_{z_1},\, |\mathcal{Z}_L| \leq K\right) \leq K\mathfrak{P}\left(\mathcal{E}_{z_1}\right) =: Kp_L,
\end{aligned}$$

where z_ℓ is the ℓ-th vertex in $\mathcal{Z}_L$ (say in lexicographic order), and p_L is the probability that the product of L i.i.d. copies of the degrees exceeds $(\log L)^{\delta_L L}$. In the last inequality we need not worry about the correlation between $\mathcal{E}_{z_1}$ and the event $|\mathcal{Z}_L| \leq K$ because we drop the latter. Thus, for any $K \in \mathbb{N}$ we have

$$\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z\right) \leq \frac{e_L}{K} + Kp_L.$$

Now minimise over K. The minimising value is $K = \sqrt{e_L/p_L}$ (to be rounded off to an integer), so that we get

$$\mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z\right) \leq 2\sqrt{e_L p_L}.$$

Since $e_L = \mathrm{e}^{\vartheta L + o(L)}$ and $p_L = \mathrm{e}^{-L\delta_L \log\log L + O(L)}$, it follows that

$$\sum_{L\in\mathbb{N}} \mathfrak{P}\left(\cup_{z\in\mathcal{Z}_L}\mathcal{E}_z\right) < \infty$$

by the assumption on δ_L. □

Lemma 5.4.5 implies that $\mathfrak{P}$-almost surely eventually as $L \to \infty$, any path y_z must satisfy

$$\prod_{i=1}^{L} \frac{1}{\deg(y_{z,i})} \geq \frac{1}{(\log L)^{\delta_L L}}. \tag{5.115}$$

CHAPTER 5

Proof of the lower bound in (5.11). Let z be as in Proposition 5.4.4. Write τ_z for the hitting time of z by the random walk X. For $s \in (0,t)$, we estimate

$$\begin{aligned} U(t) &\geq \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^t \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{\tau_z\leq s\}}\, \mathbb{1}_{\{X_u\in B_R(z)\,\forall u\in[\tau_z,t]\}}\Big] \\ &= \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^{\tau_z} \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{\tau_z\leq s\}}\, \mathbb{E}_z\Big[\mathrm{e}^{\int_0^{v} \xi(X_u)\,\mathrm{d}u}\, \mathbb{1}_{\{X_u\in B_R(z)\,\forall u\in[0,v]\}}\Big]\Big|_{v=t-\tau_z}\Big], \end{aligned} \tag{5.116}$$

where we use the strong Markov property at time τ_z. We first bound the last term in the integrand in (5.116). Since $\xi \geq \varrho \log\log|B_\ell| + q$ in $B_R(z)$,

$$\begin{aligned} \mathbb{E}_z\Big[\mathrm{e}^{\int_0^{v} \xi(X_u)\,\mathrm{d}u} \mathbb{1}_{\{X_u\in B_R(z)\,\forall u\in[0,v]\}}\Big] &\geq \mathrm{e}^{v\varrho\log\log|B_\ell|}\mathbb{E}_{\mathcal{Y}}\Big[\mathrm{e}^{\int_0^{v} q(X_u)\,\mathrm{d}u} \mathbb{1}_{\{X_u\in Q_R\,\forall u\in[0,v]\}}\Big] \\ &\geq \mathrm{e}^{v\varrho\log\log|B_\ell|}\mathrm{e}^{v\lambda_{Q_R}(q;T)}\phi^1_{Q_R}(\mathcal{Y})^2 \\ &> \exp\big\{v\,(\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon)\,\big\} \end{aligned} \tag{5.117}$$

for large v, where we use that $B_{R+1}(z)$ is isomorphic to Q_{R+1} for the indicators in the first inequality, and apply Lemma 5.2.3 and (5.111) to obtain the second and third inequalities respectively. On the other hand since $\xi \geq 0$, we have

$$\mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^{\tau_z} \xi(X_u)\,\mathrm{d}u} \mathbb{1}\{\tau_z \leq s\}\Big] \geq \mathbb{P}_{\mathcal{O}}(\tau_z \leq s), \tag{5.118}$$

and we can bound the latter probability from below by the probability that the random walk runs along a shortest path from the root $\mathcal{O}$ to z within a time at most s. This gives

$$\mathsf{P}_{\mathcal{O}}(\tau_z\leq s) \geq \Big(\prod_{i=1}^{|z|}\frac{1}{\deg(y_{z,i})}\Big)P\Big(\sum_{i=1}^{|z|}E_i\leq s\Big) \geq (\log|z|)^{-\delta_{|z|}|z|}\mathrm{Poi}_{d_{\min}s}([|z|,\infty)), \tag{5.119}$$

where Poi_γ is the Poisson distribution with parameter γ, and P is the generic symbol for probability. The final inequality uses Lemma 5.4.5. Summarising, we obtain

$$\begin{aligned} U(t) &\geq (\log|z|)^{-\delta_{|z|}|z|}\mathrm{e}^{-d_{\min}s}\frac{(d_{\min}s)^{|z|}}{|z|!}\mathrm{e}^{(t-s)\left[\varrho\log\log|B_\ell|-\widetilde{\chi}(\varrho)-\varepsilon\right]} \\ &\geq \exp\left\{-d_{\min}s + (t-s)\left[\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon\right] - |z|\log\left(\frac{(\log|z|)^{\delta_{|z|}}}{d_{\min}}\frac{|z|}{s}\right)\right\} \\ &\geq \exp\left\{-d_{\min}s + (t-s)\left[\varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon\right] - \ell\log\left(\frac{(\log\ell)^{\delta_\ell}}{d_{\min}}\frac{\ell}{s}\right)\right\}, \end{aligned} \tag{5.120}$$

where in the last inequality we use that $s \leq |z|$ and $\ell \geq |z|$. Further assuming that $\ell = o(t)$, we see that the optimum over s is obtained at

$$s = \frac{\ell}{d_{\min} + \varrho\log\log|B_\ell| - \widetilde{\chi}(\varrho) - \varepsilon} = o(t). \tag{5.121}$$

Note that, by Proposition 5.4.4, this s indeed satisfies $s \le |z|$. Applying (5.4) we get, after a straightforward computation, $(\mathfrak{P} \times \mathsf{P})$-a.s. eventually as $t \to \infty$,

$$\frac{1}{t}\log U(t) \ge \varrho \log\log|B_\ell| - \frac{\ell}{t}\log\log\ell - \frac{\ell}{t}\delta_\ell \log\log\ell - \widetilde{\chi}(\varrho) - \varepsilon + O\left(\frac{\ell}{t}\right). \tag{5.122}$$

