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

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

PhD Thesis booklet

Júlia Komjáthy

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

2012

1 Mixing times of random walks on wreath product graphs

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 [42].

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

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

We study the *total variation* or TV and the *uniform* mixing time of the models described below, corresponding to mixing in the ℓ_1 and ℓ_∞ 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 spectral gap of the chain by

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

where λ_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

$$t_{\text{rel}}(G) \leq t^{TV}_{\text{mix}}(G, 1/4) \leq t^{u}_{\text{mix}}(G, 1/4),$$

see [2] or [43]. 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 [3] introduced random stopping times achieving stationary measure. They were studied more by Lovász, Winkler [48, 49], (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 [3] and the recent papers connecting hitting times to mixing times and stopping rules by Sousi and Peres [53] and independently by Oliveira [50], or blanket times and cover times to the maxima of Gaussian free fields by Ding, Lee and Peres [28]. For a more comprehensive overview of Markov Chain mixing we refer the reader to the indispensable book [2] by Aldous and Fill or [43] by Levin, Peres and Wilmer as our main references.

In the first chapter of my thesis we investigate 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, see Figure 1. If the machines are just on-off lamps (Figure 2), 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 [1] 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 $X^{\diamond}$ associated with X and Z is the random walk on the wreath product $H \wr G$, the graph whose vertices consist of pairs $(\underline{f}, x)$ where $\underline{f} = (f_v)_{v \in V(G)}$

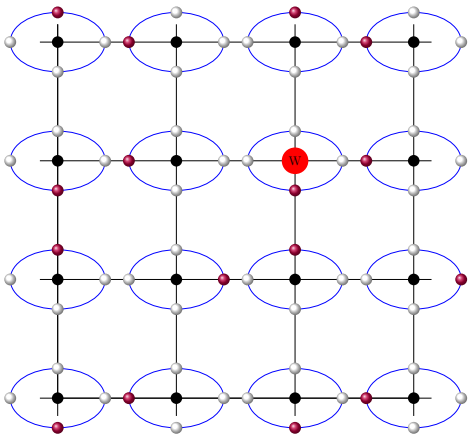


Figure 1: A typical state of the generalized lamplighter walk. Here $H = \mathbf{Z}_4$ and $G = \mathbf{Z}_4^2$; the red bullets on each copy of H represents the state of the lamps over each vertex $v \in G$ and the walker is drawn as a red W bullet.

is a labeling of the vertices of G by elements of H and x is a vertex in G. In each step, $X^\diamond$ moves from a configuration $(\underline{f}, x)$ by updating x to y using the transition rule of X and then independently updating both f_x and f_y according to the transition probabilities on H; f_z for $z \neq x, y$ remains unchanged.

Relaxation time and TV-mixing on general graphs with $\mathbf{Z}_2 = 0-1$ 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 $\mathbf{Z}_2 \wr G$, 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 $\mathbf{Z}_2 \wr G$ is related to the maximal expected hitting time of the graph, $t_{\text{hit}}(G)$, defined as $t_{\text{hit}}(G) = \max_{x,y \in G} \mathbf{E}(\tau_y | X_0 = x)$, τ_y denoting the time needed to reach vertex $y \in G$. The total variation mixing of $\mathbf{Z}_2 \wr G$ 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 $t_{\text{cov}}(G)$ 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 $\mathbf{Z}_2 \wr G$ one needs to understand the exponential moment $\mathbf{E}[2^{\mathcal{U}(t)}]$ of the not-yet-visited vertices $\mathcal{U}(t)$. One needs to determine the time when this quantity drops below $1+\varepsilon$,

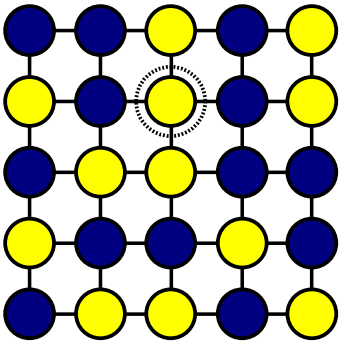


Figure 2: A typical configuration of the lamplighter over a 5×5 planar grid. The colors indicate the state of the lamps and the dashed circle gives the position of the lamplighter.

which is much harder to analyze; so it was a gap left between the lower and upper bound on the uniform mixing time for $\mathbf{Z}_2 \wr G$ in [52].

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 the construct all the eigenfunctions of $H \wr K_n$ from the eigenfunctions of H, see [43]. Nathan Levi [44] in his thesis investigated general lamplighters with $H = \mathbf{Z}_2^d$, the d-dimensional hypercube, but his mixing time bounds did not match in general. Further, Fill and Schoolfield [35] 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, in my thesis we study uniform mixing with $\mathbf{Z}_2$ 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.)

Based on a paper with Yuval Peres 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. To state our main theorem, we need one definition:

Definition 1.1. *A randomized stopping time τ is called a strong stationary time for the Markov chain X_t on G if*

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

that is, the position of the walk when it stops at τ is independent of the value of τ.

Further, a state $h(v) \in V(G)$ is called a halting state for a stopping time τ 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. *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*

$$c_1 \leq \frac{t_{rel}(H \wr G)}{t_{hit}(G) + |G| t_{rel}(H)} \leq C_1, \tag{1.1}$$

Theorem 1.3. *Assume that the conditions of Theorem 1.2 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*

$$\begin{aligned} &c_2 \left(t_{cov}(G) + |G|(t_{rel}(H) \log |G| + t_{mix}(H))\right) \leq t_{mix}(H \wr G), \\ &t_{mix}(H \wr G) \leq C_2 \left(t_{cov}(G) + |G| t_{mix}(H, \frac{1}{|G|})\right). \end{aligned} \tag{1.2}$$

If further the Markov chain is such that

(A) *There is a strong stationary time τ_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 is sharp.

