Asymmetric Particle Systems and Last-Passage Percolation in One and Two Dimensions

by

Philipp Schmidt

Keble College

A dissertation submitted for the degree of

Doctor of Philosophy

Department of Statistics, 1 South Parks Road, OX1 3TG, September 2011

Asymmetric Particle Systems and Last-Passage Percolation in One and Two Dimensions

Philipp Schmidt, Keble College, University of Oxford

A dissertation submitted for the degree of Doctor of Philosophy

Trinity Term, 2011

Abstract

This thesis studies three models: Multi-type TASEP in discrete time, long-range last-passage percolation on the line and convoy formation in a travelling servers model. All three models are relatively easy to state but they show a very rich and interesting behaviour.

The TASEP is a basic model for a one-dimensional interacting particle system with non-reversible dynamics. We study some aspects of the TASEP in discrete time and compare the results to recently obtained results for the TASEP in continuous time. In particular we focus on stationary distributions for multi-type models, speeds of second-class particles, collision probabilities and the speed process. We consider various natural update rules.

The second model we study is directed last-passage percolation on the random graph $G = (V, E)$ where $V = \mathbb{Z}$ and each edge (i, j), for $i < j \in \mathbb{Z}$, is present in E independently with some probability $p \in (0, 1]$. To every $(i, j) \in E$ we attach i.i.d. random weights $v_{i,j} > 0$. We are interested in the behaviour of $w_{0,n}$, which is the maximum weight of all directed paths from 0 to n, as $n \to \infty$. We see two very different types of behaviour, depending on whether $\mathbb{E}\left[v_{i,j}^2\right] < \infty$ or $\mathbb{E}\left[v_{i,j}^2\right] = \infty$. In the case where $\mathbb{E}\left[v_{i,j}^2\right] < \infty$ we show that the process has a certain regenerative structure, and prove a strong law of large numbers and, under an extra assumption, a functional central limit theorem. In the situation where $\mathbb{E}\left[v_{i,j}^2\right] = \infty$ we obtain scaling laws and asymptotic distributions expressed in terms of a continuous last-passage percolation model on $[0, 1]$.

In the last model customers arrive on the non-negative half-line as a Poisson process of rate $\lambda \in (0, \infty)$ and n servers start at the origin at time 0. After completing some initial holding times, each server jumps to the first free customer it sees to its right and serves this customer. All service times are i.i.d. exponentials with parameter ν. After completion of a service the customer leaves the system and the server jumps to the next customer to the right, ignoring customers that are currently being served. We study the formation of convoys, groups of servers that travel together at the same speed, and the asymptotic behaviour of $X_t^{(j)}$, see position of server j at time t.

Preface

I would like to thank my supervisor James Martin for his guidance and support during my doctoral studies. His enthusiasm and imaginativeness were inspiring and made my studies very enjoyable. I am grateful for the amount of time he spent introducing me to new ideas and open problems and his patience with my work.

For Chapters 3 and 4 I want to thank Sergey Foss not only for telling me about the models and open problems but also for his help with my work and the invitations to Cambridge, Oberwolfach and Edinburgh.

I would also like to thank the Department of Statistics for their financial support and for providing such an excellent environment for my doctoral studies and the DAAD whose scholarship made this DPhil possible. My thanks also go to Keble College where I met so many wonderful people and made many good friends.

This dissertation is not substantially the same as any that I have submitted for any degree or diploma or other qualification at any university.

Chapter 2 of this thesis is joint work with James Martin. Chapter 3 and Chapter 4 are joint work with James Martin and Sergey Foss.

Philipp Schmidt

Keble College, Oxford

June 23, 2012

Contents

Chapter 1

Introduction

Particle systems have been used for many years to describe a huge variety of physical, biological and chemical processes. The study of hydrodynamics and the modelling of gases as particle systems, where the particles represent individual molecules, goes back to Daniel Bernoulli's work Hydrodynamica [3] in the 18th century. Whereas early applications focused on statistical mechanics and thermodynamics, today there is a wide variety of models and applications. These applications range from competition models in biology, electronic networks, movement of molecules and traffic models to computer science, queueing networks and many more.

The work presented here focuses on three very different but related models: The TASEP, or *totally asymmetric simple exclusion process*, in discrete time, long-range *last-passage percolation* on the line and *convoy formation in a travelling servers model.* Although not itself a particle system like the first and the last model, the last-passage percolation model has a strong connection to the TASEP that will be examined in detail in Section 2.2.2. We will study a two-dimensional last-passage percolation model in connection with the TASEP in Chapter 2, while Chapter 3 deals with a more general model in one dimension. Apart from the one-dimensionality, a common feature of all three models is some form of asymmetry. In the TASEP and the travelling servers model the particles, and the servers respectively, can only move to the right which makes the models asymmetric. In our last-passage percolation model, the underlying graph will be a directed graph on the vertices of $\mathbb{Z}$ and this can also be seen as a type of asymmetry. All three models are also relatively simple to state but despite this simplicity they show a very rich, interesting and sometimes surprising behaviour. We will now give a brief introduction to the three models under consideration and their connections. The models will then be more formally introduced in the corresponding chapters.

Our focus in Chapter 2 will be the TASEP in discrete time and we will analyse its connection to last-passage percolation in two dimensions. The core of this chapter is the article [40], to appear in ALEA, which is joint work with James Martin. The TASEP is a basic model for a one-dimensional interacting particle system with non-reversible dynamics. The TASEP in continuous time was introduced by Spitzer in 1970 ([48]) and can be described as follows. It is a Markov process $(\eta_t)_{t\geq 0}$ on the state space $E = \{0,1\}^{\mathbb{Z}}$ where for $x \in \mathbb{Z}$ we say that site x is occupied with a particle at time t iff $\eta_t(x) = 1$. Otherwise we say that site x is empty at time t. Starting from some initial configuration $\eta_0 \in E$, *updates* occur at each site as a Poisson process of rate 1, independently; when an update occurs at site x, if there is a particle at site x and a hole to its right at site $x+1$, the particle jumps from site x to site $x+1$. If site x is empty, or if site $x+1$ is already occupied, the update has no effect. This is the *exclusion* property of the process. Since the particles can only move to the right and only one step at a time the process is *totally asymmetric* and *simple*.

In the model in discrete time, updates occur with some probability $\beta \in (0,1)$ at each site at each time-step. Since updates occur simultaneously, we now have to choose an order in which to update the sites. We will consider sequential updates (from right to left or from left to right) and sublattice-parallel updates (even sites first then odd sites). This makes the TASEP in discrete time substantially different from the model in continuous time. Here we derive results for the TASEP in discrete time that correspond to recently obtained results for the continuous-time model. These include stationary distributions for multi-type systems (e.g. [17, 18]), laws of large numbers for the path of a second class particle and their connection to competition interfaces in competition growth models (e.g. [19, 16]), and the *TASEP speed process* recently studied by Amir, Angel and Valkó ([1]).

In Chapter 3 we study a model of directed last-passage percolation on the integer line $\mathbb{Z}$. The content of this chapter is joint work with Sergey Foss and James Martin and the results have been submitted for publication to the Annals of Applied Probability, see [21]. A simple long-range last-passage percolation model can be defined as follows. Consider the random directed graph $G = (\mathbb{Z}, E)$, where every directed edge (i, j) from vertex i to vertex $j > i$ is present independently with probability $p \in (0, 1]$. Random structures of this kind have been used to study community food webs or task graphs for parallel processing in computer science. In the first case a link between i and j means that species j preys upon species i. The computational interpretation would be that task i must be completed before task j can start. Such models have been studied in [12], [24], [31] and [42], for example.

Here we consider a model in which weights $v_{i,j}$ are attached to all edges $(i, j) \in E$. In [14] the authors study the same model with constant weights $v_{i,j} \equiv 1$. We are interested in the asymptotic behaviour of the random variable $w_{0,n}$ that is defined as follows. For any increasing path $\pi = ((i_0, i_1), (i_1, i_2), \ldots, (i_{l-1}, i_l))$ from $i = i_0$ to $j = i_l$ (for some $l \geq 0$) the *weight* of the path is the sum of the edge weights $\sum_{k=1}^{l} v_{i_{k-1},i_k}$. We define the weight $w_{i,j}$ to be the maximal weight of a path from i to j. A maximizing path between two points is called a *geodesic*. If we let $\Pi_{i,j}$ be the set of all paths from i to j, then

$$w_{i,j} = \max_{\pi \in \Pi_{i,j}} \sum_{e \in \pi} v_e.$$

If there is no path between i and j (which is possible if $p < 1$) then $w_{0,n}$ will be $-\infty$. We will analyse the behaviour of $w_{0,n}$ as n tends to infinity. In the parallel processing model, $v_{i,j}$ represents a delay required between the start of task i and the start of task j, and $w_{0,n}$ represents the overall constraint on the time between the start of task 0 and the start of task n. We will write v for a generic random variable whose distribution is that of $v_{i,j}$, and we write F for the distribution function of v.

A very interesting feature of the TASEP is its connection to last-passage percolation. If we start a single-type TASEP from a specific initial configuration, then this model corresponds to two-dimensional nearest neighbour last-passage percolation in the first quadrant of $\mathbb{Z}^2$. The holding times in the TASEP correspond to vertex-weights in the last-passage percolation model (exponential weights in continuous time and geometric weights in discrete time). The time it takes particle k in the TASEP to jump n times corresponds to $w_{0,(n,k)}$ which is the weight of the heaviest path from the origin to the vertex (n, k) in the last-passage percolation model. This connection will be explained in more detail in Section 2.2.2 and we will show how this connection can be exploited to transfer results about one model into results for the other model. A shape theorem for the last-passage percolation model corresponds, for example, to a hydrodynamic limit in the TASEP. Due to this connection, it is therefore natural to study last-passage percolation models in connection with the TASEP. The last-passage percolation model that we consider in Chapter 3 has, however, a few different features: it is a one-dimensional instead of a two-dimensional model; it is not a nearest neighbour model but allows long-range edges; the weights are not geometric or exponential but come from a more general class of probability distributions.

In Chapter 4 we will analyse the formation of convoys in a model with travelling servers. The results in this chapter are joint work with James Martin and Sergey Foss. Customers arrive on the non-negative real half-line as a Poisson process with parameter $\lambda \in (0, \infty)$ (Poisson rain). We let n servers start at the origin at time 0. After completing some initial holding times, each server jumps to the first free customer it sees to its right and serves this customer. All service times are i.i.d. exponentials with parameter ν. After completion of a service the customer leaves the system and the server jumps to the next customer to the right, ignoring customers that are currently being served. We are interested in the asymptotic behaviour of $X_t^{(j)}$, the position of server j at time t, and its inverse $T_x^{(j)}$, the passage time of server j to x. We assign the labels 1 to n arbitrarily to the n servers. As we will see in Section 2.3.1, queueing systems play an important role in connection with stationary distributions for the TASEP.

For $n = 1$ and a speed $v < \infty$ this model has been studied in [36], [47] and [35]. A speed $v < \infty$ means that the server does not jump to the next customer. Instead it travels at speed v until it encounters a customer. Our model corresponds to taking $v = \infty$. [35] looks at the model with more general service time distributions. The result in Chapter 4 for $n = 1$ is slightly stronger than the results obtained in [36] and [47]. However, the assumption $v = \infty$ makes the model easier to analyse. The results concerning the model with more than one server, which we derive here, are new and show an interesting behaviour of the model. We will show that the servers form convoys of strictly decreasing size. The formation of these convoys happens randomly and all possible sequences of convoys occur with strictly positive probability. The speed of a server corresponds to the size of its convoy. There exists a vast amount of literature about related models, for example in the context of polling systems, which have applications in telecommunication systems, machine repair, manufacturing or transportation. In [27], [37] and [22, 23] the authors study polling models with fixed queueing stations. A single server visits the queueing stations according to some algorithm (for example a greedy algorithm). These models have been studied for various algorithms and service time distributions. The difference to our model is that customers can only arrive at queueing stations and not as a process in continuous space. In [34] and [32, 11] models more closely related to our model are studied: customers arrive according to a Poisson (or more general arrival) process on the line or a circle. The server either follows some algorithm or a Brownian motion with drift and stops to serve customers. Although the set-up is similar to our model, we look at a very different type of results. Whereas the literature mainly focuses on the question of stationarity for these models, our model is transient and our main results concern the arrangement of the n servers in some random sequence of convoys.

Chapter 2

Multi-type TASEP in discrete time

2.1 Introduction

For the model in continuous time, that we first introduced in the introduction, there exists a vast amount of literature. For an introduction and background to the topic see Liggett's books [38] (pp. 361-417) and [39] (pp. 209-316). However, in some physical models of interest it might be more natural to use a discrete time scale. For example in traffic models we can consider the reaction time of individuals as a smallest time scale (Blythe and Evans [7], Chowdhury, Santen and Schadschneider [10] and Helbing [29]) and this suggests modelling traffic with a model in discrete time. The ASEP (*asymmetric simple exclusion process*, particles jump to the right at rate p and to the left at rate $q < p$) in discrete time was studied for example in Schütz [46], Hinrichsen [30], Rajewsky, Santen, Schadschneider and Schreckenberg [44] and Blythe and Evans [7]. However, the behaviour of the models in discrete time has not been analysed in as much depth as the model in continuous time. The papers mentioned above are mainly concerned with the model on a finite interval with open boundary conditions and just one type of particles and analyse density profiles and stationary distributions.

Looking at the model in discrete time with update parameter $\beta \in (0, 1)$, we find that the multi-type invariant distributions for the models with sequential updates are identical with those for the model in continuous time, and do not depend on the parameter β. This has the surprising consequence that various collision probabilities for different particles in a multi-type processes started out of equilibrium, of the sort considered in [16] and [1], are also independent of β and coincide with the values for a continuous-time process. These probabilities correspond to survival probabilities of clusters in the associated multi-type competition growth models. At the moment, the only argument we have for this property is indirect, using the fact that the set of invariant measures is identical for all β; we do not know of a more direct argument based on local dynamics or couplings.

By contrast, in the case of sublattice-parallel updates, the value of β plays an important role in the set of stationary distributions. All the subsequent results will also heavily depend on the value of the parameter β. We extend the queue-based construction of the multi-type stationary distributions from [17, 18] by incorporating queues whose arrival and service rates are different at even and odd times. As Hinrichsen [30] points out, models with parallel dynamics arise in traffic and reaction-diffusion models.

The chapter is organized as follows. In Section 2.2 we will give a more formal definition of the model and introduce the multi-type TASEP. The main results are described in Section 2.3, including results concerning invariant measures and hydrodynamic limits for single-type models which are required in order to state and understand the multi-type models described above. The proofs or proof sketches for the novel results are found in Section 2.4. In Section 2.5 we make some brief remarks about a related discrete-time TASEP model with "fully parallel updates".

2.2 Model

2.2.1 Models in continuous and discrete time

The TASEP in continuous time can be described by its generator L. Recall that our state space is $E = \{0,1\}^{\mathbb{Z}}$. For cylinder functions $f : E \to \mathbb{R}$ we have

$$Lf(\eta) = \sum_{x\in\mathbb{Z}} \eta(x)\left(1-\eta(x+1)\right)\left[f\left(\eta^{x,x+1}\right) - f\left(\eta\right)\right]$$

with the configuration $\eta^{x,x+1}$ defined by

$$\eta^{x,x+1}(y) = \begin{cases} \eta(y) & y \notin \{x, x+1\} \\ \eta(x+1) & y = x \\ \eta(x) & y = x+1 \end{cases}$$

Following ideas of Harris (1978) [28] we can use the following graphical construction for the TASEP. Let $\{(P^x_t)_{t\geq 0} : x \in \mathbb{Z}\}$ be a family of independent mean 1 Poisson processes on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For $x \in \mathbb{Z}$ the process P^x marks possible jumps from site x: If $P^x_t - P^x_{t-} = 1$ and $\eta_{t-}(x) = 1$ then the particle at x tries to jump one step to the right at time t. The jump is successful if the adjacent site $x+1$ was unoccupied, i.e. $\eta_{t-}(x+1) = 0$. Note that for every $t > 0$ with positive probability (e^{-t}) there was no jump in the Poisson process P^x up to time t. Since all the Poisson processes are independent there will be infinitely many sites x such that there were no jumps in P^x. These sites separate $\mathbb{Z}$ into intervals of finite length. Since no particle can have crossed the boundaries of these intervals, it is enough to be able to construct the process separately on each of these finite intervals.

We can use the same graphical construction to define the TASEP in discrete time. All we have to do is replace the family of Poisson processes with a family $\{(B^x_n)_{n\geq 0} : x \in \mathbb{Z}\}$ of independent Bernoulli processes with parameter $\beta \in (0,1)$ and decide on an update rule for the sites. As mentioned in the introduction we will mainly consider the following three update rules:

- Rule R1: Updates are processed in order from right to left.
- Rule R2: Updates are processed in order from left to right.
- Rule R3: All updates at even sites are processed before all updates at odd sites.

To highlight the difference between the three update rules we can look at the following example.

Say we are at time n in the configuration displayed in Figure 2.1, with particles at sites -1 and 0 and holes at sites 1 and 2. There are jump attempts at the sites marked with a $*$. These jump attempts or updates can be thought of as tokens which are consumed once in the order given by the update rule. The resulting configurations under the three different update rules are as shown in Figures 2.2 - 2.4.

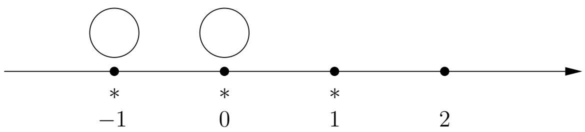

Figure 2.1: Configuration at time n and jump marks

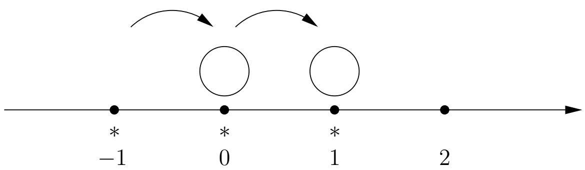

Figure 2.2: Configuration at time $n+1$ if we apply R1

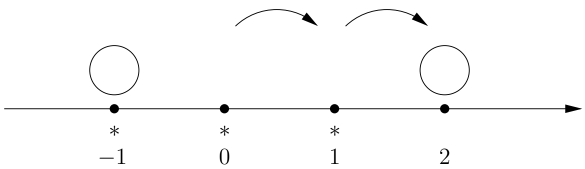

Figure 2.3: Configuration at time $n+1$ if we apply R2

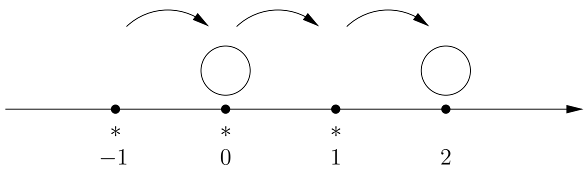

Figure 2.4: Configuration at time $n+1$ if we apply R3

Note that in R2, a single particle may jump several times at the same time-step (but jumps are only possible onto sites that were already empty at the beginning of the time-step). In R1, several neighbouring particles may jump together at the same time-step. There is a natural symmetry between systems R1 and R2 - one is transformed into the other by exchanging left and right and exchanging the roles of particle and hole. For the last example (R3) the parity of the sites is important.

In connection with the *speed process* we will also mention the model with odd/even updates (R4). Again this can be obtained from R3 by a simple transformation.

As seen above, each of these models shows a slightly different behaviour, but if we rescale time by a factor β^{-1} and let $\beta \to 0$ then they converge to the model in continuous time. In this sense the model in discrete time is more general than the model in continuous time (which in the following we will denote by R0) since we can recover the model in continuous time from the model in discrete time. In discrete time we can also consider the model with (fully) parallel updates where all sites are updated simultaneously. However, many of the methods developed for the model in continuous time that work in the models R1-R3 fail in this case. We will mention some questions connected to this model in Section 2.5.

2.2.2 Percolation representations

Both in continuous and in discrete time, one important feature of the TASEP is its connection to *last-passage percolation* and the *corner growth model*. Here we consider a special case which corresponds to a particular initial condition of the TASEP, in which, at time 0, all non-positive sites $x \leq 0$ contain a particle and all positive sites

$x > 0$ are empty. We label the particles from right to left, so that for $i \geq 1$, particle i starts at site $-i+1$ at time 0 (and always remains to the right of particle $i+1$).

For $n, k \geq 1$, let $T(n,k)$ be the time that particle k jumps to its right for the nth time. Then it is well-known that the variables $T(n,k)$ satisfy the recursions

$$T(n,k) = \max\{T(n-1,k), T(n,k-1)\} + v(n,k) \qquad n,k \geq 1 \tag{2.1}$$

with boundary conditions $T(0,k) = T(n,0) = 0$ for all n, k, where $v(n,k)$ are i.i.d. exponential random variables with mean 1. The interpretation is that before particle k can make its nth jump, both particle k must have made its $(n-1)$st jump, and particle $k-1$ must have made its nth jump. Once these two events have happened, an amount of time which is exponentially distributed with rate 1 passes before particle k makes its nth jump; this is the random variable $v(n,k)$.

The random variables $T(n,k)$ have an interpretation in terms of last-passage percolation times. For an increasing path π from $z \in \mathbb{Z}_+^2$ to $z' \in \mathbb{Z}_+^2$, i.e. a path with increments in $\{(0,1),(1,0)\}$, define the weight of π by

$$S(\pi) = \sum_{z'' \in \pi} v(z'').$$

Write $\Pi(z,z')$ for the set of all increasing paths from z to z'; then

$$R(z,z') = \max_{\pi \in \Pi(z,z')} S(\pi) \tag{2.2}$$

is the weight of the heaviest path from z to z'. Then, via the recursions (2.1), it is easy to see that $T(n,k) = R((1,1),(n,k))$. In this setting we may interpret the random variable $v(n,k)$ as a weight at the lattice point (n,k). This connection between the TASEP and last-passage percolation was the motivation for the study of the last-passage percolation model that we consider in Chapter 3.

We turn to the discrete-time case. Now let $w(n,k)$ be i.i.d. random variables whose distribution is geometric with parameter $\beta \in (0,1)$ (by which we mean that

$\mathbb{P}[w(z) = k] = (1-\beta)^k\beta$ for $k = 0, 1, 2, \dots$). We define passage-times $\tilde{T}(n,k)$ analogous to $T(n,k)$ above by the recursions

$$\tilde{T}(n,k) = \max\left\{\tilde{T}(n-1,k), \tilde{T}(n,k-1)\right\} + w(n,k) \qquad n,k \geq 1.$$

We will describe three variants on these recursions, which pertain to the different update rules R1, R2 and R3. As above, $w(n,k)$ will correspond to the delay before particle k makes its nth jump, once it is free to do so. For $i = 1, 2, 3$, let $T^{(i)}(n,k)$ be the nth jump of particle k under update rule Ri with boundary conditions $T^{(i)}(0,k) = T^{(i)}(n,0) = -1$ for all n, k.

Rule R1 (updates from right to left)

- Recursions:

$$\begin{aligned} T^{(1)}(n,k) &= \max\left\{T^{(1)}(n-1,k)+1, T^{(1)}(n,k-1)\right\} + w(n,k) \\ &= \tilde{T}(n,k) + n - 1 \end{aligned} \tag{2.3}$$

- In accordance with the updates from right to left, particles k and $k-1$ can make their nth jumps at the same time-step, but two jumps by the same particle must be separated by at least one time-step.

- This corresponds to a percolation model in which, in addition to weights $w(n,k)$ at the vertices $(n,k) \in \mathbb{Z}_+^2$, we have weights of size 1 on each horizontal edge between $(n-1,k)$ and (n,k).

Rule R2 (updates from left to right)

- Recursions:

$$\begin{aligned} T^{(2)}(n,k) &= \max\left\{T^{(2)}(n-1,k), T^{(2)}(n,k-1)+1\right\} + w(n,k) \\ &= \tilde{T}(n,k) + k - 1 \end{aligned} \tag{2.4}$$

- With updates from left to right, a particle may make several jumps at the same time-step, but at least one time-step must separate the nth jump of particles $k-1$ and k.

- In the corresponding percolation model, the weights of size 1 are now on the vertical edges of the lattice.

Rule R3 (even updates then odd updates)

- Recursions:

$$T^{(3)}(n,k) = \begin{cases} \max\left\{T^{(3)}(n-1,k)+1, T^{(3)}(n,k-1)+1\right\} + w(n,k) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad n+k \text{ even} \\ \max\left\{T^{(3)}(n-1,k), T^{(3)}(n,k-1)\right\} + w(n,k) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad n+k \text{ odd} \end{cases}$$

$$= \begin{cases} \tilde{T}(n,k) + \frac{n+k-2}{2} & n+k \text{ even} \\ \tilde{T}(n,k) + \frac{n+k-3}{2} & n+k \text{ odd} \end{cases} \tag{2.5}$$

- Now the edge weights of size 1 are added to all edges with an upper/right point (n,k) such that $n+k$ is even.

For the model in continuous time we have, for $x>0$,

$$\lim_{n\to\infty} \frac{T([xn],n)}{n} = \left(\sqrt{x}+1\right)^2 \qquad \text{a.s.} \tag{2.6}$$

This was essentially first shown in [45]. Replacing the exponential weights by geometric weights gives

$$\lim_{n\to\infty} \frac{\tilde{T}([xn],n)}{n} = \frac{(1-\beta)x + 2\sqrt{(1-\beta)x} + (1-\beta)}{\beta} \qquad \text{a.s.}; \tag{2.7}$$

see for example [43]. Using (2.3)-(2.5), this can easily be used to give similar laws of large numbers for $T^{(i)}([xn],n)$, $i=1,2,3$.

We may also view the system as a (corner) growth model. For the continuous-time case,

$$G_t = \big\{(x,y) \in \mathbb{Z}_+^2 : T(x,y) \le t\big\} \tag{2.8}$$

be the set of vertices whose passage-time is less than t. This gives a cluster in $\mathbb{Z}_+^2$ which grows over time; there is a 1-1 correspondence relating G_t to the configuration of the TASEP at time t; the length of the row at height $k \in \mathbb{Z}_+$ is the number of jumps particle k has made in the TASEP. In a similar way we can define $G^{(1)}(t)$, $G^{(2)}(t)$ and $G^{(3)}(t)$ by replacing T in (2.8) by $T^{(1)}$, $T^{(2)}$ or $T^{(3)}$ respectively.

2.2.3 Multi-type models

In the *multi-type TASEP* each particle belongs to a class $y \in \mathbb{Z}$ (or more generally $y \in \mathbb{R}$). All particles can still jump into unoccupied sites. When a particle of class k tries to jump into a site which is occupied by a particle of class j two things can happen: If $k \ge j$ the jump is suppressed and if $k < j$ then the two particles swap. This means that the lower the class of a particle the higher is its priority.

An n-type TASEP (containing n classes of particles and holes) can be regarded as a coupling of n ordered single-type TASEPs. If $\eta_0^1, \ldots, \eta_0^n$ are n TASEP configurations such that $\eta_0^1(x) \le \ldots \le \eta_0^n(x)$ for all $x \in \mathbb{Z}$, we can use the same Poisson or Bernoulli processes (this is called *basic coupling*) to get a joint realization of the TASEPs η^1, $\ldots$, η^n.

The basic coupling preserves the ordering between the processes (since the updates are processed one by one, this is true for the discrete-time models just as in the continuous-time case). Thus we can define a multi-type process ξ by

$$\xi_t(x) = n + 1 - \sum_{k=1}^{n} \eta_t^k(x).$$

We write $\xi_t = R\eta_t$. Particles of class k occur at sites x where $\xi(x) = k$. For $k > 1$, these sites represent discrepancies between the processes η^{k-1} and η^k. We may regard particles of type $n+1$ as holes. Then ξ behaves like a multi-type TASEP with n classes of particles and holes. See for example [17] for further details.

2.3 Results

We will divide this section into three subsections: The first deals with invariant measures for single- and multi-type models (Section 2.3.1), the second with hydrodynamic limits (Section 2.3.2) and the third with multi-type models out of equilibrium (Section 2.3.3).

2.3.1 Invariant measures

Proposition 2.1. *For the TASEP in continuous time as well as the discrete time TASEPs R1 and R2, the Bernoulli product measures ν_ρ with marginals $\rho \in [0,1]$ are the only translation invariant stationary ergodic measures with constant marginals. For the TASEP R3, the Bernoulli product measures μ_ρ with marginals $\rho \in [0,1]$ on even sites and marginals $\frac{\rho(1-\beta)}{1-\rho\beta}$ on odd sites are the only stationary ergodic measures with marginals that are translation invariant under even shifts.*

Remark 2.2. *Interestingly, the marginals of the invariant Bernoulli product measures for the models R1 and R2 do not depend on the model parameter β, and coincide with the invariant measures for the model in continuous time. In the model R3 however, the densities at even and odd sites differ (with a specific relation between them) and the measure depends on the parameter β.*

Proof. For references see for example Liggett [38] for R0, Blythe and Evans [7] for R1, R2 and Rajewsky, Santen, Schadschneider and Schreckenberg [44] for R3. The uniqueness statements can be proved following the approach of Mountford and Prabhakar [41]. □

We now turn to the construction of invariant measures for systems with more than one class of particles. We use the construction based on a system of queues in tandem developed in [17], and begin by recalling notation from that paper.

Given two processes α_1 and α_2, taking values in $\{0,1\}^{\mathbb{Z}}$ and representing the arrival and service processes of a discrete-time queue respectively, let $D(\alpha_1,\alpha_2)$ be the process of departures from the queue. Here the service process is the process of potential services. A service is lost and regarded as unused service if there is no customer waiting to be served at the time when the service is offered. Now define $D^{(1)}(\alpha) = \alpha$, $D^{(2)}(\alpha_1,\alpha_2) = D(\alpha_1,\alpha_2)$, and recursively $D^{(n)}(\alpha_1,\ldots,\alpha_n) = D\left(D^{(n-1)}(\alpha_1,\ldots,\alpha_{n-1}),\alpha_n\right)$ for $n>2$. (The process $D^{(n)}$ can be seen as the departure process from a system of $n-1$ queues in tandem). Now for $\alpha=(\alpha_1,\ldots,\alpha_n)$ we can define a system of n ordered single-type TASEP configurations, denoted $T\alpha=\eta=(\eta^1,\ldots,\eta^n)$ by $\eta^k = D^{(n-k+1)}(\alpha_k,\ldots,\alpha_n)$. Then the corresponding multi-type configuration $\xi=\xi^{(1,\ldots,n)}$ is given by $\xi = R\eta = RT\alpha$, with $\xi(x) = n+1-\sum_{k=1}^n \eta^k(x)$ (as in the last paragraph of Section 2.2.3). See Remark 2.4 below for further explanation of the construction.

We can now state the main result.

To state this result, we work with systems with jumps *from right to left.* To return to the systems defined before, one simply takes the space-reversal ($\tilde{\eta}_t(x)=\eta_t(-x)$). Note that time in the queueing system corresponds to space in the particle system.

Theorem 2.3. *If $\alpha=(\alpha_1,\ldots,\alpha_n)$ has distribution $\nu=\nu_{\rho_1}\times\ldots\times\nu_{\rho_n}$ ($\mu=\mu_{\rho_1}\times\ldots\times\mu_{\rho_n}$ respectively for model R3) with $\rho_1<\ldots<\rho_n$, then the law of $T\alpha=\eta$ is invariant for the coupled multi-line TASEPs R0,R1 and R2 (R3 respectively) and the law of $RT\alpha=R\eta=\xi$ is invariant for the multi-type TASEPs R0, R1 and R2 (R3 respectively) with jumps from right to left. These are the unique stationary translation invariant (invariant under even shifts respectively) ergodic measures with density ρ_1 of first class particles (density ρ_1 of first class particles on even sites), density $\rho_2-\rho_1$ of second class particles (density $\rho_2-\rho_1$ of second class particles on even sites), etc.*

Remark 2.4. *The mechanism to construct an invariant distribution as described above can be depicted in the following way: Take α_1 as the arrival process and α_2 as the service process of a queue. Using α_1 and α_2 we can construct a process consisting* *of the departures from this queue (first class particles), unused services (second class particles) and times when no service was offered (holes). We then use this process as the arrival process for a queue with service process α_3 where first class particles have priority over second class particles: If there is a service and a first and a second class particle are waiting in the queue then the first class particles gets served first. In this way we get a resulting process consisting of departures of first class particles (first class particles), departures of second class particles (second class particles), unused services (third class particles) and holes. Now we can feed this process into a queue with*

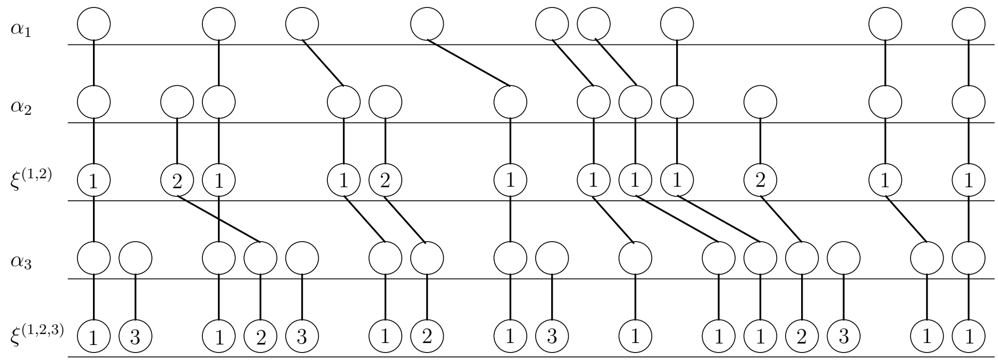

Figure 2.5: Queues in tandem and multi-type configurations $\xi^{(1,2)}$ and $\xi^{(1,2,3)}$

service process α_4 and so on. If $\alpha = (\alpha_1, \ldots, \alpha_n)$ has distribution $\nu = \nu_{\rho_1} \times \ldots \times \nu_{\rho_n}$ ($\mu = \mu_{\rho_1} \times \ldots \times \mu_{\rho_n}$ respectively) then the distribution of the resulting multi-type configuration is invariant for the multi-type TASEP. See Figure 2.5 for an illustration. Note that for models R0, R1, R2, the queues involved are simply $M/M/1$ queues in discrete-time; the same is almost true for R3, except that we have different arrival and service rates at odd and even times.

Remark 2.5. *We observe again that the invariant measures for the multi-type TASEPs R1 and R2 are the same as the invariant measures for the multi-type TASEP in continuous time and that they do not depend on β. Since the invariant measures for the single-type TASEP R3 depend on β the same is true for the invariant measures for the multi-type TASEP R3.*

2.3.2 Hydrodynamic limits

We now move to considering systems out of equilibrium. We consider the particular initial configuration given by

$$\eta_0(x) = \begin{cases} 1 & x \le 0 \\ 0 & x \ge 1 \end{cases}$$

This "step" initial condition corresponds to the corner growth model and to the particular initial conditions for the percolation models described in Section 2.2.2. We define the following functions f_0, f_1, f_2 and f_3, which will describe the evolving density profile for the continuous-time TASEP and for the TASEPs R1-R3 in discrete time:

$$f_0(u) = \mathbb{1}_{(-\infty,-1]}(u) + \frac{1}{2}(1-u)\cdot \mathbb{1}_{[-1,1]}(u)$$

$$f_1(u) = \mathbb{1}_{\left(-\infty,-\frac{\beta}{1-\beta}\right]}(u) + \frac{1}{\beta}\left(1-\sqrt{\frac{1-\beta}{1-u}}\right)\cdot \mathbb{1}_{\left[-\frac{\beta}{1-\beta},\beta\right]}(u)$$

$$f_2(u) = \mathbb{1}_{(-\infty,-\beta]}(u) + \left(1-\frac{1}{\beta}\left(1-\sqrt{\frac{1-\beta}{1+u}}\right)\right)\cdot \mathbb{1}_{\left[-\beta,\frac{\beta}{1-\beta}\right]}(u) = 1 - f_1(-u)$$

$$f_3(u) = \mathbb{1}_{\left(-\infty,-\frac{2\beta}{2-\beta}\right]}(u) + \left(\frac{1}{2}-\frac{u}{\beta}\sqrt{\frac{1-\beta}{4-u^2}}\right)\cdot \mathbb{1}_{\left[-\frac{2\beta}{2-\beta},\frac{2\beta}{2-\beta}\right]}(u)$$

We let a_3 be defined by $f_3(u) = \frac{1}{2}\left(a_3(u) + \frac{a_3(u)(1-\beta)}{1-a_3(u)\beta}\right)$, so

$$a_3(u) = \mathbb{1}_{\left(-\infty,-\frac{2\beta}{2-\beta}\right]}(u) + \frac{1}{\beta}\left(1-(2+u)\sqrt{\frac{1-\beta}{4-u^2}}\right)\cdot \mathbb{1}_{\left[-\frac{2\beta}{2-\beta},\frac{2\beta}{2-\beta}\right]}(u).$$

For $i = 0,1,2,3$ let $\tau_i(k,t)$ ($t \in \mathbb{R}_+$ or $t \in \mathbb{N}$ respectively) be the distribution of $(\eta_t(k+l), l \in \mathbb{Z})$ in the corresponding model. We have the following result for the TASEP in continuous time and the discrete TASEPs R1-R3.

Theorem 2.6. *For any $u \in \mathbb{R}$ and $i = 0,1,2$ the measure $\tau_i([ut],t)$ converges weakly to the Bernoulli product measure with marginals $f_i(u)$ and $\tau_3([ut],t)$ converges weakly to the Bernoulli product measure with marginals $a_3(u)$ on even sites and $\frac{a_3(u)(1-\beta)}{1-a_3(u)\beta}$ on* *odd sites. In particular we have that for any $u \in \mathbb{R}$ the limit $\lim_{t\to\infty} \mathbb{E}\left[\eta_t(k)\right]$ exists and is equal to $f_i(u)$, $i = 0, 1, 2$ depending on which model we are considering, whenever $\frac{k}{t}$ tends to u, and $\lim_{t\to\infty} \mathbb{E}\left[\eta_t(2[\frac{k}{2}])\right] = a_3(u)$ and $\lim_{t\to\infty} \mathbb{E}\left[\eta_t(2[\frac{k}{2}]+1)\right] = \frac{a_3(u)(1-\beta)}{1-a_3(u)\beta}$ in the model R3. Furthermore, for $i = 0, 1, 2, 3$, the quantities $\frac{1}{t}\sum_{ut<k<vt} \eta_t(k)$ converge a.s. to the constant value $\int_u^v f_i(w)dw$, for $u < v$.*

The first part of Theorem 2.6 states convergence to local equilibrium: suitably rescaled the models converge locally to the unique invariant measures from Theorem 2.1. This implies the other statements of Theorem 2.6. However, in the models R0, R1 and R2 we can prove the second part without proving convergence to local equilibrium first, while in the model R3 our proof for the second part requires convergence to local equilibrium. The statements for the model in continuous time were proved for the first time by Rost [45]. O'Connell [43] used the connection between the TASEP and last-passage percolation to prove an equivalent result about the asymptotic shape of the corner growth model (as defined in Section 2.2.2). The parts of Theorem 2.6 concerning the models in discrete time can be proved using exactly the same methods as Rost [45] and O'Connell [43]. In Section 2.4.2 we will outline the proof for the model R3.