Inserting $\log|B_\ell| \sim \vartheta\ell$, we get

$$\frac{1}{t}\log U(t) \ge F_\ell - \widetilde{\chi}(\varrho) - \varepsilon + o(1) + O\left(\frac{\ell}{t}\right) \tag{5.123}$$

with

$$F_\ell = \varrho\log(\vartheta\ell) - \frac{\ell}{t}\log\log\ell - \frac{\ell}{t}\delta_\ell\log\log\ell. \tag{5.124}$$

The optimal ℓ for F_ℓ satisfies

$$\varrho t = \ell[1 + \delta_\ell + l\tfrac{\mathrm{d}}{\mathrm{d}\ell}\delta_\ell]\log\log\ell + \frac{\ell\delta_\ell}{\log\ell} + \frac{\ell}{\log\ell}, \tag{5.125}$$

i.e. $\ell = \mathfrak{r}_t[1+o(1)]$. For this choice we obtain

$$\frac{1}{t}\log U(t) \ge \varrho\log(\vartheta\mathfrak{r}_t) - \varrho - \widetilde{\chi}(\varrho) - \varepsilon + o(1). \tag{5.126}$$

Hence $(\mathfrak{P} \times \mathsf{P})$-a.s.

$$\liminf_{t\to\infty}\left\{\frac{1}{t}\log U(t) - \frac{1}{t}\log U^*(t)\right\} \ge -\varepsilon. \tag{5.127}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the lower bound in (5.11). □

§5.5 Analysis of the variational formula

This section is dedicated the analysis of variational formula. Proposition 5.2.1 is proven in Section 5.5.1. Theorem 5.1.2 is proven in Section 5.5.2, which is done by adapting the gluing argument in [12].

§5.5.1 Alternative representations for χ

The inequality is clear. For the equality we first prove that for any graph $G = (V, E)$ and $\Lambda \subset V$ finite,

$$\widehat{\chi}_\Lambda(\varrho; G) = \inf_{\substack{p\in\mathcal{P}(V):\\ \mathrm{supp}(p)\subset\Lambda}} [I_E(p) + \varrho J_V(p)]. \tag{5.128}$$

For this we follow [25, Lemma 2.17]. By the Rayleigh-Ritz formula,

$$\lambda_\Lambda(q; G) = \sup_{\substack{\mathrm{supp}(\phi)\subset\Lambda\\ \|\phi\|=1}} \langle(\Delta_G + q)\phi, \phi\rangle = \sup_{\|\phi\|=1}\left\{\sum_{x\in\Lambda}\deg(x)[(\Delta\phi)(x) + q(x)\phi(x)]\phi(x)\right\}$$
$$= \sup_{\|\phi\|=1}\left\{-\sum_{\{x,y\}\in E_\Lambda}[\phi(x) - \phi(y)]^2 + \sum_{x\in\Lambda}\deg(x)q(x)\phi(x)^2\right\}.$$

Chapter 5

By Lemma 5.2.2(2), the eigenfunction corresponding to $\lambda_\Lambda(q;G)$ may be taken to be non-negative, and we may therefore make the substitution $\phi(x) = \sqrt{\frac{p(x)}{\deg(x)}}$ so that p with $p(x) = \phi(x)^2 \deg(x)$ is a probability measure supported on Λ. So

$$\lambda_\Lambda(q;G) = \sup_{\substack{p\in\mathcal{P}(V)\\ \text{supp}(p)\subset\Lambda}} \left\{ -I_{E_\Lambda}(p) + \sum_{x\in\Lambda} q(x)p(x) \right\},$$

and therefore

$$\begin{aligned}\widehat{\chi}_\Lambda(\varrho;G) &= -\sup_{\substack{q:V\to[-\infty,\infty)\\ \mathcal{L}_V(q;\varrho)=1}} \left[\sup_{p\in\mathcal{P}(\Lambda)} \left\{ -I_{E_\Lambda}(p) + \sum_{x\in\Lambda} q(x)p(x) \right\} \right] \\ &= -\sup_{p\in\mathcal{P}(\Lambda)} \left[\sup_{q:\ \mathcal{L}(q;\varrho)=1} \left\{ \sum_{x\in\Lambda} q(x)p(x) - \varrho \log \sum_{x\in\Lambda} \mathrm{e}^{q(x)/\varrho} \right\} - I_{E_\Lambda}(p) \right].\end{aligned}$$

As the expression in the curly brackets does not change by adding a constant to $q(x)$, the inner supremum may be taken over all $q : \Lambda \to \mathbb{R}$.

For $z \in \Lambda$, differentiating with respect to $q(z)$ and setting equal to 0 we get that the supremum is attained at $\bar{q}$ satisfying

$$p(z) = \frac{\mathrm{e}^{\bar{q}(z)/\varrho}}{\sum_{x\in\Lambda} \mathrm{e}^{\bar{q}(x)/\varrho}}$$

for all z. Or equivalently,

$$\bar{q}(z) = \varrho \log p(z) + \varrho \log \sum_{x\in\Lambda} \mathrm{e}^{\bar{q}(x)/\varrho}.$$

This gives that the value of the inner supremum is $-\varrho J_V(p)$ and (5.128) follows.

Recall the definition of $B_r(\mathcal{O})$ from (5.4). By (5.128), $\widehat{\chi}_{B_r(\mathcal{O})}(\varrho;G)$ is non-increasing in r and therefore,

$$\lim_{r\to\infty} \widehat{\chi}_{B_r(\mathcal{O})}(\varrho;G) \ge \chi_G(\varrho). \tag{5.129}$$

It remains to show the opposite inequality. For that we show that for any $p \in \mathcal{P}(V)$ and $r \in \mathbb{N}$, there exists a $p_r \in \mathcal{P}(V)$ with support in $B_r(\mathcal{O})$ such that

$$\liminf_{r\to\infty} \{I_E(p_r) + \varrho J_V(p_r)\} \le I_E(p) + \varrho J_V(p). \tag{5.130}$$

