Budapest University of Technology and Economics
Institute of Mathematics
Department of Stochastics

Asymptotic Behavior of Markov chains and Networks:
Fluctuations, mixing properties and modeling hierarchical networks

PhD Thesis

Júlia Komjáthy

Supervisor: Dr. Balázs Márton
Advisor: Prof. Károly Simon

2012

Contents

Acknowledgements

First and foremost I would like to express my sincere gratitude to both of my supervisors Márton Balázs and Károly Simon. I am thankful for their help and advices, without their support I would not have been able to write my thesis. I am greatly indebted to Károly Simon for encouragement in mathematical matters in the last two years.
I would like to thank Bálint Tóth for introducing me the world of probability during my undergraduate years, and later for his help and support. Trying to reach his high standards encouraged me all over the past years.
I am grateful to my high school math teacher Lázár Kertes. His inspiring way of teaching mathematics influenced me to become a professional mathematician.
I would like to thank Yuval Peres for his guidance during my numerous visits to Microsoft Research.
I thank my coauthors, Jason Miller, Timo Seppäläinen for our fruitful and enlightening collaboration.
I thank the people in the Department of Stochastics for the inspiring atmosphere for research.
Last but not least I would like to thank the constant support and the devotion of my family and friends.

Introduction

In this thesis we investigate three different, interesting topics of probability: Markov chain mixing, properties of network models and interacting particle systems. We briefly summarize the content of the different chapters here.

The mixing time of Markov chains have been an active research topic in the past three decades, as large databases required better and better understanding of the finite time behavior of nonstationary Markov chains. In particular, consider a sequence of Markov chains on larger and larger state spaces, and set a finite threshold (1/4, say). The question is, how long the chain has to be run to reach the threshold distance from stationary measure as a function of the size of the state space. If the distance is measured in l_1 or in the uniform metric then we call this time the *total-variation* and *the uniform mixing time*, respectively. For precise definitions see (1.2.12) and (1.2.2). A more algebraic way of measuring the correlation decay - i.e. how fast the chain forgets about its starting state - is to investigate the eigenvalues and thus the spectral gap of the chain. The *relaxation time* (1.2.3) is the reciprocal of the spectral gap, again considered as a function of the size of the state space along the sequence.

In the first chapter of the thesis we consider the mixing times of lamplighter groups. Based on the joint paper with Jason Miller and Yuval Peres [77], in the first section we investigate the uniform mixing time of the lamplighter walks. Heuristically, a lamplighter walk can be visualized as follows: imagine a random walker walking on the graph G, where there are lamps attached to each vertex of the graph. Wherever the walker is, he randomizes the state of the lamp. More precisely, suppose that G is a finite, connected graph and X is a lazy random walk on G. The lamplighter chain $X^\diamond$ associated with X is the random walk on the wreath product $\mathcal{G}^\diamond = \mathbf{Z}_2 \wr G$, the graph whose vertices consist of pairs $(\underline{f}, x)$ where $\underline{f}$ is a labeling of the vertices of G by elements of $\mathbf{Z}_2 = \{0, 1\}$ and x is a vertex in G. For example, see Figure 1.1, where G is a 5×5 grid and the $0-1$ lamps are illustrated as blue and yellow, respectively. In each step, $X^\diamond$ moves from a configuration $(\underline{f}, x)$ by updating x to y using the transition rule of X and then sampling both $f(x)$ and $f(y)$ according to the uniform distribution on $\mathbf{Z}_2$; $f(z)$ for $z \neq x, y$ remains unchanged. We give matching upper and lower bounds

on the uniform mixing time of $X^\diamond$ provided G satisfies mild hypotheses. In particular, when G is the hypercube $\mathbf{Z}_2^d$, we show that the uniform mixing time of $X^\diamond$ is of order $d2^d$. More generally, we show that when G is a torus $\mathbf{Z}_n^d$ for $d \geq 3$, the uniform mixing time of $X^\diamond$ is of order dn^d uniformly in n and d. A critical ingredient for our proof is a concentration estimate for the local time of random walk in a subset of vertices and Green's function estimates. This work closes the gap on the estimates on the uniform mixing time in [96].

Then, in the second section of the first chapter we move to larger lamp spaces and consider the wreath product $H \wr G$, the graph whose vertices consist of pairs $(\underline{f}, x)$ where $\underline{f} = (f(v))_{v \in V(G)}$ is a labeling of the vertices of G by elements of H and x is a vertex in G. Heuristically, the generalized lamplighter walk can be visualized as follows: imagine the random walker again, walking on the base graph G. At each vertex of the graph G there is an identical, complicated machine, with possible states represented by the graph H. While moving one step on the base graph G, the walker changes one step according to the transition rule of the machine on his departure and also on his arrival vertex, respectively. See Figure 1.3 where the base graph G is a torus and the lamp graphs are cycles. More precisely, the generalized lamplighter chain $X^\diamond$ associated with the Markov chains X on G and Z on H is the random walk on the wreath product $H \wr G$: In each step, $X^\diamond$ moves from a configuration $(\underline{f}, x)$ by updating x to y using the transition rule of X and then independently updating both $f(x)$ and $f(y)$ according to the transition probabilities on H; $f(z)$ for $z \neq x, y$ remains unchanged. We estimate the total variation mixing time of $X^\diamond$ in terms of the parameters of H and G. Further, we show that the relaxation time of $X^\diamond$ is the same order as the maximal expected hitting time of G plus $|G|$ times the relaxation time of H. Various methods are used in this chapter to prove the bounds, including strong stationary times, Dirichlet form techniques, distinguishing set methods and mean optimal stopping rules. The area of considering general lamp graphs is new in the literature, and is still our ongoing research joint with Yuval Peres.

In the second chapter we switch to consider mathematical properties of graph and network models. Random graphs are in the main stream of research interest since the late 50s, starting with the seminal random graph model introduced by Erdős and Rényi [61]. A wide spectrum of literature investigates graph models with a fixed number of vertices (i.e some generalization of the Erdős-Rényi (ER) graphs), we refer the reader to the books of [71] or [35] as an introduction. In the last two decades there have been a considerable amount of attention paid to the study of complex networks like the World Wide Web, social networks, or biological networks. The properties of these networks turned out to be way too different from models based on some variation of the ER graph. This resulted in the construction of numerous new, more dynamical and growing network models, see e.g. [27], [35],

[39], [59], [76]. Most of them use a version of preferential attachment and are of probabilistic nature. A different approach was initiated by Barabási, Ravasz, and Vicsek [25] based on the observation that real network often obey some hierarchical structure. They introduced deterministic network models generated by a method which is common in constructing fractals. Their model both exhibits hierarchical structure and the degree sequence obeys power law decay, and with a slight modification of the model [100] the local clustering coefficient is also decaying as in real networks. A similar, fractal based deterministic model were introduced by Zhang, Comellas, Fertin and Rong [109], and called the high-dimensional Apollonian network. The graph is generated from the cylinder sets of the fractal of the Apollonian circle packing or the Sierpinsky carpet.

Motivated by the hierarchical network model of E. Ravasz, A.-L. Barabási and T. Vicsek [25], we introduce deterministic scale-free networks derived from a graph directed self-similar fractal Λ. Starting from an arbitrary initial bipartite graph G on N vertices, we construct a hierarchical sequence of deterministic graphs G_n. The embedding of the adjacency matrix of the graph sequence G_n is carried out in the most straightforward way: A vertex with code $\underline{x} = (x_1 \dots x_n) \in G_n$ is identified with the corresponding N-adic interval $I_{\underline{x}}$, and Λ_n is the union of those $N^{-n} \times N^{-n}$ squares $I_{\underline{x}} \times I_{\underline{y}}$ for which the vertices $\underline{x}, \underline{y}$ are connected by an edge in G_n. Λ_n turns out to be a nested sequence of compact sets, which can be considered as the n-th approximation of a graph directed self-similar fractal Λ on the plane, see Figure 2.1(c). We discuss connection between the graph theoretical properties of G_n and properties of the limiting fractal Λ. In particular, we express the power law exponent of the degree distribution with the ratio of the Hausdorff dimensions of some slices of Λ (Theorem 2.3.6).

Further, we verify that our model captures some of the most important features of many real networks: we show that the degree distribution of vertices has a power law decay and thus the model obeys the scale free property. We also prove that the diameter is the logarithm of the size of the system. There are no triangles in G_n. Hence, in order to model the clustering properties of many real networks, we need to extend the set of edges of our graph sequence to destroy the bipartite property. Motivated by [100], we add some additional edges to G_1 to obtain the (no longer bipartite) graph $\widehat{G}_1$. Then we build up the graph sequence $\widehat{G}_n$ in a similar manner than it was done for G_n, and show that the average local clustering coefficient of $\widehat{G}_n$ does not depend on the size and the local clustering coefficient of a node with degree k is of order $1/k$.

The third chapter investigates fluctuations of one dimensional interacting particle systems. The motivation comes mainly from statistical mechanics: the surface growth or the fluctuations of a stream wants to be understood on the microscopical level. For a good and thorough introduction to the field we refer to the two books by Liggett [86, 87].

We consider Markov processes that describe the motion of particles and antiparticles in the one dimensional integer lattice $\mathbf{Z}$, or equivalently, growth of a surface by depositing or removing individual bricks of unit length and height over $\mathbf{Z}$. We examine the net particle current seen by an observer moving at the characteristic speed of the process. The characteristic speed is the speed at which perturbations travel in the system and can be determined e.g. via the hydrodynamic limit. The process is assumed to be asymmetric (i.e. the rates of removal and deposition, or the particle to jump to the right and to the left in the particle-picture, are not the same) and in one of its extremal stationary distributions, which is a product measure parameterized by the density of particles ϱ. We set a system of conditions called *microscopic concavity* or *convexity* and prove that under these conditions, the net particle current across the characteristic has variance of order $t^{2/3}$. The net particle current counts the number of particles that pass the observer from left to right minus the number that pass from right to left during time interval $(0, t]$. As a bi-product, we also obtain Law of Large Numbers for the second class particle and Central Limit Theorem for the particle current.

Earlier proofs of $t^{1/3}$fluctuations e.g. [8, 74, 75, 99, 44] have been quite rigid in the sense that they work only for particular cases of the models where special combinatorial properties emerge as if through some fortuitous coincidences. There is basically no room for perturbing the rules of the process. By contrast, the proof given here works for the whole class of processes. The hypothesis of microscopic concavity that is required is certainly nontrivial. But it does not seem to rigidly exclude all but a handful of the processes in the broad class.

The chapter is based on two papers, both of them joint with Márton Balázs and Timo Seppäläinen. The first one is [17], which describes microscopic concavity, the general proof under this system of conditions and investigates totally asymmetric zero range processes with a concave jump rate function whose slope decreases geometrically, and may be eventually constant. Section 3.7 is based on the paper [16], where we show that the strategy works for the exponential bricklayers process, a process obeying convex flux function.

Chapter 1

Mixing times of random walks on wreath product graphs

1.1 Introduction

In 1906 Andrey Markov introduced the random processes that would later be named after him. The classical theory of Markov chains was mostly concerned with long-time behavior of Markov chains: The goal is to understand the stationary distribution and the rate of convergence of a fixed chain. Many introductory books on stochatic processes include an introduction to Markov chains, see for example the book by Lawler [80].

However, in the past three decades, a different asymptotical analysis has emerged: in theoretical computer science, physics and biology, the growing interest in large state spaces required a better understanding of the finite time behavior of Markov chains in terms of the size of the state space. Thus, some target distance from the stationary measure in some metric on the space of measures is usually prescribed and the question is to determine the required number of steps to reach this distance as the size of the state space increases. *Mixing time* refers to this notion. For a comprehensive overview of Markov Chain mixing we refer the reader to the indispensable book [49] by Aldous and Fill or [83] by Levin, Peres and Wilmer as our main references. See also the books by Häggstöm [70], Jerrum [73], or the recent survey by Montenegro and Tetali [94].

To understand the behavior of Markov chain sequences, different notions of distances on probability measures and mixing times emerged, each capturing some different aspect or property of the chain. Aldous [3] introduced random stopping times achieving equilibrium measure. They were studied more by Lovász, Winkler [89, 90]. To find the relation between different notions of mixing is a challenging problem, see [3] and the recent papers

connecting hitting times to mixing times and stopping rules by Sousi and Peres [98] and independently by Oliveira [95], or blanket times and cover times to the maxima of Gaussian free fields by Ding, Lee and Peres [56].

A further understanding of the Markov Chain sequence is to see whether there is any "concentration" of the mixing time, i.e., if the ratio between mixing times up to different thresholds has a limit. Such behavior is called *cutoff*. In general, it was conjectured that the total variation mixing time has a cutoff as long as the necessary condition, that its ratio with the relaxation time is is tending to infinity, holds. However, the conjecture fails to be true in the highest generality, see [83, Example 18.7]. Cutoffs are proven recently for random walks on random regular graphs by Lubetzky and Sly [91] and for birth and death chains by Ding, Lubetzky and Peres [58]. The cutoff phenomenon is discussed further in Chen and Saloff-Coste [46] and Diaconis and Saloff-Coste [55].

In this chapter we are interested in the mixing properties of random walks on wreath product graphs. The intuitive representation of the walk is the following: A lamplighter or an engineer is doing simple random walk on the vertices of a *base graph* G. Further, to each vertex $v \in G$ there is an identical lamp or machine attached, and each of the machines is in some state $f(v)$. Then, as the lamplighter walks along the base graph, he can make changes in the state of the machines touched, according to the transition probabilities of the machines. If the machines are just on-off lamps, we get the well-known lamplighter problem, but if the machines (the lamp-graphs) have some more complicated structure, possibly even growing together with the size of the base, then we are in the setting of generalized lamplighter walks. If the underlying graphs H and G are Cayley-graphs of groups generated by some finite number of generators, then the graph $H \wr G$ is the graph of the wreath product of the two groups. This relates our work to understand the behavior of random walk on groups, analyzed by many authors, we refer the reader for references [1] Aldous.

More precisely, suppose that G and H are finite, connected graphs, G regular, X is a lazy random walk on G and Z is a reversible ergodic Markov chain on H. The generalized lamplighter chain $X^\diamond$ associated with X and Z is the random walk on the wreath product $H \wr G$, the graph whose vertices consist of pairs $(\underline{f}, x)$ where $\underline{f} = (f(v))_{v \in V(G)}$ is a labeling of the vertices of G by elements of H and x is a vertex in G. In each step, $X^\diamond$ moves from a configuration $(\underline{f}, x)$ by updating x to y using the transition rule of X and then independently updating both $f(x)$ and $f(y)$ according to the transition probabilities on H; $f(z)$ for $z \neq x, y$ remains unchanged. In Section 1.2, based on the joint paper with Miller an Peres [77] we give matching upper bound for the mixing time in the uniform metric of $X^\diamond$ on $\mathbf{Z}_2 \wr G$ up to universal constants in terms of the parameters of G to the lower bound given in [96, Theorem 1.4] by Peres and Revelle. Then in Section 1.3 we give bounds on the total variation mixing time and estimate the relaxation

time of $H \wr G$ for general H and G up to universal constants.

Before we proceed to the particular models, we will mention some other work on mixing times for lamplighter chains. The mixing time of $\mathbf{Z}_2 \wr G$ was first studied by Häggström and Jonasson in [69] in the case of the complete graph K_n and the one-dimensional cycle $\mathbf{Z}_n$. Their work implies a total variation cutoff with threshold $\frac{1}{2}t_{\mathrm{cov}}(K_n)$ in the former case and that there is no cutoff in the latter. Here, $t_{\mathrm{cov}}(G)$ for a graph G denotes the expected number of steps required by lazy random walk to visit every vertex in G. The connection between $t_{\mathrm{mix}}(\mathbf{Z}_2 \wr G)$ and $t_{\mathrm{cov}}(G)$ is explored further in [96], in addition to developing the relationship between the relaxation time of $\mathbf{Z}_2 \wr G$ and the maximal expected hitting time $t_{\mathrm{hit}}(G)$, and the mixing time in the uniform metric $t_{\mathrm{u}}(\mathbf{Z}_2 \wr G)$ and $\mathbf{E}[2^{|\mathcal{U}(t)|}]$, the exponential moment of the not-yet-touched vertices of the base graph G. The results of [96] include a proof of total variation cutoff for $\mathbf{Z}_2 \wr \mathbf{Z}_n^2$ with threshold $t_{\mathrm{cov}}(\mathbf{Z}_n^2)$. In [93], it is shown that $t_{\mathrm{mix}}(\mathbf{Z}_2 \wr \mathbf{Z}_n^d) \sim \frac{1}{2}t_{\mathrm{cov}}(\mathbf{Z}_n^d)$ when $d \geq 3$ and more generally that $t_{\mathrm{mix}}(\mathbf{Z}_2 \wr G_n) \sim \frac{1}{2}t_{\mathrm{cov}}(G_n)$ whenever (G_n) is a sequence of graphs satisfying some uniform local transience assumptions.

The mixing time of $X^\diamond = (\underline{F}, X)$ on $\mathbf{Z}_2 \wr G$ is typically dominated by the first coordinate $\underline{F}$ since the amount of time it takes for X to mix is negligible compared to that required by $X^\diamond$. We can sample from $\underline{F}(t)$ by:

1. sampling the range $\mathcal{C}(t)$ of lazy random walk run for time t, then
2. marking the vertices of $\mathcal{C}(t)$ by i.i.d. fair coin flips.

Determining the mixing time of $X^\diamond$ is thus typically equivalent to computing the threshold t where the corresponding marking becomes indistinguishable from a uniform marking of $V(G)$ by i.i.d. fair coin flips. This in turn can be viewed as a statistical test for the uniformity of the not covered set $\mathcal{U}(t)$ of X — if $\mathcal{U}(t)$ exhibits any sort of non-trivial systematic geometric structure then $X^\diamond(t)$ is not mixed. This connects Section 1.2 work to the literature on the geometric structure of the last visited points by random walk [51, 50, 43, 93].

Moving towards larger lamp spaces, if the base is the complete graph K_n and $|H_n| = o(n)$ one can determine the order of mixing time from [83, Theorem 20.7], since in this case the lamplighter chain is a product chain on $\prod_{i=1}^n H_n$. Levi [84] investigated random walks on wreath products when $H \neq \mathbf{Z}_2$. In particular, he determined the order of the mixing time of $K_{n^\lambda} \wr K_n$, $0 \leq \lambda \leq 1$, and he also had upper and lower bounds for the case $H_d \wr \mathbf{Z}_n$, i.e. H is the d-dimensional hypercube and the base is a cycle of length n. However, the bounds failed to match for general d and n.

Similarly as the mixing time of $H_n = \mathbf{Z}_2$ is closely related to the cover time of the base graph, larger lamp graphs give more information on the local time structure of the base graph G. This relates Section 1.3 to the literature on blanket time (when all the local times of vertices are within a constant factor of each other) [28, 56, 107].

1.2 Uniform mixing time for Random Walk on Lamplighter Graphs

1.2.1 The model

Let us start with the precise description of the model adapted to the setting $H = \mathbf{Z}_2$. Suppose that $\mathcal{G}$ is a finite graph with vertices $V(\mathcal{G})$ and edges $E(\mathcal{G})$, respectively. Let $\mathcal{X}(\mathcal{G}) = \{f \colon V(\mathcal{G}) \to \mathbf{Z}_2\}$ be the set of markings of $V(\mathcal{G})$ by elements of $\mathbf{Z}_2$. The wreath product $\mathcal{G}^\diamond = \mathbf{Z}_2 \wr \mathcal{G}$ is the graph whose vertices are pairs (f, x) where $f \in \mathcal{X}(\mathcal{G})$ and $x \in V(\mathcal{G})$. There is an edge between (f, x) and (g, y) if and only if $\{x, y\} \in E(\mathcal{G})$ and $f(z) = g(z)$ for all $z \notin \{x, y\}$. Suppose that P is a transition matrix for a Markov chain on $\mathcal{G}$. The lamplighter walk $X^\diamond$ (with respect to the transition matrix P) is the Markov chain on $\mathcal{G}^\diamond$ which moves from a configuration (f, x) by

1. picking y adjacent to x in $\mathcal{G}$ according to P, then
2. updating each of the values of $f(x)$ and $f(y)$ independently according to the uniform measure on $\mathbf{Z}_2$.

The lamp states at all other vertices in $\mathcal{G}$ remain fixed. It is easy to see that if P is ergodic and reversible with stationary distribution π_P then the unique stationary distribution of $X^\diamond$ is the product measure

$$\pi\big((f, x)\big) = \pi_P(x) 2^{-|\mathcal{G}|},$$

and $X^\diamond$ is itself reversible. In this section, we will be concerned with the special case that P is the transition matrix for the *lazy random walk* on $\mathcal{G}$ in order to avoid issues of periodicity. That is, P is given by

$$P(x, y) = \begin{cases} \frac{1}{2} \text{ if } x = y, \\ \frac{1}{2d(x)} \text{ if } \{x, y\} \in E(\mathcal{G}), \end{cases} \tag{1.2.1}$$

for $x, y \in V(\mathcal{G})$ and where $d(x)$ is the degree of x.

1.2.2 Main Results

Let P be the transition kernel for lazy random walk on a finite, connected graph $\mathcal{G}$ with stationary distribution π. The *ϵ-uniform mixing time* of $\mathcal{G}$ is given by

$$t_{\mathrm{u}}(\epsilon, \mathcal{G}) = \min \left\{ t \geq 0 : \max_{x,y \in V(\mathcal{G})} \left| \frac{P^t(x, y) - \pi(y)}{\pi(y)} \right| \leq \epsilon \right\}. \tag{1.2.2}$$

Throughout, we let $t_{\mathrm{u}}(\mathcal{G}) = t_{\mathrm{u}}((2e)^{-1}, \mathcal{G})$. The main result of the article [77] is a general theorem which gives matching upper and lower bounds of $t_{\mathrm{u}}(\mathcal{G}^\diamond)$ provided $\mathcal{G}$ satisfies several mild hypotheses. One important special case of

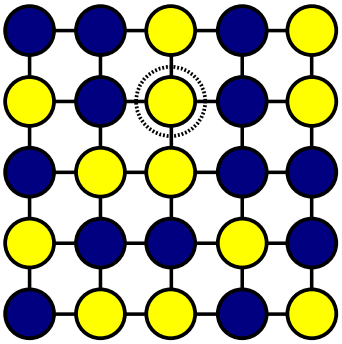


Figure 1.1: A typical configuration of the lamplighter over a 5×5 planar grid. The colors indicate the state of the lamps and the dashed circle gives the position of the lamplighter.

this result is the hypercube $\mathbf{Z}_2^d$ and, more generally, tori $\mathbf{Z}_n^d$ for $d \geq 3$. These examples are sufficiently important that we state them as our first theorem in the chapter.

Theorem 1.2.1. *There exists constants $C_1, C_2 > 0$ such that*

$$C_1 \leq \frac{t_u((\mathbf{Z}_2^d)^\diamond)}{d2^d} \leq C_2 \text{ for all } d.$$

More generally,

$$C_1 \leq \frac{t_u((\mathbf{Z}_n^d)^\diamond)}{dn^{d+2}} \leq C_2 \text{ for all } n \geq 2 \text{ and } d \geq 3.$$

Prior to this work, the best known bound [96] for $t_{\mathrm{u}}((\mathbf{Z}_2^d)^\diamond)$ was

$$C_1 d2^d \leq t_{\mathrm{u}}((\mathbf{Z}_2^d)^\diamond) \leq C_2 \log(d) d2^d$$

for $C_1, C_2 > 0$.

In order to state our general result, we first need to review some basic terminology from the theory of Markov chains. The *relaxation time* of P is

$$t_{\mathrm{rel}}(\mathcal{G}) = \frac{1}{1 - \lambda_0} \tag{1.2.3}$$

where λ_0 is the second largest eigenvalue of P. The *maximal hitting time* of P is

$$t_{\mathrm{hit}}(\mathcal{G}) = \max_{x,y \in V(\mathcal{G})} \mathbf{E}_x[\tau_y], \tag{1.2.4}$$

where τ_y denotes the first time t that $X(t) = y$ and $\mathbf{E}_x$ stands for the expectation under the law in which $X(0) = x$. The Green's function $G(x, y)$ for P is

$$G(x,y) = \mathbf{E}_x \left[\sum_{t=0}^{t_{\mathrm{u}}(\mathcal{G})} \mathbf{1}_{\{X(t)=y\}} \right] = \sum_{t=0}^{t_{\mathrm{u}}(\mathcal{G})} P^t(x,y), \tag{1.2.5}$$

i.e. the expected amount of time X spends at y up to time t_{u} given $X(0) = x$. For each $1 \leq n \leq |\mathcal{G}|$, we let

$$G^*(n) = \max_{\substack{S \subseteq V(\mathcal{G}) \\ |S|=n}} \max_{z \in S} \sum_{y \in S} G(z, y). \tag{1.2.6}$$

This is the maximal expected time X spends in a set $S \subseteq V(\mathcal{G})$ of size n before the uniform mixing time. This quantity is related to the hitting time of subsets of $V(\mathcal{G})$. Finally, recall that $\mathcal{G}$ is said to be vertex transitive if for every $x, y \in V(\mathcal{G})$ there exists an automorphism φ of $\mathcal{G}$ with $\varphi(x) = y$. Our main result requires the following hypothesis.

Assumption 1.2.2. *$\mathcal{G}$ is a finite, connected, vertex transitive graph and X is a lazy random walk on $\mathcal{G}$. There exists constants $K_1, K_2, K_3 > 0$ such that*

1. $t_{hit}(\mathcal{G}) \leq K_1|\mathcal{G}|$,
2. $2K_2(5/2)^{K_2}(G(x,y))^{K_2} \leq \exp\left(-\frac{t_u(\mathcal{G})}{t_{rel}(\mathcal{G})}\right)$,
3. $G^*(n^*) \leq K_3(t_{rel}(\mathcal{G}) + \log|\mathcal{G}|)/(\log n^*)$

where $n^* = 4K_2 t_u(\mathcal{G})/G(x,y)$ for $x, y \in V(\mathcal{G})$ adjacent.

The general theorem is:

Theorem 1.2.3. *Let $\mathcal{G}$ be any graph satisfying Assumption 1.2.2. There exists constants C_1, C_2 depending only on K_1, K_2, K_3 such that*

$$C_1 \leq \frac{t_u(\mathcal{G}^\diamond)}{|\mathcal{G}|\big(t_{rel}(\mathcal{G}) + \log|\mathcal{G}|\big)} \leq C_2 \tag{1.2.7}$$

The lower bound is proved in [96, Theorem 1.4]. The proof of the upper bound is based on the observation from [96] that the uniform distance to stationarity can be related to $\mathbf{E}[2^{|\mathcal{U}(t)|}]$ where $\mathcal{U}(t)$ is the set of vertices in $\mathcal{G}$ which have not been visited by X by time t. Indeed, suppose that f is any initial configuration of lamps, let $f(t)$ be the state of the lamps at time t, and let g be an arbitrary lamp configuration. Let W be the set of vertices where $f \neq g$. Let $\mathcal{C}(t) = V(\mathcal{G}) \setminus \mathcal{U}(t)$ be the set of vertices which have been visited by X by time t. With $\mathbf{P}_{(f,x)}$ the probability under which $X^\diamond(0) = (f, x)$, we have that

$$\mathbf{P}_{(f,x)}[f(t) = g | \mathcal{C}(t)] = 2^{-|\mathcal{C}(t)|}\mathbf{1}_{\{W \subseteq \mathcal{C}(t)\}}.$$

Since the probability of the configuration g under the uniform measure is $2^{-|\mathcal{G}|}$, we therefore have

$$\frac{\mathbf{P}_{(f,x)}[f(t) = g]}{2^{-|\mathcal{G}|}} = \mathbf{E}_{(f,x)}\big[2^{|\mathcal{U}(t)|}\mathbf{1}_{\{W \subseteq \mathcal{C}(t)\}}\big]. \tag{1.2.8}$$

The right hand side is clearly bounded from above by $\mathbf{E}[2^{|\mathcal{U}(t)|}]$ (the initial lamp configuration and position of the lamplighter no longer matters). On the other hand, we can bound (1.2.8) from below by

$$\mathbf{P}_{(f,x)}[W \subseteq \mathcal{C}(t)] \geq \mathbf{P}[|\mathcal{U}(t)| = 0] \geq 1 - (\mathbf{E}\big[2^{|\mathcal{U}(t)|}\big] - 1).$$

Consequently, to bound $t_{\mathrm{u}}(\epsilon, \mathcal{G}^{\diamond})$ it suffices to compute

$$\min\{t \geq 0 : \mathbf{E}[2^{|\mathcal{U}(t)|}] \leq 1 + \epsilon\} \tag{1.2.9}$$

since the amount of time it requires for X to subsequently uniformly mix after this time is negligible.

In order to establish (1.2.9), we will need to perform a rather careful analysis of the process by which $\mathcal{U}(t)$ is decimated by X. The key idea is to break the process of coverage into two different regimes, depending on the size of $\mathcal{U}(t)$. The main ingredient to handle the case when $\mathcal{U}(t)$ is large is the following concentration estimate of the local time

$$\mathcal{L}_S(t) = \sum_{s=0}^{t} \mathbf{1}_{\{X(s) \in S\}}$$

for X in $S \subseteq V(\mathcal{G})$.

Proposition 1.2.4. *Let λ_0 be the second largest eigenvalue of P. Assume $\lambda_0 \geq \frac{1}{2}$ and fix $S \subseteq V(\mathcal{G})$. For $C_0 = 1/50$, we have that*

$$\mathbf{P}_{\pi}\left[\mathcal{L}_S(t) \leq t\frac{\pi(S)}{2}\right] \leq \exp\left(-C_0 t \frac{\pi(S)}{t_{rel}(\mathcal{G})}\right). \tag{1.2.10}$$

Proposition 1.2.4 is a corollary of [82, Theorem 1]; we consider this sufficiently important that we state it here. By invoking Green's function estimates, we are then able to show that the local time is not concentrated on a small subset of S. The case when $\mathcal{U}(t)$ is small is handled via an estimate (Lemma 1.2.9) of the hitting time $\tau_S = \min\{t \geq 0 : X(t) \in S\}$ of S.

Earlier results on the uniform mixing time.

Suppose that μ, ν are probability measures on a finite measure space. The *total variation distance* between μ, ν is given by

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A} |\mu(A) - \nu(A)| = \frac{1}{2}\sum_{x} |\mu(x) - \nu(x)|. \tag{1.2.11}$$

The *ϵ-total variation mixing time* of P is

$$t_{\mathrm{mix}}(\epsilon, \mathcal{G}) = \min\left\{t \geq 0 : \max_{x \in V(\mathcal{G})} \|P^t(x, \cdot) - \pi\|_{\mathrm{TV}} \leq \epsilon\right\}. \tag{1.2.12}$$

Let $t_{\text{mix}}(\mathcal{G}) = t_{\text{mix}}((2e)^{-1}, \mathcal{G})$. It was proved [96, Theorem 1.4] by Peres and Revelle that if $\mathcal{G}$ is a regular graph such that $t_{\text{hit}}(\mathcal{G}) \leq K|\mathcal{G}|$, there exists constants C_1, C_2 depending only on K such that

$$C_1|\mathcal{G}|(t_{\text{rel}}(\mathcal{G}) + \log|\mathcal{G}|) \leq t_{\text{u}}(\mathcal{G}^{\diamond}) \leq C_2|\mathcal{G}|(t_{\text{mix}}(\mathcal{G}) + \log|\mathcal{G}|).$$

These bounds fail to match in general. For example, for the hypercube $\mathbf{Z}_2^d$, $t_{\text{rel}}(\mathbf{Z}_2^d) = \Theta(d)$ [83, Example 12.15] while $t_{\text{mix}}(\mathbf{Z}_2^d) = \Theta(d \log d)$ [83, Theorem 18.3]. Theorem 1.2.3 says that the lower from [96, Theorem 1.4] is sharp.

Outline

The remainder of this section is structured as follows. In Section 1.2.3 we will collect a number of estimates regarding the amount of X spends in and requires to cover sets of vertices in $\mathcal{G}$ of various sizes. Then, in Section 1.2.4, we will complete the proof of Theorem 1.2.3. Finally, in Section 1.2.5, we will give the proof of Theorem 1.2.1 by checking the hypotheses of Theorem 1.2.3.

1.2.3 Coverage Estimates

Throughout, we assume that $\mathcal{G}$ is a finite, connected, vertex transitive graph and X is lazy random walk on $\mathcal{G}$ with transition matrix P and stationary measure π. For $S \subseteq V(\mathcal{G})$, we let $\mathcal{C}_S(t)$ be the set of vertices in S visited by X by time t and let $\mathcal{U}_S(t) = S \setminus \mathcal{C}_S(t)$ be the subset of S which X has not visited by time t. We let $\mathcal{C}(t) = \mathcal{C}_{V(\mathcal{G})}(t)$ and $\mathcal{U}(t) = \mathcal{U}_{V(\mathcal{G})}(t)$. We will use $\mathbf{P}_x, \mathbf{E}_x$ to denote the probability measure and expectation under which $X(0) = x$. Likewise, we let $\mathbf{P}_\pi, \mathbf{E}_\pi$ correspond to the case that X is initialized at stationarity. The purpose of this section is to develop a number of estimates which will be useful for determining the amount of time required by X in order to cover subsets S of $V(\mathcal{G})$. We consider two different regimes depending on the size of S. If S is large, we will estimate the amount of time it takes for X to visit $t_{\text{u}}(\mathcal{G})$ distinct vertices in S. If S is small, we will estimate the amount of time it takes for X to visit $1/2$ of the vertices in S.

Large Sets

In this subsection, we will prove that the amount of time it takes for X to visit $t_{\text{u}}(\mathcal{G})$ distinct elements of a large set of vertices $S \subseteq V(\mathcal{G})$ is stochastically dominated by a geometric random variable whose parameter depends on $t_{\text{u}}(\mathcal{G})/t_{\text{rel}}(\mathcal{G})$. The main result is:

Proposition 1.2.5. *Assume X satisfies part (2) of Assumption 1.2.2 with constants K_2, K_3. Let $S \subseteq V(\mathcal{G})$ consist of at least $2K_2 t_u(\mathcal{G})/G(x,y)$ elements for $x, y \in V(\mathcal{G})$ adjacent and let*

$$t = \frac{2(K_2+2)t_u(\mathcal{G})}{\pi(S)}.$$

There exists a universal constant $C > 0$ such that for every $x \in V(\mathcal{G})$, we have that

$$\mathbf{P}_x\left[\mathcal{C}_S(t) \leq t_u(\mathcal{G})\right] \leq \exp\left(-C\frac{t_u(\mathcal{G})}{t_{rel}(\mathcal{G})}\right).$$

Recall that

$$\mathcal{L}_S(t) = \sum_{s=0}^{t} \mathbf{1}_{\{X(s)\in S\}}$$

is the amount of time that X spends in S up to time t. The proof consists of several steps. The first is Proposition 1.2.4, which we will deduce from [82, Theorem 1] shortly, which gives that the probability $\mathcal{L}_S(t)$ is less than $1/2$ its mean is exponentially small in t. Once we show that $\mathcal{L}_S(t)$ is large with high probability, in order to show that X visits many vertices in S, we need to rule out the possibility of X concentrating most of its local time in a small subset of S. This is accomplished in Lemma 1.2.6. We now proceed to the proof of Proposition 1.2.4.

Proof of Proposition 1.2.4. We rewrite the event

$$\left\{\mathcal{L}_S(t) \leq t\frac{\pi(S)}{2}\right\} = \left\{\sum_{s=0}^{t} f(X_s) > t\left(1-\pi(S)+\frac{\pi(S)}{2}\right)\right\} \tag{1.2.13}$$

where $f(x) = \mathbf{1}_{S^c}(x)$. Let $\epsilon = \pi(S)/2$ and $\mu = \mathbf{E}_\pi[f(X(t))] = 1-2\epsilon$. The case $\epsilon \geq 1/4$ follows immediately from [82, Equation 3] in the statement of [82, Theorem 1], so we will only consider the case $\epsilon \in (0, 1/4)$ here. Let $\overline{\mu} = 1-\mu = 2\epsilon$. For $x \in (0,1)$, let

$$I(x) = -x\log\left(\frac{\mu+\overline{\mu}\lambda_0}{1-2\overline{x}/(\sqrt{\Delta}+1)}\right) - \overline{x}\log\left(\frac{\overline{\mu}+\mu\lambda_0}{1-2x/(\sqrt{\Delta}+1)}\right)$$

where $\overline{x} = 1-x$ and

$$\Delta = 1 + \frac{4\lambda_0 x\overline{x}}{\mu\overline{\mu}(1-\lambda_0)^2}. \tag{1.2.14}$$

For $x \in [\mu, \mu+\epsilon] = [1-2\epsilon, 1-\epsilon]$, $\epsilon \in (0,1/4)$, and $\lambda_0 \geq 1/2$, we note that

$$\frac{1/2}{(1-\lambda_0)^2} \leq \Delta \leq \frac{20}{(1-\lambda_0)^2} \tag{1.2.15}$$

By [82, Theorem 1] and using the representation (1.2.13), we have that

$$\mathbf{P}_\pi\left[\mathcal{L}_S(t) \leq t\epsilon\right] \leq \exp(-I(\mu+\epsilon)t).$$

Since $I(\mu) = I'(\mu) = 0$ and $I''(x) = (\sqrt{\Delta}x\overline{x})^{-1}$ (see [82, Appendix B]), we can write

$$I(\mu+\epsilon) = \int_{\mu}^{\mu+\epsilon}\int_{\mu}^{x} \frac{1}{\sqrt{\Delta}y\overline{y}} dydx \tag{1.2.16}$$

where $\overline{y} = 1 - y$. Inserting the bounds from (1.2.15), we thus see that the right side of (1.2.16) admits the lower bound

$$\frac{1-\lambda_0}{\sqrt{20}}\int_{1-2\epsilon}^{1-\epsilon}\int_{1-2\epsilon}^{x}\frac{1}{2\epsilon}dydx \geq \frac{(1-\lambda_0)\epsilon}{16\sqrt{5}}$$

for all $\epsilon \in (0, 1/4)$ and $\lambda_0 \geq \frac{1}{2}$. □

As in the proof of Lemma 1.2.13, we couple X with a non-lazy random walk Y so that $X(t) = Y(N_t)$ where $N_t = \sum_{i=0}^{t}\xi_i$ and the (ξ_i) are iid with $\mathbf{P}[\xi_i = 0] = \mathbf{P}[\xi_i = 1] = \frac{1}{2}$ and are independent of Y. We let $\mathcal{L}_S^Y(t)$ denote the amount of time that $Y|_{[0,N_t]}$ spends in S (note that this differs slightly from the definition of $\mathcal{L}_x^Y(t)$ which appeared in Section 1.2.5). In other words, $\mathcal{L}_S^Y$ is the amount of that X spends in S by time t, not including those times where X does not move. The next lemma gives a lower bound on the probability that the number $\mathcal{C}_S(t)$ of distinct vertices X visits in a given set $S \subseteq V(\mathcal{G})$ by time t is proportional to $\mathcal{L}_S^Y(t)$. The lower bound for this probability will be given in terms of the Green's function $G(x,y)$ for X. Recall its definition from (1.2.5). Since X is a lazy random walk, we also have that

$$G(x,y) \leq G(x,x) \text{ for all } x, y \in V(\mathcal{G}). \tag{1.2.17}$$

This is a consequence of (1.2.38).

Lemma 1.2.6. *Fix $S \subseteq V(\mathcal{G})$. For each positive integer k, we have that*

$$\mathbf{P}_\pi\left[\mathcal{C}_S(t) \geq \frac{\mathcal{L}_S^Y(t) - t_u(\mathcal{G})}{k}\right] \geq 1 - \frac{t\pi(S)q^k(t)}{t_u(\mathcal{G})}, \tag{1.2.18}$$

where

$$q(t) = (G(x,y)-1) + (1+(2e)^{-1})\frac{t - t_u(\mathcal{G})}{|\mathcal{G}|}\mathbf{1}_{\{t > t_u(\mathcal{G})\}}. \tag{1.2.19}$$

and x is adjacent to y.

Proof. For $t \geq t_{\mathrm{u}}(\mathcal{G})$, we have $P^t(x,y) \leq (1+(2e)^{-1})\pi(y)$ by the definition of $t_{\mathrm{u}}(\mathcal{G})$. Thus by a union bound,

$$\mathbf{P}_x[\mathcal{L}_x^Y(t) > 1] \leq q(t).$$

Hence by the strong Markov property,

$$\mathbf{P}_x[\mathcal{L}_x^Y(t) > k] \leq q^k(t).$$

Observe

$$\mathbf{P}_\pi[\tau_x = s] \le \mathbf{P}_\pi[X_s = x] \le \pi(x). \tag{1.2.20}$$

Let

$$\mathcal{L}^Y_{S,k}(t) = \sum_{x\in S} \mathcal{L}^Y_x(t)\mathbf{1}_{\{\mathcal{L}^Y_x(t)>k\}}$$

be the total time that Y spends at points in S which it visits more than k times by time N_t. By (1.2.20), we have that

$$\mathbf{E}_\pi[\mathcal{L}^Y_{S,k}(t)] \le \sum_{x\in S}\sum_{s=0}^{t} \mathbf{P}_\pi[\tau_x = s]q^k(t) \le t\pi(S)q^k(t).$$

Applying Markov's inequality we have that

$$\mathbf{P}_\pi\left[\mathcal{L}^Y_{S,k}(t) \ge t_{\mathrm{u}}(\mathcal{G})\right] \le \frac{\mathbf{E}_\pi[\mathcal{L}^Y_{S,k}(t)]}{t_{\mathrm{u}}(\mathcal{G})} \le \frac{t\pi(S)q^k(t)}{t_{\mathrm{u}}(\mathcal{G})}$$

Observe

$$\mathcal{C}_S(t) = \sum_{x\in S} \mathbf{1}_{\{\mathcal{L}^Y_x(t)\ge 1\}} \ge \sum_{x\in S}\left(\mathbf{1}_{\{\mathcal{L}^Y_x(t)\ge 1\}} - \mathbf{1}_{\{\mathcal{L}^Y_x(t)>k\}}\right) \ge \frac{\mathcal{L}^Y_S(t) - \mathcal{L}^Y_{S,k}(t)}{k}.$$

Thus

$$\left\{\mathcal{L}^Y_{S,k}(t) < t_{\mathrm{u}}(\mathcal{G})\right\} \subseteq \left\{\mathcal{C}_S(t) \ge \frac{\mathcal{L}^Y_S(t) - t_{\mathrm{u}}(\mathcal{G})}{k}\right\}.$$

We arrive at

$$\mathbf{P}_\pi\left[\mathcal{C}_S(t) \ge \frac{\mathcal{L}^Y_S(t) - t_{\mathrm{u}}(\mathcal{G})}{k}\right] \ge 1 - \mathbf{P}_\pi\left[\mathcal{L}^Y_{S,k}(t) \ge t_{\mathrm{u}}(\mathcal{G})\right] \ge 1 - \frac{t\pi(S)q^k(t)}{t_{\mathrm{u}}(\mathcal{G})},$$

which completes the proof of the lemma. □

Proposition 1.2.4 gives a lower bound on the probability $\mathcal{L}_S(t)$ is proportionally lower than its expectation, Lemma 1.2.6 gives a lower bound on the probability X visits less than a positive fraction of $\mathcal{L}^Y_S(t) - t_{\mathrm{u}}(\mathcal{G})$ vertices in S by time t, and standard large deviations estimates bound the probability that $\mathcal{L}^Y_S(t)$ is proportionally smaller than $\mathcal{L}_S(t)$. By combining these two lemmas, we obtain the following result, which gives a lower bound on the rate at which X covers vertices in S.

Lemma 1.2.7. *Fix $S \subseteq V(\mathcal{G})$. Then*

$$\begin{aligned}\mathbf{P}_\pi&\left[\mathcal{C}_S(t) \le \frac{t\pi(S) - 8t_u(\mathcal{G})}{8k}\right] \\ &\le \exp\left(-C_0 t\frac{\pi(S)}{t_{rel}(\mathcal{G})}\right) + \exp\left(-\frac{1}{16}t\pi(S)\right) + \frac{t\pi(S)q^k(t)}{t_u(\mathcal{G})}\end{aligned} \tag{1.2.21}$$

where the constant C_0 is as in Proposition 1.2.4 and the function q is as in (1.2.19).

Proof. We trivially have that

$$\mathbf{P}_\pi\left[\mathcal{C}_S(t) \geq \frac{t\pi(S) - 8t_{\mathrm{u}}(\mathcal{G})}{8k}\right] \geq \mathbf{P}_\pi\left[\mathcal{C}_S(t) \geq \frac{t\pi(S) - 8t_{\mathrm{u}}(\mathcal{G})}{8k}, \mathcal{L}_S^Y(t) > \frac{t\pi(S)}{8}\right]$$
$$\geq \mathbf{P}_\pi\left[\mathcal{C}_S(t) > \frac{\mathcal{L}_S^Y(t) - t_{\mathrm{u}}(\mathcal{G})}{k},\ \mathcal{L}_S^Y(t) > \frac{t\pi(S)}{8}\right].$$

Therefore

$$\mathbf{P}_\pi\left[\mathcal{C}_S(t) \leq \frac{t\pi(S) - 8t_{\mathrm{u}}(\mathcal{G})}{8k}\right] \leq \mathbf{P}_\pi\left[\mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8}\right] + \mathbf{P}_\pi\left[\mathcal{C}_S(t) \leq \frac{\mathcal{L}_S^Y(t) - t_{\mathrm{u}}(\mathcal{G})}{k}\right]$$

We can bound the second term from above by Lemma 1.2.6. The first term is bounded from above by

$$\mathbf{P}_\pi\left[\mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8}\right] \leq \mathbf{P}_\pi\left[\mathcal{L}_S(t) \leq \frac{t\pi(S)}{2}\right] + \mathbf{P}_\pi\left[\mathcal{L}_S(t) > \frac{t\pi(S)}{2}, \mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8}\right].$$

We can bound the first term using Proposition 1.2.4. Conditionally on $\{\mathcal{L}_S(t) > \frac{t}{2}\pi(S)\}$, we note that $\{\mathcal{L}_S^Y(t) \leq \frac{t}{8}\pi(S)\}$ occurs if X stays in place for at least $\frac{3t}{8}\pi(S)$ time steps. Consequently, standard large deviations estimates imply that the second term above is bounded by $\exp(-\frac{1}{16}t\pi(S))$. □

We can now easily complete the proof of Proposition 1.2.5 by ignoring the first $t_{\mathrm{u}}(\mathcal{G})$ units of time in order to reduce to the stationary case, then apply Assumption 1.2.2 in order to match the error terms in Lemma 1.2.7.

Proof of Proposition 1.2.5. We first observe that

$$\mathbf{P}_x\left[\mathcal{C}_S(t) \leq t_{\mathrm{u}}(\mathcal{G})\right] \leq (1 + (2e)^{-1})\mathbf{P}_\pi[\mathcal{C}_S(t - t_{\mathrm{u}}(\mathcal{G})) \leq t_{\mathrm{u}}(\mathcal{G})].$$

With $\widetilde{t} = 2K_2 t_{\mathrm{u}}(\mathcal{G})/\pi(S)$ and using $|S| \geq 2K_2 t_{\mathrm{u}}(\mathcal{G})/(G(x,y) - 1)$ for $x, y \in V(\mathcal{G})$ adjacent, we see that

$$G(x,y) - 1 \leq q(\widetilde{t}) \leq \frac{5}{2}(G(x,y) - 1).$$

Combining this with part (2) of Assumption 1.2.2 implies

$$\frac{\widetilde{t}\pi(S)q^{K_2}(\widetilde{t})}{t_{\mathrm{u}}(\mathcal{G})} \leq 2K_2 q^{K_2}(\widetilde{t}) \leq \exp\left(-\frac{t_{\mathrm{u}}(\mathcal{G})}{t_{\mathrm{rel}}(\mathcal{G})}\right). \tag{1.2.22}$$

Applying Lemma 1.2.7 gives the result. □

Small Sets

We will now give an upper bound on the rate at which X covers $1/2$ the elements of a set of vertices $S \subseteq V(\mathcal{G})$, provided $|S|$ is sufficiently small.

Proposition 1.2.8. *Fix $S \subseteq V(\mathcal{G})$, let $s = |S|$, and assume that*

$$t_u(\mathcal{G}) \leq \frac{|\mathcal{G}|}{4s}.$$

There exists constants $C_2, C_3 > 0$ such that

$$\mathbf{P}_x \left[\mathcal{C}_S(C_2|\mathcal{G}|G^*(s)) \leq \frac{s}{2} \right] \leq \exp(-C_3 s)$$

for all $x \in V(\mathcal{G})$.

The main step in the proof of Proposition 1.2.8 is the next lemma, which gives an upper bound on the hitting time for S. Its proof is based on the following observation. Suppose that $S \subseteq V(\mathcal{G})$ and $\tau_S = \min\{t \geq 0 : X(t) \in S\}$. Let Z be a non-negative random variable with $Z\mathbf{1}_{\{\tau_S > t\}} = 0$ and $\mathbf{E}_x[Z\mathbf{1}_{\{\tau_S \leq t\}}] > 0$. Then we have that

$$\mathbf{P}_x[\tau_S \leq t] = \frac{\mathbf{E}_x[Z]}{\mathbf{E}_x[Z|\tau_S \leq t]}. \tag{1.2.23}$$

We will take Z to be the amount of time X spends in S.

Lemma 1.2.9. *Fix $S \subseteq V(\mathcal{G})$ and let $s = |S|$. Assume that*

$$t_u(\mathcal{G}) \leq \frac{|\mathcal{G}|}{2s}.$$

There exists a universal constant $\rho_0 > 0$ such that $x \in V(\mathcal{G})$ we have

$$\mathbf{P}_x \left[\tau_S \leq \frac{|\mathcal{G}|}{s} \right] \geq \frac{\rho_0}{G^*(s)}.$$

Proof. Let us introduce $E = \left\{ \tau_S \leq \frac{|\mathcal{G}|}{s} \right\}$. Observe that

$$\mathbf{P}_x[E] \geq \frac{\mathbf{E}_x \left[\mathcal{L}_S \left(\frac{|\mathcal{G}|}{s} \right) \right]}{\mathbf{E}_x \left[\mathcal{L}_S \left(\frac{|\mathcal{G}|}{s} \right) | E \right]}$$

We can bound the numerator from below as follows:

$$\begin{aligned} \mathbf{E}_x \left[\mathcal{L}_S \left(\frac{|\mathcal{G}|}{s} \right) \right] &\geq (1 - (2e)^{-1}) \mathbf{E}_\pi \left[\mathcal{L}_S \left(\frac{|\mathcal{G}|}{s} - t_{\mathrm{u}}(\mathcal{G}) \right) \right] \\ &\geq (1 - (2e)^{-1}) \pi(S) \left(\frac{|\mathcal{G}|}{s} - t_{\mathrm{u}}(\mathcal{G}) \right) \geq \frac{1}{4}. \end{aligned} \tag{1.2.24}$$

Let $\mathcal{L}_S(u,t) = \mathcal{L}_S(t) - \mathcal{L}_S(u-1)$ be the number of times in the set $\{u,\dots,t\}$ that X spends in S. Then we can express the denominator as the sum

$$\begin{aligned}&\mathbf{E}_x\left[\mathcal{L}_S\left(\tau_S, \tau_S + t_{\mathrm{u}}(\mathcal{G})\right)|E\right] + \mathbf{E}_x\left[\mathcal{L}_S\left(\tau_S + t_{\mathrm{u}}(\mathcal{G}) + 1, \frac{|\mathcal{G}|}{s}\right)|E\right]\\ =:& D_1 + D_2.\end{aligned}$$

We have

$$D_2 \le (1+(2e)^{-1})\mathbf{E}_\pi\left[\mathcal{L}_S\left(\frac{|\mathcal{G}|}{s}\right)\right] \le 2.$$

We will now bound D_1. By the strong Markov property, we have that

$$\begin{aligned}D_1 &\le \max_{z\in S}\mathbf{E}_z[\mathcal{L}_S(t_{\mathrm{u}}(\mathcal{G}))] = \max_{z\in S}\mathbf{E}_z\sum_{t=0}^{t_{\mathrm{u}}(\mathcal{G})}\mathbf{1}_{\{X(t)\in S\}}\\ &= \max_{z\in S}\sum_{y\in S}G(z,y) \le G^*(s).\end{aligned}$$

Putting everything together completes the proof. □

The remainder of the proof of Proposition 1.2.8 is based on a simple stochastic domination argument.

