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

PhD Thesis

Júlia Komjáthy

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

2012
3 Fluctuation bounds in a class of deposition processes

3.1 Introduction ............................................. 106
3.2 Definitions and results ................................. 110
  3.2.1 A family of deposition processes ................. 110
  3.2.2 Examples ........................................ 113
  3.2.3 Basic coupling .................................... 116
  3.2.4 Translation invariant stationary product distributions 117
  3.2.5 Hydrodynamics and some exact identities ........ 118
  3.2.6 Microscopic concavity ............................ 120
  3.2.7 Results .......................................... 123
  3.2.8 Two examples that satisfy microscopic concavity .. 126
3.3 Upper bound of the main theorem .................... 127
3.4 Lower bound of the main theorem ...................... 132
  3.4.1 Perturbing a segment initially .................... 132
  3.4.2 Completion of the proof of the lower bound ....... 135
3.5 Strong Law of Large Numbers for the second class particle 139
3.6 Microscopic concavity for a class of totally asymmetric concave exponential zero range processes .......... 140
3.7 Microscopic convexity of the exponential bricklayers process 146
  3.7.1 Proof of microscopic convexity .................... 152
  3.7.2 A tail bound for the second class particle ....... 159
3.A Monotonicity of measures ............................. 165
3.B Regularity properties of the hydrodynamic flux function 168
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, and special thanks to Gábor Pete for very useful comments on this thesis. 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. These topics are so far away that we do not even try to give a thorough overview of the results and literature in the topics here: for more detailed introduction and for the results see the introduction of each chapter. However, we summarize the content of the different chapters very briefly and without (too many) formulas 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 (some constant, 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 $l_{\infty}$ metric then we call this time the total-variation and the uniform mixing time, respectively. For precise definitions see (1.2.2) and (1.3.2). 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 of the transition matrix of the chain. The spectral gap is the distance between the absolute value of the first and second largest eigenvalue of this matrix. 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. As the size of the system tends to infinity, both the relaxation time and the mixing times are tending to infinity: the question is to determine the rates at which these happen, as a function of the size of the Markov chain.

In the first chapter of the thesis we consider the mixing times of lamp-lighter groups. In the first section of the first chapter we consider the generalized lamplighter walk, i.e. random walk on the wreath product $H \wr G$, the graph whose vertices consist of pairs $(f, x)$ where $f = (f_v)_{v \in V(G)}$ 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 doing simple random walk 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 on the machines at both ends of the edge he passes through, according to the transition rule of the machine. I.e. he can modify the state of the machine both on his departure and also on his arrival vertex by doing one transition on these machines. See Figure 1.1 where the base graph $G$ is a torus and the lamp graphs are cycles. More precisely, the generalized lamplighter chain $X^\circ$ 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, $X^\circ$ moves from a configuration $(f, x)$ 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 ($l_1$) mixing time and the relaxation time of $X^\circ$ in terms of the parameters of $H$ and $G$. 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. This section is based on a paper joint with Yuval Peres [22].

Based on the joint paper with Jason Miller and Yuval Peres [21], in the second section we investigate the uniform mixing time of the “usual” lamplighter walks. Heuristically, a lamplighter walk can be visualized as follows: imagine a random walker walking on the graph $G$, with on-off lamps attached to each vertex of the graph. The walker randomizes the lamps along his path. More precisely, suppose that $G$ is a finite, connected graph and $X$ is a lazy random walk on $G$. The lamplighter chain $X^\circ$ associated with $X$ is the random walk on the wreath product $G^\circ = Z_2 \wr G$, the graph whose vertices consist of pairs $(f, x)$ where $f$ is a labeling of the vertices of $G$ by elements of $Z_2 = \{0, 1\}$ and $x$ is a vertex in $G$. For an example see Figure 1.2 where $G$ is a 5x5 grid and the 0−1 lamps are illustrated as blue and yellow. In each step, $X^\circ$ moves from a configuration $(f, x)$ 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 $Z_2$; $f(z)$ for $z \neq x, y$ remains unchanged. We give matching upper and lower bounds on the uniform mixing time of $X^\circ$ provided $G$ satisfies some sort of local transience criterions. In particular, when $G$ is the hypercube $Z_2^d$, we show that the uniform mixing time of $X^\circ$ is of order $d2^d$. More generally, we show that when $G$ is a torus $Z_n^d$ for $d \geq 3$, the uniform mixing time of $X^\circ$ 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 [35].

In the second chapter we switch to considering mathematical properties of graph and network models. This chapter is based on the paper joint with Károly Simon [126].

Random graphs are in the main stream of research interest since the
late 50s, starting with the seminal random graph model introduced independently by Solomonoff, Rapoport (1951) [79] and by Gilbert (1959) [65], and by Erdős and Rényi (1960) [61] 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 [68] or [50] 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. [42], [50], [53], [60], [71]. Most of them use a version of preferential attachment and are of probabilistic nature. A different approach was initiated by Barabási, Ravasz, and Vicsek [41] 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 exhibits both hierarchical structure and power law decay for the degree sequence; and with a slight modification of the model, the local clustering coefficient is also decaying as in real networks (this is done in [76]). A similar, fractal based deterministic models were introduced by Zhang, Comellas, Fertin and Rong [86], and called the high-dimensional Apollonian networks. The graph sequences are generated from the cylinder sets of the fractal of the Apollonian circle packing or the Sierpiński carpet via slightly different methods in a series of papers. [83] [84] [89] [87].

Motivated by the hierarchical network model of E. Ravasz, A.-L. Barabási and T. Vicsek [41], 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 $G_n$, to be described later, in some way using codes for vertices. The embedding of the adjacency matrix of the graph sequence $G_n$ in the unit square $[0,1]^2$ is carried out in the most straightforward way: a vertex with code $\underline{x} = (x_1 \ldots x_n) \in G_n$ is identified with the corresponding $N$-adic interval $I_{\underline{x}}$, and $\Lambda_n$ 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 $G_n$. The sequence $\Lambda_n$ 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 $\Lambda$ on the plane, see Figure 2.1(c). We discuss connections between the graph theoretical properties of $G_n$ and properties of the limiting fractal $\Lambda$. In particular, we express the power law exponent of the degree distribution with the ratio of the Hausdorff dimensions of some slices of $\Lambda$ (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 [76], we add some additional edges to \( G_1 \) to obtain the (no longer bipartite) graph \( \hat{G}_1 \). Then we build up the graph sequence \( \hat{G}_n \) in a similar manner than it was done for \( G_n \), and show that the average local clustering coefficient of \( \hat{G}_n \) 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 [121, 127].

We consider Markov processes that describe the motion of particles and antiparticles in the one dimensional integer lattice \( \mathbb{Z} \), or equivalently, growth of a surface by depositing or removing individual bricks of unit length and height over \( \mathbb{Z} \). 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 \( \rho \). 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 byproduct, 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., [93, 118, 117, 122, 110], 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 and is a generalization of the one given in [106]. 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 present chapter is based on two papers, both of them joint with Márton Balázs and Timo Seppäläinen. The first one is [101], 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 [100], 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 stochastic processes include an introduction to Markov chains, see for example the book by Lawler [23].

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. Thus, in a metric \( m \) we can define the \( m \)-mixing time of the random walk with transition matrix \( P \) on graph \( G \) as

\[
t^m_{\text{mix}}(G, \varepsilon) := \min \left\{ t \geq 0 : \max_{x \in V(G)} \| P^t(x, \cdot) - \pi(\cdot) \|_m \leq \varepsilon \right\}.
\]

We will study the total variation or TV and the uniform mixing time of the models described below, corresponding to mixing in the \( \ell_1 \) and \( \ell_\infty \) norms. A more algebraic point of view of mixing is to look at the spectral behavior of the transition matrix \( P \). Namely, since \( P \) is a stochastic matrix, 1 is the main eigenvalue and all the other eigenvalues of it lie in the complex unit
disk. If further the chain is reversible, then the eigenvalues are real and it makes sense to define the relaxation time of the chain by

\[ t_{\text{rel}}(G) := \frac{1}{1 - \lambda_2}, \]

where \( \lambda_2 \) is the second largest eigenvalue of the chain. The relation and the ordering between the three quantities can be heuristically understood by the following argument: to see the order of the relaxation time, it is enough to understand how fast the chain "forgets its starting position". The TV-mixing time is related to understand the probabilities of hitting large sets, i.e. those which are at least of constant times the size of the graph \( G \). The uniform mixing time is the hardest to analyze, since for that one has to understand the transition probabilities to a single state more precisely.

In general it is known that for a reversible Markov chain the asymptotic behavior of the relaxation time, the TV and uniform mixing times can significantly differ, i.e. in terms of the size of the graph \( G \) they can have different asymptotics. More precisely, we have

\[ t_{\text{rel}}(G) \leq t_{\text{TV}}^{\text{mix}}(G, 1/4) \leq t_{\text{mix}}^{\text{u}}(G, 1/4), \]

see [3] or [26]. The lamplighter models described below is an example where these three quantities differ.

To understand the behavior of Markov chain sequences, other different notions of mixing times emerged as well, each capturing some different aspect or property of the chain. Aldous [4] introduced random stopping times achieving stationary measure. They were studied more by Lovász, Winkler [28, 29]. (E.g. they studied maximum-length-optimal or expectation-optimal stopping times reaching stationary distribution, strong stationary times and forget times.) To find the relation between different notions of mixing is a challenging problem, see [4] and the recent papers connecting hitting times to mixing times and stopping rules by Sousi and Peres [36] and independently by Oliveira [33], or blanket times and cover times to the maxima of Gaussian free fields by Ding, Lee and Peres [14]. For a more comprehensive overview of Markov Chain mixing we refer the reader to the indispensable book [3] by Aldous and Fill or [26] by Levin, Peres and Wilmer as our main references. See also the books by Häggström [19], Jerrum [20], or the recent survey by Montenegro and Tetali [32].

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 tending to infinity, holds. However, the conjecture fails to be true in the highest generality, see [26] Example 18.7]. Cutoffs are proven recently
for random walks on random regular graphs by Lubetzky and Sly \cite{30} and
for birth and death chains by Ding, Lubetzky and Peres \cite{15}. The cutoff
phenomenon is discussed further in Chen and Saloff-Coste \cite{7} and Diaconis
and Saloff-Coste \cite{13}.

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 a lamp or
machine attached, and each of these identical machines is in some state $f_v(t)$.
Then, as the lamplighter walks along the base graph, he can make changes
in the state of the machines or lamps touched, according to the transition
probabilities of the states 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 the behavior of random walk on groups, analyzed by many authors; we
refer the reader for references on this topic to \cite{2} by Aldous.

To describe the model in a precise way, 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
$X^0$ associated with $X$ and $Z$ is the random walk on the wreath product
$H \wr G$, the graph whose vertices consist of pairs $(f, x)$ where $f = (f_v)_{v \in V(G)}$
is a labeling of the vertices of $G$ by elements of $H$ and $x$ is a vertex in
$G$. In each step, $X^0$ moves from a configuration $(f, x)$ 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.

Relaxation time and TV-mixing on general base graphs $G$ with $\mathbb{Z}_2 =
0 - 1$ lamps was already well-understood, even the constant factor in the
asymptotic behavior, we will give the precise references below. Heuristically
speaking, to get the correct order of the relaxation time of the chain $\mathbb{Z}_2 \wr G$,
one needs to hit far-away vertices on the base graph to be able the "forget
about" the starting position of the chain. Thus, the relaxation time of $\mathbb{Z}_2 \wr G$
is related to the maximal expected hitting time of the graph, $t_{\text{hit}}(G)$ (see
definition \ref{defn:hit_time} below). The total variation mixing of $\mathbb{Z}_2 \wr G$ is understood
by the fact that we want to run the chain until the $0 - 1$ labeling of vertices
becomes indistinguishable from a uniform $0 - 1$ labeling. Thus, the normal
fluctuations of the $0 - 1$ lamps allow us to visit all except $\sqrt{|G|}$ vertices on
the base graph, if these last vertices does not exhibit too much nontrivial
geometric structure. From this heuristics one can see that the TV-mixing
time is related to the asymptotic behavior of the expected cover time of the
base graph $G$ (the expected time it takes the walker to visit every vertex in the graph from a worst case starting position). On the other hand, to understand the behavior of the uniform mixing time of $\mathbb{Z}_2 \wr G$ one needs to understand the exponential moment $\mathbb{E}[2^{U(t)}]$ of the not-yet-visited vertices $U(t)$. One needs to determine the time when this quantity drops below $1 + \varepsilon$, which is much harder to analyze; so it was a gap left between the lower and upper bound on the uniform mixing time for $\mathbb{Z}_2 \wr G$ in \cite{35}.

General lamp graphs $H$ were only considered before in special cases. If the base graph is a complete graph $K_n$, then the lamplighter turns into a "product-chain", which is well understood by being able to construct all the eigenfunctions of $H \wr K_n$ from the eigenfunctions of $H$, see \cite{26}. Nathan Levi \cite{27} in his thesis investigated general lamplighters with $H = \mathbb{Z}_d^2$, the d-dimensional hypercube, but his mixing time bounds did not match in general. Further, Fill and Schoolfield \cite{17} investigated the total variation and $l_2$ mixing time of $K_n \wr S_n$, where the base graph is the Cayley graph of the symmetric group $S_n$ with transpositions chosen as the generator set, and the stationary distribution on $K_n$ is not necessarily uniform.

Thus, now we study uniform mixing with $\mathbb{Z}_2$ lamps, and TV-mixing and relaxation time with general lamps, giving exact results up to constant factors in almost all cases. (The uniform mixing time on general lamp graphs $H$, for the reasons previously mentioned, can be a subject of possible future work.)

In Section 1.2, based on a paper with Yuval Peres \cite{22} 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. Namely, we show that

$$t_{\text{rel}}(H \wr G) \sim t_{\text{hit}}(G) + |G| t_{\text{rel}}(H),$$

where $t_{\text{hit}}(G)$ denotes the maximal expected hitting time of a vertex on $G$.

Further, we give upper and lower bounds on $t_{\text{TV}}(H \wr G, 1/4)$: The order is

$$t_{\text{TV}}^{\text{mix}}(H \wr G, 1/4) \sim t_{\text{cov}}(G) + |G| \cdot f(H, |G|),$$

where $t_{\text{cov}}(G)$ is the cover time of $G$ and $f(H, |G|)$ represents a mixing term on the lamp graph $H$, and equals $t_{\text{mix}}^{\text{TV}}(H, 1/4)$ in the upper bound and $t_{\text{mix}}(H, 1/4) + t_{\text{rel}}(H) \log |G|$ in the lower bound. These two bounds match for the most natural cases, e.g. for $H$ being hypercube, tori, some reversible random walks on the symmetric group or random walk on matrices over the full linear group.

In Section 1.3, based on the joint paper with Miller an Peres \cite{21} we give matching upper bound for the mixing time in the uniform metric of $\mathbb{Z}_2 \wr G$ up to universal constants in terms of the parameters of $G$ to the lower bound given in \cite{35, Theorem 1.4} by Peres and Revelle. We show that

$$t_{u}^{\text{mix}}(\mathbb{Z}_2 \wr G, 1/4) \sim |G|(t_{\text{rel}}(G) + \log |G|)$$
under some conditions which capture the local transience of the base graph $G$. Further we show that these conditions are satisfied by the hypercube $Z^d_2$ or in general the d-dimensional tori $Z^d_n$ with $d$ and $n$ both possibly tending to infinity.

Before we proceed to the particular models, we will now mention some earlier work on mixing times for lamplighter chains. The mixing time of $Z^2_2 \wr G$ was first studied by Häggström and Jonasson in [18] in the case of the complete graph $K_n$ and the one-dimensional cycle $Z_n$. Their work implies a total variation cutoff with threshold $\frac{1}{2}t_{\text{cov}}(K_n)$ in the former case and that there is no cutoff in the latter. Generalizing their results, Peres and Revelle [35, Theorem 1.2, 1.3] proved that there exists constants $c_i, C_i$ depending on $\varepsilon$ such that for any transitive graph $G$,

$$c_1t_{\text{hit}}(G) \leq t_{\text{rel}}(Z^2_2 \wr G) \leq C_1t_{\text{hit}}(G),$$

$$c_2t_{\text{cov}}(G) \leq t_{\text{mix}}(Z^2_2 \wr G, \varepsilon) \leq C_2t_{\text{cov}}(G).$$

The vertex transitivity condition was dropped in [26, Theorem 19.1, 19.2]. These bounds match with our Theorems 1.2.3 and 1.2.4 since $H_n = Z^2_2$ implies that the terms not containing $H_n$ in the denominator of (1.2.6) and in the bounds in (1.2.7) dominate.

Further, [35] also includes a proof of total variation cutoff for $Z^2_2 \wr Z^2_n$ with threshold $t_{\text{cov}}(Z^2_n)$. In [31], it is shown that $t_{\text{mix}}(Z^2_2 \wr Z^d_n) \sim \frac{1}{2}t_{\text{cov}}(Z^d_n)$ when $d \geq 3$ and more generally that $t_{\text{mix}}(Z^2_2 \wr G_n) \sim \frac{1}{2}t_{\text{cov}}(G_n)$ whenever $(G_n)$ is a sequence of graphs satisfying some uniform local transience assumptions. Thus, TV-mixing of $Z_2$ lamps is well understood up to constant.

For the mixing time in the uniform metric, we know [35, Theorem 1.4] that if $G$ is a regular graph such that $t_{\text{hit}}(G) \leq K|G|$, then there exists constants $c, C$ depending only on $K$ such that

$$c|G|(t_{\text{rel}}(G) + \log|G|) \leq t_u(Z^2_2 \wr G) \leq C|G|(t_{\text{mix}}(G) + \log|G|). \quad (1.1.1)$$

These bounds fail to match in general. For example, for the hypercube $Z^d_2$, $t_{\text{rel}}(Z^d_2) = \Theta(d)$ [26, Example 12.15] while $t_{\text{mix}}(Z^d_2) = \Theta(d \log d)$ [26, Theorem 18.3]. Then in the paper [21] we showed that the lower bound is sharp in (1.1.1) under conditions which are satisfied by the $d(n)$ dimension tori $G_n = Z^n_{d(n)}$ for arbitrary chosen $n$ and $d(n)$.

The mixing time of $Z^2_2 \wr G$ is typically dominated by the first coordinate $F$ since the amount of time it takes for $X$ to mix is negligible compared to that required to mix for the labeling. We can sample from $F(t)$ by:

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

Determining the mixing time of $Z^2_2 \wr G$ 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 the lamplighter chain is not mixed. This connects Section 1.3 to the literature on the geometric structure of the last visited points by random walk [9, 8, 6, 31].

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 [26, Theorem 20.7], since in this case the lamplighter chain is a product chain on $\prod_{i=1}^n H_n$. Levi [27] investigated random walks on wreath products when $H \neq \mathbb{Z}_2$. In particular, he determined the order of the mixing time of $K_n \wr K_n, 0 \leq \lambda \leq 1$, and he also had upper and lower bounds for the case $H_d \wr \mathbb{Z}_n$, 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 $H_n = \mathbb{Z}_2$ 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 1.2 to the literature on blanket time (when all the local times of vertices are within a constant factor of each other) [5, 14, 37].

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

1.2.1 The generalized lamplighter model

Let us describe 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 $\mathcal{X}(G) = \{f: V(G) \to H\}$ 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 $(f, x)$ where $f = (f_v)_{v \in V(G)} \in \mathcal{X}(G)$ and $x \in V(G)$. There is an edge between $(f, x)$ and $(g, y)$ 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 $X^\circ$ (with respect to the transition matrices $P$ and $Q$) is the Markov chain on $H \wr G$ which moves from a configuration $(f, x)$ 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 $\pi_G$ and $\pi_H$, respectively, then the unique stationary distribution of $X^\diamond$ is the product measure

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

and $X^\diamond$ is itself reversible. In this article, we will be concerned with the special case that $P$ is the transition matrix for the lazy random walk on $G$. In particular, $P$ is given by

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

for $x, y \in V(G)$ and where $d(x)$ is the degree of $x$. We further assume that the transition matrix $Q$ on $H$ is irreducible and aperiodic. This and the assumption (1.2.1) guarantees that we avoid issues of periodicity.

Figure 1.1: A typical state of the generalized lamplighter walk. Here $H = \mathbb{Z}_4$ and $G = \mathbb{Z}_2^4$; 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.

### 1.2.2 Main Results

In order to state our general result, we first need to review some basic terminology from the theory of Markov chains. Let $P$ be the transition kernel for a lazy random walk on a finite, connected graph $G$ with stationary distribution $\pi$.

The $\varepsilon$-mixing time of $P$ on $G$ in total variation distance is given by

$$t_{\text{mix}}(G, \varepsilon) := \min \left\{ t \geq 0 : \max_{x \in V(G)} \frac{1}{2} \sum_y \left| P^t(x, y) - \pi(y) \right| \leq \varepsilon \right\}. \quad \text{(1.2.2)}$$
Throughout, we set $t_{\text{mix}}(G) := t_{\text{mix}}(G, \frac{1}{4})$.

The relaxation time of a reversible Markov Chain with transition matrix $P$ is

$$t_{\text{rel}}(G) := \frac{1}{1 - \lambda_2} \quad (1.2.3)$$

where $\lambda_2$ is the second largest eigenvalue of $P$.

The maximal hitting time of $P$ is

$$t_{\text{hit}}(G) := \max_{x,y \in V(G)} \mathbb{E}_x[\tau_y], \quad (1.2.4)$$

where $\tau_y$ denotes the first time $t$ that $X(t) = y$ and $\mathbb{E}_x$ stands for the expectation under the law in which $X(0) = x$. The random cover time $\tau_{\text{cov}}$ is the first time when all vertices have been visited by the walker $X$, and the cover time $t_{\text{cov}}(G)$ is

$$t_{\text{cov}}(G) := \max_{x \in V(G)} \mathbb{E}_x[\tau_{\text{cov}}]. \quad (1.2.5)$$

The next needed concept is that of strong stationary times.

**Definition 1.2.1.** A randomized stopping time $\tau$ is called a strong stationary time for the Markov chain $X_t$ on $G$ if

$$P_x[X_\tau = y, \tau = t] = \pi(y)P_x[\tau = t],$$

that is, the position of the walk when it stops at $\tau$ is independent of the value of $\tau$.

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 [26, Section 6.2.2].

**Definition 1.2.2.** A state $h(v) \in V(G)$ is called a halting state for a stopping time $\tau$ and initial state $v \in V(G)$ if $\{X_t = h(v)\}$ implies $\{\tau \leq t\}$.

Our main results are summarized in the following theorems:

**Theorem 1.2.3.** 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 $c_1, C_1$ such that the relaxation time of the generalized lamplighter walk on $H \wr G$ satisfies

$$c_1 \leq \frac{t_{\text{rel}}(H \wr G)}{t_{\text{hit}}(G) + |G|t_{\text{rel}}(H)} \leq C_1, \quad (1.2.6)$$

**Theorem 1.2.4.** Assume that the conditions of Theorem 1.2.3 hold and further assume that the chain with transition matrix $Q$ on $H$ is lazy, i.e.
\[ Q(x,x) \geq \frac{1}{2} \forall x \in H. \] Then there exist universal constants \( c_2, C_2 \) such that the mixing time of the generalized lamplighter walk on \( H \wr G \) satisfies

\[
c_2 (t_{\text{cov}}(G) + |G|(t_{\text{rel}}(H) \log |G| + t_{\text{mix}}(H))) \leq t_{\text{mix}}(H \wr G),
\]

\[
t_{\text{mix}}(H \wr G) \leq C_2 \left( t_{\text{cov}}(G) + |G|t_{\text{mix}}(H, \frac{1}{|G|}) \right).
\] (1.2.7)

If further the Markov chain is such that

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

then the upper bound of 1.2.7 is sharp.

**Remark 1.2.5.** The laziness assumption on the transition matrix \( Q \) on \( H \) is only used to get the term \( c_2 |G| t_{\text{mix}}(H) \) in 1.2.7. All the other bounds hold without the laziness assumption.

**Remark 1.2.6.** If the Markov Chain on \( H \) is such that

\[ t_{\text{mix}}(H, \varepsilon) \leq t_{\text{mix}}(H, 1/4) + t_{\text{rel}}(H) \log \varepsilon, \]

then the upper bound matches the lower bound. This holds for many natural chains such as lazy random walk on hypercube \( Z_d^2 \), tori \( Z_n^d \), some walks on the permutation group \( S_n \) (the random transpositions or random adjacent transpositions shuffle, and the top-to-random shuffle, for instance).

**Remark 1.2.7.** Many examples where Assumption (A) holds are given in the thesis of Pak [34], including the cycle \( Z_n \), the hypercube \( Z_d^2 \) and more generally tori \( Z_n^d \), \( n, d \in \mathbb{N} \) and dihedral groups \( Z_2 \times Z_n \), \( n \in \mathbb{N} \) are also obtained by the construction of strong stationary times with halting states on direct and semidirect product of groups. Further, Pak constructs strong stationary times possessing halting states for the random walk on \( k \)-sets of \( n \)-sets, i.e. on the group \( S_n/(S_k \times S_{n-k}) \), and on subsets of \( n \times n \) matrices over the full linear group, i.e. on \( GL(n, \mathbb{F}_q)/(GL(k, \mathbb{F}_q) \times GL(n-k, \mathbb{F}_q)) \).

**Outline**

The remainder of this article is structured as follows. In Section 1.2.3 we state a few necessary theorems and lemmas about the Dirichlet form, strong stationary times, different notions of distances and their relations. In Lemmas 1.2.10 and 1.2.12 we construct a crucial stopping time \( \tau^5 \) and a strong stationary time \( \tau^5_2 \) 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 1.2.4 and the mixing time bounds in Section 1.2.5.
Notations
Throughout the paper, objects related to the base or the lamp graph will be indexed by $G$ and $H$, respectively, and $\phi$ always refers to an object related to the whole $H \wr G$. Unless misleading, $G$ and $H$ refer also to the vertex set of the graphs, i.e. $v \in G$ means $v \in V(G)$. $P\mu, E\mu$ denote probability and expectation under the conditional law where the initial distribution of the Markov chain under investigation is $\mu$. Similarly, $P_x$ is the law under which the chain starts at $x$.

1.2.3 Preliminaries
In this section 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 jump forward to Lemmas [1.2.10] and [1.2.12] 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 $\pi$ on the state space $\Omega$ and let $E_\pi[\phi] := \sum_{y \in \Omega} \phi(y)\pi(y)$. The Dirichlet form associated to the pair $(P, \pi)$ is defined for functions $\phi$ and $\eta$ on $\Omega$ by

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

It is not hard to see [26, Lemma 13.11] that

$$E(\phi) := E(\phi, \phi) = \frac{1}{2} E_\pi[ (\phi(X_1) - \phi(X_0))^2 ]$$

The next lemma relates the spectral gap of the chain to the Dirichlet form (for a short proof see [3] or [26, Lemma 13.12]):

**Lemma 1.2.8** (Variational characterization of the spectral gap). The spectral gap $\gamma = 1 - \lambda_2$ of a reversible Markov Chain satisfies

$$\gamma = \min_{\phi : \forall \phi \neq 0} \frac{E[\phi]}{\text{Var}_\pi[\phi]},$$

where $\text{Var}_\pi[\phi] = E_\pi[\phi^2] - (E_\pi[\phi])^2$.

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 [1.2.1]. It is not hard to see ([26, Lemma 6.9]) that this is equivalent to

$$P_x[X_t = y, \tau \leq t] = \pi(y)P_x[\tau \leq t].$$

15
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:

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

(1.2.11)

The relation between the separation distance and any strong stationary time \( \tau \) is the following inequality from \cite{3} or \cite[Lemma 6.11]{26}:

\[ \forall x \in \Omega: s_x(t) \leq P_x(\tau > t). \]  

(1.2.12)

Throughout the paper, we will need a slightly stronger result than (1.2.12). Namely, by \cite[Remark 3.39]{11} or from the proof of (1.2.12) in \cite[Lemma 6.11]{26} it follows that it follows that in (1.2.12) equality holds if \( \tau \) has a halting state \( h(x) \) for \( x \). Unfortunately, we just point out that the \cite[Remark 6.12]{26} 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 (1.2.12).

On the other hand, one can always construct \( \tau \) such that (1.2.12) 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 paper).

**Theorem 1.2.9.** \cite{Aldous, Diaconis} \cite{1}, Proposition 3.2] Let \((X_t, t \geq 0)\) be an irreducible aperiodic Markov chain on a finite state space \( \Omega \) with initial state \( x \) and stationary distribution \( \pi \), and let \( s_x(t) \) be the separation distance defined as in (1.2.11). Then

1. if \( \tau \) is a strong stationary time for \( X_t \), then \( s_x(t) \leq P_x(\tau > t) \) for all \( t \geq 0 \).

2. Conversely, there exists a strong stationary time \( \tau \) such that \( s_x(t) = P_x(\tau > t) \) holds with equality.

Combining these, we will call a strong stationary time \( \tau \) separation optimal if it achieves equality in (1.2.12). Mind that every stopping time possessing halting states is separation optimal, but the reversed statement is not necessarily true. The next two lemmas, which we will use several times, construct two stopping times for the graph \( H \uplus G \). 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

\[ L_v(t) = 2 \sum_{i=0}^{t} 1(X_i = v) - \delta_{X_0,v} - \delta_{X_t,v} \]  

(1.2.13)
for the number of moves on the lamp graph \( H_v, v \in G \) by the walker up to time \( t \). Slightly abusing terminology, we call it the local time at vertex \( v \in G \).

Let us further denote the random walk with transition matrix \( Q \) on \( H \) by \( Z \). Since the moves on the different lamp graphs \( H_v, v \in G \) are taken independently given \( L_v(t), v \in G \), we can define for each \( v \in G \) an independent copy of the chain \( Z \), denoted by \( Z_v \), running on \( H_v \). Thus, the position of the lamplighter chain at time \( t \) can be described as

\[
(E_t, X_t) = ((Z_v(L_v(t)))_{v \in G}, X_t)
\]

Below we will use copies of a strong stationary time \( \tau_H \) for the number of moves on the lamp graph \( H \).

This quotient is less than 1 since both the numerator and the denominator are halting states then the vectors \( (Z_v(L_v(t)))_{v \in G} \) are taken independently and the local times \( L_v(t) \), \( \tau_H(v) \)-s are independent of each other.

**Lemma 1.2.10.** Let \( \tau_H \) be any strong stationary time for the Markov chain on \( H \). Take the conditionally independent copies of \( (\tau_H(v))_{v \in G} \) given the local times \( L_v(t) \), realized on the lampgraphs \( H_v \)-s and define the stopping time \( \tau^o \) for \( X^o \) by

\[
\tau^o := \inf \{ t : \forall v \in G : \tau_H(v) \leq L_v(t) \} .
\]

Then, for any starting state \((f_0, x_0)\) we have

\[
P_{(f_0, x_0)}[X_t^o = (f, x), \tau^o = t] = \prod_{v \in G} \pi_H(f_v) \cdot P_{(f_0, x_0)}[X_t = x, \tau^o = t].
\]  

(1.2.15)

If further \( \tau_H \) has halting states then the vectors \( (h(f_v(0)), y) \) are halting state vectors for \( \tau^o \) and initial state \((f_0, x_0)\) for every \( y \in G \).

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

**Corollary 1.2.11.** Let \( \tau_H \) be a strong stationary time for the Markov chain on \( H \) which has a halting state \( h(z) \) for any \( z \in H \). Then define \( \tau^o \) as in Lemma 1.2.10. Then for the separation distance on the lamplighter chain \( H \) \( \tau^o \) the following lower bound holds:

\[
s_{(f_0, x_0)}(t) \geq P_{(f_0, x_0)}[\tau^o > t].
\]

*Proof.* Observe that reaching the halting state vector \((h(f_v(0)), x)\) implies the event \( \tau^o \leq t \) so we have

\[
1 - \frac{P_{(f_0, x_0)}[X_t^o = (h(f_v(0)), x)]}{\pi_G(x) \prod_{v \in G} \pi_H(h(f_v(0)))} = 1 - \frac{P_{(f_0, x_0)}[X_t^o = (h(f_v(0)), x), \tau^o \leq t]}{\pi_G(x) \prod_{v \in G} \pi_H(h(f_v(0)))}
\]

(1.2.16)

Now pick a vertex \( x_{x_0,t} \in G \) which minimizes \( P[X_t = x_{x_0,t} | \tau^o \leq t] / \pi_G(x_{x_0,t}) \).

This quotient is less than 1 since both the numerator and the denominator

17
are probability distributions on $G$. Then, using this and Lemma \[1.2.10\] the right hand side of \[1.2.16\] equals
\[
1 - \frac{P_{(\emptyset, x_0)}[X_t = x_{0:t}, \tau^\circ \leq t]}{\pi_G(x_{0:t})} P_{(\emptyset, x_0)}[\tau^\circ \leq t] \geq 1 - P_{(\emptyset, x_0)}[\tau^\circ \leq t].
\]
Clearly the separation distance is larger than the left hand side of \[1.2.16\], and the proof of the claim follows. Note that the proof only works if $\tau_H$ has a halting state and thus it is separation-optimal.

**Proof of Lemma 1.2.10.** First we show that \[1.2.15\] 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 $P$ on initial state $(\emptyset), x_0)$ for notational convenience. The left hand side of condition \[1.2.10\] equals
\[
P[X_t^\circ = (f, x), \tau^\circ \leq t] = \sum_{X[1:t]} P[X_t^\circ = (f, x), \tau^\circ \leq t|X[1:t]} P[X[1:t]].
\]
Recall that $Z_v$ stands for the Markov chain on the lamp graph $H_v$, and their conditional independence given $L_v(t)$-s. Due to \[1.2.10\] and $\tau_H$ being strong stationary for $H$ we have for all $v \in G$ that
\[
P[Z_v(L_v(t)) = f_v, \tau_H(v) \leq L_v(t)|X[1:t]} = \pi_H(f_v) \cdot P[\tau_H(v) \leq L_v(t)|X[1:t]}.
\]
Now we use that $\tau_H(v)$-s are conditionally independent given the local times to see that
\[
P[X_t^\circ = (f, x), \tau^\circ \leq t|X[1:t]} = \prod_{v \in G} \pi_H(f_v) \prod_{v \in G} P[\tau_H(v) \leq L_v(t)|X[1:t]}
\]
Note that the second product gives exactly $P[\tau^\circ \leq t|X[1:t]}$, yielding
\[
P[X_t^\circ = (f, x), \tau^\circ \leq t] = \prod_{v \in G} \pi_H(f_v) \sum_{X[1:t]} P[\tau^\circ \leq t|X[1:t]} P[X[1:t]] \quad (1.2.17)
\]
As $X_t = x$ remains fixed over the summation, thus summing over all possible $X[1, t]$ trajectories yields
\[
P[X_t^\circ = (f, x), \tau^\circ \leq t] = \prod_{v \in G} \pi_H(f_v) P[X_t = x, \tau^\circ \leq t].
\]
To turn the inequality $\tau^\circ \leq t$ inside the probability to equality can be done the same way as in \[1.2.10\] and left to the reader. To see that the vector
of halting states \((h(f_v(0)), y)\) is a halting state for \(\tau^o\) for any \(y \in G\) is based on the simple fact that reaching the halting state vector \((h(f_v), y)\) 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, \(\tau^o\) has happened as well. 

Recall the definition \((1.2.14)\) of \(\tau^o\). Then we can construct a strong stationary time for \(H \lor G\), described in the next lemma.

**Lemma 1.2.12.** Let \(\tau^o\) be the stopping time defined as in Lemma 1.2.10, and let \(\tau_G(x)\) be a strong stationary time for \(G\) starting from \(x \in G\) and define \(\tau^o_2\) by

\[
\tau^o_2 := \tau^o + \tau_G(X_{\tau^o}),
\]

(1.2.18)

where the chain is re-started at \(\tau_G\) is started from \((E_{\tau}, X_{\tau^o})\), run independently of the past and \(\tau_G\) is measured in this walk. Then, \(\tau^o_2\) is a strong stationary time for \(H \lor G\).

**Proof of Lemma 1.2.12.** The intuitive idea of the proof is based on the fact that \(\tau_G\) is conditionally independent of \(\tau_H\)-s and thus the lamp graphs stay stationary after reaching \(\tau^o\), and stationarity on \(G\) is reached by adding the term \(\tau_G(X_{\tau^o})\). The proof is not very difficult but it needs a delicate sequence of conditioning. To have shorter formulas, we write shortly \(P\) for \(P_{(f_v, x_0)}\). First we condition on the events \(\{\tau^o = s, X^o_s = (g, y)\}\) and make use of \((1.2.15)\) from Lemma 1.2.10

\[
P[X^o_t = (f, x), \tau^o_2 = t] = \sum_{s \leq t/(g, y)} P[X^o_t = (f, x), \tau^o_2 = t|\tau^o = s, X^o_s = (g, y)] 
\cdot \prod_{v \in G} \pi_H(g_v) \cdot P[\tau^o = s, X_s = y].
\]

(1.2.19)

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

\[
P[X^o_t = (f, x), \tau^o_2 = t|\tau^o = s, X^o_s = (g, y)]
= P[X^o_t = (f, x); \tau_G(y) \circ \theta_s = t - s|\tau^o = s, X^o_s = (g, y)]
\]

where \(\tau_G(y) \circ \theta_s\) means the time-shift of \(\tau_G(y)\) by \(s\), and we also used that \(\tau_G\) is only depending on \(y\). We claim that

\[
\sum_v \left( \sum_{s \leq t/(g, y)} P_{(g, y)}[X^o_{t-s} = (f, x), \tau_G(y) = t-s] \prod_{v \in G} \pi_H(g_v) \right)
= P_{y}[X_t = x, \tau_G(y) = t-s] \prod_{v \in G} \pi_H(f_v)
\]

\[
= \pi_G(x)P_{y}[\tau_G = t-s] \prod_{v \in G} \pi_H(f_v).
\]

19
The first equality holds true since \( \tau_G(y) \) is independent of the lampgraphs and the transition rules of \( X^\circ \) 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 \( \tau_G \). Thus, using this and rearranging the order of terms on the right hand side of (1.2.19) we end up with

\[
\sum_{s \leq t, y \in G} P_y[\tau^o = t - s, X_s = y] \cdot \tau_G(x) \prod_{v \in G} \pi_H(f_v).
\]

Then, realizing that the sum is just \( P[\tau^o + \tau_G(X^\circ) = t] \) finishes the proof.

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

\[
d_x(t) := \|P^t(x, \cdot) - \pi(\cdot)\|_{\text{TV}} = \frac{1}{2} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)|. \tag{1.2.20}
\]

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 1.2.13.** For any reversible Markov chain and any state \( x \in \Omega \), the separation distance from initial vertex \( x \) satisfies:

\[
d_x(t) \leq s_x(t) \tag{1.2.21} \\
s_x(2t) \leq 4d(t) \tag{1.2.22}
\]

**Proof.** For a short proof of (1.2.21) see [3] or [26, Lemma 6.13], and combine [26, Lemma 19.3] with a triangle inequality to conclude (1.2.22).

We will also make use of the following lemma: ([26, Corollary 12.6])

**Lemma 1.2.14.** For a reversible, irreducible and aperiodic Markov chain,

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

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

The two fundamental steps to prove Lemma 1.2.14 are the inequalities stating that for all \( x \in \Omega \) we have

\[
d_x(t) \leq s_x(t) \leq \frac{\lambda_1^t}{\pi_{\text{min}}}, \tag{1.2.23}
\]

\[
|\lambda_2|^t \leq 2d(t)
\]

with \( \pi_{\text{min}} = \min_{y \in \Omega} \pi(y) \). This inequality follows from [26, Equation (12.11), (12.13)].

We note that Lemma 1.2.13 implies that the assertion of Lemma 1.2.14 stays valid if we replace \( d(t)^{1/t} \) by the separation distance \( s(t)^{1/t} \).
1.2.4 Relaxation time bounds

Proof of the lower bound of Theorem 1.2.3

We prove \( \frac{c_1}{16 \log 2} \) in the lower bound of the statement of Theorem 1.2.3. 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 \( \lambda_H \) and the corresponding eigenfunction by \( \psi \). It is clear that \( \mathbb{E}_{\pi_H}(\psi) = 0 \) and we can normalize it such that \( \text{Var}_{\pi_H}(\psi) = \mathbb{E}_{\pi_H}(\psi^2) = 1 \) holds. Let us define \( \phi : V(H \wr G) \to \mathbb{R}, \phi((f, x)) = \sum_{w \in G} \psi(f_w) \), thus \( \phi \) is actually not depending on the position of the walker, only on the configuration of the lamps. Let \( X_0 = (F_t, X_t) \) be the lamplighter chain with stationary initial distribution \( \pi_\Diamond \). In the sequel we will calculate the Dirichlet form (1.2.8) for \( \phi \) at time \( t \), first conditioning on the path \( X[0, t] \) of the walker:

\[
\mathcal{E}_t[\phi] = \frac{1}{2} \mathbb{E}_{\pi_\Diamond}[((\phi(X_t^\Diamond)) - \phi(X_0^\Diamond))^2]
= \frac{1}{2} \mathbb{E}_{\pi_\Diamond}(\mathbb{E}_{\pi_\Diamond}([((\phi(X_t^\Diamond)) - \phi(X_0^\Diamond))^2]|X[0, t]))
\] (1.2.24)

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 \( L_v(t) \), the number of steps made by the Markov chain \( Z_v \) on \( H_v \). Moreover, the collection of random walks \( (Z_v)_{v \in G} \) on the lamp graphs are independent given \( L_v(t)-s \).