We follow [12, Lemma A.2]. Simply take

$$p_r(x) = \frac{p(x)\mathbb{1}_{B_r(\mathcal{O})}(x)}{p(B_r(\mathcal{O}))}, \qquad x \in V, \tag{5.131}$$

i.e. the normalized restriction of p to $B_r(\mathcal{O})$. Then we easily see that

$$\begin{aligned}J_V(p_r) - J_V(p) &= -\frac{1}{p(B_r(\mathcal{O}))} \sum_{x\in B_r(\mathcal{O})} p(x)\log p(x) + \log p(B_r(\mathcal{O})) + \sum_{x\in V} p(x)\log p(x) \\ &\le \frac{J_V(p)}{p(B_r(\mathcal{O}))}(1 - p(B_r(\mathcal{O}))) \underset{r\to\infty}{\longrightarrow} 0,\end{aligned} \tag{5.132}$$

where we use that $\log p(B_r(\mathcal{O})) \leq 0$ and $p(x)\log p(x) \leq 0$ for every x. As for the I-term,

$$I_E(p_r) = \frac{1}{p(B_r(\mathcal{O}))} \sum_{\{x,y\}\in E\colon\, x,y\in B_r(\mathcal{O})} \left(\sqrt{\tfrac{p(x)}{\deg(x)}} - \sqrt{\tfrac{p(y)}{\deg(y)}}\right)^2 \\ + \frac{1}{p(B_r(\mathcal{O}))} \sum_{\{x,y\}\in E\colon\, x\in B_r(\mathcal{O}),\, y\in B_r^{\mathrm{c}}} \frac{p(x)}{\deg(x)} \leq \frac{I_E(p)}{p(B_r(\mathcal{O}))} + \frac{p(B_{r-1}(\mathcal{O})^{\mathrm{c}})}{d_{\min}\, p(B_r(\mathcal{O}))}, \tag{5.133}$$

and therefore

$$I_E(p_r) - I_E(p) \leq \frac{I_E(p)}{p(B_r(\mathcal{O}))}(1 - p(B_r(\mathcal{O}))) + \frac{p(B_{r-1}(\mathcal{O})^{\mathrm{c}})}{d_{\min}\, p(B_r(\mathcal{O}))} \underset{r\to\infty}{\longrightarrow} 0. \tag{5.134}$$

§5.5.2 Identification of the minimiser

This section follows [12, Appendix A]. We adapt the techniques to the new I_E function defined in (5.7).

Lemma 5.5.1. *[Glue two] Let $G_i = (V_i, E_i)$, $i = 1, 2$, be two disjoint connected simple graphs, and let $x_i \in V_i$, $i = 1, 2$. Denote by G the union graph of G_1, G_2 with one extra edge between x_1 and x_2, i.e. $G = (V, E)$ with $V := V_1 \cup V_2$, $E := E_1 \cup E_2 \cup \{(x_1, x_2)\}$. Then*

$$\chi_G \geq \min\{\chi_{G_1}, \chi_{G_2}\}. \tag{5.135}$$

Proof. We follow [12, Lemma A.3]. Given $p \in \mathcal{P}(V)$, let $a_i = p(V_i)$, $i = 1, 2$, and define $p_i \in \mathcal{P}(V_i)$ by putting

$$p_i(x) := \begin{cases} \frac{1}{a_i} p(x) 1\!\!1_{V_i}(x) & \text{if } a_i > 0, \\ 1\!\!1_{x_i}(x) & \text{otherwise.} \end{cases} \tag{5.136}$$

Straightforward manipulations show that

$$I_E(p) = \sum_{i=1}^2 a_i I_{E_i}(p_i) + \left(\sqrt{\tfrac{p(x_1)}{\deg(x_1)}} - \sqrt{\tfrac{p(x_2)}{\deg(x_2)}}\right)^2, \qquad J_V(p) = \sum_{i=1}^2 \left[a_i J_{V_i}(p_i) - a_i \log a_i\right], \tag{5.137}$$

and so

$$I_E(p) + \varrho J_V(p) \geq \sum_{i=1}^2 a_i \Big[I_{E_i}(p_i) + \varrho J_{V_i}(p_i)\Big] \geq \min\{\chi_{G_1}, \chi_{G_2}\}. \tag{5.138}$$

The proof is completed by taking the infimum over $p \in \mathcal{P}(V)$. □

Below it will be useful to define, for $x \in V$,

$$\chi_G^{x,b} = \inf_{\substack{p\in\mathcal{P}(V),\\ p(x)=b}} [I_E(p) + \varrho J_V(p)], \tag{5.139}$$

Chapter 5

i.e. a version of χ_G with "boundary condition" b at x. It is clear that $\chi_G^{x,b} \geq \chi_G$. Next we glue several graphs together and derive representations and estimates for the corresponding χ. For $k \in \mathbb{N}$, let $G_i = (V_i, E_i)$, $1 \leq i \leq k$, be a collection of disjoint graphs. Let x be a point not belonging to $\bigcup_{i=1}^k V_i$. For a fixed choice $y_i \in V_i$, $1 \leq i \leq k$, we denote by $\overline{G}_k = (\overline{V}_k, \overline{E}_k)$ the graph obtained by adding an edge from each $y_1, \dots, y_k$ to x, i.e. $\overline{V}_k = V_1 \cup \dots \cup V_k \cup \{x\}$ and $\overline{E}_k = E_1 \cup \dots \cup E_k \cup \{(y_1, x), \dots, (y_k, x)\}$.

Lemma 5.5.2. *[Glue many plus vertex] For any $\varrho > 0$, any $k \in \mathbb{N}$, and any $G_i = (V_i, E_i)$, $y_i \in V_i$, $1 \leq i \leq k$,*

$$\chi_{\overline{G}_k} = \inf_{\substack{0 \leq c_i \leq a_i \leq 1, \\ a_1 + \dots + a_k \leq 1}} \Big\{ \sum_{i=1}^k a_i \left(\chi_{G_i}^{y_i, c_i/a_i} - \varrho \log a_i \right) + \sum_{i=1}^k \left(\sqrt{\frac{c_i}{\deg(y_i)}} - \sqrt{\frac{1 - \sum_{i=1}^k a_i}{\deg(x)}} \right)^2 - \varrho \Big(1 - \sum_{i=1}^k a_i\Big) \log \Big(1 - \sum_{i=1}^k a_i\Big) \Big\}. \tag{5.140}$$

Proof. We follow [12, Lemma A.4]. The claim follows from straightforward manipulations with (5.7). ☐

Lemma 5.5.2 leads to the following comparison lemma. For $j \in \mathbb{N}$, let

$$(G_i^j, y_i^j) = \begin{cases} (G_i, y_i) & \text{if } i < j, \\ (G_{i+1}, y_{i+1}) & \text{if } i \geq j, \end{cases} \tag{5.141}$$

i.e. $(G_i^j)_{i \in \mathbb{N}}$ is the sequence $(G_i)_{i \in \mathbb{N}}$ with the j-th graph omitted. Let $\overline{G}_k^j$ be the analogue of $\overline{G}_k$ obtained from G_i^j, $1 \leq i \leq k$, $i \neq j$, instead of G_i, $1 \leq i \leq k$.