Remark 2.7. *From the convergence to the density profiles f_0, f_1,f_2 and f_3 we can easily deduce a shape theorem for the corner growth model defined in Section 2.2.2. The asymptotic shape in the models R1, R2 and R3 (after rescaling by t) are for example given by the functions*

$$c_1(x) = \frac{1}{1-\beta}\left(\sqrt{\beta} - \sqrt{x}\right)^2 \tag{2.9}$$

for $x \in [0, \beta]$,

$$c_2(x) = \left(\sqrt{\beta} - \sqrt{(1-\beta)x}\right)^2 \tag{2.10}$$

for $x \in [0, \frac{\beta}{1-\beta}]$, see Figure 2.6 (simulation with $\beta = 0.5$ up to time $n = 1000$), and

$$c_3(x) = \frac{1}{(2-\beta)^2}\left(\sqrt{4x(1-\beta)} - \sqrt{4\beta - \beta^2 x - 2\beta^2}\right)^2 \tag{2.11}$$

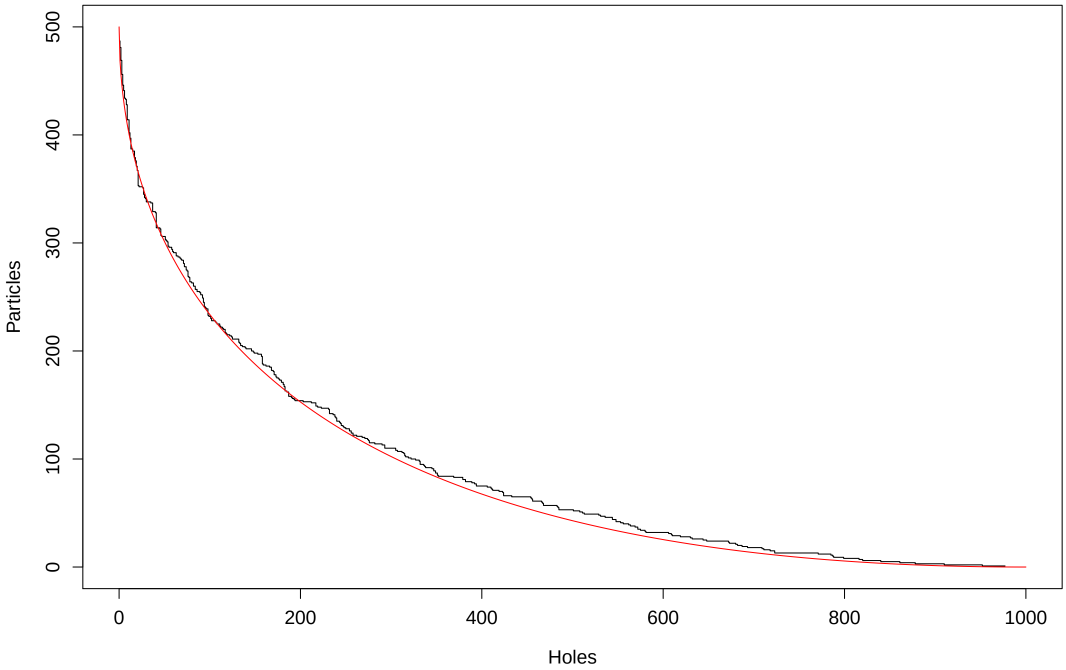

Figure 2.6: Black: Simulation of the corner growth model (R2) for $\beta = 0.5$ and time $n = 1000$; Red: Limiting shape given by a rescaled version of c_2

for $x \in \left[0, \frac{2\beta}{2-\beta}\right]$*. We have convergence in the following sense, see [19]. For* $i = 1, 2, 3$ *we have that almost surely, for every* $\varepsilon > 0$*, there exists a* t_0 *such that for all* $t \geq t_0$

$$t\,(1-\varepsilon)\,\mathcal{M}_i \subset \overline{G}^{(i)}(t) \subset t\,(1+\varepsilon)\,\mathcal{M}_i$$

where $\overline{G}^{(i)}(t) \subset (\mathbb{R}_+)^2$ *is the set with boundary* $G^{(i)}(t)$ *and*

$$\mathcal{M}_i = \left\{(u, v) \in (\mathbb{R}_+)^2 : c_i(u) \leq v\right\}.$$

Remark 2.8. *We may rescale time as well as space in Theorem 2.6, and look at the limit* $f(u,t) = \lim_{N\to\infty} \mathbb{E}\left[\eta_{Nt}(k)\right]$ *for* $\lim_{N\to\infty} \frac{k}{N} = u$*. For the continuous-time model,* *this density profile is governed by Burgers' equation (see [15]),*

$$\frac{\partial f}{\partial t} + \frac{\partial (f(1-f))}{\partial u} = 0; \tag{2.12}$$

the solution, with initial condition $f(u,0) = \mathbb{1}_{(-\infty,0]}(u)$, *is*

$$f(u,t) = \mathbb{1}_{(-\infty,-t]}(u) + \frac{t-u}{2t} \cdot \mathbb{1}_{[-t,t]}(u).$$

The differential equation also governs the evolution of the density profile for more general initial configurations than the "step" initial condition. We can get equations analogous to (2.12) for the models in discrete time. For example, for model R1,

$$f_1(u,t) = \mathbb{1}_{(-\infty,-\frac{\beta t}{1-\beta}]}(u) + \frac{1}{\beta}\left(1 - \sqrt{\frac{t(1-\beta)}{t-u}}\right) \cdot \mathbb{1}_{[-\frac{\beta t}{1-\beta},\beta t]}(u)$$

solves

$$\begin{cases} \frac{\partial f_1}{\partial t} + \frac{\partial}{\partial u}\frac{\beta f_1(1-f_1)}{1-\beta f_1} = 0 \\ f_1(u,0) = \mathbb{1}_{(-\infty,0]}(u) \end{cases} \tag{2.13}$$

Here $\frac{\beta f_1(1-f_1)}{1-\beta f_1} = \sum_{n=1}^{\infty} \beta^n f_1^n (1-f_1)$ *is the probability that a particle jumps from a given site to its neighbour in a model in equilibrium with marginal density* f_1.

2.3.3 Multi-type models out of equilibrium

In this section we consider multi-type TASEPs $\xi_t \in \mathbb{Z}^{\mathbb{Z}}$ similar to Section 2.2.3. With the results from Theorem 2.6 we can calculate the distribution of the asymptotic speed of a single second class particle in the TASEP with initial configuration

$$\xi_0(x) = \begin{cases} 1 & x \leq -1 \\ 2 & x = 0 \\ 3 & x \geq 1 \end{cases}$$

As the particles of class 3 are weaker than all other particles in the model we can think of these particles as holes. So the second-class particle sees only particles to its left and only holes to its right. The second-class particle can be seen as a discrepancy between two copies of the "step" initial condition considered in the Section 2.3.2, one of which is shifted by one step to the right. Hence the path of the second-class particle corresponds to the propagation of the discrepancy under the basic coupling. The results for the models in discrete time correspond to the result for the model in continuous time first obtained in Ferrari and Kipnis [15]. They prove convergence in distribution. In order to prove a.s. convergence we can use the connection to last-passage percolation and the growth model: As in Ferrari and Pimentel [19] the path of the second class particle corresponds to a competition interface in the growth model which has a.s. an asymptotic direction.

Theorem 2.9. *For $i = 0,1,2,3$ let $X^{(i)}(t)$ denote the position of the second class particle at time t in the corresponding model. Then we have*

$$\frac{X^{(i)}(t)}{t} \xrightarrow[t\to\infty]{a.s.} U^{(i)} \in \begin{cases} [-1,1] & i=0 \\ \left[-\frac{\beta}{1-\beta},\beta\right] & i=1 \\ \left[-\beta,\frac{\beta}{1-\beta}\right] & i=2 \\ \left[-\frac{2\beta}{2-\beta},\frac{2\beta}{2-\beta}\right] & i=3 \end{cases}$$

for random variables $U^{(i)}$ with distribution functions $1-f_i$ for $i=0,1,2$ and a_3 for $i=3$.

Remark 2.10. *The proofs for convergence in distribution are analogous to those for the model in continuous time (see for example [15]) apart from a small complication in the model R3 with the particle-particle coupling. We will give an account of this proof in Section 2.4.3. For the almost sure convergence we will explain the construction of the competition interface for the models in discrete time and prove that the second class particle has an asymptotic speed almost surely by using results about semi-infinite geodesics in the percolation models similar to [19]. An interesting observation will be* *that the distribution of the asymptotic direction of the competition interface corresponding to the path of the second class particle is the same in the models R1 and R3, see Remark 2.34. However, this does not imply that the distribution of the speed of the second class particle is the same in the two models.*

Remark 2.11. *In the continuous model the distribution of the asymptotic speed of the second class particle turns out to be uniform on* $[-1, 1]$*. The distributions in the models in discrete time are more complicated.*

Using Theorem 2.9 we can define the following so-called speed process: Consider the multi-type TASEP with initial configuration $\xi_0(n) = n$. By Theorem 2.9 we know that each particle has a.s. an asymptotic speed U_n: Particle n has only stronger particles to its left and only weaker particles (that can be seen as holes) to its right just like the second class particle in the initial configuration of Theorem 2.9 and therefore we can apply Theorem 2.9 to the speed of every particle. We call the process $\{U_n\}_{n\in\mathbb{Z}}$ the speed process and denote its distribution by μ. This process is stationary, and its marginals (for the various models) are given by the distributions in Theorem 2.9. Furthermore, we write $Y_n(m)$ instead of $\xi_m(n)$ and denote the position of particle n at time m by $X_n(m)$ in order to be consistent with the notation introduced by Amir, Angel and Valkó [1]. They have studied the speed process for the model in continuous time. Note that both Y_n and X_n can be seen as permutations of the set $\mathbb{Z}$, and are inverse to each other. We define $g_i = 1 - f_i$ for $i = 0, 1, 2$ and $g_3 = 1 - a_3$, $g_4 = 1 - \frac{a_3(1-\beta)}{1-a_3\beta}$ and our first result is the following Theorem 2.12 corresponding to Theorem 1.5 in [1] (note that the labels of particles can now be in $\mathbb{R}$ instead of just $\mathbb{Z}$):

Theorem 2.12. *For* $i = 0, 1, 2$, $\mu^{(i)}$*, the distribution of the speed process in model Ri, is the unique stationary ergodic measure for the TASEP R0 whose marginals have distribution function* g_i*. Correspondingly, for* $i = 3, 4$, $\mu^{(i)}$ *is the unique stationary measure for the TASEP* $Rj(i)$*, which has marginals distributed according to* g_i *on even sites and* $g_{j(i)}$ *on odd sites, where* $j(3) = 4$ *and* $j(4) = 3$.

For $i = 0$ this gives the result from [1] saying that the distribution of the speed process is itself a stationary ergodic measure for the TASEP in continuous time (the marginals are uniform on $[-1,1]$ in this case). The other parts of Theorem 2.12 follow from nice dualities between the models R1 and R2 and between the models R3 and R4, and the fact that R0, R1 and R2 all have the same set of stationary distributions, whatever the value of β (as given in Theorem 2.3). The dualities are given by the following result:

Theorem 2.13. *Consider the starting configuration $Y_n(0) = \xi_0(n) = n$. For $i = 0,1,2,3,4$ and any fixed $m > 0$ the process $\{X_n^{(i)}(m)\}_{n\in\mathbb{Z}}$ has the same distribution as the process $\{Y_n^{(j(i))}(m)\}_{n\in\mathbb{Z}}$ where $j(0) = 0$, $j(1) = 2$, $j(2) = 1$, $j(3) = 4$ and $j(4) = 3$.*

The following Theorems 2.14 - 2.19 provide some explicit results about the joint distributions of the speeds of adjacent particles (and particles 0 and 2 in model R3). The first result is Theorem 1.7 in [1]. Theorems 2.15 - 2.19 and Remark 2.16 give analogous results for the TASEPs R1, R2 and R3.

Theorem 2.14 (TASEP R0)**.** *The joint distribution of (U_0, U_1), supported on $[-1,1]^2$, is*

$$s(x,y)dxdy + r(x)\mathbb{1}_{\{x=y\}}dx$$

with

$$s(x,y) = \begin{cases} \frac{1}{4} & x > y \\ \frac{y-x}{4} & x \le y \end{cases} \qquad \text{and} \qquad r(x) = \frac{1-x^2}{8}.$$

In particular, $\mathbb{P}[U_0 > U_1] = \frac{1}{2}$, $\mathbb{P}[U_0 = U_1] = \frac{1}{6}$ and $\mathbb{P}[U_0 < U_1] = \frac{1}{3}$.

Theorem 2.15 (TASEP R1)**.** *The joint distribution of (U_0, U_1), supported on the set $\left[-\frac{\beta}{1-\beta}, \beta\right]^2$, is*

$$s_1(x,y)dxdy + r_1(x)\mathbb{1}_{\{x=y\}}dx$$

with

$$s_1(x,y) = \begin{cases} \frac{1-\beta}{4\beta^2}(1-x)^{-\frac{3}{2}}(1-y)^{-\frac{3}{2}} = g_1'(x)g_1'(y) & x > y \\ \frac{1-\beta}{2\beta^3}(1-x)^{-\frac{3}{2}}(1-y)^{-\frac{3}{2}}\left(\sqrt{\frac{1-\beta}{1-y}} - \sqrt{\frac{1-\beta}{1-x}}\right) & x \le y \end{cases}$$

and

$$r_1(x) = \left(\frac{\sqrt{1-\beta}}{2\beta^2(1-u)^{\frac{3}{2}}} \left(1 - \frac{1}{\beta}\right) + \frac{1-\beta}{2\beta^2(1-u)^2} \left(\frac{2}{\beta} - 1\right) - \frac{\sqrt{1-\beta}(1-\beta)}{2\beta^3(1-u)^{\frac{5}{2}}} \right).$$

In particular, $\mathbb{P}[U_0 > U_1] = \frac{1}{2}$, $\mathbb{P}[U_0 = U_1] = \frac{1}{6}$ *and* $\mathbb{P}[U_0 < U_1] = \frac{1}{3}$.

Remark 2.16. *By symmetry, the joint distribution of* (U_0, U_1) *in the model R2 is the same as that of* $(-U_1, -U_0)$ *in the model R1.*

Theorem 2.17 (TASEP R3)**.** *The joint distribution of* (U_0, U_1)*, supported on the set* $\left[-\frac{2\beta}{2-\beta}, \frac{2\beta}{2-\beta}\right]^2$*, is*

$$s_2(x,y)dxdy + r_2(x)\mathbb{1}_{\{x=y\}}dx$$

with

$$s_2(x,y) = \begin{cases} g_3'(x)g_4'(y) & x > y \\ g_3'(x)g_4'(y)\left(g_4(y) - g_4(x)\right) & \\ \qquad \cdot \left(2 - g_4(x)\beta - g_4(y)\beta\right)\left(\frac{2+y}{2-y}\right) & x \le y \end{cases}$$

and

$$r_2(x) = \frac{1-\beta}{\beta^3}\left(\frac{2(2-\beta)}{4-x^2} - \frac{8}{4-x^2}\sqrt{\frac{1-\beta}{4-x^2}}\right).$$

In particular,

$$\mathbb{P}[U_0 > U_1] = \frac{1}{\beta^2}\left(\beta - (1-\beta)\log\left(\frac{1}{1-\beta}\right)\right)$$

$$\mathbb{P}[U_0 = U_1] = \frac{(1-\beta)(2-\beta)}{\beta^3}\log\left(\frac{1}{1-\beta}\right) - \frac{2(1-\beta)}{\beta^2}$$

and

$$\mathbb{P}[U_0 < U_1] = \frac{(1-\beta)(2-\beta)}{\beta^2} + \frac{2(1-\beta)^2}{\beta^3}\log(1-\beta).$$

Remark 2.18. *Again by symmetry, we have that under rule R3,* (U_1, U_2) *has the same distribution as* $(-U_1, -U_0)$*.*

Theorem 2.19 (TASEP R3)**.** *The joint distribution of* (U_0, U_2)*, supported on the set* $\left[-\frac{2\beta}{2-\beta}, \frac{2\beta}{2-\beta}\right]^2$*, is*

$$s_3(x,y)dxdy + r_3(x)\mathbb{1}_{\{x=y\}}dx$$

with

$$s_3(x,y) = \begin{cases} g_3'(x)g_3'(y) & x > y \\ g_3'(x)g_3'(y)\Big(g_4(x)^2 - g_4(y)^2 \\ \quad -\frac{2(g_4(x)-1)(g_4(y)^2-2g_4(x)g_4(y)+g_4(x)-1}{1-\beta g_4(x)} \\ \quad\quad +\frac{2(g_4(y)-1)(g_4(x)^2-2g_4(x)g_4(y)+g_4(y)-1}{1-\beta g_4(y)} - 1\Big) & x \le y \end{cases}$$

and

$$r_3(x) = \frac{g_3(u)(1-g_3(u))(1-g_4(u)(1-g_4(u)))}{\frac{(2-u)\beta}{2}\sqrt{\frac{4-u^2}{1-\beta}}}.$$

In particular,

$$\mathbb{P}[U_0 > U_2] = \frac{1}{2}$$

$$\mathbb{P}[U_0 = U_2] = \frac{1}{6} + \frac{1}{3\beta} - \frac{13}{3\beta^2} + \frac{8}{\beta^3} - \frac{4}{\beta^4} - \frac{(1-\beta)^2}{\beta^3}\left(\log\left(\frac{1}{1-\beta}\right)\right)\left(\frac{2}{\beta} - \frac{4}{\beta^2}\right)$$

and

$$\mathbb{P}[U_0 < U_2] = \frac{1}{3} - \frac{1}{3\beta} + \frac{13}{3\beta^2} - \frac{8}{\beta^3} + \frac{4}{\beta^4} + \frac{(1-\beta)^2}{\beta^3}\left(\log\left(\frac{1}{1-\beta}\right)\right)\left(\frac{2}{\beta} - \frac{4}{\beta^2}\right).$$

We see that in every model we have that the speeds are independent on the set where $U_0 > U_1$ ($U_0 > U_2$ respectively). This agrees with the result in continuous time. The striking result, shown in [1] for the continuous model, that with positive probability the two continuous random variables U_0 and U_1 are equal, holds also in the discrete models.

Interestingly, the probabilities $\mathbb{P}[U_0 > U_1]$, $\mathbb{P}[U_0 = U_1]$ and $\mathbb{P}[U_0 < U_1]$ are the same for models R0, R1 and R2, and do not depend on the parameter β. This is rather surprising since β is not just a scaling parameter (i.e. we cannot produce models with different values of β by just applying a time change). In fact, much more is true. From the first part of Theorem 2.3, we see that, although the marginal distribution of each U_i depends on the model and the value of β, we can obtain the distribution for either of R1 and R2 and any value of β by applying an appropriate monotone function to each entry U_i (see the proof of Theorem 2.15 for further details). Hence the relative ordering of the variables U_i is not affected by the model or the value of β.

To go further, consider particles i and j with $i < j$. It's clear that if $U_i < U_j$ then particle i can never overtake particle j, while if $U_i > U_j$ then particle i must overtake particle j. In [1], it's shown that for the continuous-time model, with probability 1, if $U_i = U_j$ then particle i overtakes particle j. The same result can be shown for the discrete-time models, although the calculations involved in the argument are rather more complicated than those used to prove Theorem 1.14 of [1], and we omit them here. So, for example, the probability that particle i overtakes particle j is the same for models R0, R1 and R2. Indeed, more completely one can define an ordering $\prec$ on $\mathbb{Z}$ by $i \prec j$ iff particle j is eventually to the right of particle i. Then we have the following result:

Corollary 2.20. *The ordering $\prec$ has the same distribution for R0, R1 and R2 and for any value of β.*

It would certainly be interesting to have a more direct understanding of this property, based for example on couplings or local dynamics, as well as the indirect argument based on the equivalence of multi-type equilibrium distributions.

Overtaking probabilities in the multi-type TASEP can also be interpreted in terms of questions of survival or extinction in multi-type growth models. In [16] a coupling is given between the multi-type TASEP and a three-type version of the corner growth model, under which a given cluster survives for ever if and only if particle 0 never overtakes particle 1. (The extinction of the cluster occurs if two interfaces in the growth model meet – the paths of these interfaces are related to the paths of the two particles). Different overtaking events in the TASEP can be represented by varying the initial condition in the competition growth model. Using the results above, we find that the survival probabilities in the growth model will remain unchanged if we move from the continuous-time model to natural discrete-time models which correspond to models R1 or R2 in the TASEP. Again, this is certainly not obvious from the local dynamics of the processes.

Unlike in models R1 and R2, in the model R3 the probabilities $\mathbb{P}[U_0 > U_1]$, $\mathbb{P}[U_0 = U_1]$ and $\mathbb{P}[U_0 < U_1]$ do depend on β and the behaviour of the model is qualitatively different for different values of β (Theorem 2.17): For small β we have $\mathbb{P}[U_1 < U_0] > \mathbb{P}[U_1 > U_0]$, but $\mathbb{P}[U_1 < U_0] < \mathbb{P}[U_1 > U_0]$ for large β (the transition occurs at $\beta = 0.38860064568\ldots$).

Note however that for $\beta \to 1$ the probabilities relating U_0 and U_2 in Theorem 2.19 converge to $\frac{1}{2}$, $\frac{1}{6}$ and $\frac{1}{3}$, i.e. to the probabilities we get in the continuous model and R1 and R2 for the speeds of particle 0 and 1. In a sense, for large β the particles 0 and 2 in the model R3 behave like adjacent particles in the models R1 and R2. This can heuristically be seen in the following way: We consider the particles in the model R3 (with large β close to 1) starting on even sites. In general, particles starting on an even site will move two steps to the right in each time-step since β is large and we update even sites first. If a particle does not jump either during the even or the odd update (which happens with probability $2\beta(1-\beta)$) it ends up on an odd site and starts moving left until either

- (A) it hits a weaker particle to the left by which it cannot be overtaken
- (B) it does not get jumped over either during an even or odd update because the adjacent particle to the left did not try to jump

In both cases the particle itself will return to an even site (with high probability) and resume moving to the right. The particle that caused the stop (either because it was weaker or because it did not try to jump) will itself start moving to the left until (A) or (B) happens. Now consider the model R1 with large β. Most particles will move one step to the right in each time-step, but some particles do not jump and therefore get overtaken until again either (A) or (B) happens (where we remove the part "either during and even or odd update"). Particles in these two models have different speeds, but the probabilities $\mathbb{P}[U_0 > U_1]$, $\mathbb{P}[U_0 = U_1]$, $\mathbb{P}[U_0 < U_1]$ in R1 and $\mathbb{P}[U_0 > U_2]$, $\mathbb{P}[U_0 = U_2]$ and $\mathbb{P}[U_0 < U_2]$ in R3 are (almost) the same.

2.4 Proofs

2.4.1 Invariant Measures

Proof of Theorem 2.3: The idea of the proof for Theorem 2.3 for the discrete-time models is the same as for the continuous-time model R0 (see [17]), but here we will use a slightly different definition for the so-called dual points: In [17] the authors used an explicit construction that transformed a Bernoulli process ω into a corresponding set $\triangle(\omega)$ of dual points. Here we provide a more general method to obtain these dual points. We will not provide an explicit set of dual points, but only explain how to obtain one that has the required properties. When thinking about the model R3 bear in mind that we have different densities on even and odd sites. We start with the following proposition, that asserts that for every particle density ρ and every set of Bernoulli points ω there is essentially only one stationary and space-ergodic TASEP trajectory. This corresponds to Proposition 8 in [17].

Proposition 2.21. *For any particle density $\rho \in (0,1)$ and Bernoulli marks ω we can get an essentially unique stationary and space-ergodic trajectory $(\eta_n, n \in \mathbb{Z})$ for the TASEP governed by ω with time-marginals ν_ρ.*

Remark 2.22. *Essentially unique means that if we have two stationary and space-ergodic trajectories then they are the same with probability 1.*

Proof of Proposition 2.21: Proposition 2.21 can be proved using an argument similar to the one used in [41], see also [17]. □

Remark 2.23. *In the following we consider Bernoulli processes on $\left(\mathbb{Z} + \frac{1}{2}\right) \times \mathbb{Z}$ instead of $\mathbb{Z}^{\mathbb{Z}}$. A Bernoulli point $\left(x + \frac{1}{2}, m\right)$ corresponds to a jump attempt from site x to $x+1$ at time m.*

Now we define for each of the three models a set of dual points $\triangle_\rho(\omega)$ such that the time reflections of these dual points govern the time reversal of the TASEP. Given the Bernoulli process ω and the trajectory $(\eta_n, n \in \mathbb{Z})$ we define $C_1 \subset \omega$ as the set of jump marks that actually lead to a particle jumping one step to the left (i.e. the site that was updated was occupied and the particle was not blocked). Now we look at the time reversal $(\widetilde{\eta}_n, n \in \mathbb{Z}) = (\eta_{-n}, n \in \mathbb{Z})$ of $(\eta_n, n \in \mathbb{Z})$. Because we want the time reflections of the dual points to govern the time reversal of the TASEP we have to have $C_1 \subset \triangle_\rho(\omega)$.

The time reversal has the same distribution as the space reversal. The space reversal is governed by a Bernoulli process (the space reversal of ω) and particles jump to the right. In particular, all possible jumps happen at rate β. That means that in the time reversal all possible jumps happen at rate β, as well.

Let C_2 be the set of points $x \in \left(\left(\mathbb{Z} + \frac{1}{2}\right) \times \mathbb{Z}\right) \backslash C_1$ for which adding x to $\triangle_\rho(\omega)$ would not change the trajectory of the time reversal (i.e. if we let the time reversal be governed by the time reflections of the dual points then the trajectory does not change if we add or remove x from $\triangle_\rho(\omega)$). Let $\widetilde{\omega}$ be a Bernoulli process with rate β on C_2, independent of everything else. That means $\widetilde{\omega} \subset C_2$ and each point from C_2 is present in $\widetilde{\omega}$ independently with probability β. Now put $\triangle_\rho(\omega) = C_1 \cup \widetilde{\omega}$. The properties of these dual points are stated in the following proposition that corresponds to Proposition 10 in [17].

Proposition 2.24. *The time reflections of the dual points $\triangle_\rho(\omega)$ defined above govern the time reversal of $(\eta_n, n \in \mathbb{Z})$ and they form a Bernoulli process with rate β on $\left(\mathbb{Z} + \frac{1}{2}\right) \times \mathbb{Z}$. Furthermore, the set of dual points before time m is independent of the configuration η_m.*

Proof. The first two statements are true by the construction of the dual points. The dual points $\{(x, n) \in \triangle_\rho(\omega) : n < m\}$ govern the evolution of the time reversal on the time interval $(-m, \infty)$ starting from the configuration $\widetilde{\eta}_{-m}$ and they are independent of this configuration. But $\widetilde{\eta}_{-m} = \eta_m$. So the set of dual points $\{(x, n) \in \triangle_\rho(\omega) : n < m\}$ is independent of η_m. □

We now take $\rho_1 < \ldots < \rho_n$ and let $\alpha_m = (\alpha_m^1, \ldots, \alpha_m^n)$ be the multiline TASEP trajectory governed by ω. This means that $\omega^n = \omega$, $\omega^k = \triangle_{\rho_{k+1}}(\omega^{k+1})$ and $(\alpha_m^k)_{m \in \mathbb{Z}}$

is the TASEP trajectory governed by ω^k with density ρ_k. Then by the independence of the dual points before time m from the configuration η_m we get that the multiline process is stationary with product measure $\nu = \nu_{\rho_1} \times \ldots \times \nu_{\rho_n}$ (Proposition 11 in [17]). As in the paragraph preceding Theorem 2.3 we define $\eta = (\eta^1, \ldots, \eta^n)$ by $\eta^k = D^{(n-k+1)}(\alpha_k, \ldots, \alpha_n)$. Then induction arguments and some case-by-case checking for $n = 2$ as done in [18] show that $(\eta^k_n)_{n\in\mathbb{Z}}$ is the TASEP trajectory governed by ω with particle density ρ_k (Proposition 12 in [17]) and this implies Theorem 2.3. □

Remark 2.25. *As mentioned in the beginning of this section, for the model R3 we have to think of the ρ_k as densities on even sites and we have to replace the ν_{ρ_k} by μ_{ρ_k}.*

Remark 2.26. *Inherent in the tandem queue construction for the multi-type stationary distribution in model R3 is a version of Burke's theorem for the queues with different arrival and service rates on even and odd sites. Consider a queue with arrival process A_n, service process S_n and departure process D_n. Let A_n be a Bernoulli process with rate $\rho_1 = (\gamma_1, \gamma_2) \in (0,1)^2$ which means that on even sites arrivals happen with probability γ_1 and on odd sites they happen with probability γ_2. Motivated by the invariant distributions for the TASEP R3 with just one type of particles we want γ_1 and γ_2 to satisfy*

$$\gamma_2 = \frac{\gamma_1 (1-\beta)}{1-\gamma_1\beta} \tag{2.14}$$

where $\beta \in (0,1)$ is the rate at which jumps in the TASEP happen. Analogously, we let S_n be a Bernoulli process with rate $\rho_2 = (\delta_1, \delta_2) \in (0,1)^2$ where

$$\delta_2 = \frac{\delta_1 (1-\beta)}{1-\delta_1\beta} \tag{2.15}$$

(and $\gamma_1 < \delta_1$, $\gamma_2 < \delta_2$). The main observation in Burke's theorem (see for example [9]) that shows that arrival and departure process have the same distribution is that the queue length process is reversible and that departures look in the reversed process

like arrivals in the original process. Interestingly, it turns out that in the queueing model described above there exists a stationary reversible distribution π for the queue length process which is independent of whether we just observed arrivals and services at even sites or at odd sites. π is given by

$$\pi(j) = \left(1 - \frac{\gamma_1 (1-\delta_1)}{(1-\gamma_1)\,\delta_1}\right)\left(\frac{\gamma_1 (1-\delta_1)}{(1-\gamma_1)\,\delta_1}\right)^j \qquad j = 0, 1, \dots$$

and it is reversible because it satisfies the two systems of detailed balance equations

$$\pi(j)\gamma_i (1-\delta_i) = \pi(j+1)(1-\gamma_i)\,\delta_i \qquad j = 0, 1, \dots$$

for $i = 1, 2$ ($i = 1$ corresponds to even sites, $i = 2$ corresponds to odd sites). This follows from the relations (2.14) and (2.15). As in Burke's theorem it follows from the reversibility of the queue length process that the departure process has the same distribution as the arrival process, i.e. D_n is a Bernoulli process with rate $\rho_1 = (\gamma_1, \gamma_2)$.

Indeed, the multi-type construction yields extensions of this result which give input-output theorems for the priority queues with more than one type of customer. For a discussion of the analogous result in the context of constant arrival and service rates, see for example Section 6 of [17].

2.4.2 Hydrodynamic limits

Proof outline for Theorem 2.6 for the TASEP R3: The following Propositions 2.27, 2.29 and 2.30 correspond to Propositions 2, 3 and 5 in [45] and the proofs are essentially the same as in [45].

Proposition 2.27. *For all $u \in \mathbb{R}$ the random variables $\frac{1}{n}S([un], n) = \frac{1}{n}\sum_{k>[un]} \eta_n(k)$ converge a.s. and in L^1 to a constant $h_3(u)$, as n goes to infinity. The function h_3 is decreasing, convex; one has $h_3(u) = 0$ for $u > \frac{2\beta}{2-\beta}$ and $h_3(u) = -u$ for $u < -\frac{2\beta}{2-\beta}$.*

Remark 2.28. *Since h_3 is decreasing, it follows from Lebesgue's differentiation theorem that h_3 is differentiable almost everywhere, see for example [25].*

Proposition 2.29. *If h_3 is differentiable at u, one has*

$$\lim_{n\to\infty} \mathbb{E}\left[\eta_n\left(2\left[\frac{k}{2}\right]\right) + \eta_n\left(2\left[\frac{k}{2}\right]+1\right)\right] = -2h_3^{'}(u)$$

whenever $\frac{k}{n}$ tends to u.

Proposition 2.30. *Let $\mu_3(k,n)$ be the distribution of $(\eta_n(k+l), l \in \mathbb{Z})$. If $h_3^{'}(u)$ exists, any weak limit μ_3^* of the measures $\mu_3\left(2\left[\frac{un}{2}\right], n\right)$ for $n \to \infty$ is of the form*

$$\mu_3^* = \int_0^1 \tau_x \sigma(dx)$$

with some probability σ on $[0,1]$. τ_x is the Bernoulli product measure with density $b(x)$ on even sites and density $\frac{b(x)(1-\beta)}{1-b(x)\beta}$ on odd sites where $b(x)$ is such that the average density is given by $x = \frac{1}{2}\left(b(x) + \frac{b(x)(1-\beta)}{1-b(x)\beta}\right)$. That means that from Proposition 2.29 it follows that the measure σ satisfies

$$\int_0^1 x\sigma(dx) = f_3(u) = -h_3^{'}(u).$$

We can use the results from O'Connell [43] about last-passage percolation (see (2.7)) to calculate the function h_3:

$$h_3(u) = \begin{cases} -u & u \leq -\frac{2\beta}{2-\beta} \\ \frac{1}{\beta}\left(2 - \frac{u\beta}{2} - \sqrt{(4-u^2)(1-\beta)}\right) - 1 & -\frac{2\beta}{2-\beta} \leq u \leq \frac{2\beta}{2-\beta} \\ 0 & u \geq \frac{2\beta}{2-\beta} \end{cases}$$

Since h_3 is differentiable we can identify $f_3 = -h_3^{'}$ from Proposition 2.30 as

$$f_3(u) = -h_3^{'}(u) = \begin{cases} 1 & u \leq -\frac{2\beta}{2-\beta} \\ \frac{1}{2} - \frac{u}{\beta}\sqrt{\frac{1-\beta}{4-u^2}} & -\frac{2\beta}{2-\beta} \leq u \leq \frac{2\beta}{2-\beta} \\ 0 & u \geq \frac{2\beta}{2-\beta} \end{cases}$$

As mentioned after Theorem 2.6 in the models R0, R1 and R2 this is enough to prove the convergence statements at the end of Theorem 2.6; this is done using the monotonicity of the distribution of $\eta_n(k)$ in k. However, due to the different behaviour at odd and even sites, this monotonicity does not hold for the model R3, and the average density f_3 does not pick up the density fluctuations between even and odd sites. Hence without knowing that the model converges to local equilibrium (which would allow us to calculate the function a_3 from f_3) we cannot prove the last part of Theorem 2.6. The essential step for proving convergence to local equilibrium is the following proposition, the proof of which is again the same as in [45] (Proposition 6). Let $\rho(k;F;n) = \mathbb{P}\left[\eta_n(k+i) = 1, i \in F\right]$ for a set $F \subset \mathbb{Z}$.

Proposition 2.31. *For any finite set F and any $\epsilon > 0$ there exists a $\delta > 0$, n_0 such that*

$$\left|\rho\left(2\left[\frac{un}{2}\right];F;n\right) - \rho\left(2\left[\frac{\overline{u}n}{2}\right];F;n\right)\right| \leq \epsilon$$

for $|u - \overline{u}| \leq \delta$ and $n \geq n_0$. Also

$$\left|\rho\left(2\left[\frac{\overline{u}n}{2}\right];F;n+l\right) - \rho\left(2\left[\frac{\overline{u}n}{2}\right];F;n\right)\right| \leq \epsilon$$

for $0 \leq l \leq [\delta n]$, $n \geq n_0$.

