

DEPARTMENT OF
STATISTICS

Convergence of exclusion processes and related phenomena

A thesis submitted in partial fulfilment of the requirements for the degree
Doctorate of Philosophy in Statistics

January 2023

James Ayre
St Catherine's College

Department of Statistics
University of Oxford

Abstract

We examine convergence to equilibrium for the facilitated simple exclusion process (FSEP) on the segment with closed boundaries, as well as the symmetric FSEP on the circle. The FSEP on these spaces has degenerate rates, and exhibits transient states which the process will eventually leave to reach an ergodic component if the macroscopic density of particles is at least $1/2$. If the macroscopic density is less than $1/2$ the process will hit one of its absorbing states.

We first consider the symmetric FSEP on the segment $[N] = \{1, \dots, N\}$ with closed boundaries and $k > N/2$ particles. We show that the *pre–cutoff* phenomenon occurs and the mixing time has order $k^2 \log(N-k)$. This agrees with the order of the mixing time for the symmetric simple exclusion process on the segment $[k]$ with $N-k$ particles under suitable initial conditions. We also show that the *cutoff* phenomenon occurs for the symmetric FSEP on the segment restricted to the ergodic component. Here the mixing time has order $k^2 \log \min(N-k, 2k-N)$.

For the asymmetric FSEP on the segment, we show that the behaviour of the mixing time is qualitatively different from that of the asymmetric simple exclusion process on the segment; there exists initial conditions for which the hitting time of the ergodic component is exponentially slow in the number of holes $N-k$. In particular, when $N-k$ is large enough, the hitting time of the ergodic component determines the mixing time. However, for the asymmetric FSEP restricted to the ergodic component, the cutoff phenomenon occurs and the mixing time has order k. This agrees with the order of the mixing time for the asymmetric simple exclusion process on the segment $[k]$.

For the symmetric FSEP on the circle of size N, and macroscopic particle density $\alpha \in (1/2, 1)$, we establish that the pre–cutoff phenomenon occurs on the ergodic component and that the order of the mixing times is $N^2 \log N$. This agrees with the order of the mixing time of the symmetric simple exclusion process on the circle of size N. We also give an upper bound on the hitting time of the ergodic component of order $N^2 \log N$ for a large class of initial conditions. In particular, the mixing time of the FSEP restricted to this class of initial conditions also agrees with the order of the mixing time of the symmetric simple exclusion process on the circle of size N.

The dynamical constraint of the FSEP makes applying standard techniques for bounding the mixing time challenging. All results for the segment are obtained via comparisons to the simple exclusion process, the open boundary exclusion process and the zero–range process. We will also compare to these processes for the FSEP on the circle, although the comparisons here are less straightforward.

We then move on to consider the problem of efficiently producing a uniform random sample from the set of permutations on N objects with exactly k cycles for some $1 < k < N$. We present an algorithm, with expected run–time $O(k * \text{polylog}N)$, that samples from the distribution of the multiset of cycle lengths of a uniform random permutation with N objects, conditioned to have k cycles.

Our approach is based on the Feller representation of a random permutation. We design a Markov chain (specifically, an exclusion process) whose stationary distribution projects to the distribution of the Feller representation conditioned on having a given number of cycles. We show how to efficiently obtain an exact sample from this stationary distribution using an approach based on the *Coupling From The Past* algorithm.

We present an update scheme which allows us to update our Markov chain started from all initial conditions simultaneously, i.e. a grand coupling. We show that this update scheme has an interesting structure: applying the sequence of updates $U_1, \ldots, U_{t-1}, U_t$ in the update scheme gives coalescence if and only if applying the reversed sequence $U_t, U_{t-1}, \ldots, U_1$ in the update scheme gives coalescence. This gives rise to an alternative form of the Coupling From The Past algorithm for a class of Markov chains.

We present a coupling for the distributions of the multiset of cycle lengths of uniform random permutations on N objects with i cycles, for all $1 \leq i \leq k$. Samples from this coupling have the following consistency property: the sample with $i+1$ cycles is obtained from the sample with i cycles by breaking one of the cycles into two parts of integer size. In other words, we explicitly construct the trajectory of a *fragmentation process*. We then give an algorithm which obtains an exact sample from the coupling with expected run–time $O(k^2 \text{polylog} N)$. This algorithm is based on an application of the Coupling From The Past algorithm to a version of the aforementioned exclusion process in which particles are assigned a priority.

Finally, we construct a directed percolation model termed *intermediate–passage percolation*. We show that, on the same random environment, the intermediate–passage percolation model is bounded above and below by the well–studied cases of directed first and last–passage percolation respectively. We then investigate the growth of this model and give sufficient conditions for linear growth, as a function of time on a linear scale, in all directions. The conditions we give include cases for which first or last–passage percolation do not have linear growth in all directions. We then state some open problems and give several simulations of the process.

Contents

1 Introduction **1**

1.1 Organisation of the chapter 1

1.2 Interacting particle systems 1

1.2.1 Coupling and monotonicity 3

1.3 Exclusion processes 4

1.3.1 The open boundary exclusion process 8

1.3.2 The facilitated exclusion process 9

1.4 The zero–range process 12

1.4.1 Coupling with the FSEP 13

1.5 Mixing times 14

1.5.1 The spectral gap and log–Sobolev constant 17

1.6 Mixing times for exclusion processes 20

1.6.1 Mixing times for the OBEP 22

1.7 Coupling from the past 23

1.7.1 Some examples 24

1.8 Directed percolation 27

1.8.1 Intermediate–passage percolation. 31

2 Mixing times of the facilitated exclusion process **34**

2.1 Notation and results 36

2.1.1 Results for the segment 36

2.1.2 Results for the circle 37

2.2 A key observation 39

2.3 FSEP on the segment 39

2.3.1 Mapping to lattice paths 40

2.3.2 Monotonicity 44

2.3.3 Comparison to the OBEP 46

2.3.4 Proofs . . . 51
2.4 FSEP on the circle . . . 63
2.4.1 Lower bound on the mixing time . . . 64
2.4.2 Upper bound on the mixing time . . . 66
2.4.3 Hitting time of the ergodic component . . . 69
2.4.4 Lower bound for the log–Sobolev constant . . . 72
2.4.5 Concentration of particle density in the half–circle . . . 82
2.5 Open problems . . . 84

3 Sampling the cycle structure of a uniform random permutation 87
3.1 Background and motivation . . . 87
3.1.1 Uniform random permutations . . . 88
3.1.2 Fragmentation processes . . . 90
3.2 Results . . . 91
3.2.1 Processes . . . 93
3.2.2 Tracking sets for the Labelled process . . . 96
3.2.3 Reversing the order of updates for the Labelled process . . . 97
3.3 Remarks about sampling . . . 100
3.4 Proofs . . . 101
3.4.1 Proof of Theorem 3.4 . . . 104
3.4.2 Proof of Theorem 3.3 . . . 109

4 Intermediate–passage percolation 113
4.1 Introduction . . . 113
4.2 Motivation . . . 115
4.3 Results . . . 116
4.4 Proof of Theorem 4.1 . . . 119
4.5 Open questions and simulations . . . 123

5 Conclusions and further research 130

Acknowledgements

There are several people to whom I owe a great debt of gratitude for their continued support throughout my DPhil. First and foremost I am immensely grateful to my supervisors Prof. James Martin and Prof. Paul Chleboun for their guidance and supervision. Thank you both for your commitment, particularly during periods of global turmoil. Thanks are also due to my transfer, confirmation and viva examiner Prof. Julien Berestycki, my viva examiner Prof. Stephen Connor and my transfer and confirmation examiner Prof. Alison Etheridge. Thank you all for your feedback.

I would also like to thank my partner, Molly. I'm glad we have shared this time, and this degree would have been far harder without your company. Finally, I would like to thank several family members and friends. Thank you to my parents, Audrey and Kevin, my sister Rachael and her fiancé Tom, my grandfather John, my grandmother Sheila and my friends Eddie and Autrin. Thank you for supporting me from afar and providing me with the opportunities I needed in order to undertake a doctorate at the University of Oxford.

Chapter 1

Introduction

1.1 Organisation of the chapter

In this chapter we will introduce various topics and give a literature review. In Section 1.2 we introduce finite Markov chains in the context of interacting particle systems, and in Section 1.3 we introduce a collection of the interacting particle systems we consider in later chapters. In particular we introduce the simple exclusion process, the open boundary exclusion process and the facilitated simple exclusion process. We also introduce the zero–range process in Section 1.4, as well as a coupling between the facilitated simple exclusion process and the zero–range process.

In Section 1.5 we will give a general introduction to mixing times, as well as some related objects like the spectral gap and log–Sobolev constant. We then discuss the mixing time literature for the simple exclusion process and open boundary exclusion process in Section 1.6. The material from Section 1.2 to Section 1.6 serves as an introduction to Chapter 2.

In Section 1.7 we introduce the Coupling from the past algorithm, and give some relevant examples. This section serves as a short introduction to Chapter 3.

Finally, in Section 1.8 we introduce some directed percolation models. We introduce the well–studied cases of directed first and last–passage percolation, as well as a new process termed intermediate–passage percolation. Section 1.8 serves as an introduction to Chapter 4.

1.2 Interacting particle systems

Interacting particle systems are continuous–time Markov chains on spaces of configurations of particles. Typically, the dynamics of these processes are described locally by infinitesimal rates at which transitions between configurations occur.

We introduce continuous–time Markov chains on a finite state space Ω following the notes of Saloff–Coste [96] and Liggett [78]. The space Ω is endowed with the discrete topology, with respect to which all functions $f : \Omega \to \mathbb{R}$ are continuous.

The infinitesimal transition rates of a Markov chain are given by a Q–matrix.

Definition 1.2.1. *A Q–matrix is a collection of real numbers $Q(\zeta, \zeta')$ indexed by $\zeta, \zeta' \in \Omega$*

satisfying

$$Q(\zeta, \zeta') \geq 0,$$

for all $\zeta, \zeta' \in \Omega_{N,k}$ *with* $\zeta' \neq \zeta$, *and*

$$\sum_{\zeta' \in \Omega} Q(\zeta, \zeta') = 0.$$

In particular $Q(\zeta, \zeta) \leq 0$.

For a Markov chain in state ζ with a given Q–matrix, the chain moves to state $\zeta' \neq \zeta$ with probability $-Q(\zeta, \zeta')/Q(\zeta, \zeta)$ after an Exponential$(-Q(\zeta, \zeta))$ waiting time. The generator of a Markov chain $\mathcal{L}$ is an operator acting on functions $f : \Omega \to \mathbb{R}$ as follows:

$$\mathcal{L}f(\zeta) = \sum_{\zeta' \in \Omega} Q(\zeta, \zeta')\big[f(\zeta') - f(\zeta)\big]. \tag{1.2.1}$$

The continuous–time semigroup associated with the generator $\mathcal{L}$ is defined by

$$P_t f = \exp(t\mathcal{L})f = \sum_{n=0}^{\infty} \frac{t^n \mathcal{L}^n f}{n!}.$$

Observe that this is indeed a semigroup of operators, that is

$$\begin{aligned} P_{t+s} &= P_t P_s, \\ \lim_{t \downarrow 0} P_t &= I. \end{aligned}$$

We set

$$P_t(\zeta, \zeta') = P_t^{\zeta}(\zeta') = e^{-t} \sum_{n=0}^{\infty} \frac{t^n Q^n(\zeta, \zeta')}{n!}.$$

In particular $P_t^{\zeta}(\cdot)$ is a probability measure on Ω which represents the distribution at time t of the continuous–time Markov chain $(\zeta_t)_{t \geq 0}$ associated with the generator $\mathcal{L}$ and initial condition $\zeta \in \Omega$.

In the setting of finite state spaces, Q–matrices may be derived from the transition functions $P_t(\zeta, \zeta')$: $\exp(tQ)$ is the unique solution of the *finite* system of differential equations

$$\frac{d}{dt} P_t(\zeta, \zeta') = \sum_{\eta \in \Omega} Q(\zeta, \eta) P_t(\eta, \zeta'). \tag{1.2.2}$$

These are the backwards Kolmogorov equations.

A probability distribution π is stationary for a Markov chain, on a finite state space Ω, if its Q–matrix satisfies

$$\pi Q = 0. \tag{1.2.3}$$

We call Q irreducible if for all $\zeta, \zeta' \in \Omega$, with $\zeta \neq \zeta'$, there exists an $n > 0$ such that $Q^n(\zeta, \zeta') > 0$. When Q is irreducible there exists a unique stationary measure; see, for example, [96, Lemma 1.2.2].

A pair (Q, π) is reversible if it satisfies the detailed balance equations

$$\pi(\zeta) Q(\zeta, \zeta') = \pi(\zeta') Q(\zeta', \zeta), \tag{1.2.4}$$

for all $\zeta, \zeta' \in \Omega$. Summing the detailed balance equations over $\zeta \in \Omega$ gives (1.2.3). In particular if the pair (Q, π) satisfies the detailed balance equations and Q is irreducible, π is the unique stationary measure for the Markov chain associated to the Q–matrix.

1.2.1 Coupling and monotonicity

Following [78, Chapter 4], we introduce coupling and monotonicity for the example of the contact process. The state space is $\{0,1\}^S$ for some finite S. Let ζ^x be the configuration obtained by 'flipping' the value of ζ at x from $\zeta(x)$ to $1-\zeta(x)$, leaving the other coordinates unchanged. In particular

$$\zeta^x(y) = \begin{cases} 1-\zeta(x) & \text{if } x = y, \\ \zeta(y) & \text{otherwise.} \end{cases}$$

The Q–matrix is only positive for entries with the form $Q(\zeta, \zeta^x)$, and the generator $\mathcal{L}$ acts on functions $f : \Omega \to \mathbb{R}$ via

$$\mathcal{L}f(\zeta) = \sum_x Q(\zeta, \zeta^x)\big[f(\zeta^x) - f(\zeta)\big]. \tag{1.2.5}$$

Suppose that S is a simple graph, and let $x \sim y$ denote that the vertices x and y are neighbours in the graph S. The contact process with parameter $\lambda > 0$ has flip rates

$$Q(\zeta, \zeta^x) = \begin{cases} 1 & \text{if } \zeta(x) = 1, \\ \lambda\big|\{y \sim x : \zeta(y) = 1\}\big| & \text{if } \zeta(x) = 0. \end{cases}$$

The interpretation is that sites $x \in S$ are infected in the contact process if $\zeta(x) = 1$ and healed if $\zeta(x) = 0$. A healed site x becomes infected at rate proportional to the number of infected neighbours $y \sim x$, where the constant of proportionality is λ. An infected site heals independently at rate 1.

For our purposes, it will often be useful to give comparisons between processes using the *coupling technique*. A coupling of two Markov chains is simply a construction of the two processes on a common probability space. This allows us to compare two (or more) processes provided the joint evolution of the two processes is chosen in a useful way. The following example gives a coupling of two copies of the contact process started from different initial configurations.

Example 1.2.2. *For two configurations $\zeta, \xi \in \{0,1\}^S$ we say that $\zeta \le \xi$ if $\zeta(x) \le \xi(x)$ for all $x \in S$. Let $S = \mathbb{Z}/N\mathbb{Z} =: \mathbb{T}_N$ and we denote the elements of $\mathbb{T}_N$ by $\{0, 1, \dots, N-1\}$. We consider the simple graph with vertex set S and edges $\{x, x+1\}$ for all $0 \le x \le N-1$, i.e. the circle.*

We construct a coupling (ζ_t, ξ_t) for two contact processes $(\zeta_t)_{t\ge0}$ and $(\xi_t)_{t\ge0}$ on $\{0,1\}^{\mathbb{T}_N}$ with initial conditions $\zeta \le \xi$. The coupling is defined by specifying the rates at which transitions occur for the coordinates $(\zeta(x), \xi(x))$ at each x, if $\zeta \le \xi$, as follows:

$$(0,0) \to \begin{cases} (1,1) & \text{at rate } \lambda\zeta(x+1) + \lambda\zeta(x-1), \\ (0,1) & \text{at rate } \lambda\sum_{i=\pm1}\big[\xi(x+i) - \zeta(x+i)\big], \end{cases}$$

$$(0,1) \to \begin{cases} (0,0) & \text{at rate } 1, \\ (1,1) & \text{at rate } \lambda\zeta(x+1) + \lambda\zeta(x-1), \end{cases}$$

$$(1,1) \to (0,0) \text{ at rate } 1.$$

the assumption $\zeta \le \xi$ guarantees that all these rates are nonnegative. The marginal rates are such that the marginal processes ζ_t and ξ_t evolve as contact processes.

More formally (ζ_t, ξ_t) is a Markov chain on the state space $\{(0,0),(0,1),(1,1)\}^{\mathbb{T}_N}$ with a generator analogous to (1.2.5), but with more terms, since there are more types of transitions.

Moreover, the coupling satisfies

$$P_t^{(\zeta,\xi)}\big[(\zeta',\xi') : \zeta' \leq \xi'\big] = 1, \tag{1.2.6}$$

for all $t \geq 0$.

The coupling of two copies of the contact process in Example 1.2.2 preserves the monotonicity. The coupling technique will be particularly important in Chapter 2 and we will introduce a monotonicity for the facilitated simple exclusion process that is preserved under a coupling.

1.3 Exclusion processes

The simple exclusion process (SEP) is an interacting particle system introduced by Spitzer [101] to model a system of particles in which any site can be occupied by at most one particle at a time. The SEP is one of the must fundamental and well–studied processes in the interacting particle systems literature, and is closely tied to card shuffling [60, 63, 106]. The state space we consider is $\Omega = \{0,1\}^S$ for some finite S; the interpretation is that there is a particle at site $x \in S$ if $\zeta(x) = 1$, and there is a hole at site $x \in S$ if $\zeta(x) = 0$. In particular at most one particle can occupy a site $x \in S$ at any given time. A particle at site x attempts to move to site y at rate $p(x,y)$, only doing so if there is a hole at y.

For $\zeta \in \Omega$ we let

$$\zeta^{x,y}(z) = \begin{cases} \zeta(y) & \text{if } z = x, \\ \zeta(x) & \text{if } z = y, \\ \zeta(z) & \text{otherwise,} \end{cases}$$

i.e. the configuration with the values at site x and y swapped. The generator $\mathcal{L}_{sep}$ of the SEP with rates $p(x,y)$ acts on functions $f : \Omega \to \mathbb{R}$ via

$$\mathcal{L}_{sep} f(\zeta) = \sum_{(x,y)\in\Omega\times\Omega} p(x,y)\zeta(x)\big(1-\zeta(y)\big)\big[f(\zeta^{x,y}) - f(\zeta)\big].$$

Exclusion processes on infinite graphs have a rich literature (see, for example, [76, 78]), but we restrict ourselves to the setting of finite graphs. In particular, we consider the nearest neighbour SEP on the segment with closed boundaries.

For $[N] := \{1,\dots,N\}$ we let

$$\Omega_N := \{0,1\}^{[N]},$$

denote the state space. In the SEP on the segment a particle at site $x \in [N]$ attempts to jump right at rate $p \in [0,1]$, only doing so if $x < N$ and site $x+1$ is not occupied by a particle. A particle a site $x \in [N]$ attempts to jump left at rate $q = 1-p$, only doing so if $x > 1$ and site $x-1$ is not occupied by a particle.

The generator $\mathcal{L}_{ex}$ of the SEP on the segment $[N]$ with parameter $p \in [0,1]$ acts on function $f : \Omega_N \to \mathbb{R}$ via

$$\begin{aligned} \mathcal{L}_{ex} f(\zeta) = &\sum_{x=1}^{N-1} p\zeta(x)\big(1-\zeta(x+1)\big)\big[f(\zeta^{x,x+1}) - f(\zeta)\big] \\ &+ \sum_{x=2}^{N} q\zeta(x)\big(1-\zeta(x-1)\big)\big[f(\zeta^{x,x-1}) - f(\zeta)\big]. \end{aligned} \tag{1.3.1}$$

One may view particles as performing independent, continuous–time, simple random walks on the segment in which particles attempt to jump right at rate p, and left at rate $q = 1 - p$, only doing so if the target site is not occupied by a particle.

The number of particles in the SEP on the segment is conserved and we may restrict to the state space

$$\Omega_{N,k} = \left\{ \zeta \in \{0,1\}^N \;\middle|\; \sum_{x=1}^{N} \zeta(x) = k \right\}, \tag{1.3.2}$$

i.e. the set of exclusion configurations on the segment of size N with exactly k particles. Observe that the SEP on $\Omega_{N,k}$ is irreducible for $p \in (0,1)$.

By taking the mirror image of the segment, it is sufficient to consider the parameters $p \geq 1/2$. We call the simple exclusion process on the segment with parameter $p = 1/2$, $p \in (1/2, 1)$ and $p = 1$ the symmetric simple exclusion process (SSEP), the asymmetric simple exclusion process (ASEP) and the totally asymmetric simple exclusion process (TASEP) on the segment respectively.

Particle and holes play symmetric roles in the SEP on the segment, and holes may be viewed as performing independent random walks, jumping right with rate q and left with rate p, only doing so if the target site exists and there is no hole at the target site. Without loss of generality we may take $k \leq N/2$ by particle–hole symmetry.

To identify the stationary distribution π of the SEP on the segment, we check the detailed balance equations (1.2.4). For $p = 1/2$, we have

$$\pi(\zeta) \times \frac{1}{2} = \pi(\zeta') \times \frac{1}{2}, \tag{1.3.3}$$

for all pairs (ζ, ζ') such that $Q(\zeta, \zeta') > 0$. Therefore π is uniform on the set $\Omega_{N,k}$, i.e.

$$\pi(\zeta) = \frac{1}{|\Omega_{N,k}|}.$$

For $p > 1/2$, we label particles from left to right by $i = 1, \ldots, k$ and let x_i denote the position of the i–th particle from the left in the configuration ζ. Let

$$A(\zeta) = \sum_{i=1}^{k} x_i,$$

and $\lambda = p/q$. Suppose that ζ is such that $\zeta(x_i + 1) = 0$ for some $i \in [k]$ and let $\zeta' = \zeta^{x_i, x_i+1}$, then by the detailed balance equations

$$\pi(\zeta) \times p = \pi(\zeta') \times q.$$

In fact this gives all of the detailed balance equations. Noting that

$$\lambda^{-A(\zeta)} \times p = \left(\frac{q}{p}\right)^{\sum x_j} \times p = \left(\frac{q}{p}\right)^{\sum_{j \neq i} x_j} \times \left(\frac{q}{p}\right)^{x_i + 1} \times q = \lambda^{-A(\zeta')} \times q,$$

it follows that the unique stationary distribution for the SEP on the segment is given by

$$\pi(\zeta) = \frac{\lambda^{-A(\zeta)}}{\sum_{\zeta' \in \Omega_{N,k}} \lambda^{-A(\zeta')}}.$$

An equivalent description of the SEP on the segment is given by the following *height function*. To each $\zeta \in \Omega_{N,k}$ we associate a lattice path η^ζ in the following manner:

$$\eta^\zeta(x) := \sum_{y=1}^{x} \zeta(y) - \frac{kx}{N}, \text{ for all } 0 \leq x \leq N. \tag{1.3.4}$$

This description may be found in [63, Figure 1], and was inspired by [106, Figure 3]. See Figure 1.1 for an example of the height function.

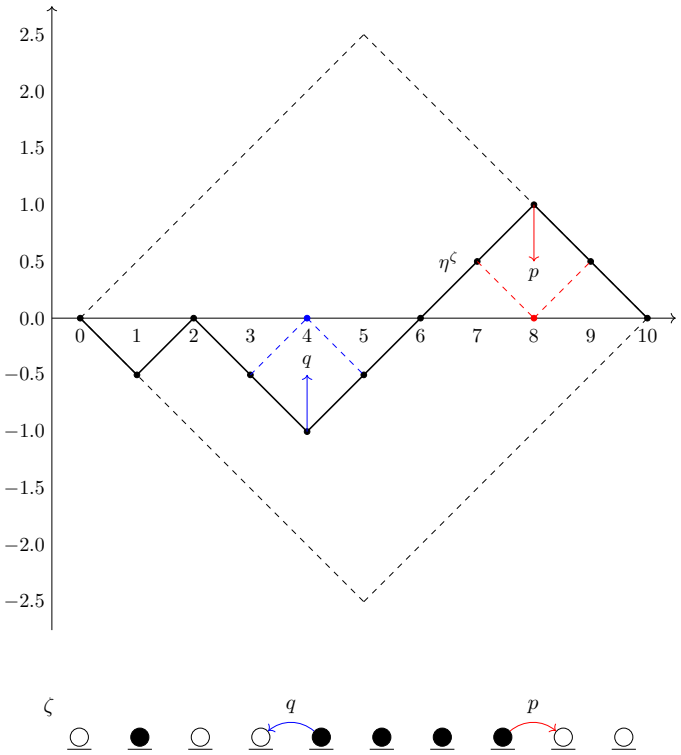


Figure 1.1: The bottom figure gives a configuration $\zeta \in \Omega_{10,5}$ with particles represented by black sites and holes represented by white sites. The solid black path in the top figure gives the corresponding height function η^ζ from (1.3.4); we draw a line between adjacent points in the lattice path. The dynamics of the SEP correspond to flipping corners, i.e. local minima and maxima, in the lattice path. In particular the blue and red jumps in the SEP configuration correspond to the blue and red corner flips in the lattice path respectively. The lower and upper dashed paths in the top figure give the minimal and maximal lattice paths, in the sense that all lattice paths η^ζ for $\zeta \in \Omega_{10,5}$ lie between these paths. The lower path arises from the configuration with $\vee(x) := \mathbb{1}(x \geq 6)$ for all $x \in [10]$, and the upper path arises from the configuration $\wedge(x) = \mathbb{1}(x \leq 5)$ for all $x \in [10]$.

The mapping to lattice paths is injective, and therefore bijective onto the image of configurations in $\Omega_{N,k}$. It is therefore sufficient to consider the dynamics on lattice paths. With an abuse

of notation we denote the set of lattice paths by $\Omega_{N,k}$. Let

$$I = \Big\{(x,z) : x \in [N-1], \\ z \in \Big(\max(0, x-N+k) - \frac{xk}{N} + \mathbb{Z} \cap \big[0, \min(x,k) - \max(0, x-N+k)\big]\Big)\Big\},$$

the set of possible points in the lattice paths in $\Omega_{N,k}$.

Let $\mathcal{T}^{\uparrow}$ and $\mathcal{T}^{\downarrow}$ be two independent Poisson processes indexed by I of rate q and p respectively. In particular $\mathcal{T}^{\uparrow}_{(x,z)}$ is a Poisson clock of rate q, and $\mathcal{T}^{\downarrow}_{(x,z)}$ is a Poisson clock of rate p, for all $(x,z) \in I$.

We construct a trajectory $(\eta_t)_{t\geq 0}$ with initial condition η, for all $\eta \in \Omega_{N,k}$ simultaneously as follows:

If $\mathcal{T}^{\uparrow}_{(x,z)}$ rings at time t then:

- if $\eta_{t-}(x) = z$, and η_{t-} has a local minimum at x then $\eta_t(x) = z+1$, and all other coordinates remain unchanged;
- otherwise no changes are made to the lattice path.

If $\mathcal{T}^{\downarrow}_{(x,z)}$ rings at time t then:

- if $\eta_{t-}(x) = z$, and η_{t-} has a local maxmimum at x then $\eta_t(x) = z-1$, and all other coordinates remain unchanged;
- otherwise no changes are made to the lattice path.

This construction appears in Lacoin, Simenhaus and Toninelli [66, Figure 3] as a construction for the graphical construction of the zero–temperature Ising model in a $k \times (N-k)$ rectangle with mixed boundary conditions. It is straightforward to see that the dynamics obtained from this construction give the dynamics of the SEP, and we let $\mathbb{P}$ denote the law of the construction. The graphical construction gives a *grand coupling* of the SEP, i.e. a coupling of the process started from all initial conditions simultaneously. See, for example, [70, Section 5.4] for more details on grand couplings.

A useful property of the graphical construction $\mathbb{P}$ is that it preserves a partial order on lattice paths. For two lattice paths $\eta, \eta' \in \Omega_{N,k}$, we say that

$$\eta \leq \eta' \quad \text{if } \eta(x) \leq \eta'(x) \text{ for all } 0 \leq x \leq N. \tag{1.3.5}$$

In particular, if $\eta \leq \eta'$ and $t \geq 0$ then

$$\mathbb{P}\big(\eta_t \leq \eta'_t\big) = 1.$$

See [63, Proposition 3.1] for the proof of an analogous result. This idea will be especially important in Chapter 2.

Remark 1.3.1. *The height function description is not necessary to give a grand coupling; there are perfectly nice grand couplings given in terms of the original particle description. See, for example, the 'basic coupling' [76, Chapter VIII] in which Poisson clocks are attached to each site in the particle system. In particular the graphical construction $\mathbb{P}$ is not the same as the basic coupling.*

We will also consider the SEP on the circle $\mathbb{T}_N := \mathbb{Z}/N\mathbb{Z}$. For the dynamics of the SEP on the circle with parameter $p = 1/2$, the number of particles is conserved and we may restrict to the state space

$$\Omega'_{N,k} := \left\{ \xi \in \{0,1\}^{\mathbb{T}_N} \;\Big|\; \sum_{x \in \mathbb{T}_N} \xi(x) = k \right\}. \tag{1.3.6}$$

The generator $\mathcal{L}^{\circ}_{ex}$ of the SEP on the circle $\mathbb{T}_N$ with parameter $p = 1/2$ acts on function $f : \Omega \to \mathbb{R}$ via

$$\mathcal{L}^{\circ}_{ex} f(\zeta) = \sum_{x \in \mathbb{T}_N} \frac{1}{2} [f(\zeta^{x,x+1}) - f(\zeta)] \tag{1.3.7}$$

When $p = 1/2$, a straightforward check of the detailed balance equations analogous to (1.3.3) gives that the stationary measure π' of the SEP on the circle is uniform on the set $\Omega'_{N,k}$. We do not consider the ASEP ($p \neq 1/2$) on the circle; the dynamics are not reversible and there is a drift (although the uniform distribution is still stationary).

1.3.1 The open boundary exclusion process

In recent years there has been some interest in exclusion processes attached to reservoirs which input and extract particles into the system [15, 41, 49, 68, 95]. We consider the open boundary exclusion process (OBEP) on the segment $[N]$, with infinite reservoirs attached to the endpoints of the segment: we think of site 1 as being attached to an infinite reservoir of particles which attempts to input particles at site 1 with rate α, only doing so if there is a hole at site 1, and attempts to remove particles from site 1 with rate γ, only doing so if there is a particle at site 1. Similarly site N is attached to an infinite reservoir which inputs and removes particles at site N with rates δ and β respectively.

Particles on the segment move like particles in the SEP: a particle at site x attempts to jump right at rate $p \in (0,1)$ to site $x+1$, only doing so if $x < N$ and $x+1$ is not occupied by another particle, and particle at site x attempts to jump left at rate $q = 1 - p$ to site $x - 1$, only doing so if $x > 1$ and $x - 1$ is not occupied by another particle. See Figure 1.2.

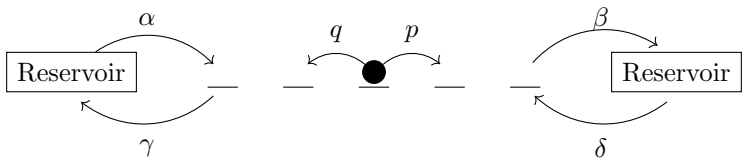


Figure 1.2: Transition rates for the OBEP on Ω_5.

More precisely, the OBEP with parameters $p \in (0,1)$ and $\alpha, \beta, \gamma, \delta \geq 0$, is the Markov process on the state space $\Omega_N := \{0,1\}^N$ with generator $\mathcal{L}_{\text{op}}$ acting on functions $f : \Omega_N \to \mathbb{R}$ via

$$\begin{aligned} \mathcal{L}_{\text{op}} f(\zeta) = \mathcal{L}_{ex} f(\zeta) &+ \alpha\big(1 - \zeta(1)\big)\big[f(\zeta^1) - f(\zeta)\big] + \gamma\zeta(1)\big[f(\zeta^1) - f(\zeta)\big] \\ &+ \delta\big(1 - \zeta(N)\big)\big[f(\zeta^N) - f(\zeta)\big] + \beta\zeta(N)\big[f(\zeta^N) - f(\zeta)\big], \end{aligned} \tag{1.3.8}$$

where $\mathcal{L}_{ex}$ is the generator of the SEP in (1.3.1), and $\zeta^x \in \Omega_N$ denotes the configuration obtained by flipping the value at position x in $\zeta \in \Omega_N$.

The OBEP is one of the most basic examples of a non–equilibrium particle system in statistical mechanics, exhibiting a variety of interesting behaviours in spite of its simplicity. Unlike the SEP, the number of particles in the OBEP is not conserved, and the dynamics are not reversible in general. Mixing times of the OBEP are examined by Gantert, Nestoridi and Schmid [41], and by Gonçalves et al. [49]. In a recent work Salez [95] considers the mixing time of a more general exclusion process on networks with reservoirs. We discuss the results of [41, 49] in greater detail in Section 1.6.1.

1.3.2 The facilitated exclusion process

The main object of study in Chapter 2 is the facilitated simple exclusion process (FSEP). The FSEP was originally introduced by Rossi, Pastor–Satorras and Vespignani [92] as a system with active–absorbing phase transition in the presence of a conserved field and, in the context of the study of non–equilibrium phase transitions, belongs to the class of conserved lattice gases. It is a one–dimensional exclusion process with a dynamical constraint that prevents a particle at site x from jumping to site $x \pm 1$ unless there is a particle at site $x \mp 1$.

Gabel, Krapivsky and Redner [39] investigate the steady state current and distribution of cluster sizes of the totally asymmetric FSEP on the circle. On the infinite line, suitable initial conditions give rise to a shock wave or a rarefaction wave with a jump discontinuity at the front. Baik et al. [10] investigate statistics of particle positions in the totally asymmetric FSEP on the integers with step initial condition. The translation invariant stationary states of the totally asymmetric FSEP on the integers have been investigated by Goldstein, Lebowitz and Speer [45, 46]. The translation invariant stationary states are examined for the asymmetric FSEP by Ayyer et al. [47].

Recent progress has been made by Blondel et al. in establishing hydrodynamic limits for the symmetric FSEP on the circle [18, 19]. In particular the macroscopic density profile evolves under a diffusive time scaling and the behaviour in the hydrodynamic limit is determined by a free boundary problem.

We consider the FSEP on both the segment and circle. It was observed in Blondel et al. [18] that the symmetric FSEP on the circle has degenerate jump rates and admits transient states, which the process will eventually leave to reach an *ergodic component* provided the macroscopic density is larger than $1/2$, or one of its absorbing states otherwise. The same observation holds for the FSEP on the segment in both the symmetric and asymmetric cases.

In the dynamics of the FSEP a particle at site x attempts to jump right to site $x+1$ at rate p, only doing so if there's a hole at site $x+1$ and a particle at site $x-1$. Moreover, a particle at site x attempts to jump left to site $x-1$ at rate q, only doing so if there's a hole at site $x-1$ and a particle at site $x+1$. See, for example, Figure 1.3.

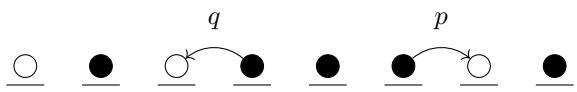


Figure 1.3: Transition rates for the FSEP on the segment. The indicated jumps are the only ones which may occur in the dynamics. This configuration is not ergodic as the leftmost endpoint is not occupied.

The generator $\mathcal{L}$ of FSEP on the segment acts on functions $f : \{0,1\}^N \to \mathbb{R}$ via

$$\begin{aligned}\mathcal{L}f(\xi) = &\sum_{x=2}^{N-1} p\xi(x)\xi(x-1)\big(1-\xi(x+1)\big)\big[f(\xi^{x,x+1}) - f(\xi)\big] \\ &+ \sum_{x=2}^{N-1} q\xi(x)\xi(x+1)\big(1-\xi(x-1)\big)\big[f(\xi^{x,x-1}) - f(\xi)\big].\end{aligned} \tag{1.3.9}$$

As the number of particles is conserved in the FSEP we may restrict to the state space $\Omega_{N,k}$ from (1.3.2). Henceforth we assume that $k > N/2$ so that the FSEP on the segment (and circle) will always reach an ergodic component.

Definition 1.3.2. *We define the ergodic component $\mathcal{E}_{N,k} \subset \Omega_{N,k}$ for the FSEP on the segment by*

$$\mathcal{E}_{N,k} = \big\{\xi \in \Omega_{N,k} : \xi(1)\xi(N) = 1,\, \xi(x) + \xi(x+1) \geq 1 \;\forall\, x \in [N-1]\big\}.$$

In particular the endpoints of the segment are occupied and no two holes are adjacent.



Figure 1.4: An ergodic configuration in $\mathcal{E}_{8,6}$.

The FSEP on the segment with parameter $p \in [1/2, 1)$ is reversible, and its stationary measure $\mu_{N,k}$ is supported on the ergodic component. For $\xi \in \mathcal{E}_{N,k}$ and $i \in [N-k]$ we let x_i denote the position of the i–th leftmost hole on the segment. If $p > 1/2$,

$$p\left(\frac{q}{p}\right)^{-x_i} = q\left(\frac{q}{p}\right)^{-1-x_i},$$

and the detailed balance equations yield

$$\mu_{N,k}(\xi) = \frac{\lambda^{-A(\xi)}}{\sum_{\xi' \in \mathcal{E}_{N,k}} \lambda^{-A(\xi')}}\,, \tag{1.3.10}$$

with $\lambda = q/p$ and

$$A(\xi) = \sum_{i=1}^{N-k} x_i.$$

When $p = 1/2$ a similar check of the detailed balance equations yields that $\mu_{N,k}$ is the uniform measure on $\mathcal{E}_{N,k}$.

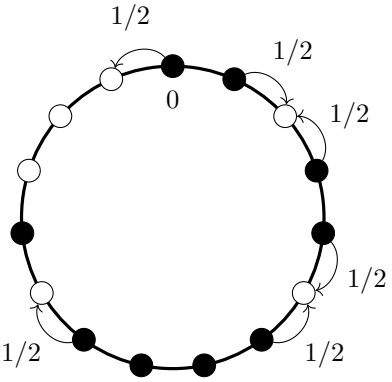


Figure 1.5: Transition rates for the symmetric FSEP on the circle. The indicated jumps are the only ones which may occur in the dynamics. The configuration in $\Omega'_{15,9}$ does not belong to the ergodic component $\mathcal{G}_{N,k}$ as there are two adjacent holes.

The generator $\mathcal{L}^\circ$ of the FSEP on the circle acts on functions $f : \{0,1\}^{\mathbb{T}_N} \to \mathbb{R}$ via

$$\begin{aligned}\mathcal{L}^\circ f(\xi) = &\sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x-1)\big((1-\xi(x+1)\big)\big[f(\xi^{x,x+1}) - f(\xi)\big] \\ &+ \sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x+1)\big((1-\xi(x-1)\big)\big[f(\xi^{x,x-1}) - f(\xi)\big]\,,\end{aligned}$$

and as the number of particles is conserved in the dynamics we may restrict to the state space $\Omega'_{N,k}$ in (1.3.6), i.e. the FSEP configurations with $k > N/2$ particles on a circle of size N.

Definition 1.3.3. *We define the ergodic component $\mathcal{G}_{N,k}$ for the FSEP on the circle by*

$$\mathcal{G}_{N,k} = \big\{\xi \in \Omega'_{N,k} : \forall\, x \in \mathbb{T}_N\,, \xi(x) + \xi(x+1) \geq 1\big\}\,.$$

In particular no two holes are adjacent on the circle in the ergodic component.

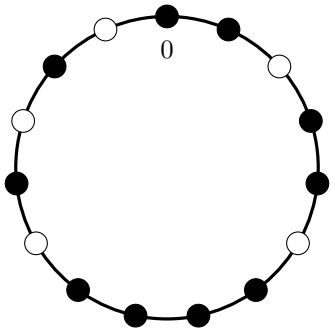


Figure 1.6: An ergodic configuration in $\mathcal{G}_{15,10}$.

Here we only consider the symmetric FSEP on the circle which is reversible on the ergodic component and has an invariant measure $\nu_{N,k}$ which is uniform on $\mathcal{G}_{N,k}$. We do not consider the asymmetric FSEP on the circle, which has a drift and is not reversible.

In Chapter 2 we will examine the mixing time of both the symmetric and asymmetric FSEP on the segment with closed boundaries. We also investigate the mixing time of the symmetric FSEP on the circle.

1.4 The zero–range process

Let $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote the set of nonnegative integers and real numbers respectively, and let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of natural numbers.

The zero–range process, first introduced by Spitzer [101], is a process in which indistinguishable particles move on a finite set S according to the following rules: if a configuration $\omega \in \mathbb{Z}_+^S$ satisfies $\omega(x) \geq 1$ for $x \in S$, then a particle moves from site x to site y at rate $g(\omega(x))p(x,y)$ for some function $g : \mathbb{N} \to \mathbb{R}_+$ and $p(x,y) \geq 0$ for $y \neq x$. Setting $g(0) = 0$, the generator $\mathcal{L}_{zr}$ of the zero–range process acts on functions $f : \mathbb{Z}_+^S \to \mathbb{R}$ via

$$\mathcal{L}_{zr} f(\omega) = \sum_{x \in S} \sum_{y \neq x} g\big(\omega(x)\big) p(x,y) \big[f(\omega^{x,y}) - f(\omega) \big], \tag{1.4.1}$$

with

$$\omega^{x,y}(z) = \begin{cases} \omega(x) - 1 & \text{if } z = x, \\ \omega(y) + 1 & \text{if } z = y, \\ \omega(z) & \text{if } z \notin \{x, y\}, \end{cases} \tag{1.4.2}$$

the configuration obtained by moving a single particle from the stack of particles at x to the stack of particles at y.

The stationary measures of the zero–range process are well–known and can be found, for example, in [5, 34, 57]. Spitzer [101] proved that if S is finite, with $g(k) > 0$ for all $k > 0$, and if $\pi : S \to \mathbb{R}_+$ satisfies $\sum_x \pi(x) p(x,y) = \pi(y)$ and $\pi(x) \leq \liminf g(k)$ for all $x \in S$, then the product measure μ_π on the state space $\mathbb{Z}_+^S$ with marginals

$$\mu_\pi\big(\omega \in \mathbb{Z}_+^S : \omega(x) = k\big) = \begin{cases} \gamma_x \frac{\pi(x)^k}{g(1) \cdots g(k)} & \text{if } k > 0, \\ \gamma_x & \text{otherwise,} \end{cases} \tag{1.4.3}$$

where γ_x is a normalising constant, is stationary for the zero–range process.

The following zero–range process will be particularly useful for our purposes. We consider the zero–range process on the circle $\mathbb{T}_n$ with $g(k) = \mathbb{1}(k \geq 2)$ and $p(x, x+1) = p(x, x-1) = 1/2$. This rate function is degenerate, and if a site has at least one particle it will never become empty. In particular the generator $\mathcal{L}_{\mathrm{ZC}}^n$, acting on functions $f : \mathbb{Z}_+^{\mathbb{T}_n} \to \mathbb{R}$, is given by

$$\mathcal{L}_{\mathrm{ZC}}^n f(\omega) = \sum_{x \in \mathbb{T}_n} \sum_{z = \pm 1} \frac{1}{2} \mathbb{1}(\omega(x) \geq 2) \big[f(\omega^{x, x+z}) - f(\omega) \big]. \tag{1.4.4}$$

We let Q_t^ω denote the law of this zero–range process at time $t \geq 0$ with initial condition ω.

Definition 1.4.1. *For $\omega, \omega' \in \mathbb{Z}_+^{\mathbb{T}_n}$ we say that $\omega \leq \omega'$ if $\omega(i) \leq \omega'(i)$ for all $i \in \mathbb{T}_n$. Moreover, we call an event $E \subset \mathbb{Z}_+^{\mathbb{T}_n}$ increasing if $\omega \in E$ implies $\omega' \in E$ for all $\omega' \geq \omega$.*

A key feature of the zero–range process is the following monotonicity property; see, for example, [57][Section 2.5]. Let $(\omega_t)_{t \geq 0}$ and $(\omega'_t)_{t \geq 0}$ be two trajectories driven by the generator

$\mathcal{L}^n_{\mathrm{ZC}}$, with initial conditions ω and ω' respectively. If $\omega \leq \omega'$ then there exists a coupling such that $\omega_t \leq \omega'_t$ for all $t \geq 0$. In particular given an increasing event $E \subset \mathbb{Z}_+^{\mathbb{T}_n}$ we have that

$$Q_t^{\omega}(E) \leq Q_t^{\omega'}(E). \tag{1.4.5}$$

We also consider the zero–range process on the segment $[n]$ with $g(k) = \mathbb{1}(k \geq 2)$ and

$$p(x,y) = \begin{cases} p & \text{if } y = x+1, \\ q & \text{if } y = x-1. \end{cases}$$

The generator $\mathcal{L}^n_{ZS}$ acts on functions $f : \mathbb{Z}_+^{[n]} \to \mathbb{R}$ via

$$\mathcal{L}^n_{\mathrm{ZS}} f(\omega) = \sum_{x=1}^{n-1} p\mathbb{1}\big(\omega(x) \geq 2\big)\big[f(\omega^{x,x+1}) - f(\omega)\big] + \sum_{x=2}^{n} q\mathbb{1}\big(\omega(x) \geq 2\big)\big[f(\omega^{x,x-1}) - f(\omega)\big]. \tag{1.4.6}$$

1.4.1 Coupling with the FSEP

Throughout Chapter 2 it will often be useful to couple the zero–range process with the FSEP. We now give an example of a coupling between the zero–range process with generator (1.4.4) and the FSEP on the circle. This coupling was introduced by Basu and Mahanty [12], and in the context of the facilitated exclusion process was exploited by Blondel et al. [18, 19].

We map the initial configuration $\xi \in \Omega'_{N,k}$ of the FSEP to a configuration $\omega \in \mathbb{Z}_+^{\mathbb{T}_{N-k}}$ of the zero–range process as follows. We look for the first empty site strictly to the left of site 0 and label it e_0. Moving clockwise around the circle, we label all empty sites in ξ by $e_0, e_1, \ldots, e_{N-k-1}$, and define $\omega(e_i)$ to be the number of particles between the e_i–th and e_{i+1}–th empty site. Here the site e_{N-k} is identified with site e_0. For any $\xi \in \Omega_{N,k}$ we denote the corresponding zero–range configuration by $\Pi_{N,k}(\xi)$. See Figure 1.7 for an example.

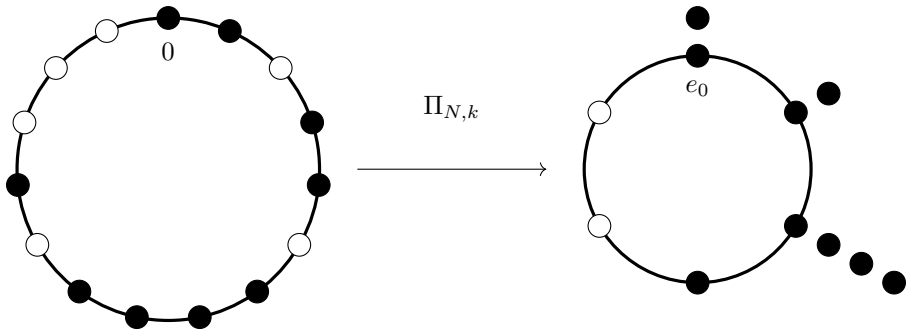


Figure 1.7: The mapping $\Pi_{N,k}$ from an FSEP configuration in $\Omega'_{15,9}$ to a zero-range configuration in $\mathbb{Z}_+^{\mathbb{T}_6}$. The FSEP configuration on the left is not in the ergodic component $\mathcal{G}_{15,9}$ since it has two adjacent holes, and therefore the zero–range configuration on the right has an empty site.

We may couple the FSEP $(\xi_t)_{t\geq 0}$ with the zero–range process with $\omega_0 = \Pi_{N,k}(\xi_0)$, so that whenever a particle jumps in ξ_t a particle in the corresponding pile in ω_t jumps in the same direction. In particular the dynamical constraint means that a particle may jump to an empty site in the FSEP if and only if the corresponding pile in the zero-range has at least two particles.

We note that the mapping $\Pi_{N,k}$ is not bijective, and it is not the case that $\omega_t = \Pi_{N,k}(\xi_t)$ for all $t \geq 0$, although it does hold that ω_t is a rotation $\Pi_{N,k}(\xi_t)$.

The image set of the ergodic component $\mathcal{G}_{N,k}$ under the mapping $\Pi_{N,k}$ is given by

$$\Pi_{N,k}\big(\mathcal{G}_{N,k}\big) = \Big\{\omega \in \mathbb{Z}_+^{\mathbb{T}_{N-k}} : \omega(e_i) \geq 1 \text{ for all } i, \sum_i \omega(e_i) = k\Big\}. \tag{1.4.7}$$

Remark 1.4.2. *We define the ergodic component for the zero–range process by*

$$G_{N-k} := \Big\{\omega \in \mathbb{Z}_+^{\mathbb{T}_{N-k}} : \omega(e_i) \geq 1 \textit{ for all } i\Big\}. \tag{1.4.8}$$

We note that G_{N-k} is increasing in the sense of Definition 1.4.1.

As remarked in [18, Section 3] the FSEP on the circle $(\xi_t)_{t\geq 0}$ is in the ergodic component $\mathcal{G}_{N,k}$ if and only if the zero–range process $(\omega_t)_{t\geq 0}$ with initial condition $\Pi_{N,k}(\xi_0)$ is in G_{N-k}. In particular

$$P_t^{\xi}(\mathcal{G}_{N,k}) = Q_t^{\Pi_{N,k}(\xi)}(G_{N-k}), \tag{1.4.9}$$

for all $\xi \in \Omega'_{N,k}$ and $t \geq 0$.

Similarly we may couple the FSEP on the segment $[N]$ to the zero–range process on the segment $[N-k+1]$. As before, we map initial conditions $\xi \in \Omega_{N,k}$ to a zero–range configuration $\omega \in \mathbb{Z}_+^{[N-k+1]}$. Labelling holes from left to right by $e_1, \ldots, e_{N-k}$, we set $\omega(e_1)$ to be the number of particles between the left boundary and the first hole, $\omega(e_i)$ to be the number of particles between the e_{i-1}–th and e_i–th hole for $i = 2, \ldots, N-k$ and $\omega(e_{N-k+1})$ to be the number of particles between the e_{N-k}–th hole and the right boundary. See Figure 1.8 for an example. For any $\xi \in \Omega_{N,k}$ we denote the corresponding zero–range configuration by $\widetilde{\Pi}_{N,k}$. We note that the mapping $\widetilde{\Pi}_{N,k}$ is bijective.

The FSEP on the segment $(\xi_t)_{t\geq 0}$ may be coupled with the zero–range process $\omega_0 = \widetilde{\Pi}_{N,k}(\xi_0)$ so that whenever a particle in ξ_t jumps a particle in the corresponding pile in ω_t jumps in the same direction.

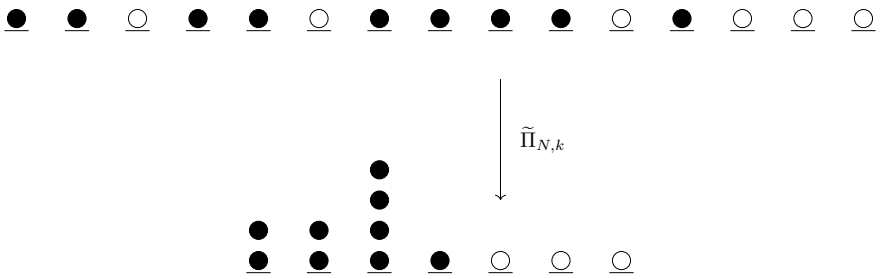


Figure 1.8: The bijective mapping $\widetilde{\Pi}_{N,k}$ for $N = 15$ and $k = 9$. The top figure is an FSEP configuration in $\Omega_{15,9}$ and the bottom figure is a zero–range configuration in $\mathbb{Z}_+^{[7]}$ with 9 particles.

1.5 Mixing times

In the classical theory of Markov chains, we fix a Markov chain and try to estimate the rate of convergence to the stationary distribution as time goes to infinity. For mixing times, we prescribe

some target distance from the stationary distribution of a finite Markov chain, and examine the time taken for the chain to reach this target distance as the size of the state space grows.