Proof of Proposition 1.2.8. Let $C_2 > 0$; we will fix its precise value at the end of the proof. That X visits at least $s/2$ points in S by the time $C_2|\mathcal{G}|G^*(s)$ with probability exponentially close to 1 in s follows from a simple large deviation estimate of a binomial random variable. Namely, we run the chain for $C_2G^*(s)s$ rounds, each of length $|\mathcal{G}|/s$. We let $S_0 = S$ and inductively let $S_i = S_{i-1}\setminus\{x\}$ if X hits x in the ith round for $i \ge 1$. If $|S_i| \ge s/2$, the hypotheses of Lemma 1.2.9 hold. In this case, the probability that X hits a point in S_i in the ith round is at least $\rho_0/G^*(s) > 0$. Thus by stochastic domination, we have that

$$\mathbf{P}[\mathcal{C}_S(C_2|\mathcal{G}|G^*(s)) < s/2] \le \mathbf{P}\left[Z < s/2\right]$$

where $Z \sim \mathrm{BIN}(C_2G^*(s)s, \rho_0/G^*(s))$. By picking C_2 large enough ($C_2 > 1/\rho_0$ will do, say) and applying the Chernoff bound, we see that

$$\mathbf{P}\left[\mathcal{C}_S(C_2|\mathcal{G}|G^*(s)) < s/2\right] \le \exp(-C_3 s) \tag{1.2.25}$$

for some constant C_3 (one can check that $C_3 = \frac{1}{8}$ suffices). This estimate also holds if $s = 1$. In this case we cover the point with constant probability in $C_2|\mathcal{G}|$ steps. □

1.2.4 Proof of Theorem 1.2.3

Throughout this section, we shall assume that X is a lazy random walk on a graph $\mathcal{G}$ which satisfies Assumption 1.2.2. Recall that $\mathcal{U}(t)$ is the set of vertices of $\mathcal{G}$ which X has not visited by time t. We will use the notation $\mathbf{P}_x, \mathbf{E}_x$ for the probability measure and expectation under which $X(0) = x$. Likewise, we let $\mathbf{P}_\pi, \mathbf{E}_\pi$ correspond to the case that X is initialized at stationarity. We will now work towards completing the proof of Theorem 1.2.3 by applying the results of the previous section to describe the process by which X covers $V(\mathcal{G})$. We will study the process of coverage in two different regimes: before and after $\mathcal{U}(t)$ contains at least n^* vertices (recall the definition of n^* from part (3) of Assumption 1.2.2). To this end, we let

$$\begin{aligned} r &= \max\{i : |\mathcal{G}| - it_{\mathrm{u}}(\mathcal{G}) \geq n^*\}, \\ \widetilde{r} &= \lfloor \log_2(|\mathcal{G}| - rt_{\mathrm{u}}(\mathcal{G}))\rfloor \end{aligned}$$

and

$$\begin{aligned} s_i &= |\mathcal{G}| - it_{\mathrm{u}}(\mathcal{G}), \quad i = 0, \ldots, r, \\ s_{r+i} &= \left\lfloor \frac{s_r}{2^i} \right\rfloor \quad i = 1, \ldots, \widetilde{r} - 1, \\ s_{r+\widetilde{r}} &= 0. \end{aligned}$$

We also define the stopping times

$$T_i = \min\{t \geq 1 : |\mathcal{U}(t)| \leq s_i\}, \quad i = 1, \ldots, r + \widetilde{r}.$$

Lemma 1.2.10. *There exists constants C_4, C_5 such that for each $1 \leq i \leq r$ and all $x \in V(\mathcal{G})$, we have that*

$$\mathbf{P}_x\left[|\mathcal{U}(t)| > s_i\right] \leq \exp\left(\frac{s_i}{t_{rel}(\mathcal{G})}\left(C_4 \log|\mathcal{G}| - \frac{C_5}{|\mathcal{G}|}t\right)\right). \tag{1.2.26}$$

Proof. For each $i \in \{1, \ldots, r\}$, we let

$$t_i = \frac{2(K_2 + 2)t_{\mathrm{u}}(\mathcal{G})|\mathcal{G}|}{s_i}$$

Proposition 1.2.5 implies that

$$\mathbf{P}_x[|\mathcal{U}(t + t_i)| \leq s_{i+1} \mid |\mathcal{U}(t)| \in (s_{i+1}, s_i]] \geq 1 - \exp\left(-C\frac{t_{\mathrm{u}}(\mathcal{G})}{t_{\mathrm{rel}}(\mathcal{G})}\right).$$

Consequently, it follows that there exists independent variables $Z_j \sim \mathrm{GEO}(1 - \exp(-Ct_{\mathrm{u}}(\mathcal{G})/t_{\mathrm{rel}}(\mathcal{G})))$ such that $T_j - T_{j-1}$ is stochastically dominated by

$t_j Z_j$ for all $j \in \{1, \dots, r\}$. Thus for $\theta_i > 0$, we have that

$$\begin{aligned}\mathbf{P}_x[|\mathcal{U}(t)| > s_i] = \mathbf{P}_x[T_i > t] &= \mathbf{P}_x\left[\sum_{j=1}^{i} T_j - T_{j-1} > t\right] \\ &\leq e^{-\theta_i t} \prod_{j=1}^{i} \mathbf{E}_x[e^{\theta_i t_j Z_j}]. \end{aligned} \tag{1.2.27}$$

Note that for every $\beta \in (0,1)$ there exists $\alpha = \alpha(\beta) > 0$ such the moment generating function of a GEO(p) random variable satisfies

$$\frac{pe^x}{1-(1-p)e^x} \leq e^{\alpha x} \text{ provided } (1-p)e^x \leq \beta. \tag{1.2.28}$$

Choosing

$$\theta_i = \frac{Ct_{\mathrm{u}}(\mathcal{G})}{2t_i t_{\mathrm{rel}}(\mathcal{G})}$$

we have that

$$\theta_i t_j = \frac{Ct_{\mathrm{u}}(\mathcal{G})}{t_{\mathrm{rel}}(\mathcal{G})} \cdot \frac{t_j}{t_i} = \frac{Ct_{\mathrm{u}}(\mathcal{G})}{2t_{\mathrm{rel}}(\mathcal{G})} \cdot \frac{s_i}{s_j}.$$

Hence as $s_i \leq s_j$ for all $i, j \in \{1, \dots, r\}$ with $j \leq i$, we have

$$\exp\left(\frac{Ct_{\mathrm{u}}(\mathcal{G})}{2t_{\mathrm{rel}}(\mathcal{G})} \cdot \frac{s_i}{s_j} - \frac{Ct_{\mathrm{u}}(\mathcal{G})}{t_{\mathrm{rel}}(\mathcal{G})}\right) \leq \exp\left(-\frac{Ct_{\mathrm{u}}(\mathcal{G})}{2t_{\mathrm{rel}}(\mathcal{G})}\right) \leq \exp(-C/2).$$

Let $\alpha = \alpha(e^{-C/2})$ as in (1.2.28). Consequently, we can bound the product of exponential moments in (1.2.27) by

$$\begin{aligned}\log \prod_{j=1}^{i} \mathbf{E}_x[e^{\theta_i t_j Z_j}] \leq \alpha \sum_{j=1}^{i} \theta_i t_j &= \frac{\alpha C t_{\mathrm{u}}(\mathcal{G}) s_i}{2t_{\mathrm{rel}}(\mathcal{G})} \sum_{j=1}^{i} \frac{1}{s_j} \\ &= \frac{\alpha C s_i}{2t_{\mathrm{rel}}(\mathcal{G})} \sum_{j=1}^{i} \frac{1}{|\mathcal{G}|/t_{\mathrm{u}}(\mathcal{G}) - j} \leq \frac{\alpha C s_i}{2t_{\mathrm{rel}}(\mathcal{G})} \log |\mathcal{G}|. \end{aligned}$$

Inserting this expression into (1.2.27) gives (1.2.26). □

Lemma 1.2.11. *There exists constants C_6, C_7 such that for all $1 \leq i \leq \widetilde{r}$ and $x \in V(\mathcal{G})$, we have that*

$$\begin{aligned}&\mathbf{P}_x\left[|\mathcal{U}(t)| > s_{r+i}\right] \\ \leq &\mathbf{P}_x[|\mathcal{U}(t/2)| > s_r] + \exp\left(s_{r+i-1}\left(C_6 i - \frac{C_7}{|\mathcal{G}|G^*(n^*)}t\right)\right). \end{aligned} \tag{1.2.29}$$

Proof. Let

$$q_{r+j} = C_2|\mathcal{G}|G^*(s_{r+j})$$

where C_2 is as in Proposition 1.2.8. Proposition 1.2.8 implies that

$$\mathbf{P}_x[|\mathcal{U}(t+q_{r+j})| \leq s_{r+j+1} \mid |\mathcal{U}(t)| \in (s_{r+j+1}, s_{r+j}]] \geq 1 - \exp(-C_3 s_{r+j})$$

for $j \in \{1, \cdots, \widetilde{r}\}$. Consequently, there exists independent random variables $Z_{r+j} \sim \mathrm{GEO}(1 - \exp(-C_3 s_{r+j}))$ such that $T_{r+j} - T_{r+j-1}$ is stochastically dominated by $q_{r+j} Z_{r+j}$. We have that

$$\begin{aligned}
&\mathbf{P}_x[|\mathcal{U}(t)| > s_{r+i}] = \mathbf{P}_x[T_{r+i} > t] \\
\leq &\mathbf{P}_x\left[T_r > \frac{t}{2}\right] + \mathbf{P}_x\left[\sum_{j=1}^{i} T_{r+j} - T_{r+j-1} > \frac{t}{2}\right] =: I_1 + I_2
\end{aligned} \tag{1.2.30}$$

Using that $I_1 = \mathbf{P}[|\mathcal{U}(t/2)| > s_r]$ gives the first term in (1.2.29). We now turn to bound I_2. Fixing $\theta_{r+i} > 0$, we have

$$I_2 \leq e^{-\theta_{r+i} t/2} \prod_{j=1}^{i} \mathbf{E}_x\left[e^{\theta_{r+i} q_{r+j} Z_{r+j}}\right]. \tag{1.2.31}$$

With the particular choice

$$\theta_{r+i} = \frac{C_3}{2C_2} \frac{s_{r+i}}{|\mathcal{G}|G^*(n^*)}$$

we have that

$$\exp(\theta_{r+i} q_{r+j} - C_3 s_{r+j}) \leq \exp(-C_3/2) =: \beta < 1.$$

Here, we used that if $n \leq m$ then $G^*(n) \leq G^*(m)$. Thus by (1.2.28) there exists $\alpha = \alpha(\beta) > 0$ such that we can bound the exponential moments in (1.2.31) by

$$\log \prod_{j=1}^{i} \mathbf{E}_x\left[e^{\theta_{r+i} q_{r+j} Z_{r+j}}\right] \leq \alpha \theta_{r+i} \sum_{j=1}^{i} q_{r+j} = \frac{\alpha C_3}{2} i s_{r+i}$$

Inserting this bound into (1.2.31) gives the second term in (1.2.29). □

Lemma 1.2.12. *There are constants $C_8, C_9, C_{10} > 0$ such that for*

$$t = (1+a) C_8 |\mathcal{G}|(t_{rel}(\mathcal{G}) + \log|\mathcal{G}|)$$

and every $x \in V(\mathcal{G})$ we have

$$\mathbf{E}_x\left[2^{|\mathcal{U}(t)|}\right] \leq 1 + C_9 \exp\left(-a C_{10} \log(n^*)\right). \tag{1.2.32}$$

Proof. We can write

$$\mathbf{E}_x\left[2^{|\mathcal{U}(t)|}\right] \leq 1 + \sum_{i=1}^{r+\widetilde{r}} 2^{s_{i-1}} \mathbf{P}\left[|\mathcal{U}(t)| > s_i\right].$$

For $i \leq r$, we have that $s_{i-1} = s_i + t_{\mathrm{u}}(\mathcal{G})$. By Lemma 1.2.10, we have that

$$2^{s_i + t_{\mathrm{u}}(\mathcal{G})} \mathbf{P}[|\mathcal{U}(t)| > t] \leq \exp\left\{(s_i + t_{\mathrm{u}}(\mathcal{G}))\log 2 + \frac{s_i}{t_{\mathrm{rel}}(\mathcal{G})}\left(C_4 \log|\mathcal{G}| - \frac{C_5}{|\mathcal{G}|}t\right)\right\}.$$

By taking C_8 (in the statement) large enough, this is in turn bounded from above by

$$\exp\left(-as_i\left(1 + \frac{\log|\mathcal{G}|}{t_{\mathrm{rel}}(\mathcal{G})}\right)\right). \tag{1.2.33}$$

For $r + i \in \{r+1, \ldots, r+\widetilde{r}\}$ we have from (1.2.29) that

$$\begin{aligned} &2^{s_{r+i-1}} \mathbf{P}_x[|\mathcal{U}(t)| > s_{r+i}] \\ \leq &2^{s_{r+i-1}} \mathbf{P}_x[|\mathcal{U}(t)| > \tfrac{t}{2}] + \exp\left(s_{r+i-1}\left((C_6 + \log 2)i - \frac{C_7}{|\mathcal{G}|G^*(n^*)}t\right)\right). \end{aligned}$$

The first term admits the same bound as (1.2.33) with $i = r$, possibly by increasing C_8 if necessary. Using that $i \leq \log_2|n^*|$, by increasing C_8 if necessary, from condition (3) it is easy to see that the second term admits the bound

$$\exp\left(-as_{r+i}\frac{\log|\mathcal{G}| + t_{\mathrm{rel}}(\mathcal{G})}{G^*(n^*)}\right). \tag{1.2.34}$$

Applying condition (3) again, we see that (1.2.34) is bounded from above by

$$\exp(-as_{r+i}\log(n^*)).$$

Putting together the estimates we get that for $i \in \{1 \ldots \widetilde{r}\}$

$$\begin{aligned} &2^{s_{r+i-1}} \mathbf{P}_x[|\mathcal{U}(t)| > s_{r+i}] \\ \leq &\exp\left(-as_r\left(1 + \frac{\log|\mathcal{G}|}{t_{\mathrm{rel}}(\mathcal{G})}\right)\right) + \exp\left(-as_{r+i}\log(n^*)\right) \end{aligned} \tag{1.2.35}$$

Summing (1.2.33) and (1.2.35) gives (1.2.32) (the dominant term in the summation comes from when $s_{r+i} = 1$) which proves the lemma. □

Proof of Theorem 1.2.3. This is a consequence of Lemma 1.2.12 and the relationship between $t_{\mathrm{u}}(\mathcal{G}^\diamond)$ and $\mathbf{E}[2^{|\mathcal{U}(t)|}]$ given in (1.2.9). □

1.2.5 Proof of Theorem 1.2.1

We are going to prove Theorem 1.2.1 by checking the hypotheses of Theorem 1.2.3. We begin by noting that by [83, Corollary 12.12] and [83, Section 12.3.1], we have that

$$t_{\text{rel}}(\mathbf{Z}_n^d) = \Theta(dn^2). \tag{1.2.36}$$

By [54, Example 2, Page 2155], we know that $t_{\text{u}}(\mathbf{Z}_n) = O(n^2)$. Hence by [52, Theorem 2.10], we have that

$$t_{\text{u}}(\mathbf{Z}_n^d) = O((d \log d)n^2). \tag{1.2.37}$$

The key to checking parts (1)–(3) of Assumption 1.2.2 are the Green's function estimates which are stated in Proposition 1.2.14 (low degree) and Proposition 1.2.18 (high degree). In order to establish these we will need to prove several intermediate technical estimates. We begin by recording the following facts about the transition kernel P for lazy random walk on a vertex transitive graph $\mathcal{G}$. First, we have that

$$P^t(x, y) \leq P^t(x, x) \text{ for all } x, y. \tag{1.2.38}$$

To see this, we note that for t even, the Cauchy-Schwarz inequality and the semigroup property imply

$$P^t(x, y) = \sum_z P^{t/2}(x, z)P^{t/2}(z, y) \leq \sqrt{P^t(x, x)P^t(y, y)} = P^t(x, x).$$

The inequality and final equality use the vertex transitivity of $\mathcal{G}$ so that $P(x, z) = P(z, x)$ and $P(x, x) = P(y, y)$. To get the same result for t odd, one just applies the same trick used in the proof of [83, Proposition 10.18(ii)]. Moreover, by [83, Proposition 10.18], we have that

$$P^t(x, x) \leq P^s(x, x) \text{ for all } s \leq t. \tag{1.2.39}$$

The main ingredient in the proof of Proposition 1.2.14, our low degree Green's function estimate, is the following bound for the return probability of a lazy random walk on $\mathbf{Z}^d$.

Lemma 1.2.13. *Let $P(x, y; \mathbf{Z}^d)$ denote the transition kernel for lazy random walk on $\mathbf{Z}^d$. For all $t \geq 1$, we have that*

$$P^t(x, x; \mathbf{Z}^d) \leq \sqrt{2} \left(\frac{4d}{\pi}\right)^{d/2} \frac{1}{t^{d/2}} + e^{-t/8}. \tag{1.2.40}$$

Proof. To prove the lemma we first give an upper bound on the transition probabilities for a (non-lazy) simple random walk Y on $\mathbf{Z}^d$. One can easily give an exact formula for the return probability of Y to the origin of $\mathbf{Z}^d$ in

$2t$ steps by counting all of the possible paths from 0 back to 0 of length $2t$ (here and hereafter, $P_{\mathrm{NL}}(x,y;\mathbf{Z}^d)$ denotes the transition kernel of Y):

$$\begin{aligned} P_{\mathrm{NL}}^{2t}\left(x,x;\mathbf{Z}^d\right) &= \sum_{n_1+\cdots+n_d=t} \frac{(2t)!}{(n_1!)^2(n_2!)^2\cdots(n_d!)^2}\cdot\frac{1}{(2d)^{2t}} \\ &= \frac{1}{(2d)^{2t}}\binom{2t}{t}\sum_{n_1+\cdots+n_d=t}\left(\frac{t!}{n_1!n_2!\cdots n_d!}\right)^2 \end{aligned}$$

We can bound the sum above as follows, using the multinomial theorem in the second step:

$$\begin{aligned} P_{\mathrm{NL}}^{2t}\left(x,x;\mathbf{Z}^d\right) &\le \frac{1}{(2d)^{2t}}\binom{2t}{t}\left(\max_{n_1+\cdots+n_d=t}\frac{t!}{n_1!\cdots n_d!}\right)\sum_{n_1+\cdots+n_d=t}\frac{t!}{n_1!\cdots n_d!} \\ &\le \frac{1}{(2d)^{2t}}\binom{2t}{t}\frac{t!}{[(\lfloor t/d\rfloor)!]^d}\cdot d^t. \end{aligned}$$

Applying Stirlings formula to each term above, we consequently arrive at

$$P_{\mathrm{NL}}^{2t}\left(x,x;\mathbf{Z}^d\right) \le \frac{\sqrt{2}}{(2\pi)^{d/2}}\cdot\frac{d^{d/2}}{t^{d/2}} \tag{1.2.41}$$

We are now going to deduce from (1.2.41) a bound on the return probability for a lazy random walk X on $\mathbf{Z}^d$. We note that we can couple X and Y so that X is a random time change of Y: $X(t)=Y(N_t)$ where $N_t=\sum_{i=0}^t\xi_i$ and the (ξ_i) are iid with $\mathbf{P}[\xi_i=0]=\mathbf{P}[\xi_i=1]=\frac{1}{2}$ and are independent of Y. Note that N_t is distributed as a binomial random variable with parameters t and $1/2$. Thus,

$$\begin{aligned} P^t(x,x;\mathbf{Z}^d) &= \sum_{i=0}^{t/2} P_{\mathrm{NL}}^{2i}(x,x;\mathbf{Z}^d)\mathbf{P}(N_t=2i) \\ &\le \mathbf{P}(N_t<t/4)+\sqrt{2}\left(\frac{4d}{\pi}\right)^{d/2}\frac{1}{t^{d/2}}, \end{aligned}$$

where in the second term we used the monotonicity of the upper bound in (1.2.41) in t. The first term can be bounded from above by using the Hoeffding inequality. This yields the term $e^{-t/8}$ in (1.2.40). □

Throughout the rest of this section, we let $|x-y|$ denote the L^1 distance between $x,y\in\mathbf{Z}_n^d$.

Proposition 1.2.14. *Let $G(x,y)$ denote the Green's function for lazy random walk on $\mathbf{Z}_n^d$. For each $\delta\in(0,1)$, there exists constants $C_1,C_2,C_3>0$ independent of n,d for $d\ge 3$ such that*

$$\begin{aligned} G(x,y) \le{}& \frac{C_1}{d}\left(\frac{4d}{\pi}\right)^{d/2}|x-y|^{1-d/2} + C_2(d\log d)\left(\frac{4d}{\pi}\right)^{d/2}n^{2-d(1-\delta/2)} \\ &+ C_3\big(d^2\log d\big)n^2e^{-n^\delta/2} \end{aligned}$$

for all $x, y \in \mathbf{Z}_n^d$ distinct.

Proof. Fix $\delta \in (0,1)$. We first observe that the probability that there is a coordinate in which the random walk wraps around the torus within $t < n^2$ steps can be estimated by using Hoeffding's inequality and a union bound by

$$d \cdot \mathbf{P}(Z(t) > n) = de^{-\frac{n^2}{2t}}$$

where $Z(t)$ is a one dimensional simple random walk on $\mathbf{Z}$. Let $k = |x - y|$. Applying (1.2.38) and (1.2.39) in the second step, and estimating the probability of wrapping around in time $n^{2-\delta}$ in the third term, we see that

$$G(x,y) = \sum_{t=k}^{t_{\mathrm{u}}} P^t(x,y) \leq \sum_{t=k}^{n^{2-\delta}} P^t(x,x;\mathbf{Z}^d) + t_{\mathrm{u}} P^{n^{2-\delta}}(x,x;\mathbf{Z}^d) \qquad (1.2.42)$$
$$+ dt_{\mathrm{u}} e^{-\frac{n^\delta}{2}}.$$

We can estimate the sum on the right hand side above using Lemma 1.2.13, yielding the first term in the assertion of the lemma. Applying Lemma 1.2.13 again, we see that there exists a constant C_2 which does not depend on n, d such that the second term in the right side of (1.2.42) is bounded by

$$C_2(d \log d) \left(\frac{4d}{\pi}\right)^{d/2} n^{2-d(1-\delta/2)}. \qquad (1.2.43)$$

Indeed, the factor $(d \log d)n^2$ comes from (1.2.37) and the other factor comes from Lemma 1.2.13. Combining proves the lemma. □

Proposition 1.2.14 is applicable when n is much larger than d. We now turn to prove Proposition 1.2.18, which gives us an estimate for the Green's function which we will use when d is large. Before we prove Proposition 1.2.18, we first need to collect the following estimates.

Lemma 1.2.15. *Suppose that X is a lazy random walk on $\mathbf{Z}_n^d$ for $d \geq 8$ and that $|X(0)| = k \leq \frac{d}{8}$. For each $j \geq 0$, let τ_j be the first time t that $|X(t)| = j$. There exists a constant $C_k > 0$ depending only on k such that $\mathbf{P}[\tau_0 < \tau_{2k}] \leq C_k d^{-k}$. If, instead, $|X(0)| = 1$, then there exists a universal constant $p > 0$ such that $\mathbf{P}[\tau_0 < \tau_{2k}] \geq p$.*

Proof. It clearly suffices to prove the result when X is non-lazy. Assume that $|X(t)| = j \in \{k, \ldots, 2k\}$. It is obvious that the probability that $|X|$ moves to $j+1$ in its next step is at least $1 - \frac{2k}{d}$. The reason is that the probability that the next coordinate to change is one of the coordinates of $X(t)$ whose value is 0 is at least $1 - \frac{2k}{d}$. Similarly, the probability that $|X|$ next moves to $j-1$ is at most $\frac{2k}{d}$. Consequently, the first result of the lemma follows from the Gambler's ruin problem (see, for example, [83, Section 17.3.1]). The second assertion of the lemma follows from the same argument. □

Lemma 1.2.16. *Assume that $k \in \mathbf{N}$ and that $d = 2k \vee 3$. Suppose that X is a lazy random walk on $\mathbf{Z}^d$ and that $|X(0)| = 2k$. Let τ_k be the first time t that $|X(t)| = k$. There exists $p_k > 0$ depending only on k such that $\mathbf{P}[\tau_k = \infty] \geq p_k > 0$.*

Proof. Let $\mathbf{P}_y$ denote the law under which X starts at y. Assume that $\mathbf{P}_y[\tau_k = \infty] = 0$ for some $y \in \mathbf{Z}^d$ with $|y| = 2k$. Suppose that $z \in \mathbf{Z}^d$ with $|z| = 2k$ and let τ_z be the first time that X hits z. Then since $\mathbf{P}_y[\tau_z < \tau_k] > 0$, it follows from the strong Markov property that $\mathbf{P}_z[\tau_k = \infty] = 0$. From this, it follows that the expected amount of time that X spends in $B(0,k)$ is infinite because it implies that on each successive hit to $\partial B(0,2k)$, X returns to $B(0,k)$ with probability 1. Since X is transient [81, Theorem 4.3.1], the expected amount of time that X spends in $B(0,k)$ is finite. This is a contradiction. □

Lemma 1.2.17. *Assume that $k \in \mathbf{N}$ and $d \geq 2k \vee 3$. Suppose that X is a lazy random walk on $\mathbf{Z}_n^d$ and that $|X(0)| = 2k$. Let τ_k be the first time t that $|X(t)| = k$. There exists $p_k, c_k > 0$ depending only on k such that $\mathbf{P}[\tau_k > c_k d n^2] \geq p_k > 0$.*

Proof. We first assume that $d = 2k \vee 3$. It follows from Lemma 1.2.16 that there exists a constant $p_{k,1} > 0$ depending only on k such that $\mathbf{P}[\tau_k > \tau_{n/4}] \geq p_{k,1}$. The local central limit theorem (see [81, Chapter 2]) implies that there exists constants $c_{k,1}, p_{k,2} > 0$ such that the probability that a random walk on $\mathbf{Z}_n^d$ moves more than distance $\frac{n}{4}$ in time $c_{k,1}n^2$ is at most $1 - p_{k,2}$. Combining implies the result for $d = 2k \vee 3$.

Now we suppose that $d \geq 2k \vee 3$. Let $(X_1(t), \ldots, X_d(t))$ be the coordinates of $X(t)$. By re-ordering if necessary, we may assume without loss of generality that $X_{2k+1}(0), \ldots, X_d(0) = 0$. Let $Y(t) = (X_1(t), \ldots, X_{2k}(t))$. Then Y is a random walk on $\mathbf{Z}_n^{2k}$. Clearly, $|Y(0)| = 2k$ because $X(0)$ cannot have more than $2k$ non-zero coordinates. For each j, let τ_j^Y be the first time t that $|Y(t)| = j$. Then $\tau_k^Y \leq \tau_k$. For each t, let N_t denote the number of steps that X takes in the time interval $\{1, \ldots, t\}$ in which one of its first $2k$ coordinates is changed (in other words, N_t is the number of steps taken by Y). The previous paragraph implies that $\mathbf{P}[N_{\tau_k^Y} \geq c_{k,1}n^2] \geq p_{k,3} > 0$ for a constant $p_{k,3} > 0$ depending only on k. Since the probability that the first $2k$ coordinates are changed in any step is k/d (recall that X is lazy), the final result holds from a simple large deviations estimate. □

Now we are ready to prove our estimate of $G(x,y)$ when d is large.

Proposition 1.2.18. *Suppose that $d \geq 8$. Let $G(x,y)$ denote the Green's function for lazy random walk on $\mathbf{Z}_n^d$. For each $k \in \mathbf{N}$ with $k \leq \frac{d}{8}$, there exists a constant $C_k > 0$ which does not depend on n, d such that*

$$G(x,y) \leq \frac{C_k}{d^k} \text{ for all } x, y \in \mathbf{Z}_n^d \text{ with } |x - y| \geq k.$$

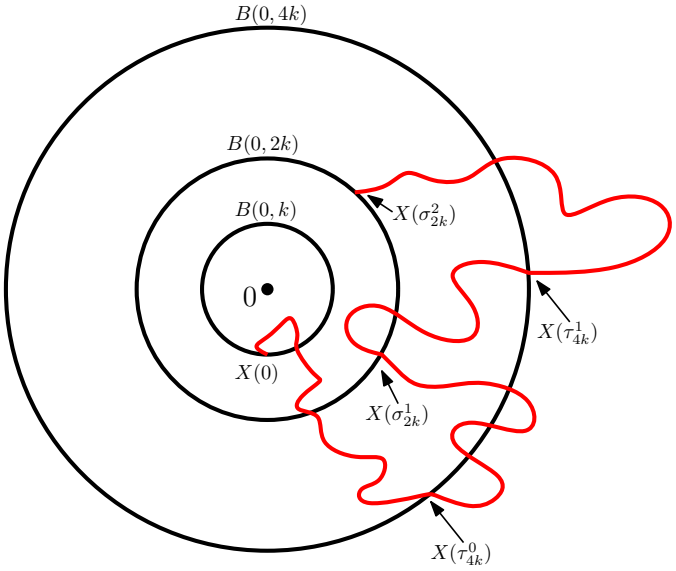


Figure 1.2: Assume that $d \geq 8$ and that $k \in \mathbf{N}$ with $d \geq 8k$. Let X be a lazy random walk on $\mathbf{Z}_n^d$ and that $X(0) = x$ with $|x - y| = k$. In Proposition 1.2.18, we show that $G(x, y) \leq C_k d^{-k}$ where $C_k > 0$ is a constant depending only on k . By translation, we may assume without loss of generality that $|x| = k$ and $y = 0$. The idea of the proof is to first invoke Lemma 1.2.15 to show that X escapes to $\partial B(0, 4k)$ with probability at least $1 - C_{k,1} d^{-k}$. We then decompose the path of X into successive excursions $\{X(\sigma_{2k}^j), \ldots, X(\tau_{4k}^j), \ldots, X(\sigma_{2k}^{j+1})\}$ between $\partial B(0, 2k)$ back to itself through $\partial B(0, 4k)$. By Lemma 1.2.15, we know that each excursion hits 0 with probability bounded by $C_{2k,1} d^{-2k}$ and Lemma 1.2.17 implies that each excursion takes length $c_k d n^2$ with probability at least $p_k > 0$. Consequently, the result follows from a simple stochastic domination argument.

Proof. See Figure 1.2 for an illustration of the proof. By translation, we may assume without loss of generality that $y = 0$; let $k = |x|$. Let τ_0 be the first time t that $|X(t)| = 0$. The strong Markov property implies that

$$G(x, y) \leq \mathbf{P}[\tau_0 > t_{\mathrm{u}}] + (1 - \mathbf{P}[\tau_0 > t_{\mathrm{u}}])G(x, x).$$

Consequently, it suffices to show that for each $k \in \mathbf{N}$, there exists constants $C_k, C_0 > 0$ such that

$$\mathbf{P}[\tau_0 > t_{\mathrm{u}}] \leq \frac{C_k}{d^k} \text{ and} \tag{1.2.44}$$

$$G(x, x) \leq C_0. \tag{1.2.45}$$

We will first prove (1.2.44); the proof of (1.2.45) will be similar.

Let N be a geometric random variable with success probability $C_{2k}d^{-2k}$ where C_{2k} is the constant from Lemma 1.2.15. Let (ξ_j) be a sequence of independent random variables with $\mathbf{P}[\xi_j = c_{2k}dn^2] = p_{2k}$ and $\mathbf{P}[\xi_j = 0] = 1 - p_{2k}$ where c_{2k}, p_{2k} are the constants from Lemma 1.2.17 independent of N. We claim that τ_0 is stochastically dominated from below by $\sum_{j=1}^{N\zeta-1} \xi_j$ where ζ is independent of N and (ξ_j) with $\mathbf{P}[\zeta = 0] = C_k d^{-k} = 1 - \mathbf{P}[\zeta = 1]$. Indeed, to see this we let $\sigma_k^0 = 0$ and let τ_{4k}^0 be the first time t that $|X(t)| = 4k$. For each $j \geq 1$, we inductively let σ_{2k}^j be the first time t after τ_{4k}^{j-1} that $|X(t)| = 2k$ and let τ_{4k}^j be the first time t after σ_{2k}^j that $|X(t)| = 4k$. Let $\mathcal{F}_t$ be the filtration generated by X. Lemma 1.2.15 implies that the probability that X hits 0 in $\{\sigma_{2k}^j, \dots, \tau_{4k}^j\}$ given $\mathcal{F}_{\sigma_{2k}^j}$ is at most $C_{2k}d^{-2k}$ for each $j \geq 1$ where $C_{2k} > 0$ only depends on $2k$. This leads to the success probability in the definition of N above. The factor ζ is to take into account the probability that X reaches distance $2k$ before hitting 0. Moreover, Lemma 1.2.17 implies that $\mathbf{P}[\sigma_{2k}^j - \tau_{4k}^{j-1} \geq c_{2k}dn^2 | \mathcal{F}_{\tau_{4k}^j}] \geq p_{2k}$. This leads to the definition of the (ξ_j) above. This implies our claim.

To see (1.2.44) from our claim, an elementary calculation yields that

$$\mathbf{P}[N\zeta \leq C_{2k}^{-1}d^k] \leq \mathbf{P}[N \leq C_{2k}^{-1}d^k \text{ or } \zeta = 0] \leq 2d^{-k} + C_k d^{-k}.$$

We also note that

$$\mathbf{P}\left[\sum_{j=1}^{m} \xi_j \leq \frac{p_k c_k m n^2}{2}\right] \leq e^{-cm}$$

for some constant $c > 0$. Combining these two observations along with a union bound implies (1.2.44). To see (1.2.45), we apply a similar argument using the second assertion of Lemma 1.2.15. □

Now that we have proved Proposition 1.2.14 and Proposition 1.2.18, we are ready to check the criteria of Assumption 1.2.2.

Part (1)

By [83, Proposition 1.14] with $\tau_x^+ = \min\{t \geq 1 : X(t) = x\}$, we have that $\mathbf{E}_x[\tau_x^+] = |\mathbf{Z}_n^d|$. Applying Proposition 1.2.18, we see that there exists constants $d_0, r > 0$ such that if $d \geq d_0$, then

$$G(x, y) \leq 1/2 \text{ for all } |x - y| \geq r. \tag{1.2.46}$$

Proposition 1.2.14 implies that there exists n_0 such that if $n \geq n_0$ and $3 \leq d < d_0$ then (1.2.46) likewise holds, possibly by increasing r (clearly, part (1) holds when $d \leq d_0$ and $n \leq n_0$; note also that we may assume without loss of generality that d_0, n_0 are large enough so that the diameter

of the graph is at least $2r$). Let τ_r be the first time t that $|X(t)-X(0)| = r$. We observe that there exists $\rho_0 = \rho_0(r) > 0$ such that

$$\mathbf{P}_x[\tau_r < \tau_x^+] \geq \rho_0 \tag{1.2.47}$$

uniform in n, d since in each time step there are d directions in which $X(t)$ increases its distance from $X(0)$. By combining (1.2.46) with (1.2.47), we see that $\mathbf{P}_x[\tau_x^+ \geq t_{\mathrm{u}}(\mathcal{G})] \geq \rho_1 > 0$ uniform $d \geq d_0$. Let $\mathcal{F}_t$ be the filtration generated by X. We consequently have that

$$\begin{aligned}\mathbf{E}_x[\tau_x^+] &\geq \mathbf{E}_x[\tau_x^+ \mathbf{1}_{\{\tau_x^+ \geq t_{\mathrm{u}}(\mathcal{G})\}}] = \mathbf{E}_x\left[\mathbf{E}_x[\tau_x^+ | \mathcal{F}_{t_{\mathrm{u}}(\mathcal{G})}]\mathbf{1}_{\{\tau_x^+ \geq t_{\mathrm{u}}(\mathcal{G})\}}\right] \\ &\geq \mathbf{E}_x\left[\mathbf{E}_{X(t_{\mathrm{u}}(\mathcal{G}))}[\tau_x]\mathbf{1}_{\{\tau_x^+ \geq t_{\mathrm{u}}(G)\}}\right] \geq \rho_1 \left(1 - \frac{1}{2e}\right) \mathbf{E}_\pi[\tau_x].\end{aligned}$$

That is, there exists $\rho_2 > 0$ uniform in $d \geq d_0$ such that $\mathbf{E}_x[\tau_x^+] \geq \rho_2 \mathbf{E}_\pi[\tau_x]$. Hence by [83, Lemma 10.2], we have that $t_{\mathrm{hit}}(\mathbf{Z}_n^d) \leq K_1 |\mathbf{Z}_n^d|$ where $K_1 = 2/\rho_2$ is a uniform constant.

Remark 1.2.19. *There is another proof of Part 1 which is based on eigenfunctions. In particular, we know that*

$$t_{hit}(\mathbf{Z}_n^d) \leq 2\mathbf{E}_\pi[\tau_x] = 4 \sum_i \frac{1}{1-\lambda_i}$$

where the λ_i are the eigenvalues of simple random walk on $\mathbf{Z}_n^d$ distinct from 1; the extra factor of 2 in the final equality accounts for the laziness of the chain. The λ_i can be computed explicitly using [83, Lemma 12.11] and the form of the λ_i when $d = 1$ which are given in [83, Section 12.3]. The assertion follows by performing the summation which can be accomplished by approximating it by an appropriate integral.

Part (2)

It follows from Proposition 1.2.18 that there exist constants $C > 0$ and $d_0 \geq 3$ such that

$$G(x,y) \leq \frac{C}{d} \text{ for } x, y \in \mathbf{Z}_n^d \text{ with } |x-y| = 1 \tag{1.2.48}$$

provided $d \geq d_0$. Consequently, there exists $K \in \mathbf{N}$ which does not depend on $d \geq d_0$ such that

$$2K(5/2)^K G^K(x,y) = O\left(2K\left(\frac{5/2}{d}\right)^K\right) \tag{1.2.49}$$

It follows by combining (1.2.36) and (1.2.37) that we have that

$$\frac{t_{\mathrm{u}}(\mathbf{Z}_n^d)}{t_{\mathrm{rel}}(\mathbf{Z}_n^d)} = O(\log d). \tag{1.2.50}$$

Combining (1.2.49) with (1.2.50) shows that part (2) of Assumption 1.2.2 is satisfied provided we take $K_2 = K$ large enough. Moreover, (1.2.50) clearly holds if $3 \leq d < d_0$ by Proposition 1.2.14.

Part (3)

We first note that it follows from (1.2.36), (1.2.37), Proposition 1.2.14, and Proposition 1.2.18 that there exists constants $C > 0$ such that n^* for $\mathbf{Z}_n^d$ is at most $Cd^2n^2 \log d$ for all $d \geq 3$. To check this part, we need to show that there exists $K_3 > 0$ such that

$$G^*(n^*) \leq K_3 \left(\frac{dn^2 + d \log n}{\log d + \log n} \right). \tag{1.2.51}$$

We are going to prove the result by considering the regimes of $d \leq \sqrt{\log n}$ and $d > \sqrt{\log n}$ separately.

Case 1: $d < \sqrt{\log n}$.

From (1.2.51) it is enough to show that $G^*(n^*) \leq Kdn^2/\log n$. We can bound $G^*(n^*)$ in this case as follows. Let $D = (d \log d \log n)^{1/(\frac{1}{2}d-1)}$. By Proposition 1.2.14, we can bound from above the expected amount of time that X starting at 0 in $\mathbf{Z}_n^d$ spends in the L^1 ball of radius D by summing radially:

$$\begin{aligned}
&\sum_{k=1}^{D} \frac{C_1}{d} \left(\frac{4d}{\pi} \right)^{d/2} k^{1-d/2} \cdot 2d(2k)^{d-1} \\
\leq &C_1 \left(\frac{16d}{\pi} \right)^{d/2} \sum_{k=1}^{D} k^{d/2} \leq \frac{C_2}{d} \left(\frac{16d}{\pi} \right)^{d/2} D^{1+d/2} \leq C_3 n (d \log d \log n)^5
\end{aligned}$$

for constants $C_1, C_2, C_3 > 0$, where we used that $d^{d/2} \leq n$. We also note that $2d(2k)^{d-1}$ is the size of the L^∞ ball of radius k. The exponent of 5 comes from the inequality

$$\frac{\frac{1}{2}d + 1}{\frac{1}{2}d - 1} \leq 5 \text{ for all } d \geq 3.$$

We can estimate $G^*(n)$ by dividing between the set of points which have distance at most D to 0 and those whose distance to 0 exceeds D by:

$$\begin{aligned}
G^*(n^*) \leq &C_3 n (d \log d \log n)^5 + C_4 D^{1-\frac{1}{2}d} n^* \\
\leq &C_3 n (d \log d \log n)^5 + \frac{C_4 \cdot Cd^2 n^2 \log d}{d \log d \log n},
\end{aligned}$$

where $C_4 > 0$ is a constant and we recall that $C > 0$ is the constant from the definition of n^*. This implies the desired result.

Case 2: $d \geq \sqrt{\log n}$.

In this case, we are going to employ Proposition 1.2.18 to bound $G^*(n^*)$. The number of points which have distance at most k to 0 is clearly $1+(2d)^k$. Consequently, by Proposition 1.2.18, we have that

$$\begin{aligned} G^*(n^*) \leq & \left(C_0 + \sum_{k=1}^{3} C_k d^{-k} (2d)^k \right) + C_4 d^{-4} n^* \\ \leq & C_5 + \frac{C_6 (\log d) n^2}{d^2} \end{aligned}$$

for some constants $C_5, C_6 > 0$. Since $d^2 \geq \log n$, this is clearly dominated by the right hand side of (1.2.51) (with a large enough constant), which completes the proof in this case.

□

1.3 Mixing and relaxation time for random walk on wreath product graphs

1.3.1 The generalized lamplighter model

Let us recall the general setting of the random walk on the wreath product $H \wr G$ first. Suppose that G and H are finite connected graphs with vertices $V(G)$, $V(H)$ and edges $E(G)$, $E(H)$, respectively. We refer to G as the base and H as the lamp graph, respectively. Let $\mathcal{X}(G) = \{\underline{f} \colon V(G) \to H\}$ be the set of markings of $V(G)$ by elements of H. The wreath product $H \wr G$ is the graph whose vertices are pairs $(\underline{f}, x)$ where $\underline{f} = (f(v))_{v \in V(G)} \in \mathcal{X}(G)$ and $x \in V(G)$. There is an edge between $(\underline{f}, x)$ and $(\underline{g}, y)$ if and only if $(x, y) \in E(G)$, $(f(x), g(x)), (f(y), g(y)) \in E(H)$ and $f(z) = g(z)$ for all $z \notin \{x, y\}$. Suppose that P and Q are transition matrices for Markov chains on G and on H, respectively. The generalized lamplighter walk $X^\diamond$ (with respect to the transition matrices P and Q) is the Markov chain on $H \wr G$ which moves from a configuration $(\underline{f}, x)$ by

1. picking y adjacent to x in G according to P, then
2. updating each of the values of $f(x)$ and $f(y)$ independently according to Q on H.

The state of lamps $f(z)$ at all other vertices $z \in G$ remain fixed. It is easy to see that if P and Q are irreducible, aperiodic and reversible with stationary distribution π_G and π_H, respectively, then the unique stationary distribution of $X^\diamond$ is the product measure

$$\pi^\diamond\big((\underline{f}, x)\big) = \pi_G(x) \cdot \prod_{x \in V(G)} \pi_H\left(f(x)\right),$$

and $X^\diamond$ is itself reversible. In this section, we will be concerned with the special case that P is the transition matrix for the *lazy random walk* on G, see (1.2.1). We further assume that the transition matrix Q on H is irreducible and aperiodic. This and lazyness on G guarantees that we avoid issues of periodicity.

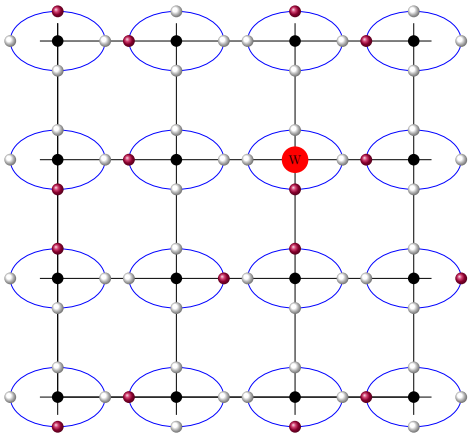


Figure 1.3: A typical state of the generalized lamplighter walk. Here $H = \mathbf{Z}_4$ and $G = \mathbf{Z}_4^2$, the red bullets on each copy of H represents the state of the lamps over each vertex $v \in G$ and the walker is drawn as a red W bullet.

1.3.2 Main Results

In order to state our general result, we first need to review some basic terminology from the theory of Markov chains. Recall the definitions of ε-*mixing time* (1.2.12), the *relaxation time* (1.2.3), the *maximal hitting time* (1.2.4) from Section 1.2. A few changes, we will set $t_{\text{mix}}(G) := t_{\text{mix}}(G, \frac{1}{4})$ throughout this section and we denote the base graph by simply G.

We further define the random cover time τ_{cov} is the first time when all vertices have been visited by the walker X, and the *cover time* $t_{\text{cov}}(G)$ is

$$t_{\text{cov}}(G) := \max_{x \in V(G)} \mathbf{E}_x[\tau_{\text{cov}}]. \tag{1.3.1}$$

At this moment, for technical reasons (we just found a mistake in the theorems we cited before) we state our main theorems under a weaker assumption than the highest generality. First we need to define the concept of strong stationary times.

Definition 1.3.1. *A randomized stopping time τ is called strong stationary time for the Markov chain X_t on G if*

$$\mathbf{P}_x\left[X_\tau = y, \tau = t\right] = \pi(y)\mathbf{P}_x[\tau = t],$$

that is, the place where τ stops is independent of the time when it stops.

The adjective randomized means that the stopping time can depend on some extra randomness, not just purely the trajectories of the Markov chain, for a precise definition see [83, Section 6.2.2].

Definition 1.3.2. *A state $h(x) \in V(G)$ is called a halting state for a stopping time τ and initial state x if $\{X_t = h(x)\}$ implies $\{\tau \leq t\}$.*

Our main results are summarized in the following theorem:

Theorem 1.3.3. *Let us assume that G and H are connected graphs with G regular and the Markov chain on H is ergodic and reversible. Then there exist universal constants c_1, C_1, c_2, C_2 such that the mixing and the relaxation time of the generalized lamplighter walk on $H \wr \mathcal{G}$ satisfies*

$$c_1 \leq \frac{t_{rel}(H \wr G)}{t_{hit}(G) + |G| t_{rel}(H)} \leq C_1, \tag{1.3.2}$$

$$\begin{aligned} &c_2 \left(t_{cov}(G) + t_{rel}(H)|G| \log |G| + |G| t_{mix}(H)\right) \leq t_{mix}(H \wr G) \\ &t_{mix}(H \wr G) \leq C_2 \left(t_{cov}(G) + |G| t_{mix}(H, \frac{1}{|G|})\right). \end{aligned} \tag{1.3.3}$$

If further the Markov chain is such that

(A) *There is a strong stationary time τ_H for the Markov chain on H which has a halting state $h(x)$ for every initial starting point $x \in H$,*

then the upper bound of 1.3.3 is sharp.

Outline

The remainder of this section is structured as follows. In Section 1.3.3 we state a few necessary theorems and lemmas about the Dirichlet form, strong stationary times, different notions of distances and their relations. In Lemmas 1.3.6 and 1.3.8 we construct a crucial stopping time $\tau^\diamond$ and a strong stationary time $\tau_2^\diamond$ on $H \wr G$ which we will use several times throughout the proofs later. Then we prove the main theorem about the relaxation time in Section 1.3.4, and the mixing time bounds in Section 1.3.5.

Notations

Throughout the section, objects related to the base or the lamp graph will be indexed by G and H, respectively, and $\diamond$ always refers to an object related to the whole $H \wr G$. Unless misleading, G and H refers also to the vertex set of the graphs, i.e. $v \in G$ means $v \in V(G)$. $\mathbf{P}_\mu, \mathbf{E}_\mu$ denotes probability and expectation under the conditional law where the initial distribution of the Markov chain under investigation is μ. Similarly, $\mathbf{P}_x(.) = \mathbf{P}(.|X_0 = x)$.

1.3.3 Preliminaries

In this subsection we collect the preliminary lemmas to be able to carry through the proofs quickly afterwards. The reader familiar with notions of strong stationary times, separation distance, and Dirichlet forms might want jump forward to Lemmas 1.3.6 and 1.3.8 immediately, and check the other lemmas here only when needed.

The first lemma is a common useful tool to prove lower bounds for relaxation times, by giving the variational characterization of the spectral gap. First we start with a definition.

Let P be a reversible transition matrix with stationary distribution π on the state space Ω and let $\mathbf{E}_\pi[\phi] := \sum_{y\in\Omega}\phi(y)\pi(y)$. The Dirichlet form associated to the pair (P,π) is defined for functions ϕ and η on Ω by

$$\mathcal{E}(\phi,\eta) := \langle (I-P)\phi,\eta\rangle_\pi = \sum_{y\in\Omega}(I-P)\phi(y)\eta(y)\pi(y).$$

It is not hard to see [83, Lemma 13.11] that

$$\mathcal{E}(\phi) := \mathcal{E}(\phi,\phi) = \frac{1}{2}\mathbf{E}_\pi\left[(\phi(X_1)-\phi(X_0))^2\right] \tag{1.3.4}$$

The next lemma relates the spectral gap of the chain to the Dirichlet form (for a short proof see [49] or [83, Lemma 13.12]):

Lemma 1.3.4 (Variational characterization of the spectral gap)**.** *The spectral gap $\gamma = 1-\lambda_2$ of a reversible Markov Chain satisfies*

$$\gamma = \min_{\phi:\, \boldsymbol{Var}_\pi\phi\neq 0}\frac{\mathcal{E}[\phi]}{\boldsymbol{Var}_\pi\phi}, \tag{1.3.5}$$

where $\boldsymbol{Var}_\pi\phi = \mathbf{E}_\pi[\phi^2] - (\mathbf{E}_\pi[\phi])^2$.

A very useful object to prove the upper bound on t_{rel} and both bounds for t_{mix} is the concept of strong stationary times. Recall the definition from (1.3.1). It is not hard to see ([83, Lemma 6.9]) that the defining equality is equivalent to

$$\mathbf{P}_x\left[X_t = y, \tau\leq t\right] = \pi(y)\mathbf{P}_x[\tau\leq t]. \tag{1.3.6}$$

To be able to relate the tail of the strong stationary times to the mixing time of the graphs, we need another distance from stationary measure, called the separation distance:

$$s_x(t) := \max_{y\in\Omega}\left[1-\frac{P^t(x,y)}{\pi(y)}\right]. \tag{1.3.7}$$

The relation between the separation distance and any strong stationary time τ is the following inequality from [49] or [83, Lemma 6.11]:

$$\forall x\in\Omega : s_x(t)\leq \mathbf{P}_x(\tau > t). \tag{1.3.8}$$

Throughout the section, we will need a slightly stronger result than (1.3.8), namely from the proof of (1.3.8) [83, Lemma 6.11] it follows that in (1.3.8) equality holds *if τ has a halting state $h(x)$ for x*. Unfortunately, we just point out that the [83, Remark 6.12] is not true and the statement can not be reversed: the state $h(x,t)$ maximizing the separation distance at time t can also depend on t and thus the existence of a halting state is not necessarily needed to get equality in (1.3.8).

On the other hand, one can always construct τ such that (1.3.8) holds with equality for every $x \in \Omega$. This is a key ingredient to our proofs, so we cite it as a Theorem (with adjusted notation to the present section).

Theorem 1.3.5. *[Aldous, Diaconis]*[2, Proposition 3.2] *Let $(X_t, t \geq 0)$ be an irreducible aperiodic Markov chain on a finite state space Ω with initial state x and stationary distribution π, and let $s_x(t)$ be the separation distance defined as in* (1.3.7)*. Then*

1. *if τ is a strong stationary time for X_t, then $s_x(t) \leq \mathbf{P}_x(\tau > t)$ for all $t \geq 0$.*
2. *Conversely, there exists a strong stationary time τ such that $s_x(t) = \mathbf{P}_x(\tau > t)$ holds with equality.*

Combining these, we will call a strong stationary time τ *separation optimal* if it achieves equality in (1.3.8). Mind that every stopping time obeying halting states is separation optimal, but the reversed statement is not necessarily true. The next two lemmas, which we will use several times, constructs two stopping times for the graph $H \wr G$. The first one will be used to lower bound the separation distance and the second one upper bounds it.

We start with introducing the notation

$$L_v(t) = 2\sum_{i=0}^{t} \mathbf{1}(X_i = v) - \delta_{X_0,v} - \delta_{X_t,v} \tag{1.3.9}$$

for the number of moves on the lamp graph $H_v, v \in G$ by the walker up to time t. Slightly confusing, we call it the local time at vertex $v \in G$.

Lemma 1.3.6. *Let τ_H be any strong stationary time for the Markov chain on H. Take independent copies of $(\tau_H(v))_{v\in G}$ and define the stopping time $\tau^\diamond$ for $X^\diamond$ by*

$$\tau^\diamond := \inf\left\{t : \forall v \in G : \tau_H(v) \leq L_v(t)\right\}. \tag{1.3.10}$$

Then, for $\tau^\diamond$ we have

$$\mathbf{P}_{(\underline{f}_0,x_0)}\left[X_t^\diamond = (\underline{f}, x), \tau^\diamond = t\right] = \prod_{v\in G}\pi_H(f(v))\mathbf{P}_{(\underline{f}_0,x_0)}\left[X_t = x, \tau^\diamond = t\right]. \tag{1.3.11}$$

If further τ_H has halting states then the vectors $(\underline{h}(f_0(v)), y)$ are halting state vectors for $\tau^\diamond$ for every $y \in G$.

We postpone the proof and continue with a corollary of the lemma:

Corollary 1.3.7. *Let τ_H be a strong stationary time for the Markov chain on H which has a halting state $h(d)$ for any $d \in H$. Then define $\tau^\diamond$ as in Lemma 1.3.6. Then for the separation distance on the lamplighter chain $H \wr G$ the following lower bound holds for any starting state $(\underline{f}_0, x_0)$:*

$$s_{(\underline{f}_0,x_0)}(t) \geq \mathbf{P}_{(\underline{f}_0,x_0)}\left[\tau^\diamond > t\right].$$

Proof. Observe that reaching the halting state vector $(\underline{h}(f_0(v)), x)$ implies the event $\tau^\diamond \leq t$ so we have

$$1 - \frac{\mathbf{P}_{(\underline{f}_0,x_0)}\left[X_t^\diamond = (\underline{h}(f_0(v)), x)\right]}{\pi_G(x)\prod_{v\in G}\pi_H\left(h(f_0(v))\right)} = 1 - \frac{\mathbf{P}_{(\underline{f}_0,x_0)}\left[X_t^\diamond = (\underline{h}(f_0(v)), x), \tau^\diamond \leq t\right]}{\pi_G(x)\prod_{v\in G}\pi_H\left(h(f_0(v))\right)} \tag{1.3.12}$$

Now pick a vertex $x_{x_0,t} \in G$ which minimizes $\mathbf{P}\left[X_t = x_{x_0,t} | \tau^\diamond \leq t\right]/\pi_G(x_{x_0,t})$. This quotient is less than 1 since both the numerator and the denominator are probability distributions on G. Then, using this and Lemma 1.3.6, the right hand side of (1.3.12) equals

$$1 - \frac{\mathbf{P}_{(\underline{f}_0,x_0)}\left[X_t = x_{x_0,t} | \tau^\diamond \leq t\right]\mathbf{P}_{(\underline{f}_0,x_0)}[\tau^\diamond \leq t]}{\pi_G(x_{x_0,t})} \geq 1 - \mathbf{P}_{(\underline{f}_0,x_0)}\left[\tau^\diamond \leq t\right].$$

Clearly the separation distance is larger than the left hand side of (1.3.12), and the proof of the claim follows. Note that the proof only works if τ_H has a halting state and thus it is separation-optimal. □

Lemma 1.3.8. *Let τ_H be the stopping time defined as in Lemma 1.3.6, and let $\tau_G(x)$ be a strong stationary time for G starting from $x \in G$ and define $\tau_2^\diamond$ by*

$$\tau_2^\diamond := \tau^\diamond + \tau_G(X_{\tau^\diamond}) \tag{1.3.13}$$

Then, $\tau_2^\diamond$ is a strong stationary time for $H \wr G$.

Proof of Lemma 1.3.6. First we show that (1.3.11) holds using the conditional independence of $\tau_H(v)$-s given the number of moves $L_v(t)$ on the lamp graphs $H(v), v \in G$. Clearly, conditioning on the trajectory of the walker $\{X_1, \ldots, X_{t-1}, X_t = x\} := X[1,t]$ contains the knowledge of $L_v(t)$-s as well. We will omit to note the dependence of $\mathbf{P}$ on initial state $(\underline{f}_0, x_0)$ for notational convenience. The left hand side of condition (1.3.6) equals

$$\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau^\diamond \leq t\right] = \sum_{X_{[1,t]}} \mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau^\diamond \leq t | X_{[1,t]}\right]\mathbf{P}\left[X_{[1,t]}\right].$$

Recall that Z stands for the Markov chain on the lamp graph H. Due to (1.3.6) and τ_H being strong stationary for H we have for all $v \in G$ that

$$\mathbf{P}[Z_{L_v(t)} = f(v), \tau_H(v) \leq L_v(t) | X_{[1,s]}] = \pi_H(f(v)) \cdot \mathbf{P}[\tau_H(v) \leq L_v(t) | X_{[1,t]}].$$

Now we use that $\tau_H(v)$-s are conditionally independent given the local times to see that

$$\begin{aligned}
&\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau^\diamond \le t | X_{[1,t]}\right] \\
&\quad = \mathbf{P}\left[\forall v \in G : Z_{L_v(t)} = f(v), \tau_H(v) \le L_v(t), X_t = x, |X_{[1,t]}\right] \\
&\quad = \prod_{v\in G} \pi_H(f(v)) \prod_{v\in G} \mathbf{P}\left[\tau_H(v) \le L_v(t) | X_{[1,t]}\right]
\end{aligned}$$

Note that the second product gives exactly $\mathbf{P}\left[\tau^\diamond \le t | X_{[1,t]}\right]$, yielding

$$\mathbf{P}\left[X_t = (\underline{f}, x), \tau^\diamond \le t\right] = \prod_{v\in G} \pi_H(f(v)) \sum_{X_{[1,t]}} \mathbf{P}\left[\tau^\diamond \le t | X_{[1,t]}\right] \mathbf{P}[X_{[1,t]}] \tag{1.3.14}$$

As $X_t = x$ remains fixed over the summation, summing over all possible trajectories yields

$$\mathbf{P}[X_t = (\underline{f}, x), \tau^\diamond \le t] = \prod_{v\in G} \pi_H(f(v)) \mathbf{P}[X_t = (\underline{f}, x), \tau^\diamond \le t].$$

To turn the inequality $\tau^\diamond \le t$ inside the probability to equality can be done the same way as in (1.3.6) and left to the reader. To see that the vector of halting states $(\underline{h}(f_0(v)), y)$ is a halting state for $\tau^\diamond$ for any $y \in G$ is based on the simple fact that reaching the halting state vector $(h(f(v))_{v\in G}, y)$ means that all the halting states $h(f(v)), v \in G$ have been reached on all the lamp graphs $H(v), v \in G$-s. Thus, by definition of the halting states, all the strong stationary times $\tau_H(v)$ have happened. Then, by its definition, $\tau^\diamond$ has happened as well. □

Proof of Lemma 1.3.8. The intuitive idea of the proof is based on the fact that τ_G is conditionally independent of τ_H-s and thus the lamp graphs stay stationary after reaching $\tau^\diamond$, and stationarity on G is reached by adding the term $\tau_G(X_{\tau^\diamond})$. The proof is not very difficult but it needs a delicate sequence of conditioning. To have shorter formulas, we write shortly $\mathbf{P}$ for $\mathbf{P}_{(\underline{f}_0, x_0)}$. First we condition on the events $\{\tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\}$ and make use of (1.3.11) from Lemma 1.3.6.

$$\begin{aligned}
\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t\right] = \sum_{s\le t; (\underline{g}, y)} &\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t | \tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\right] \\
&\cdot \prod_{v\in G} \pi_H(g(v)) \cdot \mathbf{P}\left[\tau^\diamond = s, X_s = y\right].
\end{aligned} \tag{1.3.15}$$

Now for the conditional probability inside the sum on the right hand side we have

$$\begin{aligned}
&\mathbf{P}\left[X_t^\diamond = (\underline{f}, x), \tau_2^\diamond = t | \tau^\diamond = s, X_s^\diamond = (\underline{g}, y)\right] \\
&= \mathbf{P}\left[X_t^\diamond = (\underline{f}, x); \tau_G(y) = t - s | X_s^\diamond = (\underline{g}, y)\right]
\end{aligned}$$

where we used in the last step that τ_G is only depending on y. We claim that

$$\begin{aligned}\sum_{\underline{g}} \left(\mathbf{P}_{(\underline{g},y)} \left[X^{\diamond}_{t-s} = (\underline{f}, x), \tau_G(y) = t-s \right] \prod_{v \in G} \pi_H(g(v)) \right) \\ = \mathbf{P}_y[X_{t-s} = x, \tau_G(y) = t-s] \prod_{v \in G} \pi_H(f(v)) \\ = \pi_G(x) \mathbf{P}_y[\tau_G = t-s] \prod_{v \in G} \pi_H(f(v)).\end{aligned}$$

The first equality holds true since $\tau_G(y)$ is independent of the lampgraphs and the transition rules of $X^{\diamond}$ on $H \wr G$ tells us that the lamp-chains stay stationary. We omit the details of the proof. The second equality is just the strong stationarity property of τ_G. Thus, using this and rearranging the order of terms on the right hand side of (1.3.15) we end up with

$$\sum_{s \leq t, y \in G} \mathbf{P}_y[\tau_G = t-s] \mathbf{P}[\tau^{\diamond} = s, X_s = y] \cdot \pi_G(x) \prod_{v \in G} \pi_H(f(v)).$$

Then, realizing that the sum is just $\mathbf{P}[\tau^{\diamond} + \tau_G(X_{\tau^{\diamond}}) = t]$ finishes the proof. □

We continue with a lemma which relates the separation distance to the total variation distance: Let us define first

$$d_x(t) := \|P^t(x, .) - \pi(.)\|_{\mathrm{TV}} = \frac{1}{2} \sum_{y \in \Omega} \left| P^t(x,y) - \pi(y) \right|. \tag{1.3.16}$$

The total variation distance of the chain from stationarity is defined as:

$$d(t) := \max_{x \in \Omega} d_x(t).$$

The next lemma relates the total and the separation distance:

Lemma 1.3.9. *For any reversible Markov chain and any state $x \in \Omega$, the separation distance from initial vertex x satisfies:*

$$d_x(t) \leq s_x(t) \tag{1.3.17}$$

$$s_x(2t) \leq 4d(t) \tag{1.3.18}$$

Proof. For a short proof of (1.3.17) see [49] or [83, Lemma 6.13], and combine [83, Lemma 19.3] with a triangle inequality to conclude (1.3.18). □

We will also make use of the following lemma: ([83, Corollary 12.6])

Lemma 1.3.10. *For a reversible, irreducible and aperiodic Markov chain,*

$$\lim_{t \to \infty} d(t)^{1/t} = \lambda_*,$$

with $\lambda_* = \max\{|\lambda| : \lambda$ eigenvalue of P, $\lambda \neq 1\}$.

A fundamental step to prove Lemma 1.3.10 is the inequality stating that for all $x \in \Omega$ we have

$$d_x(t) \leq s_x(t) \leq \frac{\lambda_*^t}{\pi_{\min}}, \tag{1.3.19}$$

with $\pi_{\min} = \min_{y\in\Omega} \pi(y)$. Follows from [83, Equation (12.11)].

We note that Lemma 1.3.9 implies that the assertion of Lemma 1.3.10 stays valid if we replace $d(t)^{1/t}$ by the separation distance $s(t)^{1/t}$.

1.3.4 Relaxation time bounds

Proof of the lower bound

We prove $c_1 = 1/(16 \log 2)$. First note that it is enough to prove that $t_{\text{hit}}(G)$ and $|G|t_{\text{rel}}(H)$ are both lower bounds, hence their average is a lower bound as well. First we start showing the latter.

Let us denote the second largest eigenvalue of Q by λ_H and the corresponding eigenfunction by ψ. It is clear that $\mathbf{E}_{\pi_H}(\psi) = 0$ and we can normalize it such that $\mathbf{Var}_{\pi_H}(\psi) = \mathbf{E}_{\pi_H}(\psi^2) = 1$ holds. Let us define

$$\phi : V(H \wr G) \to \mathbb{R}, \quad \phi((\underline{f}, x)) = \sum_{w\in G} \psi(f(w)),$$

thus ϕ is actually not depending on the position of the walker, only on the configuration of the lamps. Let $X_t^\diamond = (\underline{F}_t, X_t)$ be the lamplighter chain with stationary initial distribution $\pi^\diamond$. In the sequel we will calculate the Dirichlet form (1.3.4) for ϕ at time t, first conditioning on the path of the walker:

$$\begin{aligned} \mathcal{E}_t[\phi] &= \frac{1}{2}\mathbf{E}_{\pi^\diamond}[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2] \\ &= \frac{1}{2}\mathbf{E}_{\pi^\diamond}\big(\mathbf{E}_{\pi^\diamond}[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 | X_0, \dots X_t]\big) \end{aligned} \tag{1.3.20}$$

We remind the reader that in each step of the lamplighter walk, the state of the lamp graph $H(v)$ is refreshed both at the departure and arrival site of the walker. Thus, knowing the trajectory of the walker implies that we also know the number of steps made by the Markov chain $Z(v)$ on $H(v)$ equals $L_v(t)$, (twice the number of visits of the vertex $v \in G$). Moreover, conditioning on the number of moves made by $Z(v)$ on each $H(v), v \in G$, the collection of random walks $(Z(v))_{v\in G}$ on the lamp graphs are independent.

We can calculate the conditional expectation on the right hand side of (1.3.20) by using the argument above and the fact that $\mathbf{E}_{\pi_H}(\psi) = 0$ as follows:

$$\mathbf{E}_{\pi^\diamond}\left[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 | X_0, \dots X_t\right] = \sum_{w\in G} \mathbf{E}_{\pi^\diamond}\Big[\big(\psi(Z_{L_w(t)}) - \psi(Z_0)\big)^2 \big| L_w(t)\Big] \tag{1.3.21}$$

Next, the product form of the stationary measure $\pi^\diamond$ ensures that we can move to π_H inside the sum and condition on the starting state Z_0:

$$\mathbf{E}_{\pi^\diamond}\Big[\big(\psi(Z_{L_w(t)}) - \psi(Z_0)\big)^2 \big| L_w(t)\Big] = 2\mathbf{E}_{\pi_H}\psi^2 - 2\mathbf{E}_{\pi_H}\big[\psi(Z_0)\mathbf{E}_{Z_0}[\psi(Z_{L_w(t)})|Z_0]\big],$$

Since ψ was chosen to be the second eigenfunction for Q, clearly $\mathbf{E}_{Z_0}[\psi(Z_{L_w(t)})]$ $= \lambda_H^{L_w(t)}\psi(Z_0)$. Using the normalization $\mathbf{E}_{\pi_H}[\psi^2] = 1$, we arrive at

$$\mathbf{E}_{\pi^\diamond}\left[(\phi(X_t^\diamond) - \phi(X_0^\diamond))^2 \,|X_0, \dots X_t\right] = 2|G| - 2\sum_{w\in G}\lambda_H^{L_w(t)} \tag{1.3.22}$$

Since $\sum_{w\in G} L_w(t) = 2t$ and the function λ_H^x is convex, Jensen's inequality implies that

$$\sum_{w\in G}\lambda_H^{L_w(t)} \geq |G|\cdot\lambda_H^{2t/|G|}.$$

Combining this with (1.3.22) and (1.3.20) and setting $t := t^* = |G|t_{\mathrm{rel}}(H) = |G|/(1-\lambda_H)$ we arrive at

$$\mathcal{E}_t(\phi) \leq |G|\left(1 - e^{2\frac{\log\lambda_H}{1-\lambda_H}}\right) \leq |G|\left(1 - 2^{-4-1}\right),$$

where in the last step we assumed $\lambda_H > 1/2$, since in this case we have $(1-\lambda_H)^{-1}\log\lambda_H > -2\log 2$. On the other hand, if $\lambda_H < 1/2$, than we mix immediately after visiting the vertex w, thus, we will use the other lower bound $t_{\mathrm{hit}}(G)$. Dividing by $\mathbf{Var}_{\pi^\diamond}\phi = |G|$, and using the variational characterization of the spectral gap in Lemma 1.3.4 yields

$$\gamma_{t^*} \leq 1 - 2^{-5}.$$

Since γ_t is by definition the spectral gap of the chain at time t, we have

$$1 - \lambda_2(H \wr G)^{t^*} \leq 1 - 2^{-5}. \tag{1.3.23}$$

Thus

$$5\log 2 \geq t^*\left(1 - \lambda_2(H\wr G)\right),$$

so we get a lower bound $t_{\mathrm{rel}}(H\wr G) \geq \frac{1}{5\log 2}|G|t_{\mathrm{rel}}(H)$.