Using this proposition and Jensen's inequality we can prove that the measure σ from Proposition 2.30 is the unit mass on $f_3(u)$ and since $b(f_3(u)) = a_3(u)$ this implies convergence to local equilibrium (see [45], Section 4 for the details). □

2.4.3 Multi-type models out of equilibrium

Proof of Theorem 2.9: First we want to outline the proof for convergence of $\frac{X^{(i)}(t)}{t}$ in distribution (as before we have $t \in \mathbb{R}_+$ or $t \in \mathbb{N}$ depending on the model). This follows the ideas in [15]. We want to couple two TASEPs with initial configurations

$$\eta_0^1(x) = \begin{cases} 1 & x \leq 0 \\ 0 & x \geq 1 \end{cases} \text{ and } \eta_0^2(x) = \begin{cases} 1 & x \leq -1 \\ 0 & x \geq 0 \end{cases}$$

in two different ways and calculate the difference $\mathbb{E}\left[S^1([rt], t)\right] - \mathbb{E}\left[S^2([rt], t)\right]$ in both couplings. $S^1([rt], t)$ and $S^2([rt, t])$ are the number of particles to the right of $[rt]$ at time t in η^1 and η^2. Using basic coupling (i.e. using the same Poisson processes $\{(P_t^x)_{t\geq 0} : x \in \mathbb{Z}\}$ or Bernoulli processes $\{(B_n^x)_{n\in\mathbb{N}} : x \in \mathbb{Z}\}$ for η^1 and η^2) gives

$$\mathbb{E}\left[S^1([rt], t)\right] - \mathbb{E}\left[S^2([rt], t)\right] = \mathbb{P}\left[X^{(i)}(t) > [rt]\right] \tag{2.16}$$

since we can interpret the discrepancy between η^1 and η^2 as second class particle (see Section 2.2.3). This works for all models R0, R1, R2 and R3. The second coupling we want to use is called particle-particle coupling. We label the particles in η_0^1 and η_0^2 from right to left and let particles with the same label jump at the same time. Then under this coupling

$$\mathbb{E}\left[S^1([rt], t)\right] - \mathbb{E}\left[S^2([rt], t)\right] = \mathbb{P}\left[\eta_t^1([rt] + 1) = 1\right] \tag{2.17}$$

in the models R0, R1 and R2. By Theorem 2.6 the right hand side of (2.17) converges to $f_i(r)$, so together with (2.16) we have

$$\mathbb{P}\left[X^{(i)}(t) > [rt]\right] \xrightarrow[t\to\infty]{} f_i(r)$$

which proves convergence in distribution for $i = 0, 1, 2$. However, in the model R3 we cannot use the particle-particle coupling as before because this is no longer a real coupling. If we let particles with the same label jump at the same time in η^1 and η^2 then the dynamics of the η^2 process are different from the η^1 process: In η^2 we update odd sites first and then even sites. We denote by $\mathbb{E}_{PP}$ the expectation in η^1 and η^2 if particles with the same label jump at the same time in η^1 and η^2 where we update in such a way that η^1 is still a TASEP with update rule R3. Then we have

$$\mathbb{E}_{PP}\left[S^1([rt],t)\right] = \mathbb{E}\left[S^1([rt],t)\right].$$

Notice that starting the second process with updating even sites does not change anything as there is no particle on an even site with an adjacent empty site in the initial configuration. If we remove the last update of even sites at time t this changes the value of $S^2([rt],t)$ if there is a jump from site $[rt]$ to site $[rt]+1$ during this update. But there can only be a jump from $[rt]$ to $[rt]+1$ while updating the even sites if $[rt]$ is even. Let us therefore consider odd sites $2[\frac{rt}{2}]+1$ first. We get

$$\mathbb{E}_{PP}\left[S^2\left(2\left[\frac{rt}{2}\right]+1,t\right)\right] = \mathbb{E}\left[S^2\left(2\left[\frac{rt}{2}\right]+1,t\right)\right] \tag{2.18}$$

and we still have

$$\mathbb{E}_{PP}\left[S^1\left(2\left[\frac{rt}{2}\right]+1,t\right)\right] - \mathbb{E}_{PP}\left[S^2\left(2\left[\frac{rt}{2}\right]+1,t\right)\right] = \mathbb{P}\left[\eta_t^1\left(2\left[\frac{rt}{2}\right]+2\right)=1\right]$$

as before. Hence we get

$$\begin{aligned}\mathbb{P}\left[X^{(3)}(t) > 2\left[\frac{rt}{2}\right]+1\right] &= \mathbb{E}\left[S^1\left(2\left[\frac{rt}{2}\right]+1,t\right)\right] - \mathbb{E}\left[S^2\left(2\left[\frac{rt}{2}\right]+1,t\right)\right]\\ &= \mathbb{P}\left[\eta_t^1\left(2\left[\frac{rt}{2}\right]+2\right)=1\right]\\ &\xrightarrow[t\to\infty]{} a_3(r)\end{aligned}$$

by the convergence to local equilibrium (Theorem 2.6). But by monotonicity we have

$$\mathbb{P}\left[X^{(3)}(t) > 2\left[\frac{rt}{2}\right]-1\right] \geq \mathbb{P}\left[X^{(3)}(t) > 2\left[\frac{rt}{2}\right]\right] \geq \mathbb{P}\left[X^{(3)}(t) > 2\left[\frac{rt}{2}\right]+1\right]$$

Hence

$$\lim_{t\to\infty}\mathbb{P}\left[\frac{X^{(3)}(t)}{t} > r\right] = a_3(r).$$

Remark 2.32. *If the second class particle starts on an odd site (with first class particles to the left and holes/third class particles to the right) then*

$$\frac{\widetilde{X}^{(3)}(t)}{t} \xrightarrow[t\to\infty]{a.s.} \widetilde{U}^{(3)}$$

and $\widetilde{U}^{(3)}$ *has distribution function* $\frac{a_3(1-\beta)}{1-a_3\beta}$ *accordingly.*

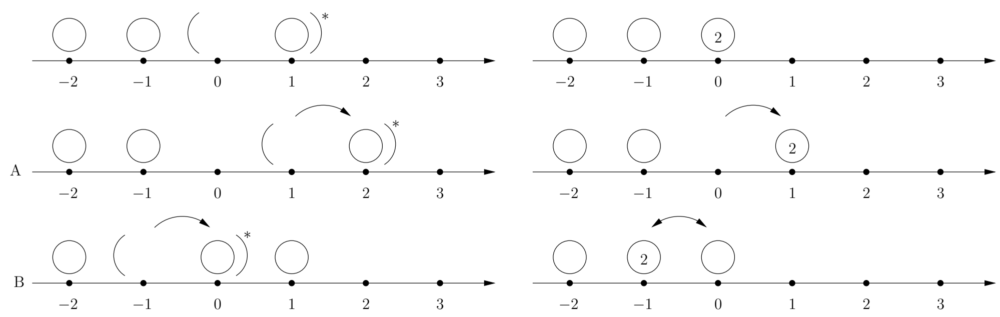

Figure 2.7: Pair representation of the second class particle: The figures on the left show the system with the pair; the figures on the right show the corresponding multi-type system

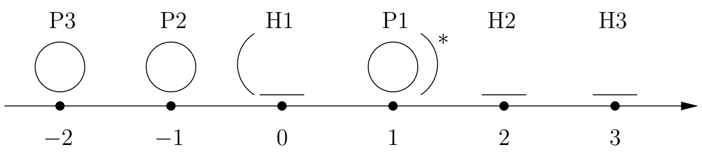

Figure 2.8: Pair representation of the second class particle with particles labelled from right to left and holes labelled from left to right

Now we want to prove almost sure convergence of the speed of a second class particle. Our methods follow the approach in [19]. The idea is to establish a connection between the path of the second class particle and a competition interface in the corresponding growth model. The cluster in the growth model can be divided into two clusters corresponding to events happening to the right and to the left of the second class particle. The interface between these two clusters is called the competition interface. Using results about semi-infinite geodesics it can be shown that this competition interface has almost surely an asymptotic direction. This can be used to deduce that

the second class particle has almost surely an asymptotic speed and since we know the distribution of this speed we can also calculate the distribution of the random angle of the competition interface. In the following we will describe this method first for the TASEP in continuous time, as given in [19], and then explain the adjustments that have to be made in the TASEPs in discrete time.
In order to establish the connection between the second class particle and the competition interface we represent the second class particle as a pair consisting of a hole and a particle. This reduces our multi-type model to a model consisting only of particles and holes and allows us to use the connection to last-passage percolation developed in Section 2.2.2. We let the pair move as follows: If the particle of the pair jumps to the right the pair moves to the right (A) and if a particle jumps from the left into the hole of the pair the pair moves to the left (B), see Figure 2.8. Then the pair behaves indeed like a second class particle. If we label the particles from right to left and the holes from left to right as in Figure 2.8 then we can consider the process $(\varphi_n)_{n\geq 0}$ giving the labels of the pair after the nth jump involving the pair. We have $\varphi_0 = (1,1)$, as initially the pair consists of hole 1 and particle 1, and $\varphi_{n+1} - \varphi_n \in \{(0,1),(1,0)\}$. φ_n satisfies the recursion formula

$$\varphi_{n+1} = \begin{cases} \varphi_n + (1,0) & T(\varphi_n + (1,0)) < T(\varphi_n + (0,1)) \\ \varphi_n + (0,1) & T(\varphi_n + (1,0)) > T(\varphi_n + (0,1)) \end{cases} \tag{2.19}$$

If the first label increases the second class particle moves one step to the right and if the second label increases the second class particle moves one step to the left.

Remark 2.33. *Note that by inserting an extra site we changed the parity of the sites to the right of and including the particle of the pair: The first hole to the right of the second class particle is at an odd site while the first hole to the right of the pair is at an even site. The second class particle itself is at an even site while the particle in the pair is at an odd site. This will be important for model R3.*

Now we want to define geodesics in the last-passage percolation model to define a competition interface in the growth model. For $z, z' \in \mathbb{Z}_+^2$ the heaviest increasing path from z to z' (i.e. the path that achieves the maximum in $R(z, z')$) is called the geodesic from z to z'. Note that in the model in continuous time geodesics are unique. (In the models in discrete time we will need a rule to break ties to achieve uniqueness of the geodesics). A semi-infinite geodesic starting at z is a path $\pi = (z, z_1, z_2, \ldots)$ in $\mathbb{Z}_+^2$ such that for every $z' = z_k, z'' = z_l \in \pi$ the geodesic from z' to z'' is $(z_k, z_{k+1}, z_{k+2}, \ldots, z_l) \subset \pi$. For $\alpha \in [0, 90°]$ a α-geodesic is a semi-infinte geodesic with direction α, i.e. a geodesic $\pi = (z, z_1, z_2, \ldots)$ such that

$$\lim_{n\to\infty} \frac{z_n}{|z_n|} = (\cos\alpha, \sin\alpha)\,.$$

Now we colour every block $Q(i,j) = (i-1, i] \times (j-1, j]$ in $(\mathbb{R}_+)^2 \backslash [0,1]^2$ either red if the geodesic from $(1,1)$ to (i,j) passes through $(1,2)$ or blue if it passes through $(2,1)$. The interface between these clusters is called the competition interface and an induction argument together with the recursion (2.19) shows that it is given by the process φ_n, see Proposition 3 in [19]. Using results about the existence and uniqueness of α-geodesics (Propositions 7,8 and 9 in [19]) it can be shown that φ_n has almost surely an asymptotic direction and we can conclude that the second class particle has almost surely an asymptotic speed (Propositions 4 and 5 in [19]). Now we want to apply these methods to the discrete time TASEPs R1, R2 and R3.

R1

The last-passage percolation model for rule R1 was described in Section 2.2.2, and in particular just after (2.3). Now to adapt to the initial configuration

$$\eta_0(x) = \begin{cases} 1 & x \le -1 \\ 0 & x = 0 \\ 1 & x = 1 \\ 0 & x \ge 2 \end{cases}$$

we have to remove the '+1' weight from the edge between $(1,1)$ and $(2,1)$ and the weight from the vertex $(1,1)$ as in the initial configuration we are considering particle 1 has already jumped over hole 1. We colour $Q(1,2)$ red, $Q(2,1)$ blue and every other block $Q(i,j)$ in $(\mathbb{R}_+)^2\setminus[0,1]^2$ either red if

$$\widetilde{R}^{(1)}((1,2),(i,j)) > \widetilde{R}^{(1)}((2,1),(i,j))$$

and blue if

$$\widetilde{R}^{(1)}((1,2),(i,j)) \leq \widetilde{R}^{(1)}((2,1),(i,j))$$

(recall the definition of R in (2.2); $\widetilde{R}^{(1)}$ is the corresponding quantity in model R1 with the changes mentioned above). This implies that if $Q(i,j+1)$ is red and $Q(i+1,j)$ is blue, then $Q(i+1,j+1)$ is red iff

$$\widetilde{T}^{(1)}(i,j+1) \geq \widetilde{T}^{(1)}(i+1,j) \tag{2.20}$$

and blue iff

$$\widetilde{T}^{(1)}(i,j+1) < \widetilde{T}^{(1)}(i+1,j) \tag{2.21}$$

where $\widetilde{T}^{(1)}$ is defined as in (2.3) but now in the model R1 with the modifications described above. The line $\left(\varphi_n^{(1)}, n \geq 0\right)$ separating the two clusters is again called the competition interface and due to the way we defined the red and blue cluster we have again that the competition interface corresponds to the path of the second class particle. We can rewrite (2.20) and (2.21) in terms of $\varphi_n^{(1)}$ as

$$\varphi_{n+1}^{(1)} = \begin{cases} \varphi_n^{(1)} + (1,0) & \widetilde{T}^{(1)}(\varphi_n^{(1)} + (1,0)) \leq \widetilde{T}^{(1)}(\varphi_n^{(1)} + (0,1)) \\ \varphi_n^{(1)} + (0,1) & \widetilde{T}^{(1)}(\varphi_n^{(1)} + (1,0)) > \widetilde{T}^{(1)}(\varphi_n^{(1)} + (0,1)) \end{cases} \tag{2.22}$$

Note the similarity of (2.19) with (2.22); the difference comes from the fact that in the models in discrete time ties are possible. As in [19] we want to prove that this competition interface has an asymptotic direction almost surely. First we note that the results in [19] about geodesics still hold with geometric weights attached to the vertices instead of exponential weights. Secondly, the results still hold in a model where '+1' weights are attached to *every* horizontal edge in $\mathbb{Z}_+^2$ as these weights do not affect the geodesics (they just give a constant weight to every path from z to z', $z, z' \in \mathbb{Z}_+^2$). The only difference in our case is that there is no weight attached to the edge from $(1, 1)$ to $(2, 1)$. But this local change does not affect the almost sure statements in Propositions 7, 8 and 9 in [19]. We conclude that the competition interface in our model has an asymptotic direction almost surely and it follows from arguments analogous to the ones in [19] that the speed of the second class particle converges almost surely.

R2

The result for model R2 follows again from the symmetry between R1 and R2.

R3

For the purpose of this section it is convenient to change the last-passage percolation model corresponding to model R3 a little bit. Instead of updating the even sites and then the odd sites during a single time-step, we separate the two batches of updates by a half time-step. Then the percolation model corresponding to the initial configuration

$$\eta_0(x) = \begin{cases} 1 & x \leq 0 \\ 0 & x \geq 1 \end{cases}$$

has no weights attached to the edges, while the weights at the vertices are geometric with an extra $\frac{1}{2}$ added. As noticed in the beginning of this section, introducing an extra site into this model changes the parity of some sites. In order to deal with this we attach a single '$+\frac{1}{2}$' weight to the edge from $(1, 1)$ to $(1, 2)$ in the percolation model corresponding to the initial configuration

$$\eta_0(x) = \begin{cases} 1 & x \leq -1 \\ 0 & x = 0 \\ 1 & x = 1 \\ 0 & x \geq 2 \end{cases}$$

(where we also remove the weight from the vertex $(1,1)$). This ensures that we apply even/odd updates to the left of the particle of the pair and odd/even updates to the right. Then the movement of the pair $\varphi_n^{(3)}$ corresponds to the movement of the second class particle in a model with even/odd updates.

Again we colour $Q(1,2)$ red, $Q(2,1)$ blue and now every other block $Q(i,j)$ in $(\mathbb{R}_+)^2 \setminus [0,1]^2$ either red if

$$\widetilde{R}^{(3)}((1,2),(i,j)) \geq \widetilde{R}^{(3)}((2,1),(i,j)) \tag{2.23}$$

and blue if

$$\widetilde{R}^{(3)}((1,2),(i,j)) < \widetilde{R}^{(3)}((2,1),(i,j)). \tag{2.24}$$

The interface between these clusters is again the competition interface and is given by $\varphi_n^{(3)}$. In terms of $\varphi_n^{(3)}$ we get from (2.23) and (2.24) that

$$\varphi_{n+1}^{(3)} = \begin{cases} \varphi_n^{(3)} + (1,0) & \widetilde{T}^{(3)}(\varphi_n^{(3)} + (1,0)) < \widetilde{T}^{(3)}(\varphi_n^{(3)} + (0,1)) \\ \varphi_n^{(3)} + (0,1) & \widetilde{T}^{(3)}(\varphi_n^{(3)} + (1,0)) > \widetilde{T}^{(3)}(\varphi_n^{(3)} + (0,1)) \end{cases} \tag{2.25}$$

$\widetilde{R}^{(3)}$ and $\widetilde{T}^{(3)}$ are defined as in (2.2) and (2.5) but now we are considering the model R3 with the modifications described above. Due to the additional '$+\frac{1}{2}$' weight on the edge from $(1,1)$ to $(1,2)$ the competition interface never encounters any ties in this model, i.e. $\widetilde{T}^{(3)}(\varphi_n^{(3)} + (1,0)) = \widetilde{T}^{(3)}(\varphi_n^{(3)} + (0,1))$ does not occur. As in R1, the local change given by the extra '$+\frac{1}{2}$' edge-weight in this model does not change the almost sure statements in Propositions 7, 8 and 9 in [19]. The rest of the argument is the same as for model R1. □

Remark 2.34. *We can use the known distributions of the speeds of the second class particles (see Theorem 2.9) together with the hydrodynamic limit results (see Theorem 2.6 and Remark 2.7) to prove the interesting result mentioned in Remark 2.10 that the distributions of the asymptotic direction of the competition interfaces in the models R1 and R3 are the same:*

Proof. The proof exploits the connections made in Proposition 5 in [19]. Similar to [19] we let $\psi_t^{(1)} = (I^{(1)}(t), J^{(1)}(t))$, $\psi_t^{(3)} = (I^{(3)}(t), J^{(3)}(t))$ be the position of the competition interface (i.e. the labels of the pair) at time t and denote by $\theta^{(1)}, \theta^{(3)} \in [0, 90°]$ the random angle of the competition interface for model R1 and R3, i.e.

$$\lim_{t\to\infty} \frac{\psi_t^{(i)}}{\left|\psi_t^{(i)}\right|} = \left(\cos(\theta^{(i)}), \sin(\theta^{(i)})\right) \text{ for } i = 1, 3.$$

By the arguments in the previous sections we know that these limits exist almost surely. Using the asymptotic shape of the growth models given in Remark 2.7 we also have

$$\lim_{t\to\infty} \frac{\psi_t^{(i)}}{t} = j^{(i)}(\theta^{(i)}) \left(\cos(\theta^{(i)}), \sin(\theta^{(i)})\right) \text{ a.s. for } i = 1, 3$$

where $j^{(i)}(\theta^{(i)})$ is the distance from the origin to the intersection of the line given by $\{(u, v) \in \mathbb{R}_+^2 : \tan(\theta^{(i)}) = \frac{v}{u}\}$ and the asymptotic *growth interface* $(x, g_i(x))$ for $i = 1, 3$. With the formulas for c_1 and c_3 in (2.9) and (2.11) we can calculate $j^{(1)}(\theta^{(1)})$ and $j^{(3)}(\theta^{(3)})$ explicitly:

$$j^{(1)}(\theta^{(1)}) = \frac{\beta}{\left(\sqrt{(1-\beta)\sin(\theta^{(1)})} + \sqrt{\cos(\theta^{(1)})}\right)^2}$$

and

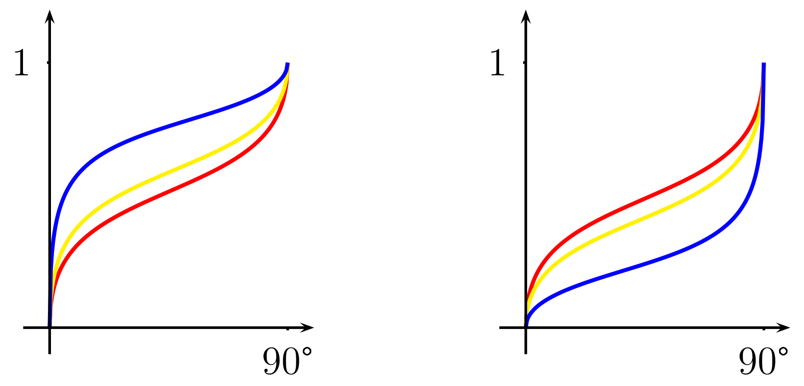

Figure 2.9: Distribution function for the random angle of the competition interface in the models R1 and R3 (left) and R2 (right) for $\beta = 0.1$ (red), $\beta = 0.5$ (yellow) and $\beta = 0.9$ (blue)

$$j^{(3)}(\theta^{(3)}) = \frac{2\beta\,(2-\beta)}{\left((2-\beta)\,\sqrt{\sin(\theta^{(3)})} + 2\sqrt{(1-\beta)\cos(\theta^{(3)})}\right)^2 + \beta^2\cos(\theta^{(3)})}.$$

By the connection between the path of the second class particle and the competition interface ($X^{(i)}(t) = I^{(i)}(t) - J^{(i)}(t)$, $i = 1, 3$, since the second class particle moves to the right iff the first label of the pair increases and to the left iff the second label of the pair increases) it follows that

$$\lim_{t\to\infty} \frac{X^{(i)}(t)}{t} = \lim_{t\to\infty} \frac{I^{(i)}(t) - J^{(i)}(t)}{t} = j^{(i)}(\theta^{(i)})\left(\cos(\theta^{(i)}) - \sin(\theta^{(i)})\right) \stackrel{\text{def}}{=} l_i(\theta^{(i)})$$

almost surely for $i = 1, 3$. Using the known distributions of the speed of the second class particle in model R1 and R3 (see Theorem 2.9) we can calculate

$$\mathbb{P}\left[\theta^{(1)} \le \alpha\right] = \mathbb{P}\left[l_1(\theta^{(1)}) \ge l_1(\alpha)\right] = f_1(l_1(\alpha))$$

and

$$\mathbb{P}\left[\theta^{(3)} \le \alpha\right] = \mathbb{P}\left[l_3(\theta^{(3)}) \ge l_3(\alpha)\right] = a_3(l_3(\alpha)).$$

A calculation shows that

$$f_1(l_1(\alpha)) = a_3(l_3(\alpha)).$$

□

Proof of Theorem 2.13: Let the operators σ_n be defined by

$$\sigma_n Y = \begin{cases} \tau_n Y & Y_n < Y_{n+1} \\ Y & \text{otherwise} \end{cases}$$

where τ_n exchanges Y_n and Y_{n+1} in Y. The proof of the following general lemma is the same as in [1] (Lemma 3.1).

Lemma 2.35. *For a fixed sequence $i_1, \ldots, i_k$ in $\mathbb{Z}$ we have*

$$\sigma_{i_k} \cdots \sigma_{i_1} \stackrel{d}{=} (\sigma_{i_1} \cdots \sigma_{i_k})^{-1}.$$

We have the relation $Y_{X_n(m)}(m) = n$ between the X and the Y process. Since $\beta < 1$ each site has a positive probability that no jump occurs at that site at any given time. At each time-step these sites separate $\mathbb{Z}$ into finite intervals and the events on these intervals during that time-step are independent. In the model with updates from right to left we apply a finite sequence of operators $\sigma_{i_1} \cdots \sigma_{i_k}$ where $i_1, \ldots, i_k$ is an increasing sequence (since we update from right to left). Lemma 2.35 states that applying $\sigma_{i_1} \cdots \sigma_{i_k}$ is the same (in distribution) as applying $\sigma_{i_k} \cdots \sigma_{i_1}$ (i.e. updating from left to right) and taking the inverse permutation. But given the configuration $Y_n(0)$ and performing the updates from left to right we get $Y_n^{(2)}(1)$ and this is the inverse permutation of $X_n^{(2)}(1)$. So

$$Y_n^{(1)}(1) \stackrel{d}{=} X_n^{(2)}(1).$$

Inductively we get that this holds for all $m > 0$. The other three parts follow in the same way. In the case with even/odd updates, $i_1, \ldots, i_k$ is a sequence such that there exists a $1 \leq j \leq k+1$ such that i_l is odd for $l < j$ and i_l is even for $l \geq j$. □

In order to prove Theorem 2.12 we will need the following lemma. We state it here for the model R1, but analogous results hold for the other models as well. The lemma corresponds to Lemma 4.1 in [1].

Lemma 2.36. *Consider two TASEPs, $Y^{(1)}$ and $\widetilde{Y}^{(1)}$, defined as functions of the same Bernoulli points on $\mathbb{Z} \times \mathbb{N}$ (i.e. under basic coupling). We set $Y_n^{(1)}(0) = n$ and $\widetilde{Y}_n^{(1)}(0) = \sigma_j \cdots \sigma_{j+k} Y_n^{(1)}(0)$ for some $j \in \mathbb{Z}$ and $k \geq 0$. Let $\{U_n^{(1)}\}$ be the speed process of $Y^{(1)}$ and $\{\widetilde{U}_n^{(1)}\}$ be the speed process of $\widetilde{Y}^{(1)}$. Then $\widetilde{U}^{(1)} = \sigma_{j+k} \cdots \sigma_j U^{(1)}$.*

Proof. Every particle other than $\{j, \dots j+k+1\}$ is either stronger than all particles $\{j, \dots, j+k+1\}$ or weaker than all particles $\{j, \dots, j+k+1\}$. Any swap of a particle other than $\{j, \dots, j+k+1\}$ will happen in both $Y^{(1)}$ and $\widetilde{Y}^{(1)}$. So for any $i \notin \{j, \dots, j+k+1\}$ we have $X_i^{(1)}(m) = \widetilde{X}_i^{(1)}(m)$ for all $m \geq 0$ and therefore $U_i^{(1)} = \widetilde{U}_i^{(1)}$ for those i. In $\widetilde{Y}^{(1)}$ particle $j+k+1$ is always to the left of all other particles $\{j, \dots, j+k\}$. So $\widetilde{U}_{j+k+1}^{(1)} = \min\{U_j^{(1)} \dots U_{j+k+1}^{(1)}\}$. Define $j \leq r \leq j+k+1$ by

$$\min\left\{U_j^{(1)} \dots U_{j+k+1}^{(1)}\right\} = U_r^{(1)}.$$

Then for $i = r, r+1, \dots, j+k$ we have $\widetilde{U}_i^{(1)} = U_{i+1}^{(1)}$ and for $i = j, j+1, \dots, r-1$

$$\begin{aligned}
\widetilde{U}_j^{(1)} &= \max\left\{U_j^{(1)}, U_{j+1}^{(1)}\right\} \\
\widetilde{U}_{j+1}^{(1)} &= \max\left\{\min\left\{U_j^{(1)}, U_{j+1}^{(1)}\right\}, U_{j+2}^{(1)}\right\} \\
\widetilde{U}_{j+2}^{(1)} &= \max\left\{\min\left\{\min\left\{U_j^{(1)}, U_{j+1}^{(1)}\right\}, U_{j+2}^{(1)}\right\}, U_{j+3}^{(1)}\right\} \\
&\dots
\end{aligned}$$

This shows that $\widetilde{U}^{(1)} = \sigma_{j+k} \cdots \sigma_j U^{(1)}$. □

Proof of Theorem 2.12: Consider a Bernoulli process on $\mathbb{Z} \times \mathbb{Z}$. Half of this process ($\mathbb{Z} \times \mathbb{N}$) is used to construct the TASEP $Y^{(1)}$. For any $l \in \mathbb{Z}$ we can translate the Bernoulli process by l (i.e. take points of the form $(n, m+l)$ where (n, m) is

in the original process). We can restrict this translated process to $\mathbb{Z} \times \mathbb{N}$ and use this restricted process to construct another TASEP. Let $U^{(1)}(l) = \{U_n^{(1)}(l)\}$ be the speed process for the TASEP that has been constructed using the Bernoulli process translated by l. For every l, $U^{(1)}(l)$ has distribution $\mu^{(1)}$. So we have to show that $\{U_n^{(1)}(l)\}$ behaves like a TASEP with updates from left to right. In order to do this we look at a transition $\{U_n^{(1)}(l)\} \to \{U_n^{(1)}(l+1)\}$. The effect on the original TASEP of changing from translating by l to translating by $l+1$ is that some finite sequences of σ operators of the form $\sigma_j \cdots \sigma_{j+k}$ are added to be applied to the TASEP before the original sequence of operations. At each location a σ operator is added with probability β. Lemma 2.36 shows, that applying each of these finite sequences has the same effect on the speeds as applying each sequence in reverse order to the speed process. This shows that $\{U_n^{(1)}(l)\}$ behaves like a TASEP with updates from left to right and therefore the measure $\mu^{(1)}$ is stationary for the TASEP with updates from left to right.
The proofs for the other three models are essentially the same (using the appropriate versions of Lemma 2.36). $\square$

The following lemma will allow us to do the explicit calculations for the joint densities of the speeds in Theorems 2.15 - 2.19 using the connection between queueing models and the invariant measures introduced in Theorem 2.3. Here $D(g_i)$ is the domain of the distribution function g_i $(i = 0, 1, 2, 3, 4)$.

Lemma 2.37. *If $F : D(g_i) \to \{1, \dots, N\}$ is a non-decreasing function then for the TASEP $\{Y_n^{(i)}(m)\}_{n,m}$ the distribution of $\{F(U_n^{(i)})\}_n$ is the unique ergodic stationary measure of the multi-type TASEP model $Rj(i)$ with types $\{1, \dots, N\}$ and densities $\lambda_l = g_i(\sup\{F^{-1}(l)\}) - g_i(\inf\{F^{-1}(l)\})$ for type $l = 1, \dots, N$ (j as in Theorem 2.12).*

Proof. The proof is analogous to the proof of Corollary 5.4 in [1]. $\square$

With the help of this lemma we can do all the calculations needed for the results in Theorems 2.15 - 2.19. Depending on the model we are considering we will choose the function F from Lemma 2.37 to be $F_i = \min\{j : g_i(u) < x_j\}$ for some increasing sequence $(x_1, \ldots, x_{N-1})$ in $[0,1]$. Then we put $V_n = F(U_n)$ and the distribution of the Vs is given by the invariant measure for the multi-type models and can be calculated explicitly using the queueing representation.

Proof of Theorem 2.15: By Lemma 2.37 we have with $N = 3$ that $F_1(U_n)$ is distributed according to the unique ergodic stationary measure of a 3-type TASEP with updates from left to right and densities

$$\lambda_1 = \mathbb{P}\left[x_0 < g_1(U_n) < x_1\right] = x_1 - x_0 = x_1$$
$$\lambda_2 = \mathbb{P}\left[x_1 < g_1(U_n) < x_2\right] = x_2 - x_1$$
$$\lambda_3 = \mathbb{P}\left[x_2 < g_1(U_n) < x_3\right] = x_3 - x_2 = 1 - x_2$$

λ_1 is the density of first class particles, λ_2 is the density of second class particles and λ_3 is the density of third class particles (or holes). Recall that $V_n = F_1(U_n)$. Using the queueing representation for the unique ergodic stationary measure of a 3-type TASEP we can calculate the joint distribution (V_0, V_1) explicitly. This distribution depends on the x_i. Taking suitable derivatives with respect to these x_i we get the density of the corresponding speeds. We have for example

$$\begin{aligned}\mathbb{P}\left[U_0 < g_1^{-1}(x_1) < U_1 < g_1^{-1}(x_2)\right] &= \mathbb{P}\left[V_0 = 1, V_1 = 2\right] \\ &= x_1 x_2 \left(x_2 - x_1\right)\end{aligned}$$

since the probability of having a second class particle at position 1 is $x_2 - x_1$ (since this is the density of second class particles) and to have a first class particle at position 0 we then have to have an arrival (probability x_1) and a service (probability x_2) because having a second class particle at site 1 means that the queue was empty at that time (so in order to have a departure at site 0 we need an arrival at site 0). Remember that in Theorem 2.3 the particles in the TASEP jumped from the right to the left. If we want to consider the TASEP with jumps from the left to the right (and that is what we are doing here) we have to read the queues from right to left. So position 1 comes before position 0 and the probability of having a second class particle at position 1 is independent of arrivals and services at position 0. So for $u_0 < u_1$ we put $x_1 = g_1(u_0)$ and $x_2 = g_1(u_1)$ and get as density

$$\begin{aligned}\mathbb{P}\left[U_0 \in du_0, U_1 \in du_1\right] &= \frac{dx_1}{du_0}\frac{dx_2}{du_1}\frac{d}{dx_1}\frac{d}{dx_2} x_1 x_2 \left(x_2 - x_1\right) \\ &= \frac{1-\beta}{4\beta^2}\left(1-u_0\right)^{-\frac{3}{2}}\left(1-u_1\right)^{-\frac{3}{2}}\left(2g_1(u_1) - 2g_1(u_0)\right) \\ &= \frac{1-\beta}{2\beta^3}\left(1-u_0\right)^{-\frac{3}{2}}\left(1-u_1\right)^{-\frac{3}{2}} \\ &\qquad \cdot\left(\sqrt{\frac{1-\beta}{1-u_1}} - \sqrt{\frac{1-\beta}{1-u_0}}\right)\end{aligned}$$

Similarly, we have

$$\begin{aligned}\mathbb{P}\left[g_1^{-1}(x_1) < U_1 < g_1^{-1}(x_2) < U_0\right] &= \mathbb{P}\left[V_0 = 3, V_1 = 2\right] \\ &= \left(1-x_2\right)\left(x_2 - x_1\right)\end{aligned}$$

and therefore we get as density for $u_0 > u_1$ (putting $x_1 = g_1(u_1)$, $x_2 = g_1(u_0)$)

$$\begin{aligned}\mathbb{P}\left[U_0 \in du_0, U_1 \in du_1\right] &= \left(-\frac{dx_1}{du_0}\right)\left(-\frac{dx_2}{du_1}\right)\frac{d}{dx_1}\frac{d}{dx_2}\left(1-x_2\right)\left(x_2-x_1\right) \\ &= \frac{1-\beta}{4\beta^2}\left(1-u_0\right)^{-\frac{3}{2}}\left(1-u_1\right)^{-\frac{3}{2}} \\ &= g_1^{'}(u_0) g_1^{'}(u_1).\end{aligned}$$

To get the density for $u_0 = u_1$ we consider

$$\begin{aligned}\mathbb{P}\left[g_1^{-1}(x_1) < U_0, U_1 < g_1^{-1}(x_2)\right] &= \mathbb{P}\left[V_0 = 2, V_1 = 2\right] \\ &= \left(1-x_1\right) x_2 \left(x_2 - x_1\right)\end{aligned}$$

and let $x_1, x_2 \to g_1(u)$. We get

$$\mathbb{P}\left[U_0, U_1 \in du\right] = \lim_{x_1, x_2 \to g_1(u)} \frac{\left(1-x_1\right) x_2 \left(x_2 - x_1\right)}{g_1^{-1}(x_2) - g_1^{-1}(x_1)}$$

$$
\begin{aligned}
&= \frac{\sqrt{1-\beta}\,(1-g_1(u))\,g_1(u)}{2\beta\,(1-u)^{\frac{3}{2}}} \\
&= \frac{\sqrt{1-\beta}}{2\beta^2\,(1-u)^{\frac{3}{2}}}\left(1-\frac{1}{\beta}\right)+\frac{1-\beta}{2\beta^2\,(1-u)^2}\left(\frac{2}{\beta}-1\right) \\
&\qquad -\frac{\sqrt{1-\beta}\,(1-\beta)}{2\beta^3\,(1-u)^{\frac{5}{2}}}.
\end{aligned}
$$

To get the probabilities in Theorem 2.15 we only have to integrate the densities over the appropriate ranges of u_0, u_1 and u. Alternatively we can use the following:

$$
\begin{aligned}
\mathbb{P}\left[U_0^{(1)} < U_1^{(1)}\right] &= \mathbb{P}\left[g_1(U_0^{(1)}) < g_1(U_1^{(1)})\right] \\
&\overset{(*)}{=} \mathbb{P}\left[g_0(U_0^{(0)}) < g_0(U_1^{(0)})\right] \\
&= \mathbb{P}\left[U_0^{(0)} < U_1^{(0)}\right] \\
&= \frac{1}{3}
\end{aligned}
$$

$(*)$ follows from the fact that the distribution of $\{g_1(U_n^{(1)})\}$ is the unique translation invariant stationary ergodic measure for the TASEP R2 with marginals uniform on $[0,1]$. The distribution of $\{g_0(U_n^{(0)})\}$ is the unique translation invariant stationary ergodic measure for the TASEP in continuous time. Since the stationary distributions for the multi-type TASEPs R0 and R2 are the same (see Theorem 2.3) $\{g_1(U_n^{(1)})\}$ has the same distribution as $\{g_0(U_n^{(0)})\}$. □

The proofs for Theorems 2.17 - 2.19 work in exactly the same way.

2.5 Fully parallel updates

Finally we mention the model with "fully parallel updates". If an update occurs at site x at time t (which happens with probability β as usual), this update causes a jump from x to $x+1$ if and only if $\eta_{t-1}(x)=1$ and $\eta_{t-1}(x+1)=0$ (that is, the jump is already possible before any other updates at the current time-step are performed).

There are several important differences between this model and the models we have studied earlier. The Bernoulli product measures ν_ρ are no longer invariant. Furthermore, the basic coupling no longer preserves an ordering between different initial configurations, and so it is no longer clear how to define a multi-class system. If we use basic coupling to couple two systems which start with one discrepancy at the origin then this single discrepancy can generate additional discrepancies. It would already be interesting to know how the leftmost and rightmost discrepancies behave. Do they have asymptotic speeds, and if so are the speeds random or deterministic? There is still a natural percolation representation, and we can still obtain a hydrodynamic limit result in the sense that $\frac{1}{n}\sum_{un<k<vn}\eta_t(n)$ converges a.s. to the constant value $\int_u^v f(w)dw$, for $u < v$ and some function f, but the stronger result that $\lim_{n\to\infty}\mathbb{E}\left[\eta_n(k)\right]$ exists and is equal to $f(u)$ whenever $\frac{k}{n}$ tends to u does not follow using the same methods as in the proof of Theorem 2.6.

Chapter 3

Long-range last-passage percolation on the line

3.1 Introduction

As described in the introduction and seen in Section 2.2.2, the TASEP has strong connections to last-passage percolation. In this chapter we study the one-dimensional random graph model described in the introduction with general weight distributions. The equivalent model with constant weights was studied in [20] and [14]. Many features of the model with constant weights remain if the random weights have a sufficiently light tail. In the case where the weights have finite variance, we prove a strong law of large numbers for the passage time $w_{0,n}$, and under the stronger assumption of a finite third moment we give a functional central limit theorem. The strategy of the proof is similar to that of [14]; we construct a renewal process with the property that no geodesic uses an edge which crosses a renewal point. In this way much of the analysis can be carried out by considering the behaviour of the length and weight of individual renewal intervals. In [14], the definition of renewal points could be made rather simply in terms of the connectivity properties of the graph, but here the disorder induced by the random weights requires a rather more intricate construction. We then define a set of auxiliary random variables to construct an upper bound for Γ_0, the first renewal point to the right of the origin, and use them to conclude that $\mathbb{E}[\Gamma_0] < \infty$ if $\mathbb{E}[v^3] < \infty$. This provides a bound on the second moment of the length of a typical renewal interval, and we can deduce that the variance that appears in the central limit theorem is finite.

This regenerative structure means that the problem retains an essentially one-dimensional nature. Most of the edges used in an optimal path are short, and the behaviour is qualitatively the same as one would see in a model with edges of bounded length. We find an entirely different behaviour when $\mathbb{E}[v^2] = \infty$. Now the passage time $w_{0,n}$ grows super-linearly in n, and the dominant contribution to the passage time is given by the weights of edges whose length is on the order of n. The appropriate comparison with a bounded-length model is now with a *two-dimensional* last-passage percolation problem. Under the assumption of a regularly varying tail, we prove scaling laws and asymptotic distributions in terms of a "continuum long-range last-passage percolation" model on the interval $[0, 1]$. The construction is closely related to that used by Hambly and Martin [26], who studied (nearest-neighbour) last-passage percolation in two dimensions with heavy-tailed weights. There are interesting relationships between such models and the theory of random matrices with heavy-tailed entries (see for example [6, 5]).

The difference between the behaviour of the model in the cases $\mathbb{E}[v^2] < \infty$ and $\mathbb{E}[v^2] = \infty$ can be seen in the simulations in Figure 3.1 in Section 3.2.2.

The chapter is organized in the following way: in Section 3.2 we will present the main results. We will split the results up into two main sections: one for the case where the weights have a second moment (Section 3.2.1) and one where they do not (Section 3.2.2). The main results in Section 3.2.1 are the strong law of large numbers and a functional central limit theorem for the random variable $w_{0,n}$ giving the weight of the heaviest path from 0 to n. Another main result (Theorem 3.6) describes scaling laws for the length of the longest edge and the weight of the heaviest edge used on the maximizing path from 0 to n, and we can use these results to conclude that in certain situations where $\mathbb{E}[v^3] = \infty$ a central limit theorem cannot hold.

The main results for the case $\mathbb{E}[v^2] = \infty$ are then given in Section 3.2.2, along with simulations illustrating the scaling limit and the difference in behaviour from the case of weights with finite variance.

The proofs for the model with finite second moments can then be found in Section 3.3. There we will also briefly discuss a variant of the model where the edge probabilities are not constant, but depend on the lengths of the edges (Section 3.3.5). This model was studied in [14] but without random weights. The proofs for the model with infinite second moments will be given in Section 3.4; the structure is closely related to that of the corresponding results in [26].

3.2 Main results

In this section we introduce the main results. Like the rest of the chapter, the results will be split up into two parts: results about the model with finite second moments and results about the model with infinite second moment.

3.2.1 Weights with finite second moment

Here we present the results for the model where the weights $v_{i,j}$ have a finite second moment, i.e. $\mathbb{E}[v^2] < \infty$. The aim is to prove a strong law of large numbers and a functional central limit theorem for $w_{0,n}$. Recall that $p \in (0,1]$.

Theorem 3.1. *If $\mathbb{E}[v^2] < \infty$ then there exists a constant $C \in (0,\infty)$ such that*

$$\frac{w_{0,n}}{n} \xrightarrow[n\to\infty]{} C \ \text{a.s.}$$

and

$$\frac{w_{0,n}^{+}}{n} \xrightarrow[n\to\infty]{} C \ \text{in } \mathcal{L}^1.$$

In order to state the central limit theorem we will need some more notation. This will also give an idea of how we want to prove the SLLN and CLT. We define so-called *renewal points* which will give the model a regenerative structure. In order to do this we need the following three events, which depend on a constant c to be chosen later. The constant c will have to be sufficiently small but still satisfy $\mathbb{P}[v < c] > 0$. We define the random variables $\alpha_{i,j}$, $i < j \in \mathbb{Z}$ to be 1 if the edge (i, j) is present and $-\infty$ otherwise. For $x \in \mathbb{Z}$, define

$$A_x^{++} = \bigcap_{l=1}^{\infty} \{w_{x,x+l} \geq cl\}, \tag{3.1}$$

$$A_x^{-+} = \bigcap_{j,l=1}^{\infty} \{\alpha_{x-j,x+l} v_{x-j,x+l} < c(l+j)\} \tag{3.2}$$

and

$$A_x^{--} = \bigcap_{l=1}^{\infty} \{w_{x-l,x} \geq cl\}. \tag{3.3}$$

We say that x is a renewal point if $A_x^{++} \cap A_x^{-+} \cap A_x^{--}$ holds. Write $A_x = A_x(c)$ for this combined event, and write $\mathcal{R} = \mathcal{R}(c)$ for the set of points such that A_x holds.