Remark 1.4. *The laziness assumption on the transition matrix Q on H is only used to get the term $c_2|G|t_{mix}(H)$ in (1.2). All the other bounds hold without the laziness assumption.*

Remark 1.5. *If the Markov Chain on H is such that*

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

then the upper bound matches the lower bound. This holds for many natural chains such as lazy random walk on hypercube $\mathbf{Z}_2^d$, tori $\mathbf{Z}_n^d$, 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 1.6. *Many examples where Assumption (**A**) holds are given in the thesis of Pak [51], including the cycle $\mathbf{Z}_n$, the hypercube $\mathbf{Z}_2^d$ and more generally tori $\mathbf{Z}_n^d, n, d \in \mathbf{N}$ and dihedral groups $\mathbf{Z}_2 \ltimes \mathbf{Z}_n, n \in \mathbf{N}$ 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 $GL(n, \mathbb{F}_q)/(GL(k, \mathbb{F}_q) \times GL(n-k, \mathbb{F}_q))$.*

Then, based on the joint paper with Miller an Peres [40] we give matching upper bound for the mixing time in the uniform metric of $\mathbf{Z}_2 \wr G$ up to universal constants in terms of the parameters of G to the lower bound given in [52, Theorem 1.4] by Peres and Revelle. We show that for vertex transitive base graph G, the uniform mixing time of the lamplighter chain on G satisfies

$$t^u_{\text{mix}}(\mathbf{Z}_2 \wr G, 1/4) = O\Big(|G|(t_{\text{rel}}(H) + \log|G|)\Big)$$

under some conditions which capture the local transience of the base graph G. Further we show that these conditions are satisfied by the hypercube $\mathbf{Z}_2^d$ or in general the d-dimensional tori $\mathbf{Z}_n^d$ with d and n both possibly tending to infinity:

Theorem 1.7. *There exists constants $C_1, C_2 > 0$ such that*

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

More generally,

$$C_1 \leq \frac{t_u(\mathbf{Z}_2 \wr \mathbf{Z}_n^d)}{dn^{d+2}} \leq C_2 \text{ for all } n \geq 2 \text{ and } d \geq 3.$$

Prior to this work, the best known bound [52] for $t_{\text{u}}(\mathbf{Z}_2 \wr \mathbf{Z}_2^d)$ was

$$C_1 2^d d \leq t_{\text{u}}(\mathbf{Z}_2 \wr \mathbf{Z}_2^d) \leq C_2 2^d d \log d$$

for $C_1, C_2 > 0$.

1.0.1 A few words about the proofs

The methods for proving Theorems 1.2 and 1.3 are a mixture of various methods on the field. The lower bound on the relaxation time is based on Dirichlet-form methods. The upper bound uses the following steps:
1. Construct a strong stationary time τ on $H \wr G$,
2. Give a precise estimate on the tail $\mathbf{P}(\tau > t)$ of the strong stationary time,
3. Use that the second eigenvalue $\lambda_2 \leq \lim_{t\to\infty} \mathbf{P}(\tau > t)^{1/t}$, calculate the order of the limit.

The proof of Theorem 1.3 also uses strong stationary times: to get the upper bound, one needs to estimate the time t^* when $\mathbf{P}(\tau > t^*) < 1/4$ for the strong stationary time constructed for the chain on $H \wr G$, but to be able to do so we need to use the equivalence of blanket and cover times, a recent result [28]. The proof with the assumption 1.3 is based on giving a lower bound on the probability for $\tau > t$. The proof of the lower bound

without the assumption uses distinguishing set method and also the relation of mean-optimal stopping times and mixing times.

The proof of uniform mixing time result performs a rather careful analysis of the process by which $\mathcal{U}(t)$, the set of not-yet-covered vertices of the base graph G is decimated by the simple random walk. The key idea is to break the process of coverage into two different regimes, depending on the size of $\mathcal{U}(t)$. In both regimes, we use stochastic domination arguments: for large $\mathcal{U}(t)$ we show that in the time depending on the current size of $\mathcal{U}(t)$, the chain covers at least some fixed amount of vertices with large probability. In the small regime we show that again, during some time interval depending on the size of $U(t)$, the chain covers at least half of the vertices in the set with large probability. Then, we can build a careful estimate on $\mathbf{E}(2^{\mathcal{U}(t)})$. Throughout the proof, we use the behavior of the Green's function of the chain.

Now we turn to describe the second chapter of the thesis.

2 Generating hierarchical scale-free graphs from fractals

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) [57] and by Gilbert (1959) [36], and by Erdős and Rényi (1960) [31]. 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 [37] or [20] as an introduction.

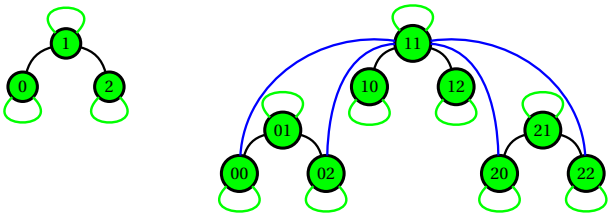
Parallel to the discussion of the ER and related models, there have been a considerable amount of attention paid to the study of complex networks like the World Wide Web, social networks, or biological networks in the last two decades.