To get the lower bound $t_{\mathrm{hit}}(G)/4$ we adjust the proof for $0-1$ lamps ($H = \mathbf{Z}_2$) [83, Theorem 19.1] to our setting. First pick a vertex $w \in G$ which maximizes the expected hitting time $E_\pi(\tau_w)$. Similarly as before, we will use the second eigenfunction ψ with eigenvalue λ_H with $\mathbf{E}_\pi(\psi) = 0, \mathbf{E}_\pi(\psi^2) = 1$ and define

$$\phi\left((\underline{f}, x)\right) := \psi(f(w)).$$

Easy to see with the same conditioning argument we used in (1.3.21) and (1.3.22) that the Dirichlet form at time t equals

$$\mathcal{E}_t(\phi) = 1 - \mathbf{E}\left[\lambda_H^{L_w(t)}\right]$$

Now we will show that $\mathbf{E}\left[\lambda_H^{L_w(t)}\right] \geq 1/4$. To see this we first note that for any t we have

$$\mathbf{E}_v(\tau_w) \leq t + t_{\text{hit}}\mathbf{P}_v[\tau_w > t]$$
$$\mathbf{E}_\pi(\tau_w) \leq t + t_{\text{hit}}\mathbf{P}_\pi[\tau_w > t]$$

To see the first line: either the walk hits w before time t, or the expected additional time it takes to arrive at w is bounded by t_{hit} regardless of where it is at time t. The second line follows by averaging over π_H.

Next, [83, Lemma 10.2] states that $t_{\text{hit}} \leq 2\max_v \mathbf{E}_\pi[\tau_v]$ holds for every irreducible Markov chain. We exactly picked w such that it maximizes $\mathbf{E}_\pi(\tau_v)$, so we have $t_{\text{hit}} \leq 2\mathbf{E}_\pi[\tau_w]$, so multiplying the previous equation by 2 and combining gives

$$t_{\text{hit}} \leq 2t + 2t_{\text{hit}}\mathbf{P}_\pi[\tau_w > t]$$

Now substituting $t = t_{\text{hit}}/4$ and rearranging terms results in

$$\mathbf{P}_\pi\left[\tau_w > \frac{t_{\text{hit}}}{4}\right] \geq \frac{1}{4}.$$

Since $\{L_w(t_{\text{hit}}/4) = 0\} = \{\tau_w > t_{\text{hit}}/4\}$, we can use this inequality to obtain the upper bound

$$\mathcal{E}_{t_{\text{hit}}/4}(\phi) = 1 - \mathbf{E}_{\pi^\diamond}\left[\lambda_H^{L_w(t_{\text{hit}}/4)}\right] \leq 1 - \mathbf{P}[\tau_w > t_{\text{hit}}/4] \leq 1 - \frac{1}{4} = \frac{3}{4}.$$

Analogous to the last lines of the proof of the lower bound above, (see (1.3.23)) we obtain the other desired lower bound:

$$t_{\text{rel}}(H) \geq \frac{1}{2\log 2}\frac{1}{4}t_{\text{hit}}(G).$$

Putting together the two bounds we get

$$\begin{aligned} t_{\text{rel}}(H \wr G) &\geq \max\left\{\frac{1}{8\log 2}t_{\text{hit}}(G), \frac{1}{5\log 2}|G|t_{\text{rel}}(H)\right\} \\ &\geq \frac{1}{16\log 2}\left(t_{\text{hit}}(G) + |G|t_{\text{rel}}(H)\right). \end{aligned}$$

Proof of the upper bound

To prove the upper bound, we will estimate the tail behavior of the strong stationary time $\tau_2^\diamond$ in Lemma 1.3.8, relate it to $s^\diamond(t)$, the separation distance on $H \wr G$, and then use Lemmas 1.3.10 and 1.3.9 to see that $s^\diamond(t)^{1/t} \to \lambda_2$. Use separation-optimal τ_H and τ_G in the construction of $\tau_2^\diamond$. The existence is guaranteed by Theorem 1.3.5.

So, combining (1.3.8) and the fact that $\tau^\diamond$ happens when all the stopping times $\tau_H(v), v \in G$ have happened on the lamp graphs, by union bound we have for any choice of $0 < \alpha < 1$

$$\begin{aligned} s^\diamond_{(\underline{f},x)}(t) &\le \mathbf{P}_{(\underline{f},x)}\left[\tau_2^\diamond > t\right] \le \mathbf{P}_{(\underline{f},x)}\left[\tau^\diamond > \alpha t\right] + \mathbf{P}_{(\underline{f},x)}\left[\tau_2^\diamond > t | \tau^\diamond < \alpha t\right] \\ &\le \mathbf{P}[\tau_{\mathrm{cov}} > \alpha t/3] && (1.3.24) \\ &+ \mathbf{P}\left[\exists w \in G : L_w(\alpha t) < \tfrac{\alpha t}{2|G|} \big| \tau_{\mathrm{cov}} \le \alpha t/3\right] && (1.3.25) \\ &+ \mathbf{P}\left[\exists w \in G : \tau_H(w) \ge L_w(\alpha t) \ \big| \forall v \in G : L_v(\alpha t) \ge \frac{\alpha t}{2|G|}\right] && (1.3.26) \\ &+ \max_{(\underline{g},y)} \mathbf{P}_{(\underline{g},y)}\left[\tau_G > (1-\alpha)t\right] && (1.3.27) \end{aligned}$$

Namely, there are four possibilities: The first option is that there is a state $w \in G$ which is not hit yet, i.e. the cover time of the chain is greater than $\alpha t/3$: giving the term (1.3.24). The constant $1/3$ could have been chosen differently, we picked $\alpha t/3$ such that the remaining $2\alpha t/3$ time is still enough to gain large enough local time on the vertices $v \in G$. Secondly, even though any state w on the graph G is reached before time $\alpha t/3$, the remaining time was not enough to have at least $\alpha t/2|G|$ many moves on some lamp graph $H(w)$, term (1.3.25). The third option is that even though there have been done many moves on all the lamp graphs, there is a vertex $w \in G$ where $\tau_H(w)$ has not happened yet, yielding the term (1.3.26). We will handle the three terms separately. The fourth term handles the case where the strong stationary time τ_G is too large. (We will drop the factor α for convenience in the first three formulas.)

We can estimate the first term (1.3.24) by a union bound:

$$\mathbf{P}[\tau_{\mathrm{cov}} > t/3] \le \mathbf{P}[\exists w : \tau_w > t/3] \le |G| 2 e^{-\frac{\log 2}{6}\frac{t}{t_{\mathrm{hit}}}}, \tag{1.3.28}$$

where t_{hit} is the maximal hitting time of the graph G, see (1.2.4). To see this, use Markov's inequality on the hitting time of $w \in G$ to obtain that for all starting states $v \in G$ we have $\mathbf{P}_v[\tau_w > 2t_{\mathrm{hit}}] \le 1/2$, and then run the chain in blocks of $2t_{\mathrm{hit}}$. In each block we hit w with probability at least $1/2$, so we have

$$\mathbf{P}_v[\tau_w > K(2t_{\mathrm{hit}})] \le \frac{1}{2^K}.$$

To get it for general t, we can move from $\lfloor t/t_{\mathrm{hit}} \rfloor$ to t/t_{hit} by adding an extra factor of 2, and (1.3.28) immediately follows by a union bound.

For the third term (1.3.26) we claim the following upper bound holds:

$$\mathbf{P}\left[\exists w : \tau_H(w) \ge L_w(t) \big| \forall v : L_v(t) > \tfrac{t}{2|G|}\right] \le |G| \frac{1}{\pi_{\min}(H)} e^{-\frac{t}{2|G| t_{\mathrm{rel}}(H)}}. \tag{1.3.29}$$

To see this we estimate the probability of the event $\{\tau_H(w) \geq L_w(t) \,\big|L_w(t) \geq \frac{t}{2|G|}\}$ on a single lamp graph and then use union bound to loose the factor $|G|$ on the right hand side. First note that according to Lemma 1.3.10, the tail of the strong stationary times τ_H are driven by λ_H^t. More precisely, using the inequality (1.3.19) we have that for any initial state $h \in H$:

$$\begin{aligned}\mathbf{P}_h\left[\tau_H(w) \geq \frac{t}{2|G|}\right] &\leq s_H\left(\tfrac{t}{2|G|}\right) \leq \frac{1}{\pi_{\min}(H)}\lambda_H^{t/2|G|}. \\ &\leq \frac{1}{\pi_{\min}(H)}\exp\left\{-\frac{(1-\lambda_H)t}{2|G|}\right\}.\end{aligned}$$

Since we have made at least $L_w(t) \geq \frac{t}{2|G|}$ steps on each coordinate, the claim (1.3.29) follows. The fourth term (1.3.27) can be handled analogously and yields an upper bound $\exp\{-c(1-\lambda_G)t\}$ which would lead to a term of order $t_{\text{rel}}(G)$, clearly dominated by $t_{\text{hit}}(G)$.

The intuition behind the estimates below for the second term (1.3.25) is that since the total time was at least $2t/3$ after hitting, regularity of G implies that the average number of moves on each lamp graph equals $4t/3|G|$ by the double refreshment at any visit to the vertex. Thus, the probability of having less than $t/2|G|$ moves must be small.

More precisely, we introduce the excursion-lengths to a vertex $w \in G$: Let us define for all $w \in G$ the first return time to state w as

$$R(w) = \inf\{t > 0 : X_t = w | X_0 = w\}.$$

The strong Markov property implies that the length of the i-th excursion $R_i(w)$, defined as the time spent between the $(i-1)$th and ith visit to w, are i.i.d random variables distributed as the first return time $R(w)$.

Thus, having not enough local time on some site $w \in G$ can be expressed in terms of the excursion lengths $R_i(w)$-s as follows:

$$\mathbf{P}\left[\exists w : L_w(t) \leq \tfrac{t}{2|G|} \big| t_{\text{cov}} \leq \frac{t}{3}\right] \leq |G|\mathbf{P}_w\left[\sum_{i=1}^{t/4|G|} R_i(w) \geq \frac{2t}{3}\right], \quad (1.3.30)$$

since conditioning on hitting before $t/3$ ensures that we had at least $2t/3$ steps to gain the $t/4|G|$ visits to w, and by the definition (1.3.9) of $L_v(t)$, this guarantees that $L_w(t) < t/2|G|$.

We aim to estimate the right hand side of (1.3.30) using the moment generating function of the first return time $R(w)$. To be able to carry out the estimates we need a tail behavior on the return times. A very similar argument can be used than the one we used for the tail of the cover time (1.3.28), namely the following holds:

$$\begin{aligned}\mathbf{P}_w\left[R(w) > 2t_{\text{hit}}\right] &= \mathbf{P}_w\left[X_1 \neq w\right]\mathbf{E}\left[\mathbf{P}_{X_1}(\tau_w > 2t_{\text{hit}} | X_1)\right] \\ &\leq \frac{\mathbf{E}[\mathbf{E}_{X_1}(\tau_w)]}{2t_{\text{hit}}} \leq \frac{1}{2}.\end{aligned}$$

Running the chains in blocks of $2t_{\rm hit}$, one can see that in each block the chain has a chance at least $1/2$ to return to w, so we have for each $t > 2t_{\rm hit}$

$$\mathbf{P}\left[R_w > t\right] \le 2\left(\frac{1}{2}\right)^{\frac{t}{2t_{\rm hit}}} = 2\exp\left\{-\frac{\log 2}{2}\frac{t}{t_{\rm hit}}\right\}, \tag{1.3.31}$$

where the factor 2 comes from ignoring to take the integer part of $t/t_{\rm hit}$.

We can use this tail behavior to estimate the moment generating function

$$\begin{aligned}\mathbf{E}\left[e^{\beta R_w}\right] &\le e^{2\beta|G|} + \mathbf{E}[e^{\beta R_w}\mathbf{1}\{R_w > 2|G|\}] \\ &= e^{2\beta|G|} + \int_{e^{2\beta|G|}}^{\infty} \mathbf{P}[e^{\beta R_w} > z]\mathrm{d}z\end{aligned}$$

where we cut the expectation at $2|G|$. Changing variables and using the bounds in (1.3.31) yields:

$$\begin{aligned}\mathbf{E}\left[e^{\beta R_w}\right] &\le e^{2\beta|G|} + \int_{e^{2\beta|G|}}^{\infty} \mathbf{P}[R_w > \frac{1}{\beta}\log z]\mathrm{d}z \\ &\le e^{2\beta|G|} + 2\int_{e^{2\beta|G|}}^{\infty} z^{-\frac{\log 2}{2\beta t_{\rm hit}}}\mathrm{d}z.\end{aligned}$$

Setting arbitrary $\beta < \log 2/(2t_{\rm hit})$ makes the second term integrable, and with the special choice of $\beta = \frac{\log 2}{4t_{\rm hit}}$ we obtain the following estimate:

$$\mathbf{E}\left[e^{\beta R_w}\right] \le e^{2\beta|G|} + 2e^{-2\beta|G|} \le e^{(2+\delta)\beta|G|} \tag{1.3.32}$$

with an appropriately chosen $0 < \delta < 1/3$. Now we apply Markov's inequality to the function $e^{\beta x}$ to estimate the right hand side of (1.3.30):

$$\mathbf{P}_w\left[\sum_{i=1}^{t/4|G|} R_i(w) \ge 2t/3\right] \le e^{-\frac{2}{3}\beta t}\cdot\mathbf{E}\left[e^{\beta R_w}\right]^{\frac{t}{4|G|}}, \tag{1.3.33}$$

where we also used the independence of the excursions $R_i(w)$-s. Using the estimate in (1.3.32) to bound the right hand side we gain that

$$\begin{aligned}\mathbf{P}_w\left[\sum_{i=1}^{t/4|G|} R_i(w) \ge 2t/3\right] &\le e^{-\frac{2}{3}\beta t}\cdot e^{(2+\delta)\beta|G|\cdot\frac{t}{4|G|}} \\ &\le e^{-\frac{1-\tilde{\delta}}{6}\beta t} = \exp\left\{-\frac{(1-\tilde{\delta})\log 2}{24}\frac{t}{t_{\rm hit}}\right\},\end{aligned} \tag{1.3.34}$$

where we used $\beta = \log 2/(4t_{\rm hit})$, and modified $\tilde{\delta} := 3\delta/2 \le 1/2$. Using the relation of the local time to the excursion lengths in (1.3.30) we finally get that the second term (1.3.25) is bounded from above by

$$\mathbf{P}\left[\exists w : L_w(t) \le \frac{t}{2|G|}\Big|t_{\rm cov} \le t/3\right] \le |G|\exp\left\{-\frac{\log 2}{48}\frac{t}{t_{\rm hit}}.\right\} \tag{1.3.35}$$

Mind that all the estimates (1.3.28), (1.3.29) and (1.3.35) were independent of the initial state $(\underline{f}, x) \in H \wr G$, so maximizing over all possible initial states yields us

$$|\lambda_2|^t \leq 2d^\diamond(t) \leq 2s^\diamond(t) \leq \frac{4|G|}{\pi_{\min}(H)} \exp\left\{-\frac{t}{2|G|t_{\rm rel}(H)}\right\} + 4\exp\left\{-\frac{(1-\tilde{\delta})\log 2}{24}\frac{t}{t_{\rm hit}}\right\} + 4|G|\exp\left\{-\frac{\log 2}{6}\frac{t}{t_{\rm hit}}\right\} \tag{1.3.36}$$

In the final step we apply Lemma 1.3.10: we take the power $1/t$ and limit as t tends to infinity with fixed graph sizes $|G|$ and $|H|$ on the right hand side of (1.3.4) to get an upper bound on λ_2. Then we use that $(1-e^{-x}) \leq 1/x+O(1)$ and obtain the bound on $t_{\rm rel}$ finally:

$$t_{\rm rel}(H \wr G) \leq \max\left\{2|G|t_{\rm rel}(H), \frac{48}{\log 2}t_{\rm hit}\right\}.$$

This finishes the proof of the upper bound on the relaxation time.

1.3.5 Mixing time bounds

Once H has a strong stationary time τ_H with halting states, the idea of the proofs is based on relating the separation distance to the tail behavior of the stopping times $\tau^\diamond$ and $\tau_2^\diamond$ constructed in Lemmas 1.3.6 and 1.3.8, respectively, then turn the estimates to estimates for the total variation distance using the relations in Lemma 1.3.9. For the lower bound in the general case, we will need slightly different methods.

Proof of the upper bound

The idea of the proof is to use appropriate quantiles of the strong stationary time τ_H on H, and give an upper bound on the tail of the strong stationary time $\tau_2^\diamond$ defined in Lemma 1.3.8. Throughout, we (only) need that τ_H and τ_G in the construction of $\tau_2^\diamond$ are separation-optimal. The existence is guaranteed by Theorem 1.3.5. (Thus, τ_H does not necessarily must have halting states.)

We start with the definition of the blanket time:

$$\mathcal{B}_2 := \inf_t \left\{\forall v, w \in G : \frac{L_w(t)}{L_v(t)} \leq 2\right\}. \tag{1.3.37}$$

Let us denote

$$\mathbf{B}_2 := \max_{y\in\Omega} \mathbf{E}_y(\mathcal{B}_2) \tag{1.3.38}$$

It is known from [56] that there exist universal constants C and C' such that $C't_{\rm cov} \leq \mathbf{B}_2 \leq Ct_{\rm cov}$.

Thus, our first goal is to show that at time

$$t^{\diamond} = 8\mathbf{B}_2 + |G| t^{\mathrm{quant}}_{1/16|G|}(\tau_H) + t^{\mathrm{quant}}_{1/16}(\tau_G) := 8\mathbf{B}_2 + |G| t^u_H + t_G$$

we have for any starting state $(\underline{f}, x)$ that

$$\mathbf{P}_{(\underline{f},x)}[\tau_2^{\diamond} > t^{\diamond}] \leq \frac{1}{4}. \tag{1.3.39}$$

We remind the reader that $\tau_2^{\diamond} = \tau^{\diamond} + \tau_G(X_{\tau^{\diamond}})$ and thus the following union bound holds:

$$\begin{aligned} \mathbf{P}\left[\tau_2^{\diamond} > t^{\diamond}\right] \leq \mathbf{P}\left[\mathcal{B}_2 > 8\mathbf{B}_2\right] + \mathbf{P}\left[\tau^{\diamond} > |G| t^u_H + 8\mathbf{B}_2 | \mathcal{B}_2 \leq 8\mathbf{B}_2\right] \\ + \max_{v \in G} \mathbf{P}_v\left[\tau_G > t_G | \mathcal{B}_2 \leq 8\mathbf{B}_2, \tau^{\diamond} < 8\mathbf{B}_2 + |G| t^u_H\right] \end{aligned} \tag{1.3.40}$$

The first term on the right hand side is less than $1/8$ by Markov's inequality, the third is less than $1/16$ by the definitoon of the worst case quantile. The second term can be handled by conditioning on the local time sequence of vertices and of the blanket time: (for shorter notation we introduce $t_1 := |G| t^u_H + 8\mathbf{B}_2$)

$$\begin{aligned} &\mathbf{P}\left[\tau^{\diamond} > |G| t^u_H + 8\mathbf{B}_2 | \mathcal{B}_2 \leq 8\mathbf{B}_2\right] \leq \\ &= \sum_{s \leq 8\mathbf{B}_2, (L_v(t_1))_v} \mathbf{P}\left[\exists w : \{\tau_H > L_w(t_1)\} \middle| \left(L_v(t_1)\right)_v, \mathcal{B}_2 = s\right] \cdot \mathbf{P}\left[\left(L_v(t_1)\right)_v, \mathcal{B}_2 = s\right] \end{aligned} \tag{1.3.41}$$

The fact that $\mathcal{B}_2 \leq 8\mathbf{B}_2$ means that the number of visits of every vertex $v \in G$ must be greater than half of the average, which is at least $\frac{1}{2} t^u_H$. Since $L_v(t)$ is twice the number of visits by (1.3.9), $\{\tau_H > L_w(t_1)\} \subseteq \{\tau_H > t^u_H\}$. By the definition of the quantiles,

$$\mathbf{P}_h\left[\tau_H > t^u_H\right] \leq \frac{1}{16|G|}$$

holds for every $h \in H$ and $v \in G$, respectively. Applying a simple union bound on the conditional probability on the right hand side of (1.3.41) yields

$$\begin{aligned} \mathbf{P}_{(\underline{f},x)}\left[\tau^{\diamond} > t_1 | \mathcal{B}_2 \leq 8\mathbf{B}_2\right] &\leq \sum_{s \leq 8\mathbf{B}, (L_v(t_1))_v} \left(|G| \frac{1}{16|G|}\right) \mathbf{P}\left[\left(L_v(t_1)\right)_v, \mathcal{B}_2 = s\right] \\ &\leq \frac{1}{16}, \end{aligned}$$

where we used that the sum of the probabilities on the right hand side is at most 1. Combining these estimates with (1.3.40) yields (1.3.39). It remains to relate the worst-case quantiles to the total variation mixing times. Here we will make use of the separation-optimal property of τ_H and τ_G. Now just consider the walk on G. Let us start the position of the walker on G from

an initial state $x_0 \in G$ for which the maximum is attained in the definition (1.3.45) of the quantile $t^{\text{quant}}_{1/16}(\tau_G)$. Then, by (1.3.18) we have that one step before the quantile we have

$$\begin{aligned}\frac{1}{16} \le \mathbf{P}_{x_0}\left[\tau_G > t^{\text{quant}}_{1/16}(\tau_G) - 1\right] &= s_{x_0}\left(t^{\text{quant}}_{1/16}(\tau_G) - 1\right) \\ &\le 4d\left(\frac{1}{2}(t^{\text{quant}}_{1/16}(\tau_G) - 1)\right).\end{aligned}$$

This immediately implies that $\frac{1}{2}(t^{\text{quant}}_{1/16}(\tau_G) - 1) \le t_{\text{mix}}\left(G, \frac{1}{64}\right)$. By the sub-multiplicative property of the total variation distance $d(kt) \le 2^k d(t)^k$ we have that $t_{\text{mix}}(G, \frac{1}{64}) \le 6t_{\text{mix}}\left(G, \frac{1}{4}\right)$. So we arrive at

$$t^{\text{quant}}_{1/16}(\tau_G) - 1 \le 12 t_{\text{mix}}(G) \tag{1.3.42}$$

Similarly, starting all the lamps from the position h_0 where the maximum is attained in the definition of $t^u_H = t^{\text{quant}}_{1/16|G|}(\tau_H)$, one step before the quantile we have

$$\frac{1}{16|G|} \le \mathbf{P}_{h_0}\left[\tau_H > t^u_H - 1\right] = s_{h_0}\left(t^u_H - 1\right) \le 4d\left((t^u_H - 1)/2\right)$$

So we have

$$\frac{1}{2}(t^{\text{quant}}_{1/16|G|}(\tau_H) - 1) \le t_{\text{mix}}\left(H, \tfrac{1}{64|G|}\right). \tag{1.3.43}$$

On the other hand, on the whole lamplighter chain $H \wr G$ we need the other direction: For every starting state $(\underline{f}, x)$ (1.3.17) and (1.3.39) implies that

$$d_{(\underline{f},x)}(t) \le s_{(\underline{f},x)}(t) \le \mathbf{P}_{(\underline{f},x)}\left[\tau^\diamond_2 > t^\diamond\right] \le 1/4$$

Maximizing over all states $(\underline{f}, x)$ yields

$$t_{\text{mix}}(H \wr G) \le t^\diamond. \tag{1.3.44}$$

Putting the estimates in (1.3.42) and (1.3.43) to (1.3.44), we get that

$$t_{\text{mix}}(H \wr G) \le t^\diamond \le 8\mathbf{B}_2(G) + 12 t_{\text{mix}}(G) + 1 + 2|G|\left(t_{\text{mix}}\left(H, \frac{1}{64|G|}\right) + \frac{1}{2}\right).$$

Since $\mathbf{B}_2(G) \le C t_{\text{cov}}(G)$, and $t_{\text{mix}}(G) \le 2t_{\text{hit}} \le 2t_{\text{cov}}(G)$ for any G, the assertion of Theorem 1.3.3 follows with $C_2 = 8(C+3)$, where C is the universal constant relating the blanket time $\mathbf{B}_2$ to the cover time t_{cov} in [56].

We remark why we did not make the constant C_2 explicit: If the blanket time $\mathbf{B}_2$ were not used in our estimates, the error probability that some vertex $w \in G$ does not have enough local time would need to be added. This, similarly as the term (1.3.25) behaves as $|G|e^{-c\left(t_{\text{cov}} + |G| t_{\text{mix}}(H, \frac{1}{G})\right)/t_{\text{hit}}}$.

If we do not assume anything about the relation of t_{hit} and t_{cov} and on $t_{\text{mix}}(H, \frac{1}{G})$, then this error term will not necessarily be small. For example, if G_n is a cycle of length n, H_n is a sequence of expander graphs, then $t_{\text{cov}} = t_{\text{hit}} = \Theta(n^2)$, and $t_{\text{mix}}(H, \frac{1}{G}) = \log|H| \cdot \log|G| = \log|H| \log n$, and we see that the term is not small if $\log|H| = o(n/\log n)$.

Proof of the lower bound

Similarly as we did with the relaxation time, it is enough to prove that all the bounds are lower bounds separately, then take averaging. First we start showing that the upper bound is sharp in 1.3.3 under the assumption that there is a strong stationary time τ_H with halting states.

Lower bound under Assumption (A)

Consider the stopping time $\tau^\diamond$ constructed in Lemma 1.3.6. Corollary 1.3.7 tells us that the tail of $\tau^\diamond$ lower bounds the separation distance at time t. We again emphasize that this bound holds only if τ_H in the construction of $\tau^\diamond$ is not only separation optimal but it also *has a halting state*. Our first goal is to lower bound the tail of $\tau^\diamond$, then relate it to the total variation distance.

We will denote the worst-case initial state upper ε-quantile of a stopping time τ as

$$t_\varepsilon^{\text{quant}}(\tau) := \max_{y\in\Omega} \inf\{t : \mathbf{P}_y[\tau > t] \le \varepsilon\} \tag{1.3.45}$$

First set

$$t_H := t^{\text{quant}}_{|G|^{-1/2}/2}(\tau_H) - 1, \quad t^\diamond := \frac{1}{4}|G| t_H, \tag{1.3.46}$$

clearly this can be done if $t^{\text{quant}}_{|G|^{-1/2}/2}(\tau_H) \neq 1$. We will handle the case if it equals 1 separately. We can estimate the upper tail of $\tau^\diamond$ by conditioning on the number of moves on the lamp graphs $H(v), v \in G$:

$$\begin{aligned}\mathbf{P}\left[\tau^\diamond > t^\diamond\right] &\ge \mathbf{P}\left[\exists w \in G : \tau_H(w) > L_w(t^\diamond)\right] \\ &\ge \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in G : \tau_H(w) > L_w(t^\diamond) \middle| (L_v(t^\diamond))_v\right] \mathbf{P}\left[(L_v(t^\diamond))_v\right]\end{aligned} \tag{1.3.47}$$

For each sequence $(L_v(t^\diamond))_{v\in G}$ we define the random set

$$S_{(L_v)_v} := \{w \in G : L_w(t^\diamond) \le t_H\}$$

Since $\sum_v L_v(t^\diamond) = 2t^\diamond = \frac{1}{2}|G| t_H$, we have that for arbitrary local time configuration $(L_v(t^\diamond))_v$,

$$|S_{(L_v)_v}| \ge |G|/2. \tag{1.3.48}$$

Thus we can lower bound (1.3.47) by restricting the event only to those $w \in G$ coordinates which belong to this set, i.e. their local time is small:

$$\begin{aligned}\mathbf{P}\left[\tau^\diamond > t^\diamond\right] &\geq \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > L_w(t^\diamond) \middle| \left(L_v\left(t^\diamond\right)\right)_v\right] \mathbf{P}\left[\left(L_v(t^\diamond)\right)_v\right] \\ &\geq \sum_{(L_v(t^\diamond))_v} \mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > t_H \middle| \left(L_v(t^\diamond)\right)_v\right] \mathbf{P}\left[\left(L_v(t^\diamond)\right)_v\right],\end{aligned} \tag{1.3.49}$$

where in the second line we used that for $w \in S_{(L_v)_v}$ we have $\{\tau_H(w) > L_w(t^\diamond)\} \supseteq \{\tau_H(w) > t_H\}$. Conditioned on the sequence $(L_v(t^\diamond))_v$, the times $\tau_H(w)$ for $w \in S_{(L_v)_v}$ are independent. Moreover, on each lamp graph $H(v)$ let us pick the starting state to be $h_0 \in H$ where the maximum is attained in the definition of $t^{\text{quant}}_{|G|^{-1/2}/2}(\tau_H)$. Since t_H is one step *before* the quantile, we have

$$\mathbf{P}_{h_0}\left[\tau_H(w) > t^{\text{quant}}_{|G|^{-1/2}}(\tau_H) - 1\right] \geq |G|^{-1/2}/2. \tag{1.3.50}$$

We need to start the lamp-chains from the worst-case scenario $h_0 \in H$ for two reasons: First, we needed to define the quantile as in (1.3.45) to be able to relate it to the total variation mixing time on H, see below. Then, the fact that $t^{\text{quant}}_{\varepsilon}$ was defined as the worst-case starting state quantile means that for other starting states the quantile may be smaller, and the lower bound can possibly fail.

Combining (1.3.50) with (1.3.48) and the conditional independence gives us the stochastic domination from below to the event in (1.3.49)

$$\mathbf{P}\left[\exists w \in S_{(L_v)_v} : \tau_H(w) > t_H \middle| (L_w(t^\diamond))_w\right] \geq \mathbf{P}[Z > 0]$$

where Z is a Binomial random variable with parameters $\left(\frac{|G|}{2}, \frac{|G|^{-1/2}}{2}\right)$. Clearly for $|G| > 8 > 16(\log 2)^2$ we have

$$\mathbf{P}\left[Z > 0\right] = 1 - \left(1 - \frac{1}{2|G|^{1/2}}\right)^{|G|/2} \geq 1 - e^{-\frac{|G|^{1/2}}{4}},$$

Combining this with (1.3.49) and summing over all possible $(L_v(t^\diamond))_{v \in G}$ sequences we easily get that

$$\mathbf{P}\left[\tau^\diamond > t^\diamond\right] \geq 1 - e^{-\frac{|G|^{1/2}}{4}}.$$

Then, by Corollary 1.3.7 we have

$$s^\diamond_{(\underline{h}_0, x)}(t^\diamond) \geq 1 - e^{-\frac{|G|^{1/2}}{4}}.$$

In the next few steps we relate the tail of $\tau^\diamond$ and τ_H to the mixing time of the graphs. First, combining the previous inequality with (1.3.18) implies that for the starting state $(\underline{h}_0, x)$ the following inequalities hold:

$$1 - e^{-|G|^{1/2}/4} \leq s^\diamond_{(\underline{h}_0, x)}(t^\diamond) \leq 4d^\diamond(t^\diamond/2).$$

These immediately imply

$$t_{\mathrm{mix}}(H \wr G, \frac{1}{8}) \geq \frac{1}{2} t^{\diamond} = \frac{1}{8}|G| t_H \tag{1.3.51}$$

Now we will relate $t_H = t^{\mathrm{quant}}_{|G|^{-1/2}/2}(\tau_H) - 1$ to the mixing time on H. Since t_H investigates the worst case initial-state scenario, by inequality (1.3.8) for any starting state $h \in H$ we have

$$s_h(t_H + 1) \leq \mathbf{P}_h \left[\tau_H \geq t_H + 1\right] \leq |G|^{-1/2}/2$$

Using $d_h(t) \leq s_h(t)$ (see Lemma 1.3.9) and maximizing over all $h \in H$ we get that

$$d_H(t_H + 1) \leq |G|^{-1/2}/2. \tag{1.3.52}$$

On the other hand, the total variation distance for any reversible Markov chain has the following sub-multiplicative property for any integer k, see [83, Section 4.5]:

$$d(kt) \leq 2^k d(t)^k. \tag{1.3.53}$$

Taking $t = t_H + 1$ and combining with (1.3.52) we have that

$$d_H(2(t_H + 1)) \leq 4 d_H(t_H + 1)^2 \leq 4 \frac{1}{4|G|},$$

which immediately implies

$$t_{\mathrm{mix}}(H, 1/|G|) \leq 2(t_H + 1).$$

Combining this with (1.3.51) yields the desired lower bound:

$$\frac{1}{16}|G| \left(t_{\mathrm{mix}} \left(H, \tfrac{1}{|G|} \right) - 2 \right) \leq t_{\mathrm{mix}}(H \wr G, \frac{1}{8}).$$

Mind that the term -2 in the brackets can be dropped when picking a possibly smaller constant and take the graph large enough. The case when $t^{\mathrm{quant}}_{|G|^{-1/2}/2}(\tau_H) = 1$ can be handled the following way: first mind that we can exchange the quantile for arbitrary $0 < \alpha < 1$, and look the proof with $t^{\mathrm{quant}}_{|G|^{-\alpha}/2}(\tau_H)$. If this is still $= 1$ for all α, that means that $\tau_H \equiv 1$. In this case, it is enough to hit the vertices to mix immediately and thus the mixing time $|G| t_{\mathrm{mix}}(H)$ is of smaller order than the cover time $t_{\mathrm{cov}}(G)$. The case when $|G| \leq 8$ but $|H| \to \infty$ is easy to see since in this case $t_{\mathrm{mix}}(H, \frac{1}{|G|}) \leq 2 t_{\mathrm{mix}}(H)$ and one can argue that mixing on $H \wr G$ requires mixing on a single lampgraph H_w for a fixed $w \in G$. Thus the lower bound remains valid.

The cover time is a lower bound for the $0-1$ lamps case, but for completeness we adjust the proof in [83, Theorem 19.2] to our setting. By Lemma

1.3.6 we can estimate the separation distance on $H \wr G$ as

$$\begin{aligned} s^{\diamond}_{(\underline{f},x)}(t) &= \mathbf{P}_{(\underline{f},x)}\left[\tau^{\diamond} > t\right] \\ &\geq \mathbf{P}_{(\underline{f},x)}\left[\exists w \in G : \tau_H(w) > L_w(t)\right] \\ &\geq \mathbf{P}_{(\underline{f},x)}\left[\exists w \in G : L_w(t) = 0\right] = \mathbf{P}_{(\underline{f},x)}\left[\tau_{\mathrm{cov}} > t\right]. \end{aligned} \tag{1.3.54}$$

Now, using the submultiplicativity of $d(t)$ in (1.3.53) and the relation of the separation distance and the total variation distance in (1.3.18), we have that at time $4t_{\mathrm{mix}}(H \wr G, 1/4)$:

$$s^{\diamond}_{(\underline{f},x)}\left(8t_{\mathrm{mix}}(H \wr G, \tfrac{1}{4})\right) \leq 4d^{\diamond}\left(4t_{\mathrm{mix}}(H \wr G, \tfrac{1}{4})\right) \leq 4\frac{2^4}{4^2} \leq \frac{1}{4}$$

Combining with (1.3.54) yields that for every starting state we have

$$\mathbf{P}_{(\underline{f},x)}\left[\tau_{\mathrm{cov}} > 4t_{\mathrm{mix}}(H \wr G, 1/4)\right] \leq 1/4.$$

Thus, run the chain in blocks of $8t_{\mathrm{mix}}(H \wr G, 1/4)$ and conclude that in each block it covers with probability at least $3/4$. Thus, the cover time is dominated by $8t_{\mathrm{mix}}(H \wr G, 1/4)$ times a geometric random variable with success probability $3/4$, so we have

$$\mathbf{E}_{(\underline{f},x)}\left[\tau_{\mathrm{cov}}\right] \leq 11t_{\mathrm{mix}}(H \wr G, 1/4).$$

Maximizing the right hand side over all possible starting states yields $t_{\mathrm{cov}}(G)$ $\leq 11t_{\mathrm{mix}}(H \wr G, 1/4)$, finishing the proof.

Proof of the lower bound, general case

Now we turn to the general case and show that $ct_{\mathrm{rel}}(H)|G|\log|G|$ is a lower bound. To see this we will use a distinguishing function method. Namely, suppose first that there is a single second eigenfunction ϕ_2 on H corresponding to the second eigenvalue λ_H. Then let us define $\psi : H \wr G \to \mathbb{C}$:

$$\psi((\underline{f},x)) := \sum_{v\in G} \phi_2(f(v)). \tag{1.3.55}$$

One can always normalize such that

$$\mathbf{E}_{\pi^{\diamond}}(\psi) = \sum_{v\in G} \mathbf{E}_{\pi}[\phi_2] = 0 \quad \mathbf{Var}_{\pi^{\diamond}}(\psi) = \sum_{v\in G} \mathbf{Var}_{\pi}(\phi_2) = |G| \cdot 1$$

This normalization has two useful consequences: First, by Chebyshev's inequality, the set $A = \{\psi < 2G^{1/2}\}$ has measure at least $3/4$ under stationarity. Second, $\phi_2(g_0) := \max_{g\in H} \phi_2(g) > 1$, otherwise the variance would be less than 1. We aim to show that the set A has measure less then $1/2$ at time

$ct_{\mathrm{rel}}|G|\log|G|$ and then we are done by using the following characterization of the total variation distance, see [49, 83]:

$$\|\nu-\mu\|_{\mathrm{TV}}=\sup_{A\subset\Omega}\{\nu(A)-\mu(A)\}.$$

Let us start all the lamp graphs from $g_0\in H$ where the maximum is attained for ϕ_2. Then we can condition on the local time sequence and use the eigenvalue property of ϕ_2

$$\begin{aligned}\mathbf{E}_{(\underline{g}_0,x)}\left[\psi((F_t,X_t))\right]&=\mathbf{E}\left[\mathbf{E}\left[\sum_{w\in G}\phi_2(F_t(w))|\left(L_v(t)\right)_v\right]\right]\\&=\phi_2(g_0)\mathbf{E}_x\left[\sum_{w\in G}\lambda_H^{L_w(t)}\right].\end{aligned}\tag{1.3.56}$$

Since $\sum_w L_v(t)=2t$, we can apply Jensen's inequality on the function $x\mapsto\lambda_H^x$ to get a lower bound on the expectation:

$$\mathbf{E}_x\left[\sum_{w\in G}\lambda_H^{L_w(t)}\right]\geq|G|\lambda_H^{\frac{2t}{G}}.$$

Now set $t=ct_{\mathrm{rel}}(H)|G|\log|G|$ to see that

$$\mathbf{E}_{(\underline{g}_0,x)}\left[\psi((F_t,X_t))\right]\geq|G|^{1-2c-o(1-\lambda_H)}\phi_2(g_0)$$

We can easily upper bound the conditional variance as follows:

$$\mathbf{Var}\left[\psi_t|(L_v(t))_{v\in G}\right]\leq\sum_{w\in G}\mathbf{E}_{g_0}\left[\phi_2^2(F_t(v))|L_v(t)\right]\leq|G|\phi_2^2(g_0).$$

Now, let us estimate the measure of set A at time t by using the lower bound on the expectation:

$$\mathbf{P}_{(\underline{g}_0,x)}\left[\psi_t\leq 2|G|^{1/2}\right]\leq\mathbf{P}_{(\underline{g}_0,x)}\left[|\psi_t-\mathbf{E}(\psi_t)|\geq\phi_2(g_0)|G|^{1-2c-O(1-\lambda_H)}-2|G|^{1/2}\right]$$

Now we use that $\phi_2(g_0)>1$ and set $c<1/4$. Then we see that on the right hand side, the term $\phi_2(g_0)|G|^{1-2c-O(\lambda_H)}$ dominates, so for $|G|$ large enough we can drop the negative term and compensate it with a multiplicative factor of $1/2$, say. Thus, condition on the local time sequence first and see that for any sequence $(L_v(t))_{v\in G}$ Chebysev's inequality yields:

$$\mathbf{P}_{(\underline{g}_0,x)}\left[\psi_t\in A\big|\left(L_v(t)\right)_{v\in G}\right]\leq\frac{\mathbf{Var}\left[\psi_t|(L_v(t))_{v\in G}\right]}{1/4\phi_2^2(g_0)|G|^{2-4c-o(1)}}$$

Combining this with the estimate on the conditional variance above yields that

$$\mathbf{P}_{(\underline{g}_0,x)}\left[\psi_t\in A|\left(L_v(t)\right)_v\right]\leq\frac{4}{|G|^{1-4c-o(1-\lambda_H)}}.$$

This is independent of the local time sequence, so law of total probability says we have the same upper bound without conditioning. Now setting $c := 1/8$ and $|G|$ large enough we see that the large hand side can be made smaller than $1/2$, finishing the proof.

Next we show that $c|G|t_{\text{mix}}(H)$ is a lower bound. First define

$$t_{\text{stop}}(G) := \min\{\mathbf{E}[\tau]; \tau\text{stopping time s.t.}\mathbf{P}[X_\tau = y] = \pi(y)\forall y \in G\}$$

and show that

$$\frac{1}{2}|G| \cdot t_{\text{stop}}(H) \leq t_{\text{stop}}(H \wr G).$$

Take a mean optimal stopping time τ^* reaching minimal expectation, i.e. $\mathbf{E}_{(\underline{f}_*, x_*)}[\tau^*] = t_{\text{stop}}(H \wr G)$ for some $(\underline{f}_*, x*) \in H \wr G$ and $\mathbf{E}_{(\underline{f}, x)}[\tau^*] \leq t_{\text{stop}}(H \wr G)$ for $(\underline{f}, x) \neq (\underline{f}_*, x_*)$.

We use this τ_* to define a stopping rule $\tau_H(v)$ on H_v, for every $v \in G$. Namely, do the following: look at one coordinate $v \in G$ and look at the chain restricted to the lamp graph H_w, i.e. only the moves which are done on the coordinate H_v. Then, stop the chain on H_v when τ^* stops on the whole lamplighter.

Start the chain from any $(\underline{f}_0, x_0)$. Since the number of moves done on different sites are just summing up to twice the total time, we have

$$\sum_{v \in G} \mathbf{E}_{f_0}[\tau_H(v)] = \mathbf{E}_{(\underline{f}_0, x_0)}\left[\sum_{v \in G} L_v(\tau^*)\right] = 2\mathbf{E}_{(\underline{f}_0, x_0)}[\tau^*]$$

Take the site w (which can depend on the starting vertex x_0), which minimizes the expectation $\mathbf{E}_{f_0}[\tau_H(w)]$. Clearly for this vertex the expected value is less than the average:

$$\mathbf{E}_{f_0}[\tau_H] \leq \frac{2}{|G|}\mathbf{E}_{(\underline{f}_0, x_0)}[\tau^*]$$

Now clearly the left hand side is at least as large as what a mean-optimal stopping rule on H can achive, and the right hand side is at most $t_{\text{stop}}(H \wr G)$. Thus we arrive at

$$\frac{1}{2}|G|t_{\text{stop}}(H) \leq t_{\text{stop}}(H \wr G).$$

Now, we use the equivalence from the paper [98, Corollary 2.5] stating that t_{stop} and t_{mix} are equivalent up to universal constants for lazy reversible chains and get that

$$c_1|G|t_{\text{mix}}(H) \leq t_{\text{mix}}(H \wr G).$$

An important example - the torii

Mind that cases such as $H = \mathbf{Z}_k^{d(k)}, d \geq 3$ can not be handled as above, since their have many second eigenfunctions. However, the result still carries through. What matters here is the product structure of the chain, i.e. the eigenfunctions corresponding to the coordinates are independent once we condition on how many steps are taken in the given coordinate.

In this case the second eigenvalue $\lambda_H = 1 - \frac{1}{d}\left(1 - \cos(\frac{2\pi}{k})\right)$ has $d = d(k)$ eigenfunctions, namely for $\underline{x} \in \mathbf{Z}_k^d$ we have $\phi_{2,i}(\underline{x}) = \sqrt{2}\cos(\frac{2\pi}{k}x_i), i = 1 \dots d$, where we normalized them such that the stationary variance is 1. Then define the sum of eigenfunctions by

$$\phi_*(\underline{x}) := d^{-1/2}\sum_{i=1}^{d}\phi_{2,i}, \quad \psi := \sum_{v\in G}\phi_*(\underline{f}(v)),$$

then clearly $\mathbf{E}_\pi^\diamond[\psi] = 0, \mathbf{Var}_\pi^\diamond[\psi] = |G|$, so $A = \{\psi < 2\sqrt{|G|}\}$ has measure at least $3/4$ under $\pi^\diamond$. Start each lamp-chain from a state $g_0 = \underline{0}$, i.e. the state which maximizes ϕ_*, clearly we have again $\phi_*(g_0) = \sqrt{2} > 1$. To estimate the expected value of ψ_t at time t we condition also on the number of steps taken on different coordinates $L_v(t)$, use linearity and the eigenfunction property to get

$$\mathbf{E}_t\left[\psi|\left(L_v(t)\right)_{v,i}\right] = \sum_v Q^t\left(\frac{1}{\sqrt{d}}\phi_{2,i}(g_0))\right) \geq |G|\lambda_H^{2t/|G|},$$

where we used Jensen's inequality to estimate the expression from below in the last step. If we set in $t = ct_{\mathrm{rel}}(H)|G|\log|G|$ then we get the lower bound $\phi_*(g_0)|G|^{1-2c-O(1-\lambda_H)}$. It makes sense to reduce the conditional variance at time t by using the product structure of the torus and conditioning also on how many steps were done in each coordinate:

$$\mathbf{Var}[\psi_t|\left(L_{v,i}(t)\right)] = \sum_v\sum_i d^{-1}\mathbf{Var}_t[\phi_2, i(G_t)] \leq |G|\max_{\underline{x},i}\phi_{2,i}(\underline{x})^2.$$

This is the point were we use that there is a state g_0 where each eigenfunctions are bounded and maximized at the same time with $\max_{x,i}\phi_{2,i}(\underline{x})^2 = \phi_{2,i}(g_0) = \phi_*(g_0)$ and conclude that $\mathbf{Var}[\psi_t|\left(L_{v,i}(t)\right)] \leq |G|\phi_*(g_0)^2$. Mind that the variance will not drop much below this even under more precise calculations. From this we can get the conditional Chebysev estimate similar as for the single eigenfunction case to get that for small enough c, $t = ct_{\mathrm{rel}}(\mathbf{Z}_k^d)|G|\log|G| = c|G|(dk^2\log|G|)$ is a lower bound. One can see via various methods that $t_{\mathrm{mix}}(\mathbf{Z}_k^d, |G|^{-1}) = O(d\log dk^2 + dk^2\log|G|)$. Thus, the method gives a matching lower bound if $|G| = o(d)$.

1.3.6 Further directions

As one can see, once we realized that the statement [83, Remark 6.12], that the separation distance equals the tail of the strong stationary time *if and only if* there is a halting state $y(x)$ for x, is false, we lost the matching upper and lower bounds in the highest generality. We have different other lower bounds which are not necessarily matching for all reversible chains. Thus, right now we are working on to relate the mixing time of $H \wr G$ to the Cesaro or geometric mixing time on H up to $\varepsilon = 1/G$. However, the equivalence of these mixing times and the total variation mixing time is only known [3, 98] for $\varepsilon = 1/4$ but not when $\varepsilon \to 0$. Thus it is still an open problem of its own interest, to find under what conditions the two mixing notions are equivalent.

The next step of understanding generalized lamplighters walks might be to investigate which properties on G and H are needed to exhibit cutoff (for a definition see [49, 83]), or to determine the mixing time in the uniform metric. For a comprehension what has already been done in this direction see the Introduction of this chapter.

Chapter 2

Generating hierarchical scale-free graphs from fractals

2.1 Introduction

Random graphs are in the main stream of research interest since the late 50s, starting with the seminal random graph model introduced by Erdős and Rényi [61]. A wide spectrum of literature investigates graph models with a fixed number of vertices (i.e some generalization of the Erdős-Rényi (ER) graphs), we refer the reader to the books of [71] or [35] as an introduction.

More recently in [40] Bollobás, Janson and Riordan introduced a general inhomogeneous random graph model, which also includes the ER random graphs as a special case. The vertices are assigned different types and the edge probabilities depend on these types given by a 'kernel' function. The authors characterized the emergence of the giant component, i.e. the phase transition, the typical distances and the diameter. In the supercritical regime they also proved what the typical graph distance is between two randomly chosen vertices of the giant component. Typical distances have also been studied in other models, see for example [47], [106]. A possible generalization of these models is to give the edges different edge weights.

This leads us to the problem of first passage percolation (FPP) in random environment: Let the environment be a random graph model and give each edge a random edge weight, typically independent and identically distributed (i.i.d.) positive random variables. Now think of fluid percolating through the graph from some source at a constant rate. First passage percolation refers to the time when vertices are reached by the fluid, i.e. the shortest path between vertices under the given edge-weights. As the environment grows one is interested in the asymptotics of various quantities of the flow.

In [32] Bhamidi, van der Hofstad and Hooghiemstra analyzed FPP on

the ER random graph with i.i.d. exponential edge weights. They proved that the hopcount, i.e. the number of edges on the shortest-weight path between two randomly chosen vertices in the giant component, follows a central limit theorem. Furthermore, they show convergence in distribution for the weight of the shortest-weight path. Related results for FPP with exponential edge weights can be found in [30], [31], [106], and newly the diameter with edge-weights were investigated in [4, 57].

Parallel to this directions, in the last two decades there have been a considerable amount of attention paid to the study of complex networks like the World Wide Web, social networks, or biological networks. This resulted in the construction of numerous new, more dynamical and growing network models, see e.g. [27], [35], [39], [59], [76]. Most of them use a version of preferential attachment and are of probabilistic nature.

In particular, the scale free property - the graph obeying a degree sequence with power law decay - raised interest and many models were introduced to capture this property, such as the *Preferential Attachment Models*. The history of similar models goes back to the 1920's [108, 104, 45]. The model was heuristically introduced by Barabási and Albert [24], and the first who investigated the model rigorously were Bollobás, Riordan, Spencer and Tusnády [37], and the mathematically rigorous construction was done by Bollobás and Riordan [36]. In the preferential attachment model discussed by Bollobás, Riordan, Spencer and Tusnády [37], starting from an initial graph, at each discrete time step a new vertex is added to the graph with some edges connected to it. These edges are attached sequentially to the existing vertices with a probability proportional to the degree of the receiving vertex at that time, thus favoring vertices with large degrees. The model obeys a *power-law* degree distribution similarly to many real life networks. Since then, many versions of preferential attachment models appeared in the literature. Let us mention some of them without the pursuit of completeness: [38] considers also directed edges, and non-linear preferential attachment model appears in [78]. Rudas, Tóth and Valkó [102] determined the asymptotic degree distribution for a wide range of weight functions in a continuous time non-linear model. Another direction of research on this field is to add some individual character to vertices, which we refer to as fitness. A new vertex at time t connects to vertex v_i with a conditional probability which is proportional to $\zeta_i D_i(t) + \eta_i$. A model where vertices only obey additive fitness is discussed in [62]. Variations of multiplicative fitness models were introduced by Bianconi and Barabási [34, 26], and studied further in [41]. The degree distribution both for the additive and multiplicative models was found by Bhamidi [33]. A new direction is to change the growth rule, such that it also takes account the structure of the existing graph. A model based on triangle-interactions appears in the work of Backhausz and Móri [6]. The wide-range literature is summarized in [35] or in [71].

A completely different approach was initiated by Barabási, Ravasz, and

Vicsek [25] based on the observation that real network often obey some hierarchical structure. They introduced deterministic network models generated by a method which is common in constructing fractals. Their model both exhibits hierarchical structure and the degree sequence obeys power law decay. To model also the clustering behavior of real networks, Ravasz and Barabási [100] developed the original model in [25] so that their deterministic network model preserved the same power law decay and has similar clustering behavior to many real networks. Namely, the average local clustering coefficient is independent of the size of the network and the local clustering coefficient decays inversely proportional to the degree of the node. A similar, fractal based deterministic model were introduced by Zhang, Comellas, Fertin and Rong [109], and called the high-dimensional Apollonian network. The graph is generated from the cylinder sets of the fractal of the Apollonian circle packing or the Sierpinsky carpet. Slightly different randomized version were introduced in [111, 112, 115, 114, 116, 110, 113].

In this section we generalize both of the models of [25] and [100]. Starting from an arbitrary initial bipartite graph G on N vertices, we construct a hierarchical sequence of deterministic graphs G_n. Namely, $V(G_n)$, the set of vertices of G_n is $\{0, 1, \dots, N-1\}^n$. To construct G_n from G_{n-1}, we take N identical copies of G_{n-1}, each of them identified with a vertex of G. Then we connect these components in a complicated way described in (2.2.1). In this way, G_n contains N^{n-1} copies of G_1, which are connected in a hierarchical manner, see Figures 2.1(a), 2.1(b) and 2.3 for two examples.

There are no triangles in G_n. Hence, in order to model the clustering properties of many real networks, we need to extend the set of edges of our graph sequence to destroy the bipartite property. Motivated by [100], we add some additional edges to G_1 to obtain the (no longer bipartite) graph $\widehat{G}_1$. Then we build up the graph sequence $\widehat{G}_n$ as follows: $\widehat{G}_n$ consist of N^{n-1} copies of $\widehat{G}_1$, which copies are connected to each other in the same way as they were in G_n. So, $\widehat{G}_n$ and G_n have the same vertex set and their edges only differ at the lowest hierarchical level, that is, within the N^{n-1} copies of G_1 and $\widehat{G}_1$, see Figures 2.3 and 2.5. We give a rigorous proof of the fact that the average local clustering coefficient of $\widehat{G}_n$ does not depend on the size and the local clustering coefficient of a node with degree k is of order $1/k$.

The embedding of the adjacency matrix of the graph sequence G_n is carried out as follows: A vertex $\underline{x} = (x_1 \dots x_n)$ is identified with the corresponding N-adic interval $I_{\underline{x}}$ (see (2.2.4)). Λ_n is the union of those $N^{-n} \times N^{-n}$ squares $I_{\underline{x}} \times I_{\underline{y}}$ for which the vertices $\underline{x}, \underline{y}$ are connected by an edge in G_n. So, Λ_n is the most straightforward embedding of the adjacency matrix of G_n into the unit square. Λ_n turns out to be a nested sequence of compact sets, which can be considered as the n-th approximation of a graph directed self-similar fractal Λ on the plane, see Figure 2.1(c). We discuss connection between the graph theoretical properties of G_n and properties of the limiting

fractal Λ. In particular, we express the power law exponent of the degree distribution with the ratio of the Hausdorff dimensions of some slices of Λ (Theorem 2.3.6).

Furthermore, using Λ we generate a random graph sequence G_n^{r} in a way which was inspired by the W-random graphs introduced by Lovász and Szegedy [88]. See also Diaconis, Janson [53], which paper contains a list of corresponding references. We show that the degree sequence has power law decay with the same exponent as the deterministic graph sequence G_n. Thus we can define a random graph sequence with a prescribed power law decay in a given range. Bollobás, Janson and Riordan [39] considered inhomogeneous random graphs generated by a kernel. Our model is not covered by their construction, since Λ is a fractal set of zero two dimensional Lebesgue measure.

The section is organized as follows: In Section 2.2 we define the deterministic model and the associated fractal set Λ. In Section 2.3, we verify the scale free property of G_n (Theorem 2.3.1). We compare the Hausdorff dimension of Λ to the power law exponent of the degree sequence of G_n. Our next result is that both of the diameter of G_n and the average length of shortest path between two vertices are of order of the logarithm of the size of G_n (Corollary 2.3.9 and Theorem 2.3.10). In Section 2.3.4 we prove the above mentioned properties of the clustering coefficient of $\widehat{G}_n$ (Theorem 2.3.16 and 2.3.14). In Section 2.4 we describe the randomized model, and in Section 2.4.1 we prove that the model exhibits the same power law decay as the corresponding deterministic version.

2.2 Deterministic model

The model was motivated by the hierarchical graph sequence model in [25], and is given as follows.

2.2.1 Description of the model

Let G, our *base graph*, be any labeled bipartite graph on the vertex set $\Sigma_1 = \{0, \dots, N-1\}$. We partition Σ_1 into the non-empty sets V_1, V_2 and one of the end points of any edge is in V_1, and the other is in V_2. We write $n_i := |V_i|$, $i = 1, 2$ for the cardinality of V_i. The edge set of G is denoted by $E(G)$. If the pair $x, y \in \Sigma_1$ is connected by an edge, then this edge is denoted by $\binom{x}{y}$, since this notation makes it convenient to follow the labels of the vertices along a path.

Now we define our graph sequence $\{G_n\}_{n\in\mathbb{N}}$ generated by the base graph G.

The vertex set is $\Sigma_n = \{(x_1 x_2 \dots x_n) : x_i \in \Sigma_1\}$, all words of length n above the alphabet Σ_1. To be able to define the edge set, we need some

further definitions.

Definition 2.2.1.

1. *We assign a type to each element of* Σ_1*. Namely,*

$$typ(x) = \begin{cases} 1, & \text{if } x \in V_1; \\ 2, & \text{if } x \in V_2. \end{cases}$$

2. *We define the* ***type*** *of a word* $\underline{z} = (z_1 z_2 \dots z_n) \in \Sigma_n$ *as follows: if all the elements* $z_j, j = 1, \dots, n$ *of* $\underline{z}$ *fall in the same* V_i*,* $i = 1, 2$ *then* $typ(\underline{z})$ *the type of* $\underline{z}$ *is* i*. Otherwise* $typ(\underline{z}) := 0$*.*

3. *For* $\underline{x} = (x_1 \dots x_n), \underline{y} = (y_1 \dots y_n) \in \Sigma_n$ *we denote the* ***common prefix*** *by*

$$\underline{x} \wedge \underline{y} = (z_1 \dots z_k) \text{ s.t. } x_i = y_i = z_i, \forall i = 0, \dots, k \text{ and } x_{k+1} \neq y_{k+1}.$$

4. *Given* $\underline{x} = (x_1 \dots x_n), \underline{y} = (y_1 \dots y_n) \in \Sigma_n$*, the* ***postfixes*** $\tilde{\underline{x}}, \tilde{\underline{y}} \in \Sigma_{n-|\underline{x} \wedge \underline{y}|}$ *are determined by*

$$\underline{x} = (\underline{x} \wedge \underline{y})\tilde{\underline{x}}, \ \underline{y} = (\underline{x} \wedge \underline{y})\tilde{\underline{y}},$$

where the concatenation of the words $\underline{a}, \underline{b}$ *is denoted by* $\underline{ab}$*.*

Now we can define the edge set $E(G_n)$. Two vertices $\underline{x}$ and $\underline{y}$ in G_n are connected by an edge if and only if the following assumptions hold:

(a) One of the postfixes $\tilde{\underline{x}}, \tilde{\underline{y}}$ is of type 1, the other is of type 2,

(b) for each $i > |x \wedge y|$, the coordinate pair $\binom{x_i}{y_i}$ forms an edge in G.