Let us explain in words the meaning of the three events used in the definition of the set $\mathcal{R}$. The event A_x^{++} occurs if for every $y > x$, the optimal path from x to y has weight at least c times $y - x$, the physical length of the path. The event A_x^{--} says the equivalent thing about paths from y to x for $y < x$. Finally, the event A_x^{-+} says that every edge that contains x in its interior has weight *less* than c times its length.

We immediately obtain the following property:

Lemma 3.2. *If $x \in \mathcal{R}$ and $i < x < j$, then the optimal path from i to j passes through the point x. In particular,*

$$w_{i,j} = w_{i,x} + w_{x,j}. \tag{3.4}$$

For suppose a path includes an edge (u, y) with x in its interior. Then we can increase the weight of the path by replacing that edge by the union of the optimal paths from u to x and from x to y.

A priori the set $\mathcal{R}$ might be empty, finite or infinite, but the following result will be proved in Section 3.3.1 (for $p = 1$) and Section 3.3.2 (for $p < 1$):

Lemma 3.3. *Suppose that* $\mathbb{E}[v^2] < \infty$. *There exists* $c > 0$ *such that the set* $\mathcal{R}(c)$ *is infinite with probability 1.*

The result is stated explicitly with a sufficient condition on c in Lemma 3.16.

We can then denote the points in $\mathcal{R}$ by $(\Gamma_n)_{n\in\mathbb{Z}}$ where Γ_0 is the smallest non-negative element of $\mathcal{R}$.

Remark 3.4. *If* $\mathbb{E}[v^2] = \infty$ *then the set* $\mathcal{R}$ *is almost surely the empty set. This will follow from the proof of Lemma 3.3, see Remark 3.14.*

The final consequence of the definition is that, as suggested by the name, the set of points $\mathcal{R}$ forms a renewal process. Furthermore, conditional on the points of $\mathcal{R}$, the weights of edges contained within different renewal intervals are independent. These properties are proved in Lemma 3.18, and will be central to the structure of the argument that follows.

We now have the necessary notation to state the functional central limit theorem for $w_{0,n}$. To ensure that the variance is finite, we need the stronger condition $\mathbb{E}[v^3] < \infty$. (see Proposition 3.26 below).

Theorem 3.5. *Suppose* $\mathbb{E}[v^3] < \infty$. *Then there exists* $c > 0$ *such that the following holds. Let* $\sigma^2 = \mathrm{Var}(w_{\Gamma_0,\Gamma_1} - C(\Gamma_1 - \Gamma_0))$ *where* C *is as in Theorem 3.1 and* $\lambda = \mathbb{P}[A_0(c)] = \mathbb{P}(0 \in \mathcal{R}(c))$. *Then* $\sigma^2 < \infty$ *and*

$$\left(l_n(t) = \frac{w_{0,[nt]} - Cnt}{\sigma\sqrt{\lambda n}}, t \geq 0 \right)$$

converges weakly as $n \to \infty$ *to a standard Brownian motion.*

A sufficient condition on c is given in Lemma 3.21.

The main idea for the proofs of both the SLLN and the CLT is to use the regenerative structure induced by the renewal points to represent $w_{0,n}$ as a random sum of i.i.d. random variables in the following way

$$w_{0,n} = w_{0,\Gamma_0} + \sum_{i=1}^{r(n)} w_{\Gamma_{i-1},\Gamma_i} + w_{\Gamma_{r(n)},n} \tag{3.5}$$

where $r(n)$ is such that $\Gamma_{r(n)}$ is the largest renewal point to the left of n. We will show in Proposition 3.18 that the random variables $w_{\Gamma_{i-1},\Gamma_i}$ form an i.i.d. sequence, for $i \geq 1$.

Let ℓ_n be the length of the longest edge and h_n the weight of the heaviest edge used on the geodesic from 0 to n. The final result of this section concerns the asymptotic behaviour of ℓ_n and h_n, under the assumption that the tail of the distribution of v is regularly varying with index $s > 2$. When $2 < s < 3$, we can deduce that the fluctuations of the passage time are of order larger than $\sqrt{n}$, and so the central limit theorem cannot be extended to this case.

Theorem 3.6. *Suppose that the tail of v is regularly varying with index $s > 2$, in the sense that*

$$\frac{1-F(tx)}{1-F(x)} \to t^{-s} \text{ as } x \to \infty, \text{ for every } t > 0. \tag{3.6}$$

Then we have

$$\frac{\log \ell_n}{\log n} \to \frac{1}{s-1} \text{ in probability as } n \to \infty \tag{3.7}$$

and the same holds with ℓ_n replaced by h_n.

Furthermore, the fluctuations of $w_{0,n}$ are of larger order than n^β for any $\beta < 1/(s-1)$, in the sense that for any sequence y_n,

$$\mathbb{P}(w_{0,n} \in [y_n, y_n + n^\beta]) \to 0 \text{ as } n \to \infty.$$

In particular, if $2 < s < 3$ then

$$\frac{\text{Var}\,(w_{0,n})}{n} \to \infty,$$

and a central limit theorem such as that in Theorem 3.5 cannot hold, even for individual values of t.

We prove Theorem 3.6 in Section 3.3.4 and also give some examples that show how the behaviour of ℓ_n depends on the tail of the distribution.

3.2.2 Weights with infinite second moment

Here we look at weight distributions that do not have a finite second moment, i.e. $\mathbb{E}[v^2] = \infty$. We assume that the tail of the weight distribution is regularly varying with index $s \in (0,2)$, in the sense of (3.6).

We introduce two useful ways to construct our model in discrete space and explain how the second construction can be used to define a corresponding model in continuous space on $[0,1]$. We show that the passage time w for the continuous model is finite and show convergence of an appropriately rescaled version of $w_{0,n}$ to w. This will give us the asymptotic behaviour of $w_{0,n}$.

Discrete model

We start with the case $p = 1$. Since $w_{0,n}$ depends only on $v_{i,j}$ for $0 \le i, j \le n$ it suffices to consider only the interval $[0, n]$. We can then rescale and consider the set $\left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\right\}$ instead of the interval $[0, n]$. For each $n \in \mathbb{N}$ and $0 \le i < j \le n$, let $v_{i,j}^{(n)}$ be i.i.d. with distribution F. The weight of the edge between $\frac{i}{n}$ and $\frac{j}{n}$ is now given by $v_{i,j}^{(n)}$. We introduce some new notation: for two edges $x = (i, j)$, $y = (i', j')$ we write $x \sim y$ and say x and y are compatible if $j \le i'$ or $j' \le i$. The edges x and y being compatible means that they do not overlap and that they can both be used on a path from 0 to 1. As before we define

$$w_{0,n} = \max_{\pi \in \Pi_n} \sum_{e \in \pi} v_e^{(n)} \tag{3.8}$$

where Π_n is the set of all paths from 0 to 1 in $\left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\right\}$.

We can think of the same model in the following alternative way: let $M_1^{(n)} \ge M_2^{(n)} \ge \dots \ge M_{\binom{n+1}{2}}^{(n)}$ be the order statistics in decreasing order of the $v_{i,j}^{(n)}$. Let $Y_1^{(n)}, Y_2^{(n)}, \dots, Y_{\binom{n+1}{2}}^{(n)}$ be a random ordering of those edges, chosen uniformly from all the $\binom{n+1}{2}!$ possibilities. $Y_i^{(n)}$ is the location of the i-th largest weight $M_i^{(n)}$. Now

$$\mathcal{C}_{0,n} = \left\{ A \subset \left\{1, \dots, \binom{n+1}{2}\right\} : Y_i^{(n)} \sim Y_j^{(n)} \text{ for all } i, j \in A \right\} \tag{3.9}$$

is the random set of admissible paths. Then we have

$$w_{0,n} = \max_{A \in \mathcal{C}_{0,n}} \sum_{i \in A} M_i^{(n)} \tag{3.10}$$

which is equivalent to the previous definition of $w_{0,n}$ in (3.8).

Continuous model

Following the second approach above, we can define a corresponding continuous model. Let $W_1, W_2, \ldots$ be a sequence of i.i.d. exponential random variables with mean 1 and define, for $k = 1, 2, \ldots$, $M_k = (W_1 + \ldots + W_k)^{-\frac{1}{s}}$. Let $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$ be two sequences of i.i.d. uniform random variables on $[0,1]$ (independent of the W_k). Put $Y_i = (\min(U_i, V_i), \max(U_i, V_i))$ for $i = 1, 2, \ldots$. The ith largest weight M_i will be attached to the ith edge Y_i. Similar to (3.9) we define

$$\mathcal{C} = \{A \subset \{1, 2, \ldots\} : Y_i \sim Y_j \text{ for all } i, j \in A\}.$$

Then we can define a last-passage time for this continuous model analogously to (3.10) by

$$w = \sup_{A \in \mathcal{C}} \sum_{i \in A} M_i. \tag{3.11}$$

A priori the random variable w could be infinite, but we will see in Theorem 3.7 below that it is almost surely finite.

Convergence results

The intuition behind the approximation of the discrete model by the continuous one is the following pair of convergence results. First, for any finite $k \in \mathbb{N}$ we have

$$\left(Y_1^{(n)}, Y_2^{(n)}, \ldots, Y_k^{(n)}\right) \xrightarrow{d} (Y_1, Y_2, \ldots, Y_k) \tag{3.12}$$

as $n \to \infty$, where we use the product topology on $([0,1]^2)^k$.

Second, define $b_n = a_{\binom{n+1}{2}} = F^{(-1)}\left(1 - \frac{1}{\binom{n+1}{2}}\right)$ and put $\widetilde{M}_i^{(n)} = \frac{M_i^{(n)}}{b_n}$. (As an example, if the weight distribution F is Pareto(s), with $F(x) = 1 - x^{-s}$ for $x \geq 1$, then b_n grows like $n^{2/s}$. More generally under assumption (3.6), $\lim_{n\to\infty} \frac{\log b_n}{\log n} = 2/s$)). Then from classical results in extreme value theory (see for example Section 9.4 of [13]) we have for any $k \in \mathbb{N}$ that

$$\left(\widetilde{M}_1^{(n)}, \widetilde{M}_2^{(n)}, \dots, \widetilde{M}_k^{(n)}\right) \xrightarrow{d} (M_1, M_2, \dots, M_k) \text{ as } n \to \infty. \tag{3.13}$$

In this way both the locations and weights of the heaviest edges in the discrete model (after appropriate rescaling) are approximated by their equivalents in the continuous model. We will show that it is the heaviest edges, which make the dominant contribution to the passage time, and obtain the following convergence result:

Theorem 3.7. *The random variable w in (3.11) is almost surely finite. If $p = 1$ and (3.6) holds, then $\frac{w_{0,n}}{b_n} \to w$ in distribution as $n \to \infty$.*

These heavy edges have length on the order of n. This is in strong contrast to the behaviour in the case $\mathbb{E}[v^2] < \infty$, where the important contribution to the passage time is given by edges of order 1. See Figure 3.1 for an illustration of the two types of behaviour.

This scaling limit extends in a simple way to the case $p < 1$, after taking account of the fact that the total number of edges available in the interval $[0, n]$ is now on the order of $pn^2/2$ rather than on the order of $n^2/2$:

Theorem 3.8. *Let $p \in (0, 1]$ and suppose that (3.6) holds. Then $\frac{p^{-1/s} w_{0,n}}{b_n} \to w$ in distribution as $n \to \infty$.*

Remark 3.9. *Although we don't pursue it in detail here, one can also prove convergence of the optimal path itself in the discrete model to that of the continuous model, using an approach similar to that in [26]. For convenience, assume that F is continuous. Then with probability 1, there exists a unique path $A^{(n)*} \in \mathcal{C}_{0,n}$ which realises the maximal passage time in (3.10). One can show that in the continuous model there*

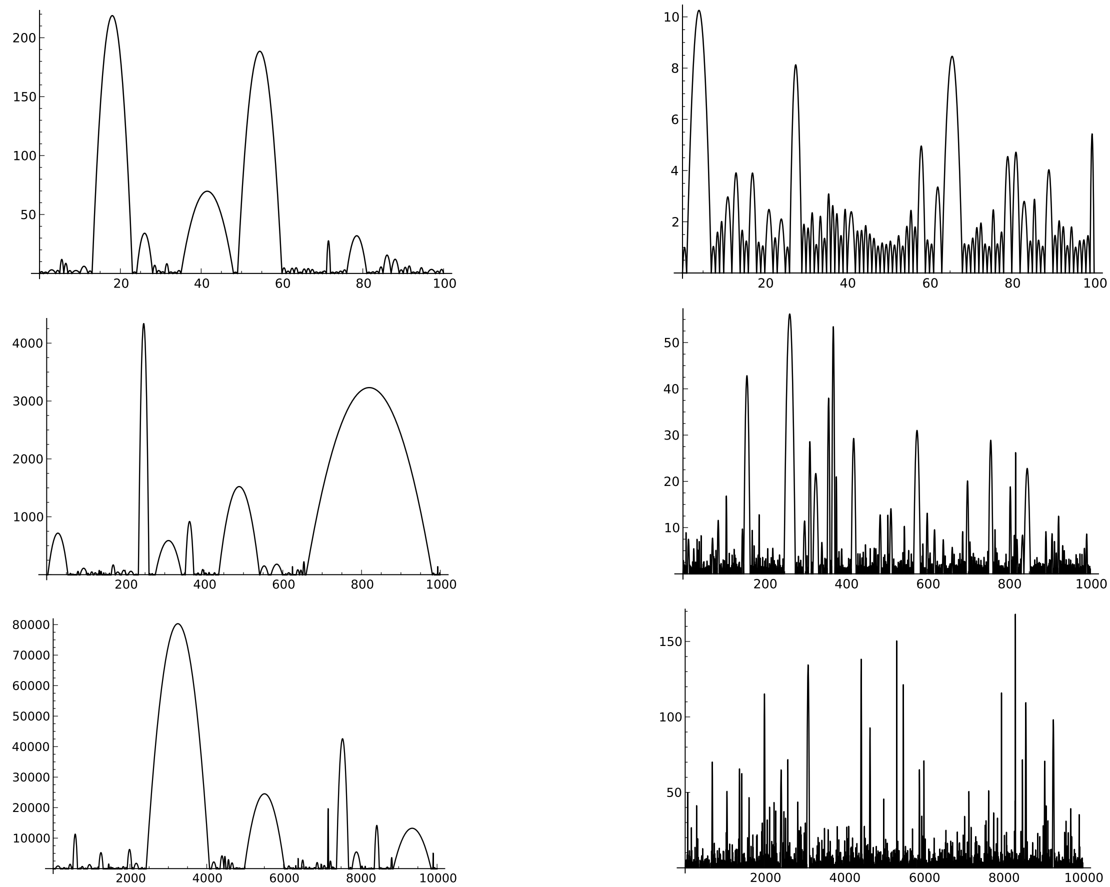

Figure 3.1: Simulations for $n = 100, 1000$ and 10000 for two weight distributions with $\mathbb{P}(v > x) = x^{-s}$, $x \geq 1$; on the left $s = 1.5$ and on the right $s = 2.5$. On the left we are in the setting of Section 3.2.2. The scaling limit is clearly visible; visually one can hardly distinguish the cases $n = 100$ and $n = 10000$ (see Remark 3.9 about convergence of the path distribution). The heaviest edges make an important contribution to the total weight of the geodesic; their length is on the order of n, and their weight is on the order of $n^{2/s}$ which is also the order of the total weight of the path. On the right, the variance of the weight distribution is finite; the heaviest edges have both length and weight approximately on the order of $n^{1/(s-1)}$, while the total weight of the path is on the order of n and obeys a law of large numbers.

exists a unique set $A^* \in \mathcal{C}$ achieving the supremum in (3.11) (which is therefore in fact a maximum), and

$$\left(\left(Y_i^{(n)}\right)_{i\in\mathbb{N}}, A^{(n)*}\right) \xrightarrow[n\to\infty]{d} \left((Y_i)_{i\in\mathbb{N}}, A^*\right),$$

where we consider the Euclidean distance on $\mathbb{R}^2$ and the product topology for the convergence of $\left(Y_i^{(n)}\right)_{i\in\mathbb{N}}$, and say that a sequence $(A_k)_{k\in\mathbb{N}}$ of subsets of $\mathbb{N}$ converges to a set $A \subset \mathbb{N}$ if for every $m \in \mathbb{N}$ there exists a $K \in \mathbb{N}$ such that $A_k \cap \{1,\ldots,m\} = A \cap \{1,\ldots,m\}$ for all $k \geq K$.

See Theorem 4.2 of [26] for an analogous result in the two-dimensional last-passage case. One can then proceed to show that in fact the set of endpoints of edges used in the optimal path from 0 to n (rescaled by n) converges in distribution under the Hausdorff metric to the equivalent object in the continuous model. (Compare Theorem 4.4 of [26]).

3.3 Proofs for the model with $\mathbb{E}\left[v^2\right] < \infty$

In this section we consider the case where the weights $v_{i,j}$ have a finite second moment, i.e. $\mathbb{E}[v^2] < \infty$. To avoid degeneracies we assume throughout that v is not constant almost surely. Our main aim is to prove Theorems 3.1, 3.5 and 3.6. We start with the model where $p = 1$; that is, all edges are present. First, we show that the set $\mathcal{R}$ of renewal points is infinite almost surely. Then we generalize this result to the case where $p \leq 1$. In the following subsection we will use this result to prove the strong law of large numbers (Theorem 3.1) and the central limit theorem (Theorem 3.5) for $w_{0,n}$ for general $p \in (0,1]$. The next subsection will look at the behaviour of the random variables ℓ_n, giving the lengths of the longest edge, and h_n, giving the weight of the heaviest edge, used on the geodesic from 0 to n, see Theorem 3.6. We will use these results to comment on the behaviour of the model if $\mathbb{E}[v^2] < \infty$, but $\mathbb{E}[v^3] = \infty$. In the last subsection we briefly discuss the case where the edge probabilities are not constant, but depend on the length of the edges.

3.3.1 Proof of Lemma 3.3 for $p = 1$

When $p = 1$, we have $\alpha_{i,j} = 1$ for all i, j; that is, all edges (i, j), $i, j \in \mathbb{Z}$, are present, and in particular there is a path between any two points.

Let $A_x = A_x^{++} \cap A_x^{-+} \cap A_x^{--}$ be the event that x is a renewal point. We start with the following lemma which is simply Lemma 3.3 with the additional condition that $\mathbb{P}[A_0] > 0$. After this lemma we will prove in Propositions 3.11, 3.12 and 3.13 that $\mathbb{P}[A_0] > 0$.

Lemma 3.10. *If $\mathbb{P}[A_0] > 0$ then $\mathcal{R}$ is almost surely an infinite set.*

Proof. Let $\lambda = \mathbb{P}[A_0]$, which is strictly positive by assumption. We can approximate the event A_0 by an event $A_0^{'}$ that depends only on finitely many of the $v_{i,j}$. In particular, for every $\varepsilon > 0$ there exists $m \in \mathbb{N}$ such that $A_0^{'}$ depends only on $v_{i,j}$ for $-m \le i < j \le m$ and $\mathbb{P}\left[A_0 \Delta A_0^{'}\right] < \varepsilon$. By translation invariance of our model we get that the same is true for any event A_x, where $A_x^{'}$ is defined as the translation of $A_0^{'}$ in the natural way. Then the events $A_0^{'}$, $A_{2m}^{'}$, $A_{4m}^{'}$, $\ldots$ are i.i.d., and we have

$$\begin{aligned}\mathbb{P}\left[A_0^{'} \cup A_{2m}^{'} \cup \ldots \cup A_{2(R-1)m}^{'}\right] &= 1 - \mathbb{P}\left[\left(A_0^{'}\right)^c\right]^R \\ &\ge 1 - \left(\mathbb{P}\left[A_0^c\right] + \varepsilon\right)^R \\ &= 1 - (1 - \lambda + \varepsilon)^R .\end{aligned}$$

With this we get

$$\begin{aligned}\mathbb{P}\left[A_0 \cup A_{2m} \cup \ldots \cup A_{2(R-1)m}\right] &= 1 - \mathbb{P}\left[A_0^c \cap A_{2m}^c \cap \ldots \cap A_{2(R-1)m}^c\right] \\ &\ge 1 - \left(\mathbb{P}\left[\left(A_0^{'}\right)^c \cap \left(A_{2m}^{'}\right)^c \cap \ldots \cap \left(A_{2(R-1)m}^{'}\right)^c\right] + R\varepsilon\right) \\ &\ge 1 - (1 - \lambda + \varepsilon)^R - R\varepsilon.\end{aligned}$$

For any $\delta > 0$ we can now first choose R large enough such that $(1 - \lambda + \varepsilon)^R < \frac{\delta}{2}$ for all small enough ε, and then further choose $\varepsilon > 0$ small enough such that also $R\varepsilon < \frac{\delta}{2}$, to get

$$\mathbb{P}\left[A_0 \cup A_{2m} \cup \ldots \cup A_{2(R-1)m}\right] \geq 1 - \delta.$$

Since δ was arbitrary this shows that at least one of the events A_x, for $x \geq 0$, holds. In the same way we can show that with probability 1 for any fixed $y \in \mathbb{Z}$ there exists $x \geq y$ such that A_x holds. This implies that with probability 1 infinitely many of the A_x hold and therefore $\mathcal{R}$ is almost surely an infinite set. □

Now it remains to show that the condition of Lemma 3.10 is satisfied, i.e. that $\mathbb{P}[A_0] > 0$. To be precise, we have to show that for sufficiently small $c > 0$ with $\mathbb{P}[v < c] > 0$ we have $\mathbb{P}[A_0] > 0$. We will do this in four steps: first we show that the events A_0^{++}, A_0^{-+} and A_0^{--} are independent and then we will show for each of them that they hold with positive probability for a suitable $c > 0$.

Proposition 3.11. *For any fixed $x \in \mathbb{Z}$ the events A_x^{++}, A_x^{-+} and A_x^{--} are independent.*

Proof. As already mentioned above, the event A_x^{++} depends only on edges whose left endpoint is at least x, A_x^{-+} depends only on edges with their left endpoint to the left of x and their right endpoint to the right of x, and A_x^{--} depends only on edges whose right endpoint is at most x. Since all the weights are i.i.d. this implies the required independence of the events A_x^{++}, A_x^{-+} and A_x^{--}. □

Proposition 3.12. *If $\mathbb{E}[v] < \infty$, then for any $c < \mathbb{E}[v]$ we have that $\mathbb{P}[A_x^{++}] > 0$ and $\mathbb{P}[A_x^{--}] > 0$.*

Proof. Since all the nearest neighbour edges are present we can bound $w_{x,x+l}$ for any $l \in \mathbb{N}$ from below by $\sum_{j=0}^{l-1} v_{x+j,x+j+1}$ and the $v_{x+j,x+j+1}$ are i.i.d. By the strong law of large numbers we have that $\mathbb{P}\left[\bigcap_{l=L}^{\infty} \sum_{j=0}^{l-1} v_{x+j,x+j+1} \geq cl\right] \geq \frac{1}{2}$ for large enough L. Since only finitely many of the $v_{x+j,x+j+1}$ are involved in the events $\sum_{j=0}^{l-1} v_{x+j,x+j+1} \geq cl$ for $l < L$ and the $v_{x+j,x+j+1}$ are i.i.d. there is positive probability that all events $\sum_{j=0}^{l-1} v_{x+j,x+j+1} \geq cl$ hold for $l < L$ as well. So $\mathbb{P}[A_x^{++}] > 0$. The proof for A_x^{--} is exactly the same. □

Proposition 3.13. *Assume that v is not a constant. If $\mathbb{E}\left[v^2\right] < \infty$, then for every c such that*

$$\operatorname{ess\,inf}[v] < c < \mathbb{E}[v],$$

we have $\mathbb{P}\left[A_x^{-+}\right] > 0$.

Proof. Note that $\operatorname{ess\,inf}[v] < \mathbb{E}[v]$ since we assume that v is not a.s. constant. We have

$$\begin{aligned}
\mathbb{P}\left[A_x^{-+}\right] &= \prod_{j,l=1}^{\infty} \mathbb{P}\left[v_{x-j,x+l} < c\,(l+j)\right] \\
&= \prod_{i=2}^{\infty} \mathbb{P}\left[v < ci\right]^{i-1} \\
&= e^{\sum_{i=2}^{\infty}(i-1)\ln(1-\mathbb{P}[v \geq ci])}.
\end{aligned}$$

The exponent is negative, and the RHS is positive if and only the sum converges to a finite quantity rather than to $-\infty$. Since $\mathbb{P}(v < ci) > 0$ for all i (because $c > \operatorname{ess\,inf}[v]$), and since $\log(1-x) \sim -x$ as $x \to 0$, this holds if and only if $\sum i\mathbb{P}(v \geq ci)$ is finite, which in turn holds if and only if the variance of v is finite. □

Remark 3.14. *If, on the other hand, $\mathbb{E}\left[v^2\right] = \infty$ then $\mathbb{P}\left[A_x^{-+}\right] = 0$ and therefore $\mathcal{R}$ is empty almost surely.*

Proof of Lemma 3.3: This follows directly from Lemma 3.10 and Propositions 3.11, 3.12 and 3.13. □

3.3.2 Proof of Lemma 3.3 for $p < 1$

Let us now consider the case where $p < 1$. We say that a point $x \in \mathbb{Z}$ is a *strongly connected point* if x is connected to every other point by a path. Here we do not consider the weights of the edges, so the paths do not have to be optimal. We denote the set of strongly connected points by $\mathcal{S}$. In the previous section every point was a strongly connected point since $p = 1$. The following results about the strongly connected points have all been shown in [14]:

- the probability that 0 is a strongly connected point is strictly positive for any $p > 0$
- there are almost surely infinitely many strongly connected points
- the sequence of strongly connected points forms a stationary renewal process
- if we let $\dots, \tau_{-1}, \tau_0, \tau_1, \tau_2, \dots$ be the sequence of strongly connected points, where τ_0 is the smallest non-negative element of $\mathcal{S}$, then for some $\alpha > 0$,

$$\mathbb{E}\left[e^{\alpha\tau_0}\right] < \infty, \text{ and } \mathbb{E}\left[e^{\alpha(\tau_{i+1}-\tau_i)}\right] < \infty \text{ for all } i. \tag{3.14}$$

By $w_{k,l}$ we denote again the weight of the geodesic from k to l. This might now be $-\infty$ if there exists no path between k and l and we are therefore taking the supremum over an empty set. However, if x is a strongly connected point, then $w_{x-j,x+l} > 0$ for all $j, l \in \mathbb{N}$ since we know that there exists a path from any $x - j$ to x and from x to any $x + l$. For $x \in \mathbb{Z}$ let $m(x)$ be the index of the largest strongly connected point such that $\tau_{m(x)} < x$.

The definition of the renewal points is the same as before – see (3.1), (3.2) and (3.3) (now $\alpha_{i,j} = -\infty$ if the edge (i, j) is not present). By definition we have that if x is not a strongly connected point then $w_{x,x+l} = -\infty$ for some $l \geq 1$ or $w_{x-j,x} = -\infty$ for some $j \geq 1$. So x can only be a renewal point if it is a strongly connected point. An equivalent of Lemma 3.10 still holds in the case where $p < 1$ and we want to prove again that the condition for Lemma 3.10 ($\mathbb{P}[A_0] > 0$) holds. This will give us Lemma 3.3. Again it will be enough to show that the three events A_x^{++}, A_x^{-+} and A_x^{--} are independent and that all of them happen with positive probability. The independence follows directly from the same argument as in Proposition 3.11. Let $\gamma > 0$ be the density of strongly connected points and $\delta = \mathbb{E}[w_{\tau_0,\tau_1}]$.

Proposition 3.15. *If* $0 < c < \gamma\delta$, *then* $\mathbb{P}[A_x^{++}] > 0$, $\mathbb{P}[A_x^{-+}] > 0$ *and* $\mathbb{P}[A_x^{--}] > 0$.

Proof. First look at the events A_x^{++} and A_x^{--}. Now not all the nearest neighbour edges are present, but we can use the strongly connected points to get a similar bound to the one in the proof of Proposition 3.12. Without loss of generality assume that $x = 0$ and note that $m(0) = -1$ from the definition. For any $l > \tau_0$ we can write

$$w_{0,l} \geq w_{0,\tau_0} + \sum_{j=1}^{m(l)} w_{\tau_{j-1},\tau_j} + w_{\tau_{m(l)},l}. \tag{3.15}$$

Fix a $c < \gamma\delta$. Since the strongly connected points form a stationary renewal process, independent of the weights, and the density γ of strongly connected points is strictly positive, the terms in the sum are i.i.d. and we have both $\frac{m(l)}{l} \to \gamma$ almost surely as $l \to \infty$, and $\frac{1}{M}\sum_{j=1}^{M} w_{\tau_{j-1},\tau_j} \to \delta$ as $M \to \infty$. So in fact

$$\frac{1}{l}\sum_{j=1}^{m(l)} w_{\tau_{j-1},\tau_j} \to \gamma\delta \text{ a.s. as } l \to \infty.$$

Then since $c < \gamma\delta$ by assumption, we have that for some L, the event

$$w_{0,l} \geq cl \text{ for all } l \geq L$$

has positive probability. But if this event occurs, then we can obtain a realisation for which

$$w_{0,l} \geq cl \text{ for all } l \geq 1$$

occurs by altering the values of only finitely many edges. Hence that event also has positive probability, and so $\mathbb{P}(A_x^{++}) > 0$ as desired. In exactly the same way, also $\mathbb{P}(A_x^{--}) > 0$.

Now look at the event A_x^{-+}. With the same arguments as in the previous section we get that for large L

$$\mathbb{P}\left[\bigcap_{j,l=L}^{\infty} \{\alpha_{x-j,x+l} v_{x-j,x+l} \leq c(l+j)\}\right] > 0.$$

Since there is a probability of $1 - p$ for each edge not to be present, i.e. $\alpha_{i,j} = -\infty$, we get

$$\mathbb{P}\left[\bigcap_{j,l=1}^{L-1} \{\alpha_{x-j,x+l} v_{x-j,x+l} \leq c(l+j)\}\right] \geq (1-p)^{(L-1)^2}.$$

Hence $\mathbb{P}[A_x^{-+}] = \mathbb{P}\left[\bigcap_{j,l=1}^{\infty} \{\alpha_{x-j,x+l} v_{x-j,x+l} \leq c(l+j)\}\right] > 0$ also. □

So we have shown that the condition in Lemma 3.10 is still satisfied in the case where $p < 1$ and therefore Lemma 3.3 holds for $p < 1$ as well.

To unify the conditions on c for the cases $p = 1$ and $p < 1$, note that if $p = 1$ then $\gamma = 1$, and that $\mathbb{E}[v] \leq \delta$. Then we can put together the results of the last two sections to give the following:

Lemma 3.16. *Let $p \in (0, 1]$. If*

$$\gamma \operatorname{ess\,inf}[v] < c < \gamma \mathbb{E}[v] \tag{3.16}$$

then $\lambda = \mathbb{P}(A_0) > 0$ and the set $\mathcal{R}$ is infinite with probability 1.

3.3.3 Proofs of the SLLN and CLT for general $p \in (0, 1]$

In the previous two sections we have shown that under the condition (3.16) on c, the set $\mathcal{R}$ of renewal points is infinite. Now we want to prove a strong law of large numbers and a central limit theorem for the random variable $w_{0,n}$, see Theorems 3.1 and 3.5. As before we denote the points in $\mathcal{R}$ by $\ldots, \Gamma_{-1}, \Gamma_0, \Gamma_1, \ldots$, where Γ_0 is the smallest non-negative element of $\mathcal{R}$. Evaluating the function w at the renewal points Γ_n gives the following equation, related to (3.5):

Proposition 3.17. *For all $m < n$ we have*

$$w_{\Gamma_m,\Gamma_n} = w_{\Gamma_m,\Gamma_{m+1}} + \ldots + w_{\Gamma_{n-1},\Gamma_n}.$$

Proof. This follows directly from the definition of the renewal points and (3.4) □

We now want to use the fact stated in this proposition to prove a strong law of large numbers and a central limit theorem for the random variable $w_{0,n}$. If $w_{\Gamma_m,\Gamma_{m+1}}$, $m \geq 0$ are independent, then for $n \geq \Gamma_0$ we can write

$$w_{0,n} = w_{0,\Gamma_0} + \sum_{i=1}^{r(n)} w_{\Gamma_{i-1},\Gamma_i} + w_{\Gamma_{r(n)},n} \tag{3.17}$$

where $r(n) = \max\{m : \Gamma_m < n\}$ and, since $w_{\Gamma_{i-1},\Gamma_i}$, $i \geq 1$, are i.i.d. use then the standard strong law of large numbers and central limit theorem (under moment conditions for the variance of $w_{\Gamma_{i-1},\Gamma_i}$) applied to the sum in (3.17) to get corresponding results for $w_{0,n}$. Note that since the density of renewal points $\lambda = \mathbb{P}[A_0]$ is strictly positive we have that $r(n) \sim \lambda n$ for large n. So first we want to show that $w_{\Gamma_{i-1},\Gamma_i}$, $i \geq 1$ are indeed independent.

Define $\mathcal{C}_k = \big(\Gamma_k - \Gamma_{k-1}, v_{\Gamma_{k-1}+n,\Gamma_{k-1}+i}, \alpha_{\Gamma_{k-1}+n,\Gamma_{k-1}+i} : 0 \leq n < i \leq \Gamma_k - \Gamma_{k-1}\big)$, $k \in \mathbb{Z}$. Then these cycles have a regenerative structure in the following sense:

Lemma 3.18. *The cycles $(\mathcal{C}_k, k \in \mathbb{Z})$ are independent and $(\mathcal{C}_k, k \in \mathbb{Z} - \{0\})$ are identically distributed. The process $(\Gamma_n)_{n\in\mathbb{Z}}$ forms a stationary renewal process.*

Proof. We start with the following observation about the effect the presence of a renewal point at site $x \in \mathbb{Z}$ has on the weights to the left and to the right of x. Let $\mathcal{F}_x^+$ be the sigma-algebra generated by the $(v_{i,j}, \alpha_{i,j} : x \leq i < j)$ and let $\mathcal{F}_x^-$ be the sigma-algebra generated by the $(v_{i,j}, \alpha_{i,j} : i < j \leq x)$. These two sigma-algebras are independent as all our weights are independent. But this is still true even if we know that there is a renewal point at x: for any $B^- \in \mathcal{F}_x^-$, $B^+ \in \mathcal{F}_x^+$ we have

$$\begin{aligned}
&\mathbb{P}\left[B^- \cap B^+ \middle| A_x^{--} \cap A_x^{-+} \cap A_x^{++}\right] \\
&\qquad = \frac{\mathbb{P}\left[B^- \cap B^+ \cap A_x^{--} \cap A_x^{-+} \cap A_x^{++}\right]}{\mathbb{P}\left[A_x^{--} \cap A_x^{-+} \cap A_x^{++}\right]} \\
&\qquad = \frac{\mathbb{P}\left[B^- \cap A_x^{--}\right] \mathbb{P}\left[A_x^{-+}\right] \mathbb{P}\left[B^+ \cap A_x^{++}\right]}{\mathbb{P}\left[A_x^{--}\right] \mathbb{P}\left[A_x^{-+}\right] \mathbb{P}\left[A_x^{++}\right]} \\
&\qquad = \mathbb{P}\left[B^- \left| A_x^{--}\right.\right] \mathbb{P}\left[B^+ \left| A_x^{++}\right.\right] \\
&\qquad = \mathbb{P}\left[B^- \left| A_x^{--} \cap A_x^{-+} \cap A_x^{++}\right.\right] \mathbb{P}\left[B^+ \left| A_x^{--} \cap A_x^{-+} \cap A_x^{++}\right.\right]
\end{aligned}$$

This shows that having a renewal point at x does not introduce any dependence between the weights to the left and the weights to the right of x. Now we want to show that if A_x holds we can determine where all the renewal points to the right of x are only by looking at edges with both endpoints to the right of x. So assume again that A_x holds. For $y > x$ (and fixed x) define the event $\widetilde{A}_y = A_y^{++} \cap \widetilde{A}_y^{-+} \cap \widetilde{A}_y^{--}$ with

$$\widetilde{A}_y^{-+} = \bigcap_{l \geq 1, 1 \leq j \leq y-x} \{\alpha_{y-j,y+l} v_{y-j,y+l} \leq c(l+j)\}$$

and

$$\widetilde{A}_y^{--} = \bigcap_{1 \leq j \leq y-x} \{w_{y-j,y} \geq cj\}.$$

The events A_y^{++}, $\widetilde{A}_y^{-+}$ and $\widetilde{A}_y^{--}$ all depend only on edges to the right of x. Now we want to show that conditioned on A_x the event A_y holds if and only if the event $\widetilde{A}_y$ holds. On A_x we have

$$w_{x-j,x} \geq cj \text{ and } \alpha_{x-j,x+l} v_{x-j,x+l} \leq c(l+j) \text{ and } w_{x,x+l} \geq cl \text{ for all } j, l \geq 1. \tag{3.18}$$

Assume that $\widetilde{A}_{x+k}$ holds. Then we have

$$\begin{aligned} w_{x+k-j,x+k} \geq cj \text{ and } \alpha_{x+k-j,x+k+l} v_{x+k-j,x+k+l} \leq c(l+j) \\ \text{and } w_{x+k,x+k+l} \geq cl \text{ for all } 1 \leq j \leq k, l \geq 1. \end{aligned} \tag{3.19}$$

We have to show that we can conclude from this that A_{x+k} holds, i.e.

$$\begin{aligned} w_{x+k-j,x+k} \geq cj \text{ and } \alpha_{x+k-j,x+k+l} v_{x+k-j,x+k+l} \leq c(l+j) \\ \text{and } w_{x+k,x+k+l} \geq cl \text{ for all } j, l \geq 1. \end{aligned} \tag{3.20}$$

So take $j > k$. Then we have

$$w_{x+k-j,x+k} = w_{x+k-j,x} + w_{x,x+k} \qquad \text{(since } x \text{ is a renewal point)}$$

$$\geq c(k-j)+kj = cj \qquad \text{(by (3.18) and (3.19))}$$

and also for any $l \geq 1$

$$\begin{aligned} \alpha_{x+k-j,x+k+l}v_{x+k-j,x+k+l} &= \alpha_{x-(j-k),x+k+l}v_{x-(j-k),x+k+l} \\ &\leq c(k+l+j-k) \qquad \text{(by (3.18))} \\ &= c(l+j). \end{aligned}$$

So (3.20) holds. This implies that A_{x+k} holds if $\widetilde{A}_{x+k}$ holds. The other implication is obvious.

This shows that for any $m \geq 1$ the cycles $\mathcal{C}_m$, $\mathcal{C}_{m+1}$, $\ldots$ are independent of the position of Γ_{m-1} and everything to the left of Γ_{m-1}. With similar arguments to the ones above we can also show that for any $m \geq 1$ the cycles $\mathcal{C}_{-m}$, $\mathcal{C}_{-m-1}$, $\ldots$ are independent of the position of Γ_{-m} and everything to the right of Γ_{-m}. Overall we get that the cycles $(\mathcal{C}_k, k \in \mathbb{Z})$ are independent and, by symmetry, that the cycles $(\mathcal{C}_k, k \in \mathbb{Z} - \{0\})$ are identically distributed.