We can calculate the conditional expectation on the right hand side of (1.2.24) by using the argument above and the fact that \( \mathbb{E}_{\pi_H}(\psi) = 0 \) as follows:

\[
\mathbb{E}_{\pi_\Diamond}[(\phi(X_t^\Diamond)) - \phi(X_0^\Diamond))^2|X[0, t]] = \sum_{v \in G} \mathbb{E}_{\pi_\Diamond}[((\psi(Z_v(L_v(t))) - \psi(Z_v(0)))^2|L_v(t)]
\] (1.2.25)

Next, the product form of the stationary measure \( \pi_\Diamond \) ensures that we can move to \( \pi_H \) inside the sum and condition on the starting state \( Z_v(0) \):

\[
\mathbb{E}_{\pi_\Diamond}[(\psi(Z_v(L_v(t))) - \psi(Z_v(0)))^2|L_v(t)]
= 2\mathbb{E}_{\pi_H}\psi^2 - 2\mathbb{E}_{\pi_H}[\psi(Z_v(0))]\mathbb{E}_{Z_v(0)}[\psi(Z_v(L_v(t)))|Z_v(0)]
\]

Since \( \psi \) was chosen to be the second eigenfunction for \( Q \), clearly \( \mathbb{E}_{Z_v(0)}[\psi(Z_v(L_v(t)))] = \lambda_H^{L_v(t)} \psi(Z_v(0)) \). Using the normalization
\( E_{\pi_H}[\psi^2] = 1 \), we arrive at
\[
E_{\pi_H} \left[ (\phi(X^\ast_t) - \phi(X^0_0))^2 \mid X[0, t] \right] = 2|G| - 2 \sum_{v \in G} \lambda_H^{L_v(t)} \tag{1.2.26}\]

Since \( \sum_{v \in G} L_v(t) = 2t \) and the function \( \lambda_H^t \) is convex, Jensen’s inequality implies that
\[
\sum_{v \in G} \lambda_H^{L_v(t)} \geq |G| \cdot \lambda_H^{2t/|G|}.
\]
Combining this with (1.2.26) and (1.2.24) and setting \( t := t^* = |G|t_{rel}(H) = |G|/(1 - \lambda_H) \) we arrive at
\[
E_t(\phi) \leq |G| \left( 1 - e^{-\frac{2 \log \lambda_H}{1 - \lambda_H}} \right) \leq |G| \left( 1 - 2^{-4} \right),
\]
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 \( t_{rel}(H) < 2 \) and we will use the other lower bound \( t_{hit}(G) \) which is at least of order \(|G|\). Dividing by \( \text{Var}_{\pi^H} \phi = |G| \), and using the variational characterization of the spectral gap in Lemma 1.2.8, we get that the spectral gap \( \gamma_{t^*} \) at time \( t^* \) satisfies
\[
\gamma_{t^*} \leq 1 - 2^{-5}.
\]
Since \( \gamma_t \) is by definition the spectral gap of the chain at time \( t \), we have
\[
1 - \lambda_2(H \wr G)^{t^*} \leq 1 - 2^{-5}. \tag{1.2.27}
\]
Thus
\[
5 \log 2 \geq t^* \left( 1 - \lambda_2(H \wr G) \right),
\]
so we get a lower bound \( t_{rel}(H \wr G) \geq \frac{1}{5 \log 2} |G|t_{rel}(H) \).
To get the lower bound \( t_{hit}(G)/4 \) we adjust the proof for 0–1 lamps (\( H = \mathbb{Z}_2 \)) \cite{26} Theorem 19.1] to our setting. First pick a vertex \( w \in G \) which maximizes the expected hitting time \( E_{\pi_G}(\tau_w) \). As before, we will use the second eigenfunction \( \psi \) with eigenvalue \( \lambda_H \) with \( E_{\pi_H}(\psi) = 0, E_{\pi_H}(\psi^2) = 1 \) and define
\[
\phi((f, x)) := \psi(f_w).
\]
Easy to see with the same conditioning argument we used in (1.2.25) and (1.2.26) that the Dirichlet form at time \( t \) equals
\[
\mathcal{E}_t(\phi) = 1 - E_{\pi_{H^t}} \left[ \lambda_{H^t}^{L_w(t)} \right].
\]
Now we will show that \( E_{\pi_{H^t}} \left[ \lambda_{H^t}^{L_w(t)} \right] \geq 1/4 \). To see this we first note that for any \( t \) we have for the hitting time \( \tau_w \) of \( w \in G \)
\[
E_v(\tau_w) \leq t + t_{hit} P_v[\tau_w > t]
\]
\[
E_{\pi_G}(\tau_w) \leq t + t_{hit} P_{\pi_G}[\tau_w > t]
\]
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 $t_{\text{hit}}$ regardless of where it is at time $t$. The second line follows by averaging over $\pi_G$.

Next, [26, Lemma 10.2] states that $t_{\text{hit}} \leq 2 \max_v \mathbb{E}_{\pi}[\tau_v]$ holds for every irreducible Markov chain. We exactly picked $w$ such that it maximizes $\mathbb{E}_{\pi_G}[\tau_w]$, so we have $t_{\text{hit}} \leq 2 \mathbb{E}_{\pi_G}[\tau_w]$, so multiplying the previous displayed inequality by 2 gives

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

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

$$\mathbb{P}_{\pi_G}[\tau_w > t_{\text{hit}}/4] \geq \frac{1}{4}.$$

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

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

Analogous to the last lines of the proof of the lower bound above, (see (1.2.27)) 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

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

**Proof of the upper bound of Theorem (1.2.3)**

To prove the upper bound, we will estimate the tail behavior of the strong stationary time $\tau^\phi_2$ in Lemma (1.2.12) relate it to $s^\phi(t)$, the separation distance on $H \wr G$, and then use Lemmas (1.2.14) and (1.2.13) to see that $s^\phi(t)^{1/t} \to \lambda^\phi$. We will use separation-optimal $\tau_H$ and $\tau_G$ in the construction of $\tau^\phi_2$. The existence is guaranteed by Theorem (1.2.9). We will use $\mathbf{P}$ for $\mathbf{P}_{(f,x)}$ for notational convenience. Combining (1.2.12) and the fact that $\tau^\phi$ 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 $0 < \alpha < 1$

$$S^\alpha(t) \leq P(f, x)[\tau^\alpha > t] \leq P(f, x)[\tau^\alpha > \alpha t] + P(f, x)[\tau^\alpha > t|\tau^\alpha < \alpha t] \leq P[\tau_{cov} > \alpha t/3]$$  
(1.2.28)

$$+ P[\exists w \in G : L_w(\alpha t) < \alpha t/2|\tau_{cov} \leq \alpha t/3]$$  
(1.2.29)

$$+ P[\exists w \in G : \tau_H(w) > L_w(\alpha t)|\forall v \in G : L(v, \alpha t) \geq \alpha t/2|G]$$  
(1.2.30)

$$+ \max_{(g,y)} P(g, y)[\tau_G > (1 - \alpha)t]$$  
(1.2.31)

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 (1.2.28). The constant $1/3$ could have been chosen differently, we picked $\alpha t/3$ such that the remaining $2\alpha t/3$ time still should be 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 (1.2.29). The third option is that even though there have been 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 (1.2.30). We will handle the three terms separately. The fourth term handles the case where the strong stationary time $\tau_G$ is too large. (For convenience, we will write $t$ instead of $\alpha t$ in estimating the first three formulas.)

We can estimate the first term (1.2.28) by a union bound:

$$P[\tau_{cov} > t/3] \leq P[\exists w : \tau_w > t/3] \leq |G|2e^{\frac{-\log 2}{6} \frac{t}{t_{hit}}},$$  
(1.2.32)

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 $P_v[\tau_w > 2t_{hit}] \leq 1/2$, and then run the chain in blocks of $2t_{hit}$. In each block we hit $w$ with probability at least $1/2$, so we have

$$P_v[\tau_w > K(2t_{hit})] \leq \frac{1}{2K}.$$  

To get it for general $t$, we can move from $\lfloor t/t_{hit} \rfloor$ to $t/t_{hit}$ by adding an extra factor of 2, and (1.2.32) immediately follows by a union bound.

For the third term (1.2.30) we claim the following upper bound holds:

$$P[\exists w : \tau_H(w) \geq L_w(t)|\forall v : L_v(t) > \frac{t}{2|G|}] \leq |G|\frac{1}{\pi_{min}(H)}e^{\frac{-t}{2|G|t_{rel}(H)}}.$$  
(1.2.33)

To see this we estimate the probability of the event $\{\tau_H(w) \geq L_w(t) | L_w(t) \geq \frac{t}{2|G|}\}$ on a single lamp graph and then use a union bound to lose a factor $|G|$.
and arrive at the right hand side. First note that according to Lemma 1.2.14, the tail of the strong stationary time $\tau_H$ is driven by $\lambda_H^t$. More precisely, using the inequality (1.2.23) we have that for any initial state $h \in H$:

$$P_h \left[ \tau_H(w) \geq \frac{t}{2|G|} \right] \leq s_H \left( \frac{t}{2|G|} \right) \leq \frac{1}{\pi_{\text{min}}(H)} \lambda_H^{t/2|G|} \leq \frac{1}{\pi_{\text{min}}(H)} \exp \left\{ -\left(1 - \lambda_H^t \right) \right\}.$$  

Since we have made at least $L_w(t) \geq \frac{t}{2|G|}$ steps on each coordinate, the claim (1.2.33) follows. The fourth term (1.2.31) can be handled analogously and yields an error probability $\exp \left\{ -\frac{ct}{t_{\text{rel}}(G)} \right\}$ which then, taking the power of $1/t$ and limit as in Lemma 1.2.14, will lead to a term of order $t_{\text{rel}}(G)$.

Then, taking into account that $t_{\text{rel}}(G) \leq c t_{\text{mix}}(G) \leq C t_{\text{hit}}(G)$ holds for any lazy reversible chain (see e.g. [26, Chapter 11.6,12.4]), we can ignore this term.

The intuition behind the estimates below for the second term (1.2.29) is that since the total time was at least $2t/3$ after hitting, regularity of $G$ implies that the average number of moves on a lamp graph equals $4t/(3|G|)$ by the double refreshment at any visit to the vertex. Thus, the probability of having less than $t_{\text{rel}}(G) \leq c t_{\text{mix}}(G) \leq C t_{\text{hit}}(G)$ holds for any lazy reversible chain (see e.g. [26, Chapter 11.6,12.4]), we can ignore this term.

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 $i$th 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:

$$\mathbb{P} \left[ \exists w : L_w(t) \leq \frac{t}{2|G|} | \tau_{\text{cov}} \leq \frac{t}{3} \right] \leq |G| \max_{w \in G} \mathbb{P}_w \left[ \sum_{i=1}^{t/4|G|} R_i(w) \geq \frac{2t}{3} \right],$$

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 (1.2.13) of $L_w(t)$, this guarantees that $L_w(t) < t/2|G|$.  

We aim to estimate the right hand side of (1.2.34) using the moment generating function of the first return time $R(w)$. To be able to carry out the estimates we need a bound on the tail behavior of the return times. A very similar argument can be used to the one we used for the tail of the
cover time \textsuperscript{1.2.32}, namely the following holds:

\[ P_w [R(w) > 2t_{hit} + 1] = P_w [X_1 \neq w] E[P_{X_1}(\tau_w > 2t_{hit} | X_1)] \leq \frac{E[E_{X_1}(\tau_w)]}{2t_{hit}} \leq \frac{1}{2}. \]

Running the chains in blocks of \(2t_{hit} + 1\), 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 \(t > 2t_{hit} + 1\)

\[ P_w[R > t] \leq 2 \left( \frac{1}{2} \right)^{\frac{t}{t_{hit}}} = 3 \exp \left\{ -\frac{\log 2}{2} \frac{t}{t_{hit}} \right\}, \tag{1.2.35} \]

where the factor 3 comes from ignoring to take the integer part of \(t/t_{hit}\) and neglecting the +1 term in the denominator.

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

\[ E[e^{\beta R_w}] \leq e^{2\beta |G|} + \int_{e^{2\beta |G|}}^{\infty} P[e^{\beta R_w} > z]dz \]

where we cut the expectation at \(2|G|\). Using the bounds in \textsuperscript{1.2.35} yields:

\[ E[e^{\beta R_w}] \leq e^{2\beta |G|} + \int_{e^{2\beta |G|}}^{\infty} P[R_w > \frac{1}{\beta} \log z]dz \leq e^{2\beta |G|} + 2 \int_{e^{2\beta |G|}}^{\infty} z^{\frac{\log 2}{2t_{hit}}} dz. \]

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

\[ E[e^{\beta R_w}] \leq e^{2\beta |G|} + 2e^{-2\beta |G|} \leq e^{(2+\delta)|G|} \tag{1.2.36} \]

with an appropriately chosen \(0 < \delta < 1/3\). Now we apply Markov’s inequality to the function \(e^{\beta \sum_{i=1}^{t/4|G|} R_i(w)}\) to estimate the right hand side of \textsuperscript{1.2.34}:

\[ P_w \left[ \sum_{i=1}^{t/4|G|} R_i(w) \geq 2t/3 \right] \leq e^{-\frac{2}{3} \beta t} \cdot E[e^{\beta R_w}]^{t/4|G|}, \tag{1.2.37} \]

where we also used the independence of the excursions \(R_i(w)\)-s. Using the estimate in \textsuperscript{1.2.36} to bound the right hand side we gain that

\[ P_w \left[ \sum_{i=1}^{t/4|G|} R_i(w) \geq 2t/3 \right] \leq e^{-\frac{2}{3} \beta t} \cdot e^{(2+\delta)|G| \frac{t}{4|G|}} \leq e^{-\frac{1-\delta}{6} \beta t} = \exp \left\{ -\frac{(1-\delta) \log 2}{24} \frac{t}{t_{hit}} \right\}, \tag{1.2.38} \]

26
where we used $\beta = \log 2/(4t_{\text{hit}})$, and modified $\tilde{\delta} := 3\delta/2 \leq 1/2$. Using the relation of the local time to the excursion lengths in (1.2.34) we finally get that the second term \(\text{(1.2.29)}\) is bounded from above by

$$\mathbb{P} \left[ \exists w : L_w(t) \leq \frac{t}{2|G|} t_{\text{cov}} \leq t/3 \right] \leq |G| \exp \left\{ -\frac{\log 2}{48} \frac{t}{t_{\text{hit}}} \right\}. \quad \text{(1.2.39)}$$

Mind that all the estimates \(\text{(1.2.32)}\), \(\text{(1.2.33)}\) and \(\text{(1.2.39)}\) were independent of the initial state \((f, x) \in H \wr G\), so using the second inequality in \(\text{(1.2.23)}\) and maximizing over all possible initial states yields us

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

\(\text{(1.2.40)}\)

In the final step we apply Lemma \(\text{1.2.14}\) 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 \(\text{(1.2.4)}\) to get an upper bound on \(\lambda_2\). Then we use that \((1 - e^{-x}) \leq x + o(x)\) for small \(x\) and obtain the bound on \(t_{\text{rel}}\) finally:

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

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

\[ \square \]

1.2.5 Mixing time bounds

Based on the fact that \(H\) has a separation-optimal strong stationary time \(\tau_{\text{H}}\), the idea of the proofs is to relate the separation distance to the tail behavior of the stopping times \(\tau^o\) and \(\tau_2^o\) constructed in Lemmas \(\text{1.2.10}\) and \(\text{1.2.12}\) respectively. Then these estimates are turned into bounds of the total variation distance using the relations in Lemma \(\text{1.2.13}\). This method gives us the upper bound in \(\text{(1.2.7)}\) and the corresponding lower bound under the assumption \((A)\). For the lower bound without the assumption, we will need slightly different methods.

Proof of the upper bound of Theorem \(\text{1.2.4}\)

The idea of the proof is to use appropriate top quantiles of the strong stationary time \(\tau_{\text{H}}\) on \(H\), and give an upper bound on the tail of the strong stationary time \(\tau_2^o\) defined in Lemma \(\text{1.2.12}\). Throughout, we (only) need that \(\tau_{\text{H}}\) and \(\tau_2^o\) in the construction of \(\tau_2^o\) are separation-optimal. The existence is guaranteed by Theorem \(\text{1.2.9}\) (Thus, \(\tau_{\text{H}}\) does not necessarily possess halting states.)
Let us denote the worst-case initial state top $\varepsilon$-quantile of a stopping time $\tau$ as

$$t_\varepsilon^{\text{quant}}(\tau) := \max_{y \in \Omega} \inf \{ t : P_y[\tau > t] \leq \varepsilon \} \quad (1.2.41)$$

We continue with the definition of the blanket time:

$$B_2 := \inf_t \left\{ \forall v, w \in G : \frac{L_w(t)}{L_v(t)} \leq 2 \right\}. \quad (1.2.42)$$

Let us further denote

$$B_2 := \max_{v \in G} E_v(B_2) \quad (1.2.43)$$

It is known from [14] that there exist universal constants $C$ and $C'$ such that

$$C' \text{cov} \leq B_2 \leq Ct_{\text{cov}}. \quad (1.2.44)$$

Thus, our first goal is to show that at time $8B_2 + |G|^{t_{\text{quant}}(\tau_H)} + t_{\text{quant}}(\tau_G) = 8B_2 + |G|^{t_H} + t_G =: t^\circ$

we have for any starting state $(f, x)$ that

$$P_{(f, x)}[\tau_2^{\circ} > t^\circ] \leq \frac{1}{4}. \quad (1.2.45)$$

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

$$P[\tau_2^{\circ} > t^\circ] \leq P[B_2 > 8B_2] + P[\tau^\circ > |G|^{t_H} + 8B_2 | B_2 \leq 8B_2] + \max_{v \in G} P_v[\tau_G > t_G | B_2 \leq 8B_2, \tau^\circ < 8B_2 + |G|^{t_H}], \quad (1.2.46)$$

where in the third term we mean that we restart the chain after time $8B_2 + |G|^{t_H}$, and measure $\tau_G$ starting from there. 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 definition of the worst case quantile. The second term can be handled by conditioning on the local time sequence of vertices and on the blanket time: (for shorter notation we introduce $t_1 := |G|^{t_H} + 8B_2$)

$$P[\tau^\circ > |G|^{t_H} + 8B_2 | B_2 \leq 8B_2]$$

$$= \sum_{s \leq 8B_{2,(L_v(t_1))_v}} P[\exists w : \{ \tau_H(w) > L_w(t_1) \} \cap (L_v(t_1))_v, B_2 = s] \cdot P[(L_v(t_1))_v, B_2 = s]. \quad (1.2.47)$$

The fact that $B_2 \leq 8B_2$ means that the number of visits to every vertex $v \in G$ must be greater than half of the average, which is at least $\frac{1}{2}t_H^n$. Since $L_w(t)$ is twice the number of visits by $\{ \tau_H(w) > L_w(t_1) \} \subseteq \{ \tau_H(w) > t_H^n \}$. By the definition of the quantiles,

$$P_h[\tau_H(w) > t_H^n] \leq \frac{1}{16|G|}.$$
holds for every $h \in H$ and $w \in G$. Applying a simple union bound on the conditional probability on the right hand side of (1.2.46) yields

$$
P_{(f,x)} [\tau^\diamond > t_1 | B_2 \leq 8B_2] \leq \sum_{s \leq 8B_2} \left( \frac{|G|}{16|G|} \right) P \left[ (L_v(t_1))_w, B_2 = s \right] \leq \frac{1}{16},$$

where we used that the sum of the probabilities on the right hand side is at most 1. Combining these estimates with (1.2.45) yields (1.2.44). 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 $\tau_H$ and $\tau_G$. Now just consider the walk on $G$. Let us start the walker on $G$ from an initial state $x_0 \in G$ for which the maximum is attained in the definition (1.2.41) of the quantile $t_{1/16}^\text{quant} (\tau_G)$. Then, by (1.2.22) we have that one step before the quantile we have

$$
\frac{1}{16} \leq P_{x_0} \left[ \tau_G > t_{1/16}^\text{quant} (\tau_G) - 1 \right] = \text{s}_{x_0} \left( t_{1/16}^\text{quant} (\tau_G) - 1 \right) \leq 4d \left( \frac{1}{2} t_{1/16}^\text{quant} (\tau_G) - 1 \right).
$$

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

$$
t_{1/16}^\text{quant} (\tau_G) - 1 \leq 12t_{\text{mix}} (G) \tag{1.2.47}
$$

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

$$
\frac{1}{16|G|} \leq P_{h_0} \left[ \tau_H > t_H^u - 1 \right] = \text{s}_{h_0} \left( t_H^u - 1 \right) \leq 4d \left( (t_H^u - 1)/2 \right)
$$

So we have

$$
\frac{1}{2} (t_{1/16|G}(\tau_H) - 1) \leq t_{\text{mix}} \left( H, \frac{1}{16|G|} \right). \tag{1.2.48}
$$

On the other hand, on the whole lamplighter chain $H \wr G$ we need the other direction: For every starting state $(f, x)$ (1.2.21) and (1.2.44) implies that

$$
d_{(f,x)}(t^\diamond) \leq s_{(f,x)}(t^\diamond) \leq P_{(f,x)} [\tau_2^\diamond > t^\diamond] \leq 1/4
$$

Maximizing over all states $(f, x)$ yields

$$
t_{\text{mix}} (H \wr G) \leq t^\diamond. \tag{1.2.49}
$$
Putting the estimates in (1.2.47) and (1.2.48) to (1.2.49), we get that
\[ t_{\text{mix}}(H \ wr G) \leq t^\diamond \leq 8B_2(G) + 12t_{\text{mix}}(G) + 1 + 2|G| \left( t_{\text{mix}}\left( H, \frac{1}{64|G|} \right) + \frac{1}{2} \right). \]

Since \( B_2(G) \leq C_{\text{cov}}(G) \), and \( t_{\text{mix}}(G) \leq 2t_{\text{hit}}(G) \leq 2t_{\text{cov}}(G) \) for any \( G \) (see for instance [26]), the assertion of Theorem 1.2.4 follows with \( C_2 = 8(C + 3) \), where \( C \) is the universal constant relating the blanket time \( B_2(G) \) to the cover time \( t_{\text{cov}}(G) \) in [14].

We remark why we did not make the constant \( C_2 \) explicit: If the blanket time \( B_2(G) \) 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 to the term (1.2.29) behaves like \(|G| e^{-c(t_{\text{cov}}(G) + t_{\text{mix}}(H, \frac{1}{G}))}/t_{\text{hit}}(G)\). If we do not assume anything about the relation of \( t_{\text{hit}}(G) \) and \( t_{\text{cov}}(G) \) 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 \( t_{\text{cov}}(G_n) = t_{\text{hit}}(G_n) = \Theta(n^2) \), 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 bounds of Theorem 1.2.4**

As we did with the relaxation time, it is enough to prove that all the bounds are lower bounds separately, then take an average. First we start showing that the upper bound is sharp in 1.2.7 under the assumption that there is a strong stationary time \( \tau_H \) with halting states.

**Lower bound under Assumption (A)**

We first aim to show that
\[ c|G|t_{\text{mix}}(H, \frac{1}{G}) \leq t_{\text{mix}}(H \ wr G). \]

Consider the stopping time \( \tau^\diamond \) constructed in Lemma 1.2.10. Corollary 1.2.11 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 \( \tau_H \) 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.

First set
\[ t_H^f := t_{\text{quant}}^f(G)^{-1/2}(\tau_H) - 1, \quad t^o := \frac{1}{2}|G|t_H^f, \quad (1.2.50) \]

clearly this time is nontrivial if \( t_{\text{quant}}^f(G)^{-1/2}(\tau_H) \neq 1 \). We handle the case if it equals 1 later. We can estimate the upper tail of \( \tau^o \) by conditioning on the
number of moves on the lamp graphs $H_v, v \in G$:

$$
P \left[ \tau^o > t^o \right] \geq P \left[ \exists w \in G : \tau_H(w) > L_w(t^o) \right] \\
\geq \sum_{(L_v(t^o))_v} P \left[ \exists w \in G : \tau_H(w) > L_w(t^o) \left| (L_v(t^o))_v \right. \right] P \left[ (L_v(t^o))_v \right]
$$

(1.2.51)

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

$$
S_{(L_v)} := \left\{ w \in G : L_w(t^o) \leq t_H^f \right\}
$$

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

$$
|S_{(L_v)}| \geq |G|/2.
$$

(1.2.52)

Thus we can lower bound (1.2.51) by restricting the event only to those $w \in G$ coordinates which belong to this set, i.e. whose local time is small:

$$
P \left[ \tau^o > t^o \right] \geq \sum_{(L_v(t^o))_v} P \left[ \exists w \in S_{(L_v)} : \tau_H(w) > L_w(t^o) \left| (L_v(t^o))_v \right. \right] P \left[ (L_v(t^o))_v \right] \\
\geq \sum_{(L_v(t^o))_v} P \left[ \exists w \in S_{(L_v)} : \tau_H(w) > t_H^f \left| (L_v(t^o))_v \right. \right] P \left[ (L_v(t^o))_v \right],
$$

(1.2.53)

where in the second line we used that for $w \in S_{(L_v)}$, we have $\left\{ \tau_H(w) > L_w(t^o) \right\} \supseteq \left\{ \tau_H(w) > t_H^f \right\}$. Conditioned on the sequence $(L_v(t^o))_v$, the times $\tau_H(w)$ for $w \in S_{(L_v)}$ are independent. 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 $t_{\text{quant}}^f_{|G|^{-1/2}/2}(\tau_H)$. Since $t_H$ is one step before the quantile, we have

$$
P \left[ \tau^o \left| S_{(L_v)} \right. : \tau_H(w) > t_{\text{quant}}^f_{|G|^{-1/2}/2}(\tau_H) - 1 \right] \geq |G|^{-1/2}/2.
$$

(1.2.54)

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 (1.2.41) to be able to relate it to the total variation mixing time on $H$, see below. Then, the fact that $t_{\text{quant}}^f$ 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 (1.2.54) with (1.2.52) and the conditional independence gives us the following stochastic domination from below to the event in (1.2.53)

$$
P \left[ \exists w \in S_{(L_v)} : \tau_H(w) > t_H^f \left| (L_w(t^o))_w \right. \right] \geq P[V > 0],
$$

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

$$
P \left[ V > 0 \right] = 1 - \left( 1 - \frac{1}{2|G|^{1/2}} \right)^{|G|/2} \geq 1 - e^{-|G|^{1/2}}/4.
$$
Combining this with (1.2.53) and summing over all possible \((L_v(t^o))_{v \in G}\) sequences we easily get that

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

Then, by Corollary 1.2.11 we have

\[
s^o(h_0,x)(t^o) \geq 1 - e^{-|G|^{1/2}/4}.
\]

In the next few steps we relate the tail of \(\tau^o\) and \(\tau_H\) to the mixing time of the graphs. First, combining the previous inequality with (1.2.22) implies that for the starting state \((h_0,x)\) the following inequalities hold:

\[
1 - e^{-|G|^{1/2}/4} \leq s^o(h_0,x)(t^o) \leq 4d^o(t^o/2).
\]

These immediately imply

\[
t_{\text{mix}}(H \lvert G, \frac{1}{8}) \geq \frac{1}{2} t^o = \frac{1}{8}|G|^t_H
\]

(1.2.55)

Now we will relate \(t'_H = t_{\text{quant}}^{\tau_H - 1/2/2}(\tau_H) - 1\) to the mixing time on \(H\). Since \(t_H\) investigates the worst case initial-state scenario, by inequality (1.2.12) for any starting state \(h \in H\) we have

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

Using \(d_h(t) \leq s_h(t)\) (see Lemma 1.2.13) and maximizing over all \(h \in H\) we get that

\[
d_H(t_H + 1) \leq |G|^{-1/2}/2.
\]

(1.2.56)

On the other hand, the total variation distance for any Markov chain has the following sub-multiplicative property for any integer \(k\), see [26, Section 4.5]:

\[
d(kt) \leq 2^k d(t)^k.
\]

(1.2.57)

Taking \(t = t_H + 1\) and combining with (1.2.56) we have that

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

which immediately implies

\[
t_{\text{mix}}(H, 1/|G|) \leq 2(t_H + 1).
\]

Combining this with (1.2.55) yields the desired lower bound:

\[
\frac{1}{16} |G| \left( t_{\text{mix}} \left( H, \frac{1}{|G|} \right) - 2 \right) \leq t_{\text{mix}}(H \lvert G, \frac{1}{8}).
\]

32
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 $t_{\text{quant}}^{(\tau_H)} = 1$ can be handled the following way: first mind that we can exchange the quantile for arbitrary $0 < \alpha < 1$, and look at the proof with $t_{\text{quant}}^{(\tau_H)} < \alpha$. If this is still $1$ for all $\alpha$, that means that $\tau_H = 1$ a.s. In this case, it is enough to hit the vertices to mix immediately and thus the mixing time $\left|G\right| t_{\text{mix}}(H)$ is of smaller order than the cover time $t_{\text{cov}}(G)$. The case when $\left|G\right| < 8$ but $\left|H\right| \to \infty$ is easy to see since in this case $t_{\text{mix}}(H, \frac{1}{\left|G\right|}) \leq 2t_{\text{mix}}(H)$ and one can argue that mixing on $H \bowtie G$ requires mixing on a single lamp graph $H_w$ for a fixed $w \in G$. Thus the lower bound remains valid.

The cover time of $G$ is already a lower bound for the $0 - 1$ lamps case by [35], hence also for general lamps, but, for completeness, we adjust the proof in [26, Theorem 19.2] to our setting. By Lemma 1.2.10 we can estimate the separation distance on $H \bowtie G$ as

$$s_{(f,x)}(t) \geq P_{(f,x)}[\tau^f > t]$$

$$\geq P_{(f,x)}[\exists w \in G : \tau_H(w) > L_w(t)]$$

$$\geq P_{(f,x)}[\exists w \in G : L_w(t) = 0] = P_{(f,x)}[\tau_{\text{cov}} > t].$$

Now, using the submultiplicativity of $d(t)$ in (1.2.57) and the relation of the separation distance and the total variation distance in (1.2.22), we have that at time $8t_{\text{mix}}(H \bowtie G, 1/4)$:

$$s_{(f,x)}(8t_{\text{mix}}(H \bowtie G, 1/4)) \leq 4d^f(4t_{\text{mix}}(H \bowtie G, 1/4)) \leq 4 \frac{2^4}{4^4} \leq \frac{1}{4}$$

Combining with (1.2.58) yields that for every starting state we have

$$P_{(f,x)}[\tau_{\text{cov}} > 8t_{\text{mix}}(H \bowtie G, 1/4)] \leq 1/4.$$

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

$$E_{(f,x)}[\tau_{\text{cov}}] \leq 11t_{\text{mix}}(H \bowtie G, 1/4).$$

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

**Proof of the lower bound of Theorem 1.2.4 without assumption (A)**

Now we turn to the general case and first show that $c t_{\text{rel}}(H)|G| \log |G|$ is a lower bound. No laziness assumption on the chain on $H$ is needed to get
this bound. We will use a distinguishing function method. Namely, take an eigenfunction \( \phi_2 \) of the transition matrix \( Q \) on \( H \) corresponding to the second eigenvalue \( \lambda_H \). Then let us define \( \psi : H \leftrightarrow G \rightarrow \mathbb{C} \):

\[
\psi((f, x)) := \sum_{v \in G} \phi_2(f_v). \tag{1.2.59}
\]

One can always normalize such that

\[
\mathbb{E}_\pi^{\psi} = 0 \quad \text{Var}_\pi^{\psi} = |G| \cdot 1
\]

This normalization has two useful consequences: First, by Chebyshev’s inequality, the set \( A = \{ \psi < 2|G|^{1/2} \} \) 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 than \( 1/2 \) at time \( ct_{\text{rel}}(H)|G| \log |G| \) and then we are done by using the following characterization of the total variation distance, see [3, 26]:

\[
\|\nu - \mu\|_{TV} = \sup_{\Delta \subset \Omega} \{ \nu(\Delta) - \mu(\Delta) \}.
\]

Let us start all the lamp graphs from \( g_0 \in H \) where the maximum is attained for \( \phi_2 \). Then we can condition on the local time sequence and use the eigenvalue property of \( \phi_2 \) to obtain

\[
\mathbb{E}_{(g_0, x)}[\psi(F_t, X_t)] = \mathbb{E}[\sum_{w \in G} \phi_2(F_w(t)) (L_w(t))_v]
\]

\[
= \phi_2(g_0) \mathbb{E}_x \left[ \sum_{w \in G} \lambda_H L_w(t) \right]. \tag{1.2.60}
\]

Since \( \sum_v L_v(t) = 2t \), we can apply Jensen’s inequality on the function \( y \rightarrow \lambda_H^y \) to get a lower bound on the expectation:

\[
\mathbb{E}_x \left[ \sum_{w \in G} \lambda_H^{L_w(t)} \right] \geq |G|^{1/2} e^{2t} \left( 1 - \frac{1}{t_{\text{rel}}(H)} \right) \frac{2t}{|G|}. \tag{1.2.61}
\]

By giving a lower bound on the right hand side we must assume here that \( \lambda_H > 0 \), or equivalently \( t_{\text{rel}}(H) > C > 1 \). Thus, first we handle the other case, i.e. when \( t_{\text{rel}}(H) < 2 \). Then the lower bound we are about to show is of order \( |G| \log |G| \) which is always at most the order of \( t_{\text{cov}}(G) \), due to a result by Feige [16] stating that for simple random walk on any connected graph \( G \), \( t_{\text{cov}}(G) \geq (1 + o(1))|G| \log |G| \).

When \( t_{\text{rel}}(H) > 2 \), we can use that \( 1 - x > e^{-1.5x} \) when \( 0 < x < 0.5 \) to get a lower bound on the right hand side of (1.2.61). Then set \( t = ct_{\text{rel}}(H)|G| \log |G| \) turning the estimate in (1.2.60) into

\[
\mathbb{E}_{(g_0, x)}[\psi(F_t, X_t)] \geq |G|^{1-3c} \phi_2(g_0).
\]

34
We can easily upper bound the conditional variance as follows:

\[
\text{Var} \left[ \psi_t \big| (L_v(t))_{v \in G} \right] \leq \sum_{w \in G} \mathbb{E}_{g_0} \left[ \phi_2^2(F_w(t)) | L_w(t) \right] \leq |G| \phi_2^2(g_0).
\]

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

\[
P_{(g_0,x)} \left[ \psi_t \leq 2|G|^{1/2} \right] \leq P_{(g_0,x)} \left[ |\psi_t - \mathbb{E}(\psi_t)| \geq \phi_2(g_0)|G|^{1-3c} - 2|G|^{1/2} \right]
\]

Now we use that \( \phi_2(g_0) > 1 \) and if \( c < 1/6 \) then on the right hand side, the term \( \phi_2(g_0)|G|^{1-3c} \) 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 \( (L_v(t))_{v \in G} \) Chebyshev’s inequality yields:

\[
P_{(g_0,x)} \left[ \psi_t \in A \big| (L_v(t))_{v \in G} \right] \leq \frac{\text{Var} \left[ \psi_t \big| (L_v(t))_{v \in G} \right]}{1/4 \phi_2^2(g_0)|G|^{2-6c}}
\]

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

\[
P_{(g_0,x)} \left[ \psi_t \in A \big| (L_v(t))_{v \in G} \right] \leq \frac{4}{|G|^{1-c}}.
\]

This bound is independent of the local time sequence, so the law of total probability says we have the same upper bound without conditioning on the local times. Now setting \( c < 1/6 \) an \( |G| \) large enough we see that the right hand side can be made smaller than 1/2, finishing the proof.

To see that the cover time is a lower bound in the general case, couple the chain on \( H \bot G \) to \( Z_2 \bot G \), i.e. jump to stationary distribution on \( H_v \) once the walker on the base hits vertex \( v \) and use [35] or [26] to see that \( t_{\text{cov}}(G) \leq t_{\text{mix}}(Z_2 \bot G) \leq t_{\text{mix}}(H \bot G) \).

Next we show that \( c|G|t_{\text{mix}}(H) \) is a lower bound if the chain on \( H \) is lazy.

Let us start with a definition for general Markov chain \( X \) on \( \Omega \)

\[ t_{\text{stop}}(G) := \max_{x \in \Omega} \min \{ \mathbb{E}[\tau] ; \tau \text{ stopping time s.t. } P_x[X_\tau = y] = \pi(y) \ \forall y \in \Omega \}. \]

We call a stopping time mean-optimal if \( \mathbb{E}[\tau] = t_{\text{stop}}(G) \). Lovász and Winkler [28] show that optimal stopping rules always exist for irreducible Markov chains. We aim to show that

\[
\frac{1}{2} |G| \cdot t_{\text{stop}}(H) \leq t_{\text{stop}}(H \bot G).
\]

Take a mean optimal stopping time \( \tau^* \) on \( H \bot G \) reaching minimal expectation, i.e. \( \mathbb{E}_f(f^*[x^*])[\tau^*] = t_{\text{stop}}(H \bot G) \) for some \((f^*, x^*) \in H \bot G \) and \( \mathbb{E}(f^*[x])[\tau^*] \leq t_{\text{stop}}(\bar{H} \bot G) \) for \((f, x) \neq (f^*, x^*)\).
We use this $\tau^*$ to define a stopping rule $\tau_H(v)$ on $H$, for every $v \in G$. Namely, do the following: look at a coordinate $v \in G$ and at the chain restricted to the lamp graph $H_v$, i.e. only the moves which are done on the coordinate $H_v$. Then, stop the chain on $H_v$ when $\tau^*$ stops on the whole $H \wr G$.

Start the chain from any $(f_0, x_0)$. Since $\sum_{v \in G} L_v(t) = 2t$, we have

$$\sum_{v \in G} E_{f_0(0)}[\tau_H(v)] = E_{(f_0, x_0)} \left[ \sum_{v \in G} L_v(\tau^*) \right] = 2 E_{(f_0, x_0)}[\tau^*].$$

Take the vertex $w \in G$ (which can depend on $x_0$), which minimizes the expectation $E_{f_0(0)}[\tau_H(w)]$. Clearly for this vertex the expected value must be less than the average:

$$E_{f_0}[\tau H] \leq \frac{2}{|G|} E_{(f_0, x_0)}[\tau^*].$$

The left hand side is at least as large as what a mean-optimal stopping rule on $H$ can achieve, and the right hand side is at most $\frac{2}{|G|} t_{\text{stop}}(H \wr G)$. Thus we arrive at

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

In the last step we use the equivalence from the paper [36, Corollary 2.5] stating that $t_{\text{stop}}$ and $t_{\text{mix}}$ are equivalent up to universal constants for lazy reversible chains and get that

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

### 1.2.6 Further directions

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 [3, 26]), or to determine the mixing time in the uniform metric. For $\mathbb{Z}_2 \wr G$, already [18] implies a total variation cutoff with threshold $\frac{1}{2} t_{\text{cov}}(K_n)$ for $G$ being the complete graph and that there is no cutoff if $G$ is a cycle of length $n$. The results of [35] include a proof of total variation cutoff for $\mathbb{Z}_2 \wr \mathbb{Z}_n^2$ with threshold $t_{\text{cov}}(\mathbb{Z}_n^2)$. The results in [31] also includes cutoff at $\frac{1}{2} t_{\text{cov}}(G)$, with some uniform local transience assumptions on $G_n$. Further, Levi [27] proved that the wreath product of two complete graphs $K_{n^\lambda} \wr K_n$, $0 \leq \lambda \leq 1$ exhibits a cutoff at $\frac{1+\lambda}{2} n \log n$.