That is, $E(G_n) \subset \Sigma_n \times \Sigma_n$:

$$E(G_n) = \left\{ \binom{\underline{x}}{\underline{y}} \Big| \underline{x} = \underline{y} \text{ or } \{\text{typ}(\tilde{\underline{x}}), \text{typ}(\tilde{\underline{y}})\} = \{1, 2\}, \forall |\underline{x} \wedge \underline{y}| < i \leq n, \binom{x_i}{y_i} \in E(G) \right\} \quad (2.2.1)$$

Remark 2.2.2. *Note that we artificially added all loops to the (otherwise bipartite) graph sequence* G_n*, implying easier calculations later without loss of the important properties. In particular,* G_1 *differs from* G *only in the loops.*

Remark 2.2.3 (Hierarchical structure of G_n)**.** *For every initial digit* $x \in \{0, 1, \dots, N-1\}$*, consider the set* W_x *of vertices* $(x_1 \dots x_n)$ *of* G_n *with* $x_1 = x$*. Then the induced subgraph on* W_x *is identical to* G_{n-1}*.*

We write $\deg_n(\underline{x})$ for the *degree of a vertex* in G_n, including the loop which increases the degree by 2. However, for an $x \in \Sigma_1$, $\deg x$ denotes degree of x in G. In particular $\deg_1(x) = \deg(x) + 2$. In what follows, we will frequently use $\ell(\underline{x})$, the length of the longest block from backwards in $\underline{x}$ which has a nonzero type,

$$\ell(\underline{x}) := \max_{i \in \mathbb{N}} \{\mathrm{typ}(x_{n-i+1}, \dots x_n) \in \{1, 2\}\} \tag{2.2.2}$$

Remark 2.2.4. *The degree of a node* $\underline{x} \in \Sigma_n$

$$\deg_n(\underline{x}) = 2 + S(\underline{x}) \cdot \deg(x_n),$$

where

$$\begin{aligned} S(\underline{x}) : &= 1 + \deg(x_{n-1}) + \cdots + \deg(x_{n-1}) \deg(x_{n-2}) \cdots \deg(x_{n-\ell(\underline{x})+1}) \\ &= \sum_{r=0}^{\ell(\underline{x})-1} \left(\prod_{j=1}^{r} \deg(x_{n-j}) \right), \end{aligned} \tag{2.2.3}$$

where the empty sum is meant to be 1.

The following two examples satisfy the requirements of our general model.

Example 2.2.5 (Cherry)**.** *Barabási, Ravasz and Vicsek [25] introduced the "cherry" model presented on Figures 2.1(a) and 2.1(b): Let* $V_1 = \{1\}$ *and* $V_2 = \{0, 2\}$, $E(G) = \{(1, 0), (1, 2)\}$.

Example 2.2.6 (Fan)**.** *Our second example is called "fan", and is defined on Figure 2.3. Note that here* $|V_1| > 1$.

2.2.2 The embedding of the adjacency matrices into $[0, 1]^2$

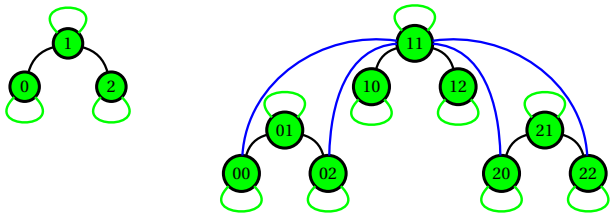
In this Section, we investigate the sequence of adjacency matrices corresponding to $\{G_n\}_{n \in \mathbb{N}}$. Roughly speaking, we will map them in the unit square, see Figure 2.1(c).

To represent the adjacency matrix of G_n as a subset of the unit square, first partition $[0, 1]^2$ into N^{2n} congruent boxes, i.e. divide $[0, 1]$ into equal subintervals of length $\frac{1}{N^n}$, corresponding to the first n digits of the N-adic expansion of elements of $[0, 1]$:

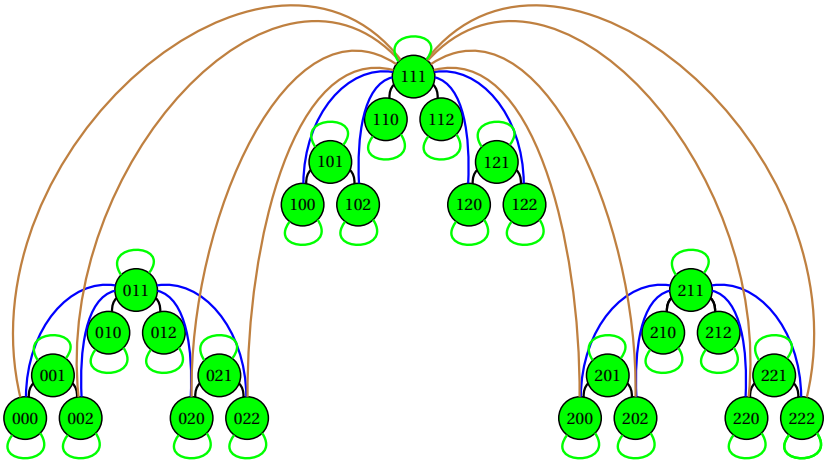
$$I_{x_1 \dots x_n} = \left[\sum_{r=1}^{n} \frac{x_r}{N^r}, \sum_{r=1}^{n} \frac{x_r}{N^r} + \frac{1}{N^n} \right], \forall (x_1 \dots x_n) \in \Sigma_n. \tag{2.2.4}$$

We partition $[0, 1]^2$ with the corresponding level-n squares:

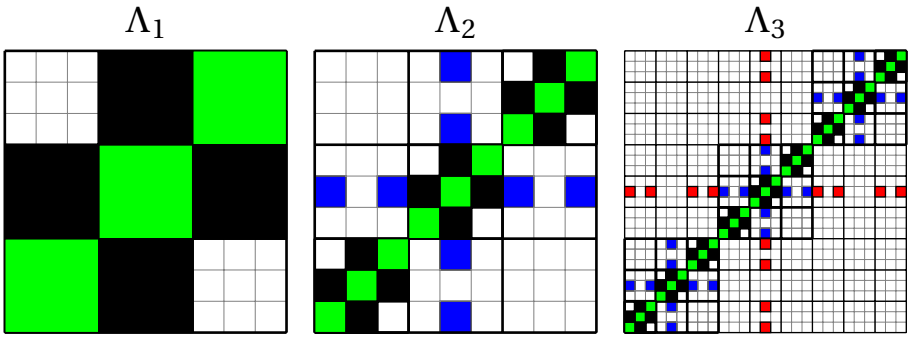
$$Q_{\binom{\underline{x}}{\underline{y}}} := I_{\underline{x}} \times I_{\underline{y}}, \quad \binom{\underline{x}}{\underline{y}} \in \Sigma_n \times \Sigma_n. \tag{2.2.5}$$



(a) G_1 and G_2 with loops



(b) G_3



(c) The sets $\Lambda_1, \Lambda_2, \Lambda_3$

Figure 2.1: $G_1, G_2, G_3, \Lambda_1, \Lambda_2, \Lambda_3$ for the cherry Example 2.2.5

A natural embedding of the adjacency matrix of G_n in the unit square is as follows:

$$\Lambda_n(a,b) := \begin{cases} 1, & \text{if } (a,b) \in Q_{\binom{x}{y}}, \binom{x}{y} \in E(G_n); \\ 0, & \text{otherwise.} \end{cases} \tag{2.2.6}$$

That is,

$$\Lambda_n(a,b) = \sum_{\substack{x,y \in \Sigma_n \\ \binom{x}{y} \in E(G_n)}} \mathbf{1}_{Q_{\binom{x}{y}}}(a,b).$$

We write Λ_n for the support of the function $\Lambda_n(a,b)$, see Figure 2.1(c). Observe that Λ_n is a compact set and $\Lambda_{n+1} \subset \Lambda_n$ holds for all n. So we can define the non-empty compact set

$$\Lambda := \bigcap_{n=1}^{\infty} \Lambda_n. \tag{2.2.7}$$

Clearly,

$$\mathbf{1}_{\Lambda}(a,b) = \lim_{n\to\infty} \Lambda_n(a,b).$$

Remark 2.2.7. *This representation obviously depends on the labeling of the graph G. For an arbitrary permutation π of $\{0,\dots,N-1\}$, the corresponding representation of G_n is denoted by $\Lambda_n^{\pi}(a,b)$. The relation between these two representations is given by the formula*

$$\begin{aligned} \Lambda_n^{\pi}(a,b) &= \Lambda_n(\varphi_{\pi^{-1}}(a), \varphi_{\pi^{-1}}(b)), \text{ and} \\ \mathbf{1}_{\Lambda^{\pi}}(a,b) &= \mathbf{1}_{\Lambda}(\varphi_{\pi^{-1}}(a), \varphi_{\pi^{-1}}(b)), \end{aligned}$$

where the measurable function $\varphi_{\pi}(x) : [0,1] \to [0,1]$ is defined by

$$\varphi_{\pi}\left(\sum_{i=1}^{\infty} \frac{x_i}{N^i}\right) = \sum_{i=1}^{\infty} \frac{\pi(x_i)}{N^i}.$$

2.2.3 Graph-directed structure of Λ

Now we prove that the limit Λ (defined in (2.2.7)) can be considered as the attractor of a not irreducible graph-directed self-similar iterated function system, (for the definition see [64]), with the directed graph $\mathcal{G}$ defined below.

Definition 2.2.8. *The **vertex set** $V(\mathcal{G})$ is partitioned into three subsets:*

$$\begin{aligned} V_{dd} &= \left\{ \binom{z}{z}, z \in \Sigma_1 \right\} \\ V_{12} &= \left\{ \binom{x}{y}, x \in V_1, y \in V_2 \right\} \\ V_{21} &= \left\{ \binom{x}{y}, x \in V_2, y \in V_1 \right\}. \end{aligned} \tag{2.2.8}$$

Then

$$V(\mathcal{G}) = V_{dd} \cup V_{12} \cup V_{21}.$$

The ***set of directed edges*** $E(\mathcal{G})$ *of* $\mathcal{G}$ *is as follows: First we connect all vertices in both directions within each of the three sets* V_{dd}, V_{12} *and* V_{21} *(loops included). Then there is an outgoing edge for each vertex in* V_{dd} *to all vertices in* V_{12} *and* V_{21}.

For every directed edge $e = (v_1, v_2) \in E(\mathcal{G})$ *we define a homothety:*

$$f_e : \; Q_{v_2} \to Q_{v_1}, \;\; f_e(a,b) := \frac{1}{N}(a,b) + \frac{1}{N}(x_1, y_1), \;\; \textit{with } v_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \quad (2.2.9)$$

where $Q_v := Q_{\binom{x}{y}}$ *is the level-1 square for* $v = \binom{x}{y} \in V(\mathcal{G})$.

The graph $\mathcal{G}$ corresponding to the graph sequence in the "cherry" example is given by Figure 2.2.

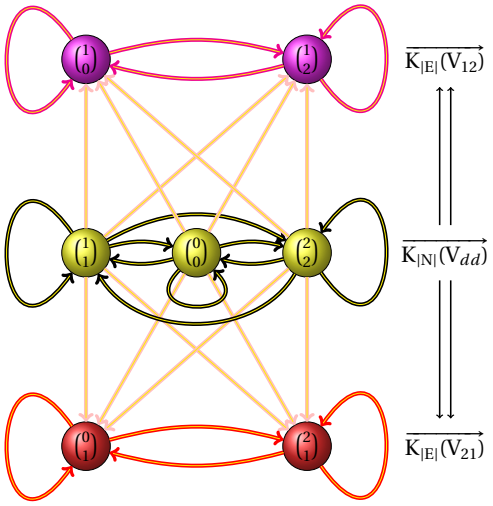


Figure 2.2: The graph $\mathcal{G}$ for the "cherry", Example 2.2.5.

In general, $\mathcal{G}$ is given by the schematic picture on the right hand side of Figure 2.2, where the double arrow in between the complete directed graphs $\overrightarrow{K_{\cdot}(V_{\cdot\cdot})}$ illustrates that we connect all pairs of vertices in the given direction.

Let $\mathcal{P}_n$ be the set of all paths of length n in $\mathcal{G}$, i.e.

$$\mathcal{P}_n := \{\underline{v} = (v_1 \dots v_n) | \forall \; 1 \leq i < n \; (v_i, v_{i+1}) \in E(\mathcal{G})\}.$$

For a $\underline{v} = (v_1 \dots v_n) = \binom{x_1 \dots x_n}{y_1 \dots y_n} \in \mathcal{P}_n$ it immediately follows from definitions (2.2.5) and (2.2.9) that

$$Q_{\underline{v}} = f_{\underline{v}}\left([0,1]^2\right) = I_{x_1 \dots x_n} \times I_{y_1 \dots y_n}, \tag{2.2.10}$$

where

$$\begin{aligned} f_{\underline{v}}(.) :&= f_{(v_1,v_2)} \circ \dots \circ f_{(v_{n-1},v_n)}(.) \text{ if } n \geq 2, \\ f_v(a,b) :&= \frac{1}{N}(a,b) + \frac{1}{N}(x,y), \text{ if } n = 1,\ v = \binom{x}{y}. \end{aligned} \tag{2.2.11}$$

The key observation of connecting $\mathcal{G}$ to the graph sequence G_n is the following:

Claim 2.2.9. *For all n we have*

$$E(G_n) = \mathcal{P}_n.$$

Proof. Let $\underline{v} = (v_1 \dots v_n) = \binom{a_1 \dots a_n}{b_1 \dots b_n} \in \Sigma_n \times \Sigma_n$, thus $\underline{a} = (a_1 \dots a_n)$ and $\underline{b} = (b_1 \dots b_n)$ are vertices in G_n. First we assume that $\underline{v} \in E(G_n)$. Observe that by (2.2.1), $\binom{a_i}{b_i}$ are vertices in $\mathcal{G}$. We would like to prove that the sequence

$$\binom{a_1}{b_1} \dots \binom{a_n}{b_n} \in \mathcal{P}_n. \tag{2.2.12}$$

If $k := |\underline{a} \wedge \underline{b}| \geq 1$, then for $i \leq k$, $a_i = b_i$ holds, thus the sequence of points $\binom{a_1}{b_1} \dots \binom{a_k}{b_k}$ forms a path in $\overrightarrow{K_{|N|}(V_{dd})}$. By (2.2.1), the pairs $\binom{a_{k+1}}{b_{k+1}}, \dots, \binom{a_n}{b_n}$ are all edges in G thus vertices in $\mathcal{G}$. Furthermore, either they all belong to V_{12} or they are all contained in V_{21}, see (2.2.8). This implies that this postfix also forms a path in $\overrightarrow{K_{|N|}(V_{12})}$ or in $\overrightarrow{K_{|N|}(V_{21})}$. By definition of $E(\mathcal{G})$, $\left(\binom{a_k}{b_k}, \binom{a_{k+1}}{b_{k+1}}\right)$ is an edge in $\mathcal{G}$, so $\binom{a_1}{b_1} \dots \binom{a_n}{b_n}$ is a path in $\mathcal{G}$. If $k = 0$ then the whole path is contained either in V_{12} or in V_{21}. This completes the proof of (2.2.12).

On the other hand, if $\binom{a_1}{b_1} \dots \binom{a_n}{b_n}$ is a path of length n in $\mathcal{G}$, then we claim that for $\underline{a} = (a_1 \dots a_n), \underline{b} = (b_1 \dots b_n) \in V(G_n)$

$$(\underline{a}, \underline{b}) \in E(G_n).$$

The proof is very similar to the previous one. □

In this way we can characterize Λ_n as follows:

Corollary 2.2.10.

$$\Lambda_n = \bigcup_{\underline{v} \in \mathcal{P}_n} Q_{\underline{v}} = \bigcup_{\underline{v} \in \mathcal{P}_n} f_{\underline{v}}\left([0,1]^2\right).$$

Proof. Immediately follows from (2.2.6) and (2.2.10) and the assertion of the Claim 2.2.9. □

Let us define

$$\mathcal{P}_\infty := \{\underline{v} = (v_1 v_2 \dots) | \forall i \in \mathbb{N},\ (v_i, v_{i+1}) \in E(\mathcal{G})\}.$$

Now for every $\underline{v} \in \mathcal{P}_\infty$ we have $\bigcap_{n=1}^{\infty} Q_{(v_1 \dots v_n)}$ is a point in $[0,1]^2$, which will be denoted by $\Pi_{\underline{v}}$. That is,

$$\Pi : \mathcal{P}_\infty \to [0,1]^2,\ \Pi(\underline{v}) := \bigcap_{n=1}^{\infty} Q_{(v_1 \dots v_n)} = \lim_{n\to\infty} f_{v_1 \dots v_n}(0,0).$$

It is an immediate consequence of Corollary 2.2.10, that

$$\Pi(\mathcal{P}_\infty) = \Lambda,\ \text{i.e. } \Lambda = \bigcup_{\underline{v}\in\mathcal{P}_\infty} \Pi_{\underline{v}}. \tag{2.2.13}$$

This means that Λ_n, the embedded adjacency matrix of G_n, can be considered as the n-th approximation of the fractal set Λ.

In this way we coded the elements of Λ by the elements of $\mathcal{P}_\infty$. This coding is not $1-1$ for the same reason as the N-adic expansion is not $1-1$. However, if neither of the two coordinates of a point $(a,b) \in \Lambda$ are N-adic rational numbers, then (a,b) has a unique code.

2.2.4 Fractal geometric characterization of Λ.

For notational convenience we define the set of finite words above the alphabet V_{dd} (including the empty word as well):

$$V_{dd}^* := \{\underline{v} |\ \exists n \in \mathbb{N} \cup \{0\}, \underline{v} = (v_1 \dots v_n) \text{ and } v_i \in V_{dd}\}.$$

The three subgraphs $\overrightarrow{K_{|E|}(V_{12})}$, $\overrightarrow{K_{|E|}(V_{21})}$ and $\overrightarrow{K_{|E|}(V_{dd})}$ of $\mathcal{G}$ are complete directed graphs. We consider the three corresponding self-similar iterated function systems (IFS):

$$\begin{aligned}
\mathcal{F}_{dd} &:= \{f_v\}_{v\in V_{dd}},\\
\mathcal{F}_{12} &:= \{f_v\}_{v\in V_{12}},\\
\mathcal{F}_{21} &:= \{f_v\}_{v\in V_{21}},
\end{aligned}$$

where the functions $f_v, v \in V(\mathcal{G})$ were defined in (2.2.11). The attractors of these IFS-s (see [64, p.30]) are the unique nonempty compact sets satisfying

$$\begin{aligned}
\Lambda_{dd} &:= \bigcup_{v\in V_{dd}} f_v(\Lambda_{dd}) = \{\Pi(\underline{v}) | \underline{v} = (v_1, v_2 \dots) \text{ and } v_i \in V_{dd}\}\\
\Lambda_{12} &:= \bigcup_{v\in V_{12}} f_v(\Lambda_{12}) = \{\Pi(\underline{v}) | \underline{v} = (v_1, v_2 \dots) \text{ and } v_i \in V_{12}\}\\
\Lambda_{21} &:= \bigcup_{v\in V_{21}} f_v(\Lambda_{21}) = \{\Pi(\underline{v}) | \underline{v} = (v_1, v_2 \dots) \text{ and } v_i \in V_{21}\}.
\end{aligned} \tag{2.2.14}$$

The Open Set Condition (see e.g. [64, p.35]) holds for these IFS-s, so we can easily compute the Hausdorff-dimension of the attractors. Clearly, Λ_{dd} is the diagonal of the unit square.

Now we prove that Λ is a countable union of homothetic copies of these attractors.

Theorem 2.2.11.

$$\Lambda = \underbrace{\mathrm{Diag}}_{\Lambda_{dd}} \cup \bigcup_{\underline{v} \in V_{dd}^*} \left(f_{\underline{v}}(\Lambda_{12}) \cup f_{\underline{v}}(\Lambda_{21}) \right),$$

where $\mathrm{Diag} = \{(x,x) : x \in [0,1]\}$.

Remark 2.2.12. *Observe that Λ_{21} is the image of Λ_{12} by the reflection through the diagonal, hence Λ is symmetric to the diagonal. The same is true for the n-th approximation Λ_n of Λ. This can be seen immediately by using the embedded adjacency matrix characterization of Λ_n.*

Proof of Theorem 2.2.11. We start by showing that

$$\Lambda \subset \mathrm{Diag} \cup \bigcup_{\underline{v} \in V_{dd}^*} \left(f_{\underline{v}}(\Lambda_{12}) \cup f_{\underline{v}}(\Lambda_{21}) \right). \tag{2.2.15}$$

Pick an arbitrary point $(a,b) \in \Lambda$. As a consequence of (2.2.13) there exists a $\underline{v} = (v_1 v_2 \dots) \in \mathcal{P}_\infty$ such that $\Pi(\underline{v}) = (a,b)$. Let $k := \max\{\ell : v_\ell \in \Lambda_{dd}\}$. We distinguish three cases: $k = 0$, $k = \infty$ or $0 < k < \infty$. Mind that for all $i \le k, v_i \in V_{dd}$ since once the path left the component V_{dd}, there is no way to return. Since V_{12} and V_{21} are closed, for $k < \infty$ all $v_i, i > k$ are in the same component V_{12} or V_{21}.

Case $k = 0$ Clearly either all v_i are in V_{12} or in V_{21}, so $\Pi(\underline{v}) \in \Lambda_{12} \cup \Lambda_{21}$.

Case $k = \infty$ For the same reason, $\Pi(\underline{v}) = \lim_{n\to\infty} f_{v_1 \dots v_n}(0,0) \in \Lambda_{dd} = \mathrm{Diag}$. This is so because $f_{v_1 \dots v_n}(0,0)$ is in the $\frac{1}{N^n}$ neighborhood of the diagonal $\{(x,x) : x \in [0,1]\}$.

Case $0 < k < \infty$ Let $\underline{v}_k = (v_1 \dots v_k)$. For symmetry, without loss of generality we may assume that $v_{k+1} \in V_{12}$. As in the first case, we can see that for $\underline{w} := (v_{k+1} v_{k+2} \dots)$, $\Pi(\underline{w}) \in \Lambda_{12}$. Hence $\Pi(\underline{v}) = f_{\underline{v}_k}(\Pi_{\underline{w}}) \in f_{\underline{v}_k}(\Lambda_{12})$.

Now we have verified (2.2.15). To prove the opposite direction, that is

$$\Lambda \supset \mathrm{Diag} \cup \bigcup_{\underline{v} \in V_{dd}^*} \left(f_{\underline{v}}(\Lambda_{12}) \cup f_{\underline{v}}(\Lambda_{21}) \right), \tag{2.2.16}$$

we will use the symbolic representation of Λ given in (2.2.13).

Pick an $x \in [0,1]$ and take the N-adic code $(x_1x_2\dots)$ of x. That is, $x = \sum_{n=1}^{\infty} \frac{x_i}{N^i}$, $x_i \in \{0,\dots,N-1\}$. Then

$$\underline{v} := \left(\underbrace{\binom{x_1}{x_1}}_{v_1}, \underbrace{\binom{x_2}{x_2}}_{v_2}, \dots \right) \in \mathcal{P}_\infty,$$

it is easy to see that $\Pi(\underline{v}) = (x,x)$. So by (2.2.13), $(x,x) \in \Lambda$.

Now we assume that $(a,b) \in \bigcup_{\underline{v}\in V_{dd}^*} \big(f_{\underline{v}}(\Lambda_{12}) \cup f_{\underline{v}}(\Lambda_{21})\big)$. Without loss of generality we may further assume that $(a,b) \in f_{\underline{v}}(\Lambda_{12})$ for some $\underline{v} \in V_{dd}^*$. That is, $(a,b) = f_{\underline{v}}(a',b')$ where $(a',b') \in \Lambda_{12}$. By (2.2.14) there exists a $\underline{w} := (w_1w_2\dots)$, $w_i \in V_{12}$ such that $\Pi(\underline{w}) = (a',b')$. In this way, for the concatenation $\underline{t} := \underline{v}\underline{w} \in \mathcal{P}_\infty$ we have $(a,b) = \Pi(\underline{t})$ which implies $(a,b) \in \Lambda$. This completes the proof of (2.2.16). □

2.2.5 The same model without loops.

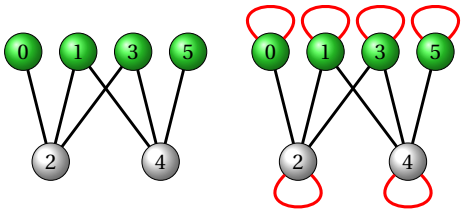
Let G_n' be the same graph as G_n but without loops, i.e. $V(G_n') = V(G_n)$ and $E(G_n') \subset \Sigma_n \times \Sigma_n$:

$$E(G_n') = \left\{ \binom{\underline{x}}{\underline{y}} \;\middle|\; \{\mathrm{typ}(\tilde{\underline{x}}), \mathrm{typ}(\tilde{\underline{y}})\} = \{1,2\} \text{ and } \forall |\underline{x} \wedge \underline{y}| < i \le n, \binom{x_i}{y_i} \in E(G) \right\}$$

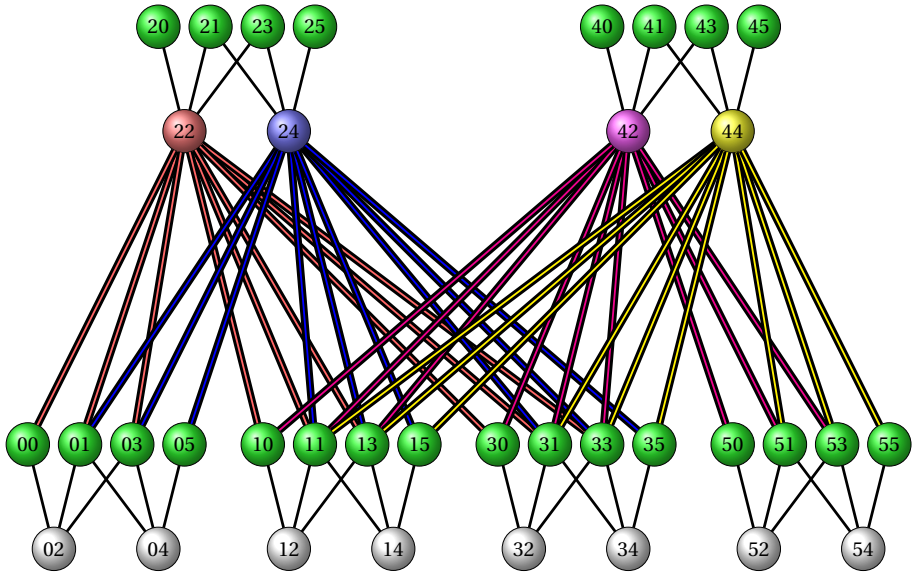
In this case $\Lambda_n' = \Lambda_n \backslash \mathrm{Diag}_n$, where Diag_n is the union of the level n squares that have nonempty intersection with the diagonal. The sequence Λ_n' is not a nested sequence of compact sets. However, it is easy to see that the characteristic function of Λ_n' tends to characteristic function of $\Lambda \setminus \mathrm{Diag}$. Further, Λ_n' tends to Λ in the Hausdorff metric, see [64].

2.3 Properties of the sequence $\{G_n\}$ and Λ

In this section we compute the degree distribution of G_n, and relate it to the Hausdorff dimension of Λ. We also compute the length of the average shortest path in G_n. To get interesting result about the local clustering coefficient we need to modify our graph sequence G_n in the line as it was done in [25].



(a) G on the left and G_1 on the right hand side.
Here $V_1 = \{2, 4\}$ and $V_2 = \{0, 1, 3, 5\}$



(b) The graph G_2 (contains additionally all loops).

Figure 2.3: Example "fan".

2.3.1 Degree distribution of $\{G_n\}$

Here we compute the degree distribution under the following regularity assumption on the base graph G:

$$\begin{aligned} &\deg(x) := d_1, \quad \forall x \in V_1 \\ &\max_{j \in V_2} \deg(y) := d_2 \leq d_1 - 1, \quad \forall y \in V_2 \end{aligned} \tag{A1}$$

Recall that we defined $\ell(\underline{x})$ in (2.2.2) as the length of the longest block from backwards of the node $\underline{x}$ such that the last $\ell(\underline{x})$ digits of $\underline{x}$ belong to the same V_i. Put $\Sigma_n^i := \{\underline{x} \in \Sigma_n | x_n \in V_i\}, i = 1, 2$. It follows from **A1** and Remark 2.2.4 that the degree of a node $\underline{x} \in \Sigma_n^1$ is $\frac{d_1^{\ell(\underline{x})+1}-1}{d_1-1} + 1$, and the number of such nodes with $\ell(\underline{x}) = \ell$ is exactly $N^{n-\ell+1} \cdot n_2 \cdot n_1^\ell$.

Under assumption **A1**, the decay of the degree distribution is determined

by the set of high degree nodes denoted by

$$HD_n := \left\{ \underline{x} \in \Sigma_n^1 | \deg_n(\underline{x}) > \max_{\underline{y} \in \Sigma_n^2} \deg_n(\underline{y}) \right\}.$$

An equivalent characterisation of HD_n is

$$HD_n = \left\{ \underline{x} \in \Sigma_n^1 | \ell(\underline{x}) > \frac{1}{\log d_1} \max\left\{(n+1)\log(d_2), \log n\right\} \right\}.$$

This is so because the degree of any $\underline{y} \in \Sigma_n^2$ is at most $\max\{d_2^{n+1}, n\}$. The tail of the cumulative degree distribution is

$$\mathbf{P}\left[\deg_n(\underline{X}) > \frac{d_1^{\ell+1}-1}{d_1-1} + 1\right] = \frac{n_1^{\ell+1} N^{n-\ell-1}}{N^n} = \left(\frac{n_1}{N}\right)^{\ell+1},$$

where $\underline{X}$ is a uniformly chosen node of G_n. Mind that as long as $\ell < n$, this probability does not depend on n. Writing $\widetilde{F}(t) = \mathbf{P}(\deg(\underline{X}) > t)$ for the tail of the cumulative distribution function we get the power law decay

$$\widetilde{F}(t) = t^{-\frac{\log(N/n_1)}{\log d_1}} \cdot c(d_1) \quad \text{for } t = \frac{d_1^{\ell+1}-1}{d_1-1}.$$

So we have proved

Theorem 2.3.1. *The degree distribution of the graph sequence G_n satisfying assumption **A1**, has a power law decay with exponent*

$$\tilde{\gamma} = \gamma - 1 = \frac{\log(N/n_1)}{\log d_1}. \tag{2.3.1}$$

Using this, a simple calculation shows that the largest decay γ we can possibly get from our model is $1 + \frac{\log 3}{\log 2}$, and this maximum is attained when $n_1 = 1$ and $d_1 = 2 = n_2$. This is exactly the graph sequence in Example 2.2.5, see Figures 2.1(a) and 2.1(b). We will later see that the case $n_1 = 1$ is important in another sense as well, see Section 2.3.2.

2.3.2 Hausdorff dimension of Λ

In Theorem 2.2.11 we decomposed Λ into the diagonal of the square and countably many homothetic copies of the self-similar sets Λ_{12} and Λ_{21}. By definition, we obtain Λ_{21} from Λ_{12} by interchanging the coordinates. Hence

$$\dim_{\mathrm{H}}\left(\Lambda \backslash \mathrm{Diag}\right) = \dim_{\mathrm{H}}(\Lambda_{12}). \tag{2.3.2}$$

We have seen that Λ_{12} is the attractor of the self-similar IFS $\mathcal{F}_{12}$ (defined in Section 2.2.4) which consists of $|E|$ similarities of contraction ratio $\frac{1}{N}$, and $\mathcal{F}_{12}$ satisfies the Open Set Condition. Hence

$$\dim_{\mathrm{H}}\left(\Lambda_{12}\right) = \frac{\log|E|}{\log N}. \tag{2.3.3}$$

Combining (2.3.2) and (2.3.3) yields the assertion of the following theorem:

Theorem 2.3.2. *The Hausdorff dimension of* Λ *is*

$$\dim_H \Lambda = \max\left\{\frac{\log|E|}{\log N}, 1\right\},$$

furthermore,

$$\dim_H\big(\Lambda\backslash\mathrm{Diag}\big) = \dim_H(\Lambda_{12}) = \frac{\log|E|}{\log N}. \tag{2.3.4}$$

Relation between the Hausdorff dimension and the power-law exponent.

Now we discuss the relation between the Hausdorff dimension of $\Lambda \setminus \mathrm{Diag}$ and the decay exponent $\tilde{\gamma}$ of the degree distribution in G_n. To shorten the notation, in this section we write

$$\mathrm{HD} := \dim_{\mathrm{H}}(\Lambda \setminus \mathrm{Diag}).$$

First we consider the simplest case when $n_1 = 1$ which is a generalization of the Cherry Example 2.2.5.

Corollary 2.3.3. *If* $|V_1| = n_1 = 1$*, then* ***(A1)*** *holds with* $d_1 = |E|$ *in the bipartite* G*. Hence the degree distribution exponent (defined in equation (2.3.1)) equals*

$$\tilde{\gamma} = \frac{\log N}{\log|E|} = \frac{1}{\dim_H\big(\Lambda\backslash\mathrm{Diag}\big)}.$$

Proof. This follows immediately from (2.3.1) and (2.3.4) . □

In the more general setting where we assume only that (**A1**) holds we still have a simple relation in between $\dim_{\mathrm{H}}(\Lambda \setminus \mathrm{Diag})$ and $\tilde{\gamma}$. Namely, putting together (2.3.1) and (2.3.4) and using that $|E| = n_1 \cdot d_1$ we obtain the following Corollary:

Corollary 2.3.4. *Assume that* ***(A1)*** *holds. Then*

$$N^{\mathrm{HD}-1} = d_1^{1-\tilde{\gamma}}. \tag{2.3.5}$$

Hence

$$\tilde{\gamma} = 1 - (\mathrm{HD}-1)\cdot\frac{\log N}{\log d_1}. \tag{2.3.6}$$

In particular using our model one can study the generally problematic area, where the degree distribution of the graph sequence has diverging averages. Namely, when $\gamma = 1 + \tilde{\gamma} \in (1,2)$. Using (2.3.6) we have

$$\tilde{\gamma} \in (0,1] \text{ if and only if } \mathrm{HD} \in [1,2). \tag{2.3.7}$$

By (2.3.4) this happens exactly when $|E| \geq N$, i.e. the number of edges is at least as large as the number of vertices in the base graph G. In particular,

in Example 2.2.6 we have HD $= \tilde{\gamma} = 1$, attaining the upper bound for $\tilde{\gamma}$ in Equation (2.3.7). On the other hand, our model produces graph sequences with degree exponent $\tilde{\gamma}$ arbitrarily close to zero when our base graph G is the complete bipartite graph on $|V_1| = n_1$ and $|V_2| = n_2 = n_1 + 1$ vertices, for large n_1.

In this case $N \sim 2n_1$ and $d_1 \sim n_1$. So, for large n_1 we have $\frac{\log N}{\log d_1} \sim 1$. Using (2.3.6) this yields $\tilde{\gamma} \sim 0$. More precisely, in this case (**A1**) holds with $d_1 = n_1 + 1$ and $d_2 = n_1$, thus the Hausdorff dimension and $\tilde{\gamma}$ are equal to

$$\text{HD} = \frac{\log n_1(n_1+1)}{\log N} = 2 - \frac{\log 4}{\log N},$$
$$\tilde{\gamma} = \frac{\log(2 - \frac{1}{n_1})}{\log(n_1+1)}.$$

As a consequence of our discussions above we can conclude that our model can produce graph sequences with power law exponent $\gamma \in \left(1, \frac{\log 3}{\log 2} + 1\right]$.

The geometric interpretation of the connection between HD **and** $\tilde{\gamma}$ Here we always assume that assumption (**A1**) is satisfied and in the rest of this section we focus on the case when (2.3.7) holds. That is, we assume that in our bipartite base graph G

$$N = |V(G)| < |E| = n_1 \cdot d. \tag{2.3.8}$$

Example 2.3.5 (Triplets)**.** *On Figure 2.4 we introduce a base graph G, that we call Triplets, which satisfies (**A1**) and the Hausdorff dimension is greater than* 1*.*

The following theorem describes the geometric meaning of $\tilde{\gamma}$.

Theorem 2.3.6. *Assume that both (**A1**) and (2.3.8) holds. Let ℓ_{vert} be an arbitrary vertical line that intersects Λ_{12} and let ℓ_{rand} be a randomly chosen line on the plane that intersects Λ_{12}. Then almost surely,*

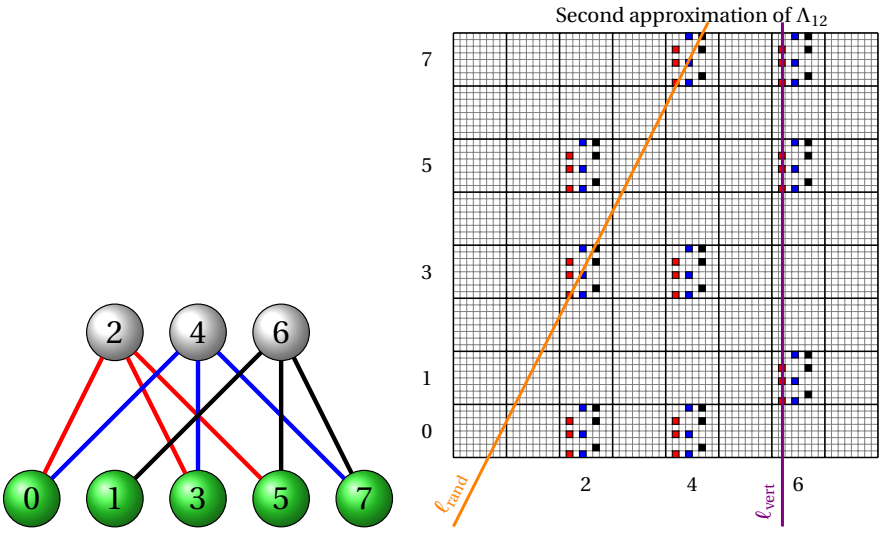
$$1 - \tilde{\gamma} = \frac{\dim_H(\ell_{\text{rand}} \cap \Lambda_{12})}{\dim_H(\ell_{\text{vert}} \cap \Lambda_{12})}. \tag{2.3.9}$$

More precisely, any straight line ℓ on the plane can be described as

$$\ell = \ell_{(a,b)} = \{(x,y) | y = a \cdot x + b\}. \tag{2.3.10}$$

Let $\mathcal{A} := \left\{(a,b) : \ell_{(a,b)} \cap \Lambda_{12} \neq \emptyset\right\}$. Then for Lebesgue almost all $(a,b) \in \mathcal{A}$ the assertion of (2.3.9) holds with $\ell_{\text{rand}} = \ell_{(a,b)}$.

That is, among those lines that intersect Λ_{12}, a randomly chosen one intersects Λ_{12} in a set of smaller Hausdorff dimension than a vertical line does. The ratio of the Hausdorff dimensions of these two intersections is equal to $1 - \tilde{\gamma}$, which gives a nice geometric characterization of $\tilde{\gamma}$.



(a) The triplets graph G.

(b) The x axis corresponds to $V_1 = \{2, 4, 6\}$ the y axis corresponds to $V_2 = \{0, 1, 3, 5, 7\}$.

Figure 2.4: The triplets example which satisfies both (**A1**) and (2.3.8)

Proof of the Theorem. From (2.3.6) we get:

$$1 - \tilde{\gamma} = \frac{\mathrm{HD} - 1}{\frac{\log d_1}{\log N}}. \tag{2.3.11}$$

It is enough to prove that

(a) $\mathrm{HD} - 1 = \dim_{\mathrm{H}} (\ell_{\mathrm{rand}} \cap \Lambda_{12})$,

(b) $\frac{\log d_1}{\log N} = \dim_{\mathrm{H}} (\ell_{\mathrm{vert}} \cap \Lambda_{12})$.

Proof of Part (a) Let $L_{P,\theta}$ be a line through a typical (in the sense of appropriate dimensional Hausdorff measure) point P of Λ_{12} with Lebesgue typical direction θ. Then a well-known Theorem of Marstrand [92, Theorem III] states that

$$\dim_{\mathrm{H}}(L_{P,\theta} \cap \Lambda_{12}) = HD - 1. \tag{2.3.12}$$

However, in Theorem 2.3.6 we phrased the notion of typicality of a line in a seemingly different (perhaps more natural) way. Below we prove that in spite of this, Marstrand theorem implies Part (a) of Theorem 2.3.6.

Let ν be the natural (evenly distributed) measure on Λ_{12} which is known [72] to be equal to constant times the appropriate dimensional Hausdorff measure:

$$\nu(\cdot) = \mathrm{const} \cdot \mathcal{H}^{HD}|_{\Lambda_{12}}(\cdot).$$

Let $W_a = \{(x, y) : y = a \cdot x\}$ and $W_a^{\perp}$ be the line through origin which is perpendicular to W_a and we call Λ_{12}^a the orthogonal projection of Λ_{12} to the

line $W_a^{\perp}$. We write ν_a for the projection of ν to $W_a^{\perp}$. It follows from the PROOF of [63, Theorem 6.8 (b)] that

$$\nu_a \ll \mathcal{L}\text{eb for } \mathcal{L}\text{eb almost every } a. \tag{2.3.13}$$

On the other hand, observe that ν_a is a self-similar measure in itself. Using the previous observation and [97, Proposition 3.1 (ii)] for ν_a we obtain that

$$\text{either } \nu_a \perp \mathcal{L}\text{eb}|_{\Lambda_{12}^a} \text{ or } \nu_a \sim \mathcal{L}\text{eb}|_{\Lambda_{12}^a}. \tag{2.3.14}$$

Putting together (2.3.13) and (2.3.14) we get

$$\nu_a \sim \mathcal{L}\text{eb}|_{\Lambda_{12}^a} \text{ for } \mathcal{L}\text{eb almost every } a. \tag{2.3.15}$$

Combining Fubini Theorem and the Theorem of Marstrand [92, Theorem III] mentioned above it follows that for ν_a-almost every c

$$\dim_{\mathrm{H}}([W_a + c] \cap (\Lambda_{12})) = \mathrm{HD} - 1. \tag{2.3.16}$$

From this and (2.3.15) we obtain that for Lebesgue almost every a satisfying $[W_a + c] \cap (\Lambda_{12}) \neq \emptyset$, the equation $\dim_{\mathrm{H}}([W_a + c] \cap (\Lambda_{12})) = \mathrm{HD} - 1$ holds also for Lebesgue almost every c. This completes the proof of Part (a).

Proof of Part (b) Let ℓ_{vert} be a vertical line that intersects Λ_{12}. Put $\Lambda' := \ell_{\mathrm{vert}} \cap \Lambda_{12}$. Then the n-th approximation of Λ' consists of d_1^n intervals with disjoint interior of length N^{-n}. The assertion of Part (b) is an immediate corollary of [29, Theorem 3.1.1]. □

2.3.3 Average shortest path in G_n

In many real networks, the typical distance between two randomly chosen points is of order $\log(|G|)$, the logarithm of the size of the network. We will see that our model also shares this property as well as the power law decay and the hierarchical structure, combining all these important features.

In this section we calculate the average length of shortest path between two nodes in G_n. First we give a deterministic way to construct one of the shortest paths between any two nodes in the graph. To do so, we need to introduce some notation. Recall that the graph G is a bipartite graph with partition V_1, V_2, see the beginning of Section 2.2. We remind the reader that for $\underline{x}, \underline{y} \in \Sigma_n$, $\mathrm{typ}(\underline{x})$, the common prefix $\underline{x} \wedge \underline{y}$ and the postfixes $\tilde{\underline{x}}, \tilde{\underline{y}}$ were defined in Definition 2.2.1.

Definition 2.3.7.

For two arbitrary vertices $\underline{x}, \underline{y} \in \Sigma_n$ we denote the length of their common prefix by $k = k(\underline{x}, \underline{y}) := |\underline{x} \wedge \underline{y}|$. Furthermore, let us decompose the postfixes $\tilde{\underline{x}}, \tilde{\underline{y}}$ into blocks of digits of the same type:

$$\tilde{\underline{x}} = \underline{b}_1 \underline{b}_2 \dots \underline{b}_r, \ \tilde{\underline{y}} = \underline{c}_1 \underline{c}_2 \dots \underline{c}_q, \tag{2.3.17}$$

such that all of the blocks have a nonzero type and the consecutive blocks are of different types. That is, for $i = 1, \dots, r-1,\ j = 1, \dots q-1$ *we have*

$$typ(b_i) \neq typ(b_{i+1}) \in \{1,2\},\ \textit{and}\ typ(c_j) \neq typ(c_{j+1}) \in \{1,2\}.$$

Note, that we denoted the number of blocks in $\tilde{\underline{x}}, \tilde{\underline{y}}$ *by* r *and* q*, respectively. If* $\underline{X}$ *and* $\underline{Y}$ *are two random vertices of* G_n*, then the same notation as in (2.3.17) is used with capital letters.*

Now we fix an arbitrary self-map p of Σ_1 such that

$$(x, p(x)) \in E(G)\ \forall x \in G.$$

Most commonly, $p(p(x)) \neq x$. Note that x and $p(x)$ have different types since G is bipartite. For a word $\underline{z} = (z_1 \dots z_m)$ with $\mathrm{typ}(\underline{z}) \in \{1,2\}$ we define $p(\underline{z}) := (p(z_1) \dots p(z_m))$.. Then,

$$(\underline{tz}, \underline{t}p(\underline{z})) \text{ is an edge in } G_{\ell+m}, \forall \underline{t} = (t_1 \dots t_\ell), \tag{2.3.18}$$

follows from (2.2.1).

As usual we write $\mathrm{Diam}(G)$ for the maximal graph-distance in the graph G within components of G. Clearly $\mathrm{Diam}(G) \leq N-1$.

Lemma 2.3.8. *Let* $\underline{x}, \underline{y}$ *be arbitrary vertices in the same connected component of* G_n*. Using the notation above, the length of the shortest path between them is at least* $r+q-1$ *and at most* $r+q+Diam(G)-2$*.*

Considering the worst case scenario, i.e. choosing all blocks of length 1 yields:

Corollary 2.3.9. *The diameter of the graph* G_n *is at most* $2n+Diam(G)-2$*. Since the size of the graph is* N^n*, therefore*

$$Diam(G_n) = \frac{2}{\log N} \log(|G_n|) + O(1).$$

Proof of Lemma 2.3.8. First we construct a path $P(\underline{x}, \underline{y})$ of minimal length. Starting from $\underline{x}$ the first half of the path $P(\underline{x}, \underline{y})$ is as follows:

$$\begin{aligned}
\hat{\underline{x}}^0 &= \underline{x} = (\underline{x} \wedge \underline{y})\underline{b}_1 \dots \underline{b}_{r-1}\underline{b}_r \\
\hat{\underline{x}}^1 &= (\underline{x} \wedge \underline{y})\underline{b}_1 \dots \underline{b}_{r-1}p(\underline{b}_r) \\
&\dots \\
\hat{\underline{x}}^{r-1} &= (\underline{x} \wedge \underline{y})\underline{b}_1 p(\underline{b}_2 \dots p(\underline{b}_{r-1}p(\underline{b}_r))),
\end{aligned}$$

Starting from $\underline{y}$ the first half of the path $P(\underline{x}, \underline{y})$ is as follows:

$$\begin{aligned}
\hat{\underline{y}}^0 &= \underline{y} = (\underline{x} \wedge \underline{y})\underline{c}_1\underline{c}_2 \dots \underline{c}_r \\
\hat{\underline{y}}^1 &= (\underline{x} \wedge \underline{y})\underline{c}_1 \dots \underline{c}_{r-1}p(\underline{c}_r) \\
&\dots \\
\hat{\underline{y}}^{q-1} &= (\underline{x} \wedge \underline{y})\underline{c}_1 p(\underline{c}_2 \dots p(\underline{c}_{r-1}p(\underline{c}_q))).
\end{aligned}$$

It follows from (2.3.18) that

$$P_x := (\hat{\underline{x}}^0, \hat{\underline{x}}^1, \dots, \hat{\underline{x}}^{r-1})$$
$$P_y := (\hat{\underline{y}}^{q-1}, \cdots \hat{\underline{y}}^1, \hat{\underline{y}}^0)$$

are two paths in G_n. To construct $P(\underline{x}, \underline{y})$ the only thing remained is to connect $\hat{\underline{x}}^{r-1}$ and $\hat{\underline{y}}^{q-1}$. Using (2.3.18) it is easy to see that this can be done with a path P_c of length at most $\mathrm{Diam}(G)$. In this way,

$$P(\underline{x}, \underline{y}) := P_x P_c P_y.$$

Clearly,

$$r + q - 1 \leq \mathrm{Length}(P(\underline{x}, \underline{y})) \leq r + q + \mathrm{Diam}(G) - 2$$

On the other hand, now we prove that no shorter paths exists than $P(\underline{x}, \underline{y})$. Recall that it follows from (2.2.1) that for any path $Q(\underline{x}, \underline{y}) = (\underline{x} = \underline{q}^0, \dots, \underline{q}^\ell = \underline{y})$, the consecutive elements of the path only differ in their postfixes, which have different types. That is,

$$\forall i, \underline{q}^i = \underline{w}^i \underline{z}^i,\ \underline{q}^{i+1} = \underline{w}^i \tilde{\underline{z}}^i,\ \text{with } \mathrm{typ}(\underline{z}^i) \neq \mathrm{typ}(\tilde{\underline{z}}^i) \in \{1, 2\}.$$

This implies that in each step on the path, the number of blocks in (2.3.17) changes by at most one. Recall that $|\underline{x} \wedge \underline{y}| = k$, so $x_{k+1} \neq y_{k+1}$. Since the digit on the $k+1$-th position changes on the path, we have to reach a point where all the digits to the right from the k-th position are of the same type. Starting from $\tilde{\underline{p}}^0 = \underline{x}$, to reach the first vertex $\underline{a}$ of this property, we need at least $r - 1$ steps on any path $\tilde{P}$, where r was defined in formula (2.3.17). Similarly, starting from $\underline{y}$, we need at least $q - 1$ steps to reach the first vertex $\underline{b}$ where all the digits after the k-th position are of the same type. Because $x_{k+1} \neq y_{k+1}$, we need at least one more edge and at most $\mathrm{Diam}(G)$ edges. □

Theorem 2.3.10. *The expectation of the length of a shortest path between two uniformly chosen vertices $\underline{X}, \underline{Y} \in G_n$ can be bounded by*

$$\frac{4n_1n_2}{N^2}(n-1) < \mathbf{E}(|P(\underline{X}, \underline{Y})|) < N + \frac{4n_1n_2}{N^2}(n-1).$$

Corollary 2.3.11. *The magnitude of the average length of a shortest path between two uniformly chosen vertices in G_n is the logarithm of the size of G_n, which is the same order as Diam(G_n).*

Proof of Theorem 2.3.10. Let $\underline{X}, \underline{Y}$ be independent, uniformly chosen vertices of G_n. In this proof we use the notation introduced in Definitions 2.2.1

and 2.3.7. The digits of the code of a uniformly chosen vertex are independent and uniform in $\{0, \dots, N-1\}$, hence $K(\underline{X}, \underline{Y}) := |\underline{X} \wedge \underline{Y}|$ has a truncated geometric distribution with parameter $\frac{N-1}{N}$. That is

$$\mathbf{P}(K(\underline{X}, \underline{Y}) = k) = \begin{cases} \left(\frac{1}{N}\right)^k \cdot \frac{N-1}{N}, \text{ if } 0 \le k < n, \\ \left(\frac{1}{N}\right)^n \text{ if } k = n. \end{cases}$$

Furthermore, given that the length of the prefix is $k = K(\underline{X}, \underline{Y})$, the random variables R and Q (see Definition 2.2.1) can be represented as the sum of indicators corresponding to the start of a new block:

$$R = 1 + \sum_{i=1}^{n-k-1} \mathbf{1}_{\mathrm{typ}(X_{k+i}) \neq \mathrm{typ}(X_{k+i+1})},$$
$$Q = 1 + \sum_{i=1}^{n-k-1} \mathbf{1}_{\mathrm{typ}(Y_{k+i}) \neq \mathrm{typ}(Y_{k+i+1})}.$$

Taking expectation yields

$$\begin{aligned} \mathbf{E}(Q | K(\underline{X}, \underline{Y}) = k) &= \mathbf{E}(R | K(\underline{X}, \underline{Y}) = k) \\ &= 1 + \mathbf{E}\left(\sum_{i=1}^{n-k-1} \mathbf{1}_{\mathrm{typ}(X_{k+i}) \neq \mathrm{typ}(X_{k+i+1})} \right) \\ &= 1 + \sum_{i=1}^{n-k-1} \mathbf{P}(\mathrm{typ}(X_{k+i}) \neq \mathrm{typ}(X_{k+i+1})) \\ &= 1 + (n-k-1)\frac{2n_1n_2}{N^2}. \end{aligned}$$

So weighting this with the geometric weights of the length of the prefix, we get

$$\begin{aligned} \mathbf{E}(Q) = \mathbf{E}(R) &= \mathbf{E}\big(\mathbf{E}(R | K(\underline{X}, \underline{Y}))\big) \\ &= \mathbf{E}\big(1 + (n - K(\underline{X}, \underline{Y}) - 1)\frac{2n_1n_2}{N^2}\big) \\ &= 1 + \left(n - \frac{1}{N-1}\left(1 - \frac{1}{N^n}\right) - 1\right)\frac{2n_1n_2}{N^2}. \end{aligned}$$

Using this and the following immediate consequence of Lemma 2.3.8

$$-1 \le \mathbf{E}(|P(\underline{X}, \underline{Y})| - (R+Q)) \le \mathrm{Diam}(G) - 2,$$

finally we obtain that

$$1 - \frac{1}{N-1} + \frac{4n_1n_2}{N^2}(n-1) \le \mathbf{E}(|P(\underline{X}, \underline{Y})| < \mathrm{Diam}(G) + \frac{4n_1n_2}{N^2}(n-1).$$

□

2.3.4 Decay of local clustering coefficient of the modified sequence $\left\{\hat{G}_n\right\}$

An important property of most real networks is the high degree of clustering. In general, the local clustering coefficient of a node v having n_v neighbors is defined as

$$C_v := \frac{\#\{\text{links between neighbors of } v\}}{\binom{n_v}{2}}.$$

Note that the numerator in the formula is the number of triangles containing v and C_v is the portion of the pairs of neighbors of v which form a triangle with v in the graph.

Observe that without the loops the graph sequence G_n is bipartite, i.e. there are no triangles in the graph G_n. However, we can modify the graph sequence G_n in a natural way, like in [100], to get a new sequence $\hat{G}_n$ preserving the hierarchical structure of G_n, still reflecting the dependence of clustering coefficient on node degree observed in several real networks. Namely, the local clustering coefficient of a vertex v is of order $1/\deg(v)$.

Definition 2.3.12.

- *We obtain the graph $\hat{G}$ adding a set of extra edges $RE(\hat{G})$ to G satisfying the following property:*

  Property R

  $\forall x \in \Sigma_1,\ \exists y, z \in \Sigma_1$, such that two among the edges of the triangle $(x, y, z)_\Delta$ are contained in $E(G)$ and one of the edges is in $RE(\hat{G})$.

  So,

  $$V(\hat{G}) = V(G) \text{ and } E(\hat{G}) = E(G) \cup RE(\hat{G}).$$

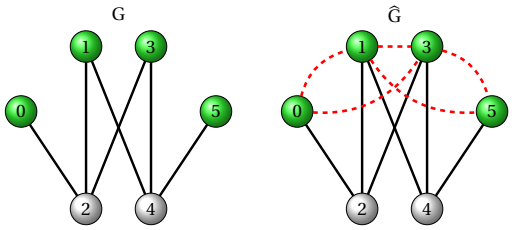
  In the example presented on Figure 2.5 the edges from $RE(\hat{G})$ are the dashed red edges.

- *Similarly we define the graph sequence $\left\{\hat{G}_n\right\}_{n=1}^{\infty}$ by deleting all loops in G_n and adding extra edges to G_n. That is, the vertices $V(\hat{G}_n) = V(G_n) = \Sigma_n$, and with the definition of the simple graph G'_n in Section 2.2.5, the edge set is extended by the following rule*

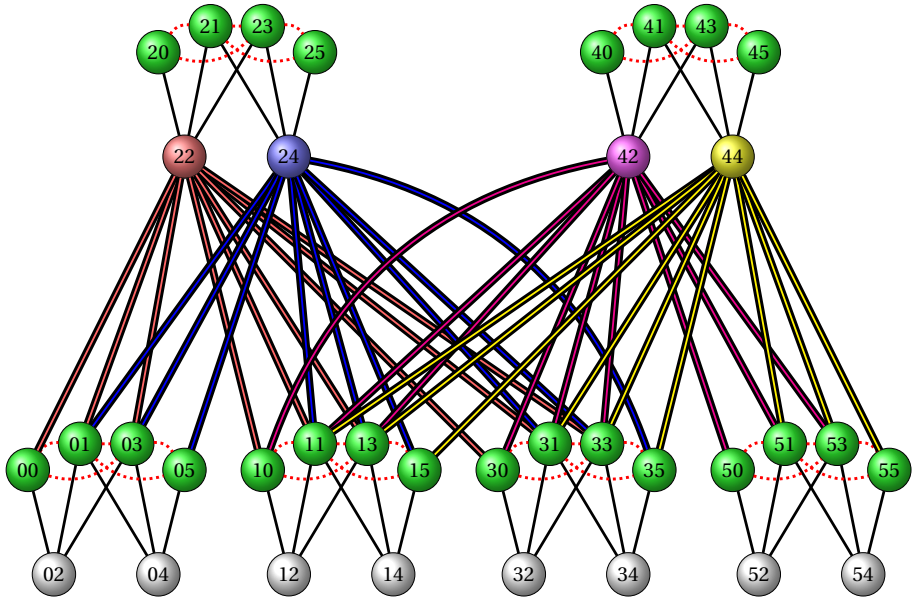
  $$E(\hat{G}_n) = E(G'_n) \bigcup RE(\hat{G}_n), \tag{2.3.19}$$

  where

  $$RE(\hat{G}_n) = \left\{ \begin{pmatrix} x_1 \dots x_n \\ y_1 \dots y_n \end{pmatrix} : x_i = y_i, i \leq n-1, \begin{pmatrix} x_n \\ y_n \end{pmatrix} \in RE(\hat{G}) \right\}. \tag{2.3.20}$$



(a) We obtain $\widehat{G}$ by adding the dashed (red) edges to G.



(b) $\widehat{G}_2$: The edges of $\widehat{G}_2$ and G_2 differ only at the lowest hierarchical level (cf. Figure 2.3)

Figure 2.5: Clustering extended "fan".

It is clear from Property **R** that

$$\hat{C}_{\min} := \min_{x \in \hat{G}} C_x > 0. \tag{2.3.21}$$

Further, using (2.2.1) and (2.3.20) one can easily see that the degree of a vertex $\underline{x} \in \hat{G}_n$ is

$$\widehat{\deg}_n(\underline{x}) = S(\underline{x}) \cdot \deg(x_n) + \left(\widehat{\deg}(x_n) - \deg(x_n)\right), \tag{2.3.22}$$

where $\widehat{\deg}(.)$ denotes the degree of a vertex in $\hat{G}$, while $\deg(.)$ stands for the degree in G.