Let μ and ν be two probability distributions on some finite state space Ω, then the total variation distance between μ and ν on Ω is defined by

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)|, \tag{1.5.1}$$

the maximum difference between the probabilities assigned to a single event by the probability distributions μ and ν. There are many equivalent characterisations of the total variation distance: by [70, Proposition 4.2]

$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \sum_{\zeta \in \Omega} |\mu(\zeta) - \nu(\zeta)| = \sum_{\substack{\zeta \in \Omega \\ \mu(\zeta) \geq \nu(\zeta)}} [\mu(\zeta) - \nu(\zeta)], \tag{1.5.2}$$

by [70, Proposition 4.5]

$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \sup \left\{ \sum_{\zeta \in \Omega} f(\zeta)\mu(\zeta) - \sum_{\zeta \in \Omega} f(\zeta)\nu(\zeta) : f : \Omega \to \mathbb{R}, \max_{\zeta \in \Omega} |f(\zeta)| \leq 1 \right\}, \tag{1.5.3}$$

and by [70, Proposition 4.7]

$$\|\mu - \nu\|_{\mathrm{TV}} = \inf \big\{ P(X \neq Y) : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu \big\}. \tag{1.5.4}$$

There exists an optimal coupling which obtains the infimum in (1.5.4). The total variation satisfies the triangle inequality [70, Remark 4.4]. If π is a probability distribution on Ω then

$$\|\mu - \nu\|_{\mathrm{TV}} \leq \|\mu - \pi\|_{\mathrm{TV}} + \|\pi - \nu\|_{\mathrm{TV}}.$$

If π is the unique stationary distribution for an irreducible Markov semigroup $P_t = e^{t\mathcal{L}}$, we define the distance

$$d(t) = \max_{\zeta \in \Omega} \big\| P_t^{\zeta}(\cdot) - \pi(\cdot) \big\|_{\mathrm{TV}}. \tag{1.5.5}$$

It may be shown that $d(t)$ is nonincreasing in t; that is advancing the Markov chain can only move it closer to stationarity. See, for example, [70, Exercise 20.2]. Let $\mathcal{M}$ denote the collection of all probability distributions on Ω. Then another characterisation of the distance $d(t)$ is given by

$$d(t) = \sup_{\pi' \in \mathcal{M}} \big\| P_t^{\pi'}(\cdot) - \pi(\cdot) \big\|_{\mathrm{TV}}. \tag{1.5.6}$$

It is often possible to bound $\|P_t^{\zeta}(\cdot) - P_t^{\zeta'}(\cdot)\|_{\mathrm{TV}}$ uniformly over all pairs $(\zeta, \zeta') \in \Omega$, so we define the distance

$$\bar{d}(t) := \max_{\zeta, \zeta' \in \Omega} \big\| P_t^{\zeta}(\cdot) - P_t^{\zeta'}(\cdot) \big\|_{\mathrm{TV}}.$$

The relationship between $d(t)$ and $\bar{d}(t)$ is as follows:

$$d(t) \leq \bar{d}(t) \leq 2d(t). \tag{1.5.7}$$

See, for example, [70, Lemma 4.10]. Other distances between distributions are also useful. For two real–valued functions $f, g : \Omega \to \mathbb{R}$ we define the scalar product

$$\langle f, g \rangle_{\pi} = \sum_{\zeta \in \Omega} f(\zeta) g(\zeta) \pi(\zeta). \tag{1.5.8}$$

For $1 \le p < \infty$, we define the $\ell^p(\pi)$ norm of a function f to be

$$\|f\|_{p,\pi} = \left(\sum_{\zeta \in \Omega} |f(\zeta)|^p \pi(\zeta) \right)^{1/p}, \tag{1.5.9}$$

and

$$\|f\|_{\infty,\pi} = \max_{\zeta \in \Omega} |f(\zeta)|. \tag{1.5.10}$$

The ℓ^p–distance $d^{(p)}$ is defined as

$$d^{(p)}(t) := \max_{\zeta \in \Omega} \left\| \frac{P_t^\zeta(\cdot)}{\pi(\cdot)} - 1 \right\|_{p,\pi}.$$

By (1.5.2) $d^{(1)}(t) = 2d(t)$, and for any $f : \Omega \to \mathbb{R}$ the function $p \mapsto \|f\|_p$ is nondecreasing in $p \ge 1$. In particular

$$2d(t) = d^{(1)}(t) \le d^{(2)}(t) \le d^{(\infty)}(t).$$

Moreover, the distance $d^{(p)}$ is submultiplicative and

$$d^{(p)}(s+t) \le d^{(p)}(s) d^{(p)}(t),$$

see, for example, [70, Lemma 4.18]. Similarly the distance $\bar{d}$ is submultiplicative with

$$\bar{d}(s+t) \le \bar{d}(s)\bar{d}(t).$$

If c and t are positive integers, it follows from (1.5.7) that

$$d(ct) \le \bar{d}(ct) \le \bar{d}(t)^c. \tag{1.5.11}$$

For $\epsilon \in (0,1)$ the ϵ–mixing time $T_{\text{mix}}(\epsilon)$ is defined by

$$T_{\text{mix}}(\epsilon) = \inf\{t \ge 0 : d(t) \le \epsilon\}. \tag{1.5.12}$$

In particular the ϵ–mixing time is the first time that the distance $d(t)$ passes below the ϵ threshold. We set

$$T_{\text{mix}} := T_{\text{mix}}(1/4). \tag{1.5.13}$$

By (1.5.7) and (1.5.11) we note that

$$d\big(cT_{\text{mix}}(\epsilon)\big) \le \bar{d}\big(T_{\text{mix}}(\epsilon)\big)^c \le (2\epsilon)^c, \tag{1.5.14}$$

and choosing $\epsilon = 1/4$ yields that

$$d\big(cT_{\text{mix}}\big) \le 2^{-c},$$

and

$$T_{\text{mix}}(\epsilon) \le \lceil \log_2 \epsilon^{-1} \rceil T_{\text{mix}}. \tag{1.5.15}$$

The choice of $\epsilon = 1/4$ in (1.5.13) is somewhat arbitrary, although a value of ϵ less than $1/2$ is required to make the upper bound in (1.5.14) meaningful and achieve an inequality of the form (1.5.15).

Suppose we have a sequence of Markov chains indexed by $n = 1, 2, \ldots$ and the ϵ–mixing time of the nth chain is denoted by $T^n_{\text{mix}}(\epsilon)$. We say that the sequence of chain has *cutoff* if, for all $\epsilon \in (0, 1)$,

$$\lim_{n\to\infty} \frac{T^n_{\text{mix}}(\epsilon)}{T^n_{\text{mix}}(1-\epsilon)} = 1. \tag{1.5.16}$$

In particular cutoff characterises an 'abrupt' convergence to equilibrium for the sequence of Markov chains. More precisely, if $d_n(t)$ is the distance in (1.5.5) for the nth chain, then the sequence of chains has cutoff with a *window* of size $O(w_n)$ if

$$\lim_{n\to\infty} \frac{w_n}{T^n_{\text{mix}}} = 0,$$

and

$$\lim_{\alpha\to-\infty} \liminf_{n\to\infty} d_n\Big(T^n_{\text{mix}} + \alpha w_n\Big) = 1,$$

$$\lim_{\alpha\to\infty} \limsup_{n\to\infty} d_n\Big(T^n_{\text{mix}} + \alpha w_n\Big) = 0.$$

In particular the total variation distance from equilibrium drops from $1-\epsilon$ to ϵ in a window of time which is smaller than the order of the mixing time. We say that the sequence of chains satisfies *pre–cutoff* if

$$\sup_{\epsilon\in(0,1/2)} \limsup_{n\to\infty} \frac{T^n_{\text{mix}}(\epsilon)}{T^n_{\text{mix}}(1-\epsilon)} < \infty. \tag{1.5.17}$$

Pre–cutoff is a weaker condition than cutoff in the sense that all sequences of chains with cutoff have pre–cutoff, but there are examples of chains with pre–cutoff which do not have cutoff. See, for example, [70][Figure 18.2] for a sequence of chains with pre–cutoff but not cutoff. Recently Schmid and Sly [98, Theorem 1.1] showed that the TASEP on the circle with size N with k particles mixes on the order $N^2 \min(k, N-k)^{-1/2}$ and satisfies pre–cutoff but not cutoff.

1.5.1 The spectral gap and log–Sobolev constant

In this section we introduce the spectral gap and log–Sobolev constant, largely following the lecture notes of Saloff–Coste [96]. We consider irreducible reversible continuous–time Markov chains on a finite state space Ω, with semigroup $P_t = e^{t\mathcal{L}}$ and unique stationary measure π. For a real–valued function $f : \Omega \to \mathbb{R}$ we use the notation

$$\pi(f) = \sum_{\zeta\in\Omega} f(\zeta)\pi(\zeta) \text{ and } \text{Var}_\pi(f) = \sum_{\zeta\in\Omega} |f(\zeta) - \pi(f)|^2 \pi(\zeta).$$

The Dirichlet form associated to the semigroup $P_t = e^{t\mathcal{L}}$ is defined by

$$\mathcal{D}(f, g) = \langle -\mathcal{L}f, g\rangle_\pi, \tag{1.5.18}$$

and we write $\mathcal{D}(f) = \mathcal{D}(f, f)$.

Lemma 1.5.1 ([96, Lemma 2.1.2])**.** *The Dirichlet form $\mathcal{D}(f)$ satisfies*

$$\mathcal{D}(f) = \frac{1}{2} \sum_{\zeta,\zeta'\in\Omega} Q(\zeta, \zeta')\pi(\zeta)\big|f(\zeta) - f(\zeta')\big|^2, \tag{1.5.19}$$

and

$$\frac{\partial}{\partial t}\big\|P_t f\big\|^2_{2,\pi} = -2\mathcal{D}(P_t f).$$

The characterisation of the Dirichlet form in (1.5.19) is straightforward to bound below by removing transitions between sets.

Definition 1.5.2. *The spectral gap $\lambda = \lambda(\mathcal{L})$ is defined by*

$$\lambda = \inf\left\{ \frac{\mathcal{D}(f)}{\mathrm{Var}_\pi(f)}; \mathrm{Var}_\pi(f) \neq 0 \right\},$$

or equivalently

$$\lambda = \inf\big\{ \mathcal{D}(f) : \|f\|_{2,\pi}^2 = 1, \pi(f) = 0 \big\}.$$

In particular λ is the smallest non–zero eigenvalue of $-\mathcal{L}$. The spectral gap may be used to bound the distance from stationarity. The *relaxation time* is defined to be the inverse of the spectral gap.

Lemma 1.5.3 ([96, Lemma 2.1.4])**.** *The generator $\mathcal{L}$ with spectral gap λ and associated semigroup P_t satisfies*

$$\big\|P_t f - \pi(f)\big\|_2^2 \leq \exp(-2\lambda t)\mathrm{Var}_\pi(f),$$

for all $f : \Omega \to \mathbb{R}$ with $\|f\|_{2,\pi} < \infty$.

Moreover, if $p_t^\zeta(\cdot) = P_t^\zeta(\cdot)/\pi(\cdot)$ then the following result holds.

Lemma 1.5.4 ([96, Theorem 2.1.7])**.** *For all $1 \leq p \leq \infty$*

$$\lim_{t\to\infty} \frac{-1}{t} \log\Big(\max_{\zeta\in\Omega} \|p_t^\zeta - 1\|_{p,\pi} \Big) = \lambda.$$

Furthermore, we define the ℓ^p–mixing time with threshold $1/e$ by

$$T_p = T_p(1/e) := \inf\Big\{ t \geq 0 : \max_{\zeta\in\Omega} \big\|p_t^\zeta - 1\big\|_{p,\pi} \leq 1/e \Big\}, \tag{1.5.20}$$

and we let

$$\pi_* = \min_{\zeta\in\Omega} \big(\pi(\zeta)\big). \tag{1.5.21}$$

For $1 \leq p \leq 2$,

$$\frac{1}{\lambda} \leq T_p \leq \frac{1}{2\lambda}\left(2 + \log\frac{1}{\pi_*} \right),$$

and, for $2 < p \leq \infty$,

$$\frac{1}{\lambda} \leq T_p \leq \frac{1}{\lambda}\left(1 + \log\frac{1}{\pi_*} \right).$$

We define the entropy w.r.t π by

$$\mathrm{Ent}_\pi(f) = \sum_{\zeta\in\Omega} f(\zeta)\pi(\zeta) \log\Big(\frac{f(\zeta)}{\pi(f)}\Big) = \pi\left(f \log\Big(\frac{f}{\pi(f)}\Big) \right). \tag{1.5.22}$$

By an application of Jensen's inequality to the convex function $\phi(t) = t\log t$, the entropy $\mathrm{Ent}_\pi(f) \geq 0$. Furthermore $\mathrm{Ent}_\pi(f) = 0$ if and only if f is a constant function.

Definition 1.5.5. *The log–Sobolev constant $\rho = \rho(\mathcal{L})$ is defined by*

$$\rho = \inf\left\{ \frac{\mathcal{D}(f)}{\mathrm{Ent}_\pi(f^2)} : \mathrm{Ent}_\pi(f^2) \neq 0 \right\}.$$

It follows from Definition 1.5.5 that ρ is the largest constant c such that

$$c\,\mathrm{Ent}_\pi(f^2) \leq \mathcal{D}(f),$$

holds for all f. Letting $|f|$ be the function $|f|(\zeta) = |f(\zeta)|$, we observe that

$$\mathrm{Ent}_\pi(f^2) = \mathrm{Ent}_\pi(|f|^2) \text{ and } \mathcal{D}(|f|) \leq \mathcal{D}(f),$$

and we may take f to be nonnegative in Definition 1.5.5.

The log–Sobolev constant may be used to bound the ℓ^p–mixing time in (1.5.20) as follows:

Lemma 1.5.6 ([96, Corollary 2.2.7])**.** *For $1 \leq p \leq 2$, it holds that*

$$\frac{1}{2\rho} \leq T_p \leq \frac{1}{4\rho}\left(4 + \log^+ \log \frac{1}{\pi_*}\right), \tag{1.5.23}$$

with $\log^+ t = \max(0, \log t)$ and π_* as in (1.5.21)*. For $2 < p \leq \infty$ it holds that*

$$\frac{1}{2\rho} \leq T_p \leq \frac{1}{2\rho}\left(3 + \log^+ \log \frac{1}{\pi_*}\right).$$

The upper bound in (1.5.23) will be particularly important for our purposes. Since the ℓ^p–distance is submultiplicative we have that

$$d^{(2)}\big(kT_2\big) \leq e^{-k},$$

and it follows that

$$T_2(\epsilon) \leq \lceil \log \epsilon^{-1} \rceil T_2,$$

for all $\epsilon \in (0,1)$. Moreover, $2d(t) \leq d^{(2)}(t)$ and the ℓ^2–mixing time is an upper bound for the total variation mixing time. Therefore,

$$T_{\mathrm{mix}}(\epsilon) \leq \frac{\lceil \log \epsilon^{-1} \rceil}{4\rho}\left(4 + \log^+ \log \frac{1}{\pi_*}\right). \tag{1.5.24}$$

For our purposes, it will be beneficial to relate the entropy to the conditional entropy, which we define as follows:

Definition 1.5.7. *Let $\mathcal{P}(\Omega)$ denote the power set of Ω, and let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two sub–σ–algebras of $\mathcal{P}(\Omega)$. If $\pi_i(f) := \pi(f \mid \mathcal{F}_i)$ for $i \in [2]$ and f is a nonnegative function, then we define the conditional entropy by*

$$\mathrm{Ent}_{\pi_i}(f) = \pi_i\left(f \log\left(\frac{f}{\pi_i(f)}\right)\right). \tag{1.5.25}$$

The following result is a corollary of a result of Cesi [23] and gives a quasi–factorisation of the entropy. In particular if the two sub–σ–algebras are 'close' to being independent, the entropy is bounded by the average of the conditional entropies, up to some constant.

Proposition 1.5.8 ([23, Proposition 2.1]). *Let*

$$\vartheta(\epsilon) := \frac{84\epsilon}{(1-\epsilon)^2}.$$

If, for some $\epsilon \in [0,1)$,

$$\big\|\pi_2(g) - \pi(g)\big\|_{\infty,\pi} \leq \epsilon \big\|g\|_{1,\pi}, \tag{1.5.26}$$

for all real–valued $\mathcal{F}_1$*–measurable functions* g*, then*

$$\mathrm{Ent}_\pi(f^2) \leq \pi\big[\mathrm{Ent}_{\pi_1}(f^2) + \mathrm{Ent}_{\pi_2}(f^2)\big] + \vartheta(\epsilon)\mathrm{Ent}_\pi(f^2). \tag{1.5.27}$$

We note that Proposition 1.5.8 is the restriction of [23, Proposition 2.1] to finite state spaces, and the statement of [23, Proposition 2.1] is far more general.

1.6 Mixing times for exclusion processes

The mixing time for the simple exclusion process has been examined extensively across a large number of works. See, for example, [13, 36, 52, 60, 64, 62, 63, 65, 87, 89, 98, 106, 69, 107]. In particular sharp estimates on the mixing time and related quantities have been obtained on the complete graph [65, 69], the discrete tori $(\mathbb{Z}/L\mathbb{Z})^d$ [87], the segment [60, 63], the circle [64, 62] and several random graphs [89].

We now highlight and discuss a collection of results for the SEP on the segment and circle. These results will be key to many arguments in Chapter 2. Let $T^{N,k}_{\mathrm{seg}}(\epsilon)$ denote the ϵ–mixing time of the SEP with parameter $p \geq 1/2$ on $\Omega_{N,k}$, then the following result of Lacoin gives cutoff for the SSEP ($p = 1/2$) on the segment.

Theorem 1.1 ([63, Theorem 2.4]). *For any* $\epsilon \in (0,1)$*, given a sequence* $k(N)$ *which is such that* k *and* $N-k$ *tend to infinity we have the following asymptotics for the mixing time:*

$$\lim_{N\to\infty} \frac{\pi^2 T^{N,k}_{\mathrm{seg}}(\epsilon)}{N^2 \log\min(k, N-k)} = 1.$$

Recalling the graphical construction $\mathbb{P}$ in Section 1.3, a crucial ingredient in the proof of Theorem 1.1 is an upper bound on the coalescence time of the minimal and maximal lattice paths on $\Omega_{N,k}$ under $\mathbb{P}$. The SEP configurations in $\{0,1\}^{[N]}$ that give rise to the minimal and maximal lattice paths are given by

$$\begin{aligned} \vee(x) &= \mathbb{1}(x \geq N-k+1), \\ \wedge(x) &= \mathbb{1}(x \leq k), \end{aligned} \tag{1.6.1}$$

respectively. This notation is chosen to represent the shape of the lattice path; see, for example, Figure 1.1. The following proposition bounds the coalescence time of the minimal and maximal lattice paths under the coupling $\mathbb{P}$.

Proposition 1.6.1 ([63][Proposition 8.1]). *Given $\delta > 0$, set*

$$t_1 = \frac{N^2(1+\delta)}{\pi^2} \log \min(k, N-k).$$

Then for any $\epsilon \in (0,1)$ we have

$$\mathbb{P}\big(\eta_{t_1}^{\vee} \neq \eta_{t_1}^{\wedge}\big) \leq \epsilon,$$

for all N sufficiently large, depending on ϵ.

Cutoff was shown for the ASEP on the segment by Labbé and Lacoin in the regime where $k/N \to \alpha$ for some fixed $\alpha \in [0,1]$. We take $1 \leq k \leq N-1$ to avoid trivial state spaces, and use the notation

$$\lim_{\substack{N\to\infty \\ k/N\to\alpha}} f(k,N) = L \tag{1.6.2}$$

to express that the limit of the real–valued function $f(k,N)$ is L along all sequences such that k/N tends to α. In particular (1.6.2) denotes

$$\lim_{\epsilon\to 0} \limsup_{N\to\infty} \sup_{\substack{k\in[N-1] \\ |k/N-\alpha|\leq\epsilon}} \big|f(k,N) - L\big| = 0.$$

Theorem 1.2 ([60, Theorem 2]). *For all $p \in (1/2, 1]$, $\alpha \in [0,1]$ and $\epsilon \in (0,1)$ it holds that*

$$\lim_{\substack{N\to\infty \\ k/N\to\alpha}} \frac{T_{\text{seg}}^{N,k}(\epsilon)}{N} = \frac{\big(\sqrt{\alpha} + \sqrt{1-\alpha}\big)}{p-q}.$$

This result is achieved by relating the mixing time to hydrodynamics, however an upper bound on the mixing time may be obtained using a height function argument and the monotonicity in (1.3.5). In particular, if $\lambda = p/q$, $\varrho = (\sqrt{p} - \sqrt{q})^2$ and

$$\gamma_N = 2\sqrt{pq}\Big[1 - \cos\Big(\frac{\pi}{N}\Big)\Big],$$

then

$$\mathbb{P}\big(\eta_t^{\vee} \neq \eta_t^{\wedge}\big) \leq C_\lambda N \lambda^{N/2} \exp\big(-(\gamma_N + \varrho)t\big). \tag{1.6.3}$$

See [60, Section 3.4] for some details of the argument.

Let $T_{\text{cir}}^{N,k}(\epsilon)$ denote the ϵ–mixing time of the SSEP on the state space $\Omega'_{N,k}$. The following result of Lacoin gives cutoff for the SSEP on the circle.

Theorem 1.3 ([64, Theorem 1.1]). *Let $k(N)$ be a sequence such that both $k(N)$ and $N - k(N)$ tend to infinity. For every $\epsilon \in (0,1)$,*

$$\lim_{N\to\infty} \frac{4\pi^2 T_{\text{cir}}^{N,k}(\epsilon)}{N^2 \log \min(k, N-k)} = 1.$$

Moreover, it was shown in [64] that the SSEP on the circle has a cutoff window of size $O(N^2)$.

1.6.1 Mixing times for the OBEP

Recall the OBEP with parameters $(p, \alpha, \beta, \gamma, \delta)$ from Section 1.3.1. Recently, significant progress has been made in characterising the mixing times of the OBEP [41, 49, 95]. For the OBEP on Ω_N with parameters $(p, \alpha, \beta, \gamma, \delta) = (1/2, \rho, 1-\rho, \rho, 1-\rho)$ for some $\rho \in (0,1)$, it was shown in Gonçalves et al. [49] that cutoff occurs on the order $N^2 \log N$, with cutoff window of size $O(N^2)$.

We now state a selection of results due to Gantert, Nestoridi and Schmid [41]. These results will be key to many arguments in Chapter 2. The following theorem is a pre–cutoff result for the symmetric OBEP with general rates at the boundary

Theorem 1.4 ([41][Theorem 1.1]). *For $p = 1/2$ the ϵ–mixing time of the OBEP satisfies*

$$\frac{1}{\pi^2} \leq \liminf_{N\to\infty} \frac{T^N_{\text{mix}}(\epsilon)}{N^2 \log N} \leq \limsup_{N\to\infty} \frac{T^N_{\text{mix}}(\epsilon)}{N^2 \log N} \leq C,$$

for all $\epsilon \in (0,1)$ and some constant $C = C(\alpha, \beta, \gamma, \delta) > 0$.

The following theorem is a cutoff result for the symmetric OBEP with exactly one open boundary, i.e. $\alpha = \gamma = 0$, and $\min(\beta, \delta) > 0$.

Theorem 1.5 ([41, Theorem 1.2]). *For $p = 1/2$, suppose that $\alpha = \gamma = 0$ and $\min(\beta, \delta) > 0$. Then for all $\epsilon \in (0,1)$, the ϵ–mixing time of the OBEP satisfies*

$$\lim_{N\to\infty} \frac{T^N_{\text{mix}}(\epsilon)}{N^2 \log N} = \frac{4}{\pi^2}. \tag{1.6.4}$$

By symmetry (1.6.4) holds when $\beta = \delta = 0$ and $\min(\alpha, \gamma) > 0$.

The following theorem gives cutoff for the asymmetric OBEP with one blocked entry, i.e. $\alpha > 0$ and $\beta = 0$ or $\alpha = 0$ and $\beta > 0$ by particle–hole symmetry.

Theorem 1.6 ([41][Theorem 1.3]). *Suppose that $p > 1/2$ and let $\gamma, \delta \geq 0$ be arbirtrary. If $\alpha > 0$ and $\beta = 0$, then for all $\epsilon \in (0,1)$, the ϵ–mixing time of the OBEP satisfies*

$$\lim_{N\to\infty} \frac{T^N_{\text{mix}}(\epsilon)}{N} = \frac{\big(\max(a,1)+1\big)^2}{(2p-1)\max(a,1)},$$

for some constant $a = a(p, \alpha, \gamma) > 0$ not depending on ϵ. Similarly, for $\alpha = 0$ and $\beta > 0$, we have that

$$\lim_{N\to\infty} \frac{T^N_{\text{mix}}(\epsilon)}{N} = \frac{\big(\max(b,1)+1\big)^2}{(2p-1)\max(b,1)},$$

for some constant $b = b(p, \beta, \delta) > 0$ not depending on ϵ.

The following theorem is for the 'reverse–bias' phase of the OBEP. Intuitively the boundary condition forces particles to move against their natural drift and the systems mixes exponentially slowly.

Theorem 1.7 ([41][Theorem 1.4]). *Suppose $\alpha = \beta = 0$ and $p \in (1/2, 1)$. For all $\epsilon \in (0, 1/2)$ we have that*

$$\lim_{N\to\infty} \frac{\log T^N_{\text{mix}}(\epsilon)}{N} = \log\left(\frac{p}{q}\right),$$

whenever $\min(\gamma, \delta) = 0$ and $\max(\gamma, \delta) > 0$. If $\min(\gamma, \delta) > 0$ then

$$\lim_{N\to\infty} \frac{\log T^N_{\text{mix}}(\epsilon)}{N} = \frac{1}{2}\log\left(\frac{p}{q}\right).$$

1.7 Coupling from the past

In Markov chain Monte Carlo (MCMC) we typically want to understand the properties of some probability distribution. However, the probability distribution may be hard to deal with explicitly, for example the probability distribution may be specified indirectly or the state space may be very complicated. To this end we design a suitable Markov chain whose long–run equilibrium distribution is the target probability distribution. We can then sample from a distribution which is approximately the target distribution by running the Markov chain for enough time that the chain is close to equilibrium. If the mixing time is known, this is a suitable choice for the length of time to run the chain.

One natural question is whether MCMC algorithms can be modified to output exact samples from the target probability distribution. Propp and Wilson [91] gave a modification of MCMC known as the *Coupling from the past* (CFTP) algorithm or Propp–Wilson algorithm which gives exact samples from the equilibrium distribution. One of the main features of the CFTP algorithm that distinguishes it from ordinary MCMC is that it involves coupling several copies of a Markov chain with different initial conditions. In addition, the Markov chain is not run forwards in time from time 0, but from some time in the past until time 0.

We now give a brief summary of the CFTP algorithm following the notes of Häggström [50] and Kendall [56].

We assume that we have access to some random number generator which produces a sequence $(U_t)_{t\in\mathbb{Z}}$ of integer–valued i.i.d. random variables. We refer to U_t as a t–update.

Definition 1.7.1. *Let S be a finite state space and P be an irreducible and aperiodic transition matrix on S. An update function $\phi : S \times \mathbb{Z} \to S$ is a function such that for $s, s' \in S$ fixed*

$$\sum_{\substack{u\in\mathbb{Z}\\ \phi(s,u)=s'}} \mathbb{P}\big(U_0 = u\big) = P(s, s').$$

Setting $Y_0 = s \in S$ we may simulate the Markov chain $(Y_t)_{t\geq 0}$ with transition matrix P by setting $Y_t = \phi(Y_{t-1}, U_t)$ for all $t \geq 1$.

Let $(N_1, N_2, \ldots)$ be an increasing sequence of positive integers, then the CFTP algorithm runs as follows:

(1) Set $m = 1$.

(2) For each $s \in S$, we simulate the Markov chain starting at time $-N_m$ in state s until time 0 using the update function ϕ and random numbers $U_{-N_m+1}, U_{-N_m+2}, \ldots, U_{-1}, U_0$. We use the same random numbers for each of the $|S|$ chains.

(3) If all $|S|$ chains in Step 2 are in the same state $\tilde{s}$ at time 0 then the algorithm outputs $\tilde{s}$. Otherwise we increase m by 1 and repeat Step 2.

If μ is the unique stationary distribution for P, then [91][Theorem 1] states that the output of the CFTP algorithm is a sample from μ. It is essential that, at the m–th time we perform Step 2, we re–use the updates $U_{-N_{m-1}+1}, \ldots, U_0$. The algorithm gives a grand coupling, i.e. a coupling of the chain started from all initial conditions, by applying the same collection of t–updates. The CFTP algorithm is not guaranteed to terminate, and Step 2 may be repeated infinitely.

Remark 1.7.2. *A common and sensible choice for the increasing sequence* $(N_n)_{n=1}^{\infty}$ *is* $N_n = 2^{n-1}$*. In particular, if*

$$T = \min\big\{t \geq 0 : \text{ the chains starting at time } -t \text{ coalesce by time } 0\big\}, \tag{1.7.1}$$

then using the CFTP algorithm with $(N_1, N_2, \ldots) = (1, 2, 4, \ldots)$ *gives that the total number of applications of the update function in the CFTP algorithm is bounded by* $4T$*.*

1.7.1 Some examples

The time T in (1.7.1) has the same distribution as the forwards coalescence time of all copies of the chain; it is tempting to discard the sequence of starting times going further and further back into the past, and perform a 'Coupling into the future' algorithm where one simply runs every copy of the chain forwards in time until coalescence, outputting the common value. The following example shows that this approach gives biased samples in general.

Example 1.7.3. *Let* $(Y_0, Y_1, \ldots)$ *be a Markov chain with state space* $S = \{s_1, s_2\}$ *and transition matrix*

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix}, \tag{1.7.2}$$

see Figure 1.9.

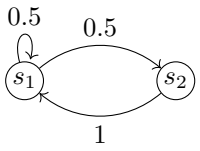


Figure 1.9: The transition graph for the Markov chain with transition matrix in (1.7.2).

The chain is reversible with stationary distribution

$$\mu = \big(\mu(1), \mu(2)\big) = \Big(\frac{2}{3}, \frac{1}{3}\Big).$$

We run two copies of the chain starting at time 0*, with one copy starting at state* s_1 *and the other starting at state* s_2*. The chains will coalesce for the first time at some random time* τ*. At time* $\tau - 1$ *the chains are in different states by definition, therefore exactly one copy of the chain is in state* s_2 *at time* $\tau - 1$*.*

The transition matrix P *tells us that this copy of the chain is in state* s_1 *at time* τ *with probability* 1*, and both chains are in state* s_1 *with probability* 1 *at the coalescence time. Clearly this does not agree with the stationary distribution, and the naive Coupling into the future algorithm does not output samples from the stationary distribution in general.*

It is often very computationally expensive to keep track of all $|S|$ copies of the chain. In some cases this can be simplified; it is shown in [91] that if the chain in question obeys certain monotonicity properties, then all copies of the chain may be sandwiched between the chains started from the minimal and maximal elements. In this case, it is sufficient to only keep track of the chains started from the minimal and maximal elements. A classical chain to which this applies is the Stochastic Ising model.

Of particular importance to us is a technique applied to Markov random fields by Huber [54] and Häggström and Nelander [51]. Let $S = \Sigma^V$ for some finite set Σ and some finite vertex set V. We consider an augmented state space

$$\widetilde{S} = \big(\Sigma \cup \{\omega\}\big)^V. \tag{1.7.3}$$

Here $\zeta \in \widetilde{S}$ is interpreted to represent a set of configurations in S given by

$$\big\{s \in S : s(v) = \zeta(v) \text{ for all } v \in V \text{ such that } \zeta(v) \in \Sigma\big\},$$

i.e. the set of configurations $s \in S$ which agree with ζ at all points where it does not take value ω.

It is sometimes possible to construct an augmented Markov chain on $\widetilde{S}$, started from time $-t$ with initial condition ω^V, whose evolution tracks the evolution of all copies of the original Markov chain on S started from time $-t$. If such a Markov chain exists, and can be updated in a computationally inexpensive way, this is useful as we need only keep track of a single chain.

We now given an example of this approach for the Hardcore model.

Example 1.7.4. *The following description of the Hardcore model may be found in [70][Chapter 3]. Let $G = (V, E)$ be a graph. A hardcore configuration on the graph G is a placement of particles on the vertex set V so that each vertex is occupied by at most one particle and no two particles are adjacent. If $s(v) = 1$ we say that vertex v is occupied by a particle and if $s(v) = 0$ we say that v is vacant, so the set of hardcore configurations $\mathcal{X}$ is given by*

$$\mathcal{X} = \big\{s \in \{0,1\}^V : s(v)s(u) = 0 \text{ if } \{v,u\} \in E\big\}.$$

The Hardcore model with fugacity λ is the probability distribution π on $\{0,1\}^V$ defined by

$$\pi(s) = \begin{cases} \frac{1}{Z(\lambda)} \lambda^{\sum_{v \in V} s(v)} & \text{if } s \in \mathcal{X}, \\ 0 & \text{otherwise}, \end{cases}$$

where

$$Z(\lambda) = \sum_{s \in \mathcal{X}} \lambda^{\sum_{v \in V} s(v)},$$

is the normalisation factor for π. For the Glauber dynamics of the Hardcore model a configuration Y_t is updated to a new configuration Y_{t+1} as follows: first, select a vertex $u \in V$ uniformly at random. If there exists a neighbour u' of u such that $Y_t(u') = 1$, then set $Y_{t+1}(u) := 0$; otherwise, let

$$Y_{t+1}(u) := \begin{cases} 1 & \text{with probability } \lambda/(1+\lambda), \\ 0 & \text{with probability } 1/(1+\lambda). \end{cases}$$

Furthermore, set $Y_{t+1}(v) = Y_t(v)$ for all $v \neq u$.

We consider an augmented chain on the state space

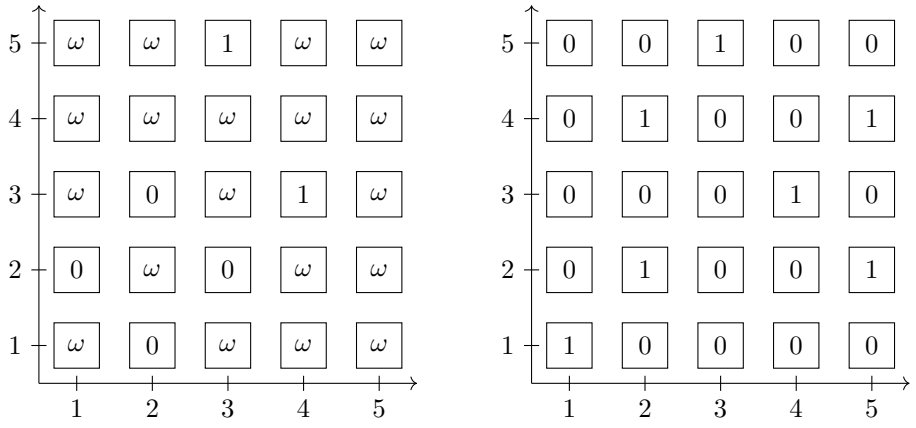
$$\widetilde{\mathcal{X}} = \Big\{\zeta \in \{0,1,\omega\}^V : \big(\zeta(v), \zeta(u)\big) \neq (1,1) \text{ if } \{v,u\} \in E\Big\},$$

where we interpret $\zeta \in \widetilde{\mathcal{X}}$ to represent the set of configurations

$$\big\{s \in \mathcal{X} : s(v) = \zeta(v) \text{ for all } v \in V \text{ such that } \zeta(v) \neq \omega\big\}.$$

For the augmented chain $(Z_t)_{t\geq 0}$ with $Z_0 = \omega^V$, the configuration Z_t is updated to Z_{t+1} as follows: first, select a vertex $u \in V$ uniformly at random. With probability $1/(1+\lambda)$ we place a vacancy at vertex u, i.e. $Z_{t+1}(u) = 0$, and with probability $\lambda/(1+\lambda)$ we attempt to place a particle at vertex u, only doing so if all neighbouring vertices are vacant. In this case set $Z_{t+1}(u) := 1$, and otherwise set $Z_{t+1}(u) := \omega$.

The augmented chain mimics the dynamics of all copies of the Hardcore model, as placing a vacancy at some vertex always succeeds in every copy of the chain, whereas an attempt to place a particle at some vertex will only succeed in those copies which have vacancies at all neighbouring sites. See Figure 1.10 for example.



(a) A configuration ζ on the augmented state space $\widetilde{\mathcal{X}}$, where the graph G is the 5×5 square grid with periodic boundary conditions.

(b) A hardcore configuration $s \in \mathcal{X}$ on the 5×5 square grid with periodic boundary conditions.

Figure 1.10: The hardcore configuration $s \in \mathcal{X}$ in Figure 1.10b belongs to the set of configurations represented by $\zeta \in \widetilde{\mathcal{X}}$ in Figure 1.10a. In the augmented chain dynamics, we may only place a particle at the vertex $(2, 2)$ in ζ, whereas we may place a vacancy at any other vertex.

We say that $f(n) = O(g(n))$ as $n \to \infty$ if $|f|$ is bounded above by g, up to a constant factor, asymptotically. That is,

$$\limsup_{n\to\infty} \frac{|f(n)|}{g(n)} < \infty. \tag{1.7.4}$$

In Chapter 3 we will design some algorithms to sample the cycle structure of a uniform random permutation on N objects with k cycles based on an application of the CFTP algorithm to an exclusion process. The main result of the chapter is that these algorithms have expected run–time $O(k * \text{polylog}N)$.

We present a coupling of the cycle structures of a uniform random permutation on $[N]$ with i cycles, for all $i \in [k]$, whose samples have the following consistency property: the cycle structure with $i + 1$ cycles is obtained from the cycle structure with i cycles by breaking one of the cycles into two parts. In particular we give an explicit construction of a *fragmentation process*; see, for example Pitman's Saint–Flour lecture notes [90]. We then design an algorithm based on an application of CFTP which exactly samples from this coupling in expected run–time

$O(k^2 \text{polylog} N)$. In particular, we will apply CFTP to an exclusion process in which particles are assigned a priority.

1.8 Directed percolation

We consider the directed last–passage percolation model on the nonnegative lattice $\mathbb{Z}_+^d$. For $\boldsymbol{z} \in \mathbb{Z}_+^d$ we write z_i for the i–th coordinate of $\boldsymbol{z}$, and write $\boldsymbol{z} \leq \boldsymbol{z}'$ if $z_i \leq z_i'$ for $i = 1, \ldots, d$. We use the norm $\|\boldsymbol{z}\| = \sum |z_i|$.

We let $\mathbf{0}$ denote the origin in $\mathbb{Z}_+^d$, and to each point $\boldsymbol{z} \in \mathbb{Z}_+^d \setminus \mathbf{0}$ we associate the *weight* $X(\boldsymbol{z})$. We assume the weights $\{X(\boldsymbol{z}) : \boldsymbol{z} \in \mathbb{Z}_+^d\}$ are i.i.d. random variables, with common distribution function F, where $F(s) = \mathbb{P}(X(\boldsymbol{z}) \leq s)$. By convention we set $X(\mathbf{0}) = 0$.

A directed path in $\mathbb{Z}_+^d$ is a path in which each step increases a single coordinate by 1. For $\boldsymbol{z}, \boldsymbol{z}' \in \mathbb{Z}_+^d$, with $\boldsymbol{z} \leq \boldsymbol{z}'$, we let $\Pi(\boldsymbol{z}, \boldsymbol{z}']$ denote the set of directed paths beginning at $\boldsymbol{z}$ and ending at $\boldsymbol{z}'$. We identify each path in $\Pi(\boldsymbol{z}, \boldsymbol{z}']$ by the set of points it contains, excluding $\boldsymbol{z}$ by convention.

For $\boldsymbol{z} \leq \boldsymbol{z}'$, we define the last–passage time $T_\ell(\boldsymbol{z}, \boldsymbol{z}')$ from $\boldsymbol{z}$ to $\boldsymbol{z}'$ by

$$T_\ell(\boldsymbol{z}, \boldsymbol{z}') = \max_{\pi \in \Pi(\boldsymbol{z}, \boldsymbol{z}']} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}), \tag{1.8.1}$$

with $T_\ell(\boldsymbol{z}, \boldsymbol{z}) = 0$. We define $T_\ell(\boldsymbol{z}) := T_\ell(\mathbf{0}, \boldsymbol{z})$ for all $\boldsymbol{z} \in \mathbb{Z}_+^d$.

If $\boldsymbol{z}_1 \leq \boldsymbol{z}_2 \leq \boldsymbol{z}_3 \in \mathbb{Z}_+^d$ then the following superadditivity property holds for the last–passage time:

$$T_\ell(\boldsymbol{z}_1, \boldsymbol{z}_2) + T_\ell(\boldsymbol{z}_2, \boldsymbol{z}_3) \leq T_\ell(\boldsymbol{z}_1, \boldsymbol{z}_3). \tag{1.8.2}$$

The last–passage time may be used to define a growth model in $\mathbb{R}_+^d$. In particular let

$$B_t = \big\{\boldsymbol{x} \in \mathbb{R}_+^d : T_\ell(\lfloor \boldsymbol{x} \rfloor) \leq t\big\},$$

be the last–passage shape at time $t \geq 0$. See Figure 1.11 for a simulation of B_t when $d = 2$ and the weights are Exp(1). In this case, the last–passage time has a one–to–one correspondence with the totally asymmetric simple exclusion process on $\mathbb{Z}$ with initial condition $\mathbb{1}(x < 0) + \mathbb{1}(x = 1)$. Labelling particles $i = 1, 2, \ldots$ from right to left, if $x = \sup\{y \in \mathbb{R}_+ : T_\ell(i, y) \leq t\}$ then particle i is in position $x + 1 - i$ at time t if $i \geq 2$, and if $i = 1$ particle 1 is in position $x + 1$. In particular x is the displacement, at time t, of particle i from it's original position. See Figure 1.12 for an illustration of this correspondence.

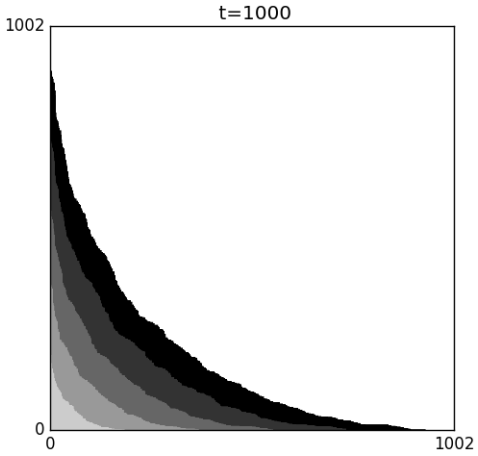
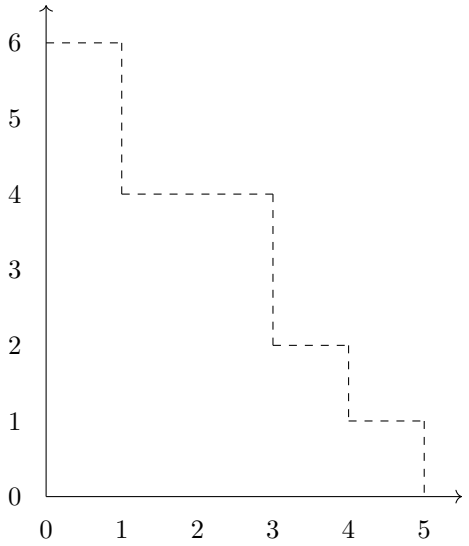
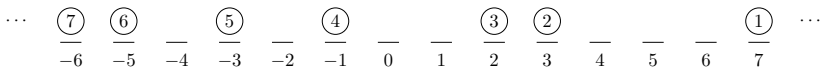


Figure 1.11: A simulation of B_{1000} with $d = 2$ and Exp(1) weights. The box contains all real numbers $\boldsymbol{x} \in \mathbb{R}^2_+$ with $x_1, x_2 \leq 1002$. The shapes $B_{200}, B_{400} \setminus B_{200}, \ldots, B_{1000} \setminus B_{800}$ are shaded from lightest to darkest.



(a) The last–passage percolation shape B_t with Exp(1) weights. All points in $\mathbb{R}^2_+$ strictly below the dashed line have last–passage time at most t.



(b) A configuration for the TASEP at time t with initial condition $\mathbb{1}(x < 0) + \mathbb{1}(x = 1)$, with particles labelled from right to left.

Figure 1.12: The growth model in Figure 1.12a corresponds to the TASEP percolation in Figure 1.12b.

The one–to–one correspondence in Figure 1.12 is an example of a height function; recall (1.3.4) and Figure 1.1.

For $\boldsymbol{x} \in \mathbb{R}_+^d$ we define the shape function

$$g(\boldsymbol{x}) = \sup_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E} T_\ell\big(\lfloor n\boldsymbol{x} \rfloor\big). \tag{1.8.3}$$

The form of the shape function when $d = 2$ and the weights are Exp(1) was first given by Rost [93]. In particular

$$g(x_1, x_2) = \big(\sqrt{x_1} + \sqrt{x_2}\big).$$

Johansson [55] showed that for $\alpha > 0$, the distribution of $n^{-1/3}[T_\ell((n, \lfloor \alpha n \rfloor)) - ng((1, \alpha))]$ converges as $n \to \infty$ to the Tracy–Widom distribution. Two–dimensional last–passage percolation with more general weight distributions have been examined in the context of tandem queueing systems; see, for example [9, 42, 82].

Analogously to (1.8.2), if $\mathbb{E}|X(\boldsymbol{z})| < \infty$, the shape function g satisfies

$$g(\boldsymbol{x}) + g(\boldsymbol{y}) \le g(\boldsymbol{x} + \boldsymbol{y}), \tag{1.8.4}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^d$. This is a consequence of the superadditive version of Kingman's subadditive ergodic theorem; see, for example [83][Proposition 2.1]. Continuity of the shape function g on all of $\mathbb{R}_+^d$ was proved by Martin [83][Theorem 2.3] under some general moment conditions.

Similarly, for $\boldsymbol{z} \le \boldsymbol{z}' \in \mathbb{Z}_+^d$ we define the first–passage time $T_f(\boldsymbol{z}, \boldsymbol{z}')$ from $\boldsymbol{z}$ to $\boldsymbol{z}'$ by

$$T_f(\boldsymbol{z}, \boldsymbol{z}') = \min_{\pi \in \Pi(\boldsymbol{z}, \boldsymbol{z}']} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}). \tag{1.8.5}$$

The first–passage shape C_t at time $t \ge 0$ is defined by

$$C_t = \big\{\boldsymbol{x} \in \mathbb{R}_+^d : T_f\big(\lfloor \boldsymbol{x} \rfloor\big) \le t\big\}.$$

See Figure 1.13 for a simulation of C_t when $d = 2$ and the weights are Exp(1).

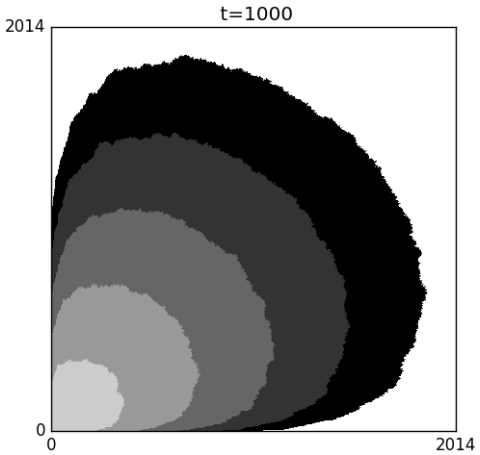


Figure 1.13: A simulation of C_{1000} with $d = 2$ and Exp(1) weights. The box contains all real numbers $\boldsymbol{x} \in \mathbb{R}_+^2$ with $x_1, x_2 \le 2014$. The shapes $C_{200}, C_{400} \setminus C_{200}, \ldots, C_{1000} \setminus C_{800}$ are shaded from lightest to darkest.

The shape function h is given by

$$h(\boldsymbol{x}) = \inf_{n\in\mathbb{N}} \frac{1}{n}\mathbb{E}T_f\big(\lfloor n\boldsymbol{x}\rfloor\big). \tag{1.8.6}$$

If $\boldsymbol{z}_1 \le \boldsymbol{z}_2 \le \boldsymbol{z}_3 \in \mathbb{Z}_+^d$ then the following subadditivity property holds for the first–passage time:

$$T_f(\boldsymbol{z}_1, \boldsymbol{z}_2) + T_f(\boldsymbol{z}_2, \boldsymbol{z}_3) \ge T_f(\boldsymbol{z}_1, \boldsymbol{z}_3). \tag{1.8.7}$$

By the superadditive version of Kingman's subadditive ergodic theorem, if $\mathbb{E}|X(\boldsymbol{z})| < \infty$, the shape function h satisfies

$$h(\boldsymbol{x}) + h(\boldsymbol{y}) \ge h(\boldsymbol{x}+\boldsymbol{y}), \tag{1.8.8}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^d$.

We note that both B_t and C_t are increasing in t, in the sense that $B_s \subseteq B_t$ and $C_s \subseteq C_t$ for $0 \le s \le t$. Let B and C be the subsets of $\mathbb{R}_+^d$ given by

$$B = \big\{\boldsymbol{x} : g(\boldsymbol{x}) \le 1\big\},$$
$$C = \big\{\boldsymbol{x} : h(\boldsymbol{x}) \le 1\big\}.$$

B is concave and C is convex by (1.8.4) and (1.8.8) respectively. The following theorem says that B and C are the asymptotic shapes for the processes $(B_t)_{t\ge0}$ and $(C_t)_{t\ge0}$ respectively.

Theorem 1.8 ([83][Theorem 5.1])**.** *Suppose that the weight distribution F satisfies*

$$\int_{-\infty}^{0} F(s)^{1/d} ds < \infty, \tag{1.8.9}$$

and

$$\int_{-\infty}^{0} \big(1 - F(s)\big)^{1/d} ds < \infty. \tag{1.8.10}$$

(i) Last–passage shape theorem.
If $\mathbb{E}X > 0$, then for any $\epsilon > 0$,

$$(1-\epsilon)B \subseteq \frac{B_t}{t} \subseteq (1+\epsilon)B, \tag{1.8.11}$$

for all t sufficiently large, with probability 1.

(ii) First–passage shape theorem.
If

$$\inf_{\boldsymbol{x}\in\mathbb{R}_+^d\setminus\mathbf{0}} \frac{h(\boldsymbol{x})}{\|\boldsymbol{x}\|} > 0,$$

then for any $\epsilon > 0$,

$$(1-\epsilon)C \subseteq \frac{C_t}{t} \subseteq (1+\epsilon)C, \tag{1.8.12}$$

for all t sufficiently large, with probability 1.

Analogous results are known for undirected first–passage percolation. See Cox and Durrett [27] for a shape theorem in two dimensions.

It is known [30, 81] that regions in which the limiting shape for undirected first–passage percolation is linear at the boundary, i.e. there is a flat edge, occur when the weight distribution is supported on $[1, \infty)$ and attains its minimum 1 with sufficiently high probability. That is, with probability greater than the critical probability for oriented percolation in $\mathbb{Z}^d$; see, for example, Durrett [31]. However, it is believed that the the shape is strictly convex outside of a region known as the percolation cone, under very general assumptions on the weight distribution.

In last–passage percolation, regions in which the shape function g is linear occur when the weight distribution is nonnegative and bounded, attaining its maximum with sufficiently high probability. It is not clear that there are any other ways in which g can fail to be strictly concave. See Martin [84] for a survey of last–passage percolation. There are examples of inhomogeneous environments which give rise to shapes with flat edges for last–passage percolation, and points in the boundary which are not differentiable. See, for example, Ciech and Georgiou [26].

1.8.1 Intermediate–passage percolation.

We now introduce a two–dimensional process termed intermediate–passage percolation. We consider a random environment with i.i.d. weights $X(\boldsymbol{z})$ for $\boldsymbol{z} \in \mathbb{Z}^2_+ \setminus \mathbf{0}$, and $X(\mathbf{0}) = 0$.

Definition 1.8.1. *The neighbourhood $\mathcal{N}(\boldsymbol{z})$ of $\boldsymbol{z} \in \mathbb{N}^2$ is defined to be the set*

$$\mathcal{N}(\boldsymbol{z}) = \big\{(z_1 - 1, z_2), (z_1, z_2 - 1), (z_1 - 1, z_2 - 1)\big\} \subseteq \mathbb{Z}^2_+.$$

Given a random environment $\{X(\boldsymbol{z})\}_{\boldsymbol{z} \in \mathbb{Z}^2_+}$, we define the intermediate–passage time $T(\boldsymbol{z})$ for each site $\boldsymbol{z} \in \mathbb{Z}^2_+$ via the following boundary conditions and recursion.

Definition 1.8.2. *If $\boldsymbol{z} = (z_1, z_2) \in \mathbb{Z}^2_+$ with $z_1 z_2 = 0$ we say that $\boldsymbol{z}$ is a boundary site. For $z_1, z_2 \geq 0$, we define the intermediate–passage times for the boundary sites via*

$$\begin{aligned} T(z_1, 0) &= \sum_{i=1}^{z_1} X(i, 0), \\ T(0, z_2) &= \sum_{i=1}^{z_2} X(0, i). \end{aligned} \tag{1.8.13}$$

Suppose the intermediate–passage times for the boundary sites are given, then the intermediate–passage times for the sites $\boldsymbol{z} \in \mathbb{N}^2$ may be defined recursively via

$$T(\boldsymbol{z}) = X(\boldsymbol{z}) + \text{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\big\}. \tag{1.8.14}$$

Here the median refers to the middle value when the elements of $\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\}$ are ordered from smallest to largest.

The intermediate–passage time may be used to define a growth model in the plane. We call this growth model intermediate–passage percolation.

Definition 1.8.3. *We define the shape at time t for intermediate–passage percolation by*

$$\Gamma_t = \big\{\boldsymbol{x} \in \mathbb{R}^2_+ : T(\lfloor \boldsymbol{x} \rfloor) < t\big\}.$$

See Figure 1.14 for a simulation of the shape Γ_t.

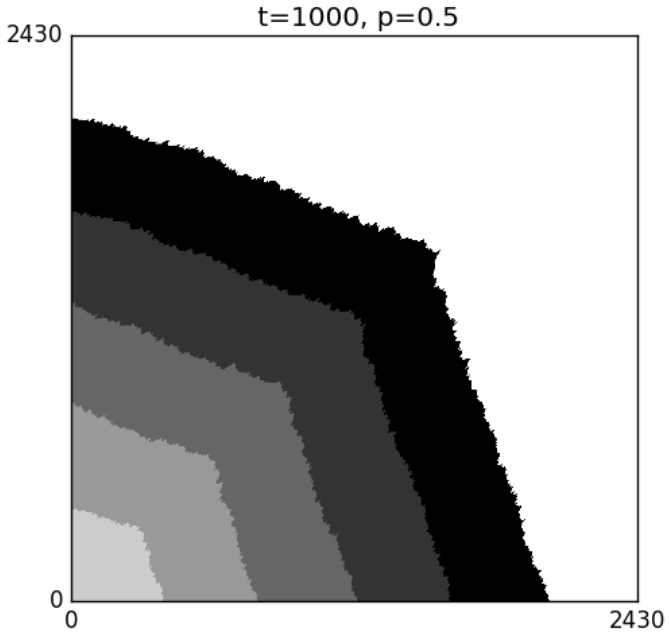


Figure 1.14: A simulation of the shape Γ_{100} with Bernoulli(0.5) weights satisfying $\mathbb{P}[X(\boldsymbol{z}) = 1] = 1 - \mathbb{P}[X(\boldsymbol{z}) = 0] = 0.5$. The box contains all real numbers $\boldsymbol{x} \in \mathbb{R}^2_+$ with $x_1, x_2 \leq 2430$, and the shapes $\Gamma_{200}, \Gamma_{400} \setminus \Gamma_{200}, \ldots, \Gamma_{1000} \setminus \Gamma_{800}$ are shaded from lightest to darkest.