Lemma 5.5.3. *[Comparison] For any $\varrho > 0$ and any $k \in \mathbb{N}$,*

$$\begin{aligned} \chi_{\overline{G}_{k+1}} = \inf_{1 \leq j \leq k+1} \inf_{0 \leq c \leq u \leq \frac{1}{k+1}} \inf_{\substack{0 \leq c_i \leq a_i \leq 1, \\ a_1 + \dots + a_k \leq 1}} \Bigg\{ & (1-u) \Big[\sum_{i=1}^k a_i \big(\chi_{G_{\sigma_j(i)}}^{y_{\sigma_j(i)}, c_i/a_i} - \varrho \log a_i \big) \\ & + \sum_{i=1}^k \left(\sqrt{\frac{c_i}{\deg(y_i)}} - \sqrt{\frac{1 - \sum_{i=1}^k a_i}{\deg(x)}} \right)^2 - \varrho \Big(1 - \sum_{i=1}^k a_i\Big) \log \Big(1 - \sum_{i=1}^k a_i\Big) \Big] \\ & + u \chi_{G_j}^{y_j, c/u} + \left(\sqrt{\frac{c}{\deg(y_j)}} - \sqrt{\frac{(1-u)\Big(1 - \sum_{i=1}^k a_i\Big)}{\deg(x)}} \right)^2 \\ & - \varrho \left[u \log u + (1-u) \log(1-u) \right] \Bigg\}. \end{aligned} \tag{5.142}$$

Moreover,

$$\chi_{\overline{G}_{k+1}} \geq \inf_{1\leq j\leq k+1} \inf_{0\leq u\leq \frac{1}{k+1}} \Bigg\{(1-u)\chi_{\overline{G}_k^j} + \inf_{v\in[0,1]} \Big\{u\chi_{G_j}^{(y_j,v)} + \mathbb{1}_{\{u(1+v)\geq 1\}}\Big[\sqrt{\tfrac{vu}{\deg(y_j)}} - \sqrt{\tfrac{1-u}{\deg(x)}}\,\Big]^2\Big\} - \varrho\,[u\log u + (1-u)\log(1-u)]\Bigg\}. \tag{5.143}$$

Proof. See [12, Lemma A.5]. The argument still applies with the definition of I_E given in (5.7). □

Lemma 5.5.4. *[Propagation of lower bounds] If $\varrho > 0$, $M \in \mathbb{R}$, $C > 0$ and $k \in \mathbb{N}$ satisfy $\varrho \geq C/\log(k+1)$ and*

$$\inf_{1\leq j\leq k+1} \chi_{\overline{G}_k^j} \geq M, \qquad \inf_{1\leq j\leq k+1} \inf_{v\in[0,1]} \chi_{G_j}^{y_j,v} \geq M - C, \tag{5.144}$$

then $\chi_{\overline{G}_{k+1}} \geq M$.

Proof. See [12, Lemma A.6]. The proof carries over directly since I_E does not appear. □

The above results will be applied in the next section to minimise χ over families of trees with minimum degrees.

Trees with minimum degrees

Fix $d \in \mathbb{N}$. Let $\mathring{\mathcal{T}}_d$ be an infinite tree rooted at $\mathcal{O}$ such that the degree of $\mathcal{O}$ equals $d-1$ and the degree of every other vertex in $\mathring{\mathcal{T}}_d$ is d. Let $\mathring{\mathscr{T}}_d^0 = \{\mathring{\mathcal{T}}_d\}$ and, recursively, let $\mathring{\mathscr{T}}_d^{n+1}$ denote the set of all trees obtained from a tree in $\mathring{\mathscr{T}}_d^n$ and a disjoint copy of $\mathring{\mathcal{T}}_d$ by adding an edge between a vertex of the former and the root of the latter. Write $\mathring{\mathscr{T}}_d = \bigcup_{n\in\mathbb{N}_0} \mathring{\mathscr{T}}_d^n$. Assume that all trees in $\mathring{\mathscr{T}}_d$ are rooted at $\mathcal{O}$.

Recall that $\mathcal{T}_d$ is the infinite regular d-tree. Observe that $\mathcal{T}_d$ is obtained from $(\mathring{\mathcal{T}}_d, \mathcal{O})$ and a disjoint copy $(\mathring{\mathcal{T}}_d', \mathcal{O}')$ by adding one edge between $\mathcal{O}$ and $\mathcal{O}'$. Consider $\mathcal{T}_d$ to be rooted at $\mathcal{O}$. Let $\mathscr{T}_d^0 = \{\mathcal{T}_d\}$ and, recursively, let $\mathscr{T}_d^{n+1}$ denote the set of all trees obtained from a tree in $\mathscr{T}_d^n$ and a disjoint copy of $\mathring{\mathcal{T}}_d$ by adding an edge between a vertex of the former and the root of the latter. Write $\mathscr{T}_d = \bigcup_{n\in\mathbb{N}_0} \mathscr{T}_d^n$, and still consider all trees in $\mathscr{T}_d$ to be rooted at $\mathcal{O}$. Note that $\mathscr{T}_d^n$ contains precisely those trees of $\mathring{\mathscr{T}}_d^{n+1}$ that have $\mathcal{T}_d$ as a subgraph rooted at $\mathcal{O}$. In particular, $\mathscr{T}_d^n \subset \mathring{\mathscr{T}}_d^{n+1}$ and $\mathscr{T}_d \subset \mathring{\mathscr{T}}_d$.

Our objective is to prove the following.

Proposition 5.5.5. *[Minimal tree is optimal] If $\varrho \geq \frac{1}{d\log(d+1)}$, then*

$$\chi_{\mathcal{T}_d}(\varrho) = \min_{T\in\mathscr{T}_d} \chi_T(\varrho).$$

For the proof of Proposition 5.5.5, we will need the following.

Lemma 5.5.6. *[Minimal half-tree is optimal] For all $\varrho \in (0,\infty)$,*

$$\chi_{\mathring{\mathcal{T}}_d}(\varrho) = \min_{T \in \mathring{\mathscr{T}}_d} \chi_T(\varrho).$$

Proof. See [12, Lemma A.8]. The proof carries over directly since I_E does not appear. □

Lemma 5.5.7. *[A priori bounds] For any $d \in \mathbb{N}$ and any $\varrho \in (0,\infty)$,*

$$\chi_{\mathring{\mathcal{T}}_d}(\varrho) \le \chi_{\mathcal{T}_d}(\varrho) \le \chi_{\mathring{\mathcal{T}}_d}(\varrho) + \frac{1}{d}. \tag{5.145}$$

Proof. We follow [12, Lemma A.9]. The first inequality follows from Lemma 5.5.6. For the second inequality, note that $\mathcal{T}_d$ contains as subgraph a copy of $\mathring{\mathcal{T}}_d$, and restrict the minimum in (5.8) to $p \in \mathcal{P}(\mathring{\mathcal{T}}_d)$. □

Proof of Proposition 5.5.5. We follow [12, Proposition A.7]. Fix $\varrho \ge \frac{1}{d\log(d+1)}$. It will be enough to show that

$$\chi_{\mathcal{T}_d} = \min_{T \in \mathscr{T}_d^n} \chi_T, \qquad n \in \mathbb{N}_0. \tag{5.146}$$

We will prove this by induction in n. The case $n = 0$ is trivial. Assume that, for some $n_0 \ge 0$, (5.146) holds for all $n \le n_0$. Let $T \in \mathscr{T}_d^{n_0+1}$. Then there exists a vertex x of T with degree $k+1 \ge d+1$. Let $y_1, \dots, y_{k+1}$ be set of neighbours of x in T. When we remove the edge between y_j and x, we obtain two connected trees; call G_j the one containing y_j, and $\overline{G}_k^j$ the other one. With this notation, T may be identified with $\overline{G}_{k+1}$.