Remark 2.3.13. *The difference between the degree of any node $\underline{x} \in \Sigma_n$ in G_n and in $\hat{G}_n$ is bounded, thus the degree sequence of $\hat{G}_n$ has the same power law exponent as G_n.*

Theorem 2.3.14. *There exists $K_1, K_2 > 0$ such that the local clustering coefficient $C_{\underline{x}}$ of an arbitrary node $\underline{x} \in \hat{G}_n$ satisfies.*

$$\frac{K_1}{\widehat{\deg}_n(\underline{x})} \leq C_{\underline{x}} \leq \frac{K_2}{\widehat{\deg}_n(\underline{x})}.$$

Proof. We write $\mathcal{T}_n(\underline{x})$ for the set of all triangles in $\hat{G}_n$ containing the node $\underline{x} \in \Sigma_n$. We say that a triangle $(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{T}_n(\underline{x})$ is regular if and only if exactly two of its edges are from $E(G_n)$. The triangle $(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{T}_n(\underline{x})$ is called irregular if it is not regular. The set of irregular triangles containing $\underline{x}$ is denoted by $\mathcal{IRT}_n(\underline{x})$. We partition the set of regular triangles $\mathcal{RT}_n(\underline{x})$ into the classes:

$$\mathcal{RT}_n(\underline{x}) = \mathcal{RT}^1_n(\underline{x}) \cup \mathcal{RT}^2_n(\underline{x})$$

in the following way: A triangle $(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{RT}_n(\underline{x})$ belongs to $\mathcal{RT}^1_n(\underline{x})$ if and only if $\underline{x}$ is NOT an endpoint of the edge contained in $RE(\hat{G}_n)$. That is

$$\mathcal{RT}^1_n(\underline{x}) := \left\{ (\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{RT}_n(\underline{x}) : \binom{\underline{x}}{\underline{y}}, \binom{\underline{x}}{\underline{z}} \in E(G_n). \right\}$$

Hence, $\mathcal{RT}^2_n(\underline{x})$ is the set of those $(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{RT}_n(\underline{x})$ for which either $\binom{\underline{x}}{\underline{y}} \in E(G_n)$ and $\binom{\underline{x}}{\underline{z}} \in RE(\hat{G}_n)$ or vice versa. Summarizing these partitions:

$$\mathcal{T}_n(\underline{x}) = \mathcal{RT}_n(\underline{x}) \cup \mathcal{IRT}_n(\underline{x}) = \mathcal{RT}^1_n(\underline{x}) \cup \mathcal{RT}^2_n(\underline{x}) \cup \mathcal{IRT}_n(\underline{x})$$

Now we define the cardinality of these classes:

$$\Delta^1_n(\underline{x}) := \#\mathcal{RT}^1(\underline{x}), \Delta^2_n(\underline{x}) := \#\mathcal{RT}^2(\underline{x}) \text{ and } \Delta^{\text{ir}}_n(\underline{x}) := \#\mathcal{IRT}(\underline{x}).$$

When $n = 1$ then we suppress the index n. Observe that by Property **R**,

$$\Delta^{\text{r}}_n(\underline{x}) := \Delta^1_n(\underline{x}) + \Delta^2_n(\underline{x}) \geq 1, \quad \forall n \geq 1, \underline{x} \in \Sigma_n.$$

Now we compute $\Delta_n^i(\underline{x})$, $i \in \{1, 2, \mathrm{ir}\}$, for an arbitrary fixed $\underline{x} \in \Sigma_n$. To do so the notation $\ell(\underline{x})$ will be used. First we verify that

$$\Delta_n^1(\underline{x}) = \sum_{r=0}^{\ell(\underline{x})-1} \prod_{j=1}^{r} \deg(x_{n-j}) \cdot \Delta^1(x_n) = S(\underline{x}) \cdot \Delta^1(x_n), \tag{2.3.23}$$

where $S(\underline{x})$ was defined in (2.2.3). To see this, observe that it follows from (2.2.1), (2.3.19) and (2.3.20) that

$$(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{RT}_n^1(\underline{x})$$

holds if and only if all of the following three assertions are satisfied:

1. $\exists 0 \leq r \leq \ell(\underline{x}) - 1,\ |\underline{y} \wedge \underline{z}| = n - 1$ and $|\underline{x} \wedge \underline{y}| = |\underline{x} \wedge \underline{z}| = n - r - 1$
2. $\binom{x_k}{y_k} \in E(G)$ whenever $n - r \leq k \leq n - 1$
3. $(x_n, y_n, z_n)_\Delta \in \mathcal{RT}^1(x_n)$.

Hence (2.3.23) is obtained by an immediate calculation.

Now we prove that

$$\Delta_n^2(\underline{x}) = \sum_{r=0}^{\ell(\underline{x})-1} \prod_{j=1}^{r} \deg(x_{n-j}) \cdot \Delta^2(x_n) = S(\underline{x}) \cdot \Delta^2(x_n). \tag{2.3.24}$$

This is so because by (2.2.1), (2.3.19) and (2.3.20) we have

$$(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{RT}_n^2(\underline{x})$$

holds if and only if all of the following three assertions are satisfied:

1. $\exists 0 \leq r \leq \ell(\underline{x}) - 1,\ |\underline{x} \wedge \underline{y}| = n - 1$ and $|\underline{x} \wedge \underline{z}| = |\underline{y} \wedge \underline{z}| = n - r - 1$,
2. $\binom{x_k}{z_k} \in E(G)$ whenever $n - r \leq k \leq n - 1$
3. $(x_n, y_n, z_n)_\Delta \in \mathcal{RT}^2(x_n)$.

Hence, using the same argument as above we get (2.3.24).

Finally, we determine the number of irregular triangles containing $\underline{x}$:

$$\Delta_n^{\mathrm{ir}}(\underline{x}) = \Delta^{\mathrm{ir}}(x_n). \tag{2.3.25}$$

This follows from the fact that

$$(\underline{x}, \underline{y}, \underline{z})_\Delta \in \mathcal{IRT}_n(\underline{x})$$

is equivalent to

$$\forall 1 \leq i \leq n - 1,\ x_i = y_i = z_i \text{ and } (x_n, y_n, z_n)_\Delta \in \mathcal{IRT}(x_n).$$

We write $Z_\Delta(\underline{x})$ for the number of all triangles in $\hat{G}_n$ containing $\underline{x}$:

$$Z_\Delta(\underline{x}) := \underbrace{\Delta_n^1(\underline{x}) + \Delta_n^2(\underline{x})}_{\Delta^{\mathrm{r}}(\underline{x})} + \Delta_n^{\mathrm{ir}}(\underline{x}).$$

Using (2.3.22), (2.3.23), (2.3.24) and (2.3.25) we get

$$C_{\underline{x}} = \frac{Z_\Delta(\underline{x})}{\binom{\widehat{\deg_n}(\underline{x})}{2}} = \frac{2\Delta^{\mathrm{r}}(x_n)\cdot S(\underline{x}) + 2\Delta^{\mathrm{ir}}(x_n)}{\widehat{\deg_n}(\underline{x})(\widehat{\deg_n}(\underline{x})-1)}, \tag{2.3.26}$$

where $S(\underline{x})$ was defined in (2.2.3). Now we estimate $C_{\underline{x}}$.

Claim 2.3.15.

(i) *If $\ell(\underline{x}) = 1$, then $C_{\underline{x}} = C_{x_n}$.*

(ii) *If $\ell(\underline{x}) \geq 2$, then we have*

$$\left| C_{\underline{x}} - \frac{2\Delta^{\mathrm{r}}(x_n)}{\deg(x_n)} \cdot \frac{1}{\widehat{\deg_n}(\underline{x})} \right| \leq \frac{\mathrm{const}}{\widehat{\deg_n}^2(\underline{x})}. \tag{2.3.27}$$

Proof of the Claim. Part **(i)** immediately follows from (2.2.1). To prove **(ii)** we fix an arbitrary $\underline{x} \in \Sigma_n$ with $\ell(\underline{x}) \geq 2$. Since t, u, v introduced below depend only on x_n there exists a constant C_* independent of n and $\underline{x}$ such that

$$0 \leq t := \frac{\Delta^r(\underline{x})}{\deg(x_n)}, \quad u := \widehat{\deg}(x_n) - \deg(x_n), \quad v := 2\Delta^{\mathrm{ir}}(x_n) < C_*. \tag{2.3.28}$$

To prove (2.3.27) it is enough to verify that

$$Q := \left(\widehat{\deg_n}(\underline{x})\right)\left(\widehat{\deg_n}(\underline{x}) - 1\right) \cdot C_{\underline{x}} - 2t \cdot (\widehat{\deg_n}(\underline{x}) - 1)$$

is bounded in n and $\underline{x} \in \Sigma_n$. This so, because by (2.3.22) and (2.3.26) we have

$$\begin{aligned} Q &= 2\Delta^r(x_n)\cdot S + v - 2t\big(\underbrace{S\cdot\deg(x_n) + u}_{\widehat{\deg}(\underline{x})} - 1\big) \\ &= 2\Delta^r(x_n)\cdot S + v - \underbrace{2\Delta^r(x_n)\cdot S}_{2tS\deg(x_n)} - 2t(u-1) \\ &= v - 2t(u-1), \end{aligned}$$

which is bounded by (2.3.28). □

Property **R** implies that both C_{x_n} and $\frac{\Delta^{\mathrm{r}}(x_n)}{\deg(x_n)}$ are bounded away from zero. This completes the proof of the Theorem 2.3.14. □

The following theorem shows that the graph sequence $\hat{G}_n$ displays similar features to that of considered in [100], namely, the average local clustering coefficient of the graphs $\hat{G}_n$ is not tending to zero with the size of $\hat{G}_n$.

Theorem 2.3.16. *The average local clustering coefficient $\bar{C}(\hat{G}_n)$ of the graph $\hat{G}_n$ is bounded by two positive constants, more precisely*

$$\frac{2n_1n_2\hat{C}_{\min}}{N^2} \le \bar{C}(\hat{G}_n) \le \bar{C}(\hat{G}), \tag{2.3.29}$$

where $\hat{C}_{\min}$ was defined in (2.3.21).

Proof. We will use the notation introduced in the proof of Theorem 2.3.14. It easily follows from the proof of Theorem 2.3.14 that

$$C_{\underline{x}} \le C_{x_n}. \tag{2.3.30}$$

Namely, if $\ell(\underline{x}) = 1$ then by (2.3.20), $C_{\underline{x}} = C_{x_n}$. If $\ell(\underline{x}) \ge 2$ then $S(\underline{x}) \ge 1$ thus using (2.3.26) we obtain

$$\begin{aligned} C_{\underline{x}} &\le \underbrace{\frac{\Delta^{\mathrm{r}}(x_n) + \Delta^{\mathrm{ir}}(x_n)}{\binom{\deg(x_n)}{2}}}_{C_{x_n}} \cdot \underbrace{\binom{\deg(x_n)}{2} \cdot \frac{S(\underline{x})}{\binom{\widehat{\deg}_n(\underline{x})}{2}}}_{\le 1} \cdot \\ &\le C_{x_n}. \end{aligned}$$

This completes the proof of (2.3.30) from which the upper estimate of (2.3.29) follows by averaging. On the other hand to see that the lower estimate holds we take into consideration only the contribution of $\underline{x} \in \Sigma_n$ with $\ell(\underline{x}) = 1$.

$$\bar{C}(\hat{G}_n) > \frac{1}{N^n}\left(\sum_{z\in V_1} N^{n-2}n_2C_z + \sum_{z\in V_2} N^{n-2}n_1C_z\right)$$

Using $C_z > \hat{C}_{\min}$, the lower bound of (2.3.29) follows. □

2.4 The randomized model

In this section we randomize the deterministic model in Section 2.2 by using Λ in $[0,1]^2$. The random graph sequence G_n^{r} is generated in a way which was inspired by the W-random graphs introduced by Lovász and Szegedy [88]. See also [39].

Fix a deterministic model with a base graph G, $|V(G)| = N$. This determines $\Lambda(a,b)$ the limit of the sequence of scaled adjacency matrices,

see the definition (2.2.7) and (2.2.6) in Section 2.2.2. Now for each n, we throw $M_n + 1$ independent, uniform random numbers over $[0, 1]$:

$$X^{(1)}, X^{(2)}, \dots, X^{(M_n+1)} \sim U[0,1], \text{ i.i.d.}$$

We denote the N-adic expansion of each of these numbers by

$$\underline{X}^{(i)} = (X_1^i, X_2^i, \dots), \quad \text{i.e. } X^{(i)} = \sum_{k=1}^{\infty} \frac{X_k^i}{N^k},$$

where the X_k^i-s are uniform over the set $\{0, 1, \dots, N-1\}$. The n-th approximation of $X^{(i)}$ is

$$X_{[n]}^{(i)} = \sum_{k=1}^{n} \frac{X_k^i}{N^k}, \quad \underline{X}_n^{(i)} = (X_1^i, \dots, X_n^i).$$

Now we construct the random graph G_n^{r} as follows: $|V(G_n^{\mathrm{r}})| = \{1, \dots, M_n\}$, and $E(G_n^{\mathrm{r}})$ is given by

$$E(G_n^{\mathrm{r}}) = \left\{ (i,j) \middle| \text{ int} \left(I_{X_{[n]}^{(i)}} \times I_{X_{[n]}^{(j)}} \right) \cap \Lambda \neq \emptyset \right\},$$

where int denotes the interior of a set. Clearly,

$$E(G_n^{\mathrm{r}}) = \left\{ (i,j) | \Lambda_n(X^{(i)}, X^{(j)}) = 1 \right\}.$$

Note that

$$\Lambda_n(X^{(i)}, X^{(j)}) = 1 \;\Leftrightarrow\; \begin{pmatrix} X_1^i \dots X_n^i \\ X_1^j \dots X_n^j \end{pmatrix} \in E(G_n).$$

Namely, we can think of the first n digits $(X_1^i, \dots, X_n^i)$ and $(X_1^j, \dots, X_n^j)$ of the N-adic expansion of $X^{(i)}$ and $X^{(j)}$ as vertices in G_n. We draw an edge between the two vertices i and j in G_n^{r} if the vertices $(X_1^i \dots X_n^i)$ and $(X_1^j \dots X_n^j)$ are connected by an edge in the deterministic model G_n. This gives the following probabilistic interpretation of the random model:

Remark 2.4.1. *Consider the deterministic graph sequence G_n with urns sitting at each vertex $v \in G_n$. Now throw $M_n + 1$ balls independently and uniformly into the urns, and connect vertex i to vertex j by an edge in the random graph G_n^{r} if and only if the urns of ball i and j are connected by and edge in G_n.*

We need to introduce some further notation.

Frequently used definitions. Under assumption **A1**, for an $\underline{x} \in G_n$ with $\ell(\underline{x}) = k$ the degree of $\underline{x}$ is

$$t_k := \frac{d_1^{k+1} - 1}{d_1 - 1} + 1,$$

independently of the length of n.

In the random graph G_n^{r}, the conditional probability of the degree distribution of a random node $V \in \{0, \dots, M_n\}$ conditioned on the first n digits of the N-adic expansion of the corresponding code $X^{(V)}$ follows a Binomial distribution:

$$\left(\deg(V) | (X_1^V \dots X_n^V) = \underline{x}\right) \sim BIN\left(M_n, \frac{t_{\ell(\underline{x})}}{N^n}\right). \tag{2.4.1}$$

This follows from the characterization of G_n^{r} described in Remark 2.4.1. Namely, assume that the V-th ball has landed in urn with label $\underline{x} \in \Sigma_n$. In G_n there are exactly $\deg_n(\underline{x}) - 1 = t_{\ell(\underline{x})}$ vertices $\underline{y} \in \Sigma_n$ that are connected to $\underline{x}$. All the balls landing into urns corresponding to these vertices $\underline{y}$ will be connected to V in G_n^{r}.

2.4.1 Properties of the randomized model

In this section we determine the proportion of isolated vertices and characterize the degree sequence.

Isolated vertices

Theorem 2.4.2. *If $M_n = c_n N^n$ with $\lim_{n\to\infty} c_n = \infty$, then the fraction of isolated vertices tends to zero as $n \to \infty$. More precisely, for a uniformly chosen node $V \in G_n^r$,*

$$\mathbf{P}\left(\deg(V) = 0\right) \leq e^{-d_{\min} c_n},$$

where $d_{\min}$ stands for the minimal degree in the base graph G, and in $\deg(.)$ we do not count the loops.

The following corollary is an immediate consequence of the Borel-Cantelli lemma.

Corollary 2.4.3. *If $\sum_{n=1}^{\infty} c_n N^n e^{-d_{\min} c_n} < \infty$, then almost surely there will be only finitely many n-s, for which the graph G_n^{r} has isolated vertices.*

The assumption of the Corollary is satisfied if e.g. $c_n > n \log(N+1)$.

Proof of Theorem 2.4.2. Given the N-adic expansion of $X^{(V)}$, the probability that a vertex is isolated depends on how many neighbors the vertex $(X_1^V \dots X_n^V)$ has in the deterministic model. So we can write

$$\mathbf{P}\left(\deg(V) = 0\right) = \sum_{\underline{x} \in \Sigma_n} \mathbf{P}(\deg(V) = 0 | (X_1^V \dots X_n^V) = \underline{x}) \cdot \frac{1}{N^n}$$

As we have already seen, $\big(\deg(V)|(X_1^V \dots X_n^V) = \underline{x}\big)$ follows a Binomial distribution with parameters M_n and $\frac{\deg_n(\underline{x})-1}{N^n}$, so the conditional probability of isolation is

$$\begin{aligned}\mathbf{P}(\deg(V) = 0|(X_1^V \dots X_n^V) = \underline{x}) &= \left(1 - \frac{t_{\ell(\underline{x})}}{N^n}\right)^{M_n} \\ &\le e^{-\deg_n(\underline{x})c_n}(1 + o(1)).\end{aligned}$$

Obviously $e^{-\deg_n(\underline{x})c_n} \le e^{-d_{\min}c_n}$ holds for all $\underline{x} \in \Sigma_n$, which completes the proof. □

Decay of degree distribution

Fix a constant K such that for a standard normal variable Z, $\mathbf{P}(|Z| > K) < e^{-10}$. We write

$$I_{k,n} := [c_n t_k - K\sqrt{c_n t_k}, c_n t_k + K\sqrt{c_n t_k}],$$

and

$$k_0(n) := \max\left\{(n+1)\frac{\log d_2}{\log d_1}, \frac{\log n}{\log d_1}\right\}.$$

Now we describe the degree distribution for the random model.

Theorem 2.4.4. *Let $k > k_0(n)$ and $u \in I_{k,n}$. Then for a uniformly chosen node V in G_n^r*

$$\mathbf{P}\left(\deg(V) = u\right) = \Big(\frac{n_1}{N}\Big)^k \frac{n_2}{N} \cdot \frac{1}{\sqrt{c_n t_k}} \phi\Big(\frac{u - c_n t_k}{\sqrt{c_n t_k(1 - \frac{t_k}{N^n})}}\Big)\big(1 + O(\frac{1}{\sqrt{c_n t_k}})\big),$$

where ϕ denotes the density function of a standard Gaussian variable.

This immediately implies

Corollary 2.4.5. *The degree distribution of the random model is given by the following formula for $a, b \in [-K, K]$:*

$$\begin{aligned}\mathbf{P}\left(\deg(V) \in [c_n t_k + a\sqrt{c_n t_k}, c_n t_k + b\sqrt{c_n t_k}]\right) &= \Big(\frac{n_1}{N}\Big)^k \frac{n_2}{N} \cdot (\Phi(b) - \Phi(a)) \\ &\quad + O\Big(\Big(\frac{n_1}{N}\Big)^k \frac{1}{\sqrt{c_n t_k}}\Big),\end{aligned}$$

where $k > k_0(n)$ and Φ denotes the distribution function of a standard Gaussian variable. So, for $u \in I_{k,n}$, $k > k_0(n)$ the tail of the probability distribution is:

$$\begin{aligned}\mathbf{P}(\deg(V) > u) = \Big(\frac{n_1}{N}\Big)^{k+1} &+ \Big(\frac{n_1}{N}\Big)^k \frac{n_2}{N}\left(1 - \Phi\Big(\frac{u - c_n t_k}{\sqrt{c_n t_k(1 - \frac{t_k}{N^n})}}\Big)\right) \\ &+ \Big(\frac{n_1}{N}\Big)^{k+1} O\Big(\frac{1}{\sqrt{c_n t_k}}\Big).\end{aligned} \tag{2.4.2}$$

This holds because $\mathbf{P}(\deg(V) > u)$ equals the sum of all probability mass that is concentrated around t_l-s for $l \geq k+1$, resulting in the first term, plus the second term coming from the part greater than u of the binomial mass around t_k. As a consequence, the decay of the degree distribution follows a power law. Namely, the following holds

Theorem 2.4.6. *Let*

$$\gamma := 1 + \frac{\log(\frac{N}{n_1})}{\log d_1}.$$

Then the decay of the degree distribution is:

$$\mathbf{P}(\deg(V) > u) = u^{-\gamma+1} \cdot L(u),$$

where $L(u)$ is a bounded function:

$$\frac{n_1}{N} \leq L(u) \leq \frac{N}{n_1}.$$

The idea of the proof of Theorem 2.4.4. The conditional distribution of the degree of a node V conditioned on the n-digit N-adic expansion of $X_n^{(V)} = \underline{x}$ follows a $BIN(c_n N^n, \frac{t_{\ell(\underline{x})}}{N^n})$ law. This is close to a $POI(c_n t_{\ell(\underline{x})})$ random variable, because c_n and $t_{\ell(\underline{x})}$ tend to infinity in a much smaller order than N^n. Now for the $POI(c_n t_{\ell(\underline{x})})$ variable, the Central Limit Theorem holds with an error term of order $1/\sqrt{c_n t_{\ell(\underline{x})}}$. Now the unconditional degree distribution comes from the law of total probability and from the fact that all other errors are negligible. □

Proof of Theorem 2.4.4. We determined the degree distribution of the deterministic model under assumption (**A1**), see Section 2.3.1 for details. Recall that if $k > k_0(n)$, then the mass at t_k is

$$p_k := \mathbf{P}(\ell(\underline{x}) = k) = \left(\frac{n_1}{N}\right)^k \frac{n_2}{N}.$$

We show that in the random model G_n^{r}, these Dirac masses are turned into Gaussian masses centered at $c_n t_k$. Suppose $u \in I_{k,n}$. By the law of total probability, we have

$$\begin{aligned}\mathbf{P}(\deg(V) = u) = \mathbf{P}(\deg(V) = u | \left(X_1^V \dots X_n^V\right) = \underline{x}, \ell(\underline{x}) = k) \cdot p_k \\ + S_1 + S_2,\end{aligned} \tag{2.4.3}$$

where

$$S_1 = \sum_{j=1}^{k-1} \mathbf{P}(\deg(V) = u | \left(X_1^V \dots X_n^V\right) = \underline{x}, \ell(\underline{x}) = j) \cdot p_j$$

$$S_2 = \sum_{j=k+1}^{n} \mathbf{P}(\deg(V) = u | \left(X_1^V \dots X_n^V\right) = \underline{x}, \ell(\underline{x}) = j) \cdot p_j$$

S_1 and S_2 combines the total contribution of cases when $\ell(X_1^V \dots X_n^V) \neq k$, i.e. referring to the urn model of our random graph, $S_1 + S_2$ settles the cases when the random ball V falls into an urn which has degree different from t_k in G_n. As a first step in our proof we show that the right hand side in the first line of (2.4.3) gives the formula in Theorem 2.4.4, then as a second step we verify that $S_1 + S_2$ is negligible.

First step: Following the standard proof of the local form of de Moivre-Laplace CLT, we obtain that for $u \in I_{k,n}$

$$\mathbb{P}\Big(\deg(V) = u | \left(X_1^V \dots X_n^V\right) = \underline{x}\Big)$$
$$= \frac{1}{\sqrt{c_n t_{\ell(\underline{x})}(1 - \frac{t_{\ell(\underline{x})}}{N^n})}} \phi\left(\frac{u - c_n t_{\ell(\underline{x})}}{\sqrt{c_n t_{\ell(\underline{x})}(1 - \frac{t_{\ell(\underline{x})}}{N^n})}}\right) \cdot \left(1 + O(\frac{1}{\sqrt{c_n t_{\ell(\underline{x})}}})\right).$$

We can neglect $1 - \frac{t_{\ell(\underline{x})}}{N^n}$. This completes the first step.

Second step: Since $u \in I_{k,n}$ we have:

$$\begin{aligned} S_1 &\leq \sum_{j=1}^{k-1} \mathbf{P}(\deg(V) > t_k - K\sqrt{t_k} | (X_1^V \dots X_n^V) = \underline{x}, \ell(\underline{x}) = j) \cdot p_j \\ S_2 &\leq \sum_{j=k+1}^{n} \mathbf{P}(\deg(V) < t_k + K\sqrt{t_k} | (X_1^V \dots X_n^V) = \underline{x}, \ell(\underline{x}) = j) \cdot p_j \end{aligned} \tag{2.4.4}$$

Now we use the fact known from Chernoff-bounds: for an $Z \sim BIN(m, p)$ variable

$$\mathbf{P}(Z \geq (1 + \delta)\mathbf{E}(Z)) \leq e^{-\frac{1}{2}\delta^2 \mathbf{E}(Z)},$$

and the same bound holds for $\mathbf{P}(Z \leq (1 - \delta)\mathbf{E}(S))$. By (2.4.1), to estimate each summand in (2.4.4) we can apply these inequalities for $Z_j \sim BIN(c_n N^n, \frac{t_j}{N^n})$, $j \in \{1, \dots, n\} \setminus \{k\}$, yielding an upper bound

$$\begin{aligned} S_1 + S_2 &\leq \sum_{j=1}^{k-1} e^{-\frac{1}{2} d_1^{2k-j} c_n} \cdot p_j + \sum_{j=k+1}^{n} e^{-\frac{1}{2}(1 - d_1^{k-j})^2 d_1^j c_n} \cdot p_j \\ &\leq e^{-\frac{1}{8} d_1^k c_n}. \end{aligned}$$

Since $e^{-\frac{1}{8} d_1^k c_n} = o(\frac{1}{\sqrt{c_n t_k}})$, the statement of Theorem 2.4.4 follows. □

Now we are ready to prove the main result of the section.

Proof of Theorem 2.4.6. If $u \in I_{k,n}$, then

$$u = d_1^k \cdot \left(1 + O\left(\frac{1}{d}\right)\right).$$

Using (2.4.2) we obtain that there exists $C(u) \in [\frac{n_1}{N}, 1]$ such that

$$\mathbf{P}(\deg(V) > u) = \left(\frac{n_1}{N}\right)^k C(u).$$

The last two formulas immediately imply the assertion of the Theorem whenever $u \in I_{k,n}$. Actually in this case we have $\frac{n_1}{N} \leq L(u) \leq 1$. If $u \notin \cup_k I_{n,k}$, then there exists $k = k(u)$ such that $u \in (c_n t_k, c_n t_{k+1})$. By monotonicity of the distribution function we have

$$\mathbf{P}(\deg(V) > c_n t_{k+1}) \leq \mathbf{P}(\deg(V) > u) \leq \mathbf{P}(\deg(V) > c_n t_k).$$

Applying the theorem for $c_n t_{k+1}$ and $c_n t_k$, we loose a factor of $\frac{N}{n_1}$ in the upper bound of $L(u)$ and the assertion of the Theorem follows. □

2.5 Conclusion

Using the attractor Λ of a self-similar graph-directed iterated function system, which consists of homotheties of contraction ratio $1/N$, we construct both deterministic and random graph sequences. In this way we obtain random and deterministic network models which share some of the important properties of real networks: scale-free property, clustering and the length of average shortest path.

Acknowledgement

The research was supported by the NKTH OTKA grant # 7778. We would like to say thanks to Béla Bollobás and Tamás Vicsek for useful conversations.

Chapter 3

Fluctuation bounds in a class of deposition processes

3.1 Introduction

This chapter studies fluctuations in deposition processes of the following type. An integer-valued height function

$$\underline{h}(t) = \{h_i(t)\}_{i\in\mathbb{Z}}$$

evolves via random deposition and removal of individual bricks of unit length and height. The Poisson rates of deposition and removal at point i are allowed to depend on the neighboring increments $h_{i-1} - h_i$ and $h_i - h_{i+1}$. Assumptions are made on these rates to guarantee stochastic monotonicity (attractivity) and the existence of a family of product-form stationary distributions μ^ϱ for the increments $\{h_{i-1} - h_i : i \in \mathbb{Z}\}$. The family of invariant measures is indexed by the average slope $\varrho = \mathbf{E}^\varrho(h_{i-1} - h_i)$. The flux function $\mathcal{H}(\varrho) = t^{-1}\mathbf{E}^\varrho(h_i(t) - h_i(0))$ gives the average velocity of the height as a function of the slope ϱ. In this chapter we consider *asymmetric* systems for which $\mathcal{H}''(\varrho) < 0$ at least in a neighborhood of a particular density value ϱ.

Height increments are conserved because every deposition and removal event causes a change of $+1$ in one increment and a change of -1 in a neighboring increment. The increments (when non negative) are naturally regarded as occupation numbers of particles. Figure 3.1 shows a configuration and a possible step with both walls and particles. It is in the particle guise that many of these processes appear in the literature: simple exclusion processes, zero range processes and misanthrope processes are examples included in the class studied in this chapter. In the particle picture the parameter ϱ that indexes invariant distributions is the mean particle density per site. Height increment $h_i(t) - h_i(0)$ is the cumulative net particle current across the edge $(i,\, i+1)$ during time $(0,\, t]$.

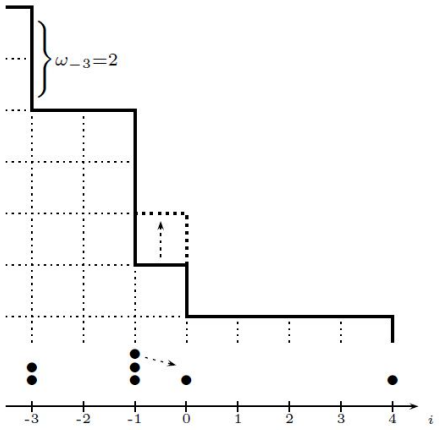


Figure 3.1: The wall and the particles with a possible step

Fix ϱ and consider the process $\underline{h}(t)$ with stationary increments at average slope ϱ, normalized so that $h_0(0) = 0$. Interesting fluctuations can be found by observing the height $h_{\lfloor V^\varrho t \rfloor}(t)$ in the characteristic direction $V^\varrho = \mathcal{H}'(\varrho)$. In the particle picture the height fluctuations in the characteristic direction become fluctuations of the cumulative net particle current seen by an observer traveling at the characteristic velocity.

Rigorous results on these fluctuations exist for examples that fall in two categories.

Order $t^{1/4}$ fluctuations. When $\mathcal{H}$ is linear the fluctuations are of order $t^{1/4}$ and converge to Gaussian processes related to fractional Brownian motion. This has been proved for independent particles [60, 79, 103] and the random average process [18, 66].

Order $t^{1/3}$ fluctuations. When $\mathcal{H}''(\varrho) \neq 0$ the fluctuations are of order $t^{1/3}$ and converge to distributions and processes related to the Tracy-Widom distributions from random matrix theory. The most-studied examples are the totally asymmetric simple exclusion process (TASEP), the polynuclear growth model (PNG) and the Hammersley process. Two types of mathematical work should be distinguished.

(a) Exact limit distributions have been derived with techniques of asymptotic analysis applied to determinantal representations of the probabilities of interest. Most of this work has dealt with particular deterministic initial conditions, and the stationary situation has been less studied. The seminal results appeared in [8] for the last-passage version of the Hammersley process and in [74] for the last-passage model associated with TASEP. Current fluctuations for stationary TASEP were analyzed in [67]. Here is a selection of further results in this direction: [9, 42, 68, 75, 99].

(b) Probabilistic approaches exist to prove fluctuation bounds of the correct order. The seminal work [44] was on the last-passage version of the Hammersley process, and then the approach was adapted to the last-passage model associated with TASEP [13]. The next step was the development of a proof that works for particle systems: the asymmetric simple exclusion process (ASEP) was treated in [23] and the totally asymmetric zero range process with constant jump rate in [15]. The ASEP work [23] was the first to prove $t^{1/3}$ order of fluctuations for a process where particle motion is not restricted to totally asymmetric.

The present chapter is based on two papers, both of them joint with Márton Balázs and Timo Seppäläinen. The first one is [17], which take a further step toward universality of the $t^{1/3}$ order for fluctuations in the case $\mathcal{H}''(\varrho) \neq 0$. In [17] we develop a general strategy for proving that in a stationary process fluctuations in the characteristic direction have order of magnitude $t^{1/3}$, then in [16] we show that the strategy works for a process obeying convex flux function. In its present form the argument rests on a nontrivial hypothesis that involves control of second class particles. This control of second class particles that we require is a microscopic counterpart of the macroscopic effect that convexity or concavity of $\mathcal{H}$ has on characteristics. Throughout the first part of the chapter we consider the concave case $\mathcal{H}''(\varrho) < 0$, hence we name the property *microscopic concavity*, then in Section 3.7 we show that the proof works for convex flux functions as well.

Once the microscopic concavity assumption is made (in a form that we make technically precise in Section 3.2.6) the proof works for the entire class of processes. This then is the sense in which we take a step toward universality. As a bi-product, we also obtain superdiffusivity of the second class particle in the stationary process. Earlier proofs of $t^{1/3}$ fluctuations have been quite rigid in the sense that they work only for particular cases of the models where special combinatorial properties emerge as if through some fortuitous coincidences. There is basically no room for perturbing the rules of the process. By contrast, the proof given in the present chapter works for the whole class of processes. The hypothesis of microscopic concavity that is required is certainly nontrivial. But it does not seem to rigidly exclude all but a handful of the processes in the broad class. The estimates that it requires can probably be proved in different ways for different subclasses of the processes. And the proof itself may evolve further and weaken the hypothesis required.

We are currently able to verify the required hypothesis of microscopic concavity for the following three subclasses of processes.

(i) The asymmetric simple exclusion process (ASEP). Full details of this case are reported by Balazs and Seppäläinen [22] and we give a brief informal description in Section 3.2.8. This proof is somewhat simpler than the earlier one given in [23].

(ii) Totally asymmetric zero range processes with a concave jump rate function whose slope decreases geometrically, and may be eventually constant. This example is developed fully here, and thus the proofs are simpler than the one given in my Masters thesis [15].

(iii) The totally asymmetric bricklayers process with convex, exponential jump rate. This system satisfies the analogous *microscopic convexity*. Due to the fast growth of the jump rate function this example needs more preliminary work and so the result is shown in Section 3.7. We postpone a more thorough introduction to bricklayer processes until then.

We expect that a broader class of totally asymmetric concave zero range processes should be amenable to further progress because a key part of the hypothesis can be verified, and only a certain tail estimate is missing. We explain this in Section 3.2.8.

This chapter has three parts. In the main part we prove the general fluctuation bound under the assumptions needed for membership in the class of processes and the assumption of microscopic concavity. The second and the third part shows that the assumptions required by the general result are satisfied by a class of zero range processes, and the exponential bricklayers process, respectively. Here is a section by section outline.

In Section 3.2 we define the general family of processes under consideration, describe the microscopic concavity property and other assumptions used, and state the general results. Partly as corollaries to the fluctuation bound along the characteristic we obtain a law of large numbers for a second class particle and limits that show how fluctuations in non-characteristic directions on the diffusive scale come directly from fluctuations of the initial state. Section 3.2.8 describes two examples. First, it gives a brief description of how the asymmetric simple exclusion process (ASEP) satisfies the assumptions of our general theorem. (Full details for this example are reported in [22].) Then it describes a class of totally asymmetric zero range processes with concave jump rates that increase with exponentially decaying slope.

The general theorem is proved in two parts: the upper bound in Section 3.3 and the lower bound in Section 3.4. Section 3.5 proves a strong law for the second class particle, partly as a corollary of the main fluctuation bounds. We then return to the zero range example and give a complete proof for this class of processes in Section 3.6. Finally, Section 3.7 handles the microscopic convexity property of the exponential bricklayers process.

The appendices contain auxiliary computations for the stationary distribution and hydrodynamic flux function. In particular we show monotonicity of the measures μ and $\hat{\mu}$ in ϱ and regularity properties of the flux function. Further, if the jump rate function of a zero range process is concave and not linear then the hydrodynamic flux $\mathcal{H}$ is smooth and satisfies $\mathcal{H}''(\varrho) < 0$ for all densities $0 < \varrho < \infty$. We omit to include the very first part of the

Appendix of [17] showing that the hydrodinamic flux function is a convex function of the density ϱ, and we refer the reader to [20].

Notation

We summarize here some notation for easy reference. $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$, $\mathbb{R}^+ = [0, \infty)$. Centering a random variable is denoted by $\widetilde{X} = X - EX$. Constants $C_\cdot$, $\alpha_\cdot$ do not depend on time, but may depend on the density parameter ϱ and their values can change from line to line. The numbering of these constants is of no particular significance and is meant only to facilitate following the arguments.

3.2 Definitions and results

We define the class of processes studied in this chapter, give a list of examples, and discuss some of basic properties. Then come the hypotheses and main results of this chapter, followed by two examples of subclasses of processes for which the hypotheses can be verified.

3.2.1 A family of deposition processes

The family of processes we consider is the one described in [21], and we repeat the definition here. We start with the interface growth picture, but we end up using the height and particle languages interchangeably. For extended-integer-valued boundaries $-\infty \le \omega^{\min} \le 0$ and $1 \le \omega^{\max} \le \infty$ define the single-site state space

$$I := \left\{ z \in \mathbb{Z} \,:\, \omega^{\min} - 1 < z < \omega^{\max} + 1 \right\}$$

and the increment configuration space

$$\Omega := \{\underline{\omega} = (\omega_i)_{i \in \mathbb{Z}} \;:\; \omega_i \in I\} = I^{\mathbb{Z}}.$$

At times it will be convenient to have notation for the increment configuration $\underline{\delta}_i \in \Omega$ with exactly one nonzero entry equal to 1:

$$(\underline{\delta}_i)_j = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{for } i \neq j. \end{cases} \tag{3.2.1}$$

For each pair of neighboring sites i and $i+1$ of $\mathbb{Z}$ imagine a column of bricks over the interval $(i, i+1)$. The height h_i of this column is integer-valued. The components of a configuration $\underline{\omega} \in \Omega$ are the negative discrete gradients of the heights: $\omega_i = h_{i-1} - h_i \in I$.

The evolution is described by jump processes whose rates p and q are nonnegative functions on $I \times I$. Two types of moves are possible. A brick can be deposited:

$$\left.\begin{aligned}(\omega_i,\, \omega_{i+1}) &\longrightarrow (\omega_i - 1,\, \omega_{i+1} + 1)\\ h_i &\longrightarrow h_i + 1\end{aligned}\right\} \text{with rate } p(\omega_i,\, \omega_{i+1}), \tag{3.2.2}$$

or removed:

$$\left.\begin{aligned}(\omega_i,\, \omega_{i+1}) &\longrightarrow (\omega_i + 1,\, \omega_{i+1} - 1)\\ h_i &\longrightarrow h_i - 1\end{aligned}\right\} \text{with rate } q(\omega_i,\, \omega_{i+1}). \tag{3.2.3}$$

Conditionally on the present state, these moves happen independently at all sites i. We can summarize this information in the formal infinitesimal generator L of the process $\underline{\omega}(\cdot)$:

$$\begin{aligned}(L\varphi)(\underline{\omega}) = &\sum_{i\in\mathbb{Z}} p(\omega_i,\, \omega_{i+1}) \cdot [\varphi(\dots,\, \omega_i - 1,\, \omega_{i+1} + 1,\, \dots) - \varphi(\underline{\omega})]\\ &+ \sum_{i\in\mathbb{Z}} q(\omega_i,\, \omega_{i+1}) \cdot [\varphi(\dots,\, \omega_i + 1,\, \omega_{i+1} - 1,\, \dots) - \varphi(\underline{\omega})]\,.\end{aligned} \tag{3.2.4}$$

L acts on bounded cylinder functions $\varphi : \Omega \to \mathbb{R}$ (this means that φ depends only on finitely many ω_i-values).

Thus we have a Markov process $\{\underline{\omega}(t) : t \in \mathbb{R}^+\}$ of an evolving increment configuration and a Markov process $\{\underline{h}(t) : t \in \mathbb{R}^+\}$ of an evolving height configuration. The initial increments $\underline{\omega}(0)$ specify the initial height $\underline{h}(0)$ up to a vertical translation. We shall always normalize the height process so that $h_0(0) = 0$.

In the particle picture the variable $\omega_i(t)$ represents the number of particles at site i at time t. Step (3.2.2) represents a rightward jump of a particle over the edge $(i,\, i+1)$, while step (3.2.3) represents a leftward jump. (If negative ω-values are permitted, one needs to consider particles and antiparticles, with antiparticles jumping in the opposite direction.) It will be useful to see that

$$\begin{aligned}h_i(t) = h_i(t) - h_0(0) = &\text{ the net number of particles that have passed,}\\ &\text{from left to right, the straight-line space-time path}\\ &\text{that connects } (1/2,\, 0) \text{ to } (i + 1/2,\, t).\end{aligned} \tag{3.2.5}$$

We impose the following four assumptions (3.2.6)–(3.2.9) on the rates.

- The rates $p,\, q : I \times I \to \mathbb{R}^+$ must satisfy

$$p(\omega^{\min},\, \cdot\,) \equiv p(\,\cdot\,,\, \omega^{\max}) \equiv q(\omega^{\max},\, \cdot\,) \equiv q(\,\cdot\,,\, \omega^{\min}) \equiv 0 \tag{3.2.6}$$

  whenever either $\omega^{\min}$ or $\omega^{\max}$ is finite. Either both p and q are strictly positive in all other cases, or one of them is identically zero. The process is called *totally asymmetric* if either $q \equiv 0$ or $p \equiv 0$.

- The dynamics has a smoothing effect when we assume the following monotonicity:

$$\begin{aligned} p(z+1,\, y) &\geq p(z,\, y), \qquad & p(y,\, z+1) &\leq p(y,\, z) \\ q(z+1,\, y) &\leq q(z,\, y), \qquad & q(y,\, z+1) &\geq q(y,\, z) \end{aligned} \tag{3.2.7}$$

  for $y,\, z,\, z+1 \in I$. Under this property the higher the neighbors of a column, the faster it grows and the longer it waits for a brick removal, on average. This is the notion of *attractivity*.

- The next two assumptions guarantee the existence of translation-invariant product-form stationary measures. (Similar assumptions were employed by Cocozza-Thivent [48].)

  – For any $x,\, y,\, z \in I$

$$\begin{aligned} &p(x,\, y) + p(y,\, z) + p(z,\, x) \\ &+ q(x,\, y) + q(y,\, z) + q(z,\, x) = p(x,\, z) + p(z,\, y) + p(y,\, x) \\ &\qquad\qquad + q(x,\, z) + q(z,\, y) + q(y,\, x). \end{aligned} \tag{3.2.8}$$

  – There are symmetric functions s_p and s_q on $I \times I$, and a function f on I such that $f(\omega^{\min}) = 0$ whenever $\omega^{\min}$ is finite, $f(z) > 0$ for $z > \omega^{\min}$, and for any $y,\, z \in I$,

$$\begin{aligned} p(y,\, z) &= s_p(y,\, z+1) f(y) \\ \text{and} \quad q(y,\, z) &= s_q(y+1,\, z) f(z). \end{aligned} \tag{3.2.9}$$

  (Interpret $s_p(y,\, z) = s_q(y,\, z) = 0$ if y or $z > \omega^{\max}$.) Condition (3.2.7) implies that f is nondecreasing on I.

An attempt at covering this broad class of processes raises the uncomfortable point that there is no unified existence proof for this entire class. Different constructions in the literature place various boundedness or growth conditions on p and q and the space I, and result in various degrees of regularity for the semigroup. (Among key references are Liggett's monograph [86], and articles [5], [19] and [85].) These existence matters are beyond the scope of this thesis. Yet we wish to give a general proof for fluctuations that in principal works for all processes in the family, subject to the more serious assumptions we explain in Section 3.2.6. To avoid extraneous technical issues we make the following blanket assumptions on the rates p and q to be considered.

- We assume that the increment process $\underline{\omega}(t)$, and the corresponding height process $\underline{h}(t)$ with normalization $h_0(0) = 0$, that obey Poisson rates p and q as described by (3.2.2) and (3.2.3), can be constructed with cadlag paths in a subspace $\widetilde{\Omega}$ of tempered increment configurations (i.e. configurations that obey some restrictive growth conditions).

- The subspace $\widetilde{\Omega}$ is of full measure under the invariant distributions μ^θ defined in Section 3.2.4.

- It is also possible to construct jointly several versions of the process with initial configurations from the space $\widetilde{\Omega}$ and with joint evolution obeying basic coupling (described in Section 3.2.3).

- Rates p and q have all moments under the invariant distributions μ^θ. In fact arguments like Lemma 3.B.2 of the Appendix provide this when f does not grow faster than exponential on $\mathbb{Z}^+$ and does not decrease faster to zero than exponential on $\mathbb{Z}^-$.

The reader will see that our proofs in Sections 3.3, 3.4, 3.5 and 3.6 do not make any analytic demands on the semigroup and its relation to the generator. We only use couplings, counting of particle currents and simple Poisson bounds.

Two identities from article [21] play a key role in this chapter, given as (3.2.19) and (3.2.20) in Section 3.2.5. These identities hold for all processes in the family under study. The proofs given in [21] use generator calculations which may not be justified for all these processes. However, these identities can also be proved by counting particles and taking limits of finite-volume processes ([22] contains an example). Such a proof should be available with any reasonable construction of a process. Hence we shall not hesitate to use the results of [21].

3.2.2 Examples

To give concrete meaning to the general formulation of the previous section we describe some basic examples. The type of state space I distinguishes three cases that we call generalized exclusion, misanthrope and bricklayers processes. In all cases there are two parameters $0 \le p, q \le 1$ such that $p + q = 1$. *Asymmetric processes* have $p \ne q$. These are the processes for which our results are relevant.

1. **Generalized exclusion processes.** These are the cases where both $\omega^{\min}$ and $\omega^{\max}$ are finite.

   - **The asymmetric simple exclusion process (ASEP)** introduced by F. Spitzer [105] is defined by $\omega^{\min} = 0$, $\omega^{\max} = 1$, $f(z) = \mathbf{1}\{z = 1\}$, $s_p(y, z) = p \cdot \mathbf{1}\{y = z = 1\}$ and $s_q(y, z) = q \cdot \mathbf{1}\{y = z = 1\}$. This produces the familiar rates

     $$p(y, z) = p \cdot \mathbf{1}\{y = 1, z = 0\} \quad \text{and} \quad q(y, z) = q \cdot \mathbf{1}\{y = 0, z = 1\}.$$

     Here $\omega_i \in \{0, 1\}$ is the occupation number for site i, $p(\omega_i, \omega_{i+1})$ is the rate for a particle to jump from site i to $i+1$, and $q(\omega_i, \omega_{i+1})$

is the rate for a particle to jump from site $i+1$ to i. These rates have values p and q, respectively, whenever there is a particle to perform the above jumps, and there is no particle on the terminal site of the jumps. Conditions (3.2.7) and (3.2.8) are also satisfied by these rates.

- **Particle-antiparticle exclusion process.** Let $\omega^{\min} = -1$, $\omega^{\max} = 1$. Take $f(-1) = 0$, $f(0) = c$ (*creation*), $f(1) = a$ (*annihilation*) where c and a are positive rates with $c \leq a/2$,

$$\begin{aligned} s_p(0,\,1) = s_p(1,\,0) = p, \quad s_p(0,\,0) = \frac{pa}{2c}, \quad s_p(1,\,1) = \frac{p}{2}, \\ s_q(0,\,1) = s_q(1,\,0) = q, \quad s_q(0,\,0) = \frac{qa}{2c}, \quad s_q(1,\,1) = \frac{q}{2} \end{aligned}$$

and $s_p,\ s_q$ zero in all other cases. These result in rates

$$\begin{aligned} p(0,\,0) = pc, \quad p(0,\,-1) = p(1,\,0) = \frac{pa}{2}, \quad p(1,\,-1) = pa, \\ q(0,\,0) = qc, \quad q(-1,\,0) = q(0,\,1) = \frac{qa}{2}, \quad q(-1,\,1) = qa \end{aligned}$$

and zero in all other cases. If ω_i is the number of particles at site i, with $\omega_i = -1$ meaning the presence of an antiparticle, then this model describes an asymmetric exclusion process of particles and antiparticles with annihilation and particle-antiparticle pair creation. These rates also satisfy our conditions.

One can imagine other generalizations with bounded numbers of particles and/or antiparticles per site.

2. **Generalized misanthrope processes** have $\omega^{\min} > -\infty$, $\omega^{\max} = \infty$.

   - **Zero range process.** Take $\omega^{\min} = 0$, $\omega^{\max} = \infty$, an arbitrary nondecreasing function $f : \mathbb{Z}^+ \to \mathbb{R}^+$ such that $f(0) = 0$,

$$\begin{aligned} s_p(y,\,z) \equiv p \qquad &\text{and} \qquad s_q(y,\,z) \equiv q, \\ p(y,\,z) = pf(y) \qquad &\text{and} \qquad q(y,\,z) = qf(z). \end{aligned}$$

   Again, ω_i represents the number of particles at site i. Depending on this number, a particle jumps from i to the right with rate $pf(\omega_i)$, and to the left with rate $qf(\omega_i)$. These rates trivially satisfy conditions (3.2.7) and (3.2.8).

3. **General deposition processes** have $\omega^{\min} = -\infty$ and $\omega^{\max} = \infty$. The height differences between adjacent columns can be arbitrary integers. Antiparticles are needed for a particle representation of the process.

- **Bricklayers process.** Let $f : \mathbb{Z} \to \mathbb{R}^+$ be non-decreasing and satisfy
$$f(z) \cdot f(1-z) = 1 \qquad \text{for all } z \in \mathbb{Z}.$$
The values of f for positive z's thus determine the values for non-positive z's. Let
$$s_p(y,\, z) = p + \frac{p}{f(y)f(z)} \qquad \text{and} \qquad s_q(y,\, z) = q + \frac{q}{f(y)f(z)},$$
which results in
$$p(y,\, z) = pf(y) + pf(-z) \qquad \text{and} \qquad q(y,\, z) = qf(-y) + qf(z).$$
The following picture motivates the name bricklayers process. At each site i stands a bricklayer who lays a brick on the column to his left at rate $pf(-\omega_i)$ and on the column to his right at rate $pf(\omega_i)$. Each bricklayer also removes a brick from his left at rate $qf(\omega_i)$ and from his right at rate $qf(-\omega_i)$. Conditions (3.2.7) and (3.2.8) hold for the rates.

These were examples for which our theorem holds, *provided the hypotheses on microscopic concavity to be described below can be verified.*

While this chapter has nothing to say about symmetric processes, let us point out that the general class defined in Section 3.2.1 contains also such processes. Symmetric processes are characterized by the identity $p(y,\, z) = q(z,\, y)$. In this case (3.2.8) holds automatically and we only need to take care of (3.2.7) and (3.2.9). Here is an interesting example.

- **The symmetric K-exclusion process** is obtained if we set $\omega^{\min} = 0$, $\omega^{\max} = K$, $f(z) = \mathbf{1}\{z > 0\}$,
$$s_p(y,\, z) = s_q(y,\, z) = \mathbf{1}\{z,\, y \le K\}.$$
These result in
$$p(y,\, z) = q(z,\, y) = \mathbf{1}\{y > 0,\ z < K\}.$$
Thus this process also has a family of product-form invariant distribution, as described below.

The interesting point about this example is that the asymmetric version does not have product-form invariant distributions, and indeed the existence of spatially ergodic invariant distributions for all density values $\varrho \in [0,\, K]$ has been an open problem for many years. And of course, current fluctuations for the asymmetric process are also open.

3.2.3 Basic coupling

In *basic coupling* the joint evolution of n processes $\underline{\omega}^m(\cdot),\ m = 1, \dots, n$, is defined in such a manner that the processes "jump together as much as possible." The joint rates are determined as follows, given the current configurations $\underline{\omega}^1, \underline{\omega}^2, \dots, \underline{\omega}^n \in \widetilde{\Omega}$. Consider a step of type (3.2.2) over the edge $(i, i+1)$. Let $m \mapsto \ell(m)$ be a permutation that orders the rates of the individual processes for this move:

$$r(m) \equiv p(\omega_i^{\ell(m)}, \omega_{i+1}^{\ell(m)}) \le p(\omega_i^{\ell(m+1)}, \omega_{i+1}^{\ell(m+1)}) \equiv r(m+1), \quad 1 \le m < n.$$

Set also the dummy value $r(0) = 0$. Now the rule is that independently for each $m = 1, \dots, n$, at rate $r(m) - r(m-1)$, precisely processes $\underline{\omega}^{\ell(m)}, \underline{\omega}^{\ell(m+1)}, \dots, \underline{\omega}^{\ell(n)}$ execute move (3.2.2), and processes $\underline{\omega}^{\ell(1)}, \underline{\omega}^{\ell(2)}, \dots, \underline{\omega}^{\ell(m-1)}$ do not. The combined effect of these joint rates creates the correct marginal rates, that is, process $\underline{\omega}^{\ell(m)}$ executes this move with rate $r(m)$.

Notice also that, due to (3.2.7), a jump of $\underline{\omega}^a$ without $\underline{\omega}^b$ can only occur if $p(\omega_i^b, \omega_{i+1}^b) < p(\omega_i^a, \omega_{i+1}^a)$ which implies $\omega_i^a > \omega_i^b$ or $\omega_{i+1}^a < \omega_{i+1}^b$. The result of this step (3.2.2) then cannot increase the number of discrepancies between the two processes, hence the name *attractivity* for (3.2.7). In particular, a sitewise ordering $\omega_i^a \le \omega_i^b \quad \forall i \in \mathbb{Z}$ is preserved by the basic coupling.

One can check that moves of type (3.2.3) with rates q obey the same attractivity property.

The differences between two processes are called *second class particles*. Their number is nonincreasing. In particular, if $\omega_i^a \ge \omega_i^b$ for each $i \in \mathbb{Z}$, then the second class particles are conserved. In view of (3.2.5), in this case the net number of second class particles that pass from left to right across the straight-line space-time path from $(1/2,\ 0)$ to $(i + 1/2,\ t)$ equals the growth difference

$$\big(h_i^a(t) - h_0^a(0)\big) - \big(h_i^b(t) - h_0^b(0)\big) = h_i^a(t) - h_i^b(t) \tag{3.2.10}$$

between the two processes $\underline{\omega}^a(\cdot)$ and $\underline{\omega}^b(\cdot)$.

A special case that is of key importance to us is the situation where only one second class particle is present between two processes.

3.2.4 Translation invariant stationary product distributions

The results of this chapter concern stationary processes with particular product-form marginal distributions that we define in this section. For many cases it has been proved that these measures are the only extremal translation-invariant stationary distributions. Following some ideas in Cocozza-Thivent [48], we first consider the nondecreasing function f whose existence was assumed in (3.2.9). For $I \ni z > 0$ define

$$f(z)! := \prod_{y=1}^{z} f(y),$$

while for $I \ni z < 0$ let

$$f(z)! := \frac{1}{\prod_{y=z+1}^{0} f(y)},$$

and then $f(0)! := 1$. This definition satisfies $f(z)! \cdot f(z+1) = f(z+1)!$ for all $z \in I$. Let

$$\bar{\theta} := \begin{cases} \log\left(\liminf_{z\to\infty} (f(z)!)^{1/z}\right) = \lim_{z\to\infty} \log(f(z)), & \text{if } \omega^{\max} = \infty \\ \infty, & \text{else} \end{cases}$$

and

$$\underline{\theta} := \begin{cases} \log\left(\limsup_{z\to\infty} (f(-z)!)^{1/z}\right) = \lim_{z\to\infty} \log(f(-z)), & \text{if } \omega^{\min} = -\infty \\ -\infty, & \text{else.} \end{cases}$$

By monotonicity of f, we have $\bar{\theta} \geq \underline{\theta}$. The case $\bar{\theta} = \underline{\theta}$ would imply that $\omega^{\min} = -\infty$, $\omega^{\max} = \infty$, and f is a constant. Notice that (3.2.7) and (3.2.9) imply that s_p is non-increasing in its variables, but p is non-decreasing in its first variable. Hence a constant f results in an s_p that does not depend on its first variable. But then by its symmetric property it does not depend on its second variable either, and we conclude that a constant f implies constant rates p (and, similarly, q). We exclude this uninteresting case by postulating

$$\text{Assume } f \text{ to be such that } \underline{\theta} < \bar{\theta}. \tag{3.2.11}$$

For $\theta \in \left(\underline{\theta}, \bar{\theta}\right)$ define the state sum

$$Z(\theta) := \sum_{z\in I} \frac{\mathrm{e}^{\theta z}}{f(z)!} < \infty. \tag{3.2.12}$$

Let the product-distribution $\underline{\mu}^{\theta}$ on $\Omega = I^{\mathbb{Z}}$ have marginals

$$\mu^{\theta}(z) = \underline{\mu}^{\theta}\{\underline{\omega} : \omega_i = z\} := \frac{1}{Z(\theta)} \cdot \frac{\mathrm{e}^{\theta z}}{f(z)!} \qquad (z \in I). \tag{3.2.13}$$

Assumptions (3.2.6), (3.2.7), (3.2.8), (3.2.9) imply that for $\theta \in \left(\underline{\theta}, \bar{\theta}\right)$ the product distribution $\underline{\mu}^{\theta}$ is stationary for the process generated by (3.2.4) (see [21]). Note that the family $\{\mu^{\theta}\}$ can be obtained by exponentially weighting a probability measure μ^{θ_0} for a fixed value $\theta_0 \in (\underline{\theta}, \bar{\theta})$, see [20].

$\mathbf{P}^{\theta}$, $\mathbf{E}^{\theta}$, $\mathbf{Var}^{\theta}$, $\mathbf{Cov}^{\theta}$ will refer to laws of a process evolving in this stationary distribution. In the Appendix we show that the *density*

$$\varrho(\theta) := \mathbf{E}^{\theta}(\omega)$$

is a strictly increasing, infinitely differentiable function of the parameter θ that maps the interval $(\underline{\theta}, \bar{\theta})$ onto the interval $(\omega^{\min}, \omega^{\max})$. (The following

point should cause no confusion: the single-site state space I consists of the integers between $\omega^{\min}$ and $\omega^{\max}$, including endpoints if finite, but for density values the interval $(\omega^{\min}, \omega^{\max})$ is an interval of real numbers.) For most cases we shall use the density ϱ, rather than θ, for parameterizing the stationary distributions. Accordingly, μ^ϱ, $\mathbf{P}^\varrho$, $\mathbf{E}^\varrho$, $\mathbf{Var}^\varrho$, $\mathbf{Cov}^\varrho$ will refer to laws of a density ϱ stationary process.

3.2.5 Hydrodynamics and some exact identities

The *hydrodynamic flux* is defined as

$$\mathcal{H}(\varrho) := \mathbf{E}^\varrho(p(\omega_0, \omega_1) - q(\omega_0, \omega_1)). \tag{3.2.14}$$

$\mathcal{H}(\varrho)$ is the expected net rate at which a given column grows, or at which particles pass any fixed lattice edge from left to right in a stationary density-ϱ process. We show smoothness of $\mathcal{H}$ in Section 3.B of the Appendix. It is expected, and in many instances proved, that asymmetric members of our class satisfy the conservation law

$$\partial_T \varrho(T, X) + \partial_X \mathcal{H}(\varrho(T, X)) = 0$$

in the Eulerian-scaled time and space variables T and X, see e.g. Rezakhanlou [101] or Bahadoran, Guiol, Ravishankar and Saada [7]. The *characteristic speed* is the velocity with which small perturbations of the solution of this PDE propagate, and is given by

$$V^\varrho := \mathcal{H}'(\varrho). \tag{3.2.15}$$

A particular expectation we shall need several times is

$$\mathbf{E}^\varrho(h_i(t)) = \mathcal{H}(\varrho)t - \varrho i, \quad t \geq 0,\ i \in \mathbb{Z}. \tag{3.2.16}$$

For $i = 0$ this follows from (3.2.5), and in general from the $i = 0$ case together with $\omega_j(t) = h_{j-1}(t) - h_j(t)$.

When a stationary process is perturbed by adding a second class particle at the origin at time zero, we obtain two processes, $\underline{\omega}^-(\cdot)$ and $\underline{\omega}(\cdot)$. It is not a priori clear what the initial joint distribution of the occupation variables $\omega_0^-(0)$, $\omega_0(0)$ should be. For ASEP there is no ambiguity due to the simplicity of the single-site state space: the only way to have a discrepancy is to set $\omega_0^-(0) = 0$, $\omega_0(0) = 1$. A useful generalization of this distribution to the broader class of processes involves the following family of probability measures on I introduced in [21]:

$$\widehat{\mu}^\varrho(y) := \frac{1}{\mathbf{Var}^\varrho(\omega_0)} \sum_{z=y+1}^{\omega^{\max}} (z - \varrho)\mu^\varrho(z), \qquad y \in I. \tag{3.2.17}$$

An empty sum is zero by convention and so if $\omega^{\max} < \infty$, $\widehat{\mu}^{\varrho}(\omega^{\max}) = 0$. Consequently there is room for an additional particle under the $\widehat{\mu}^{\varrho}$ distribution, in the sense that if $\omega \sim \widehat{\mu}^{\varrho}$ then also $\omega + 1 \in I$.