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

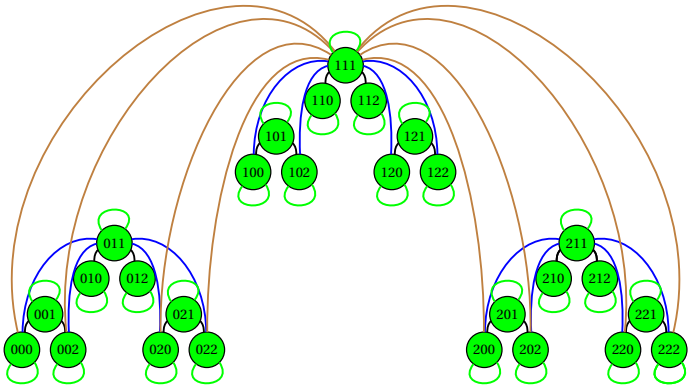
high clustering. Here and later, the local clustering coefficient of a vertex refers to the proportion of closed triangles and all edge-pair starting from the given vertex.

The Watts and Strogatz model [59] 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 β 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.

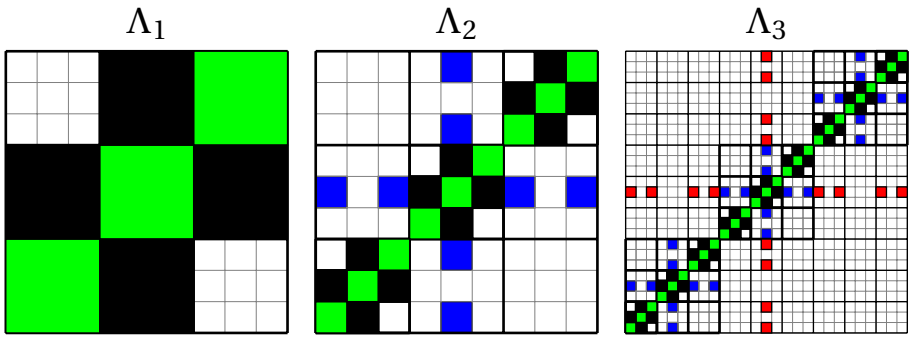
A different attempt to model real networks resulted in the construction of numerous new, more dynamical and growing network models, see e.g. [19], [20], [23], [29], [39]. Most of them use a version of preferential attachment and are of probabilistic nature. In particular, the scale free property - the 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 [60, 56, 25]. The model was heuristically introduced by Barabási and Albert [17], and the first who investigated the model rigorously were Bollobás, Riordan, Spencer and Tusnády [22], and the mathematically rigorous construction was done by Bollobás and Riordan [21]. In the preferential attachment model (sometimes also called Barabási Albert model) discussed by Bollobás, Riordan, Spencer and Tusnády [22], 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. The literature on this field has a wide range and is summarized e.g. in [20] or in [37]. A completely different approach than preferential attachment was initiated by Barabási, Ravasz, and Vicsek [18] 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



(a) G_1 and G_2 with loops



(b) G_3



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

Figure 3: $G_1, G_2, G_3, \Lambda_1, \Lambda_2, \Lambda_3$ for the "cherry" example. 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.

clustering behavior of real networks, Ravasz and Barabási [54] developed the original model in [18] 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 [61], 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.

In the second chapter of the thesis we generalize both of the models of [18] and [54]. Starting from an arbitrary initial bipartite graph G on N vertices, we construct a hierarchical sequence of deterministic graphs G_n. Namely, $V(G_n)$, the set of vertices of G_n is $\{0, 1, \dots, N-1\}^n$. To construct G_n from G_{n-1}, we take N identical copies of G_{n-1}, each of them identified with a vertex of G. Then we connect these components in a complicated way based on the coding of vertices and the postfix the codes have. In this way, G_n contains N^{n-1} copies of G_1, which are connected in a hierarchical manner, see Fig. 3(a), 3(b) for examples.

The main advantage of our generalization is that our construction provides easily analyzable unbounded average degree examples: namely, the extreme-end exponent γ 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, see Fig.4. Further, we explicitly calculate the diameter and the average shortest path length between two uniformly chosen vertices and show that they scale as the logarithm of the size of the graph.

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 [54], we add some additional edges to G_1 to obtain the (no longer bipartite) graph $\widehat{G}_1$. Then we build up the graph sequence $\widehat{G}_n$ as follows: $\widehat{G}_n$ consist of N^{n-1} copies of $\widehat{G}_1$, which copies are connected to each other in the same way as they were in G_n. So, $\widehat{G}_n$ and G_n have the same vertex set and their edges only differ at the lowest hierarchical level, that is, within the N^{n-1} copies of G_1 and $\widehat{G}_1$, see Fig. 5(b) and 5. We give a rigorous proof of the fact that local clustering coefficient of a node with degree k is of order $1/k$ in $\widehat{G}_n$, thus the average is uniformly bounded and bounded away from zero.

The embedding of the adjacency matrix of the graph sequence G_n into the unit square is carried out as follows: A vertex $\underline{x} = (x_1 \dots x_n)$ is identified