In Chapter 4 we show that, on the same random environment, Γ_t is bounded above and below by B_t and C_t respectively. See Figure 1.15 for an example. We also give sufficient conditions for the shape Γ_t to grow linearly, as a function of t, in all directions. These conditions contain cases for which linear growth in all directions does not hold for first or last–passage percolation. Finally, we will discuss some open questions for intermediate–passage percolation based on a collection of simulations.

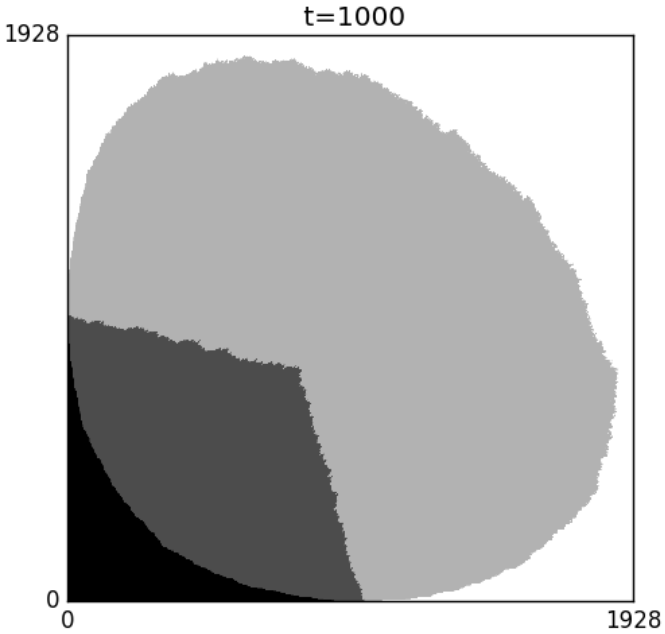
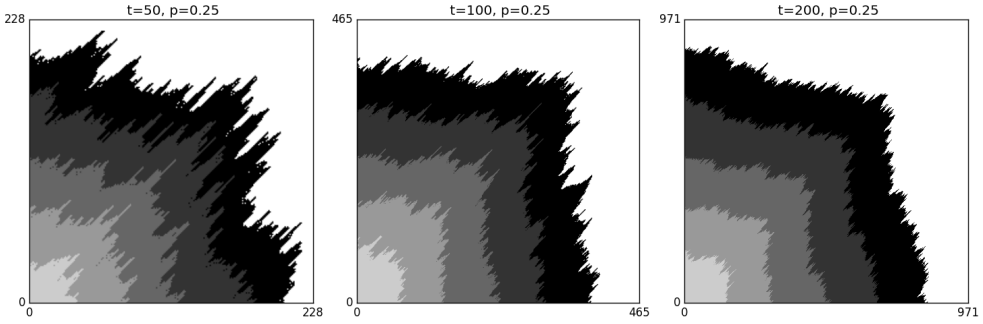


Figure 1.15: A simulation of the shapes B_{1000}, Γ_{1000} and C_{1000} on the same random environment with Exp(1) weights. The box contains all real numbers $\boldsymbol{x} \in \mathbb{R}^2_+$ with $x_1, x_2 \leq 1928$. The shapes $B_t, \Gamma_t \setminus B_t$ and $C_t \setminus \Gamma_t$ are shaded from darkest to lightest.

See Figure 1.16 for some simulations of intermediate–passage percolation with Bernoulli weights.



(a) (b) (c)

Figure 1.16: Figures 1.16a,1.16b and 1.16c are simulations of Γ_t with Bernoulli(0.25) weights. Figure 1.16a gives the shapes $\Gamma_{10}, \Gamma_{20} \setminus \Gamma_{10}, \ldots, \Gamma_{50} \setminus \Gamma_{40}$ shaded from lightest to darkest. Figure 1.16b gives the shapes $\Gamma_{20}, \Gamma_{40} \setminus \Gamma_{20}, \ldots, \Gamma_{100} \setminus \Gamma_{80}$ shaded from lightest to darkest. Figure 1.16c gives the shapes $\Gamma_{40}, \Gamma_{80} \setminus \Gamma_{40}, \ldots, \Gamma_{200} \setminus \Gamma_{160}$.

Chapter 2

Mixing times of the facilitated exclusion process

The facilitated simple exclusion process (FSEP), introduced in Section 1.3.2, is a one–dimensional exclusion process with a dynamical constraint. In the dynamics of the FSEP a particle at site x attempts to jump right to site $x+1$ at rate p, only doing so if there's a hole at site $x+1$ and a particle at site $x-1$. Moreover, a particle at site x attempts to jump left to site $x-1$ at rate q, only doing so if there's a hole at site $x-1$ and a particle at site $x+1$.

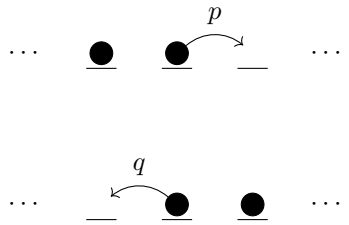


Figure 2.1: Transition rates for the FSEP.

For the FSEP on the segment, with closed boundaries, we consider the state space

$$\Omega_{N,k} = \left\{ \xi \in \{0,1\}^{[N]} \,\middle|\, \sum_{x=1}^{N} \xi(x) = k \right\},$$

with $k > N/2$. The generator $\mathcal{L}$ of the FSEP on the segment acts on functions $f : \Omega_{N,k} \to \mathbb{R}$ via

$$\begin{aligned} \mathcal{L}f(\xi) = &\sum_{x=2}^{N-1} p\xi(x)\xi(x-1)\big(1-\xi(x+1)\big)\big[f(\xi^{x,x+1}) - f(\xi)\big] \\ &+ \sum_{x=2}^{N-1} q\xi(x)\xi(x+1)\big(1-\xi(x-1)\big)\big[f(\xi^{x,x-1}) - f(\xi)\big]. \end{aligned}$$

Recall from Section 1.3.2 that the FSEP on the segment has degenerate rates and admits transient

states, which the process will eventually leave to reach an ergodic component

$$\mathcal{E}_{N,k} = \big\{\xi \in \Omega_{N,k} : \xi(1)\xi(N) = 1,\, \xi(x) + \xi(x+1) \geq 1 \;\forall\, x \in [N-1]\big\}.$$

See Definition 1.3.2. For the FSEP on the circle $\mathbb{T}_N$ we consider the state space

$$\Omega_{N,k} = \left\{\xi \in \{0,1\}^{\mathbb{T}_N} \;\middle|\; \sum_{x\in\mathbb{T}_N} \xi(x) = k\right\},$$

with $k > N/2$. The generate $\mathcal{L}^\circ$ acts on functions $f : \Omega'_{N,k} \to \mathbb{R}$ via

$$\begin{aligned}\mathcal{L}^\circ f(\xi) = &\sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x-1)\big((1-\xi(x+1)\big)\big[f(\xi^{x,x+1}) - f(\xi)\big] \\ &+ \sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x+1)\big((1-\xi(x-1)\big)\big[f(\xi^{x,x-1}) - f(\xi)\big]\,.\end{aligned}$$

Recalling Definition 1.3.3, the ergodic component on the circle is given by

$$\mathcal{G}_{N,k} = \big\{\xi \in \Omega'_{N,k} : \xi(x) + \xi(x+1) \geq 1 \;\forall\, x \in \mathbb{T}_N\big\}\,.$$

In this chapter we establish some bounds on the mixing time for the FSEP on both the segment and the circle. To avoid confusion, we will separate the discussions for the segment and circle as far as possible.

In Section 2.1.1 we state formally the main results for the FSEP on the segment: the symmetric FSEP on the segment $[N]$ with $k > N/2$ particles satisfies pre–cutoff (1.5.17) on the order of $k^2\log(N-k)$ under the conditions that both $2k-N$ and $N-k$ go to infinity, and the asymmetric FSEP on the segment $[N]$ with $k > N/2$ particles mixes exponentially slowly in the number of holes $N-k$ provided that $N-k$ is sufficiently large.

In Section 2.1.2 we state formally the main result for the FSEP on the circle: the mixing time of the FSEP on the circle $\mathbb{T}_N$ with particle density $\alpha \in (1/2, 1)$ started from any 'reasonable' initial condition is $N^2 \log N$, up to universal constant prefactors.

Section 2.3 will focus entirely on the FSEP on the segment. We describe the dynamics in terms of lattice paths and show that the dynamics are monotone w.r.t. a partial order. We then compare the minimal and maximal lattice paths to both the simple exclusion process (SEP) and the open boundary exclusion process (OBEP).

Section 2.4 will focus on the FSEP on the circle, although we will sometimes compare to the FSEP on the segment. We use a relaxation of the 'first' Fourier coefficient argument for a lower bound on the mixing time. For the upper bound, we show that 'reasonable' initial conditions hit the ergodic component on the order of $N^2 \log N$ with high probability, and give a lower bound on the log–Sobolev constant (Definition 1.5.5). The lower bound on the log–Sobolev is obtained using entropy arguments.

There is significant interest in interacting particle systems with degenerate rates; bounds on the spectral gap for interacting particle systems with kinetic constraints can be found, for example, in [4, 16, 21, 38, 85, 88].

2.1 Notation and results

Recalling (1.7.4), we say that $f(n) = O(g(n))$ as $n \to \infty$ if $|f|$ is bounded above by g, up to a constant factor, asymptotically. That is,

$$\limsup_{n\to\infty} \frac{|f(n)|}{g(n)} < \infty.$$

Moreover, we say that $f(n) = \Omega(g(n))$ as $n \to \infty$ if $g(n) = O(f(n))$ as $n \to \infty$. We will write $f(n) \ll g(n)$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0, \tag{2.1.1}$$

and $f(n) \asymp g(n)$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1. \tag{2.1.2}$$

In Section 1.3.2 we saw that the stationary distribution $\mu_{N,k}$ of the FSEP on the segment $[N]$, with $k > N/2$ particles, is uniform on $\Omega_{N,k}$ when $p = 1/2$. For $p > 1/2$, if x_i denotes the position of the i–th leftmost hole on the segment for a configuration $\xi \in \Omega_{N,k}$, $\lambda = q/p$ and

$$A(\xi) = \sum_i x_i,$$

then the stationary distribution (see (1.3.10)) is given by

$$\mu_{N,k}(\xi) = \frac{\lambda^{-A(\xi)}}{\sum_{\xi' \in \mathcal{E}_{N,k}} \lambda^{-A(\xi')}}.$$

Recalling (1.5.5) and (1.5.12) we define the distance

$$d_{N,k}(t) := \max_{\xi\in\Omega_{N,k}} \left\| P_t^\xi - \mu_{N,k} \right\|_{\mathrm{TV}},$$

and denote the ϵ–mixing time of the FSEP on the segment $[N]$ with $k > N/2$ particles by $T_{\mathrm{seg}}^{N,k}(\epsilon)$.

For the FSEP on the circle, the stationary measure $\nu_{N,k}$ is uniform and we let

$$d'_{N,k}(t) := \max_{\xi\in\Omega'_{N,k}} \left\| P_t^\xi - \nu_{N,k} \right\|_{\mathrm{TV}},$$

and denote the ϵ–mixing time of the FSEP on the circle $\mathbb{T}_N$ with $k > N/2$ particles by $T_{\mathrm{cir}}^{N,k}(\epsilon)$.

2.1.1 Results for the segment

Theorem 2.1 (Symmetric FSEP)**.** *For the FSEP on the segment with parameter $p = 1/2$, let $k(N)$ be a sequence satisfying $N/2 < k(N) < N$ such that $N - k$ and $k - N/2$ go to infinity. For all $\epsilon \in (0, 1)$ there exists positive constants $0 < C_1 < C_2$ which do not depend on ϵ such that*

$$C_1 \le \liminf_{N\to\infty} \frac{T_{\mathrm{seg}}^{N,k}(\epsilon)}{N^2 \log(N-k)} \le \limsup_{N\to\infty} \frac{T_{\mathrm{seg}}^{N,k}(\epsilon)}{N^2 \log(N-k)} \le C_2. \tag{2.1.3}$$

In particular the symmetric FSEP mixes on the order of $N^2 \log(N-k)$. Moreover, pre–cutoff (1.5.17) holds for the symmetric FSEP under the conditions of Theorem 2.1. We note that Theorem 2.1 covers the regime where the density of particles $k/N \to \alpha \in [1/2, 1]$ as $N \to \infty$.

Remark 2.1.1. *We conjecture that cutoff* (1.5.16) *holds for the symmetric FSEP under the conditions of Theorem 2.1, i.e. we may take* $C_1 = C_2$.

On the ergodic component $\mathcal{E}_{N,k}$*, the mixing time of the FSEP is determined by the mixing time of the SSEP on the segment. We will show that cutoff of order* $k^2 \log \min(N-k, 2k-N)$ *holds for the symmetric FSEP restricted to the ergodic component as a consequence of [63, Theorem 2.4]. In particular when* $\log(2k-N) \ll \log(N-k)$ *the mixing time is determined by the hitting time of the ergodic component.*

For the asymmetric FSEP on the segment we show that the mixing time is exponentially slow in the number of holes $N-k$, provided the number of holes is large enough.

Theorem 2.2 (Asymmetric FSEP)**.** *For the FSEP on the segment with parameter* $p \in (1/2, 1)$*, let* $\epsilon \in (0,1)$ *and* $k(N)$ *be a sequence such that* $N - k \gg \log N$*, then the* ϵ*–mixing time satisfies*

$$\lim_{N\to\infty} \frac{\log T_{\text{seg}}^{N,k}(\epsilon)}{N-k} = \log\Big(\frac{p}{q}\Big). \tag{2.1.4}$$

We note the similarity between Theorem 2.2 and Theorem 1.7 for the reverse–bias phase of the OBEP. This is not a coincidence; the result will essentially follow from a comparison between the reverse–bias phase of the OBEP and a class of transient initial conditions in $\mathcal{E}_{N,k}^c$ for the asymmetric FSEP. In particular, some initial conditions must work against a drift before hitting the ergodic component.

On the ergodic component $\mathcal{E}_{N,k}$, the mixing time of the FSEP is determined by the mixing time of the ASEP on the segment. We will show that cutoff of order N holds for the asymmetric FSEP restricted to the ergodic component as a consequence of cutoff for the ASEP on the segment [60, Theorem 2]. Therefore the asymmetric FSEP on the segment mixes very rapidly, with order at most N, once it reaches the ergodic component, and the the mixing time is determined by the hitting time of the ergodic component.

2.1.2 Results for the circle

Before stating the main result for the FSEP on the circle we introduce some notation.

Definition 2.1.2. *Let* $\xi \in \mathcal{G}_{N,k}^c$ *and define* $[x,y] = \{x, x+1, \dots, y\} \subset \mathbb{T}_N$ *to be a (clockwise) interval on the circle. We say that* $[x,y]$ *is an ergodic interval if* $\xi(x) = 1$*,* $\xi(y) = 1$ *and* $\xi(z) + \xi(z+1) \geq 1$ *for all* $z \in [x,y]$ *such that* $z+1 \in [x,y]$*. In particular the singleton* $\{x\}$ *is an ergodic interval if* $\xi(x) = 1$*. We say that an ergodic interval* I *for the configuration* ξ *contains* k_I *particles if*

$$\sum_{x\in I} \xi(x) = k_I.$$

We define the set of ergodic regions of ξ *to be the smallest set of ergodic intervals* $\{I_1, \dots, I_n\}$ *which contain all particles in* ξ*. In particular, any ergodic interval* $[x,y]$ *satisfies* $[x,y] \subseteq I_i$ *for some* $i \in [n]$*, and* $\xi(x) = 0$ *for all* $x \notin \{I_1, \dots, I_n\}$.

We define $\mathcal{I}^m_{N,k} \subset \mathcal{G}^c_{N,k}$ *to be the configurations for which there exists a set of at most* m *ergodic regions containing at least* $N-k$ *particles collectively, i.e.*

$$\mathcal{I}^m_{N,k} = \Bigg\{ \xi \in \mathcal{G}^c_{N,k} : \text{ there exists a set of ergodic regions } \{I_1, \dots, I_{m'}\},$$
$$m' \le m \text{ such that } \sum_{i=1}^{m'} \sum_{x \in I_i} \xi(x) \ge N-k \Bigg\}.$$

See Figure 2.2 for an example of ergodic regions.

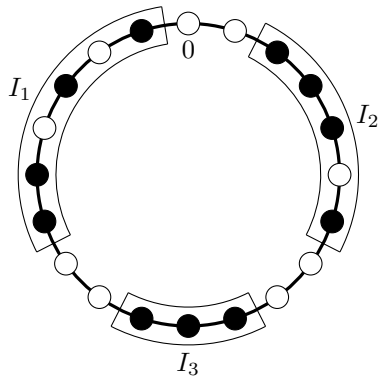


Figure 2.2: The above figure represents a configuration in $\mathcal{G}^c_{20,11} \subset \Omega'_{20,11}$, where black circles denote particles and white circles denote holes. The figure is oriented so that site 0 is at the top, and the ergodic regions from Definition 2.1.2 are given by the intervals $I_1 = [14, 19], I_2 = [2, 6]$ and $I_3 = [9, 11]$.

Theorem 2.3. *Let* $\epsilon \in (0,1)$ *be given, and let* $k(N) > N/2$ *be a sequence such that* $k(N)/N \to \alpha \in (1/2, 1)$ *as* $N \to \infty$.

(a) The ϵ*–mixing time of the symmetric FSEP on the circle satisfies*

$$T^{N,k}_{\text{cir}}(\epsilon) = \Omega(N^2 \log N), \tag{2.1.5}$$

as $N \to \infty$.

(b) Fix $m > 0$ *to be a positive integer that does not depend on* N. *If*

$$d^m_{N,k}(t) := \max_{\xi \in \mathcal{I}^m_{N,k} \cup \mathcal{G}_{N,k}} \|P^\xi_t - \nu_{N,k}\|_{\text{TV}},$$

we define the ϵ*–mixing time of the FSEP on the circle restricted to initial conditions in* $\mathcal{I}^m_{N,k} \cup \mathcal{G}_{N,k}$ *by*

$$\widehat{T}^{N,k,m}_{\text{cir}}(\epsilon) := \inf\big\{ t \ge 0 \mid d^m_{N,k}(t) \le \epsilon \big\}.$$

The restricted mixing time satisfies

$$\widehat{T}^{N,k,m}_{\text{cir}}(\epsilon) = O(N^2 \log N), \tag{2.1.6}$$

as $N \to \infty$.

If we further restrict to the FSEP on the ergodic component $\mathcal{G}_{N,k}$, then Theorem 2.3 shows that pre–cutoff holds. The conditions on the sequence $k(N)$ in Theorem 2.3 are required in the proof of the upper bound, but a lower bound analogous to (2.1.5) may be shown under more general conditions.

Remark 2.1.3. *The result states that the mixing time started from any 'reasonable' initial condition is $N^2 \log N$, up to universal constant prefactors. Obtaining good upper bounds on the hitting time of the ergodic component for general initial conditions in $\mathcal{G}_{N,k}^c$ appears to be quite challenging. We conjecture that the initial conditions that maximise the hitting time of the ergodic component, and control the mixing time, are those in which there is one ergodic region of length k. These initial conditions are contained in the set $\mathcal{I}_{N,k}^m$ for each m.*

We also conjecture that the symmetric FSEP on the circle has cutoff under suitable conditions.

2.2 A key observation

The intuitive idea we will use throughout the rest of the chapter is that particle–hole (or by symmetry hole–particle) objects in the FSEP move like extended exclusion objects of size 2, see Figure 2.3.

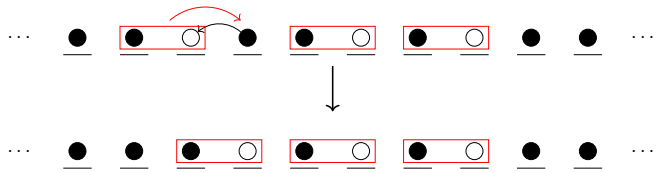


Figure 2.3: In the above figure black circles denote particles and white circles denote holes. Particle–hole objects are placed in red boxes. In the dynamics, the third leftmost particle in the top figure jumps left, which corresponds to the leftmost particle–hole object jumping one space to the right. In particular the boxes do not overlap in the dynamics and particle–hole objects behave like particles in the SEP. By symmetry this also holds for hole–particle objects.

This behaviour is observed in Gabel et al. [39]. Baik et al. [10] use this observation to couple a totally asymmetric facilitated exclusion process to a totally asymmetric simple exclusion process (TASEP) on both the half–line and $\mathbb{Z}$. Similarly Ayyer et al. [47, Theorem 4.3] couple the asymmetric FSEP with the asymmetric simple exclusion process (ASEP) on $\mathbb{Z}$.

2.3 FSEP on the segment

Before proceeding with the proofs of Theorem 2.1 and Theorem 2.2 in Section 2.3.4 we present some useful tools and properties particular to the FSEP on the segment. In Section 2.3.1 we describe the dynamics of the FSEP in terms of lattice paths. The lattice paths have a natural partial order, and in Section 2.3.2 we show that this partial order is preserved under a graphical construction similar to the construction we saw for the SEP in Section 1.6. We then show how to relate the dynamics started from the the minimal and maximal lattice paths to the OBEP in Section 2.3.3.

2.3.1 Mapping to lattice paths

In this section we introduce a mapping to lattice paths for the FSEP on the segment analogous to the height function in (1.3.4) and Figure 1.1.

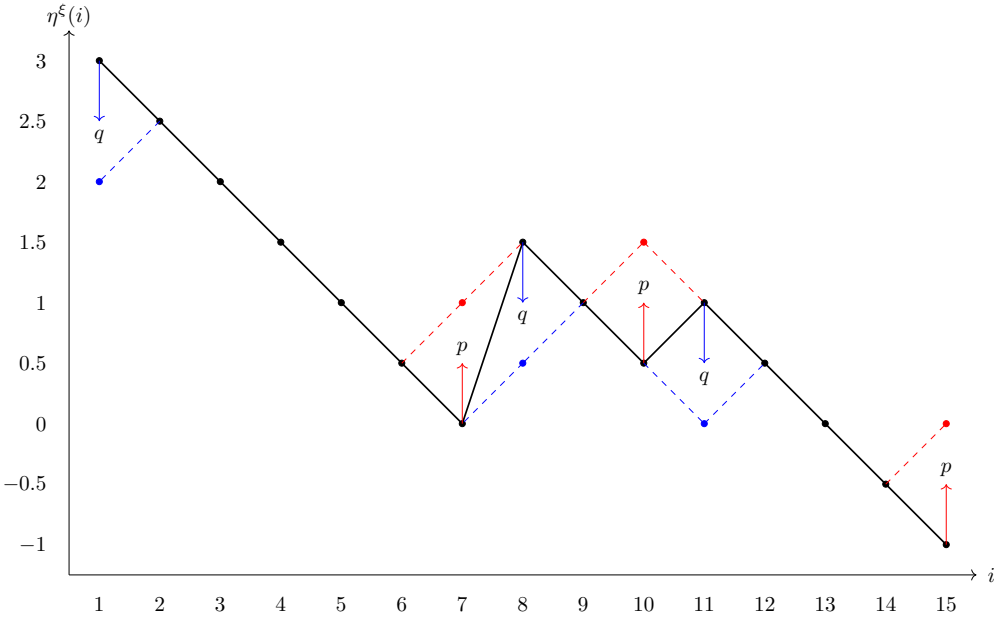
For any $\xi \in \Omega_{N,k}$ we label the particles from left to right and let y_i denote the position of the i–th leftmost particle in the configuration ξ. To each $\xi \in \Omega_{N,k}$ we may associate a lattice path η^ξ in the following manner:

$$\eta^\xi(i) = y_i - \frac{(N-1)i + k - N}{k-1}, \text{ for all } i \in [k]. \tag{2.3.1}$$

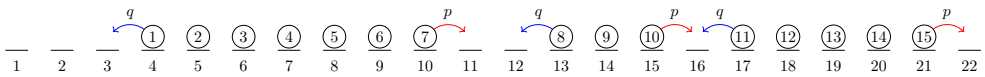
Equivalently, if $\eta^\xi(1) = y_1 - 1$, then the lattice path may be described via the following recursion:

$$\eta^\xi(i+1) = \eta^\xi(i) + y_{i+1} - (y_i + 1) - \frac{N-k}{k-1}. \tag{2.3.2}$$

We note that $y_{i+1} - (y_i + 1)$ is the number of holes between the i–th and $(i+1)$–th particles. See Figure 2.4 for an example of the lattice path.



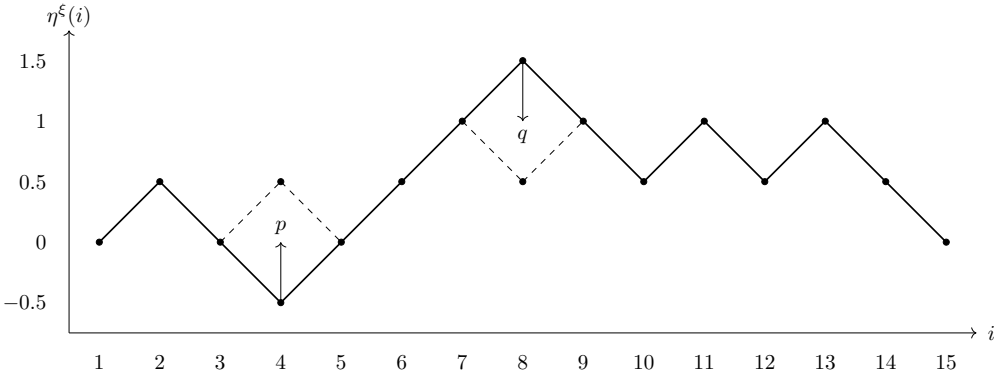
(a) The black path gives the lattice path η^ξ for the configuration $\xi \in \Omega_{22,15}$ given in Figure 2.4b. Read from left to right, the red and blue arrows give the only allowed transitions in the lattice path dynamics. These correspond to the clock rings $\mathcal{T}^{\downarrow}_{(1,3)}$, $\mathcal{T}^{\uparrow}_{(7,0)}$, $\mathcal{T}^{\downarrow}_{(8,1.5)}$, $\mathcal{T}^{\uparrow}_{(10,0.5)}$, $\mathcal{T}^{\downarrow}_{(11,1)}$ and $\mathcal{T}^{\uparrow}_{(15,-1)}$ respectively in the graphical construction. The dashed lines indicate the changes to the lattice path in the dynamics.



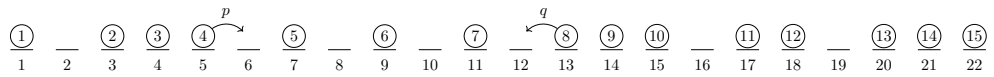
(b) A configuration $\xi \in \Omega_{22,15}$. The i–th particle from the left has label i, and sites are labelled below. The red and blue arrows, read from left to right, indicate the particle jumps corresponding to the clock rings $\mathcal{T}^{\downarrow}_{(1,3)}$, $\mathcal{T}^{\uparrow}_{(7,0)}$, $\mathcal{T}^{\downarrow}_{(8,1.5)}$, $\mathcal{T}^{\uparrow}_{(10,0.5)}$, $\mathcal{T}^{\downarrow}_{(11,1)}$ and $\mathcal{T}^{\uparrow}_{(15,-1)}$ respectively in the graphical construction. These are the only allowed transitions in the FSEP dynamics for ξ.

Figure 2.4: For configurations $\xi \in \mathcal{E}^c_{N,k}$, the lattice path may include 'steep' segments where $\eta^\xi(i+1) - \eta^\xi(i) > (2k - N - 1)/(k - 1)$ for $i = 1, \ldots, k-1$. See, for example, the segment between $\eta^\xi(7)$ and $\eta^\xi(8)$ in Figure 2.4a. Crucially the dynamics cannot create steep segments and steep segments cannot be made more steep. Steep segments are made less steep when particles in the FSEP make jumps which are not reversible, such as those indicated for particle 7 and 8 in Figure 2.4b. The left and right endpoints of the lattice path may only move down and up respectively, stopping at 0. This is because particle 1 and k in the FSEP move left and right respectively until they become stuck at the endpoints.

The constants in (2.3.1) are chosen so that lattice paths for configurations in the ergodic component satisfy $\eta^\xi(1) = \eta^\xi(k) = 0$, i.e. lattice paths in ergodic components are normalised. See, for example, Figure 2.5.



(a) The lattice path η^ξ for the configuration $\xi \in \mathcal{E}_{22,15}$ in Figure 2.5b. The up arrow and down arrow are examples of allowed transitions in the lattice path dynamics, and correspond to the clock rings $\mathcal{T}^{\uparrow}_{4,-0.5}$ and $\mathcal{T}^{\downarrow}_{8,1.5}$ respectively in the graphical construction. The dashed lines indicate the changes made to the lattice path in the indicated transitions. By construction the endpoints $\eta^\xi(1)$ and $\eta^\xi(k)$ cannot move in the lattice path dynamics.



(b) A configuration $\xi \in \mathcal{E}_{22,15}$. The i–th particle from the left has label i, and sites are labelled below. The arrows, read from left to right, indicate the particle jumps corresponding to the clock rings $\mathcal{T}^{\uparrow}_{4,-0.5}$ and $\mathcal{T}^{\downarrow}_{8,1.5}$ respectively in the graphical construction.

Figure 2.5: Unlike the lattice path in Figure 2.4a, the endpoints of the lattice path in Figure 2.5a are at zero. Moreover, there are no steep segments in the path. Since the dynamics cannot create steep segments in the path, the dynamics of the lattice paths restricted to the ergodic component are exactly the corner flip dynamics from Section 1.6 given by the graphical construction for the SEP.

The mapping $\xi \mapsto \eta^\xi$ is injective, and therefore bijective onto its image. With an abuse of notation we write $\Omega_{N,k}$ for the image set of $\xi \mapsto \eta^\xi$. Similarly we write $\mathcal{E}_{N,k}$ for the image set of the FSEP configurations in the ergodic component.

We now introduce a graphical construction for the dynamics on lattice paths. This graphical construction is analogous to the graphical construction given in Section 1.6.

- To each site (i, z), with $2 \le i \le k$ and
$$z \in I_i := -\frac{(N-1)i + k - N}{k-1} + \mathbb{Z} \cap [i, N-i],$$
we independently associate two independent Poisson clocks $\mathcal{T}^{\uparrow}_{(i,z)}$ and $\mathcal{T}^{\downarrow}_{(i,z)}$ of rate p and q respectively.
- To each site $(1, z)$ with $z \in I_1 := \mathbb{Z} \cap [1, N-k]$, we associate independent Poisson clocks $\mathcal{T}^{\downarrow}_{(1,z)}$ of rate q.
- To each site (k, z) with $z \in I_k := \mathbb{Z} \cap [k-N, -1]$ we associate independent Poisson clocks $\mathcal{T}^{\uparrow}_{(k,z)}$ of rate p.

Given $\xi \in \Omega_{N,k}$ we construct a trajectory $(\eta_t^\xi)_{t\geq 0}$ with initial condition η^ξ, for all $\xi \in \Omega_{N,k}$ simultaneously, as follows:

If $\mathcal{T}^{\uparrow}_{(i,z)}$ rings at time t, then

- if $\eta^\xi_{t-}(i) = z$ and η^ξ_{t-} has a local minimum at z then $\eta^\xi_t(i) = z+1$ and all other coordinates remain unchanged;
- otherwise no changes are made to the lattice path at time t.

If $\mathcal{T}^{\downarrow}_{(i,z)}$ rings at time t then

- if $\eta^\xi_{t-}(i) = z$ and η^ξ_{t-} has a local maximum at z then $\eta^\xi_t(i) = z-1$ and all other coordinates remain unchanged;
- otherwise no changes are made to the lattice path at time t.

The graphical construction gives a grand coupling of the FSEP on the segment. To see that this construction gives the Markov chain with the generator in (1.3.9), observe that local maxima in the lattice path occur at $i = 2, \ldots, k-2$ if and only if the positions of the particles satisfy $y_{i+1} = y_i + 1$ and $y_{i-1} < y_i - 1$. Similarly local minima occur at $i = 2, \ldots, k-2$ if and only if $y_{i-1} = y_i - 1$ and $y_{i+1} > y_i + 1$. A local maximum occurs at $i = 1$ with $\eta^\xi(i) > 0$ if and only if $y_1 > 1$ and $y_1 = y_2 - 1$, and a local minimum occurs at $i = k$ with $\eta^\xi(k) < 0$ if and only if $y_k < N$ and $y_{k-1} + 1 = y_k < N$.

Remark 2.3.1. *By comparing the above graphical construction and the graphical construction for the SEP in Section 1.6, the FSEP with parameter p restricted to the ergodic component $\mathcal{E}_{N,k}$ (see Figure 2.5) is equivalent to the SEP with parameter q on the state space $\Omega_{k-1,N-k}$. Moreover, each FSEP configuration on the ergodic component $\mathcal{E}_{N,k}$ may be uniquely identified with a SEP configuration in $\Omega_{k-1,N-k}$.*

Recalling the intuitive particle–hole idea from Section 2.2, the following figure illustrates the relation between the FSEP on the ergodic component and the SEP on the segment. The change in parameters from p for the FSEP to q for the SEP in Remark 2.3.1 occurs because a particle jumping right in the FSEP causes a particle–hole object to jump left and vice versa.

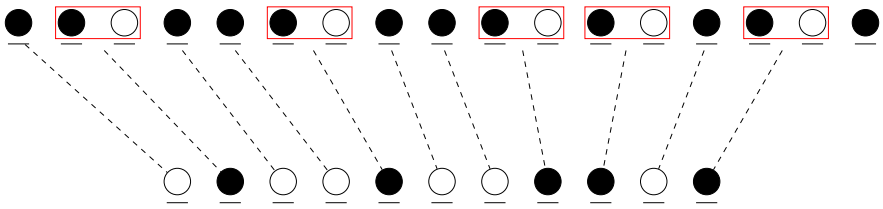


Figure 2.6: The top figure is an FSEP configuration in the ergodic component $\mathcal{E}_{17,12}$ and the bottom gives the corresponding SEP configuration in $\Omega_{11,5}$. The dashed lines indicate the role that each object in the SEP plays in the FSEP picture: each particle in the SEP plays the role of a particle–hole object in the FSEP and each hole in the SEP plays the role of a particle which is not followed by a hole in the FSEP. The rightmost particle in the FSEP configuration has no role in the SEP configuration. This is because the right most site a hole can occupy in the FSEP on $\Omega_{N,k}$ is $N-1$, and we chose the convention that particle–hole objects represent particles in the SEP configuration. If we instead choose the convention that hole–particle objects in the FSEP represent particles in the SEP, then the leftmost particle in the FSEP has no role in the SEP.

With an abuse of notation, we denote the law of the graphical construction by $\mathbb{P}$, and denote the marginal law of the trajectory at time t with initial condition $\eta \in \Omega_{N,k}$ by P_t^η.

2.3.2 Monotonicity

In this section we show that the graphical construction given in Section 2.3.1 preserves a partial order for the lattice paths in $\Omega_{N,k}$. The lattice paths admit a natural partial order; for $\eta, \eta' \in \Omega_{N,k}$ we say that $\eta \leq \eta'$ if

$$\eta(i) \leq \eta'(i) \text{ for all } i \in [k]. \tag{2.3.3}$$

Let $\eta = \eta^{\xi_1}$ and $\eta' = \eta^{\xi_2}$ for two particle configurations $\xi_1, \xi_2 \in \Omega_{N,k}$, and let y_i^1 and y_i^2 denote the positions of the i–th leftmost particles in ξ_1 and ξ_2 respectively. It is straightfoward to see from (2.3.1) that (2.3.3) is equivalent to $y_i^1 \leq y_i^2$ for all $i \in [k]$.

We now define some special configurations.

Definition 2.3.2. *Let η^- and η^+ denote the minimal and maximal lattice paths on $\Omega_{N,k}$ w.r.t. the partial order* (2.3.3)*. The minimal and maximal lattice paths arise from the FSEP configurations given by*

$$\begin{aligned} \xi^-(x) &= \mathbb{1}(x \leq k), \\ \xi^+(x) &= \mathbb{1}(x \geq N - k + 1), \end{aligned} \tag{2.3.4}$$

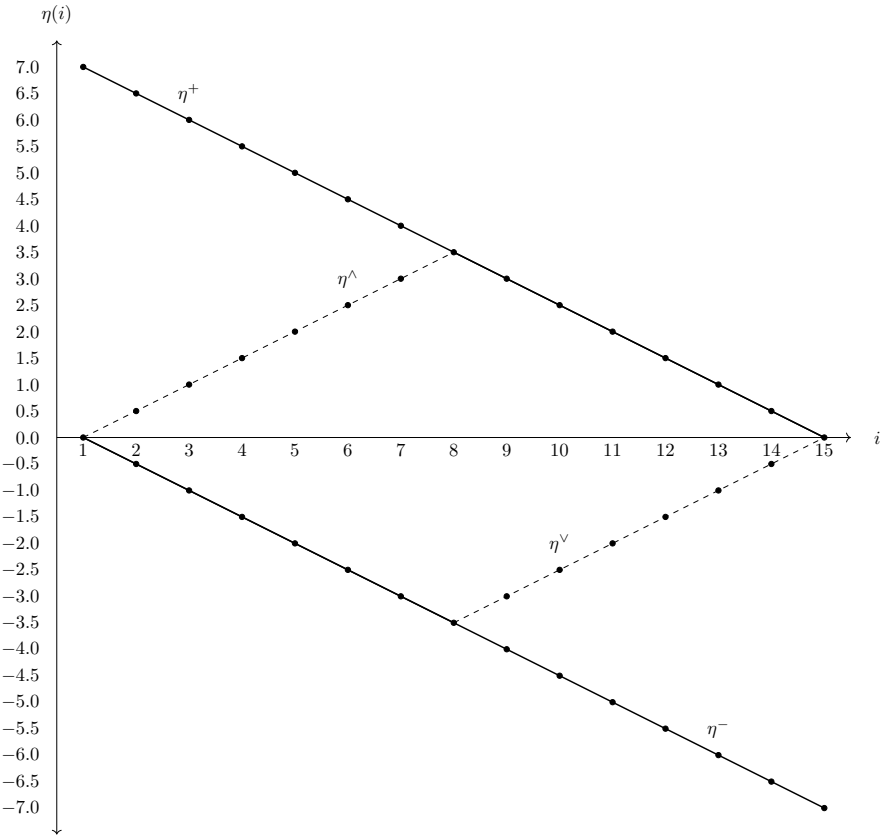
respectively.

Let $\eta^\vee$ and $\eta^\wedge$ denote the minimal and maximal configurations w.r.t. the partial order restricted to the ergodic component $\mathcal{E}_{N,k}$. This notation is chosen to be consistent with [63], i.e. the notation $\eta^\vee$ and $\eta^\wedge$ is chosen to represent the shapes of the corresponding lattice paths. The lattice paths $\eta^\vee$ and $\eta^\wedge$ arise from the FSEP configurations

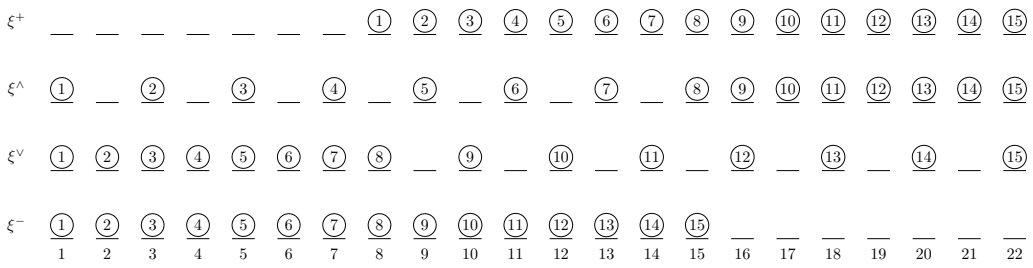
$$\xi^\vee(x) = \mathbb{1}\big(x \notin 2\mathbb{Z} \cap [1, 2N - 2k]\big) \text{ and } \xi^\wedge(x) = \xi^\vee(N + 1 - x),$$

respectively.

See Figure 2.7a for an example of the lattice paths from Definition 2.3.2 and Figure 2.7b for their corresponding FSEP configurations.



(a) The lattice paths $\eta^+, \eta^-, \eta^\wedge$ and $\eta^\vee$ on the state space $\Omega_{22,15}$. The upper and lower bold lines correspond to the lattice paths η^+ and η^- respectively. The upper dashed line indicates the part of the lattice path $\eta^\wedge$ not contained in η^+. The lower dashed line indicates the part of the lattice path $\eta^\vee$ not contained in η^-.



(b) The configurations $\xi^+, \xi^\wedge, \xi^\vee$ and ξ^- on the state space $\Omega_{22,15}$ with particles labelled $i = 1, \ldots, 15$ from left to right.

Figure 2.7: Figure 2.7a contains the minimal and maximal lattice paths w.r.t. the partial order (2.3.3) on the state space $\Omega_{22,15}$ and the ergodic component $\mathcal{E}_{22,15}$. Figure 2.7b gives the FSEP configurations which give rise to the lattice paths in Figure 2.7a.

The following proposition gives that the partial order (2.3.3) is conserved under the coupling

$\mathbb{P}$, i.e. the lattice path dynamics are monotone.

Proposition 2.3.3. *Let $\eta, \eta' \in \Omega_{N,k}$ be two lattice paths with $\eta \leq \eta'$ then the lattice path trajectories $(\eta_t)_{t\geq 0}$ and $(\eta'_t)_{t\geq 0}$ satisfy*

$$\mathbb{P}\big(\eta_t \leq \eta'_t\big) = 1, \tag{2.3.5}$$

for all $t \geq 0$.

The proof of Proposition 2.3.3 is analogous to the proof of [63, Proposition 3.1].

Proof of Proposition 2.3.3. It is sufficient to check that (2.3.5) is conserved on the clock rings in the graphical construction $\mathbb{P}$. Suppose that $\eta_{t-} \leq \eta'_{t-}$ and there is a clock ring in the graphical construction at time t. Assume WLOG that $T^{\uparrow}_{(i,z)}$ rings at time t for some $i \in [k-1]$.

If $\eta_{t-}(i)$ is not a local minimum, no change is made to the lattice path in the dynamics and (2.3.5) must hold at time t since $T^{\uparrow}_{(i,z)}$ cannot make coordinates of the lattice path η'_{t-} smaller. Henceforth we assume that $\eta_{t-}(i)$ is a local minimum.

If $\eta_{t-}(i) \leq \eta'_{t-}(i) - 1$ then $\eta_t(i) \leq \eta'_t(i)$ since the clock ring $T^{\uparrow}_{(i,z)}$ can increase coordinates in both lattice paths by at most 1.

The only remaining case is $\eta_{t-}(i) = \eta'_{t-}(i)$, and the lattice paths can only change if $\eta_{t-}(i) = \eta'_{t-}(i) = z$. By assumption $\eta_{t-}(i) = z$ is a local minimum and it holds that $\eta_{t-}(i-1) = z + (N-k)/(k-1)$, if $i \geq 2$. Moreover, it holds that $\eta_{t-}(i+1) > z$ if $i \leq k-2$, but by the partial order (2.3.5) at time $t-$ we have that $\eta'_{t-}(i-1) \geq \eta_{t-}(i-1)$, if $i \geq 2$, and $\eta'_{t-}(i+1) \geq \eta_{t-}(i+1)$ if $i \leq k-2$. Therefore $\eta'_{t-}(i)$ must also be a local minimum, and in the dynamics $\eta_t(i) = \eta'_t(i) = z+1$. No other coordinates in the lattice paths are changed and (2.3.5) holds at time t. □

2.3.3 Comparison to the OBEP

For the OBEP, introduced in Section 1.3.1, we consider the state space

$$\Omega_n := \{0,1\}^{[n]}.$$

In the OBEP a particle at site x attempts to jump right to site $x+1$ at rate $p \in (0,1)$, only doing so if $x < n$ and $x+1$ is not occupied by another particle. A particle at site x attempts to jump left at rate $q = 1-p$, only doing so if $x > 1$ and $x-1$ is not occupied by another particle. The endpoints of the segment are attached to infinite reservoirs which input and remove particles.

We think of site 1 as being attached to an infinite reservoir of particles which attempts to input particles at site 1 with rate $\alpha \geq 0$, only doing so if there is a hole at site 1, and attempts to remove particles from site 1 with rate $\gamma \geq 0$, only doing so if there is a particle at site 1. Similarly site N is attached to an infinite reservoir which inputs and removes particles at site N with rates $\delta \geq 0$ and $\beta \geq 0$ respectively. We say that the OBEP has parameters $(p, \alpha, \beta, \gamma, \delta)$.

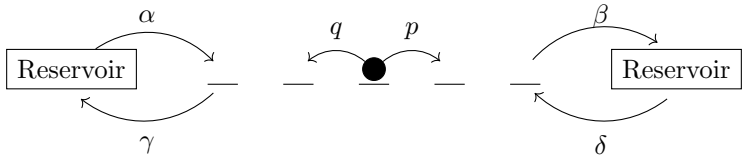


Figure 2.8: Transition rates for the OBEP on Ω_5.

In this section we relate the FSEP on the segment, started from the special configurations $\eta^-, \eta^+ \in \Omega_{N,k}$ from Definition 2.3.2, to an OBEP with a single open boundary.

We first give an equivalent description for the dynamics of the OBEP with a single open boundary using lattice paths. Let N and k be given with $N > k/2$, and let $(q, 0, \beta, 0, \delta)$ be the parameters of the OBEP. For $\zeta \in \Omega_{k-1}$, we define the height function h^ζ recursively. Let $h^\zeta(1) = 0$ and, given $h^\zeta(i)$ for $i \in [k-1]$, set

$$h^\zeta(i+1) = h^\zeta(i) + \zeta(i) - \frac{N-k}{k-1}. \tag{2.3.6}$$

See Figure 2.9a for an example. The mapping is injective and, with an abuse of notation, we denote the image set by Ω_{k-1}.

We note that we have included N in the height function, when N has no relation to the OBEP on Ω_{k-1}. This is to more easily relate the height functions in (2.3.6) to the lattice paths (2.3.1) for the FSEP in Section 2.3.1. In particular, this description of the OBEP is isomorphic to the height function in [41, Section 2.1.2]. We now give a graphical construction for the OBEP. Let

$$I'_i = -\frac{(N-k)(i-1)}{k-1} + \mathbb{Z} \cap [0, i-1],$$

for $i = 2, 3, \ldots, k$.

- To each pair (i, z) with $i \in \{2, \ldots, k-1\}$ and $z \in I'_i$ we attach independent Poisson clocks $T^{\uparrow}_{(i,z)}$ and $T^{\downarrow}_{(i,z)}$ of rate p and q respectively.

- To each pair (k, z) with $z \in I'_k$ we attach independent Poisson clocks $T^{\uparrow}_{(k,z)}$ and $T^{\downarrow}_{(k,z)}$ of rate δ and β respectively.

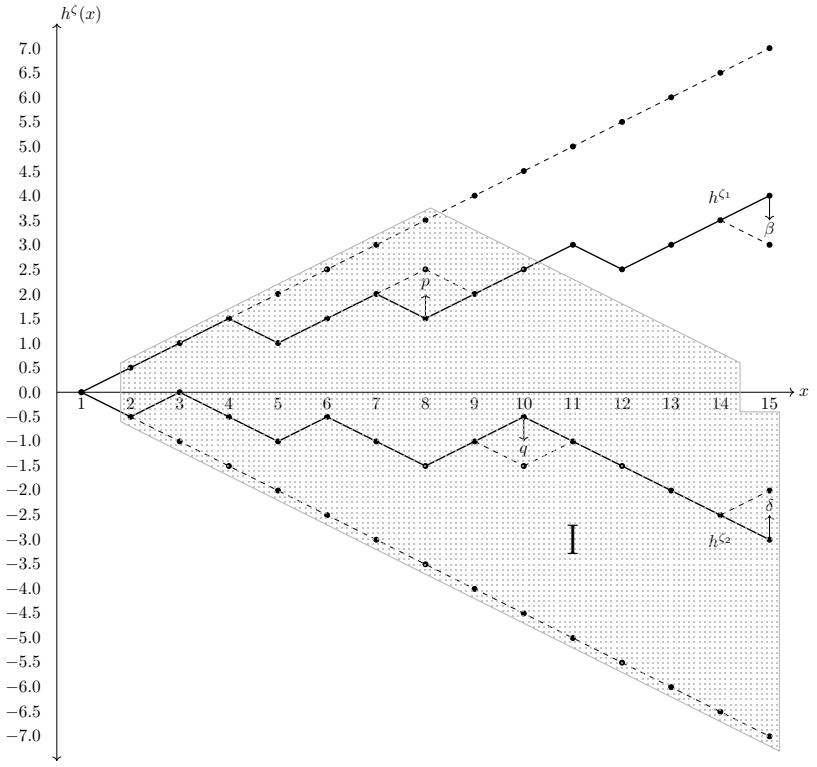
The lattice paths trajectories $(h^\zeta_t)_{t \geq 0}$ are then updated according to the same rules as for the graphical construction for the FSEP $\mathbb{P}$, see Figure 2.9 for an example. We let $\mathbb{P}_{RB}$ denote the law of the graphical construction for the OBEP on Ω_{k-1} with parameters $(q, 0, \beta, 0, \delta)$, i.e. with open *right boundary*.

By symmetry for the OBEP on Ω_{k-1} with parameters $(q, \alpha, 0, \gamma, 0)$ we may consider the reflection of the height functions in the line $z = (k-1)/2$ given by $\tilde{h}(i) := h(k-i)$ and the reflected index sets $\widetilde{I}'_i := I'_{k-i}$.

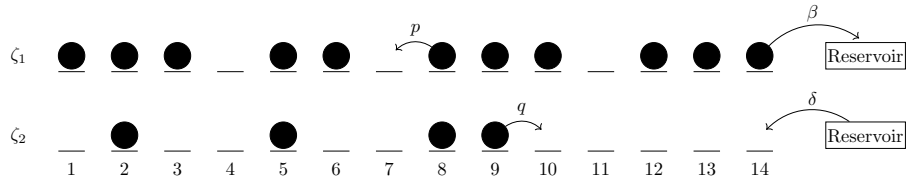
- To each pair (i, z) with $i \in \{2, \ldots, k-1\}$ and $z \in \widetilde{I}'_i$ we attach independent Poisson clocks $T^{\uparrow}_{(i,z)}$ and $T^{\downarrow}_{(i,z)}$ of rate p and q respectively.

- To each pair $(1, z)$ with $z \in \widetilde{I}'_1$ we attach independent Poisson clocks $T^{\uparrow}_{(1,z)}$ and $T^{\downarrow}_{(1,z)}$ of rate α and γ respectively.

We denote the law of the graphical construction for the OBEP on Ω_{k-1} with parameters $(q, \alpha, 0, \gamma, 0)$, i.e. with open *left boundary*, by $\mathbb{P}_{LB}$.



(a) The graphical construction with $k = 15$ and $N = 22$ for the OBEP on Ω_{14} with parameters $(q, 0, \beta, 0, \delta)$. The upper and lower dashed lines correspond to the minimal and maximal lattice paths for the OBEP on Ω_{14} respectively. The two bold lattice paths h^{ζ_1} and h^{ζ_2} correspond to the configurations ζ_1 and ζ_2 in Figure 2.9b respectively. Some allowed transitions in the lattice path dynamics are indicated; the rate p, q, β and δ transitions correspond to the clock rings $T^{\uparrow}_{(8,1.5)}, T^{\downarrow}_{(10,-0.5)}, T^{\downarrow}_{(15,4)}$ and $T^{\uparrow}_{(15,-3)}$ in the graphical construction $\mathbb{P}_{RB}$ respectively.



(b) The configuration $\zeta_1, \zeta_2 \in \Omega_{14}$ which give rise to the height functions h^{ζ_1} and h^{ζ_2} in Figure 2.9a. The transitions of rate p, q, β and δ in the lattice paths are indicated for the OBEP configurations.

Figure 2.9: If $\delta = p$ and $\beta = 0$ then each of the Poisson clocks, $\mathcal{T}^{\uparrow}_{(i,z)}$ and $\mathcal{T}^{\downarrow}_{(i,z)}$,whose index (i, z) is contained in the shaded region I in Figure 2.9a, have the same rate as the Poisson clocks $\mathcal{T}^{\uparrow}_{(i,z)}$ and $\mathcal{T}^{\downarrow}_{(i,z)}$ in the graphical construction $\mathbb{P}$ on $\Omega_{N,k}$. The shaded region I (strictly) contains all possible points in the lattice paths belonging to the set $\{h \in \Omega_{k-1} : h(1) = 0, h(k) \leq -1\}$.

It is straightforward to check that the above construction gives the dynamics of the OBEP, and we may consider the dynamics on lattice paths.