Then $\Gamma_0, \Gamma_1, \ldots$ and $\Gamma_{-1}, \Gamma_{-2}, \ldots$ are non-stationary (delayed) renewal processes and translation invariance implies that $(\Gamma_n)_{n\in\mathbb{Z}}$ is a stationary renewal process. □

With this result we can already prove the strong law of large numbers.

Proof of Theorem 3.1: As above, let $r(n)$ be the label of the last renewal point to the left of n, so that $\Gamma_{r(n)} < n \leq \Gamma_{r(n)+1}$. Then if $n \geq \Gamma_0$,

$$w_{0,\Gamma_0} + \sum_{i=1}^{r(n)} w_{\Gamma_{i-1},\Gamma_i} \leq w_{0,n} \leq w_{0,\Gamma_0} + \sum_{i=1}^{r(n)+1} w_{\Gamma_{i-1},\Gamma_i}. \tag{3.21}$$

First we find a linear upper bound for $w_{0,n}$. Since the edges in the path from 0 to n cannot overlap, and the sum of their lengths is n, we have

$$w_{0,n} \leq n + \sum_{0\leq x<y\leq n} [v_{x,y} - (y-x)]_+$$

$$\leq n + \sum_{0 \leq x < n} Z_x$$

where we define $Z_x = \sum_{y>x}[v_{x,y} - (y-x)]_+$. Note that Z_x are i.i.d. and non-negative with

$$\begin{aligned}\mathbb{E}Z_x &= \mathbb{E}\sum_{y>0}[v_{0,y} - y]_+ \\ &\leq \frac{1}{2}\mathbb{E}v^2 \\ &< \infty.\end{aligned}$$

So $\limsup w_{0,n}/n < \infty$ a.s. and so, from the left-hand inequality in (3.21), we also have

$$\limsup \frac{1}{n}\sum_{i=1}^{r(n)} w_{\Gamma_{i-1},\Gamma_i} < \infty \qquad \text{a.s.}$$

But $r(n)/n \to \lambda$ a.s. as $n \to \infty$, and the terms $w_{\Gamma_{i-1},\Gamma_i}$ are i.i.d. and non-negative for $i \geq 1$. So $\mathbb{E}w_{\Gamma_{i-1},\Gamma_i}$ must be finite. Then finally using again the fact that $r(n)/n \to \lambda$ a.s., and the law of large numbers on both sides of (3.21), we have $w_{0,n}/n \to \lambda^{-1}\mathbb{E}w_{\Gamma_{i-1},\Gamma_i}$ a.s. □

Remark 3.19. *For $p = 1$ the result follows directly from Kingman's subadditive ergodic theorem, since $w_{0,n}$ is superadditive (but for $p < 1$ we may have $w_{0,n} = -\infty$).*

Remark 3.20. *One can show that for non-constant weights there is a strict inequality $C > \widehat{C}\mathbb{E}[v]$ where $\widehat{C}$ is the constant corresponding to C in the case where $v \equiv 1$.*

In order to prove the central limit theorem, we will need to establish that $\Gamma_1 - \Gamma_0$, the length of a typical renewal interval, has finite variance. By general results about renewal processes (see for example Chapter 1, Section 4 in [2], in particular Remark 4.2.1), this is equivalent to the property that the "residual renewal time" Γ_0 has finite expectation. In order to obtain that $\mathbb{E}[\Gamma_0]$ is finite, an additional condition on the distribution of v is required; instead of just a second moment we need that the third moment of v is finite.

Lemma 3.21. *Suppose* $\mathbb{E}[v^3] < \infty$. *If*

$$\gamma \operatorname{ess\,inf}[v] < c < \gamma \mathbb{E}\left[\min_{\tau_0 \le i < j \le \tau_1} v_{i,j}\right], \tag{3.22}$$

then $\mathbb{E}[\Gamma_0] < \infty$.

Proof. Recall that the τ_r are the points of the renewal process of strongly connected points, defined at the beginning of Section 3.3.2, with $\dots < \tau_{-1} < 0 \le \tau_0 < \tau_1 < \dots$. So (τ_0, τ_1) is a typical renewal interval. γ is the density of strongly connected points. Since the process of strongly connected points is independent of the weights $v_{i,j}$, and the weight distribution is not a.s. constant, the RHS of (3.22) is strictly greater than the LHS so the set of "good" values of c is non-empty. Also note that (3.22) implies (3.16), so the conclusion of Lemma 3.16 applies.

We will use an algorithmic construction of Γ_0 similar to the construction in [14] to prove that the expectation $\mathbb{E}[\Gamma_0]$ is finite. Here we will not construct Γ_0 itself, but an upper bound for it. We will use the following events $A^{++}_{x,d}$, $A^{-+}_{x,d}$ and $A^{--}_{x,d}$ that are similar to A^{++}_x, A^{-+}_x and A^{--}_x but restricted to certain regions:

$$A^{++}_{x,d} = \bigcap_{l=1}^{d} \{w_{x,x+l} \ge cl\},$$

$$A^{-+}_{x,d} = \bigcap_{1 \le l \le d, j \ge 1} \{\alpha_{x-j,x+l} v_{x-j,x+l} < c(l+j)\}$$

and

$$A^{--}_{x,d} = \bigcap_{j=1}^{d} \{w_{x-j,x} \ge cj\}.$$

We now introduce another process $\mathcal{U}$ related to the renewal process $\mathcal{R}$. Define

$$\mathcal{U} = \{x \in \mathbb{Z} : A^{--}_x \text{ holds}\}. \tag{3.23}$$

A point in $\mathcal{U}$ clearly has to be connected to every point to its left. In [14] the authors refer to points that are connected to every point to their left as *silver points*. We immediately have $\mathcal{R} \subseteq \mathcal{U}$.

We will write $\ldots < \rho_{-2} < \rho_{-1} < 0 \leq \rho_0 < \rho_1 < \ldots$ for the sequence of points in $\mathcal{U}$, where ρ_0 is the smallest non-negative element of $\mathcal{U}$.

The following result about $\mathcal{U}$ is analogous to Lemma 3.18 about $\mathcal{R}$, but is much more straightforward to prove. For $k \in \mathbb{Z}$, define

$$\mathcal{D}_k = \left(\rho_k - \rho_{k-1}, v_{\rho_{k-1}+n,\rho_{k-1}+i}, \alpha_{\rho_{k-1}+n,\rho_{k-1}+i} : 0 \leq n < i \leq \rho_k - \rho_{k-1}\right).$$

Lemma 3.22. *The cycles* $(\mathcal{D}_k, k \in \mathbb{Z})$ *are independent and* $(\mathcal{D}_k, k \in \mathbb{Z} - \{0\})$ *are identically distributed. The process* $\mathcal{U} = (\rho_n)_{n\in\mathbb{Z}}$ *forms a stationary renewal process.*

Proof. Note that if A_x^{--} holds, and $y > x$, then A_y^{--} holds if and only if $A_{y,y-x}^{--}$ holds. Hence given $x \in \mathcal{U}$, we can find the next $y > x$ such that $y \in \mathcal{U}$ by finding the smallest $y > x$ such that $A_{y,y-x}^{--}$ holds, and to determine whether the event $A_{y,y-x}^{--}$ holds we only have to consider edges with both endpoints in the interval $[x, y]$. The regenerative structure described in Lemma 3.22 follows immediately. □

Next we define

$$\mu = \inf\left\{d > 0 : \mathbb{1}_{A_{0,d}^{-+}\cap A_{0,d}^{++}} = 0\right\}.$$

The random variable μ is the smallest distance $d > 0$ such that at least one of $A_{0,d}^{-+}$ and $A_{0,d}^{++}$ fails. Note that μ may be infinite; this is the case precisely if A_0^{-+} and A_0^{++} hold, so that

$$\beta \overset{def}{=} \mathbb{P}[\mu = \infty] = \mathbb{P}[A_0^{-+} \cap A_0^{++}] > 0.$$

The idea of the proof can best be explained using Figure 3.2 below. We define $\sigma_0 = \rho_0$.

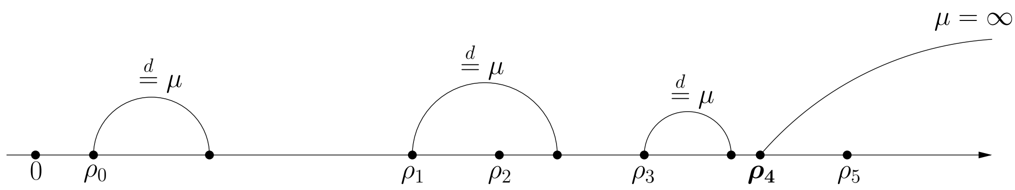

Figure 3.2: Construction of the process $(\rho_n)_{n\in\mathbb{Z}}$ and eventually, using μ, an upper bound for Γ_0, which is in this case given by ρ_4

Now recursively, for each $k \geq 0$ we define

$$\mu_k = \theta_{\sigma_k}\mu = \inf\left\{d > 0 : \mathbb{1}_{A^{-+}_{\sigma_k,d} \cap A^{++}_{\sigma_k,d}} = 0\right\},$$

and

$$\sigma_{k+1} = \inf\{x \in \mathcal{U} : x \geq \sigma_k + \mu_k\}.$$

The set $\{\sigma_0, \sigma_1, \ldots\}$ is a subset of $\{\rho_0, \rho_1, \ldots\}$. We continue until we reach a K such that μ_K is infinite. Then the corresponding σ_K must be a point of $\mathcal{R}$. For certainly $\sigma_K \in \mathcal{U}$, so the event $A^{--}_{\sigma_K}$ holds. But also $\mu_K = \infty$, which by definition of μ_k implies also that $A^{-+}_{\sigma_K}$ and $A^{++}_{\sigma_K}$ hold.

In particular $\Gamma_0 \leq \sigma_K$, which will serve as the upper bound we require.

Now it also follows from the regenerative properties in Lemma 3.22 above that the random variables μ_k are i.i.d., and their common distribution is the same as that of μ. So $K = \inf\{k : \mu_k = \infty\}$ is a geometric random variable with parameter β. Also, given K, the random variables $\mu_k, 0 \leq k < K$ are i.i.d. and their common distribution is that of μ conditioned on $\mu < \infty$ (in particular this does not depend on K).

Since each renewal interval (ρ_{j-1}, ρ_j) has length at least one, we also have that $\sigma_k \leq \rho_L$ where $L = \sum_{j=0}^{K-1} \mu_j$.

We can write ρ_L in the following way

$$\rho_L = \rho_0 + \sum_{j=1}^{L} \rho_j - \rho_{j-1} \tag{3.24}$$

with i.i.d. $\rho_j - \rho_{j-1}$, $j = 1, 2, \ldots$. We will use the following proposition to show that the expectation of ρ_L is finite.

Proposition 3.23. *Let $X_1, X_2, X_3, \ldots$ be an i.i.d. sequence of non-negative random variables with finite variance and let N be a non-negative integer valued random variable with finite mean. Then the expectation of $S_N = X_1 + \ldots + X_N$ is finite.*

Proof. For $a > \mathbb{E}[X_1]$, the expectation of

$$R_a = \sup_{n\in\mathbb{N}} (S_n - an)$$

is finite whenever X_1 has finite variance. This result is familiar in the context of queueing theory, saying that the expected waiting time in a single-server queue is finite if the service time distribution has finite variance (see for example Section 2.2 in [2]). Therefore,

$$\mathbb{E}[S_N] \leq \mathbb{E}[R_a] + a\mathbb{E}[N]$$

is finite. □

It is therefore enough to show that $\mathbb{E}[L] < \infty$ and $\mathbb{E}\left[(\rho_1 - \rho_0)^2\right] < \infty$ (note that $\mathbb{E}[\rho_0] < \infty$ if $\mathbb{E}\left[(\rho_1 - \rho_0)^2\right] < \infty$; this is the same renewal process result we quoted just before Lemma 3.21.). The expectation of L is finite if $\mathbb{E}[\mu|\, \mu < \infty] < \infty$ and this will be proved in Proposition 3.25. In order to show that $\mathbb{E}\left[(\rho_1 - \rho_0)^2\right] < \infty$ we will show that the following random variable ν, which satisfies $\nu \overset{d}{=} \rho_1 - \rho_0$, has exponential moments for appropriate c:

$$\nu = \inf\left\{x > 0 : \mathbb{1}_{A^{--}_{x,x}} = 1\right\}.$$

Proposition 3.24. *If c satisfies (3.22) then*

$$\mathbb{E}[e^{\alpha\nu}] < \infty \qquad \textit{for some } \alpha > 0.$$

Proof. As above the τ_k are the strongly connected points with $\tau_{-1} < 0 \leq \tau_0$, and $m(x)$ satisfies $\tau_{m(x)} < x \leq \tau_{m(x)+1}$. We immediately have

$$\begin{aligned}\mathbb{P}[\nu > x] &\leq \mathbb{P}\left[\nu > \tau_{m(x)}\right] \\ &\leq \mathbb{P}\left[\left(A^{--}_{\tau_0,\tau_0}\right)^c \cap \ldots \cap \left(A^{--}_{\tau_{m(x)},\tau_{m(x)}}\right)^c\right]. \end{aligned} \tag{3.25}$$

Now we claim that if none of the events $A^{--}_{\tau_0,\tau_0}, \ldots, A^{--}_{\tau_{m(x)},\tau_{m(x)}}$ occur, then for $k =$

$0, 1, \ldots, m(x)$,

$$\sum_{r=0}^{k} \min_{\tau_{r-1} \le i \le j \le \tau_r} v_{i,j} < c(\tau_k - \tau_{-1}). \tag{3.26}$$

For suppose (3.26) fails for some value $k \ge 0$ (but is true for all smaller values). Then by subtraction,

$$\sum_{r=a}^{k} \min_{\tau_{r-1} \le i \le j \le \tau_r} v_{i,j} \ge c(\tau_k - \tau_{a-1}) \qquad \forall 0 \le a \le k. \tag{3.27}$$

In that case suppose $0 \le x < \tau_k$. For some a with $0 \le a \le k$ we have $\tau_{a-1} \le x < \tau_a$.

Since the τ_r are strongly connected points, there exists a path from x to τ_k which passes through all of $\tau_a, \tau_{a+1}, \ldots, \tau_k$, and which therefore includes at least one edge within each interval $[\tau_{a-1}, \tau_a]$, $[\tau_a, \tau_{a+1}], \ldots, [\tau_{k-1}, \tau_k]$.

From (3.27), this path must have weight at least $c(\tau_k - \tau_{a-1})$, which is at least $c(\tau_k - x)$.

Since this holds for all $0 \le x \le \tau_k$, it follows that the event $A^{--}_{\tau_k, \tau_k}$ would have to hold.

So indeed the event on the RHS of (3.25) implies (3.26), and so in particular we have

$$\mathbb{P}(\nu > x) \le \mathbb{P}\left(\sum_{r=0}^{m(x)} \min_{\tau_{r-1} \le i \le j \le \tau_r} v_{i,j} < c\big(\tau_{m(x)} - \tau_{-1}\big) \right). \tag{3.28}$$

Since the strongly connected points form a renewal process whose intervals have exponential moments (see (3.14)), we have, for any $\epsilon > 0$,

$$\mathbb{P}\left(\frac{m(x)}{x} < \gamma - \epsilon \right) \le c_1 e^{-c_2 x}, \tag{3.29}$$

$$\mathbb{P}\left(\frac{\tau_m(x) - \tau_{-1}}{x} > 1 + \epsilon \right) \le c_2 e^{-c_4 x}, \tag{3.30}$$

$$\tag{3.31}$$

for some constants c_1, c_2, c_3, c_4 and all $x \in \mathbb{Z}$.

But the quantities $\min_{\tau_{r-1} \le i < j \le \tau_r} v_{i,j}$ are non-negative, and i.i.d. for $r \ge 1$, and we have assumed that $c < \gamma\mathbb{E}[\min_{\tau_0 \le i < j \le \tau_1} v_{i,j}]$. Hence for sufficiently small ϵ and some c_5, c_6,

$$\mathbb{P}\left(\sum_{r=0}^{\lfloor(\gamma-\epsilon)x\rfloor} \min_{\tau_{r-1} \le i < j \le \tau_r} v_{i,j} < c(1+\epsilon)x\right) \le c_5 e^{-c_6 x}.$$

Putting all these together with (3.28), we get that $\mathbb{P}[\nu > x]$ decays exponentially in x, as desired. □

Next we want to prove that the expectation of μ, conditioned on $\{\mu < \infty\}$ is also finite under suitable moment conditions for v.

Proposition 3.25. *If $\mathbb{E}[v^3] < \infty$ and c satisfies (3.22), then*

$$\mathbb{E}[\mu | \, \mu < \infty] < \infty.$$

Proof. We have for $d > 0$

$$\begin{aligned}
\mathbb{P}[\mu = d] &= \mathbb{P}\left[\left(A^{-+}_{0,d} \cap A^{++}_{0,d}\right)^c \cap \left(A^{-+}_{0,d-1} \cap A^{++}_{0,d-1}\right)\right] \\
&\le \mathbb{P}\left[\left(A^{-+}_{0,d}\right)^c \cap A^{-+}_{0,d-1} \cap A^{++}_{0,d-1}\right] \\
&\quad + \mathbb{P}\left[\left(A^{++}_{0,d}\right)^c \cap A^{-+}_{\tau_0,d-1} \cap A^{++}_{\tau_0,d-1}\right] \\
&\le \mathbb{P}\left[\left(A^{-+}_{0,d}\right)^c \cap A^{-+}_{0,d-1}\right] + \mathbb{P}\left[\left(A^{++}_{0,d}\right)^c \cap A^{++}_{0,d-1}\right] \\
&\le \mathbb{P}\left[\sup_{i \ge 1}(v_{-i,d} - ci) > cd\right] + \mathbb{P}[w_{0,d} < cd] \\
&\le \sum_{i=1}^{\infty} \mathbb{P}[v_{-i,d} > c(d+i)] + \mathbb{P}[w_{0,d} < cd].
\end{aligned}$$

By the same arguments as in the proof of Proposition 3.24 we have that the second probability decays exponentially in d. Looking at the first probability we get

$$\begin{aligned}
\sum_{d=1}^{\infty} d \sum_{i=1}^{\infty} \mathbb{P}[v_{-i,d} > c(d+i)] &= \sum_{l=2}^{\infty} \sum_{j=1}^{l-1} j\mathbb{P}[v > cl] \\
&\le \sum_{l=2}^{\infty} \frac{l^2}{2}\mathbb{P}[v > cl],
\end{aligned}$$

which is finite if $\mathbb{E}\left[v^3\right]$ is finite. Therefore, we get that the condition $\mathbb{E}\left[v^3\right] < \infty$ implies that $\mathbb{E}\left[\mu \mid \mu < \infty\right] < \infty$. □

This completes the proof that $\mathbb{E}\left[\Gamma_0\right] < \infty$ whenever c satisfies (3.22). □

Now we are ready to prove the central limit theorem for $w_{0,n}$ (Theorem 3.5).

Proof of Theorem 3.5: Take any c satisfying (3.22). Since under the condition $\mathbb{E}\left[v^3\right] < \infty$ we have $\mathbb{E}\left[\Gamma_0\right] < \infty$ we also get $\mathbb{E}\left[|\Gamma_{-1}|\right] < \infty$. This implies that the variance of $\Gamma_1 - \Gamma_0$ is finite (since the Γ_n form a stationary renewal process, see Remark 4.2.1 in [2]). Now we want to show that $\sigma^2 = \mathrm{Var}\left(w_{\Gamma_0,\Gamma_1} - C\left(\Gamma_1 - \Gamma_0\right)\right)$ is finite. We will prove this in a separate proposition.

Proposition 3.26. *If* $\mathbb{E}\left[v^3\right] < \infty$ *then* $\mathrm{Var}\left(w_{\Gamma_0,\Gamma_1} - C\left(\Gamma_1 - \Gamma_0\right)\right) < \infty$.

Proof. In order to show that the variance of $w_{\Gamma_0,\Gamma_1} - C\left(\Gamma_1 - \Gamma_0\right)$ is finite, it is enough to show that the second moment of this random variable is finite.

$$\begin{aligned}
&\mathbb{E}\left[\left(w_{\Gamma_0,\Gamma_1} - C(\Gamma_1 - \Gamma_0)\right)^2\right] \\
&\quad = \mathbb{E}\left[\left(w_{\Gamma_0,\Gamma_1} - C(\Gamma_1 - \Gamma_0)\right)^2 \mathbb{1}_{\left\{w_{\Gamma_0,\Gamma_1} \geq C(\Gamma_1 - \Gamma_0)\right\}}\right] \\
&\qquad\quad + \mathbb{E}\left[\left(w_{\Gamma_0,\Gamma_1} - C(\Gamma_1 - \Gamma_0)\right)^2 \mathbb{1}_{\left\{w_{\Gamma_0,\Gamma_1} < C(\Gamma_1 - \Gamma_0)\right\}}\right] \\
&\quad \leq \mathbb{E}\left[\left(\max_{\Gamma_0 = i_0 < j_0 = i_1 < j_1 = \ldots < j_m = \Gamma_1} \sum_{l=0}^{m} \left[v_{i_l,j_l} - C(j_l - i_l)\right]_+\right)^2\right] \\
&\qquad\quad + \mathbb{E}\left[C^2(\Gamma_1 - \Gamma_0)^2\right].
\end{aligned}$$

Under the assumption $\mathbb{E}\left[v^3\right] < \infty$ we know that the second expectation is finite. Therefore we will only consider the first expectation in the following. For the first expectation we get

$$\mathbb{E}\left[\left(\max_{\Gamma_0 = i_0 < j_0 = i_1 < j_1 = \ldots < j_m = \Gamma_1} \sum_{l=0}^{m} \left[v_{i_l,j_l} - C(j_l - i_l)\right]_+\right)^2\right] \tag{3.32}$$

$$\leq \mathbb{E}\left[\sum_{\Gamma_0 \leq x < y \leq \Gamma_1} [v_{x,y} - C(y-x)]_+^2\right] \tag{3.33}$$

$$+ 2\mathbb{E} \sum_{\Gamma_0 \leq x < y \leq u < z \leq \Gamma_1} [v_{x,y} - C(y-x)]_+ [v_{u,z} - C(z-u)]_+ . \tag{3.34}$$

We will look at the expectations in (3.33) and (3.34) separately. For the first one, we can use that the expected length of a typical renewal interval is λ^{-1} to give

$$\begin{aligned}
\mathbb{E}\left[\sum_{\Gamma_0 \leq x < y \leq \Gamma_1} [v_{x,y} - C(y-x)]_+^2\right] &\leq \mathbb{E}\left[\sum_{\Gamma_0 \leq x < \Gamma_1} \sum_{y > x} [v_{x,y} - C(y-x)]_+^2\right] \\
&= \lambda^{-1} \sum_{y>0} \mathbb{E}\left[[v_{0,y} - Cy]_+^2\right] \\
&\leq const \cdot \mathbb{E}\left[v^3\right],
\end{aligned}$$

and so the expectation in (3.33) is finite.

Let us now look at the second expectation, for which we have to sum over pairs of edges that are in the same renewal interval. For $i \leq j$ let $R_{i,j}$ be the event that the set $\{i, i+1, \ldots, j\}$ contains at least one renewal point. Note that $\mathbb{P}(R_{i,j}^c) = \mathbb{P}(R_{0,j-i}^c) = \mathbb{P}(\Gamma_0 > j - i)$.

Then define

$$s_{r,n} = \sum_{r \leq x < y \leq u < z \leq n} [v_{x,y} - C(y-x)]_+ [v_{u,z} - C(z-u)]_+ I(R_{x+1,z-1}^c).$$

Notice that the expression in (3.34) is precisely $\mathbb{E} s_{\Gamma_0,\Gamma_1}$; we need to show that this is finite. We first aim to show that the expectation of $s_{0,n}$ grows only linearly with n. To do so we make the following claim, to be proved below: for any $x < y \leq u < z$, and any $s, t \geq 0$,

$$\mathbb{P}\left(v_{x,y} \geq t, v_{u,z} \geq s\right) \geq \mathbb{P}\left(v_{x,y} \geq t, v_{u,z} \geq s \middle| R_{y,u}^c\right). \tag{3.35}$$

In that case

$$\begin{aligned}
\mathbb{E}s_{0,n} &= \mathbb{E}\sum_{0\le x<y\le u<z\le n} [v_{x,y}-C(y-x)]_+ [v_{u,z}-C(z-u)]_+ I(R^c_{x+1,z-1}) \\
&\le \mathbb{E}\sum_{0\le x<y\le u<z\le n} [v_{x,y}-C(y-x)]_+ [v_{u,z}-C(z-u)]_+ I(R^c_{y,u}) \\
&= \sum_{0\le x<y\le u<z\le n} \mathbb{E}\left([v_{x,y}-C(y-x)]_+ [v_{u,z}-C(z-u)]_+ \middle| R^c_{y,u}\right)\mathbb{P}(R^c_{y,u}) \\
&\le \sum_{0\le x<y\le u<z\le n} \mathbb{E}\left([v_{x,y}-C(y-x)]_+ [v_{u,z}-C(z-u)]_+\right)\mathbb{P}(R^c_{y,u}) \\
&\le n\sum_{0<y\le u<z} \mathbb{E}\,[v_{0,y}-Cy]_+\,\mathbb{E}\,[v_{u,z}-C(z-u)]_+\,\mathbb{P}(\Gamma_0>u-y) \\
&= n\mathbb{E}\Gamma_0\left(\mathbb{E}\sum_{y>0}[v_{0,y}-Cy]_+\right)^2.
\end{aligned}$$

By Lemma 3.21, this gives $\mathbb{E}s_{0,n}=O(n)$ whenever $\mathbb{E}(v^3)<\infty$.

Now note that for $n\ge\Gamma_0$ we have

$$\frac{1}{n}\left(s_{0,\Gamma_0}+\sum_{i=1}^{r(n)} s_{\Gamma_{i-1},\Gamma_i}\right)\le\frac{1}{n}s_{0,n} \tag{3.36}$$

where as before, we write $r(n)$ for the label of the last renewal point to the left of n, so that $\Gamma_{r(n)}<n\le\Gamma_{r(n)+1}$.

By Lemma 3.18, the quantities $s_{\Gamma_{i-1},\Gamma_i}$ are i.i.d. for $i\ge 1$. Hence, if s_{Γ_0,Γ_1} had infinite mean, then the left-hand side of (3.36) would converge to infinity almost surely, since $r(n)/n\to\lambda$ a.s. But then also $s_{0,n}/n$ would converge to infinity almost surely, which contradicts the fact that $\mathbb{E}s_{0,n}/n$ is bounded. Hence s_{Γ_0,Γ_1} has finite mean, which is to say that the expectation in (3.34) is finite.

This completes the proof, subject to the claim (3.35) which we now justify. We consider the dependence of the event $R_{y,u}$ on the weights $v_{x,y}$ and $v_{u,z}$.

Take $r\in\{y,y+1,\dots,u\}$. From the definition of the set of renewal points $\mathcal{R}$, it's easy to see that the event $\{r\in\mathcal{R}\}$ is an increasing event as a function of $v_{x,y}$ and

$v_{u,z}$; that is, if $r \in \mathcal{R}$ and we increase the values of $v_{x,y}$ or $v_{u,z}$ while leaving all other weights the same, then it remains the case that $r \in \mathcal{R}$. But $R_{y,u} = \bigcup_{y \le r \le u}\{r \in \mathcal{R}\}$, so $R_{y,u}$ is also an increasing event as a function of $v_{x,y}$ and $v_{u,z}$.

Since these events depend only on the weights $v_{i,j}$ and the indicator variables $\alpha_{i,j}$ determining which edges are present, and since these quantities are all independent, it follows that the distribution of $(v_{x,y}, v_{u,z})$ conditioned on $R_{y,u}$ dominates the unconditioned distribution, which is equivalent to (3.35). □

With Proposition 3.26 established, the rest of the argument to prove the central limit theorem in Theorem 3.5 is analogous to that in [14], see proof of Theorem 2 (pp. 20-22), using Donsker's theorem and the continuous mapping theorem (and the fact that the fraction of renewal points between 0 and $[nt]$ converges to the deterministic function λt). □

3.3.4 Length of the longest edge

In this section we analyse the asymptotic behaviour of ℓ_n and h_n, the length of the longest edge and the weight of the heaviest edge used on the geodesic between 0 and n, and prove Theorem 3.6.

We are working under the assumption (3.6) that F is regularly varying with index s. Define $f(x)$ by $1 - F(x) = x^{-s}f(x)$, so that f is a slowly varying function, i.e. $f(tx)/f(x) \to 1$ as $x \to \infty$, for any $t > 0$.

We start with some general results about regularly varying functions that will be useful throughout this section. Let $g(z) = z^{-s}f(z)$ be a regularly varying function with index $s > 1$ (i.e. f is slowly varying). Then we have

$$\int_x^\infty g(z)dz \sim \frac{x^{-s+1}}{s-1}f(x). \tag{3.37}$$

See for example Proposition 1.5.10 in [4]. From the representation theorem (Theorem 1.3.1 in [4]) it follows that we can choose a function $r_0(x)$, depending on f, that increases to infinity but does so slowly enough that

$$\sup_{x\le y\le xr_0(x)} \frac{f(y)}{f(x)} \xrightarrow{x\to\infty} 1;$$

then for this r_0

$$\int_x^{xr_0(x)} g(z)dz \sim \frac{x^{-s+1}}{s-1} f(x). \tag{3.38}$$

From (3.37) and (3.38) we get that for any function $r(x) \le \infty$ such that $r(x) \to \infty$ as $x \to \infty$ the following holds

$$\int_x^{xr(x)} g(z)dz \sim \frac{x^{-s+1}}{s-1} f(x). \tag{3.39}$$

Proof of Theorem 3.6: We start by proving the limit in (3.7). We start with an upper bound, and aim to show that if $\beta > 1/(s-1)$, the optimal path is unlikely to use an edge as long as n^β.

Upper bound. For any edge $e = (x, y)$, write $|e| = y - x$ for the length of the edge. Write $w^-_{i,j}$ for the maximal weight of a path from i to j not using the edge (i, j) itself.

Lemma 3.27. *Fix $\beta \in (0, 1)$.*

(i) For some c_1 and $M > 0$,

$$\mathbb{P}(w^-_{0,m} \le mM) \le e^{-c_1 m} \text{ for all } m.$$

(ii) For some c_2 and $M > 0$, for all n,

$$\mathbb{P}\left(w^-_{x,y} \le M(y-x) \text{ for some } 0 \le x < y \le n \text{ with } y - x \ge n^\beta\right) \le e^{-c_2 n^\beta},$$

and

$$\mathbb{P}\big(\text{The geodesic from 0 to } n \text{ uses an edge } e \text{ with } |e| \ge n^\beta \text{ and } v_e \le M|e|\big) \le e^{-c_2 n^\beta}. \tag{3.40}$$

Proof. Property (i) is immediate for $p = 1$, since the quantity $w^-_{0,m}$ is bounded below by the sum of m i.i.d. non-negative random variables. For $p < 1$ we can do something analogous using the strongly connected points. From (3.14), the distance between successive strongly connected points has an exponentially decaying tail, and so the probability that there exist fewer than $m\gamma/2$ strongly connected points between 0 and m decays exponentially in m, where γ is the density of strongly connected points. If there exist at least $m\gamma/2$ such points, then there is a path from 0 to m containing at least $m\gamma/2$ edges.

But the weights are i.i.d. bounded below and with positive mean, so for appropriately chosen M the probability that their sum is less than Mm decays exponentially, as required for (i).

For (ii), simply sum (i) over all appropriate values of x and y. The second part of (ii) also follows, since an edge (x, y) with $v_{x,y} < w^-_{x,y}$ will never be used in an optimal path. □

Choose M according to Lemma 3.27. Now let N_β be the number of edges e in the interval $[0, n]$ such that $|e| \geq n^\beta$ and $v_e > M|e|$. From the last part of Lemma 3.27,

$$\mathbb{P}(N_\beta = 0 \text{ and } \ell_n \geq n^\beta) \leq e^{-c_2 n^\beta}. \tag{3.41}$$

But also

$$\mathbb{E}(N_\beta) \leq \sum_{k=n^\beta}^{n} n\mathbb{P}\left(v > Mk\right),$$

since there are at most n edges of any given length k in the interval $[0, n]$. Since $\mathbb{P}(v > Mk) \sim k^{-s} f(k)$ by assumption, we can use (3.39) with $x = n^\beta$ and $r(x) = x^{1/\beta}$ to get

$$\mathbb{E}(N_\beta) \leq const \cdot n^{1-\beta(s-1)} f(n^\beta).$$

Since f is slowly varying, this tends to 0 as $n \to \infty$ whenever $\beta > 1/(s-1)$. Hence for all such β, $\mathbb{P}(N_\beta > 0) \to 0$ as $n \to \infty$; combining with (3.41) we have

$$\mathbb{P}\left(\frac{\log \ell_n}{\log n} \geq \beta\right) \to 0 \text{ as } n \to \infty \text{ for all } \beta > \frac{1}{s-1},$$

as required for the upper bound.

Lower bound. Fix $K > 0$ (to be chosen later) and let R_β be the number of edges within the interval $\left(\left\lceil \frac{2n}{5} \right\rceil, \left\lfloor \frac{3n}{5} \right\rfloor\right)$ which satisfy $v_e \geq K|e|$ and $|e| \geq n^\beta$. Then

$$\begin{aligned}\mathbb{E}(R_\beta) &\geq \sum_{k=n^\beta}^{n/12} \frac{n}{12} \mathbb{P}\left(\alpha_{0,1} v \geq Kk\right) \\ &\geq \mathit{const} \cdot n^{1-\beta(s-1)} f(n^\beta). \end{aligned} \tag{3.42}$$

The first inequality holds since for any k with $n^\beta \leq k \leq n/12$, there are at least $n/12$ edges of length k within $\left(\left\lceil \frac{2n}{5} \right\rceil, \left\lfloor \frac{3n}{5} \right\rfloor\right)$, and the second follows again from (3.39).

The RHS of (3.42) tends to infinity as $n \to \infty$ if $\beta < 1/(s-1)$. Since the corresponding events for different edges e are independent, we obtain that $\mathbb{P}(R_\beta \geq 1) \to 1$, i.e. with high probability, at least one such edge exists.

If so, let $e^* = (x^*, y^*)$ be the longest such edge. Then define the interval $I^* = (x^* - 2|e^*|, y^* + 2|e^*|)$, which is centred on e but is five times as long. Note that I^* is still contained in $[0, n]$. Finally, let w^* be the maximal weight of a path contained in the interval I^* which uses only edges shorter than e^*.

We claim that, if $R_\beta \geq 1$, then at least one of the following events must hold:

(a) Some edge at least as long as e^* (maybe e^* itself) is used in the optimal path from 0 to n.

(b) $w^* \geq K|e^*|$.

(c) There is either no strongly connected point in $(x^* - |e^*|, x^*)$ or there is no strongly connected point in $(y^*, y^* + |e^*|)$.

For if (b) does not hold, then using the edge e^* is preferable to any combination of edges in I^* which are shorter than e^*. If in addition (c) fails, then using appropriate strongly connected points one can include the edge e^* simultaneously with any edge set of compatible edges which are wholly to the left of $x^* - |e^*|$ or wholly to the right of $y^* + |e^*|$. Then the only reason not to use e^* is if the optimal path contains an edge (r, s) where either $r < x^* - 2|e^*|$, $s > x^* - |e^*|$ or $r < y^* + |e^*|$, $s > y^* + 2|e^*|$. But such an edge has length at least e^*. So indeed (a) then holds.

We already have $\mathbb{P}(R_\beta \geq 1) \to 1$ as $n \to \infty$. Since the strongly connected points form a renewal process with positive density and the renewal intervals have exponential tails (see (3.14)), and $|e^*| \geq n^\beta$, event (c) has probability tending to 0 as $n \to \infty$.

If we can show that the probability of event (b) also goes to 0, then with probability tending to 1, event (a) occurs. Then indeed $\ell_n \geq n^\beta$, and we will have shown that for any $\beta < 1/(s-1)$, $\mathbb{P}(\ell_n \geq n^\beta) \to 1$ as $n \to \infty$ as required.

So, we need to prove the following:

Claim: for appropriate K, $\mathbb{P}(R_\beta \geq 1, w^* \geq K|e^*|) \to 0$ as $n \to \infty$.

Suppose $R_\beta \geq 1$, and condition on the identity of the edge e^*. Let $m = |e^*|$. From the definition of e^*, knowing the identity of e^* has given us no information about the weights of edges shorter than m. Then since I^* has length $5m$, the distribution of w^* is dominated by the distribution of $w_{0,5m}$ in the case $p = 1$. But the SLLN in that case gives $\frac{1}{5m} w_{0,5m} \to C^{(p=1)}$ in probability, for some constant $C^{(p=1)}$. Hence the claim holds for any $K > C^{(p=1)}$.

This completes the argument for the longest edge ℓ_n, and we can use those results to give the corresponding statements for the heaviest weight h_n.

The lower bound follows immediately from the bound for ℓ_n and property (3.40). For the upper bound, suppose $\beta > 1/(s-1)$. Take $\beta' \in (1/(s-1), \beta)$. We know that as $n \to \infty$, the probability that the optimal path uses an edge as long as $n^{\beta'}$ tends to 0. But also the probability that there exists an edge of length less than $n^{\beta'}$ with weight as high as n^β is bounded above by

$$\sum_{k=1}^{n^{\beta'}} n\mathbb{P}\left(v \geq n^\beta\right) \leq const \cdot n^{1+\beta'} n^{-\beta s} f(n^\beta)$$

which converges to 0 as $n \to \infty$. So indeed the probability that an edge as heavy as n^β is used goes to 0, as required.

Now we turn to the fluctuations of $w_{0,n}$ when $2 < s < 3$. Suppose $\beta < 1/(s-1)$, and choose $\beta' \in (\beta, 1/(s-1))$. Let $\bar{e}$ be the heaviest edge used in the optimal path from 0 to n. We know from above that with high probability $v_{\bar{e}} > n^{\beta'}$.

Condition on the identity of $\bar{e}$ and the weight of all the other edges in $[0, n]$, but not the weight of $\bar{e}$ itself. Write $\mathcal{A}$ for the collection of all this information.

Given $\mathcal{A}$, we have a lower bound, $V_{\min}$ say, for the weight of $\bar{e}$ (since given the weights of all other edges, $\bar{e}$ will be the heaviest weight in the optimal path if and only if its weight exceeds some threshold). Now the conditional distribution of $v_{\bar{e}}$ given $\mathcal{A}$ is the distribution of a typical weight v conditioned on $v > V_{\min}$. Given $\mathcal{A}$, the value of $w_{0.n} - v_{\bar{e}}$ is constant.

Certainly we will have $V_{\min} > n^{\beta'}/2$ with high probability. For if $V_{\min} \leq n^{\beta'}/2$, then $\mathbb{P}(v_{\bar{e}} \leq n^{\beta'} | \mathcal{A}) \geq \mathbb{P}(v \leq n^{\beta'} | v \geq n^{\beta'}/2)$, which does not go to 0 as $n \to \infty$ (since the tail of the distribution of v is regularly varying, so that $P(v \leq 2x | v \geq x)$ converges to a non-zero limit as $x \to \infty$).

But again since v has a regularly varying tail, we have that $\mathbb{P}(v \in [x_n, x_n + n^\beta] | v \geq v_{\min})$ goes to 0 as $n \to \infty$ uniformly in x_n and in $v_{\min} > n^{\beta'}/2$, since $n^\beta = o(n^{\beta'})$.