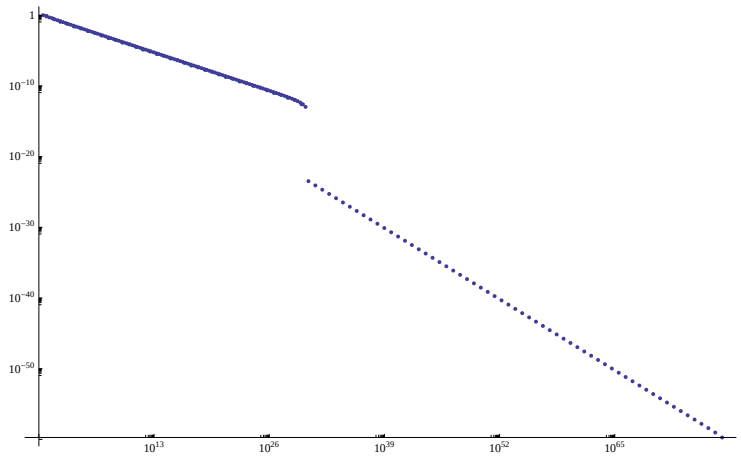
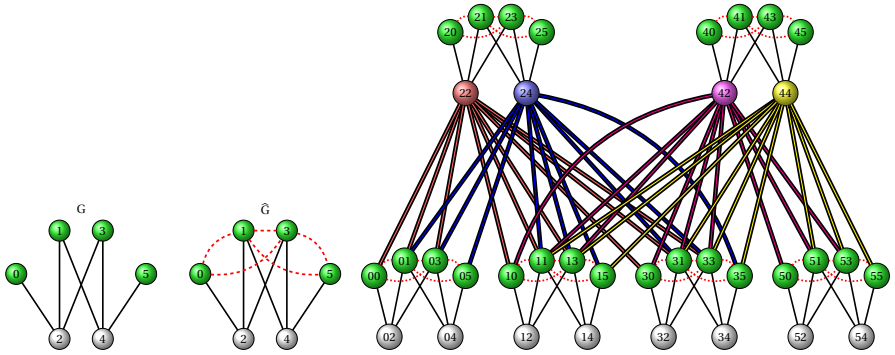


Figure 4: The log-log-plot of the tail distribution of the graph G_{100} for a bipartite base graph with $d_1 = 6, d_2 = 2, E = 12$. We can see the discontinuity of the slope of the curve where the type-2 postfix vertices vanish.

with the corresponding N-adic interval $I_{\underline{x}}$. Λ_n is the union of those $N^{-n} \times N^{-n}$ squares $I_{\underline{x}} \times I_{\underline{y}}$ for which the vertices $\underline{x}, \underline{y}$ are connected by an edge in G_n. So, Λ_n is the most straightforward embedding of the adjacency matrix of G_n into the unit square. Λ_n turns out to be a nested sequence of compact sets, which can be considered as the n-th approximation of a graph-directed self-similar fractal Λ on the plane, see Fig. 3(c).

We prove that the limit Λ can be considered as the attractor of a not irreducible graph-directed self-similar iterated function system, with the directed graph $\mathcal{G}$ similar to the one on Fig. 6. Heuristically speaking, the n-th adjacency matrix Λ_n can be written as the union of iterated maps formed by the maps $f_i, i = 1 \dots |E(G)|$ 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 Λ_n can be determined as the paths of length n $i_1 i_2 \dots i_n$ in the directed graph $\mathcal{G}$ as on the schematic picture on Fig. 6. We discuss connections between the graph theoretical properties of G_n and properties of the limiting fractal Λ. In particular, we express the power law exponent of the degree distribution with the ratio of the Hausdorff dimensions of some slices of Λ.

Furthermore, using Λ we generate a random graph sequence G_n^{r} in a way which was inspired by the W-random graphs introduced by Lovász and Szegedy [47], see also Diaconis, Janson [27], 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 [23] considered inhomogeneous random graphs generated by a kernel. Our model is not



(a) We obtain $\widehat{G}$ by adding the dashed (red) edges to G.

(b) $\widehat{G}_2$: We obtain the edges of $\widehat{G}_2$ from G_2 by adding the same extra edges as those added to G in $\widehat{G}$ everywhere at the lowest hierarchical level

Figure 5: Clustering extended "fan".

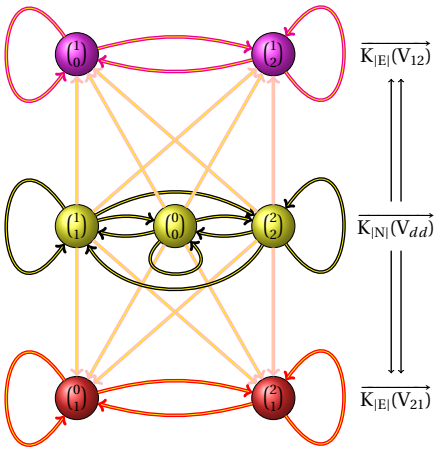


Figure 6: The graph $\mathcal{G}$ for the "cherry". Each edge of this graph corresponds to a homothetic map of the unit square into itself. Λ_n, the adjacency matrix of G_n, is the n-th approximation of the graph-directed fractal Λ, and can be obtained as the union of all iterated maps corresponding to paths of length n in this picture.

covered by their construction, since Λ is a fractal set of zero two dimensional Lebesgue measure. We remark that the fractal limit Λ of our embedded

adjacency matrices of G_n or G_n^{r} is not stable under 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 $\{1, \dots N\}$ gives different fractal limit Λ-s with the same Hausdorff dimension.

3 Fluctuation bounds in a class of deposition processes

The third chapter of the thesis studies fluctuations in deposition processes of the following type. An integer-valued height function

$$\underline{h}(t) = \{h_i(t)\}_{i \in \mathbb{Z}}$$

evolves via random deposition and removal of individual bricks of unit length and height. The Poisson rates of deposition and removal at point i are allowed to depend on the neighboring increments $h_{i-1} - h_i$ and $h_i - h_{i+1}$. Assumptions are made on these rates to guarantee stochastic monotonicity (attractivity) and the existence of a family of product-form stationary distributions μ^ϱ for the increments $\{h_{i-1} - h_i : i \in \mathbb{Z}\}$. The family of invariant measures is indexed by the average slope $\varrho = \mathbf{E}^\varrho(h_{i-1} - h_i)$. The flux function $\mathcal{H}(\varrho) = t^{-1}\mathbf{E}^\varrho(h_i(t) - h_i(0))$ gives the average velocity of the height as a function of the slope ϱ. In this chapter we consider *asymmetric* systems, for which $\mathcal{H}''(\varrho) < 0$ holds additionally at least in a neighborhood of a particular density value ϱ. Asymmetry here always mean spatial asymmetry, i.e. models in which the jump rates to the right differ from those to the left.