Now, for each j, the rooted tree (G_j, y_j) is isomorphic (in the obvious sense) to a tree in $\mathring{\mathscr{T}}_d^{\ell_j}$, where $\ell_j \in \mathbb{N}_0$ satisfy $\ell_1 + \dots + \ell_{k+1} \le n_0$, while $\overline{G}_k^j$ belongs to $\mathscr{T}_d^{(n_j)}$ for some $n_j \le n_0$. Therefore, by the induction hypothesis,

$$\chi_{\overline{G}_k^j} \ge \chi_{\mathcal{T}_d}, \tag{5.147}$$

while, by (5.139), Lemma 5.5.6 and Lemma 5.5.7,

$$\inf_{v\in[0,1]} \chi_{G_j}^{(y_j,v)} \ge \chi_{G_j} \ge \chi_{\mathring{\mathcal{T}}_d} \ge \chi_{\mathcal{T}_d} - \frac{1}{d}. \tag{5.148}$$

Thus, by Lemma 5.5.4 applied with $M = \chi_{\mathcal{T}_d}$ and $C = \frac{1}{d}$,

$$\chi_T = \chi_{\bar{G}_{k+1}} \ge \chi_{\mathcal{T}_d}, \tag{5.149}$$

which completes the induction step. □

Proof of Theorem 5.1.2. We follow [12, Theorem 1.2]. First note that, since $\mathcal{T}_{d_{\min}}$ has degrees in $\mathrm{supp}(D_g)$, $\widetilde{\chi}(\varrho) \le \chi_{\mathcal{T}_{d_{\min}}}(\varrho)$. For the opposite inequality, we proceed as follows. Fix an infinite tree T with degrees in $\mathrm{supp}(D_g)$, and root it at a vertex $\mathcal{Y}$. For

Chapter 5

$r \in \mathbb{N}$, let $\widetilde{T}_r$ be the tree obtained from $B_r(\mathcal{O}) = B_r^T(\mathcal{Y})$ by attaching to each vertex $x \in B_r(\mathcal{O})$ with $|x| = r$ a number $d_{\min} - 1$ of disjoint copies of $(\mathring{\mathcal{T}}_{d_{\min}}, \mathcal{O})$, i.e. adding edges between x and the corresponding roots. Then $\widetilde{T}_r \in \mathscr{T}_{d_{\min}}$ and, since $B_r(\mathcal{O})$ has more out-going edges in T than in $\widetilde{T}_r$, we may check using (5.128) that

$$\widehat{\chi}_{B_r}(\varrho; T) \geq \widehat{\chi}_{B_r}(\varrho; \widetilde{T}_r) \geq \chi_{\widetilde{T}_r}(\varrho) \geq \chi_{\mathcal{T}_{d_{\min}}}(\varrho). \tag{5.150}$$

Taking $r \to \infty$ and applying Proposition 5.2.1, we obtain $\chi_T(\varrho) \geq \chi_{\mathcal{T}_{d_{\min}}}(\varrho)$. Since T is arbitrary, the proof is complete. □

Chapter 5

Appendix of Part II

APPENDIX B

Appendix: Chapter 4

§B.1 Largest eigenvalue

We recall the Rayleigh-Ritz formula for the principal eigenvalue of the Anderson Hamiltonian. For $\Lambda \subset V$ and $q\colon V \to [-\infty, \infty)$, let $\lambda_\Lambda(q;G)$ denote the largest eigenvalue of the operator $\Delta_G + q$ in Λ with Dirichlet boundary conditions on $V \setminus \Lambda$, i.e.,

$$\lambda_\Lambda(q;G) := \sup\big\{\langle(\Delta_G+q)\phi,\phi\rangle_{\ell^2(V)}\colon\ \phi\in\mathbb{R}^V,\ \operatorname{supp}\phi\subset\Lambda,\ \|\phi\|_{\ell^2(V)}=1\big\}. \tag{B.1}$$

Lemma B.1.1. *[Spectral bounds]*

(1) *For any $\Gamma \subset \Lambda \subset V$,*

$$\max_{z\in\Gamma} q(z) - D_{\bar z} \le \lambda_\Gamma(q;G) \le \lambda_\Lambda(q;G) \le \max_{z\in\Lambda} q(z) \tag{B.2}$$

with $\bar z = \arg\max_{z\in\Gamma} q(z)$ and $D_{\bar z}$ the degree of $\bar z$.

(2) *The eigenfunction corresponding to $\lambda_\Lambda(q;G)$ can be taken to be non-negative.*

(3) *If q is real-valued and $\Gamma \subsetneq \Lambda$ is finite and connected in G, then the second inequality in* (B.2) *is strict and the eigenfunction corresponding to $\lambda_\Lambda(q;G)$ is strictly positive.*

Proof. Write

$$\begin{aligned}\langle(\Delta_G+q)\phi,\phi\rangle_{\ell^2(V)} &= \sum_{x\in\Lambda}\left[(\Delta_G\phi)(x)+q(x)\phi(x)\right]\phi(x)\\ &= \sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda:\\ \{x,y\}\in E_\Lambda}}[\phi(y)-\phi(x)]\phi(x) + \sum_{x\in\Lambda} q(x)\phi(x)^2\\ &= -\tfrac12\sum_{\substack{x,y\in\Lambda:\\ \{x,y\}\in E_\Lambda}}[\phi(x)-\phi(y)]^2 + \sum_{x\in\Lambda} q(x)\phi(x)^2,\end{aligned} \tag{B.3}$$

where the first sum in the last line runs over all ordered pairs (x,y) with $(x,y)\neq(y,x)$, which gives rise to the factor $\frac12$. The upper bound in (B.2) follows from the estimate

$$\langle(\Delta_G+q)\phi,\phi\rangle \le \sum_{x\in\Lambda} q(x)\phi(x)^2 \le \max_{z\in\Lambda} q(z)\sum_{x\in\Lambda}\phi(x)^2 = \max_{z\in\Lambda} q(z). \tag{B.4}$$