To our knowledge these distributions $\widehat{\mu}^{\varrho}$ do not possess any invariance properties. Their virtue is that they make identities (3.2.19) and (3.2.20) below true. We show in Section 3.A of the Appendix that both μ^{ϱ} and $\widehat{\mu}^{\varrho}$ are stochastically monotone in the density ϱ. (There is, however, no stochastic domination between μ^{ϱ} and $\widehat{\mu}^{\varrho}$ in general.)

Denote by $\mathbf{E}$ the expectation w.r.t. the evolution of a pair $(\underline{\omega}^{-}(\cdot), \underline{\omega}(\cdot))$ started with initial data (recall (3.2.1))

$$\underline{\omega}^{-}(0) = \underline{\omega}(0) - \underline{\delta}_0 \sim \Big(\bigotimes_{i \neq 0} \mu^{\varrho}\Big) \otimes \widehat{\mu}^{\varrho}, \tag{3.2.18}$$

and evolving under the basic coupling. This pair will always have a single second class particle whose position is denoted by $Q(t)$. In other words, $\underline{\omega}^{-}(t) = \underline{\omega}(t) - \underline{\delta}_{Q(t)}$. Corollaries 2.4 and 2.5 of [21] state that

$$\mathbf{Var}^{\varrho}(h_i(t)) = \mathbf{Var}^{\varrho}(\omega) \cdot \mathbf{E}|Q(t) - i| \tag{3.2.19}$$

and

$$\mathbf{E}(Q(t)) = V^{\varrho} \cdot t \tag{3.2.20}$$

for any $i \in \mathbb{Z}$ and $t \geq 0$. Note in particular that in (3.2.19) the variances are taken in a stationary process, while the expectation of $Q(t)$ is taken in the coupling with initial distribution (3.2.18).

3.2.6 Microscopic concavity

From now on fix the jump rates $p, q : I \times I \to \mathbb{R}^{+}$ that define the process in question, assumed to satisfy all the assumptions discussed thus far. The $t^{1/3}$ current or height fluctuations are expected when the hydrodynamic flux $\mathcal{H}(\varrho)$ is strictly concave or convex. In this section we discuss only the concave case. Concavity implies that the characteristic speed $V^{\varrho} = \mathcal{H}'(\varrho)$ is a nonincreasing function of density ϱ:

$$\lambda < \varrho \implies V^{\lambda} \geq V^{\varrho}. \tag{3.2.21}$$

The microscopic counterpart of a characteristic is the motion of a second class particle. Our key assumption that we term *microscopic concavity* is that the ordering (3.2.21) can also be realized at the particle level as an ordering between two second class particles introduced into two processes at densities λ and ϱ. Since this is now a probabilistic notion, there are several possible formulations, ranging from almost sure ($Q^{\lambda}(t) \geq Q^{\varrho}(t)$ in a

coupling) to distributional formulations. Assumption 3.2.1 below gives the precise technical form in which this section utilizes this notion of microscopic concavity. It stipulates that the ordering of second class particles is achieved by processes that evolve on the labels of auxiliary second class particles, and also requires some control of the tails of these random labels.

We do not imagine that this precise formulation will be the right one for all processes. We take it as a starting point and future work may lead to alternative formulations. Assumption 3.2.1 has the virtue that its requirements can be verified for some interesting processes.

Let $\lambda < \varrho$ be two densities. Proposition 3.A.4 in the Appendix gives the stochastic domination $\widehat{\mu}^{\lambda} \le \widehat{\mu}^{\varrho}$. Define $\widehat{\mu}^{\varrho} + 1$ as the measure that gives weight $\widehat{\mu}^{\varrho}(z-1)$ to an integer z such that $\omega^{\min} < z < \omega^{\max} + 1$. Let $\widehat{\mu}^{\lambda,\varrho}$ be a coupling measure with marginals $\widehat{\mu}^{\lambda}$ and $\widehat{\mu}^{\varrho} + 1$ and with the property

$$\widehat{\mu}^{\lambda,\varrho}\{(y,\, z) \,:\, \omega^{\min} - 1 < y < z < \omega^{\max} + 1\} = 1. \tag{3.2.22}$$

Let also $\mu^{\lambda,\varrho}$ be a coupling measure of site-marginals μ^{λ} and μ^{ϱ} of the invariant distributions, with

$$\mu^{\lambda,\varrho}\{(y,\, z) \,:\, \omega^{\min} - 1 < y \le z < \omega^{\max} + 1\} = 1. \tag{3.2.23}$$

Note the distinction that under $\widehat{\mu}^{\lambda,\varrho}$ the second coordinate is strictly above the first.

To have notation for inhomogeneous product measures on $I^{\mathbb{Z}}$, let $\underline{\lambda} = (\lambda_i)_{i\in\mathbb{Z}}$ and $\underline{\varrho} = (\varrho_i)_{i\in\mathbb{Z}}$ denote sequences of density values, with λ_i and ϱ_i assigned to site i. The product distribution with marginals $\widehat{\mu}^{\lambda_0,\varrho_0}$ at the origin and $\mu^{\lambda_i,\varrho_i}$ at other sites is denoted by

$$\underline{\widehat{\mu}}^{\underline{\lambda},\underline{\varrho}} := \Big(\bigotimes_{i\neq 0} \mu^{\lambda_i,\varrho_i}\Big) \otimes \widehat{\mu}^{\lambda_0,\varrho_0}. \tag{3.2.24}$$

Measure $\underline{\widehat{\mu}}^{\underline{\lambda},\underline{\varrho}}$ gives probability one to the event

$$\{(\underline{\eta}(0),\, \underline{\omega}(0)) \,:\, \eta_0(0) < \omega_0(0), \text{ and } \eta_i(0) \le \omega_i(0) \text{ for } 0 \neq i \in \mathbb{Z}\}.$$

The initial configuration $(\underline{\eta}(0),\, \underline{\omega}(0))$ will always be assumed a member of this set, and the pair process $(\underline{\eta}(t),\, \underline{\omega}(t))$ evolves in basic coupling. In general $\underline{\widehat{\mu}}^{\underline{\lambda},\underline{\varrho}}$ is *not* stationary for this joint evolution.

The discrepancies between these two processes are called the $\omega - \eta$ *(second class) particles*. The number of such particles at site i at time t is $\omega_i(t) - \eta_i(t)$. In the basic coupling the $\omega - \eta$ particles are conserved, in the sense that none are created or annihilated. We label the $\omega - \eta$ particles with integers, and let $X_m(t)$ denote the position of particle m at time t. The initial labeling is chosen to satisfy

$$\cdots \le X_{-1}(0) \le X_0(0) = 0 < X_1(0) \le \cdots .$$

We can specify that $X_0(0) = 0$ because under $\widehat{\underline{\mu}}^{\underline{\lambda},\underline{\varrho}}$ there is an $\omega - \eta$ particle at site 0 with probability 1. During the evolution we keep the positions $X_i(t)$ of the $\omega - \eta$ particles ordered. To achieve this we stipulate that

$$\begin{aligned} &\text{whenever an } \omega - \eta \text{ particle jumps from a site,} \\ &\text{if the jump is to the right the highest label moves,} \\ &\text{and if the jump is to the left the lowest label moves.} \end{aligned} \tag{3.2.25}$$

Here is the precise form of microscopic concavity. The assumption states that a certain joint construction of processes (that is, a coupling) can be performed for a range of densities in a neighborhood of a fixed density ϱ. Recall (3.2.1) for the definition of the configuration $\underline{\delta}$.

Assumption 3.2.1. *Given a density $\varrho \in (\omega^{min}, \omega^{max})$, there exists $\gamma_0 > 0$ such that the following holds. For any $\underline{\lambda}$ and $\underline{\varrho}$ such that $\varrho - \gamma_0 \le \lambda_i \le \varrho_i \le \varrho + \gamma_0$ for all $i \in \mathbb{Z}$, a joint process $(\underline{\eta}(t), \underline{\omega}(t), y(t), z(t))_{t\ge 0}$ can be constructed with the following properties.*

- *Initially $(\underline{\eta}(0), \underline{\omega}(0))$ is $\widehat{\underline{\mu}}^{\underline{\lambda},\underline{\varrho}}$-distributed and the joint process $(\underline{\eta}(\cdot), \underline{\omega}(\cdot))$ evolves in basic coupling.*
- *Processes $y(\cdot)$ and $z(\cdot)$ are integer-valued. Initially $y(0) = z(0) = 0$. With probability one*

$$y(t) \le z(t) \text{ for all } t \ge 0. \tag{3.2.26}$$

- *Define the processes*

$$\underline{\omega}^-(t) := \underline{\omega}(t) - \underline{\delta}_{X_{y(t)}(t)} \quad \text{and} \quad \underline{\eta}^+(t) := \underline{\eta}(t) + \underline{\delta}_{X_{z(t)}(t)}. \tag{3.2.27}$$

  Then both pairs $(\underline{\eta}, \underline{\eta}^+)$ and $(\underline{\omega}^-, \underline{\omega})$ evolve marginally in basic coupling.

- *For each $\gamma \in (0, \gamma_0)$ and large enough $t \ge 0$ there exists a probability distribution $\nu^{\varrho,\gamma}(t)$ on $\mathbb{Z}^+$ satisfying the tail bound*

$$\nu^{\alpha,\gamma}(t)\{y : y \ge y_0\} \le Ct^{\kappa-1}\gamma^{2\kappa-3}y_0^{-\kappa} \tag{3.2.28}$$

  for some fixed constants $3/2 \le \kappa < 3$ and $C < \infty$, and such that if $\varrho - \gamma \le \lambda_i \le \varrho_i \le \varrho + \gamma$ for all $i \in \mathbb{Z}$, then we have the stochastic bounds

$$y(t) \overset{d}{\le} \nu^{\varrho,\gamma}(t) \quad \text{and} \quad z(t) \overset{d}{\ge} -\nu^{\varrho,\gamma}(t). \tag{3.2.29}$$

Let us clarify some of the details in this assumption.

Equation (3.2.27) says that $Q^\eta(t) := X_{z(t)}(t)$ is the single second class particle between $\underline{\eta}$ and $\underline{\eta}^+$, while $Q(t) := X_{y(t)}(t)$ is the one between $\underline{\omega}^-$ and $\underline{\omega}$. The first three bullets say that it is possible to construct jointly four

processes $(\underline{\eta}, \underline{\eta}^+, \underline{\omega}^-, \underline{\omega})$ with the specified initial conditions and so that each pair $(\underline{\eta}, \underline{\omega})$, $(\underline{\eta}, \underline{\eta}^+)$ and $(\underline{\omega}^-, \underline{\omega})$ has the desired marginal distribution, and most importantly so that

$$Q^\eta(t) = X_{z(t)}(t) \geq X_{y(t)}(t) = Q(t). \tag{3.2.30}$$

This is a consequence of (3.2.26) because the $\omega - \eta$ particles $X_i(t)$ stay ordered.

The tail bound (3.2.28) is formulated in this somewhat complicated fashion because this appears to be the weakest form our present proof allows. In our currently available examples $\nu^{\varrho,\gamma}(t)$ is actually a fixed geometric distribution. However, we expect that other examples will require more complicated bounds and so including this generality is sensible.

The assumptions made imply $\underline{\eta}(t) \leq \underline{\omega}(t)$ a.s., and by (3.2.27)

$$\underline{\eta}(t) \leq \underline{\eta}^+(t) \leq \underline{\omega}(t) \quad \text{and} \quad \underline{\eta}(t) \leq \underline{\omega}^-(t) \leq \underline{\omega}(t) \quad \text{a.s.}$$

In our actual constructions of the processes $\underline{\eta}$, $\underline{\eta}^+$, $\underline{\omega}^-$, $\underline{\omega}$ for ASEP (Section 3.2.8 and [22]), for a class of totally asymmetric zero range processes (Section 3.6) and for the totally asymmetric bricklayers process with exponential rates (future work) it turns out that the triples $(\underline{\eta}, \underline{\eta}^+, \underline{\omega})$ and $(\underline{\eta}, \underline{\omega}^-, \underline{\omega})$ evolve also in basic coupling, but the full joint evolution $(\underline{\eta}, \underline{\eta}^+, \underline{\omega}^-, \underline{\omega})$ does not.

As already explained, the microscopic concavity idea is contained in inequality (3.2.26). There is also a sense in which the tail bounds (3.2.29) relate to concavity of the flux. Consider the situation $\lambda_i \equiv \lambda < \varrho \equiv \varrho_i$. We would expect the $\omega - \eta$ particle $X_0(\cdot)$ to have average and long-term velocity

$$R(\lambda, \varrho) = \frac{\mathcal{H}(\varrho) - \mathcal{H}(\lambda)}{\varrho - \lambda},$$

the Rankine-Hugoniot or shock speed. By concavity

$$\mathcal{H}'(\varrho) = V^\varrho \leq R(\lambda, \varrho) \leq V^\lambda = \mathcal{H}'(\lambda).$$

A strict microscopic counterpart would be $y(t) \leq 0 \leq z(t)$. But this condition is overly restrictive. The only cases we know to satisfy it are the totally asymmetric simple exclusion process and the totally asymmetric zero range process with constant rate. The distributional bounds (3.2.29) are natural relaxations of $y(t) \leq 0 \leq z(t)$.

By the same token, perhaps the way to covering more examples with our approach involves a similar distributional weakening of (3.2.26), but this seems less straightforward.

3.2.7 Results

We need a few more assumptions and then we can state the main result. Constants $C_\cdot$, $\alpha_\cdot$ will not depend on time, but might depend on the density parameter ϱ, and their values can change from line to line. We are now working with a fixed member of the class of processes described in Section 3.2.1 with rate functions $p, q : I \times I \to \mathbb{R}^+$. Recall that $\mathcal{H}$ is the hydrodynamic flux defined in (3.2.14). In the Appendix we show $\mathcal{H}$ is infinitely differentiable under the restrictions on the rates placed in Section 3.2.1.

Assumption 3.2.2. *The rates p, q and density $\varrho \in (\omega^{min}, \omega^{max})$ have the following properties.*

- *The jump rate functions p and q satisfy assumptions* (3.2.6), (3.2.7), (3.2.8), (3.2.9) *and* (3.2.11) *discussed in Sections 3.2.1 and 3.2.4.*
- $\mathcal{H}''(\varrho) < 0$.
- *Let $(\underline{\omega}^-, \underline{\omega})$ be a pair of processes in basic coupling, started from distribution* (3.2.18), *with second class particle $Q(t)$. Then there exist constants $0 < \alpha_0$, $C_0 < \infty$ such that*

$$\mathbf{P}\{|Q(t)| > K\} \le C_0 \cdot \frac{t^2}{K^3} \tag{3.2.31}$$

whenever $K > \alpha_0 t$ and t is large enough.

As mentioned, our results are valid only for asymmetric processes. The assumption of asymmetry is implicitly contained in $\mathcal{H}''(\varrho) < 0$. Symmetric processes have $\mathcal{H}(\varrho) \equiv 0$. Exponential tail bounds for $|Q(t)|$ that imply assumption (3.2.31) hold automatically if the rates p, q have bounded increments because the rates for Q come from these increments of p and q. Here is the main result.

Theorem 3.2.3. *Let Assumptions 3.2.1 and 3.2.2 hold for density ϱ. Let the processes $(\underline{\omega}^-(t), \underline{\omega}(t))$ evolve in basic coupling with initial distribution* (3.2.18) *and let $Q(t)$ be the position of the second class particle between $\underline{\omega}^-(t)$ and $\underline{\omega}(t)$. Then there is a constant $C_1 = C_1(\varrho) \in (0, \infty)$ such that for all $1 \le m < 3$,*

$$\frac{1}{C_1} < \liminf_{t\to\infty} \frac{\mathbf{E}|Q(t) - V^\varrho t|^m}{t^{2m/3}} \le \limsup_{t\to\infty} \frac{\mathbf{E}|Q(t) - V^\varrho t|^m}{t^{2m/3}} < \frac{C_1}{3-m}. \tag{3.2.32}$$

Superdiffusivity of the second class particle is best seen with the choice $m = 2$: the variance of its position is of order $t^{4/3}$. Next some corollaries. Notation $\lfloor X \rfloor$ stands for the lower integer part of X.

Corollary 3.2.4 (Current variance)**.** *Under Assumptions 3.2.1 and 3.2.2, there is a constant $C_1 = C_1(\varrho) > 0$, such that*

$$\frac{1}{C_1} < \liminf_{t\to\infty} \frac{\boldsymbol{Var}^{\varrho}(h_{\lfloor V^{\varrho}t\rfloor}(t))}{t^{2/3}} \le \limsup_{t\to\infty} \frac{\boldsymbol{Var}^{\varrho}(h_{\lfloor V^{\varrho}t\rfloor}(t))}{t^{2/3}} < C_1.$$

This follows from (3.2.19) with the choice $m = 1$.

Corollary 3.2.5 (Law of Large Numbers for the second class particle)**.** *Under Assumptions 3.2.1 and 3.2.2, the Weak Law of Large Numbers holds in a density-ϱ stationary process:*

$$\frac{Q(t)}{t} \xrightarrow{d} V^{\varrho}. \tag{3.2.33}$$

If the rates p and q have bounded increments, then almost sure convergence also holds in (3.2.33) *(Strong Law of Large Numbers).*

The Weak Law is a simple consequence of Theorem 3.2.3. The Strong Law will be proved in Section 3.5.

Corollary 3.2.6 (Dependence of current on the initial configuration)**.** *Under Assumptions 3.2.1 and 3.2.2, for any $V \in \mathbb{R}$ and $\alpha > 1/3$ the following limit holds in the L^2 sense for a density-ϱ stationary process:*

$$\lim_{t\to\infty} \frac{h_{\lfloor Vt\rfloor}(t) - h_{\lfloor Vt\rfloor - \lfloor V^{\varrho}t\rfloor}(0) - t(\mathcal{H}(\varrho) - \varrho\mathcal{H}'(\varrho))}{t^{\alpha}} = 0. \tag{3.2.34}$$

Recall that

$$h_{\lfloor Vt\rfloor - \lfloor V^{\varrho}t\rfloor}(0) = \begin{cases} \displaystyle\sum_{i=\lfloor Vt\rfloor - \lfloor V^{\varrho}t\rfloor + 1}^{0} \omega_i(0), & \text{if } V < V^{\varrho}, \\ 0, & \text{if } V = V^{\varrho}, \\ -\displaystyle\sum_{i=1}^{\lfloor Vt\rfloor - \lfloor V^{\varrho}t\rfloor} \omega_i(0), & \text{if } V > V^{\varrho} \end{cases} \tag{3.2.35}$$

only depends on a finite segment of the initial configuration. Limit (3.2.34) shows that on the diffusive time scale $t^{1/2}$ only fluctuations from the initial distribution are visible: these fluctuations are translated rigidly at the characteristic speed V^{ϱ}. Proof of (3.2.34) follows by translating $h_{\lfloor Vt\rfloor}(t) - h_{\lfloor Vt\rfloor - \lfloor V^{\varrho}t\rfloor}(0)$ to $h_{\lfloor V^{\varrho}t\rfloor}(t) - h_0(0) = h_{\lfloor V^{\varrho}t\rfloor}(t)$ and by applying Corollary 3.2.4. From (3.2.34), (3.2.35) and the i.i.d. initial $\{\omega_i\}$ follow a limit for the variance and a central limit theorem (CLT), which we record in our final corollary.

Corollary 3.2.7 (Central Limit Theorem for the current)**.** *Under Assumptions 3.2.1 and 3.2.2, for any $V \in \mathbb{R}$ in a density-ϱ stationary process*

$$\lim_{t\to\infty} \frac{\boldsymbol{Var}^{\varrho}(h_{\lfloor Vt\rfloor}(t))}{t} = \boldsymbol{Var}^{\varrho}(\omega)\cdot|V^{\varrho}-V| =: D, \tag{3.2.36}$$

and the Central Limit Theorem also holds: the centered and normalized height $\widetilde{h}_{\lfloor Vt\rfloor}(t)/\sqrt{t\cdot D}$ converges in distribution to a standard normal.

For ASEP the CLT, the limiting variance (3.2.36) and the appearance of initial fluctuations on the diffusive scale were proved by P. A. Ferrari and L. R. G. Fontes [65]. For convex rate zero range and bricklayers processes Corollary 3.2.7 was proved by M. Balázs [11].

Remark on the convex case

Our results and proofs work in the analogous way in the case where the flux is convex and the corresponding *microscopic convexity* is assumed. This case is carried out in more detail in Section 3.7

3.2.8 Two examples that satisfy microscopic concavity

Presently we have verified all the hypotheses of Theorem 3.2.3 for two classes of processes.

The asymmetric simple exclusion process

The asymmetric simple exclusion process (ASEP) was the first example described in Section 3.2.2. It has two parameters $0 \le p \ne q \le 1$ such that $p+q=1$. To be specific let us take $p > q$ so that on average particles prefer to drift to the right. The invariant measure μ^{ϱ} is the Bernoulli distribution with parameter $0 \le \varrho \le 1$, while $\widehat{\mu}^{\varrho}$ is concentrated on zero for any ϱ. The hydrodynamic flux is strictly concave: $\mathcal{H}(\varrho) = (p-q)\varrho(1-\varrho)$.

The detailed construction of the processes $y(t)$ and $z(t)$ needed for Assumption 3.2.1 can be found in [22]. Here it is in a nutshell.

Given the background process $(\underline{\eta}(\cdot), \underline{\omega}(\cdot))$ and the second class particles $\{X_m(\cdot)\}$ between them, the processes $y(\cdot)$ and $z(\cdot)$ are nearest-neighbor random walks on the labels $\{m\}$ with rates p and q. Walk $y(\cdot)$ has bias to the left (rate p to the left, rate q to the right) and walk $z(\cdot)$ has bias to the right (rate p to the right, rate q to the left). Their jumps are restricted so that jumps between labels m and $m+1$ are permitted only when X_m and X_{m+1} are adjacent. The clocks governing these jumps are coupled so that the ordering $y \le z$ is preserved.

Since a second class particle in ASEP is bounded by a rate one Poisson process, (3.2.31) holds.

Balázs and Seppäläinen gave an earlier proof of Theorem 3.2.3 for ASEP in [23]. The present general proof evolved from that earlier one.

Totally asymmetric zero range process with jump rates that increase with exponentially decaying slope

As explained in Section 3.2.2, in a totally asymmetric zero range process (TAZRP) one particle is moved from site i to site $i+1$ at rate $f(\omega_i)$, and no particle jumps to the left (our convention for total asymmetry is $p = 1-q = 1$). The jump rate $f : \mathbb{Z}^+ \to \mathbb{R}^+$ is nondecreasing, $f(0) = 0$, and $f(z) > 0$ for $z > 0$. Assume further that f is concave.

As we shall see later in Section 3.6, one aspect of microscopic concavity, namely the ordering of second class particles, can be achieved for any TAZRP with a nondecreasing concave jump rate. Indeed, up to Lemma 3.6.2 in Section 3.6 we only use monotonicity and concavity of the rates f. Thus for concave TAZRP only the tail control (3.2.28)–(3.2.29) of the label processes remains to be provided. For this part we currently need a stronger hypothesis, detailed in the next assumption.

Assumption 3.2.8. *Let $p = 1 - q = 1$. Assume the jump rate function f of a totally asymmetric zero range process has these properties:*

- $f(0) = 0 < f(1)$,
- *f is nondecreasing:* $f(z+1) \geq f(z)$,
- *f is concave with an exponentially decreasing slope: there is an $0 < r < 1$ such that for each $z \geq 1$ such that $f(z) - f(z-1) > 0$,*

$$\frac{f(z+1) - f(z)}{f(z) - f(z-1)} \leq r. \tag{3.2.37}$$

The case where f becomes constant above some z_0 is included.

Theorem 3.2.9. *Under Assumption 3.2.8, a stationary totally asymmetric zero range process satisfies the conclusions of Theorem 3.2.3, and the conclusions of Corollaries 3.2.4, 3.2.5, 3.2.6 and 3.2.7.*

A class of examples of rates that satisfy Assumption 3.2.8 are

$$f(z) = 1 - \exp(-\beta z^{\vartheta}), \quad \beta > 0,\ \vartheta \geq 1.$$

Another example is the most basic, constant rate TAZRP with $f(z) = \mathbf{1}\{z > 0\}$. For this last case a proof has already been given in [15].

To prove Theorem 3.2.9 we need to check Assumptions 3.2.1 and 3.2.2 of Theorem 3.2.3. The construction of the label processes $y(t)$ and $z(t)$ and verification of Assumption 3.2.1 are done in Section 3.6. Assumption 3.2.2 requires only a few comments. The properties of the rates required in the first bullet of Assumption 3.2.2 are straightforward. Since f is concave and cannot be linear due to (3.2.37), Proposition 3.B.1 in the Appendix implies that $\mathcal{H}''(\varrho) < 0$ for each $\varrho > 0$. Concavity of f implies bounded jump

rates for the second class particle $Q(t)$, hence a simple Poisson bound gives (3.2.31).

The next two sections prove Theorem 3.2.3, after that we prove the Strong Law for the second class particle, and then we return to finish the proof of Theorem 3.2.9.

3.3 Upper bound of the main theorem

In this section we prove the upper bound of (3.2.32). Density ϱ is fixed. Let $\lambda \in (\rho,\, \rho - \gamma_0)$ and apply Assumption 3.2.1 with constant sequences $\varrho_i \equiv \varrho$ and $\lambda_i \equiv \lambda$ for all $i \in \mathbb{Z}$. Notations $\mathbf{P}$, $\mathbf{E}$, $\mathbf{Var}$, $\mathbf{Cov}$ will refer to the coupled four-process evolution described in Assumption 3.2.1, while $\mathbf{P}^\varrho$, $\mathbf{E}^\varrho$, $\mathbf{Var}^\varrho$, $\mathbf{Cov}^\varrho$ will refer to a density ϱ stationary process. Abbreviate

$$\Psi(t) := \mathbf{E}|Q(t) - \lfloor V^\varrho t \rfloor|. \tag{3.3.1}$$

The requirement that $(\underline{\omega}^-, \underline{\omega})$ obey the basic coupling was included in Assumption 3.2.1. Consequently $\Psi(t)$ is the $m = 1$ expectation of (3.2.32).

The following lemma does the main work towards the upper bound.

Lemma 3.3.1. *There exist positive constants α_1, α_2, t_0 such that for each $t > t_0$ and integer u such that $\alpha_2\sqrt{t} < u < \alpha_1 t$,*

$$\mathbf{P}\{Q(t) > \lfloor V^\varrho t \rfloor + u\} \le C_5 \frac{t^2 \mathcal{H}''(\varrho)^2}{u^4} \big\{\Psi(t) + u\big\} + C_4 \frac{t^2}{u^3}. \tag{3.3.2}$$

Proof. We start with an integer $u > 0$, and write

$$\mathbf{P}\{Q(t) > \lfloor V^\varrho t \rfloor + u\} \le \mathbf{P}\{y(t) \ge k\} + \mathbf{P}\{X_k(t) \ge Q(t) > \lfloor V^\varrho t \rfloor + u\}. \tag{3.3.3}$$

The event $\{X_k(t) > \lfloor V^\varrho t \rfloor + u\}$ implies that among the X_m's at most particles $X_1, \dots, X_{k-1}$ have passed the path $\big(s(\lfloor V^\varrho t \rfloor + u) + 1/2\big)_{0 \le s \le 1}$ from right to left. Each such passing decreases $h^\omega_{\lfloor V^\varrho t \rfloor + u}(t) - h^\eta_{\lfloor V^\varrho t \rfloor + u}(t)$ by one (recall the statement around (3.2.10)). Hence we can bound the probability in (3.3.3) by

$$\mathbf{P}\{y(t) \ge k\} + \mathbf{P}\{h^\omega_{\lfloor V^\varrho t \rfloor + u}(t) - h^\eta_{\lfloor V^\varrho t \rfloor + u}(t) > -k\}.$$

We introduce two more processes: $\underline{\eta}^{\mathrm{eq}}$ is a stationary process started with initial data $\eta^{\mathrm{eq}}_i(0) = \eta_i(0)$ for $i \neq 0$, while $\eta^{\mathrm{eq}}_0(0)$ is μ^λ distributed independently of everything. $\underline{\omega}^{\mathrm{eq}}$ is a stationary process started with $\omega^{\mathrm{eq}}_i(0) = \omega_i(0)$

for $i \neq 0$, and $\omega_0^{\mathrm{eq}}(0)$ is μ^ϱ distributed independently of everything. Include these in the basic coupling of $(\underline{\eta}, \underline{\omega})$ and write

$$\begin{aligned} h^{\omega}_{\lfloor V^\varrho t\rfloor+u}(t) - h^{\eta}_{\lfloor V^\varrho t\rfloor+u}(t) &= h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) - h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) \\ &\quad + h^{\omega}_{\lfloor V^\varrho t\rfloor+u}(t) - h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) \\ &\quad - h^{\eta}_{\lfloor V^\varrho t\rfloor+u}(t) + h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t). \end{aligned}$$

Basic coupling implies

$$\begin{aligned} & h^{\omega}_{\lfloor V^\varrho t\rfloor+u}(t) - h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) \le |\omega_0(0) - \omega_0^{\mathrm{eq}}(0)| \le |\omega_0(0)| + |\omega_0^{\mathrm{eq}}(0)| \\ \text{and} \quad & h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) - h^{\eta}_{\lfloor V^\varrho t\rfloor+u}(t) \le |\eta_0^{\mathrm{eq}}(0) - \eta_0(0)| \le |\eta_0^{\mathrm{eq}}(0)| + |\eta_0(0)|. \end{aligned}$$

We bound the stationary expectations using (3.2.16), (3.2.15) and Taylor's formula:

$$\begin{aligned} & \mathbf{E}^\varrho h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) - \mathbf{E}^\lambda h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) \\ &= \mathcal{H}(\varrho)t - (\lfloor V^\varrho t\rfloor + u)\varrho - \mathcal{H}(\lambda)t + (\lfloor V^\varrho t\rfloor + u)\lambda \\ &\le t\big(\mathcal{H}(\varrho) - \mathcal{H}(\lambda) + \mathcal{H}'(\varrho)(\lambda-\varrho)\big) + u(\lambda-\varrho) + C_1 \\ &\le -\frac{t}{2}\mathcal{H}''(\varrho)(\varrho-\lambda)^2 + u(\lambda-\varrho) + C_2 t(\varrho-\lambda)^3 + C_1. \end{aligned}$$

$\mathcal{H}$ can be differentiated arbitrarily many times, as we show in Section 3.B of the Appendix. Constant C_1 above bounds errors from discarded integer parts. Recall that tilde stands for the centered random variable. Collecting terms we continue from (3.3.3) as follows.

$$\begin{aligned} & \mathbf{P}\{Q(t) > \lfloor V^\varrho t\rfloor + u\} \\ &\quad \le \mathbf{P}\{y(t) \ge k\} \\ &\qquad + \mathbf{P}\{\widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) - \widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) > -k + \frac{t}{2}\mathcal{H}''(\varrho)(\varrho-\lambda)^2 + u(\varrho-\lambda) \\ &\qquad\qquad - C_2 t(\varrho-\lambda)^3 - C_1 - |\eta_0(0)| - |\eta_0^{\mathrm{eq}}(0)| - |\omega_0(0)| - |\omega_0^{\mathrm{eq}}(0)|\} \\ &\quad \le \mathbf{P}\{y(t) \ge k\} \\ &\qquad + \mathbf{P}\{\widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) - \widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor+u}(t) > \frac{t}{2}\mathcal{H}''(\varrho)(\varrho-\lambda)^2 + \frac{u}{2}(\varrho-\lambda)\} \\ &\qquad + \mathbf{P}\{|\eta_0(0)| + |\eta_0^{\mathrm{eq}}(0)| + |\omega_0(0)| + |\omega_0^{\mathrm{eq}}(0)| \\ &\qquad\qquad > -k + \frac{u}{2}(\varrho-\lambda) - C_2 t(\varrho-\lambda)^3 - C_1\}. \end{aligned}$$

From now on we use the specific assumption $\mathcal{H}''(\varrho) < 0$. We maximize the terms on the right-hand side of the probability of h's by the choice

$$\varrho - \lambda = \frac{-u}{2t\mathcal{H}''(\varrho)}.$$

To stay within the range of densities covered by Assumption 3.2.1 we must ensure that $\lambda > \varrho - \gamma_0$. So we introduce a small constant $\alpha_1 > 0$ and restrict our calculations to the case $u < \alpha_1 t$. Then

$$\begin{aligned}\mathbf{P}\{Q(t) > \lfloor V^\varrho t\rfloor + u\} &\le \mathbf{P}\{y(t) \ge k\} \\ &\quad + \mathbf{P}\{\widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t) - \widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t) > \frac{-u^2}{8t\mathcal{H}''(\varrho)}\} \\ &\quad + \mathbf{P}\{|\eta_0(0)| + |\eta_0^{\mathrm{eq}}(0)| + |\omega_0(0)| + |\omega_0^{\mathrm{eq}}(0)| \\ &\qquad > -k - \frac{1}{4\mathcal{H}''(\varrho)} \cdot \frac{u^2}{t} + \frac{C_2}{\mathcal{H}''(\varrho)^3} \cdot \frac{u^3}{t^2} - C_1\}.\end{aligned}$$

Now we set

$$k = \lfloor \frac{-1}{8\mathcal{H}''(\varrho)} \cdot \frac{u^2}{t} \rfloor,$$

and assume $\alpha_2\sqrt{t} < u < \alpha_1 t$ for a possibly smaller α_1 and a large enough α_2. That allows us to unify the right-hand side of the inequality in the last line. Thus for all large u and t with $\alpha_2\sqrt{t} < u < \alpha_1 t$

$$\begin{aligned}\mathbf{P}\{Q(t) > \lfloor V^\varrho t\rfloor + u\} &\le \mathbf{P}\{y(t) \ge \lfloor \frac{-1}{8\mathcal{H}''(\varrho)} \cdot \frac{u^2}{t} \rfloor\} \\ &\quad + \mathbf{P}\{\widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t) - \widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t) > \frac{-u^2}{8t\mathcal{H}''(\varrho)}\} \\ &\quad + \mathbf{P}\{|\eta_0(0)| + |\eta_0^{\mathrm{eq}}(0)| + |\omega_0(0)| + |\omega_0^{\mathrm{eq}}(0)| > C_3\frac{u^2}{t}\}.\end{aligned}$$

Assumption 3.2.28 allows us to bound the first probability on the right by $C_4 t^2/u^3$ (take $\gamma = \varrho - \lambda$). Apply Chebyshev's inequality on the second line and Markov's inequality on the third one:

$$\begin{aligned}&\mathbf{P}\{Q(t) > \lfloor V^\varrho t\rfloor + u\} \\ &\qquad \le 64\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\mathbf{Var}(h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t) - h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t)) + C_3\frac{t}{u^2} + C_4\frac{t^2}{u^3} \\ &\qquad \le 128\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\Big\{\mathbf{Var}^\varrho(h^{\omega^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t)) + \mathbf{Var}^\lambda(h^{\eta^{\mathrm{eq}}}_{\lfloor V^\varrho t\rfloor + u}(t))\Big\} \\ &\qquad\quad + C_4\frac{t^2}{u^3}.\end{aligned}$$

The term $C_3 t/u^2$ was subsumed under $C_4 t^2/u^3$ due to the condition $u < \alpha_1 t$. The variances here are taken under the stationary distributions of the processes $\underline{\eta}^{\mathrm{eq}}$ and $\underline{\omega}^{\mathrm{eq}}$. That allows us to apply (3.2.19), whose right-hand side takes us back to the four-process coupling under measure $\mathbf{P}$. Recall

(3.3.1).

$$
\begin{aligned}
&\mathbf{P}\{Q(t) > \lfloor V^{\varrho}t\rfloor + u\} \\
&\quad \le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\Big\{\mathbf{E}|Q(t) - \lfloor V^{\varrho}t\rfloor - u| + \mathbf{E}|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor - u|\Big\} + C_4\frac{t^2}{u^3} \\
&\quad \le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\Big\{\mathbf{E}|Q(t) - \lfloor V^{\varrho}t\rfloor| + \mathbf{E}|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor| + 2u\Big\} + C_4\frac{t^2}{u^3} \\
&\quad = C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\Big\{\Psi(t) + 2u + \mathbf{E}|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor|\Big\} + C_4\frac{t^2}{u^3}.
\end{aligned}
$$

The variable $Q^{\eta}(t)$ above is the location of a single discrepancy between the process $\underline{\eta}$ and one started initially with $\underline{\eta}^{+}(0) = \underline{\eta}(0) + \underline{\delta}_0$.

It remains to relate $\mathbf{E}|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor|$ to $\Psi(t)$. This is where part (3.2.30) of Assumption 3.2.1 is a key point. Compute now in the four-process coupling of $\underline{\eta}$, $\underline{\eta}^{+}$, $\underline{\omega}^{-}$, $\underline{\omega}$ described in Assumption 3.2.1. Use (3.2.30) and Taylor expansion of $\mathcal{H}$ again:

$$
\begin{aligned}
\mathbf{E}|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor| &\le \mathbf{E}(Q^{\eta}(t) - Q(t)) + \Psi(t) \\
&= (\mathcal{H}'(\lambda) - \mathcal{H}'(\varrho))t + \Psi(t) \\
&\le \mathcal{H}''(\varrho)\cdot(\lambda - \varrho)t + C_6(\varrho - \lambda)^2 t + \Psi(t) \\
&= \frac{u}{4} + C_6\frac{u^2}{t} + \Psi(t) \le (\tfrac{1}{4} + C_6\alpha_1)u + \Psi(t).
\end{aligned}
\tag{3.3.4}
$$

The last inequality used $u < \alpha_1 t$. Substitute this back into the previous display and rename constants. This finishes the proof of (3.3.2) and completes the Lemma. □

Completely analogous arguments lead to the same upper bound for the lower tail of $Q(t)$, and together we get the following bound on the tail of the absolute deviation, still for $\alpha_2\sqrt{t} < u < \alpha_1 t$:

$$
\mathbf{P}\{|Q(t) - \lfloor V^{\varrho}t\rfloor| > u\} \le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\big\{\Psi(t) + u\big\} + C_4\frac{t^2}{u^3}.
$$

Next we relax the restriction to integer u and the upper limit on it:

Lemma 3.3.2. *There are positive constants α_2, t_0 such that for all $t > t_0$ and all* real *$u > \alpha_2\sqrt{t}$,*

$$
\mathbf{P}\{|Q(t) - \lfloor V^{\varrho}t\rfloor| > u\} \le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\big\{\Psi(t) + u\big\} + C_4\frac{t^2}{u^3}.
$$

Proof. Any $u \ge 1$ is less than twice its integer part. Hence by simply increasing the constants C_i, for all large t and all *real* $u \in (\alpha_2\sqrt{t},\, \alpha_1 t)$,

$$
\mathbf{P}\{|Q(t) - \lfloor V^{\varrho}t\rfloor| > u\} \le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\big\{\Psi(t) + u\big\} + C_4\frac{t^2}{u^3}. \tag{3.3.5}
$$

Recall (3.2.31). When $\alpha_1 < \alpha_0 + 2|V^\varrho| + 2$, assume $\alpha_1 t \le u < (\alpha_0 + 2|V^\varrho| + 2)t$. Then $\alpha_2\sqrt{t} < u \cdot \alpha_1/(\alpha_0 + 2|V^\varrho| + 2) < \alpha_1 t$ for large enough t, and (3.3.5) still holds for u replaced by $u \cdot \alpha_1/(\alpha_0 + 2|V^\varrho| + 2)$:

$$\begin{aligned}\mathbf{P}\{|Q(t) - \lfloor V^\varrho t\rfloor| > u\} &\le \mathbf{P}\Big\{|Q(t) - \lfloor V^\varrho t\rfloor| > u \cdot \frac{\alpha_1}{\alpha_0 + 2|V^\varrho| + 2}\Big\}\\ &\le C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\big\{\Psi(t) + u\big\} + C_4\frac{t^2}{u^3}\end{aligned}$$

via modifying the constants by factors of $\alpha_1/(\alpha_0 + 2|V^\varrho| + 2)$.

Finally, when $u \ge (\alpha_0 + 2|V^\varrho| + 2)t$, the fact that $u - |\lfloor V^\varrho t\rfloor| > \alpha_0 t$ allows us to use (3.2.31):

$$\begin{aligned}\mathbf{P}\{|Q(t) - \lfloor V^\varrho t\rfloor| > u\} &\le \mathbf{P}\{|Q(t)| > u - |\lfloor V^\varrho t\rfloor|\}\\ &\le C_7\frac{t^2}{(u - |\lfloor V^\varrho t\rfloor|)^3} \le C_8\frac{t^2}{u^3}.\end{aligned} \tag{3.3.6}$$

Combining the above cases we get the statement for all $u > \alpha_2\sqrt{t}$. □

Proof of the upper bound of Theorem 3.2.3. We now fix $r > 0$, $1 \le m < 3$, and write

$$\begin{aligned}&\mathbf{E}\big(|Q(t) - \lfloor V^\varrho t\rfloor|^m\big)\\ &= \int_0^\infty \mathbf{P}\{|Q(t) - \lfloor V^\varrho t\rfloor|^m > v\}\,\mathrm{d}v\\ &\le r^m t^{\frac{2}{3}m} + m\int_{rt^{2/3}}^\infty \Big(C_5\frac{t^2\mathcal{H}''(\varrho)^2}{u^4}\big\{\Psi(t) + u\big\} + C_4\frac{t^2}{u^3}\Big)u^{m-1}\,\mathrm{d}u\\ &= r^m t^{\frac{2}{3}m} + \frac{mC_5\mathcal{H}''(\varrho)^2}{4-m}r^{m-4}t^{\frac{2}{3}m-\frac{2}{3}}\Psi(t) + \frac{mC_5\mathcal{H}''(\varrho)^2 + C_4}{3-m}r^{m-3}t^{\frac{2}{3}m}.\end{aligned}$$

First choose $m = 1$ and r large enough to get $\Psi(t) \le Ct^{2/3}$. Then insert this bound back into the last line of the display to get the bound for general $1 \le m < 3$. □

3.4 Lower bound of the main theorem

In this section we prove the lower bound of (3.2.32). Density ϱ is fixed again, and $\lambda \in (\varrho - \gamma_0,\, \varrho)$ is a varying auxiliary density. We let the jointly defined four processes $(\underline{\eta},\, \underline{\eta}^+,\, \underline{\omega}^-,\, \underline{\omega})$ be exactly as defined in the upper bound proof of Section 3.3, namely, as given by Assumption 3.2.1 with constant densities $\lambda_i \equiv \lambda$ and $\varrho_i \equiv \varrho$. The initial distribution of $(\underline{\eta},\, \underline{\omega})$ is $\widehat{\underline{\mu}}^{\lambda,\varrho}$ of (3.2.24). Two second class particles start from the origin: Q^η between processes $\underline{\eta}$ and $\underline{\eta}^+$, and Q between processes $\underline{\omega}^-$ and $\underline{\omega}$. The quantity of primary interest is abbreviated, as before, by $\Psi(t) = \mathbf{E}|Q(t) - \lfloor V^\varrho t\rfloor|$.

To prove the lower bound of (3.2.32) it suffices, by Jensen's inequality, to prove the case $m = 1$. This means showing that $\Psi(t) \ge Ct^{2/3}$ for large t and a constant $C > 0$.

3.4.1 Perturbing a segment initially

For this proof we need to introduce another coupled system and invoke Assumption 3.2.1 once more. By concavity of the flux characteristic speeds $V^\varrho = \mathcal{H}'(\varrho)$ and $V^\lambda = \mathcal{H}'(\lambda)$ satisfy $V^\varrho \le V^\lambda$. Throughout this section $u > 0$ denotes a fixed positive integer, and

$$n = \lfloor V^\lambda t \rfloor - \lfloor V^\varrho t \rfloor + u.$$

Recall definitions (3.2.22) and (3.2.23) of the single-site coupling measures. Let $(\underline{\xi}(\cdot), \underline{\zeta}(\cdot))$ be a pair of processes that obeys the basic coupling, and whose initial distribution is the product measure

$$\Big(\bigotimes_{i<-n} \mu^{\lambda,\varrho}\Big)\Big(\bigotimes_{i=-n} \widehat{\mu}^{\lambda,\varrho}\Big)\Big(\bigotimes_{-n<i\le 0} \mu^{\lambda,\lambda}\Big)\Big(\bigotimes_{0<i} \mu^{\lambda,\varrho}\Big).$$

This initial measure complies with the pattern in (3.2.24), but translated n sites to the left so that $\widehat{\mu}^{\lambda,\varrho}$ is the distribution at site $-n$ instead of the origin. A few points about this initial state: $\underline{\xi}(0)$ has the stationary density-λ product distribution except at site $-n$ where it is $\widehat{\mu}^\lambda$-distributed. $\underline{\zeta}(0)$ has the product distribution with marginals μ^ϱ, except at sites $\{-n+1, \dots, 0\}$ where the parameter ϱ switches to λ, and at site $-n$ where it has distribution $\widehat{\mu}^\varrho + 1$. At sites $-n < i \le 0$ $\mu^{\lambda,\lambda}$ forces $\xi_i(0) = \zeta_i(0)$.

We add a second class particle to the process $\underline{\xi}(\cdot)$, start it at site $-n$ and denote its position at time t by $Q^{(-n)}(t)$. Let $\underline{\xi}^+(t) := \underline{\xi}(t) + \underline{\delta}_{Q^{-n}(t)}$.

As described in Section 3.2.6 the $\zeta - \xi$ second class particles are labeled and their ordered positions denoted by $\{X_m(t)\}$. The labeling is chosen to satisfy initially

$$\cdots \le X_{-1}(0) \le X_0(0) = -n < 0 < X_1(0) \le X_2(0) \le \dots \tag{3.4.1}$$

Thus initially $X_0(0) = -n = Q^{(-n)}(0)$. We invoke Assumption 3.2.1 to have a label process $z(t)$ with tail bound (3.2.29) such that $Q^{(-n)}(t) = X_{z(t)}(t)$. (Here $\underline{\xi}$ plays the role of $\underline{\eta}$ and $\underline{\zeta}$ plays the role of $\underline{\omega}$ of Assumption 3.2.1).

As before, the heights (or currents, recall (3.2.5)) of the processes $\underline{\xi}(\cdot)$ and $\underline{\zeta}(\cdot)$ are denoted by $h^{\xi}_{\lfloor Vt \rfloor}$ and $h^{\zeta}_{\lfloor Vt \rfloor}$, respectively. The first observation is that $Q^{(-n)}$ gives one-sided control over the difference of these currents.

Lemma 3.4.1. *For any $V \in \mathbb{R}$*

$$Q^{(-n)}(t) \le \lfloor Vt \rfloor \quad \textit{implies} \quad h^{\zeta}_{\lfloor Vt \rfloor}(t) - h^{\xi}_{\lfloor Vt \rfloor}(t) \le -z(t).$$

Proof. Recall again, from (3.2.5) and the statement around (3.2.10), that the height difference $h^{\zeta}_{\lfloor Vt \rfloor}(t) - h^{\xi}_{\lfloor Vt \rfloor}(t)$ equals the net number of second class particle passings of the path $\big(s\lfloor Vt \rfloor + 1/2\big)_{0 \le s \le 1}$ from left to right. That is,

each left-to-right passing increases $h^{\zeta}_{\lfloor Vt\rfloor}(t) - h^{\xi}_{\lfloor Vt\rfloor}(t)$ while each right-to-left passing decreases it.

Suppose $z(t) \le 0$. Then (3.4.1) and $X_{z(t)}(t) = Q^{(-n)}(t) \le \lfloor Vt\rfloor$ imply that only those second class particles with labels $z(t)+1$, $z(t)+2$, …, 0 could have passed the path $\big(s\lfloor Vt\rfloor + 1/2\big)_{0\le s\le 1}$ from left to right. The claim follows.

If $z(t) > 0$, then $X_{z(t)}(t) = Q^{(-n)}(t) \le \lfloor Vt\rfloor$ implies that at least those second class particles with labels 1, 2, …, $z(t)$ have crossed the path $\big(s\lfloor Vt\rfloor + 1/2\big)_{0\le s\le 1}$ from right to left. Again the claim follows. □

Let $\widehat{\underline{\omega}}(\cdot)$ be a process started from the product distribution $\big(\bigotimes_{i\ne -n} \mu^{\varrho}\big) \otimes (\widehat{\mu}^{\varrho}+1)$. The next lemma compares the distributions of $\underline{\zeta}$ and $\widehat{\underline{\omega}}$. No coupling of $\underline{\zeta}$ and $\widehat{\underline{\omega}}$ is proposed or required.

Lemma 3.4.2. *There exist constants $\gamma = \gamma(\varrho) > 0$ and $C_1(\varrho) < \infty$ such that for all $\lambda \in (\varrho-\gamma,\, \varrho)$ and all events A the following inequality holds:*

$$\mathbf{P}\{\underline{\zeta} \in A\} \le \mathbf{P}\{\widehat{\underline{\omega}} \in A\}^{\frac12} \cdot \exp\big\{C_1(\varrho) n (\varrho-\lambda)^2\big\}.$$

Proof. We use the Cauchy-Schwarz inequality below to perform a change of measure on the distribution of the $\underline{\zeta}$ process. First we condition on the initial $\underline{\zeta}$-configuration at sites $\{-n+1, -n+2, \dots, -1, 0\}$.

$$\begin{aligned}
\mathbf{P}\{\underline{\zeta} \in A\} &= \sum_{z_{-n+1},\dots z_{-1}, z_0} \mathbf{P}\{\,\underline{\zeta} \in A \,|\, \zeta_{-n+1}(0) = z_{-n+1}, \dots, \zeta_0(0) = z_0\} \\
&\qquad \times \Big[\prod_{i=-n+1}^{0} \mu^{\varrho}(z_i)\Big]^{\frac12} \prod_{i=-n+1}^{0} \frac{\mu^{\lambda}(z_i)}{[\mu^{\varrho}(z_i)]^{\frac12}} \\
&\le \Big[\sum_{z_{-n+1},\dots,z_0} [\mathbf{P}\{\,\underline{\zeta} \in A \,|\, \zeta_{-n+1}(0) = z_{-n+1}, \dots, \zeta_0(0) = z_0\}]^2 \prod_{i=-n+1}^{0} \mu^{\varrho}(z_i)\Big]^{\frac12} \\
&\qquad \times \Big[\sum_{z_{-n+1},\dots,z_0} \prod_{i=-n+1}^{0} \frac{[\mu^{\lambda}(z_i)]^2}{\mu^{\varrho}(z_i)}\Big]^{\frac12} \\
&\le \Big[\sum_{z_{-n+1},\dots,z_0} \mathbf{P}\{\,\underline{\zeta} \in A \,|\, \zeta_{-n+1}(0) = z_{-n+1}, \dots, \zeta_0(0) = z_0\} \prod_{i=-n+1}^{0} \mu^{\varrho}(z_i)\Big]^{\frac12} \\
&\qquad \times \Big[\sum_{z_{-n+1},\dots,z_0} \prod_{i=-n+1}^{0} \frac{[\mu^{\lambda}(z_i)]^2}{\mu^{\varrho}(z_i)}\Big]^{\frac12} \\
&= \mathbf{P}\{\widehat{\underline{\omega}} \in A\}^{\frac12} \cdot \Big[\sum_{z_{-n+1},\dots,z_0} \prod_{i=-n+1}^{0} \frac{[\mu^{\lambda}(z_i)]^2}{\mu^{\varrho}(z_i)}\Big]^{\frac12}.
\end{aligned}$$

The last inequality came from dropping the square. For the last equality note that the distributions of the initial configurations $\{\widehat{\omega}_i(0)\}$ and $\{\zeta_i(0)\}$

are product-form and agree outside the interval $\{-n+1,\, -n+2, \ldots, -1,\, 0\}$. Thus conditioned on the initial values in $\{-n+1,\, -n+2, \ldots, -1,\, 0\}$ these processes have identical conditional probabilities.

To complete the proof we bound the last factor in brackets. Recall formulas (3.2.12) and (3.2.13) for the state sum and the site-marginals. Without the power 1/2 the factor in brackets equals

$$\sum_{z_{-n+1},\ldots z_0} \Big(\frac{Z(\theta(\varrho))}{Z(\theta(\lambda))^2}\Big)^n \prod_{i=-n+1}^{0} \frac{\mathrm{e}^{(2\theta(\lambda)-\theta(\varrho))z_i}}{f(z_i)!} = \Big(\frac{Z(2\theta(\lambda)-\theta(\varrho))Z(\theta(\varrho))}{Z(\theta(\lambda))^2}\Big)^n.$$

In the Appendix we show that $\log Z(\theta)$ and $\theta(\varrho)$ are infinitely differentiable. Let $\varepsilon = \theta(\varrho) - \theta(\lambda)$. By local Lipschitz continuity of the function $\theta(\varrho)$, the interval $(\theta(\lambda) - \varepsilon,\, \theta(\lambda) + \varepsilon)$ is in $(\underline{\theta},\, \bar{\theta})$ with a small enough choice of γ. There exists some $\theta \in (\theta(\lambda) - \varepsilon,\, \theta(\lambda) + \varepsilon)$ such that

$$\begin{aligned}\log \Big(\frac{Z(2\theta(\lambda)-\theta(\varrho))Z(\theta(\varrho))}{Z(\theta(\lambda))^2}\Big) &= \log Z(\theta(\lambda)-\varepsilon) + \log Z(\theta(\lambda)+\varepsilon) \\ &\quad - 2\log Z(\theta(\lambda)) \\ &= \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\log Z(\theta)\varepsilon^2 \le C_1(\varrho)\cdot(\varrho-\lambda)^2.\end{aligned}$$

Thus we get the bound

$$\Big(\frac{Z(2\theta(\lambda)-\theta(\varrho))Z(\theta(\varrho))}{Z(\theta(\lambda))^2}\Big)^n \le \exp\{C_1(\varrho)\cdot n(\varrho-\lambda)^2\}. \qquad \square$$

3.4.2 Completion of the proof of the lower bound

The gist of the proof is to get upper bounds on the complementary probabilities $\mathbf{P}\{Q^{(-n)}(t) > \lfloor V^{\varrho}t\rfloor\}$ and $\mathbf{P}\{Q^{(-n)}(t) \le \lfloor V^{\varrho}t\rfloor\}$. As stated u is an arbitrary but fixed positive integer and $n = \lfloor V^{\lambda}t\rfloor - \lfloor V^{\varrho}t\rfloor + u$.

Lemma 3.4.3.

$$\mathbf{P}\{Q^{(-n)}(t) > \lfloor V^{\varrho}t\rfloor\} \le \frac{\Psi(t)}{u} + \frac{C_2 t(\varrho-\lambda)}{u} + \frac{2}{u}.$$

Proof. Distributionwise the system $(\underline{\xi},\, \underline{\xi}^{+},\, Q^{(-n)})$ is a translate of $(\underline{\eta},\, \underline{\eta}^{+},\, Q^{\eta})$, and so

$$\begin{aligned}\mathbf{P}\{Q^{(-n)}(t) > \lfloor V^{\varrho}t\rfloor\} &= \mathbf{P}\{Q^{(-n)}(t) + n - \lfloor V^{\lambda}t\rfloor > u\} \\ &= \mathbf{P}\{Q^{\eta}(t) - \lfloor V^{\lambda}t\rfloor > u\} \le \frac{\mathbf{E}(|Q^{\eta}(t) - \lfloor V^{\lambda}t\rfloor|)}{u} \\ &\le \frac{\mathbf{E}(|Q^{\eta}(t) - Q(t)|)}{u} + \frac{\mathbf{E}(|Q(t) - \lfloor V^{\varrho}t\rfloor|)}{u} + \frac{\lfloor V^{\lambda}t\rfloor - \lfloor V^{\varrho}t\rfloor}{u}.\end{aligned}$$

Use (3.2.30) precisely as was done in (3.3.4) to conclude that the first term equals

$$u^{-1}\mathbf{E}(Q^\eta(t) - Q(t)) = u^{-1}t\big(H'(\lambda) - H'(\varrho)\big) = -u^{-1}H''(\nu)t(\varrho - \lambda)$$

for some $\nu \in (\lambda, \varrho)$. The second term is $\Psi(t)/u$, and the third term is similarly estimated by $-u^{-1}H''(\nu)t(\varrho-\lambda)+2/u$, the last part coming from discarded integer parts. Setting $C_2 := 2 \max_{\nu\in[\varrho-\gamma,\,\varrho]} -H''(\nu)$ finishes the proof. □

Notice that $H''(\varrho) < 0$ was crucial in the previous proof, as well as in the following lemma, and the final proof thereafter. These points show where the proof fails for symmetric systems – recall that these would have lower-order current fluctuations on the characteristics.