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 shows a configuration and a possible step with both walls and particles. It is in the particle guise that many of these processes appear in the literature: simple exclusion processes, zero range processes and misanthrope processes are examples included in the class studied in this chapter. In the particle picture the parameter ϱ that indexes invariant distributions is the mean particle density per site. Height increment $h_i(t) - h_i(0)$ is the cumulative net particle current across the edge $(i,\, i+1)$ during time $(0,\, t]$.

Fix ϱ and consider $\underline{h}(t)$ with stationary increments at average slope ϱ, normalized so that $h_0(0) = 0$. Interesting fluctuations can be found by observing the height $h_{\lfloor V^\varrho t \rfloor}(t)$ in the characteristic direction $V^\varrho := \mathcal{H}'(\varrho)$. (Characteristics is a line $X(T)$ where the density $\varrho(T, X(T))$ is constant. The *characteristic speed* V^ϱ is the velocity with which small perturbations of the solution of the PDE obtained by hydrodynamic limit propagate, i.e.

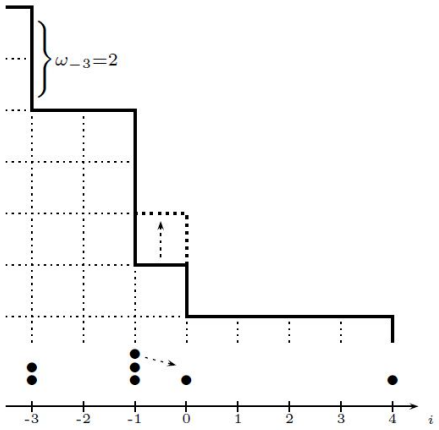


Figure 7: The wall and the particles with a possible step

the slope of constant density lines.) We show 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 [30, 41, 55] and the random average process [11, 33].

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 [5] for the last-passage version of the Hammersley process and in [38] for the last-passage model associated with TASEP. Current fluctuations for stationary TASEP were analyzed in [34]. (b) Probabilistic approaches exist to prove fluctuation bounds of the correct order. The sem-

inal work [24] was on the last-passage version of the Hammersley process, and then the approach was adapted to the last-passage model associated with TASEP [7]. The next step was the development of a proof that works for particle systems: the asymmetric simple exclusion process (ASEP) was treated in [14] and the totally asymmetric zero range process with constant jump rate in [8]. The ASEP work [14] was the first to prove $t^{1/3}$ order of fluctuations for a process where particle motion is not restricted to totally asymmetric.

The chapter is based on two papers, both of them joint with Márton Balázs and Timo Seppäläinen. The first one is [10], which takes a further step toward universality of the $t^{1/3}$ order for fluctuations in the case $\mathcal{H}''(\varrho) \neq 0$. In [10] 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 [9] we show that the strategy works for a process obeying convex flux function. In its present form the argument rests on a nontrivial hypothesis that involves control of second class particles. This control of second class particles that we require is a microscopic counterpart of the macroscopic effect that convexity or concavity of $\mathcal{H}$ has on characteristics. Throughout the first part of the chapter we consider the concave case $\mathcal{H}''(\varrho) < 0$, hence we name the property *microscopic concavity*, then in the second part 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 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 [14]) 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 [13]. This proof is somewhat simpler than the earlier one given in [14].

(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 [8], as the first generalization of the proof in [14] for processes with more than one allowed particle per site. The proof given here is simpler than the one in [8]. 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.

(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 so the result was shown in [9] and in the second part of the chapter in the thesis.

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 [14]. 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 on the label of the second class particle is missing.

The chapter has three parts. First we define the general family of processes under consideration, describe the microscopic concavity property and other assumptions used, and state the general results. 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. 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). Then, we give 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 [13].) Then we prove that our microscopic concavity criterion works for class of totally asymmetric zero range processes with concave jump rates that increase with exponentially decaying slope. Last, we show that the microscopic convexity counterpart of the assumptions required by the general result are satisfied by the exponential bricklayers process.

3.1 A family of deposition processes

The family of processes we consider is the one described in [12]. We start with the interface growth picture, but we end up using the height and particle languages interchangeably. For extended-integer-valued boundaries $-\infty \leq$

$\omega^{\min} \le 0$ and $1 \le \omega^{\max} \le \infty$ define the single-site state space

$$I := \left\{ z \in \mathbb{Z} \,:\, \omega^{\min} - 1 < z < \omega^{\max} + 1 \right\}$$

and the increment configuration space

$$\Omega := \{\underline{\omega} = (\omega_i)_{i \in \mathbb{Z}} \;:\; \omega_i \in I\} = I^{\mathbb{Z}}.$$

For each pair of neighboring sites i and $i+1$ of $\mathbb{Z}$ imagine a column of bricks over the interval $(i,\, i+1)$. The height h_i of this column is integer-valued. The components of a configuration $\underline{\omega} \in \Omega$ are the negative discrete gradients of the heights: $\omega_i = h_{i-1} - h_i \in I$.