To get the lower bound in (B.2), we use the fact that λ_Λ is non-decreasing in q. Hence, replacing $q(z)$ by $-\infty$ for every $z \neq \bar{z}$ and taking as test function $\phi = \bar{\phi} = \delta_{\bar{z}}$, we get from (B.3) that

$$\begin{aligned} \lambda_\Lambda(q;G) &\geq -\tfrac{1}{2} \sum_{\substack{x,y\in\Lambda:\\ \{x,y\}\in E_\Lambda}} \left[\bar{\phi}(x) - \bar{\phi}(y)\right]^2 + \sum_{x\in\Lambda} q(x)\bar{\phi}(x)^2 \\ &= -\tfrac{1}{2} \sum_{\substack{y\in\Lambda:\\ \{\bar{z},y\}\in E_\Lambda}} 1 + q(\bar{z}) = -D_{\bar{z}} + \max_{z\in\Lambda} q(z), \end{aligned} \tag{B.5}$$

which settles the claim in (1). The claims in (2) and (3) are standard. □

Inside $\mathcal{GW}$, fix a finite connected subset $\Lambda \subset V$, and let H_Λ denote the Anderson Hamiltonian in Λ with zero Dirichlet boundary conditions on $\Lambda^c = V\backslash\Lambda$ (i.e., the restriction of the operator $H_G = \Delta_G + \xi$ to the class of functions supported on Λ). For $y \in \Lambda$, let u^y_Λ be the solution of

$$\begin{array}{rcll} \partial_t u(x,t) &=& (H_\Lambda u)(x,t), & x\in\Lambda,\ t>0, \\ u(x,0) &=& \delta_y(x), & x\in\Lambda, \end{array} \tag{B.6}$$

and set $U^y_\Lambda(t) := \sum_{x\in\Lambda} u^y_\Lambda(x,t)$. The solution admits the Feynman-Kac representation

$$u^y_\Lambda(x,t) = \mathbb{E}_y\left[\exp\left\{\int_0^t \xi(X_s)\mathrm{d}s\right\} \mathbb{1}\{\tau_{\Lambda^c} > t, X_t = x\}\right], \tag{B.7}$$

where τ_{Λ^c} is the hitting time of Λ^c. It also admits the spectral representation

$$u^y_\Lambda(x,t) = \sum_{k=1}^{|\Lambda|} \mathrm{e}^{t\lambda^k_\Lambda} \phi^k_\Lambda(y)\phi^k_\Lambda(x), \tag{B.8}$$

where $\lambda^1_\Lambda \geq \lambda^2_\Lambda \geq \cdots \geq \lambda^{|\Lambda|}_\Lambda$ and $\phi^1_\Lambda, \phi^2_\Lambda, \ldots, \phi^{|\Lambda|}_\Lambda$ are, respectively, the eigenvalues and the corresponding orthonormal eigenfunctions of H_Λ. These two representations may be exploited to obtain bounds for one in terms of the other, as shown by the following lemma.

Lemma B.1.2. *[Bounds on the solution] For any $y \in \Lambda$ and any $t > 0$,*

$$\begin{aligned} \mathrm{e}^{t\lambda^1_\Lambda}\phi^1_\Lambda(y)^2 &\leq \mathbb{E}_y\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{\Lambda^c}>t, X_t=y\}}\right] \\ &\leq \mathbb{E}_y\left[\mathrm{e}^{\int_0^t \xi(X_s)\mathrm{d}s} \mathbb{1}_{\{\tau_{\Lambda^c}>t\}}\right] \leq \mathrm{e}^{t\lambda^1_\Lambda}|\Lambda|^{1/2}. \end{aligned} \tag{B.9}$$

Proof. The first and third inequalities follow from (B.7)–(B.8) after a suitable application of Parseval's identity. The second inequality is elementary. □

Bibliography

[1] P. W. Anderson. Absence of diffusion in certain random lattices. *Phys. Rev.*, 109:1492–1505, Mar 1958.

[2] E. Archer and A. Pein. Parabolic Anderson model on critical Galton-Watson trees in a Pareto environment. *Stochastic Process. Appl.*, 159:34–100, 2023.

[3] A. Astrauskas. From extreme values of i.i.d. random fields to extreme eigenvalues of finite-volume Anderson Hamiltonian. *Probab. Surv.*, 13:156–244, 2016.

[4] K. B. Athreya. Large deviation rates for branching processes. I. Single type case. *Ann. Appl. Probab.*, 4(3):779–790, 1994.

[5] L. Avena, O. Gün, and M. Hesse. The parabolic Anderson model on the hypercube. *Stochastic Process. Appl.*, 130(6):3369–3393, 2020.

[6] J. D. Biggins and N. H. Bingham. Large deviations in the supercritical branching process. *Adv. in Appl. Probab.*, 25(4):757–772, 1993.

[7] M. Biskup and W. König. Long-time tails in the parabolic Anderson model with bounded potential. *Ann. Probab.*, 29(2):636–682, 2001.

[8] M. Biskup and W. König. Eigenvalue order statistics for random Schrödinger operators with doubly-exponential tails. *Comm. Math. Phys.*, 341(1):179–218, 2016.

[9] M. Biskup, W. König, and R. S. dos Santos. Mass concentration and aging in the parabolic Anderson model with doubly-exponential tails. *Probab. Theory Related Fields*, 171(1-2):251–331, 2018.

[10] R. A. Carmona and S. A. Molchanov. *Parabolic Anderson problem and intermittency*, volume 108 of *Mem. Amer. Math. Soc.* American Mathematical Society, Providence, RI, 1994.

[11] F. den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs.* American Mathematical Society, Providence, RI, 2000.

[12] F. den Hollander, W. König, and R. S. dos Santos. The parabolic Anderson model on a Galton-Watson tree. In *In and out of equilibrium 3. Celebrating Vladas Sidoravicius*, volume 77 of *Progr. Probab.*, pages 591–635. Birkhäuser/Springer, 2021.

[13] F. den Hollander and D. Wang. The parabolic Anderson model on a Galton-Watson tree revisited. *J. Stat. Phys.*, 189(1):Paper 8, 2022.

[14] F. den Hollander and D. Wang. Annealed parabolic Anderson model on a regular tree. *Markov Process. Related Fields*, 30:105–147, 2024.

[15] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28:1–47; 279–301, 1975.

[16] K. Fleischmann and S. A. Molchanov. Exact asymptotics in a mean field model with random potential. *Probab. Theory Related Fields*, 86(2):239–251, 1990.

[17] J. Gärtner, F. den Hollander, and G. Maillard. Intermittency on catalysts. In *Trends in stochastic analysis*, volume 353 of *London Math. Soc. Lecture Note Ser.*, pages 235–248. Cambridge Univ. Press, Cambridge, 2009.

[18] J. Gärtner and W. König. Moment asymptotics for the continuous parabolic Anderson model. *Ann. Appl. Probab.*, 10(1):192–217, 2000.

[19] J. Gärtner, W. König, and S. Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. *Ann. Probab.*, 35(2):439–499, 2007.