Hence for some function $\epsilon(n)$ tending to 0 as $n \to \infty$, we have that with high probability as $n \to \infty$,

$$\mathbb{P}(v_{\bar{e}} \in [x_n, x_n + n^\beta] | \mathcal{A}) < \epsilon(n) \text{ for all } x_n,$$

and hence also with high probability

$$\mathbb{P}(w_{0,n} \in [y_n, y_n + n^\beta] | \mathcal{A}) < \epsilon(n) \text{ for all } y_n.$$

Now we can average over $\mathcal{A}$, to give that for any sequence y_n, the unconditional distribution of $w_{0,n}$ satisfies

$$\mathbb{P}(w_{0,n} \in [y_n, y_n + n^\beta]) \to 0 \text{ as } n \to \infty.$$

as required. In particular, with $s < 3$ this implies that $\mathrm{Var}(w_{0,n})$ grows faster than n, and that no central limit theorem such as that in Theorem 3.5 can hold (even for single values of t). □

Now we want to present two examples that show that for $s = 3$ both critical cases are possible: it might happen that the longest edge is $o(\sqrt{n})$ and it is possible that the longest edge satisfies $\frac{\ell_n}{\sqrt{n}} \to \infty$ in probability as $n \to \infty$. Let us first look at the case where $f(x) = \frac{1}{\log x}$. Then we have $\mathbb{E}[v^3] = \infty$.

Example 3.28. *With $\mathbb{P}[v > k] = \frac{1}{k^3 \log k}$ we have $\mathbb{E}[v^3] = \infty$ since $\int_2^\infty \frac{1}{x \log x} dx = \infty$, but on the other hand (again using (3.39))*

$$\begin{aligned}
\mathbb{E}\left[N_{\frac{1}{2}}\right] &\le n \sum_{k=\sqrt{n}}^{n} \frac{1}{(Mk)^3 \log Mk} \\
&\le const \cdot \frac{1}{\log\left(\sqrt{n}\right)} \\
&\xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

So in this case we have that although $\mathbb{E}[v^3] = \infty$ we will not see edges of length $\sqrt{n}$.

However, if f is increasing and such that $\mathbb{E}\left[N_{\frac{1}{2}}\right] \xrightarrow[n\to\infty]{} \infty$ then we have $\frac{\ell_n}{\sqrt{n}} \to \infty$ in probability by the same arguments as in the proof of (3.7) in Lemma 3.6. An example is the case where and $f(x) = \log x$:

Example 3.29. *Let $\mathbb{P}[v > k] = \frac{\log k}{k^3}$. Then the expected number of edges of at least length $\sqrt{n} \log\log n$ and weight at least M times their length is bounded from below by*

$$\sum_{k=\sqrt{n}\log\log n}^{\frac{n}{2}} \frac{n}{2} \mathbb{P}[v > Mk] \ge const \cdot n \left(\sqrt{n} \log\log n\right)^{-2} \cdot \log\left(\sqrt{n} \log\log n\right)$$

$$= const \cdot \frac{\frac{1}{2}\log n + \log\log\log n}{(\log\log n)^2}$$

$$\xrightarrow[n\to\infty]{} \infty.$$

By the same arguments as for the lower bound in the proof of (3.7) in Lemma 3.6 we have that with positive probability we will use an edge of length $\sqrt{n}\log\log n$*, so* $\frac{\ell_n}{\sqrt{n}} \to \infty$ *in probability.*

3.3.5 Non-constant edge probabilities

In this section we want to discuss briefly the situation in which the probabilities that edges are present are not given by a constant $p \in (0,1]$, but by a sequence $(p_i)_{i\geq 1}$ where p_i is the probability that an edge of length i is present. In the case with constant edge weights this situation was analysed in [14], and it can be extended to our case as follows. As in [14] we need the following two conditions:

$$[C1]\ 0 < p_1 < 1$$

$$[C2]\ \sum_{k=1}^{\infty}(1-p_1)\dots(1-p_k) < \infty.$$

Under these conditions the set of strongly connected points is almost surely infinite and the set of strongly connected points forms a stationary renewal process. Since this is all we needed to establish that the set of renewal points is almost surely infinite, conditions $[C1]$ and $C[2]$ are sufficient to get that

$$\mathcal{R} \text{ is almost surely an infinite set.}$$

However, in the proof of the strong law of large numbers and the central limit theorem above, we used that the strongly connected points τ_i have exponential moments. This is in general no longer the case if we replace the constant p by a sequence $(p_i)_{i\in\mathbb{N}}$. In

[14] it was proven that $\mathbb{E}[\tau_0] < \infty$ if the condition

$$[C3] \sum_{k=1}^{\infty} k(1-p_1)\dots(1-p_k) < \infty$$

holds. In the proofs of Propositions 3.24 and 3.25 we used that certain errors, see (3.29) and (3.30), decay exponentially because the τ_i had exponential moments. If we are however only interested in showing that Γ_0 has a finite first moment, then it is enough if these errors decay fast enough to give us finite first moments of ν and μ conditioned on $\mu < \infty$. For these errors to decay fast enough it is sufficient to have two moments of $\tau_1 - \tau_0$ and for this it is enough to have $\mathbb{E}[\tau_0] < \infty$. So under $[C1]$, $[C2]$, $[C3]$ and the condition $\mathbb{E}[v^3] < \infty$ for the weights, we still get a SLLN and a CLT for the weight $w_{0,n}$. This agrees with the results in [14]: if the weights are constant then conditions $[C1]$, $[C2]$ and $[C3]$ give us a SLLN and CLT.

3.4 Proofs for the model with $\mathbb{E}\left[v^2\right] = \infty$

In this section the weights have a distribution which does not have a second moment, i.e. $\mathbb{E}[v^2] = \infty$. We want to prove Theorems 3.7 and 3.8. Again, we start with the case $p = 1$ (Theorem 3.7) and then look at the case $p < 1$ (Theorem 3.8). The proofs follow closely those in [26] where analogous results for directed last-passage percolation in two dimensions were established.

3.4.1 Proof of Theorem 3.7

Proof of Theorem 3.7: To prove Theorem 3.7 we use approximations of $w_{0,n}$ and w that use only the k largest weights. We define

$$\mathcal{C}^k = \{A \subset \{1, 2, \dots, k\} : Y_i \sim Y_j \text{ for all } i, j \in A\}$$

and

$$\mathcal{C}_{0,n}^k = \left\{ A \subset \left\{ 1, 2, \ldots, k \wedge \binom{n+1}{2} \right\} : Y_i^{(n)} \sim Y_j^{(n)} \text{ for all } i, j \in A \right\}$$

and put

$$w^k = \sup_{A \in \mathcal{C}} \sum_{i \in A, i \leq k} M_i,$$

$$w_{0,n}^k = \sup_{A \in \mathcal{C}_{0,n}} \sum_{i \in A, i \leq k} M_i^{(n)}.$$

We also define appropriately rescaled versions

$$\widetilde{w}_{0,n}^k = \frac{w_{0,n}^k}{b_n},$$

$$\widetilde{w}_{0,n} = \frac{w_{0,n}}{b_n}.$$

The tails of w and $w_{0,n}$ are bounded by

$$S^k = \sup_{A \in \mathcal{C}} \sum_{i \in A, i > k} M_i \quad \text{and} \quad S_{0,n}^k = \sup_{A \in \mathcal{C}_{0,n}} \sum_{i \in A, i > k} M_i^{(n)}.$$

The following lemma implies that w is almost surely finite and that $w^k \to w$ for $k \to \infty$.

Lemma 3.30. *With probability 1 we have $S^k < \infty$ for all $k \geq 0$ and $S^k \to 0$ for $k \to \infty$.*

Proof. Define $\Lambda_i = \sup_{A \in \mathcal{C}} |A \cap \{1, \ldots, i\}|$. This is the largest number of the edges $Y_1, \ldots, Y_i$ that can be included simultaneously in an admissible path. This is independent of the weights $(M_i)_{i \in \mathbb{N}}$. In the two-dimensional case in [26] the corresponding random variable L_i had the distribution of the length of the longest increasing subsequence of a random permutation of the set $\{1, \ldots, i\}$. In our case the distribution is slightly different, but we get the same asymptotic behaviour and the same bounds that we need to prove Lemma 3.30. In the two-dimensional case two points $(i, j), (i', j')$ are compatible if

- $i \le i'$ and $j \le j'$ or $i' \le i$ and $j' \le j$.

In our case two edges (represented by two points in $[0,1]^2$) are compatible if

- $i \le \min(i', j')$ and $j \le \min(i', j')$ or $i' \le \min(i, j)$ and $j' \le \min(i, j)$ (under the condition $i < j$ and $i' < j'$ this is equivalent to $j \le i'$ or $j' \le i$).

We can see that the second condition is more restrictive. Therefore, any path that is admissible in our model is also admissible in the two-dimensional model. If we therefore look at the largest number of points $Y_1, \ldots, Y_i$ that we can include in an admissible path in the two models we get that

$$L_i \ge \Lambda_i \text{ a.s.} \tag{3.43}$$

Since we have $\mathbb{E}[L_i] \le c\sqrt{i}$, $\mathbb{E}[L_i^2] \le ci$ and $\frac{L_i}{i^{\frac{1}{s}}} \xrightarrow{d} 0$, for $i \to \infty$ and $s \in (0, 2)$, we get the same results for the corresponding variable Λ_i:

$$\mathbb{E}[\Lambda_i] \le c\sqrt{i},\ \mathbb{E}\left[\Lambda_i^2\right] \le ci,\ \frac{\Lambda_i}{i^{\frac{1}{s}}} \xrightarrow{d} 0 \text{ for } i \to \infty \text{ and } s \in (0, 2).$$

Indeed, Λ_i has been studied in its own right as the "independence number of a random interval graph"; see for example [33] and [8], where, among other things, a central limit theorem and large deviations principle are obtained.

The difference between L_i and Λ_i is the main difference between our model and the two-dimensional nearest-neighbour last-passage percolation model studied in [26]. Since we have the bound (3.43), we can follow the proof of Lemma 3.1 in [26]: put $U_k = \sum_{i=k+1}^{\infty} \Lambda_i (M_i - M_{i+1})$ and for fixed $A \in \mathcal{C}$ define $R_i = |A \cap \{1, \ldots, i\}|$. Then

$$\begin{aligned}
\sum_{i \in A, i > k} M_i &= \lim_{n \to \infty} \sum_{i \in A, k < i \le n} M_i \\
&= \lim_{n \to \infty} \sum_{i=k+1}^{n} M_i \mathbb{1}_{\{i \in A\}} \\
&= \lim_{n \to \infty} \sum_{i=k+1}^{n} M_i (R_i - R_{i-1})
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \left[-M_{k+1}R_k + \sum_{i=k+1}^{n-1} R_i \left(M_i - M_{i+1}\right) + M_n R_n \right] \\
&\le \lim_{n\to\infty} \sum_{i=k+1}^{n-1} R_i \left(M_i - M_{i+1}\right) + \liminf_{n\to\infty} M_n R_n \\
&\le \lim_{n\to\infty} \sum_{i=k+1}^{n-1} \Lambda_i \left(M_i - M_{i+1}\right) + \liminf_{n\to\infty} M_n \Lambda_n \\
&= U_k + \liminf_{n\to\infty} M_n \Lambda_n \\
&= U_k.
\end{aligned}
$$

Therefore, we have $S^k \le U_k$ for all k and it suffices to show that $U_k \to 0$ as $k \to \infty$. Since U_k is the remainder of an infinite sum, it is actually enough to show that $U_k < \infty$ almost surely. By the independence of (Λ_i) and (M_i) we get that

$$
\begin{aligned}
\mathbb{E}\left[U_k\right] &= \sum_{i=k+1}^{\infty} \mathbb{E}\left[\Lambda_i\right] \left(\mathbb{E}\left[M_i\right] - \mathbb{E}\left[M_{i+1}\right]\right) \\
&\le \sum_{i=k+1}^{\infty} c\sqrt{i} \left(\mathbb{E}\left[M_i\right] - \mathbb{E}\left[M_{i+1}\right]\right).
\end{aligned}
$$

Now we can use the known distribution of the M_i (M_i has the distribution of $(V_i)^{-\frac{1}{s}}$ where $V_i \sim \text{Gamma}(i, 1)$) to get

$$
\mathbb{E}\left[U_k\right] \le \frac{c}{s} \sum_{i=k+1}^{\infty} \sqrt{i} \left(i - \frac{1}{s} - 1\right)^{-\frac{1}{s}}
$$

(see Lemma 3.1 of [26] for the details). The last sum is finite for all $k > \frac{1}{s}$ and it follows that $U_k < \infty$ almost surely for all k. □

Theorem 3.7 then follows from the next two propositions which were proved in [26] (Propositions 3.2 and 3.3). The proofs are almost identical and rely on the two following facts:

- The distribution of the weights M_i in [26] and in our model is exactly the same.

- The definition of compatible points/edges is slightly different, but such that $\Lambda_i \leq L_i$.

Proposition 3.31. *Let $\varepsilon > 0$ and k be fixed. Then for all sufficiently large n there exists a coupling of the continuous and the discrete model indexed by n such that*

$$\mathbb{P}\left[\sum_{i=1}^{k} \left|M_i - \widetilde{M}_i^{(n)}\right| > \varepsilon\right] \leq \varepsilon,$$

$$\mathbb{P}\left[\sum_{i=1}^{k} \left\|Y_i - Y_i^{(n)}\right\| > \varepsilon\right] \leq \varepsilon, \tag{3.44}$$

$$\mathbb{P}\left[\mathcal{C}_{0,n}^k \neq \mathcal{C}^k\right] \leq \varepsilon.$$

Here we use the Euclidean distance in $\mathbb{Z}^2$ as distance between two edges $Y_1 = (a, b)$ and $Y_2 = (a', b')$ in (3.44).

Sketch of the Proof: The first two statements follow straightforwardly from the convergence stated in (3.12) and (3.13). The last statement follows from the fact that with high probability, a small perturbation of the Y_i does not affect the ordering of the points. □

Proposition 3.32. *Let $\varepsilon > 0$. Then for sufficiently large k and $\widetilde{S}_{0,n}^k = \frac{S_{0,n}^k}{b_n}$,*

$$\mathbb{P}\left[\widetilde{S}_{0,n}^k > \varepsilon\right] \leq \varepsilon$$

for all n.

Remark 3.33. *A detailed proof of Proposition 3.32 can be found in Section 3.2 of [26]. The transfer to our situation follows again from the two facts stated before Proposition 3.31.*

We can then write

$$|w - \widetilde{w}_{0,n}| = \left|\left(w - w^{k_n}\right) + \left(w^{k_n} - \widetilde{w}_{0,n}^{k_n}\right) + \left(\widetilde{w}_{0,n}^{k_n} - \widetilde{w}_{0,n}\right)\right|$$

$$\leq S^{k_n} + \left|w^{k_n} - \widetilde{w}_{0,n}^{k_n}\right| + \widetilde{S}_{0,n}^{k_n}$$

and for some suitable sequence k_n we have that the first and last term tend to 0 in probability. We also have that on $\mathcal{C}_{0,n}^{k_n} = \mathcal{C}^{k_n}$,

$$\left|w^{k_n} - \widetilde{w}_{0,n}^{k_n}\right| \leq \sum_{i=1}^{k_n} \left|M_i - \widetilde{M}_i^{(n)}\right|$$

holds. Since $\mathbb{P}\left[\mathcal{C}_{0,n}^k \neq \mathcal{C}^k\right] \to 0$ and $\sum_{i=1}^{k_n} \left|M_i - \widetilde{M}_i^{(n)}\right| \to 0$ in probability, we have that $\widetilde{w}_{0,n} \to w$ as required for Theorem 3.7. □

3.4.2 Proof of Theorem 3.8

The proof in the case $p < 1$ goes through in an essentially identical way, after making a couple of appropriate observations.

First, the number of edges in the interval $[0, n]$ is no longer $\binom{n+1}{2}$, but is now a Binomial$\left(\binom{n+1}{2}, p\right)$ random variable. Since under (3.6) we have that

$$a_{\binom{n+1}{2}} / a_{p\binom{n+1}{2}} \to p^{-1/s} \text{ as } n \to \infty,$$

it's easy to obtain that equation (3.13) generalises for $p \in (0, 1]$ to

$$p^{-1/s}\left(\widetilde{M}_1^{(n)}, \widetilde{M}_2^{(n)}, \ldots, \widetilde{M}_k^{(n)}\right) \xrightarrow{d} (M_1, M_2, \ldots, M_k) \tag{3.45}$$

as $n \to \infty$, so that the asymptotics of the heaviest edges change simply by a constant factor.

The second issue concerns the set of feasible paths. Since not all edges are present, it is no longer the case that if $x = (i, j)$ and $y = (i', j')$ are two edges with $i < j \leq i' < j'$, then x and y can necessarily be used in the same path (there may be no feasible path between j and i').

However, if k is fixed and $n \to \infty$, then with high probability, any subset of the k

heaviest edges which are compatible with each other in this sense can be used together in a path (since the minimal distance between the endpoints of two such edges goes to infinity in probability). Then, since the argument above shows that we can obtain an arbitrarily close approximation to w by considering only the k heaviest edges, the result of Theorem 3.8 for $p < 1$ can be obtained just as before.

Chapter 4

Convoy formation in a travelling servers model

4.1 Introduction

We consider the following model: we have $n \in \mathbb{N}$ servers starting at the origin at time 0. A Poisson process with intensity $\lambda \in (0, \infty)$ on $(\mathbb{R}_+)^2$ generates arrivals of customers for these servers. The number of points in the Poisson process in a set $[a, b] \times [c, d]$ is the number of customers that arrive in the (space) interval $[a, b]$ between time c and d. Initially, each server waits an exponential amount of time with mean $\frac{1}{\nu}$ at the origin, independent of the other servers. Upon completion of this holding time, a server jumps to the first free customer to the right of the origin and serves this customer. All service times are i.i.d. exponentials with parameter ν. After completion of a service the customer leaves the system and the server jumps to the next free customer to the right. When a server makes a jump it ignores customers that are currently being served by other servers. Figure 4.1 below illustrates the behaviour of the model for $n = 1$.

As mentioned in the introduction, we want to analyse the asymptotic behaviour of $X_t^{(j)}$, the position of server j at time t, and its inverse $T_x^{(j)}$, the passage time of server j to x. In particular, we are interested in the number and sizes of convoys formed by

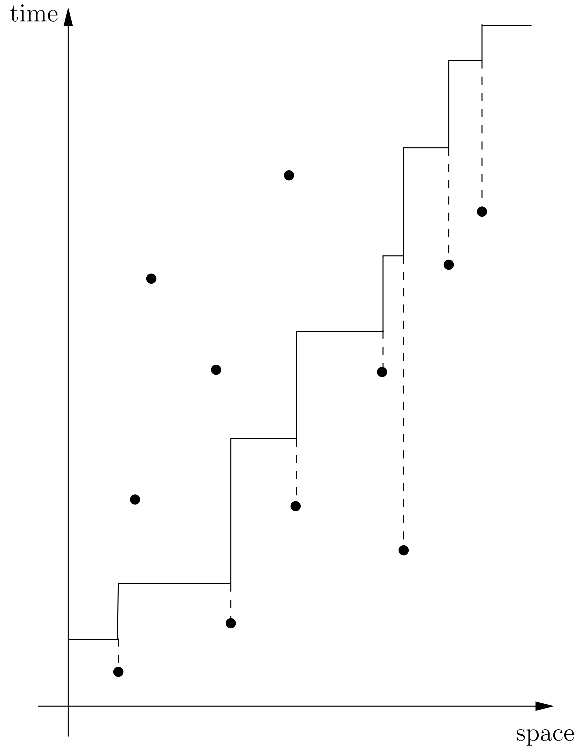

Figure 4.1: Travelling servers model for $n = 1$. The black dots symbolize the customers/arrivals in the Poisson process. The solid line is the space-time path of the server.

the servers. A convoy is formed by servers that overtake each other infinitely often. We can show that for $n = 2$ the two servers meet infinitely often with probability 1 (Proposition 4.12). This corresponds to the formation of a single convoy of size 2. For $n = 3$ two different behaviours are possible: with positive probability we will see a single convoy of size 3 and with positive probability we will see two convoys, one of size 2 and one of size 1. We prove asymptotic results for $X_t^{(j)}$ for $n = 2$ (see Theorem 4.4) and for $n = 3$ for a single convoy (see Theorem 4.6) and if the servers split up (see Theorem 4.8). It follows from the arguments used in the proofs of Theorems 4.6 and 4.8 that we can generalize the results for general n, see Theorem 4.10: the n servers will randomly arrange themselves in any sequence of convoys that are strictly decreasing in size.

We will start with the model for $n = 1$. The results are based on the treatment of this model in [36] and [47]. We will then move on to the model with $n \geq 2$ servers. This model has not been studied before and the results are original. The chapter is organized as follows: Section 4.2 contains our main results for $n = 1, 2, 3$ and general n. We will prove these results in Section 4.3.

Remark 4.1. *For $n \geq 2$, apart from referring to the servers by their labels $1, \ldots, n$, we will use the labels (b) and (f) to refer to the back/leftmost and front/rightmost server ($n = 2$) and the labels (b), (f) and (m) to refer to the back/leftmost, front/rightmost and middle server ($n = 3$).*

4.2 Results

We present our main results in five theorems: Theorem 4.2 ($n = 1$), Theorem 4.4 ($n = 2$), Theorems 4.6 and 4.8 ($n = 3$) and Theorem 4.10 (general n). We begin with the model with $n = 1$. Since we do not have to distinguish between different servers we let X_t be the position of the server at time t and T_x its passage time to site x, i.e.

$$T_x = \inf \{t \geq 0 : X_t \geq x\}.$$

We get the following result describing the asymptotic logarithmic speed of the server. We will prove this result in Section 4.3.1.

Theorem 4.2. *Let $n = 1$. Then*

$$X_t - \frac{\nu}{\lambda} \log t \tag{4.1}$$

converges almost surely to a finite random variable. In particular,

$$\lim_{t \to \infty} \frac{X_t}{\log t} = \frac{\nu}{\lambda}$$

almost surely.

Remark 4.3. *The almost sure convergence in (4.1) is equivalent to the following statement:* $T_x/e^{\frac{\lambda}{\nu}x}$ *converges almost surely to a strictly positive, finite random variable.*

Now we consider the model with two servers. Each server waits an exponential amount of time before making its first jump. The random variables $X_t^{(1)}, X_t^{(2)}, T_x^{(1)}$ and $T_x^{(2)}$ are similar to X_t and T_x above, but now we have to distinguish between the two servers (where we decide arbitrarily which server has label 1 and which server has label 2). We will prove the following result.

Theorem 4.4. *Let* $n = 2$. *Then*

$$X_t^{(1)} - \frac{2\nu}{\lambda}\log t \text{ and } X_t^{(2)} - \frac{2\nu}{\lambda}\log t$$

converge almost surely to the same finite random variable. In particular,

$$\lim_{t\to\infty}\frac{X_t^{(1)}}{\log t} = \lim_{t\to\infty}\frac{X_t^{(2)}}{\log t} = \frac{2\nu}{\lambda}$$

almost surely.

Remark 4.5. *Theorem 4.4 shows that two servers behave essentially like a single server with twice the service rate. With probability* 1 *the two servers form a single convoy of size* 2.

For $n = 3$ two different behaviours are possible. Let A be the event that the three servers overtake each other infinitely often (i.e. the back server overtakes the middle server infinitely often and the middle server overtakes the front server infinitely often).

Theorem 4.6. *Let* $n = 3$. *We have* $\mathbb{P}[A] > 0$ *and on* A

$$X_t^{(1)} - \frac{3\nu}{\lambda}\log t,\ X_t^{(2)} - \frac{3\nu}{\lambda}\log t \text{ and } X_t^{(3)} - \frac{3\nu}{\lambda}\log t$$

converge almost surely to the same finite random variable. In particular,

$$\lim_{t\to\infty}\frac{X_t^{(1)}}{\log t} = \lim_{t\to\infty}\frac{X_t^{(2)}}{\log t} = \lim_{t\to\infty}\frac{X_t^{(3)}}{\log t} = \frac{3\nu}{\lambda}$$

almost surely.

Remark 4.7. *On A the three servers behave essentially like a single server with three times the service rate and they form a single convoy of size 3.*

Theorem 4.8. *Let $n = 3$. We also have $\mathbb{P}[A^c] > 0$. On A^c there exist $i, j \in \{1, 2, 3\}$ and $k \in \{1, 2, 3\} \setminus \{i, j\}$ such that*

$$X_t^{(i)} - \frac{2\nu}{\lambda} \log t \text{ and } X_t^{(j)} - \frac{2\nu}{\lambda} \log t \tag{4.2}$$

converge almost surely to the same finite random variable. In particular,

$$\lim_{t\to\infty} \frac{X_t^{(i)}}{\log t} = \lim_{t\to\infty} \frac{X_t^{(j)}}{\log t} = \frac{2\nu}{\lambda}$$

almost surely. Furthermore,

$$X_t^{(k)} - \frac{\nu}{\lambda} \log t \tag{4.3}$$

converges almost surely to a finite random variable. In particular,

$$\lim_{t\to\infty} \frac{X_t^{(k)}}{\log t} = \frac{\nu}{\lambda}$$

almost surely.

Remark 4.9. *On A^c the front two servers behave like a single server with twice the service rate and the last server behaves just like a single server. The front two servers move so fast that their removal of some of the customers has no positive effect on the asymptotic speed of the last server. We observe the formation of two convoys: one of size 2 and one of size 1.*

For general n we have that results similar to Theorems 4.6 and 4.8 hold.

Theorem 4.10. *For any $\ell \geq 1$ and any sequence $a_1 > a_2 > \ldots > a_\ell > 0$ such that $\sum_{i=1}^{\ell} a_i = n$ we have that the following result holds. With positive probability there exist*

$i_{j,1},\dots,i_{j,a_j} \in \{1,\dots,n\} \setminus \{i_{1,1},\dots,i_{1,a_1},i_{2,1},\dots,i_{2,a_2},i_{3,1},\dots,i_{j-1,a_{j-1}}\}$, $j=1,\dots,\ell$, *such that for every* $j=1,\dots,\ell$

$$X_t^{(i_{j,k})} - \frac{a_j \nu}{\lambda}\log t \qquad k=1,\dots,a_j$$

all converge almost surely to the same finite random variable. In particular, for $j = 1,\dots,\ell$,

$$\lim_{t\to\infty} \frac{X_t^{(i_{j,k})}}{\log t} = \frac{a_j\nu}{\lambda} \qquad k=1,\dots,a_j$$

almost surely.

Remark 4.11. *Theorem 4.10 says that the* n *servers can arrange themselves in any sequence of convoys that are strictly decreasing in size. The speed of a server is proportional to the size of the group. The asymptotic behaviour of a server travelling in a group of size* m *is the same as the asymptotic behaviour of a single server with service rate* $m\nu$*. The proof of this theorem follows from the methods used in the proofs of Theorems 4.6 and 4.8 and we omit the details here. Note that we have at most* $-\frac{1}{2}+\sqrt{2n+\frac{1}{4}}$ *convoys and that this is also the minimum for the size of the first (largest) convoy.*

4.3 Proofs

In this section we will present the proofs of Theorems 4.2, 4.4, 4.6 and 4.8.

4.3.1 Proof of Theorem 4.2 ($n = 1$)

Proof of Theorem 4.2: The last statement in Theorem 4.2 has been proved in [36] and [47] for a speed $v < \infty$. In [35] more general service time distributions were considered. The first statement in Theorem 4.2 is stronger and implies the second statement.

We introduce the following notation for the service times and the jumps of the server. Let E_i be the length of the i-th service, $i = 1, 2, \dots$. Then $E_1, E_2, \dots$ are i.i.d.

exponentials with parameter ν and

$$S_k = \sum_{i=1}^{k} E_i \tag{4.4}$$

is the time of the k-th jump. Furthermore, we let U_k be the position of the server after the k-th jump and define for $i = 1, 2, \ldots$

$$F_i = S_i \left(U_i - U_{i-1} \right).$$

Then $F_1, F_2, \ldots$ are i.i.d. exponentials with parameter λ, independent of the sequence $(E_i)_{i\geq 1}$. This is because at time S_i, $U_i - U_{i-1} = F_i/S_i$ is the distance to the next customer, conditionally on S_i, F_i/S_i is exponentially distributed with parameter λS_i and conditionally on $S_1, S_2, \ldots$ the F_i/S_i, $i = 1, 2, \ldots$, are independent. This implies that the F_i are i.i.d. exponentials with parameter λ. We can then write

$$U_k = \sum_{i=1}^{k} \frac{F_i}{S_i}. \tag{4.5}$$

We can now write X_t simply as

$$X_t = \sum_{i=1}^{\infty} \frac{F_i}{S_i} \mathbb{1}_{\{S_i \leq t\}}. \tag{4.6}$$

We can use the representation of U_k in (4.5) to rewrite it in the following way

$$U_k = \sum_{i=1}^{k} \frac{\nu}{\lambda i} + \nu \sum_{i=1}^{k} \frac{\lambda F_i - 1}{\lambda i} + \sum_{i=1}^{k} F_i \left(\frac{1}{S_i} - \frac{\nu}{i} \right). \tag{4.7}$$

Our aim is to show that

$$U_k - \frac{\nu}{\lambda} \log S_k$$

converges almost surely to a finite random variable. We will do this by analysing the three sums in (4.7) separately. The first sum is entirely deterministic and we know that

$$\lim_{k\to\infty}\left(\sum_{i=1}^{k}\frac{\nu}{\lambda i}-\frac{\nu}{\lambda}\log k\right)=\frac{\nu}{\lambda}\gamma \tag{4.8}$$

where γ is the Euler constant. For the second sum we observe that

$$A_k=\nu\sum_{i=1}^{k}\frac{\lambda F_i-1}{\lambda i}$$

is a martingale with respect to the filtration $\mathcal{A}_k=\sigma\left(F_1,F_2,\ldots,F_k\right)$ because the F_i are i.i.d. and $\mathbb{E}\left[\frac{\lambda F_i-1}{\lambda i}\right]=0$. Furthermore, we have that

$$\begin{aligned}\mathbb{E}\left[A_k^2\right]=\mathrm{Var}\left(A_k\right)&=\nu^2\sum_{i=1}^{k}\mathrm{Var}\left(\frac{\lambda F_i-1}{\lambda i}\right)\\&=\nu^2\sum_{i=1}^{k}\frac{1}{i^2}\mathrm{Var}\left(F_i\right)\\&=\left(\frac{\nu}{\lambda}\right)^2\sum_{i=1}^{k}\frac{1}{i^2}.\end{aligned}$$

Therefore

$$\sup_{k\geq 0}\mathbb{E}\left[A_k^2\right]=\left(\frac{\nu}{\lambda}\right)^2\frac{\pi^2}{6}<\infty.$$

This implies that the martingale A_k converges almost surely (and in L^1) to a finite random variable A_∞, i.e.

$$A_k=\nu\sum_{i=1}^{k}\frac{\lambda F_i-1}{\lambda i}\xrightarrow[k\to\infty]{}A_\infty\in(-\infty,\infty)\quad\text{a.s.} \tag{4.9}$$

Let us now look at the last sum in (4.7). We put

$$M_k=\sum_{i=1}^{k}F_i\left(\frac{1}{S_i}-\frac{\nu}{i}\right).$$

If M_k is absolute convergent, i.e. if the sum

$$\sum_{i=1}^{k} F_i \left| \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \right|.$$

converges almost surely to a finite random variable, then we are done. We make the following two observations. For any $\varepsilon > 0$ we have

$$\mathbb{P}\left[\exists I_1 \text{ s.t. } \frac{i}{\nu} - i^{\frac{1}{2}+\varepsilon} \leq S_i \leq \frac{i}{\nu} + i^{\frac{1}{2}+\varepsilon} \;\forall i \geq I_1 \right] = 1$$

and

$$\mathbb{P}\left[\exists I_2 \text{ s.t. } F_i \leq i^{\frac{1}{4}} \;\forall i \geq I_2 \right] = 1.$$

This follows from the facts that S_i is the sum of i.i.d. random variables with mean $\frac{1}{\nu}$ and the F_i are i.i.d. with

$$\sum_{i=1}^{\infty} \mathbb{P}\left[F_i > i^{\frac{1}{4}} \right] = \sum_{i=1}^{\infty} \exp\left\{ -\lambda i^{\frac{1}{4}} \right\} < \infty.$$

The two observations together imply that

$$\mathbb{P}\left[\exists I \text{ s.t. } \sum_{i=1}^{k} F_i \left| \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \right| \leq \sum_{i=1}^{I-1} F_i \left| \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \right| + \sum_{i=I}^{k} i^{\frac{1}{4}} \frac{\nu^2}{i^{\frac{3}{2}-\varepsilon}} \;\forall k \geq I \right] = 1$$

and therefore, with $c = \sum_{i=1}^{\infty} i^{\frac{1}{4}} \frac{\nu^2}{i^{\frac{3}{2}-\varepsilon}} < \infty$,

$$\mathbb{P}\left[\exists I \text{ s.t. } \sum_{i=1}^{k} F_i \left| \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \right| \leq \sum_{i=1}^{I-1} F_i \left| \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \right| + c \;\forall k \geq I \right] = 1 \tag{4.10}$$

and this implies that M_k converges almost surely to a finite random variable M_∞ as $k \to \infty$, that is

$$M_k = \sum_{i=1}^{k} F_i \left(\frac{1}{S_i} - \frac{\nu}{i} \right) \xrightarrow[k\to\infty]{} M_\infty \in (-\infty, \infty) \quad \text{a.s.} \tag{4.11}$$

Putting (4.8), (4.9) and (4.11) together gives us

$$\lim_{k\to\infty}\Big(U_k - \frac{\nu}{\lambda}\log k\Big) = \frac{\nu}{\lambda}\gamma + A_\infty + M_\infty \in (-\infty,\infty) \quad \text{a.s.}$$

Since

$$\frac{S_k}{k} \xrightarrow[k\to\infty]{} \frac{1}{\nu} \quad \text{a.s.}$$

we also have

$$\begin{aligned}\lim_{k\to\infty}\Big(U_k - \frac{\nu}{\lambda}\log S_k\Big) &= \lim_{k\to\infty}\Big(U_k - \frac{\nu}{\lambda}\log k + \frac{\nu}{\lambda}\log k - \frac{\nu}{\lambda}\log S_k\Big)\\ &= \lim_{k\to\infty}\Big(U_k - \frac{\nu}{\lambda}\log k + \frac{\nu}{\lambda}\log\frac{k}{S_k}\Big)\\ &= \frac{\nu}{\lambda}\gamma + \frac{\nu}{\lambda}\log\nu + A_\infty + M_\infty \in (-\infty,\infty) \quad \text{a.s.}\end{aligned}$$

With $\frac{\nu}{\lambda}\gamma + \frac{\nu}{\lambda}\log\nu + A_\infty + M_\infty \overset{def}{=} Y_\infty$ we get

$$\lim_{k\to\infty}\Big(U_k - \frac{\nu}{\lambda}\log S_k\Big) = Y_\infty \in (-\infty,\infty) \quad \text{a.s.}$$

and this implies

$$\lim_{t\to\infty}\Big(X_t - \frac{\nu}{\lambda}\log t\Big) = Y_\infty \quad \text{a.s.}$$

This proves Theorem 4.2. □

4.3.2 Proof of Theorem 4.4 ($n = 2$)

Before we start with the proof itself we want to give a brief overview over the main ideas of the proof.

Sketch of the Proof of Theorem 4.4: The idea of the proof is very similar to the proof of Theorem 4.2. Since we now have two servers the total service rate is 2ν. We let

$$S_k = \sum_{i=1}^{k} E_i$$

be the time when the k-th service starts (by either server). The E_i are now i.i.d. exponentials with parameter 2ν. Note that now E_i does not correspond to the length of the i-th service but only to the length of time between the beginning of the $(i-1)$st service and the beginning of the i-th service.

In a first step we will prove that the two servers overtake each other infinitely often, see Proposition 4.12.

We then say that, for $i \geq 2$, a service E_i is "bad" if the customer whose service starts at time S_{i-1} is strictly to the left of the front server. A jump to such a customer does not increase the position of the rightmost server. All other services E_i are "good" and do increase the position of the front server. Note that E_2 is always "good". Recall the definitions of the labels (b) and (f) in Remark 4.1. According to those definitions $X_t^{(f)}$ is the position of the front server at time t.

Similar to the notation in (4.6) we can now write

$$X_t^{(f)} = \sum_{i=1}^{\infty} \frac{F_i}{S_i} \mathbb{1}_{\{E_{i+1} \text{ good}, \, S_i \leq t\}} \tag{4.12}$$

where the F_i are i.i.d. exponentials with parameter λ, independent of the E_i. If E_{i+1} is good then F_i/S_i is the amount by which $X_t^{(f)}$ increases (either by the front server jumping or the back server overtaking the front server). If E_{i+1} is bad, then F_i/S_i does not have an interpretation. After proving some preliminary results concerning the number of arrivals between the two servers between two overtaking events (which are well defined due to Proposition 4.12) in Proposition 4.13, Corollary 4.14 will give us a bound on the proportion of services E_i that are bad. This will allow us to rewrite (4.12) in terms of X_t for a single server with service rate 2ν and, using Theorem 4.2, this will prove Theorem 4.4. □

We start by proving the following proposition.

Proposition 4.12. *The two servers meet infinitely often with probability 1.*

Proof. Assume that with positive probability the two servers never meet. Without loss of generality assume that $X_t^{(1)} > X_t^{(2)}$ for all $t \geq E_1$ and let A be the corresponding event. Then $\mathbb{P}[A] > 0$. From the almost sure convergence in Theorem 4.2 it follows that the limit

$$Z_\infty = e^{-\frac{\lambda}{\nu}Y_\infty} = \lim_{x\to\infty} \frac{T_x^{(1)}}{e^{\frac{\lambda}{\nu}x}}$$

exists almost surely on A and $Z_\infty \in (0,\infty)$. However, conditional on A, the distribution of the limiting random variables Y_∞ and Z_∞ will be different from the unconditional limits in Theorem 4.2.

Ignoring that the first server will remove some of the customers, the second server behaves like a single server with only one difference: the initial waiting time at the origin is not exponential with parameter ν but distributed like the maximum of two exponentials with parameter ν. From the proof of Theorem 4.2 we can see that this does not change the asymptotic behaviour of $T_x^{(2)}$ (although the limiting random variable will be different). The fact that the first server removes some of the customers will only decrease $T_x^{(2)}$. Therefore

$$\limsup_{x\to\infty} \frac{T_x^{(2)}}{e^{\frac{\lambda}{\nu}x}} \le \widetilde{Z}_\infty$$

almost surely on A for some strictly positive, finite random variable $\widetilde{Z}_\infty$.

All the following statements hold almost surely on A. If we fix $\varepsilon > 0$ then we have that there exists $x_0 \in \mathbb{Q}, x_0 > 0$ such that

$$T_x^{(1)} \ge (Z_\infty - \varepsilon)\, e^{\frac{\lambda}{\nu}x}$$

for all $x \ge x_0$ and

$$T_x^{(2)} \le \left(\widetilde{Z}_\infty + \varepsilon\right) e^{\frac{\lambda}{\nu}x}$$

for all $x \ge x_0$. Note that x_0 is random.