For the graphical construction $\mathbb{P}_{RB}$ of the OBEP with parameters $(q, 0, 0, 0, p)$, i.e. $\beta = 0$ and $\delta = p$, let (i, z) be a point in a lattice path $h \in \Omega_{k-1}$ with $h(1) = 0$ and $h(k) < 0$. We observe that the clock processes in $\mathbb{P}_{RB}$ with index (i, z) have the same rate as the clock processes with index (i, z) in the graphical construction $\mathbb{P}$ on $\Omega_{N,k}$ for the FSEP with parameter p. Moreover, each FSEP lattice path with $\eta(0) = 0$ and $\eta(k) < 0$ and no 'steep' segments may be uniquely identified with an OBEP lattice path. See the shaded region I in Figure 2.9.

We note that that $\eta^- \in \Omega_{N,k} \cap \Omega_{k-1}$ from Definition 2.3.2 is the minimal lattice on both Ω_{k-1} and $\Omega_{N,k}$. Let $\eta \in \Omega_{N,k} \cap \Omega_{k-1}$ with $\eta(1) = 0$ and $\eta(k) < 0$, and define

$$\tau_{N-k} := \inf\{t \geq 0 : \eta_t^-(k) = 0\},$$

i.e. the first time the right end of the height function hits 0 started from $k - N$. For particle configurations, this corresponds to the time it takes $N-k$ particles to enter at the right boundary in the OBEP. For $t \geq 0$ it holds that

$$\mathbb{P}(\eta_t^- \notin \mathcal{E}_{N,k}) = \mathbb{P}_{RB}(\tau_{N-k} > t). \tag{2.3.7}$$

Remark 2.3.4. *The trajectory of the FSEP $(\eta_t^-)_{t\geq 0}$ with parameter $p \geq 1/2$ on $\Omega_{N,k}$ is equivalent to the trajectory of an OBEP on Ω_{k-1} with parameters $(q, 0, 0, 0, p)$ and empty initial condition, up until the time that the right boundaries of the lattice paths reach zero. For the FSEP this represents the hitting time of the ergodic component, and for the OBEP this corresponds to the first time that $N - k$ particles have entered at the right boundary. In particular $(\eta_t^-)_{t\geq 0}$ is equivalent to an OBEP with a finite reservoir, containing $N - k$ particles, attached to the right boundary.*

By symmetry the FSEP $(\eta_t^+)_{t\geq 0}$ with parameter p is equivalent to the trajectory of an OBEP on Ω_{k-1} with parameters $(q, 0, 0, q, 0)$ and full initial condition $\mathbf{1}$ up until the time that $N - k$ particles have exited at the left boundary.

Recalling Section 2.2, the following figure gives the intuition for the relation between the FSEP with initial condition ξ^- and the OBEP when we view the FSEP as having particle–hole objects.

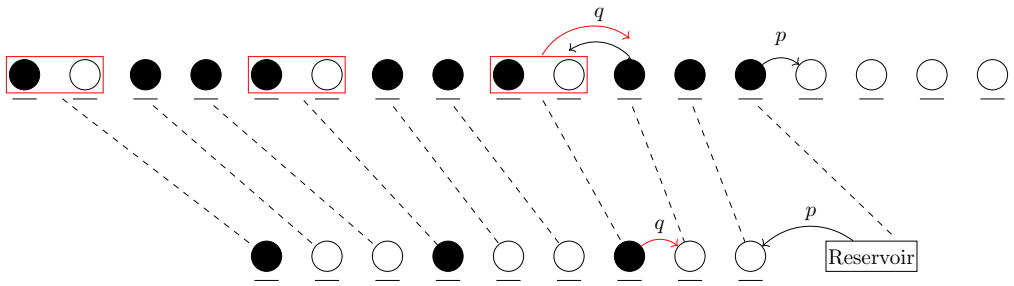


Figure 2.10: The top figure gives an FSEP configuration in $\Omega_{17,10}$ that can be reached from the initial condition ξ^-, and the bottom figure gives the corresponding OBEP configuration in Ω_9. The dashed lines indicate the role that each object in the OBEP plays in the FSEP picture: Each particle in the OBEP plays the role of a particle–hole object in the FSEP, each hole in the OBEP plays the role of a particle which is not followed by a hole in the FSEP and the reservoir plays the role of the rightmost particle in the FSEP. In this case the reservoir can be viewed as containing 4 particles. When the rightmost particle in the OBEP jumps right (at rate q), the corresponding particle–hole object jumps one step to the right. When the reservoir inputs a particle (at rate p) at the right boundary in the OBEP, the rightmost particle in the FSEP jumps right and creates another particle–hole object behind the rightmost particle.

As for the graphical construction $\mathbb{P}$, the graphical construction $\mathbb{P}_{RB}$ preserves the natural partial order (2.3.3) on lattice paths, and the lattice path dynamics are monotone. We say $h \leq h'$ if $h(i) \leq h'(i)$ for all $i \in [k]$. If $h \leq h'$ for two lattice paths $h, h' \in \Omega_{k-1}$ then

$$\mathbb{P}_{RB}\big(h_t \leq h'_t\big) = 1 \tag{2.3.8}$$

The proof of (2.3.8) is analogous the proof of Proposition 2.3.3. See also [41, Lemma 2.3]. Similarly, the graphical construction $\mathbb{P}_{LB}$ preserves the partial order on lattice paths and the lattice path dynamics are monotone.

The following lemma is a consequence of (2.3.8) and [70, Corollary 5.5].

Lemma 2.3.5. *Let η^- and h^+ denote the minimal and maximal lattice paths w.r.t. the partial order (2.3.3) for the OBEP on Ω_{k-1} with parameters $(p, 0, \beta, 0, \delta)$ respectively. Let $T^{k-1}_{\text{mix}}(\epsilon)$ denote its ϵ–mixing time and define*

$$\tau_{\text{coalesce}} := \inf\big\{t \geq 0 : \eta^-_t = h^+_t\big\},$$

to be the coalescence time of the minimal and maximal height functions. If

$$\mathbb{P}_{RB}\big(\tau_{\text{coalesce}} \geq s\big) \leq \epsilon,$$

then

$$T^{k-1}_{\text{mix}}(\epsilon) \leq s.$$

By symmetry, an analogous result holds for the OBEP with parameters $(p, \alpha, 0, \gamma, 0)$ and graphical construction $\mathbb{P}_{LB}$.

2.3.4 Proofs

In this section we prove Theorem 2.1 and Theorem 2.2, beginning with the upper bound in (2.1.3).

The idea for the proof of the upper bound in (2.1.3) is to bound the probability that η_t^+ and η_t^- have not coalesced under $\mathbb{P}$. We split the coalescence time into two parts: The time until both η_t^+ and η_t^- have arrived at the ergodic component, and the remaining coalescence time. We bound the hitting time of the ergodic component using the coupling to the OBEP in the previous section. Finally we bound the remaining coalescence time by the coalescence time of $\eta_t^\wedge$ and $\eta_t^\vee$, which is obtained via Proposition 1.6.1.

Proof of upper bound in (2.1.3). Let η_t, η_t' be given for some arbitrary configurations $\eta, \eta' \in \Omega_{N,k}$, then by Proposition 2.3.3 both η_t and η_t' are paths which are squeezed between η_t^- and η_t^+ under the graphical construction $\mathbb{P}$. Therefore

$$\mathbb{P}\big(\eta_t \neq \eta_t'\big) \leq \mathbb{P}\big(\eta_t^- \neq \eta_t^+\big),$$

as if η_t^- and η_t^+ agree then so must η_t and η_t'.

By an application of the triangle inequality, for all configurations $\eta \in \Omega_{N,k}$ we have that

$$\begin{aligned}\big\|P_t^\eta - \mu_{N,k}\big\|_{\mathrm{TV}} &= \big\|P_t^\eta - P_t^{\mu_{N,k}}\big\|_{\mathrm{TV}} \\ &\leq \frac{1}{|\Omega_{N,k}|}\sum_{\eta' \in \Omega_{N,k}} \big\|P_t^\eta - P_t^{\eta'}\big\|_{\mathrm{TV}} \leq \max_{\eta' \in \Omega_{N,k}} \big\|P_t^\eta - P_t^{\eta'}\big\|_{\mathrm{TV}}.\end{aligned}$$

An application of [70, Proposition 4.7] gives that

$$\big\|P_t^\eta - P_t^{\eta'}\big\|_{\mathrm{TV}} \leq \mathbb{P}\big(\eta_t \neq \eta_t'\big) \leq \mathbb{P}\big(\eta_t^- \neq \eta_t^+\big),$$

so that

$$d_{N,k}(t) = \max_{\eta \in \Omega_{N,k}} \big\|P_t^\eta(\cdot) - \mu_{N,k}(\cdot)\big\|_{\mathrm{TV}}$$

is bounded by the RHS. This reasoning was used by Wilson [106] to obtain an upper bound on the mixing time of the simple exclusion process.

Let

$$\tau = \inf\big\{t \geq 0 : \eta_t^-, \eta_t^+ \in \mathcal{E}_{N,k}\big\}.$$

If $t = t_1 + t_2$ then

$$\begin{aligned}\mathbb{P}\big(\eta_t^- \neq \eta_t^+\big) &\leq \mathbb{P}\big(\tau > t_1\big) + \mathbb{P}\big(\eta_t^- \neq \eta_t^+ \mid \tau < t_1\big) &(2.3.9)\\ &\leq 2\mathbb{P}\big(\eta_{t_1}^- \notin \mathcal{E}_{N,k}\big) + \mathbb{P}\big(\eta_{t_2}^\vee \neq \eta_{t_2}^\wedge\big), &(2.3.10)\end{aligned}$$

where (2.3.9) follows from the fact that $\{\tau > t_1\} \subseteq \{\eta_t^- \neq \eta_t^+\}$ and application of the total law of probability. The first term in (2.3.10) follows by union bound and symmetry, and the second term follows from an application of the Strong Markov property and the partial order (2.3.3) on $\mathcal{E}_{N,k}$.

We now bound the first term on the RHS of (2.3.10). Recalling Remark 2.3.4, $(\eta_t^-)_{t\geq 0}$ is equivalent to an OBEP on Ω_{k-1} with parameters $(q, 0, 0, 0, p)$ and initial condition given by the lattice path η^-. By (2.3.7)

$$\mathbb{P}\big(\eta_t^- \notin \mathcal{E}_{N,k}\big) = \mathbb{P}_{RB}\big(\tau_{N-k} > t\big),$$

i.e. the probability at most $N-k-1$ particles have entered at the right boundary in the OBEP.

For $\ell = N - k$ and $M = k - 1$, we bound the time taken for ℓ particles to enter an OBEP $(\zeta_t)_{t\geq 0}$ on Ω_M with parameters $(1/2, \alpha, 0, 0, 1/2)$ with $0 \leq \alpha \leq 1/2$. We assume the OBEP has

an initial condition ζ with at least ℓ holes, and we let P_t^ζ denote the law of the OBEP at time t with initial condition ζ.

The following argument may be found in [41, Section 4.1]. Since $p = 1/2$ and objects of the same type are indistinguishable, we can describe the dynamics of the OBEP along edges $\{x, x+1\}$, for $x \in [M-1]$, so that the values at the endpoints x and $x+1$ are swapped independently at rate $1/2$. In particular holes perform continuous time simple random walks with absorption at the left and right boundary at rate α and $1/2$ respectively.

Labelling holes $i = 1, 2, \dots$ in ζ from left to right, let H_i denote the event that the i–th hole is absorbed at a boundary. Using a comparison to the Gambler's ruin problem on $[M]$, with reflection at the left boundary, we see that a given hole reaches site M (or is absorbed) by time $2M^2$ with probability at least $1/2$. This bound does not depend on the starting position of the given hole. A hole at site M at time t is absorbed by time $t+1$ at the right boundary with probability at least $(1-e^{-1/2})e^{-1/2}$, and it follows that

$$P^\zeta_{2M^2+1}\big(H_i^c\big) \le 1 - \frac{1-e^{-1/2}}{2e^{1/2}},$$

for all $i \in [M]$. By an application of the Markov property, if

$$t = n\frac{2e^{1/2}}{1-e^{-1/2}}(2M^2+1),$$

for $n \in \mathbb{N}$, then

$$P_t^\zeta\big(H_i^c\big) \le \left(1 - \frac{1-e^{-1/2}}{2e^{1/2}}\right)^{n\frac{2e^{1/2}}{1-e^{-1/2}}} \le e^{-n}, \tag{2.3.11}$$

for all $i \in [M]$ and $n \in \mathbb{N}$. Let $H(\ell)$ be the event that at least ℓ holes have been absorbed at the boundary, then

$$H(\ell)^c \subseteq \bigcup_{i=1}^{\ell} H_i^c, \tag{2.3.12}$$

since the LHS of (2.3.12) is the event that at most $\ell - 1$ holes have been absorbed, and it follows that at least one of any collection of ℓ holes is not absorbed. The above bounds do not depend on the starting positions of the holes, and we may choose the collection of ℓ holes to be holes labelled $1, \dots, \ell$. The event on the RHS of (2.3.12) is the event that at least one of the of the ℓ leftmost holes has not been absorbed at the boundary.

By (2.3.11) and union bound it holds that

$$P_t^\zeta\big(H(\ell)^c\big) \le \sum_{i=1}^{\ell} P_t^\zeta\big(H_i^c\big) \le \ell e^{-n}.$$

In particular, for $\epsilon > 0$

$$P^\zeta_{CM^2\log\ell}\big(H(\ell)^c\big) \le \epsilon, \tag{2.3.13}$$

for all M sufficiently large, where the constant $C > 0$ (not depending on ϵ) is chosen appropriately.

By application of (2.3.13) there exists a constant $C > 0$, not depending on ϵ, such that

$$\mathbb{P}_{RB}\big(\tau_{N-k} > Ck^2\log(N-k)\big) = \mathbb{P}\big(\eta^-_{Ck^2\log(N-k)} \notin \mathcal{E}_{N,k}\big) \le \epsilon/3, \tag{2.3.14}$$

for N sufficiently large.

We now bound the second term on the RHS of (2.3.10). Let $\epsilon > 0$ be given, if

$$t = \frac{k^2}{\pi^2}(1+\delta)\log\min(N-k, 2k-N),$$

for some $\delta > 0$, then by application of Proposition 1.6.1 and Remark 2.3.1

$$\mathbb{P}\big(\eta_t^{\wedge} \neq \eta_t^{\vee}\big) \leq \epsilon/3, \tag{2.3.15}$$

for all N sufficiently large. Combining (2.3.14) and (2.3.15) we see that

$$\limsup_{N\to\infty} \frac{T_{\text{seg}}^{N,k}(\epsilon)}{N^2\log(N-k)} \leq C_2,$$

for some constant $C_2 > 0$, which completes the proof. □

We now prove the lower bound in Theorem 2.1. Here we compare the FSEP on the ergodic component directly to the SSEP to obtain a lower bound on the mixing time. In the regime $\log(2k-N) \ll \log(N-k)$ the bound we obtain from the SSEP is too small, and we show that the hitting time of the ergodic component determines the mixing time using the comparison to the OBEP in Section 2.3.3.

Proof of lower bound in (2.1.3). We note that

$$d_{N,k}(t) \geq \max_{\eta\in\mathcal{E}_{N,k}} \big\|P_t^{\eta} - \mu_{N,k}\big\|_{\text{TV}},$$

By Remark 2.3.1 the FSEP on $\Omega_{N,k}$ is equivalent to a SSEP on $\Omega_{k-1,N-k}$, and by an application of [106, Lemma 5] (see also [63, Section 7]), it holds that

$$d_{N,k}\Big(\frac{1}{\pi^2}k^2(1-\delta)\log\min(N-k,2k-N)\Big) \geq \epsilon, \tag{2.3.16}$$

for all $\epsilon \in (0,1)$, $\delta > 0$ and N sufficiently large. This is sufficient to give the lower bound in (2.1.3) when $\log(2k-N)$ has the same order as $\log(N-k)$.

Assume now that $\log(2k-N) \ll \log(N-k)$, then the lower bound on the mixing time given by (2.3.16) is too small. We show in this case that the hitting time of the ergodic component determines the mixing time. By Remark 2.3.4 we consider the time for $N-k$ particles to enter the OBEP on Ω_{k-1} with parameters $(1/2, 0, 0, 0, 1/2)$ and empty initial condition.

As this is an OBEP with one open boundary we wish to apply Theorem 1.5, see Section 1.6.1, to initially lower bound the probability that it takes longer than the leading order of the mixing time to completely fill the segment $[k-1]$. Technically Theorem 1.5 does not include the parameters we consider, but the lower bound holds nonetheless, see [41, Lemma 3.3].

By the contrapositive of Lemma 2.3.5 and [41, Lemma 3.3], if $\epsilon \in (0,1)$ and $0 < \delta < (1-\epsilon)/\epsilon$ then

$$\mathbb{P}_{RB}\Big(\tau_{\text{coalesce}} > \frac{2}{\pi^2}k^2\log k\Big) \geq \epsilon(1+\delta), \tag{2.3.17}$$

for all N sufficiently large.

We only require that $N-k$ particles have entered the OBEP. Recall that τ_{N-k} denotes the first time that $N-k$ particles have entered the OBEP, then $\tau_{\text{coalesce}} - \tau_{N-k}$ is the time for the

remaining $2k - N - 1$ particles to enter, or equivalently $2k - N - 1$ holes to exit at the right boundary. By application of the upper bound (2.3.13)

$$\mathbb{P}_{RB}\big(\tau_{\text{coalesce}} - \tau_{N-k} > Ck^2 \log(2k - N)\big) \le \epsilon\delta, \tag{2.3.18}$$

for all N sufficiently large. By union bound

$$\begin{aligned}&\mathbb{P}_{RB}\Big(\tau_{\text{coalesce}} > \frac{2}{\pi^2}k^2 \log k\Big)\\ &\le \mathbb{P}_{RB}\big(\tau_{\text{coalesce}} - \tau_{N-k} > Ck^2 \log(2k - N)\big) + \mathbb{P}_{RB}\Big(\tau_{N-k} > \frac{2}{\pi^2}k^2 \log k - Ck^2 \log(2k - N)\Big),\end{aligned}$$

and it follows from (2.3.17) and (2.3.18) that

$$\mathbb{P}_{RB}\Big(\tau_{N-k} > \frac{2}{\pi^2}k^2 \log k - Ck^2 \log(2k - N)\Big) \ge \epsilon,$$

for all N sufficiently large. Noting that $\log(k) > \log(N - k) \gg \log(2k - N)$, there exists a constant $C_1 > 0$ (not depending on ϵ) with

$$C_1 k^2 \log(N - k) \le \frac{2}{\pi^2}k^2 \log k - Ck^2 \log(2k - N),$$

for all N sufficiently large. Therefore, by (2.3.7)

$$\mathbb{P}_{RB}\big(\tau_{N-k} > C_1 k^2 \log(N - k)\big) = \mathbb{P}\big(\eta^-_{C_1 k^2 \log(N-k)} \notin \mathcal{E}_{N,k}\big) \ge \epsilon.$$

Since $d_{N,k}(t) \ge \mathbb{P}(\eta^-_t \notin \mathcal{E}_{N,k})$ the lower bound in (2.1.3) follows. □

We now prove the upper bound in Theorem 2.2. The idea here is that the FSEP on $\Omega_{N,k}$ with initial condition η^- behaves like the reverse–bias phase for OBEP on Ω_{k-1}, and this determines the mixing time.

Proof of upper bound in (2.1.4). By the same reasoning in the proof of the upper bound in (2.1.3), if

$$\tau = \inf\big\{t \ge 0 : \eta^-_t, \eta^+_t \in \mathcal{E}_{N,k}\big\},$$

and $t = t_1 + t_2$ then

$$\mathbb{P}\big(\eta^-_t \ne \eta^+_t\big) \le \mathbb{P}\big(\tau > t_1\big) + \mathbb{P}\big(\eta^-_t \ne \eta^+_t \mid \tau < t_1\big) \tag{2.3.19}$$

$$\le \mathbb{P}\big(\eta^+_{t_1} \notin \mathcal{E}_{N,k}\big) + \mathbb{P}\big(\eta^-_{t_1} \notin \mathcal{E}_{N,k}\big) + \mathbb{P}\big(\eta^\vee_{t_2} \ne \eta^\wedge_{t_2}\big), \tag{2.3.20}$$

where (2.3.19) follows from the fact that $\{\tau > t_1\} \subseteq \{\eta^-_t \ne \eta^+_t\}$ and application of the total law of probability. The first and second term in (2.3.20) follows by union bound, and the third term follows from an application of the Strong Markov property and the partial order (2.3.3) on $\mathcal{E}_{N,k}$.

To bound the third term on the RHS of (2.3.20) we apply the bound (1.6.3) from Section 1.6. In particular, if $\lambda = p/q$, $\varrho = (\sqrt{p} - \sqrt{q})^2$ and

$$\gamma_{k-1} = 2\sqrt{pq}\Big[1 - \cos\Big(\frac{\pi}{k-1}\Big)\Big],$$

then

$$\begin{aligned}\mathbb{P}\big(\eta_t^{\vee} \neq \eta_t^{\wedge}\big) &\leq C_\lambda k \lambda^{k/2} \exp\big(-(\gamma_{k-1} + \varrho)t\big) \\ &\leq C_\lambda k \lambda^{k/2} \exp\Big(2t\sqrt{pq}\cos\Big(\frac{\pi}{k-1}\Big) - t\Big) \\ &\leq C_\lambda k \lambda^{k/2} \exp\Big((2\sqrt{pq} - 1)t\Big).\end{aligned}$$

Since $p \neq 1/2$, $pq < 1/4$ and $2\sqrt{pq} < 1$. Choosing

$$t = 2\frac{k}{1 - 2\sqrt{pq}} \log \sqrt{\lambda},$$

gives that

$$\mathbb{P}\big(\eta_t^{\vee} \neq \eta_t^{\wedge}\big) \leq C_\lambda k \times \exp\big(-k \log \sqrt{\lambda}\big) \to 0,$$

as $k \to \infty$. It follows that the coalescence time of $\eta_t^{\wedge}$ and $\eta_t^{\vee}$ is $O(C(p)k)$, and for $\epsilon \in (0,1)$ we have

$$\mathbb{P}\big(\eta_{N^2}^{\vee} \neq \eta_{N^2}^{\wedge}\big) \leq \epsilon/3, \tag{2.3.21}$$

for all N sufficiently large.

To bound the second term on the RHS of (2.3.20) we apply Theorem 1.6. Theorem 1.6 is for $\alpha > 0$, $\beta = 0$ and $\gamma, \delta \geq 0$, but the OBEP with parameters $(q, 0, 0, 0, q)$ is equivalent to an OBEP with parameters $(p, q, 0, 0, 0)$ by reflecting the segment.

By Remark 2.3.4 we may consider the first time for $N-k$ particles to enter an OBEP on Ω_{k-1} with parameters $(q, 0, 0, 0, q)$. This is certainly dominated by the time it takes $k-1$ particles to enter the OBEP, and by Theorem 1.6 it follows that

$$\mathbb{P}\big(\eta_{N^2}^{-} \notin \mathcal{E}_{n,k}\big) \leq \epsilon/3, \tag{2.3.22}$$

for all $\epsilon \in (0,1)$ and N sufficiently large.

To bound the first term on the RHS of (2.3.20), by Remark 2.3.4, we consider the time it takes $N-k$ particles to exit the left boundary for an OBEP on Ω_{k-1} with empty initial condition and parameters $(q, 0, 0, q, 0)$. We wish to apply Theorem 1.7, however the mixing time for the OBEP is exponentially slow in k in this case, and the bound is too large for our purposes. To fix this we use the following lemma to compare to an OBEP of size $N - k$.

Lemma 2.3.6. *Let $(\zeta_t)_{t\geq 0}$ and $(\zeta'_t)_{t\geq 0}$ denote two OBEPs with parameters $(q, 0, 0, q, 0)$ and full initial conditions on the configuration spaces Ω_{k-1} and Ω_{N-k} respectively. Let*

$$\tilde{\tau}_{N-k} = \inf\left\{t \geq 0 : \sum_{x=1}^{k-1} \zeta_t(x) = 2k - N - 1\right\},$$

$$\tilde{\tau}'_{N-k} = \inf\left\{t \geq 0 : \sum_{x=1}^{N-k} \zeta'_t(x) = 0\right\},$$

i.e. the first time at which $N-k$ particles have exited at the left boundary of Ω_{k-1} and Ω_{N-k} respectively. Then there exists a coupling P of $(\zeta_t)_{t\geq 0}$ and $(\zeta'_t)_{t\geq 0}$ such that

$$P\big(\tilde{\tau}_{N-k} \leq \tilde{\tau}'_{N-k}\big) = 1. \tag{2.3.23}$$

In particular Lemma 2.3.6 says that for an OBEP in which particles can only exit at one boundary, the time for $N-k$ particles to leave segment $[k-1]$ is stochastically smaller than the time to completely empty the segment $[N-k]$. We defer the proof of Lemma 2.3.6 until after the proof of the upper bound in (2.1.4).

Let $P_t^{\mathbf{1}}$ denote the law of $(\zeta'_t)_{t\geq 0}$ on the particle configurations Ω_{N-k} with full initial condition $\mathbf{1}$. By Remark 2.3.4 and Lemma 2.3.6

$$\begin{aligned}\mathbb{P}\big(\eta_t^+ \notin \mathcal{E}_{N,k}\big) &= \mathbb{P}_{RB}\big(\tilde{\tau}_{N-k} > t\big)\\ &\leq P_t^{\mathbf{1}}\big(\tilde{\tau}'_{N-k} > t\big).\end{aligned}$$

Let $\delta_{\mathbf{0}}$ denote the Dirac measure on the empty particle configuration $\mathbf{0} \in \Omega_{N-k}$ then

$$\begin{aligned}P_t^{\mathbf{1}}\big(\tilde{\tau}'_{N-k} > t\big) &= P_t^{\mathbf{1}}\big(\Omega_{N-k}\setminus\{\mathbf{0}\}\big)\\ &= \big|P_t^{\mathbf{1}}\big(\Omega_{N-k}\setminus\{\mathbf{0}\}\big) - \delta_{\mathbf{0}}\big(\Omega_{N-k}\setminus\{\mathbf{0}\}\big)\big|\\ &\leq \big\|P_t^{\mathbf{1}} - \delta_{\mathbf{0}}\big\|_{\mathrm{TV}}\\ &\leq \max_{\zeta\in\Omega_{N-k}} \big\|P_t^{\zeta} - \delta_{\mathbf{0}}\big\|_{\mathrm{TV}},\end{aligned}$$

where the second inequality follows from the characterisation of the total variation in (1.5.1). By a comparison to the mixing time $T_{\mathrm{mix}}^{N-k}(\epsilon/3)$ for the OBEP on Ω_{N-k}, it follows from Theorem 1.7 that

$$\limsup_{N\to\infty} \frac{\log\big(\inf\{t\geq 0 : \mathbb{P}(\eta_t^+ \notin \mathcal{E}_{N,k}) \leq \epsilon/3\}\big)}{N-k} \leq \log\Big(\frac{p}{q}\Big), \tag{2.3.24}$$

for all $\epsilon \in (0,1)$. Therefore, for $\epsilon \in (0,1)$

$$\begin{aligned}T_{\mathrm{seg}}^{N,k}(\epsilon) &\leq \inf\big\{t\geq 0 : \mathbb{P}(\eta_t^- \neq \eta_t^+)\big\}\\ &\leq \inf\big\{t\geq 0 : \mathbb{P}(\eta_t^+ \notin \mathcal{E}_{N,k}) \leq \epsilon/3, \mathbb{P}(\eta_t^- \notin \mathcal{E}_{N,k}) \leq \epsilon/3, \mathbb{P}(\eta_t^\wedge \neq \eta_t^\vee) \leq \epsilon/3\big\}\\ &\leq \max\big(N^2, \inf\big\{t\geq 0 : \mathbb{P}(\eta_t^+ \notin \mathcal{E}_{N,k}) \leq \epsilon/3\big\}\big),\end{aligned}$$

for all N sufficiently large, where the second inequality follows from (2.3.20) and the third inequality follows from (2.3.21) and (2.3.22). By assumption

$$\lim_{N\to\infty} \frac{\log(N^2)}{N-k} = 0,$$

and the upper bound in (2.1.4) follows from (2.3.24). □

We now prove Lemma 2.3.6. The idea is to construct the dynamics of ζ'_t using an auxiliary height function and the clock processes in the graphical construction $\mathbb{P}_{LB}$.

Proof of Lemma 2.3.6. Given N and k, we consider the auxiliary height function

$$g(i) = N - k - \frac{(i-1)(N-k)}{k-1},$$

for all $i \in [N-k+1]$. We use the clock processes in $\mathbb{P}_{LB}$ to construct the trajectory $(g_t)_{t\geq 0}$. For $i < N-k+1$ the dynamics are the usual corner flip dynamics:

- if $g_{t-}(i)$ is a local minima and the clock $\mathcal{T}^{\uparrow}_{i,g_t(i)}$ rings at time t we set $g_t(i) = i+1$, making no changes to the other coordinates;

- if $g_{t-}(i)$ is a local maxima and the clock $\mathcal{T}^{\downarrow}_{i,g_t(i)}$ rings at time t we set $g_t(i) = i-1$, making no changes to the other coordinates.

In particular $g_t(N-k+1) = g(N-k+1)$ for all $t \geq 0$ and the auxiliary height function is frozen at its right endpoint $i = N-k+1$. See Figure 2.11 for an example.

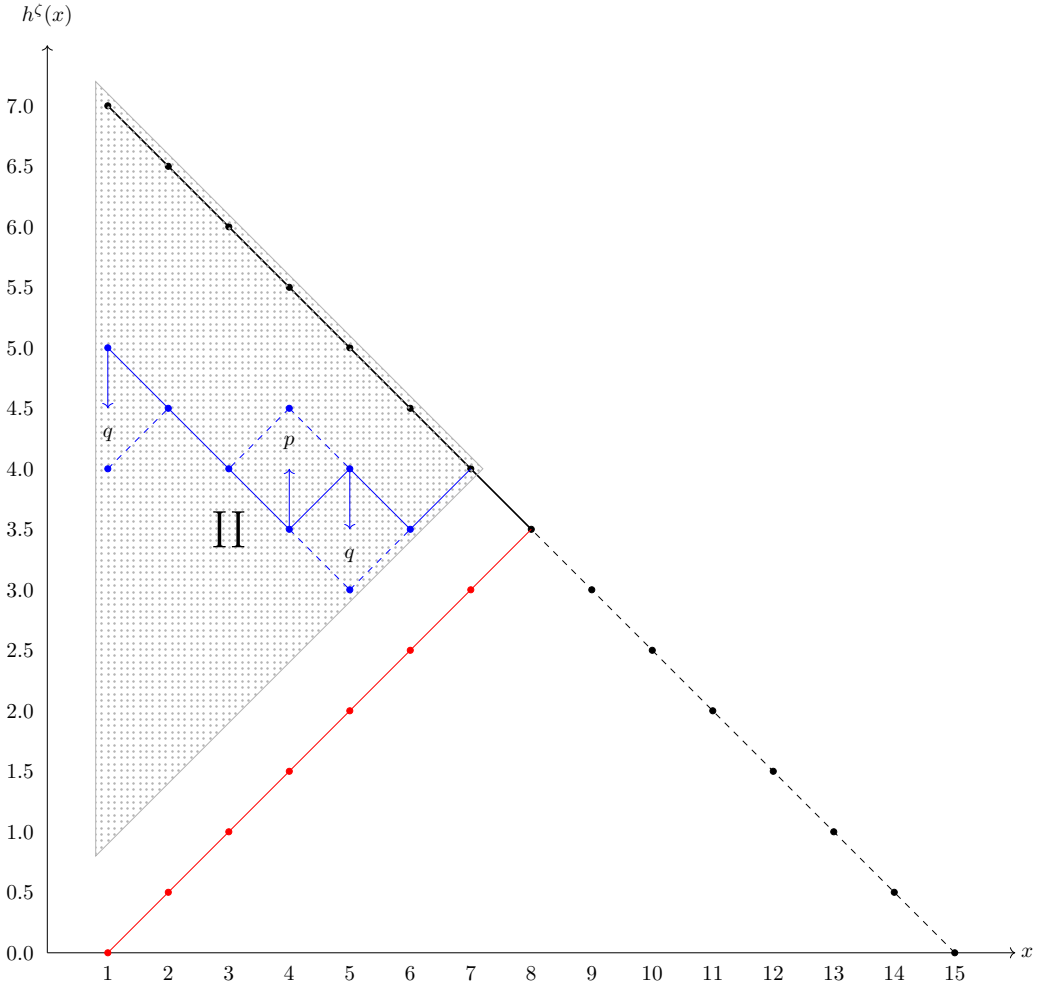


Figure 2.11: The bold black line gives the auxiliary height function g with $N = 22$ and $k = 15$, and the dashed black line is the part of the maximal height function for the OBEP on Ω_{14} with parameters $(q, 0, 0, q, 0)$, which is not contained in g. The bold red line gives the minimum auxiliary height function that may be obtained in the dynamics. In particular only the Poisson clocks whose indices are contained in the shaded region II may contribute to the dynamics of the auxiliary height function. The bold blue path together with the segment of the bold black path for $i \in [7, 8]$ give an auxiliary height function. The dashed blue lines indicate three possible transitions in the auxiliary dynamics for the blue lattice path.

It is straightforward to see that the dynamics of the auxiliary height function are equivalent (up to isomorphism) to the dynamics of the OBEP with maximal initial condition on Ω_{N-k}. In

particular this gives a coupling P of $(\zeta_t)_{t\geq 0}$ and $(\zeta'_t)_{t\geq 0}$, and it remains to check that $\tilde{\tau}_{N-k} \leq \tilde{\tau}'_{N-k}$ P–almost surely.

We note that $\eta^+ \in \Omega_{N,k} \cap \Omega_{k-1}$ from Definition 2.3.2 is the maximal lattice path for both the FSEP with parameters p and the OBEP with parameters $(q, 0, 0, q, 0)$. Noting that

$$\tilde{\tau}_{N-k} = \inf\big\{t :\geq 0 : \eta_t^+(1) = 0\big\},$$
$$\tilde{\tau}'_{N-k} = \inf\big\{t :\geq 0 : g_t(1) = 0\big\},$$

we compare the dynamics of the height functions $(\eta_t^+)_{t\geq 0}$ and $(g_t)_{t\geq 0}$. We say that $\eta_t^+ \leq g_t$ if $\eta_t^+(x) \leq g_t(x)$ for all $x \in [N-k+1]$. In particular (2.3.23) holds if

$$\mathbb{P}_{LB}\big(\eta_t^+ \leq g_t\big) = 1, \tag{2.3.25}$$

for all $t \geq 0$. The proof from here is analogous to the proof of Proposition 2.3.3; it is sufficient to check that (2.3.25) holds on clock rings.

Let $\eta_{t-}^+ \leq g_{t-}$ and assume WLOG that $\mathcal{T}_{(i,z)}^{\downarrow}$ rings for some $i \in [k]$. If $i \geq N-k+1$ then the partial order certainly holds at time t since the auxiliary function is undefined when $i > N-k+1$ and frozen when $i = N-k+1$.

If $\eta_{t-}^+(i) \leq g_{t-}(i) - 1$ then the partial order will also hold at time t, so assume further that $z = \eta_{t-}^+(i) = g_{t-}(i)$. The value $g_{t-}(i)$ is a local maximum if and only if $g_{t-}(i+1) < g_{t-}(i)$ and $g_{t-}(i-1) < g_{t-}(i)$ if $i \geq 2$, but $\eta_{t-}^+(i+1) \leq g_{t-}(i+1)$ and $\eta_{t-}^+(i-1) \leq g_{t-}(i-1)$ if $i \geq 2$ by the partial order at time $t-$. In particular $\eta_{t-}^+(i)$ must also be a local maximum and the partial order is conserved by the dynamics. □

We now prove the lower bound in Theorem 2.2 using a coupling to the zero–range process on the segment; see, for example, Section 1.4.1.

Proof of lower bound in (2.1.4). Let

$$\mathcal{H}_{N,k} = \big\{\xi \in \Omega_{N,k} : \xi(1) = 0, \xi(N) = 1,\, \xi(x) + \xi(x+1) \geq 1, \text{for all } x \in [N-1]\big\}.$$

The configurations $\mathcal{H}_{N,k}$ are almost in the ergodic component; if the particle at $x = 2$ jumps left in the dynamics the resulting configuration lies in the ergodic component. In particular $\mathcal{H}_{N,k}$ is the set of configurations with $\xi(1) = 0, \xi(2) = 1$ that may be reached by the dynamics with initial condition $\xi^+ \in \mathcal{E}_{N,k}^c$ from Definition 2.3.2.

Recall from Section 1.4.1 that the FSEP on the segment $[N]$ may be coupled to the zero–range process on the segment $[N-k+1]$. Since the number of particles is conserved, we let $\Sigma_{n,\ell}$ denote the set of zero–range configurations on the segment $[n]$ with ℓ particles.

We note that

$$\widetilde{\Pi}_{N,k}\big(\mathcal{H}_{N,k}\big) = \big\{\omega \in \Sigma_{N-k+1,k} : \omega(1) = 0,\, \omega(x) > 0, \text{ for all } x > 1\big\},$$

the set of zero–range configurations in which all sites except $x = 1$ are occupied. We may couple the zero–range process, with generator $\mathcal{L}_{\mathrm{ZS}}^{N-k+1}$ on the set $\widetilde{\Pi}_{N,k}(\mathcal{H}_{N,k})$ to the 'free' zero–range process in which there are no degenerate rates and particles cannot become stuck. The rate function of the free zero–range process is $g(\omega(x)) = \mathbb{1}(\omega(x) \geq 1)$ and we consider its dynamics on the set

$$\big\{\omega \in \Sigma_{n,\ell} : \omega(1) = 0\big\},$$

with $n = N - k + 1$ and $\ell = 2k - N$. The generator of the free zero–range $\mathcal{L}^n_{ZF}$ process acts on function $f : \mathbb{Z}_+^{[n]} \to \mathbb{R}$ via

$$\mathcal{L}^n_{\mathrm{ZF}} f(\omega) = \sum_{x=1}^{n-1} p \mathbb{1}\big(\omega(x) \geq 1\big)\big[f(\omega^{x,x+1}) - f(\omega)\big] + \sum_{x=2}^{n} q \mathbb{1}\big(\omega(x) \geq 1\big)\big[f(\omega^{x,x-1}) - f(\omega)\big]. \quad (2.3.26)$$

The coupling is achieved by deleting the stuck particles for configurations in $\widetilde{\Pi}_{N,k}(\mathcal{H}_{N,k})$ using the mapping $\Lambda : \widetilde{\Pi}_{N,k}(\mathcal{H}_{N,k}) \to \Sigma_{n,\ell}$ given by

$$\Lambda\big(\omega\big)(x) = \min(0, \omega(x) - 1).$$

See Figure 2.12 for an example.

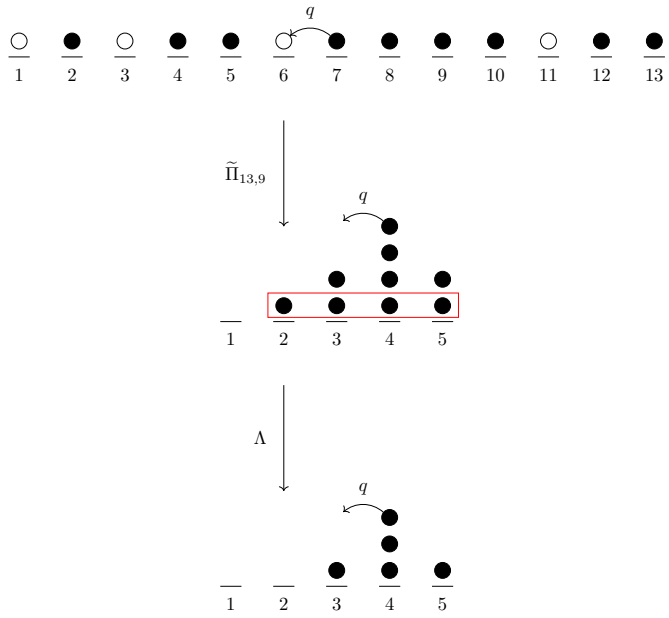


Figure 2.12: The top figure is a configuration $\xi \in \mathcal{H}_{13,9}$, the middle figure is the configuration $\omega = \widetilde{\Pi}_{13,9}(\xi) \in \Sigma_{5,9}$ and the bottom figure is the configuration $\Lambda(\omega) \in \Sigma_{5,5}$. The particles in ω contained in the red box are stuck in the dynamics of the zero–range process with generator $\mathcal{L}^5_{zs}$, and therefore the dynamics of the zero–range process $(\omega_t)_{t\geq 0}$ with initial condition ω are equivalent to the dynamics of the free zero–range process $(\omega'_t)_{t\geq 0}$ with initial condition $\Lambda(\omega)$ and generator $\mathcal{L}^n_{ZF}$ until the first time that a particle enters site 1 in both processes.

Let $\bar{Q}^\pi_t$ denote the law of the zero–range process on $\Sigma_{n,k}$ with initial distribution π and generator $\mathcal{L}^n_{ZS}$, and let $Q^{\pi'}_t$ denote the law of the zero–range process on $\Sigma_{n,\ell}$ with initial distribution π' and generator $\mathcal{L}^n_{ZF}$.

We consider the event

$$E = \big\{\omega \in \Sigma_{n,\ell} : \omega(1) > 0\big\}.$$

By the coupling between the FSEP and the zero–range process in Section 1.4.1, for all $\xi \in \mathcal{H}_{N,k}$ it holds that

$$P^\xi_t\big(\mathcal{E}^c_{N,k}\big) = \bar{Q}^{\widetilde{\Pi}_{N,k}(\xi)}_t\big(\widetilde{\Pi}_{N,k}(\mathcal{H}_{N,k})\big) = Q^{\Lambda\circ\widetilde{\Pi}_{N,k}(\xi)}_t\big(E^c\big).$$

Let Q_t^ν denote the law of the free zero–range process with initial distribution ν, and let ν be any distribution on $\Sigma_{n,\ell}$ supported on E^c. Let $\tilde{\mu}$ be the pushforward measure on $\mathcal{H}_{N,k}$ given by

$$\tilde{\mu}(A) = \nu\big(\Lambda \circ \widetilde{\Pi}_{N,k}(A)\big),$$

for $A \subseteq \mathcal{H}_{N,k}$, then it holds that

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}\big) = Q_t^\nu\big(E^c\big). \tag{2.3.27}$$

Let $\pi_{n,\ell}$ denote the stationary distribution of the free zero–range process on $[n]$ with ℓ particles. By (1.4.3), if $\lambda = q/p$ then

$$\pi_{n,\ell}(\omega) = \frac{1}{Z_{n,\ell}} \prod_{x=1}^{n} \lambda^{(n+1-x)\omega(x)}, \tag{2.3.28}$$

where

$$Z_{n,\ell} = \sum_{\omega \in \Sigma_{n,\ell}} \prod_{x=1}^{n} \lambda^{(n+1-x)\omega(x)},$$

is the normalisation constant.

We consider the conditional measure $\tilde{\pi}(\cdot) = \pi_{n,\ell}(\cdot \mid E^c)$. We will lower bound the hitting time of the ergodic component for the FSEP with initial measure

$$\tilde{\mu}(A) = \tilde{\pi}\big(\Lambda \circ \widetilde{\Pi}_{N,k}(A)\big),$$

for $A \subseteq \mathcal{H}_{N,k}$. By (2.3.27)

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}^c\big) = Q_t^{\tilde{\pi}}(E^c) = \frac{Q_t^{\pi_{n,\ell}}(E^c)}{\pi_{n,\ell}(E^c)}. \tag{2.3.29}$$

In order to give a lower bound on $Q_t^{\pi_{n,\ell}}(E^c)$ we employ a hitting time bound due to Aldous and Brown [2]. By an application of [2, Corollary 5]

$$Q_t^{\pi_{n,\ell}}(E^c) \geq \pi_{n,\ell}(E^c) \exp\left(-t \frac{q(E, E^c)}{\pi_{n,\ell}(E^c)}\right), \tag{2.3.30}$$

where $q(E, E^c)$ is the *capacity* given by

$$q(E, E^c) = \sum_{\omega \in E} \sum_{\omega' \in E^c} \pi_{n,\ell}(\omega) c(\omega, \omega'),$$

and $c(\omega, \omega')$ is the transition rate from ω to ω' in the free zero–range process. Here

$$q(E, E^c) = \pi_{n,\ell}\big(\omega : \omega(1) = 1\big) \times p.$$

Combining (2.3.29) and (2.3.30) we obtain

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}^c\big) \geq \exp\left(-t \frac{\pi_{n,\ell}\big(\omega : \omega(1) = 1\big) \times p}{\pi_{n,\ell}(E^c)}\right). \tag{2.3.31}$$

We now give an upper bound on $\pi_{n,\ell}(\omega : \omega(1) = 1)$ in order to lower bound the RHS of (2.3.31). Essentially the lower bound on the mixing time will follow from the product structure of the stationary distribution $\pi_{n,\ell}$ in (2.3.28). Strictly speaking $\pi_{n,\ell}$ is not a product measure because

of the normalisation factor $Z_{n,\ell}$, but we will remove the normalisation factor in the upper bound for $\pi_{n,\ell}(\omega : \omega(1) = 1)$. First note that

$$\begin{aligned}
\pi_{n,\ell}(\omega : \omega(1) = 1) &= \frac{1}{Z_{n,\ell}} \sum_{\substack{\omega \in \Sigma_{n,\ell} \\ \omega(1)=1}} \prod_{x=1}^{n} \lambda^{(n+1-x)\omega(x)} \\
&= \frac{1}{Z_{n,\ell}} \lambda^n \sum_{\substack{\omega \in \Sigma_{n,\ell} \\ \omega(1)=1}} \prod_{x=2}^{n} \lambda^{(n+1-x)\omega(x)} \\
&= \frac{1}{Z_{n,\ell}} \lambda^n \sum_{\omega \in \Sigma_{n-1,\ell-1}} \prod_{y=1}^{n-1} \lambda^{(n-y)\omega(y)} \\
&= \frac{1}{Z_{n,\ell}} \lambda^n Z_{n-1,\ell-1},
\end{aligned}$$

where the third equality follows from the fact that every configuration in $\omega \in \Sigma_{n,\ell}$ with $\omega(1) = 1$ may be uniquely identified with a configuration $\omega' \in \Sigma_{n-1,\ell-1}$ given by

$$\omega'(x) = \omega(x+1),$$

for all $x \in [n-1]$. Similarly we have that

$$\pi_{n,\ell}(\omega : \omega(1) = 0) = \frac{1}{Z_{n,\ell}} Z_{n-1,\ell}.$$

Therefore

$$\pi_{n,\ell}(\omega : \omega(1) = 1) \leq \frac{\pi_{n,\ell}(\omega : \omega(1) = 1)}{\pi_{n,\ell}(\omega : \omega(1) = 0)} = \lambda^{n-1} \frac{\lambda Z_{n-1,\ell-1}}{Z_{n-1,\ell}}.$$

We observe that

$$\begin{aligned}
\frac{\lambda Z_{n-1,\ell-1}}{Z_{n-1,\ell}} &= \frac{1}{Z_{n-1,\ell}} \lambda \sum_{\omega \in \Sigma_{n-1,\ell-1}} \prod_{x=1}^{n-1} \lambda^{(n-x)\omega(x)} \\
&= \frac{1}{Z_{n-1,\ell}} \sum_{\substack{\omega \in \Sigma_{n-1,\ell} \\ \omega(n-1)>0}} \prod_{x=1}^{n-1} \lambda^{(n-x)\omega(x)} \\
&= \pi_{n-1,\ell}(\omega : \omega(n-1) > 0) \leq 1,
\end{aligned}$$

where the second equality follows from the fact that every configuration $\omega \in \Sigma_{n-1,\ell-1}$ may be uniquely identified with a configuration in $\omega' \in \Sigma_{n-1,\ell}$ with $\omega'(n-1) > 0$ by adding a single particle to the site $n-1$ in ω.

In particular

$$\pi_{n,\ell}(\omega : \omega(1) = 1) \leq \lambda^{n-1}.$$

Moreover,

$$\pi_{n,\ell}(E^c) \geq 1 - \pi_{n,\ell}(\omega : \omega(1) = 1) \geq 1 - \lambda^{n-1},$$

and for $n-1 = N-k$ sufficiently large

$$1 - \Big(\frac{q}{p}\Big)^{N-k} \geq 1/2.$$

By (2.3.31) we have that

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}^c\big) \geq \exp\big(-2tp\lambda^{N-k}\big). \tag{2.3.32}$$

For $\epsilon \in (0,1)$ and

$$t \leq \frac{1}{2p}\log(\epsilon^{-1})\Big(\frac{p}{q}\Big)^{N-k},$$

it follows from (2.3.32) that

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}^c\big) \geq \epsilon.$$

Since

$$P_t^{\tilde{\mu}}\big(\mathcal{E}_{N,k}^c\big) \leq \big\|P_t^{\tilde{\mu}} - \mu_{N,k}\|_{\mathrm{TV}} \leq \sup_{\pi}\|P_t^{\pi} - \mu_{N,k}\|_{\mathrm{TV}} = d_{N,k}(t),$$

we may choose

$$t = \frac{1}{2p}\log(\epsilon^{-1})\Big(\frac{p}{q}\Big)^{N-k},$$

so that

$$T_{\mathrm{seg}}^{N,k}(\epsilon) \geq \frac{1}{2p}\log(\epsilon^{-1})\Big(\frac{p}{q}\Big)^{N-k},$$

for all $\epsilon \in (0,1)$. In particular

$$\liminf_{N\to\infty} \frac{\log T_{\mathrm{seg}}^{N,k}(\epsilon)}{N-k} \geq \log\Big(\frac{p}{q}\Big),$$

which completes the proof. □

2.4 FSEP on the circle

In this section we prove Theorem 2.3. In Section 2.4.1 we give a lower bound on the mixing time on the ergodic component, which lower bounds the full mixing time. In Section 2.4.2 we show how the upper bound on the mixing time follows from a bound on the hitting time of the ergodic component and a lower bound for the log–Sobolev constant restricted to the ergodic component. In Section 2.4.3 we show that the FSEP with initial condition in $\mathcal{I}_{N,k}^m$ lies in the ergodic component after time of order $N^2 \log N$ with high probability. The argument follows via comparisons to the zero–range process and OBEP.

In Section 2.4.4 we lower bound the log–Sobolev constant for the FSEP on the ergodic component $\mathcal{G}_{N,k}$ by the log–Sobolev of the SSEP on the *segment*, which has order N^{-2}. To do this we use the quasi–factorisation of the entropy in Proposition 1.5.8. In order to show that quasi–factorisation holds we show that the events in two sub–σ–algebras are asymptotically independent. A key ingredient in the proof of this asymptotic independence is a concentration of the number of particles in the half–circle. We show this concentration in Section 2.4.5 using a large deviation argument.

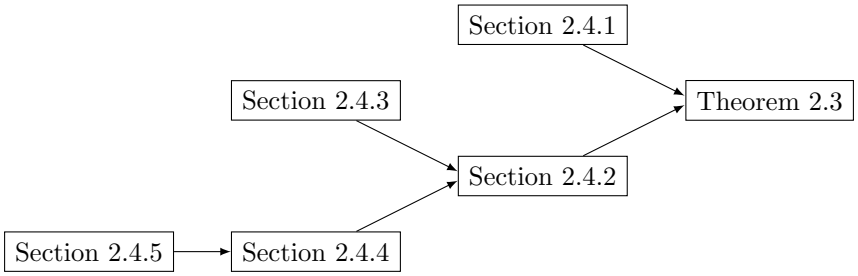


Figure 2.13: The graph of dependencies between each section in Section 2.4 w.r.t. the proof of Theorem 2.3.

2.4.1 Lower bound on the mixing time

In this section we give a proof of Theorem 2.3a. The idea of the proof is to write down an eigenfunction for the FSEP on the ergodic component $\mathcal{G}_{N,k}$ that exploits the fact that particle–hole objects move like exclusion particles (recall Section 2.2). We then duplicate the relaxation of the 'first' Fourier coefficient argument for the SSEP on the circle [64][Proposition 2.1].

Proof of Theorem 2.3a. We consider the dynamics of the FSEP on the ergodic component, which will provide a lower bound for the mixing time. Let $\xi \in \mathcal{G}_{N,k}$ and assume that $N - k \leq k/2$, i.e. $k \geq 2N/3$. Labelling particles $i = 0, \dots, k-1$ in the clockwise direction on the circle, beginning at site 0, we let x_i denote the position of the i-th particle in ξ for $i \in \mathbb{T}_k$.

We consider the function given by

$$a_1(\xi) = \sum_{i \in \mathbb{T}_k} \mathbb{1}\big(\xi(x_i + 1) = 0\big) \cos\Big(\frac{2\pi i}{k}\Big).$$

This will be an eigenfunction for the generator $\mathcal{L}^\circ$ restricted to the ergodic component $\mathcal{G}_{N,k}$.