### Acknowledgement

We thank Elisabetta Candellero, Gábor Pete and Tim Kam Wong for useful comments and Louigi Addario-Berry, Perla Sousi and Peter Winkler for a useful discussion of halting states.
1.3 Uniform mixing time for Random Walk on Lamplighter Graphs

1.3.1 The model

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

1. picking $y$ adjacent to $x$ in $G$ according to $P$,
2. updating each of the values of $f_x$ and $f_y$ independently according to the uniform measure on $\mathbb{Z}_2$.

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

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

and $X^\circ$ 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 $G$ in order to avoid issues of periodicity. That is, $P$ is given by

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

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

1.3.2 Main Results

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

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

(1.3.2)

Throughout, we let $t_u(G) = t_u(G, \frac{1}{2 \epsilon})$. The main result of the article [21] is a general theorem which gives matching upper and lower bounds of $t_u(\mathbb{Z}_2 \wr G)$ provided $G$ satisfies several mild hypotheses. One important special case of
Figure 1.2: A typical configuration of the lamplighter over a $5 \times 5$ 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 $\mathbb{Z}_2^d$ and, more generally, tori $\mathbb{Z}_n^d$ for $d \geq 3$. These examples are sufficiently important that we state them as our first theorem in the chapter.

**Theorem 1.3.1.** There exists constants $C_1, C_2 > 0$ such that

$$C_1 \leq \frac{t_u(\mathbb{Z}_2 \wr \mathbb{Z}_2^d)}{d^{2d}} \leq C_2$$

for all $d$.

More generally,

$$C_1 \leq \frac{t_u(\mathbb{Z}_2 \wr \mathbb{Z}_n^d)}{dn^{d+2}} \leq C_2$$

for all $n \geq 2$ and $d \geq 3$.

Prior to this work, the best known bound [35] for $t_u(\mathbb{Z}_2 \wr \mathbb{Z}_2^d)$ was

$$C_1 2^{d}d \leq t_u(\mathbb{Z}_2 \wr \mathbb{Z}_2^d) \leq C_2 2^{d}d \log d$$

for $C_1, C_2 > 0$.

In order to state our general result, we first need to recall some basic terminology from the theory of Markov chains. Recall the notions of relaxation time $t_{\text{rel}}$ (1.2.3), and the maximal hitting time $t_{\text{hit}}$ (1.2.4) from the previous section. Further, we define the Green's function $G(x, y)$ for the random walk $X(t)$ with transition matrix $P$ by

$$G(x, y) = \mathbb{E}_x \left[ \sum_{t=0}^{t_u(\mathcal{G})} \mathbf{1}_{\{X(t)=y\}} \right] = \sum_{t=0}^{t_u(\mathcal{G})} P^t(x, y),$$

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

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

**Assumption 1.3.2.** $G$ is a finite, connected, vertex transitive graph and $X$ is a lazy random walk on $G$. There exist constants $K_1, K_2, K_3 > 0$ such that

1. $t_{hit}(G) \leq K_1|G|$

2. $2K_2(5/2)K_2^2(\max_{y \sim x} G(x, y)K_2 \leq \exp \left( -\frac{t_u(G)}{t_{rel}(G)} \right)$

3. $G^*(n^*) \leq K_3(t_{rel}(G) + \log |G|)/(\log n^*)$

where $n^* = 4K_2t_{u}(G)/\min_{y \sim x} G(x, y)$ for $x, y \in V(G)$ adjacent.

**Remark 1.3.3.** We need to write $\max_{y \sim x} G(x, y)$ in 2 and $\min_{y \sim x} G(x, y)$ in the definition of $n^*$ since the graph $G$ is only assumed to be vertex transitive and not edge-transitive, thus $G(x, y)$ may vary along the neighbors of a vertex $x$.

These assumptions are rather technical: 2 and 3 express some sort of local transience criterion. The gap left by Peres-Revelle in [35] is only there if $t_{rel}(G) = o(t_{mix}(G))$ and $\log |G| = o(t_{mix}(G))$ both hold. Thus, the right hand side of criterion 2 is tending to 0 as $|G| \to \infty$ in cases of our main interest, and the criterion says that the Green function should behave in a similar manner. The criterion connects two error terms later in the proofs. Assumption 3 will serve to be able to establish a Binomial domination argument in the lemmas below for covering the last few left-out vertices, and establishes that the extra time we will give for the Binomial domination is still less than the total time in which we want to mix the chain.

The general theorem is:

**Theorem 1.3.4.** Let $G$ be any graph satisfying Assumption 1.3.2. There exists constants $C_1, C_2$ depending only on $K_1, K_2, K_3$ such that

\[
C_1 \leq \frac{t_u(Z_2 \wr G)}{|G|\left(t_{rel}(G) + \log |G|\right)} \leq C_2
\]  

(1.3.5)

The lower bound is proved in [35, Theorem 1.4]. The proof of the upper bound is based on the observation from [35] that the uniform distance to stationarity can be related to $\mathbf{E}[2^{U(t)}]$ where $U(t)$ is the set of vertices in $G$ 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 \( P_{(f,x)} \) the probability under which \( X^0(0) = (f,x) \), we have that

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

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

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

(1.3.6)

The right hand side is clearly bounded from above by \( E[2^{|\mathcal{U}(t)|}] \) (the initial lamp configuration and position of the lamplighter no longer matters). On the other hand, we can bound (1.3.6) from below by

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

Consequently, to bound \( t_u(\mathbb{Z}_2 \wr \mathcal{G}, \varepsilon) \) it suffices to compute

\[
\min\{t \geq 0 : E[2^{|\mathcal{U}(t)|}] \leq 1 + \varepsilon\}
\]

(1.3.7)

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

In order to establish (1.3.7), 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

\[
L_S(t) = \sum_{s=0}^{t} \mathbf{1}_{\{X(s) \in S\}}
\]

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

**Proposition 1.3.5.** Let \( \lambda_0 \) 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

\[
P_{\pi}\left[ L_S(t) \leq \frac{t \pi(S)}{2} \right] \leq \exp \left( -C_0 t \frac{\pi(S)}{t_{\text{rel}}(\mathcal{G})} \right).
\]

(1.3.8)

Proposition [1.3.5] is a corollary of [25, 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 \( S \). Since the total time of ours is limited by \( C|\mathcal{G}|(t_{\text{rel}}(\mathcal{G}) + \log |\mathcal{G}|) \), these estimates give us a small error probability only as long as the set which is not yet covered is large enough. Thus, we will handle the case when \( \mathcal{U}(t) \) is small separately, via an estimate (Lemma [1.3.10]) of the hitting time \( \tau_S = \min\{t \geq 0 : X(t) \in S\} \) of \( S \).
Outline

The remainder of this section is structured as follows. In Section 1.3.3 we will collect a number of estimates regarding the amount of time $X$ spends in and requires to cover sets of vertices in $G$ of various sizes. Then, in Section 1.3.4 we will complete the proof of Theorem 1.3.4. Finally, in Section 1.3.5 we will give the proof of Theorem 1.3.1 by checking the hypotheses of Theorem 1.3.4.

1.3.3 Coverage Estimates

Throughout, we assume that $G$ is a finite, connected, vertex transitive graph and $X$ is lazy random walk on $G$ with transition matrix $P$ and stationary measure $\pi$. For $S \subseteq V(G)$, we let $C_S(t)$ be the set of vertices in $S$ visited by $X$ by time $t$ and let $U_S(t) = S \setminus C_S(t)$ be the subset of $S$ which $X$ has not visited by time $t$. We let $C(t) = C_V(G)(t)$ and $U(t) = U_V(G)(t)$. We will use $P_x, E_x$ to denote the probability measure and expectation under which $X(0) = x$. Likewise, we let $P_\pi, E_\pi$ 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 $V(G)$. 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 $t_u(G)$ 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 $t_u(G)$ distinct elements of a large set of vertices $S \subseteq V(G)$ is stochastically dominated by a geometric random variable whose parameter depends on $t_u(G)/t_{rel}(G)$. The main result is:

**Proposition 1.3.6.** Assume $X$ satisfies part (2) of Assumption 1.3.2 with constants $K_2$. Let $S \subseteq V(G)$ consist of at least $2K_2t^*/G(x,y)$ elements for $x,y \in V(G)$ adjacent and let

$$t = \frac{2(K_2 + 2)t^*}{\pi(S)}.$$

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

$$P_x[C_S(t) \leq t^*] \leq \exp \left( -C \frac{t^*}{t_{rel}(G)} \right).$$

Recall that

$$C_S(t) = \sum_{s=0}^{t} 1_{\{X(s) \in S\}}.$$
is the amount of time that $X$ spends in $S$ up to time $t$. The proof consists of several steps. The first is Proposition 1.3.5, which we will deduce from [25, Theorem 1] shortly, which gives that the probability $L_S(t)$ is less than $1/2$ its mean is exponentially small in $t$. Once we show that $L_S(t)$ 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 1.3.7. We now proceed to the proof of Proposition 1.3.5.

**Proof of Proposition 1.3.5.** We rewrite the event

$$\left\{ L_S(t) \leq t\frac{\pi(S)}{2} \right\} = \left\{ \sum_{s=0}^{t} f(X_s) > t \left( 1 - \frac{\pi(S)}{2} \right) \right\} \quad (1.3.9)$$

where $f(x) = 1_{S_0}(x)$. Let $\epsilon = \pi(S)/2$ and $\mu = \mathbb{E}_{\pi}[f(X(t))] = 1 - 2\epsilon$. The case $\epsilon \geq 1/4$ follows immediately from [25, Equation 3] in the statement of [25, Theorem 1], so we will only consider the case $\epsilon \in (0, 1/4)$ here. Let $\mu = 1 - \mu = 2\epsilon$. For $x \in (0, 1)$, let

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

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

$$\Delta = 1 + \frac{4\lambda_0 x \bar{x}}{\mu \bar{\pi}(1 - \lambda_0)^2}. \quad (1.3.10)$$

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

$$\frac{1}{(1 - \lambda_0)^2} \leq \Delta \leq \frac{20}{(1 - \lambda_0)^2} \quad (1.3.11)$$

By [25, Theorem 1] and using the representation (1.3.9), we have that

$$\mathbb{P}_\pi [L_S(t) \leq t\epsilon] \leq \exp(-I(\mu + \epsilon)t).$$

Since $I(\mu) = I'(\mu) = 0$ and $I''(x) = (\sqrt{\Delta x \bar{x}})^{-1}$ (see [25, Appendix B]), we can write

$$I(\mu + \epsilon) = \int_{\mu}^{\mu + \epsilon} \int_{\mu}^{x} \frac{1}{\sqrt{\Delta y \bar{y}}} dy dx \quad (1.3.12)$$

where $\bar{y} = 1 - y$. Inserting the bounds from (1.3.11), we thus see that the right side of (1.3.12) admits the lower bound

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

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

42
As in the proof of Lemma 1.3.14, 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 $(\xi_i)$ are iid with $\mathbb{P}[\xi_i = 0] = \mathbb{P}[\xi_i = 1] = \frac{1}{2}$ 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 $L_Y^x(t)$ which appeared in Section 1.3.5). 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 $C_S(t)$ 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 time the non-lazy walk spends in the set $S$. 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 (1.3.3). Since $X$ is a lazy random walk, we also have that

$$G(x,y) \leq G(x,x) \text{ for all } x,y \in V(\mathcal{G}). \quad (1.3.13)$$

This is a consequence of (1.3.34).

Lemma 1.3.7. Fix $S \subseteq V(\mathcal{G})$. For each positive integer $k$ and $t^*$, we have that

$$\mathbb{P}_\pi \left[ C_S(t) \geq \frac{\mathcal{L}_S^Y(t) - t^*}{k} \right] \geq 1 - \frac{t\pi(S)q^k(t)}{t^*}, \quad (1.3.14)$$

where

$$q(t) = G(x,y) + (1 + (2e)^{-1})\frac{t}{|\mathcal{G}|} t_u(\mathcal{G}) 1_{\{t > t_u(\mathcal{G})\}}. \quad (1.3.15)$$

Proof. We want to show that the number of distinct vertices covered is large, thus, we have to exclude cases where the random walk is concentrated on a small fraction of the set $S$. Thus, we will exclude all the points from our counting which are visited more than $k$ times and see that the number of these is still small in expectation.

For $t \geq t_u(\mathcal{G})$, we have $P'_v(x,y) \leq (1 + (2e)^{-1})\pi(y)$ by the definition of $t_u(\mathcal{G})$. Thus by a union bound, to return to $x$ more than once (since the chain $Y$ is non-lazy)

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

Hence by the strong Markov property, the probability that a vertex $x$ is visited more than $k$ times can be bounded from below by

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

Observe

$$\mathbb{P}_x[\tau_x = s] \leq \mathbb{P}_x[Y_s = x] \leq \pi(x). \quad (1.3.16)$$

Let

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

43
be the total time that $Y$ spends at points in $S$ which it visits more than $k$ times by time $N_t$. By (1.3.16), we have that
\[
E_\pi[L_{S,k}^Y(t)] \leq \sum_{x \in S} \sum_{s=0}^{t} P_\pi[\tau_x = s]q^k(t) \leq t\pi(S)q^k(t).
\]
Applying Markov’s inequality we have that
\[
P_\pi[L_{S,k}^Y(t) \geq t^*] \leq \frac{E_\pi[L_{S,k}^Y(t)]}{t^*} \leq \frac{t\pi(S)q^k(t)}{t^*}.
\]
Observe
\[
C_S(t) = \sum_{x \in S} 1_{\{L_Y^x(t) \geq 1\}} \geq \sum_{x \in S} (1_{\{L_Y^x(t) \geq 1\}} - 1_{\{L_Y^x(t) > k\}}) \geq \frac{L_Y^S(t) - L_{S,k}^Y(t)}{k}.
\]
Thus
\[
\{L_{S,k}^Y(t) < t^*\} \subseteq \left\{C_S(t) \geq \frac{L_Y^S(t) - t^*}{k}\right\}.
\]
We arrive at
\[
P_\pi[C_S(t) \geq \frac{L_Y^S(t) - t^*}{k}] \geq 1 - P_\pi[L_{S,k}^Y(t) \geq t^*] \geq 1 - \frac{t\pi(S)q^k(t)}{t^*},
\]
which completes the proof of the lemma.

Proposition 1.3.5 gives a lower bound on the probability $L_S(t)$ is proportionally lower than its expectation, Lemma 1.3.7 gives a lower bound on the probability $X$ visits less than a positive fraction of $L_Y^S(t) - t^*$ vertices in $S$ by time $t$, and standard large deviations estimates bound the probability that $L_Y^S(t)$ is proportionally smaller than $L_S(t)$. 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 1.3.8. Fix $S \subseteq V(G)$. Then
\[
P_\pi[C_S(t) \leq \frac{t\pi(S) - 8t^*}{8k}] \leq \exp\left(-C_0t\frac{\pi(S)}{t_{rel}(G)}\right) + \exp\left(-\frac{1}{16}t\pi(S)\right) + \frac{t\pi(S)q^k(t)}{t^*},
\]
where the constant $C_0$ is as in Proposition 1.3.3 and the function $q$ is as in (1.3.15).
Proof. On the event \( \{ \mathcal{L}_S^Y(t) > \frac{t\pi(S)}{8} \} \) we have \( \{ C_S(t) \leq \frac{t\pi(S) - 8t^*}{8k} \} \subseteq \{ C_S(t) \leq \frac{\mathcal{L}_S^Y(t) - t^*}{k} \} \). Thus, we can write the following union bound:

\[
P_\pi \left[ C_S(t) \leq \frac{t\pi(S) - 8t^*}{8k} \right] \leq P_\pi \left[ \mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8} \right] + P_\pi \left[ C_S(t) \leq \frac{\mathcal{L}_S^Y(t) - t^*}{k} \right].
\]

We can bound the second term from above by Lemma 1.3.7. For the first term we move back to the lazy chain to get the following bound from above

\[
P_\pi \left[ \mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8} \right] \leq P_\pi \left[ \mathcal{L}_S(t) \leq \frac{t\pi(S)}{2} \right] + P_\pi \left[ C_S(t) \leq \mathcal{L}_S^Y(t) \leq \frac{t\pi(S)}{8} \right].
\]

We can bound the first term using Proposition 1.3.5. Conditionally on \( \{ \mathcal{L}_S(t) > \frac{S}{8} \pi(S) \} \), we note that \( \{ \mathcal{L}_S^Y(t) \leq \frac{S}{8} \pi(S) \} \) occurs if \( X \) stays in place for at least \( \frac{3}{8} \pi(S) \) 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 1.3.6 by ignoring the first \( t_u(G) \) units of time in order to reduce to the stationary case, then apply Assumption 1.3.2 in order to match the error terms in Lemma 1.3.8.

Proof of Proposition 1.3.6. We first observe that

\[
P_x [ C_S(t) \leq t^* ] \leq (1 + (2e)^{-1}) P_\pi [ C_S(t - t_u(G)) ] \leq t^*.\]

With \( \tilde{t} = 2K_2t^*/\pi(S) \) and using \( |S| \geq 2K_2t^*/\min_{y \sim x} G(x, y) \) for \( x, y \in V(G) \) adjacent, we see that

\[
G(x, y) \leq q(\tilde{t}) \leq \frac{5}{2} G(x, y).
\]

Combining this with part [2] of Assumption 1.3.2 implies

\[
\frac{\tilde{t}\pi(S)q^{K_2}(\tilde{t})}{t_u(G)} \leq 2K_2q^{K_2}(\tilde{t}) \leq \exp \left( -\frac{t^*}{t_{rel}(G)} \right). \tag{1.3.18}
\]

Applying Lemma 1.3.8 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 \( S \subseteq V(G) \), provided \( |S| \) is sufficiently small.
Proposition 1.3.9. 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

$$\mathbb{P}_x \left[ \mathcal{L}_S(C_2|\mathcal{G}^*(s)) \leq \frac{s^2}{2} \right] \leq \exp(-C_3s)$$

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

The main step in the proof of Proposition 1.3.9 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 $Z1_{\{\tau_S > t\}} = 0$ and $\mathbb{E}_x[Z1_{\{\tau_S \leq t\}}] > 0$. Then we have that

$$\mathbb{P}_x[\tau_S \leq t] = \frac{\mathbb{E}_x[Z]}{\mathbb{E}_x[Z1_{\{\tau_S \leq t\}}]}.$$  \hspace{1cm} (1.3.19)

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

Lemma 1.3.10. 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

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

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

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

We can bound the numerator from below as follows:

$$\mathbb{E}_x[\mathcal{L}_S \left( \frac{|\mathcal{G}|}{s} \right)] \geq (1 - (2e)^{-1})\mathbb{E}_\pi \left[ \mathcal{L}_S \left( \frac{|\mathcal{G}|}{s} - t_u(\mathcal{G}) \right) \right]$$

$$\geq (1 - (2e)^{-1})\pi(S) \left( \frac{|\mathcal{G}|}{s} - t_u(\mathcal{G}) \right) \geq \frac{1}{4}. \hspace{1cm} (1.3.20)$$

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

$$\mathbb{E}_x[\mathcal{L}_S(\tau_S, \tau_S + t_u(\mathcal{G})) |E] + \mathbb{E}_x[\mathcal{L}_S(\tau_S + t_u(\mathcal{G}) + 1, \frac{|\mathcal{G}|}{s}) |E]$$

$$=: D_1 + D_2.$$
We have
\[ D_2 \leq (1 + (2e)^{-1}) \mathbb{E}_\pi \left[ \mathcal{L}_S \left( \frac{|G|}{s} \right) \right] \leq 2. \]

We will now bound \( D_1 \). By the strong Markov property, we have that
\[
D_1 \leq \max_{z \in S} \mathbb{E}_z [\mathcal{L}_S(t_u(G))] = \max_{z \in S} \mathbb{E}_z \sum_{t=0}^{t_u(G)} 1_{\{X(t) \in S\}} = \max_{z \in S} \sum_{y \in S} G(z, y) \leq G^*(s).
\]

Putting everything together completes the proof. \( \square \)

The remainder of the proof of Proposition 1.3.9 is based on a simple stochastic domination argument.

**Proof of Proposition 1.3.9.** 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 \( C_2|G|/G^*(s) \) 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 \( C_2|G|/G^*(s) \) steps, each of length \( |G|/s \). We let \( S_0 = S \) and inductively let \( S_i = S_{i-1} \setminus \{x\} \) if \( X \) hits \( x \) in the \( i \)th round for \( i \geq 1 \). If \( |S_i| \geq s/2 \), the hypotheses of Lemma 1.3.10 hold. In this case, the probability that \( X \) hits a point in \( S_i \) in the \( i \)th round is at least \( \rho_0/G^*(s) > 0 \). Thus by stochastic domination, we have that
\[
\mathbb{P} \left[ C_S(C_2|G|/G^*(s)) < s/2 \right] \leq \mathbb{P} \left[ Z < s/2 \right]
\]
where \( Z \sim \text{BIN}(C_2G^*(s)s, \rho_0/G^*(s)) \). By picking \( C_2 \) large enough (\( C_2 > 1/\rho_0 \) will do, say) and applying the Chernoff bound, we see that
\[
\mathbb{P} \left[ C_S(C_2|G|/G^*(s)) < s/2 \right] \leq \exp(-C_3s) \tag{1.3.21}
\]
for some constant \( C_3 \) (one can check that \( C_3 = \frac{1}{4} \) suffices). This estimate also holds if \( s = 1 \). In this case we cover the point with constant probability in \( C_2|G| \) steps. \( \square \)

**1.3.4 Proof of Theorem 1.3.4**

Throughout this section, we shall assume that \( X \) is a lazy random walk on a graph \( G \) which satisfies Assumption 1.3.2. Recall that \( \mathcal{U}(t) \) is the set of vertices of \( G \) which \( X \) has not visited by time \( t \). We will use the notation \( \mathbb{P}_x, \mathbb{E}_x \) for the probability measure and expectation under which \( X(0) = x \). Likewise, we let \( \mathbb{P}_\pi, \mathbb{E}_\pi \) correspond to the case that \( X \) is initialized at stationarity. We will now work towards completing the proof of Theorem 1.3.4 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 \(U(t)\) contains at least \(n^*\) vertices (recall the definition of \(n^*\) from part (3) of Assumption 1.3.2). To this end, we let

\[
\begin{align*}
    r &= \max\{i : |\mathcal{G}| - it_u(\mathcal{G}) \geq n^*\}, \\
    \tilde{r} &= \lfloor \log_2(|\mathcal{G}| - rt_u(\mathcal{G})) \rfloor
\end{align*}
\]

and

\[
\begin{align*}
    s_i &= |\mathcal{G}| - it_u(\mathcal{G}), \quad i = 0, \ldots, r, \\
    s_{r+i} &= \left\lfloor \frac{s_i}{2^i} \right\rfloor \quad i = 1, \ldots, \tilde{r} - 1, \\
    s_{r+\tilde{r}} &= 0.
\end{align*}
\]

We also define the stopping times

\[
T_i = \min\{t \geq 1 : |U(t)| \leq s_i\}, \quad i = 1, \ldots, r + \tilde{r}.
\]

**Lemma 1.3.11.** 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

\[
P_x[|U(t)| > s_i] \leq \exp\left(\frac{s_i}{t_{rel}(\mathcal{G})} \left( C_4 \log |\mathcal{G}| - C_5 \frac{t}{|\mathcal{G}|} \right) \right).
\]

**(1.3.22)**

**Proof.** For each \(i \in \{1, \ldots, r\}\), we let

\[
t_i = \frac{2(K_2 + 2)t_u(\mathcal{G})|\mathcal{G}|}{s_i}
\]

Proposition 1.3.6 with the choice \(t^* = t_u(\mathcal{G})\) implies that

\[
P_x[|U(t)| \leq s_{i+1} \mid |U(t)| \in (s_{i+1}, s_i)] \geq 1 - \exp\left(-C \frac{t_u(\mathcal{G})}{t_{rel}(\mathcal{G})} \right).
\]

Consequently, it follows that there exists independent variables \(Z_j \sim \text{GEO}(1 - \exp(-Ct_u(\mathcal{G})/t_{rel}(\mathcal{G})))\) such that \(T_j - T_{j-1}\) is stochastically dominated by \(t_j Z_j\) for all \(j \in \{1, \ldots, r\}\). Thus for \(\theta_i > 0\), we have that

\[
P_x[|U(t)| > s_i] = P_x[T_i > t] = P_x \left[ \sum_{j=1}^{i} T_j - T_{j-1} > t \right]
\]

\[
\leq e^{-\theta_i t} \prod_{j=1}^{i} E_x[e^{\theta_j t_j Z_j}].
\]

**(1.3.23)**

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

\[
\frac{pe^x}{1 - (1 - p)e^x} \leq e^{\alpha x} \text{ provided } (1 - p)e^x \leq \beta.
\]

**(1.3.24)**
Choosing $$\theta_i = \frac{C_t G}{2_t \text{rel}(G)}$$ we have that $$\theta_i t_j = \frac{C_t G}{t_{\text{rel}}(G)} \cdot \frac{t_j}{t_i} = \frac{C_t G}{2_t \text{rel}(G)} \cdot \frac{s_i}{s_j}.$$ Hence as $$s_i \leq s_j$$ for all $$i, j \in \{1, \ldots, r\}$$ with $$j \leq i$$, we have

$$\exp \left( \frac{C_t G}{2_t \text{rel}(G)} \cdot \frac{s_i}{s_j} \right) \leq \exp \left( \frac{-C_t G}{2_t \text{rel}(G)} \right) \leq \exp(-C/2).$$

Let $$\alpha = \alpha(e^{-C/2})$$ as in (1.3.24). Consequently, we can bound the product of exponential moments in (1.3.23) by

$$\log \prod_{j=1}^{i} E_x[ e^{\theta_i t_j Z_j} ] \leq \alpha \sum_{j=1}^{i} \theta_i t_j = \frac{\alpha C_t G s_i}{2_t \text{rel}(G)} \sum_{j=1}^{i} \frac{1}{s_j} \leq \frac{\alpha C s_i}{2_t \text{rel}(G)} \log |G|.$$

Inserting this expression into (1.3.23) gives (1.3.22).

Lemma 1.3.12. There exists constants $$C_6, C_7$$ such that for all $$1 \leq i \leq \tilde{r}$$ and $$x \in V(G)$$, we have that

$$P_x[|U(t)| > s_{r+i}] \leq P_x[|U(t/2)| > s_r] + \exp \left( s_{r+i-1} \left( C_6 i - \frac{C_7}{|G| G^*(n^*)} t \right) \right).$$

Proof. Let

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

where $$C_2$$ is as in Proposition 1.3.9. Proposition 1.3.9 implies that

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

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

$$P_x[|U(t)| > s_{r+i}] = P_x[T_{r+i} > t]$$

$$\leq P_x \left[ T_r > \frac{t}{2} \right] + P_x \left[ \sum_{j=1}^{i} T_{r+j} - T_{r+j-1} > \frac{t}{2} \right] =: I_1 + I_2$$

(1.3.26)
Using that $I_1 = \mathbf{P}(|U(t/2)| > s_r)$ gives the first term in (1.3.25). We now turn to bound $I_2$. Fixing $\theta_{r+i} > 0$, we have

$$I_2 \leq e^{-\theta_{r+i}t/2} \prod_{j=1}^{i} \mathbf{E}_x \left[ e^{\theta_{r+i}q_{r+j}Z_{r+j}} \right]. \tag{1.3.27}$$

With the particular choice

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

we have that

$$\exp(\theta_{r+i}q_{r+j} - C_3s_{r+j}) \leq \exp(-C_3/2) =: \beta < 1.$$  

Here, we used that if $n \leq m$ then $G^*(n) \leq G^*(m)$. Thus by (1.3.24) there exists $\alpha = \alpha(\beta) > 0$ such that we can bound the exponential moments in (1.3.27) 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 (1.3.27) gives the second term in (1.3.25). \hfill \Box

**Lemma 1.3.13.** There are constants $C_8, C_9, C_{10} > 0$ such that for $t = (1 + a)C_8|G|(t_{rel}(G) + \log |G|)$ and every $x \in V(G)$ we have

$$\mathbf{E}_x \left[ 2^{U(t)} \right] \leq 1 + C_9 \exp \left( -a C_{10} \log (n^*) \right). \tag{1.3.28}$$

**Proof.** We can write

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

For $i \leq r$, we have that $s_{i-1} = s_i + t_u(G)$. By Lemma 1.3.11, we have that

$$2^{s_{i-1}} \mathbf{P}[|U(t)| > t] \leq \exp \left( (s_{i-1} + t_u(G)) \log 2 + \frac{s_i}{t_{rel}(G)} \left( C_4 \log |G| - \frac{C_5}{|G|^t} \right) \right).$$

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

$$\exp \left( -a s_i \left( 1 + \frac{\log |G|}{t_{rel}(G)} \right) \right). \tag{1.3.29}$$
For \( r + i \in \{ r + 1, \ldots, r + \tilde{r} \} \) we have from (1.3.25) that
\[
2^{s_{r+i-1}} \mathbb{P}_x [ |U(t)| > s_{r+i} ]
\leq 2^{s_{r+i-1}} \mathbb{P}_x [ |U(t)| > \frac{1}{2} ] + \exp \left( s_{r+i} \left( C_6 + \log 2 \right) + \frac{C_7}{G|G^*(n^*)|} t \right).
\]

The first term admits the same bound as (1.3.29) with \( i = r \), possibly by increasing \( C_8 \) if necessary. Using that \( i \leq \log |n^*| \), by increasing \( C_8 \) if necessary, from condition (3) it is easy to see that the second term admits the bound
\[
\exp \left( -a s_{r+i} \log |G| + t_{rel}(G) \right).
\]

Applying condition (3) again, we see that (1.3.30) is bounded from above by
\[
\exp ( -a s_{r+i} \log(n^*) ).
\]

Putting together the estimates we get that for \( i \in \{ 1 \ldots \tilde{r} \} \)
\[
2^{s_{r+i-1}} \mathbb{P}_x [ |U(t)| > s_{r+i} ]
\leq \exp \left( -a s_{r+i} \left( 1 + \log|G| \right) \right) + \exp ( -a s_{r+i} \log(n^*) )
\]

Summing (1.3.29) and (1.3.31) gives (1.3.28) (the dominant term in the summation comes from when \( s_{r+i} = 1 \)) which proves the lemma.

**Proof of Theorem 1.3.4.** This is a consequence of Lemma 1.3.13 and the relationship between \( t_u(G^*) \) and \( E[2^{U(t)}] \) given in (1.3.7).

**1.3.5 Proof of Theorem 1.3.1.**

We are going to prove Theorem 1.3.1 by checking the hypotheses of Theorem 1.3.4. We begin by noting that by [26, Corollary 12.12] and [26, Section 12.3.1], we have that
\[
t_{rel}(Z^d_n) = \Theta(d n^2).
\]

By [12, Example 2, Page 2155], we know that \( t_u(Z_n) = O(n^2) \). Hence by [10, Theorem 2.10], we have that
\[
t_u(Z^d_n) = O((d \log d)n^2).
\]

The key to checking parts (1)–(3) of Assumption 1.3.2 are the Green’s function estimates which are stated in Proposition 1.3.15 (low dimension) and Proposition 1.3.19 (high dimension). 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 \( G \). First, we have that
\[
P^t(x,y) \leq P^t(x,x) \text{ for all } x,y.
\]
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 \( G \) 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 [26, Proposition 10.18(ii)]. Moreover, by [26, Proposition 10.18], we have that

\[
P_t(x,x) \leq P_s(x,x) \text{ for all } s \leq t. \tag{1.3.35}
\]

The main ingredient in the proof of Proposition 1.3.15, our low dimension Green’s function estimate, is the following bound for the return probability of a lazy random walk on \( \mathbb{Z}^d \).

**Lemma 1.3.14.** Let \( P(x,y;\mathbb{Z}^d) \) denote the transition kernel for lazy random walk on \( \mathbb{Z}^d \). For all \( t \geq 1 \), we have that

\[
P_t(x,x;\mathbb{Z}^d) \leq \sqrt{2} \left( \frac{4d}{\pi} \right)^{d/2} \frac{1}{4^{d/2}} + e^{-t/8}. \tag{1.3.36}
\]

**Proof.** To prove the lemma we first give an upper bound on the transition probabilities for a (non-lazy) simple random walk \( Y \) on \( \mathbb{Z}^d \). One can easily give an exact formula for the return probability of \( Y \) to the origin of \( \mathbb{Z}^d \) in \( 2t \) steps by counting all of the possible paths from 0 back to 0 of length \( 2t \) (here and hereafter, \( P_{NL}(x,y;\mathbb{Z}^d) \) denotes the transition kernel of \( Y \)):

\[
P_{NL}^{2t}(x,x;\mathbb{Z}^d) = \sum_{n_1 + \ldots + n_d = t} \frac{(2t)!}{(n_1!)^2(n_2!)^2 \cdots (n_d!)^2} \frac{1}{(2d)^{2t}}
\]

\[
= \frac{1}{(2d)^{2t}} \left( \frac{2t}{t} \right) \sum_{n_1 + \ldots + n_d = t} \left( \frac{t!}{n_1!n_2! \cdots n_d!} \right)^2.
\]

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

\[
P_{NL}^{2t}(x,x;\mathbb{Z}^d) \leq \frac{1}{(2d)^{2t}} \left( \frac{2t}{t} \right) \max_{n_1 + \ldots + n_d = t} \frac{t!}{n_1! \cdots n_d!} \sum_{n_1 + \ldots + n_d = t} \frac{t!}{n_1! \cdots n_d!}
\]

\[
\leq \frac{1}{(2d)^{2t}} \left( \frac{2t}{t} \right) \frac{t!}{\left( \left\lceil \frac{t}{d} \right\rceil \right)!^d} \cdot d^d.
\]

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

\[
P_{NL}^{2t}(x,x;\mathbb{Z}^d) \leq \frac{\sqrt{2}}{(2\pi)^{d/2}} \cdot \frac{d^{d/2}}{t^{d/2}} \tag{1.3.37}
\]

52
We are now going to deduce from (1.3.37) a bound on the return probability for a lazy random walk $X$ on $\mathbb{Z}^d$. 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 $(\xi_i)$ are iid with $P[\xi_i = 0] = P[\xi_i = 1] = \frac{1}{2}$ and are independent of $Y$. Note that $N_t$ is distributed as a binomial random variable with parameters $t$ and $1/2$. Thus,

$$P^t(x,x;\mathbb{Z}^d) = \sum_{i=0}^{t/2} P^i_{NL}(x,x;\mathbb{Z}^d) P(N_t = 2i)$$

where in the second term we used the monotonicity of the upper bound in (1.3.37) in $t$. The first term can be bounded from above by using the Hoeffding inequality. This yields the term $e^{-t/8}$ in (1.3.36).

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

**Proposition 1.3.15.** Let $G(x,y)$ denote the Green’s function for lazy random walk on $\mathbb{Z}_n^d$. For each $\delta \in (0,1)$, there exists constants $C_1, C_2, C_3 > 0$ independent of $n,d$ for $d \geq 3$ such that

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

for all $x,y \in \mathbb{Z}_n^d$ 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

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

where $Z(t)$ is a one dimensional simple random walk on $\mathbb{Z}$. Let $k = |x-y|$. Applying (1.3.34) and (1.3.35) in the second step, and estimating the probability of wrapping around in time $n^{2-\delta}$ in the third term, we see that

$$G(x,y) = \sum_{t=k}^{t_u} P^t(x,y) \leq \sum_{t=k}^{n^{2-\delta}} P^t(x,x;\mathbb{Z}^d) + t_u P^{n^{2-\delta}}(x,x;\mathbb{Z}^d) + dt_u e^{-n^\delta/\pi}$$

We can estimate the sum on the right hand side above using Lemma 1.3.14 yielding the first term in the assertion of the lemma. Applying Lemma 1.3.14
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 (1.3.38) is bounded by

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

Indeed, the factor $(d \log d)n^2$ comes from (1.3.33) and the other factor comes from Lemma 1.3.14. Combining proves the lemma.

Proposition 1.3.15 is applicable when $n$ is much larger than $d$. We now turn to prove Proposition 1.3.19, which gives us an estimate for the Green’s function which we will use when $d$ is large. Before we prove Proposition 1.3.19, we first need to collect the following estimates.

Lemma 1.3.16. Suppose that $X$ is a lazy random walk on $\mathbb{Z}_d^n$ for $d \geq 8$ and that $|X(0)| = k \leq \frac{d}{8}$. For each $j \geq 0$, let $\tau_j$ be the first time $t$ that $|X(t)| = j$. There exists a constant $C_k > 0$ depending only on $k$ such that

$$P[\tau_0 < \tau_2k] \leq C_k d^{-k}.$$  

If, instead, $|X(0)| = 1$, then there exists a universal constant $p > 0$ such that

$$P[\tau_0 < \tau_2k] \geq p.$$  

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, [26, Section 17.3.1]). The second assertion of the lemma follows from the same argument.

Lemma 1.3.17. Assume that $k \in \mathbb{N}$ and that $d = 2k \lor 3$. Suppose that $X$ is a lazy random walk on $\mathbb{Z}^d$ and that $|X(0)| = 2k$. Let $\tau_k$ be the first time $t$ that $|X(t)| = k$. There exists $p_k > 0$ depending only on $k$ such that

$$P[\tau_k = \infty] \geq p_k > 0.$$  

Proof. Let $P_y$ denote the law under which $X$ starts at $y$. Assume that $P_y[\tau_k = \infty] = 0$ for some $y \in \mathbb{Z}^d$ with $|y| = 2k$. Suppose that $z \in \mathbb{Z}^d$ with $|z| = 2k$ and let $\tau_z$ be the first time that $X$ hits $z$. Then since $P_y[\tau_z < \tau_k] > 0$, it follows from the strong Markov property that $P_z[\tau_k = \infty] = 0$. 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 [24, Theorem 4.3.1], the expected amount of time that $X$ spends in $B(0, k)$ is finite. This is a contradiction.
**Lemma 1.3.18.** Assume that \( k \in \mathbb{N} \) and \( d \geq 2k \lor 3 \). Suppose that \( X \) is a lazy random walk on \( \mathbb{Z}_n^d \) and that \( |X(0)| = 2k \). Let \( \tau_k \) be the first time \( t \) such that \( |X(t)| = k \). There exists \( p_k, c_k > 0 \) depending only on \( k \) such that \( \mathbb{P}[\tau_k > c_k d n^2] \geq p_k > 0 \).

**Proof.** We first assume that \( d = 2k \lor 3 \). It follows from Lemma 1.3.17 that there exists a constant \( p_{k,1} > 0 \) depending only on \( k \) such that \( \mathbb{P}[\tau_k > \tau_{n/4}] \geq p_{k,1} \). The local central limit theorem (see [24, Chapter 2]) implies that there exists constants \( c_{k,1}, p_{k,2} > 0 \) such that the probability that a random walk on \( \mathbb{Z}^d \) 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 \( d = 2k \lor 3 \).

Now we suppose that \( d \geq 2k \lor 3 \). 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 \( \mathbb{Z}_{2k}^n \). Clearly, \( |Y(0)| = 2k \) because \( X(0) \) cannot have more than \( 2k \) non-zero coordinates. For each \( j \), let \( \tau_j^Y \) be the first time \( t \) such that \( |Y(t)| = j \). Then \( \tau_j^Y \leq \tau_k \). 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 \( \mathbb{P}[N_{\tau_k^Y} \geq c_{k,1} n^2] \geq p_{k,3} > 0 \) 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. \(\square\)

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

**Proposition 1.3.19.** Suppose that \( d \geq 8 \). Let \( G(x, y) \) denote the Green’s function for lazy random walk on \( \mathbb{Z}_n^d \). For each \( k \in \mathbb{N} \) with \( k \leq \frac{d}{8} \), there exists a constant \( C_k > 0 \) which does not depend on \( n, d \) such that

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

**Proof.** See Figure 1.3 for an illustration of the proof. By translation, we may assume without loss of generality that \( y = 0 \); let \( k = |x| \). Let \( \tau_0 \) be the first time \( t \) such that \( |X(t)| = 0 \). The strong Markov property implies that

\[
G(x, y) \leq \mathbb{P}[\tau_0 < t_u]G(x, x).
\]

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

\[
\mathbb{P}[\tau_0 < t_u] \leq \frac{C_k}{d^k} \text{ and } G(x, x) \leq C_0.
\]

We will first prove (1.3.40); the proof of (1.3.41) will be similar.