We now want to use these inequalities to get an upper bound on the number of customers the second server has to serve. Let P be the random set of Poisson points that determines the arrivals of customers. A customer arriving at time t at site x will have to be served by the second server if $T_x^{(1)} < t < T_x^{(2)}$. So the total number of customers served between x_0 and $w > x_0$ is

$$\left\{(x,t) \in P : x_0 < x < w, T_x^{(1)} < t < T_x^{(2)}\right\}$$

By defition of x_0 the set above is contained in the set

$$\left\{(x,t) \in P : x_0 < x < w, (Z_\infty - \varepsilon)\, e^{\frac{\lambda}{\nu}x} < t < \left(\widetilde{Z}_\infty + \varepsilon\right) e^{\frac{\lambda}{\nu}x}\right\}$$

This set is increasing in w and the expected number of points in this set is

$$\lambda \int_{x_0}^{w} \left(\widetilde{Z}_\infty - Z_\infty + 2\varepsilon\right) e^{\frac{\lambda}{\nu}x} dx \leq \nu \left(\widetilde{Z}_\infty - Z_\infty + 3\varepsilon\right) \left(e^{\frac{\lambda}{\nu}w} - e^{\frac{\lambda}{\nu}x_0}\right).$$

By the strong law of large numbers there exists a $y > x_0$ such that for every $x > y$ the number of customers served by the second server in the interval $[x_0, x]$ is at most

$$\nu \left(\widetilde{Z}_\infty - Z_\infty + 3\varepsilon\right) \left(e^{\frac{\lambda}{\nu}x} - e^{\frac{\lambda}{\nu}x_0}\right)$$

Since the average service time length for the second server converges almost surely to $\frac{1}{\nu}$, there exists $x_1 \in \mathbb{Q}, x_1 \geq y$ such that the time it takes the second server to get from x_0 to x is bounded from above by

$$\left(\widetilde{Z}_\infty - Z_\infty + 4\varepsilon\right) e^{\frac{\lambda}{\nu}x}$$

for all $x \geq x_1$. So,

$$\frac{T_x^{(2)}}{e^{\frac{\lambda}{\nu}x}} \leq \frac{T_{x_0}^{(2)}}{e^{\frac{\lambda}{\nu}x}} + \left(\widetilde{Z}_\infty - Z_\infty + 4\varepsilon\right)$$

for all $x \geq x_1$ and

$$\frac{T_{x_0}^{(2)}}{e^{\frac{\lambda}{\nu}x}} + \left(\widetilde{Z}_\infty - Z_\infty + 4\varepsilon\right) \xrightarrow[x\to\infty]{} \left(\widetilde{Z}_\infty - Z_\infty + 4\varepsilon\right).$$

If we choose x_1 large enough then we have

$$T_x^{(2)} \leq \left(\widetilde{Z}_\infty - Z_\infty + 5\varepsilon\right) e^{\frac{\lambda}{\nu}x}$$

for all $x \geq x_1$. We can now iterate this argument to see that

$$T_x^{(2)} \leq \left(\widetilde{Z}_\infty - kZ_\infty + (4k+1)\,\varepsilon\right) e^{\frac{\lambda}{\nu}x}$$

for all $x \geq x_k, x_k \in \mathbb{Q}$ and $k \in \mathbb{N}$. Since eventually $\widetilde{Z}_\infty - kZ_\infty + (4k+1)\,\varepsilon < Z_\infty$ we get that eventually $T_x^{(2)} < T_x^{(1)}$ and this implies that the two servers have to meet almost surely on A which is a contradiction. Therefore, the two servers have to meet with probability 1, i.e. $\mathbb{P}[A] = 0$.

From the time the two servers meet for the first time we can repeat the argument above to see that the two servers have to meet infinitely often. □

With Proposition 4.12 we can define a random sequence of times t_i in the following way. Let

$$\mathrm{dist}_{1,2}(t) = X_t^{(1)} - X_t^{(2)} \tag{4.13}$$

be the distance between the two servers at time t (positive if $X_t^{(1)} > X_t^{(2)}$ and negative if $X_t^{(1)} < X_t^{(2)}$). Let $t_1 = E_1 \sim \mathrm{Exp}(2\nu)$ be the time when the first server makes the first jump and put for $i \geq 2$

$$t_i = \inf\left\{t > t_{i-1} : \mathrm{sgn}\left(\mathrm{dist}_{1,2}(t_{i-1})\right) \neq \mathrm{sgn}\left(\mathrm{dist}_{1,2}(t)\right)\right\}. \tag{4.14}$$

The times t_i are exactly the times when the servers overtake each other. In order to make the process right continuous we assume that at time t_i the servers have already swapped. The following proposition will bring us another step closer to proving Theorem 4.4. Here N_i is the number of customers that arrive between the front and the back server between time t_i and t_{i+1}.

Proposition 4.13. *Fix $k \in \mathbb{N}$ and $t \in (k-1, k]$. For any $\varepsilon > 0$ we have*

$$\mathbb{P}\left[N_i \geq 2 \mid t_i = t\right] \leq const. \frac{1}{k^{2(1-\varepsilon)}} \tag{4.15}$$

and

$$\mathbb{P}\left[N_i = 1 \mid t_i = t\right] \leq const. \frac{1}{k^{1-\varepsilon}}. \tag{4.16}$$

In particular, there are almost surely only finitely many intervals $[t_i, t_{i+1}]$ such that $N_i \geq 2$ and

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{N_i = 1\}} < \infty \tag{4.17}$$

almost surely.

Proof. Let $t_i = t \in (k-1, k]$. Without loss of generality we assume that server 1 overtakes server 2 at time t_i. See Figure 4.2 for a depiction of the situation at time t_i. The second server is serving a customer and by the memoryless property of the exponential distribution, the remaining service time $C_{i,1}$ is exponential with parameter ν. The first server has jumped to the first customer to the right. If we let $C_{i,2}$ be

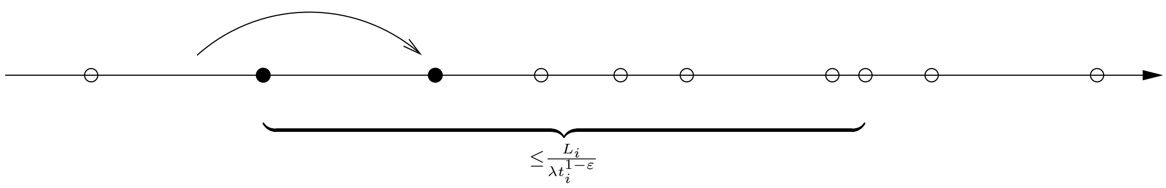

Figure 4.2: At time t_i one server has just overtaken the other server. The white circles are customers in front of the first server and behind the last server. By definition of t_i there is no customer between the two servers. On B_i^c the next L_i customers are at most $\frac{L_i}{\lambda t_i^{1-\varepsilon}}$ away (here for $L_i = 6$).

the length of the next service of the second server and put $C_i = C_{i,1} + C_{i,2}$, then by time $t_i + C_i$ the second server will have completed two services. C_i has a $\Gamma(2, \nu)$ distribution. Let $L_{i,1}$ and $L_{i,2}$ be the number of services that the first server starts between t_i and $t_i + C_{i,1}$ and between $t_i + C_{i,1}$ and $t_i + C_{i,1} + C_{i,2}$ respectively and put

$$L_i = L_{i,1} + L_{i,2}. \tag{4.18}$$

The random variables $L_{i,1}$ and $L_{i,1}$ are i.i.d. geometric random variables with parameter $\frac{1}{2}$ (i.e. $\mathbb{P}[L_{i,1} = j] = \left(\frac{1}{2}\right)^j$, $j \geq 1$). We consider the following event, see again Figure 4.2,

$$B_i = \Big\{ \text{Just before time } t_i \text{ the next } L_i \text{ customers to the right of the front server} \\ \text{are spread out over a distance greater than } \frac{L_i}{\lambda t_i^{1-\varepsilon}} \Big\}. \tag{4.19}$$

We can think of B_i as a "bad" event and we want to bound the probability of B_i from above. If B_i does not hold then the next L_i customers are sufficiently close and the distance between the two servers does not grow by too much. Conditional on $L_{i,1} = j$, $j \geq 1$, $L_{i,2} = m$, $m \geq 1$, and $t_i = t$, the probability of this event is

$$\begin{aligned} \mathbb{P}\left[B_i \mid t_i = t, L_{i,1} = j, L_{i,2} = m\right] &\leq \mathbb{P}\left[\text{Poi}\left(\lambda t \frac{j+m}{\lambda t^{1-\varepsilon}}\right) < j + m\right] \\ &= \mathbb{P}\left[\text{Poi}\left(t^{\varepsilon}\left(j+m\right)\right) < j+m\right] \\ &\leq const.e^{-const.(j+m)t^{\varepsilon}}. \end{aligned} \tag{4.20}$$

Summing over j and m and using the fact $t \in (k-1, k]$, we get

$$\begin{aligned} \mathbb{P}\left[B_i \mid t_i = t\right] &= \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} \mathbb{P}\left[B_i \mid t_i = t, L_{i,1} = j, L_{i,2} = m\right] \mathbb{P}\left[L_{i,1} = j\right] \mathbb{P}\left[L_{i,2} = m\right] \\ &\leq const. \sum_{j=1}^{\infty} \sum_{m=1}^{\infty} e^{-const.(j+m)t^{\varepsilon}} \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^m \\ &= const. \left(\frac{\frac{1}{2}e^{-const.t^{\varepsilon}}}{1 - \frac{1}{2}e^{-const.t^{\varepsilon}}}\right)^2 \\ &\leq const. \frac{1}{k^{2(1-\varepsilon)}}. \end{aligned} \tag{4.21}$$

Note that $t_{i+1} = t_i + C_{i,1}$ iff $N_i = 0$ and $t_{i+1} = t_i + C_i$ iff $N_i = 1$. On B_i^c the distance between the first and the second server at any time $t_i \leq t \leq \min\{t_{i+1}, t_i + C_i\}$ will therefore be at most $\frac{L_i}{\lambda t_i^{1-\varepsilon}}$.

We can use this to calculate

$$\begin{aligned}\mathbb{P}\left[N_i \geq 2 \,|t_i = t, B_i^c, C_i = s, L_i = j\right] &\leq \mathbb{P}\left[\text{Poi}\left(\lambda s \frac{j}{\lambda t^{1-\varepsilon}}\right) \geq 2\right] \\ &\leq const. \frac{j^2 s^2}{k^{2(1-\varepsilon)}} \qquad (4.22)\end{aligned}$$

for $s > 0$ and $j \geq 2$. Summing over $j \geq 2$, integrating over $s > 0$ and using the known distributions of L_i and C_i gives us

$$\mathbb{P}\left[N_i \geq 2 \,|t_i = t, B_i^c\right] \leq const. \frac{1}{k^{2(1-\varepsilon)}} \qquad (4.23)$$

and therefore

$$\begin{aligned}\mathbb{P}\left[N_i \geq 2 \,|t_i = t\right] &\leq \mathbb{P}\left[B_i \,|t_i = t\right] + \mathbb{P}\left[N_i \geq 2 \,|t_i = t, B_i^c\right] \\ &\leq const. \frac{1}{k^{2(1-\varepsilon)}}. \qquad (4.24)\end{aligned}$$

Similarly, we get for $s > 0$ and $j \geq 2$

$$\begin{aligned}\mathbb{P}\left[N_i = 1 \,|t_i = t, B_i^c, C_i = s, L_i = j\right] &\leq \mathbb{P}\left[\text{Poi}\left(\lambda s \frac{j}{\lambda t^{1-\varepsilon}}\right) = 1\right] \\ &\leq const. \frac{js}{k^{1-\varepsilon}} \qquad (4.25)\end{aligned}$$

and conclude

$$\begin{aligned}\mathbb{P}\left[N_i = 1 \,|t_i = t\right] &\leq \mathbb{P}\left[B_i \,|t_i = t\right] + \mathbb{P}\left[N_i = 1 \,|t_i = t, B_i^c\right] \\ &\leq const. \frac{1}{k^{1-\varepsilon}}. \qquad (4.26)\end{aligned}$$

This proves (4.15) and (4.16).

For the last two statements in Proposition 4.13 we make the following observation.

For every $k \in \mathbb{N}$, the number of $i \geq 1$ such that $t_i \in (k-1,k]$ is bounded from above by a random variable G_k. The G_k, $k = 1, 2, \ldots$ are i.i.d. geometric with parameter p where $p = \mathbb{P}[\mathrm{Exp}(\nu) > 1]$ (i.e. $\mathbb{P}[G_k = m] = p(1-p)^{m-1}$, $m \geq 1$). This is because $t_{i+1} - t_i \geq C_{i,1}$ and if $t_i \in (k-1,k]$ and $C_{i,1} > 1$ then certainly $t_{i+1} \notin (k-1,k]$. We let μ_i be the distribution of $t_i \in (0,\infty)$. Then

$$
\begin{aligned}
\sum_{i=1}^{\infty} \mathbb{P}[N_i \geq 2] &= \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \int_{k-1}^{k} \mathbb{P}[N_i \geq 2 | \, t_i = t] \, \mu_i(dt) \\
&\leq const. \sum_{k=1}^{\infty} \frac{1}{k^{2(1-\varepsilon)}} \sum_{i=1}^{\infty} \mathbb{P}[t_i \in (k-1,k]] \\
&= const. \sum_{k=1}^{\infty} \frac{1}{k^{2(1-\varepsilon)}} \sum_{i=1}^{\infty} \mathbb{E}\left[\mathbb{1}_{\{t_i \in (k-1,k]\}}\right] \\
&\leq const. \sum_{k=1}^{\infty} \frac{1}{k^{2(1-\varepsilon)}} \mathbb{E}[G_k] \\
&= const. \sum_{k=1}^{\infty} \frac{1}{k^{2(1-\varepsilon)}} \frac{1}{p} \\
&< \infty.
\end{aligned}
$$

This implies that almost surely there are only finitely many intervals $[t_i, t_{i+1}]$ such that $N_i \geq 2$. For (4.17) it is enough to show that

$$
\mathbb{E}\left[\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{N_i = 1\}}\right] = \sum_{i=1}^{\infty} \frac{1}{i} \mathbb{P}[N_i = 1] < \infty.
$$

We have for any $\delta > 0$

$$
\begin{aligned}
\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{P}[N_i = 1] &\leq \sum_{i=1}^{\infty} \frac{1}{i} \left(\mathbb{P}\left[S_i < \left(\frac{1}{2\nu} - \delta\right) i\right] + \mathbb{P}\left[N_i = 1 \middle| S_i \geq \left(\frac{1}{2\nu} - \delta\right) i\right]\right) \\
&\leq \sum_{i=1}^{\infty} \frac{1}{i} \left(const.e^{-const.i} + const. \frac{1}{\left(\left(\frac{1}{2\nu} - \delta\right) i\right)^{1-\varepsilon}} \right) \\
&< \infty.
\end{aligned}
$$

The bound

$$\mathbb{P}\left[S_i < \left(\frac{1}{2\nu} - \delta\right) i\right] \leq const.e^{-const.i}$$

for some positive constants follows from Cramér's Theorem since S_i is the sum of i.i.d. random variables with mean $\frac{1}{2\nu}$. Since $t_i \geq S_i$ and $\mathbb{P}\left[N_i = 1 | \, t_i = t\right]$ is decreasing in t we have

$$\mathbb{P}\left[N_i = 1 \middle| S_i \geq \left(\frac{1}{2\nu} - \delta\right) i\right] \leq \mathbb{P}\left[N_i = 1 \middle| t_i = \left(\frac{1}{2\nu} - \delta\right) i\right] \leq const.\frac{1}{\left(\left(\frac{1}{2\nu} - \delta\right) i\right)^{1-\varepsilon}}.$$

This proves that $\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{N_i=1\}}$ in (4.17) is almost surely finite and completes the proof of Proposition 4.13. □

Recall the notion of a "bad" service E_i from the sketch of the proof of Theorem 4.4. Proposition 4.13 implies the following corollary.

Corollary 4.14. *We have*

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \ bad\}} < \infty$$

almost surely.

Proof. Note that for every $i \geq 2$ there exists a unique $j(i) \geq 1$ such that $S_i \in \left(t_{j(i)}, t_{j(i)+1}\right]$ and that E_{i+1} has to be good if $N_{j(i)} = 0$. We can use this to write

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} &= \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad}, \ N_{j(i)} \geq 2\right\}} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad}, \ N_{j(i)} = 1\right\}} \\ &\leq \sum_{i=2}^{\infty} \mathbb{1}_{\left\{E_{i+1} \text{ bad}, \ N_{j(i)} \geq 2\right\}} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\{N_i=1\}}.\end{aligned}$$

The first sum is almost surely finite by Proposition 4.13 (almost surely there are only finitely many intervals $[t_i, t_{i+1}]$ such that $N_i \geq 2$) and Proposition 4.12. The second sum is almost surely finite by (4.17) in Proposition 4.13. □

Now we are ready to prove Theorem 4.4.

Proof of Theorem 4.4: Recall the representation of $X_t^{(f)}$ in (4.12). Similar to (4.7) we can write

$$\begin{aligned} X_t^{(f)} &= \sum_{i=1}^{\infty} \frac{F_i}{S_i} \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}} \\ &= \sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}} + 2\nu \sum_{i=1}^{\infty} \frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}} \\ &\qquad + \sum_{i=1}^{\infty} F_i \left(\frac{1}{S_i} - \frac{2\nu}{i} \right) \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}}. \end{aligned} \tag{4.27}$$

Again, we analyse the three sums separately. For the first sum we have

$$\sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}} = \sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{S_i \le t\}} - \sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ bad},\, S_i \le t\}}.$$

Writing $i(t)$ for the largest index such that $S_{i(t)} \le t$ we have

$$\begin{aligned} \lim_{t\to\infty} \left(\sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{S_i \le t\}} - \frac{2\nu}{\lambda} \log t \right) &= \lim_{t\to\infty} \left(\sum_{i=1}^{S_{i(t)}} \frac{2\nu}{\lambda i} - \frac{2\nu}{\lambda} \log S_{i(t)} + \frac{2\nu}{\lambda} \log \frac{S_{i(t)}}{t} \right) \\ &= \frac{2\nu}{\lambda} \gamma \end{aligned}$$

almost surely and

$$\lim_{t\to\infty} \sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ bad},\, S_i \le t\}} = \frac{2\nu}{\lambda} \sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} < \infty$$

almost surely by Corollary 4.14. Therefore

$$\lim_{t\to\infty} \left(\sum_{i=1}^{\infty} \frac{2\nu}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good},\, S_i \le t\}} - \frac{2\nu}{\lambda} \log t \right) = C_\infty \quad \text{a.s.} \tag{4.28}$$

for some almost surely finite random variable C_∞. For the second sum we can again use a martingale argument. Since the F_i are independent of the E_i, we have that

$$A_t = 2\nu \sum_{i=1}^{\infty} \frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good, } S_i \le t\}}$$

is a continuous time martingale with respect to the canonical filtration $\mathcal{A}_t = \sigma(A_s, s \le t)$. Similar to the calculations leading to (4.9) we have

$$\begin{aligned}
\sup_{t\ge 0} \mathbb{E}\left[A_t^2\right] &= \mathbb{E}\left[\left(2\nu \sum_{i=1}^{\infty} \frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good}\}}\right)^2\right] \\
&= 4\nu^2 \left(\sum_{i=1}^{\infty} \mathbb{E}\left[\left(\frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good}\}}\right)^2\right]\right. \\
&\qquad \left. +2 \sum_{i \ne j} \mathbb{E}\left[\frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good}\}}\right] \mathbb{E}\left[\frac{\lambda F_j - 1}{\lambda j} \mathbb{1}_{\{E_{j+1} \text{ good}\}}\right]\right) \\
&\le 4\nu^2 \sum_{i=1}^{\infty} \frac{1}{i^2} \operatorname{Var}(F_i) \\
&= \left(\frac{2\nu}{\lambda}\right)^2 \frac{\pi^2}{6} \\
&< \infty.
\end{aligned}$$

This implies that A_t converges almost surely (and in L^1) to a finite random variable A_∞, i.e.

$$A_t = 2\nu \sum_{i=1}^{\infty} \frac{\lambda F_i - 1}{\lambda i} \mathbb{1}_{\{E_{i+1} \text{ good, } S_i \le t\}} \xrightarrow[t\to\infty]{} A_\infty \in (-\infty, \infty) \quad \text{a.s.} \tag{4.29}$$

We now look at the third sum in (4.27). We put

$$M_t = \sum_{i=1}^{\infty} F_i \left(\frac{1}{S_i} - \frac{2\nu}{i}\right) \mathbb{1}_{\{E_{i+1} \text{ good, } S_i \le t\}}.$$

We have

$$\lim_{t\to\infty}\sum_{i=1}^{\infty} F_i \left|\frac{1}{S_i}-\frac{2\nu}{i}\right| \mathbb{1}_{\{E_{i+1}\text{ good},\, S_i\le t\}} = \sum_{i=1}^{\infty} F_i \left|\frac{1}{S_i}-\frac{2\nu}{i}\right| \mathbb{1}_{\{E_{i+1}\text{ good}\}} \le \sum_{i=1}^{\infty} F_i \left|\frac{1}{S_i}-\frac{2\nu}{i}\right|$$

and we know from (4.10) that the last sum is almost surely finite. Therefore

$$M_t = \sum_{i=1}^{\infty} F_i \left(\frac{1}{S_i}-\frac{2\nu}{i}\right) \mathbb{1}_{\{E_{i+1}\text{ good},\, S_i\le t\}} \xrightarrow[t\to\infty]{} M_\infty \in (-\infty,\infty) \quad \text{a.s.} \tag{4.30}$$

for some random variable M_∞ that is almost surely finite.

Putting (4.28), (4.29) and (4.30) together we get that

$$\lim_{t\to\infty}\left(X_t^{(f)} - \frac{2\nu}{\lambda}\log t\right) = Y_\infty \in (-\infty,\infty) \quad \text{a.s.}$$

for some random variable Y_∞ that is almost surely finite. This implies the convergence statements in Theorem 4.4 for the front server. But since we know that eventually the two servers are never separated by more than one customer, we have that, if we let $X_t^{(b)}$ be the position of the back server at time t, for any $d > 0$ there exists a t_d such that

$$X_t^{(f)} - X_t^{(b)} \le d \qquad \forall t \ge t_d.$$

Therefore

$$\lim_{t\to\infty}\left(X_t^{(b)} - \frac{2\nu}{\lambda}\log t\right) \ge \lim_{t\to\infty}\left(X_t^{(f)} - \frac{2\nu}{\lambda}\log t\right) - d = Y_\infty - d \text{ a.s.}$$

Since this is true for any $d > 0$, we get that the same convergence statements hold for the back server, too. This completes the proof of Theorem 4.4. $\square$

4.3.3 Proofs of Theorems 4.6 and 4.8 ($n = 3$)

Proof of Theorem 4.6: Recall that A is the event that the three servers overtake each other infinitely often. We start by proving the following proposition.

Proposition 4.15. $\mathbb{P}[A] > 0$.

Proof. The proof is similar to the arguments that we used in the proof of Proposition 4.13.

The total rate of service is now 3ν and we let again

$$S_k = \sum_{i=1}^{k} E_i$$

be the time when the k-th service starts. The E_i are now i.i.d. exponentials with parameter 3ν. We use the dynamic labels (b), (m) and (f) to refer to the back, middle and front server, e.g. $X_t^{(m)}$ is the position of the middle server at time t. We now define a sequence s_i, $i \geq 1$, in the following way. We let $s_1 = E_1 \sim \text{Exp}(3\nu)$ be the time when the first service starts and put for $i \geq 2$

$$s_i = \inf\left\{t > s_{i-1} : X_t^{(b)} \geq X_{s_{i-1}}^{(f)}\right\}.$$

The s_i are now not necessarily overtaking events. We only require that at time s_i the servers that were back and middle server at time s_{i-1} have reached the position of the front server at time s_{i-1}. We do not know anything about the position of the front server at time s_i. Below we will define a "good state" and show that with positive probability the servers will be in a good state at time s_i for all $i \geq 1$. Looking at a suitable subsequence of the s_i will then imply that the three servers overtake each other infinitely often with positive probability.

To define a "good state" we let $\text{dist}_{i,j}(t)$ be again the distance between server i and server j at time t, see (4.13), and we use the dynamic labels as described above. For example, $\text{dist}_{b,f}(t)$ is the distance between the back and the front server at time t. Fix

$\varepsilon > 0$. We say that the three servers are in a good state at time t, if

$$\text{dist}_{b,f}(t) \leq \frac{\log t}{\lambda t^{1-\varepsilon}}$$

and there is no customer between the front and the back server except the one being served by the middle server. Let G_i be the event that the system is in a good state at time s_i, see Figure 4.3. It is easy to see that with positive probability we are

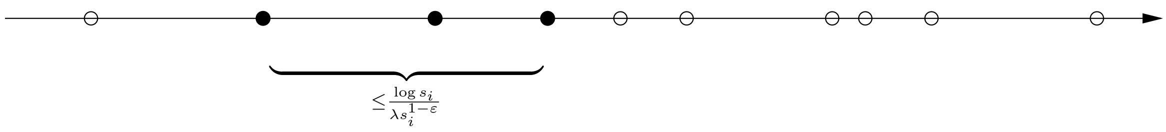

Figure 4.3: The three servers are in a good state at time s_i

in a good state at time s_1. In fact we have that given s_1 the distance $\text{dist}_{b,f}(s_1)$ is distributed like an exponential random variable with parameter λs_1 and there are no customers between the servers by definition. As before, see Proposition 4.13, we divide the half-line $\mathbb{R}_+$ into intervals $(k-1, k]$, $k = 1, 2, \ldots$. We put for $k \geq 1$

$$A_k = \left\{\exists i \geq 1 \text{ such that } s_i \in (k-1, k] \text{ and } G_i \cap G_{i+1}^c \text{ holds}\right\}.$$

The event A_k says that there is a time s_i between $k-1$ and k such that the system is in a good state at time s_i but in a bad state at time s_{i+1}. If none of the A_k hold and G_1 holds, then the system will be in a good state for all $i \geq 1$. We define R_t to be the state of the system at time t. This state contains all information about the position of the servers and customers between them at time t. From the state R_t we can tell whether we are in a good state at time t. Note that we are deliberately not including information about any customers in front of the rightmost server in the state R_t. We want to prove the following lemma.

Lemma 4.16. *Let $t \in (k-1, k]$ and let R be any good state for time t. Then*

$$\mathbb{P}\left[G_{i+1}^c \,\middle|\, s_i = t, R_t = R\right] \leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}.$$

If Lemma 4.16 holds then we can get the following bound for the probability of the event A_k. Let the index $i(k)$ be the smallest i such that $s_i \in (k-1,k]$ and G_i holds if such an i exists and define

$$D_k = \{\exists i \geq 1 \text{ such that } s_i \in (k-1,k] \text{ and } G_i \text{ holds}\}.$$

Note that $s_{i+1} - s_i \geq V_i$ for some i.i.d. exponentials V_i with parameter ν, since the last server has to finish at least one service before it can reach site $X^{(f)}_{s_i}$. Therefore

$$\begin{aligned}
\mathbb{P}[A_k] &= \mathbb{P}\left[\exists i \geq 1 \text{ such that } s_i \in (k-1,k] \text{ and } G_i \cap G^c_{i+1} \text{ holds}\right] \\
&\leq \mathbb{P}\big[D_k \cap \big\{\exists j \geq 0 \text{ such that } s_{i(k)+j} \in (k-1,k] \\
&\qquad\qquad \text{and } G_{i(k)+1} \cap \ldots \cap G_{i(k)+j} \cap G^c_{i(k)+j+1} \text{ holds}\big\}\big] \\
&\leq \sum_{j=0}^{\infty} \mathbb{P}\left[D_k \cap \left\{s_{i(k)+j} \in (k-1,k]\right\} \cap \bigcap_{l=1}^{j} \left(G_{i(k)+l} \cap \left\{V_{i(k)+l-1} \leq 1\right\}\right) \cap G^c_{i(k)+j+1}\right] \\
&\overset{(*)}{\leq} \sum_{j=0}^{\infty} \mathbb{P}[V_1 \leq 1]^j \int_{k-1}^{k} \int_{\Xi_t} \mathbb{P}\left[G^c_{i(k)+j+1} \middle| s_{i(k)+j} = t, R_t = R\right] \mu_{2,j}(dR)\mu_{1,j}(dt) \\
&\leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}. \qquad (4.31)
\end{aligned}$$

In $(*)$, $\mu_{1,j}$ is the conditional distribution of $s_{i(k)+j}$ given

$$D_k \cap \left\{s_{i(k)+j} \in (k-1,k]\right\} \cap \bigcap_{l=1}^{j} \left(G_{i(k)+l} \cap \left\{V_{i(k)+l-1} \leq 1\right\}\right)$$

and $\mu_{2,j}$ is the conditional distribution of R_t given

$$\left\{s_{i(k)+j} = t\right\} \cap \{R_t \in \Xi_t\} \cap D_k \cap \bigcap_{l=1}^{j} \left(G_{i(k)+l} \cap \left\{V_{i(k)+l-1} \leq 1\right\}\right)$$

where Ξ_t is the set of all possible good states for time t. Since Lemma 4.16 gives us a uniform bound on the probability, independent of t and R, the integrals disappear in the next line. Since (4.31) is summable over k we get

$$\sum_{k=1}^{\infty} \mathbb{P}\left[A_k\right] < \infty$$

and this implies that with positive probability none of the A_k happen. We will then show that for every $i \geq 1$ the probability that the last two servers overtake the front server between s_i and s_{i+1} is bounded from below by some constant independent of i, see (4.32) below. For infinitely many i the time s_i will therefore correspond to an overtaking event and this implies that the three servers overtake each other infinitely often with positive probability.

Now we have to prove Lemma 4.16.

Proof of Lemma 4.16: Consider $s_i = t \in (k-1, k]$ and assume $R_t = R \in \Xi_t$, i.e. we are in a good state at time t. Without loss of generality we may assume that the servers are ordered in the following way: server 1 is at the front, server 2 in the middle and server 3 at the back. We aim to show that with high probability by the time server 2 and server 3 have each completed two services, both have passed $X^{(f)}_{s_i}$ but no server has passed $X^{(f)}_{s_i} + \frac{\log s_{i+1}}{\lambda s_{i+1}^{1-\varepsilon}}$ and there are no customers between the servers. If that holds then the system is in a good state at time s_{i+1}.

After two exponential times with parameter ν server 3 will have completed two services. We can bound the number of services that server 1 and server 2 start during this time by the sum of two geometric random variables with parameter $\frac{1}{3}$. After server 3 has completed two services, we wait another two exponential times with parameter ν for server 2 to complete two services. The number of services that server 1 and server 3 start during this time can again be bounded by the sum of two geometric random variables with parameter $\frac{1}{3}$. Let C_i be the sum of the four exponential random variables with parameter ν and let L_i be the sum of the four geometric random variables with parameter $\frac{1}{3}$, see (4.18) for the corresponding random variable in the case with only two servers. Analogous to B_i in (4.19) we define

$$B_i = \{\text{At time } s_i \text{ the next } L_i \text{ customers to the right of server 1} \\ \text{are spread out over a distance greater than } \frac{L_i}{\lambda s_i^{1-\varepsilon}}\Big\}$$

and calculations analogous to (4.20) and (4.21) show that

$$\mathbb{P}\left[B_i \middle| s_i = t, R_t = R\right] \leq const. \frac{1}{k^{2(1-\varepsilon)}} \leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}.$$

Let M_i be the number of arrivals between the front and the back server between time s_i and s_{i+1}. Note that $s_{i+1} \leq s_i + C_i$ if $M_i \leq 1$. Conditional on B_i^c, we can therefore bound the probabilities of $M_i \geq 2$ and $M_i = 1$ by the probabilities of having more than one or exactly one arrival in the interval $[X_{s_i}^{(b)}, X_{s_i}^{(f)} + \frac{L_i}{\lambda s_i^{1-\varepsilon}}]$ between time s_i and time $s_i + C_i$ (similar to N_i in the proof of Proposition 4.13). Conditional on being in a good state at time s_i, i.e. $X_{s_i}^{(f)} - X_{s_i}^{(b)} \leq \frac{\log s_i}{\lambda s_i^{1-\varepsilon}}$, it is then easy to calculate that

$$\mathbb{P}\left[M_i \geq 2 \middle| s_i = t, R_t = R\right] \leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}$$

and

$$\mathbb{P}\left[M_i = 1 \middle| s_i = t, R_t = R\right] \leq const. \frac{\log k}{k^{1-\varepsilon}}.$$

See (4.22) - (4.26) for the corresponding calculations in the situation with two servers. We observe now that if $M_i \leq 1$ and B_i^c holds, then at time $s_{i+1} \leq s_i + C_i$ the distance between the (new) back and the (new) front server satisfies

$$\text{dist}_{b,f}(s_{i+1}) \leq \frac{L_i}{\lambda s_i^{1-\varepsilon}}.$$

If therefore $M_i \leq 1$, B_i^c holds and

$$\frac{L_i}{\lambda s_i^{1-\varepsilon}} \leq \frac{\log s_{i+1}}{\lambda s_{i+1}^{1-\varepsilon}}$$

then we are in a good state at time s_{i+1}. Define

$$D_i = \left\{ \frac{L_i}{\lambda s_i^{1-\varepsilon}} > \frac{\log s_{i+1}}{\lambda s_{i+1}^{1-\varepsilon}} \right\}.$$

Then

$$\begin{aligned}
\mathbb{P}\left[D_i \middle| s_i = t, R_t = R, B_i^c, M_i \leq 1\right] &= \mathbb{P}\left[L_i s_{i+1}^{1-\varepsilon} > s_i^{1-\varepsilon} \log s_{i+1} \middle| s_i = t, R_t = R\right] \\
&\leq \mathbb{P}\left[L_i \left(1 + \frac{C_i}{k-1}\right) > \log(k-1)\right] \\
&\leq \mathbb{P}\left[L_i > \sqrt{\log(k-1)}\right] \\
&\quad + \mathbb{P}\left[C_i > (k-1)\left(\sqrt{\log(k-1)} - 1\right)\right] \\
&\leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}.
\end{aligned}$$

Putting everything together we get

$$\begin{aligned}
\mathbb{P}\left[G_{i+1}^c \middle| s_i = t, R_t = R\right] &\leq \mathbb{P}\left[B_i \middle| s_i = t, R_t = R\right] + \mathbb{P}\left[M_i \geq 2 \middle| s_i = t, R_t = R, B_i^c\right] \\
&\quad + \mathbb{P}\left[D_i \middle| s_i = t, R_t = R, B_i^c, M_i \leq 1\right] \\
&\quad + \mathbb{P}\left[G_{i+1}^c \middle| s_i = t, R_t = R, B_i^c, M_i \leq 1, D_i^c\right] \\
&\leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}} + const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}} + const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}} + 0 \\
&\leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}
\end{aligned}$$

since on $\{s_i = t, R_t = R, M_i \leq 1\} \cap B_i^c \cap D_i^c$ we have to be in a good state at time s_{i+1}, i.e.

$$\mathbb{P}\left[G_{i+1}^c \middle| s_i = t, R_t = R, B_i^c, M_i \leq 1, D_i^c\right] = 0.$$

This completes the proof of Lemma 4.16. □

Now we want to prove that $\mathbb{P}[A] > 0$. We have

$$\mathbb{P}\left[\bigcap_{i=1}^{\infty} G_i\right] = \mathbb{P}\left[\bigcap_{k=1}^{\infty} A_k^c\right] \stackrel{def}{=} p > 0$$

by (4.31). If we let H_i be the event that the servers that were the back and middle servers at time s_i overtake the front server before time s_{i+1}, then we also have for any large $k \geq 1$, $t \in (k-1, k]$ and $R_t = R \in \Xi_t$

$$\begin{aligned} \mathbb{P}\left[H_i \middle| s_i = t, R_t = R\right] &\geq \frac{1}{4}\mathbb{P}\left[M_i = 0 \middle| s_i = t, R_t = R\right] \\ &\geq \frac{1}{4}\left(1 - const.\frac{\log k}{k^{1-\varepsilon}}\right) \\ &> q > 0. \end{aligned} \tag{4.32}$$

We get the factor $\frac{1}{4}$ because with probability at least $\frac{1}{4}$ the front server will not have completed its service by the time the back and the middle server complete their services. If $M_i = 0$ then this implies that the back and the middle server overtake the front server. Since conditional on s_i and R_{s_i}, $i \geq 1$, the H_i are independent, equation (4.32) implies that

$$\mathbb{P}\left[\bigcap_{m=1}^{\infty}\bigcup_{i=m}^{\infty} H_i \middle| \bigcap_{i=1}^{\infty} G_i\right] = 1.$$

Therefore,

$$\begin{aligned} \mathbb{P}[A] &\geq \mathbb{P}\left[\bigcap_{i=1}^{\infty} G_i\right]\mathbb{P}\left[A \middle| \bigcap_{i=1}^{\infty} G_i\right] \\ &= \mathbb{P}\left[\bigcap_{i=1}^{\infty} G_i\right]\mathbb{P}\left[\bigcap_{m=1}^{\infty}\bigcup_{i=m}^{\infty} H_i \middle| \bigcap_{i=1}^{\infty} G_i\right] \\ &\geq p > 0. \end{aligned}$$

This completes the proof that $\mathbb{P}[A] > 0$. □

Now we want to prove the convergence statements in Theorem 4.6. We condition on the event A. We can therefore define an infinite sequence of times $t_i^{(m,f)}$ when the middle server overtakes the front server and an infinite sequence of times $t_i^{(b,m)}$ when the back server overtakes the middle server. Let $N_i^{(m,f)}$ be the number of customers that arrive between the middle and the front server between time $t_i^{(m,f)}$ and time $t_{i+1}^{(m,f)}$ and let $N_i^{(b,m)}$ be the number of customers that arrive between the back and the middle server between time $t_i^{(b,m)}$ and time $t_{i+1}^{(b,m)}$. Similar to Proposition 4.13 we have the following results.

Proposition 4.17. *Fix $k \in \mathbb{N}$ and $t \in (k-1, k]$. For any $\varepsilon > 0$ we have*

$$\mathbb{P}\left[N_i^{(m,f)} \geq 2 \middle| t_i^{(m,f)} = t \right] \leq const. \frac{1}{k^{2(1-\varepsilon)}}$$

and

$$\mathbb{P}\left[N_i^{(m,f)} = 1 \middle| t_i^{(m,f)} = t \right] \leq const. \frac{1}{k^{1-\varepsilon}}.$$

In particular, there are almost surely only finitely many intervals $\left[t_i^{(m,f)}, t_{i+1}^{(m,f)}\right]$ such that $N_i^{(m,f)} \geq 2$ and

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{N_i^{(m,f)}=1\right\}} < \infty \tag{4.33}$$

almost surely.

Proof. For any fixed t we have

$$\mathbb{P}\left[N_i^{(m,f)} \geq 2 \middle| t_i^{(m,f)} = t \right] \leq \mathbb{P}\left[N_i \geq 2 | \, t_i = t \right]$$

and

$$\mathbb{P}\left[N_i^{(m,f)} = 1 \middle| t_i^{(m,f)} = t \right] \leq \mathbb{P}\left[N_i = 1 | \, t_i = t \right]$$

where on the right-hand side in each equation we consider the model with only two servers. N_i is the quantity defined before Proposition 4.13. This is because adding a third server can only decrease the time between two overtaking events and the maximal distance between the front and the middle server. The results in Proposition 4.13 then imply Proposition 4.17. □

Now we want to prove a similar result for the $N_i^{(b,m)}$.