We consider the SEP configuration $\zeta \in \Omega'_{k,N-k}$ given by

$$\zeta(i) = \mathbb{1}\big(\xi(x_i + 1) = 0\big),$$

for $0 \le i \le k-1$. With an abuse of notation we set $a_1(\zeta) = a_1(\xi)$, and observe that

$$\begin{aligned}
\mathcal{L}^\circ a_1(\xi) =& \sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x-1)\big(1-\xi(x+1)\big)\big[a_1\big(\xi^{x,x+1}\big) - a_1\big(\xi\big)\big] \\
&+ \sum_{x\in\mathbb{T}_N} \frac{1}{2}\xi(x)\xi(x+1)\big(1-\xi(x-1)\big)\big[a_1\big(\xi^{x,x-1}\big) - a_1\big(\xi\big)\big] \\
=& \sum_{i\in\mathbb{T}_k} \frac{1}{2}\mathbb{1}\big(\zeta(i)=1,\zeta(i+1)=0\big)\big[a_1\big(\zeta^{i,i+1}\big) - a_1\big(\zeta\big)\big] \\
&+ \sum_{i\in\mathbb{T}_k} \frac{1}{2}\mathbb{1}\big(\zeta(i)=0,\zeta(i+1)=1\big)\big[a_1\big(\zeta^{i,i+1}\big) - a_1\big(\zeta\big)\big] \\
=& \frac{1}{2}\sum_{i\in\mathbb{T}_k} a_1\big(\zeta^{i,i+1}\big) - a_1\big(\zeta\big) \\
=& \sum_{i\in\mathbb{T}_k} \zeta(i)\left[\cos\Big(\frac{2\pi(i+1)}{k}\Big) + \cos\Big(\frac{2\pi(i-1)}{k}\Big)\right] - a_1(\xi) \\
=& a_1(\xi)\Big[\cos\Big(\frac{2\pi}{k}\Big) - 1\Big].
\end{aligned}$$

where the second equality follows from the fact that $\xi \in \mathcal{G}_{N,k}$. Therefore $a_1(\xi)$ is an eigenfunction for the generator $\mathcal{L}^\circ$. We note the similarity of the third equality to the generator of the SSEP on the circle in (1.3.7).

The lower bound now follows directly from the argument for [64][Proposition 2.1]. We give the details of the argument below. Let

$$\lambda_1 = 1 - \cos\Big(\frac{2\pi}{k}\Big),$$

and

$$M_t := e^{t\lambda_1} a_1(\xi_t).$$

By the Markov property, for all positive t it holds that

$$\partial_s \mathbb{E}\big[M_{t+s} \mid \mathcal{F}_t\big] \mid_{s=0} = \lambda_1 M_t + e^{t\lambda_1}\big(\mathcal{L}^\circ a_1\big)\big(\xi_t\big) = 0,$$

so that M_t is a martingale for the filtration $\mathcal{F}$ defined by

$$\mathcal{F}_t := \sigma\big((\xi_s)_{s\le t}\big).$$

In particular it holds that

$$\mathbb{E}\big[a_1(\xi_t)\big] = e^{-t\lambda_1} a_1(\xi_0). \tag{2.4.1}$$

The maximal transition rate for the process is $N-k$, since there are $N-k$ holes and therefore at most $2(N-k)$ particles which can jump independently at rate $1/2$. If a transition occurs at time s the value of M_s varies by an amount which is at most

$$e^{\lambda_1 s} \max_{i\in\mathbb{T}_k} \Big|\cos\Big(\frac{2\pi(i+1)}{k}\Big) - \cos\Big(\frac{2\pi i}{k}\Big)\Big| \le e^{\lambda_1 s}\frac{2\pi}{k}. \tag{2.4.2}$$

Letting $\langle M\rangle_t$ be the predictable process which makes $M_t^2 - \langle M\rangle_t$ a martingale, (2.4.2) yields

$$\langle M\rangle_t \le (N-k)\int_0^t e^{2\lambda_1 s}\Big(\frac{2\pi}{k}\Big)^2 ds. \tag{2.4.3}$$

Using that $\mathrm{Var}(M_t) = \mathbb{E}[\langle M_t \rangle_t]$, for all $\xi \in \mathcal{G}_{N,k}$, $t \geq 0$ and N sufficiently large it holds that

$$\begin{aligned} \mathrm{Var}\big[a_1(\xi_t)\big] \leq e^{-2\lambda_1 t}\mathbb{E}\big[\langle M \rangle_t\big] &\leq (N-k)\int_0^t e^{2\lambda_1(s-t)}\Big(\frac{2\pi}{k}\Big)^2 ds \\ &\leq \frac{2\pi^2(N-k)}{k^2\lambda_1} \leq N-k, \end{aligned} \tag{2.4.4}$$

where the first inequality is an application of (2.4.3) and the second inequality follows from the fact that $\lambda_1 \asymp 2\pi^2 k^{-2}$.

Taking $t \to \infty$ in (2.4.4), we observe that the variance at equilibrium satisfies

$$\mathrm{Var}_\nu\big[a_1\big] \leq N-k. \tag{2.4.5}$$

By [70, Proposition 7.12] it holds that

$$\begin{aligned} \big\|P_t^\xi - \nu\big\|_{\mathrm{TV}} &\geq \frac{\big(\mathbb{E}\big[a_1(\xi_t)\big]\big)^2}{\big(\mathbb{E}\big[a_1(\xi_t)\big]\big)^2 + \big(\mathrm{Var}\big[a_1(\xi_t)\big] + \mathrm{Var}_\nu\big[a_1\big]\big)/2} \\ &\geq \frac{1}{1+(N-k)\exp(2\lambda_1 t)a_1(\xi)^{-2}}. \end{aligned} \tag{2.4.6}$$

We choose $\xi = \xi_0$ to be the configuration which maximises the eigenfunction $a_1(\xi)$, i.e.

$$\xi_0(x) = \begin{cases} \mathbb{1}\big(x \notin (2\mathbb{Z}+1) \cap [k+1-N, N-k-1]\big) & \text{if } N-k \in 2\mathbb{Z}, \\ \mathbb{1}\big(x \notin 2\mathbb{Z} \cap [k+2-N, N-k]\big) & \text{if } N-k \notin 2\mathbb{Z}. \end{cases}$$

It is straightforward to check that $a_1(\xi_0) \geq (N-k)/2$ for any $k \geq 2$ and $(N-k) \leq k/2$. Setting

$$t_N = uk^2 + \frac{k^2}{4\pi^2}\log(N-k),$$

for $u \in \mathbb{R}$ an application of (2.4.6) gives that

$$\liminf_{N\to\infty}\big\|P_{t_N}^{\xi_0} - \nu\big\|_{\mathrm{TV}} \geq \lim_{N\to\infty}\frac{1}{1+4(N-k)^{-1}\exp(-2\lambda_1 t_N)} = \frac{1}{1+\exp(-4\pi^2 \mathrm{u})},$$

which implies (2.1.5). If $N-k \geq k/2$, i.e. $2k-N \leq k/2$ we may set

$$a_1(\xi) = \sum_{i\in\mathbb{T}_k}\mathbb{1}\big(\xi(x_i+1)=1\big)\cos\Big(\frac{2\pi i}{k}\Big),$$

and repeat the argument above. □

2.4.2 Upper bound on the mixing time

In this section we prove Theorem 2.3b. Before proceeding with the proof we state two lemmas.

Lemma 2.4.1 gives that the process started from initial configurations in $\mathcal{I}_{N,k}^m$ will hit the ergodic component with high probability after a time of order $N^2 \log N$.

Lemma 2.4.1. *Let $\epsilon > 0$, $k(N)$ be a sequence, $m > 0$ a positive integer that does not depend on N. There exists a constant $C = C(m) > 0$ (not depending on ϵ) such that*

$$\max_{\xi \in \mathcal{I}^m_{N,k}} P^{\xi}_{C(m)N^2 \log N}\big(\mathcal{G}^c_{N,k}\big) < \epsilon, \tag{2.4.7}$$

for all N sufficiently large depending on ϵ.

We defer the proof of Lemma 2.4.1 to Section 2.4.3. Lemma 2.4.2 gives a lower bound on the log–Sobolev constant of the FSEP on the ergodic component $\mathcal{G}_{N,k}$.

Lemma 2.4.2. *Let $\epsilon \in (0,1)$ and $k(N)$ be such that $k(N)/N \to \alpha \in (1/2, 1)$ as $N \to \infty$. For N sufficiently large the log–Sobolev constant $\rho_{N,k}$ for the FSEP on $\mathcal{G}_{N,k}$ satisfies*

$$\rho_{N,k} \geq C' N^{-2}, \tag{2.4.8}$$

for some constant $C' > 0$ which does not depend on N.

We defer the proof of Lemma 2.4.2 to Section 2.4.4.

We now show how Lemma 2.4.1 and Lemma 2.4.1 quickly yield Theorem 2.3b.

Proof of Theorem 2.3b. For $\epsilon \in (0,1)$ we define

$$T_{\mathcal{G}_{N,k}}(\epsilon) = \inf\Big\{t \geq 0 : \max_{\xi \in \mathcal{G}_{N,k}} \big\|P^{\xi}_t - \nu_{N,k}\big\|_{\mathrm{TV}} \leq \epsilon\Big\}, \tag{2.4.9}$$

i.e. the ϵ–mixing time of the FSEP on the circle restricted to the ergodic component. By an application of (1.5.24) it holds that

$$T_{\mathcal{G}_{N,k}}(\epsilon) \leq \frac{\lceil \log_{e/2}(\epsilon^{-1}) \rceil}{4\rho_{N,k}} \Big(4 + \log^+ \log \big|\mathcal{G}_{N,k}\big|\Big). \tag{2.4.10}$$

To count the size of the set $\mathcal{G}_{N,k}$ we imagine a circle consisting of k particles, and we place $N-k$ holes between the particles on the circle, so that there is at most one hole between particles. Suppose first that we fix particle at site 1, then there are k places in which we may place $N-k$ holes and

$$\big|\{\xi \in \mathcal{G}_{N,k} : \xi(1) = 1\}\big| = \binom{k}{N-k}.$$

If we fix a hole at site 1, then there are $k-1$ places in which to place the remaining $N-k-1$ holes and

$$\big|\{\xi \in \mathcal{G}_{N,k} : \xi(1) = 0\}\big| = \binom{k-1}{N-k-1}.$$

In particular

$$\big|\mathcal{G}_{N,k}\big| = \binom{k}{N-k} + \binom{k-1}{N-k-1} = \frac{N}{k}\binom{k}{N-k}. \tag{2.4.11}$$

Using the standard bound

$$\binom{n}{\ell} \leq \left(\frac{e \cdot n}{\ell}\right)^{\ell},$$

and the fact that $N/k \leq 2$, it follows that

$$\big|\mathcal{G}_{N,k}\big| \leq 2\left(\frac{e\cdot k}{N-k}\right)^{N-k}.$$

In particular

$$\log^+\log\big|\mathcal{G}_{N,k}\big| = \log(N-k) + \log^+\log\left(2^{1/(N-k)}\frac{e\cdot k}{N-k}\right) = O(\log N), \tag{2.4.12}$$

as $N\to\infty$. By application of Lemma 2.4.2 and (2.4.12) to (2.4.10), it follows that

$$T_{\mathcal{G}_{N,k}}(\epsilon) = O(N^2\log N), \tag{2.4.13}$$

as $N\to\infty$.

Let $\xi \in \mathcal{I}^m_{N,k}\cup\mathcal{G}_{N,k}$ be arbitrary, then for $t = t_1+t_2$

$$\begin{aligned}
\big\|P^\xi_t - \nu_{N,k}\big\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{\xi'\in\Omega'_{N,k}}\big|P^\xi_t(\xi') - \nu_{N,k}(\xi')\big| \\
&= \frac{1}{2}P^\xi_t\big(\mathcal{G}^c_{N,k}\big) + \frac{1}{2}\sum_{\xi'\in\mathcal{G}_{N,k}}\big|P^\xi_t(\xi') - \nu_{N,k}(\xi')\big| && (2.4.14)\\
&\leq \frac{1}{2}P^\xi_t\big(\mathcal{G}^c_{N,k}\big) + \frac{1}{2}\sum_{\xi'\in\mathcal{G}_{N,k}}\sum_{\eta\in\Omega'_{N,k}}P^\xi_{t_1}(\eta)\big|P^\eta_{t_2}(\xi') - \nu_{N,k}(\xi')\big| && (2.4.15)\\
&= \frac{1}{2}P^\xi_t\big(\mathcal{G}^c_{N,k}\big) + \sum_{\eta\in\Omega'_{N,k}}P^\xi_{t_1}(\eta)\times\frac{1}{2}\sum_{\xi'\in\mathcal{G}_{N,k}}\big|P^\eta_{t_2}(\xi') - \nu_{N,k}(\xi')\big| && (2.4.16)\\
&\leq \frac{1}{2}P^\xi_t\big(\mathcal{G}^c_{N,k}\big) + 1\times\max_{\eta\in\mathcal{G}_{N,k}}\big\|P^\eta_{t_2} - \nu_{N,k}\big\|_{\mathrm{TV}} + P^\xi_{t_1}\big(\mathcal{G}^c_{N,k}\big)\times 1. && (2.4.17)
\end{aligned}$$

Here (2.4.14) follows from the fact that $\nu_{N,k}$ is supported on $\mathcal{G}_{N,k}$, and (2.4.15) is obtained via an application of the triangle inequality. We obtain (2.4.17) by splitting the sum in (2.4.16) into two parts, namely the sum over $\eta\in\mathcal{G}_{N,k}$ and the sum over $\eta\in\mathcal{G}^c_{N,k}$. The terms on both the LHS and RHS of the product in (2.4.16) are bounded by 1, since the term on the LHS is a probability, and the term on the RHS is bounded by

$$\max_{\eta\in\Omega'_{N,k}}\big\|P^\eta_{t_2} - \nu_{N,k}\big\|_{\mathrm{TV}} \leq 1.$$

Fixing $\epsilon\in(0,1)$, we choose $t_1 = C(m)N^2\log N$ and $t_2 = T_{\mathcal{G}_{N,k}}(\epsilon/3)$. It follows from Lemma 2.4.1 that

$$d^m_{N,k}\big(C(m)N^2\log(N) + T_{\mathcal{G}_{N,k}}(\epsilon/3)\big) \leq \epsilon,$$

for all N sufficiently large. In particular

$$\widehat{T}^{N,k,m}_{\mathrm{cir}}(\epsilon) \leq C(m)N^2\log(N) + T_{\mathcal{G}_{N,k}}(\epsilon/3) = O(N^2\log N),$$

as $N\to\infty$ by (2.4.13), which completes the proof. □

Remark 2.4.3. *The lower bound on the mixing time in Section 2.4.1 together with* (2.4.10) *gives an upper bound on the log–Sobolev constant $\rho_{N,k}$ of the FSEP on $\mathcal{G}_{N,k}$ of order N^{-2}, for N sufficiently large. Combining this with Lemma 2.4.2, we see that the order of the log–Sobolev constant of the FSEP on the ergodic component is N^{-2}, up to some multiplicative constants.*

2.4.3 Hitting time of the ergodic component

In this section we prove Lemma 2.4.1. Before proceeding with the proof we introduce some notation and a coupling between the zero–range process on the circle $\mathbb{T}_n$ and an OBEP.

Definition 2.4.4. *For a zero–range configuration $\omega \in \mathbb{Z}_+^{\mathbb{T}_n}$ we say that the (clockwise) interval $[x,y] \subsetneq \mathbb{T}_n$ with $|y-x| < n$ is an ergodic region for the configuration ω if $\omega(z) \geq 1$ for all $z \in [x,y]$, and $\omega(x-1) = \omega(y+1) = 0$.*

Recalling the mapping $\Pi_{N,k}$ from Section 1.4.1 (see Figure 1.7), we note that every ergodic region for an FSEP configuration $\xi \in \mathcal{G}_{N,k}^c$ (Definition 2.1.2) corresponds to an ergodic region from Definition 2.4.4 in the zero–range configuration $\Pi_{N,k}(\xi)$.

Recall from (1.4.8) that the ergodic component for the zero–range process is defined by

$$G_n := \Big\{\omega \in \mathbb{Z}_+^{\mathbb{T}_n} : \omega(z) \geq 1 \text{ for all } z \in \mathbb{T}_n\Big\},$$

and the generator $\mathcal{L}_{ZC}^n$ in (1.4.4) of the zero–range process on the circle $\mathbb{T}_n$ acts on functions $f : \mathbb{Z}_+^{\mathbb{T}_n} \to \mathbb{R}$ via

$$\mathcal{L}_{\mathrm{ZC}}^n f(\omega) = \sum_{x\in\mathbb{T}_n} \sum_{z=\pm 1} \frac{1}{2} \mathbb{1}(\omega(x) \geq 2)\big[f(\omega^{x,x+z}) - f(\omega)\big].$$

By Remark (1.4.2), for an FSEP configuration $\xi \in \mathcal{G}_{N,k}^c$ it holds that

$$P_t^{\xi}(\mathcal{G}_{N,k}^c) = Q_t^{\Pi_{N,k}(\xi)}(G_{N-k}^c). \tag{2.4.18}$$

Recall from Definition 1.4.1 that we say that $\omega \leq \omega'$ for $\omega, \omega' \in \mathbb{Z}_+^{\mathbb{T}_n}$ if $\omega(z) \leq \omega'(z)$ for all $z \in \mathbb{T}_n$. An event $A \subset \mathbb{Z}_+^{\mathbb{T}_n}$ is *increasing* if $\omega \in A$ implies $\omega' \in A$ for all $\omega' \geq \omega$, and note that the event G_n is increasing.

Let Q_t^{ω} denote the law of the zero–range process with generator $\mathcal{L}_{ZC}^n$ and initial condition ω at time t. By (1.4.5), if $\omega \leq \omega'$, then

$$Q_t^{\omega}(G_n) \leq Q_t^{\omega'}(G_n), \tag{2.4.19}$$

We now give a coupling between the zero–range process whose initial condition ω has a single ergodic region $[x,y] \subsetneq \mathbb{T}_n$ containing ℓ particles and an OBEP $(\zeta_t)_{t\geq 0}$ on $\Omega_{\ell-1}$ with parameters $(1/2, 1/2, 0, 0, 1/2)$.

Let $E_{n,\ell}$, denote the set of zero–range configurations on $\mathbb{T}_n$ with n particles and a single ergodic region containing $\ell \leq n$ particles, then we define a mapping $\Phi : E_{n,\ell} \to \Omega_{\ell-1}$ as follows: if the ergodic region in the zero–range configuration is a single site $\{x\}$ then we set $\Phi(\omega)$ to be the empty particle configuration $\mathbf{0} \in \Omega_{\ell-1}$. If the ergodic region is given by $[x,y]$ with $y > x$, then we label zero–range stacks (clockwise) from left to right by $i = 1, \ldots, y-x+1$. In the mapping Φ the i–th stack in the zero–range process corresponds to a particle in the OBEP for each $i \in [1, y-x]$.

We label particles from left to right in the OBEP. The number of particles less one in stack 1 of the zero–range configuration ω, i.e. $\omega(x)-1$, is the number of holes between the left boundary and particle 1 in the OBEP configuration $\Phi(\omega)$. For $i \in [2, y-x]$, the number of particles less one in the i–th stack of the zero–range configuration, i.e. $\omega(i+x-1)-1$, is the number of holes

between the $(i-1)$–th and i–th particles in the OBEP configuration $\Phi(\omega)$. Lastly the number of particles less one in stack $y-x+1$ of the zero–range configuration ω, i.e. $\omega(y)-1$, is the number of holes between the $(y-x)$–th particle and the right boundary in the OBEP configuration $\Phi(\omega)$. See Figure 2.14 for an example.

For the coupling, each time a particle in the zero–range process jumps left from the leftmost stack to an empty site, we input a particle at the left boundary in the OBEP. We then relabel the stacks in the zero–range process clockwise and the particles in the OBEP from left to right.

Similarly, each time a particle in the zero–range process jumps right from the rightmost stack to an empty site we input a particle at the right boundary in the OBEP. If a particle in the zero–range process jumps left from the i–th leftmost stack to the $(i-1)$–th leftmost stack, the $(i-1)$–th particle in the OBEP jumps one step right. If a particle in the zero–range process moves right from the i–th stack to the $(i+1)$–th stack, the i–th particle in the OBEP jumps one step left.

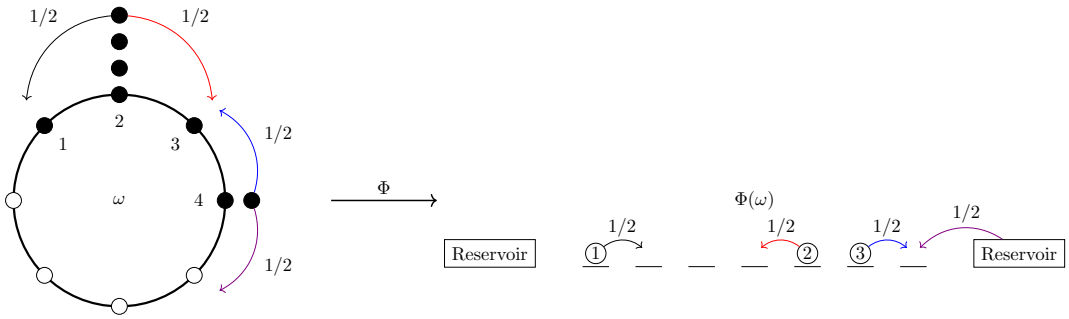


Figure 2.14: The mapping $\Phi : E_{8,8} \to \Omega_7$ for a configuration $\omega \in E_{8,8}$. All possible transitions of the zero–range process and OBEP are indicated. Matching colours indicate corresponding transitions in the coupling, and the stacks in ω are labelled $1, 2, 3$ and 4.

The mapping Φ is not bijective; every rotation of a configuration in $E_{n,\ell}$ results in the same OBEP configuration. However, under the coupling the zero–range event that all particles become stuck in the dynamics, i.e.

$$\mathcal{F}_n := \Big\{\omega \in \mathbb{Z}_+^{\mathbb{T}_n} : \omega(i) \leq 1 \text{ for all } i \in \mathbb{T}_n\Big\},$$

corresponds to the full configuration $\mathbf{1} \in \Omega_{\ell-1}$ in the OBEP. In particular, if P_t^ζ denotes the law of the OBEP with parameters $(1/2, 1/2, 0, 0, 1/2)$ and initial condition $\zeta \in \Omega_{\ell-1}$ at time t, then

$$Q_t^\omega(\mathcal{F}_n) = P_t^{\Phi(\omega)}(\mathbf{1}), \tag{2.4.20}$$

for all $\omega \in E_{n,\ell}$ and $t \geq 0$. Since $\ell \leq n$ the zero–range process will hit the set $\mathcal{F}_n$ almost surely.

We now prove Lemma 2.4.1.

Proof of Lemma 2.4.1. Let $\xi \in \mathcal{I}_{N,k}^m$ be arbitrary. By (2.4.18), it is sufficient to bound

$$Q_t^{\Pi_{N,k}(\xi)}(G_{N-k}^c).$$

We now construct a configuration $\bar{\omega} \leq \Pi_{N,k}(\xi)$ with exactly $N-k$ particles and at most m ergodic regions. Assume WLOG that the FSEP configuration ξ has exactly m ergodic regions

$\{I_1, \dots, I_m\}$ collectively containing at least $N-k$ particles, then each ergodic region I_i corresponds to an ergodic region in the zero–range configuration $\Pi_{N,k}(\xi)$, which we denote by $\{J_1, \dots, J_m\}$. To construct $\bar{\omega}$ we delete all particles not contained in $\{J_1, \dots, J_m\}$ then, if necessary, arbitrarily delete particles from the ergodic regions $\{J_1, \dots, J_m\}$ until only $N-k$ remain. By (2.4.19)

$$Q_t^{\Pi_{N,k}(\xi)}(G_{N-k}^c) \leq Q_t^{\bar{\omega}}(G_{N-k}^c).$$

We now define a collection of stopping times for the zero–range process $(\bar{\omega}_t)_{t\geq 0}$ with law $Q_t^{\bar{\omega}}$. Let $\tau_0 = 0$ and suppose that τ_{s-1} is given for some $s \geq 1$ with $\bar{\omega}_{\tau_{s-1}} \notin G_{N-k}$, then τ_s is the first time $t > \tau_{s-1}$ that one of the following three conditions is met:

(1) The number of ergodic regions in $\bar{\omega}_t$ is one less than the number of ergodic regions in $\bar{\omega}_{\tau_{s-1}}$, i.e. two ergodic regions evolve and meet to form a single ergodic region;

(2) The number of ergodic regions in $\bar{\omega}_t$ for which all particles are stuck is one greater than the number of such ergodic regions in $\bar{\omega}_{\tau_{s-1}}$, i.e. an ergodic region evolves until all of its particles become stuck;

(3) $\bar{\omega}_t \in G_{N-k}$.

The number of such stopping times is at most $2m-1$ since (1) and (2) can occur at most $m-1$ times, and (3) occurs exactly once. In particular,

$$Q_t^{\Pi_{N,k}(\xi)}(G_{N-k}^c) \leq Q_t^{\bar{\omega}}(G_{N-k}^c) = Q_t^{\bar{\omega}}\left(\sum_{s\geq 1} \tau_s - \tau_{s-1} > t\right).$$

Let $\hat{\omega} \leq \bar{\omega}$ be the configuration obtained by deleting all but one ergodic region J_1 in $\bar{\omega}$ with at least one stack containing at least two particles, i.e. $\bar{\omega} \notin \mathcal{F}_{N-k}$. Let $(\hat{\omega}_t)_{t\geq 0}$ be the zero–range process generated by $\mathcal{L}_{ZC}^{N-k}$ with initial condition $\hat{\omega}$.

By monotonicity of the zero–range process, the hitting time of the event $\mathcal{F}_{N-k}$ in the process $(\hat{\omega}_t)_{t\geq 0}$ is stochastically larger than τ_1; in the dynamics of $(\bar{\omega}_t)_t$ the ergodic region J_1 either evolves to become completely frozen or the stopping time τ_1 occurs at some earlier time. In particular, it holds that

$$Q_t^{\bar{\omega}}(\tau_1 > t) \leq Q_t^{\hat{\omega}}(\mathcal{F}_{N-k}^c).$$

If $\ell \leq N-k$ is the number of particles in $\hat{\omega}$, then by the coupling to the OBEP on $\Omega_{\ell-1}$ with parameters $(1/2, 1/2, 0, 0, 1/2)$ and (2.4.20) we have

$$Q_t^{\hat{\omega}}(\mathcal{F}_{N-k}^c) = P_t^{\Phi(\hat{\omega})}(\mathbf{1}^c).$$

The stationary distribution for the OBEP with parameters $(1/2, 1/2, 0, 0, 1/2)$ is the Dirac measure $\delta_{\mathbf{1}}$, and

$$\begin{aligned} P_t^{\Phi(\hat{\omega})}(\mathbf{1}^c) &= \left|P_t^{\Phi(\hat{\omega})}(\mathbf{1}^c) - \delta_{\mathbf{1}}(\mathbf{1}^c)\right| \\ &\leq \max_{A\subseteq\Omega_{\ell-1}} \left|P_t^{\Phi(\hat{\omega})}(A) - \delta_{\mathbf{1}}(A)\right| \\ &= \left\|P_t^{\Phi(\hat{\omega})} - \delta_{\mathbf{1}}\right\|_{\mathrm{TV}} \\ &\leq \max_{\zeta\in\Omega_{\ell-1}} \left\|P_t^{\zeta} - \delta_{\mathbf{1}}\right\|_{\mathrm{TV}}, \end{aligned}$$

where the third line follows from (1.5.1). By application of Theorem 1.4, there exists a constant $C > 0$ such that

$$Q^{\bar{\omega}}_{CN^2 \log N}(\tau_1 > CN^2 \log N) \leq P^{\Phi(\hat{\omega})}_{CN^2 \log N}(\mathbf{1}^c) \leq \frac{\epsilon}{2m-1},$$

for all N sufficiently large, since $\ell \leq N - k$ and $N - k \asymp (1-\alpha)N$ as $N \to \infty$.

By application of the Strong Markov property, and repeating the above argument for each time interval $[\tau_{s-1}, \tau_s]$, it holds that

$$P^{\xi}_{C(2m-1)N^2 \log N}(\mathcal{G}^c_{N,k}) \leq \epsilon,$$

for all N sufficiently large. As $\xi \in \mathcal{I}^m_{N,k}$ was chosen arbitrarily, this completes the proof. □

2.4.4 Lower bound for the log–Sobolev constant

In this section we prove Lemma 2.4.2. Before proceeding with the proof we introduce some notation and give the idea of the proof. Throughout Section 2.4.4 we will often drop the subscripts N and k from the notation. Moreover, we assume for notational ease that N is even. All proofs may be generalised to N odd by taking the integer part of $N/2$ whenever it appears.

We partition the ergodic component $\mathcal{G}$ into three parts for each $x \in [N]$ as follows:

$$\begin{aligned} \mathcal{G}^{(1)}_x &= \{\xi \in \mathcal{G} : \xi(x) = 1, \xi(x+1) = 1\}, \\ \mathcal{G}^{(2)}_x &= \{\xi \in \mathcal{G} : \xi(x) = 1, \xi(x+1) = 0\}, \\ \mathcal{G}^{(3)}_x &= \{\xi \in \mathcal{G} : \xi(x) = 0, \xi(x+1) = 1\}. \end{aligned} \tag{2.4.21}$$

For any configuration $\xi \in \mathcal{G}^{(2)}_x$ we must have that $\xi(x+2) = 1$, and for any configuration $\xi \in \mathcal{G}^{(3)}_x$ we must have that $\xi(x-1) = 1$, since $\mathcal{G}^{(2)}_x \cup \mathcal{G}^{(3)}_x \subset \mathcal{G}$, and no two holes are adjacent in the ergodic component.

Remark 2.4.5. *We observe that the transitions of the FSEP on $\mathcal{G}^{(1)}_x$ are the same as the transitions for the FSEP on the ergodic component for the segment $\mathcal{E}_{N,k}$, treating x and $x+1$ as the endpoints of the segment. By Remark 2.3.1 this is equivalent to the transitions of a SSEP on the segment $[k-1]$ with $N-k$ particles. Similarly the transitions of the FSEP on $\mathcal{G}^{(2)}_x$ are the same as the transitions for the FSEP on the ergodic component for the segment $\mathcal{E}_{N-1,k}$, treating x and $x+2$ as the endpoints of the segment. Lastly the transitions of the FSEP on $\mathcal{G}^{(3)}_x$ are the same as the transitions for the FSEP on the ergodic component for the segment $\mathcal{E}_{N-1,k}$, treating $x-1$ and $x+1$ as the endpoints of the segment. By Remark 2.3.1 the transitions on $\mathcal{G}^{(2)}_x$ and $\mathcal{G}^{(3)}_x$ are equivalent to the transitions of a SSEP on the segment $[k-1]$ with $N-k-1$ particles.*

Let ρ denote the log–Sobolev constant for the generator $\mathcal{L}^\circ$ restricted to the ergodic component $\mathcal{G}$, and let $\tilde{\rho}_{n,\ell}$ denote the log–Sobolev constant for the SSEP on $\Omega_{n,\ell}$.

The idea of the proof is to bound the log–Sobolev constant ρ by the log–Sobolev constants for the SSEP on the segment. Recall that the log–Sobolev ρ is the largest constant c such that

$$c \operatorname{Ent}_\nu(f^2) \leq \mathcal{D}(f).$$

We will lower bound the Dirichlet form by removing the transitions between the sets in the partition (2.4.21), and we will upper bound the entropy using the quasi–factorisation in Proposition

1.5.8. This will allow us to compare to the log–Sobolev constant for the SSEP via Remark 2.4.5, and known bounds on the log–Sobolev constant for the SSEP [107] will yield the result.

In order to apply Proposition 1.5.8, we need to construct two weakly dependent σ–algebras. Let $\mathcal{F}_1$ be the σ–algebra generated by $\{\mathcal{G}_0^{(1)}, \mathcal{G}_0^{(2)}, \mathcal{G}_0^{(3)}\}$ and let $\mathcal{F}_2$ be the σ–algebra generated by $\{\mathcal{G}_{N/2}^{(1)}, \mathcal{G}_{N/2}^{(2)}, \mathcal{G}_{N/2}^{(3)}\}$.

The idea here is that the occupation parameters of sites 0 and 1 should not depend too strongly on the occupation parameters of sites $N/2$ and $1 + N/2$ in the limit, see Figure 2.15.

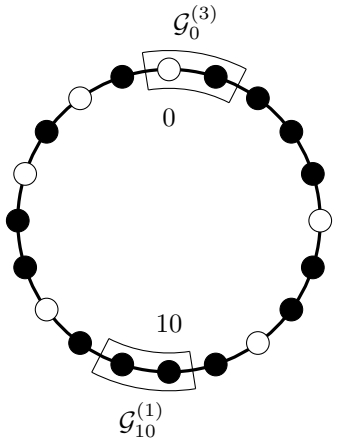


Figure 2.15: A configuration in $\mathcal{G}_0^{(3)} \cap \mathcal{G}_{10}^{(1)} \subset \mathcal{G}_{20,14}$. When k is close to $N/2$ or N the events $\mathcal{G}_0^{(3)}$ and $\mathcal{G}_{N/2}^{(1)}$ are not independent, but as we insist on a density $\alpha \in (1/2, 1)$ of particles, the events $\mathcal{G}_0^{(3)}$ and $\mathcal{G}_{N/2}^{(1)}$ will be asymptotically independent as $N \to \infty$ (see Lemma 2.4.6).

The following lemma gives that the σ–algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ are asymptotically independent.

Lemma 2.4.6. *Let $k(N) > N/2$ be a sequence such that $k(N)/N \to \alpha$ as $N \to \infty$. Then for $i, j \in [3]$ it holds that*

$$\left| \frac{\nu_{N,k}\big(\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}\big)}{\nu_{N,k}\big(\mathcal{G}_{N/2}^{(j)}\big)\nu_{N,k}\big(\mathcal{G}_0^{(i)}\big)} - 1 \right| \leq \epsilon, \tag{2.4.22}$$

for all N sufficiently large depending on ϵ.

Before proving Lemma 2.4.6, we show how it may be used to yield Lemma 2.4.2.

Proof of Lemma 2.4.2. The Dirichlet form of the FSEP restricted to the ergodic component $\mathcal{G}$ is given by

$$\mathcal{D}(f) = \frac{1}{2} \sum_{\xi, \xi' \in \mathcal{G}} \nu(\xi) c(\xi, \xi') \big(f(\xi) - f(\xi')\big)^2,$$

for $f : \mathcal{G} \to \mathbb{R}$, where $c(\xi, \xi')$ is the transition rate between ξ and ξ'. By removing transitions

between sets in the partition (2.4.21) it holds that

$$\mathcal{D}(f) \geq \frac{1}{2}\sum_{i=1}^{3}\sum_{\xi,\xi'\in\mathcal{G}_x^{(i)}} \nu(\xi)c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2, \tag{2.4.23}$$

for all $x \in \mathbb{T}_N$. Combining the lower bounds for the Dirichlet form in (2.4.23) for both $x = 0$ and $x = N/2$, it follows that

$$\mathcal{D}(f) \geq \frac{1}{4}\sum_{x\in\{0,N/2\}}\sum_{i=1}^{3}\sum_{\xi,\xi'\in\mathcal{G}_x^{(i)}} \nu(\xi)c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2. \tag{2.4.24}$$

Let $\nu_j(\cdot) = \nu(\cdot \mid \mathcal{F}_j)$ for $j \in [2]$, and observe that $\nu_1(f)$ is the function given by

$$\nu_1\big(f\big)(\xi) = \sum_{i=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_0^{(i)}\big)\sum_{\xi'\in\mathcal{G}_0^{(i)}}\frac{f(\xi')}{|\mathcal{G}_0^{(i)}|},$$

for $\xi \in \mathcal{G}$, since ν is uniform on $\mathcal{G}$. In particular the value of the function $\nu_1(f)$ for each configuration in $\mathcal{G}_0^{(i)}$ is the uniform average of the function f over $\mathcal{G}_0^{(i)}$.

$$\begin{aligned}
\nu\big(\mathrm{Ent}_{\nu_1}(f^2)\big) &= \sum_{\xi\in\mathcal{G}}\nu(\xi)\nu_1\left(f^2\log\left(\frac{f^2}{\nu_1(f^2)}\right)\right)(\xi)\\
&= \frac{1}{|\mathcal{G}|}\sum_{\xi\in\mathcal{G}}\sum_{i=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_0^{(i)}\big)\sum_{\xi'\in\mathcal{G}_0^{(i)}}\frac{f^2(\xi')}{|\mathcal{G}_0^{(i)}|}\log\left(\frac{f^2(\xi')}{\nu_1\big(f^2\big)(\xi')}\right)\\
&= \sum_{i=1}^{3}\frac{|\mathcal{G}_0^{(i)}|}{|\mathcal{G}|}\sum_{\xi'\in\mathcal{G}_0^{(i)}}\frac{f^2(\xi')}{|\mathcal{G}_0^{(i)}|}\log\left(\frac{f^2(\xi')}{\nu_1\big(f^2\big)(\xi')}\right).
\end{aligned} \tag{2.4.25}$$

By Remark 2.4.5 we have that

$$\sum_{\xi,\xi'\in\mathcal{G}_0^{(1)}}\frac{1}{|\mathcal{G}_0^{(1)}|}c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2 \geq \tilde{\rho}_{k-1,N-k}\sum_{\xi'\in\mathcal{G}_0^{(1)}}\frac{f^2(\xi')}{|\mathcal{G}_0^{(1)}|}\log\left(\frac{f^2(\xi')}{\nu_1\big(f^2\big)(\xi')}\right), \tag{2.4.26}$$

and for $i = 2, 3$ we have

$$\sum_{\xi,\xi'\in\mathcal{G}_0^{(i)}}\frac{1}{|\mathcal{G}_0^{(i)}|}c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2 \geq \tilde{\rho}_{k-1,N-k-1}\sum_{\xi'\in\mathcal{G}_0^{(i)}}\frac{f^2(\xi')}{|\mathcal{G}_0^{(i)}|}\log\left(\frac{f^2(\xi')}{\nu_1\big(f^2\big)(\xi')}\right). \tag{2.4.27}$$

Combining (2.4.26) and (2.4.27), it follows that

$$\begin{aligned}
&\sum_{i=1}^{3}\sum_{\xi,\xi'\in\mathcal{G}_0^{(i)}}\frac{1}{|\mathcal{G}|}c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2\\
&\qquad=\sum_{i=1}^{3}\frac{|\mathcal{G}_0^{(i)}|}{|\mathcal{G}|}\sum_{\xi,\xi'\in\mathcal{G}_0^{(i)}}\frac{1}{|\mathcal{G}_0^{(i)}|}c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2\\
&\qquad\geq\min(\tilde\rho_{k-1,N-k},\tilde\rho_{k-1,N-k-1})\sum_{i=1}^{3}\frac{|\mathcal{G}_0^{(i)}|}{|\mathcal{G}|}\sum_{\xi'\in\mathcal{G}_0^{(i)}}\frac{f^2(\xi')}{|\mathcal{G}_0^{(i)}|}\log\left(\frac{f^2(\xi')}{\nu_1\big(f^2\big)(\xi')}\right)\\
&\qquad=\min(\tilde\rho_{k-1,N-k},\tilde\rho_{k-1,N-k-1})\nu\big(\mathrm{Ent}_{\nu_1}(f^2)\big),
\end{aligned}\tag{2.4.28}$$

where the last equality follows from (2.4.25). By symmetry,

$$\sum_{i=1}^{3}\sum_{\xi,\xi'\in\mathcal{G}_{N/2}^{(i)}}\frac{1}{|\mathcal{G}|}c(\xi,\xi')\big(f(\xi)-f(\xi')\big)^2\geq\min(\tilde\rho_{k-1,N-k},\tilde\rho_{k-1,N-k-1})\nu\big(\mathrm{Ent}_{\nu_2}(f^2)\big).\tag{2.4.29}$$

By (2.4.24), (2.4.28) and (2.4.29) it holds that

$$\mathcal{D}(f)\geq\frac{1}{4}\min(\tilde\rho_{k-1,N-k},\tilde\rho_{k-1,N-k-1})\nu\big[\mathrm{Ent}_{\nu_1}\big(f^2\big)+\mathrm{Ent}_{\nu_2}\big(f^2\big)\big].\tag{2.4.30}$$

We wish to bound the entropy above by the mean of the sum of conditional entropies using Proposition 1.5.8. To apply Proposition 1.5.8 we need to check that (1.5.26) holds, i.e. for some $\epsilon\in[0,1)$

$$\|\nu_2(g)-\nu(g)\|_{\infty,\pi}\leq\epsilon\|g\|_{1,\pi},$$

for all real–valued $\mathcal{F}_1$–measurable functions g. If g is $\mathcal{F}_1$–measurable then g takes exactly one value on $\mathcal{G}_0^{(1)},\mathcal{G}_0^{(2)}$ and $\mathcal{G}_0^{(3)}$, and we set $g_i=g(\xi)$ for all $\xi\in\mathcal{G}_0^{(i)}$ and $i\in[3]$. Note that

$$\begin{aligned}
|\nu_2(g)(\xi)-\nu(g)|&=\left|\sum_{j=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_{N/2}^{(j)}\big)\sum_{\xi'\in\mathcal{G}_{N/2}^{(j)}}\frac{g(\xi')}{|\mathcal{G}_{N/2}^{(j)}|}-\sum_{\xi'\in\mathcal{G}}\frac{g(\xi')}{|\mathcal{G}|}\right|\\
&=\left|\sum_{j=1}^{3}\sum_{i=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_{N/2}^{(j)}\big)g_i\frac{|\mathcal{G}_{N/2}^{(j)}\cap\mathcal{G}_0^{(i)}|}{|\mathcal{G}_{N/2}^{(j)}|}-\sum_{i=1}^{3}g_i\frac{|\mathcal{G}_0^{(i)}|}{|\mathcal{G}|}\right|\\
&\leq\sum_{i=1}^{3}|g_i|\frac{|\mathcal{G}_0^{(i)}|}{|\mathcal{G}|}\max_i\left(\left|\sum_{j=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_{N/2}^{(j)}\big)\frac{|\mathcal{G}_{N/2}^{(j)}\cap\mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|}-1\right|\right)\\
&=\|g\|_{1,\pi}\max_i\left(\left|\sum_{j=1}^{3}\mathbb{1}\big(\xi\in\mathcal{G}_{N/2}^{(j)}\big)\frac{|\mathcal{G}_{N/2}^{(j)}\cap\mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|}-1\right|\right)\\
&\leq\|g\|_{1,\pi}\max_{(i,j)}\left(\left|\frac{|\mathcal{G}_{N/2}^{(j)}\cap\mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|}-1\right|\right).
\end{aligned}$$

Since $\xi \in \mathcal{G}$ and g were arbitrary, it holds that

$$\|\nu_2(g) - \nu(g)\|_{\infty,\pi} \leq \|g\|_{1,\pi} \max_{(i,j)} \left(\left| \frac{|\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|} - 1 \right| \right),$$

and it remains to check that

$$\max_{(i,j)} \left(\left| \frac{|\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|} - 1 \right| \right) \leq \epsilon, \tag{2.4.31}$$

for all N sufficiently large depending on ϵ. Noting that

$$\begin{aligned} \left| \frac{|\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|} - 1 \right| &= \left| \frac{|\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}|}{|\mathcal{G}|} \times \frac{|\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}|} \times \frac{|\mathcal{G}|}{|\mathcal{G}_0^{(i)}|} - 1 \right| \\ &= \left| \frac{\nu\big(\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}\big)}{\nu\big(\mathcal{G}_{N/2}^{(j)}\big)\nu\big(\mathcal{G}_0^{(i)}\big)} - 1 \right|, \end{aligned}$$

an application of Lemma 2.4.6 gives that (2.4.31) holds for all $\epsilon \in (0,1)$ and N sufficiently large depending on ϵ. Choosing $\epsilon \in (0,1)$ small enough so that

$$\vartheta(\epsilon) = \frac{84\epsilon}{(1-\epsilon)^2} < 1,$$

(1.5.27) gives that

$$\big(1 - \vartheta(\epsilon)\big)\mathrm{Ent}_\nu\big(f^2\big) \leq \nu\big[\mathrm{Ent}_{\nu_1}\big(f^2\big) + \mathrm{Ent}_{\nu_2}\big(f^2\big)\big],$$

for all N sufficiently large. By application of (2.4.30) it follows that

$$\mathcal{D}(f) \geq \frac{1}{4}\big(1 - \vartheta(\epsilon)\big)\min(\tilde{\rho}_{k-1,N-k}, \tilde{\rho}_{k-1,N-k-1})\mathrm{Ent}_\nu\big(f^2\big),$$

for all N sufficiently large. In particular

$$\rho_{N,k} \geq \frac{1}{4}\big(1 - \vartheta(\epsilon)\big)\min(\tilde{\rho}_{k-1,N-k}, \tilde{\rho}_{k-1,N-k-1}), \tag{2.4.32}$$

for all N sufficiently large. By an application of [107, Theorem A]

$$\tilde{\rho}_{n,\ell} \geq Cn^{-2}, \tag{2.4.33}$$

and by (2.4.32) it holds that there exists a constant $C' > 0$ depending on α (but not N) such that

$$\rho_{N,k} \geq C'N^{-2},$$

for all N sufficiently large, which completes the proof. □

In the proof of Lemma 2.4.2 we only required that

$$\max_{(i,j)} \left(\left| \frac{|\mathcal{G}_{N/2}^{(j)} \cap \mathcal{G}_0^{(i)}||\mathcal{G}|}{|\mathcal{G}_{N/2}^{(j)}||\mathcal{G}_0^{(i)}|} - 1 \right| \right) \leq \epsilon,$$

for some $\epsilon \in (0,1)$ such that $\vartheta(\epsilon) < 1$. Lemma 2.4.6 says that this holds for all $\epsilon \in (0,1)$ and N sufficiently large and is a stronger statement than required for the proof of Lemma 2.4.2.

To prove Lemma 2.4.6 we establish the limit of all the probabilities contained in (2.4.22) as $N \to \infty$. Let $L_\alpha : [3] \times [3] \to \mathbb{R}$ be the function given by

$$L_\alpha(i,j) = \begin{cases} (2\alpha-1)^2 & \text{if } i = j = 1, \\ (1-\alpha)^2 & \text{if } i \neq 1, j \neq 1, \\ (2\alpha-1)(1-\alpha) & \text{otherwise.} \end{cases}$$

The majority of the proof is devoted to showing that

$$\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)}\big) = L_\alpha(i,j),$$

for all $(i,j) \in [3] \times [3]$.

It is quite challenging to directly compute the size of the sets $\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)}$. To determine the limit we will generate a uniform random sample on $\mathcal{G}_{N,k}$ conditional on the number of particles in the half–circle $[1, N/2] \subset \mathbb{T}_N$. We do this by generating uniform random configurations on the segment $[N/2]$ with a given number of particles, then gluing these segments together to form a configuration $\Omega'_{N,k}$. We then use rejection sampling to make sure the resulting configuration lies in the ergodic component $\mathcal{G}_{N,k}$. We establish the asymptotic probability of such configurations lying in $\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)}$ when the number of particles in each half–circle is 'close' to $k/2$ using a combinatorial argument, and use a concentration result to remove the conditioning. We will also establish that

$$\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_x^{(i)}\big) = \begin{cases} 2\alpha - 1 & \text{if } i = 1, \\ 1-\alpha & \text{otherwise,} \end{cases}$$

for $x = 0$ and $x = N/2$ using a combinatorial argument.

Proof of Lemma 2.4.6. For $\xi \in \Omega'_{N,k}$, let $K_1(\xi)$ denote the number of particles in the interval $[1, N/2]$. Let $k_1(N)$ be a sequence of integers with $k_1 \in \mathcal{J}_N$, where

$$\mathcal{J}_N = [k/2 - N^{3/4}, k/2 + N^{3/4}] \cap \mathbb{Z},$$

the window of integers within $N^{3/4}$ of $k/2$. We take N large enough so that $k_1 \in [N/4, N/2]$.

We consider the set of configurations

$$\begin{aligned} S(N,k,k_1) := \big\{\xi \in \Omega'_{N,k} : & K_1(\xi) = k_1, \xi(x) + \xi(x+1) \geq 1 \\ & \text{for all } x \in \{1, \dots, N/2 - 1\} \cup \{1 + N/2, \dots N - 2\}\big\}. \end{aligned}$$

We note that $S(N,k,k_1)$ contains the set of ergodic configurations $\xi \in \mathcal{G}_{N,k}$ with $K_1(\xi) = k_1$, i.e.

$$G(N,k,k_1) := \big\{\xi \in \mathcal{G}_{N,k} : K_1(\xi) = k_1\big\},$$

however $S(N,k,k_1)$ also contains configurations with $\xi(0) = \xi(1) = 0$ or $\xi(N/2) = \xi(1+N/2) = 0$.

We also consider the set of FSEP configurations on the segment $[M]$ with $\ell > M/2$ particles given by

$$S'(M,\ell) := \big\{\xi \in \Omega_{M,\ell} : \xi(x) + \xi(x+1) \geq 1 \text{ for all } x \in [M-1]\big\}.$$

We note that $S'(M,\ell)$ contains the ergodic component on the segment $\mathcal{E}_{M,\ell}$, but also contains configurations where the endpoints have holes.

We generate a uniform sample on $G(N,k,k_1)$ using the following rejection sampling:

- generate a uniform sample from $S'(N/2, k_1)$;
- generate independently a uniform sample on $S'(N/2, k_2)$ for $k_2 := k_1 - k$;
- glue the uniform samples from $S'(N/2, k_1)$ and $S'(N/2, k_2)$ together at the endpoints so that the segment with k_1 particles is the half–circle $[1, N/2] \subset \mathbb{T}_N$, and the other segment corresponds to the half–circle $[N/2+1, N] \subset \mathbb{T}_N$. This is a uniform sample from $S(N, k, k_1)$;
- if the uniform sample from $S(N, k, k_1)$ is in $\mathcal{G}_{N,k}$ output the sample, otherwise reject the sample and repeat the above procedure.

Observe that a configuration in $S'(M, \ell)$ may be identified with a configuration in the ergodic component for the segment $\mathcal{E}_{M+2,\ell+2}$ by placing an additional particle to the left and right of the left and right boundary of the segment respectively. In particular, for $\xi \in S'(M, \ell)$ we identify ξ with a configuration $\tilde{\xi} \in \mathcal{E}_{M+2,\ell+2}$ by

$$\tilde{\xi}(x) = \begin{cases} 1 & \text{if } x \in \{1, M+2\}, \\ \xi(x-1) & \text{if } 2 \le x \le M+1. \end{cases}$$

Recall from Remark 2.3.1 that the FSEP on $\mathcal{E}_{M,\ell}$ is equivalent to a SEP on the segment $[\ell - 1]$ with $M - \ell$ particles, and

$$|\Omega_{\ell-1,M-\ell}| = \binom{\ell-1}{M-\ell},$$

since we choose $M - \ell$ sites to place the particles. Therefore

$$|S'(N/2, k_1)| = |\mathcal{E}_{2+N/2,k_1+2}| = |\Omega_{k_1+1,N/2-k_1}|,$$

and the probability we generate a given configuration $\xi \in S'(N/2, k_1)$ is

$$\binom{k_1+1}{N/2-k_1}^{-1}.$$

Similarly the probability we generate a given configuration in $S'(N/2, k_2)$ is

$$\binom{k_2+1}{N/2-k_2}^{-1}.$$

Moreover,

$$\big|S(N, k, k_1)\big| = \binom{k_1+1}{N/2-k_1}\binom{k_2+1}{N/2-k_2}. \tag{2.4.34}$$

Let π_{N,k,k_1} be the uniform measure on $S(N, k, k_1)$. The probability we do not reject a sample from π_{N,k,k_1} in the rejection sampling is $\pi_{N,k,k_1}(G(N, k, k_1))$, and the conditional measure $\pi_{N,k,k_1}(\cdot|G(N, k, k_1))$ is uniform on the set $G(N, k, k_1)$. It follows that

$$\nu_{N,k}(\cdot \mid K_1 = k_1) = \pi_{N,k,k_1}\big(\cdot \;\big|\; G(N, k, k_1)\big). \tag{2.4.35}$$

We now find the limit of the probability we do not reject a sample in the sampling procedure.