55
Figure 1.3: Assume that $d \geq 8$ and that $k \in \mathbb{N}$ with $d \geq 8k$. Let $X$ be a lazy random walk on $\mathbb{Z}^d$ and that $X(0) = x$ with $|x - y| = k$. In Proposition 1.3.19 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 1.3.16 to show that $X$ escapes to $\partial B(0, 4k)$ with probability at least $1 - C_k d^{-k}$. We then decompose the path of $X$ into successive excursions $\{X(\sigma_j^{2k}), \ldots, X(\tau_j^{4k}), \ldots, X(\sigma_{j+1}^{2k})\}$ between $\partial B(0, 2k)$ back to itself through $\partial B(0, 4k)$. By Lemma 1.3.16, we know that each excursion hits 0 with probability bounded by $C_{2k} d^{-k}$ and Lemma 1.3.18 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.

Let $N$ be a geometric random variable with success probability $C_{2k} d^{-2k}$ where $C_{2k}$ is the constant from Lemma 1.3.16. Let $(\xi_j)$ be a sequence of independent random variables with $\mathbb{P}[\xi_j = c_{2k} d n^2] = p_{2k}$ and $\mathbb{P}[\xi_j = 0] = 1 - p_{2k}$ where $c_{2k}, p_{2k}$ are the constants from Lemma 1.3.18 independent of $N$. We claim that $\tau_0$ is stochastically dominated from below by $\sum_{j=1}^N \xi_j$ where $\zeta$ is independent of $N$ and $(\xi_j)$ with $\mathbb{P}[\xi_j = 0] = C_{2k} d^{-k} = 1 - \mathbb{P}[\zeta = 1]$. Indeed, to see this we let $\sigma_k^0 = 0$ and let $\tau_{4k}^0$ be the first time $t$ that $|X(t)| = 4k$. For each $j \geq 1$, we inductively let $\sigma_{2k}^j$ be the first time $t$ after $\tau_{4k}^{j-1}$ that $|X(t)| = 2k$ and let $\tau_{4k}^j$ be the first time $t$ after $\sigma_{2k}^j$ that $|X(t)| = 4k$. Let $\mathcal{F}_t$ be the filtration generated by $X$. Lemma 1.3.16 implies that the probability that $X$ hits 0 in $\{\sigma_{2k}^j, \ldots, \tau_{4k}^j\}$ 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 \( \zeta \) is to take into account the probability that \( X \) reaches distance 2\( k \) before hitting 0. Moreover, Lemma 1.3.18 implies that \( \mathbb{P}[\sigma_{2k}^j - \tau_{2k}^{(j)} \geq c_2 k d^2 n^2; F_{\tau_{2k}^{(j)}}] \geq \rho_2 k \). This leads to the definition of the \((\xi_j)\) above. This implies our claim.

To see (1.3.40) from our claim, an elementary calculation yields that
\[
\mathbb{P}[N \zeta \leq C_{2k}^{-1} d^{k}] \leq \mathbb{P}[N \leq C_{2k}^{-1} d^{k} \text{ or } \zeta = 0] \leq 2 d^{-k} + C_k d^{-k}.
\]

We also note that
\[
\mathbb{P} \left[ \sum_{j=1}^{m} \xi_j \leq \frac{p_k c_k m n^2}{2} \right] \leq e^{-cn}
\]
for some constant \( c > 0 \). Combining these two observations along with a union bound implies (1.3.40). To see (1.3.41), we apply a similar argument using the second assertion of Lemma 1.3.16.

Now that we have proved Proposition 1.3.15 and Proposition 1.3.19, we are ready to check the criteria of Assumption 1.3.2.

**Part (1)**

By [26, Proposition 1.14] with \( \tau_x^+ = \min\{t \geq 1 : X(t) = x\} \), we have that \( \mathbb{E}_x[\tau_x^+] = |Z^d_n| \). Applying Proposition 1.3.19, we see that there exists constants \( d_0, r > 0 \) such that if \( d \geq d_0 \), then
\[
G(x, y) \leq 1/2 \text{ for all } |x - y| \geq r.
\] (1.3.42)

Proposition 1.3.15 implies that there exists \( n_0 \) such that if \( n \geq n_0 \) and \( 3 \leq d < d_0 \), then (1.3.42) 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 \( \tau r \) be the first time \( t \) that \( |X(t) - X(0)| = r \). We observe that there exists \( \rho_0 = \rho_0(r) > 0 \) such that
\[
\mathbb{P}_x[\tau_r < \tau_x^+] \geq \rho_0
\] (1.3.43)
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 (1.3.42) with (1.3.43), we see that \( \mathbb{P}_x[\tau_x^+ \geq t_u(G)] \geq \rho_1 > 0 \) uniform \( d \geq d_0 \). Let \( F_t \) be the filtration generated by \( X \). We consequently have that
\[
\mathbb{E}_x[\tau_x^+] \geq \mathbb{E}_x[\tau_x^+ 1_{\{\tau_x^+ \geq t_u(G)\}}] = \mathbb{E}_x[\mathbb{E}_x[\tau_x^+ | F_{t_u(G)}] 1_{\{\tau_x^+ \geq t_u(G)\}}] 
\geq \mathbb{E}_x[\mathbb{E}_x(t_u(G)) | \tau_x] 1_{\{\tau_x^+ \geq t_u(G)\}}] \geq \rho_1 \left( 1 - \frac{1}{2e} \right) \mathbb{E}_\pi[\tau_x].
\]
That is, there exists \( \rho_2 > 0 \) uniform in \( d \geq d_0 \) such that \( \mathbb{E}_x[\tau_x^+] \geq \rho_2 \mathbb{E}_\pi[\tau_x] \). Hence by [26, Lemma 10.2], we have that \( t_{\text{hit}}(Z^d_n) \leq K_1 |Z^d_n| \) where \( K_1 = 2/\rho_2 \) is a uniform constant.

57
Remark 1.3.20. There is another proof of Part 1 which is based on eigenfunctions. In particular, we know that
\[ t_{hit}(\mathbf{Z}_n^d) \leq 2E_x[\tau_x] = 4 \sum_i \frac{1}{1 - \lambda_i} \]
where the \( \lambda_i \) 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 \( \lambda_i \) can be computed explicitly using [26, Lemma 12.11] and the form of the \( \lambda_i \) when \( d = 1 \) which are given in [26, 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 1.3.19 that there exist constants \( C > 0 \) and \( d_0 \geq 3 \) such that
\[ G(x, y) \leq \frac{C}{d} \text{ for } x, y \in \mathbf{Z}_n^d \text{ with } |x - y| = 1 \]  
provided \( d \geq d_0 \). Consequently, there exists \( K \in \mathbf{N} \) which does not depend on \( d \geq d_0 \) such that
\[ 2K(5/2)^K G^K(x, y) = O \left( 2K \left( \frac{5/2}{d} \right)^K \right) \]  
(1.3.45)
It follows by combining (1.3.32) and (1.3.33) that we have that
\[ \frac{t_{hit}(\mathbf{Z}_n^d)}{t_{rel}(\mathbf{Z}_n^d)} = O(\log d). \]  
(1.3.46)
Combining (1.3.45) with (1.3.46) shows that part 2 of Assumption 1.3.2 is satisfied provided we take \( K_2 = K \) large enough. Moreover, (1.3.46) clearly holds if \( 3 \leq d < d_0 \) by Proposition 1.3.15.

Part 3

We first note that it follows from (1.3.32), (1.3.33), Proposition 1.3.15 and Proposition 1.3.19 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
\[ G^*(n^*) \leq K_3 \left( \frac{dn^2 + d \log n}{\log d + \log n} \right). \]  
(1.3.47)
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 (1.3.47) it is enough to show that $G^*(n^*) \leq Kd n^2 / \log n$. We can bound $G^*(n^*)$ in this case as follows. Let $D = (d \log d \log n)^{1/(\frac{d}{2}-1)}$. By Proposition 1.3.15 we can bound from above the expected amount of time that $X$ starting at 0 in $\mathbb{Z}_n^d$ spends in the $L^1$ ball of radius $D$ by summing radially:

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

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 $L^\infty$ 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:

$$G^*(n^*) \leq C_3 n(d \log d \log n)^{\frac{5}{2}} + C_4 D^{1-\frac{1}{2}d} n^*$$

$$\leq C_3 n(d \log d \log n)^{\frac{5}{2}} + C_4 \cdot C d^2 n^2 \log d \frac{d \log d \log n}{d \log d \log n},$$

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 1.3.19 to bound $G^*(n^*)$. The number of points which have distance at most $k$ to 0 is clearly $1 + (2d)^k$. Consequently, by Proposition 1.3.19 we have that

$$G^*(n^*) \leq \left( C_0 + \sum_{k=1}^{3} C_k d^{-k}(2d)^k \right) + C_4 d^{-4} n^*$$

$$\leq C_5 + \frac{C_6 (\log d)n^2}{d^2}$$

for some constants $C_5, C_6 > 0$. Since $d^2 \geq \log n$, this is clearly dominated by the right hand side of (1.3.47) (with a large enough constant), which completes the proof in this case.

\[ \square \]
Bibliography


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 independently by Solomonoff, Rapoport (1951) [79] and by Gilbert (1959) [65], and by Erdős and Rényi (1960) [61]. Given a finite set of vertices, a link between vertex \( x \) and \( y \) is formed independently of all other pair of vertices with probability \( p \). Albeit the simplicity of the model, it serves as an interesting example of phase transition: there is a threshold in the link probability, such that the network has crucially different properties above and below the threshold. A wide spectrum of literature investigates graph models with a fixed number of vertices (i.e some generalizations of the Erdős-Rényi (ER) graphs), we refer the reader to the books of [68] or [50] as an introduction.

More recently in [53] 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 phase transition: i.e. the emergence of the giant component, change of 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 [57], [67]. 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 a 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 edges of the graph - serving as pipes with lengths determined by the weights - 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 [46] Bhamidi, van der Hofstad and Hooghiemstra analyzed FPP on the ER random graph with i.i.d. exponentially distributed 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 [44], [45], [67], and newly the diameter with edge-weights was investigated in [38, 59].

Parallel to the discussion of the ER and related models, 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 in the last two decades.

The Erdős - Rényi graphs and their generalizations offer a simple and powerful model with many applications, but they fail to match some very important properties that are typical for real-world networks. First, the number of edges of a vertex follows asymptotically a Poisson-type distribution, having an exponential decay for large degrees: This fact hinders the formation of hubs, i.e. vertices with very high degree, existing in most real network. Second, one can show that the number of triangles in the graph is negligible compared to its size: the ER graphs and their generalizations have a low local clustering coefficient, unlike many real networks having a high clustering. Here and later, the local clustering coefficient of a vertex refers to the proportion of closed triangles and all edge-pair starting from the given vertex.

The Watts and Strogatz model [81] is an interpolation between the ER model and high clustering grid-based models: The vertices of the network are arranged on a grid, say, on a circle, and each of the nodes is connected to the vertices which are closer than $k$ steps in the grid. This graph has high clustering but large diameter, thus to obtain the small diameter each edge is re-wired to a uniform random vertex with some probability $0 \leq \beta \leq 1$. For $\beta = 0$ the model is just a regular grid, and for $\beta = 1$ it approaches the ER graphs. The model is often called small world model, since even for small re-wiring probability $\beta$ the diameter is significantly smaller than that in the grid and similar to the one in the ER model. The high clustering property is ensured by having the grid as an initial configuration.

A different attempt to model real networks resulted in the construction of numerous new, more dynamical and growing network models, see e.g. [42], [50], [53], [60], [71]. 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 \[82, 78, 56\]. The model was heuristically introduced by Barabási and Albert \[40\], and the first who investigated the model rigorously were Bollobás, Riordan, Spencer and Tusnády \[52\], and the mathematically rigorous construction was done by Bollobás and Riordan \[51\]. In the preferential attachment model (sometimes also called Barabási Albert model) discussed by Bollobás, Riordan, Spencer and Tusnády \[52\], 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: \[54\] considers also directed edges, and non-linear preferential attachment model appears in \[72\]. Rudas, Tóth and Valkó \[77\] 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 \(\zeta_i D_i(t) + \eta_i\), \(\zeta_i\) and \(\eta_i\) are nonnegative parameters called the multiplicative and additive fitness of vertex \(v_i\), respectively. A model where vertices only obey additive fitness \(\eta_i\) is discussed in \[62\]. Variations of multiplicative fitness models were introduced by Bianconi and Barabási \[49, 48\], and studied further in \[55\]. The degree distribution both for the additive and multiplicative models was found by Bhamidi \[47\]. Another interesting modification of the preferential attachment model is the Kim-Holmes model \[69\], where the authors extend the dynamics by a triangle-formation step. In the model, the clustering coefficient is tunable by changing a control parameter. Another 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 \[39\]. The literature on this field has a wide range and is summarized e. g. in \[50\] or in \[68\].

A completely different approach than preferential attachment was initiated by Barabási, Ravasz, and Vicsek \[41\] based on the observation that real networks often obey some hierarchical structure. They introduced deterministic network models generated by a method which is common in constructing fractals. Their model exhibits both hierarchical structure and an extreme-end power law decay of the degree sequence. This means that vertices of ”high enough” degree follow power law behavior. However, it is
a bipartite graph, hence no triangles. The clustering coefficient of a vertex is the proportion of triangles to the edge-pairs starting from the vertex, so the clustering coefficient of the model equals 0. In order to model also the clustering behavior of real networks, Ravasz and Barabási [76] developed the original model in [41] so that their deterministic network model preserved the same power law decay and had similar clustering behavior to many real networks. Namely, the local clustering coefficient decays inversely proportional to the degree of the node. As a consequence of this and the power law decay, in their model and also in real networks, the average local clustering coefficient is more or less independent of the size of the network (uniformly bounded away from both infinity and 0). A similar, fractal based deterministic model were introduced by Zhang, Comellas, Fertin and Rong [86], 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 Sierpiński carpet. Slightly different randomized version were introduced in [83, 84, 87, 85, 88].

In this section we generalize both of the models of [41] and [76]. 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 $\{0, 1, \ldots, N - 1\}^n$. 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.

The main advantage of our generalization is that our construction provides easily analyzable unbounded average degree examples: namely, the extreme-end exponent $\gamma$ in the power-law can be any log-rational number between $(1, 1 + \log 3/\log 2]$, producing graph sequences in the regime $\gamma \in (1, 2)$. If the initial bipartite graph is bi-regular, we can explicitly determine the degree exponent of the "high degree" and the "low degree" vertices and show that two different power law exponents dominate the degree distribution.

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 [76], we add some additional edges to $G_1$ to obtain the (no longer bipartite) graph $\hat{G}_1$. Then we build up the graph sequence $\hat{G}_n$ as follows: $\hat{G}_n$ consist of $N^{n-1}$ copies of $\hat{G}_1$, which copies are connected to each other in the same way as they were in $G_n$. So, $\hat{G}_n$ 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 $\hat{G}_1$, see Figures 2.3 and 2.6. We give a rigorous proof of the fact that the average local clustering coefficient of $\hat{G}_n$ 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 $G_n$ into
the unit square is carried out as follows: A vertex \( x = (x_1, \ldots, x_n) \) is identified with the corresponding \( N \)-adic interval \( I_x \) (see (2.2.4)). \( \Lambda_n \) is the union of those \( N^{-n} \times N^{-n} \) squares \( I_x \times I_y \) for which the vertices \( x, y \) are connected by an edge in \( G_n \). So, \( \Lambda_n \) is the most straightforward embedding of the adjacency matrix of \( G_n \) into the unit square. \( \Lambda_n \) 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 \( \Lambda \) on the plane, see Figure 2.1(c).

We discuss connections between the graph theoretical properties of \( G_n \) and properties of the limiting fractal \( \Lambda \). In particular, we express the power law exponent of the degree distribution with the ratio of the Hausdorff dimensions of some slices of \( \Lambda \) (Theorem 2.3.6).

Furthermore, using \( \Lambda \) we generate a random graph sequence \( G'_n \) in a way which was inspired by the \( W \)-random graphs introduced by Lovász and Szegedy [73], see also Diaconis, Janson [58], 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 [53] considered inhomogeneous random graphs generated by a kernel. Our model is not covered by their construction, since \( \Lambda \) is a fractal set of zero two dimensional Lebesgue measure. Here we remark in advance that the fractal limit \( \Lambda \) of our embedded adjacency matrices of \( G_n \) or \( G'_n \) is not stable under all the isomorphisms of the unit square into itself, thus, the Lovász-Szegedy limit theory does not apply to our graph sequences word by word. However, different encoding of vertices of the base graph \( G \) in the alphabet \( \{1, \ldots, N\} \) gives different fractal limit \( \Lambda \)-s with the same Hausdorff dimension.

The section is organized as follows: In Section 2.2 we define the deterministic model and the associated fractal set \( \Lambda \). In Section 2.3, we verify the scale free property of \( G_n \) (Theorem 2.3.1). We compare the Hausdorff dimension of \( \Lambda \) 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 \( G_n \) (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 [41], 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 $\Sigma_1 = \{0, \ldots, N - 1\}$. We partition $\Sigma_1$ 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 $(x, y)$, 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 $\Sigma_n = \{(x_1x_2\ldots x_n) : x_i \in \Sigma_1\}$, all words of length $n$ over the alphabet $\Sigma_1$. 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 $\Sigma_1$. Namely,

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

2. We define the type of a word $z = (z_1z_2\ldots z_n) \in \Sigma_n$ as follows: if all the elements $z_j, j = 1, \ldots, n$ of $z$ fall in the same $V_i, i = 1, 2$ then $\text{typ}(z)$ the type of $z$ is $i$. Otherwise $\text{typ}(z) := 0$.

3. For $x = (x_1\ldots x_n), y = (y_1\ldots y_n) \in \Sigma_n$ we denote the common prefix by

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

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

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

where the concatenation of the words $a, b$ is denoted by $ab$.

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

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

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

69
That is, $E(G_n) \subset \Sigma_n \times \Sigma_n$:

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

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 $x \in \{0, 1, \ldots, N - 1\}$, consider the set $W_x$ of vertices $(x_1 \ldots x_n)$ of $G_n$ with $x_1 = x$. Then the induced subgraph on $W_x$ is identical to $G_{n-1}$.

We write $\deg_n(x)$ 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 $\ell(x)$, the length of the longest block from backwards (longest postfix) in $x$ which has a nonzero type,

$$\ell(x) := \max_{i \in \mathbb{N}} \{\text{typ}(x_{n-i+1}, \ldots, x_n) \in \{1, 2\}\} \quad (2.2.2)$$

Remark 2.2.4. The degree of a node $x \in \Sigma_n$

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

where

$$S(x) = 1 + \deg(x_{n-1}) + \cdots + \deg(x_{n-1}) \cdot \deg(x_{n-2}) \cdots \deg(x_{n-\ell(x)+1})$$

$$= \sum_{r=0}^{\ell(x)-1} \prod_{j=1}^{r} \deg(x_{n-j}), \quad (2.2.3)$$

where the empty product 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 [41] 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 $|V_1| > 1$. 

70
Figure 2.1: \( G_1, G_2, G_3, \Lambda_1, \Lambda_2, \Lambda_3 \) for the cherry Example 2.2.5. The adjacency matrices are drawn such that the origin is at the left-bottom corner and the orientation of the two axes goes right and up, respectively. Everything which is colored belongs to the adjacency matrix, and a box of a given color corresponds to an edge of the same color in the corresponding graph.
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 $[0,1]$:

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

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

$$Q\left(\frac{x}{y}\right) := I_x \times I_y, \quad \left(\frac{x}{y}\right) \in \Sigma_n \times \Sigma_n. \quad (2.2.5)$$

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

$$\Lambda_n(a,b) := \begin{cases} 1, & \text{if } (a,b) \in Q\left(\frac{x}{y}\right), \left(\frac{x}{y}\right) \in E(G_n); \\ 0, & \text{otherwise}. \end{cases} \quad (2.2.6)$$

That is,

$$\Lambda_n(a,b) = \sum_{\left(\frac{x}{y}\right) \in \sum_n E(G_n)} 1_{Q\left(\frac{x}{y}\right)}(a,b).$$

We write $\Lambda_n$ for the support of the function $\Lambda_n(a,b)$, see Figure 2.1(c).

Observe that $\Lambda_n$ is a compact set and $\Lambda_{n+1} \subset \Lambda_n$ holds for all $n$. So we can define the non-empty compact set

$$\Lambda := \bigcap_{n=1}^{\infty} \Lambda_n. \quad (2.2.7)$$

Clearly,

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

**Remark 2.2.7.** This representation obviously depends on the labeling of the graph $G$. For an arbitrary permutation $\pi$ of $\{0, \ldots, N-1\}$, 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

$$\Lambda_n^\pi(a,b) = \Lambda_n(\varphi_{\pi^{-1}}(a), \varphi_{\pi^{-1}}(b)), \quad \text{and}$$

$$1_{\Lambda_n^\pi}(a,b) = 1_\Lambda(\varphi_{\pi^{-1}}(a), \varphi_{\pi^{-1}}(b)),$$

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 $\Lambda$

Now we prove that the limit $\Lambda$ (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 and more see the book of Falconer [63, page 49-51]), with the directed graph $G$ defined below. Heuristically speaking, the $n$-th adjacency matrix $\Lambda_n$ can be written as the union of iterated maps formed by the maps $f_i, i = 1 \ldots |E(G)|$ of the unit interval, such as the the composition of maps $f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_n}$ which are allowed in $\Lambda_n$ can be determined as the paths of length $n i_1i_2 \ldots i_n$ in the directed graph $G$ below. We make these heuristics precise in the sequel.

Recall that the alphabet was partitioned into two sets $\Sigma_1 = V_1 \cup V_2$ according to the bipartite property of the base graph $G$.

Definition 2.2.8. The vertex set $V(G)$ is partitioned into three subsets:

\begin{align*}
V_{dd} &= \left\{ \left( \begin{array}{c} z \\ z \end{array} \right) : z \in \Sigma_1 \right\} \\
V_{12} &= \left\{ \left( \begin{array}{c} x \\ y \end{array} \right) : (x, y) \in E(G), x \in V_1, y \in V_2 \right\} \\
V_{21} &= \left\{ \left( \begin{array}{c} x \\ y \end{array} \right) : (x, y) \in E(G), x \in V_2, y \in V_1 \right\}.
\end{align*}

Then

$$V(G) = V_{dd} \cup V_{12} \cup V_{21}.$$  \hfill (2.2.8)

The set of directed edges $E(G)$ of $G$ 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 $e = (v_1, v_2) \in E(G)$ we define a homothety:

$$f_e : Q_{v_2} \to Q_{v_1}, \quad f_e(a, b) := \frac{1}{N}(a, b) + \frac{1}{N}(x_1, y_1), \quad \text{with } v_i = \left( \begin{array}{c} x_i \\ y_i \end{array} \right),$$

where $Q_v := Q_{v_i}$ is the level-1 square for $v = \left( \begin{array}{c} x \\ y \end{array} \right) \in V(G)$.

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

In general, $G$ 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 $\overrightarrow{K(V_i)}$ illustrates that we connect all pairs of vertices in the given direction.

Let $P_n$ be the set of all paths of length $n$ in $G$, i.e.

$$P_n := \{ u = (v_1 \ldots v_n) : \forall 1 \leq i < n (v_i, v_{i+1}) \in E(G) \}.$$
For a $v = (v_1 \ldots v_n) = (x_{y_1} \ldots x_{y_n}) \in \mathcal{P}_n$ it immediately follows from definitions (2.2.5) and (2.2.9) that

$$Q_{v} = f_{v} ([0,1]^{2}) = I_{x_{1} \ldots x_{n}} \times I_{y_{1} \ldots y_{n}},$$

(2.2.10)

where

$$f_{v}(.) := f_{(v_{1},v_{2})} \circ \cdots \circ f_{(v_{n-1},v_{n})}(.) \text{ if } n \geq 2,$$

$$f_{v}(a,b) := \frac{1}{N}(a,b) + \frac{1}{N}(x,y), \text{ if } n = 1, \ v = \begin{pmatrix} x \\ y \end{pmatrix}.$$  \hspace{1cm} (2.2.11)

The key observation of connecting $G$ 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 \( v = (v_1 \ldots v_n) = (a_1 \ldots a_n) \in \Sigma_n \times \Sigma_n \), thus \( a = (a_1 \ldots a_n) \) and \( b = (b_1 \ldots b_n) \) are vertices in \( G_n \). First we assume that \( v \in E(G_n) \). Observe

that by (2.2.1), \( (a_i, b_i) \) are vertices in \( G \). We would like to prove that the sequence

\[
(a_1, b_1) \ldots (a_n, b_n) \in \mathcal{P}_n.
\]  

(2.2.12)

If \( k := |a \cap b| \geq 1 \), then for \( i \leq k \), \( a_i = b_i \) holds, thus the sequence of points \( (a_1, b_1) \ldots (a_k, b_k) \) forms a path in \( K_{|N|}(V_{dd}) \). By (2.2.1), the pairs \( (a_{k+1}, b_{k+1}) \ldots (a_n, b_n) \) are all edges in \( G \) thus vertices in \( G \). 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 \( K_{|N|}(V_{12}) \) or in \( K_{|N|}(V_{21}) \). By definition of \( E(G) \), \( ((a_k, b_k), (a_{k+1}, b_{k+1})) \) is an edge in \( G \), so \( (a_i) \ldots (a_n) \) is a path in \( G \). 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 \( (a_1, b_1) \ldots (a_n, b_n) \) is a path of length \( n \) in \( G \), then we claim that for \( a = (a_1 \ldots a_n) \), \( b = (b_1 \ldots b_n) \in V(G_n) \)

\[
(a, b) \in E(G_n).
\]

The proof is very similar to the previous one.

In this way we can characterize \( \Lambda_n \) as follows:

**Corollary 2.2.10.**

\[
\Lambda_n = \bigcup_{v \in \mathcal{P}_n} Q_v = \bigcup_{v \in \mathcal{P}_n} f_v([0, 1]^2).
\]

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

Let us define

\[
\mathcal{P}_\infty := \{ v = (v_1v_2 \ldots) | \forall i \in \mathbb{N}, (v_i, v_{i+1}) \in E(G) \}.
\]

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

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

It is an immediate consequence of Corollary 2.2.10 that

\[
\Pi(\mathcal{P}_\infty) = \Lambda, \text{ i.e. } \Lambda = \bigcup_{v \in \mathcal{P}_\infty} \Pi_v.
\]  

(2.2.13)
This means that $\Lambda_n$, the embedded adjacency matrix of $G_n$, can be considered as the $n$-th approximation of the fractal set $\Lambda$.

In this way we coded the elements of $\Lambda$ by the elements of $P_\infty$. 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 $\Lambda$.

In this section we show that $\Lambda$ is the union of the diagonal of the unit square and countably many homothetic copies of the attractors formed by the maps in $V_{12}$ and the ones in $V_{21}$, respectively. The homotheties “pull back these attractors to the diagonal”, i.e. they are defined by the composition of the maps corresponding to finite words over the alphabet $V_{dd}$.

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

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

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):

$$F_{dd} := \{f_v\}_{v \in V_{dd}},$$

$$F_{12} := \{f_v\}_{v \in V_{12}},$$

$$F_{21} := \{f_v\}_{v \in V_{21}},$$

where the functions $f_v, v \in V(\mathcal{G})$ were defined in (2.2.11). The attractors of these IFS-s (see [63, p.30]) are the unique nonempty compact sets satisfying

$$\Lambda_{dd} := \bigcup_{v \in V_{dd}} f_v(\Lambda_{dd}) = \{\Pi(v)|v = (v_1, v_2 \ldots) \text{ and } v_i \in V_{dd}\}$$

$$\Lambda_{12} := \bigcup_{v \in V_{12}} f_v(\Lambda_{12}) = \{\Pi(v)|v = (v_1, v_2 \ldots) \text{ and } v_i \in V_{12}\}$$

$$\Lambda_{21} := \bigcup_{v \in V_{21}} f_v(\Lambda_{21}) = \{\Pi(v)|v = (v_1, v_2 \ldots) \text{ and } v_i \in V_{21}\}.$$ (2.2.14)

The Open Set Condition (see e.g. [63, p.35]) holds for these IFS-s, so we can easily compute the Hausdorff-dimension of the attractors. Clearly, $\Lambda_{dd}$ is the diagonal of the unit square.

Now we prove that $\Lambda$ is a countable union of homothetic copies of these attractors.
Theorem 2.2.11.

\[ \Lambda = \text{Diag} \cup \bigcup_{v \in V^*} \left( f^v_1(\Lambda_{12}) \cup f^v_2(\Lambda_{21}) \right), \]

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

Remark 2.2.12. Observe that \( \Lambda_{21} \) is the image of \( \Lambda_{12} \) by the reflection through the diagonal, hence \( \Lambda \) is symmetric to the diagonal. The same is true for the \( n \)-th approximation \( \Lambda_n \) of \( \Lambda \). This can be seen immediately by using the embedded adjacency matrix characterization of \( \Lambda_n \).

Remark 2.2.13. In the cherry example, \( \Lambda_{12} \) is a vertical, \( \Lambda_{21} \) is a horizontal copy of the one dimensional canonical Cantor set located at \( \frac{1}{2} \) in each cases. The one dimensional canonical Cantor set is defined as by the two transformations: \( f_0(x) = \frac{1}{3}x, f_2(x) = \frac{1}{3}x + \frac{2}{3} \). The third approximation of these are the red squares on \( \Lambda_3 \), Figure 2.1(c).

Proof of Theorem 2.2.11. We start by showing that

\[ \Lambda \subset \text{Diag} \cup \bigcup_{v \in V^*} \left( f^v_1(\Lambda_{12}) \cup f^v_2(\Lambda_{21}) \right). \]  

(2.2.15)

Pick an arbitrary point \((a,b)\) \(\in\) \(\Lambda\). As a consequence of \(\text{(2.2.13)}\) there exists a \(v = (v_1v_2\ldots) \in \mathcal{P}_\infty\) such that \(\Pi(v) = (a, b)\). 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 \(i \leq k, v_i \in V_{dd}\) 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 \(\Pi(v) \in \Lambda_{12} \cup \Lambda_{21}\).

Case \(k = \infty\) For the same reason, \(\Pi(v) = \lim_{n \to \infty} f_{v_1\ldots v_n}(0, 0) \in \Lambda_{dd} = \text{Diag}\).

This is so because \(f_{v_1\ldots v_n}(0, 0)\) is in the \(\frac{1}{N^n}\) neighborhood of the diagonal \(\{ (x, x) : x \in [0, 1] \}\).

Case \(0 < k < \infty\) Let \(v_k = (v_1 \ldots v_k)\). 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 \(w := (v_{k+1}v_{k+2}\ldots), \Pi(w) \in \Lambda_{12}\). Hence \(\Pi(v) = f_{w_k}(\Pi(w)) \in f_{v_k}(\Lambda_{12})\).

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

\[ \Lambda \supset \text{Diag} \cup \bigcup_{v \in V^*} \left( f^v_1(\Lambda_{12}) \cup f^v_2(\Lambda_{21}) \right), \]  

(2.2.16)

we will use the symbolic representation of \(\Lambda\) given in \(\text{(2.2.13)}\).
Pick an \( x \in [0, 1] \) and take the \( N \)-adic code \( (x_1x_2 \ldots) \) of \( x \). That is, 
\[
x = \sum_{n=1}^{\infty} \frac{x_n}{N^n}, \quad x_n \in \{0, \ldots, N-1\}.
\]
Then
\[
\pi := \left( \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ v_1 \\ v_2 \end{array} \right) \in \mathcal{P}_\infty,
\]

it is easy to see that \( \Pi(\pi) = (x, x) \). So by (2.2.13), \( (x, x) \in \Lambda \).

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

\[ \square \]

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 \( E(G'_n) \subset \Sigma_n \times \Sigma_n \):
\[
E(G'_n) = \left\{ \left( \begin{array}{c} x \\ y \end{array} \right) \mid \{\text{typ}(\tilde{x}), \text{typ}(\tilde{y})\} = \{1, 2\} \text{ and } \forall |x \land y| < i \leq n, (x_i y_i) \in E(G) \right\}
\]

In this case \( \Lambda'_n = \Lambda_n \setminus \text{Diag}_n \), where \( \text{Diag}_n \) is the union of the level \( n \) squares that have nonempty intersection with the diagonal. The sequence \( \Lambda'_n \) is not a nested sequence of compact sets. However, it is easy to see that the characteristic function of \( \Lambda'_n \) tends to characteristic function of \( \Lambda \setminus \text{Diag} \). Further, \( \Lambda'_n \) tends to \( \Lambda \) in the Hausdorff metric, see [63].

2.3 Properties of the sequence \( \{G_n\} \) and \( \Lambda \)

In this section we compute the degree distribution of \( G_n \), and relate it to the Hausdorff dimension of \( \Lambda \). We also compute the length of the average shortest path in \( G_n \). To get interesting results about the local clustering coefficient we need to modify our graph sequence \( G_n \) along the lines of [41].
(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 $G$:

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

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

Under assumption \textit{A1} the extreme end of the decay of the degree dis-
Determination is determined by the set of high degree nodes denoted by

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

An equivalent characterization of \( HD_n \) is

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

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

\[ P \left[ \deg_n(X) > \frac{d_2^{\ell+1} - 1}{d_1 - 1} + 1 \right] = n^{\ell+1}N^{n-\ell-1} = \left( \frac{n_1}{N} \right)^{\ell+1} \]

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

\[ \tilde{F}(t) = t^{-\frac{\log(N/n_1)}{\log d_1}} \cdot e(d_1) \quad \text{for} \quad t = \frac{d_2^{\ell+1} - 1}{d_1 - 1}. \]

So we have proved

**Theorem 2.3.1.** The degree distribution of the graph sequence \( G_n \) satisfying assumption [A1] has an extreme-end power law decay with exponent

\[ \tilde{\gamma} = \gamma - 1 = \frac{\log(N/n_1)}{\log d_1}. \quad (2.3.1) \]

Using this, a simple calculation shows that the largest decay \( \gamma \) 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.

We note that this decay exponent only holds for the far end of the degree distribution. Namely, the definition of the high degree vertices \( HD_n \) implies that only with vertices having a type-1 postfix of minimal length \( \sim \log d_2 \log n \) are responsible for this decay exponent if \( d_2 > 1 \) and the minimal postfix length is \( \log d_1 \log n \) if \( d_2 = 1 \). Thus, as \( n \) goes to infinity, larger and larger proportion of vertices are left out in both cases, and the extreme-end degree exponent can not be verified etc. by uniform sampling.

However, if we suppose that the base graph is bi-regular bipartite graph, i.e. all vertices in \( V_2 \) have degree \( d_2 > 1 \), then the tail distribution can be explicitly calculated and equals:
\[ P[\text{deg}_n(X) > x] = \left( \frac{d_2}{d_1 + d_2} \right)^{\left\lfloor \frac{\log x - \log d_1 - \log(d_1 - 1)}{\log d_1} \right\rfloor} \mathbf{1}_{\left\{ \left\lfloor \frac{\log x - \log d_1 - \log(d_1 - 1)}{\log d_1} \right\rfloor \leq n \right\}} + \left( \frac{d_1}{d_1 + d_2} \right)^{\left\lfloor \frac{\log x - \log d_2 - \log(d_2 - 1)}{\log d_2} \right\rfloor} \mathbf{1}_{\left\{ \left\lfloor \frac{\log x - \log d_2 - \log(d_2 - 1)}{\log d_2} \right\rfloor \leq n \right\}}. \]

The first term corresponds to vertices with type-1 postfix and gives the extreme end exponent

\[ \bar{\gamma} = -\frac{\log d_2 - \log(d_1 + d_2)}{\log d_1} = \frac{\log(N/n_1)}{\log d_1} \]

before. (Mind that \( n_1/(n_1 + n_2) = d_2/(d_1 + d_2) \) in the bi-regular case.) The second term corresponds to vertices with type-2 postfix, thus the indicator function after the second term vanishes first. It is an elementary calculation that the second term is dominating until the indicator function is not equal to zero. Thus, the function on the loglogplot has a discontinuity at \( x_0 \) where \( n = \frac{x_0 - \log d_2 - \log(d_2 - 1)}{\log d_2} \). Also, the exponent before this point (call it the exponent in the bulk) is given by

\[ \bar{\gamma}_0 := -\frac{\log d_1 - \log(d_1 + d_2)}{\log d_2} = \frac{\log(N/n_2)}{\log d_2}. \]

(2.3.2)

It is not hard to see that \( \bar{\gamma}_0 < \bar{\gamma} \) thus the bulk has an exponent which is even less than the one in the extreme end. Still, this parameter can be tuned similarly as choosing \( d_1, d_2 > 1 \) appropriately. Moreover, the interesting phenomenon occurs that the degree distribution changes power law exponent, see Figure 2.4.

### 2.3.2 Hausdorff dimension of \( \Lambda \)

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

\[ \dim_H (\Lambda \setminus \text{Diag}) = \dim_H (\Lambda_{12}). \]

(2.3.3)

We have seen that \( \Lambda_{12} \) 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

\[ \dim_H (\Lambda_{12}) = \frac{\log |E|}{\log N}. \]

(2.3.4)

Combining (2.3.3) and (2.3.4) yields the assertion of the following theorem:
Figure 2.4: The log-log-plot of the tail distribution of the graph $G_{100}$ for a bipartite base graph with $d_1 = 6, d_2 = 2, E = 12$. We can see the discontinuity of the slope of the curve where the type-2 postfix vertices vanish.

**Theorem 2.3.2.** The Hausdorff dimension of $\Lambda$ is

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

furthermore,

$$\dim_H (\Lambda \setminus \text{Diag}) = \dim_H (\Lambda_{12}) = \frac{\log |E|}{\log N}. \quad (2.3.5)$$

**Relation between the Hausdorff dimension and the extreme end power-law exponent.**

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

$$\text{HD} := \dim_H (\Lambda \setminus \text{Diag}).$$

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 extreme end degree distribution exponent (defined in equation (2.3.1)) equals

$$\tilde{\gamma} = \frac{\log N}{\log |E|} = \frac{1}{\dim_H (\Lambda \setminus \text{Diag})}.$$

**Proof.** This follows immediately from (2.3.1) and (2.3.5). \qed
In the more general setting where we assume only that \((A1)\) holds we still have a simple relation in between \(\dim_H (\Lambda \setminus \text{Diag})\) and \(\tilde{\gamma}\). Namely, putting together \((2.3.1)\) and \((2.3.5)\) and using that \(|E| = n_1 \cdot d_1\) we obtain the following Corollary:

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

\[
N^{HD-1} = d_1^{1 - \tilde{\gamma}}. \quad (2.3.6)
\]

If further the base graph is bipartite with \(d_2 = \deg(x), x \in V_2\), then we also have for the power law exponent in the bulk \(\tilde{\gamma}_0\) defined in \((2.3.2)\)

\[
N^{HD-1} = d_2^{1 - \tilde{\gamma}_0}.
\]

Hence

\[
\tilde{\gamma} = 1 - (HD - 1) \cdot \frac{\log N}{\log d_1},
\]

\[
\tilde{\gamma}_0 = 1 - (HD - 1) \cdot \frac{\log N}{\log d_2}. \quad (2.3.7)
\]

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 \((2.3.7)\) we have

\[
\tilde{\gamma} \in (0, 1] \text{ if and only if } HD \in [1, 2). \quad (2.3.8)
\]

By \((2.3.5)\) 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 \((2.3.8)\). On the other hand, our model produces graph sequences with extreme-end 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 \((2.3.7)\) 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

\[
\text{HD} = \frac{\log n_1(n_1 + 1)}{\log N} = 2 - \frac{\log 4}{\log N},
\]

\[
\text{\tilde{\gamma}} = \frac{\log(2 - \frac{1}{n_1})}{\log(n_1 + 1)}.
\]

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

**Geometric interpretation of the connection between HD and \(\tilde{\gamma}\).**

Here we always assume that assumption \((A1)\) is satisfied and in the rest of
Example 2.3.5 (Triplets). On Figure 2.5 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 (2.3.9) holds. Let $\ell_{\text{vert}}$ be an arbitrary vertical line that intersects $\Lambda_{12}$ and let $\ell_{\text{rand}}$ be a randomly chosen line on the plane that intersects $\Lambda_{12}$. Then almost surely,

$$1 - \tilde{\gamma} = \frac{\dim_H (\ell_{\text{rand}} \cap \Lambda_{12})}{\dim_H (\ell_{\text{vert}} \cap \Lambda_{12})}. \quad (2.3.10)$$

More precisely, any straight line $\ell$ on the plane can be described as

$$\ell = \ell_{(a,b)} = \{(x, y) | y = a \cdot x + b\}. \quad (2.3.11)$$

Let $\mathcal{A} := \{(a, b) : \ell=(a,b) \cap \Lambda_{12} \neq \emptyset\}$. Then for Lebesgue almost all $(a, b) \in \mathcal{A}$ the assertion of (2.3.10) holds with $\ell_{\text{rand}} = \ell_{(a,b)}$.