[20] J. Gärtner, W. König, and S. A. Molchanov. Almost sure asymptotics for the continuous parabolic Anderson model. *Probab. Theory Related Fields*, 118(4):547–573, 2000.

[21] G. Grüninger and W. König. Potential confinement property of the parabolic Anderson model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):840–863, 2009.

[22] J. Gärtner. On large deviations from an invariant measure. *Theory Probab. Appl.*, 22:24–39, 1977.

[23] J. Gärtner and F. den Hollander. Correlation structure of intermittency in the parabolic Anderson model. *Probab. Theory Related Fields*, 114(1):1–54, 1999.

[24] J. Gärtner and S. Molchanov. Parabolic problems for the Anderson model I. intermittency and related problems. *Comm. Math. Phys.*, 132(3):613–655, 1990.

[25] J. Gärtner and S. Molchanov. Parabolic problems for the Anderson model. II. second-order asymptotics and structure of high peaks. *Probab. Theory Related Fields*, 111(1):15–55, 1998.

[26] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, 1952.

[27] J. Jost and M. P. Joy. Spectral properties and synchronization in coupled map lattices. *Phys. Rev. E (3)*, 65(1):016201, 9, 2002.

[28] W. König, H. Lacoin, P. Mörters, and N. Sidorova. A two cities theorem for the parabolic Anderson model. *Ann. Probab.*, 37(1):347–392, 2009.

[29] W. König. *The parabolic Anderson model*. Pathways in Mathematics. Birkhäuser/Springer, 2016.

[30] R. Lyons and Y. Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016.

[31] M. Mariani and L. Zambotti. Large deviations for the empirical measure of heavy-tailed Markov renewal processes. *Adv. in Appl. Probab.*, 48(3):648–671, 2016.

[32] B. Schmidt. On a semilinear variational problem. *ESAIM Control Optim. Calc. Var.*, 17(1):86–101, 2011.

[33] F. Spitzer. *Principles of random walk*, volume 34 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Heidelberg, 1976.

[34] A.-S. Sznitman. *Brownian motion, obstacles and random media*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[35] R. van der Hofstad, W. König, and P. Mörters. The universality classes in the parabolic Anderson model. *Comm. Math. Phys.*, 267(2):307–353, 2006.

[36] D. Wang. The parabolic Anderson model on a Galton-Watson tree with normalised Laplacian, 2023. Preprint, arXiv:2310.05602.

Samenvatting

Dit proefschrift bestudeert het parabolisch Anderson model (PAM), d.w.z. het Cauchy-probleem voor de warmtevergelijking met een toevallige potentiaal. Het PAM is een wiskundig model dat beschrijft hoe massa (= materie of energie) door een medium stroomt in aanwezigheid van een toevallige potentiaal, dat als een veld van 'bronnen' en 'putten' fungeert. Zonder deze potentiaal zorgt diffusie ervoor dat de massa zich gelijkmatig over het medium verdeelt. De potentiaal verandert dit gedrag op een ingrijpende wijze, aangezien de massa de neiging heeft zich rond de 'bronnen' te concentreren en rond de 'putten' te verdwijnen. Het doel van dit proefschrift is om het lange-termijn-gedrag van de massa te begrijpen: de asymptotische groeisnelheid ervan, evenals hoe en waar de meeste massa zich concentreert.

De vergelijkingen die het PAM aandrijven worden opgelost door een functionaal van een toevalswandeling, die door de bekende Feynman-Kac-representatie gegeven wordt. Deze representatie is het startpunt van de analyse van het PAM. Het PAM is uitgebreid bestudeerd op reguliere roosters en is daar goed begrepen. Het is bekend dat de staart van de verdeling van de potentiaal het asymptotische gedrag van de massa bepaalt, zowel voor de groeisnelheid als voor de locaties met hoge concentraties van de massa. Het reguliere rooster is echter niet altijd een geschikt model, en we zoeken derhalve naar uitbreidingen voor algemene grafen. Er is weinig over het PAM op algemene grafen bekend en de literatuur is uiterst schaars. Dit proefschrift is een bijdrage aan dit ontluikende onderzoeksgebied. Omdat grafen waarvan de knooppunten begrensde graden hebben vaak door bomen kunnen worden benader, is het bestuderen van het PAM op een boom de natuurlijke eerste stap. Dit proefschrift is dan ook vooral gewijd aan het bestuderen van het PAM op toevallige bomen met potentialen met een dubbel-exponentiële staartverdeling.

De studie van het PAM is verdeeld in twee gevallen: 'quenched' en 'annealed'. Dat wil zeggen, bijna zeker ten opzichte van de potentiaal, respectievelijk, gemiddeld over de potentiaal. Dit proefschrift bestrijkt beide en is daarom in twee delen verdeeld: deel I is gewijd aan het 'quenched' model en deel II aan het 'annealed' model.

Deel I bestaat uit de Hoofdstukken 2 en 3 en behandelt het 'quenched' model. Hoofdstuk 2 onderzoekt het PAM op een reguliere boom en leidt de eerste twee termen van de asymptotische groei van de massa af. Net als bij het rooster vereist het argument het afleiden van asymptotisch op elkaar aansluitende boven- en ondergrenzen. De grootste uitdaging is het omgaan met de exponentiële groei van de graafgrootte, die niet in het rooster optreedt. Een nieuwe techniek voor het vouwen van de paden van de toevalswandeling wordt gepresenteerd en het grote-afwijkingen-gedrag ervan wordt geanalyseerd om de bovengrens at te leiden. Bovendien wordt het concen-

tratiegedrag van de massa bepaald in termen van minimalisatie van een variationele formule. Het argument berust op een herschikkingsongelijkheid, die het gebrek aan translatie-invariantie overwint.

Hoofdstuk 3 breidt Hoofdstuk 2 uit naar een Galton-Watson-boom met een eindige periodiciteit – een generalisatie van de reguliere boom. Dit vereist zorgvuldig navigeren door de niet-homogeniteiten van de Galton-Watson-boom en het omgaan met graden die toevalsvariabelen zijn. Dit hoofdstuk is de cruciale stap in het begrijpen van het PAM op de Galton-Watson-boom, die veel toevallige graafmodellen waarvan de graden niet al te groot zijn goed benadert.

Deel II bestaat uit de Hoofdstukken 4 en 5 en behandelt het 'annealed' model. Hoofdstuk 4 onderzoekt het PAM op een Galton-Watson-boom. De asymptotische groeisnelheid van de totale massa werd afgeleid in eerder werk van den Hollander, König en dos Santos, onder de restrictieve aanname dat de graden van de boom begrensd zijn. Hoofdstuk 4 breidt hun analyse uit naar bomen met onbegrensde graden, en identificeert de zwakste condities op de gradenverdeling waaronder hun technieken nog steeds bruikbaar zijn. Dit wordt gedaan door de grote graden in sub-bomen onder controle te houden. Het bestaan en de uniciteit van de Feynman-Kac-representatie op exponentieel groeiende grafen wordt ook in dit hoofdstuk behandeld.