The evolution is described by jump processes whose rates p and q are nonnegative functions on $I \times I$. Two types of moves are possible. A brick can be deposited:

$$\left.\begin{aligned} (\omega_i,\, \omega_{i+1}) &\longrightarrow (\omega_i - 1,\, \omega_{i+1} + 1) \\ h_i &\longrightarrow h_i + 1 \end{aligned}\right\} \text{ with rate } p(\omega_i,\, \omega_{i+1}), \tag{3.1}$$

or removed:

$$\left.\begin{aligned} (\omega_i,\, \omega_{i+1}) &\longrightarrow (\omega_i + 1,\, \omega_{i+1} - 1) \\ h_i &\longrightarrow h_i - 1 \end{aligned}\right\} \text{ with rate } q(\omega_i,\, \omega_{i+1}). \tag{3.2}$$

Conditionally on the present state, these moves happen independently at all sites i. We can summarize this information in the formal infinitesimal generator L of the process $\underline{\omega}(\cdot)$:

$$\begin{aligned} (L\varphi)(\underline{\omega}) = &\sum_{i \in \mathbb{Z}} p(\omega_i,\, \omega_{i+1}) \cdot [\varphi(\dots,\, \omega_i - 1,\, \omega_{i+1} + 1,\, \dots) - \varphi(\underline{\omega})] \\ &+ \sum_{i \in \mathbb{Z}} q(\omega_i,\, \omega_{i+1}) \cdot [\varphi(\dots,\, \omega_i + 1,\, \omega_{i+1} - 1,\, \dots) - \varphi(\underline{\omega})]\,. \end{aligned} \tag{3.3}$$

L acts on bounded cylinder functions $\varphi : \Omega \to \mathbb{R}$ (this means that φ depends only on finitely many ω_i-values).

Thus we have a Markov process $\{\underline{\omega}(t) : t \in \mathbb{R}_{\ge 0}\}$ of an evolving increment configuration and a Markov process $\{\underline{h}(t) : t \in \mathbb{R}_{\ge 0}\}$ of an evolving height configuration. The initial increments $\underline{\omega}(0)$ specify the initial height $\underline{h}(0)$ up to a vertical translation. We shall always normalize the height process so that $h_0(0) = 0$.

In the particle picture the variable $\omega_i(t)$ represents the number of particles at site i at time t. Step (3.1) represents a rightward jump of a particle over the edge $(i,\, i+1)$, while step (3.2) represents a leftward jump. (If negative ω-values are permitted, one needs to consider particles and antiparticles,

with antiparticles jumping in the opposite direction.) It will be useful to see that

$$h_i(t) = h_i(t) - h_0(0) = \text{the net number of particles that have passed, from left to right, the straight-line space-time path that connects } (1/2,\, 0) \text{ to } (i + 1/2,\, t). \tag{3.4}$$

We impose four assumptions on the rates, here we only describe them heuristically.

- The rates $p,\, q : I \times I \to \mathbb{R}_{\geq 0}$ must satisfy

$$p(\omega^{\min},\, \cdot\,) \equiv p(\,\cdot\,,\, \omega^{\max}) \equiv q(\omega^{\max},\, \cdot\,) \equiv q(\,\cdot\,,\, \omega^{\min}) \equiv 0 \tag{3.5}$$

  whenever either $\omega^{\min}$ or $\omega^{\max}$ is finite. Either both p and q are strictly positive in all other cases, or one of them is identically zero. The process is called *totally asymmetric* if either $q \equiv 0$ or $p \equiv 0$.

- The dynamics has a smoothing effect: the increments of the rates are monotonous such that the more particles on a site, the faster they jump out. In the height language, 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*.

- Two technical assumptions guarantee the existence of translation-invariant product-form stationary measures. (Similar assumptions were employed by Cocozza-Thivent [26].)

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 [46], and articles [4], [15] and [45].) 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 call microscopic concavity. To avoid extraneous technical issues we make some blanket assumptions on the rates p and q to be considered.

The reader will see that our proofs 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 [12] play a key role in this chapter. These identities hold for all processes in the family under study. The proofs given in [12] 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 ([13] 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 [12].

3.2 Basic coupling

In *basic coupling* the joint evolution of n processes $\underline{\omega}^m(\cdot),\ m = 1, \dots, n$, is defined in such a manner that the processes "jump together as much as possible." The joint rates are determined as follows, given the current configurations $\underline{\omega}^1$, $\underline{\omega}^2$, ..., $\underline{\omega}^n \in \widetilde{\Omega}$. Consider a step of type (3.1) over the edge $(i,\, i+1)$. Let $m \mapsto \ell(m)$ be a permutation that orders the rates of the individual processes for this move:

$$r(m) \equiv p(\omega_i^{\ell(m)},\, \omega_{i+1}^{\ell(m)}) \le p(\omega_i^{\ell(m+1)},\, \omega_{i+1}^{\ell(m+1)}) \equiv r(m+1), \quad 1 \le m < n.$$

Set also the dummy value $r(0) = 0$. Now the rule is that independently for each $m = 1, \dots, n$, at rate $r(m) - r(m-1)$, precisely processes $\underline{\omega}^{\ell(m)}$, $\underline{\omega}^{\ell(m+1)}$, ..., $\underline{\omega}^{\ell(n)}$ execute the move (3.1), and the processes $\underline{\omega}^{\ell(1)}$, $\underline{\omega}^{\ell(2)}$, ..., $\underline{\omega}^{\ell(m-1)}$ do not. The combined effect of these joint rates creates the correct marginal rates, that is, process $\underline{\omega}^{\ell(m)}$ executes this move with rate $r(m)$.