That is, among those lines that intersect $\Lambda_{12}$, a randomly chosen one intersects $\Lambda_{12}$ 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 $\tilde{\gamma}$. 

Figure 2.5: The triplets example which satisfies both (A1) and (2.3.9)
Proof of the Theorem. From (2.3.7) we get:
\[ 1 - \tilde{\gamma} = \frac{HD - 1}{\log \frac{d_1}{\log N}}. \] (2.3.12)

It is enough to prove that
\[(a)\] \(HD - 1 = \dim_H (\ell_{\text{rand}} \cap \Lambda_{12}),\)
\[(b)\] \(\frac{\log d_1}{\log N} = \dim_H (\ell_{\text{vert}} \cap \Lambda_{12}).\)

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 \(\Lambda_{12}\) with Lebesgue typical direction \(\theta\). Then a well-known Theorem of Marstrand [74, Theorem III] states that
\[\dim_H (L_{P,\theta} \cap \Lambda_{12}) = HD - 1.\] (2.3.13)

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 \(\nu\) be the natural (evenly distributed) measure on \(\Lambda_{12}\) which is known [70] to be equal to constant times the appropriate dimensional Hausdorff measure:
\[\nu(\cdot) = \text{const} \cdot H^{HD}|_{\Lambda_{12}}(\cdot).\]

Let \(W_a = \{(x, y) : y = a \cdot x, x, y, a \in \mathbb{R}\}\), and \(W_a^\perp\) be the line through origin which is perpendicular to \(W_a\) and we call \(\Lambda_{12}^a\) the orthogonal projection of \(\Lambda_{12}\) to the line \(W_a^\perp\). We write \(\nu_a\) for the projection of \(\nu\) to \(W_a^\perp\). It follows from the PROOF of [64, Theorem 6.8 (b)] that
\[\nu_a \ll \mathcal{L}eb\] for Lebesgue almost every \(a\). (2.3.14)

On the other hand, observe that \(\nu_a\) is a self-similar measure in itself. Using this and [73] Proposition 3.1 (ii)] for \(\nu_a\) we obtain that
\[\text{either } \nu_a \ll \mathcal{L}eb|_{\Lambda_{12}^a} \text{ or } \nu_a \sim \mathcal{L}eb|_{\Lambda_{12}^a}.\] (2.3.15)

Putting together (2.3.14) and (2.3.15) we get
\[\nu_a \sim \mathcal{L}eb|_{\Lambda_{12}^a}\] for Lebesgue-almost every \(a\). (2.3.16)

Combining Fubini’s Theorem and the Theorem of Marstrand [74, Theorem III] mentioned above, it follows that for \(\nu_a\)-almost every \(c\)
\[\dim_H ([W_a + c] \cap (\Lambda_{12})) = HD - 1.\] (2.3.17)

From this and (2.3.16) we obtain that for Lebesgue-almost every \(a\) satisfying \([W_a + c] \cap (\Lambda_{12}) \neq \emptyset\), the equation \(\dim_H ([W_a + c] \cap (\Lambda_{12})) = HD - 1\) holds also for Lebesgue-almost every \(c\). This completes the proof of Part (a).

Proof of Part (b) Let \(\ell_{\text{vert}}\) be a vertical line that intersects \(\Lambda_{12}\). Put \(\Lambda' := \ell_{\text{vert}} \cap \Lambda_{12}\). Then the \(n\)-th approximation of \(\Lambda'\) consists of \(d_1^n\) intervals with disjoint interior of length \(N^{-n}\). The assertion of Part (b) is an immediate corollary of [43, Theorem 3.1.1]. \(\square\)
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 $G_n$, for arbitrary bipartite base graph $G$ (i.e. without assumption A1).

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 $x, y \in \Sigma^n$, typ$(x)$, the common prefix $x \land y$ and the postfixes $\tilde{x}, \tilde{y}$ were defined in Definition 2.2.1.

Definition 2.3.7.

For two arbitrary vertices $x, y \in \Sigma_n$ we denote the length of their common prefix by $k = k(x, y) := |x \land y|$. Furthermore, let us decompose the postfixes $\tilde{x}, \tilde{y}$ into blocks of digits of the same type:

$$\tilde{x} = b_1 b_2 \ldots b_r, \quad \tilde{y} = c_1 c_2 \ldots c_q,$$

(2.3.18)

such that all of the blocks have a nonzero type and the consecutive blocks are of different types. That is, for $i = 1, \ldots, r - 1$, $j = 1, \ldots, q - 1$ we have

$$\text{typ}(b_i) \neq \text{typ}(b_{i+1}) \in \{1, 2\}, \quad \text{and} \quad \text{typ}(c_j) \neq \text{typ}(c_{j+1}) \in \{1, 2\}.$$ 

Note that we denoted the number of blocks in $\tilde{x}, \tilde{y}$ by $r$ and $q$, respectively. If $X$ and $Y$ are two random vertices of $G_n$, then the same notation as in (2.3.18) is used with capital letters.

Now we fix an arbitrary self-map $p$ of $\Sigma_1$ such that

$$(x, p(x)) \in E(G) \quad \forall x \in G.$$ 

Mind that $p(p(x)) \neq x$ can happen in many cases (if $n_1 \neq n_2$, this must happen for some $x \in \Sigma$). Also note that $x$ and $p(x)$ have different types since $G$ is bipartite. For a word $\tilde{z} = (z_1 \ldots z_m)$ with typ$(\tilde{z}) \in \{1, 2\}$ we define $p(\tilde{z}) := (p(z_1) \ldots p(z_m))$. Then,

$$(t \tilde{z}, t p(\tilde{z})) \text{ is an edge in } G_{\ell + m}, \forall t = (t_1 \ldots t_\ell),$$

(2.3.19)

follows from (2.2.1).

As usual we write Diam$(G)$ for the maximal graph-distance in the graph $G$ within components of $G$. Clearly Diam$(G) \leq N - 1$. 86
Lemma 2.3.8. Let \( x, y \) 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 \( r + q + \text{Diam}(G) - 2 \).

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 \( 2n + \text{Diam}(G) - 2 \).

Since the size of the graph is \( N^n \), therefore

\[
\text{Diam}(G_n) = \frac{2}{\log N} \log(|G_n|) + O(1).
\]

Proof of Lemma 2.3.8. First we construct a path \( P(x, y) \) of minimal length. Starting from \( x \) the first half of the path \( P(x, y) \) is as follows:

\[
\hat{x}^0 = x = (x \land y)b_1 \ldots b_{r-1}b_r,
\hat{x}^1 = (x \land y)b_1 \ldots b_{r-1}p(b_r),
\vdots
\hat{x}^{r-1} = (x \land y)b_1p(b_2 \ldots p(b_{r-1}p(b_r))),
\]

Starting from \( y \) the first half of the path \( P(x, y) \) is as follows:

\[
\hat{y}^0 = y = (x \land y)c_1c_2 \ldots c_r,
\hat{y}^1 = (x \land y)c_1 \ldots c_{r-1}p(c_r),
\vdots
\hat{y}^{q-1} = (x \land y)c_1p(c_2 \ldots p(c_{r-1}p(c_q))),
\]

It follows from (2.3.19) that

\[
P_x := (\hat{x}^0, \hat{x}^1, \ldots, \hat{x}^{r-1})
\]

\[
P_y := (\hat{y}^{q-1}, \ldots, \hat{y}^1, \hat{y}^0)
\]

are two paths in \( G_n \). To construct \( P(x, y) \) the only thing remained is to connect \( \hat{x}^{r-1} \) and \( \hat{y}^{q-1} \). Using (2.3.19) it is easy to see that this can be done with a path \( P_c \) of length at most \( \text{Diam}(G) \). In this way,

\[
P(x, y) := P_xP_cP_y.
\]

Clearly,

\[
r + q - 1 \leq \text{Length}(P(x, y)) \leq r + q + \text{Diam}(G) - 2
\]

On the other hand, now we prove that no shorter paths exists than \( P(x, y) \).

Recall that it follows from (2.2.1) that for any path \( Q(x, y) = (\bar{q}^0, \ldots, \bar{q}^\ell = 87 \).
the consecutive elements of the path only differ in their postfixes, which have different types. That is,

$$\forall i, q^i = w^i_z^i, q^{i+1} = w^i_{\tilde{z}^i}, \text{ with } \text{typ}(z^i) \neq \text{typ}(\tilde{z}^i) \in \{1, 2\}.$$

This implies that in each step on the path, the number of blocks in (2.3.18) changes by at most one. Recall that $|x \wedge y| = k$, 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 $\tilde{P}_0 = x$, to reach the first vertex $a$ of this property, we need at least $r-1$ steps on any path $\tilde{P}$, where $r$ was defined in formula (2.3.18). Similarly, starting from $y$, we need at least $q-1$ steps to reach the first vertex $b$ 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 $\text{Diam}(G)$ edges.

**Theorem 2.3.10.** The expectation of the length of a shortest path between two uniformly chosen vertices $X, Y \in G_n$ can be bounded by

$$\frac{4n_1n_2}{N^2}(n-1) < E(|P(X, Y)|) < N + \frac{4n_1n_2}{N^2}(n-1).$$

**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 $\text{Diam}(G_n)$.

**Proof of Theorem 2.3.10.** Let $X, Y$ 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 $\{0, \ldots, N-1\}$, hence $K(X, Y) := |X \wedge Y|$ has a truncated geometric distribution with parameter $\frac{N-1}{N}$. That is

$$\mathbb{P}(K(X, Y) = k) = \begin{cases} \left(\frac{1}{N}\right)^k \cdot \frac{N-1}{N}, & \text{if } 0 \leq k < n, \\ \left(\frac{1}{N}\right)^n & \text{if } k = n. \end{cases}$$

Furthermore, given that the length of the prefix is $k = K(X, Y)$, 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:

$$R = 1 + \sum_{i=1}^{n-k-1} 1_{\text{typ}(X_{k+i}) \neq \text{typ}(X_{k+i+1})},$$

$$Q = 1 + \sum_{i=1}^{n-k-1} 1_{\text{typ}(Y_{k+i}) \neq \text{typ}(Y_{k+i+1})}.$$

88
Taking expectation yields

\[ E(Q|K(X, Y) = k) = E(R|K(X, Y) = k) = \]

\[ = 1 + E \left( \sum_{i=1}^{n-k-1} 1_{typ(X_{k+i}) \neq typ(X_{k+i+1})} \right) \]

\[ = 1 + \sum_{i=1}^{n-k-1} P(typ(X_{k+i}) \neq typ(X_{k+i+1})) \]

\[ = 1 + (n - k - 1) \frac{2n_1n_2}{N^2}. \]

So weighting this with the geometric weights of the length of the prefix, we get

\[ E(Q) = E(R) = E(E(R|K(X, Y))) = \]

\[ = E\left(1 + (n - K(X, Y) - 1) \frac{2n_1n_2}{N^2}\right) \]

\[ = 1 + \left(n - \frac{1}{N-1} \left(1 - \frac{1}{N^n}\right) - 1\right) \frac{2n_1n_2}{N^2}. \]

Using this and the following immediate consequence of Lemma 2.3.8

\[-1 \leq E(|P(X, Y)| - (R + Q)) \leq \text{Diam}(G) - 2,\]

finally we obtain that

\[ 1 - \frac{1}{N-1} + \frac{4n_1n_2}{N^2} (n - 1) \leq E(|P(X, Y)| < \text{Diam}(G) + \frac{4n_1n_2}{N^2} (n - 1). \]

\[ \Box \]

2.3.4 Decay of local clustering coefficient of the modified sequence \( \{\hat{G}_n\} \)

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 give a rather flexible way of modifying the graph sequence \( G_n \) in a natural way, like in [76], to get
(a) We obtain $\hat{G}$ by adding the dashed (red) edges to $G$.

(b) $\hat{G}_2$: The edges of $\hat{G}_2$ and $G_2$ differ only at the lowest hierarchical level (cf. Figure 2.3).

Figure 2.6: Clustering extended "fan".
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 \( \text{RE}(\hat{G}) \) to \( G \) satisfying the following property:

**Property R**

\[ \forall x \in \Sigma_1, \exists y, z \in \Sigma_1, \text{ such that two among the edges of the triangle } (x, y, z)_{\Delta} \text{ are contained in } E(G) \text{ and one of the edges is in } \text{RE}(\hat{G}). \]

So,

\[ V(\hat{G}) = V(G) \text{ and } E(\hat{G}) = E(G) \cup \text{RE}(\hat{G}). \]

In the example presented on Figure 2.6 the edges from \( \text{RE}(\hat{G}) \) are the dashed red edges.

- Similarly we define the graph sequence \( \{ \hat{G}_n \}_{n=1}^{\infty} \) 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

\[ E(\hat{G}_n) = E(G'_n) \cup \text{RE}(\hat{G}_n), \quad (2.3.20) \]

where

\[ \text{RE}(\hat{G}_n) = \left\{ \left( x_1, \ldots, x_n, y_1, \ldots, y_n \right) \mid x_i = y_i, i \leq n - 1, \left( x_n, y_n \right) \in \text{RE}(\hat{G}) \right\}. \quad (2.3.21) \]

Note that this definition describes a lot of possible ways of modifying the graph sequence \( G_n \), what is important is only that for all the vertices of the base graph, we close at least one edge-pair leaving the vertex to form a triangle. This means that the local clustering coefficient of every vertex in the base graph is modified to be positive.

Thus, it is clear from Property R that

\[ \hat{C}_{\min} := \min_{x \in \hat{G}} C_x > 0. \quad (2.3.22) \]

Further, using (2.2.1) and (2.3.21) one can easily see that the degree of a vertex \( \hat{x} \in \hat{G}_n \) is

\[ \hat{\deg}(\hat{x}) = S(\hat{x}) \cdot \deg(x_n) + \left( \hat{\deg}(x_n) - \deg(x_n) \right), \quad (2.3.23) \]

where \( \hat{\deg}(\cdot) \) denotes the degree of a vertex in \( \hat{G} \), while \( \deg(\cdot) \) stands for the degree in \( G \).
Remark 2.3.13. The difference between the degree of any node \( x \in \Sigma_n \) 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 \( C_x \) of an arbitrary node \( x \in \hat{G}_n \) satisfies

\[
\frac{K_1}{\deg_n(x)} \leq C_x \leq \frac{K_2}{\deg_n(x)}.
\]

Proof. We write \( T_n(x) \) for the set of all triangles in \( \hat{G}_n \) containing the node \( x \in \Sigma_n \). We say that a triangle \( (x, y, z) \Delta \in T_n(x) \) is regular if and only if exactly two of its edges are from \( E(G_n) \). The triangle \( (x, y, z) \Delta \in T_n(x) \) is called irregular if it is not regular. The set of irregular triangles containing \( x \) is denoted by \( IRT_n(x) \). We partition the set of regular triangles \( RT_n(x) \) into the classes:

\[
RT_n(x) = RT^1_n(x) \cup RT^2_n(x)
\]

in the following way: A triangle \( (x, y, z) \Delta \in RT_n(x) \) belongs to \( RT^1_n(x) \) if and only if \( x \) is NOT an endpoint of the edge contained in \( RE(\hat{G}_n) \). That is

\[
RT^1_n(x) := \{(x, y, z) \Delta \in RT_n(x) : (\not x, y) \in E(G_n)\}.
\]

Hence, \( RT^2_n(x) \) is the set of those \( (x, y, z) \Delta \in RT_n(x) \) for which either \( (\not x, y) \Delta \in RT_n(x) \) or vice versa. Summarizing these partitions:

\[
T_n(x) = RT_n(x) \cup IRT_n(x) = RT^1_n(x) \cup RT^2_n(x) \cup IRT_n(x)
\]

Now we define the cardinality of these classes:

\[
\Delta^1_n(x) := \#RT^1_n(x), \Delta^2_n(x) := \#RT^2_n(x) \text{ and } \Delta^1_n(x) := \#IR^n(x).
\]

When \( n = 1 \) then we suppress the index \( n \). Observe that by Property R,

\[
\Delta^1_n(x) := \Delta^1(x) + \Delta^2_n(x) \geq 1, \quad \forall n \geq 1, x \in \Sigma_n.
\]

Now we compute \( \Delta^1_n(x), i \in \{1, 2, \text{ir} \} \), for an arbitrary fixed \( x \in \Sigma_n \). To do so the notation \( \ell(x) \) will be used. First we verify that

\[
\Delta^1_n(x) = \sum_{r=0}^{\ell(x)-1} \prod_{j=1}^{r} \deg(x_{n-j}) \cdot \Delta^1(x_n) = S(x) \cdot \Delta^1(x_n), \quad (2.3.24)
\]

where \( S(x) \) was defined in (2.2.3). To see this, observe that it follows from (2.2.1), (2.3.20) and (2.3.21) that

\[
(x, y, z) \Delta \in RT^1_n(x)
\]

holds if and only if all of the following three assertions are satisfied:
1. \(0 \leq r \leq \ell(x) - 1, \ |y \wedge z| = n - 1 \) and \(|x \wedge y| = |x \wedge z| = n - r - 1\)

2. \((x_k, y_k) \in E(G)\) whenever \(n - r \leq k \leq n - 1\)

3. \((x_n, y_n, z_n)_\Delta \in \mathcal{RT}^1(x_n)\).

Hence (2.3.24) is obtained by an immediate calculation.

Now we prove that

\[
\Delta_n^2(x) = \sum_{r=0}^{\ell(x) - 1} \prod_{j=1}^{r} \deg(x_{n-j}) \cdot \Delta^2(x_n) = S(x) \cdot \Delta^2(x_n). \tag{2.3.25}
\]

This is so because by (2.2.1), (2.3.20) and (2.3.21) we have

\[
(x, y, z)_\Delta \in \mathcal{RT}_n^2(x)
\]

holds if and only if all of the following three assertions are satisfied:

1. \(0 \leq r \leq \ell(x) - 1, \ |x \wedge y| = n - 1 \) and \(|x \wedge z| = |y \wedge z| = n - r - 1,\)

2. \((x_k) \in E(G)\) 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 (2.3.25).

Finally, we determine the number of irregular triangles containing \(x\):

\[
\Delta^\text{ir}_n(x) = \Delta^\text{ir}_n(x_n). \tag{2.3.26}
\]

This follows from the fact that

\[
(x, y, z)_\Delta \in \mathcal{RT}_n(x)
\]

is equivalent to

\[
\forall 1 \leq i \leq n - 1, \ x_i = y_i = z_i \text{ and } (x_n, y_n, z_n)_\Delta \in \mathcal{RT}(x_n).
\]

We write \(Z_\Delta(x)\) for the number of all triangles in \(\hat{G}_n\) containing \(x\):

\[
Z_\Delta(x) := \Delta_1^\text{ir}(x) + \Delta_n^\text{ir}(x) + \Delta^\text{ir}_{n-1}(x).
\]

Using (2.3.23), (2.3.24), (2.3.25) and (2.3.26) we get

\[
C_x = \frac{Z_\Delta(x)}{\deg_n(x)} = \frac{2\Delta^\text{ir}(x_n) \cdot S(x) + 2\Delta^\text{ir}_n(x_n)}{\deg_n(x)(\deg_n(x) - 1)}, \tag{2.3.27}
\]

where \(S(x)\) was defined in (2.2.3). Now we estimate \(C_x\).
Claim 2.3.15.

(i) If $\ell(x) = 1$, then $C_x = C_{x^n}$.

(ii) If $\ell(x) \geq 2$, then we have

$$\left| C_x - \frac{2\Delta^r(x_n)}{\deg(x_n)} \cdot \frac{1}{\deg_n(x)} \right| \leq \frac{\text{const}}{\deg_n(x)^2}. \quad (2.3.28)$$

Proof of the Claim. Part (i) immediately follows from (2.2.1). To prove (ii) we fix an arbitrary $x \in \Sigma_n$ with $\ell(x) \geq 2$. Since $t, u, v$ introduced below depend only on $x_n$ there exists a constant $C_*$ independent of $n$ and $x$ such that

$$0 \leq t := \frac{\Delta^r(x)}{\deg(x_n)}, \quad u := \deg(x_n) - \deg(x_n), \quad v := 2\Delta^r(x_n) < C_*.$$  \quad (2.3.29)

To prove (2.3.28) it is enough to verify that

$$Q := \left( \deg_n(x) \right) \left( \deg_n(x) - 1 \right) \cdot C_x - 2t \cdot (\deg_n(x) - 1)$$

is bounded in $n$ and $x \in \Sigma_n$. This so, because by (2.3.23) and (2.3.27) we have

$$Q = 2\Delta^r(x_n) \cdot S + v - 2t \left( S \cdot \deg(x_n) + u - 1 \right)$$

$$= 2\Delta^r(x_n) \cdot S + v - 2\Delta^r(x_n) \cdot S - 2t(u - 1)$$

$$= v - 2t(u - 1),$$

which is bounded by (2.3.29). \qed

Property $R$ implies that both $C_{x^n}$ and $\frac{\Delta^r(x_n)}{\deg(x_n)}$ are bounded away from zero. This completes the proof of the Theorem 2.3.14. \qed

The following theorem shows that the graph sequence $\hat{G}_n$ displays features similar to those of the sequence considered in [76], 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

$$\frac{2n_1 n_2 \hat{C}_{\min}}{N^2} \leq \bar{C}(\hat{G}_n) \leq \bar{C}(\hat{G}), \quad (2.3.30)$$

where $\hat{C}_{\min}$ was defined in (2.3.22).
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
\[ C_{\bar{x}} \leq C_{x_n}. \] (2.3.31)
Namely, if \( \ell(x) = 1 \) then by (2.3.21), \( C_{\bar{x}} = C_{x_n} \). If \( \ell(x) \geq 2 \) then \( S(x) \geq 1 \) thus using (2.3.27) we obtain
\[ C_{\bar{x}} \leq \Delta^t(x_n) + \Delta^{ir}(x_n) \left( \frac{\deg(x_n)}{2} \right) \left( \frac{S(x)}{\deg(x_n)} \right) \leq C_{x_n}. \]
This completes the proof of (2.3.31) from which the upper estimate of (2.3.30) follows by averaging. On the other hand to see that the lower estimate holds we take into consideration only the contribution of \( x \in \Sigma_n \) with \( \ell(x) = 1 \).

\[ \bar{C}(\hat{G}_n) > \frac{1}{N^n} \left( \sum_{z \in V_1} N^{n-2} n_2 C_z + \sum_{z \in V_2} N^{n-2} n_1 C_z \right) \]
Using \( C_z > \hat{C}_{\min} \), the lower bound of (2.3.30) follows. \( \square \)

2.4 The randomized model

In this section we randomize the deterministic model in Section 2.2 by using \( \Lambda \) in \([0, 1]^2\). The random graph sequence \( G^*_n \) is generated in a way which was inspired by the \( W \)-random graphs introduced by Lovász and Szegedy [73]. See also [53].

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 \([0, 1]\) (the value of \( M_n \) will be determined later): \( X^{(1)}, X^{(2)}, \ldots, X^{(M_n+1)} \sim U[0, 1], \) i.i.d.

We denote the \( N \)-adic expansion of each of these numbers by
\[ X^{(i)} = (X_1^{(i)}, X_2^{(i)}, \ldots), \] i.e. \( X^{(i)} = \sum_{k=1}^{\infty} X_k^{(i)} / N^k \),
where the \( X_k^{(i)} \)-s are uniform over the set \( \{0, 1, \ldots, N-1\} \). The n-th approximation of \( X^{(i)} \) is
\[ X^{(i)}_{[n]} = \sum_{k=1}^{n} X_k^{(i)} / N^k, \quad X^{(i)}_n = (X_1^{(i)}, \ldots, X_n^{(i)}). \]
Now we construct the random graph $G_r^n$ as follows: $|V(G_r^n)| = \{1, \ldots, M_n\}$, and $E(G_r^n)$ is given by

$$E(G_r^n) = \left\{ (i, j) \mid \text{int} \left( I_{X_i^1} \times I_{X_j^1} \right) \cap \Lambda \neq \emptyset \right\},$$

where $\text{int}$ denotes the interior of a set. Clearly,

$$E(G_r^n) = \left\{ (i, j) \mid \Lambda_n(X^{(i)}, X^{(j)}) = 1 \right\}.$$

Note that

$$\Lambda_n(X^{(i)}, X^{(j)}) = 1 \iff \left( X_i^1 \ldots X_i^n \right) \in E(G_n).$$

Namely, we can think of the first $n$ digits $(X_i^1, \ldots, X_i^n)$ and $(X_j^1, \ldots, X_j^n)$ 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 $G_r^n$ if the vertices $(X_i^1 \ldots X_i^n)$ and $(X_j^1 \ldots X_j^n)$ 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 $G_r^n$ if and only if the urns of ball $i$ and $j$ are connected by an edge in $G_n$.

We need to introduce some further notation.

**Frequently used definitions.** Under assumption $[A1]$ for an $x \in G_n$ with $\ell(x) = k$ the degree of $x$ is

$$t_k := \frac{d_1^{k+1} - 1}{d_1 - 1} + 1,$$

independently of $n$.

In the random graph $G_r^n$, the conditional probability of the degree distribution of a random node $V \in \{0, \ldots, M_n\}$ conditioned on the first $n$ digits of the N-adic expansion of the corresponding code $X^{(V)}$ follows a Binomial distribution:

$$\left( \operatorname{deg}(V) \mid (X_1^V \ldots X_n^V) = x \right) \sim \operatorname{BIN} \left( M_n, \frac{t_{\ell(x)}}{N_n} \right). \quad (2.4.1)$$

This follows from the characterization of $G_r^n$ described in Remark 2.4.1. Namely, assume that the $V$-th ball has landed in urn with label $x \in \Sigma_n$. In $G_n$ there are exactly $\deg_n(x) - 1 = t_{\ell(x)}$ vertices $y \in \Sigma_n$ that are connected to $x$. All the balls landing into urns corresponding to these vertices $y$ will be connected to $V$ in $G_r^n$.
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 $V \in G_n^r$, we have

$$P(\text{deg}(V) = 0) \leq e^{-d_{\text{min}} c_n},$$

where $d_{\text{min}}$ stands for the minimal degree in the base graph $G$, and in $\text{deg}(.)$ we do not count the loops.

The following corollary is an immediate consequence of the Borel-Cantelli lemma.

**Corollary 2.4.3.** If $\sum_{n=1}^{\infty} c_n N^n e^{-d_{\text{min}} c_n} < \infty$, 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 $(X_1^V \ldots X_n^V)$ has in the deterministic model. So we can write

$$P(\text{deg}(V) = 0) = \sum_{\bar{x} \in \Sigma_n} P(\text{deg}(V) = 0 | (X_1^V \ldots X_n^V) = \bar{x}) \cdot \frac{1}{N^n}$$

As we have already seen, $(\text{deg}(V)| (X_1^V \ldots X_n^V) = \bar{x})$ follows a Binomial distribution with parameters $M_n$ and $\frac{\text{deg}_n(\bar{x})-1}{N^n}$, so the conditional probability of isolation is

$$P(\text{deg}(V) = 0 | (X_1^V \ldots X_n^V) = \bar{x}) = \left(1 - \frac{t_\ell(\bar{x})}{N^n}\right)^{M_n} \leq e^{-\text{deg}_n(\bar{x}) c_n} (1 + o(1)).$$

Obviously $e^{-\text{deg}_n(\bar{x}) c_n} \leq e^{-d_{\text{min}} c_n}$ holds for all $\bar{x} \in \Sigma_n$, which completes the proof. \qed
Decay of degree distribution

Fix a constant $K$ such that for a standard normal variable $Z$, $P(|Z| > K) < e^{-10}$. 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. It turns out that the degree distribution for large enough $M_n$-s is basically the same as for the deterministic, just the mass around $t_k = (d_{k+1} - 1)/(d_1 - 1) + 1$ is spread out a ”bit” in a Gaussian way. The next theorem and corollary make these heuristics more precise.

**Theorem 2.4.4.** Let $k > k_0(n)$ and $u \in I_{k,n}$. Then for a uniformly chosen node $V$ in $G^r_n$

$$P(\deg(V) = u) = \left(\frac{n_1}{N}\right)^k \frac{n_2}{N} \frac{1}{\sqrt{c_n t_k}} \phi\left(\frac{u - c_n t_k}{\sqrt{c_n t_k(1 - \frac{t_k}{N})}}\right)(1 + O\left(\frac{1}{\sqrt{c_n t_k}}\right)),$$

where $\phi$ 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 $a, b \in [-K, K]$:

$$P(\deg(V) \in [c_n t_k + a\sqrt{c_n t_k}, c_n t_k + b\sqrt{c_n t_k}]) = \left(\frac{n_1}{N}\right)^k \frac{n_2}{N} \left(\Phi(b) - \Phi(a)\right) + O\left(\frac{1}{\sqrt{c_n t_k}}\right),$$

where $k > k_0(n)$ and $\Phi$ 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:

$$P(\deg(V) > u) = \left(\frac{n_1}{N}\right)^{k+1} + \left(\frac{n_1}{N}\right)^k \frac{n_2}{N} \left(1 - \Phi\left(\frac{u - c_n t_k}{\sqrt{c_n t_k(1 - \frac{t_k}{N})}}\right)\right) + \left(\frac{n_1}{N}\right)^{k+1} O\left(\frac{1}{\sqrt{c_n t_k}}\right).$$

This holds because $P(\deg(V) > u)$ 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
\[ \gamma := 1 + \frac{\log(N)}{\log d_1}. \]
Then the decay of the degree distribution is:
\[ P(\deg(V) > u) = u^{-\gamma+1} \cdot L(u), \]
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 \( X^V_n = x \) follows a \( BIN(c_nN^n, \frac{t(\ell(x))}{N^n}) \) law. This is close to a \( POI(c_n t(\ell(x))) \) random variable, because \( c_n \) and \( t(\ell(x)) \) tend to infinity in a much smaller order than \( N^n \).

Now for the \( POI(c_n t(\ell(x))) \) variable, the Central Limit Theorem holds with an error term of order \( 1/\sqrt{c_n t(\ell(x))} \). 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
\[ p_k := P(\ell(x) = k) = \left( \frac{n_1}{N} \right)^k \frac{n_2}{N}. \]
We show that in the random model \( G^n \), 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
\[ P(\deg(V) = u) = P(\deg(V) = u | (X^V_1 \ldots X^V_n) = x, \ell(x) = k) \cdot p_k + S_1 + S_2, \tag{2.4.3} \]
where
\[ S_1 = \sum_{j=1}^{k-1} P(\deg(V) = u | (X^V_1 \ldots X^V_n) = x, \ell(x) = j) \cdot p_j \]
\[ S_2 = \sum_{j=k+1}^{n} P(\deg(V) = u | (X^V_1 \ldots X^V_n) = x, \ell(x) = j) \cdot p_j \]
\( S_1 \) and \( S_2 \) combines the total contribution of cases when \( \ell(X^V_1 \ldots X^V_n) \neq k \), 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 $u \in I_{k,n}$

$$\Pr \left( \deg(V) = u \left| (X_1^V \ldots X_n^V) = \bar{x} \right. \right) = \frac{1}{\sqrt{c_n t_{\ell}(\bar{x}) (1 - \frac{t_{\ell}(\bar{x})}{N})}} \phi \left( \frac{u - c_n t_{\ell}(\bar{x})}{\sqrt{c_n t_{\ell}(\bar{x}) (1 - \frac{t_{\ell}(\bar{x})}{N})}} \right) \cdot \left( 1 + O \left( \frac{1}{\sqrt{c_n t_{\ell}(\bar{x})}} \right) \right).$$

We can neglect $1 - \frac{t_{\ell}(\bar{x})}{N}$. This completes the first step.

**Second step:** Since $u \in I_{k,n}$ we have:

$$S_1 \leq \sum_{j=1}^{k-1} \Pr(\deg(V) > t_k - K \sqrt{t_k} | (X_1^V \ldots X_n^V) = \bar{x}, \ell(\bar{x}) = j) \cdot p_j$$

$$S_2 \leq \sum_{j=k+1}^{n} \Pr(\deg(V) < t_k + K \sqrt{t_k} | (X_1^V \ldots X_n^V) = \bar{x}, \ell(\bar{x}) = j) \cdot p_j$$

(2.4.4)

Now we use the fact known from Chernoff-bounds: for an $Z \sim BIN(m,p)$ variable

$$\Pr(Z \geq (1 + \delta)E(Z)) \leq e^{-\frac{1}{2} \delta^2 E(Z)},$$

and the same bound holds for $\Pr(Z \leq (1 - \delta)E(S))$. By (2.4.1), to estimate each summand in (2.4.4) we can apply these inequalities for $Z_j \sim BIN(c_n N^n, \frac{d_j}{N^n})$, $j \in \{1, \ldots, n\} \setminus \{k\}$, yielding an upper bound

$$S_1 + S_2 \leq \sum_{j=1}^{k-1} e^{-\frac{1}{2} d_1^2 c_n} \cdot p_j + \sum_{j=k+1}^{n} e^{-\frac{1}{2} (1 - d_1^{k-j})^2 d_1^2 c_n} \cdot p_j \leq e^{-\frac{1}{2} d_1^2 c_n}.$$

Since $e^{-\frac{1}{2} d_1^2 c_n} = o\left(\frac{1}{\sqrt{c_n t_k}}\right)$, the statement of Theorem 2.4.4 follows. \qed

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 $C(u) \in \left[\frac{n_1}{N}, 1\right]$ such that

$$\Pr(\deg(V) > u) = \left( \frac{n_1}{N} \right)^k C(u).$$
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 \( u \not\in \cup_k I_{n,k} \), 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

\[
P(\deg(V) > c_n t_{k+1}) \leq P(\deg(V) > u) \leq P(\deg(V) > c_n t_k).
\]

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. \( \square \)

**Further direction**

It is an interesting question to relate the Hausdorff dimension of the limiting fractal \( \Lambda \) to the box-cover dimension of the graph sequence, where the latter is meant as in [80] and in [66]. We plan to work out this project with a student soon.

**Acknowledgement**

The research for this paper was supported by the NKTH OTKA grant # 7778. We would like to say thanks to Béla Bollobás and Tamás Vicsek for useful conversations.

**Bibliography**


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

\[ h(t) = \{h_i(t)\}_{i \in \mathbb{Z}} \]

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 \( \mu^\rho \) for the increments \( \{h_{i-1} - h_i : i \in \mathbb{Z}\} \). The family of invariant measures is indexed by the average slope \( \rho = \mathbb{E}^\rho(h_{i-1} - h_i) \). The flux function \( H(\rho) = t^{-1}\mathbb{E}^\rho(h_i(t) - h_i(0)) \) gives the average velocity of the height as a function of the slope \( \rho \). In this chapter we consider asymmetric systems, for which \( H''(\rho) < 0 \) holds additionally at least in a neighborhood of a particular density value \( \rho \). Asymmetry here always mean spatial asymmetry, i.e. models in which the jump rates for deposition differ from those for removal.

The sum of 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 \( \rho \) 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 \( \varrho \) and consider \( \hat{h}(t) \) with stationary increments at average slope \( \varrho \), normalized so that \( h_0(0) = 0 \). Interesting fluctuations can be found by observing the height \( h_{[V \varrho t]}(t) \) in the characteristic direction \( V\varrho := \mathcal{H}'(\varrho) \). (Later we will see that this particular speed for an observer causes interesting fluctuations for the height function, and other velocities give normal fluctuations.) 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 \([112, 119, 124]\) and the random average process \([102, 114]\).

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 [93] for the last-passage version of the Hammersley process and in [118] for the last-passage model associated with TASEP. Current fluctuations for stationary TASEP were analyzed in [115]. Here is a selection of further results in this direction: [94, 109, 116, 117, 122].

(b) Probabilistic approaches exist to prove fluctuation bounds of the correct order. The seminal work [110] was on the last-passage version of the Hammersley process, and then the approach was adapted to the last-passage model associated with TASEP [98]. The next step was the development of a proof that works for particle systems: the asymmetric simple exclusion process (ASEP) was treated in [106] and the totally asymmetric zero range process with constant jump rate in [99]. The ASEP work [106] 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 [101], which takes a further step toward universality of the $t^{1/3}$ order for fluctuations in the case $\mathcal{H}''(\rho) \neq 0$. In [101] 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 [100] 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 $\mathcal{H}''(\rho) < 0$, hence we name the property microscopic concavity, then in Section 3.7 we show that the same strategy also works for a convex model, the point not being the modification from concave to convex, but to check the exact convexity assumptions in that model.

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 byproduct, we also obtain superdiffusivity of the second class particle in the stationary process. Mostly, (but not including [106]) earlier proofs of $t^{1/3}$ fluctuations have been quite rigid in the sense that they work only for particular cases of 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 a 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 a broad class. The estimates that it requires might be proved in different ways for different further subclasses of processes. And the general 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 Balázs and Seppäläinen [105] and we give a brief informal description in Section 3.2.8. This proof is somewhat simpler than the earlier one given in [106].

(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. Earlier, totally asymmetric constant rate zero range processes were handled in [99], as the first generalization of the proof in [106] for processes with more than one allowed particle per site. The proof given here is simpler than the one in [99]. We expect that a broader class of totally and not 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.

(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.

A comment on NOT totally asymmetric models: by now, the only model in this category for which $t^{1/3}$ fluctuations are proved is the asymmetric simple exclusion process, treated in [106]. Note that the general proof given the microscopic concavity would work also for these models, thus what is left is to verify the criterions of microscopic concavity for asymmetric models. In many cases, we already do have a proper coupling described below, only the distributional bound (3.2.29) below is missing.

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 the 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 (as opposed to fluctuations generated by the dynamics). 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 [105].) 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 $\mathcal{H}$. In particular we show monotonicity of the one-site marginal measures of particles $\mu$ and $\hat{\mu}$ in $\rho$ 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 $\mathcal{H}''(\rho) < 0$ for all densities $0 < \rho < \infty$. We omit the very first part of the Appendix of [101] showing that the hydrodynamic flux function is a convex function of the density $\rho$, and we refer the reader for a more probabilistic proof to [103] or to the omitted Appendix of [101].

Notation

We summarize here some notation for easy reference. $\mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$, $\mathbb{R}_{\geq 0} = [0, \infty)$. Centering a random variable is denoted by $\tilde{X} = X - EX$. Constants $C_i, \alpha_i$ do not depend on time, but may depend on the density parameter $\rho$ 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 [104], 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 $-\infty \leq \omega_{\min} \leq 0$ and $1 \leq \omega_{\max} \leq \infty$ define the single-site state space

$$I := \{ z \in \mathbb{Z} : \omega_{\min} - 1 < z < \omega_{\max} + 1 \}$$
and the increment configuration space

$$\Omega := \{ \omega = (\omega_i)_{i \in \mathbb{Z}} : \omega_i \in I \} = I^\mathbb{Z}. $$

At times it will be convenient to have notation for the increment configuration \( \delta_i \in \Omega \) with exactly one nonzero entry equal to 1:

$$(\delta_i)_j = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{for } i \neq j. \end{cases} \quad (3.2.1)$$

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 \( \omega \in \Omega \) are the negative discrete gradients of the heights: \( \omega_i = h_i - 1 - h_{i+1} \in I \).

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:

$$(\omega_i, \omega_{i+1}) \rightarrow (\omega_i - 1, \omega_{i+1} + 1) \quad h_i \rightarrow h_i + 1$$

with rate \( p(\omega_i, \omega_{i+1}) \), \( (3.2.2) \)

or removed:

$$(\omega_i, \omega_{i+1}) \rightarrow (\omega_i + 1, \omega_{i+1} - 1) \quad h_i \rightarrow h_i - 1$$

with rate \( q(\omega_i, \omega_{i+1}) \). \( (3.2.3) \)

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 \( \omega(\cdot) \):

$$(L\varphi)(\omega) = \sum_{i \in \mathbb{Z}} p(\omega_i, \omega_{i+1}) \cdot [\varphi(\omega_i - 1, \omega_{i+1} + 1, \ldots) - \varphi(\omega)] + \sum_{i \in \mathbb{Z}} q(\omega_i, \omega_{i+1}) \cdot [\varphi(\omega_i + 1, \omega_{i+1} - 1, \ldots) - \varphi(\omega)]. \quad (3.2.4)$$

\( L \) acts on bounded cylinder functions \( \varphi : \Omega \rightarrow \mathbb{R} \) (this means that \( \varphi \) depends only on finitely many \( \omega_i \)-values).

Thus we have a Markov process \( \{\omega(t) : t \in \mathbb{R}_{\geq 0}\} \) of an evolving increment configuration and a Markov process \( \{h(t) : t \in \mathbb{R}_{\geq 0}\} \) of an evolving height configuration. The initial increments \( \omega(0) \) specify the initial height \( h(0) \) 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 $\omega$-values are permitted, one needs to consider particles and antiparticles, with antiparticles jumping in the opposite direction.) It will be useful to see that

$$h_i(t) = h_i(t) - h_0(0) \text{ the net number of particles that have passed,}$$

from left to right, the straight-line space-time path that connects $(1/2, 0)$ to $(i + 1/2, t)$. 

$$(3.2.5)$$

We impose the following four assumptions (3.2.6)–(3.2.9) on the rates.