Proposition 4.18. *Fix $k \in \mathbb{N}$ and $t \in (k-1, k]$. For any $\varepsilon > 0$ we have*

$$\mathbb{P}\left[N_i^{(b,m)} \geq 2 \middle| t_i^{(b,m)} = t \right] \leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}} \tag{4.34}$$

and

$$\mathbb{P}\left[N_i^{(b,m)}=1\middle| t_i^{(b,m)}=t\right]\leq const.\frac{\log k}{k^{1-\varepsilon}}. \tag{4.35}$$

In particular, there are almost surely only finitely many intervals $\left[t_i^{(b,m)},t_{i+1}^{(b,m)}\right]$ *such that* $N_i^{(b,m)}\geq 2$ *and*

$$\sum_{i=1}^{\infty}\frac{1}{i}\mathbb{1}_{\left\{N_i^{(b,m)}=1\right\}}<\infty \tag{4.36}$$

almost surely.

Proof. Note that for every $i\geq 1$ there exists a $j(i)$ such that $t_i^{(b,m)}\in\left(t_{j(i)}^{(m,f)},t_{j(i)+1}^{(m,f)}\right]$. By Proposition 4.17 we can choose I large enough such that $N_{j(i)}^{(m,f)}\leq 1$ for all $i\geq I$. For $i\geq I$ we have

$$\mathrm{dist}_{m,f}(t)\leq\frac{L_{j(i)}}{\lambda\left(t_{j(i)}^{(m,f)}\right)^{1-\varepsilon}}$$

for all $t\in\left(t_{j(i)}^{(m,f)},t_{j(i)+1}^{(m,f)}\right)$ and for $L_{j(i)}=L_{j(i),1}+L_{j(i),2}$ for some independent random variables $L_{j(i),1}$ and $L_{j(i),2}$ that are geometric with parameter $\frac{1}{2}$, see (4.18) in the proof of Proposition 4.13. We also have $t_{j(i)+1}^{(m,f)}-t_{j(i)}^{(m,f)}\leq F_{j(i)}=F_{j(i),1}+F_{j(i),2}$ for two independent exponential random variables with parameter ν. If the above inequality holds, then at time $t_i^{(b,m)}$ the distance between the back and the middle server is at most

$$\mathrm{dist}_{b,m}(t_i^{(b,m)})\leq\frac{L_{j(i)}}{\lambda\left(t_{j(i)}^{(m,f)}\right)^{1-\varepsilon}}.$$

In the worst case there was no customer between the middle server and the front server when the back server jumped (i.e. $t_i^{(b,m)}=t_{j(i)+1}^{(m,f)}$). If there was a customer then the back server jumps to a site between the middle and the front server and the bound above still holds. Let

$$D_i=\left\{\frac{L_{j(i)}}{\lambda\left(t_{j(i)}^{(m,f)}\right)^{1-\varepsilon}}>\frac{\log t_i^{(b,m)}}{\lambda\left(t_i^{(b,m)}\right)^{1-\varepsilon}}\right\}.$$

Then

$$\begin{aligned}
\mathbb{P}\left[D_i \left| t_{j(i)}^{(m,f)} = t\right.\right] &= \mathbb{P}\left[N_{j(i)}^{(m,f)} \geq 2 \left| t_{j(i)}^{(m,f)} = t\right.\right] + \mathbb{P}\left[D_i \left| t_{j(i)}^{(m,f)} = t, N_{j(i)}^{(m,f)} \leq 1\right.\right] \\
&= \mathbb{P}\left[N_{j(i)}^{(m,f)} \geq 2 \left| t_{j(i)}^{(m,f)} = t\right.\right] \\
&\quad + \mathbb{P}\left[L_{j(i)} \left(\frac{t_i^{(b,m)}}{t_{j(i)}^{(m,f)}}\right)^{1-\varepsilon} > \log t_i^{(b,m)} \left| t_{j(i)}^{(m,f)} = t, N_{j(i)}^{(m,f)} \leq 1\right.\right] \\
&\leq \mathbb{P}\left[L_{j(i)} > \sqrt{\log (k-1)}\right] + \mathbb{P}\left[F_{j(i)} > (k-1)\sqrt{\log (k-1)}\right] \\
&\leq const. \frac{(\log k)^2}{k^{2(1-\varepsilon)}}.
\end{aligned}$$

For all large i we have therefore

$$\mathrm{dist}_{b,m}(t_i^{(b,m)}) \leq \frac{\log t_i^{(b,m)}}{\lambda \left(t_i^{(b,m)}\right)^{1-\varepsilon}}.$$

We can now proceed in the same way as in the proof of Proposition 4.13 or the proof of Lemma 4.16. Eventually, the distance between the two servers will be bounded by

$$\frac{\log t_i^{(b,m)}}{\lambda \left(t_i^{(b,m)}\right)^{1-\varepsilon}} \tag{4.37}$$

every time they meet. We can bound the distance between the servers after the completion of two services of the back server and show that with sufficiently small probability we will see more than one arrival between the back and the middle server during those two service times. If we see at most one arrival then the back server will overtake the middle server after completing two services at the latest. Because we have (4.37) instead of

$$\mathrm{dist}_{1,2}(t_i) \leq \frac{1}{\lambda t_i^{1-\varepsilon}},$$

which was the case for the model with two servers, the probabilities in (4.34) and (4.35) contain an extra $(\log k)^2$ or $\log k$. □

Now we say again that a service time E_i is bad if the customer whose service starts at time S_{i-1} was strictly to the left of the front server. Analogous to Corollary 4.14 we get

Corollary 4.19. *We have*

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} < \infty$$

almost surely.

Proof. Note that for every bad E_i the corresponding customer was either between the back and the middle server or between the middle and the front server. Let $\kappa_i^{(b,m)}$ be the indicator for the first event and $\kappa_i^{(m,f)}$ the indicator for the second event. Let $k_1(i)$ be such that $S_i \in \left(t_{k_1(i)}^{(b,m)}, t_{k_1(i)+1}^{(b,m)}\right]$ and $k_2(i)$ such that $S_i \in \left(t_{k_2(i)}^{(m,f)}, t_{k_2(i)+1}^{(m,f)}\right]$. Then

$$\begin{aligned}
\sum_{i=1}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} &= \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} \kappa_{i+1}^{(b,m)} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\{E_{i+1} \text{ bad}\}} \kappa_{i+1}^{(m,f)} \\
&= \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_1(i)}^{(b,m)} \geq 2\right\}} \kappa_{i+1}^{(b,m)} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_1(i)}^{(b,m)} = 1\right\}} \kappa_{i+1}^{(b,m)} \\
&\quad + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_2(i)}^{(m,f)} \geq 2\right\}} \kappa_{i+1}^{(m,f)} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_2(i)}^{(m,f)} = 1\right\}} \kappa_{i+1}^{(m,f)} \\
&\leq \sum_{i=2}^{\infty} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_1(i)}^{(b,m)} \geq 2\right\}} \kappa_{i+1}^{(b,m)} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{N_i^{(b,m)} = 1\right\}} \\
&\quad + \sum_{i=2}^{\infty} \mathbb{1}_{\left\{E_{i+1} \text{ bad, } N_{k_2(i)}^{(m,f)} \geq 2\right\}} \kappa_{i+1}^{(m,f)} + \sum_{i=2}^{\infty} \frac{1}{i} \mathbb{1}_{\left\{N_i^{(m,f)} = 1\right\}}.
\end{aligned}$$

The first and the third sums are almost surely finite since there are only finitely many intervals $[t_i^{(b,m)}, t_{i+1}^{(b,m)}]$ and $[t_i^{(m,f)}, t_{i+1}^{(m,f)}]$ such that $N_i^{(b,m)} \geq 2$ and $N_i^{(m,f)} \geq 2$. The second and fourth sums are finite by (4.36) and (4.33) in Propositions 4.18 and 4.17. □

Just like in the proof of Theorem 4.4 we can now look again at $X_t^{(f)}$, the position of the front server at time t. Rewriting $X_t^{(f)}$ as in (4.27) and using the same arguments as in the proof of Theorem 4.4 together with Corollary 4.19 proves the convergence statements for $X_t^{(f)}$. The same argument as the one used in the end of the proof of Theorem 4.4 shows that the convergence statements also hold for the back and the middle server.

This completes the proof of Theorem 4.6. □

Proof of Theorem 4.8: Recall that A^c is the event that the three servers do not overtake each other infinitely often. We will start by proving that $\mathbb{P}[A^c] > 0$. Note that it follows from Proposition 4.12 that in the model with $n = 3$ the middle server has to overtake the front server infinitely often. On A^c we have to have therefore that the back server overtakes the middle server only finitely often.

Proposition 4.20. $\mathbb{P}[A^c] > 0$.

Proof. We want to prove here the even stronger statement that with positive probability the back server never meets the middle server. Since the presence of the back server can only increase the speed of the middle and the front server, it follows from the almost sure convergence in Theorem 4.4 that

$$T_x^{(m)} < e^{\frac{\lambda}{2\nu}(1+\delta)x} = f(x) \tag{4.38}$$

for all $x \geq 0$ with positive probability for any $\delta > 0$. Here $T_x^{(m)}$ is the passage time of the middle server to site x. With probability 1 the inequality holds for all large x and by manipulating finitely many points in the Poisson process we can make it hold for all $x \geq 0$ with positive probability. Now we want to show that, conditional on (4.38), with positive probability the back server satisfies

$$T_x^{(b)} \geq e^{\frac{\lambda}{2\nu}(1+\delta)x}$$

for all $x \geq 0$. We can assume without loss of generality that the back server makes its first jump at time h for a large $h >> 0$ (for any h this has positive probability). Now look at Figure 4.4. Let $E_1, E_2, \ldots$ be the service times of the back server, i.e. they are i.i.d. exponentials with parameter ν. Let $F_1, F_2, \ldots$ be i.i.d. exponentials with

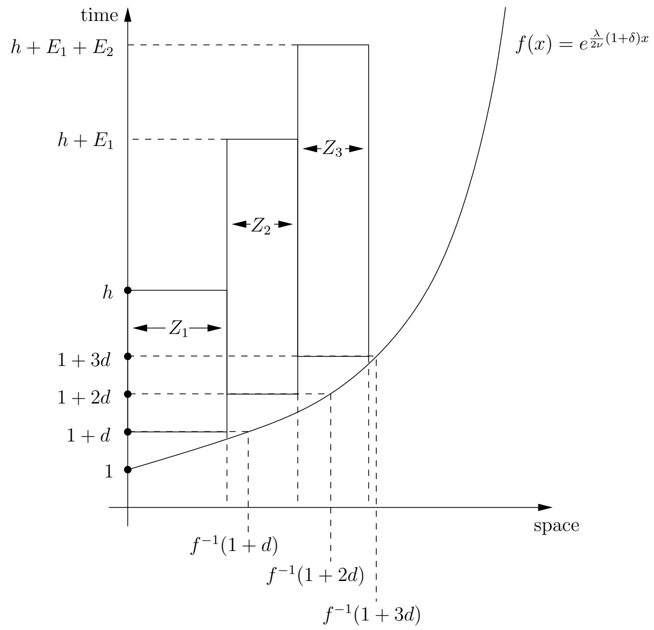

Figure 4.4: Construction of a lower bound for $T_x^{(b)}$

parameter 1 and define for $i = 1, 2, \ldots$ and some $d > 0$

$$Z_i = \frac{F_i}{\lambda\Big((h-d-1) + \sum_{j=1}^{i-1}(E_j - d)\Big)}.$$

We think of the random variable Z_i as the smallest site x such that there is a Poisson point in the rectangle $\left[1+id, h+\sum_{j=1}^{i-1} E_j\right] \times [Z_{i-1}, Z_{i-1}+x]$ (with $Z_0 = 0$), see Figure 4.4. Since $E_1, E_2, \ldots$ are all i.i.d. with mean $\frac{1}{\nu}$ we get for any $c < \frac{1}{\nu}$ that

$$p = \mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{j=1}^{m} E_j > mc\right\}\right] > 0.$$

We choose our constants c, d and h in the following way

$$c = \frac{1-\varepsilon}{\nu}, \qquad d = \frac{\varepsilon}{\nu}, \qquad h > c$$

for some fixed $\varepsilon > 0$. Then it is enough to show that

$$\mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} Z_i < f^{-1}(1+md)\right\}\right] > 0,$$

see again Figure 4.4. If $\bigcap_{m=1}^{\infty}\{\sum_{i=1}^{m} Z_i < f^{-1}(1+md)\}$ then the back server will never cross the line given by $f(x) = e^{\frac{\lambda}{2\nu}(1+\delta)x}$ and always stay behind the middle server. We get

$$\begin{aligned}
&\mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} Z_i < f^{-1}(1+md)\right\}\right] \\
&\qquad \geq \mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} Z_i < f^{-1}(1+md)\right\} \cap \bigcap_{m=1}^{\infty}\left\{\sum_{j=1}^{m} E_j > mc\right\}\right] \\
&\qquad \geq p \cdot \mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} \frac{F_i}{\lambda\left((h-d-1)+(i-1)(c-d)\right)} < f^{-1}(1+md)\right\}\right] \\
&\qquad = p \cdot \mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} \frac{F_i}{\lambda\left((h-d-1)+(i-1)(c-d)\right)} < \frac{2\nu}{\lambda(1+\delta)}\log(1+md)\right\}\right] \\
&\qquad \geq p \cdot \mathbb{P}\left[\bigcap_{m=1}^{\infty}\left\{\sum_{i=1}^{m} \frac{F_i}{i(c-d)} < \frac{2\nu}{(1+\delta)}\log(1+md)\right\}\right].
\end{aligned}$$

Looking at the left hand side of the inequality, we observe that this is the time of the m-th birth in a pure birth process with rate $c-d$. If we let S_t be the number of individuals in this birth process at time t (with $S_0 = 1$), then we can bound the last line above by

$$p \cdot \mathbb{P}\left[S_t \geq \frac{1}{d}\left(e^{\frac{(1+\delta)}{2\nu}t} - 1\right) \forall t \geq 0\right].$$

From the theory of pure birth processes we get, since $c - d = \frac{1-2\varepsilon}{\nu}$, that

$$\lim_{t\to\infty} S_t e^{-\frac{1-3\varepsilon}{\nu}t} = \infty \qquad a.s.$$

and in particular

$$S_t \geq e^{\frac{1-3\varepsilon}{\nu}t} \qquad \forall t \geq 0$$

with positive probability. If therefore

$$\frac{1-3\varepsilon}{\nu} > \frac{1+\delta}{2\nu}$$

then

$$p \cdot \mathbb{P}\left[S_t \geq \frac{1}{d}\left(e^{\frac{1+\delta}{2\nu}t} - 1\right) \forall t \geq 0\right] > 0.$$

But $\frac{1-3\varepsilon}{\nu} > \frac{1+\delta}{2\nu}$ can be achieved by choosing ε and δ appropriately. This proves Proposition 4.20. □

To complete the proof of the first part of Theorem 4.8 we make the following observation: on A^c the front two servers will eventually behave like the servers in a model with $n = 2$ because they interact with the back server only finitely often. This implies that the limit results hold for the first two servers.

It remains to prove the limit result (4.3) for the back server. We will do this in three steps (Propositions 4.21, 4.22 and 4.23) and start with the following proposition.

Proposition 4.21. *On A^c we have*

$$\limsup_{t\to\infty} \frac{X_t^{(b)}}{\log t} \in \left\{\frac{\nu}{\lambda}, \frac{2\nu}{\lambda}\right\} \tag{4.39}$$

almost surely.

Proof. Since the back server travels at least as fast as a server in the model with $n = 1$ (because the front two servers remove some of the customers) but at most as fast as the middle server (because the back server overtakes the middle server only finitely often) we have

$$\frac{\nu}{\lambda} \leq \limsup_{t\to\infty} \frac{X_t^{(b)}}{\log t} \leq \frac{2\nu}{\lambda}$$

almost surely. Let

$$B = \left\{\limsup_{t\to\infty} \frac{X_t^{(b)}}{\log t} \in \left(\frac{\alpha_1\nu}{\lambda}, \frac{\alpha_2\nu}{\lambda}\right)\right\}$$

for some $1 < \alpha_1 < \alpha_2 < 2$ and assume $\mathbb{P}[B] > 0$. In the following we always condition on B. On B we have that for any $\varepsilon > 0$, almost surely,

$$X_t^{(b)} < \left(\frac{\alpha_2 \nu}{\lambda} + \varepsilon\right) \log t \tag{4.40}$$

for all large enough t. Recall that $T_x^{(m)}$ is the passage time of the middle server to x. We know from (4.2) that, almost surely,

$$T_x^{(m)} \leq (Z_\infty + \varepsilon)\, e^{\frac{\lambda}{2\nu} x} \tag{4.41}$$

for all large enough x and some strictly positive, finite random variable Z_∞. Equations (4.40) and (4.41) together imply that, almost surely,

$$T^{(m)}_{X_t^{(b)}+a} \leq (Z_\infty + \varepsilon)\, t^{\frac{\alpha_1}{2}+\tau} e^{\frac{\lambda a}{2\nu}}$$

for all large enough t with $\tau = \frac{\lambda \varepsilon}{2\nu}$ and for any constant $a > 0$. For any fixed $\beta > 1$ we can choose $\varepsilon > 0$ such that, almost surely,

$$\lambda \left(t - (Z_\infty + \varepsilon)\, t^{\frac{\alpha_1}{2}+\tau} e^{\frac{\lambda a}{2\nu}}\right) \geq \frac{\lambda t}{\beta}$$

for all large t. For all large t, the back server will therefore have to serve more customers between $X_t^{(b)}$ and $X_t^{(b)} + a$ than a server in a model with $n = 1$ and λ replaced by $\frac{\lambda}{\beta}$ has to serve between X_t and $X_t + a$. Theorem 4.2 implies that the latter server will satisfy

$$\lim_{t\to\infty} \frac{X_t}{\log t} = \frac{\beta \nu}{\lambda} \qquad \text{a.s.}$$

We can use this as an upper bound for the back server and get

$$\limsup_{t\to\infty} \frac{X_t^{(b)}}{\log t} \leq \frac{\beta \nu}{\lambda} \qquad \text{a.s.}$$

In particular, we can choose $1 < \beta < \alpha_1$ which leads to a contradiction since we conditioned on B. This completes the proof of Proposition 4.21. □

Proposition 4.22. *On A^c we have*

$$\limsup_{t\to\infty} \frac{X_t^{(b)}}{\log k} = \frac{\nu}{\lambda} \qquad a.s.$$

Proof. Let

$$B = \left\{ \limsup_{t\to\infty} \frac{X_t^{(b)}}{\log k} = \frac{2\nu}{\lambda} \right\}$$

and assume $\mathbb{P}[B] > 0$. In the following we will always condition on B and the paths of the front two servers that satisfy that the limit

$$Y_\infty = \lim_{t\to\infty} \left(X_t^{(m)} - \frac{2\nu}{\lambda} \log t \right)$$

exists (the second event is an event of probability 1 by (4.2)). Conditional on these paths, the limit Y_∞ is deterministic. We will split the proof of Proposition 4.22 into three parts. We will prove

1. $\liminf_{t\to\infty} \left(X_t^{(m)} - X_t^{(b)} \right) \le \frac{2\nu}{\lambda} \log 2$ a.s.
2. $\liminf_{t\to\infty} \left(X_t^{(m)} - X_t^{(b)} \right)$ does not have any atoms
3. $\liminf_{t\to\infty} \left(X_t^{(m)} - X_t^{(b)} \right) < \frac{2\nu}{\lambda} \log 2$ a.s. leads to a contradiction

Statement 1 and statement 3 together imply that $\liminf_{t\to\infty} \left(X_t^{(m)} - X_t^{(b)} \right) = \frac{2\nu}{\lambda} \log 2$ with positive probability, but this contradicts statement 2. This is essentially because only at distance $\frac{2\nu}{\lambda} \log 2$ from the middle server, the back server has to serve the right number of customers to travel at speed $\frac{2\nu}{\lambda}$. At this distance, the back server has to serve half as many customers as the front two servers together and therefore it travels at the same speed. If the back server gets to close to the front two servers (distance less than $\frac{2\nu}{\lambda} \log 2$) or falls too much behind (distance greater than $\frac{2\nu}{\lambda} \log 2$) then it will either catch up (because it has to serve too few customers) or fall even further behind (because it has to serve too many customers).

Proof of statement 1:

Let C be the event that

$$\liminf_{t\to\infty}\left(X_t^{(m)}-X_t^{(b)}\right)>\frac{2\nu}{\lambda}\log 2$$

and assume $\mathbb{P}\left[C\right]>0$ (conditional on B and the paths of the front two servers). Then we can condition on C, too. On C we have that for some $\varepsilon>0$ we have, almost surely,

$$X_t^{(m)}-X_t^{(b)}\geq\frac{2\nu}{\lambda}\log 2+\varepsilon$$

for all large enough t. From the convergence statements for the front two servers in the first part of Theorem 4.8 we get

$$X_t^{(m)}\leq Y_\infty+\frac{2\nu}{\lambda}\log t+\frac{\varepsilon}{2}$$

and

$$T_x^{(m)}\leq\left(Z_\infty+\varepsilon\right)e^{\frac{\lambda}{2\nu}x}$$

for all large t and large x. Since we condition on the paths of the front two servers we can think of Y_∞ and Z_∞ as fixed. We have $Y_\infty=-\frac{2\nu}{\lambda}\log Z_\infty$. Putting the last three equations together we get that for any constant $a>0$, almost surely,

$$T^{(m)}_{X_t^{(b)}+a}\leq\left(1+\frac{\varepsilon}{Z_\infty}\right)\frac{t}{2}e^{-\frac{\lambda(\varepsilon-a)}{2\nu}}$$

for all large enough t. For a suitable $\beta<2$, $a>0$, $\varepsilon>0$ and large enough t we have

$$\lambda\left(t-\left(1+\frac{\varepsilon}{Z_\infty}\right)\frac{t}{2}e^{-\frac{\lambda(\varepsilon-a)}{2\nu}}\right)\geq\frac{\lambda t}{\beta}$$

and, using the same arguments as in the proof of Proposition 4.21, this implies

$$\limsup_{t\to\infty}\frac{X_t^{(b)}}{\log t}\leq\frac{\beta\nu}{\lambda}<\frac{2\nu}{\lambda}$$

and this is a contradiction.

Proof of statement 2:

To prove this statement we will change the model slightly in the following way. Instead of letting customers leave the system we mark them if they have been served by one of the front two servers but leave them in the system. Now recall the notation we used in (4.4) and (4.5). Here we can use a similar notation for the back server. Let

$$S_m^{(b)} = \sum_{i=1}^{m} E_i I_i \tag{4.42}$$

be the time of the m-th jump of the back server where I_i is the indicator function on the event that the customer corresponding to service E_i has not been marked. The E_i are again i.i.d. exponentials with parameter ν. Note that now we might have $S_m^{(b)} = S_{m+1}^{(b)}$ if the m-th customer was marked. However, by definition, we will always have $S_1^{(b)} > 0$. With this notation we can still write

$$U_m^{(b)} = \sum_{i=1}^{m} \frac{2\nu}{\lambda i} + 2\nu \sum_{i=1}^{m} \frac{\lambda F_i - 1}{\lambda i} + \sum_{i=1}^{m} F_i \left(\frac{1}{S_i^{(b)}} - \frac{2\nu}{i} \right) \tag{4.43}$$

where the F_i are again i.i.d. exponentials with parameter λ, independent of the E_i. However, the indicator functions in (4.42) depend in some complicated way on both the E_i and the F_i as well as on the paths of the front two servers. We know that the second sum converges almost surely to a finite random variable and

$$\lim_{m\to\infty} \left(\sum_{i=1}^{m} \frac{2\nu}{\lambda i} - \frac{2\nu}{\lambda} \log m \right) = \frac{2\nu}{\lambda}\gamma.$$

Conditional on B and the paths of the front two servers, the complementary event of statement 2 implies that the limit

$$\limsup_{m\to\infty} \left(U_m^{(b)} - \frac{2\nu}{\lambda} \log m \right)$$

has at least one atom. We can now also condition on the F_i. Since E_1 has a continuous distribution and increasing E_1 strictly decreases $U_m^{(b)}$, the $\limsup$ cannot have any atoms. This holds despite the dependence of the I_i on the E_i: increasing E_1 leads to server 3 having to serve more customers (more of the I_i will be equal to 1) and this strictly decreases $U_m^{(b)}$ and therefore the lim sup. This shows that $\liminf_{t\to\infty}\left(X_t^{(m)} - X_t^{(b)}\right)$ does not have any atoms and proves statement 2.

Proof of statement 3:

Assume now

$$\liminf_{t\to\infty}\left(X_t^{(m)} - X_t^{(b)}\right) < \frac{2\nu}{\lambda}\log 2 \qquad \text{a.s.}$$

This implies that for some $\varepsilon > 0$, almost surely,

$$X_t^{(m)} - X_t^{(b)} \le \frac{2\nu}{\lambda}\log 2 - \varepsilon$$

for arbitrarily large t. With

$$Y_\infty + \frac{2\nu}{\lambda}\log t + \frac{\varepsilon}{2} \ge X_t^{(m)} \ge Y_\infty + \frac{2\nu}{\lambda}\log t - \frac{\varepsilon}{2} \tag{4.44}$$

for all large t and

$$T_x^{(m)} \ge \left(Z_\infty - \frac{\varepsilon}{2}\right) e^{\frac{\lambda}{2\nu}x}$$

for all large x we get

$$X_t^{(b)} \ge Y_\infty + \frac{2\nu}{\lambda}\log\frac{t}{2} + \frac{\varepsilon}{2}$$

and

$$T^{(m)}_{X_t^{(b)}} \ge \left(1 - \frac{\varepsilon}{2Z_\infty}\right)\frac{t}{2}e^{\frac{\lambda\varepsilon}{4\nu}}$$

for arbitrarily large t. Now fix any t such that the inequalities above all hold. We want to get an upper bound on the time that it takes the back server to get from $X_t^{(b)}$ to site $X_t^{(b)} + a$ for some constant a that we will choose later. Note that we have

$$\begin{aligned} t - T^{(m)}_{X_t^{(b)}} &\le t - \left(1 - \frac{\varepsilon}{2Z_\infty}\right)\frac{t}{2}e^{\frac{\lambda\varepsilon}{4\nu}} \\ &= t\left(1 - \left(1 - \frac{\varepsilon}{2Z_\infty}\right)\frac{1}{2}e^{\frac{\lambda\varepsilon}{4\nu}}\right) \end{aligned}$$

$$= \beta t$$

with

$$\beta = 1 - \left(1 - \frac{\varepsilon}{2Z_\infty}\right)\frac{1}{2}e^{\frac{\lambda\varepsilon}{4\nu}} < \frac{1}{2}.$$

Therefore, the number of customers in the interval $\left[X_t^{(b)}, X_t^{(b)} + a\right]$ at time t is bounded from above by a Poisson random variable with parameter $\lambda\beta ta$. We bound the time it takes the back server to get to $X_t^{(b)} + a$ in the following way: let V be the time it takes the server to serve all the customers that are in the interval $\left[X_t^{(b)}, X_t^{(b)} + a\right]$ at time t plus all the customers that arrive in the interval while the server is busy. Since the expectation of a busy period with service rate ν and arrival rate λa is $\frac{1}{\nu - \lambda a}$ we get

$$\mathbb{E}\left[V\right] = \lambda\beta ta\frac{1}{\nu - \lambda a}$$

and for any $\delta > 0$ we have

$$\mathbb{P}\left[V \leq (1+\delta)\,\mathbb{E}\left[V\right]\right] > \frac{\delta}{1+\delta}.$$

This means that with probability of at least $\frac{\delta}{1+\delta}$ the back server will be to the right of $X_t^{(b)} + a$ at time $t + \lambda\beta ta\frac{1+\delta}{\nu-\lambda a}$. For the middle server we use inequality (4.44) to get

$$\begin{aligned}
X^{(m)}_{t+\lambda\beta ta\frac{1+\delta}{\nu-\lambda a}} - X_t^{(m)} &\leq Y_\infty + \frac{2\nu}{\lambda}\log\left(t + \lambda\beta ta\frac{1+\delta}{\nu-\lambda a}\right) + \frac{\varepsilon}{2} - Y_\infty - \frac{2\nu}{\lambda}\log t + \frac{\varepsilon}{2} \\
&= \frac{2\nu}{\lambda}\log\left(1 + \lambda\beta a\frac{1+\delta}{\nu-\lambda a}\right) + \varepsilon \\
&\leq \frac{2\nu}{\lambda}\lambda\beta a\frac{1+\delta}{\nu-\lambda a} + \varepsilon \\
&= 2\beta\left(1+\delta\right)\frac{1}{1-\frac{\lambda a}{\nu}}a + \varepsilon.
\end{aligned}$$

Because δ, ε and a can be chosen arbitrarily small and $\beta < \frac{1}{2}$ we can get

$$2\beta\left(1+\delta\right)\frac{1}{1-\frac{\lambda a}{\nu}}a + \varepsilon \leq a - \tau$$

for some $\tau > 0$. This implies that with positive probability we have

$$X^{(m)}_{t+\lambda\beta ta\frac{1+\delta}{\nu-\lambda a}} - X^{(b)}_{t+\lambda\beta ta\frac{1+\delta}{\nu-\lambda a}} \leq \frac{2\nu}{\lambda}\log 2 - \varepsilon - \tau.$$

Since the probability is bounded from below by a constant ($\frac{\delta}{1+\delta}$) independent of t, we get that, almost surely,

$$X^{(m)}_t - X^{(b)}_t \leq \frac{2\nu}{\lambda}\log 2 - \varepsilon - \tau$$

for arbitrarily large t. We can now iterate this argument to see that eventually

$$X^{(m)}_t - X^{(b)}_t < 0$$

i.e. the back server overtakes the middle server almost surely. But since we had conditioned on A^c this leads to a contradiction. □

It remains to show that if $\limsup_{t\to\infty} \frac{X^{(b)}_t}{\log t} = \frac{\nu}{\lambda}$ we actually have convergence of $X^{(b)}_t - \frac{\nu}{\lambda}\log t$.

Proposition 4.23. *On A^c we have that*

$$X^{(b)}_t - \frac{\nu}{\lambda}\log t$$

converges almost surely to a finite random variable.

Proof. Because the asymptotic speed of the third server cannot be less than the asymptotic speed of the server in the model with $n = 1$ we have

$$\liminf_{t\to\infty} \frac{X^{(b)}_t}{\log t} \geq \frac{\nu}{\lambda} \qquad \text{a.s.} \tag{4.45}$$

Propositions 4.21, 4.22 and (4.45) imply

$$\lim_{t\to\infty} \frac{X^{(b)}_t}{\log t} = \frac{\nu}{\lambda} \qquad \text{a.s.}$$

Therefore, almost surely,

$$X_t^{(b)} \leq \left(\frac{\nu}{\lambda} + \varepsilon\right) \log t$$

for all large enough t and, almost surely,

$$T^{(m)}_{X_t^{(b)}} \leq (Z_\infty + \varepsilon)\, e^{\frac{\lambda}{2\nu}\left(\frac{\nu}{\lambda}+\varepsilon\right)\log t} = (Z_\infty + \varepsilon)\, t^{\frac{1}{2}+\tau}$$

for all large enough t and with $\tau = \frac{\lambda\varepsilon}{2\nu}$. We now use the slightly changed model that we introduced in the proof of statement 2 again, see (4.42) and the discussion before it. Write

$$U_m^{(b)} = \sum_{i=1}^{m} \frac{\nu}{\lambda i} + \nu \sum_{i=1}^{m} \frac{\lambda F_i - 1}{\lambda i} + \sum_{i=1}^{m} F_i \left(\frac{1}{S_i^{(b)}} - \frac{\nu}{i} \right).$$

The convergence of the first two sums follows again from the same arguments as the ones used in the proofs of Theorems 4.2 and 4.4. For the third sum we observe that it follows from (4.3.3) that

$$m - \sum_{i=1}^{m} I_i \leq O\left(m^{\frac{1}{2}+\tau}\right)$$

almost surely and this implies

$$\frac{m}{\nu} - O\left(m^{\frac{1}{2}+\tau}\right) \leq S_m \leq \frac{m}{\nu} + O\left(m^{\frac{1}{2}+\tau}\right)$$

almost surely. Now we can proceed in the same way as in the proof of Theorem 4.2 to see that the last sum in (4.43) converges almost surely to a finite random variable. □

Propositions 4.21, 4.22 and 4.23 together complete the proof of the second part of Theorem 4.8. □

Bibliography

[1] G. Amir, O. Angel, and B. Valko. The TASEP speed process. *ArXiv e-prints, arXiv:0811.3706 [math.PR]*, November 2008.

[2] François Baccelli and Pierre Brémaud. *Elements of queueing theory*, volume 26 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2003.

[3] Daniel Bernoulli. *Danielis Bernoulli Joh. Fil. Hydrodynamica, sive, De viribus et motibus fluidorum commentarii [microform] / opus academicum ab auctore, dum Petropoli ageret, congestum.* Sumptibus Johannis Reinholdi Dulseckeri, typis Joh. Henr. Deckeri, Typographi Basiliensis, Argentorati, 1738.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1989.

[5] G. Biroli, J.-P. Bouchaud, and M. Potters. On the top eigenvalue of heavy-tailed random matrices. *Europhys. Lett. EPL*, 78(1):Art. 10001, 5, 2007.

[6] Giulio Biroli, Jean-Philippe Bouchaud, and Marc Potters. Extreme value problems in random matrix theory and other disordered systems. *J. Stat. Mech. Theory Exp.*, (7):P07019, 15 pp. (electronic), 2007.

[7] R. A. Blythe and M. R. Evans. Nonequilibrium steady states of matrix-product form: a solver's guide. *J. Phys. A*, 40(46):R333–R441, 2007.

[8] S. Boucheron and W. Fernandez de la Vega. On the independence number of random interval graphs. *Combin. Probab. Comput.*, 10(5):385–396, 2001.

[9] Paul J. Burke. The output of a queuing system. *Operations Res.*, 4:699–704, 1956.

[10] Debashish Chowdhury, Ludger Santen, and Andreas Schadschneider. Statistical physics of vehicular traffic and some related systems. *Phys. Rep.*, 329(4-6):199–329, 2000.

[11] E. G. Coffman, Jr. and E. N. Gilbert. Polling and greedy servers on a line. *Queueing Systems Theory Appl.*, 2(2):115–145, 1987.

[12] Joel E. Cohen, Frédéric Briand, and Charles M. Newman. *Community food webs*, volume 20 of *Biomathematics*. Springer-Verlag, Berlin, 1990.

[13] Herbert Aron David. *Order statistics*. John Wiley & Sons Inc., New York, second edition, 1981.

[14] D. Denisov, S. Foss, and T. Konstantopoulos. Limit theorems for a random directed slab graph. *ArXiv e-prints, arXiv:1005.4806 [math.PR]*, May 2010.

[15] P. A. Ferrari and C. Kipnis. Second class particles in the rarefaction fan. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(1):143–154, 1995.

[16] Pablo A. Ferrari, Patrícia Gonçalves, and James B. Martin. Collision probabilities in the rarefaction fan of asymmetric exclusion processes,. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(4):1048–1064, 2009.

[17] Pablo A. Ferrari and James B. Martin. Multi-class processes, dual points and $M/M/1$ queues. *Markov Process. Related Fields*, 12(2):175–201, 2006.

[18] Pablo A. Ferrari and James B. Martin. Stationary distributions of multi-type totally asymmetric exclusion processes. *Ann. Probab.*, 35(3):807–832, 2007.

[19] Pablo A. Ferrari and Leandro P. R. Pimentel. Competition interfaces and second class particles. *Ann. Probab.*, 33(4):1235–1254, 2005.

[20] S. Foss and T. Konstantopoulos. Extended renovation theory and limit theorems for stochastic ordered graphs. *Markov Process. Related Fields*, 9(3):413–468, 2003.

[21] S. Foss, J. Martin, and P. Schmidt. Long-range last-passage percolation on the line. *ArXiv e-prints, arXiv:1104.2420 [math.PR]*, April 2011.

[22] Serguei Foss and Günter Last. Stability of polling systems with exhaustive service policies and state-dependent routing. *Ann. Appl. Probab.*, 6(1):116–137, 1996.

[23] Serguei Foss and Günter Last. On the stability of greedy polling systems with general service policies. *Probab. Engrg. Inform. Sci.*, 12(1):49–68, 1998.

[24] E. Gelenbe, R. Nelson, T. Philips, and A. Tantawi. An approximation of the processing time for a random graph model of parallel computation. In *Proceedings of 1986 ACM Fall joint computer conference*, ACM '86, pages 691–697, Los Alamitos, CA, USA, 1986. IEEE Computer Society Press.

[25] John W. Hagood. The Lebesgue differentiation theorem via nonoverlapping interval covers. *Real Anal. Exchange*, 29(2):953–956, 2003/04.

[26] Ben Hambly and James B. Martin. Heavy tails in last-passage percolation. *Probab. Theory Related Fields*, 137(1-2):227–275, 2007.

[27] Arie Harel and Alan Stulman. Polling, greedy and horizon servers on a circle. *Operations Research*, 43(1):pp. 177–186, 1995.

[28] T. E. Harris. Additive set-valued Markov processes and graphical methods. *Ann. Probability*, 6(3):355–378, 1978.

[29] Dirk Helbing. Traffic and related self-driven many-particle systems. *Rev. Mod. Phys.*, 73(4):1067–1141, Dec 2001.

[30] Haye Hinrichsen. Matrix product ground states for exclusion processes with parallel dynamics. *J. Phys. A*, 29(13):3659–3667, 1996.

[31] Marco Isopi and Charles M. Newman. Speed of parallel processing for random task graphs. *Comm. Pure Appl. Math.*, 47(3):361–376, 1994.

[32] Edward G. Coffman Jr. and Edgar N. Gilbert. A continuous polling system with constant service times. *IEEE Transactions on Information Theory*, 32(4):584–591, 1986.

[33] Joyce Justicz, Edward R. Scheinerman, and Peter M. Winkler. Random intervals. *Amer. Math. Monthly*, 97(10):881–889, 1990.

[34] Dirk P. Kroese and Volker Schmidt. Single-server queues with spatially distributed arrivals. *Queueing Systems*, 17:317–345, 1994. 10.1007/BF01158698.

[35] L. Kultaeva. The server's asymptotics in a poisson rain. *Master dissertation, Novisibirsk State University*, 2001.

[36] I. A. Kurkova. A sequential clearing process. *Fundam. Prikl. Mat.*, 2(2):619–624, 1996.

[37] I.A. Kurkova and M.V. Menshikov. Greedy algorithm, z^1 case. *Markov Processes and Related Fields*, 3(2):243–259, 1997.

[38] Thomas M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[39] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[40] J. Martin and P. Schmidt. Multi-type TASEP in discrete time. *ArXiv e-prints, arXiv:1002.3539 [math.PR]*, February 2010.

[41] T. Mountford and B. Prabhakar. On the weak convergence of departures from an infinite series of $\cdot/M/1$ queues. *Ann. Appl. Probab.*, 5(1):121–127, 1995.

[42] Charles M. Newman. Chain lengths in certain random directed graphs. *Random Structures Algorithms*, 3(3):243–253, 1992.

[43] Neil M. O'Connell. Directed percolation and tandem queues. HP Labs technical report, HPL-BRIMS-2000-28, `http://www.hpl.hp.com/techreports/2000/`, 2000.

[44] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg. The asymmetric exclusion process: comparison of update procedures. *J. Statist. Phys.*, 92(1-2):151–194, 1998.

[45] H. Rost. Nonequilibrium behaviour of a many particle process: density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete*, 58(1):41–53, 1981.

[46] Günther Schütz. Time-dependent correlation functions in a one-dimensional asymmetric exclusion process. *Phys. Rev*, 1993.

[47] G. Shakirova. A server in a poisson rain. *Master dissertation, Novisibirsk State University*, 1997.

[48] Frank Spitzer. Interaction of Markov processes. *Advances in Math.*, 5:246–290, 1970.