Lemma 3.4.4. *Let K be any number such that $0 < K < -\frac{1}{3}tH''(\varrho)(\varrho-\lambda)^2$. Then for small enough $\gamma > 0$, large enough t, and $\lambda \in (\varrho-\gamma,\, \varrho)$,*

$$\begin{aligned}\mathbf{P}\{Q^{(-n)}(t) \le \lfloor V^\varrho t\rfloor\} \le{}& \frac{\boldsymbol{Var}^\varrho(\omega_0)^{1/2}\Psi(t)^{1/2}}{-\frac{1}{3}tH''(\varrho)(\varrho-\lambda)^2 - K}\cdot e^{C_1 n(\varrho-\lambda)^2}\\ &+ \frac{C_4}{-\frac{1}{6}tH''(\varrho)(\varrho-\lambda)^2 - C_3 t(\varrho-\lambda)^3 - \varrho}\cdot e^{C_1 n(\varrho-\lambda)^2}\\ &+ \frac{\boldsymbol{Var}^\lambda(\eta_0)\Psi(t)}{K^2/4} + \frac{C_6 t(\varrho-\lambda)}{K^2} + \frac{C_5}{K - 4|\lambda|}\\ &+ Ct^{\kappa-1}\gamma^{2\kappa-3}K^{-\kappa}.\end{aligned}$$

Proof. Lemma 3.4.1 leads to

$$\begin{aligned}\mathbf{P}\{Q^{(-n)}(t) \le \lfloor V^\varrho t\rfloor\} &\le \mathbf{P}\{h^\zeta_{\lfloor V^\varrho t\rfloor}(t) - h^\xi_{\lfloor V^\varrho t\rfloor}(t) \le -z(t)\}\\ &\le \mathbf{P}\{-z(t) \ge K/4\} &(3.4.2)\\ &\quad + \mathbf{P}\{h^\zeta_{\lfloor V^\varrho t\rfloor}(t) \le K + t\big(H(\lambda) - \lambda H'(\varrho)\big)\} &(3.4.3)\\ &\quad + \mathbf{P}\Big\{h^\xi_{\lfloor V^\varrho t\rfloor}(t) > 3K/4 + t\big(H(\lambda) - \lambda H'(\varrho)\big)\Big\} &(3.4.4)\end{aligned}$$

To bound (3.4.2) we use the assumed distribution bound (3.2.28) on $z(t)$ and get

$$\mathbf{P}\{-z(t) \ge K/4\} \le Ct^{\kappa-1}\gamma^{2\kappa-3}K^{-\kappa}.$$

Apply Lemma 3.4.2 to line (3.4.3) to bound it by the probability of the process $\widehat{\underline{\omega}}$:

$$(3.4.3) \le \big[\mathbf{P}\{h^{\widehat{\omega}}_{\lfloor V^\varrho t\rfloor}(t) \le K + t\big(H(\lambda) - \lambda H'(\varrho)\big)\}\big]^{\frac{1}{2}}\cdot \mathrm{e}^{C_1 n(\varrho-\lambda)^2}.$$

As in the proof of Lemma 3.3.1 we switch to stationary processes to get precise bounds:

$$\begin{aligned} h^{\widehat{\omega}}_{\lfloor V^{\varrho}t\rfloor}(t) &= \widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) + [h^{\widehat{\omega}}_{\lfloor V^{\varrho}t\rfloor}(t) - h^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t)] \\ &\quad + [\mathbf{E}h^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) - t(H(\varrho) - \varrho H'(\varrho))] + t(H(\varrho) - \varrho H'(\varrho)) \\ &\geq \widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) - |\widehat{\omega}_0(0)| - |\omega_0^{\mathrm{eq}}(0)| - |\varrho| + t(H(\varrho) - \varrho H'(\varrho)). \end{aligned}$$

After the equality sign, the absolute value of the first term in brackets is not larger than $|\widehat{\omega}_0(0) - \omega_0^{\mathrm{eq}}(0)| \leq |\widehat{\omega}_0(0)| + |\omega_0^{\mathrm{eq}}(0)|$. The second term in brackets is between $-|\varrho|$ and $|\varrho|$ due to the integer part in $\lfloor V^{\varrho}t\rfloor$. Consequently

$$h^{\widehat{\omega}}_{\lfloor V^{\varrho}t\rfloor}(t) \leq K + t(H(\lambda) - \lambda H'(\varrho))$$

implies

$$\begin{aligned} \widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) - |\widehat{\omega}_0(0)| - |\omega_0^{\mathrm{eq}}(0)| &\leq K + t[H(\lambda) - H(\varrho) + H'(\varrho)(\varrho - \lambda)] + |\varrho| \\ &\leq K + \frac{1}{2}tH''(\varrho)(\varrho - \lambda)^2 + C_3t(\varrho - \lambda)^3 + |\varrho|. \end{aligned}$$

Then, we cut the event into two parts according to the value of $|\widehat{\omega}_0(0)| + |\omega_0^{\mathrm{eq}}(0)|$ and we use (3.2.19) to bound the variance of $\mathbf{Var}[h^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t)]$ by the function $\Psi(t)$.

$$\begin{aligned} (3.4.3) &\leq \big[\mathbf{P}^{\varrho}\{\widetilde{h}^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) \leq K + \frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2\}\big]^{\frac{1}{2}} \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2} \\ &\quad + \big[\mathbf{P}\{|\widehat{\omega}_0(0)| + |\omega_0^{\mathrm{eq}}(0)| > -\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|\}\big]^{\frac{1}{2}} \\ &\qquad \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2} \\ &\leq \frac{\mathbf{Var}^{\varrho}(h^{\omega^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t))^{1/2}}{-\frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2 - K} \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2} \\ &\quad + \frac{\mathbf{E}[|\widehat{\omega}_0(0)| + |\omega_0^{\mathrm{eq}}(0)|]}{-\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|} \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2} \\ &\leq \frac{\mathbf{Var}^{\varrho}(\omega_0)^{1/2}\Psi(t)^{1/2}}{-\frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2 - K} \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2} \\ &\quad + \frac{C_4}{-\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|} \cdot \mathrm{e}^{C_1n(\varrho-\lambda)^2}. \end{aligned}$$

Now we turn to (3.4.4). To reduce $h^{\xi}_{\lfloor V^{\varrho}t\rfloor}$ to the current of the density-λ equilibrium process $h^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}$ and to get rid of the integer part errors we argue as before.

$$\begin{aligned} h^{\xi}_{\lfloor V^{\varrho}t\rfloor} &= \widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor} + [h^{\xi}_{\lfloor V^{\varrho}t\rfloor} - h^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}] \\ &\quad + [\mathbf{E}^{\lambda}h^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor} - t(H(\lambda) - \lambda H'(\varrho))] + t(H(\lambda) - \lambda H'(\varrho)). \end{aligned}$$

$h^{\xi}_{\lfloor V^{\varrho}t\rfloor}(t)$ differs by at most $|\xi_0(0)-\eta_0^{\mathrm{eq}}(0)| \le |\xi_0(0)|+|\eta_0^{\mathrm{eq}}(0)|$ from $h^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t)$). Taking integer parts again into account, giving another error term $|\lambda|$, line (3.4.4) is bounded from above by

$$\mathbf{P}\Big\{\widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) + |\xi_0(0)| + |\eta_0^{\mathrm{eq}}(0)| + |\lambda| \ge 3K/4\Big\}$$

Then, we cut the event into two parts and use Markov's inequality in the second one:

$$\begin{aligned}(3.4.4) &\le \mathbf{P}^{\lambda}\Big\{\widetilde{h}^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor}(t) \ge K/2\Big\} + \mathbf{P}\Big\{|\xi_0(0)| + |\eta_0^{\mathrm{eq}}(0)| > K/4 - |\lambda|\Big\}\\ &\le \frac{\mathbf{Var}^{\lambda}(h^{\eta^{\mathrm{eq}}}_{\lfloor V^{\varrho}t\rfloor})}{K^2/4} + \frac{C_5}{K-4|\lambda|}.\end{aligned}$$

We can use (3.2.19) again to continue with

$$(3.4.4) \le \frac{\mathbf{Var}^{\lambda}(\xi_0)\mathbf{E}(|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor|)}{K^2/4} + \frac{C_5}{K-4|\lambda|}.$$

Repeating the first two steps of calculation (3.3.4) we can write

$$\begin{aligned}\mathbf{E}(|Q^{\eta}(t) - \lfloor V^{\varrho}t\rfloor|) &\le \mathbf{E}(|Q^{\eta}(t) - Q(t)|) + \mathbf{E}(|Q(t) - \lfloor V^{\varrho}t\rfloor|)\\ &\le Ct(\varrho-\lambda) + \Psi(t).\end{aligned}$$

So, we finally get

$$(3.4.4) \le \frac{\mathbf{Var}^{\lambda}(\eta_0)\Psi(t)}{K^2/4} + \frac{C_6 t(\varrho-\lambda)}{K^2} + \frac{C_5}{K-4|\lambda|}. \qquad \square$$

Proof of the lower bound of Theorem 3.2.3. As observed in the beginning of this Section, it suffices to prove that

$$\liminf_{t\to\infty} t^{-2/3}\Psi(t) > 0. \tag{3.4.5}$$

In the last two lemmas take

$$u = \lceil ht^{2/3}\rceil, \quad \varrho-\lambda = bt^{-1/3}, \quad \text{and} \quad K = bt^{1/3},$$

where h and b are large, in particular b large enough to have $b < -\frac{1}{3}H''(\varrho)b^2$ so that K satisfies the assumption of Lemma 3.4.4. Then

$$\begin{aligned}n &= \lfloor V^{\lambda}t\rfloor - \lfloor V^{\varrho}t\rfloor + u\\ &\le (H'(\lambda) - H'(\varrho))t + u + 2\\ &= -H''(\varrho)(\varrho-\lambda)t + u + C_7 t(\varrho-\lambda)^2 + 2\\ &\le (-H''(\varrho)b + h)t^{2/3} + C_7 b^2 t^{\frac{1}{3}} + 3\\ &\le C_8 t^{2/3}\end{aligned}$$

for large enough t. With these definitions we can simplify the outcomes of Lemma 3.4.3 and Lemma 3.4.4 to the inequalities

$$\mathbf{P}\{Q^{(-n)}(t) > \lfloor V^{\varrho} t \rfloor\} \leq C\frac{\Psi(t)}{t^{2/3}} + \frac{C_2 b}{h} + \frac{2}{ht^{2/3}} \tag{3.4.6}$$

and

$$\mathbf{P}\{Q^{(-n)}(t) \leq \lfloor V^{\varrho} t \rfloor\} \leq C\left(\frac{\Psi(t)}{t^{2/3}}\right)^{1/2} + C\frac{\Psi(t)}{t^{2/3}} + \frac{C_6}{b} + \frac{C_5}{bt^{\frac{1}{3}}} + Cb^{\kappa-3}. \tag{3.4.7}$$

The new constant C depends on b and h.

The lower bound (3.4.5) now follows because the left-hand sides of (3.4.6)–(3.4.7) add up to 1 for each fixed t, while we can fix b large enough and then h large enough so that $C_2 b/h + C_6/b + Cb^{\kappa-3} < 1$ (recall $\kappa < 3$). Then $t^{-2/3}\Psi(t)$ must have a positive lower bound for all large enough t. This completes the proof of Theorem 3.2.3. □

3.5 Strong Law of Large Numbers for the second class particle

This section proves the Strong Law of Large Numbers (Corollary 3.2.5). We assume that the jump rates of the second class particle are bounded, i.e.,

$$\left.\begin{matrix} p(y+1,\, z) - p(y,\, z), & p(y,\, z) - p(y,\, z+1) \\ q(y,\, z+1) - q(y,\, z), & q(y,\, z) - q(y+1,\, z) \end{matrix}\right\} \leq C \quad \forall \omega^{\min} \leq y,\, z < \omega^{\max}. \tag{3.5.1}$$

This means that the second class particle has at most rate C to jump to the right and to the left, respectively, implying that starting at any time t, it can be bounded by rate C Poisson processes that start from its position $Q(t)$.

Proof of Corollary 3.2.5. Let $\varepsilon,\, \delta > 0$. Define the events

$$A_n := \left\{\left|\frac{Q(n^{1+\delta})}{n^{1+\delta}} - V^{\varrho}\right| > \varepsilon/2\right\}$$

for $n \in \mathbb{N}$. Then, Markov's inequality and Theorem 3.2.3 imply, for $1 \leq m < 3$ and large n,

$$\begin{aligned} \mathbf{P}\{A_n\} &= \mathbf{P}\{\left|Q(n^{1+\delta}) - V^{\varrho} n^{1+\delta}\right|^m > (\varepsilon/2)^m n^{(1+\delta)m}\} \\ &\leq \frac{1}{(\varepsilon/2)^m n^{(1+\delta)m}} \cdot \mathbf{E}[|Q(n^{1+\delta}) - V^{\varrho} n^{1+\delta}|^m] \\ &\leq \frac{C_1}{(3-m)(\varepsilon/2)^m} \cdot \frac{1}{n^{m(1+\delta)/3}}, \end{aligned}$$

which is summable if $(1+\delta)m > 3$. Here δ can be chosen arbitrarily small by taking m close to 3. By the Borel-Cantelli Lemma there exists a.s. $n_0 \in \mathbb{N}$ such that

$$\forall n \geq n_0 \qquad \Big|\frac{Q(n^{1+\delta})}{n^{1+\delta}} - V^{\varrho}\Big| < \varepsilon/2. \tag{3.5.2}$$

Using this we show that a.s. there exists $n_1 \in \mathbb{N}$ such that

$$\Big|\frac{Q(t)}{t} - V^{\varrho}\Big| < \varepsilon \quad \text{for all real } t \geq n_1^{(1+\delta)}. \tag{3.5.3}$$

Let $n \geq n_0$ and suppose there exists some $t \in [n^{1+\delta}, (n+1)^{1+\delta})$ such that (3.5.3) fails: $|Q(t) - V^{\varrho}t| \geq \varepsilon t$. Together with (3.5.2) we have, if n is large,

$$\begin{aligned} |Q(t) - Q(n^{1+\delta})| &\geq |Q(t) - V^{\varrho}t| - |Q(n^{1+\delta}) - V^{\varrho}n^{1+\delta}| - |V^{\varrho}t - V^{\varrho}n^{1+\delta}| \\ &\geq \varepsilon t - \varepsilon/2 \cdot n^{1+\delta} - |V^{\varrho}|(t - n^{1+\delta}) \\ &\geq \frac{\varepsilon}{4} n^{1+\delta}. \end{aligned} \tag{3.5.4}$$

The jump rates (3.5.1) (both left and right) of Q are bounded by C. However, the event (3.5.4) implies that at least $\lfloor \frac{\varepsilon}{4} n^{1+\delta} \rfloor$ many left jumps or this many right jumps happen in the time interval $[n^{1+\delta}, (n+1)^{1+\delta})$. For large n, the length of this interval is smaller than $2(1+\delta)n^{\delta}$. Let $N(\cdot)$ be a rate C Poisson process. Then for large n the probability of the event (3.5.4) is bounded from above by

$$\begin{aligned} 2\mathbf{P}\{N(2(1+\delta)n^{\delta}) \geq \frac{\varepsilon}{4} n^{1+\delta}\} &\leq 2\mathbf{P}\{e^{N(2(1+\delta)n^{\delta})} \geq \mathrm{e}^{\varepsilon/4 \cdot n^{1+\delta}}\} \\ &\leq 2\mathrm{e}^{-\varepsilon/4 \cdot n^{1+\delta}} \mathbf{E}[e^{N(2(1+\delta)n^{\delta})}] \\ &= 2\mathrm{e}^{-\varepsilon/4 \cdot n^{1+\delta}} \cdot \mathrm{e}^{(\mathrm{e}-1)2C(1+\delta)n^{\delta}}. \end{aligned}$$

This quantity is summable over n, so the Borel Cantelli Lemma implies that a.s. (3.5.3) holds eventually. Since this is true for each $\varepsilon > 0$, the Strong Law of Large Numbers holds. □

3.6 Microscopic concavity for a class of totally asymmetric concave exponential zero range processes

In this section we verify that Assumption 3.2.1 can be satisfied under Assumption 3.2.8, and thereby complete the proof of Theorem 3.2.9.

The task is to construct the processes $y(t)$ and $z(t)$ with the requisite properties. First let the processes $(\underline{\eta}(\cdot), \underline{\omega}(\cdot))$ evolve in the basic coupling so that $\eta_i(t) \leq \omega_i(t)$ for all $i \in \mathbb{Z}$ and $t \geq 0$. We consider as a background process this pair with the labeled and ordered $\omega - \eta$ second class particles $\cdots \leq X_{-2}(t) \leq X_{-1}(t) \leq X_0(t) \leq X_1(t) \leq X_2(t) \leq \cdots$.

At each time $t \geq 0$ this background induces a partition $\{\mathcal{M}_i(t)\}$ of the label space $\mathbb{Z}$ into intervals indexed by sites $i \in \mathbb{Z}$, with partition intervals given by

$$\mathcal{M}_i(t) := \{m \,:\, X_m(t) = i\}.$$

(For simplicity we assumed infinitely many second class particles in both directions, but no problem arises in case we only have finitely many of them.) $\mathcal{M}_i(t)$ contains the labels of the second class particles that reside at site i at time t, and can be empty. The labels of the second class particles that are at the same site as the one labeled m form the set $\mathcal{M}_{X_m(t)}(t) =: \{a^m(t),\, a^m(t)+1,\, \ldots,\, b^m(t)\}$. The processes $a^m(t)$ and $b^m(t)$ are always well-defined and satisfy $a^m(t) \leq m \leq b^m(t)$.

Let us clarify these notions by discussing the ways in which $a^m(t)$ and $b^m(t)$ can change.

- A second class particle jumps from site $X_m(t-) - 1$ to site $X_m(t-)$. Then this one necessarily has label $a^m(t-) - 1$, and it becomes the lowest labeled one at site $X_m(t-) = X_m(t)$ after the jump. Hence $a^m(t) = a^m(t-) - 1$.

- A second class particle, different from X_m, jumps from site $X_m(t-)$ to site $X_m(t-) + 1$. Then this one is necessarily labeled $b^m(t-)$, and it leaves the site $X_m(t-)$, hence $b^m(t) = b^m(t-) - 1$.

- The second class particle X_m is the highest labeled on its site, that is, $m = b^m(t-)$, and it jumps to site $X_m(t-) + 1$. Then this particle becomes the lowest labeled in the set $\mathcal{M}_{X_m(t-)+1} = \mathcal{M}_{X_m(t)}$, hence $a^m(t) = m$. In this case $b^m(t)$ can be computed from $b^m(t) - a^m(t) + 1 = \omega_{X_m(t)}(t) - \eta_{X_m(t)}(t)$, the number of second class particles at the site of X_m after the jump.

We fix initially $y(0) = z(0) = 0$. The evolution of $(y,\, z)$ is superimposed on the background evolution $(\underline{\eta},\, \underline{\omega},\, \{X_m\})$ following the general rule below: Immediately after every move of the background process that involves the site where y resides before this move, y picks a new value from the labels on the site where it resides after the move. Thus y itself jumps only within partition intervals $\mathcal{M}_i$. But y joins a new partition interval whenever it is the highest X-label on its site and its "carrier" particle X_y is forced to move to the next site on the right. This is the situation when $y(t-) = b^{y(t-)}(t-)$ and at time t an $\omega - \eta$ move from this site happens. (Recall that the choice of X-particle to move is determined by rule (3.2.25). In the present case there is only one type of $\omega - \eta$ move: the highest label from a site moves to the next site on the right.) All this works for z in exactly the same way.

Next we specify the probabilities that y and z use to refresh their values. When y and z reside at separate sites, they refresh independently.

When they are together in the same partition interval, they use the joint distribution in the third bullet below.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at site $X_{y(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) \neq a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to different parts *after* the jump then, independently of everything else,

$$y(t) := \begin{cases} a^{y(t-)}(t), & \text{with pr. } \dfrac{f(\omega_{X_{y(t-)}(t)}(t)-1) - f(\eta_{X_{y(t-)}(t)}(t))}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))}, \\ b^{y(t-)}(t), & \text{with pr. } \dfrac{f(\omega_{X_{y(t-)}(t)}(t)) - f(\omega_{X_{y(t-)}(t)}(t)-1)}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))} \end{cases} \tag{3.6.1}$$

  when the denominator is non-zero, and $y(t) := a^{y(t-)}(t)$ when the denominator is zero.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at site $X_{z(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) \neq a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to different parts *after* the jump then, independently of everything else,

$$z(t) := \begin{cases} b^{z(t-)}(t)-1, \text{ with pr. } \dfrac{f(\omega_{X_{z(t-)}(t)}(t)) - f(\eta_{X_{z(t-)}(t)}(t)+1)}{f(\omega_{X_{z(t-)}(t)}(t)) - f(\eta_{X_{z(t-)}(t)}(t))}, \\ b^{z(t-)}(t), \qquad \text{with pr. } \dfrac{f(\eta_{X_{z(t-)}(t)}(t)+1) - f(\eta_{X_{z(t-)}(t)}(t))}{f(\omega_{X_{z(t-)}(t)}(t)) - f(\eta_{X_{z(t-)}(t)}(t))} \end{cases} \tag{3.6.2}$$

  when the denominator is non-zero, and $z(t) := b^{z(t-)}(t)$ when the denominator is zero. When $\omega_{X_{z(t-)}(t)}(t) = \eta_{X_{z(t-)}(t)}(t)+1$, $b^{z(t-)}(t)-1$ is not an admissible value but in this case the probability in the first line is zero.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at sites $X_{y(t-)}(t-)$ or $X_{z(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) = a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to the same part *after* the jump, that is,

$X_{y(t-)}(t) = X_{z(t-)}(t)$ then, independently of everything else,

$$\begin{pmatrix} y(t) \\ z(t) \end{pmatrix} := \begin{cases} \begin{pmatrix} a^{y(t-)}(t) \\ b^{y(t-)}(t) - 1 \end{pmatrix}, \\ \quad \text{with pr. } \dfrac{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t) + 1)}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))}, \\ \begin{pmatrix} a^{y(t-)}(t) \\ b^{y(t-)}(t) \end{pmatrix}, \\ \quad \text{with pr. } \dfrac{f(\eta_{X_{y(t-)}(t)}(t) + 1) - f(\eta_{X_{y(t-)}(t)}(t))}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))} \\ \qquad\qquad - \dfrac{f(\omega_{X_{y(t-)}(t)}(t)) - f(\omega_{X_{y(t-)}(t)}(t) - 1)}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))}, \\ \begin{pmatrix} b^{y(t-)}(t) \\ b^{y(t-)}(t) \end{pmatrix}, \\ \quad \text{with pr. } \dfrac{f(\omega_{X_{y(t-)}(t)}(t)) - f(\omega_{X_{y(t-)}(t)}(t) - 1)}{f(\omega_{X_{y(t-)}(t)}(t)) - f(\eta_{X_{y(t-)}(t)}(t))} \end{cases} \tag{3.6.3}$$

when the denominator is non-zero, and

$$(y(t),\, z(t)) := (a^{y(t-)}(t),\, b^{y(t-)}(t))$$

when the denominator is zero. When $\omega_{X_{z(t-)}(t)}(t) = \eta_{X_{z(t-)}(t)}(t) + 1$, $b^{z(t-)}(t) - 1$ is not an admissible value but in this case the probability in the first line is zero.

The fact that the numbers on the right hand-sides are probabilities follows from $\omega_i(t) > \eta_i(t)$ on the sites i in question, and from the monotonicity and concavity of f. The above moves for y and z always occur within labels at a given site. This determines whether the particle $Q(t) := X_{y(t)}(t)$ or $Q^\eta(t) := X_{z(t)}(t)$ is the one to jump if the next move out of the site is an $\omega - \eta$ move.

We prove that the above construction has the properties required in Assumption 3.2.1.

Lemma 3.6.1. *The pair* $(\underline{\omega}^-, \underline{\omega}) := (\underline{\omega} - \underline{\delta}_{X_y}, \underline{\omega})$ *obeys basic coupling, as does the pair* $(\underline{\eta}, \underline{\eta}^+) := (\underline{\eta}, \underline{\eta} + \underline{\delta}_{X_z})$.

Proof. We write the proof for $(\underline{\omega}^-, \underline{\omega})$. We need to show that, given the configuration $(\underline{\eta}, \underline{\omega}, \{X_m\}, y)$, the jump rates of $(\underline{\omega}^-, \underline{\omega})$ are the ones prescribed in basic coupling (Section 3.2.3) and by (3.2.2). Leftward jumps of type (3.2.3) do not happen in the system under discussion. Since the jump rate function p depends only on its first argument, jumps out of sites $i \neq Q$

happen for $\underline{\omega}^-$ and $\underline{\omega}$ with the same rate $p(\omega_i^-, \omega_{i+1}^-) = f(\omega_i^-) = f(\omega_i) = p(\omega_i, \omega_{i+1})$. The only point to consider is jumps out of site $i = Q$.

Since the last time any change occurred at site i, y chose values according to (3.6.1) or (3.6.3). Notice that (3.6.1) and (3.6.3) give the same marginal probabilities for this choice. Hence

$$y \text{ took on value } a^y \text{ with probability } \quad \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} \tag{3.6.4}$$

and

$$y \text{ took on value } b^y \text{ with probability } \quad \frac{f(\omega_i) - f(\omega_i - 1)}{f(\omega_i) - f(\eta_i)}, \tag{3.6.5}$$

as given in (3.6.1), or y took on value a^y in the case $f(\omega_i) = f(\eta_i)$. According to the basic coupling of $\underline{\eta}$ and $\underline{\omega}$, the following jumps can occur over the edge $(i, i+1)$:

- With rate $p(\omega_i, \omega_{i+1}) - p(\eta_i, \eta_{i+1}) = f(\omega_i) - f(\eta_i)$, when positive, $\underline{\omega}$ jumps without $\underline{\eta}$. The highest labeled second class particle, X_{b^y} jumps from site i to site $i+1$.
  - With probability (3.6.5) $X_y = Q$ jumps with X_{b^y}. In this case
    $$\omega_i^-(t-) = \omega_i(t-) - 1 = \omega_i(t) = \omega_i^-(t)$$
    since the difference Q disappears from site i. Also,
    $$\omega_{i+1}^-(t-) = \omega_{i+1}(t-) = \omega_{i+1}(t) - 1 = \omega_{i+1}^-(t),$$
    since the difference Q appears at site $i+1$. So in this case $\underline{\omega}$ undergoes a jump but $\underline{\omega}^-$ does not, and the rate is
    $$[f(\omega_i) - f(\eta_i)] \cdot \frac{f(\omega_i) - f(\omega_i - 1)}{f(\omega_i) - f(\eta_i)} = f(\omega_i) - f(\omega_i^-).$$
  - With probability (3.6.4) $X_y = Q$ does not jump with X_{b^y}, since it has label a^y and not b^y (this probability is zero if $\omega_i = \eta_i + 1$). In this case $\underline{\omega}^-$ and $\underline{\omega}$ perform the same jump and it occurs with rate
    $$[f(\omega_i) - f(\eta_i)] \cdot \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} = f(\omega_i^-) - f(\eta_i).$$
- With rate $p(\eta_i, \eta_{i+1}) = f(\eta_i)$, both $\underline{\eta}$ and $\underline{\omega}$ jump over the edge $(i, i+1)$. No change occurs in the $\omega - \eta$ particles, hence no change occurs in Q. This implies that the process $\underline{\omega}^-$ jumps as well.

Summarizing we see that the rate for $(\underline{\omega}^-, \underline{\omega})$ to jump together over $(i, i+1)$ is $f(\omega_i^-)$, and the rate for $\underline{\omega}$ to jump without $\underline{\omega}^-$ is $f(\omega_i) - f(\omega_i^-)$. This is exactly what basic coupling requires.

A very similar argument can be repeated for $(\underline{\eta}, \underline{\eta}^+)$. □

Lemma 3.6.2. *Inequality* (3.2.26) $y \le z$ *holds in the above construction.*

Proof. Since no jump of y or z moves one of them into a new partition interval, the only situation that can jeopardize (3.2.26) is the simultaneous refreshing of y and z in a common partition interval. But this case is governed by step (3.6.3) which by definition ensures that $y \le z$. □

So far in this section everything is valid for a general zero range process with nondecreasing concave jump rate. Now we use the special convexity requirement (3.2.37). With $r \in (0, 1)$ from (3.2.37), define the geometric distribution

$$\nu(m) := \begin{cases} (1-r)r^m, & m \ge 0 \\ 0, & m < 0. \end{cases} \tag{3.6.6}$$

Lemma 3.6.3. *Conditioned on the process* $(\underline{\eta}, \underline{\omega})$*, the bounds* $y(t) \overset{d}{\le} \nu$ *and* $z(t) \overset{d}{\ge} -\nu$ *hold for all* $t \ge 0$.

The proof of this lemma is achieved in three steps.

Lemma 3.6.4. *Let* Y *be a random variable with distribution* ν*, and fix integers* $a \le b$ *and* $\eta < \omega$ *so that* $\omega - \eta = b - a + 1$*. Apply the following operation to* Y*:*

(i) if $a \le Y \le b$*, apply the probabilities from* (3.6.1) *(equivalently,* (3.6.4) *and* (3.6.5)*) with parameters* a*,* b*,* η*,* ω *to pick a new value for* Y*;*

(ii) if $Y < a$ *or* $Y > b$ *then do not change* Y*.*

Then the resulting distribution ν^* *is stochastically dominated by* ν*.*

Proof. There is nothing to prove when $b = a$, hence we assume $b > a$ or, equivalently, $\omega - \eta = b - a + 1 \ge 2$. It is also clear that $\nu^*(m) = \nu(m)$ for $m < a$ or $m > b$. We need to prove, in view of the distribution functions,

$$\sum_{\ell=a}^{m} \nu^*(\ell) \ge \sum_{\ell=a}^{m} \nu(\ell) \quad \text{or, equivalently,} \quad \sum_{\ell=m}^{b} \nu^*(\ell) \le \sum_{\ell=m}^{b} \nu(\ell)$$

for all $a \le m \le b$. Notice that ν^* gives zero weight on values $a < m < b$ (if any), therefore the left hand-side of the second inequality equals $\nu^*(b)$ for $a < m \le b$. Hence the above display is proved once we show

$$\nu^*(b) \le \nu(b), \text{ that is,}$$

$$\frac{f(\omega) - f(\omega - 1)}{f(\omega) - f(\eta)} \cdot \sum_{\ell=a}^{b} \nu(\ell) \le \nu(b), \tag{3.6.7}$$

see (3.6.1). When $f(\omega) = f(\omega-1)$, there is nothing to prove. Hence assume $f(\omega) > f(\omega-1)$ which by concavity implies that f has positive increments on $\{\eta, \dots, \omega\}$. If $b < 0$ then both sides are zero. If $b \geq 0$ then we have, by (3.2.37),

$$\begin{aligned}\nu(\ell) \leq \nu(b) \cdot r^{\ell-b} &\leq \nu(b) \cdot \prod_{z=\omega-b+\ell}^{\omega-1} \frac{f(z)-f(z-1)}{f(z+1)-f(z)} \\ &= \nu(b) \cdot \frac{f(\omega-b+\ell)-f(\omega-b+\ell-1)}{f(\omega)-f(\omega-1)}\end{aligned}$$

for each $\ell \leq b$. The first inequality also takes into account possible $\nu(\ell) = 0$ values for negative ℓ's. With this we can write

$$\sum_{\ell=a}^{b} \nu(\ell) \leq \nu(b) \cdot \frac{f(\omega)-f(\omega-b+a-1)}{f(\omega)-f(\omega-1)}$$

which becomes (3.6.7) via $\omega - \eta = b - a + 1$. □

We repeat the lemma for $z(t)$.

Lemma 3.6.5. *Let Z be a random variable of distribution $-\nu$, and fix integers $a \leq b$, $\eta < \omega$ so that $\omega - \eta = b - a + 1$. Operate on Z as was done for Y in Lemma 3.6.4, but this time use the probabilities from* (3.6.2) *with parameters a, b, η, ω. Let $-\nu^*$ be the resulting distribution. Then ν^* is stochastically dominated by ν.*

Proof. Again, we assume $b > a$ or, equivalently, $\omega - \eta = b - a + 1 \geq 2$. It is also clear that $\nu^*(-m) = \nu(-m)$ for $m < a$ or $m > b$. We need to prove

$$\sum_{\ell=a}^{m} \nu^*(-\ell) \leq \sum_{\ell=a}^{m} \nu(-\ell)$$

for all $a \leq m \leq b$. Notice that $-\nu^*$ gives zero weight on values $a \leq \ell < b-1$ (if any), therefore the left hand-side of the inequality equals 0 for $a \leq m < b-1$, $\nu^*(b-1)$ for $m = b-1$, and agrees to the right hand-side for $m = b$. Hence the above display is proved once we show

$$\nu^*(-b) \geq \nu(-b), \text{ that is,}$$

$$\frac{f(\eta+1)-f(\eta)}{f(\omega)-f(\eta)} \cdot \sum_{\ell=a}^{b} \nu(-\ell) \geq \nu(-b), \tag{3.6.8}$$

see (3.6.2). We have, by (3.2.37),

$$\begin{aligned}\nu(-\ell) \geq \nu(-b) \cdot r^{b-\ell} &\geq \nu(-b) \cdot \prod_{\substack{z=\eta+1 \\ f(z)>f(z-1)}}^{\eta+b-\ell} \frac{f(z+1)-f(z)}{f(z)-f(z-1)} \\ &= \nu(-b) \cdot \frac{f(\eta+1+b-\ell)-f(\eta+b-\ell)}{f(\eta+1)-f(\eta)}\end{aligned}$$

for each $\ell \le b$. The first inequality also takes into account possible $\nu(-b) = 0$ values for positive b's. With this we can write

$$\sum_{\ell=a}^{b} \nu(-\ell) \ge \nu(-b) \cdot \frac{f(\eta+1+b-a) - f(\eta)}{f(\eta+1) - f(\eta)}$$

which becomes (3.6.8) via $\omega - \eta = b - a + 1$. □

Lemma 3.6.6. *The dynamics defined by* (3.6.1) *or* (3.6.2) *is attractive.*

Proof. Following the same realizations of (3.6.1), we see that two copies of $y(\cdot)$ under a common environment can be coupled so that whenever they get to the same part $\mathcal{M}_i$, they move together from that moment. The same holds for $z(\cdot)$. □

Proof of Lemma 3.6.3. Initially $y(0) = 0$ by definition, which is clearly a distribution dominated by ν of (3.6.6). Now we argue recursively: by time t the distribution of $y(t)$ was a.s. only influenced by finitely many jumps of the environment, which resulted in distributions ν_1, then ν_2, then ν_3, etc. Suppose $\nu_k \overset{\mathrm{d}}{\le} \nu$, and let ν^* be the distribution that would result from ν by the $k+1^{\mathrm{st}}$ jump. Then $\nu_{k+1} \overset{\mathrm{d}}{\le} \nu^*$ by $\nu_k \overset{\mathrm{d}}{\le} \nu$ and Lemma 3.6.6, while $\nu^* \overset{\mathrm{d}}{\le} \nu$ by Lemma 3.6.4. A similar argument proves the lemma for $z(\cdot)$. □

3.7 Microscopic convexity of the exponential bricklayers process

In this section we prove the microscopic convexity properties and hence the $t^{1/3}$ scaling for yet another system, the totally asymmetric exponential bricklayers process (TAEBLP). This model was introduced in [10], and its normal fluctuations off-characteristics were demonstrated in [11] (in case of general convex jump rates, not only exponential).

Very briefly, the proofs in the previous sections work if one proves the following properties of a model (see the exact formulation therein):

1. a strict domination of a second class particle of a denser system on one of a sparser system,

2. a non-strict, but tight, domination of a second class particle on a system of second class particles that are defined between the system in question and another system with a different density,

3. strictly concave or convex, in the second derivative sense, hydrodynamic flux function of the hyperbolic conservation law obtained by the Eulerian limiting procedure,

4. a tail bound of a second class particle in a(n essentially) stationary process.

Properties 1 and 2 form what we call the *microscopic concavity or convexity property*. Arguments in this chapter are worked out for the concave setting, but everything works word-for-word in the convex case.

General convex increasing rates of a totally asymmetric bricklayers process allow couplings that prove properties 1 and 3 above. The exponential jump rates have a strong enough convexity property that will allow us to show property 2. We do this by repeating an argument somewhat similar to the one applied to the concave zero range process in Section 3.6. The idea resembles much to the concave case, but we include this convex case in full details (rather than listing all the differences from previous work) due to the complexity of the method.

Finally, property 4 is highly nontrivial when the jump rates have unbounded increments. We use a coupling based on property 1 and a recent result [14] that asserts that a second class particle of the exponential bricklayers process performs a simple (drifted) random walk under appropriate shock initial conditions. It is worth noting that exponential jump rates were also of fundamental importance in [14], this technical point being the main reason for considering this particular family of jump rates in this section. Indeed, this is the only point that prevents us from proving the result for e.g. the totally asymmetric zero range process with *convex* exponential jump rates.

We emphasize at this point that we only consider nearest neighbor models. We believe that, as far as the hydrodynamic flux is strictly convex in the second derivative sense, the $t^{1/3}$ scaling should hold for a wide class of non nearest neighbor dynamics as well. However, as intricate couplings and orderings of second class particles play a crucial role in the methods, we do not see an easy way to deal with the non-nearest neighbor case.

Let us also have a few comments on explicit product invariant stationary distributions. In [17] and thus in this chapter we explicitly use them, as they make the arguments easier. The crucial points of the method are properties 2 (microscopic convexity) and 4 (tail bound of the second class particle) above. These depend on the details of the models, and the few known examples for which they could be proved indeed have product stationary distributions. Therefore we have not investigated how the arguments in [17] could be generalized to the case of other types of stationary distributions. We again believe that once microscopic convexity and the tail bound were proved, the remainder of the argument could be generalized and the scaling would remain valid for many models with non product stationary distributions as well.

The case of exponential jump rates was constructed in [19]. The results of the note [21] are used by [17]. Those require strong construction results

which are not provided by [19] and therefore, to our knowledge, are not available. To close that gap, we reproduce the results of [21] here for the TAEBLP.

The model we discuss in this Section is the totally asymmetric exponential bricklayers process (TAEBLP), see Section 3.2.2. It was introduced in [10], and also treated in [12] and [14]. The model is a member of the class in Section 3.2.1, here is a brief definition. The process describes the growth of a surface which we imagine as the top of a wall formed by columns of bricks over the interval $(i,\, i+1)$ for each pair of neighboring sites i and $i+1$ of $\mathbb{Z}$. The height h_i of this column is integer-valued. Bricklayers processes are characterized by a function $f : \mathbb{Z} \to \mathbb{R}^+$. We only consider the totally asymmetric nearest neighbor case here, in which only deposition of bricks in the following way are allowed:

$$\left.\begin{aligned} (\omega_i,\, \omega_{i+1}) &\longrightarrow (\omega_i - 1,\, \omega_{i+1} + 1) \\ h_i &\longrightarrow h_i + 1 \end{aligned}\right\} \text{ with rate } f(\omega_i) + f(-\omega_{i+1}). \quad (3.7.1)$$

Conditionally on the present state, these moves happen independently at all sites i. Attractivity of the process is essential, this is achieved by assuming that f is nondecreasing.

Finally, stationary translation-invariant product distributions for $\underline{\omega}(\cdot)$ are ensured by $f(z) \cdot f(1-z) = 1$ for each $z \in \mathbb{Z}$.

The totally asymmetric *exponential* bricklayers process (TAEBLP) is obtained by taking

$$f(z) = \mathrm{e}^{\beta(z-1/2)}. \quad (3.7.2)$$

The construction of the bricklayers process with any nondecreasing f that is bounded by an exponential function is given in [19] on a set of tempered configurations $\widetilde{\Omega}$. This set consists of configurations with bounded asymptotic slope, the precise definition is given in [19]. As certain desired semigroup properties are not fully proved, we avoid technical difficulties in the proofs of [21] by reproducing its results for the TAEBLP in the Appendix of [16]. However, we neglect to add that Appendix here.

The basic coupling for TAEBLP

We use a particularly simple form of the basic coupling which is made possible by the bricklayer representation: it is enough to define the structure of moves as described in Section 3.2.3 for a given side (left or right) of an individual bricklayer. Here is how to do it for a given bricklayer at site i. Given the present configurations $\underline{\omega}^1, \underline{\omega}^2, \ldots, \underline{\omega}^n \in \widetilde{\Omega}$, let $m \mapsto \ell(m)$ be a permutation that orders the ω_i values:

$$\omega_i^{\ell(m)} \le \omega_i^{\ell(m+1)}, \qquad 1 \le m < n.$$

For simplicity, set

$$p(m) := f(\omega_i^{\ell(m)}) \qquad \text{and} \qquad q(m) := f(-\omega_i^{\ell(m)}),$$

and the dummy variables $p(0) = q(n+1) = 0$. Recall that the function f is nondecreasing. Now the rule is that independently for each $m = 1, \dots, n$, at rate $p(m) - p(m-1)$, precisely bricklayers of $\underline{\omega}^{\ell(m)}$, $\underline{\omega}^{\ell(m+1)}$, $\dots$, $\underline{\omega}^{\ell(n)}$ place a brick on their right, and bricklayers of $\underline{\omega}^{\ell(1)}$, $\underline{\omega}^{\ell(2)}$, $\dots$, $\underline{\omega}^{\ell(m-1)}$ do not. Independently, at rate $q(m) - q(m+1)$, precisely bricklayers of $\underline{\omega}^{\ell(1)}$, $\underline{\omega}^{\ell(2)}$, $\dots$, $\underline{\omega}^{\ell(m)}$ place a brick on their left, and bricklayers of $\underline{\omega}^{\ell(m+1)}$, $\underline{\omega}^{\ell(m+2)}$, $\dots$, $\underline{\omega}^{\ell(n)}$ do not. Given the configurations $\underline{\omega}^1$, $\underline{\omega}^2$, $\dots$, $\underline{\omega}^n \in \widetilde{\Omega}$, bricklayers at different sites perform the above steps independently.

The combined effect of these joint rates creates the correct marginal rates, that is, the bricklayer of $\underline{\omega}^{\ell(m)}$ executes the move (3.7.1) with rate $p(m) = f(\omega_i^{\ell(m)})$, and the same move on column h_{i-1} with rate $q(m) = f(-\omega_i^{\ell(m)})$.

Notice also that, due to monotonicity of f, a jump of $\underline{\omega}^a$ without $\underline{\omega}^b$ on the column $[i, i+1]$ by the bricklayers at site i can only occur if $f(\omega_i^b) < f(\omega_i^a)$ which implies $\omega_i^a > \omega_i^b$. Also, a jump of $\underline{\omega}^a$ without $\underline{\omega}^b$ on the column $[i-1, i]$ by the bricklayers at site i can only occur if $f(-\omega_i^b) < f(-\omega_i^a)$ which implies $\omega_i^a < \omega_i^b$. The result of any of these steps then cannot increase the number of discrepancies between the two processes, hence the name *attractivity* for monotonicity of f. In particular, a sitewise ordering $\omega_i^a \le \omega_i^b \quad \forall i \in \mathbb{Z}$ is preserved by the basic coupling.

Hydrodynamics and exact identities for TAEBLP

Recall the jump rates (3.7.1). As described in 3.2.4, the process has product translation-invariant stationary distribution with marginals μ^θ

$$\mu^\theta(z) = \frac{\mathrm{e}^{\theta z}}{f(z)!} \cdot \frac{1}{Z(\theta)} \tag{3.7.3}$$

that turn out to be of discrete Gaussian type, see [10] for the explicit formula. The density $\varrho(\theta) := \mathbf{E}^\theta(\omega) \in \mathbb{R}$ is a strictly increasing function of the parameter $\theta \in \mathbb{R}$, and can take on any real value by the Appendices below. As before, μ^ϱ, $\mathbf{P}^\varrho$, $\mathbf{E}^\varrho$, $\mathbf{Var}^\varrho$, $\mathbf{Cov}^\varrho$ will refer to laws of a density ϱ stationary process.

The hydrodynamic flux in this case is

$$\mathcal{H}(\varrho) = \mathbf{E}^\varrho[f(\omega) + f(-\omega)] = \mathrm{e}^{\theta(\varrho)} + \mathrm{e}^{-\theta(\varrho)}.$$

As f (3.7.2) is convex and nonlinear, the Appendices apply and yields a convex hydrodynamic flux with

$$\mathcal{H}''(\varrho) > 0. \tag{3.7.4}$$

Recall the measure (3.2.17) from Section 3.2.5, The Appendices below apply to show that both μ^ϱ and $\widehat{\mu}^\varrho$ are stochastically monotone in ϱ. As before, denote by $\mathbf{E}$ the expectation w.r.t. the evolution of a pair $(\underline{\omega}^-(\cdot), \underline{\omega}(\cdot))$ started with initial data (recall (3.2.1))

$$\underline{\omega}^-(0) = \underline{\omega}(0) - \underline{\delta}_0 \sim \Big(\bigotimes_{i\neq 0} \mu^\varrho\Big) \otimes \widehat{\mu}^\varrho, \tag{3.7.5}$$

and evolving under the basic coupling. This pair will always have a single second class particle whose position is denoted by $Q(t)$. In other words, $\underline{\omega}^-(t) = \underline{\omega}(t) - \underline{\delta}_{Q(t)}$. We reprove Corollaries 2.4 and 2.5 of [21] in the Appendix of [16] that state that for any $i \in \mathbb{Z}$ and $t \geq 0$,

$$\mathbf{Var}^\varrho(h_i(t)) = \mathbf{Var}^\varrho(\omega) \cdot \mathbf{E}|Q(t) - i| \tag{3.7.6}$$

and

$$\mathbf{E}(Q(t)) = V^\varrho \cdot t,$$

where $V^\varrho = \mathcal{H}'(\varrho)$ is the characteristic speed. Note in particular that in (3.7.6) the variances are taken in a stationary process, while the expectation of $Q(t)$ is taken in the coupling with initial distribution (3.7.5).

Microscopic convexity

We start with the definition of *microscopic convexity*. This is just translated from the microscopic concavity property of Section 3.2.6, where more detailed explanations and comments can be found. We assume that the setting in Section 3.2.6 holds true, i.e. the processes start under the measure (3.2.24) and the label of second class particles follows the rule (3.2.25).

Assumption 3.7.1. *Given a density $\varrho \in \mathbb{R}$, there exists $\gamma_0 > 0$ such that the following holds. For any $\underline{\lambda}$ and $\underline{\varrho}$ such that $\varrho - \gamma_0 \leq \lambda_i \leq \varrho_i \leq \varrho + \gamma_0$ for all $i \in \mathbb{Z}$, a joint process $(\underline{\eta}(t), \underline{\omega}(t), y(t), z(t))_{t\geq 0}$ can be constructed with the following properties.*

- *Initially $(\underline{\eta}(0), \underline{\omega}(0))$ is $\underline{\widehat{\mu}}^{\underline{\lambda},\underline{\varrho}}$-distributed and the joint process $(\underline{\eta}(\cdot), \underline{\omega}(\cdot))$ evolves in basic coupling.*
- *Processes $y(\cdot)$ and $z(\cdot)$ are integer-valued. Initially $y(0) = z(0) = 0$. With probability one*

$$y(t) \geq z(t) \text{ for all } t \geq 0. \tag{3.7.7}$$

- *Define the processes*

$$\underline{\omega}^-(t) := \underline{\omega}(t) - \underline{\delta}_{X_{y(t)}(t)} \quad \text{and} \quad \underline{\eta}^+(t) := \underline{\eta}(t) + \underline{\delta}_{X_{z(t)}(t)}. \tag{3.7.8}$$

Then both pairs $(\underline{\eta}, \underline{\eta}^+)$ and $(\underline{\omega}^-, \underline{\omega})$ evolve marginally in basic coupling.

- *For each $\gamma \in (0, \gamma_0)$ and large enough $t \geq 0$ there exists a probability distribution $\nu^{\varrho,\gamma}(t)$ on $\mathbb{Z}^+$ satisfying the tail bound*

$$\nu^{\varrho,\gamma}(t)\{y : y \geq y_0\} \leq Ct^{\kappa-1}\gamma^{2\kappa-3}y_0^{-\kappa} \tag{3.7.9}$$

for some fixed constants $3/2 \leq \kappa < 3$ and $C < \infty$, and such that if $\varrho - \gamma \leq \lambda_i \leq \varrho_i \leq \varrho + \gamma$ for all $i \in \mathbb{Z}$, then we have the stochastic bounds

$$y(t) \overset{d}{\geq} -\nu^{\varrho,\gamma}(t) \quad \text{and} \quad z(t) \overset{d}{\leq} \nu^{\varrho,\gamma}(t). \tag{3.7.10}$$

Thus, the only difference is that in the convex setting we need

$$Q^\eta(t) = X_{z(t)}(t) \leq X_{y(t)}(t) = Q(t).$$

and the tail bounds of (3.2.29) hold in the reversed order (3.7.10)

Section 3.7.1 contains the proof of Assumption 3.7.1 for the TAEBLP. The proof of (3.7.10) makes use of the particular exponential form (3.7.2) of the rates. Unfortunately, we do not have an argument for more general convex rates at the moment.

There is one more assumption in Section 3.2.6 needed to state the main result, and this is the inequality (3.2.31) of Assumption 3.2.2. Such an assumption is natural and easy to prove if the jump rates have bounded increments. Since f (3.7.2) does not, this statement for the TAEBLP is nontrivial. We prove it in Section 3.7.2 for the TAEBLP.

3.7.1 Proof of microscopic convexity

In this section we verify that Assumption 3.7.1 can be satisfied. The proof is similar to that of for concave zero range processes, so we try to stick to the differences, but some repetition might occur. Recall from the beginning of Section 3.6 the partitioning $\{\mathcal{M}_i(t)\}$ of the label space $\mathbb{Z}$ into intervals indexed by sites $i \in \mathbb{Z}$, with partition intervals given by $\mathcal{M}_i(t)$ contains the labels of the second class particles that reside at site i at time t, and can be empty. The labels of the second class particles that are at the same site as the one labeled m form the set $\mathcal{M}_{X_m(t)}(t) =: \{a^m(t), a^m(t)+1, \ldots, b^m(t)\}$. The processes $a^m(t)$ and $b^m(t)$ are always well-defined and satisfy $a^m(t) \leq m \leq b^m(t)$. Notice that

$$|\mathcal{M}_{X_m(t)}(t)| = b^m(t) - a^m(t) + 1 = \omega_{X_m(t)}(t) - \eta_{X_m(t)}(t). \tag{3.7.11}$$

In the TAEBLP, the ways in which $a^m(t)$ and $b^m(t)$ can change are a bit more complicated than the same for TAZRP, and can be summarized the following list:

- A second class particle jumps from site $X_m(t-) - 1$ to site $X_m(t-)$. Then this one necessarily has label $a^m(t-) - 1$, and it becomes the lowest labeled one at site $X_m(t-) = X_m(t)$ after the jump. Hence $a^m(t) = a^m(t-) - 1$.

- A second class particle jumps from site $X_m(t-)+1$ to site $X_m(t-)$. Then this one necessarily has label $b^m(t-)+1$, and it becomes the highest labeled one at site $X_m(t-) = X_m(t)$ after the jump. Hence $b^m(t) = b^m(t-)+1$.

- A second class particle, different from X_m, jumps from site $X_m(t-)$ to site $X_m(t-)+1$. Then this one is necessarily labeled $b^m(t-)$, and it leaves the site $X_m(t-)$, hence $b^m(t) = b^m(t-)-1$.

- A second class particle, different from X_m, jumps from site $X_m(t-)$ to site $X_m(t-)-1$. Then this one is necessarily labeled $a^m(t-)$, and it leaves the site $X_m(t-)$, hence $a^m(t) = a^m(t-)+1$.

- The second class particle X_m is the highest labeled on its site, that is, $m = b^m(t-)$, and it jumps to site $X_m(t-)+1$. Then this particle becomes the lowest labeled in the set $\mathcal{M}_{X_m(t-)+1} = \mathcal{M}_{X_m(t)}$, hence $a^m(t) = m$. In this case $b^m(t)$ can be computed from (3.7.11), the number of second class particles at the site of X_m after the jump.

- The second class particle X_m is the lowest labeled on its site, that is, $m = a^m(t-)$, and it jumps to site $X_m(t-)-1$. Then this particle becomes the highest labeled in the set $\mathcal{M}_{X_m(t-)-1} = \mathcal{M}_{X_m(t)}$, hence $b^m(t) = m$. In this case $a^m(t)$ can be computed from (3.7.11), the number of second class particles at the site of X_m after the jump.

Similarly as for the TAZRP, we fix initially $y(0) = z(0) = 0$. Next we specify the probabilities that y and z use to refresh their values. Recall (3.7.2). To simplify notation, we abbreviate, given integers $\eta < \omega$,

$$p(\eta,\, \omega) = \frac{f(\omega) - f(\omega-1)}{f(\omega) - f(\eta)} = \frac{f(-\eta) - f(-\eta-1)}{f(-\eta) - f(-\omega)} = \frac{\mathrm{e}^{\beta(\omega-\eta)} - \mathrm{e}^{\beta(\omega-\eta-1)}}{\mathrm{e}^{\beta(\omega-\eta)} - 1} \tag{3.7.12}$$

and

$$q(\eta,\, \omega) = \frac{f(-\omega+1) - f(-\omega)}{f(-\eta) - f(-\omega)} = \frac{f(\eta+1) - f(\eta)}{f(\omega) - f(\eta)} = \frac{\mathrm{e}^{\beta} - 1}{\mathrm{e}^{\beta(\omega-\eta)} - 1}. \tag{3.7.13}$$

Notice that both $p(\eta,\, \omega)$ and $q(\eta,\, \omega)$ only depend on $\omega - \eta$. Therefore, with a little abuse of notation, we write $p(\omega-\eta) := p(\eta,\, \omega)$, $q(\omega-\eta) := q(\eta,\, \omega)$. Then

$$p(1) = q(1) = 1, \qquad p(d) \geq q(d), \qquad p(d)+q(d) \leq 1 \qquad \text{for } 2 \leq d \in \mathbb{Z}.$$

When y and z reside at separate sites, they refresh independently. When they are together in the same partition interval, they use the joint distribution in the third bullet below.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at site $X_{y(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) \neq a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to different parts *after* the jump, we abbreviate

$$p = p(\eta_{X_{y(t-)}(t)}(t),\, \omega_{X_{y(t-)}(t)}(t)) \qquad q = q(\eta_{X_{y(t-)}(t)}(t),\, \omega_{X_{y(t-)}(t)}(t))$$

of (3.7.12) and (3.7.13) in the formulas below. These depend on the values of the respective processes at the site where the label y can be found right after the jump. In this case, independently of everything else,

$$y(t) := \begin{cases} a^{y(t-)}(t), & \text{with prob. } q, \\ b^{y(t-)}(t) - 1, & \text{with prob. } 1-p-q, \\ b^{y(t-)}(t), & \text{with prob. } p, \end{cases} \tag{3.7.14}$$

except for $y(t) := a^{y(t-)}(t) = b^{y(t-)}(t)$ when the difference

$$\omega_{X_{y(t-)}(t)}(t) - \eta_{X_{y(t-)}(t)}(t)$$

is 1. Notice that the second line in (3.7.14) has probability zero when this difference is 2.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at site $X_{z(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) \neq a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to different parts *after* the jump, we abbreviate

$$p = p(\eta_{X_{z(t-)}(t)}(t),\, \omega_{X_{z(t-)}(t)}(t)) \qquad q = q(\eta_{X_{z(t-)}(t)}(t),\, \omega_{X_{z(t-)}(t)}(t))$$

of (3.7.12) and (3.7.13) in the formulas below. These depend on the values of the respective processes at the site where the label z can be found right after the jump. In this case, independently of everything else,

$$z(t) := \begin{cases} a^{z(t-)}(t), & \text{with prob. } p, \\ a^{z(t-)}(t) + 1, & \text{with prob. } 1-p-q, \\ b^{z(t-)}(t), & \text{with prob. } q, \end{cases} \tag{3.7.15}$$

except for $z(t) := a^{z(t-)}(t) = b^{z(t-)}(t)$ when the difference

$$\omega_{X_{z(t-)}(t)}(t) - \eta_{X_{z(t-)}(t)}(t)$$

is 1. Notice that the second line in (3.7.15) has probability zero when this difference is 2.

- Whenever any change occurs in either $\underline{\omega}$ or $\underline{\eta}$ at sites $X_{y(t-)}(t-)$ or $X_{z(t-)}(t-)$ and, as a result of the jump, $a^{y(t-)}(t) = a^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to the same part *after* the jump, that is, $X_{y(t-)}(t) = X_{z(t-)}(t)$ then we have

$$\omega_{X_{y(t-)}(t)}(t) = \omega_{X_{z(t-)}(t)}(t) \qquad \text{and} \qquad \eta_{X_{y(t-)}(t)}(t) = \eta_{X_{z(t-)}(t)}(t),$$

and we abbreviate

$$p = p(\eta_{X_{y(t-)}(t)}(t),\, \omega_{X_{y(t-)}(t)}(t)) \qquad q = q(\eta_{X_{y(t-)}(t)}(t),\, \omega_{X_{y(t-)}(t)}(t))$$

of (3.7.12) and (3.7.13) in the formulas below. These depend on the values of the respective processes at the site where both the labels y and z can be found right after the jump. In this case, independently of everything else,

$$\begin{pmatrix} y(t) \\ z(t) \end{pmatrix} := \begin{cases} \begin{pmatrix} a^{y(t-)}(t) \\ a^{y(t-)}(t) \end{pmatrix}, \text{ with prob. } q, \\ \begin{pmatrix} b^{y(t-)}(t) - 1 \\ a^{y(t-)}(t) \end{pmatrix}, \text{ with prob. } (p-q) \wedge (1-p-q), \\ \begin{pmatrix} b^{y(t-)}(t) \\ a^{y(t-)}(t) \end{pmatrix}, \text{ with prob. } [2p-1]^+, \\ \begin{pmatrix} b^{y(t-)}(t) - 1 \\ a^{y(t-)}(t) + 1 \end{pmatrix}, \text{ with prob. } [1-2p]^+, \\ \begin{pmatrix} b^{y(t-)}(t) \\ a^{y(t-)}(t) + 1 \end{pmatrix}, \text{ with prob. } (p-q) \wedge (1-p-q), \\ \begin{pmatrix} b^{y(t-)}(t) \\ b^{y(t-)}(t) \end{pmatrix}, \text{ with prob. } q, \end{cases} \tag{3.7.16}$$

except for $y(t) = z(t) := a^{y(t-)}(t) = b^{y(t-)}(t)$ when the difference $\omega_{X_{y(t-)}(t)}(t) - \eta_{X_{y(t-)}(t)}(t)$ is 1. Notice that the second, the fourth and the fifth lines have probability zero when this difference is 2.

The above moves for y and z always occur within labels at a given site. This determines whether the particle $Q(t) := X_{y(t)}(t)$ or $Q^\eta(t) := X_{z(t)}(t)$ is the one to jump if the next move out of the site is an $\omega - \eta$ move.

We prove that the above construction has the properties required in Assumption 3.7.1. First note that the refreshing rule (3.7.16) marginally gives the same moves and probabilities as (3.7.14) or (3.7.15) for $y(\cdot)$ or $z(\cdot)$, respectively.

Lemma 3.7.2. *The pair* $(\underline{\omega}^-, \underline{\omega}) := (\underline{\omega} - \underline{\delta}_{X_y}, \underline{\omega})$ *obeys basic coupling, as does the pair* $(\underline{\eta}, \underline{\eta}^+) := (\underline{\eta}, \underline{\eta} + \underline{\delta}_{X_z})$.

Proof. We write the proof for $(\underline{\omega}^-, \underline{\omega})$. We need to show that, given the configuration $(\underline{\eta}, \underline{\omega}, \{X_m\}, y)$, the jump rates of $(\underline{\omega}^-, \underline{\omega})$ are the ones prescribed in basic coupling (Section 3.7) and by (3.7.1). As mentioned in Section 3.7, the effect of bricklayers determine the evolution of processes. Notice first that an $\omega - \eta$ particle can only jump away from a site i if a bricklayer of ω or η moves. As the moves (3.7.14) or (3.7.16) by themselves never result in a change of $X_{y(\cdot)}(\cdot)$, any move of Q from a site i is a result of a bricklayer's move at site i. Therefore, we see that moves initiated by bricklayers of $\underline{\omega}$ at sites $i \neq Q$ happen as well to $\underline{\omega}^-$, as required by the basic coupling. The only point to consider is moves by the bricklayers at site $i = Q$. We start with them putting a brick on their right. Since the last time any change occurred at site i, y chose values according to (3.7.14) or (3.7.16). Notice that (3.7.14) and (3.7.16) give the same marginal probabilities for this choice. Hence

$$y \text{ took a value } < b^y \text{ with prob.} \quad 1 - p = \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} \tag{3.7.17}$$

and

$$y \text{ took on value } b^y \text{ with prob.} \quad p = \frac{f(\omega_i) - f(\omega_i - 1)}{f(\omega_i) - f(\eta_i)}, \tag{3.7.18}$$

as given in (3.7.14). Notice that (3.7.17) happens with probability zero if $\omega_i = \eta_i + 1$. According to the basic coupling of $\underline{\eta}$ and $\underline{\omega}$, the following right moves of bricklayers at i can occur:

- With rate $f(\omega_i) - f(\eta_i)$, $\underline{\omega}$ jumps without $\underline{\eta}$. The highest labeled second class particle, X_{b^y} jumps from site i to site $i + 1$.
  - With probability (3.7.18) $X_y = Q$ jumps with X_{b^y}. In this case
    $$\omega_i^-(t-) = \omega_i(t-) - 1 = \omega_i(t) = \omega_i^-(t)$$
    since the difference Q disappears from site i. Also,
    $$\omega_{i+1}^-(t-) = \omega_{i+1}(t-) = \omega_{i+1}(t) - 1 = \omega_{i+1}^-(t),$$
    since the difference Q appears at site $i + 1$. So in this case $\underline{\omega}$ undergoes a jump but $\underline{\omega}^-$ does not, and the rate is
    $$[f(\omega_i) - f(\eta_i)] \cdot \frac{f(\omega_i) - f(\omega_i - 1)}{f(\omega_i) - f(\eta_i)} = f(\omega_i) - f(\omega_i^-).$$
  - With probability (3.7.17) $X_y = Q$ does not jump with X_{b^y}, since it has label less than b^y (this probability is zero if $\omega_i = \eta_i + 1$). In this case $\underline{\omega}^-$ and $\underline{\omega}$ perform the same jump and it occurs with rate
    $$[f(\omega_i) - f(\eta_i)] \cdot \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} = f(\omega_i^-) - f(\eta_i).$$

- With rate $f(\eta_i)$, both bricklayers of $\underline{\eta}$ and $\underline{\omega}$ at site i move. No change occurs in the $\omega - \eta$ particles, hence no change occurs in Q. This implies that the process $\underline{\omega}^-$ jumps as well.

Summarizing we see that the rate for the bricklayers of $(\underline{\omega}^-, \underline{\omega})$ at site i to lay brick on their rights together is $f(\omega_i^-)$, and the rate for the one of $\underline{\omega}$ to move without $\underline{\omega}^-$ is $f(\omega_i) - f(\omega_i^-)$. This is exactly what basic coupling requires.