The number of configurations $\xi \in S'(M, \ell)$ with $\xi(1) = 0$ is given by

$$\big|\{\xi \in S'(M, \ell) : \xi(1) = 0\}\big| = \binom{\ell}{M-\ell-1}. \tag{2.4.36}$$

To see this observe that $\xi(2) = 1$ and each configuration may be identified with a configuration $\tilde{\xi} \in \mathcal{E}_{M,\ell+1}$ via

$$\tilde{\xi}(x) = \begin{cases} 1 & \text{if } x = M, \\ \xi(x+1) & \text{if } x \in [M-1]. \end{cases}$$

By symmetry it follows that

$$\big|\{\xi \in S'(M,\ell) : \xi(M) = 0\}\big| = \binom{\ell}{M-\ell-1}. \tag{2.4.37}$$

The number of configurations $\xi \in S'(M,\ell)$ with both $\xi(1) = 0$ and $\xi(M) = 0$ is given by

$$\big|\{\xi \in S'(M,\ell) : \xi(1) = \xi(M) = 0\}\big| = \binom{\ell-1}{M-\ell-2}. \tag{2.4.38}$$

To see this observe that $\xi(2) = \xi(M-1) = 1$ and each configuration may be identified with a configuration $\tilde{\xi} \in \mathcal{E}_{M-2,\ell}$ via

$$\tilde{\xi}(x) = \xi(x+1) \text{ for all } x \in [M-2].$$

In particular, by the inclusion–exclusion principle we have

$$\begin{aligned} \big|G(N,k,k_1)\big| &= \big|S(N,k,k_1)\big| \\ &\quad - \big|\{\xi \in S'(N/2,k_1) : \xi(1) = 0\}\big| \times \big|\{\xi \in S'(N/2,k_2) : \xi(N/2) = 0\}\big| \\ &\quad - \big|\{\xi \in S'(N/2,k_1) : \xi(N/2) = 0\}\big| \times \big|\{\xi \in S'(N/2,k_2) : \xi(1) = 0\}\big| \\ &\quad + \big|\{\xi \in S'(N/2,k_1) : \xi(1) = \xi(N/2) = 0\}\big| \times \big|\{\xi \in S'(N/2,k_2) : \xi(1) = \xi(N/2) = 0\}\big|. \end{aligned}$$

The second term on the RHS corresponds to the configurations in $S(N,k,k_1)$ with $\xi(0) = 0$ and $\xi(1) = 0$, the third term corresponds to the configurations in $S(N,k,k_1)$ with $\xi(N/2) = 0$ and $\xi(1+N/2) = 0$ and the fourth term corresponds to the configurations in $S(N,k,k_1)$ with $\xi(0) = 0$, $\xi(1) = 0$, $\xi(N/2) = 0$ and $\xi(1+N/2) = 0$. It follows from (2.4.34),(2.4.36),(2.4.37) and (2.4.38) that

$$\begin{aligned} \pi_{N,k,k_1}\big(G(N,k,k_1)\big) &= \frac{|G(N,k,k_1)|}{|S(N,k,k_1)|} \\ &= 1 - 2\frac{\binom{k_1}{N/2-k_1-1}\binom{k_2}{N/2-k_2-1}}{\binom{k_1+1}{N/2-k_1}\binom{k_2+1}{N/2-k_2}} + \frac{\binom{k_1-1}{N/2-k_1-2}\binom{k_2-1}{N/2-k_2-2}}{\binom{k_1+1}{N/2-k_1}\binom{k_2+1}{N/2-k_2}} \\ &= 1 - 2\prod_{i=1}^{2}\frac{N/2-k_i}{k_i+1} + \prod_{i=1}^{2}\frac{(N/2-k_i)(N/2-k_i-1)}{k_i(k_i+1)}. \end{aligned}$$

Since $k \asymp \alpha N$ and $k_i \asymp \alpha N/2$ as $N \to \infty$ it follows that

$$\lim_{N\to\infty} \pi_{N,k,k_1}\big(G(N,k,k_1)\big) = \left(\frac{2\alpha-1}{\alpha^2}\right)^2. \tag{2.4.39}$$

We now find the limit of the probabilities

$$\pi_{N,k,k_1}\Big(\big\{\xi \in \mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)} : K_1(\xi) = k_1\big\}\Big),$$

as $N \to \infty$ for all $i, j \in [3]$. We note that

$$\big\{\xi \in S(M,\ell) : \xi(1) = 1, \xi(M) = 0\big\} = \binom{\ell}{M-\ell}. \tag{2.4.40}$$

To see this, observe that each configuration ξ may be identified with a configuration $\tilde{\xi} \in \mathcal{E}_{M+1,\ell+1}$ given by

$$\tilde{\xi}(x) = \begin{cases} \xi(x) & \text{if } x \in [M], \\ 1 & \text{if } x = M+1. \end{cases}$$

By symmetry it holds that

$$\big\{\xi \in S(M,\ell) : \xi(1) = 0, \xi(M) = 1\big\} = \binom{\ell}{M-\ell}. \tag{2.4.41}$$

Moreover,

$$\big\{\xi \in S(M,\ell) : \xi(1) = 1, \xi(M) = 1\big\} = \binom{\ell-1}{M-\ell}, \tag{2.4.42}$$

since this is the ergodic component $\mathcal{E}_{M,\ell}$. For $i = j = 1$,

$$\begin{aligned}
&\pi_{N,k,k_1}\Big(\big\{\xi \in \mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)} : K_1(\xi) = k_1\big\}\Big) \\
&= \frac{\big|\big\{\xi \in S(N/2,k_1) : \xi(1) = 1, \xi(N/2) = 1\big\}\big|}{|S(N/2,k_1)|} \times \frac{\big|\big\{\xi \in S(N/2,k_2) : \xi(1) = 1, \xi(N/2) = 1\big\}\big|}{|S(N/2,k_2)|} \\
&= \frac{\binom{k_1-1}{N/2-k_1}}{\binom{k_1+1}{N/2-k_1}} \times \frac{\binom{k_2-1}{N/2-k_2}}{\binom{k_2+1}{N/2-k_2}} \\
&= \prod_{i=1}^{2} \frac{(2k_i+1-N/2)(2k_i-N/2)}{k_i(k_i+1)} \to \left(\frac{2\alpha-1}{\alpha}\right)^4,
\end{aligned}$$

as $N \to \infty$, where the second equality follows from (2.4.42). By application of (2.4.35) and (2.4.39) it holds that

$$\begin{aligned}
&\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_0^{(1)} \cap \mathcal{G}_{N/2}^{(1)} \,\big|\, K_1 = k_1\big) \\
&\qquad = \lim_{N\to\infty} \pi_{N,k,k_1}\Big(\big\{\xi \in \mathcal{G}_0^{(1)} \cap \mathcal{G}_{N/2}^{(1)} : K_1(\xi) = k_1\big\}\,\Big|\, G(N,k,k_1)\Big) \\
&\qquad = (2\alpha-1)^2.
\end{aligned}$$

Repeating the above calculation for all other cases, (2.4.40),(2.4.41) and (2.4.42) yield

$$\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)} \,\big|\, K_1 = k_1\big) = L_\alpha(i,j). \tag{2.4.43}$$

For fixed N we have

$$\begin{aligned}
&\nu_{N,k}\big(\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)} \,\big|\, K_1 \in \mathcal{J}_N\big) \\
&\qquad = \sum_{k_1 \in \mathcal{J}_N} \nu_{N,k}\big(\mathcal{G}_0^{(i)} \cap \mathcal{G}_{N/2}^{(j)} \,\big|\, K_1 = k_1\big)\nu_{N,k}\big(K_1 = k_1 \,\big|\, K_1 \in \mathcal{J}_N\big),
\end{aligned}$$

and it follows that

$$\min_{k_1\in\mathcal{J}_N}\nu_{N,k}\big(\mathcal{G}_0^{(i)}\cap\mathcal{G}_{N/2}^{(j)}\;\big|\;K_1=k_1\big)\le\nu_{N,k}\big(\mathcal{G}_0^{(i)}\cap\mathcal{G}_{N/2}^{(j)}\;\big|\;K_1\in\mathcal{J}_N\big)\le\max_{k_1\in\mathcal{J}_N}\nu_{N,k}\big(\mathcal{G}_0^{(i)}\cap\mathcal{G}_{N/2}^{(j)}\;\big|\;K_1=k_1\big).$$

Since (2.4.43) holds for all sequences with $k_1\in\mathcal{J}_N$, we may choose $k_1(N)$ to be a sequence which minimises or maximises the probability on the LHS of (2.4.43) for each N, choosing arbitrarily on agreements. By squeezing we have that

$$\lim_{N\to\infty}\nu_{N,k}\big(\mathcal{G}_0^{(i)}\cap\mathcal{G}_{N/2}^{(j)}\;\big|\;K_1\in\mathcal{J}_N\big)=L_\alpha(i,j).\tag{2.4.44}$$

We now use the following concentration property:

Lemma 2.4.7. *Suppose that $k(N)/N\to\alpha\in(1/2,1)$ as $N\to\infty$, then*

$$\nu_{N,k}\Big(K_1\in\big[k/2-N^{3/4},k/2+N^{3/4}\big]\Big)\to1,\tag{2.4.45}$$

as $N\to\infty$.

A proof of Lemma 2.4.7 using a large deviation argument is included in Section 2.4.5 for the sake of completion. By application of Lemma 2.4.7 and (2.4.44) it follows that

$$\lim_{N\to\infty}\nu_{N,k}\big(\mathcal{G}_0^{(i)}\cap\mathcal{G}_{N/2}^{(j)}\big)=L_\alpha(i,j).\tag{2.4.46}$$

We now show that

$$\lim_{N\to\infty}\nu_{N,k}\big(\mathcal{G}_x^{(i)}\big)=\begin{cases}2\alpha-1 & \text{if } i=1,\\ 1-\alpha & \text{otherwise,}\end{cases}$$

for $x=0$ and $x=N/2$.

Recall from (2.4.11) that

$$\big|\mathcal{G}_{N,k}\big|=\frac{N}{k}\binom{k}{N-k}.$$

To count the size of $\mathcal{G}_x^{(1)}$ we place a particle at sites x and $x+1$, and there are $k-1$ remaining places to put $N-k$ holes between particles. Therefore

$$\big|\mathcal{G}_x^{(1)}\big|=\binom{k-1}{N-k},$$

and

$$\nu_{N,k}\big(\mathcal{G}_0^{(1)}\big)=\nu_{N,k}\big(\mathcal{G}_{N/2}^{(1)}\big)=\frac{2k-N}{N}\to2\alpha-1,$$

as $N\to\infty$. To count the size of $\mathcal{G}_x^{(2)}$ we place a particle at site x and a hole at site $x+1$, and there are $k-1$ remaining places to put $N-k-1$ holes between particles. Therefore

$$\big|\mathcal{G}_x^{(2)}\big|=\binom{k-1}{N-k-1},$$

and

$$\nu_{N,k}\big(\mathcal{G}_0^{(2)}\big)=\nu_{N,k}\big(\mathcal{G}_{N/2}^{(2)}\big)=\frac{N-k}{N}\to1-\alpha,$$

as $N \to \infty$. Similarly

$$\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_0^{(3)}\big) = \lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_{N/2}^{(3)}\big) = 1-\alpha,$$

and

$$\lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_0^{(i)}\big) = \lim_{N\to\infty} \nu_{N,k}\big(\mathcal{G}_{N/2}^{(i)}\big) = \begin{cases} 2\alpha - 1 & \text{if } i = 1, \\ 1-\alpha & \text{if } i \neq 1. \end{cases} \tag{2.4.47}$$

Lemma 2.4.6 now follows from (2.4.46) and (2.4.47). □

2.4.5 Concentration of particle density in the half–circle

In this section we prove Lemma 2.4.7 uisng a large deviation argument.

Proof of Lemma 2.4.7. We consider a discrete Markov chain $(X_t)_{t\geq 1}$ on $\{0,1\}$ with transition matrix

$$\begin{bmatrix} 0 & 1 \\ \beta & 1-\beta \end{bmatrix},$$

for some $\beta \in (0,1)$. We let $X_1 \sim \text{Bernoulli}(1-\beta) \in \{0,1\}$, where $P[X_1 = 1] = 1-\beta$.

We define the events

$$A_N = \{X_1 + X_N > 0\} \bigcap \left\{ \sum_{t=1}^{N} X_t = k \right\}, \tag{2.4.48}$$

$$B_N^1 = \left\{ \sum_{t=1}^{N/2} X_t < k/2 - N^{3/4} \right\}, \tag{2.4.49}$$

$$B_N^2 = \left\{ \sum_{t=1}^{N/2} X_t > k/2 + N^{3/4} \right\}, \tag{2.4.50}$$

with $B_N := B_N^1 \cup B_N^2$.

We note that any trajectory $(X_t)_{t=1}^N$ contained in the event A_N corresponds to a unique configuration contained in the ergodic component $\mathcal{G}_{N,k}$. Moreover, every trajectory in the event A_N occurs with probability $\beta^{N-k}(1-\beta)^{2k-N}$ so that

$$P(B_N \mid A_N) = 1 - \nu\big(K_1 \in [k/2 - N^{3/4}, k/2 + N^{3/4}]\big),$$

so it is sufficient to check that

$$P(B_N \mid A_N) \to 0 \quad \text{as } N \to \infty. \tag{2.4.51}$$

We first bound $P(A_N)$ from below. We consider a renewal process $(Y_N)_{N\geq 0}$ with inter-arrival times $W_n \sim 1 + \text{Geom}(\beta)$, where $P(\text{Geom}(\beta) = i) = (1-\beta)^{i-1}\beta$ for $i \geq 1$. By some standard results for renewal processes (the law of large numbers and central limit theorem respectively) we have that

$$\frac{E[Y_N]}{N} \to \frac{1}{E[W_1]} = \frac{\beta}{1+\beta},$$

$$\frac{\sigma_N^2}{N} \to \frac{\text{Var}[W_1]}{E[W_1]^3} = \frac{\beta(1-\beta)}{(1+\beta)^3},$$

as $N \to \infty$, where $\sigma_N^2 = \mathrm{Var}[Y_N]$. By Chebyshev's inequality

$$P\Big(\big|Y_N - \mathbb{E}[Y_N]\big| \le \sqrt{2}\sigma_N\Big) \ge 1/2,$$

i.e. there is a window of order σ_N around the mean which occurs with probability at least $1/2$. We set $\beta = (1-\alpha)/\alpha$ so that

$$\frac{E[Y_N]}{N} \to 1-\alpha, \quad \text{as } N \to \infty.$$

Noting that σ_N grows like $\sqrt{N}$, $N - k \asymp (1-\alpha)N$ and Y_N takes integer values, for any fixed integer $\ell \in \mathbb{Z}$ it follows that

$$P\big(Y_N = N - k + \ell\big) \ge \frac{C_1(\alpha)}{\sqrt{N}}, \tag{2.4.52}$$

for all N sufficiently large, where $C_1 = C_1(\alpha) > 0$ is some positive constant depending on α. Since $\{X_N + X_1 > 0\} \supset \{X_1 = 1\}$ it holds that

$$\begin{aligned} P(A_N) &\ge \mathbb{P}\left(\{X_1 = 1\} \bigcap \left\{\sum_{t=1}^{N} X_t = k\right\}\right) \\ &= (1-\beta) P\left(\sum_{t=1}^{N} X_t = k \;\middle|\; X_1 = 1\right). \end{aligned} \tag{2.4.53}$$

The distribution of the number of zeros in the trajectory $(X_t)_{t=1}^N$ conditional on $X_1 = 1$, has the same distribution as the number of zeros occurring in the renewal process Y_{N-1}. Combining (2.4.52) with (2.4.53) gives

$$P(A_N) \ge \frac{C(\alpha)}{\sqrt{N}}, \tag{2.4.54}$$

for some constant $C = C(\alpha) > 0$ and all N sufficiently large.

We now bound $P(B_N)$ from above. Setting $b_n = n^{3/4}$ we use that

$$\frac{1}{b_n} \sum_{i=1}^{n} \big(W_i - E[W_i]\big),$$

satisfies a large deviation principle with rate function

$$I(x) = \frac{x^2}{2\mathrm{Var}[W_1]} = \frac{x^2(1-\beta)}{2\beta^2},$$

and speed $b_n^2/n = \sqrt{n}$, see De Acosta [28, Theorem 2.3]. A necessary and sufficient condition for this to hold (see Arcones [6]) is

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}} \log\Big[nP\big(|W_1 - E[W_1]| > n^{3/4}\big)\Big] = -\infty. \tag{2.4.55}$$

A concise proof is given by Eichelsbacher and Löwe [33]. For n sufficiently large

$$P\big(|W_1 - E[W_1]| > n^{3/4}\big) = P\big(\mathrm{Geom}(\beta) > n^{3/4} + 1/\beta\big) = (1-\beta)^{\lfloor n^{3/4} + 1/\beta \rfloor},$$

where $\lfloor x \rfloor$ is the integer part of $x \geq 0$, so that condition (2.4.55) holds.

Setting $n = \lfloor N/2 - k/2 + N^{3/4} \rfloor$, it is straightforward to see that

$$P\big(B_N^1\big) = \beta P\left(\sum_{i=1}^{n-1} W_i \leq N/2 - 1\right) + (1-\beta) P\left(\sum_{i=1}^{n} W_i \leq N/2\right),$$

by conditioning on the outcome of X_1. Noting that

$$P\left(\sum_{i=1}^{n} W_i \leq N/2\right) = P\left(\frac{1}{n^{3/4}} \sum_{i=1}^{n} \big(W_i - E[W_i]\big) \leq \frac{N/2 - n/(1-\alpha)}{n^{3/4}}\right),$$

and observing that

$$\lim_{N\to\infty} \frac{N/2 - n/(1-\alpha)}{n^{3/4}} = -\frac{2^{3/4}}{(1-\alpha)^{7/4}},$$

we may set $x = -(1-\alpha)^{-7/4}$ so that for N sufficiently large

$$P\left(\sum_{i=1}^{n} W_i \leq N/2\right) \leq P\left(\frac{1}{n^{3/4}} \sum_{i=1}^{n} \big(W_i - E[W_i]\big) \leq x\right) \asymp \exp\big(-I(x)\sqrt{n}\big),$$

by [28, Theorem 2.3]. It follows that

$$P(B_N^1) \leq \exp\big(-C_2(\alpha)\sqrt{N}\big),$$

for some constant $C_2(\alpha) > 0$. In the same way we may show that

$$P(B_N^2) \leq \exp\big(-C_2(\alpha)\sqrt{N}\big),$$

and by union bound

$$P(B_N) \leq 2\exp\big(-C_2(\alpha)\sqrt{N}\big). \tag{2.4.56}$$

In particular

$$P\big(B_N \mid A_N\big) \leq \frac{P\big(B_N\big)}{P\big(A_N\big)} \leq 2\exp\big(-C_2(\alpha)\sqrt{N}\big) \times \frac{\sqrt{N}}{C(\alpha)} \to 0,$$

as $N \to \infty$, where the last inequality follows from (2.4.54) and (2.4.56). Therefore (2.4.51) holds and the proof is complete. □

2.5 Open problems

In Sections 2.1.1 and 2.1.2 we conjectured that cutoff holds for the symmetric FSEP on the segment and circle respectively. For the segment, Figure 2.4a seems like a useful object in the hunt for cutoff; our original arguments did not have this picture in mind and we believe that the upper bound in cutoff can be achieved by careful consideration of these lattice path dynamics.

On the circle, it is less clear how to tackle this problem. The argument for cutoff of the SSEP on the circle [64] is very different to the arguments we gave here, and it is possible than an analogous argument can be constructed. If we could show that an arbitrary initial condition arrives at the ergodic component in good time, or even arrives in the set of 'reasonable' initial conditions $I_{N,k}^m$ for some $m > 0$ in good time, then this would yield a full pre–cutoff result for the circle.

The following is a list of open questions for the FSEP:

- Does cutoff hold for the symmetric FSEP on the segment or circle? If so, what is the constant for the leading order of the mixing time? How does the constant depend on the density?

- What can we say about the mixing time of the asymmetric or totally asymmetric FSEP on the circle?

- How far do arguments for the FSEP on the segment carry over to the FSEP with *open boundaries* which input and remove particles at the endpoints of the segment? Can we say anything about the mixing time of the FSEP with open boundaries in general?

- A version of the facilitated exclusion process on the d–dimensional hypercube was considered in [92]. Particles may jump to a nearest neighbour provided the jump is facilitated by an occupied nearest neighbour. In two–dimensions we may also consider a facilitated exclusion process in which a particle at site (x_1, x_2) may jump to $(x_1 \pm 1, x_2)$ if the jump is facilitated by a particle at $(x_1 \mp 1, x_2)$. Similarly the particle may jump to $(x_1, x_2 \pm 1)$ if the jump is facilitated by a particle at $(x_1, x_2 \mp 1)$. What can we say for these processes?

We expect that problems involving the asymmetric FSEP on the circle and the process in higher dimensions are hard.

There is some interest in exclusion with extended objects of varying length which move like particles in the SEP. See, for example, [1, 35, 67, 80, 79, 99].

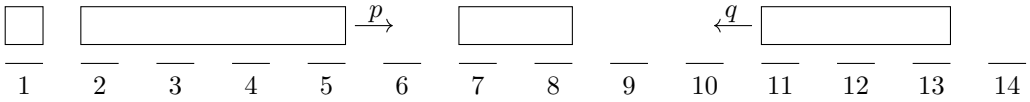


Figure 2.16: Extended objects with SEP dynamics on the segment of length 14. Read from left to right, the objects have length $1, 4, 2$ and 3. The dynamics are such that distinct extended objects cannot occupy the same site.

For extended objects on the segment, the process is in bijection with the usual SEP, since the ordering of the objects does not change in the dynamics (i.e. objects do not jump over each other) and cutoff is an immediate consequence of Theorem 1.1 and Theorem 1.2 in the symmetric and asymmetric case respectively.

We note that the FSEP on the circle restricted to the ergodic component is equivalent to an exclusion process with extended objects in which all objects have length 2. In particular particle–hole objects play the role of the extended objects. A corollary of Theorem 2.3 is that pre–cutoff holds for the symmetric version of this model with a macroscopic density of objects.

Natural questions that arise for the symmetric exclusion process with extended objects on the circle are as follows:

- Does the cutoff phenomenon occur for symmetric process in which all objects have length 2?

- What can we say about the mixing time for symmetric or asymmetric process in which objects have arbitrary lengths?

We expect that the method we used for the mixing time of the symmetric FSEP on the ergodic component of the circle $\mathcal{G}_{N,k}$ should yield pre–cutoff for the symmetric exclusion process with

extended objects when all particles have the same length, and the length is not too large relative to the size of the circle. It may also work when the objects alternate in size, e.g. the lengths of the objects are $1, 2, 1, 2, \ldots$ when read clockwise.

Chapter 3

Sampling the cycle structure of a uniform random permutation

3.1 Background and motivation

In this chapter we consider the problem of efficiently producing a uniform random sample from the set of permutations on N objects with exactly k cycles for some $1 < k < N$.

We present an algorithm with run–time $O(k * \text{polylog} N)$ that samples from the distribution of the multiset of cycle lengths of a uniform random permutation on $[N] = \{1, \ldots, N\}$ conditioned to have k cycles. We refer to the multiset of cycle lengths of a permutation as the cycle structure, for example the permutation $(1, 4, 5)(2)(3)(6, 7)$ has cycle structure $\{3, 2, 1, 1\}$.

Our approach is based on the Feller representation of a random permutation. We design a Markov chain (specifically, an exclusion process) whose stationary distribution projects to the distribution of the Feller representation conditioned on having a given number of cycles. We then show how to obtain efficiently an exact sample from this stationary distribution, using an approach based on *Coupling From The Past* (CFTP).

Recall that Example 1.7.4 in Section 1.7.1 gave an augmented chain whose update scheme tracked the evolution of all copies of the Hardcore model. In the style of Huber [54] and Häggström and Nelander [51] we construct an augmented chain whose update scheme tracks the evolution of all copies of our Markov chain. A version of the algorithm is then obtained via an application of the standard CFTP algorithm, given in Section 1.7, in which the time window looking back into the past is successively doubled to determine coalescence of all copies of the chain.

We also show that the update scheme for the augmented chain has an interesting structure: a sequence of updates $U_1, \ldots, U_t$ in the update scheme achieves coalescence if and only if the sequence with reverse order $U_t, U_{t-1}, \ldots, U_1$ achieves coalescence. This leads to an alternative form of the CFTP algorithm in which the augmented chain runs one way through a sequence of updates until coalescence occurs, and then runs through the same sequence of updates in reverse order to achieve an exact sample from the stationary distribution. This structure extends to a related class of models.

We present a coupling of the cycle structures of a uniform random permutation on $[N]$ with

i cycles for all $i \in [k]$ whose samples have the following consistency property: the cycle structure with $i+1$ cycles is obtained from the cycle structure with i cycles by breaking one of the cycles into two integer parts. In particular we give an explicit construction of a *fragmentation process*. We then present an algorithm which samples from this coupling in expected run–time $O(k^2 \text{polylog} N)$.

As for the previous algorithms, our approach is based on the Feller representation of a random permutation. We give a variation of our previous Markov chain in which particles are assigned priority as well as a collection of $k-1$ projections; the i–th projection from the stationary distribution of this chain projects to the Feller representation of a uniform random permutation conditioned to have $i+1$ cycles. We then construct an augmented chain and update scheme whose dynamics mimic all copies of the chain, and the algorithm is obtained via an application of the CFTP algorithm to the augmented chain.

3.1.1 Uniform random permutations

One of the earliest algorithms for the generation of a uniform random permutation with N elements was described by Durstenfeld [32]. Knuth [58] attributes this algorithm to Fisher and Yates [37], and it is often referred to as the Fisher–Yates shuffle or Knuth shuffle. One way to describe the method is as follows: starting from any initial permutation (for example the identity permutation), we successively multiply by the transpositions $(N, X_1), (N-1, X_2), \ldots, (2, X_{N-1})$, where the X_i are independent and uniform on $[N+1-i]$.

The procedure has $O(N)$ steps, and so (under reasonable assumptions on the cost of manipulations and random number generation involving numbers of size up to N) has run–time $O(N \log N)$. Similarly, the time required even to write down a permutation of $[N]$ is already of order $N \log N$ so one cannot hope to produce a uniform random permutation in a smaller order of time.

However, suppose we are only interested in the cycle structure of a typical permutation. Then we may sample in expected run-time $O(\text{polylog} N)$ using the following "stick-breaking" procedure. We initially sample from $U_1 \sim \text{Uniform}\{1, \ldots, N\}$ for the length of the first cycle. Assuming we have generated $U_1, \ldots, U_i$ for some $i \geq 1$, the length of the $(i+1)$-th cycle is sampled from

$$U_{i+1} \sim \text{Uniform}\left\{1, \ldots, N - \sum_{j=1}^{i} U_j\right\}.$$

Particularly important for our purposes will be the Feller representation of the cycle structure of a uniform random permutation as a function of a collection of independent Bernoulli random variables. Let $p_x = 1/(N+1-x)$ and $B_x \sim \text{Bernoulli}(p_x)$ independently for $x = 1, \ldots, N$. Note that $B_N = 1$ with probability 1. We write a string of length N, divided into blocks so that a block is closed at position x, and if $x < N$ a new block is started at position $x+1$, whenever $B_x = 1$.

To observe that this is the correct distribution, we may consider placing objects in a permutation in positions $x = 1, \ldots, N$ successively to generate the cycle structure. After placing the x–th object we either close a cycle in the permutation by adding a bracket, or we continue the current cycle to the next position. The probability we close a cycle after the x–th position is $(N+1-x)^{-1}$. See Figure 3.1 for an example.

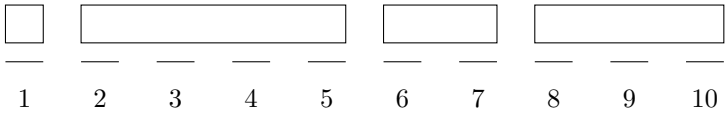


Figure 3.1: The above figure divides the interval $[1, 10]$ into blocks of integer size. In particular this is the partition obtained in the Feller representation if $B_x = 1$ for $x \in \{1, 5, 7, 10\}$ and $B_x = 0$ otherwise. This corresponds to the cycle structure $\{4, 3, 2, 1\}$.

Remark 3.1.1. *We note that several values of the collection of Bernoulli random variables $(B_x)_{x=1}^N$ can give rise to the same cycle structure. In particular the Feller representation is not an injective mapping to the set of cycle structures.*

Crucially the probability of sampling any collection of Bernoulli random variables which gives rise to a fixed cycle structure is the probability of sampling that cycle structure from a uniform random permutation. Therefore to sample the cycle structure of a uniform random permutation it is sufficient to sample from the collection of Bernoulli random variables $(B_x)_{x=1}^N$.

The number of cycles of the permutation is $\sum_{x=1}^N B_x$. In particular we see that the expected number of cycles in a uniform random permutation on N objects is the Nth harmonic number H_N.

For $2 \leq k \leq N$ let

$$\mathcal{P}_k^N = \Big\{ \Pi \subseteq [N-1] : \big|\Pi\big| = k-1 \Big\}. \tag{3.1.1}$$

We interpret $\Pi = \{x_1, \ldots, x_{k-1}\} \subseteq [N-1]$ to be the set of closing positions of the blocks, excluding the Nth site, in the Feller representation.

For $\Pi \in \mathcal{P}_k^N$ we define the probability distribution

$$\mu_{N,k}\big(\Pi\big) = \frac{1}{Z_{N,k}} \prod_{x=1}^{N-1} \Big[p_x \mathbb{1}\big(x \in \Pi\big) + (1-p_x)\mathbb{1}\big(x \notin \Pi\big) \Big], \tag{3.1.2}$$

where $p_x = (N+1-x)^{-1}$ and $Z_{N,k}$ is the normalisation factor. In particular $\mu_{N,k}(\Pi)$ is the probability of observing the set of closing positions Π in the Feller representation conditioned to have k cycles.

We turn to the question of sampling uniformly from the set of permutations on N objects with exactly k cycles. An informal discussion may be found on MathOverflow [86]. The number of such permutations is given by the Stirling number of the first kind $S(N, k)$. These numbers satisfy the recurrence relation

$$S(N, k) = S(N-1, k-1) + (N-1)S(N-1, k),$$

with $S(1, 1) = 1$. There is a natural recursive approach to sampling from this set [7, 104]. Consider the final element N. With probability $S(N-1, k-1)/S(N, k)$, let N be a fixed point, and sample the rest of the permutation uniformly from the set of permutations on $[N-1]$ with $k-1$ cycles. With the complementary probability $(N-1)S(N-1, k)/S(N, k)$, instead sample a permutation on $[N-1]$ with k cycles, and insert the element N into one of its cycles, choosing uniformly between the $N-1$ possibilities.

Such an algorithm needs an initialisation to generate the Stirling numbers $S(r, j)$, for $r \leq N$ and $j \leq k$, requiring $O(Nk)$ steps. After this initialisation, the algorithm requires $O(N)$ steps to generate each sample.

3.1.2 Fragmentation processes

We also consider the problem of efficiently generating the trajectory of a fragmentation process.

Definition 3.1.2. *Given $N \geq 1$, we say that the process $(\Pi_{N,i}, 1 \leq i \leq N)$ with $\Pi_{N,1} = \emptyset$ and $\Pi_{N,i} \in \mathcal{P}_i^N$ for $i \geq 2$ is a fragmentation process if it satisfies the following two properties:*

(1) $\Pi_{N,i} \sim \mu_{N,i}$ *for all* $2 \leq i \leq N$*;*

(2) if $2 \leq i \leq N-1$ *and* $\Pi_{N,i} = \{x_1, \ldots, x_{i-1}\}$*, then* $\Pi_{N,i+1} = \{x_1, \ldots, x_i\}$ *for some* $x_i \in [N-1] \setminus \Pi_{N,i}$*. The Feller representation corresponding to* $\Pi_{N,i+1}$ *is a refinement of the representation corresponding to* $\Pi_{N,i}$ *obtained by splitting one of the blocks into two parts.*

In particular the fragmentation process gives a coupling of the distributions of uniform random permutations with a given number i of cycles, as i varies, with the following consistency property. The permutation with $i+1$ cycles is obtained from the one with i cycles by splitting one cycle into two parts. See Figure 3.2 for an example.

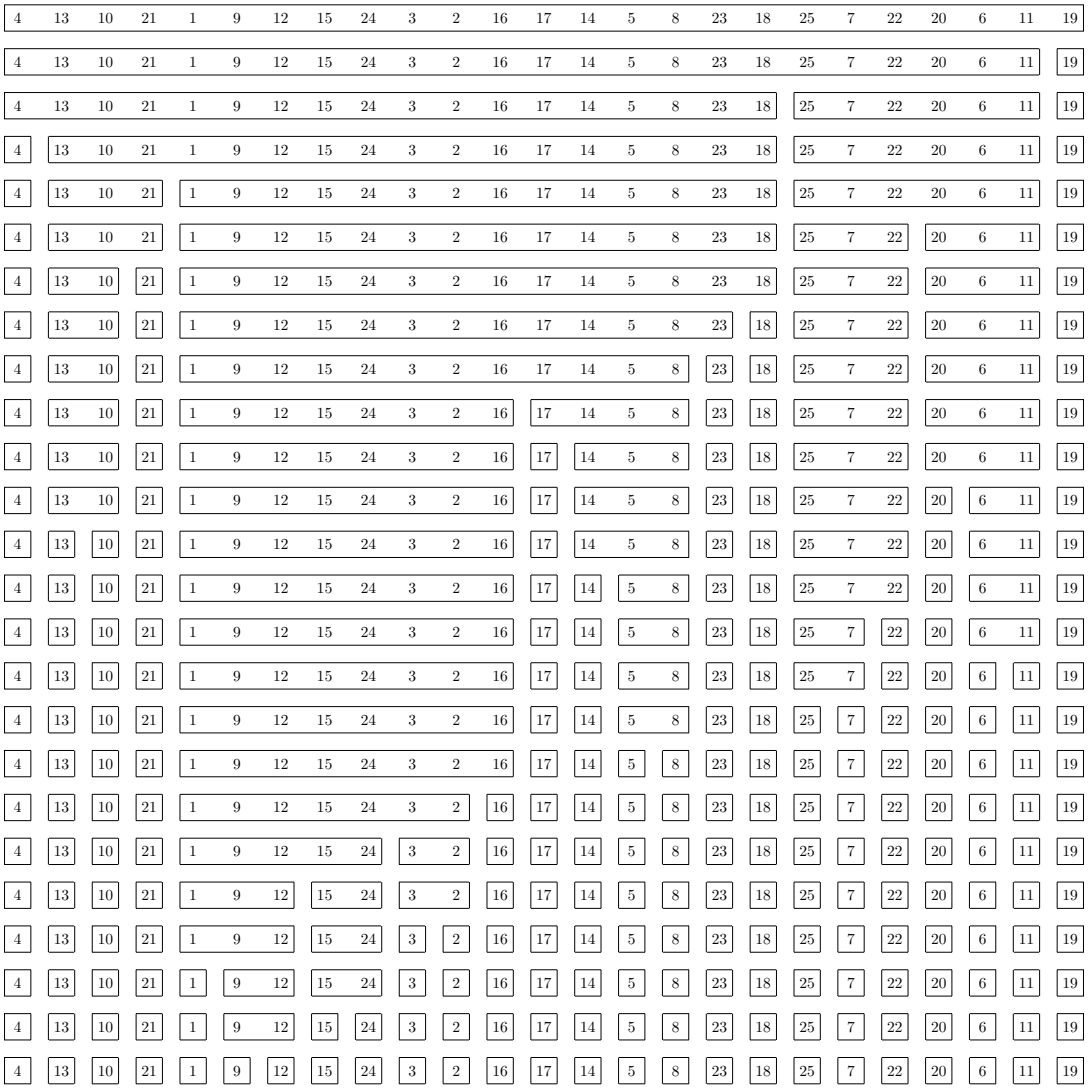


Figure 3.2: The above figure is a simulation of a fragmentation process $(\Pi_{25,i}, 1 \leq i \leq 25)$, where $\Pi_{25,i}$ is represented by the i–th row from the top of the figure. We have populated the fragmentation process with a uniform random permutation on [25] in order to get a sequence of permutations from the sequence of cycle structures. In particular the i–th row from the top is a uniform random permutation on [25] conditioned to have i cycles. The first $i-1$ entries in the set $\{24, 18, 1, 4, 21, 3, 17, 16, 12, 13, 22, 2, 14, 20, 23, 19, 15, 11, 9, 7, 10, 5, 8, 6\}$ gives the the closing positions $\Pi_{25,i} \subseteq [24]$ excluding the 25–th site in the Feller representation.

3.2 Results

Theorem 3.1. *Let $(B_x)_{x=1}^N$ be a collection of independent Bernoulli random variables with $\mathbb{P}[B_x = 1] = 1 - \mathbb{P}[B_x = 0] = 1/(N+1-x)$. There exists an algorithm which samples from*

$(B_x)_{x=1}^N$ *conditioned on* $\sum B_x = k$ *whose random run–time* R_1 *satisfies*

$$\mathbb{E}\big[R_1\big] = O(k * \text{polylog}N). \tag{3.2.1}$$

By Remark 3.1.1, the following corollary is an immediate consequence of Theorem 3.1.

Corollary 3.2. *For* $2 \leq k \leq N$ *there exists an algorithm which exactly samples from the distribution* $\mu_{N,k}$ *in* (3.1.2) *whose random run–time* R_1 *satisfies*

$$\mathbb{E}[R_1] = O\big(k * \text{polylog}N\big).$$

Let $(I_t)_{t\in\mathbb{Z}}$ be a collection of i.i.d. uniform random variables on the set $[k-1]$, and independently, let $(X_t)_{t\in\mathbb{Z}}$ be a collection of i.i.d. random variables so that $\mathbb{P}(X_t = x) = [(N-x)H_{N-1}]^{-1}$ for all $x \in [N-1]$. We call $U_t := (I_t, X_t)$ the t–update.

The following pseudocode gives an algorithm which satisfies Theorem 3.1.

Algorithm (A)

```
1 :  t ← 0
2 :  L = (L_1, ..., L_{k-1}) ← (0, ..., 0)
3 :  while L_1 × L_2 × ··· × L_{k-1} = 0 :
4 :      L ← (0, ..., 0)
5 :      generate U_{-s} = (I_{-s}, X_{-s}) for t ≤ s < max(2t, k-1)
6 :      t ← max(2t, k-1)
7 :      for s = t-1 down to 0
8 :          (i, x) ← (I_{-s}, X_{-s})
9 :          if x not in L :
10 :             L_i ← x
11 :         else:
12 :             find j : L_j = x          #a unique such j exists
13 :             L_j ← L_i
14 :             L_i ← x
15 :  output {L_1, ..., L_{k-1}}
```

A priori it is not clear that Algorithm (A) will terminate. Each nonzero entry in the list L will represent a particle position in an exclusion process, so that every nonzero entry in the list appears exactly once, and the index j in line 12 is unique. Algorithm (A) is an application of the CFTP algorithm; we repeatedly double the number of t–updates, storing and re–using them. We discuss implementation of Algorithm (A) in Section 3.3.

Algorithm (A) outputs a set of indices $\{L_1, \ldots, L_{k-1}\}$. These are the indices of the Bernoulli random variables which take value 1. To obtain a cycle structure:

- put the indices in increasing order as $L'_1, \ldots, L'_{k-1}$;
- take the successive differences $L'_1, L'_2 - L'_1, \ldots, L'_{k-1} - L'_{k-2}, N - L'_{k-1}$, to get a multiset of cycle lengths.

The following theorem is for fragmentation processes.

Theorem 3.3. *For $2 \leq k \leq N$ there exists an algorithm which samples a joint distribution $(\Pi_{N,i})_{i=1}^{k}$ with the properties of a fragmentation process in Definition 3.1.2. The random run–time R_2 of this algorithm satisfies*

$$\mathbb{E}[R_2] = O\big(k^2 * \text{polylog}N\big).$$

We give an algorithm which satisfies Theorem 3.3 in Section 3.4.2.

3.2.1 Processes

We consider three discrete–time Markov chains, representing systems consisting of ℓ particles moving between M different sites. We will call them the Unlabelled, Labelled and Priority processes. The Labelled process and the Priority process will be projected to the Unlabelled process.

For the Unlabelled process we consider the state space

$$\Omega_{M,\ell} = \Big\{ \xi \in \{0,1\}^{[M]} \;\Big|\; \sum_{x=1}^{M} \xi(x) = \ell \Big\}.$$

If $\xi(x) = 1$ we say there is a particle at site x, and if $\xi(x) = 0$ we say there is a hole at site x.

The dynamics of the Unlabelled process are as follows. At each time step:

- Choose a particle uniformly at random, and choose independently a site x with probability q_x;
- if site x is empty, then move the chosen particle from its current position to site x.

See Figure 3.3a for an example. A straightforward check of the detailed balance equations yields that the stationary distribution $\nu_{M,\ell}$ of the Unlabelled process is given by

$$\nu_{M,\ell}(\xi) = \frac{1}{Z'_{M,\ell}} \prod_{x=1}^{M} \big(q_x \xi(x) + (1-q_x)[1-\xi(x)]\big),$$

where $Z'_{M,\ell}$ is the normalisation factor. Setting $\ell = k-1$, $M = N-1$, $q_x = [(M+1-x)H_M]^{-1}$ and $\Pi = \{x : \xi(x) = 1\}$ we observe that

$$\nu_{M,\ell}(\xi) = \mu_{N,k}\big(\Pi\big),$$

where $\mu_{N,k}$ is the distribution from (3.1.2).

For the Labelled and Priority process we consider the state space

$$\Sigma_{M,\ell} = \Big\{ \zeta \in \{0,1,\ldots,\ell\}^{[M]} \;\Big|\; \sum_{x=1}^{M} \mathbb{1}(\zeta(x) = r) = 1 \text{ for all } r \in [\ell] \Big\}.$$

If $\zeta(x) = r \in [\ell]$ we say that particle r is at site x, and if $\zeta(x) = 0$ we say there is a hole at site x.

The dynamics of the Labelled process are as follows. At each time step:

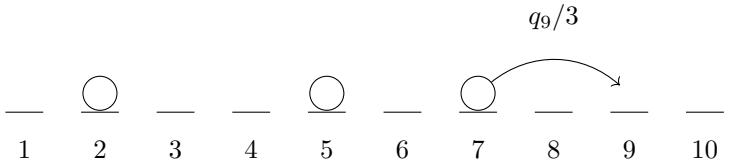
- Choose a particle label r uniformly at random, and choose independently a site x with probability q_x.
- Move particle r to site x. If site x was previously occupied by another particle $r' \neq r$, move that particle to the previous location of particle r.

See Figure 3.3b for an example.

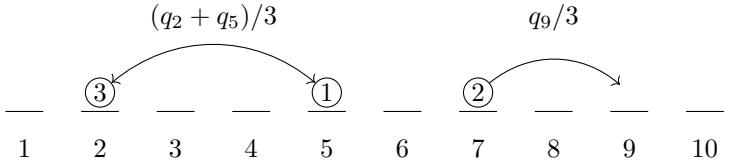
The dynamics of the Priority process are as follows. At each time step:

- Choose a particle label r uniformly at random, and choose independently a site x with probability q_x.
- If site x is vacant, or occupied by a particle with label $r' > r$, move particle r to site x, and move particle r' (if it exists) to the previous location of particle r.

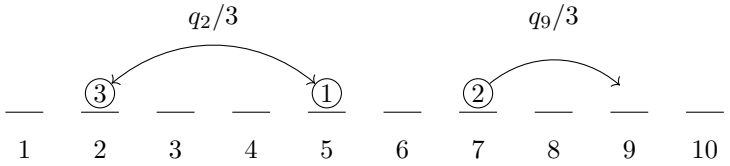
Note the difference between the Labelled and the Priority processes: if the target site x contains a particle of label $r' < r$, then the two particles swap in the Labelled process, but no change occurs in the Priority process. In particular particles are assigned a priority, and higher priority particles can't be displaced by lower priority particles. See Figure 3.3c for an example.



(a) A configuration for the Unlabelled process on $\Omega_{10,3}$. In the dynamics the particle at site 7 is chosen to jump to site 9 with probability $q_9/3$.



(b) A configuration for the Labelled process on $\Sigma_{10,3}$. In the dynamics, particle 2 and site 9 are chosen with probability $q_9/3$ and particle 2 is moved to site 9. If particle 1 and site 2 are chosen or particle 3 and site 5 are chosen in the dynamics with probability $(q_2 + q_5)/3$, then particles 1 and 3 are swapped.



(c) A configuration for the Priority process on $\Sigma_{10,3}$. In the dynamics, particle 2 and site 9 are chosen with probability $q_9/3$ and particle 2 is moved to site 9. If particle 1 and site 2 are chosen with probability $q_2/3$, then particle 1 and 3 are swapped. If particle 3 is chosen in the dynamics, then neither particle 1 or particle 2 can be moved.

Figure 3.3: Figures 3.3a, 3.3b and 3.3c give the dynamics of the Unlabelled, Labelled and Priority processes respectively.

Remark 3.2.1. *If we run the Labelled process on $\Sigma_{M,\ell}$, ignoring particle labels, then we recover the dynamics of the Unlabelled process $\Omega_{M,\ell}$. Moreover, if we run the Priority process on $\Sigma_{M,\ell}$, regarding all particles with label at most i as particles and all particles with label greater than i as holes, then we recover the dynamics of the Unlabelled process on $\Omega_{M,i}$.*

In particular, to sample from the stationary distribution $\nu_{M,\ell}$ of the Unlabelled process, it is sufficient to sample from the stationary distribution of the Labelled process and project. The key point is that the Labelled process has an easier CFTP–type construction.

Similarly we may sample from $\nu_{M,i}$ by sampling from the stationary distribution of the Priority process and applying the projection from Remark 3.2.1. The point here is that the consistency property for the fragmentation process from Definition 3.1.2 holds for the projections of the Priority process.

We generalise the t–updates from the previous section. Let $(I_t)_{t\in\mathbb{Z}}$ be a collection of i.i.d uniform random variables on $[\ell]$, and independently, let $(X_t)_{t\in\mathbb{Z}}$ be a collection of i.i.d. random variables so that $\mathbb{P}(X_t = x) = [(M+1-x)H_M]^{-1}$ for all $x \in [M]$. As before let $U_t = (I_t, X_t) \in$

$[\ell] \times [M]$ denote the t–update.

Remark 3.2.2. *We may couple the Labelled process starting from all initial configurations simultaneously by applying the same collection of t–updates $(U_t)_{t\geq 0}$ to every copy of the chain. This is an example of a grand coupling, see [70][Chapter 5] for more details.*

3.2.2 Tracking sets for the Labelled process

We now introduce the idea of a *tracking set*, which represents the information acquired as different copies of the Labelled process evolve according to the same sequence of updates.

Definition 3.2.3. *Let ℓ and M be positive integers, and let A be a subset of $[\ell] \times [M]$. We say that A is an (ℓ, M)–tracking set if whenever (r_1, x_1) and (r_2, x_2) are distinct members of A, then $r_1 \neq r_2$ and $x_1 \neq x_2$.*

In particular no element of A shares its first or second coordinate with any other element in A, and we say that the elements of A do not overlap. We let $A_1 = \{r : (r, x) \in A \text{ for some } x\}$ and $A_2 = \{x : (r, x) \in A \text{ for some } r\}$ be the projections of A onto the first and second coordinate respectively.

The empty set and any singleton $\{\alpha\}$ with $\alpha \in [\ell] \times [M]$ are trivially (ℓ, M)–tracking sets. Moreover, any (ℓ, M)–tracking set has size at most $\min(\ell, M)$ and henceforth we assume that $\ell < M$.

Since the elements of an (ℓ, M)–tracking set A do not overlap, it holds that $|A| = |A_1| = |A_2|$.

In the style of Huber [54] and Häggström and Nelander [51], we consider an augmented state space and a Markov chain on the augmented space whose dynamics mimic the evolution of the grand coupling of the Labelled process (see Remark 3.2.2). Let

$$\widetilde{\Sigma}_{M,\ell} = \big\{A \subset [\ell] \times [M] : A \text{ is an } (\ell, M)\text{–tracking set}\big\}, \tag{3.2.2}$$

where $A \in \widetilde{\Sigma}_{M,\ell}$ represents the set of configurations in $\Sigma_{M,\ell}$ given by

$$\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in A\big\},$$

i.e. the set of possible configurations in the Labelled process with particle r at site x for all $(r, x) \in A$.

We now define the effect of an update on an (ℓ, M)–tracking set. The interpretation is that an update $\alpha = (r, x)$ causes particle r to jump to site x. The effect of this update on the tracking set depends on whether we currently know the position of particle r, and whether we currently know the contents of site x.

Definition 3.2.4. *Let A be an (ℓ, M)–tracking set and $\alpha = (r, x) \in [\ell] \times [M]$. We update A with α to obtain the set $A \otimes \alpha$ as follows:*

***Good**: If $r \notin A_1$ and $x \notin A_2$ then we define $A \otimes (r, x) = A \cup \{(r, x)\}$. We say that (r, x) is an $\otimes$–good update for A;*

***Jump**: if $r \in A_1$ and $x \notin A_2$, so that $(r, x') \in A$ with $x' \neq x$, then we define*

$$A \otimes (r, x) = \{(r, x)\} \cup \big(A \setminus \{(r, x')\}\big).$$

We say that (r, x) is an $\otimes$–jump update for A;

***Label**: If $r \notin A_1$ and $x \in A_2$, so that $(r', x) \in A$ with $r' \neq r$, then we define*

$$A \otimes (r, x) = \{(r, x)\} \cup \big(A \setminus \{(r', x)\}\big).$$

We say that (r, x) is an $\otimes$–label update for A;

***Swap**: If $r \in A_1$ and $x \in A_2$ but $(r, x) \notin A$, so that $(r, x'), (r', x) \in A$ with $r' \neq r$ and $x' \neq x$, then we define*

$$A \otimes (r, x) = \{(r, x), (r', x')\} \cup \big(A \setminus \{(r, x'), (r', x)\}\big).$$

We say that (r, x) is an $\otimes$–swap update for A;

***Null**: If $(r, x) \in A$ we set $A \otimes (r, x) = A$. We say that (r, x) is an $\otimes$–null update for A.*

The set $A \otimes \alpha$ is necessarily an (ℓ, M)–tracking set.

We consider the Markov chain on $\widetilde{\Sigma}_{M,\ell}$ given by

$$\Theta_{s,t} = \emptyset \otimes U_{s+1} \otimes \cdots \otimes U_t, \tag{3.2.3}$$

for all $s, t \in \mathbb{Z}$ with $s \leq t$. Let $\Theta^1_{s,t}$ and $\Theta^2_{s,t}$ denote the projections onto the first and second coordinates of $\Theta_{s,t}$ respectively. The following lemma says that the chain $\Theta_{s,t}$ precisely represents the evolution of all copies of the Labelled process.

Lemma 3.2.5. *Let the collection of t–updates $U_{s+1}, \ldots, U_t$ be given and let $(\zeta_n)_s^t$ denote the trajectory of the Labelled process with initial condition $\zeta_s \in \Sigma_{M,\ell}$ run through the updates $U_{s+1}, \ldots, U_t$. Then*

$$\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t}\big\} = \big\{\zeta_t : \zeta_s \in \Sigma_{M,\ell}\big\}. \tag{3.2.4}$$

We will prove Lemma 3.2.5 in Section 3.4.

Remark 3.2.6. *We note that Algorithm (A) is an application of the CFTP algorithm to the Labelled process. In particular, after each while loop, the set $\{(i, L_i) : L_i \neq 0\}$ is given by*

$$\Theta_{-t,0} = \emptyset \otimes U_{-t+1} \otimes U_{-t+2} \otimes \cdots \otimes U_0.$$

If $|\Theta_{-t,0}| = \ell$, all copies of the Labelled process have coalesced and, by Remark 3.2.1 and [91][Theorem 1], the output $L = \Theta^2_{-t,0}$ of Algorithm (A) is a sample from $\nu_{M,\ell}$.

We note that when $\ell = M$ coalescence actually occurs when $|\Theta_{-t,0}| = M - 1$, since there is only a single location that the remaining particle (not contained in the tracking set) can occupy. When $\ell < M$ and $|\Theta_{s,t}| < \ell$, the set on the LHS of (3.2.4) has size at least 2, although the statement of Lemma 3.2.5 also holds when $\ell = M$.

3.2.3 Reversing the order of updates for the Labelled process

We note a surprising property of the update procedure for the Labelled process. Suppose we fix a finite sequence of updates. The next result says that the number of particles whose position is known after applying the given sequence is the same as when the sequence is applied in reverse.

Theorem 3.4. *Let $s < t$ and consider a sequence of updates $U_{s+1}, \ldots, U_t \in [\ell] \times [M]$. Recall that*

$$\Theta_{s,t} = \emptyset \otimes U_{s+1} \otimes \cdots \otimes U_{t-1} \otimes U_t.$$

Define also

$$\widehat{\Theta}_{s,t} = \emptyset \otimes U_t \otimes U_{t-1} \otimes \cdots \otimes U_{s+1}. \tag{3.2.5}$$

Then

$$\big|\Theta_{s,t}\big| = \big|\widehat{\Theta}_{s,t}\big|. \tag{3.2.6}$$

If we run all copies of the Labelled process through the same t–updates $U_{s+1}, U_2, \ldots, U_t$, then the number of particles whose locations agree in every copy is $|\Theta_{s,t}|$. Theorem 3.4 says that $|\Theta_{s,t}|$ is also the number of particles whose locations agree when all copies of the Labelled process are run through the t–updates in reverse order $U_t, U_{t-1}, \ldots, U_1$.

In particular, $|\Theta_{s,t}|$ is equal to ℓ if and only if coalescence has occurred by time t for the grand coupling started at time s. Theorem 3.4 tells us that this happens precisely when the reverse sequence of updates gives coalescence, i.e. when $|\widehat{\Theta}_{s,t}| = \ell$.

Recall from Example 1.7.3 in Section 1.7.1 that running all copies of the chain forwards in time until coalescence does *not* necessarily result in a sample from the stationary distribution. However, Theorem 3.4 says something related and surprising. If

$$T^{\otimes} = \inf\big\{t \geq 0 : |\Theta_{0,t}| = \ell\big\}, \tag{3.2.7}$$

is the coalescence time of all copies of the Labelled process, then $\widehat{\Theta}_{0,T^{\otimes}}$ does give a sample from the stationary distribution of the Labelled process, by an application of the CFTP algorithm. The following pseudocode gives an alternative form of Algorithm (A) via an application of Theorem 3.4.

Algorithm (A^*):

```
1 :  t ← 1
2 :  L = (L_1, ..., L_{k-1}) ← (0, ..., 0)
3 :  while L_1 × L_2 × ··· × L_{k-1} = 0 :
4 :      generate U_t = (I_t, X_t)
5 :      (i, x) ← (I_t, X_t)
6 :      t ← t + 1
7 :      if x not in L :
8 :          L_i ← x
9 :      else :
10 :         find j : L_j = x        #a unique such j exists
11 :         L_j ← L_i
12 :         L_i ← x
13 :
14 :
15 :     (i, x) ← (I_s, X_s)
16 :     if x not in L :
17 :         L_i ← x
18 :     else :
19 :         find j : L_j = x        #a unique such j exists
20 :         L_j ← L_i
21 :         L_i ← x
22 :
```

The while loop in lines 3–12 of Algorithm (A^*) generates $\Theta_{0,t}$, and the for loop in lines 14–21 generates $\widehat{\Theta}_{0,t}$. In fact, line 13 may be removed as $\Theta_{0,T^\otimes}$ represents a configuration in $\Sigma_{M,\ell}$ and therefore

$$\widehat{\Theta}_{0,T^\otimes} = \Theta_{0,T^\otimes} \otimes U_{T^\otimes} \otimes U_{T^\otimes-1} \otimes \cdots \otimes U_1,$$

as the initial condition given by $\Theta_{0,T^\otimes}$ must also coalesce to the configuration represented by $\widehat{\Theta}_{0,T^\otimes}$ in the reverse sequence.