- The rates $p, q : I \times I \to \mathbb{R}_{\geq 0}$ must satisfy

$$p(\omega_{\text{min}}, \cdot) \equiv p(\cdot, \omega_{\text{max}}) \equiv q(\omega_{\text{max}}, \cdot) \equiv q(\cdot, \omega_{\text{min}}) \equiv 0 \quad (3.2.6)$$

whenever either $\omega_{\text{min}}$ or $\omega_{\text{max}}$ 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:

$$p(z + 1, y) \geq p(z, y), \quad p(y, z + 1) \leq p(y, z)$$

$$q(z + 1, y) \leq q(z, y), \quad q(y, z + 1) \geq q(y, z) \quad (3.2.7)$$

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.)

- For any $x, y, z \in I$

$$p(x, y) + p(y, z) + p(z, x) + q(x, y) + q(y, z) + q(z, x) = p(x, z) + p(z, y) + p(y, x) + q(x, z) + q(z, y) + q(y, x). \quad (3.2.8)$$

- There are symmetric functions $s_p$ and $s_q$ on $I \times I$, and a function $f$ on $I$ such that $f(\omega_{\text{min}}) = 0$ whenever $\omega_{\text{min}}$ is finite, $f(z) > 0$ for $z > \omega_{\text{min}}$, and for any $y, z \in I$,

$$p(y, z) = s_p(y, z + 1)f(y)$$

and

$$q(y, z) = s_q(y + 1, z)f(z). \quad (3.2.9)$$

(Interpret $s_p(y, z) = s_q(y, z) = 0$ if $y$ or $z > \omega_{\text{max}}$.) Condition (3.2.7) implies that $f$ is nondecreasing on $I$. 

112
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 [121], and articles [91], [107] and [120].) 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 $\omega(t)$, and the corresponding height process $h(t)$ 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 $\Omega$ of tempered increment configurations (i.e. configurations that obey some restrictive growth conditions).
- The subspace $\Omega$ is of full measure under the invariant distributions $\mu$ defined in Section 3.2.4.
- It is also possible to construct jointly several versions of the process with initial configurations from the space $\Omega$ and with joint evolution obeying basic coupling (described in Section 3.2.3).
- Rates $p$ and $q$ have all moments under the invariant distributions $\mu$.

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 [104] 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 [104] 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 ([105] 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 [104].

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 \(0 \leq p, q \leq 1\) such that \(p + q = 1\). \textbf{Asymmetric processes} have \(p \neq q\). These are the processes for which our results are relevant.

1. \textbf{Generalized exclusion processes}. These are the cases where both \(\omega_{\text{min}}\) and \(\omega_{\text{max}}\) are finite.

   - \textbf{The asymmetric simple exclusion process (ASEP)} introduced by F. Spitzer \[125\] is defined by \(\omega_{\text{min}} = 0, \omega_{\text{max}} = 1\), \(f(z) = 1\{z = 1\}\), \(s_p(y, z) = p \cdot 1\{y = z = 1\}\) and \(s_q(y, z) = q \cdot 1\{y = z = 1\}\). This produces the familiar rates \(p(y, z) = p \cdot 1\{y = 1, z = 0\}\) and \(q(y, z) = q \cdot 1\{y = 0, z = 1\}\).

   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.

   - \textbf{Particle-antiparticle exclusion process}. Let \(\omega_{\text{min}} = -1, \omega_{\text{max}} = 1\). Take \(f(-1) = 0, f(0) = c\) (creation), \(f(1) = a\) (annihilation) where \(c\) and \(a\) are positive rates with \(c \leq a/2\),
     \[s_p(0, 1) = s_p(1, 0) = p, \quad s_p(0, 0) = \frac{pa}{2c}, \quad s_p(1, 1) = \frac{p}{2},\]
     \[s_q(0, 1) = s_q(1, 0) = q, \quad s_q(0, 0) = \frac{qa}{2c}, \quad s_q(1, 1) = \frac{q}{2},\]
     and \(s_p, s_q\) zero in all other cases. This result in rates
     \[p(0, 0) = pc, \quad p(0, -1) = p(1, 0) = \frac{pa}{2}, \quad p(1, -1) = pa,\]
     \[q(0, 0) = qc, \quad q(-1, 0) = q(0, 1) = \frac{qa}{2}, \quad q(-1, 1) = qa\]
     and zero in all other cases. If \(\omega_i = -1\) 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. \textbf{Generalized misanthrope processes} have \(\omega_{\text{min}} > -\infty, \omega_{\text{max}} = \infty\).
• **Zero range process.** Take $\omega_{\text{min}} = 0, \omega_{\text{max}} = \infty$, an arbitrary nondecreasing function $f : \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $f(0) = 0$,
\[
s_p(y, z) \equiv p \quad \text{and} \quad s_q(y, z) \equiv q,
\]
\[
p(y, z) = pf(y) \quad \text{and} \quad q(y, z) = qf(z).
\]
Again, $\omega_i$ 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 $\omega_{\text{min}} = -\infty$ and $\omega_{\text{max}} = \infty$. The height differences between adjacent columns can be arbitrary integers. Antiparticles are needed for a particle representation of the process.

• **Bricklayers process.** Let $f : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ be non-decreasing and satisfy
\[
f(z) \cdot f(1-z) = 1 \quad \text{for all } z \in \mathbb{Z}.
\]
The values of $f$ for positive $z$’s thus determine the values for non-positive $z$’s. Let
\[
s_p(y, z) = p + \frac{p}{f(y)f(z)} \quad \text{and} \quad s_q(y, z) = q + \frac{q}{f(y)f(z)},
\]
which results in
\[
p(y, z) = pf(y) + pf(-z) \quad \text{and} \quad q(y, z) = qf(-y) + qf(z).
\]
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 $\omega_{\text{min}} = 0$, $\omega_{\text{max}} = K$, $f(z) = 1\{z > 0\}$,
\[
s_p(y, z) = s_q(y, z) = 1\{z, y \leq K\}.
\]
These result in
\[ p(y, z) = q(z, y) = 1 \{ y > 0, \ z < K \} \]

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 \( \rho \in [0, K] \) 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 \( \omega^m(\cdot), \ m = 1, \ldots, n \), 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 \( \omega^1, \omega^2, \ldots, \omega^n \in \Omega \). Consider a step of type \eqref{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:

\[
 r(m) \equiv p(\omega^{\ell(m)}_i, \omega^{\ell(m)}_{i+1}) \leq p(\omega^{\ell(m+1)}_i, \omega^{\ell(m+1)}_{i+1}) \equiv r(m+1), \quad 1 \leq m < n.
\]

Set also the dummy value \( r(0) = 0 \). Now the rule is that independently for each \( m = 1, \ldots, n \), at rate \( r(m) - r(m-1) \), precisely processes \( \omega^{\ell(m)}_i \) and \( \omega^{\ell(m+1)}_i \), \( \ldots, \omega^{\ell(n)}_i \) execute the move \eqref{3.2.2}, and the processes \( \omega^{\ell(1)}_i, \omega^{\ell(2)}_i, \ldots, \omega^{\ell(m-1)}_i \) do not. The combined effect of these joint rates creates the correct marginal rates, that is, process \( \omega^{\ell(m)}_i \) executes this move with rate \( r(m) \).

Notice also that, due to \eqref{3.2.7}, a jump of \( \omega^a \) without \( \omega^b \) can only occur if \( p(\omega^a_i, \omega^a_{i+1}) < p(\omega^b_i, \omega^a_{i+1}) \) which implies \( \omega^a_i > \omega^b_i \) or \( \omega^a_{i+1} < \omega^b_{i+1} \). The result of this step \eqref{3.2.2}, then cannot increase the number of discrepancies between the two processes, hence the name *attractivity* for \eqref{3.2.7}. Further, \eqref{3.2.7} and the coupling implies that a sitewise ordering \( \omega^a_i \leq \omega^b_i \forall i \in \mathbb{Z} \) is preserved by the basic coupling.

One can check that moves of type \eqref{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 \( \omega^a_i \geq \omega^b_i \) for each \( i \in \mathbb{Z} \), then the second class particles are conserved. In view of \eqref{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

\[
 (h^a_i(t) - h^a_0(0)) - (h^b_i(t) - h^b_0(0)) = h^a_i(t) - h^b_i(t) \quad (3.2.10)
\]
between the two processes $\omega^a(\cdot)$ and $\omega^b(\cdot)$.

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 [111], 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

$$\bar{\theta} := \begin{cases} \log \left( \liminf_{z \to \infty} (f(z)!)^{1/z} \right) = \lim_{z \to \infty} \log(f(z)), & \text{if } \omega^\text{max} = \infty \\ \infty, & \text{else} \end{cases}$$

and

$$\theta := \begin{cases} \log \left( \limsup_{z \to \infty} (f(-z)!)^{1/z} \right) = \lim_{z \to \infty} \log(f(-z)), & \text{if } \omega^\text{min} = -\infty \\ -\infty, & \text{else}. \end{cases}$$

By monotonicity of $f$, we have $\bar{\theta} \geq \theta$. The case $\bar{\theta} = \theta$ would imply that $\omega^\text{min} = -\infty$, $\omega^\text{max} = \infty$, 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

Assume $f$ to be such that $\bar{\theta} < \bar{\theta}$. (3.2.11)

For $\theta \in (\bar{\theta}, \bar{\theta})$ define the state sum

$$Z(\theta) := \sum_{z \in I} \frac{e^{\theta z}}{(f(z)!)^z} < \infty. \quad (3.2.12)$$

117
Let the product-distribution \( \mu^\theta \) on \( \Omega = I^\mathbb{Z} \) have marginals

\[
\mu^\theta(z) = \mu^\theta(\{\omega : \omega_i = z\}) = \frac{1}{Z(\theta)} \cdot \frac{e^{\theta z}}{f(z)!} \quad (z \in I).
\] (3.2.13)

Assumptions (3.2.6), (3.2.7), (3.2.8), (3.2.9) imply that for \( \theta \in (\tilde{\theta}, \bar{\theta}) \) the product distribution \( \mu^\theta \) is stationary for the process generated by (3.2.4) (see [104]). Note that the family \( \{\mu^\theta\} \) can be obtained by exponentially weighting a probability measure \( \mu^{\theta_0} \) for a fixed value \( \theta_0 \in (\tilde{\theta}, \bar{\theta}) \), see [103].

\( P^\theta, E^\theta, \text{Var}^\theta, \text{Cov}^\theta \) will refer to laws of a process evolving in this stationary distribution. In the Appendix we show that the density

\[
\varrho(\theta) := E^\theta(\omega_0)
\]

is a strictly increasing, infinitely differentiable function of the parameter \( \theta \) that maps the interval \( (\tilde{\theta}, \bar{\theta}) \) 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 \( \varrho \), rather than \( \theta \), for parameterizing the stationary distributions. Accordingly, \( \mu^\varrho, P^\varrho, E^\varrho, \text{Var}^\varrho, \text{Cov}^\varrho \) will refer to laws of a density \( \varrho \) stationary process.

### 3.2.5 Hydrodynamics and some exact identities

The *hydrodynamic flux* is defined as

\[
\mathcal{H}(\varrho) := E^\varrho(p(\omega_0, \omega_1) - q(\omega_0, \omega_1)).
\] (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-\( \varrho \) 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. Rezakhani-lou [123] or Bahadoran, Guiol, Ravishankar and Saada [92]. Characteristics is a line \( X(T) \) where the density \( \varrho(T, X(T)) \) is constant. The *characteristic speed* is the velocity with which small perturbations of the solution of this PDE propagate, i.e. the slope of constant density lines, and is given by

\[
V^\varrho := \mathcal{H}'(\varrho).
\] (3.2.15)

A particular expectation we shall need several times is

\[
E^\varrho(h_i(t)) = \mathcal{H}(\varrho)t - \varrho i, \quad t \geq 0, \; i \in \mathbb{Z}.
\] (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 \( \underline{\omega} (\cdot) \). (This is a slightly confusing notation, we denote by \( \underline{\omega}^- \) the lower process only because throughout the proofs, the higher one will be our reference point, so we keep the notation \( \underline{\omega} \) for that process.)

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 [104]:

\[
\hat{\mu}^\varrho (y) := \frac{1}{\text{Var}^\varrho (\omega_0)} \sum_{z=y+1}^{\omega_{\max}} (z - \varrho) \mu^\varrho (z), \quad y \in I.
\]  

(3.2.17)

An empty sum is zero by convention and so if \( \omega_{\max} < \infty \), \( \hat{\mu}^\varrho (\omega_{\max}) = 0 \). Consequently there is room for an additional particle under the \( \hat{\mu}^\varrho \) distribution, in the sense that if \( \omega \sim \hat{\mu}^\varrho \) then also \( \omega + 1 \in I \).

To our knowledge these distributions \( \hat{\mu}^\varrho \) 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 \( \mu^\varrho \) and \( \hat{\mu}^\varrho \) are stochastically monotone in the density \( \varrho \). (There is, however, no stochastic domination between \( \mu^\varrho \) and \( \hat{\mu}^\varrho \) in general.)

Denote by \( E \) the expectation w.r.t. the evolution of a pair \( (\omega^- (\cdot), \omega (\cdot)) \) started with initial data (recall (3.2.1))

\[
\omega^- (0) = \omega (0) - \delta_0 \sim \bigotimes_{i \neq 0} \mu^\varrho \otimes \hat{\mu}^\varrho,
\]  

(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, \( \omega^- (t) = \omega (t) - \hat{\delta}Q(t) \). Corollaries 2.4 and 2.5 of [104] state that

\[
\text{Var}^\varrho (h_i(t)) = \text{Var}^\varrho (\omega_0) \cdot E|Q(t) - i|
\]  

(3.2.19)

and

\[
E(Q(t)) = V^\varrho \cdot t
\]  

(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 $p, q : I \times I \to \mathbb{R}_{\geq 0}$ 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 $\mathcal{H}(\rho)$ is strictly concave or convex. In this section we discuss only the concave case. Concavity implies that the characteristic speed $V^\rho = \mathcal{H}'(\rho)$ is a nonincreasing function of density $\rho$:

$$\lambda < \rho \implies V^\lambda \geq V^\rho. \quad (3.2.21)$$

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 $\lambda$ and $\rho$. Since this is now a probabilistic notion, there are several possible formulations, ranging from almost sure ($Q^\lambda(t) \geq Q^\rho(t)$ 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 $\lambda < \rho$ be two densities. Proposition 3.2.4 in the Appendix gives the stochastic domination $\hat{\mu}^\lambda \leq \hat{\mu}^\rho$. Define $\hat{\mu}^\rho + 1$ as the measure that gives weight $\hat{\mu}^\rho(z - 1)$ to an integer $z$ such that $\omega_{\text{min}} < z < \omega_{\text{max}} + 1$. Let $\hat{\mu}^{\lambda,\rho}$ be a coupling measure with marginals $\hat{\mu}^\lambda$ and $\hat{\mu}^\rho + 1$ and with the property

$$\hat{\mu}^{\lambda,\rho}\{(y, z) : \omega_{\text{min}} - 1 < y < z < \omega_{\text{max}} + 1\} = 1. \quad (3.2.22)$$

Let also $\mu^{\lambda,\rho}$ be a coupling measure of site-marginals $\mu^\lambda$ and $\mu^\rho$ of the invariant distributions, with

$$\mu^{\lambda,\rho}\{(y, z) : \omega_{\text{min}} - 1 < y \leq z < \omega_{\text{max}} + 1\} = 1. \quad (3.2.23)$$

Note the distinction that under $\hat{\mu}^{\lambda,\rho}$ the second coordinate is strictly above the first.

To have notation for inhomogeneous product measures on $I^\mathbb{Z}$, let $\lambda = (\lambda_i)_{i \in \mathbb{Z}}$ and $\rho = (\rho_i)_{i \in \mathbb{Z}}$ denote sequences of density values, with $\lambda_i$ and $\rho_i$ assigned to site $i$. The product distribution with marginals $\hat{\mu}^{\lambda_0,\rho_0}$ at the origin and $\mu^{\lambda_i,\rho_i}$ at other sites is denoted by

$$\hat{\mu}^{\lambda,\rho} := \left(\bigotimes_{i \neq 0} \mu^{\lambda_i,\rho_i}\right) \otimes \hat{\mu}^{\lambda_0,\rho_0}. \quad (3.2.24)$$
Measure $\hat{\mu}_{\lambda, \varrho}$ gives probability one to the event
\[ \{ (\eta(0), \omega(0)) : \eta_0(0) < \omega_0(0), \text{ and } \eta_i(0) \leq \omega_i(0) \text{ for } 0 \neq i \in \mathbb{Z} \}. \]

The initial configuration $(\eta(0), \omega(0))$ will always be assumed to be a member of this set, and the pair process $(\eta(t), \omega(t))$ evolves in basic coupling. In general $\hat{\mu}_{\lambda, \varrho}$ 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
\[ \cdots \leq X_{-1}(0) \leq X_0(0) = 0 < X_1(0) \leq \cdots . \]

We can specify that $X_0(0) = 0$ because under $\hat{\mu}_{\lambda, \varrho}$ 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 whenever an $\omega - \eta$ particle jumps from a site,
- if the jump is to the right the highest label moves,
- and if the jump is to the left the lowest label moves.

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 $\varrho$. Recall (3.2.1) for the definition of the configuration $\hat{\delta}$.

**Assumption 3.2.1.** Given a density $\varrho \in (\omega^\text{min}, \omega^\text{max})$, there exists $\gamma_0 > 0$ such that the following holds. For any $\lambda$ and $\varrho$ such that $\varrho - \gamma_0 \leq \lambda_i \leq \varrho_i \leq \varrho + \gamma_0$ for all $i \in \mathbb{Z}$, a joint process $(\eta(t), \omega(t), y(t), z(t))_{t \geq 0}$ can be constructed with the following properties.

- Initially $(\eta(0), \omega(0))$ is $\hat{\mu}_{\lambda, \varrho}$-distributed and the joint process $(\eta(\cdot), \omega(\cdot))$ evolves in basic coupling.
- The processes $y(\cdot)$ and $z(\cdot)$ are integer-valued. Initially $y(0) = z(0) = 0$. With probability one
  \[ y(t) \leq z(t) \text{ for all } t \geq 0. \]

- Define the processes
  \[ \omega^-(t) := \omega(t) - \hat{\delta}_{X_m(t)}(t) \quad \text{and} \quad \eta^+(t) := \eta(t) + \hat{\delta}_{X_m(t)}(t). \]

Then both pairs $(\eta, \eta^+)$ and $(\omega^-, \omega)$ evolve marginally in basic coupling.
For each $\gamma \in (0, \gamma_0)$ and large enough $t \geq 0$ there exists a probability distribution $\nu^{\rho, \gamma}(t)$ on $Z_{\geq 0}$ satisfying the tail bound
\[ \nu^{\rho, \gamma}(t)\{y : y \geq y_0\} \leq C t^{\kappa - 1} \gamma^{2\kappa - 3} y_0^{-\kappa} \] (3.2.28)
for some fixed constants $3/2 \leq \kappa < 3$ and $C < \infty$, and such that if $\rho - \gamma \leq \lambda_i \leq \rho + \gamma$ for all $i \in \mathbb{Z}$, then we have the stochastic bounds
\[ y(t) \leq \nu^{\rho, \gamma}(t) \text{ and } z(t) \sim -\nu^{\rho, \gamma}(t). \] (3.2.29)

Let us clarify some of the details in this assumption.

Equation (3.2.27) says that $Q^\rho(t) := X_{z(t)}(t)$ is the single second class particle between $\eta$ and $\eta^+$, while $Q(t) := X_{y(t)}(t)$ is the one between $\omega^-$ and $\omega$. The first three bullets say that it is possible to construct jointly four processes $(\eta, \eta^+, \omega^-, \omega)$ with the specified initial conditions and so that each pair $(\eta, \omega), (\eta, \eta^+)$ and $(\omega^-, \omega)$ has the desired marginal distribution, and most importantly so that
\[ Q^\rho(t) = X_{z(t)}(t) \geq X_{y(t)}(t) = Q(t). \] (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 $\nu^{\rho, \gamma}(t)$ 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 $\eta(t) \leq \omega(t)$ a.s., and by (3.2.27)
\[ \eta(t) \leq \eta^+(t) \leq \omega(t) \quad \text{and} \quad \eta(t) \leq \omega^-(t) \leq \omega(t) \quad \text{a.s.} \]
In our actual constructions of the processes $\eta, \eta^+, \omega^-, \omega$ for ASEP (Section 3.2.8 and [105]), for a class of totally asymmetric zero range processes (Section 3.6) and for the totally asymmetric bricklayers process with exponential rates it turns out that the triples $(\eta, \eta^+, \omega)$ and $(\eta, \omega^-, \omega)$ evolve also in basic coupling, but the full joint evolution $(\eta, \eta^+, \omega^-, \omega)$ 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 $\lambda_i \equiv \lambda < \rho \equiv \rho_i$. We would expect the $\omega - \eta$ particle $X_0(\cdot)$ to have average and long-term velocity
\[ R(\lambda, \rho) = \frac{\mathcal{H}(\rho) - \mathcal{H}(\lambda)}{\rho - \lambda}, \]
the Rankine-Hugoniot or shock speed. By concavity
\[ \mathcal{H}'(\rho) = V^\rho \leq R(\lambda, \rho) \leq V^\lambda = \mathcal{H}'(\lambda), \]
Figure 3.2: Convexity of the flux function implies that the speed of the single second class particle $Q^\lambda$ is the largest, then comes $X_0$, the second class particle with label 0 and the slowest is $Q$, the single second class particle on a $\varrho$ density system. The corresponding microscopic ordering would be thus $Q(t) \leq X_0(t) \leq Q^n(t)$, but this assumption is too strict: we allow some error and require only $Q(t) \leq Q^n(t)$ plus a distributional bound on the labels of them on the second class particles $X_i$-s.

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 (3.2.20) 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 (3.2.20), but this seems less straightforward.

### 3.2.7 Results

We need a few more assumptions and then we can state the main result. Constants $C$, $\alpha$ will not depend on time, but might depend on the density parameter $\varrho$, 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 $p, q : I \times I \to \mathbb{R}_\geq 0$. 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 $\varrho \in (\varrho^{\min}, \varrho^{\max})$ 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.
- $\mathcal{H}''(\varrho) < 0$.  

123
• Let \((\omega^-, \omega)\) 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

\[
P\{|Q(t)| > K\} \leq C_0 \cdot \frac{t^2}{K^3} \tag{3.2.31}
\]

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 \(H''(\varrho) < 0\). Symmetric processes have \(H(\varrho) \equiv 0\). 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 \(q\). More precisely, if \(p(y + 1, z) - p(y, z)\) and \(q(y, z - 1) - q(y, z)\) are bounded, than the rate of the second class particle is bounded, thus its movement can be dominated by a biased continuous time simple random walk on \(\mathbb{Z}\) and the large deviation result for its position follows.

Here is the main result.

**Theorem 3.2.3.** Let Assumptions \(3.2.1\) and \(3.2.2\) hold for density \(\varrho\). Let the processes \((\omega^-, \omega(t))\) evolve in basic coupling with initial distribution \((3.2.18)\) and let \(Q(t)\) be the position of the second class particle between \(\omega^-\) and \(\omega\). Then there is a constant \(C_1 = C_1(\varrho) \in (0, \infty)\) such that for all \(1 \leq m < 3\),

\[
\frac{1}{C_1} < \liminf_{t \to \infty} \frac{\mathbb{E}|Q(t) - V_{\varrho} t|^m}{t^{2m/3}} \leq \limsup_{t \to \infty} \frac{\mathbb{E}|Q(t) - V_{\varrho} t|^m}{t^{2m/3}} < \frac{C_1}{3 - m}. \tag{3.2.32}
\]

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

\[
\frac{1}{C_1} < \liminf_{t \to \infty} \frac{\mathrm{Var}^e(h_{\lfloor V_{\varrho} t \rfloor}(t))}{q^{2/3}} \leq \limsup_{t \to \infty} \frac{\mathrm{Var}^e(h_{\lfloor V_{\varrho} t \rfloor}(t))}{q^{2/3}} < C_1.
\]

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-\(\varrho\) stationary process:

\[
\frac{Q(t)}{t} \xrightarrow{d} V^{\varrho}. \tag{3.2.33}
\]

If the rates \(p\) and \(q\) have bounded increments, (i.e. \(p(y + 1, z) - p(y, z)\) and \(q(y, z - 1) - q(y, z)\) are bounded), 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-$\varrho$ stationary process:

$$
\lim_{t \to \infty} \frac{h_{\lfloor Vt \rfloor}(t) - h_{\lfloor Vt \rfloor}(0) - t(\mathcal{H}(\varrho) - \varrho \mathcal{H}'(\varrho))}{t^\alpha} = 0.
$$

(3.2.34)

Recall that

$$
h_{\lfloor Vt \rfloor}(0) = \begin{cases} 
\sum_{i=\lfloor Vt \rfloor - \lfloor V\varrho t \rfloor + 1}^{\lfloor Vt \rfloor} \omega_i(0), & \text{if } V < V^e, \\
0, & \text{if } V = V^e, \\
- \sum_{i=1}^{\lfloor Vt \rfloor - \lfloor V\varrho t \rfloor} \omega_i(0), & \text{if } V > V^e
\end{cases}
$$

(3.2.35)

only depends on a finite segment of the initial configuration. The limit (3.2.34) shows that, on the diffusive scale $t^{1/2}$, only fluctuations from the initial distribution are visible: these fluctuations are translated rigidly at the characteristic speed $V^e$. I.e. all spatial fluctuations not coming from the initial configuration are smaller than $t^{1/2}$, moreover, $t^{1/3+\varepsilon}$.

The proof of (3.2.34) follows by translating $h_{\lfloor Vt \rfloor}(t) - h_{\lfloor Vt \rfloor}(0)$ 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. Recall that $\bar{X}$ stands for centering the random variable $X$.

**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-$\varrho$ stationary process

$$
\lim_{t \to \infty} \frac{\text{Var}^\varrho(h_{\lfloor Vt \rfloor}(t))}{t} = \text{Var}^\varrho(\omega) \cdot |V^e - V| = : D,
$$

(3.2.36)

and the Central Limit Theorem also holds: for the centered and normalized height we have $h_{\lfloor Vt \rfloor}(t)/\sqrt{t \cdot \bar{D}}$ 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 [113]. For convex rate zero range and bricklayers processes, Corollary 3.2.7 was proved by M. Balázs [95].
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 $0 \leq p \neq q \leq 1$ 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 $\mu^\varrho$ is the Bernoulli distribution with parameter $0 \leq \varrho \leq 1$, while $\hat{\mu}^\varrho$ is concentrated on zero for any $\varrho$. The hydrodynamic flux is strictly concave: $H(\varrho) = (p - q)\varrho(1 - \varrho)$.

The detailed construction of the processes $y(t)$ and $z(t)$ needed for Assumption 3.2.1 can be found in [105]. Here it is in a nutshell.

Given the background process $(\eta(\cdot), \omega(\cdot))$ 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 $y \leq z$ 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 [106]. 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 $f: \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ 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
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 \( f(z) - f(z - 1) > 0 \),
  \[
  \frac{f(z + 1) - f(z)}{f(z) - f(z - 1)} \leq r.
  \]
  \[ (3.2.37) \]

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 \( f(z) = 1 \{z > 0\} \). For this last case a proof has already been given in [99].

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 \( H''(\varrho) < 0 \) 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 \( \varrho \) is fixed. Let \( \lambda \in (\rho, \rho - \gamma_0) \) and apply Assumption 3.2.1 with constant sequences
Lemma 3.3.1. \( \varrho_i \equiv \varrho \) and \( \lambda_i \equiv \lambda \) for all \( i \in \mathbb{Z} \). Notations \( \mathbf{P}, \mathbf{E}, \text{Var}, \text{Cov} \) will refer to the coupled four-process evolution described in Assumption 3.2.1 while \( \mathbf{P}^\varrho, \mathbf{E}^\varrho, \text{Var}^\varrho, \text{Cov}^\varrho \) will refer to a density \( \varrho \) stationary process. Abbreviate

\[
\Psi(t) := \mathbf{E}[Q(t) - |V^\varrho t|].
\]  

(3.3.1)

The requirement that \( (\omega^-, \omega) \) 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 \( \alpha_1, \alpha_2, t_0 \) such that for each \( t > t_0 \) and integer \( u > 0 \) such that \( \alpha_2 \sqrt{t} < u < \alpha_1 t \),

\[
\mathbf{P}\{Q(t) > |V^\varrho t| + u\} \leq C_5 \frac{t^2 \mathbf{H}^\varrho(\varrho)^2}{u^4} \left\{ \Psi(t) + u \right\} + C_4 \frac{t^2}{u^3}.
\]  

(3.3.2)

**Proof.** We start with an integer \( u > 0 \), and write

\[
\mathbf{P}\{Q(t) > |V^\varrho t| + u\} \leq \mathbf{P}\{y(t) \geq k\} + \mathbf{P}\{X_k(t) \geq Q(t) > |V^\varrho t| + u\}.
\]  

(3.3.3)

The event \( \{X_k(t) > |V^\varrho t| + u\} \) implies that among the \( X_m \)'s at most particles \( X_1, \ldots, X_{k-1} \) have passed the path \( (s([V^\varrho t] + u) + 1/2)_{0 \leq s \leq 1} \) from right to left. Each such passing decreases \( h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\eta(t) \) by one (recall the statement around (3.2.10)). Hence we can bound the probability in (3.3.3) by

\[
\mathbf{P}\{y(t) \geq k\} + \mathbf{P}\{h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\eta(t) > -k\}.
\]

We introduce two more processes: \( \omega_{i}^{eq} \) is a stationary process started with initial data \( \eta_i^{eq}(0) = \eta_i(0) \) for \( i \neq 0 \), while \( \eta_0^{eq}(0) \) is \( \mu^\lambda \) distributed independently of everything. \( \omega_{i}^{eq} \) is a stationary process started with \( \omega_{i}^{eq}(0) = \omega_i(0) \) for \( i \neq 0 \), and \( \omega_0^{eq}(0) \) is \( \mu^\varrho \) distributed independently of everything. Include these in the basic coupling of \( (\eta, \omega) \) and write

\[
\begin{align*}
&h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\eta(t) = h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\eta(t) \\
&\quad + h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\omega(t) \\
&\quad - h_{[V^\varrho t]+u}^\eta(t) + h_{[V^\varrho t]+u}^\omega(t).
\end{align*}
\]

Basic coupling implies

\[
\begin{align*}
&h_{[V^\varrho t]+u}^\varrho(t) - h_{[V^\varrho t]+u}^\omega(t) \leq |\omega_0(0) - \omega_0^{eq}(0)| \leq |\omega_0(0)| + |\omega_0^{eq}(0)| \\
\text{and } &h_{[V^\varrho t]+u}^\omega(t) - h_{[V^\varrho t]+u}^\eta(t) \leq |\eta_0^{eq}(0) - \eta_0(0)| \leq |\eta_0^{eq}(0)| + |\eta_0(0)|.
\end{align*}
\]
We bound the stationary expectations using (3.2.16), (3.2.15) and Taylor’s formula:

\[
E^{\varphi \omega_{\varphi t} + u}(t) - E^{\lambda \omega_{\varphi t} + u}(t) \\
= \mathcal{H}(\varphi) t - (|V^{\varphi t}| + u) \varphi - \mathcal{H}(\lambda) t + (|V^{\varphi t}| + u) \lambda \\
\leq t (\mathcal{H}(\varphi) - \mathcal{H}(\lambda) + \mathcal{H}'(\varphi)(\lambda - \varphi)) + u(\lambda - \varphi) + C_1 \\
\leq -\frac{t}{2} \mathcal{H}''(\varphi)(\varphi - \lambda)^2 + u(\varphi - \lambda) + C_2 t(\varphi - \lambda)^3 + C_1.
\]

\(\mathcal{H}\) can be differentiated arbitrarily many times, as we show in Section 3.1 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.

\[
P\{Q(t) > |V^{\varphi t}| + u\} \\
\leq P\{y(t) \geq k\} \\
+ P\{\tilde{h}^{\varphi}_{\varphi t} + u(t) - \tilde{h}^{\mu}_{\varphi t} + u(t) > -k + \frac{t}{2} \mathcal{H}''(\varphi)(\varphi - \lambda)^2 + u(\varphi - \lambda) \\
- C_2 t(\varphi - \lambda)^3 - C_1 - |\eta_0(0)| - |\eta_0^{eq}(0)| - |\omega_0(0)| - |\omega_0^{eq}(0)|\} \\
\leq P\{y(t) \geq k\} \\
+ P\{\tilde{h}^{\varphi}_{\varphi t} + u(t) - \tilde{h}^{\mu}_{\varphi t} + u(t) > \frac{t}{2} \mathcal{H}''(\varphi)(\varphi - \lambda)^2 + \frac{u}{2}(\varphi - \lambda)\} \\
+ P\{|\eta_0(0)| + |\eta_0^{eq}(0)| + |\omega_0(0)| + |\omega_0^{eq}(0)|\} \\
> -k + \frac{u}{2}(\varphi - \lambda) - C_2 t(\varphi - \lambda)^3 - C_1\}.
\]

From now on we use the specific assumption \(\mathcal{H}''(\varphi) < 0\). We maximize the terms on the right-hand side of the probability of \(\tilde{h}\)'s by the choice \(\varphi - \lambda = \frac{-u}{2t\mathcal{H}''(\varphi)}\).

To stay within the range of densities covered by Assumption 3.2.1 we must ensure that \(\lambda > \varphi - \gamma_0\). So we introduce a small constant \(\alpha_1 > 0\) and restrict our calculations to the case \(u < \alpha_1 t\). Then

\[
P\{Q(t) > |V^{\varphi t}| + u\} \leq P\{y(t) \geq k\} \\
+ P\{\tilde{h}^{\varphi}_{\varphi t} + u(t) - \tilde{h}^{\mu}_{\varphi t} + u(t) > \frac{-u^2}{8t\mathcal{H}''(\varphi)}\} \\
+ P\{|\eta_0(0)| + |\eta_0^{eq}(0)| + |\omega_0(0)| + |\omega_0^{eq}(0)|\} \\
> -k - \frac{1}{4\mathcal{H}''(\varphi)} \cdot \frac{u^2}{t} + C_2 \cdot \frac{u^3}{t^2} - C_1\}.
\]

Now we set

\[
k = \left\lfloor \frac{1}{8\mathcal{H}''(\varphi)} \cdot \frac{u^2}{t} \right\rfloor,
\]

129
and assume $\alpha_2 \sqrt{t} < u < \alpha_1 t$ for a possibly smaller $\alpha_1$ and a large enough $\alpha_2$. That allows us to unify the right-hand side of the inequality in the last line. Thus for all large $u$ and $t$ with $\alpha_2 \sqrt{t} < u < \alpha_1 t$

$$\mathbb{P}\{ Q(t) > \lfloor V^q_t \rfloor + u \} \leq \mathbb{P}\{ y(t) \geq \lfloor -\frac{1}{8\mathcal{H}''(q)} \cdot \frac{u^2}{t} \rfloor \}$$

$$+ \mathbb{P}\{ \tilde{\eta}_{\lfloor V^q_t \rfloor + u}^\eta(t) - \tilde{\eta}_{\lfloor V^q_t \rfloor + u}^\eta(t) > \frac{-u^2}{8t\mathcal{H}''(q)} \}$$

$$+ \mathbb{P}\{ \lfloor \eta_0(0) \rfloor + \lfloor \eta_0^\eta(0) \rfloor + \lfloor \omega_0(0) \rfloor + \lfloor \omega_0^\eta(0) \rfloor > C_3 \frac{u^2}{t} \}.$$  

Assumption 3.2.28 allows us to bound the first probability on the right by $C_4t^2/u^3$ (take $\gamma = \varrho - \lambda$). Apply Chebyshev’s inequality on the second line and Markov’s inequality on the third one:

$$\mathbb{P}\{ Q(t) > \lfloor V^q_t \rfloor + u \} \leq \frac{64}{u^2} \mathcal{H}''(q)^2 \mathbb{E}[\eta_{\lfloor V^q_t \rfloor + u}^\eta(t) - \eta_{\lfloor V^q_t \rfloor + u}^\eta(t)] + C_3 \frac{t}{u^2} + C_4 \frac{t^2}{u^3}$$

$$\leq \frac{128}{u^2} \mathcal{H}''(q)^2 \left\{ \mathbb{E}[\eta_{\lfloor V^q_t \rfloor + u}^\eta(t)] + \mathbb{E}[\omega_{\lfloor V^q_t \rfloor + u}^\eta(t)] \right\}$$

$$+ C_4 \frac{t^2}{u^3}.$$  

The term $C_3t/u^2$ was subsumed under $C_4t^2/u^3$ due to the condition $u < \alpha_1 t$. The variances here are taken under the stationary distributions of the processes $\eta_0^\eta$ and $\omega_0^\eta$. That allows us to apply (3.2.19), whose right-hand side takes us back to the four-process coupling under measure $\mathbb{P}$. Recall (3.3.1).

$$\mathbb{P}\{ Q(t) > \lfloor V^q_t \rfloor + u \}$$

$$\leq C_5 \left( \frac{t^2}{u^4} \mathcal{H}''(q)^2 \left\{ \mathbb{E}[Q(t) - \lfloor V^q_t \rfloor - u] + \mathbb{E}[Q^\eta(t) - \lfloor V^q_t \rfloor - u] \right\} + C_4 \frac{t^2}{u^3} \right)$$

$$\leq C_5 \left( \frac{t^2}{u^4} \mathcal{H}''(q)^2 \left\{ \mathbb{E}[Q(t) - \lfloor V^q_t \rfloor] + \mathbb{E}[Q^\eta(t) - \lfloor V^q_t \rfloor] + 2u \right\} + C_4 \frac{t^2}{u^3} \right)$$

$$= C_5 \left( \frac{t^2}{u^4} \mathcal{H}''(q)^2 \left\{ \Psi(t) + 2u + \mathbb{E}[Q^\eta(t) - \lfloor V^q_t \rfloor] \right\} + C_4 \frac{t^2}{u^3} \right).$$

The variable $Q^\eta(t)$ above is the location of a single discrepancy between the process $\eta$ and one started initially with $\eta_0^+(0) = \eta(0) + \delta_0$.

It remains to relate $\mathbb{E}[Q^\eta(t) - \lfloor V^q_t \rfloor]$ 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 $\eta$, $\eta^+$, $\omega^-$, $\omega$ described in Assumption 3.2.1. Use (3.2.30) and Taylor
The expansion of $\mathcal{H}$ again:

\[
\begin{align*}
\mathbb{E}|Q^u(t) - [V^et]| &\leq \mathbb{E}(Q^u(t) - Q(t)) + \Psi(t) \\
&= (\mathcal{H}'(\lambda) - \mathcal{H}'(\rho))t + \Psi(t) \\
&\leq \mathcal{H}''(\rho) \cdot (\lambda - \rho)t + C_6(\rho - \lambda)^2t + \Psi(t) \\
&= \frac{u}{4} + C_6 \frac{u^2}{t} + \Psi(t) \leq (\frac{1}{4} + C_6 \alpha_1)u + \Psi(t).
\end{align*}
\]

The last inequality used $u < \alpha_1t$. 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 $\alpha_2 \sqrt{t} < u < \alpha_1 t$:

\[
\mathbb{P}\{ |Q(t) - [V^et]| > u \} \leq C_5 \frac{t^2 \mathcal{H}''(\rho)^2}{u^4} \{\Psi(t) + u\} + C_4 \frac{t^2}{u^3}.
\]

Next we relax the restriction to integer $u$ and the upper limit on it:

**Lemma 3.3.2.** There are positive constants $\alpha_2, t_0$ such that for all $t > t_0$ and all real $u > \alpha_2 \sqrt{t}$,

\[
\mathbb{P}\{ |Q(t) - [V^et]| > u \} \leq C_5 \frac{t^2 \mathcal{H}''(\rho)^2}{u^4} \{\Psi(t) + u\} + C_4 \frac{t^2}{u^3}.
\]

**Proof.** Any $u \geq 1$ is less than twice its integer part. Hence by simply increasing the constants $C_i$, for all large $t$ and all real $u \in (\alpha_2 \sqrt{t}, \alpha_1 t)$,

\[
\mathbb{P}\{ |Q(t) - [V^et]| > u \} \leq C_5 \frac{t^2 \mathcal{H}''(\rho)^2}{u^4} \{\Psi(t) + u\} + C_4 \frac{t^2}{u^3}.
\]

Recall (3.2.31). When $\alpha_1 < \alpha_0 + 2|V^e| + 2$, assume $\alpha_1 t \leq u < (\alpha_0 + 2|V^e| + 2) t$. Then $\alpha_2 \sqrt{t} < u \cdot \alpha_1/(\alpha_0 + 2|V^e| + 2) < \alpha_1 t$ for large enough $t$, and (3.3.5) still holds for $u$ replaced by $u \cdot \alpha_1/(\alpha_0 + 2|V^e| + 2)$:

\[
\mathbb{P}\{ |Q(t) - [V^et]| > u \} \leq \mathbb{P}\{ |Q(t) - [V^et]| > u \cdot \frac{\alpha_1}{\alpha_0 + 2|V^e| + 2} \} \\
\leq C_5 \frac{t^2 \mathcal{H}''(\rho)^2}{u^4} \{\Psi(t) + u\} + C_4 \frac{t^2}{u^3}.
\]

via modifying the constants by factors of $\alpha_1/(\alpha_0 + 2|V^e| + 2)$.