Consider now bricklayers at site $i = Q$ putting a brick on their left. Since the last time any change occurred at site i, y chose values according to (3.7.14) or (3.7.16). Hence

$$y \text{ took on value } a^y \text{ with prob.} \quad q = \frac{f(-\omega_i + 1) - f(-\omega_i)}{f(-\eta_i) - f(-\omega_i)} \tag{3.7.19}$$

and

$$y \text{ took a value } > a^y \text{ with prob.} \quad 1 - q = \frac{f(-\eta_i) - f(-\omega_i + 1)}{f(-\eta_i) - f(-\omega_i)}, \tag{3.7.20}$$

as given in (3.7.14). Notice that (3.7.20) happens with probability zero if $\omega_i = \eta_i + 1$. According to the basic coupling of $\underline{\eta}$ and $\underline{\omega}$, the following left moves of bricklayers at i can occur:

- With rate $f(-\eta_i) - f(-\omega_i)$, $\underline{\eta}$ jumps without $\underline{\omega}$. The lowest labeled second class particle, X_{a^y} jumps from site i to site $i - 1$.
  - With probability (3.7.19) $X_y = Q$ jumps with X_{a^y}. In this case

    $$\omega_i^-(t-) = \omega_i(t-) - 1 = \omega_i(t) - 1 = \omega_i^-(t) - 1$$

    since the difference Q disappears from site i. Also,

    $$\omega_{i-1}^-(t-) = \omega_{i-1}(t-) = \omega_{i-1}(t) = \omega_{i-1}^-(t) + 1,$$

    since the difference Q appears at site $i + 1$. So in this case $\underline{\omega}^-$ undergoes a jump but $\underline{\omega}$ does not, and the rate is

    $$[f(-\eta_i) - f(-\omega_i)] \cdot \frac{f(-\omega_i + 1) - f(-\omega_i)}{f(-\eta_i) - f(-\omega_i)} = f(-\omega_i^-) - f(-\omega_i).$$

  - With probability (3.7.20) $X_y = Q$ does not jump with X_{a^y}, since it has label more than a^y (this probability is zero if $\omega_i = \eta_i + 1$). In this case none of $\underline{\omega}^-$ or $\underline{\omega}$ move; this occurs with rate

    $$[f(-\eta_i) - f(-\omega_i)] \cdot \frac{f(-\eta_i) - f(-\omega_i + 1)}{f(-\eta_i) - f(\omega_i)} = f(-\eta_i) - f(-\omega_i^-).$$

- With rate $f(-\omega_i)$, both bricklayers of $\underline{\eta}$ and $\underline{\omega}$ at site i move. No change occurs in the $\omega - \eta$ particles, hence no change occurs in Q. This implies that the process $\underline{\omega}^-$ jumps as well.

Summarizing we see that the rate for the bricklayers of $(\underline{\omega}^-, \underline{\omega})$ at site i to lay brick on their rights together is $f(-\omega_i)$, and the rate for the one of $\underline{\omega}^-$ to move without $\underline{\omega}$ is $f(-\omega_i^-) - f(-\omega_i)$. This is exactly what basic coupling requires.

A very similar argument can be repeated for $(\underline{\eta}, \underline{\eta}^+)$. □

Lemma 3.7.3. *Inequality* (3.7.7) $y \geq z$ *holds in the above construction.*

Proof. Since no jump of y or z moves one of them into a new partition interval, the only situation that can jeopardize (3.7.7) is the simultaneous refreshing of y and z in a common partition interval. But this case is governed by step (3.7.16) which by definition ensures that $y \geq z$. (When $b^{y(t-)}(t) = a^{y(t-)}(t) + 1$, we have, by (3.7.11), $\omega_{X_{y(t-)}(t)}(t) - \eta_{X_{y(t-)}(t)}(t) = 2$, and hence p of (3.7.16) is more than $1/2$. Therefore the probability of the step in line 4 of (3.7.16) is zero.) □

Define the geometric distribution

$$\nu(m) := \begin{cases} \mathrm{e}^{-\beta m}(1 - \mathrm{e}^{-\beta}), & m \geq 0 \\ 0, & m < 0. \end{cases} \tag{3.7.21}$$

Lemma 3.7.4. *Conditioned on the process* $(\underline{\eta}, \underline{\omega})$*, the bounds* $y(t) \overset{d}{\geq} -\nu$ *and* $z(t) \overset{d}{\leq} \nu$ *hold for all* $t \geq 0$*.*

To avoid unnecessary complications with negative values, we show the proof for $z(t)$. Notice that both the statement and the behavior of $y(t)$ is reflected compared to $z(t)$, hence the proof is the same for the two processes. The proof is the same manner as for the TAZRP but with different calculations, thus consists of three steps.

Lemma 3.7.5. *Let* Z *be a random variable with distribution* ν*, and fix integers* $a \leq b$ *and* $\eta < \omega$ *so that* $\omega - \eta = b - a + 1$*. Apply the following operation to* Z*:*

(i) if $a \leq Z \leq b$*, apply the probabilities from* (3.6.2) *with parameters* a*,* b*,* η*,* ω *to pick a new value for* Z*;*

(ii) if $Z < a$ *or* $Z > b$ *then do not change* Z*.*

Then the resulting distribution ν^* *is stochastically dominated by* ν*.*

Proof. There is nothing to prove when $b = a$, hence we assume $b > a$ or, equivalently, $\omega - \eta = b - a + 1 \geq 2$. It is also clear that $\nu^*(m) = \nu(m)$ for $m < a$ or $m > b$. We need to prove, in view of the distribution functions,

$$\sum_{\ell=a}^{m} \nu^*(\ell) \geq \sum_{\ell=a}^{m} \nu(\ell)$$

for all $a \leq m \leq b$. Notice that ν^* gives zero weight on values $a+1 < m < b$ (if any), and also that the display becomes an equality if $m = b$. Therefore, it is enough to prove the inequality for $m = a$:

$$\nu^*(a) \geq \nu(a), \tag{3.7.22}$$

and $m = b - 1$:

$$\sum_{\ell=a}^{b-1} \nu^*(\ell) \geq \sum_{\ell=a}^{b-1} \nu(\ell) \qquad \text{that is,} \qquad \nu^*(b) \leq \nu(b). \tag{3.7.23}$$

Notice that (3.7.22) is trivially true for $a < 0$. For $a \geq 0$ we start with rewriting the left hand-side of (3.7.22) with the use of (3.7.15), (3.7.12), and the abbreviation $d = \omega - \eta = b - a + 1$:

$$\begin{aligned}
\nu^*(a) &= p(d) \cdot \sum_{\ell=a}^{b} \nu(\ell) \\
&= \frac{\mathrm{e}^{\beta d} - \mathrm{e}^{\beta(d-1)}}{\mathrm{e}^{\beta d} - 1} \cdot (\mathrm{e}^{-\beta a} - \mathrm{e}^{-\beta(b+1)}) \\
&= \mathrm{e}^{-\beta a} \cdot \frac{\mathrm{e}^{\beta d} - \mathrm{e}^{\beta(d-1)}}{\mathrm{e}^{\beta d} - 1} \cdot (1 - \mathrm{e}^{-\beta d}) = \nu(a).
\end{aligned}$$

As for (3.7.23), both sides become zero if $b < 0$. For $b \geq 0$ we have

$$\begin{aligned}
\nu^*(b) &= q(d) \cdot \sum_{\ell=a}^{b} \nu(\ell) \\
&\leq \frac{\mathrm{e}^{\beta} - 1}{\mathrm{e}^{\beta d} - 1} \cdot (\mathrm{e}^{-\beta a} - \mathrm{e}^{-\beta(b+1)}) \\
&= \mathrm{e}^{-\beta b} \cdot \frac{\mathrm{e}^{\beta} - 1}{\mathrm{e}^{\beta d} - 1} \cdot (\mathrm{e}^{\beta(d-1)} - \mathrm{e}^{-\beta}) = \nu(b).
\end{aligned}$$

□

Lemma 3.7.6. *The dynamics defined by* (3.7.15) *is attractive.*

Proof. Following the same realizations of (3.7.15), we see that two copies of $z(\cdot)$ under a common environment can be coupled so that whenever they get to the same part $\mathcal{M}_i$, they move together from that moment. □

Proof of Lemma 3.7.4. Follows from the previous three lemmas analogously as the proof of Lemma 3.6.3. □

3.7.2 A tail bound for the second class particle

In this section we prove that Assumption 3.2.2 holds for the TAEBLP model. The difficulty comes from the fact that jump rates of the second class particle, being the increments of the growth rates (3.7.2), are unbounded. First recall the coupling measure $\mu^{\lambda,\varrho}$ of (3.2.23) and notice that it gives weight one on pairs of the form $(y,\, y)$ if $\lambda = \varrho$. Define also $\mu^{\text{shock}\,\varrho}$ by

$$\mu^{\text{shock}\,\varrho}(y,\, z) = \begin{cases} \mu^{\varrho}(y), & \text{if } z = y+1, \\ 0, & \text{otherwise.} \end{cases}$$

With these marginals we define the shock product distribution

$$\underline{\mu}^{\text{shock}\,\varrho} := \bigotimes_{i<0} \mu^{\varrho+1,\varrho+1} \cdot \bigotimes_{i=0} \mu^{\text{shock}\,\varrho} \cdot \bigotimes_{i>0} \mu^{\varrho,\varrho}, \tag{3.7.24}$$

a measure on a pair of coupled processes with a single second class particle at the origin.

Lemma 3.7.7. *The first marginal of $\underline{\mu}^{\text{shock}\,\varrho}$ is the product distribution*

$$\bigotimes_{i<0} \mu^{\varrho+1} \cdot \bigotimes_{i\geq 0} \mu^{\varrho},$$

while the second marginal is

$$\bigotimes_{i\leq 0} \mu^{\varrho+1} \cdot \bigotimes_{i<0} \mu^{\varrho}. \tag{3.7.25}$$

Proof. The first part of the statement and the second part, apart from $i = 0$, follow from the definitions. The nontrivial part is

$$\mu^{\varrho+1}(z) = \mu^{\varrho}(z-1), \qquad z \in \mathbb{Z},$$

valid for the second marginal at $i = 0$. This is specific to the definition (3.2.13) of μ^{ϱ}, and of the exponential rates (3.7.2), and to prove it we write, with $\theta = \theta(\varrho)$,

$$\mu^{\varrho}(z-1) = \frac{f(z)}{\mathrm{e}^{\theta}} \cdot \frac{\mathrm{e}^{\theta z}}{f(z)!} \cdot \frac{1}{Z(\theta)} = \frac{\mathrm{e}^{(\theta+\beta)z}}{f(z)!} \cdot \frac{1}{\mathrm{e}^{\theta+\beta/2} Z(\theta)}.$$

Summing this up for all $z \in \mathbb{Z}$ gives one on the left hand-side, hence leads to

$$Z(\theta+\beta) = \sum_{z=-\infty}^{\infty} \frac{\mathrm{e}^{(\theta+\beta)z}}{f(z)!} = \mathrm{e}^{\theta+\beta/2} Z(\theta),$$

which also implies

$$\varrho(\theta+\beta) = \frac{\mathrm{d}}{\mathrm{d}\theta} \log(Z(\theta+\beta)) = \varrho(\theta) + 1.$$

We conclude that

$$\mu^{\varrho}(z-1) = \frac{\mathrm{e}^{(\theta+\beta)z}}{f(z)!} \cdot \frac{1}{Z(\theta+\beta)} = \mu^{\varrho(\theta+\beta)}(z) = \mu^{\varrho+1}(z),$$

which finishes the proof of the lemma. □

The translation of $\underline{\mu}^{\mathrm{shock}\,\varrho}$ is denoted by

$$\tau_k \underline{\mu}^{\mathrm{shock}\,\varrho} := \bigotimes_{i<k} \mu^{\varrho+1,\varrho+1} \cdot \bigotimes_{i=k} \mu^{\mathrm{shock}\,\varrho} \cdot \bigotimes_{i>k} \mu^{\varrho,\varrho}.$$

The main tool we use is Theorem 1 from [14], which we reformulate here. $\mu S(t)$ will just denote the time evolution of a measure μ under the process dynamics:

Theorem 3.7.8. *In the sense of bounded test functions on $\Omega \times \Omega$,*

$$\begin{aligned}\frac{d}{dt}(\tau_k \underline{\mu}^{shock\,\varrho}) S(t) = \left(e^{\theta(\varrho+1)} - e^{\theta(\varrho)}\right) \cdot (\tau_{k+1}\underline{\mu}^{shock\,\varrho} - \underline{\mu}^{shock\,\varrho}) \\ + \left(e^{-\theta(\varrho)} - e^{-\theta(\varrho+1)}\right) \cdot (\tau_{k-1}\underline{\mu}^{shock\,\varrho} - \underline{\mu}^{shock\,\varrho}).\end{aligned} \tag{3.7.26}$$

The first interesting consequence of this theorem is that the measure $\underline{\mu}^{\mathrm{shock}\,\varrho}$ on a coupled pair evolves into a linear combination of its shifted versions. Second, notice that (3.7.26) is the Kolmogorov equation for an asymmetric simple random walk. Indeed, this theorem implies the following

Corollary 3.7.9. *Let the pair $(\underline{\xi}^-(0), \underline{\xi}(0))$ have initial distribution $\underline{\mu}^{shock\,\varrho}$ defined by* (3.7.24)*. Then its later distribution evolves into a linear combination of translated versions of $\underline{\mu}^{shock\,\varrho}$: at time t the pair $(\underline{\xi}^-(t), \underline{\xi}(0))$ has distribution*

$$\underline{\mu}^{shock\,\varrho} S(t) = \sum_{k=-\infty}^{\infty} P_k(t) \cdot \tau_k \underline{\mu}^{shock\,\varrho},$$

where $P_k(t)$ is the transition probability at time t from the origin to k of a continuous time asymmetric simple random walk with jump rates

$$e^{\theta(\varrho+1)} - e^{\theta(\varrho)} \text{ to the right and } e^{-\theta(\varrho)} - e^{-\theta(\varrho+1)} \text{ to the left.}$$

In particular, $Q^{\xi}(\cdot)$, started from an environment $\underline{\mu}^{shock\,\varrho}$, is a continuous time asymmetric simple random walk with these rates.

Although the corollary is quite natural, let us give a formal proof here. First some notation. $(\underline{\xi}^-(\cdot), \underline{\xi}(\cdot))$ will denote a pair of processes evolving under the basic coupling, g will be a bounded function on the path space of such a pair, and for shortness we introduce Θ_t for the whole random path, shifted to time t: $\Theta_t = (\underline{\xi}^-(t+\cdot), \underline{\xi}(t+\cdot))$. Expectation of the process,

started from $\tau_k \underline{\mu}^{\text{shock}\,\varrho}$, will be denoted by $\mathbf{E}^{(k)}$. Notice that under $\mathbf{E}^{(k)}$ we a.s. have a single position $Q^\xi(t)$ where the coupled pair differ by one, this is the position of the single conserved second class particle. With some abuse of notation we also use $\mathbf{E}^{(\underline{\xi}^-,\underline{\xi})}$ for the evolution of the pair $(\underline{\xi}^-(\cdot), \underline{\xi}(\cdot))$, started from the specific initial state $(\underline{\xi}^-, \underline{\xi})$.

We aim for proving the semigroup property of $S(\cdot)$. The first step is

Lemma 3.7.10. *Given times $0 < s < t$ and $k \in \mathbb{Z}$,*

$$\mathbf{E}^{(0)}\big[g(\Theta_t)\,|\,Q^\xi(s)=k\big] = \mathbf{E}^{(k)}[g(\Theta_{t-s})].$$

Proof. The left hand-side is

$$\begin{aligned}
&\frac{\mathbf{E}^{(0)}\big[g(\Theta_t)\,;\,Q^\xi(s)=k\big]}{\mathbf{P}^{(0)}\{Q^\xi(s)=k\}} \\
&\qquad = \frac{\mathbf{E}^{(0)}\big[\mathbf{E}^{(\underline{\xi}^-(s),\,\underline{\xi}(s))}g(\Theta_{t-s})\,;\,Q^\xi(s)=k\big]}{\mathbf{P}^{(0)}\{Q^\xi(s)=k\}} \\
&\qquad = \frac{\sum_{j\in\mathbb{Z}} \mathbf{P}^{(0)}\{Q^\xi(s)=j\}\mathbf{E}^{(j)}\big[\mathbf{E}^{(\underline{\xi}^-(0),\,\underline{\xi}(0))}g(\Theta_{t-s})\,;\,Q^\xi(0)=k\big]}{\mathbf{P}^{(0)}\{Q^\xi(s)=k\}} \\
&\qquad = \frac{\mathbf{P}^{(0)}\{Q^\xi(s)=k\}\mathbf{E}^{(k)}\big[\mathbf{E}^{(\underline{\xi}^-(0),\,\underline{\xi}(0))}g(\Theta_{t-s})\,;\,Q^\xi(0)=k\big]}{\mathbf{P}^{(0)}\{Q^\xi(s)=k\}} \\
&\qquad = \mathbf{E}^{(k)}[g(\Theta_{t-s})],
\end{aligned}$$

where in the second equality we used that the distribution at time s is a linear combination of shifted versions of $\underline{\mu}^{\text{shock}\,\varrho}$. □

Next we prove the Markov property for $Q^\xi(\cdot)$.

Lemma 3.7.11. *Let $n > 0$ be an integer, φ_i, $i = 0, \dots, n$ bounded functions on $\mathbb{Z}$, and $0 = t_0 < t_1 < \cdots < t_n$. Then*

$$\mathbf{E}^{(0)} \prod_{i=1}^{n} \varphi_i\big(Q^\xi(t_i) - Q^\xi(t_{i-1})\big) = \prod_{i=1}^{n} \mathbf{E}^{(0)}\varphi_i\big(Q^\xi(t_i - t_{i-1})\big).$$

Proof. The statement is trivially true for $n = 1$. We proceed by induction,

and assume the statement is true for $n-1$. Then

$$\begin{aligned}
&\mathbf{E}^{(0)}\prod_{i=1}^{n}\varphi_i\big(Q^\xi(t_i)-Q^\xi(t_{i-1})\big)\\
&=\sum_{j\in\mathbb{Z}}\mathbf{P}^{(0)}\{Q^\xi(t_1)=j\}\varphi_1(j)\cdot\mathbf{E}^{(0)}\Big[\prod_{i=2}^{n}\varphi_i\big(Q^\xi(t_i)-Q^\xi(t_{i-1})\big)\,|\,Q^\xi(t_1)=j\Big]\\
&=\sum_{j\in\mathbb{Z}}\mathbf{P}^{(0)}\{Q^\xi(t_1)=j\}\varphi_1(j)\cdot\mathbf{E}^{(j)}\prod_{i=2}^{n}\varphi_i\big(Q^\xi(t_i-t_1)-Q^\xi(t_{i-1}-t_1)\big)\\
&=\sum_{j\in\mathbb{Z}}\mathbf{P}^{(0)}\{Q^\xi(t_1)=j\}\varphi_1(j)\cdot\mathbf{E}^{(0)}\prod_{i=2}^{n}\varphi_i\big(Q^\xi(t_i-t_1)-Q^\xi(t_{i-1}-t_1)\big)\\
&=\sum_{j\in\mathbb{Z}}\mathbf{P}^{(0)}\{Q^\xi(t_1)=j\}\varphi_1(j)\cdot\prod_{i=2}^{n}\mathbf{E}^{(0)}\varphi_i\big(Q^\xi(t_i-t_{i-1})\big)\\
&=\prod_{i=1}^{n}\mathbf{E}^{(0)}\varphi_i\big(Q^\xi(t_i-t_{i-1})\big).
\end{aligned}$$

The second equality uses Lemma 3.7.10, the third one uses the fact that ϕ's only depend on Q^ξ-differences, and the fourth one follows from the induction hypothesis. □

Proof of Corollary 3.7.9. We know that at any fixed time $t>0$ the distribution of $(\underline{\xi}^-(t),\,\underline{\xi}(t))$ is a linear combination of shifted versions of $\underline{\mu}^{\text{shock}\,\varrho}$. The shift is traced by the second class particle $Q^\xi(t)$, therefore the differential equation

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}^{(0)}\{Q^\xi(t)=k\}\\
&\quad=\big(\mathrm{e}^{\theta(\varrho+1)}-\mathrm{e}^{\theta(\varrho)}\big)\cdot\big(\mathbf{P}^{(0)}\{Q^\xi(t)=k+1\}-\mathbf{P}^{(0)}\{Q^\xi(t)=k\}\big)\\
&\qquad+\big(\mathrm{e}^{-\theta(\varrho)}-\mathrm{e}^{-\theta(\varrho+1)}\big)\cdot\Big(\mathbf{P}^{(0)}\{Q^\xi(t)=k-1\}-\mathbf{P}^{(0)}\{Q^\xi(t)=k\}\Big)
\end{aligned}\tag{3.7.27}$$

follows from (3.7.26). In the above lemmas, we also proved that $Q^\xi(t)$ is Markovian (annealed w.r.t. the initial distribution of $(\underline{\xi}^-,\,\underline{\xi})$). As there exists only one Markovian process with Kolmogorov equation (3.7.27) of the simple asymmetric random walk, we conclude that the process $Q^\xi(\cdot)$ with initial environment $\underline{\mu}^{\text{shock}\,\varrho}$ is an asymmetric simple random walk with rates as stated in the Corollary. □

Lemma 3.7.12. *Let $(\underline{\omega}^-,\,\underline{\omega})$ be a pair of processes in basic coupling, started from distribution* (3.2.18)*, with second class particle $Q(t)$. Then there exist constants $0<\alpha_0$, $C<\infty$ such that*

$$\mathbf{P}\{|Q(t)|>K\}\le e^{-CK}$$

whenever $K > \alpha_0 t$ and t is large enough.

Notice that this implies that Assumption 3.2.2 holds for the TAEBLP.

Proof. The proof uses auxiliary processes to connect the above arguments to the setting of Assumption 3.2.2. Define the pair

$$(\lambda_i,\, \varrho_i) := \begin{cases} (\varrho,\, \varrho+1), & \text{for } i \le 0, \\ (\varrho,\, \varrho), & \text{for } i > 0. \end{cases}$$

Draw the pair $(\underline{\zeta}(0),\, \underline{\xi}(0))$ from the product distribution of coupling measures (3.2.23)

$$\bigotimes_{i\in\mathbb{Z}} \mu^{\lambda_i,\varrho_i}.$$

Then $\underline{\xi}(0)$ has distribution

$$\bigotimes_{i\le 0} \mu^{\varrho+1} \cdot \bigotimes_{i<0} \mu^{\varrho},$$

in agreement with (3.7.25).

Let now the pair $(\underline{\zeta}(\cdot),\, \underline{\xi}(\cdot))$ evolve in the basic coupling, and let them play the role of $(\underline{\eta}(\cdot),\, \underline{\omega}(\cdot))$ of Section 3.7.1. This results in the pair $(\underline{\zeta}(\cdot),\, \underline{\zeta}^{+}(\cdot))$ with a second class particle $Q^{\zeta}(\cdot)$ and the pair $(\underline{\xi}^{-}(\cdot),\, \underline{\xi}(\cdot))$ with a second class particle $Q^{\xi}(\cdot)$ such that $Q^{\zeta}(t) \le Q^{\xi}(t)$, see Lemma 3.7.3. Therefore the random walk result in Corollary 3.7.9 on $Q^{\xi}(\cdot)$ yields the desired estimate for $Q^{\zeta}(t)$. Finally, notice that the distribution of $\underline{\omega}^{-}(0)$ in Assumption 3.2.2 and of $\underline{\zeta}(0)$ above only differ by $\omega_0^{-}(0) \sim \widehat{\mu}^{\varrho}$, while $\zeta_0(0) \sim \mu^{\varrho}$. Therefore

$$\begin{aligned}
\mathbf{P}\{Q(t) > K\} &= \sum_{z=-\infty}^{\infty} \mathbf{P}\{Q(t) > K \,|\, \omega_0^{-}(0) = z\} \cdot \mu^{\varrho}(z)^{\frac{1}{2}} \Big(\frac{\widehat{\mu}^{\varrho}(z)^2}{\mu^{\varrho}(z)}\Big)^{\frac{1}{2}} \\
&= \sum_{z=-\infty}^{\infty} \mathbf{P}\{Q^{\zeta}(t) > K \,|\, \zeta_0(0) = z\} \cdot \mu^{\varrho}(z)^{\frac{1}{2}} \Big(\frac{\widehat{\mu}^{\varrho}(z)^2}{\mu^{\varrho}(z)}\Big)^{\frac{1}{2}} \\
&\le \Big[\sum_{z=-\infty}^{\infty} \mathbf{P}\{Q^{\zeta}(t) > K \,|\, \zeta_0(0) = z\} \cdot \mu^{\varrho}(z)\Big]^{\frac{1}{2}} \cdot \Big[\sum_{y=-\infty}^{\infty} \frac{\widehat{\mu}^{\varrho}(y)^2}{\mu^{\varrho}(y)}\Big]^{\frac{1}{2}} \\
&= \mathbf{P}\{Q^{\zeta}(t) > K\}^{\frac{1}{2}} \cdot \Big[\sum_{y=-\infty}^{\infty} \frac{\widehat{\mu}^{\varrho}(y)^2}{\mu^{\varrho}(y)}\Big]^{\frac{1}{2}}.
\end{aligned}$$

We are done as soon as we show that $\widehat{\mu}^{\varrho}(y)/\mu^{\varrho}(y)$ is uniformly bounded in y. With the exponential rates (3.7.2) one obtains from (3.2.17)

$$\frac{\widehat{\mu}^{\varrho}(y)}{\mu^{\varrho}(y)} = C \sum_{z=y+1}^{\infty} (z-\varrho)\mathrm{e}^{-\frac{\beta}{2}(z-\frac{\theta}{\beta})^2+\frac{\beta}{2}(y-\frac{\theta}{\beta})^2} = C\sum_{k=1}^{\infty}(k+y-\varrho)\mathrm{e}^{-\frac{\beta}{2}k^2-\beta k y+\theta k}.$$

This is uniformly bounded for large y's since then $ye^{-\beta y} < 1$. For large negative y's one uses the equivalent form

$$\widehat{\mu}^{\varrho}(y) := \frac{1}{\mathbf{Var}^{\varrho}(\omega_0)} \sum_{z=-\infty}^{y} (\varrho - z)\mu^{\varrho}(z)$$

of (3.2.17) and writes

$$\frac{\widehat{\mu}^{\varrho}(y)}{\mu^{\varrho}(y)} = C \sum_{z=-\infty}^{y} (\varrho - z)e^{-\frac{\beta}{2}(z-\frac{\theta}{\beta})^2 + \frac{\beta}{2}(y-\frac{\theta}{\beta})^2} = C\sum_{k=1}^{\infty}(k - y + \varrho)e^{-\frac{\beta}{2}k^2 + \beta k y - \theta k}$$

which is again uniformly bounded for large negative y values.

To show a lower bound on $Q(t)$, start with

$$(\lambda_i,\, \varrho_i) := \begin{cases} (\varrho,\, \varrho), & \text{for } i < 0, \\ (\varrho - 1,\, \varrho), & \text{for } i = 0, \\ (\varrho,\, \varrho - 1), & \text{for } i > 0, \end{cases}$$

and the coupled pair $(\underline{\zeta}(0),\, \underline{\xi}(0))$ in distribution

$$\bigotimes_{i\in\mathbb{Z}} \mu^{\lambda_i, \varrho_i}.$$

Now the roles of the pair $(\underline{\zeta}(\cdot),\, \underline{\zeta}^{+}(\cdot))$ with a second class particle $Q^{\zeta}(\cdot)$ and the pair $(\underline{\xi}^{-}(\cdot),\, \underline{\xi}(\cdot))$ with a second class particle $Q^{\xi}(\cdot)$ are interchanged and we have $Q^{\zeta}(t) \geq Q^{\xi}(t)$. The random walk estimate on Q^{ξ} and a Radon-Nikodym estimate similar to the one above completes the proof of the lower bound. □

3.A Monotonicity of measures

In this first section of the Appendix we show that the measures μ^{ϱ} and $\widehat{\mu}^{\varrho}$ defined in (3.2.13) and (3.2.17), respectively, are stochastically monotone as functions of ϱ. We start with a simple

Lemma 3.A.1. *Fix a function $\varphi(\omega)$ on $\mathbb{Z}$, bounded by a polynomial. Then $\mathbf{E}^{\theta}(\varphi(\omega))$ is differentiable in θ on $(\underline{\theta},\, \bar{\theta})$, and*

$$\frac{d}{d\theta}\mathbf{E}^{\theta}(\varphi(\omega)) = \boldsymbol{Cov}^{\theta}(\varphi(\omega),\, \omega).$$

Proof. Convergence of the series involved in $\mathbf{E}^{\theta}(\varphi(\omega))$ can be verified via the ratio test, even after differentiating the terms. Since μ^{θ} is the exponentially

weighted version of μ^{θ_0} for some θ_0, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbf{E}^{\theta}\varphi(\omega)\\
&\quad=\frac{\mathrm{d}}{\mathrm{d}\theta}\,\frac{\mathbf{E}^{\theta_0}(\varphi(\omega)\cdot \mathrm{e}^{(\theta-\theta_0)\omega})}{\mathbf{E}^{\theta_0}\mathrm{e}^{(\theta-\theta_0)\omega}}\\
&\quad=\frac{\mathbf{E}^{\theta_0}(\varphi(\omega)\cdot\omega\cdot \mathrm{e}^{(\theta-\theta_0)\omega})}{\mathbf{E}^{\theta_0}\mathrm{e}^{(\theta-\theta_0)\omega}}-\mathbf{E}^{\theta_0}(\varphi(\omega)\cdot \mathrm{e}^{(\theta-\theta_0)\omega})\cdot\frac{\mathbf{E}^{\theta_0}(\omega\cdot \mathrm{e}^{(\theta-\theta_0)\omega})}{[\mathbf{E}^{\theta_0}\mathrm{e}^{(\theta-\theta_0)\omega}]^2}\\
&\quad=\mathbf{Cov}^{\theta}(\varphi(\omega),\,\omega).\quad\square
\end{aligned}$$

Corollary 3.A.2. *For any $\underline{\theta}<\theta<\bar{\theta}$, the state sum* (3.2.12) *satisfies*

$$\frac{d}{d\theta}\log Z(\theta)=\frac{1}{Z(\theta)}\sum_{z=\omega^{min}}^{\omega^{max}} z\frac{e^{\theta z}}{f(z)!}=\mathbf{E}^{\theta}(\omega)=:\varrho(\theta),\tag{3.A.1}$$

$$\frac{d^2}{d\theta^2}\log Z(\theta)=\frac{d}{d\theta}\varrho(\theta)=\boldsymbol{Var}^{\theta}(\omega).\tag{3.A.2}$$

The function $\varrho(\theta)$ is strictly increasing and maps $(\underline{\theta},\,\bar{\theta})$ onto $(\omega^{min},\,\omega^{max})$.

Proof. Everything is already covered except the last surjectivity statement. Due to the monotonicity and continuity one only needs to show convergence at the boundaries $\underline{\theta}$, $\bar{\theta}$ to $\omega^{\min}$, $\omega^{\max}$. First let us consider the case when $\bar{\theta}<\infty$. Then $\omega^{\max}=\infty$ and Fatou's lemma implies

$$\liminf_{\theta\nearrow\bar{\theta}} Z(\theta)=\liminf_{\theta\nearrow\bar{\theta}}\sum_{z\in I}\frac{\mathrm{e}^{\theta z}}{f(z)!}\geq\sum_{z\in I}\liminf_{\theta\nearrow\bar{\theta}}\frac{\mathrm{e}^{\theta z}}{f(z)!}=\sum_{z\in I}\frac{\mathrm{e}^{\bar{\theta}z}}{f(z)!}=\infty$$

since for $z>0$

$$\frac{\mathrm{e}^{\bar{\theta}z}}{f(z)!}=\prod_{y=1}^{z}\frac{\mathrm{e}^{\bar{\theta}}}{f(y)}\geq 1$$

by definition of $\bar{\theta}$ and f being nondecreasing. This shows that $\log Z(\theta)$ takes on arbitrarily large values as $\theta\nearrow\bar{\theta}$. We also know that it is a smooth and convex function on $(\underline{\theta},\,\bar{\theta})$ (see (3.A.2)). This implies that its derivative (3.A.1) is not bounded from above i.e., arbitrarily large ϱ values can be achieved. The same reasoning works in case $\underline{\theta}>-\infty$ for arbitrarily large negative ϱ values.

When $\bar{\theta}=\infty$ then, regardless whether $\omega^{\max}$ is finite or infinite, fix any $0\leq y<\omega^{\max}$ and write

$$\begin{aligned}
\varrho(\theta)&=\mathbf{E}^{\theta}(\omega\cdot\mathbf{1}\{\omega>y\})+\mathbf{E}^{\theta}([\omega]^{+}\cdot\mathbf{1}\{\omega\leq y\})-\mathbf{E}^{\theta}([\omega]^{-}\cdot\mathbf{1}\{\omega\leq y\})\\
&\geq(y+1)\cdot\mathbf{P}^{\theta}(\omega>y)-\mathbf{E}^{\theta}([\omega]^{-}\cdot\mathbf{1}\{\omega\leq y\})\\
&\geq(y+1)-(y+1)\cdot\mathbf{P}^{\theta}(\omega\leq y)-\sqrt{\mathbf{E}^{\theta}(([\omega]^{-})^2)}\cdot\sqrt{\mathbf{P}^{\theta}(\omega\leq y)}\\
&\geq(y+1)-(y+1)\cdot\mathbf{P}^{\theta}(\omega\leq y)-\sqrt{\mathbf{E}^{\theta_0}(([\omega]^{-})^2)}\cdot\sqrt{\mathbf{P}^{\theta}(\omega\leq y)}
\end{aligned}\tag{3.A.3}$$

for a fixed $\underline{\theta} < \theta_0 < \theta$. The last inequality follows by monotonicity of μ^θ in θ and $([\omega]^-)^2$) being a nonincreasing function of ω. For any $\omega^{\min} - 1 < z \le y$ and $\theta > \underline{\theta}$,

$$\frac{\mu^\theta(z)}{\mu^\theta(y+1)} = \prod_{x=z}^{y} \frac{\mu^\theta(x)}{\mu^\theta(x+1)} = \prod_{x=z}^{y} \frac{f(x+1)}{\mathrm{e}^\theta} \le \Big(\frac{f(y+1)}{\mathrm{e}^\theta}\Big)^{y-z+1}.$$

Given $0 \le y < \omega^{\max}$ and $1 > \varepsilon > 0$, there is a large enough θ which makes the last fraction smaller than ε. With such a choice we have

$$\mathbf{P}^\theta\{\omega \le y\} = \sum_{z=\omega^{\min}}^{y} \mu^\theta(z) \le \mu^\theta(y+1) \sum_{z=\omega^{\min}}^{y} \varepsilon^{y-z+1} \le \varepsilon \cdot \frac{1-\varepsilon^{y-\omega^{\min}+1}}{1-\varepsilon}.$$

Therefore, for the case of a finite $\omega^{\max}$, choosing $y = \omega^{\max} - 1$ and large θ makes (3.A.3) arbitrarily close to $\omega^{\max}$. When $\omega^{\max} = \infty$, the argument shows that $\varrho(\theta) \ge y+1$ can be achieved for any $y \ge 0$. A similar computation demonstrates that any density towards $\omega^{\min}$ can be reached when $\underline{\theta} = -\infty$. □

Corollary 3.A.3. *The measures μ^ϱ are stochastically nondecreasing in ϱ.*

Proof. Since ϱ and θ are strictly increasing functions of each other, it is equivalent to show monotonicity of μ^θ. This follows if we can show $0 \le \frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{E}^\theta(\varphi(\omega))$ for an arbitrary bounded nondecreasing function φ. Lemma 3.A.1 transforms this derivative into the covariance of $\varphi(\omega)$ and ω, which is non-negative due to φ being nondecreasing. □

Monotonicity of $\widehat{\mu}^\varrho$ requires somewhat more of a convexity argument.

Proposition 3.A.4. *The family of measures $\widehat{\mu}^\varrho$, defined in* (3.2.17)*, is stochastically nondecreasing in ϱ.*

Proof. Start by rewriting the definition:

$$\widehat{\mu}^\varrho(y) = \frac{\mathbf{E}^\varrho\big([\omega - \varrho] \cdot \mathbf{1}\{\omega > y\}\big)}{\mathbf{Var}^\varrho(\omega)} = \frac{\mathbf{Cov}^\varrho(\omega,\, \mathbf{1}\{\omega > y\})}{\mathbf{Cov}^\varrho(\omega,\, \omega)}$$
$$= \frac{\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{P}^\theta\{\omega > y\}}{\frac{\mathrm{d}}{\mathrm{d}\theta}\varrho(\theta)}\bigg|_{\theta=\theta(\varrho)} = \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{P}^\varrho\{\omega > y\}.$$

Let us denote the $\widehat{\mu}^\varrho$-expectation by $\widehat{\mathbf{E}}^\varrho$. Fix a bounded nondecreasing function φ. We need to show

$$0 \le \frac{\mathrm{d}}{\mathrm{d}\varrho}\widehat{\mathbf{E}}^\varrho\varphi(\omega).$$

We compute a different expression for this derivative. Passing the derivative through the sum in the third equality below is justified because the series

involved are dominated by certain geometric series, uniformly over θ in small open neighborhoods. This follows from the definitions of $\underline{\theta}$ and $\bar{\theta}$ and the assumption $\underline{\theta} < \theta(\varrho) < \bar{\theta}$.

$$\begin{aligned}
\widehat{\mathbf{E}}^{\varrho}\varphi(\omega) &= \sum_{y=\omega^{\min}}^{\omega^{\max}} \varphi(y) \cdot \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{P}^{\varrho}\{\omega > y\} \\
&= \sum_{y=\omega^{\min}}^{\omega^{\max}} \varphi(y) \cdot \frac{\mathrm{d}}{\mathrm{d}\varrho}[\mathbf{P}^{\varrho}\{\omega > y\} - \mathbf{1}\{0 \geq y\}] \\
&= \frac{\mathrm{d}}{\mathrm{d}\varrho} \sum_{y=\omega^{\min}}^{\omega^{\max}} \varphi(y) \cdot [\mathbf{P}^{\varrho}\{\omega > y\} - \mathbf{1}\{0 \geq y\}] \\
&= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\varrho} \sum_{y=\omega^{\min}}^{\omega^{\max}} \varphi(y) \cdot [\mathbf{1}\{\omega > y\} - \mathbf{1}\{0 \geq y\}] \\
&= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\varrho} \sum_{y=\omega^{\min}}^{\omega^{\max}} \varphi(y) \cdot [\mathbf{1}\{\omega > y > 0\} - \mathbf{1}\{0 \geq y \geq \omega\}] \\
&= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\varrho}\Big[\sum_{y=1}^{\omega-1} \varphi(y) - \sum_{y=\omega}^{0} \varphi(y)\Big] = \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\varrho}\Phi(\omega).
\end{aligned}$$

Above we introduced the function

$$\Phi(x) = \sum_{y=1}^{x-1} \varphi(y) - \sum_{y=x}^{0} \varphi(y),$$

with the convention that empty sums are zero. To conclude the proof, notice that $\Phi(x+1) - \Phi(x) = \varphi(x)$. Thus a nondecreasing function φ determines a (non-strictly) convex function Φ with $\Phi(1) = 0$, and vice-versa. Hence the convexity theorem [20, Theorem 2.1] establishes that

$$\frac{\mathrm{d}}{\mathrm{d}\varrho}\widehat{\mathbf{E}}^{\varrho}\varphi(\omega) = \frac{\mathrm{d}^2}{\mathrm{d}\varrho^2}\mathbf{E}^{\varrho}\Phi(\omega) \geq 0. \qquad \square$$

3.B Regularity properties of the hydrodynamic flux function

For the zero range process defined among the examples in Section 3.2.2, the hydrodynamic (macroscopic) flux function $\mathcal{H} : \mathbb{R}^+ \to \mathbb{R}^+$ of (3.2.14) is given by

$$\mathcal{H}(\varrho) = \mathbf{E}^{\varrho} f(\omega).$$

The results of [20] for f now read as follows:

Proposition 3.B.1. *If the jump rate f of the zero range process is convex (or concave), then the flux $\mathcal{H}$ is also convex (or concave, respectively). Moreover, in this case $\mathcal{H}''(\varrho) > 0$ (or $\mathcal{H}''(\varrho) < 0$, respectively) for all $\varrho > 0$ if and only if f is not a linear function.*

Parts of this proposition were proved with coupling methods in [11].

Next we show in the general case that $\mathcal{H}(\varrho)$ is well defined, and is infinitely differentiable. (We use third derivatives in the proof of Theorem 3.2.3.) The function $\mathcal{H}(\varrho)$ is, in general, the expected net growth rate w.r.t. μ^ϱ as defined in (3.2.14). We show that the series making up this expectation is finite, even after differentiating its terms. This will then lead to smoothness of $\mathcal{H}(\varrho)$.

Lemma 3.B.2. *Let $g(y, z) \geq 0$ be any function on $\mathbb{Z} \times \mathbb{Z}$, bounded by a polynomial in $|y|$ and $|z|$. Then for any $\underline{\theta} < \theta < \bar{\theta}$,*

$$\mathbf{E}^\theta\big[(p(\omega_0, \omega_1) + q(\omega_0, \omega_1))g(\omega_0, \omega_1)\big] < \infty.$$

Proof. We deal with the first part that contains p, the one with q can be treated analogously. The sum we are looking at is

$$\sum_{y=\omega^{\min}+1}^{\omega^{\max}} \sum_{z=\omega^{\min}}^{\omega^{\max}-1} p(y, z) \cdot g(y, z) \cdot \frac{e^{\theta(y+z)}}{f(y)! \cdot f(z)!} \cdot \frac{1}{Z(\theta)^2}.$$

These sums are certainly convergent if $\omega^{\min}$ and $\omega^{\max}$ are both finite. When this is not the case we split both summations at zero, and convergence is established on the four quadrants of the plane. We use (3.2.7) and the corollary

$$p(y, z) = p(z+1, y-1) \cdot \frac{f(y)}{f(z+1)} \quad \text{for } \omega^{\min} < y \leq \omega^{\max},\ \omega^{\min} \leq z < \omega^{\max}$$

of (3.2.9), and we consider empty sums to be zero.

- $y > 0$, $z > 0$: In this case

$$p(y, z) \leq p(y, 0) = p(1, y-1) \cdot \frac{f(y)}{f(1)} \leq p(1, 0) \cdot \frac{f(y)}{f(1)},$$

  and the corresponding part of the summation is bounded by

$$\frac{p(1, 0)}{f(1)} \cdot \sum_{y=1}^{\omega^{\max}} \sum_{z=1}^{\omega^{\max}-1} g(y, z) \cdot \frac{e^{\theta(y+z)}}{f(y-1)! \cdot f(z)!} \cdot \frac{1}{Z(\theta)^2}.$$

- $y \leq 0$, $z > 0$: In this case

$$p(y, z) \leq p(1, 0),$$

and the corresponding part of the summation is bounded by

$$p(1,\,0)\cdot\sum_{y=\omega^{\min}+1}^{0}\sum_{z=1}^{\omega^{\max}-1}g(y,\,z)\cdot\frac{e^{\theta(y+z)}}{f(y)!\cdot f(z)!}\cdot\frac{1}{Z(\theta)^2}.$$

- $y\leq 0$, $z\leq 0$: In this case

$$p(y,\,z)\leq p(1,\,z)=p(z+1,\,0)\cdot\frac{f(1)}{f(z+1)}\leq p(1,\,0)\cdot\frac{f(1)}{f(z+1)},$$

and the corresponding part of the summation is bounded by

$$p(1,\,0)f(1)\cdot\sum_{y=\omega^{\min}+1}^{0}\sum_{z=\omega^{\min}}^{0}g(y,\,z)\cdot\frac{e^{\theta(y+z)}}{f(y)!\cdot f(z+1)!}\cdot\frac{1}{Z(\theta)^2}.$$

- $y>0$, $z\leq 0$: In this case

$$p(y,\,z)=p(z+1,\,y-1)\cdot\frac{f(y)}{f(z+1)}\leq p(1,\,0)\cdot\frac{f(y)}{f(z+1)},$$

and the corresponding part of the summation is bounded by

$$p(1,\,0)\cdot\sum_{y=1}^{\omega^{\max}}\sum_{z=\omega^{\min}}^{0}g(y,\,z)\cdot\frac{e^{\theta(y+z)}}{f(y-1)!\cdot f(z+1)!}\cdot\frac{1}{Z(\theta)^2}.$$

Convergence of each of these bounds for $\underline{\theta}<\theta<\bar{\theta}$ is established e.g. by the ratio test. □

Notice that a similar argument gives finite higher moments of the rates when $\log(f)$ is at most linear in both directions on $\mathbb{Z}$.

Corollary 3.B.3. *$\mathcal{H}(\varrho)$ is infinitely differentiable at all $\varrho\in(\omega^{min},\omega^{max})$.*

Proof. By the previous lemma the series

$$F(\theta):=\mathcal{H}(\varrho(\theta))=\frac{1}{Z(\theta)^2}\cdot\sum_{y,\,z=\omega^{\min}}^{\omega^{\max}}(p(y,\,z)-q(y,\,z))\frac{e^{\theta(y+z)}}{f(y)!\cdot f(z)!},$$

is convergent and infinitely differentiable. Since $\mathcal{H}(\varrho)=F(\theta(\varrho))$ and $\varrho\mapsto\theta(\varrho)$ is infinitely differentiable as well, the claim follows. □

Bibliography

[1] D. Aldous. Random walks on finite groups and rapidly mixing markov chains. *Seminar on probability, XVII*, pages 243–297, 1983.

[2] D. Aldous and P. Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93:243–297, 1986.

[3] David J. Aldous. Some inequalities for reversible markov chains. *Journal of the London Mathematical Society*, s2-25(3):564–576, 1982.

[4] Hamed Amini and Marc Lelarge. The diameter of weighted random graphs. *http://arxiv.org/abs/1112.6330*, 2011.

[5] E. D. Andjel. Invariant measures for the zero range processes. *Ann. Probab.*, 10(3):525–547, 1982.

[6] Ágnes Backhausz and Tamás F. Móri. A random graph model based on 3-interactions. 2011.

[7] C. Bahadoran, H. Guiol, K. Ravishankar, and E. Saada. Euler hydrodynamics of one-dimensional attractive particle systems. *Ann. Probab.*, 34(4):1339–1369, 2006.

[8] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[9] J. Baik and E. M. Rains. Limiting distributions for a polynuclear growth model with external sources. *J. Stat. Phys.*, 100(3-4):523–541, 2000.

[10] M. Balázs. Microscopic shape of shocks in a domain growth model. *J. Stat. Phys.*, 105(3/4):511–524, 2001.

[11] M. Balázs. Growth fluctuations in a class of deposition models. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):639–685, 2003.

[12] M. Balázs. Multiple shocks in bricklayers' model. *J. Stat. Phys.*, 117:77–98, 2004.

[13] M. Balázs, E. Cator, and T. Seppäläinen. Cube root fluctuations for the corner growth model associated to the exclusion process. *Electron. J. Probab.*, 11:no. 42, 1094–1132 (electronic), 2006.

[14] M. Balázs, Gy. Farkas, P. Kovács, and A. Rákos. Random walk of second class particles in product shock measures. *J. Stat. Phys.*, 139(2):252–279, 2010.

[15] M. Balázs and J. Komjáthy. Order of current variance and diffusivity in the rate one totally asymmetric zero range process. *J. Stat. Phys.*, 133(1):59–78, 2008.

[16] M. Balázs, J. Komjáthy, and T. Seppäläinen. Fluctuation bounds in the exponential bricklayers process. *Journal of Statistical Physics*, 147(1):35–62, 2012.

[17] M. Balázs, J. Komjáthy, and T. Seppäläinen. Microscopic concavity and fluctuation bounds in a class of deposition processes. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 48(1):151–187, 2012.

[18] M. Balázs, F. Rassoul-Agha, and T. Seppäläinen. The random average process and random walk in a space-time random environment in one dimension. *Comm. Math. Phys.*, 266(2):499–545, 2006.

[19] M. Balázs, F. Rassoul-Agha, T. Seppäläinen, and S. Sethuraman. Existence of the zero range process and a deposition model with superlinear growth rates. *Ann. Probab.*, 35(4):1201–1249, 2007.

[20] M. Balázs and T. Seppäläinen. A convexity property of expectations under exponential weights. *http://arxiv.org/abs/0707.4273*, 2007.

[21] M. Balázs and T. Seppäläinen. Exact connections between current fluctuations and the second class particle in a class of deposition models. *J. Stat. Phys.*, 127(2):431–455, 2007.

[22] M. Balázs and T. Seppäläinen. Fluctuation bounds for the asymmetric simple exclusion process. *ALEA Lat. Am. J. Probab. Math. Stat.*, VI:1–24, 2009.

[23] M. Balázs and T. Seppäläinen. Order of current variance and diffusivity in the asymmetric simple exclusion process. *Ann. of Math.*, 171(2):1237–1265, 2010.

[24] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[25] A. L. Barabási, E. Ravasz, and T. Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3-4):559 – 564, 2001.

[26] G. Bianconi A. L. Barabasi. Competition and multiscaling in evolving networks, 2000.

[27] Albert-Lszl Barabsi and Rka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[28] Martin T. Barlow, Jian Ding, Asaf Nachmias, and Yuval Peres. The evolution of the cover time. *Combin. Probab. Comput.*, 20(3):331–345, 2011.

[29] L. Barreira. *Dimension and recurrence in hyperbolic dynamics.* Birkhäuser, 1991.

[30] S. Bhamidi. First passage percolation on locally tree-like networks. i. dense random graphs. *Journal of Mathematical Physics*, 49(12):125218, 2008.

[31] S. Bhamidi, R. van der Hofstad, and G. Hooghimstra. First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.*, 20(5):1907–1965, 2010.

[32] S. Bhamidi, R. van der Hofstad, and G. Hooghimstra. First passage percolation on the erds-rnyi random graph. *Combinatorics, Probability and Computing*, 20:683–707, 2011.

[33] Shankar Bhamidi. Universal techniques to analyze preferential attachment trees: Global and local analysis. 2007.

[34] Ginestra Bianconi and Albert-László Barabási. Bose-einstein condensation in complex networks. *Phys. Rev. Lett.*, 86:5632–5635, Jun 2001.

[35] B. Bollobás. *Random Graphs.* Cambridge University Press, 2001.

[36] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[37] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scalefree random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.

[38] Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan. Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '03, pages 132–139, Philadelphia, PA, USA, 2003. Society for Industrial and Applied Mathematics.

[39] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, August 2007.

[40] B. Bollobs, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct Algor*, 31:3–122, 2007.

[41] C. Borgs, J. Chayes, C. Daskalakis, and S. Roch. First to Market is not Everything: an Analysis of Preferential Attachment with Fitness. *ArXiv e-prints*, oct 2007.

[42] A. Borodin, P. L. Ferrari, M. Prähofer, and T. Sasamoto. Fluctuation properties of the TASEP with periodic initial configuration. *J. Stat. Phys.*, 129(5-6):1055–1080, 2007.

[43] M. Brummelhuis and H. Hilhorst. Covering of a finite lattice by a random walk. *Physica A.*, 176:387–408, 1991.

[44] E. Cator and P. Groeneboom. Second class particles and cube root asymptotics for Hammersley's process. *Ann. Probab.*, 34(4):1273–1295, 2006.

[45] D. Champernowne. A model for income distribution. *Economic Journal*, 63:318–351, 1953.

[46] Guan-Yu Chen and Laurent Saloff-Coste. The cutoff phenomenon for ergodic markov processes. 2008.

[47] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math.*, 1:91–113, 2003.

[48] C. Cocozza-Thivent. Processus des misanthropes. *Z. Wahrsch. Verw. Gebiete*, 70(4):509–523, 1985.

[49] J. Fill D. Aldous. *Reversible Markov Chains and Random Walks on Graphs*. University of California, Berkeley, 2002.

[50] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. Math.*, 160(2):433–464, 2004.

[51] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Late points for random walk in two dimensions. *Ann. Probab.*, 34:219–263, 2006.

[52] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[53] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rendiconti di Matematica (7)*, 28:33–61, 2008.

[54] Persi Diaconis and Laurent Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.*, 21(4):2131–2156, 1993.

[55] Persi Diaconis and Laurent Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Probab.*, 16(4):2098–2122, 2006.

[56] J. Ding, Y. Peres, and J. R. Lee. Cover times, blanket times, and majorizing measures. *http://arxiv.org/pdf/1004.4371v5.pdf*, 2011.

[57] Jian Ding, Jeong Han Kim, Eyal Lubetzky, and Yuval Peres. Diameters in supercritical random graphs via first passage percolation. *Combinatorics, Probability and Computing*, 19(Special Issue 5-6):729–751, 2010.

[58] Jian Ding, Eyal Lubetzky, and Yuval Peres. Total variation cutoff in birth-and-death chains. *Probability Theory and Related Fields*, 146:61–85, 2010. 10.1007/s00440-008-0185-3.

[59] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Rev. Mod. Phys.*, 80:1275–1335, Oct 2008.

[60] D. Dürr, S. Goldstein, and J. Lebowitz. Asymptotics of particle trajectories in infinite one-dimensional systems with collisions. *Comm. Pure Appl. Math.*, 38(5):573–597, 1985.

[61] P. Erds and A. Rnyi. On the evolution of random graphs. pages 17–61, 1960.

[62] G. Ergün and G.J. Rodgers. Growing random networks with fitness. *Physica A: Statistical Mechanics and its Applications*, 303(1-2):261 – 272, 2002.

[63] K. Falconer. *The geometry of fractal sets.* Wiley, 1985.

[64] K. Falconer. *Techniques in Fractal Geometry.* Wiley, 1985.

[65] P. A. Ferrari and L. R. G. Fontes. Current fluctuations for the asymmetric simple exclusion process. *Ann. Probab.*, 22(2):820–832, 1994.

[66] P. A. Ferrari and L. R. G. Fontes. Fluctuations of a surface submitted to a random average process. *Electron. J. Probab.*, 3:no. 6, 34 pp. (electronic), 1998.

[67] P. L. Ferrari and H. Spohn. Scaling limit for the space-time covariance of the stationary totally asymmetric simple exclusion process. *Comm. Math. Phys.*, 265(1):1–44, 2006.

[68] J. Gravner, C. A. Tracy, and H. Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Stat. Phys.*, 102(5-6):1085–1132, 2001.

[69] O. Häggström and J. Jonasson. Rates of convergence of lamplighter processes. *Stochastic Processes and their Applications*, 67:227–249, 1997.

[70] Olle Häggström. *Finite Markov Chains and Algorithmic Applications.* Cambridge University Press, 2000.

[71] Remco Van Der Hofstad. *Random Graphs and Complex Networks.* 2010.

[72] J. E. Hutchinson. Fractals and self-similarity. *Indiana University Mathemathics Journal*, 30:713–747, 1981.

[73] Mark Jerrum and Mark Jerrum. Sampling and counting. In Oscar E. Lanford, editor, *Counting, Sampling and Integrating: Algorithm and Complexity*, Lectures in Mathematics ETH Zrich, pages 25–32. Birkhauser Basel, 2003.

[74] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[75] K. Johansson. Discrete polynuclear growth and determinantal processes. *Comm. Math. Phys.*, 242(1-2):277–329, 2003.

[76] S. Jung, S. Kim, and B. Kahng. Geometric fractal growth model for scale-free networks. *Phys. Rev. E*, 65:056101, Apr 2002.

[77] J. Komjathy, J. Miller, and Y. Peres. Uniform mixing time for random walk on lamplighter graphs. *http://arxiv.org/pdf/1109.4281.pdf*, 2011.

[78] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85:4629–4632, Nov 2000.

[79] R. Kumar. Space-time current process for independent random walks in one dimension. *ALEA Lat. Am. J. Probab. Math. Stat.*, IV:307–336, 2008.

[80] Gregory F. Lawler. *Introduction to Stochastic Processes, Second Edition.* Chapman and Hall, CRC Probability Series, 2006.

[81] Gregory F. Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2010.

[82] C. A. Leon and F. Perron. Optimal hoeffding bounds for discrete reversible Markov chains. *Annals of Applied Probability*, 14(2):958–970, 2004.

[83] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.

[84] Nathan Levy. Mixing time for lamplighter graphs. Master's thesis, University of California, Berkeley, 2006.

[85] T. M. Liggett. An infinite particle system with zero range interactions. *Ann. Probab.*, 1:240–253, 1973.

[86] T. M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[87] T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*. Springer-Verlag, 1999.

[88] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.

[89] László Lovász and Peter Winkler. Efficient stopping rules for markov chains. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, STOC '95, pages 76–82, New York, NY, USA, 1995. ACM.

[90] László Lovász and Peter Winkler. Mixing times. *AMS DIMACS SERIES*, 41:189–204, 1998.

[91] Eyal Lubetzky and Allan Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.*, 153(3):475–510, 2010.

[92] J. M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proceedings of the London Mathematical Society*, s3-4(1):257–302, 1954.

[93] J. Miller and Y. Peres. Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *Annals of Probability*, 2011.

[94] R. Montenegro and P. Tetali. Mathematical aspects of mixing times in markov chains. *Found. Trends Theor. Comput. Sci.*, 1(3):237–354, May 2006.

[95] R. I. Oliveira. Mixing and hitting times for finite markov chains. *http://front.math.ucdavis.edu/1108.1708*.

[96] Y. Peres and D. Revelle. Mixing times for random walks on finite lamplighter groups. *Electronic Journal of Probability*, 9:825–845, 2004.

[97] Y. Peres, W. Schlag, and B. Solomyak. Sixty years of bernoulli convolutions. In *Conference poceedings: Fractal Geometry and Stochastics II*, pages 39–65. Birkhauser, 2000.

[98] Y. Peres and P. Sousi. Mixing times are hitting times of large sets. *http://arxiv.org/abs/1108.0133*.

[99] M. Prähofer and H. Spohn. Scale invariance of the PNG droplet and the Airy process. *J. Stat. Phys.*, 108(5-6):1071–1106, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[100] Erzsébet Ravasz and Albert-László Barabási. Hierarchical organization in complex networks. *Phys. Rev. E*, 67:026112, Feb 2003.

[101] F. Rezakhanlou. Hydrodynamic limit for attractive particle systems on $\mathbb{Z}^d$. *Comm. Math. Phys.*, 140(3):417–448, 1991.

[102] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Struct. Alg.*, 31:186–202, 2007.

[103] T. Seppäläinen. Second-order fluctuations and current across characteristic for a one-dimensional growth model of independent random walks. *Ann. Probab.*, 33(2):759–797, 2005.

[104] Herbert A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3–4):425–440, 1955.

[105] F. Spitzer. Interaction of Markov processes. *Advances in Math.*, 5:246–290 (1970), 1970.

[106] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures and Algorithms*, 27:76–123, 2005.

[107] Peter Winkler and David Zuckerman. Multiple cover time. *Random Structures Algorithms*, 9(4):403–411, 1996.

[108] Udny G. Yule. A mathematical theory of evolution, based on the conclusions of dr. j. c. willis, f.r.s. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character:*, 213:21–87, 1925.

[109] Zhongzhi Zhang, Francesc Comellas, Guillaume Fertin, and Lili Rong. High dimensional apollonian networks, 2005.

[110] Zhongzhi Zhang, Jihong Guan, Bailu Ding, Lichao Chen, and Shuigeng Zhou. Contact graphs of disk packings as a model of spatial planar networks. *New Journal of Physics*, 11(8):083007, 2009.

[111] Zhongzhi Zhang, Lili Rong, and Francesc Comellas. High-dimensional random apollonian networks. *Physica A: Statistical Mechanics and its Applications*, 364(0):610 – 618, 2006.

[112] Zhongzhi Zhang, Lili Rong, and Shuigeng Zhou. Evolving apollonian networks with small-world scale-free topologies. *Phys. Rev. E*, 74:046105, Oct 2006.

[113] Zhongzhi Zhang, Shuigeng Zhou, Lujun Fang, Jihong Guan, and Yichao Zhang. Maximal planar scale-free sierpinski networks with small-world effect and power law strength-degree correlation. *EPL (Europhysics Letters)*, 79(3):38007, 2007.

[114] Zhongzhi Zhang, Shuigeng Zhou, Zhan Su, Tao Zou, and Jihong Guan. Random sierpinski network with scale-free small-world and modular structure. *The European Physical Journal B - Condensed Matter and Complex Systems*, 65:141–147, 2008. 10.1140/ epjb/ e2008-00305-8.

[115] Zhongzhi Zhang, Shuigeng Zhou, Wenlei Xie, Lichao Chen, Yuan Lin, and Jihong Guan. Standard random walks and trapping on the koch network with scale-free behavior and small-world effect. *Phys. Rev. E*, 79:061113, Jun 2009.

[116] Zhongzhi Zhang, Shuigeng Zhou, Tao Zou, Lichao Chen, and Jihong Guan. Incompatibility networks as models of scale-free small-world graphs. *The European Physical Journal B - Condensed Matter and Complex Systems*, 60:259–264, 2007. 10.1140/ epjb/ e2007-00344-7.