Hoofdstuk 5 beschouwt opnieuw het PAM op een Galton-Watson-boom, en breidt het eerdere werk van den Hollander, König en dos Santos uit naar de versie van het model waarin de Laplaciaan door de graad-genormaliseerde Laplaciaan vervangen wordt. Het komt erop neer dat de toevalswandeling in de Feynman-Kac-representatie een sprongsnelheid van 1 heeft, in plaats van de graad van het knooppunt waarop het zich bevindt. Deze normalisatie zorgt ervoor dat de Laplaciaan niet langer symmetrisch is, wat in andere spectrale eigenschappen resulteert. We laten zien dat de asymptotische groeisnelheid van de totale massa hetzelfde is, terwijl de variationele formule die de verdeling van de massa beschrijft anders is. Ook wordt de zwakste conditie op de gradenverdeling geïdentificeerd.

Summary

This thesis investigates the parabolic Anderson model (PAM), which is the Cauchy problem for the heat equation with random potential. The PAM is a mathematical model that describes how mass (i.e. matter or energy) flows in a medium in the presence of a random potential, which acts as a field of sources and sinks. Without the potential, diffusion causes the mass to be evenly distributed across the medium. However, the potential vastly changes this behaviour as mass tends to concentrate around the sources and deplete around the sinks. The goal of the thesis is to understand the long-term behaviour of the mass: its asymptotic growth rate, as well as how and where it concentrates.

The equation that drives the PAM is solved by a functional of a random walk that is given by the well-known Feynman-Kac representation. This representation is the starting point for the analysis of the PAM. The PAM has been extensively studied on regular lattices and is well understood there. It is known that the upper tail of the distribution of the potential fully determines the asymptotic behaviour of the mass for both the growth rate and the locations of high concentration. However, the lattice is not always a suitable model and we look for extensions to random graphs. Very little is known for general graphs and the literature is extremely sparse. The present thesis is a contribution to this developing area. Because sparse random graphs can often be approximated by trees, the natural first step is to consider the PAM on a tree. In particular, this thesis is devoted to studying the PAM on random trees with potentials having a double-exponential distribution.

The study of the PAM is naturally divided into two cases: quenched and annealed, i.e. almost surely with respect to the potential and averaged over the potential, respectively. The thesis covers both and is therefore divided into two parts: Part I is dedicated to the annealed model and Part II to the quenched model.

Part I consists of Chapters 2 and 3 and considers the annealed model. Chapter 2 investigates the PAM on a regular tree and derives the first two terms of the asymptotic growth of the mass. As was the case on the lattice, the argument requires finding asymptotically matching upper and lower bounds. The main challenge is to deal with the exponential growth of the graph size, which is not present in the lattice. A novel technique of folding random walk paths is devised and its large deviation behaviour is analysed to achieve the upper bound. Furthermore, the concentration behaviour of the mass is determined and is given in terms of a minimiser of a variational formula. The argument relies on a rearrangement inequality that overcomes the lack of translation invariance.

Chapter 3 extends Chapter 2 to a Galton-Watson tree with large periodicity – a generalisation of the regular tree. This requires carefully navigating the non-homogeneity of the Galton-Watson tree and dealing with degrees that are random. This chapter is the crucial step in understanding the PAM on the regular Galton-Watson tree, which approximates many sparse random graph models.

Part II consists of Chapters 4 and 5 and considers the quenched model. Chapter 4 investigates the PAM on a Galton-Watson tree. The asymptotic growth rate of the total mass was derived in previous work by den Hollander, König and dos Santos under the restrictive assumption that the degrees are bounded. Chapter 4 extends their analysis to trees with unbounded degrees, and identifies the weakest condition on the degree distribution under which their arguments still hold. This is done by uniformly controlling the appearance of large degrees in subtrees. The existence and uniqueness of the Feynman-Kac representation on exponentially growing graphs is also shown in this chapter.

Chapter 5 again considers the PAM on a Galton-Watson tree, and extends the previous work by den Hollander, König and dos Santos to the version of model in which the Laplacian is replaced by with the degree-normalised Laplacian. This amounts to the random walk in the Feynman-Kac representation having jump rate equal to 1 instead of the degree of the vertex it resides on. The normalisation causes the Laplacian to no longer be symmetric, which results in different spectral properties. We find that the asymptotic growth rate of the total mass is the same, while the variational formula describing the distribution of mass is different. The weakest condition required on the degree distribution is also identified.

Acknowledgements

This thesis is nothing short of a collection of miracles that have happened to me, some of the people I mention here:

My supervisor Frank den Hollander, who will forever be the tallest giant whose shoulders I have had the privilege of standing on.

My collaborators Renato dos Santos and Wolfgang König, whom I regretfully could not visit due to COVID. No one is more knowledgeable about the PAM than them.

Márton Balázs, who got and kept me interested in Probability during my time at Bristol, and who guided me towards starting this PhD.

My parents, with their unconditional love that I constantly forget to appreciate. 谁言寸草心, 报得三春晖。

"*I thank you, I am not of many words, but I thank you.*"
Much Ado About Nothing, I.i.154-155
Shakespeare

Curriculum Vitae

Daoyi Wang was born on November 1st 1996 in Beijing, China. He moved to The Netherlands in 1997 where he attended the *British School in the Netherlands*. He obtained his GCSEs in 2013, and his A-Levels in Mathematics, Further Mathematics, Physics and Biology in 2015.

Afterwards, he attended the University of Bristol to pursue a Master of Science in Mathematics where he decided to focus on probability and statistics. He wrote his Master's project titled 'Ergodicity of Stochastic Processes and Markov Chain Central Limit Theorem' under the supervision of Prof. Márton Balázs, for which he was awarded the Howell Peregrine prize for the best undergraduate project. He graduated with first class honours in June 2019.

In October 2019, Daoyi moved to Leiden to start his PhD under the supervision of Prof. Frank den Hollander at Mathematical Institute of the University of Leiden. He was part of the NETWORKS Gravitation Programme funded by the Dutch Ministry of Education, Culture and Science. During this time, he has lectured at the Leiden University College, supervised students for the Leiden PRE-university programme and assisted in various Bachelor and Master units.

Publications

1. F. den Hollander and D. Wang. Annealed parabolic Anderson model on a regular tree. *Markov Process. Related Fields*, 30:105–147, 2024.
2. F. den Hollander and D. Wang. The parabolic Anderson model on a Galton-Watson tree revisited. *J. Stat. Phys.*, 189(1):Paper 8, 2022.
3. D. Wang. The parabolic Anderson model on a Galton-Watson tree with normalised Laplacian, 2023. Preprint, arXiv:2310.05602, submitted.