Due to the second assumption on the rate function, a jump of $\underline{\omega}^a$ without $\underline{\omega}^b$ can only occur if $p(\omega_i^b,\, \omega_{i+1}^b) < p(\omega_i^a,\, \omega_{i+1}^a)$ which implies $\omega_i^a > \omega_i^b$ or $\omega_{i+1}^a < \omega_{i+1}^b$. The result of this step (3.1) then cannot increase the number of discrepancies between the two processes, hence the name *attractivity*. Further, the monotonous increments and the coupling implies that a sitewise ordering $\omega_i^a \le \omega_i^b \quad \forall i \in \mathbb{Z}$ is preserved by the basic coupling.

The differences between two processes are called *second class particles*. Their number is nonincreasing. In particular, if $\omega_i^a \ge \omega_i^b$ for each $i \in \mathbb{Z}$, then the second class particles are conserved. In view of (3.4), in this case the net number of second class particles that pass from left to right across the straight-line space-time path from $(1/2,\, 0)$ to $(i + 1/2,\, t)$ equals the growth difference

$$\big(h_i^a(t) - h_0^a(0)\big) - \big(h_i^b(t) - h_0^b(0)\big) = h_i^a(t) - h_i^b(t) \tag{3.6}$$

between the two processes $\underline{\omega}^a(\cdot)$ and $\underline{\omega}^b(\cdot)$.

A special case that is of key importance to us is the situation where only one second class particle is present between two processes $\underline{\omega}^-(t)$ and $\underline{\omega}(t)$, we denote its position by $Q(t)$.

3.3 Results

Very briefly, the proofs in the chapter sections work if one proves the following properties of a model, which we call *microscopic concavity* (see the exact formulation in the thesis):

1. a strict domination between the second class particle $Q^\omega(t)$ of a denser system $\underline{\omega}(t)$ and one Q^η of a sparser system $\underline{\eta}(t)$,

2. a non-strict, but tight, domination between the single second class particle $Q(t)$ 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 $\mathcal{H}$,

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 the chapter are worked out for the concave setting, but everything works word-for-word in the convex case.

Theorem 3.1. *Let the microscopic concavity assumptions hold for density ϱ with the tail bound 4. Let the processes $(\underline{\omega}^-(t), \underline{\omega}(t))$ evolve in basic coupling with initial distribution stationary on each site except at 0 (defined precisely in the thesis), and let $Q(t)$ be the position of the single second class particle between $\underline{\omega}^-(t)$ and $\underline{\omega}(t)$. Then there is a constant $C_1 = C_1(\varrho) \in (0, \infty)$ such that for all $1 \le m < 3$,*

$$\frac{1}{C_1} < \liminf_{t\to\infty} \frac{\mathbf{E}|Q(t) - V^\varrho t|^m}{t^{2m/3}} \le \limsup_{t\to\infty} \frac{\mathbf{E}|Q(t) - V^\varrho t|^m}{t^{2m/3}} < \frac{C_1}{3-m}. \quad (3.7)$$

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 (Current variance)**.** *Under the microscopic concavity assumptions, there is a constant $C_1 = C_1(\varrho) > 0$, such that*

$$\frac{1}{C_1} < \liminf_{t\to\infty} \frac{\boldsymbol{Var}^\varrho(h_{\lfloor V^\varrho t\rfloor}(t))}{t^{2/3}} \le \limsup_{t\to\infty} \frac{\boldsymbol{Var}^\varrho(h_{\lfloor V^\varrho t\rfloor}(t))}{t^{2/3}} < C_1.$$

This follows from $\mathbf{Var}^\varrho(h_i(t)) = \mathbf{Var}^\varrho(\omega_0) \cdot \mathbf{E}|Q(t) - i|$ with the choice $m = 1,\ i = \lfloor V^\varrho t \rfloor$.

Corollary 3.3 (Law of Large Numbers for the second class particle)**.** *Under the microscopic concavity assumptions, the Weak Law of Large Numbers holds in a density-ϱ stationary process:*

$$\frac{Q(t)}{t} \xrightarrow{d} V^\varrho. \quad (3.8)$$

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 (3.8) *(Strong Law of Large Numbers).*

The Weak Law is a simple consequence of Theorem 3.1. The Strong Law has to be proved separately.

Corollary 3.4 (Dependence of current on the initial configuration)**.** *Under the microscopic concavity assumptions, for any $V \in \mathbb{R}$ and $\alpha > 1/3$ the following limit holds in the L^2 sense for a density-ϱ stationary process:*

$$\lim_{t\to\infty} \frac{h_{\lfloor Vt\rfloor}(t) - h_{\lfloor Vt\rfloor-\lfloor V^\varrho t\rfloor}(0) - t(\mathcal{H}(\varrho) - \varrho\mathcal{H}'(\varrho))}{t^\alpha} = 0. \tag{3.9}$$

The limit (3.9) 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^ϱ. 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 (3.9) follows by translating $h_{\lfloor Vt\rfloor}(t) - h_{\lfloor Vt\rfloor-\lfloor V^\varrho t\rfloor}(0)$ to $h_{\lfloor V^\varrho t\rfloor}(t) - h_0(0) = h_{\lfloor V^\varrho t\rfloor}(t)$ and by applying Corollary 3.2. From (3.9) 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 $\tilde{X}$ stands for centering the random variable X.

Corollary 3.5 (Central Limit Theorem for the current)**.** *Under the microscopic concavity assumptions, for any $V \in \mathbb{R}$ in a density-ϱ stationary process*

$$\lim_{t\to\infty} \frac{\boldsymbol{Var}^\varrho(h_{\lfloor Vt\rfloor}(t))}{t} = \boldsymbol{Var}^\varrho(\omega)\cdot|V^\varrho - V| =: D, \tag{3.10}$$

and the Central Limit Theorem also holds: for the centered and normalized height we have $\widetilde{h}_{\lfloor Vt\rfloor}(t)/\sqrt{t\cdot D}$ converges in distribution to a standard normal.