Remark 3.2.7. *In Algorithm (A^*) we will read each update exactly twice, once when we generate $\Theta_{0,t}$ and once when we generate $\widehat{\Theta}_{0,t}$. An advantage of Algorithm (A^*) is that it generates exactly the number of samples we need for coalescence, whereas Algorithm (A) may generate up to twice more than needed.*

We note that although $\Theta_{s,t}$ and $\widehat{\Theta}_{s,t}$ have the same size, $\Theta^1_{s,t}$ need not agree with $\widehat{\Theta}^1_{s,t}$, and $\Theta^2_{s,t}$ need not agree with $\widehat{\Theta}^2_{s,t}$.

Example 3.2.8. *If*

$$\big(U_1, U_2, U_3\big) = \big((1,a),(2,a),(2,b)\big),$$

then

$$\Theta_{0,3} = \big\{(2,b)\big\},\ \widehat{\Theta}_{0,3} = \big\{(1,a)\big\}.$$

Remark 3.2.9. *Both Lemma 3.2.5 and Theorem 3.4 hold regardless of the distribution of the t–updates, and we may assume that $(U_t)_{t\in\mathbb{Z}}$ is an i.i.d. collection of random variables so that $\mathbb{P}[U_t = (r,x)] = \pi(r,x)$ for some distribution π supported on $[\ell]\times[M]$.*

This gives a class of exclusion processes in which particle r is chosen to jump to site x (with replacement) at each time–step with probability $\pi(r,x)$, and Theorem 3.4 together with Lemma 3.2.5 says that all copies of such an exclusion process run through a sequence of t–updates coalesce if and only if they coalesce when run through the same sequence of t–updates in reverse order.

3.3 Remarks about sampling

In order to check that Algorithm (A) (or Algorithm (A^*)) satisfies Theorem 3.1, we check the following conditions:

(a) the output of the Algorithm (A) and Algorithm (A^*) is a sample from $\mu_{N,k}$;

(b) we may generate a t–update in expected run–time $O(\text{polylog}N)$;

(c) We may update the list L in Algorithm (A) and (A^*) in run–time $O(\text{polylog}N)$;

(d) the expected number of updates we must generate to terminate the while loop in Algorithm (A) and (A^*) is $O(k * \text{polylog}N)$.

It is straightforward to see that the above conditions give that Algorithm (A) satisfies Theorem 3.1.

Condition (a) follows by Remark 3.2.6. To check condition (b), we assume that we have access to some random number generator which produces a sequence of i.i.d. Uniform$[0,1]$ random variables in $O(1)$ time. In practice this may not be true; random number generators (RNG) are often not truly random and numbers are produced by some deterministic procedure.

It is straightforward to see that we may obtain a sample from I_t in time $O(\log k)$. If we take a sample u from a Uniform$[0,1]$ random variable, then $\lceil \ell u \rceil$, the least integer greater than ℓu, is a sample from I_t. To determine $\lceil \ell u \rceil$, the expected number of bits of u we need to inspect is of order $\log(k)$.

To obtain a sample from X_t we may use rejection sampling. We consider the function $f : [0,M] \to \mathbb{R}$ given by $f(s) = 1/s$. We generate a sample from the random variable $S \sim$ Uniform$[0,M]$, and conditional on $S = s$ we generate a sample y from the random variable $Y \sim$ Uniform$[0, f(s)]$. If $y < f(\lceil s \rceil)$ then we may take $X_t = M + 1 - \lceil s \rceil$. Otherwise we reject and repeat the procedure. The probability we reject is $1 - (H_M / \log M)$, which converges to zero as $M \to \infty$. In particular the expected number of trials of the above procedure is $O(1)$ (in fact it is $1 + o(1)$), and each trial takes expected time $O(\text{polylog}N)$ time. Therefore we may generate a t–update in expected time $O(\text{polylog}N)$.

To check condition (c), we need a method to quickly lookup whether there exists a j such that $L_j = x$; see, for example, line 9 and 11 in Algorithm (A). In particular we want a data structure, indexed by x, which indicates all nonzero entries of the list L. We will also need to (possibly) delete and insert elements to this data structure according to how we update the list L. One data structure that does this is a self–balancing binary search tree, and each lookup, insertion and deletion operation in the self–balancing binary search tree takes at most order $\log k$ operations when the structure contains at most k objects. See, for example, [59, Section 6.2] and the references therein.

To prove Theorem 3.1 it remains to check condition (d). In Section 3.4 we check condition (d) holds.

3.4 Proofs

Proof of Theorem 3.1. By the remarks in Section 3.3, it remains to show that the expected number of applications of t–updates is $O(\ell * \text{polylog}M)$, and by Remark 1.7.2 the number of applications of t–updates in Algorithm (A) is bounded by $4T^{\otimes}$, where $T^{\otimes}$ is the coalescence time in (3.2.7). In particular, it is sufficient to bound $\mathbb{E}[T^{\otimes}]$.

Suppose that $|\Theta_{0,t}| = m < \ell$. Observe that $I_{t+1} \notin \Theta^1_{0,t}$ with probability $(\ell - m)/\ell$ and independently $X_{t+1} \notin \Theta^2_{0,t}$ with probability at least $\sum_{x=1}^{M-m} q_x$. Then the probability that U_{t+1} is an $\otimes$–good update for $\Theta_{0,t}$ is at least

$$\lambda_m = \frac{\ell - m}{\ell} \sum_{x=1}^{M-m} q_x.$$

Noting that every $\otimes$–good update increases the size of the tracking set by 1, and no update ever decreases the size of the tracking set, the time for the $|\Theta_{0,t}|$ to increase from size m to size $m+1$ is dominated by a geometric random variable $\text{Geom}(\lambda_m)$ with mean $1/\lambda_m$. Therefore

$$\mathbb{E}\big[T^{\otimes}\big] \le \sum_{m=0}^{\ell-1} \frac{1}{\lambda_m} = \sum_{m=0}^{\ell-1} \frac{\ell}{\ell - m} \left(\sum_{x=1}^{M-m} q_x \right)^{-1}.$$

In particular

$$\begin{aligned}
\mathbb{E}\big[4T^{\otimes}\big] &\le 4\ell \sum_{m=0}^{\ell-1} \frac{H_M}{\ell - m} \left(\sum_{x=1}^{M-m} \frac{1}{M+1-x} \right)^{-1} \\
&= 4\ell \sum_{m=0}^{\ell-1} \frac{H_M}{(\ell - m)(H_M - H_m)} \\
&\le \frac{4\ell H_\ell H_M}{H_M - H_\ell} = O\big(\ell * \log(M)^2\big),
\end{aligned}$$

as $M \to \infty$, which completes the proof of Theorem 3.1. □

We now prove Lemma 3.2.5.

Proof of Lemma 3.2.5. We proceed by induction on $t \ge s$. For the base case $t = s$, $\Theta_{s,s} = \emptyset$ and both the LHS and RHS of (3.2.4) are given by $\Sigma_{M,\ell}$. Assume for an induction that

$$\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t}\big\} = \big\{\zeta_t : \zeta_s \in \Sigma_{M,\ell}\big\}, \tag{3.4.1}$$

for $t \ge s$.

Let $U_{t+1} = (r_1, x_1)$, then we will check that (3.4.1) holds for $t+1$ when (r_1, x_1) is an $\otimes$–null $\otimes$–swap, $\otimes$–label, $\otimes$–jump and $\otimes$–good update separately.

⊗–null update:

It is straightforward to see that (3.4.1) holds at time $t+1$ when (r_1, x_1) is an ⊗–null update: the tracking set $\Theta_{s,t}$ is unchanged when U_{t+1} is applied and, by the induction hypothesis at time t, particle r is at site x in all copies of the Labelled process so that no changes occur when U_{t+1} is applied.

⊗–swap update:

If (r_1, x_1) is an ⊗–swap update then $(r_1, y), (r', x_1) \in \Theta_{s,t}$, and in every copy of the Labelled process $\zeta_t(y) = r_1$ and $\zeta_t(x_1) = r'$ by the induction hypothesis at time t. In the Labelled process dynamics $\zeta_{t+1} = \zeta_t^{x_1,y}$, i.e. the configuration given by swapping the values in ζ_t at x_1 and y.

Moreover, by Definition 3.2.4

$$\Theta_{s,t+1} = \{(r_1, x_1), (r', y)\} \cup \big(\Theta_{s,t} \setminus \{(r_1, y), (r', x_1)\}\big),$$

and

$$\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t+1}\big\} = \big\{\zeta_t^{x_1,y} : \zeta_s \in \Sigma_{M,\ell}\big\} = \big\{\zeta_{t+1} : \zeta_s \in \Sigma_{M,\ell}\big\},$$

which gives (3.4.1) at time $t+1$.

⊗–label update:

If (r_1, x_1) is in an ⊗–label update, then in every copy of the Labelled process $\zeta_t(x_1) = r'$ for $r' \neq r_1$ with $(r', x_1) \in \Theta_{s,t}$ by the induction hypothesis at time t. Since $r_1 \notin \Theta^1_{s,t}$ there exists a set of configurations with $\zeta_t(y) = r_1$ for all $y \in [M] \setminus \Theta^2_{s,t}$, and $\zeta_{t+1} = \zeta_t^{x_1,y}$ for each configuration in this set.

In particular,

$$\begin{aligned}\big\{\zeta_{t+1} : \zeta_s \in \Sigma_{M,\ell}\big\} &= \bigcup_{y \in [M] \setminus \Theta^2_{s,t}} \big\{\zeta_t^{x_1,y} : \zeta_t(y) = r_1\big\} \\ &= \bigcup_{y \in [M] \setminus \Theta^2_{s,t}} \big\{\zeta^{x_1,y} \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t},\, \zeta(y) = r_1\big\},\end{aligned} \tag{3.4.2}$$

where the last equality follows from the induction hypothesis at time t. By Definition 3.4.2

$$\Theta_{s,t+1} = \{(r_1, x_1)\} \cup \big(\Theta_{s,t} \setminus \{(r', x_1)\}\big).$$

In particular $\Theta^2_{s,t} = \Theta^2_{s,t+1}$, and it holds that

$$\begin{aligned}&\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t+1}\big\} \\ &\quad= \bigcup_{y \in [M] \setminus \Theta^2_{s,t+1}} \big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t+1},\, \zeta(y) = r'\big\} \\ &\quad= \bigcup_{y \in [M] \setminus \Theta^2_{s,t}} \big\{\zeta^{x_1,y} \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r, x) \in \Theta_{s,t},\, \zeta(y) = r_1\big\}.\end{aligned} \tag{3.4.3}$$

Therefore (3.4.1) holds at time $t+1$ by (3.4.2) and (3.4.3).

⊗–jump update:

If (r_1, x_1) is an ⊗–jump update then in every copy of the Labelled process $\zeta_t(r_1) = y$ for $y \neq x_1$ and $(r_1, y) \in \Theta_{s,t}$ by the induction hypothesis at time t. Moreover, $x_1 \notin \Theta^2_{s,t}$ and there

exists a set of configurations with $\zeta_t(x_1) = r'$ for all $r' \in \{0\} \cup [\ell] \setminus \Theta^1_{s,t}$. In the Labelled process dynamics, each configuration belonging to this set is updated to $\zeta_{t+1} = \zeta_t^{x_1,y}$, and

$$\begin{aligned}\big\{\zeta_{t+1} : \zeta_s \in \Sigma_{M,\ell}\big\} &= \bigcup_{r' \in \{0\}\cup[\ell]\setminus\Theta^1_{s,t}} \big\{\zeta_t^{x_1,y} : \zeta_t(x_1) = r'\big\} \\ &= \bigcup_{r' \in \{0\}\cup[\ell]\setminus\Theta^1_{s,t}} \big\{\zeta^{x_1,y} \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t},\, \zeta(x_1) = r'\big\},\end{aligned} \tag{3.4.4}$$

where the last equality follows from the induction hypothesis at time t. By Definition 3.2.4

$$\Theta_{s,t+1} = \big\{(r_1, x_1)\big\} \cup \big(\Theta_{s,t} \setminus \big\{(r_1, y)\big\}\big).$$

In particular $\Theta^1_{s,t} = \Theta^1_{s,t+1}$ and it holds that

$$\begin{aligned}&\big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t+1}\big\} \\ &\quad= \bigcup_{r' \in \{0\}\cup[\ell]\setminus\Theta^1_{s,t+1}} \big\{\zeta \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t+1},\, \zeta(y) = r'\big\} \\ &\quad= \bigcup_{r' \in \{0\}\cup[\ell]\setminus\Theta^1_{s,t}} \big\{\zeta^{x_1,y} \in \Sigma_{M,\ell} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t},\, \zeta(x_1) = r'\big\}.\end{aligned} \tag{3.4.5}$$

Therefore (3.4.1) holds at time $t+1$ by (3.4.4) and (3.4.5).

$\otimes$–good update:

If (r_1, x_1) is an $\otimes$–good update then $r_1 \notin \Theta^1_{s,t}$ and $x_1 \notin \Theta^2_{s,t}$. By the induction hypothesis at time t there exists a set of configurations with $\zeta_t(x_1) = r_1$; in the Labelled process dynamics these configurations are unchanged and $\zeta_{t+1} = \zeta_t$. This set of configurations is exactly represented by $\Theta_{s,t+1}$ by the induction hypothesis at time t and Definition 3.2.4.

Moreover, for all $y \in [M] \setminus (\Theta^2_{s,t} \cup \{x_1\})$ and all $r' \in \{0\} \cup [\ell] \setminus (\Theta^1_{s,t} \cup \{r_1\})$, there exists a set of configurations with $\zeta_t(y) = r_1$ and $\zeta_t(x_1) = r'$, and in the Labelled process dynamics each of these configurations is updated so that $\zeta_{t+1} = \zeta_t^{x_1,y}$. Let

$$\begin{aligned}M_1 &= [M] \setminus \big(\Theta^2_{s,t} \cup \{x_1\}\big), \\ L_1 &= \{0\} \cup [\ell] \setminus \big(\Theta^1_{s,t} \cup \{r_1\}\big),\end{aligned}$$

then

$$\begin{aligned}&\big\{\zeta_{t+1} : \zeta_s \in \Sigma_{M,\ell},\, \zeta_t(x_1) \neq r_1\big\} \\ &\quad= \bigcup_{y \in M_1} \bigcup_{r' \in L_1} \big\{\zeta_t^{x_1,y} : \zeta_t(y) = r_1,\, \zeta_t(x_1) = r'\big\} \\ &\quad= \bigcup_{y \in M_1} \bigcup_{r' \in L_1} \big\{\zeta^{x_1,y} : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t},\, \zeta(y) = r_1,\, \zeta(x_1) = r'\big\} \\ &\quad= \bigcup_{y \in M_1} \bigcup_{r' \in L_1} \big\{\zeta : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t},\, \zeta(x_1) = r_1,\, \zeta(y) = r'\big\} \\ &\quad= \big\{\zeta : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t},\, \zeta(x_1) = r_1\big\} \\ &\quad= \big\{\zeta : \zeta(x) = r \text{ for all } (r,x) \in \Theta_{s,t+1}\big\},\end{aligned} \tag{3.4.6}$$

where the second equality follows from the induction hypothesis at time t, and the last equality follows from Definition 3.2.4. □

3.4.1 Proof of Theorem 3.4

Let A be an (ℓ, M)–tracking set and $\beta_1, \dots, \beta_n \in [\ell] \times [M]$. To avoid the proliferation of brackets we write

$$A \otimes \beta_1 \otimes \cdots \otimes \beta_n := \{\cdots\{\{A \otimes \beta_1\} \otimes \beta_2\} \cdots\} \otimes \beta_i\} \cdots\} \otimes \beta_n,$$

for multiple applications of the $\otimes$–updates in Definition 3.2.4. Before proving Theorem 3.4 we state two lemmas and a definition.

Lemma 3.4.1. *let $1 \leq n \leq \ell \leq M$, and A and $B = \{\beta_1, \dots, \beta_n\}$ be (ℓ, M)–tracking sets. If $\sigma \in S_n$ is a permutation on $[n]$ then*

$$A \otimes \beta_1 \otimes \cdots \otimes \beta_n = A \otimes \beta_{\sigma(1)} \otimes \cdots \otimes \beta_{\sigma(n)}. \tag{3.4.7}$$

We note that (3.4.7) does not necessarily hold if the elements of B are allowed to overlap (i.e. if we remove the assumption that the set B is a tracking set). Recalling Example 3.2.8, if $A = \emptyset$ and $B = \{(1,a), (2,a), (2,b)\}$ then

$$\emptyset \otimes (1,a) \otimes (2,a) \otimes (2,b) = \{(2,b)\},$$
$$\emptyset \otimes (2,b) \otimes (2,a) \otimes (1,a) = \{(1,a)\},$$

and (3.4.7) does not hold.

Lemma 3.4.1 motivates the following definition.

Definition 3.4.2. *Let A and $B = \{\beta_1, \dots, \beta_n\}$ be two (ℓ, M)–tracking sets with $n \geq 1$. We define*

$$A \otimes B = A \otimes \beta_1 \otimes \cdots \otimes \beta_n,$$

and set $A \otimes \emptyset = \emptyset \otimes A = A$.

By Lemma 3.4.1 this is well–defined, as updating A using the elements of the tracking set B, in any order, results in the same set.

Lemma 3.4.3. *Let A and B be two (ℓ, M)–tracking sets and $\gamma \in [\ell] \times [M]$, then*

$$|(A \otimes \gamma) \otimes B| = |A \otimes (B \otimes \gamma)|. \tag{3.4.8}$$

Before proving Lemma 3.4.1 and Lemma 3.4.3, we first show how they yield the result of Theorem 3.4.

Proof of Theorem 3.4. Let $s \leq t$ and $U_{s+1}, \dots, U_t$ be given, and define

$$\Theta_{s,s} = \widehat{\Theta}_{t,t} = \emptyset.$$

For all $n \in [s+1, t]$

$$\begin{aligned} \left|\Theta_{s,n} \otimes \widehat{\Theta}_{n,t}\right| &= \left|(\Theta_{s,n-1} \otimes U_n) \otimes \widehat{\Theta}_{n,t}\right| \\ &= \left|\Theta_{s,n-1} \otimes (\widehat{\Theta}_{n,t} \otimes U_n)\right| \\ &= \left|\Theta_{s,n-1} \otimes \widehat{\Theta}_{n-1,t}\right|, \end{aligned} \tag{3.4.9}$$

where the second equality is an application of Lemma 3.4.3. By successive application of (3.4.9) for $n = t, t-1, \dots, s+1$ it follows that

$$\left|\Theta_{s,t}\right| = \left|\Theta_{s,t} \otimes \widehat{\Theta}_{t,t}\right| = \left|\Theta_{s,t-1} \otimes \widehat{\Theta}_{t-1,t}\right| = \cdots = \left|\Theta_{s,s+1} \otimes \widehat{\Theta}_{s+1,t}\right| = \left|\Theta_{s,s} \otimes \widehat{\Theta}_{s,t}\right| = \left|\widehat{\Theta}_{s,t}\right|,$$

which completes the proof. □

We now prove Lemma 3.4.1. The idea is to check that (3.4.7) holds for $n = 2$, then use that any permutation can be written as a composition of transpositions. Since $A \cup B$ is not necessarily a tracking set, i.e. some elements of B may overlap with A, the proof requires some case checking.

Proof of Lemma 3.4.1. We show that (3.4.7) holds for $n = 2$. Since $B = \{\beta_1, \beta_2\}$ is a tracking set, WLOG suppose that $\beta_1 = (1, a)$ and $\beta_2 = (2, b)$ where $a \neq b$.

We consider the bipartite graph $G = (A, B, E)$ where $((i, x), (j, y)) \in E$ if $(i, x) \in A$, $(j, y) \in B$ and $i = j$ or $x = y$. Moreover, let

$$\Gamma_1 = \big\{\alpha \in A : (\alpha, \beta_1) \in E\big\}, \tag{3.4.10}$$
$$\Gamma_2 = \big\{\alpha \in A : (\alpha, \beta_2) \in E\big\}. \tag{3.4.11}$$

See Figure 3.4 for an example.

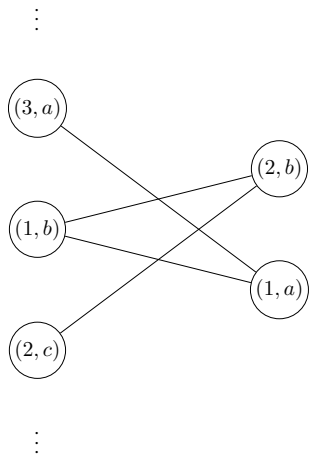


Figure 3.4: The above figure gives the bipartite graph (A, B, E) with vertex sets $A \supseteq \{(1, b), (2, c), (3, a)\}$, $B = \{(1, a), (2, b)\}$, where $a, b, c \in [M]$ are distinct. Here $\Gamma_1 = \{(1, b), (3, a)\}$ and $\Gamma_2 = \{(1, b), (2, c)\}$.

Since A is an (ℓ, M)–tracking set we have that $|\Gamma_1|, |\Gamma_2| \leq 2$.

We note that

$$\begin{aligned} A \otimes \beta_1 &= (A \setminus \Gamma_1) \cup \big\{\Gamma_1 \otimes \beta_1\big\}, \\ A \otimes \beta_2 &= (A \setminus \Gamma_2) \cup \big\{\Gamma_2 \otimes \beta_2\big\}, \end{aligned} \tag{3.4.12}$$

i.e. applying the $\otimes$–updates β_1 or β_2 leaves all elements in $A \setminus \Gamma_1$ or $A \setminus \Gamma_2$ unchanged.

To see (3.4.12), observe from Definition 3.2.4 that $A \setminus (A \otimes \beta_1) = \Gamma_1$, and $(A \otimes \beta_1) \setminus A = \Gamma_1 \otimes \beta_1$ for each of the five update types. For example, if β_1 is a $\otimes$–good update for A then $\Gamma_1 = \emptyset$ and $A \otimes \beta_1 = A \cup \{\beta_1\}$. In particular

$$\begin{aligned} A \setminus (A \otimes \beta_1) &= \emptyset, \\ (A \otimes \beta_1) \setminus A &= \{\beta_1\}. \end{aligned}$$

Since

$$\begin{aligned} A \otimes \beta_1 &= A \setminus \big(A \setminus (A \otimes \beta_1)\big) \cup \big(A \otimes \beta_1 \setminus A\big) \\ &= \big(A \setminus \Gamma_1\big) \cup \big\{\Gamma_1 \otimes \beta_1\big\}, \end{aligned}$$

(3.4.12) follows immediately. By symmetry this holds for β_2.

We now divide the proof into cases depending on the size of $|\Gamma_1 \cap \Gamma_2|$.

Case 1: $|\Gamma_1 \cap \Gamma_2| = 0$.

In this case $\Gamma_2 \subseteq (A \setminus \Gamma_1) \subseteq A \otimes \beta_1$, and $\Gamma_1 \subseteq (A \setminus \Gamma_2) \subseteq A \otimes \beta_2$ by application of (3.4.12). In particular,

$$\begin{aligned} A \otimes \beta_1 \otimes \beta_2 &= (A \otimes \beta_1 \setminus \Gamma_2) \cup \{\Gamma_2 \otimes \beta_2\} \\ &= \big[A \setminus\ (\Gamma_1 \cup \Gamma_2)\big] \cup \big\{\Gamma_1 \otimes \beta_1\big\} \cup \big\{\Gamma_2 \otimes \beta_2\big\} \\ &= (A \otimes \beta_2 \setminus \Gamma_1) \cup \{\Gamma_1 \otimes \beta_1\} \\ &= A \otimes \beta_2 \otimes \beta_1, \end{aligned}$$

where the first and third inequalities are an application of (3.4.12).

Case 2: $|\Gamma_1 \cap \Gamma_2| = 2$.

In this case $\Gamma_1 = \Gamma_2 = \{(1, b), (2, a)\}$ and it is straightforward to check that

$$\begin{aligned} A \otimes \beta_1 \otimes \beta_2 &= \{(1, a), (2, b)\} \cup \big(A \setminus \{(1, b), (2, a)\}\big) \\ &= A \otimes \beta_2 \otimes \beta_1. \end{aligned}$$

Case 3: $|\Gamma_1 \cap \Gamma_2| = 1$.

WLOG $\Gamma_1 \cap \Gamma_2 = \{(1, b)\}$, and we divide the proof of Case 3 into two cases.

Case 3(a): $\Gamma_1 = \{(1, b)\}$ or $\Gamma_2 = \{(1, b)\}$.

Suppose WLOG that $\Gamma_1 = \{(1, b)\}$. In this case

$$A \otimes \beta_1 = \big(A \setminus \{(1, b)\}\big) \cup \{(1, a)\},$$

by (3.4.12). Since $A \otimes \beta_1$ is a tracking set $b \notin (A \otimes \beta_1)_2$, the projection onto the second coordinate of $A \otimes \beta_1$.

Therefore $(2, b)$ is an $\otimes$–good or $\otimes$–jump update for $A \otimes \beta_1$ and, if $(2, b)$ is an $\otimes$–jump update for $A \otimes \beta_1$, let $(2, c) \in \Gamma_2$ for $c \notin \{a, b\}$. In particular,

$$A \otimes \beta_1 \otimes \beta_2 = \begin{cases} (A \setminus \{(1, b)\}) \cup \{(1, a), (2, b)\} & \text{if } (2, c) \in A, \\ (A \setminus \{(1, b), (2, c)\}) \cup \{(1, a), (2, b)\} & \text{if } (2, c) \notin A. \end{cases}$$

If $(2, c) \in \Gamma_2$ then $(2, b)$ is an $\otimes$–swap update for A and by Definition 3.2.4

$$A \otimes \beta_2 = \big(A \setminus \{(1, b), (2, c)\}\big) \cup \{(2, b), (1, c)\},$$

so that

$$A \otimes \beta_2 \otimes \beta_1 = \big(A \setminus \{(1, b), (2, c)\}\big) \cup \{(2, b), (1, a)\}.$$

If $(2, c) \notin \Gamma_2$ then it is straightforward to check that

$$A \otimes \beta_2 \otimes \beta_1 = \big(A \setminus \{(1, b)\}\big) \cup \{(1, a), (2, b)\},$$

and (3.4.7) holds.

Case 3(b): WLOG let $\Gamma_1 = \{(1,b),(3,a)\}$ and $\Gamma_2 = \{(1,b),(2,c)\}$.

Here both β_1 and β_2 are both $\otimes$–swap updates for A. Therefore

$$A \otimes \beta_1 = \big(A \setminus \{(1,b),(3,a)\}\big) \cup \{(1,a),(3,b)\},$$

and in particular $(2,c),(3,b) \in A \otimes \beta_1$, so that β_2 is an $\otimes$–swap update for $A \otimes \beta_1$. It follows that

$$A \otimes \beta_1 \otimes \beta_2 = \big(A \setminus \{(1,b),(3,a),(2,c)\}\big) \cup \{(1,a),(2,b),(3,c)\}.$$

Moreover,

$$\begin{aligned} A \otimes \beta_2 \otimes \beta_1 &= \Big[\big(A \setminus \{(1,b),(2,c)\}\big) \cup \{(2,b),(1,c)\}\Big] \otimes \beta_1 \\ &= \big(A \setminus \{(1,b),(2,c),(3,a)\}\big) \cup \{(2,b),(1,a),(3,c)\}, \end{aligned}$$

and (3.4.7) holds.

Supposing that $\{\beta_1, \beta_2, \dots, \beta_n\}$ is an (ℓ, M)–tracking set for $n \geq 3$, then by the argument for $n = 2$ we have that

$$\big(A \otimes \beta_1 \otimes \beta_2\big) \otimes \beta_3 \otimes \cdots \otimes \beta_n = \big(A \otimes \beta_2 \otimes \beta_1\big) \otimes \beta_3 \otimes \cdots \otimes \beta_n.$$

In particular we may transpose the application of any two adjacent elements from a tracking set. Any permutation may be written as a composition of adjacent transpositions, and the proof generalises to all $n \geq 1$. □

We now prove Lemma 3.4.3, the idea here is to consider the maximal subset of B for which γ is $\otimes$–good, and reduce the problem to a few cases.

Proof of Lemma 3.4.3. WLOG let $\gamma = (1,a)$ and $B = \{\beta_1, \dots, \beta_n\}$.

Case 1: Suppose $(1,a)$ is a $\otimes$–good update for B.

By Definition 3.2.4, it holds that

$$B \otimes \gamma = B \cup \{\gamma\},$$

and $B \otimes \gamma$ is an (ℓ, M)–tracking set. In particular, the LHS of (3.4.8) is given by

$$A \otimes \gamma \otimes \beta_1 \otimes \cdots \otimes \beta_n,$$

and the RHS of (3.4.8) is given by

$$A \otimes \beta_1 \otimes \cdots \otimes \beta_n \otimes \gamma,$$

but by an application of Lemma 3.4.1 these result in the same set and (3.4.8) follows.

Case 2: Suppose now that $(1,a)$ is not $\otimes$–good for B. We define $\Gamma_+ \subseteq B$ to be the maximal subset of B such that $(1,a)$ is $\otimes$–good for Γ_+, and let $\Gamma_- = B \setminus \Gamma_+$.

WLOG it holds that $\Gamma_- = \{\beta_n\} = \{(1,b)\}$ (allowing the case $b = a$) and $\Gamma_+ = \{\beta_1, \dots, \beta_{n-1}\}$, or $\Gamma_- = \{\beta_{n-1}, \beta_n\} = \{(1,b),(2,a)\}$ with $b \neq a$ and $\Gamma_+ = \{\beta_1, \dots, \beta_{n-2}\}$.

Case 2(a): $\Gamma_- = \{\beta_n\} = \{(1,b)\}$:

In this case,

$$\begin{aligned}\big(A\otimes(1,a)\big)\otimes B &= \big(A\otimes(1,a)\big)\otimes\beta_1\otimes\cdots\otimes\beta_n\\ &= A\otimes\big((\emptyset\otimes\gamma)\otimes\beta_1\otimes\cdots\otimes\beta_{n-1}\big)\otimes\Gamma_-\\ &= A\otimes\big(\Gamma_+\otimes\gamma\big)\otimes\Gamma_-,\end{aligned}\tag{3.4.13}$$

where the second equality follows from the fact that $\{\gamma\}\otimes\Gamma_+ = \Gamma_+\cup\{\gamma\} = \Gamma_+\otimes\gamma$ is an (ℓ,M)–tracking set. Moreover, $B\otimes(1,a) = \{\beta_1,\ldots,\beta_{n-1},(1,a)\}$ and

$$A\otimes\big(B\otimes\gamma\big) = A\otimes\big(\Gamma_+\otimes\gamma\big).$$

Noting that $\gamma=(1,a)\in A\otimes\big(\Gamma_+\otimes\gamma\big)$, it holds that $(1,b)$ is a $\otimes$–jump, $\otimes$–swap or $\otimes$–null update for the set $A\otimes\big(B\otimes\gamma\big)$. These updates do not change the size of a tracking set and (3.4.8) holds.

Case 2(b): $\Gamma_- = \{(1,b),(2,a)\}$:

In this case,

$$\big(A\otimes(1,a)\big)\otimes B = A\otimes\big(\Gamma_+\otimes\gamma\big)\otimes\Gamma_-.$$

Since $(1,a)$ is an $\otimes$–swap update for the set $B\supseteq\Gamma_-$,

$$B\otimes\gamma = \{\beta_1,\ldots,\beta_{n-2},(1,a),(2,b)\} = \big(\Gamma_+\otimes\gamma\big)\otimes(2,b),$$

and

$$A\otimes\big(B\otimes\gamma\big) = A\otimes\big(\Gamma_+\otimes\gamma\big)\otimes\{(2,b)\}.$$

It's left to check that applying the updates $(1,b)$ then $(2,a)$ or applying only the update $(2,b)$ to the (ℓ,M)–tracking set

$$\bar{A} := A\otimes\big(\Gamma_+\otimes\gamma\big),$$

results in sets which have the same size.

Noting that $(1,a)\in\bar{A}$, the update $(1,b)$ is not $\otimes$–good for $\bar{A}$. In particular,

$$\big|\bar{A}\otimes(1,b)\big| = \big|\bar{A}\big|,$$

and since $b\neq a$ we have that $(1,b)$ is either an $\otimes$–jump or $\otimes$–swap update for $\bar{A}$. By Definition 3.2.4, the projections onto the first coordinates of $\bar{A}$ and $[\bar{A}\otimes(1,b)]$ satisfy $\bar{A}_1 = [\bar{A}\otimes(1,b)]_1$.

Suppose that $2\in\bar{A}_1$, then $(2,a)$ is not an $\otimes$–good update for $\bar{A}\otimes(1,b)$ and $(2,b)$ is not an $\otimes$–good update for $\bar{A}$. It follows that

$$|\bar{A}\otimes(1,b)\otimes(2,a)| = |\bar{A}| = |\bar{A}\otimes(2,b)|.\tag{3.4.14}$$

Suppose that $b\in\bar{A}_2$ then $(1,b)$ is an $\otimes$–swap update for $\bar{A}$ and it holds that $a\in[\bar{A}\otimes(1,b)]_2$. In particular, $(2,a)$ is not an $\otimes$–good update for $\bar{A}\otimes(1,b)$, and $(2,b)$ is not an $\otimes$–good update for $\bar{A}$, since we assumed that $b\in\bar{A}_2$, so (3.4.14) holds.

Lastly, suppose that $2\notin\bar{A}_1$ and $b\notin\bar{A}_2$, i.e. $(2,b)$ is $\otimes$–good for $\bar{A}$. Since $b\notin\bar{A}_2$ it holds that $(1,b)$ is an $\otimes$–jump update for $\bar{A}$. Therefore $a\notin[\bar{A}\otimes(1,b)]_2$, and $2\notin\bar{A}_1 = [\bar{A}\otimes(1,b)]_1$ by assumption, i.e. $(2,a)$ is an $\otimes$–good update for $\bar{A}\otimes(1,b)$. In this case

$$\big|\bar{A}\otimes(1,b)\otimes(2,a)\big| = 1+\big|\bar{A}\big| = \big|\bar{A}\otimes(2,b)\big|,$$

and (3.4.8) holds. In all cases (3.4.8) holds, which completes the proof. □

3.4.2 Proof of Theorem 3.3

In this section we prove Theorem 3.3, which concerns fragmentation processes. The following pseudocode gives an algorithm which satisfies Theorem 3.3.

Algorithm (B):

```
1 :  t ← 0
2 :  L = (L_1, L_2, ..., L_{k-1}) ← (0, ..., 0)
3 :  while L_{k-1} = 0 :
4 :      L ← (0, ..., 0)
5 :      p ← 0
6 :      generate (U_{-s} = (I_{-s}, X_{-s}) : t ≤ s < max(2t, k − 1))
7 :      t ← max(2t, k − 1)
8 :      for s = t − 1 down to 0 :
9 :          (i, x) ← (I_{-s}, X_{-s})
10 :         if i ≤ p + 1 and x not in L :
11 :             L_i ← x
12 :             if i = p + 1 :
13 :                 p ← p + 1
14 :         if i ≤ p and L_j = x  and i ≤ j ≤ p :     # a unique such j exists
15 :             L_j ← L_i
16 :             L_i ← x
17 :  output {L_1, ..., L_{k-1}}
```

A priori it is not clear that Algorithm (B) will terminate. Similarly to Algorithm (A), every nonzero entry in the list L will represent a particle position in the Priority process. Algorithm (B) is an application of the CFTP algorithm; we repeatedly double the number of t–updates, storing and re–using them. In addition, p represents the lowest priority amongst set of particles whose locations agree in all copies of the Priority process.

The output of Algorithm (B) is a set of indices $\{L_1, \ldots, L_{k-1}\}$. The first i entries $\{L_1, \ldots, L_i\}$ of the output of Algorithm (B) gives rise to a cycle structure with $i+1$ cycles as follows:

- put the indices in increasing order as $L'_1, \ldots, L'_i$;
- take the successive differences $L'_1, L'_2 - L'_1, \ldots, L'_i - L'_{i-1}, N - L'_i$ to get a multiset of cycle lengths.

We adapt the ideas of tracking sets and updates to the setting of the Priority process. As before, an update (r, x) will represent an attempt by particle r to jump to site x. To know whether this attempt succeeds, we need to know whether site x contains a particle with label in $[r-1]$.

In particular we acquire information about the particles in increasing order of their labels. We can never determine the location of particle r unless we already have information about all the particles with lower labels. As a result, the tracking sets A we work with will have the property that A_1 is empty or equal to $[p]$ for some $p \in [\ell]$.

We now define the effect of an update of the Priority process on a tracking set.

Definition 3.4.4. *Let A be an (ℓ, M)–tracking set from Definition 3.2.3 and suppose that $A_1 = [p]$ for some $p \in [\ell]$ or $A = \emptyset$, in which case we set $p = 0$. Let $\alpha = (r, x) \in [\ell] \times [M]$ be arbitrary. Then we update A with α to obtain the set $A \oplus \alpha$ as follows:*

***Good**: If $r = p + 1$ and $x \notin A_2$ then we define $A \oplus (r, x) = A \cup \{(r, x)\}$. We say that (r, x) is an $\oplus$–good update for A;*

***Jump**: if $r \leq p$ and $x \notin A_2$ then $(r, x') \in A$ and we set*

$$A \oplus (r, x) = \{(r, x)\} \cup \big(A \setminus \{(r, x')\}\big).$$

We say that (r, x) is an $\oplus$–jump update for A;

***Swap**: if $r \leq p$ and $(r, x'), (r', x) \in A$ with $r < r'$, then we set*

$$A \oplus (r, x) = \{(r, x), (r', x')\} \cup \big(A \setminus \{(r', x), (r, x')\}\big).$$

We say that (r, x) is an $\oplus$–swap update for A;

***Null**: otherwise we set $A \oplus (r, x) = A$. We say that (r, x) is an $\oplus$–null update for A.*

The set $A \oplus \alpha$ is necessarily an (ℓ, M)–tracking with projection $(A \oplus \alpha)_1 = [p]$ for some $p \in [\ell]$ (or $(A \oplus \alpha)_1 = \emptyset$).

As for the Labelled process we consider a Markov chain on the augmented state space $\widetilde{\Sigma}_{M,\ell}$ whose dynamics mimic the evolution of all copies of the Priority process. For two integers $s \leq t$ let

$$\Phi_{s,t} = \emptyset \oplus U_{s+1} \oplus \cdots \oplus U_t,$$

with $\Phi_{s,s} = \emptyset$. Let $\Phi^1_{s,t}$ and $\Phi^2_{s,t}$ denote the projections onto the first and second coordinate respectively.

Lemma 3.4.5. *Let $U_{s+1}, \ldots, U_t$ be given, and let $(\chi_n)_s^t$ denote the Priority process with initial condition $\chi_s \in \Sigma_{M,\ell}$ run through the updates $U_{s+1}, \ldots, U_t$. Then*

$$\chi_t(x) = r \text{ for all } (r, x) \in \Phi_{s,t}, \tag{3.4.15}$$

and all $t \geq s$ and $\chi_s \in \Sigma_{M,\ell}$.

Proof of Lemma 3.4.5. We proceed by induction on $t \geq s$ For the base case $t = s$, $\Phi_{s,s} = \emptyset$ and (3.4.15) holds trivially. Assume for an induction that (3.4.15) holds for some $t \geq s$ and $\Phi^1_{s,t} = [p]$ for some $p \in [\ell]$.

Let $U_{t+1} = (r, x)$. If (r, x) with $r \leq p + 1$ is an $\oplus$–good or $\oplus$–jump update for $\Phi_{s,t}$ then $x \notin \Phi^2_{s,t}$. Since $x \notin \Phi^2_{s,t}$ no particle $r' \leq p$ occupies the site x by (3.4.15) at time t. Therefore particle r moves to site x in the dynamics of the Priority process, i.e. $\chi_{t+1}(x) = r$, and (3.4.15) holds at time $t + 1$.

If $r \leq p$ and $x \in \Phi^2_{s,t}$ then (r, x) is either an $\oplus$–swap or $\oplus$–null update. In this case the particles in $[p]$ perform the same swap or do nothing in all copies of the Priority process, and (3.4.15) holds at time $t + 1$.

If $r = p + 1$ and $x \notin \Phi^2_{s,t}$ or $r > p + 1$ then (r, x) is an $\oplus$–null update. In the dynamics no particle in $[p]$ moves and (3.4.15) holds at time $t + 1$. □

Remark 3.4.6. *By Lemma 3.4.5 Algorithm (B) is an application of the CFTP algorithm to the Priority process. In particular, after each while loop, the set* $\{(i, L_i) : L_i \neq 0\}$ *is given by*

$$\Phi_{-t,0} = \emptyset \oplus U_{-t+1} \oplus U_{-t+2} \oplus \cdots \oplus U_0.$$

If $|\Phi_{-t,0}| = \ell$*, all copies of the Priority process have coalesced and, by Remark 3.2.1 and [91][Theorem 1], the first* i *entries in the output* $L = \Phi^2_{-t,0}$ *of Algorithm (B) is a sample from* $\nu_{M,i}$*.*

Remark 3.4.7. *The analogue of Theorem 3.4 does not hold the for the process* $\Phi_{s,t}$*, i.e.*

$$\widehat{\Phi}_{s,t} := \emptyset \oplus U_t \oplus U_{t-1} \oplus \cdots \oplus U_{s+1},$$

is not in general the same size as $\Phi_{s,t}$*. For example*

$$\emptyset \oplus (1, a) \oplus (2, b) = \{(1, a), (2, b)\},$$
$$\emptyset \oplus (2, b) \oplus (1, a) = \{(1, a)\},$$

for $a \neq b$*. In particular there is no analogue of Algorithm (A*) for the Priority process.*

We now prove Theorem 3.3.

Proof of Theorem 3.3. By Remark 3.4.6 the first i entries of the output L are a sample from the distribution $\nu_{M,i}$ for each $i \in [\ell]$. Since $\mu_{N,i+1} \sim \nu_{M,i}$ for $M = N - 1$, property (1) from Definition 3.1.2 holds and property (2) is immediate.

Remarks analogous to those made in Section 3.3 for Algorithm (A) hold for Algorithm (B), and it is sufficient to check that the expected number of applications of t–updates in Algorithm (B) is at most $O(\ell^2 \text{polylog} M)$. Let

$$T^{\oplus} = \inf\big\{t \geq 0 : |\Phi_{0,t}| = \ell\big\},$$

then by Remark 1.7.2 the number of applications of t–updates in Algorithm (B) is at most $4T^{\oplus}$.

Suppose that $|\Phi_{0,t}| = m < \ell$, then $I_{t+1} = m + 1$ with probability $1/\ell$ and $X_{t+1} \notin \Phi^2_{0,t}$ with probability at least $\sum_{x=1}^{M-m} q_x$ independent of I_{t+1}. Therefore U_{t+1} is a $\oplus$–good update for $\Phi_{s,t}$ with probability at least

$$\lambda'_m = \frac{1}{\ell} \sum_{x=1}^{M-m} q_x.$$

Note that every $\oplus$–good update increases the size of the tracking set by 1, and no update ever decreases the size of the tracking set. Hence the time for the $|\Phi_{0,t}|$ to increase from size m to size $m + 1$ is dominated by a geometric random variable $\text{Geom}(\lambda'_m)$ with mean $1/\lambda'_m$.

Therefore

$$\mathbb{E}\big[T^{\oplus}\big] \leq \sum_{m=0}^{\ell-1} \frac{1}{\lambda'_m} \leq \ell H_M \sum_{m=0}^{\ell-1} \left(\sum_{x=1}^{M-m} \frac{1}{M + 1 - x} \right)^{-1},$$

and

$$\begin{aligned}\mathbb{E}\big[4T^{\oplus}\big] &\leq 4\ell H_M \sum_{m=0}^{\ell-1}\left(\sum_{x=1}^{M-m}\frac{1}{M+1-x}\right)^{-1}\\ &= 4\ell H_M \sum_{m=0}^{\ell-1}\frac{1}{(H_M-H_m)}\\ &\leq \frac{4\ell^2 H_M}{H_M-H_\ell} = O\big(\ell^2 * \log M\big),\end{aligned}$$

as $M \to \infty$, which completes the proof of Theorem 3.3. □

Chapter 4

Intermediate–passage percolation

4.1 Introduction

In this chapter we examine properties of a two–dimensional directed percolation model, namely intermediate–passage percolation, introduced in Section 1.8.1. The model may be thought of as lying 'between' the well–studied cases of directed first and last–passage percolation. To each $\boldsymbol{z} \in \mathbb{Z}_+^2 \setminus \{\mathbf{0}\}$ we associate an i.i.d. random variable $X(\boldsymbol{z})$, which we call the *weight*. By convention we set the weight at the origin $X(\mathbf{0}) = 0$.

In Definition 1.8.1 we defined the neighbourhood $\mathcal{N}(\boldsymbol{z})$ of $\boldsymbol{z} \in \mathbb{N}^2$ to be

$$\mathcal{N}(\boldsymbol{z}) = \big\{(z_1 - 1, z_2), (z_1, z_2 - 1), (z_1 - 1, z_2 - 1)\big\} \subseteq \mathbb{Z}_+^2.$$

We now give a recursive definition for first and last–passage times for each $\boldsymbol{z} \in \mathbb{Z}_+^2$, and reiterate Definition 1.8.2 for intermediate–passage times.

Definition 4.1.1. *Let the collection of weights $\{X(\boldsymbol{z})\}_{\boldsymbol{z} \in \mathbb{Z}_+^2}$ be given. If $\boldsymbol{z} = (z_1, z_2) \in \mathbb{Z}_+^2$ with $z_1 z_2 = 0$ we say that $\boldsymbol{z}$ is a boundary site. For each $z_1, z_2 \geq 0$, we define the passage times for the boundary sites via*

$$\begin{aligned} T_\ell(z_1, 0) = T(z_1, 0) = T_f(z_1, 0) &= \sum_{i=0}^{z_1} X(i, 0), \\ T_\ell(0, z_2) = T(0, z_2) = T_f(0, z_2) &= \sum_{i=0}^{z_2} X(0, i). \end{aligned} \tag{4.1.1}$$

The passage times in $\mathbb{N}^2$ are defined using the following recursions:

$$T_\ell(\boldsymbol{z}) = X(\boldsymbol{z}) + \max\big\{T_\ell(z_1 - 1, z_2), T_\ell(z_1, z_2 - 1)\big\}, \tag{4.1.2}$$

$$T(\boldsymbol{z}) = X(\boldsymbol{z}) + \operatorname{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\big\}, \tag{4.1.3}$$

$$T_f(\boldsymbol{z}) = X(\boldsymbol{z}) + \min\big\{T_f(z_1 - 1, z_2), T_f(z_1, z_2 - 1), \big\}, \tag{4.1.4}$$

where the median refers to the middle value when the elements of $\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\}$ are ordered from smallest to largest. We call $T_\ell(\boldsymbol{z}), T(\boldsymbol{z})$ and $T_f(\boldsymbol{z})$ the first, intermediate and last–passage time for $\boldsymbol{z} \in \mathbb{Z}_+^2$ respectively.

The equivalence of the recursive passage time $T_\ell(\boldsymbol{z})$ to the passage time in (1.8.1) for boundary sites is straightforward; there is exactly one path in the set $\Pi(\mathbf{0}, \boldsymbol{z}]$. To see the equivalence for $\boldsymbol{z} \in \mathbb{N}^2$, observe that the last–passage times in (1.8.1) satisfy the recursion (4.1.2). In particular, a path $\pi \in \Pi(\mathbf{0}, \boldsymbol{z}]$, which attains the maximum

$$\max_{\pi \in \Pi(\mathbf{0}, \boldsymbol{z}]} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}),$$

must pass through $(z_1 - 1, z_2)$ or $(z_1, z_2 - 1)$ and

$$\max_{\pi \in \Pi(\mathbf{0}, \boldsymbol{z}]} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}) = X(\boldsymbol{z}) + \max\left\{ \max_{\pi \in \Pi(\mathbf{0}, (z_1 - 1, z_2)]} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}), \max_{\pi \in \Pi(\mathbf{0}, (z_1, z_2 - 1)]} \sum_{\boldsymbol{v} \in \pi} X(\boldsymbol{v}) \right\}.$$

There are finitely many recursions in the rectangle $[1, z_1] \times [1, z_2]$, and $T_\ell(\boldsymbol{z})$ is the unique solution. A similar argument yields the equivalence for the first–passage times.

The passage times in Definition 4.1.1 may be used to define growth models in $\mathbb{R}^2_+$.

Definition 4.1.2. *Given the passage times $T_\ell(\boldsymbol{z}), T(\boldsymbol{z})$ and $T_f(\boldsymbol{z})$, we set*

$$B_t := \big\{ \boldsymbol{x} \in \mathbb{R}^2_+ : T_\ell(\lfloor \boldsymbol{x} \rfloor) \le t \big\}, \tag{4.1.5}$$

$$\Gamma_t := \big\{ \boldsymbol{x} \in \mathbb{R}^2_+ : T(\lfloor \boldsymbol{x} \rfloor) \le t \big\}, \tag{4.1.6}$$

$$C_t := \big\{ \boldsymbol{x} \in \mathbb{R}^2_+ : T_f(\lfloor \boldsymbol{x} \rfloor) \le t \big\}, \tag{4.1.7}$$

for all $t \ge 0$.

We will show that $T_f(\boldsymbol{x}) \le T(\boldsymbol{x}) \le T_\ell(\boldsymbol{x})$ when they are defined on the same random environment. Moreover, we give conditions under which the shape Γ_t grows linearly with probability 1, as a function of t, in every direction. The conditions include cases for which linear growth in every direction does not hold for first or last–passage percolation. In Section 4.5 we give several simulations of the intermediate–passage process $(\Gamma_t)_{t \ge 0}$ and state some open questions motivated by simulations.

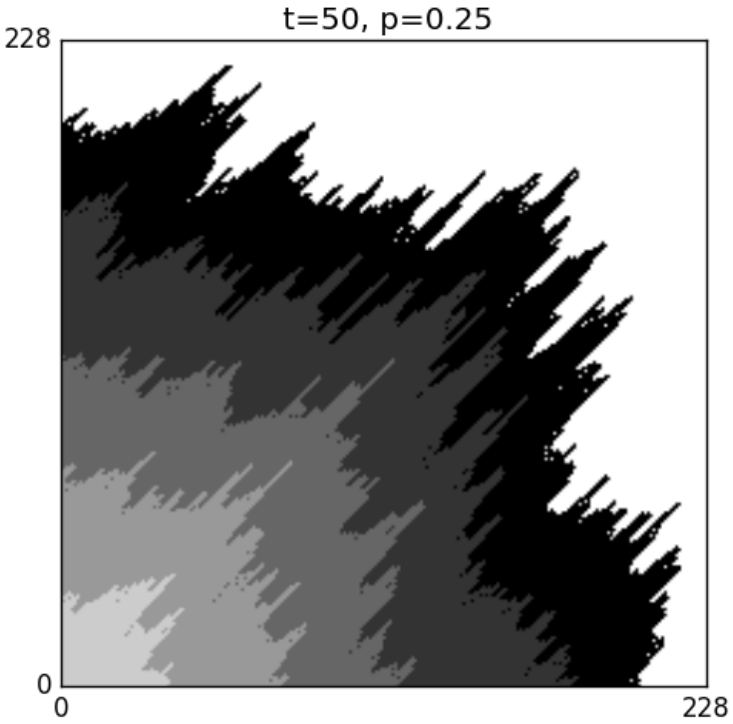


Figure 4.1: A simulation of the shape Γ_{50} with Bernoulli(0.25) weights satisfying $\mathbb{P}(X(\boldsymbol{z}) = 1) = 1 - \mathbb{P}(X(\boldsymbol{z}) = 0) = 0.25$. For these weights, Γ_t is increasing in t and the shapes $\Gamma_{10}, \Gamma_{20} \setminus \Gamma_{10}, \ldots, \Gamma_{50} \setminus \Gamma_{40}$ are shaded from lightest to darkest.

4.2 Motivation

One motivation for intermediate–passage percolation came from kinetically constrained models (KCM). KCMs are interacting particle system on $\mathbb{Z}^d$ and were introduced in the Physics literature by Frederickson and Andersen [38] to model the liquid–glass transition, which is an important area of condensed matter physics that still has many open problems.

KCMs are characterised by a finite collection of finite subsets $U_i \subset \mathbb{Z}^d \setminus \{\mathbf{0}\}$ called *update sets*. We call the collection $\mathcal{U} = \{U_1, \ldots, U_m\}$ an *update family*. Let $\omega \in \{0,1\}^{\mathbb{Z}^d}$ and $\boldsymbol{z} + A$ denote the set $\{\boldsymbol{z} + \boldsymbol{a} : \boldsymbol{a} \in A\}$. The dynamics of KCMs may be described as follows: each site $\boldsymbol{z} \in \mathbb{Z}^d$ waits a mean one exponential time and the configuration ω is updated so that $\omega(\boldsymbol{z}) = 1$ with probability p, or $\omega(\boldsymbol{z}) = 0$ with probability $q = 1 - p$, if and only if there exists a $U_i \in \mathcal{U}$ with $\omega(\boldsymbol{y}) = 0$ for all $\boldsymbol{y} \in \boldsymbol{z} + U_i$. The equilibrium distribution of KCMs is product Bernoulli with parameter p.

KCMs are closely related to $\mathcal{U}$–bootstrap percolation introduced by Chalupa, Leath and Reich [24]; KCMs may be thought of as the non–monotone stochastic counterpart of $\mathcal{U}$–bootstrap percolation. The update families in $\mathcal{U}$–bootstrap percolation may be divided into universality classes which exhibit different behaviours. We refer the interested reader to Balister et al. [11] and Bollobás et al. [20]. A natural question is whether there are also universality classes for KCMs which exhibit different equilibrium behaviours. Significant progress has been made

towards establishing the correct university classes in two dimensions; see for example Martinelli, Morris and Toninelli [85].