Finally, when $u \geq (\alpha_0 + 2|V^e| + 2)t$, the fact that $u - |[V^et]| > \alpha_0 t$ allows us to use (3.2.31):

\[
\mathbb{P}\{ |Q(t) - [V^et]| > u \} \leq \mathbb{P}\{ |Q(t)| > u - |[V^et]| \} \\
\leq C_7 \frac{t^2}{(u - |[V^et]|)^3} \leq C_8 \frac{t^2}{u^3}.
\]

Combining the above cases we get the statement for all $u > \alpha_2 \sqrt{t}$. \qed
Proof of the upper bound of Theorem 3.2.3. We now fix $r > 0$, $1 \leq m < 3$, and write
\[
E( |Q(t) - [V^q t]|^m ) = \int_0^\infty P \{ |Q(t) - [V^q t]|^m > v \} \, dv 
\leq r m t^{3 \frac{2}{3} m + \frac{m}{4} m^{-1} \Psi(t) + C_4 t^{2} u^{m-3} t^{3 \frac{2}{3} m}.
\]

First choose $m = 1$ and $r$ large enough to get $\Psi(t) \leq C t^{2/3}$. Then insert this bound back into the last line of the display to get the bound for general $1 \leq m < 3$.

3.4 Lower bound of the main theorem

In this section we prove the lower bound of (3.2.32). Density $\varrho$ is fixed again, and $\lambda \in (\varrho - \gamma_0, \varrho)$ is a varying auxiliary density. We let the jointly defined four processes $(\eta, \eta^+, \omega^-, \omega)$ 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 $(\eta, \omega)$ is $\hat{\mu}_{\lambda,\varrho}$ of (3.2.24). Two second class particles start from the origin: $Q_\eta$ between processes $\eta$ and $\eta^+$, and $Q_\omega$ between processes $\omega^-$ and $\omega$. The quantity of primary interest is abbreviated, as before, by $\Psi(t) = E[|Q(t) - [V^\varrho t]|]$. 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 $\Psi(t) \geq C t^{2/3}$ 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 $V^q = H'(\varrho)$ and $V^\lambda = H'(\lambda)$ satisfy $V^q \leq V^\lambda$. Throughout this section $u > 0$ denotes a fixed positive integer, and
\[
n = [V^\lambda t] - [V^q t] + u.
\]

Recall definitions (3.2.22) and (3.2.23) of the single-site coupling measures. Let $(\xi(\cdot), \zeta(\cdot))$ be a pair of processes that obeys the basic coupling, and whose initial distribution is the product measure
\[
\bigotimes_{i < -n} \mu^{\lambda,q} \bigotimes_{i = -n} \hat{\mu}^{\lambda,q} \bigotimes_{-n < i \leq 0} \mu^{\lambda,\lambda} \bigotimes_{0 < i} \mu^{\lambda,q}.
\]

132
This initial measure complies with the pattern in (3.2.24), but translated $n$ sites to the left so that $\hat{\mu}^{\lambda,\theta}$ is the distribution at site $-n$ instead of the origin. A few points about this initial state: $\zeta(0)$ has the stationary density-$\lambda$ product distribution except at site $-n$ where it is $\hat{\mu}^{\lambda}$-distributed. $\zeta(0)$ has the product distribution with marginals $\mu^\theta$, except at sites $\{n+1, \ldots, 0\}$ where the parameter $\theta$ switches to $\lambda$, and at site $-n$ where it has distribution $\hat{\mu}^{\theta} + 1$. At sites $-n < i \leq 0$ $\mu^{\lambda,\lambda}$ forces $\xi_i(0) = \zeta_i(0)$.

We add a second class particle to the process $\xi(\cdot)$, start it at site $-n$ and denote its position at time $t$ by $Q^{(-n)}(t)$. Let $\tilde{\xi}(t) := \xi(t) + \delta_{Q^{(-n)}(t)}$.

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

$$\cdots \leq X_{-1}(0) \leq X_0(0) = -n < 0 < X_1(0) \leq X_2(0) \leq \cdots \quad (3.4.1)$$

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 $\xi$ plays the role of $\eta$ and $\zeta$ plays the role of $\omega$ of Assumption 3.2.1)

As before, the heights (or currents, recall (3.2.5)) of the processes $\xi(\cdot)$ and $\zeta(\cdot)$ are denoted by $h_{[V[t]}^\xi$ and $h_{[V[t]}^\zeta$, respectively. The first observation is that $Q^{(-n)}$ gives one-sided control over the difference of these currents.

**Lemma 3.4.1.** For any $V \in \mathbb{R}$

$$Q^{(-n)}(t) \leq |V[t| \quad \text{implies} \quad h_{[V[t]}^\xi(t) - h_{[V[t]}^\zeta(t) \leq -z(t).$$

**Proof.** Recall again, from (3.2.5) and the statement around (3.2.10), that the height difference $h_{[V[t]}^\xi(t) - h_{[V[t]}^\zeta(t)$ equals the net number of second class particle passings of the path $(s[V[t]+1/2]_{0 \leq s \leq 1}$ from left to right. That is, each left-to-right passing increases $h_{[V[t]}^\xi(t) - h_{[V[t]}^\zeta(t)$ while each right-to-left passing decreases it.

Suppose $z(t) \leq 0$. Then (3.4.1) and $X_{z(t)}(t) = Q^{(-n)}(t) \leq |V[t|$ imply that only those second class particles with labels $z(t) + 1$, $z(t) + 2, \ldots, 0$ could have crossed the path $(s[V[t]+1/2]_{0 \leq s \leq 1}$ from left to right. The claim follows.

If $z(t) > 0$, then $X_{z(t)}(t) = Q^{(-n)}(t) \leq |V[t|$ implies that at least those second class particles with labels $1, 2, \ldots, z(t)$ have crossed the path $(s[V[t]+1/2]_{0 \leq s \leq 1}$ from right to left. Again the claim follows. \hfill $\square$

Let $\hat{\omega}(\cdot)$ be a process started from the product distribution $(\bigotimes_{i \neq -n} \mu^\theta) \otimes (\hat{\mu}^{\theta+1})$. The next lemma compares the distributions of $\zeta$ and $\hat{\omega}$. No coupling of $\zeta$ and $\hat{\omega}$ is proposed or required.
Lemma 3.4.2. There exist constants $\gamma = \gamma(q) > 0$ and $C_1(q) < \infty$ such that for all $\lambda \in (q - \gamma, q)$ and all events $A$ the following inequality holds:

$$P(\zeta \in A) \leq P(\hat{\zeta} \in A)^{1/2} \cdot \exp\{C_1(q)n(q - \lambda)^2\}.$$ 

Proof. We use the Cauchy-Schwarz inequality below to perform a change of measure on the distribution of the $\zeta$ process. First we condition on the initial $\zeta$-configuration at sites $\{-n+1, -n+2, \ldots, -1, 0\}$.

$$P(\zeta \in A) = \sum_{z_{-n+1}, \ldots, z_0} P(\zeta \in A | \zeta_{-n+1}(0) = z_{-n+1}, \ldots, \zeta_0(0) = z_0) \times \left[ \prod_{i=-n+1}^{0} \mu^\lambda(z_i) \right]^{1/2} \prod_{i=-n+1}^{0} \mu^\theta(z_i) \leq \left[ \sum_{z_{-n+1}, \ldots, z_0} P(\zeta \in A | \zeta_{-n+1}(0) = z_{-n+1}, \ldots, \zeta_0(0) = z_0) \right]^{1/2} \prod_{i=-n+1}^{0} \mu^\lambda(z_i) \times \left[ \sum_{z_{-n+1}, \ldots, z_0} \prod_{i=-n+1}^{0} \left[ \frac{\mu^\lambda(z_i)}{\mu^\theta(z_i)} \right]^{1/2} \right] \leq \left[ \sum_{z_{-n+1}, \ldots, z_0} P(\zeta \in A | \zeta_{-n+1}(0) = z_{-n+1}, \ldots, \zeta_0(0) = z_0) \right]^{1/2} \prod_{i=-n+1}^{0} \mu^\lambda(z_i) \times \left[ \sum_{z_{-n+1}, \ldots, z_0} \prod_{i=-n+1}^{0} \left[ \frac{\mu^\lambda(z_i)}{\mu^\theta(z_i)} \right]^{1/2} \right] = P(\hat{\zeta} \in A)^{1/2} \cdot \left[ \sum_{z_{-n+1}, \ldots, z_0} \prod_{i=-n+1}^{0} \left[ \frac{\mu^\lambda(z_i)}{\mu^\theta(z_i)} \right]^{1/2} \right].$$

The last inequality came from dropping the square. For the last equality note that the distributions of the initial configurations $\{\hat{\zeta}_i(0)\}$ 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

$$\sum_{z_{-n+1}, \ldots, z_0} \left( \frac{Z(\theta(q))}{Z(\theta(\lambda))} \right)^n \prod_{i=-n+1}^{0} \frac{\omega^{(2\theta(\lambda) - \theta(q))z_i}}{f(z_i)!} = \left( \frac{Z(2\theta(\lambda) - \theta(q))Z(\theta(q))}{Z(\theta(\lambda))^2} \right)^n.$$

In the Appendix we show that $\log Z(\theta)$ and $\theta(q)$ are infinitely differentiable. Let $\varepsilon = \theta(q) - \theta(\lambda)$. By local Lipschitz continuity of the function $\theta(q)$, the
interval \((\theta(\lambda) - \varepsilon, \theta(\lambda) + \varepsilon)\) is in \((\tilde{\theta}, \bar{\theta})\) with a small enough choice of \(\gamma\). There exists some \(\theta \in (\theta(\lambda) - \varepsilon, \theta(\lambda) + \varepsilon)\) such that

\[
\log \left( \frac{Z(2\theta(\lambda) - \theta(\varrho))Z(\theta(\varrho))}{Z(\theta(\lambda))^2} \right) = \log Z(\theta(\lambda) - \varepsilon) + \log Z(\theta(\lambda) + \varepsilon) - 2 \log Z(\theta(\lambda)) = \frac{1}{2} \frac{d^2}{d\theta^2} \log Z(\theta) \varepsilon^2 \leq C_1 (\varrho - \lambda)^2.
\]

Thus we get the bound

\[
\left( \frac{Z(2\theta(\lambda) - \theta(\varrho))Z(\theta(\varrho))}{Z(\theta(\lambda))^2} \right)^n \leq \exp \{ C_1 (\varrho - \lambda)^2 \}. 
\]

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 \(P\{Q^{(-n)}(t) > [V^\varrho t]\}\) and \(P\{Q^{(-n)}(t) \leq [V^\varrho t]\}\). As stated \(u\) is an arbitrary but fixed positive integer and \(n = [V^\lambda t] - [V^\varrho t] + u\).

Lemma 3.4.3.

\[
P\{Q^{(-n)}(t) > [V^\varrho t]\} \leq \frac{\Psi(t)}{u} + \frac{C_2 t (\varrho - \lambda)}{u} + \frac{2}{u}.
\]

Proof. Distributionwise the system \((\xi, \xi^+, Q^{(-n)})\) is a translate of \((n, n^+, Q^\varrho)\), and so

\[
P\{Q^{(-n)}(t) > [V^\varrho t]\} = P\{Q^{(-n)}(t) + n - [V^\lambda t] > u\} = \frac{E(|Q^\varrho(t) - [V^\lambda t]|)}{u} + \frac{E(|Q(t) - [V^\varrho t]|)}{u} + \frac{|V^\lambda t| - [V^\varrho t]}{u}.
\]

Use (3.2.30) precisely as was done in (3.3.4) to conclude that the first term equals

\[
u^{-1}E(Q^\varrho(t) - Q(t)) = \nu^{-1}t(H'(\lambda) - H'(\varrho)) = -\nu^{-1}H''(\nu)t(\varrho - \lambda)
\]

for some \(\nu \in (\lambda, \varrho)\). The second term is \(\Psi(t)/u\), and the third term is similarly estimated by \(-\nu^{-1}H''(\nu)t(\varrho - \lambda) + 2/u\), the last part coming from discarded integer parts. Setting \(C_2 := 2 \max_{\nu \in [\varrho - \gamma, \varrho]} -H''(\nu)\) finishes the proof.

\[\square\]
Notice that $H''(\rho) < 0$ 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 $0 < K < -\frac{1}{3}tH''(\rho)(\rho - \lambda)^2$. Then for small enough $\gamma > 0$, large enough $t$, and $\lambda \in (\rho - \gamma, \rho)$,

$$
P\{Q^{(n)}(t) \leq [V^eq_t]\} \leq \frac{\text{Var}^e(\omega_0)^{1/2}\Psi(t)^{1/2}}{\frac{1}{3}tH''(\rho)(\rho - \lambda)^2 - K} \cdot e^{C_1n(e-\lambda)^2}
+ \frac{\text{Var}^e(\xi_0)(\eta_0)\Psi(t)}{K^2/4} + \frac{C_5}{K^2} + C_{\gamma}n \cdot e^{C_{\lambda}n(e-\lambda)^2}
\leq C \cdot e^{C_{\lambda}n(e-\lambda)^2}.
$$

**Proof.** Lemma 3.4.1 leads to

$$
P\{Q^{(n)}(t) \leq [V^eq_t]\} \leq P\{h^\xi_{[V^eq_t]}(t) - h^\xi_{[V^eq_t]}(t) \leq -z(t)\}
\leq P\{-z(t) \geq K/4\} \quad (3.4.2)
+ P\{h^\xi_{[V^eq_t]}(t) \leq K + t(\lambda H - \lambda H'(\rho))\} \quad (3.4.3)
+ P\{h^\xi_{[V^eq_t]}(t) > 3K/4 + t(\lambda H - \lambda H'(\rho))\}
\leq C \cdot e^{C_{\lambda}n(e-\lambda)^2}.
$$

To bound (3.4.2) we use the assumed distribution bound (3.2.26) on $z(t)$ and get

$$
P\{-z(t) \geq K/4\} \leq C \cdot e^{C_{\lambda}n(e-\lambda)^2}.
$$

Apply Lemma 3.4.2 to line (3.4.3) to bound it by the probability of the process $\hat{\omega}$:

$$
\leq \left[ P\{h^\xi_{[V^eq_t]}(t) \leq K + t(\lambda H - \lambda H'(\rho))\}\right]^{1/2} \cdot e^{C_{\lambda}n(e-\lambda)^2}.
$$

As in the proof of Lemma 3.3.1 we switch to stationary processes to get precise bounds:

$$
h^\xi_{[V^eq_t]}(t) = \hat{h}^\xi_{[V^eq_t]}(t) + \left[h^\xi_{[V^eq_t]}(t) - \hat{h}^\xi_{[V^eq_t]}(t)\right]
+ [\mathbb{E}h^\xi_{[V^eq_t]}(t) - t(\lambda H - \lambda H'(\rho))] + t(\lambda H - \lambda H'(\rho))
\geq \hat{h}^\xi_{[V^eq_t]}(t) - |\hat{\omega}_0(0)| - |\omega_0^eq(0)| - |\rho| + t(\lambda H - \lambda H'(\rho)).
$$

After the equality sign, the absolute value of the first term in brackets is not larger than $|\hat{\omega}_0(0) - \omega_0^eq(0)| \leq |\hat{\omega}_0(0)| + |\omega_0^eq(0)|$. The second term in brackets is between $-|\rho|$ and $|\rho|$ due to the integer part in $[V^eq_t]$. Consequently

$$
h^\xi_{[V^eq_t]}(t) \leq K + t(\lambda H - \lambda H'(\rho))
\leq C \cdot e^{C_{\lambda}n(e-\lambda)^2}.
$$

136
implies
\[ \hat{h}^{\omega_{eq}}_{[\mathcal{V}_{et}]}(t) - |\hat{\omega}_0(0)| - |\omega_0^{eq}(0)| \leq K + t[H(\lambda) - H(\varrho) + H'(\varrho)(\varrho - \lambda)] + |\varrho| \leq K + \frac{1}{2}tH''(\varrho)(\varrho - \lambda)^2 + C_3t(\varrho - \lambda)^3 + |\varrho|. \]

Then, we cut the event into two parts according to the value of $|\hat{\omega}_0(0)| + |\omega_0^{eq}(0)|$ and we use (3.2.19) to bound the variance of $\text{Var}[h^{\omega_{eq}}_{[\mathcal{V}_{et}]}(t)]$ by the function $\Psi(t)$.

\[ (3.4.3) \leq \left[ P^\varrho(\hat{h}^{\omega_{eq}}_{[\mathcal{V}_{et}]}(t) \leq K + \frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2) \right]^\frac{1}{2} \cdot e^{C_1n(\varrho - \lambda)^2} + \left[ P\{|\hat{\omega}_0(0)| + |\omega_0^{eq}(0)| > -\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|\} \right]^\frac{1}{2} \cdot e^{C_1n(\varrho - \lambda)^2} \leq \frac{\text{Var}^\varrho(h^{\omega_{eq}}_{[\mathcal{V}_{et}]}(t))^{1/2}}{-\frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2 - K} \cdot e^{C_1n(\varrho - \lambda)^2} + \frac{\text{E}[|\hat{\omega}_0(0)| + |\omega_0^{eq}(0)|]}{-\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|} \cdot e^{C_1n(\varrho - \lambda)^2} \leq \frac{\text{Var}^\varrho(\omega_0)^{1/2}\Psi(t)^{1/2}}{-\frac{1}{3}tH''(\varrho)(\varrho - \lambda)^2 - K} \cdot e^{C_1n(\varrho - \lambda)^2} + \frac{C_4}{-\frac{1}{6}tH''(\varrho)(\varrho - \lambda)^2 - C_3t(\varrho - \lambda)^3 - |\varrho|} \cdot e^{C_1n(\varrho - \lambda)^2}. \]

Now we turn to (3.4.4). To reduce $h^{\xi}_{[\mathcal{V}_{et}]}$ to the current of the density-$\lambda$ equilibrium process $h^{\eta_{eq}}_{[\mathcal{V}_{et}]}$ and to get rid of the integer part errors we argue as before.

\[ h^{\xi}_{[\mathcal{V}_{et}]} = \hat{h}^{\eta_{eq}}_{[\mathcal{V}_{et}]} + [h^{\xi}_{[\mathcal{V}_{et}]} - h^{\eta_{eq}}_{[\mathcal{V}_{et}]}] \]

\[ + \left[ \text{E}^\lambda h^{\eta_{eq}}_{[\mathcal{V}_{et}]} - t(H(\lambda) - \lambda H'(\varrho)) + t(H(\lambda) - \lambda H'(\varrho)) \right]. \]

$h^{\xi}_{[\mathcal{V}_{et}]}(t)$ differs by at most $|\xi_0(0) - \xi_0^{eq}(0)| = \xi_0(0) + |\xi_0^{eq}(0)|$ from $h^{\eta_{eq}}_{[\mathcal{V}_{et}]}(t)$. Taking integer parts again into account, giving another error term $|\lambda|$, line (3.4.4) is bounded from above by

\[ P\left\{ \hat{h}^{\eta_{eq}}_{[\mathcal{V}_{et}]}(t) + |\xi_0(0)| + |\xi_0^{eq}(0)| + |\lambda| \geq 3K/4 \right\} \]

Then, we cut the event into two parts and use Markov’s inequality in the second one:

\[ (3.4.4) \leq P^\lambda\left\{ \hat{h}^{\eta_{eq}}_{[\mathcal{V}_{et}]}(t) \geq K/2 \right\} + P\left\{ |\xi_0(0)| + |\xi_0^{eq}(0)| > K/4 - |\lambda| \right\} \leq \frac{\text{Var}^\lambda(h^{\eta_{eq}}_{[\mathcal{V}_{et}]} \cdot C_5}{K^2/4} + \frac{K}{K - 4|\lambda|}. \]
We can use (3.2.19) again to continue with
\[ (3.4.4) \leq \text{Var}^{\lambda}(\xi_0)E(\left|Q^q(t) - [V^q t]\right|) + \frac{C_5}{K - 4|\lambda|}. \]

Repeating the first two steps of calculation (3.3.4) we can write
\[ E(\left|Q^q(t) - [V^q t]\right|) \leq E(\left|Q^q(t) - Q(t)\right|) + E(\left|Q(t) - [V^q t]\right|) \]
\[ \leq Ct(\varrho - \lambda) + \Psi(t). \]

So, we finally get
\[ (3.4.4) \leq \frac{\text{Var}^{\lambda}(\eta_0)\Psi(t)}{K^2/4} + \frac{C_5}{K} - \frac{C_5}{4|\lambda|}. \]

Proof of the lower bound of Theorem 3.2.3. As observed in the beginning of this Section, it suffices to prove that
\[ \liminf_{t \to \infty} t^{-2/3}\Psi(t) > 0. \] (3.4.5)

In the last two lemmas take
\[ u = \left[ht^{2/3}\right], \quad \varrho - \lambda = bt^{-1/3}, \quad \text{and} \quad K = bt^{1/3}, \]
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
\[ n = [V^\lambda t] - [V^q t] + u \]
\[ \leq (H'(\lambda) - H'(\varrho))t + u + 2 \]
\[ = -H''(\varrho)(\varrho - \lambda)t + u + Ct(\varrho - \lambda)^2 + 2 \]
\[ \leq (-H''(\varrho)b + h)t^{2/3} + C_7t^{2/3} + 3 \]
\[ \leq C_8t^{2/3} \]
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
\[ P\{Q^{(-n)}(t) > [V^q t]\} \leq C\frac{\Psi(t)}{t^{2/3}} + \frac{C_2b}{h} + \frac{2}{ht^{2/3}} \] (3.4.6)
and
\[ P\{Q^{(-n)}(t) \leq [V^q t]\} \leq C\left(\frac{\Psi(t)}{t^{2/3}}\right)^{1/2} + C\frac{\Psi(t)}{t^{2/3}} + \frac{C_6}{b} + \frac{C_5}{bt^{1/2}} + Cb^{\kappa - 3}. \] (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 \( C_2b/h + C_6/b + Cb^{\kappa - 3} < 1 \) (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.,

\[
\begin{align*}
    p(y + 1, z) - p(y, z), \quad p(y, z) - p(y, z + 1) \leq C & \quad \forall \omega_{\text{min}} \leq y, \ z < \omega_{\text{max}}. \\
    q(y, z + 1) - q(y, z), \quad q(y, z) - q(y + 1, z) \leq C \\
\end{align*}
\]  

(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

\[
A_n := \left\{ \left| \frac{Q(n^{1+\delta})}{n^{1+\delta}} - V^e \right| > \varepsilon/2 \right\}
\]

for \( n \in \mathbb{N} \). Then, Markov’s inequality and Theorem 3.2.3 imply, for \( 1 \leq m < 3 \) and large \( n \),

\[
\begin{align*}
P\{A_n\} &= P\left\{ \left| Q(n^{1+\delta}) - V^e n^{1+\delta} \right|^m > \left( \varepsilon/2 \right)^m n^{(1+\delta)m} \right\} \\
&\leq \frac{1}{\left( \varepsilon/2 \right)^m n^{(1+\delta)m}} \cdot \mathbb{E}[\left| Q(n^{1+\delta}) - V^e n^{1+\delta} \right|^m] \\
&\leq \frac{C_1}{(3-m)(\varepsilon/2)^m} \cdot \frac{1}{n^{m(1+\delta)/3}},
\end{align*}
\]

which is summable if \((1+\delta)m > 3\). Here \( \delta \) 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

\[
\forall n \geq n_0 \quad \left| \frac{Q(n^{1+\delta})}{n^{1+\delta}} - V^e \right| < \varepsilon/2. \tag{3.5.2}
\]

Using this we show that a.s. there exists \( n_1 \in \mathbb{N} \) such that

\[
\left| \frac{Q(t)}{t} - V^e \right| < \varepsilon \quad \text{for all real } t \geq n_1^{(1+\delta)}. \tag{3.5.3}
\]

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: \( |Q(t) - V^e t| \geq \varepsilon t \). Together with (3.5.2) we have, if \( n \) is large,

\[
\begin{align*}
|Q(t) - Q(n^{1+\delta})| &\geq |Q(t) - V^e t| - |Q(n^{1+\delta}) - V^e n^{1+\delta}| - |V^e t - V^e n^{1+\delta}| \\
&\geq \varepsilon t - \varepsilon/2 \cdot n^{1+\delta} - |V^e| (t - n^{1+\delta}) \\
&\geq \frac{\varepsilon}{4} n^{1+\delta}.
\end{align*}
\]  

(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 \(\left\lfloor \frac{\varepsilon}{4} n^{1+\delta} \right\rfloor\) 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 \(\mathcal{N}(\cdot)\) be a rate \(C\) Poisson process. Then for large \(n\) the probability of the event \(3.5.4\) is bounded from above by

\[
2\mathbb{P}\{\mathcal{N}(2(1+\delta)n^\delta) \geq \frac{\varepsilon}{4} n^{1+\delta}\} \leq 2\mathbb{P}\{e^{\mathcal{N}(2(1+\delta)n^\delta)} \geq e^{\varepsilon/4 n^{1+\delta}}\} \\
\leq 2e^{-\varepsilon/4 n^{1+\delta}} \mathbb{E}[e^{\mathcal{N}(2(1+\delta)n^\delta)}] \\
= 2e^{-\varepsilon/4 n^{1+\delta}} \cdot e^{(e-1)2C(1+\delta)n^\delta}.
\]

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 \((\eta(t), \omega(t))\) 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 $\mathcal{M}_m(t-+1) = \mathcal{M}_m(t)$, 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 $(\eta, \omega, \{X_m\})$ 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 with a distribution described below. Thus $y$ itself jumps only within partition intervals $\mathcal{M}_t$ if it was not the highest label in the partition interval. 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 $\omega$ or $\eta$ 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,

$$
\begin{align*}
y(t) :&= \begin{cases} 
  a^y(t-)(t), & \text{with pr.} \frac{f(\omega X_y(t-)(t)(t)-1) - f(\eta X_y(t-)(t)(t))}{f(\omega X_y(t-)(t)(t))-f(\eta X_y(t-)(t)(t))}, \\
  b^y(t-)(t), & \text{with pr.} \frac{f(\omega X_y(t-)(t)(t))-f(\omega X_y(t-)(t)(t)-1)}{f(\omega X_y(t-)(t)(t))-f(\eta X_y(t-)(t)(t))}
\end{cases}
\end{align*}
$$

(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 $\omega$ or $\eta$ 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 $z(t-)$

141
belong to different parts after the jump then, independently of everything else,

\[
z(t) := \begin{cases} 
\frac{b^z(t^-)(t) - 1}{b^z(t^-)(t)}, & \text{with pr.} \quad \frac{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t)) + 1}{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t))}, \\
\frac{b^z(t^-)(t)}{b^z(t^-)(t)}, & \text{with pr.} \quad \frac{f(\eta_{X_{y(t^-)}(t)}(t)) + 1}{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t))}
\end{cases}
\]

(3.6.2)

when the denominator is non-zero, and \( z(t) := b^z(t^-)(t) \) when the denominator is zero. When \( \omega_{X_{y(t^-)}(t)}(t) = \eta_{X_{y(t^-)}(t)}(t) + 1 \), \( b^z(t^-)(t) - 1 \) is not an admissible value but in this case the probability in the first line is zero.

- Whenever any change occurs in either \( \omega \) or \( \eta \) 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,

\[
\begin{pmatrix} y(t) \\ z(t) \end{pmatrix} := \begin{cases} 
\begin{pmatrix} a^y(t^-)(t) \\ b^y(t^-)(t) - 1 \end{pmatrix}, & \text{with pr.} \quad \frac{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t)) + 1}{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t))}, \\
\begin{pmatrix} a^y(t^-)(t) \\ b^y(t^-)(t) \end{pmatrix}, & \text{with pr.} \quad \frac{f(\eta_{X_{y(t^-)}(t)}(t)) + 1}{f(\omega_{X_{y(t^-)}(t)}(t)) - f(\eta_{X_{y(t^-)}(t)}(t))}
\end{cases}
\]

(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 \( \omega_{X_{y(t^-)}(t)}(t) = \eta_{X_{y(t^-)}(t)}(t) + 1 \), \( b^z(t^-)(t) - 1 \) 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 \( Q^*(t) := X_{z(t)}(t) \) 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 \((\omega^-, \omega) := (\omega - \delta X_y, \omega)\) obeys basic coupling, as does the pair \((\eta, \eta^+) := (\eta, \eta + \delta X_z)\).

**Proof.** We write the proof for \((\omega^-, \omega)\). We need to show that, given the configuration \((\eta, \omega, \{X_m\}, y)\), the jump rates of \((\omega^-, \omega)\) 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 \( \omega^- \) and \( \omega \) 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

\[
y \text{ took on value } a^y \text{ with probability } \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)}, \tag{3.6.4}
\]

and

\[
y \text{ took on value } b^y \text{ with probability } \frac{f(\omega_i) - f(\omega_i - 1)}{f(\omega_i) - f(\eta_i)}, \tag{3.6.5}
\]

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 \( \eta \) and \( \omega \), 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, \( \omega \) jumps without \( \eta \). 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 $\omega$ undergoes a jump but $\omega^-$ 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) $X_\gamma = Q$ does not jump with $X_{b\beta}$, since it has label $a^\beta$ and not $b^\beta$ (this probability is zero if $\omega_i = \eta_i + 1$). In this case $\omega^-$ and $\omega$ perform the same jump and it occurs with rate

$$[f(\omega_i) - f(\eta_i)] \cdot \frac{f(\omega_i) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} = f(\omega_i^-) - f(\eta_i).$$

- With rate $p(\eta_i, \eta_i+1) = f(\eta_i)$, both $\eta$ and $\omega$ 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 $\omega^-$ jumps as well.

Summarizing we see that the rate for $(\omega^-, \omega)$ to jump together over $(i, i + 1)$ is $f(\omega_i^-)$, and the rate for $\omega$ to jump without $\omega^-$ is $f(\omega_i) - f(\omega_i^-)$. This is exactly what basic coupling requires.

A very similar argument can be repeated for $(\eta, \eta^+)$. \hfill \square

**Lemma 3.6.2.** Inequality (3.2.26) $y \leq z$ 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 $y \leq z$. \hfill \square

So far in this section everything is valid for a general zero range process with nondecreasing concave jump rate. Now we use the special concavity requirement (3.2.37). With $r \in (0, 1)$ from (3.2.37), define the geometric distribution

$$\nu(m) := \begin{cases} (1-r)^m, & m \geq 0 \\ 0, & m < 0. \end{cases} \quad (3.6.6)$$

**Lemma 3.6.3.** Conditioned on the process $(\eta, \omega)$, the bounds $y(t) \overset{d}{\leq} \nu$ and $z(t) \overset{d}{\geq} -\nu$ hold for all $t \geq 0$.

The proof of this lemma is achieved in three steps.

**Lemma 3.6.4.** Let $Y$ be a random variable with distribution $\nu$, and fix integers $a \leq b$ and $\eta < \omega$ so that $\omega - \eta = b - a + 1$. Apply the following operation to $Y$:

(i) if $a \leq Y \leq b$, apply the probabilities from (3.6.1) (equivalently, (3.6.4) and (3.6.5)) with parameters $a, b, \eta, \omega$ to pick a new value for $Y$;

(ii) if $Y < a$ or $Y > b$ then do not change $Y$.

Then the resulting distribution $\nu^*$ is stochastically dominated by $\nu$.
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,

\[
\sum_{\ell=a}^{m} \nu^*(\ell) \geq \sum_{\ell=a}^{m} \nu(\ell) \quad \text{or, equivalently,} \quad \sum_{\ell=m}^{b} \nu^*(\ell) \leq \sum_{\ell=m}^{b} \nu(\ell)
\]

for all \( a \leq m \leq b \). Notice that \( \nu^* \) gives zero weight on values \( a < m < b \) (if any), therefore the left hand-side of the second inequality equals \( \nu^*(b) \) for \( a < m \leq b \). Hence the above display is proved once we show

\[
\nu^*(b) \leq \nu(b), \quad \text{that is,}
\]

\[
\frac{f(\omega) - f(\omega - 1)}{f(\omega) - f(\eta)} \cdot \sum_{\ell=a}^{b} \nu(\ell) \leq \nu(b),
\]

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 \( \{\eta, \ldots, \omega\} \). If \( b < 0 \) then both sides are zero. If \( b \geq 0 \) then we have, by (3.2.37),

\[
\nu(\ell) \leq \nu(b) \cdot \ell^{b-a} \leq \nu(b) \cdot \prod_{z=\omega+b}^{\omega-1} \frac{f(z) - f(z - 1)}{f(z + 1) - f(z)} = \nu(b) \cdot \frac{f(\omega - b + \ell) - f(\omega - b + \ell - 1)}{f(\omega) - f(\omega - 1)}
\]

for each \( \ell \leq b \). The first inequality also takes into account possible \( \nu(\ell) = 0 \) values for negative \( \ell \)'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 \( a, b, \eta, \omega \). Let \( -\nu^* \) be the resulting distribution. Then \( \nu^* \) is stochastically dominated by \( \nu \).

**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

\[
\nu^*(-b) \geq \nu(-b), \quad \text{that is},
\]

\[
\frac{f(n+1) - f(\eta)}{f(\omega) - f(\eta)} \cdot \sum_{\ell=a}^b \nu(-\ell) \geq \nu(-b),
\] (3.6.8)

see (3.6.2). We have, by (3.2.37),

\[
\nu(-\ell) \geq \nu(-b) \cdot \nu^{b-\ell} \geq \nu(-b) \cdot \prod_{z=\eta+1}^{\eta+b-\ell} \frac{f(z+1) - f(\eta)}{f(\eta+1) - f(z)}
\]

\[
= \nu(-b) \cdot \frac{f(\eta + 1 + b - \ell) - f(\eta + b - \ell)}{f(\eta + 1) - f(\eta)}
\]

for each \( \ell \leq b \). The first inequality also takes into account possible \( \nu(-b) = 0 \) values for positive \( b \)'s. With this we can write

\[
\sum_{\ell=a}^b \nu(-\ell) \geq \nu(-b) \cdot \frac{f(\eta + 1 + b - a) - f(\eta)}{f(\eta + 1) - f(\eta)}
\]

which becomes (3.6.8) via \( \omega - \eta = b - a + 1 \).

\[\square\]

**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 \( M_i \), they move together from that moment. The same holds for \( z(\cdot) \).

\[\square\]

**Proof of Lemma 3.6.3.** Initially \( y(0) = 0 \) by definition, which is clearly a distribution dominated by \( \nu \) 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 \( \nu_1 \), then \( \nu_2 \), then \( \nu_3 \), etc.

Suppose \( \nu_k \overset{d}{\leq} \nu \), and let \( \nu^* \) be the distribution that would result from \( \nu \) by the \( k+1^{\text{st}} \) jump. Then \( \nu_{k+1} \overset{d}{\leq} \nu^* \) by \( \nu_k \overset{d}{\leq} \nu \) and Lemma 3.6.6, while \( \nu^* \overset{d}{\leq} \nu \) by Lemma 3.6.4. A similar argument proves the lemma for \( z(\cdot) \).

\[\square\]

### 3.7 Microscopic convexity of the exponential brick-layers 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 [96], and its normal fluctuations away from the characteristics were demonstrated in [95] (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 between the second class particle of a denser system and one of a sparser system,
2. a non-strict, but tight, domination between the single second class particle and a set of second class particles that are defined between the system in question and another system with a different density, (this means that the label of the second class particle must not be too much away from what it should be)
3. strictly concave or convex, in the second derivative sense, hydrodynamic flux function,
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 the concave case, but we include this convex case in full detail (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 [108] 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 [108], this technical point being the main reason for considering this particular family of jump rates in this section. Indeed, the lack of the proof of property 4 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 [101] 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 [101] 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 existence of the model with exponential jump rates was proved in [107]. The exact connection results between the second class particle and the current variance of the note [104] are used by [101]. Those require strong construction results (certain desired semigroup properties) which are not provided by [107] and therefore, to our knowledge, are not available. To close that gap, we reproduce the exact connection results of [104] in the Appendices of [100] 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 [96], and also treated in [97] and [108]. 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 \(f: \mathbb{Z} \to \mathbb{R}_{\geq 0}\). We only consider the totally asymmetric nearest neighbor case here, in which only deposition of bricks in the following way is allowed:

\[
\begin{align*}
(\omega_i, \omega_{i+1}) &\rightarrow (\omega_i - 1, \omega_{i+1} + 1) \\
h_i &\rightarrow h_i + 1
\end{align*}
\]

with rate \(f(\omega_i) + f(-\omega_{i+1})\). (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 \(\omega(\cdot)\) 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

\[
f(z) = e^{\beta(z-1/2)}.
\]

(3.7.2)
The construction of the bricklayers process with any nondecreasing $f$ that is bounded by an exponential function is given in [107] on a set of tempered configurations $\tilde{\Omega}$. This set consists of configurations with bounded asymptotic slope, the precise definition is given in [107]. As certain desired semigroup properties are not fully proved, we avoid technical difficulties in the proofs of [104] by reproducing its results for the TAEBLP in the Appendix of [100]. 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 $\omega_1, \omega_2, \ldots, \omega_n \in \tilde{\Omega}$, let $m \mapsto \ell(m)$ be a permutation that orders the $\omega_i$ values:

$$\omega_\ell(m) \leq \omega_\ell(m+1), \quad 1 \leq m < n.$$  

For simplicity, set $p(m) := f(\omega_\ell(m))$ and $q(m) := f(-\omega_\ell(m))$,

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 $m = 1, \ldots, n$, at rate $p(m) - p(m - 1)$, precisely bricklayers of $\omega_{\ell(m)}$, $\omega_{\ell(m+1)}$, $\ldots$, $\omega_{\ell(n)}$ place a brick on their right, and bricklayers of $\omega_{\ell(1)}$, $\omega_{\ell(2)}$, $\ldots$, $\omega_{\ell(m-1)}$ do not. Independently, at rate $q(m) - q(m+1)$, precisely bricklayers of $\omega_{\ell(1)}$, $\omega_{\ell(2)}$, $\ldots$, $\omega_{\ell(m)}$ place a brick on their left, and bricklayers of $\omega_{\ell(m+1)}$, $\omega_{\ell(m+2)}$, $\ldots$, $\omega_{\ell(n)}$ do not. Given the configurations $\omega_1, \omega_2, \ldots, \omega_n \in \tilde{\Omega}$, 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 $\omega_\ell(m)$ executes the move (3.7.1) with rate $p(m) = f(\omega_\ell(m))$, and the same move on column $h_{i-1}$ with rate $q(m) = f(-\omega_\ell(m))$.

Notice also that, due to monotonicity of $f$, a jump of $\omega_a$ without $\omega_b$ on the column $[i, i+1]$ by the bricklayers at site $i$ can only occur if $f(\omega_a^i) < f(\omega_b^i)$ which implies $\omega_a^i > \omega_b^i$. Also, a jump of $\omega_a$ without $\omega_b$ on the column $[i-1, i]$ by the bricklayers at site $i$ can only occur if $f(-\omega_b^i) < f(-\omega_a^i)$ which implies $\omega_a^i < \omega_b^i$. The result of any of these steps then cannot increase the number of discrepancies between the two processes, hence the name attractivity. Also, a sitewise ordering $\omega_a^i \leq \omega_b^i \ \forall i \in \mathbb{Z}$ is preserved by the basic coupling.
Hydrodynamics and exact identities for TAEBLP

Recall the jump rates (3.7.1). As described in Section 3.2.4, the process has product translation-invariant stationary distribution with marginals $\mu^\theta$

$$\mu^\theta(z) = \frac{e^{\theta z} f(z)}{Z(\theta)}$$  

that turn out to be of discrete Gaussian type, see [96] for the explicit formula. The density $\varrho(\theta) := E^\theta(\omega) \in \mathbb{R}$ is a strictly increasing function of the parameter $\theta \in \mathbb{R}$, and can take on any real value by the Appendix 3.A below. As before, $\mu^\varrho$, $P^\varrho$, $E^\varrho$, $\text{Var}^\varrho$, $\text{Cov}^\varrho$ will refer to laws of a density $\varrho$ stationary process.