For ASEP, the CLT, the limiting variance (3.10) and the appearance of initial fluctuations on the diffusive scale were proved by P. A. Ferrari and L. R. G. Fontes [32]. For convex rate zero range and bricklayers processes, Corollary 3.5 was proved by M. Balázs [6].

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 for the exponential bricklayer process.

3.4 Three examples that satisfy microscopic concavity

Presently we have verified all the hypotheses of Theorem 3.1 for three classes of processes. **The asymmetric simple exclusion process**

The asymmetric simple exclusion process (ASEP) was introduced by F. Spitzer [58], is defined by $\omega^{\min} = 0,\ \omega^{\max} = 1$, The rate functions are given by

$$p(y,\, z) = p \cdot \mathbf{1}\{y = 1,\, z = 0\} \quad \text{and} \quad q(y,\, z) = q \cdot \mathbf{1}\{y = 0,\, z = 1\}.$$

Here $\omega_i \in \{0,\, 1\}$ is the occupation number for site i, $p(\omega_i,\, \omega_{i+1})$ is the rate for a particle to jump from site i to $i + 1$, and $q(\omega_i,\, \omega_{i+1})$ is the rate for a particle to jump from site $i + 1$ to i. These rates have values p and q, respectively, whenever there is a particle to perform the above jumps, and there is no particle on the terminal site of the jumps. To be specific let us take $p > q$ so that on average particles prefer to drift to the right. The invariant measure μ^ϱ is the Bernoulli distribution with parameter $0 \le \varrho \le 1$, while $\widehat{\mu}^\varrho$ is concentrated on zero for any ϱ. The hydrodynamic flux is strictly concave: $\mathcal{H}(\varrho) = (p - q)\varrho(1 - \varrho)$. The detailed construction of the processes $y(t)$ and $z(t)$ needed for Assumption **??** can be found in [13].

Balázs and Seppäläinen gave an earlier proof of Theorem 3.1 for ASEP in [14]. The present general proof evolved from that earlier one.

Totally asymmetric zero range process with jump rates that increase with exponentially decaying slope

In a totally asymmetric zero range process (TAZRP), $\omega^{\min} = 0,\ \omega^{\max} = \infty$, 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 $f : \mathbb{Z}_{\ge 0} \to \mathbb{R}_{\ge 0}$ is nondecreasing, $f(0) = 0$, and $f(z) > 0$ for $z > 0$. Assume further that f is concave. Again, ω_i represents the number of particles at site i. Depending on this number, a particle jumps from i to the right with rate $pf(\omega_i)$, and to the left with rate $qf(\omega_i)$. These rates trivially satisfy the needed conditions.

We show that, one aspect of microscopic concavity, namely the ordering of second class particles, can be achieved for any TAZRP with a nondecreasing concave jump rate, we only use monotonicity and concavity of the rates f. Thus for concave TAZRP only the tail control on the single second class particle remains to be provided. For this part we currently need a stronger hypothesis, detailed in the next assumption.

Assumption 3.6. *Let $p = 1 - q = 1$. Assume the jump rate function f of a totally asymmetric zero range process has these properties:*
1. $f(0) = 0 < f(1)$ and f is nondecreasing: $f(z + 1) \ge f(z)$,
2. f is concave with an exponentially decreasing slope: there is an $0 < r < 1$

such that for each $z \geq 1$ *such that* $f(z) - f(z-1) > 0$,

$$\frac{f(z+1) - f(z)}{f(z) - f(z-1)} \leq r. \tag{3.11}$$

The case where f *becomes constant above some* z_0 *is included.*

Theorem 3.7. *Under Assumption 3.6, a stationary totally asymmetric zero range process satisfies the conclusions of Theorem 3.1, and the conclusions of Corollaries 3.2, 3.3, 3.4 and 3.5.*

A class of examples of rates that satisfy Assumption 3.6 are

$$f(z) = 1 - \exp(-\beta z^{\vartheta}), \quad \beta > 0,\ \vartheta \geq 1.$$

Another example is the most basic, constant rate TAZRP with $f(z) = \mathbf{1}\{z > 0\}$. For this last case a proof has already been given in [8].

Totally Asymmetric Exponential Bricklayers process.

Let $f : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ be non-decreasing and satisfy $f(z) \cdot f(1-z) = 1$ for all $z \in \mathbb{Z}$. The values of f for positive z's thus determine the values for non-positive z's. The jump rates of the process are given by

$$p(y, z) = pf(y) + pf(-z) \qquad \text{and} \qquad q(y, z) = qf(-y) + qf(z).$$

The following picture motivates the name bricklayers process. At each site i stands a bricklayer who lays a brick on the column to his left at rate $pf(-\omega_i)$ and on the column to his right at rate $pf(\omega_i)$. Each bricklayer also removes a brick from his left at rate $qf(\omega_i)$ and from his right at rate $qf(-\omega_i)$. The needed conditions hold for the rates. The totally asymmetric *exponential* bricklayers process (TAEBLP) is obtained by taking

$$f(z) = \mathrm{e}^{\beta(z-1/2)}. \tag{3.12}$$

The increments of the rate function are not bounded, hence the jump rate of the second class particle cannot be dominated by the jump rates of a biased random walk. Thus, the hardest is to verify the tail bound on the single second class particle. The main tool we use to show this bound is Theorem 1 from [16].