The hydrodynamic flux in this case is

$$\mathcal{H}(\varrho) = E^\varrho[f(\omega) + f(-\omega)] = e^{\theta(\varrho)} + e^{-\theta(\varrho)}.$$  

As $f$ (3.7.2) is convex and nonlinear, the Appendix 3.B apply and yields a convex hydrodynamic flux with

$$\mathcal{H}''(\varrho) > 0.$$  

Recall the measure (3.2.17) from Section 3.2.5. The Appendix 3.A below apply to show that both $\mu^\varrho$ and $\hat{\mu}^\varrho$ are stochastically monotone in $\varrho$. As before, denote by $E$ the expectation w.r.t. the evolution of a pair $(\omega^-(\cdot), \omega^+(\cdot))$ started with initial data (recall (3.2.1))

$$\omega^-(0) = \omega(0) - \delta_0 \sim \bigotimes_{i \neq 0} \mu^\varrho \otimes \hat{\mu}^\varrho,$$  

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, $\omega^-(t) = \omega(t) - \delta_{Q(t)}$. We reprove Corollaries 2.4 and 2.5 of [104] in the Appendix of [100] that state that for any $i \in \mathbb{Z}$ and $t \geq 0$,

$$\text{Var}^\varrho(h_i(t)) = \text{Var}^\varrho(\omega) \cdot E[Q(t) - i]$$  

and

$$E(\omega(t)) = V^\varrho \cdot t,$$

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).

150
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 $\lambda$ and $\varrho$ such that $\varrho - \gamma_0 \leq \lambda_i \leq \varrho + \gamma_0$ for all $i \in \mathbb{Z}$, a joint process $(\eta(t), \omega(t), y(t), z(t))_{t \geq 0}$ can be constructed with the following properties.

- Initially $(\eta(0), \omega(0))$ is $\hat{\mu}_{\lambda, \varrho}$-distributed and the joint process $(\eta(\cdot), \omega(\cdot))$ 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

$$\omega^-(t) := \omega(t) - \delta X_{y(t)}(t) \quad \text{and} \quad \eta^+(t) := \eta(t) + \delta X_{z(t)}(t). \tag{3.7.8}$$

Then both pairs $(\eta, \eta^+)$ and $(\omega^-, \omega)$ 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 $\mathbb{Z}_{\geq 0}$ satisfying the tail bound

$$\nu_{\varrho, \gamma}(t)\{y : y \geq y_0\} \leq C t^{\kappa - 1} \gamma^{2\kappa - 3} y_0^{-\kappa} \tag{3.7.9}$$

for some fixed constants $3/2 \leq \kappa < 3$ and $C < \infty$, and such that if $\varrho - \gamma \leq \lambda_i \leq \varrho + \gamma$ 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. (Recall that the jump rates of $Q$ are given by the increments of the rate function, not by the rate function itself.) Since $f$ (3.7.2) does not have bounded increments, 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 $\{M_i(t)\}$ of the label space $\mathbb{Z}$ into intervals indexed by sites $i \in \mathbb{Z}$, with partition intervals given by $M_i(t)$ 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 $M_{X_m(t)}(t) = \{a_m(t), a_m(t) + 1, \ldots, b_m(t)\}$. 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

$$|M_{X_m(t)}(t)| = b^m(t) - a^m(t) + 1 = \omega_{X_m(t)}(t) - \eta_{X_m(t)}(t). \quad (3.7.11)$$

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 $X_m(t-)$ 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 $X_m(t-)$ 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 $X_m(t-)$. 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 $M_{X_m(t-)+1} = M_{X_m(t)}$, 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 $M_{X_m(t)-1} = M_{X_m(t)}$, 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 $\eta < \omega$,

$$p(\eta, \omega) = \frac{f(\omega) - f(\omega - 1)}{f(\omega) - f(\eta)} = \frac{f(-\eta) - f(-\eta - 1)}{f(-\eta) - f(-\omega)} = \frac{e^{\beta(\omega-\eta)} - e^{\beta(\omega-\eta-1)}}{e^{\beta(\omega-\eta)} - 1}$$

and

$$q(\eta, \omega) = \frac{f(-\omega + 1) - f(-\omega)}{f(-\eta) - f(-\omega)} = \frac{f(\eta + 1) - f(\eta)}{f(\omega) - f(\eta)} = \frac{e^{\beta} - 1}{e^{\beta(\omega-\eta)} - 1}.$$  \hspace{1cm} (3.7.12)

(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

$$p(1) = q(1) = 1, \quad p(d) \geq q(d), \quad p(d) + q(d) \leq 1 \quad \text{for } 2 \leq d \in \mathbb{Z}.$$  \hspace{1cm} (3.7.14)

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 $\omega$ or $\eta$ at site $X_{y(t-)}(t-)$ and, as a result of the jump, $\alpha^{y(t-)}(t) \neq \alpha^{z(t-)}(t)$, that is, $y(t-)$ and $z(t-)$ belong to different parts after the jump, we abbreviate

$$p = p(\eta_{X_{y(t-)}(t)}(t), \omega_{X_{y(t-)}(t)}(t)) \quad q = q(\eta_{X_{y(t-)}(t)}(t), \omega_{X_{y(t-)}(t)}(t))$$

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,

$$y(t) := \begin{cases} 
\alpha^{y(t-)}(t), & \text{with prob. } q, \\
b^{y(t-)}(t) - 1, & \text{with prob. } 1 - p - q, \\
b^{y(t-)}(t), & \text{with prob. } p,
\end{cases}$$

(3.7.14)
except for \( y(t) := a^{y(t-)}(t) = b^{y(t-)}(t) \) when the difference

\[
\omega_{X_y(t-)}(t) - \eta_{X_y(t-)}(t)(t)
\]

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 \( \omega \) or \( \eta \) 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

\[
p = p(\eta_{X_{y(t-)}(t)}, \omega_{X_{z(t-)}(t)}(t)) \quad q = q(\eta_{X_{z(t-)}(t)}(t), \omega_{X_{z(t-)}(t)}(t))
\]

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,

\[
z(t) := \begin{cases} 
a^z(t-)(t), & \text{with prob. } p, \\
a^z(t-)(t) + 1, & \text{with prob. } 1-p-q, \\
b^z(t-)(t), & \text{with prob. } q,
\end{cases}
\]

except for \( z(t) := a^z(t-)(t) = b^z(t-)(t) \) when the difference

\[
\omega_{X_{z(t-)}(t)}(t) - \eta_{X_{z(t-)}(t)}(t)(t)
\]

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 \( \omega \) or \( \eta \) 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

\[
\omega_{X_{y(t-)}(t)}(t) = \omega_{X_{z(t-)}(t)}(t) \quad \text{and} \quad \eta_{X_{y(t-)}(t)}(t) = \eta_{X_{z(t-)}(t)}(t),
\]

and we abbreviate

\[
p = p(\eta_{X_{y(t-)}(t)}(t), \omega_{X_{y(t-)}(t)}(t)) \quad q = q(\eta_{X_{y(t-)}(t)}(t), \omega_{X_{y(t-)}(t)}(t))
\]

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,

\[
\begin{pmatrix}
(a_y(t) - (t)) \\
(a_y(t) - (t))
\end{pmatrix}, \text{ with prob. } q,
\]

\[
\begin{pmatrix}
(b_y(t) - 1) \\
(a_y(t) - (t))
\end{pmatrix}, \text{ with prob. } (p - q) \land (1 - p - q),
\]

\[
\begin{pmatrix}
(b_y(t) - (t)) \\
(a_y(t) - (t))
\end{pmatrix}, \text{ with prob. } [2p - 1] + ,
\]

\[
\begin{pmatrix}
(b_y(t) - (t)) \\
(a_y(t) - (t)) + 1
\end{pmatrix}, \text{ with prob. } (p - q) \land (1 - p - q),
\]

\[
\begin{pmatrix}
(b_y(t) - (t)) \\
(b_y(t) - (t))
\end{pmatrix}, \text{ with prob. } q,
\]

except for \( y(t) = z(t) = a_y(t) - (t) = b_y(t) - (t) \) when the difference
\( \omega X_y(t) - \eta X_y(t) \) 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 \( Q(t) := X_y(t) \) or \( Q(t) := X_z(t) \) 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 \( (\omega^-, \omega) := (\omega - \delta X_y, \omega) \) obeys basic coupling, as
does the pair \( (\eta, \eta^+) := (\eta, \eta + \delta X_y). \)

**Proof.** We write the proof for \( (\omega^-, \omega) \). We need to show that, given the
configuration \( (\eta, \omega, \{X_m\}, y) \), the jump rates of \( (\omega^-, \omega) \) 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 \( \omega \)
or \( \eta \) moves. As the moves (3.7.14) or (3.7.16) by themselves never result in
a change of \( X_y(\cdot) \), 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 \( \omega \) at
sites \( i \neq Q \) happen as well to \( \omega^- \), 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
\[ y \text{ took a value } < b^y \text{ with prob. } 1 - p = \frac{f(\omega_i - 1) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} \quad (3.7.17) \]
and
\[ y \text{ took on value } b^y \text{ with prob. } p = \frac{f(\omega_i) - f(\eta_i)}{f(\omega_i) - f(\eta_i)} \quad (3.7.18) \]
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 \( \eta \) and \( \omega \), the following right moves of bricklayers at \( i \) can occur:

- With rate \( f(\omega_i) - f(\eta_i) \), \( \omega \) jumps without \( \eta \). 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 \( \omega \) undergoes a jump but \( \omega^- \) 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 \( \omega^- \) and \( \omega \) 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 \( \eta \) and \( \omega \) at site \( i \) move. No change occurs in the \( \omega^- - \eta \) particles, hence no change occurs in \( Q \). This implies that the process \( \omega^- \) jumps as well.

Summarizing we see that the rate for the bricklayers of \( (\omega^-, \omega) \) at site \( i \) to lay brick on their rights together is \( f(\omega_i^-) \), and the rate for the one of \( \omega \) to move without \( \omega^- \) 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

\[
y \text{ took on value } a^y \text{ with prob. } q = \frac{f(-\omega_i + 1) - f(-\omega_i)}{f(-\eta_i) - f(-\omega_i)} \quad (3.7.19)
\]

and

\[
y \text{ took a value } > a^y \text{ with prob. } 1 - q = \frac{f(-\eta_i) - f(-\omega_i + 1)}{f(-\eta_i) - f(-\omega_i)} \quad (3.7.20)
\]

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 \( \eta \) and \( \omega \), the following left moves of bricklayers at \( i \) can occur:

- With rate \( f(-\eta_i) - f(-\omega_i) \), \( \eta \) jumps without \( \omega \). 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
    \[
    \omega_i^-(-) = \omega_i(-) - 1 = \omega_i(t) - 1 = \omega_i^-(t) - 1
    \]
    since the difference \( Q \) disappears from site \( i \). Also,
    \[
    \omega_{i-1}^-(-) = \omega_{i-1}(-) = \omega_{i-1}(t) = \omega_{i-1}^-(t) + 1,
    \]
    since the difference \( Q \) appears at site \( i + 1 \). So in this case \( \omega^- \) undergoes a jump but \( \omega \) 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 \( \omega^- \) or \( \omega \) 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 \( \eta \) and \( \omega \) at site \( i \) move. No change occurs in the \( \omega - \eta \) particles, hence no change occurs in \( Q \). This implies that the process \( \omega^- \) jumps as well.

Summarizing we see that the rate for the bricklayers of \( (\omega^-, \omega) \) at site \( i \) to lay brick on their rights together is \( f(-\omega_i) \), and the rate for the one of \( \omega^- \) to move without \( \omega \) is \( f(-\omega_i^-) - f(-\omega_i) \). This is exactly what basic coupling requires.

A very similar argument can be repeated for \( (\eta, \eta^+) \). \qed
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 \( b^{(t-)}(t) = a^{(t-)}(t) + 1 \), we have, by (3.7.11), \( \omega_{X_{\eta_{(t-)}}(t)}(t) - \eta_{X_{\eta_{(t-)}}(t)}(t) = 2 \), 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.) \qed

Define the geometric distribution

\[
\nu(m) := \begin{cases} 
  e^{-\beta m}(1 - e^{-\beta}), & m \geq 0 \\
  0, & m < 0.
\end{cases}
\]  (3.7.21)

Lemma 3.7.4. Conditioned on the process \( (\eta, \omega) \), the bounds \( y(t) \geq -\nu \) and \( z(t) \leq \nu \) 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 in 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 \( \nu \), 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 \( a, b, \eta, \omega \) to pick a new value for \( Z \);

(ii) if \( Z < a \) or \( Z > b \) then do not change \( Z \).

Then the resulting distribution \( \nu^* \) is stochastically dominated by \( \nu \).

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 \( \nu^* \) 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 \( m = a \):

\[
\nu^*(a) \geq \nu(a),
\]  (3.7.22)

158
and $m = b - 1$:

$$
\sum_{\ell = a}^{b-1} \nu^*(\ell) \geq \sum_{\ell = a}^{b-1} \nu(\ell) \quad \text{that is,} \quad \nu^*(b) \leq \nu(b). \quad (3.7.23)
$$

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 $d = \omega - \eta = b - a + 1$:

$$
\nu^*(a) = p(d) \cdot \sum_{\ell = a}^{b} \nu(\ell)
$$

As for $(3.7.23)$, both sides become zero if $b < 0$. For $b \geq 0$ we have

$$
\nu^*(b) = q(d) \cdot \sum_{\ell = a}^{b} \nu(\ell)
$$

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 $M_i$, they move together from that moment.

Proof of Lemma 3.7.4. Follows from the previous three lemmas analogously to 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 $\mu^{\lambda, r}$ of $(3.2.23)$ and notice that it gives probability one on pairs of the form $(y, y)$ if $\lambda = r$. Define also $\mu^{\text{shock}, r}$ by

$$
\mu^{\text{shock}, r}(y, z) = \begin{cases} 
\mu^r(y), & \text{if } z = y + 1, \\
0, & \text{otherwise.}
\end{cases}
$$

159
With these marginals we define the shock product distribution
\[
\mu_{\text{shock}} : = \bigotimes_{i < 0} \mu_{\varrho + 1} \cdot \bigotimes_{i = 0} \mu_{\text{shock}} \cdot \bigotimes_{i > 0} \mu_{\varrho}, \tag{3.7.24}
\]
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 \(\mu_{\text{shock}}\) is the product distribution
\[
\bigotimes_{i < 0} \mu_{\varrho + 1} \cdot \bigotimes_{i = 0} \mu_{\varrho},
\]
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
\[\mu_{\varrho}(z - 1) = \mu_{\varrho}(z - 1), \quad z \in \mathbb{Z},\]
valid for the second marginal at \(i = 0\). This is specific to the definition (3.2.13) of \(\mu_{\varrho}\), and of the exponential rates (3.7.2), and to prove it we write, with \(\theta = \theta(\varrho)\),
\[
\mu_{\varrho}(z - 1) = \frac{f(z)}{e^\theta} \cdot \frac{e^{\theta z}}{f(z)!} \cdot \frac{1}{Z(\theta)} = \frac{e^{(\theta + \beta)z}}{f(z)!} \cdot \frac{1}{e^{\theta + \beta} Z(\theta)}.
\]
Summing this up for all \(z \in \mathbb{Z}\) gives one on the left hand-side, hence leads to
\[
Z(\theta + \beta) = \sum_{z = -\infty}^{\infty} \frac{e^{(\theta + \beta)z}}{f(z)!} = e^{(\theta + \beta)/Z(\theta)},
\]
which also implies
\[\varrho(\theta + \beta) = \frac{d}{d\theta} \log(Z(\theta + \beta)) = \varrho(\theta) + 1.
\]
We conclude that
\[\mu_{\varrho}(z - 1) = \frac{e^{(\theta + \beta)z}}{f(z)!} \cdot \frac{1}{Z(\theta + \beta)} = \mu_{\varrho(\theta + \beta)}(z) = \mu_{\varrho + 1}(z),\]
which finishes the proof of the lemma. \(\square\)
The translation of $\mu_\text{shock} \varrho$ is denoted by
\[
\tau_k \mu_\text{shock} \varrho := \bigotimes_{i<k} \mu_{\varrho+1}, \bigotimes_{i=k} \mu_\text{shock} \varrho, \bigotimes_{i>k} \mu_\varrho.
\]

The main tool we use is Theorem 1 from [108], which we reformulate here. $\mu S(t)$ will just denote the time evolution of a measure $\mu$ under the process dynamics:

**Theorem 3.7.8.** In the sense of bounded test functions on $\Omega \times \Omega$,
\[
\frac{d}{dt}(\tau_k \mu_{\text{shock}} \varrho) S(t) = (e^{\theta(\varrho+1)} - e^{\theta(\varrho)}) \cdot (\tau_{k+1} \mu_{\text{shock}} \varrho - \mu_{\text{shock}} \varrho) \\
+ (e^{-\theta(\varrho)} - e^{-\theta(\varrho+1)}) \cdot (\tau_{k-1} \mu_{\text{shock}} \varrho - \mu_{\text{shock}} \varrho).
\]

First some remarks. The exact role of the special exponential form of the rate function $f$ can be seen here: Theorem 3.7.8 from [108] is only valid if the rates are of these form, and this theorem is crucial for us to show that the required tail bound for $Q(t)$ holds. Without this and in the case of $f$ with unbounded increments, we do not even have a good linear bound on the second class particle.

The first interesting consequence of this theorem is that the measure $\mu_\text{shock} \varrho$ 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 $(\xi^- (0), \xi (0))$ have initial distribution $\mu_{\text{shock}} \varrho$ defined by (3.7.24). Then its later distribution evolves into a linear combination of translated versions of $\mu_{\text{shock}} \varrho$: at time $t$ the pair $(\xi^- (t), \xi (0))$ has distribution
\[
\mu_{\text{shock}} \varrho S(t) = \sum_{k=-\infty}^{\infty} P_k(t) \cdot \tau_k \mu_{\text{shock}} \varrho,
\]
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 $e^{\theta(\varrho+1)} - e^{\theta(\varrho)}$ to the right and $e^{-\theta(\varrho)} - e^{-\theta(\varrho+1)}$ to the left.

In particular, $Q\xi (\cdot)$, started from an environment $\mu_{\text{shock}} \varrho$, 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. $(\xi^- (\cdot), \xi (\cdot))$ 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 $\Theta_t$ for the whole random path, shifted to time $t$: $\Theta_t = (\xi^- (t+\cdot), \xi (t+\cdot))$. Expectation of the process,
started from $\tau_k\mu_{\text{shock}}^\varrho$, will be denoted by $E^{\langle k \rangle}$. Notice that under $E^{\langle k \rangle}$ 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 $E^{\langle \xi^- \cdot \xi \rangle}$ for the evolution of the pair $(\xi^-(\cdot), \xi(\cdot))$, started from the specific initial state $(\xi^-, \xi)$.

We aim for proving the semigroup property of $S(\cdot)$. The first step is

**Lemma 3.7.10.** Given times $0 < s < t$ and $k \in \mathbb{Z}$,

$$E^{\langle 0 \rangle}[g(\Theta_t) \mid Q^{\xi}(s) = k] = E^{\langle k \rangle}[g(\Theta_{t-s})].$$

**Proof.** The left hand-side is

$$E^{\langle 0 \rangle}[g(\Theta_t) \mid Q^{\xi}(s) = k] = \frac{E^{\langle 0 \rangle}[E^{\langle \xi^- \cdot \xi \rangle}(\xi^-(s), \xi(s))g(\Theta_{t-s}) \mid Q^{\xi}(s) = k]}{P^{\langle 0 \rangle}\{Q^{\xi}(s) = k\}}$$

$$= \frac{\sum_{j \in \mathbb{Z}} P^{\langle 0 \rangle}\{Q^{\xi}(s) = j\} E^{\langle j \rangle}[E^{\langle \xi^- \cdot (0) \cdot \xi^-(0) \rangle}(\xi^-(0), \xi(0))g(\Theta_{t-s}) \mid Q^{\xi}(0) = k]}{P^{\langle 0 \rangle}\{Q^{\xi}(s) = k\}}$$

$$= \frac{E^{\langle k \rangle}[g(\Theta_{t-s})]}{P^{\langle 0 \rangle}\{Q^{\xi}(s) = k\}},$$

where in the second equality we used that the distribution at time $s$ is a linear combination of shifted versions of $\mu_{\text{shock}}^\varrho$.

Next we prove the Markov property for $Q^{\xi}(\cdot)$.

**Lemma 3.7.11.** Let $n > 0$ be an integer, $\varphi_i$, $i = 0, \ldots, n$ bounded functions on $\mathbb{Z}$, and $0 = t_0 < t_1 < \cdots < t_n$. Then

$$E^{\langle 0 \rangle}\prod_{i=1}^n \varphi_i(Q^{\xi}(t_i) - Q^{\xi}(t_{i-1})) = \prod_{i=1}^n E^{\langle 0 \rangle}\varphi_i(Q^{\xi}(t_i - t_{i-1})).$$

**Proof.** The statement is trivially true for $n = 1$. We proceed by induction,
and assume the statement is true for \( n - 1 \). Then

\[
\mathbb{E}^{(0)} \prod_{i=1}^{n} \varphi_i(Q^\xi(t_i) - Q^\xi(t_{i-1}))
\]

\[
= \sum_{j \in \mathbb{Z}} P^{(0)}\{Q^\xi(t_1) = j\} \varphi_1(j) \cdot \mathbb{E}^{(0)} \prod_{i=2}^{n} \varphi_i(Q^\xi(t_i) - Q^\xi(t_{i-1})),
\]

\[
= \sum_{j \in \mathbb{Z}} P^{(0)}\{Q^\xi(t_1) = j\} \varphi_1(j) \cdot \mathbb{E}^{(0)} \prod_{i=2}^{n} \varphi_i(Q^\xi(t_i - t_1) - Q^\xi(t_{i-1} - t_1))
\]

\[
= \sum_{j \in \mathbb{Z}} P^{(0)}\{Q^\xi(t_1) = j\} \varphi_1(j) \cdot \mathbb{E}^{(0)} \prod_{i=2}^{n} \varphi_i(Q^\xi(t_i - t_1))
\]

\[
= \prod_{i=1}^{n} \mathbb{E}^{(0)} \varphi_i(Q^\xi(t_i - t_{i-1})).
\]

The second equality uses Lemma 3.7.10, the third one uses the fact that \( \phi \)'s only depend on \( Q^\xi \)-differences, and the fourth one follows from the induction hypothesis.

\( \Box \)

**Proof of Corollary 3.7.9.** We know that at any fixed time \( t > 0 \) the distribution of \((\xi^-(t), \xi(t))\) is a linear combination of shifted versions of \( \mu^{\text{shock}} \). The shift is traced by the second class particle \( Q^\xi(t) \), therefore the differential equation

\[
\frac{d}{dt} P^{(0)}\{Q^\xi(t) = k\}
\]

\[
= (e^{\theta(t+1)} - e^{\theta(t)}) \cdot (P^{(0)}\{Q^\xi(t) = k + 1\} - P^{(0)}\{Q^\xi(t) = k\})
\]

\[
+ (e^{-\theta(t)} - e^{-\theta(t+1)}) \cdot (P^{(0)}\{Q^\xi(t) = k - 1\} - P^{(0)}\{Q^\xi(t) = k\})
\]

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 \((\xi^-, \xi)\)). 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 \( \mu^{\text{shock}} \) is an asymmetric simple random walk with rates as stated in the Corollary.

\( \Box \)

**Lemma 3.7.12.** Let \((\omega^-, \omega)\) 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

\[
P\{|Q(t)| > K\} \leq e^{-CK}
\]

163
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
\[
(\lambda_i, \varphi_i) := \begin{cases} (\varphi, \varphi + 1), & \text{for } i \leq 0, \\ (\varphi, \varphi), & \text{for } i > 0. \end{cases}
\]

Draw the pair $(\zeta(0), \xi(0))$ from the product distribution of coupling measures (3.2.23)
\[
\bigotimes_{i \in \mathbb{Z}} \mu_{\lambda_i, \varphi_i}.
\]

Then $\xi(0)$ has distribution
\[
\bigotimes_{i \leq 0} \mu_{\varphi + 1} \cdot \bigotimes_{i < 0} \mu_\varphi,
\]
in agreement with (3.7.25).

Let now the pair $(\zeta(\cdot), \xi(\cdot))$ evolve in the basic coupling, and let them play the role of $(\eta(\cdot), \omega(\cdot))$ of Section 3.7.1. This results in the pair $(\zeta(\cdot), \xi^+(\cdot))$ with a second class particle $Q^\zeta(\cdot)$ and the pair $(\xi^-(\cdot), \xi(\cdot))$ with a second class particle $Q^\xi(\cdot)$ such that $Q^\zeta(t) \leq Q^\xi(t)$, 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 $\omega^-(0)$ in Assumption 3.2.2 and of $\zeta(0)$ above only differ by $\omega^0(0) \sim \hat{\mu}^\varphi$, while $\zeta_0(0) \sim \mu^\varphi$. Therefore

\[
\mathbb{P}\{Q(t) > K\} = \sum_{z = -\infty}^{\infty} \mathbb{P}\{Q(t) > K \mid \omega_0^-(0) = z\} \cdot \mu^\varphi(z) \left(\frac{\hat{\mu}^\varphi(z)^2}{\mu^\varphi(z)}\right)^{\frac{1}{2}}
\]

\[
= \sum_{z = -\infty}^{\infty} \mathbb{P}\{Q^\xi(t) > K \mid \zeta_0(0) = z\} \cdot \mu^\varphi(z) \left(\frac{\hat{\mu}^\varphi(z)^2}{\mu^\varphi(z)}\right)^{\frac{1}{2}}
\]

\[
\leq \left[ \sum_{z = -\infty}^{\infty} \mathbb{P}\{Q^\xi(t) > K \mid \zeta_0(0) = z\} \cdot \mu^\varphi(z) \right]^{\frac{1}{2}} \cdot \left[ \sum_{y = -\infty}^{\infty} \hat{\mu}^\varphi(y)^2 \right]^{\frac{1}{2}}
\]

\[
= \mathbb{P}\{Q^\xi(t) > K\} \cdot \left[ \sum_{y = -\infty}^{\infty} \hat{\mu}^\varphi(y)^2 \mu^\varphi(y) \right]^{\frac{1}{2}}.
\]

We are done as soon as we show that $\hat{\mu}^\varphi(y)/\mu^\varphi(y)$ is uniformly bounded in $y$. With the exponential rates (3.7.2) one obtains from (3.2.17)

\[
\frac{\hat{\mu}^\varphi(y)}{\mu^\varphi(y)} = C \sum_{z = y+1}^{\infty} (z - y) e^{-\frac{\beta}{2}(z - \frac{\beta}{2})^2 + \frac{\beta^2}{4}(y - \frac{\beta}{2})^2} = C \sum_{k = 1}^{\infty} (k + y - y) e^{-\frac{\beta}{2}k^2 - \frac{\beta}{2}ky + \theta k}.
\]

164
This is uniformly bounded for large $y$’s since then $ye^{-\beta y} < 1$. For large negative $y$’s one uses the equivalent form

$$\hat{\mu}(y) = \frac{1}{\text{Var}^\theta(\omega_0)} \sum_{z=-\infty}^{y} (q-z)\mu(z)$$

of (3.2.17) and writes

$$\frac{\hat{\mu}(y)}{\mu(y)} = C \sum_{z=-\infty}^{y} (q-z)e^{-\frac{\theta}{2}(z-y)^2+\frac{\beta}{2}(y-z)^2} = C \sum_{k=1}^{\infty} (k-y+q)e^{-\frac{\theta}{2}k^2+\beta ky-\theta k}$$

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 $(\zeta(0), \xi(0))$ in distribution

$$\otimes_{i \in \mathbb{Z}} \mu_{\lambda_i, \varrho_i}.$$ 

Now the roles of the pair $(\zeta(\cdot), \zeta^+(\cdot))$ with a second class particle $Q^\zeta(\cdot)$ and the pair $(\xi(\cdot), \xi^+(\cdot))$ 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 $Q^\xi$ 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 $\mu^\varrho$ and $\hat{\mu}^\varrho$ defined in (3.2.13) and (3.2.17), respectively, are stochastically monotone as functions of $\varrho$. We start with a simple

**Lemma 3.A.1.** Fix a function $\varphi(\omega)$ on $\mathbb{Z}$, bounded by a polynomial. Then $E^\vartheta(\varphi(\omega))$ is differentiable in $\theta$ on $(\theta, \bar{\theta})$, and

$$\frac{d}{d\theta} E^\vartheta(\varphi(\omega)) = \text{Cov}^\vartheta(\varphi(\omega), \omega).$$

**Proof.** Convergence of the series involved in $E^\vartheta(\varphi(\omega))$ can be verified via the ratio test, even after differentiating the terms. Since $\mu^\vartheta$ is the exponentially
For any \( \theta < \bar{\theta} \), the state sum \( (3.2.12) \) satisfies
\[
\frac{d}{d\theta} \log Z(\theta) = \frac{1}{Z(\theta)} \sum_{z=\omega_{\min}}^{\omega_{\max}} z e^{\theta z} f(z)! = E^\theta(\omega) =: \varrho(\theta), \quad (3.3.1)
\]
\[
\frac{d^2}{d\theta^2} \log Z(\theta) = \frac{d}{d\theta} \varrho(\theta) = VAR^\theta(\omega). \quad (3.3.2)
\]

The function \( \varrho(\theta) \) is strictly increasing and maps \((\bar{\theta}, \bar{\theta})\) 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 \( \bar{\theta}, \bar{\theta} \) to \( \omega_{\min}, \omega_{\max} \). First let us consider the case when \( \bar{\theta} < \infty \). Then \( \omega_{\max} = \infty \) and Fatou’s lemma implies
\[
\liminf_{\theta \uparrow \bar{\theta}} Z(\theta) = \liminf_{\theta \uparrow \bar{\theta}} \sum_{z=1}^{\omega_{\max}} e^{\theta z} f(z)! \geq \sum_{z=1}^{\omega_{\max}} \liminf_{\theta \uparrow \bar{\theta}} e^{\theta z} f(z)! = \sum_{z=1}^{\omega_{\max}} e^{\theta z} f(z)! = \infty
\]

since for \( z > 0 \)
\[
\frac{e^{\theta z}}{f(z)!} = \prod_{y=1}^{z} \frac{e^{\theta y}}{f(y)!} \geq 1
\]
by definition of \( \bar{\theta} \) and \( f \) being nondecreasing. This shows that \( \log Z(\theta) \) takes on arbitrarily large values as \( \theta \uparrow \bar{\theta} \). We also know that it is a smooth and convex function on \((\theta, \bar{\theta})\) (see \( (3.3.2) \)). This implies that its derivative \( (3.3.1) \) is not bounded from above i.e., arbitrarily large \( \varrho \) values can be achieved. The same reasoning works in case \( \bar{\theta} > -\infty \) for arbitrarily large negative \( \varrho \) values.

When \( \bar{\theta} = \infty \) then, regardless whether \( \omega_{\max} \) is finite or infinite, fix any \( 0 \leq y < \omega_{\max} \) and write
\[
\varrho(\theta) = E^\theta(\omega \cdot 1_{\{\omega > y\}}) + E^\theta([\omega]^+ \cdot 1_{\{\omega \leq y\}}) - E^\theta([\omega]^- \cdot 1_{\{\omega \leq y\}})
\]
\[
\geq (y + 1) \cdot P^\theta(\omega > y) - E^\theta([\omega]^-) \cdot 1_{\{\omega \leq y\}}
\]
\[
\geq (y + 1) - (y + 1) \cdot P^\theta(\omega \leq y) - \sqrt{E^\theta((\omega^-)^2)} \cdot \sqrt{P^\theta(\omega \leq y)}
\]
\[
\geq (y + 1) - (y + 1) \cdot P^\theta(\omega \leq y) - \sqrt{E^{\theta_0}([\omega^-]^2)} \cdot \sqrt{P^\theta(\omega \leq y)} \quad (3.3.3)
\]
for a fixed \( \theta < \theta_0 < \theta \). Here \([\cdot]^+\) and \([\cdot]^−\) denotes the positive and the negative part of the expression in the brackets. The last inequality follows by monotonicity of \( \mu^\theta \) in \( \theta \) and \(([\omega]^−)^2\) being a nonincreasing function of \( \omega \).

For any \( \omega^{\min} − 1 < z \leq y \) and \( \theta > \theta_0 \),

\[
\frac{\mu^\theta(z)}{\mu^\theta(y + 1)} = \prod_{x = z}^{y} \frac{\mu^\theta(x)}{\mu^\theta(x + 1)} = \prod_{x = z}^{y} \frac{f(x + 1)}{e^\theta} \leq \left( \frac{f(y + 1)}{e^\theta} \right)^{y−z+1}.
\]

Given \( 0 \leq y < \omega^{\max} \) and \( 1 > \varepsilon > 0 \), there is a large enough \( \theta \) which makes the last fraction smaller than \( \varepsilon \). With such a choice we have

\[
P^\theta{\{\omega \leq y\}} = \sum_{z = \omega^{\min}}^{y} \mu^\theta(z) \leq \mu^\theta(y + 1) \sum_{z = \omega^{\min}}^{y} \varepsilon^{y−z+1} \leq \varepsilon \cdot \frac{1 − \varepsilon^{y^{\omega^{\min}−1}}}{1 − \varepsilon}.
\]

Therefore, for the case of a finite \( \omega^{\max} \), choosing \( y = \omega^{\max} − 1 \) and large \( \theta \) makes \( 3.3.3 \) arbitrarily close to \( \omega^{\max} \). When \( \omega^{\max} = \infty \), the argument shows that \( \rho(\theta) \geq y + 1 \) can be achieved for any \( y \geq 0 \). A similar computation demonstrates that any density towards \( \omega^{\min} \) can be reached when \( \theta = -\infty \).

**Corollary 3.3.3.** The measures \( \mu^\theta \) are stochastically nondecreasing in \( \varrho \).

**Proof.** Since \( \varrho \) and \( \theta \) are strictly increasing functions of each other, it is equivalent to show monotonicity of \( \mu^\theta \). This follows if we can show \( 0 \leq \frac{d}{d \varrho} E^\theta(\varphi(\omega)) \) for an arbitrary bounded nondecreasing function \( \varphi \). Lemma \( 3.3.1 \) transforms this derivative into the covariance of \( \varphi(\omega) \) and \( \omega \), which is non-negative due to \( \varphi \) being nondecreasing.

Monotonicity of \( \hat{\mu}^\theta \) requires somewhat more of a convexity argument.

**Proposition 3.3.4.** The family of measures \( \hat{\mu}^\theta \), defined in \( 3.2.17 \), is stochastically nondecreasing in \( \varrho \).

**Proof.** Start by rewriting the definition:

\[
\hat{\mu}^\theta(y) = \frac{E^\theta([\omega − \varrho] \cdot 1\{\omega > y\})}{\text{Var}^\theta(\omega)} = \frac{\text{Cov}^\theta(\omega, 1\{\omega > y\})}{\text{Cov}^\theta(\omega, \omega)}
\]

\[
= \frac{d}{d \varrho} P^\theta{\{\omega > y\}} \bigg|_{\varrho(\theta)} = \frac{d}{d \theta} P^\theta{\{\omega > y\}} \bigg|_{\theta(\varrho)}
\]

Let us denote the \( \hat{\mu}^\theta \)-expectation by \( \hat{E}^\varrho \). Fix a bounded nondecreasing function \( \varphi \). We need to show

\[
0 \leq \frac{d}{d \varrho} \hat{E}^\varrho \varphi(\omega).
\]
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 $\theta$ in small open neighborhoods. This follows from the definitions of $\hat{\theta}$ and $\tilde{\theta}$ and the assumption $\hat{\theta} < \theta(q) < \tilde{\theta}$.

\[
\hat{E}^\theta \varphi(\omega) = \sum_{y=\omega_{\min}}^{\omega_{\max}} \varphi(y) \cdot \frac{d}{d\varrho} P^\varrho \{ \omega > y \}
= \sum_{y=\omega_{\min}}^{\omega_{\max}} \varphi(y) \cdot \frac{d}{d\varrho} P^\varrho \{ \omega > y \} - \mathbf{1} \{ 0 \geq y \}
= \frac{d}{d\varrho} \sum_{y=\omega_{\min}}^{\omega_{\max}} \varphi(y) \cdot \left[ P^\varrho \{ \omega > y \} - \mathbf{1} \{ 0 \geq y \} \right]
= \frac{d}{d\varrho} \sum_{y=\omega_{\min}}^{\omega_{\max}} \varphi(y) \cdot \left[ 1 \{ \omega > y \} - \mathbf{1} \{ 0 \geq y \} \right]
= \frac{d}{d\varrho} \sum_{y=\omega_{\min}}^{\omega_{\max}} \varphi(y) \cdot \left[ 1 \{ \omega > y > 0 \} - \mathbf{1} \{ 0 \geq y \geq \omega \} \right]
= \frac{d}{d\varrho} \left[ \sum_{y=1}^{\omega_{\max}} \varphi(y) - \sum_{y=\omega}^{0} \varphi(y) \right] = \frac{d}{d\varrho} E^\varrho \Phi(\omega).
\]

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 $\varphi$ determines a (non-strictly) convex function $\Phi$ with $\Phi(1) = 0$, and vice-versa. Hence the convexity theorem \[103\] Theorem 2.1 establishes that

\[
\frac{d}{d\varrho} \hat{E}^\varrho \varphi(\omega) = \frac{d^2}{d\varrho^2} E^\varrho \Phi(\omega) \geq 0.
\]

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 $H : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ of (3.2.14) is given by

\[
H(\varrho) = E^\varrho f(\omega).
\]

The results of \[103\] 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 \( H \) is also convex (or concave, respectively). Moreover, in this case \( H''(\varrho) > 0 \) (or \( H''(\varrho) < 0 \), 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 [95].

Next we show in the general case (i.e. not only for zero range processes) that \( H(\varrho) \) is well defined, and is infinitely differentiable. (We use third derivatives in the proof of Theorem [3.2.3].) The function \( H(\varrho) \) is, in general, the expected net growth rate w.r.t. \( \mu_\varrho \) 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 \( H(\varrho) \).

**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 \( \theta < \theta < \bar{\theta} \),

\[
E^\theta[(p(\omega_0, \omega_1) + q(\omega_0, \omega_1))g(\omega_0, \omega_1)] < \infty.
\]

**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

\[
\sum_{y=\omega_{\min}+1}^{\omega_{\max}} \sum_{z=\omega_{\min}}^{\omega_{\max}-1} p(y, z) \cdot g(y, z) \cdot e^{\theta(y+z)} \cdot \frac{f(y)! \cdot f(z)!}{f(y+1)! \cdot f(z+1)!} \cdot \frac{1}{Z(\theta)^2}.
\]

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

\[ p(y, z) = p(z+1, y-1) \cdot \frac{f(y)}{f(z+1)} \]

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

\[
p(1, 0) \cdot \frac{\omega_{\max}}{f(1)} \sum_{y=1}^{\omega_{\max}} \sum_{z=1}^{\omega_{\max}-1} g(y, z) \cdot e^{\theta(y+z)} \cdot \frac{1}{Z(\theta)^2}.
\]

- \( y \leq 0, z > 0 \): In this case

\[
p(y, z) \leq p(1, 0),
\]

169
and the corresponding part of the summation is bounded by
\[ p(1, 0) \cdot \sum_{y=\omega_{\min}+1}^{\omega_{\max}} \sum_{z=1}^{\omega_{\max}-1} g(y, z) \cdot \frac{e^{\theta(y+z)}}{f(y)! \cdot f(z)!} \cdot \frac{1}{Z(\theta)^2}. \]

- **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
  \[ p(1, 0) f(1) \cdot \sum_{y=\omega_{\min}+1}^{\omega_{\max}} \sum_{z=\omega_{\min}}^{0} g(y, z) \cdot \frac{e^{\theta(y+z)}}{f(y)! \cdot f(z)!} \cdot \frac{1}{Z(\theta)^2}. \]

- **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 \( \theta < \theta < \bar{\theta} \) 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. \( \square \)
Bibliography