The non–equilibrium behaviour of KCMs is a relatively new and active area of research (see e.g. [17, 21, 25]). We now give a relation between intermediate–passage percolation and a KCM with a non–equilibrium initial condition.

In two dimensions we consider the update family $\mathcal{U}$ given by

$$\begin{aligned} U_1 &= \big\{(-1,0),(0,-1)\big\}, \\ U_2 &= \big\{(-1,0),(-1,-1)\big\}, \\ U_3 &= \big\{(0,-1),(-1,-1)\big\}, \\ U_4 &= \big\{(-1,0),(0,-1),(-1,-1)\big\}. \end{aligned} \tag{4.2.1}$$

Let $q = 1$, $E = \{(-1,z),(z,-1) : z \in 2\mathbb{Z}_+\} \cup \{\mathbf{0}\}$ and set

$$\omega(\boldsymbol{z}) = \begin{cases} 0 & \text{if } \boldsymbol{z} \in E, \\ 1 & \text{otherwise.} \end{cases} \tag{4.2.2}$$

When $X(\boldsymbol{z}) \sim \mathrm{Exp}(1)$ and $q = 1$ the shape Γ_t is exactly the set of empty sites contained in $\mathbb{Z}_+^2$ in the KCM at time t. See, for example, Figure 4.2.

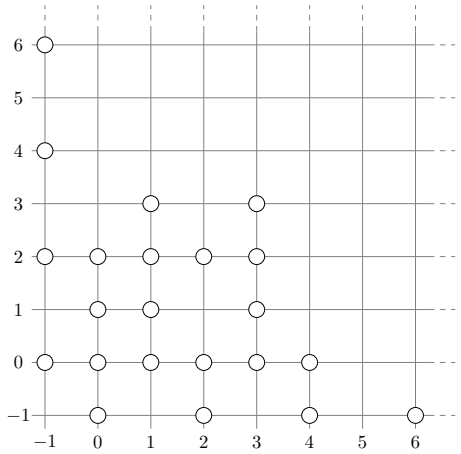


Figure 4.2: A configuration which may be reached by the KCM with update family given by (4.2.1) and initial condition ω given by (4.2.2). Each white circle represents a site that takes value 0, and all other sites take value 1. The set of sites which may be updated to 0 in the KCM with the above configuration is $\{(0,3),(2,3),(2,1),(4,3),(4,2),(4,1),(5,0)\}$.

4.3 Results

Theorem 4.1. *For $\theta \in [0,\pi/2]$ let*

$$f_t(\theta) = \sup\big\{r \in \mathbb{R}_+ : (r\cos\theta, r\sin\theta) \in \Gamma_t\big\}.$$

Suppose the weights away from the origin are i.i.d. and nonnegative with finite and strictly positive mean. Then there exist positive constants $C_1(\theta), C_2(\theta) > 0$ so that

$$C_1(\theta)\, t \le f_t(\theta) \le C_2(\theta)\, t \tag{4.3.1}$$

for all $\theta \in [0, \pi/2]$ *and* t *sufficiently large, with probability* 1.

In particular Theorem 4.1 says that Γ_t grows linearly in every direction provided that the underlying environment consists of nonnegative i.i.d. random variables which have strictly positive and finite first moment.

The following lemma serves as justification for the name intermediate–passage percolation.

Lemma 4.3.1. *On the same random environment* $\{X(\boldsymbol{z})\}_{\boldsymbol{z}\in\mathbb{Z}^2_+}$,

$$T_f(\boldsymbol{z}) \le T(\boldsymbol{z}) \le T_\ell(\boldsymbol{z}), \ \ \forall\, \boldsymbol{z} \in \mathbb{Z}^2_+. \tag{4.3.2}$$

Proof of Lemma 4.3.1. The first, last and intermediate–passage times agree on the boundary sites by definition. We check by an induction on the sets

$$D_n := \big\{\boldsymbol{z} \in \mathbb{N}^2 : z_1 \le n + 1 - z_2\},$$

that (4.3.2) holds on $\mathbb{N}^2$.

For the base case $n = 1$ we have that

$$\begin{aligned} T_f(1,1) &= X(1,1) + \min\{X(1,0), X(0,1)\},\\ T(1,1) &= X(1,1) + \text{median}\{X(1,0), X(0,1), 0\},\\ T_\ell(1,1) &= X(1,1) + \max\{X(1,0), X(0,1)\}, \end{aligned}$$

and since

$$\min\{a,b\} \le \text{median}\{a,b,c\} \le \max\{a,b\}, \tag{4.3.3}$$

for all $a, b, c \in \mathbb{R}$, (4.3.2) holds on D_1. Assume for an induction that (4.3.2) holds on D_n for some $n \ge 1$. Let $\boldsymbol{z} \in D_{n+1} \setminus D_n$, i.e. $z_1 = n + 2 - z_2$. By the induction hypothesis we have that

$$\begin{aligned} T_f(\boldsymbol{z}) &\le X(\boldsymbol{z}) + \min\{T(z_1 - 1, z_2), T(z_1, z_2 - 1)\},\\ T(\boldsymbol{z}) &= X(\boldsymbol{z}) + \text{median}\{T(z_1 - 1, z_2), T(z_1, z_2 - 1), T(z_1 - 1, z_2 - 1)\},\\ T_\ell(\boldsymbol{z}) &\ge X(\boldsymbol{z}) + \max\{T(z_1 - 1, z_2), T(z_1, z_2 - 1)\}, \end{aligned}$$

since all sites in the minimum, median and maximum lie in D_n. Therefore (4.3.2) holds on D_{n+1} by an application of (4.3.3). By induction (4.3.2) holds on $\mathbb{Z}^2_+$, which completes the proof. □

An immediate consequence of Lemma 4.3.1 is that $B_t \subseteq \Gamma_t \subseteq C_t$ when the processes are coupled by their underlying environments. See Figure 4.3.

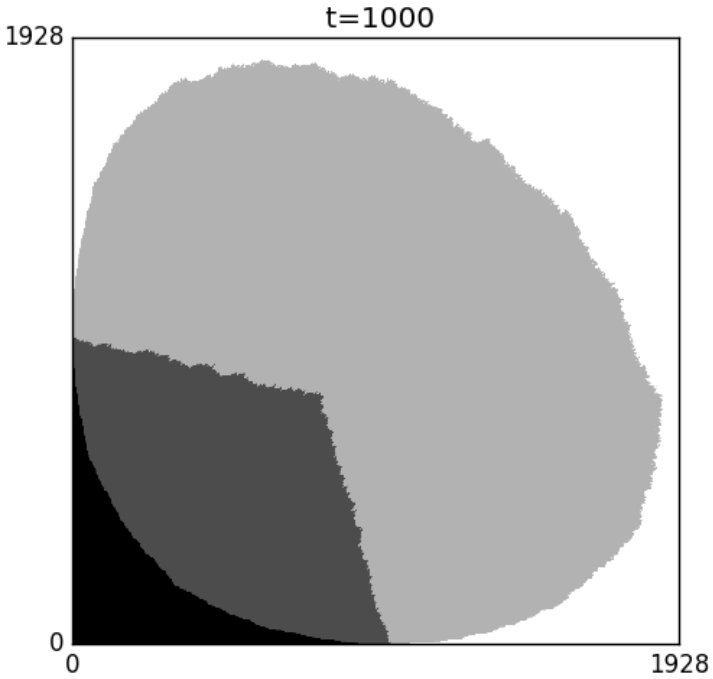


Figure 4.3: A simulation of the shapes $\boldsymbol{B_{1000}}, \boldsymbol{\Gamma_{1000}}$ and C_{1000} on the same i.i.d. random environment with Exp(1) weights.

The following Remarks may be found in [83, Section 5.2] and show that Theorem 4.1 cannot be deduced immediately from Lemma 4.3.1.

Remark 4.3.2. *Suppose that the weights are nonnegative so that the shapes B_t, Γ_t, C_t are connected. If $\mathbb{E}X(\boldsymbol{z})^2 = \infty$, then the shape function g in* (1.8.3) *is infinite throughout the interior and the growth of B_t along any interior direction, i.e. $\theta \in (0, \pi/2)$, is sublinear in t. In this case the asymptotic shape collapses into the boundary.*

If $p_c^{(2)} < \mathbb{P}(X(\boldsymbol{z}) = 0) < 1$, where $p_c^{(2)}$ is the critical probability for directed first–passage percolation in two dimensions, then the shape function h in (1.8.6) *is zero for some cone around the direction $\theta = \pi/4$. In this case, with probability* 1 *there exists an infinite path starting at the origin which has finite total weight, and C_t will have infinite size at some finite time.*

Remark 4.3.2 gives a condition on the random environment for which last–passage percolation does not exhibit linear growth in all directions; however intermediate–passage percolation, by Theorem 4.1, does grow linearly in all directions for this random environment. Similarly first–passage percolation does not have linear growth in all directions for environments with the conditions given in Remark 4.3.2, but intermediate–passage percolation does.

In Section 1.8 we saw that it is convenient to describe the first and last–passage times in terms of paths; superadditivity and subadditivity are extremely useful properties in the derivation of shape theorems for first and last–passage percolation respectively, but no similar obvious property seems to be available in the intermediate–passage case. If the intermediate–passage times have

been defined recursively, we can define a path $\pi = (\pi(1), \pi(2), \ldots)$ from $\boldsymbol{z}$ to $\mathbf{0}$ such that

$$\sum_{n \geq 1} X\big(\pi(n)\big) = T(\boldsymbol{z}).$$

We set $T(\boldsymbol{z}) = -\infty$ for all $\boldsymbol{z} \in E \setminus \mathbf{0}$, where E is the set in (4.2.2), and $T(\boldsymbol{z}) = \infty$ for all $\boldsymbol{z} \in \mathbb{Z}^2 \setminus (E \cup \mathbb{Z}^2_+)$. Then for $\boldsymbol{z} \in \mathbb{Z}^2_+ \setminus \{\mathbf{0}\}$, we set $\pi(1) = \boldsymbol{z}$ and

$$\pi(n+1) = \Big\{ \boldsymbol{y} \in \mathcal{N}\big(\pi(n)\big) : T(\boldsymbol{y}) = \text{median}\big\{T(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{N}\big(\pi(n)\big)\big\} \Big\},$$

choosing arbitrarily on agreements and stopping $\pi(n) = \mathbf{0}$. The boundary passage times in $\mathbb{Z}^2 \setminus \mathbb{Z}^2_+$ guarantee that the path lies in $\mathbb{Z}^2_+$ and $\pi(n) = \mathbf{0}$ for some $n \geq 1$. There is no reason to expect this path description to have any useful properties like subadditivity; it is not clear that it can even be defined without the recursive passage times.

4.4 Proof of Theorem 4.1

Let

$$a_1 \wedge a_2 \wedge \cdots \wedge a_n := \min\{a_1, \ldots, a_n\},$$

and

$$a_1 \vee a_2 \vee \cdots \vee a_n := \max\{a_1, \ldots, a_n\},$$

for $a_1, \ldots, a_n \in \mathbb{R}$ and $n \geq 2$.

The key ingredient in the proof of Theorem 4.1 is the following upper and lower bound on $\Delta T(z_1, z_2) := T(z_1, z_2) - T(z_1 - 1, z_2 - 1)$ for $(z_1, z_2) \in \mathbb{N}^2$. In particular, $\Delta T(\boldsymbol{z})$ is bounded above and below by a function of the weights in $\{\boldsymbol{z}\} \cup \mathcal{N}(\boldsymbol{z})$.

Lemma 4.4.1. *Let $\{X(\boldsymbol{z})\}_{\boldsymbol{z} \in \mathbb{Z}^2_+}$ be an arbitrary environment with nonnegative weights, and define*

$$W(\boldsymbol{z}) := \min\big\{0, \big[X(z_1 - 1, z_2) \wedge X(z_1, z_2 - 1)\big] - X(z_1 - 1, z_2 - 1)\big\}.$$

(a) For all $\boldsymbol{z} \in \mathbb{N}^2$,

$$\Delta T(\boldsymbol{z}) \leq X(\boldsymbol{z}) + X(z_1 - 1, z_2) \vee X(z_1, z_2 - 1); \tag{4.4.1}$$

(b) for all $\boldsymbol{z} \in \mathbb{N}^2$,

$$\Delta T(\boldsymbol{z}) \geq X(\boldsymbol{z}) + W(\boldsymbol{z}). \tag{4.4.2}$$

Before proving Lemma 4.4.1, we show how it quickly yields the result of Theorem 4.1.

Proof of Theorem 4.1. We first check the upper bound in (4.3.1). Let $\theta \in [0, \pi/4]$ and $m(r) = \lfloor r \cos\theta \rfloor - \lfloor r \sin\theta \rfloor$ for $r \geq 0$. The boundary conditions (4.1.1) give

$$T(m(r) - 1, 0) = \sum_{i=1}^{m(r)-1} X(i, 0).$$

If $m(r) = 0$ we set $T(-1, 0) = 0$. By successive application of (4.4.2),

$$\begin{aligned} T(r\cos\theta, r\sin\theta) \geq\, & T(m(r) - 1, 0) + X(\lfloor r\cos\theta\rfloor, \lfloor r\sin\theta\rfloor) \\ & + \sum_{i=0}^{\lfloor r\sin\theta\rfloor - 1} \min\big\{X(m(r) + i, i), X(m(r) + i + 1, i) \wedge X(m(r) + i, i + 1)\big\} \\ \geq\, & T(m(r) - 1, 0) \\ & + \sum_{i=0}^{\lfloor r\sin\theta\rfloor} X(m(r) + i, i) \wedge X(m(r) + i + 1, i) \wedge X(m(r) + i, i + 1). \end{aligned}$$

Writing the RHS as $T(m(r) - 1, 0) + Y_r$, for $\gamma > 0$ the strong law of large numbers yields

$$\lim_{t\to\infty} \frac{T(m(\gamma t) - 1, 0) + Y_{\gamma t}}{t} = \mathbb{E}\big[X(1, 1)\big]\gamma(\cos\theta - \sin\theta) + \mathbb{E}\big[X(1, 1) \wedge X(1, 0) \wedge X(0, 1)\big]\gamma\sin\theta,$$

and we may choose

$$C_2(\theta) \geq \frac{1}{\mathbb{E}[X(1, 1)](\cos\theta - \sin\theta) + \mathbb{E}[X(1, 1) \wedge X(1, 0) \wedge X(0, 1)]\sin\theta},$$

in (4.3.1). The upper bound for $\theta \in (\pi/4, \pi/2]$ follows by symmetry.

It remains to check the lower bound in (4.3.1). Let $\theta \in [0, \pi/4]$. Then by successive application of (4.4.1),

$$\begin{aligned} T(r\cos\theta, r\sin\theta) \leq\, & T(m(r), 0) \\ & + \sum_{j=1}^{\lfloor r\sin\theta\rfloor} \big[X(m(r) + j, j) + X(m(r) + j - 1, j) \vee X(m(r) + j, j - 1)\big]. \end{aligned}$$

Writing the RHS as $T(m(r), 0) + Z_r$, for $\gamma > 0$ the strong law of large numbers yields

$$\lim_{t\to\infty} \frac{T(m(\gamma t), 0) + Z_{\gamma t}}{t} = \mathbb{E}\big[X(1, 1)\big]\gamma\cos\theta + \mathbb{E}\big[X(1, 0) \vee X(0, 1)\big]\gamma\sin\theta$$

and we may choose

$$C_1(\theta) \leq \frac{1}{\mathbb{E}[X(1, 1)]\cos\theta + \mathbb{E}[X(1, 0) \vee X(0, 1)]\sin\theta},$$

in (4.3.1). The case $\theta \in (\pi/4, \pi/2]$ follows by symmetry. □

Lemma 4.4.1 is proved by exploiting overlapping neighbourhoods of nearby sites. We will prove part (a) and part (b) of Lemma 4.4.1 separately, beginning with part (a).

Proof of Lemma 4.4.1a. We first check (4.4.1) when $\boldsymbol{z}$ is close to the boundary sites. Let $\boldsymbol{z} = (z_1, 1)$ for some $z_1 \in \mathbb{N}$. By (4.1.1), it holds that $T(z_1, 0) = T(z_1 - 1, 0) + X(z_1, 0)$, and since the environment is nonnegative $T(z_1, 0) \geq T(z_1 - 1, 0)$. Therefore

$$\text{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(z_1, 1)\big\} \leq T(z_1, 0),$$

and

$$\begin{aligned}\Delta T(z_1,1) &\le X(z_1,1)+X(z_1,0)\\ &\le X(z_1,1)+X(z_1,0)\vee X(z_1-1,1).\end{aligned}$$

By symmetry

$$\begin{aligned}\Delta T(1,z_2) &\le X(1,z_2)+X(0,z_2)\\ &\le X(1,z_2)+X(0,z_2)\vee X(1,z_2-1),\end{aligned}$$

for all $z_2\in\mathbb{N}$. Suppose now that $\boldsymbol{z}\in\mathbb{N}^2$ with $z_1,z_2>1$. Letting $\boldsymbol{z}'=(z_1-1,z_2-1)$, if

$$T(\boldsymbol{z}')\ge \text{median}\big\{T(\boldsymbol{y}):\boldsymbol{y}\in\mathcal{N}(\boldsymbol{z})\big\},$$

then by (4.1.3)

$$\Delta T(\boldsymbol{z})\le X(\boldsymbol{z}),$$

and (4.4.1) follows by nonnegativity of the random environment. We now assume that

$$T(\boldsymbol{z}')=\min\big\{T(\boldsymbol{y}):\boldsymbol{y}\in\mathcal{N}(\boldsymbol{z})\big\}. \tag{4.4.3}$$

Since $\min\{a,b\}\le\text{median}\{a,b,c\}$ for all $a,b,c\in\mathbb{R}$, by the recursion (4.1.3) for $\boldsymbol{z}'$ we have

$$\begin{aligned}T(\boldsymbol{z}') &\ge X(\boldsymbol{z}')+T(z_1-2,z_2-1)\wedge T(z_1-1,z_2-2)\\ &\ge T(z_1-2,z_2-1)\wedge T(z_1-1,z_2-2).\end{aligned} \tag{4.4.4}$$

Noting that $(z_1-2,z_2-1)\in\mathcal{N}(z_1-1,z_2)$, $(z_1-1,z_2-2)\in\mathcal{N}(z_1,z_2-1)$ and $\boldsymbol{z}'$ lies in the intersection $\mathcal{N}(z_1-1,z_2)\cap\mathcal{N}(z_1,z_2-1)$, it follows from (4.4.4) that

$$T(\boldsymbol{z}')\ge\text{median}\big\{T(\boldsymbol{y}):\boldsymbol{y}\in\mathcal{N}(z_1-1,z_2)\big\}\wedge\text{median}\big\{T(\boldsymbol{y}):\boldsymbol{y}\in\mathcal{N}(z_1,z_2-1)\big\}. \tag{4.4.5}$$

By the recursion (4.1.3) for (z_1-1,z_2) and (z_1,z_2-1), it follows from (4.4.5) that

$$\begin{aligned}T(\boldsymbol{z}') &\ge \min\big\{T(z_1-1,z_2)-X(z_1-1,z_2),T(z_1,z_2-1)-X(z_1,z_2-1)\big\}\\ &\ge \min\big\{T(z_1-1,z_2),T(z_1,z_2-1)\big\}-X(z_1-1,z_2)\vee X(z_1,z_2-1).\end{aligned} \tag{4.4.6}$$

By assumption (4.4.3), and the recursion (4.1.3) for $\boldsymbol{z}$, we have that

$$\begin{aligned}T(\boldsymbol{z}) &= X(\boldsymbol{z})+\min\big\{T(z_1-1,z_2),T(z_1,z_2-1)\big\}\\ &\le T(\boldsymbol{z}')+X(\boldsymbol{z})+X(z_1-1,z_2)\vee X(z_1,z_2-1),\end{aligned}$$

where the last inequality follows from (4.4.6). □

Proof of Lemma 4.4.1b. We first check (4.4.2) when $\boldsymbol{z}$ is close to the boundary sites. Let $\boldsymbol{z}=(z_1,1)$ for some $z_1\in\mathbb{N}$. By (4.1.1) and nonnegativity of the random environment we have that $T(z_1-1,0)\le T(z_1,0)$ and therefore $T(z_1-1,0)\le\text{median}\{T(\boldsymbol{y}):\boldsymbol{y}\in\mathcal{N}(z_1,1)\}$. It follows that

$$T(z_1,1)-T(z_1-1,0)\ge X(z_1,1),$$

and since $W(\boldsymbol{z})\le 0$, (4.4.2) holds. By symmetry

$$T(1,z_2)-T(0,z_2-1)\ge X(1,z_2),$$

for all $z_2 \in \mathbb{N}$. Let $\boldsymbol{z} \in \mathbb{N}^2$ with $z_1, z_2 > 1$ and $\boldsymbol{z}' = (z_1 - 1, z_2 - 1)$. If

$$T(\boldsymbol{z}') \leq \text{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\big\},$$

then by the recursion (4.1.3)

$$T(\boldsymbol{z}) - T(\boldsymbol{z}') \geq X(\boldsymbol{z}),$$

and since $W(\boldsymbol{z}) \leq 0$, (4.4.2) holds. It remains to show (4.4.2) when

$$T(\boldsymbol{z}') = \max\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(\boldsymbol{z})\big\}. \tag{4.4.7}$$

Let

$$\begin{aligned} M_1 &= \text{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(z_1 - 1, z_2)\big\}, \\ M_2 &= \text{median}\big\{T(\boldsymbol{y}) : \boldsymbol{y} \in \mathcal{N}(z_1, z_2 - 1)\big\}. \end{aligned}$$

By the recursion (4.1.3) for both $(z_1 - 1, z_2)$ and $(z_1, z_2 - 1)$,

$$\begin{aligned} T(z_1 - 1, z_2) \vee T(z_1, z_2 - 1) &= \big[X(z_1 - 1, z_2) + M_1\big] \vee \big[X(z_1, z_2 - 1) + M_2\big] \\ &\geq X(z_1 - 1, z_2) \wedge X(z_1, z_2 - 1) + M_1 \vee M_2, \end{aligned} \tag{4.4.8}$$

where the last inequality follows from the fact that $\max\{a_1 + b_1, a_2 + b_2\} \geq \min\{a_1, a_2\} + \max\{b_1, b_2\}$ for all $a_1, a_2, b_1, b_2 \in \mathbb{R}$. Since $\boldsymbol{z}' \in \mathcal{N}(z_1 - 1, z_2)$ and by (4.4.7) $T(\boldsymbol{z}') \geq T(z_1 - 1, z_2)$, it holds that

$$\begin{aligned} T(z_1 - 1, z_2) &= X(\boldsymbol{z}) + M_1 \\ &= X(\boldsymbol{z}) + T(z_1 - 2, z_2) \vee T(z_1 - 2, z_2 - 1). \end{aligned}$$

In particular,

$$M_1 \geq T(z_1 - 2, z_2 - 1), \tag{4.4.9}$$

and by symmetry

$$M_2 \geq T(z_1 - 1, z_2 - 2). \tag{4.4.10}$$

Combining (4.4.9) and (4.4.10) with (4.4.8) yields

$$\begin{aligned} &T(z_1 - 1, z_2) \vee T(z_1, z_2 - 1) \geq \\ &\qquad X(z_1 - 1, z_2) \wedge X(z_1, z_2 - 1) + T(z_1 - 2, z_2 - 1) \vee T(z_1 - 1, z_2 - 2). \end{aligned} \tag{4.4.11}$$

Noting that $\max\{a, b\} \geq \text{median}\{a, b, c\}$ for all $a, b, c \in \mathbb{R}$, an application of the recursion (4.1.3) for $\boldsymbol{z}'$ gives that

$$T(\boldsymbol{z}') \leq X(\boldsymbol{z}') + T(z_1 - 2, z_2 - 1) \vee T(z_1 - 1, z_2 - 2),$$

and combining this with (4.4.11) gives

$$T(z_1 - 1, z_2) \vee T(z_1, z_2 - 1) \geq T(\boldsymbol{z}') + X(z_1 - 1, z_2) \wedge X(z_1, z_2 - 1) - X(\boldsymbol{z}'). \tag{4.4.12}$$

By assumption (4.4.7) and the recursion (4.1.3) for $\boldsymbol{z}$,

$$T(\boldsymbol{z}) = X(\boldsymbol{z}) + T(z_1 - 1, z_2) \vee T(z_1, z_2 - 1),$$

and (4.4.12) yields

$$\begin{aligned} \Delta T(\boldsymbol{z}) &\geq X(\boldsymbol{z}) + X(z_1 - 1, z_2) \wedge X(z_1, z_2 - 1) - X(\boldsymbol{z}') \\ &\geq X(\boldsymbol{z}) + W(\boldsymbol{z}), \end{aligned}$$

which completes the proof. □

4.5 Open questions and simulations

The following list is a collection of open questions for intermediate–passage percolation motivated by simulations:

- Does an asymptotic shape $\Gamma \subseteq \mathbb{R}^2_+$ exist for the process $(\Gamma_t)_{t\geq 0}$ under the conditions of Theorem 4.1? In particular, does it hold with probability 1 that

$$(1-\epsilon)\Gamma \subseteq \frac{\Gamma_t}{t} \subseteq (1+\epsilon)\Gamma,$$

  for all $\epsilon \in (0,1)$ and t sufficiently large? We conjecture that such a shape exists.

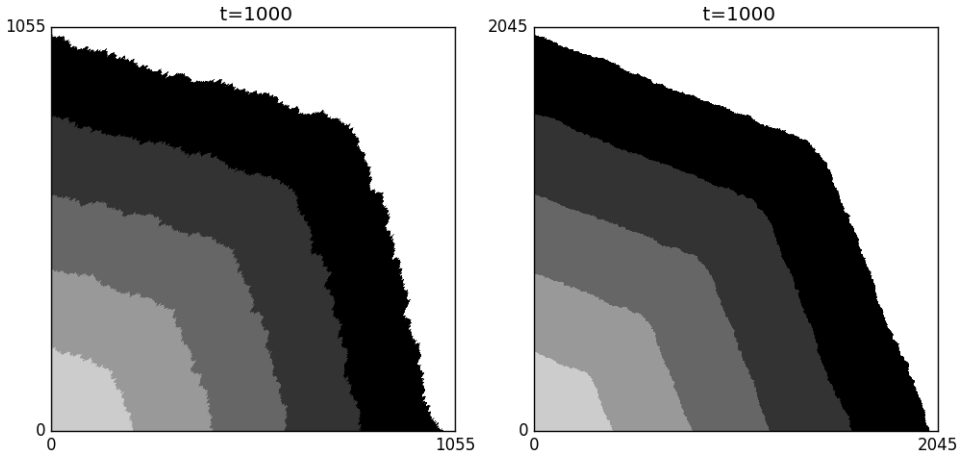
- Flat edges appear to occur very commonly for intermediate–passage percolation. Does there exist a weight distribution such that the limiting shape, if it exists, is strictly convex? If so, can we characterise the conditions under which flat edges occur?

- Can we describe the possible behaviours of the boundary of the limiting shape, if it exists, around the diagonal $\theta = \pi/4$? The simulations for Bernoulli weights (Figure 4.5) with low probability indicate a point which is not differentiable at $\theta = \pi/4$, whereas the simulations for larger p indicate a region in which the shape is strictly convex and has smooth boundary in some interval around $\pi/4$. Simulations for geometric weights (Figure 4.7) with high probability indicate a flat region centered around $\pi/4$.

- For intermediate–passage percolation with Bernoulli weights, does there exist a range of probabilities (p_1, p_2) such that the boundary of the limiting shape consists of exactly two flat edges? Moreover, if $p_2 < p < 1$, does there exist a cone around the diagonal where the limiting shape is strictly convex in the cone and flat outside of this region?

- For geometric weights (Figures 4.6 and 4.7), which values of p give rise to a flat edge in the boundary of the limiting shape? When p is small simulations indicate two flat edges away from the diagonal, whereas large values of p indicate a flat edge around the diagonal which gets larger as p approaches 1.

Figure 4.4 gives simulations of Γ_t for exponential and uniform weights. Figure 4.5 gives simulations of Γ_t with Bernoulli weights for various values of p. Figures 4.6 and 4.7 give simulations of Γ_t with Geometric weights, for various values of t and p. Figures 4.8 and 4.9 give simulations of the shape Γ_t with chi–squared and Poisson weights respectively.



(a) Exp(1) (b) Uniform$(0, 1)$

Figure 4.4: Figures 4.4a and 4.4b are simulations of Γ_{1000} for Exp(1) and Uniform$(0, 1)$ weights respectively.

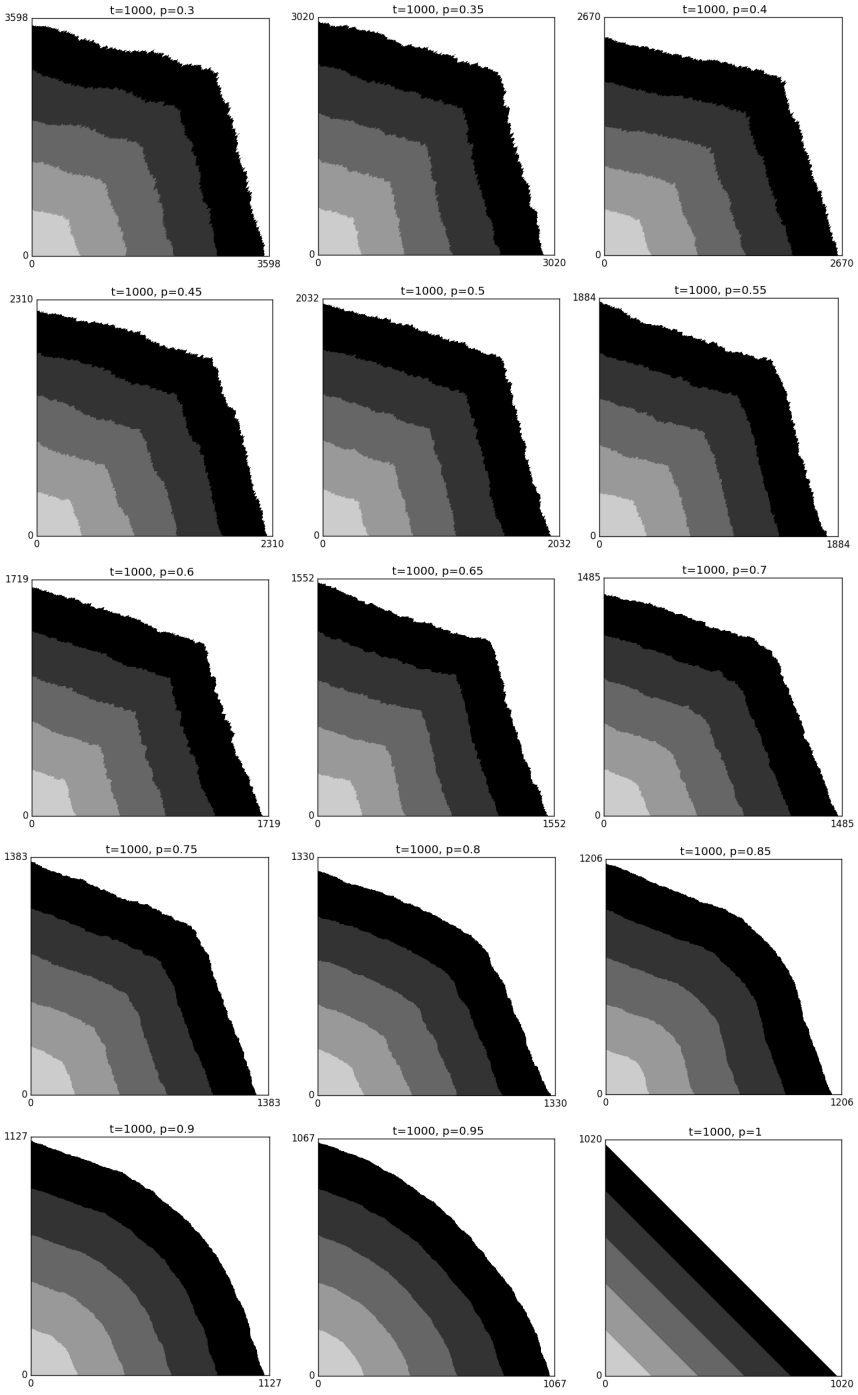


Figure 4.5: Simulations of the shape Γ_{1000} with i.i.d. Bernoulli(p) weights, satisfying $\mathbb{P}[X(\boldsymbol{z}) = 1] = 1 - \mathbb{P}[X(\boldsymbol{z}) = 0] = p$, for each $p = 0.3, 0.35, \ldots, 1$. The shapes $\Gamma_{200}, \Gamma_{400} \backslash \Gamma_{200}, \ldots, \Gamma_{1000} \backslash \Gamma_{800}$ are shaded from lightest to darkest.

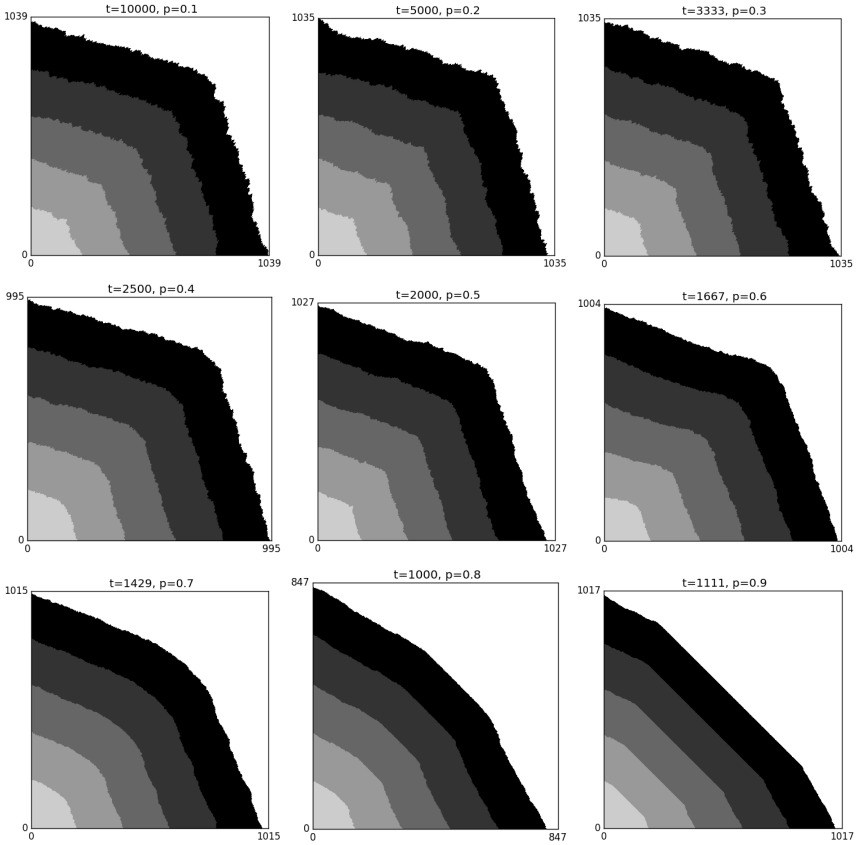


Figure 4.6: Simulations of Γ_t for i.i.d Geometric(p) weights with mean $1/p$. For each $p = 0.1, 0.2, \ldots, 0.9$ we have simulated Γ_t, where t is given by $1000/p$ rounded to the nearest integer.

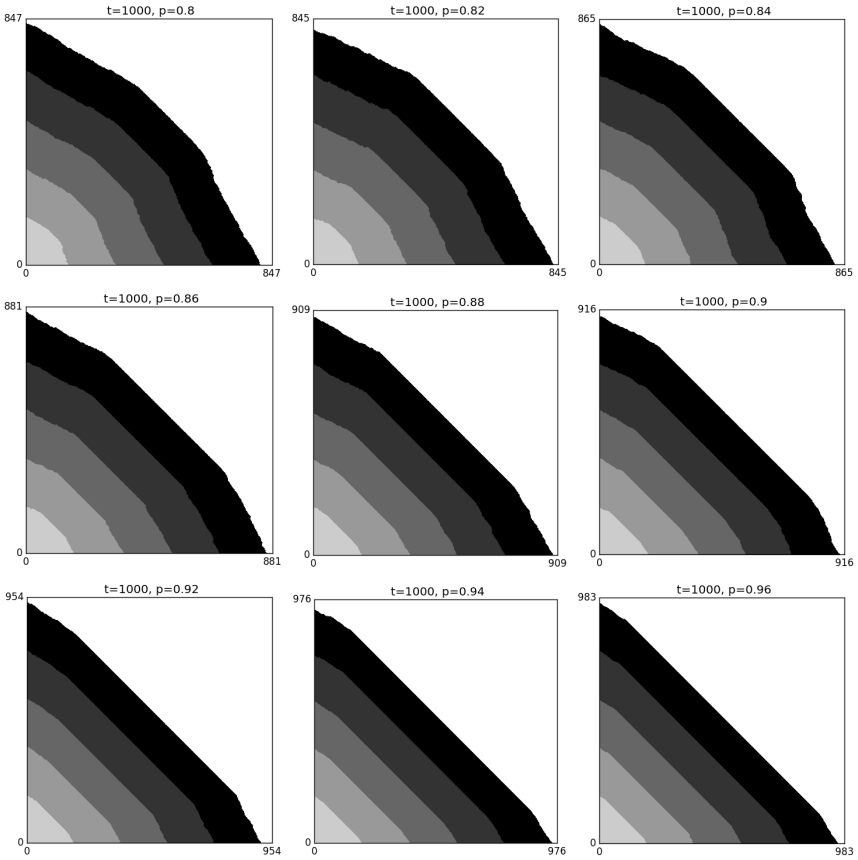


Figure 4.7: Simulations of Γ_{1000} for i.i.d Geometric(p) weights, with mean $1/p$, for each $p = 0.8, 0.82, \ldots, 0.96$.

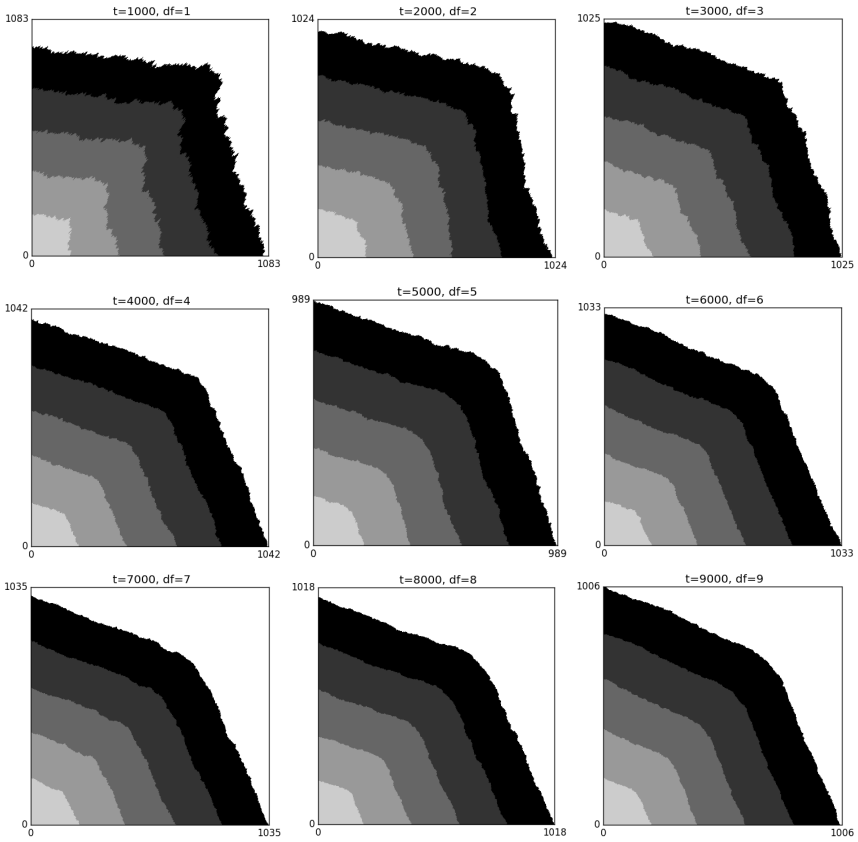


Figure 4.8: Simulations of Γ_t for i.i.d. Chi–squared weights with df degrees of freedom. For each $df = 1, 2, \ldots, 9$ we have simulated Γ_t for $t = df \times 1000$.

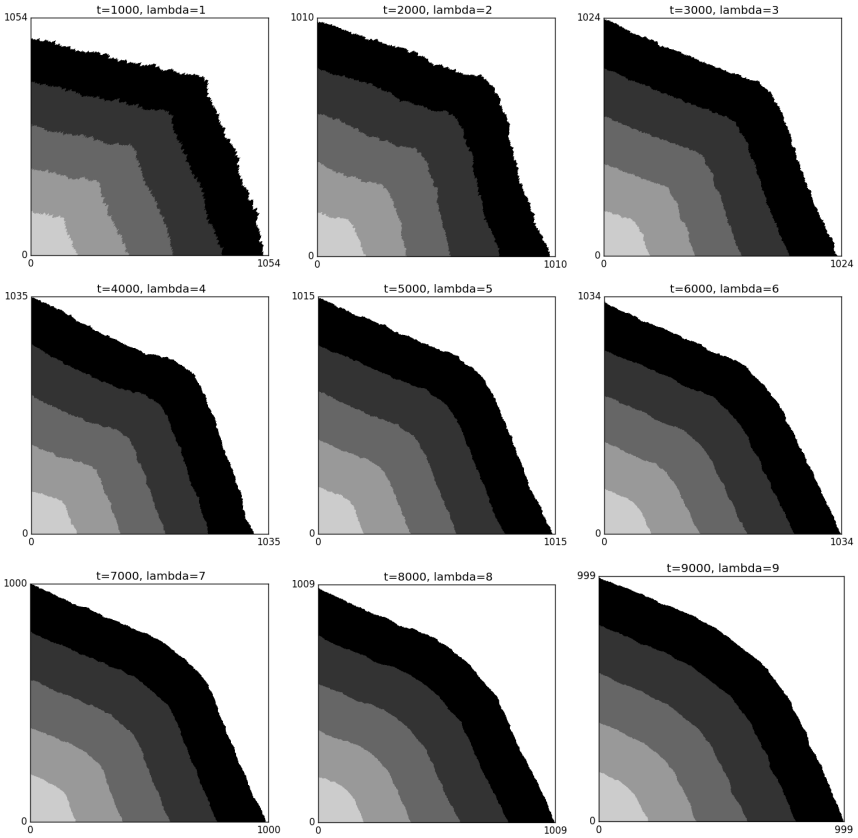


Figure 4.9: Simulations of Γ_t for i.i.d. Poisson(λ) weights. For each $\lambda = 1, 2, \ldots, 9$ we have simulated Γ_t with $t = \lambda \times 1000$.

Chapter 5

Conclusions and further research

In Chapter 2 we were able to give bounds on the mixing time of the FSEP on both the segment and circle. In some situations we showed that the transient part of the dynamics determined the mixing time, or at the very least had the same order. The mixing time bounds were achieved by comparison to the SEP and OBEP, and an application of known results and ideas for these processes. These comparisons were very successful on the segment due to the monotonicity of the lattice path dynamics. These comparisons were more difficult to directly apply on the circle, and while we were still able to give some good bounds, we were not able to give an upper bound of the desired order for the hitting time of the ergodic component for a class of initial conditions. Further research is needed to close this gap in the argument, although we expect that all of the 'worst' initial conditions lie in the set we already considered.

We conjecture that cutoff holds on the segment for the symmetric FSEP in the general setting, and we expect that Figure 2.4a will be a useful object in the hunt for cutoff. Our original arguments did not have this picture in mind, and we believe that the upper bound in cutoff can be achieved by careful consideration of these lattice path dynamics. It would be interesting to precisely characterise the cutoff constant, as when $N-k$ is close to $N/2$ the hitting time of the ergodic component determines the mixing time.

We also conjecture that cutoff holds on the circle, although it is less clear how to tackle this problem. The argument for cutoff of the SSEP on the circle is very different to the arguments we gave for the FSEP on the circle, and it is possible that an analogous argument can be constructed.

In Chapter 3 we produced two algorithms, with expected run–time $O(k * \text{polylog}N)$, which output the cycle structure of a uniform random permutation with a given number of cycles. Further investigation of the Labelled process led to the discovery of an interesting property: coalescence occurs for a sequence of updates $U_1, \ldots, U_t$ if and only if it occurs for the same sequence in reverse order, and we gave a class of models for which this holds. This property was found with little intuition, and further investigation into which processes have this property (and why) would be of interest.

We were also able to give an explicit construction of a fragmentation process, and gave an algorithm with expected run–time $O(k^2\text{polylog}N)$ which output the trajectory of a fragmentation process. The fragmentation process is not unique; there are other joint distributions which satisfy this property. Further research into the construction of processes which give rise to different joint distributions satisfying the properties of a fragmentation process would be intriguing.

Finally, in Chapter 4 we investigated properties of the intermediate–passage process. We showed that this process has linear growth when the environment is nonnegative with finite first moment, then stated several open problems motivated by simulations.

Bibliography

[1] F. Alcaraz and R. Bariev. Exact solution of the asymmetric exclusion model with particles of arbitrary size. *Physical Review E*, 60:79–88, 1999.

[2] D. Aldous and M. Brown. Inequalities for rare events in time-reversible Markov chains I. *Lecture Notes-Monograph Series*, 22:1–16, 1992.

[3] D. Aldous and P. Diaconis. Shuffling cards and stopping times. *The American Mathematical Monthly*, 93:333–348, 1986.

[4] D. Aldous and P. Diaconis. The asymmetric one–dimensional constrained Ising model: Rigorous results. *Journal of Statistical Physics*, 107:945–975, 2002.

[5] E. Andjel. Invariant measures for the zero range process. *The Annals of Probability*, 10:525–547, 1982.

[6] M. Arcones. Large deviations of empirical processes. In *High Dimensional Probability III*, volume 55, pages 205–223. Birkhäuser Basel, 2003.

[7] J. Arndt. *Generating random permutations.* PhD thesis, Australian National University, 2010.

[8] A. Auffinger and M. Damron. Differentiability at the edge of the percolation cone and related results in first–passage percolation. *Probability Theory and Related Fields*, 156:193–227, 2011.

[9] F. Baccelli, A. Borovkov, and J. Mairesse. Asymptotic results on infinite tandem queueing networks. *Probability Theory and Related Fields*, 118:365–405, 2000.

[10] J. Baik, G. Barraquand, I. Corwin, and T. Suidan. Facilitated exclusion process. In *Computation and Combinatorics in Dynamics, Stochastics and Control*, volume 13, pages 1–35. Springer International Publishing, 2018.

[11] P. Balister, B. Bollobás, M. Przykucki, and P. Smith. Subcritical $\mathcal{U}$–bootstrap percolation models have non–trivial phase transitions. *Transactions of the American Mathematical Society*, 368:7358–7411, 2016.

[12] U. Basu and P. Mohanty. Active–absorbing-state phase transition beyond directed percolation: A class of exactly solvable models. *Physical Review E*, 79:041143, 2009.

[13] I. Benjamini, N. Berger, C. Hoffman, and E. Mossel. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Transactions of the American Mathematical Society*, 357:3013–3029, 2002.

[14] N. Berestycki and J. Pitman. Gibbs distributions for random partitions generated by a fragmentation process. *Journal of Statistical Physics*, 127:381–418, 2007.

[15] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Large deviations for the boundary driven symmetric simple exclusion process. *Mathematical Physics, Analysis and Geometry*, 6:231–267, 2003.

[16] L. Bertini and C. Toninelli. Exclusion processes with degenerate rates: Convergence to equilibrium and tagged particle. *Journal of Statistical Physics*, 117:549–580, 2004.

[17] O. Blondel, N. Cancrini, F. Martinelli, C. Roberto, and C. Toninelli. Fredrickson - Andersen one spin facilitated model out of equilibrium. *Markov Processes and Related Fields*, 19:383–406, 2013.

[18] O. Blondel, C. Erignoux, M. Sasada, and M. Simon. Hydrodynamic limit for a facilitated exclusion process. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56:667–714, 2020.

[19] O. Blondel, C. Erignoux, and M. Simon. Stefan problem for a nonergodic facilitated exclusion process. *Probability and Mathematical Physics*, 2:127–178, 2021.

[20] B. Bollobás, H. Duminil–Copin, R. Morris, and P. Smith. Universality for two–dimensional critical cellular automata. *Proceedings of the London Mathematical Society*, 2022.

[21] N. Cancrini, F. Martinelli, R. H. Schonmann, and C. Toninelli. Facilitated oriented spin models: some non equilibrium results. *Journal of Statistical Physics*, 138:1109–1123, 2010.

[22] P. Caputo, T. Liggett, and T. Richthammer. Proof of Aldous' spectral gap conjecture. *Journal of the American Mathematical Society*, 23:831–851, 2010.

[23] F. Cesi. Quasi–factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120:569–584, 2001.

[24] J. Chalupa, P. Leath, and G. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C: Solid State Physics*, 12:31–35, 1979.

[25] P. Chleboun, A. Faggionato, and F. Martinelli. Mixing time and local exponential ergodicity of the East–like process in $\mathbb{Z}^d$. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 24:717–743, 2015.

[26] F. Ciech and N. Georgiou. Order of the variance in the discrete Hammersley process with boundaries. *Journal of Statistical Physics*, 176:591–638, 2019.

[27] J. Cox and R. Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *The Annals of Probability*, 9:583–603, 1981.

[28] A. De Acosta. Moderate deviations and associated laplace approximations for sums of independent random vectors. *Transactions of the American Mathematical Society*, 329:357–375, 1992.

[29] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6:695–750, 1996.

[30] R. Durrett and T. Liggett. The shape of the limit set in Richardson's growth model. *The Annals of Probability*, 9:186–193, 1981.

[31] Richard Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12:999–1040, 1984.

[32] R. Durstenfeld. Algorithm 235: Random permutation. *Communications of the association for computing machinery*, 7:420, 1964.

[33] P. Eichelsbacher and M. Löwe. Moderate deviations for I.I.D. random variables. *ESAIM: Probability and Statistics*, 7:209–218, 2003.

[34] M. Evans. Phase transitions in one-dimensional nonequilibrium systems. *Brazilian Journal of Physics*, 30:42–57, 2000.

[35] A. Ferreira and F. Alcaraz. Anomalous tag diffusion in the asymmetric exclusion model with particles of arbitrary sizes. *Physical Review E*, 65:052102, 2002.

[36] J. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *The Annals of Applied Probability*, 1:62–87, 1991.

[37] R. Fisher and F. Yates. *Statistical tables for biological, agricultural and medical research.* Edinburgh: Oliver and Boyd, 1938.

[38] G. Fredrickson and H. Andersen. Kinetic Ising model of the glass transition. *Physical Review Letters*, 53:1244–1247, 1984.

[39] A. Gabel, P. Krapivsky, and S. Redner. Facilitated asymmetric exclusion. *Physical Review Letters*, 105:210603, 2010.

[40] A. Gabel and S. Redner. Cooperativity-driven singularities in asymmetric exclusion. *Journal of Statistical Mechanics: Theory and Experiment*, 2011:P06008, 2011.

[41] N. Gantert, E. Nestoridi, and D. Schmid. Mixing times for the simple exclusion process with open boundaries. *arXiv e-prints*, page arXiv:2003.03781, 2020.

[42] P. Glynn and W. Whitt. Departures from many queues in series. *The Annals of Applied Probability*, 1:546–572, 1991.

[43] A. Gnedin and J. Pitman. Poisson representation of a Ewens fragmentation process. *Combinatorics, Probability and Computing*, 16:819–827, 2007.

[44] C. Goldschmidt, J. Martin, and D. Spano. Fragmenting random permutations. *Electronic Communications in Probability*, 13:461–474, 2008.

[45] S. Goldstein, J. Lebowitz, and E. Speer. Exact solution of the facilitated totally asymmetric simple exclusion process. *Journal of Statistical Mechanics: Theory and Experiment*, 2019:123202, 2019.

[46] S. Goldstein, J. Lebowitz, and E. Speer. The discrete-time facilitated totally asymmetric simple exclusion process. *Pure and Applied Functional Analysis*, 6:177–203, 2021.

[47] S. Goldstein, J. Lebowitz, and E. Speer. Stationary states of the one-dimensional discrete-time facilitated symmetric exclusion process. *Journal of Mathematical Physics*, 63:083301, 2022.

[48] P. Gonçalves, C. Landim, and C. Toninelli. Hydrodynamic limit for a particle system with degenerate rates. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 45:887–909, 2009.

[49] P. Gonçalves, M. Jara, R. Marinho, and O. Menezes. Sharp convergence to equilibrium for the SSEP with reservoirs. *arXiv e-prints*, page arXiv:2110.06353, 2021.

[50] O. Häggström. *Finite Markov Chains and Algorithmic Applications*. London Mathematical Society Student Texts. Cambridge University Press, 2002.

[51] O. Häggström and K. Nelander. On exact simulation of Markov random fields using coupling from the past. *Scandinavian Journal of Statistics*, 26:395–411, 1999.

[52] J. Hermon and R. Pymar. The exclusion process mixes (almost) faster than independent particles. *Annals of Probability*, 48:3077–3123, 2020.

[53] R. Holley. A class of interactions in an infinite particle system. *Advances in Mathematics*, 5:291–309, 1970.

[54] M. Huber. Exact sampling and approximate counting techniques. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, page 31–40. Association for Computing Machinery, 1998.

[55] K. Johansson. Shape fluctuations and random matrices. *Communications in Mathematical Physics*, 209:437–476, 2000.

[56] W. Kendall. *Notes on perfect simulation*, pages 93–146. World Scientific, 2005.

[57] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*. Springer-Verlag, 1999.

[58] D. Knuth. *The Art of Computer Programming, third edition, Volume 2: Seminumerical Algorithms*. Addison-Wesley, 1997.

[59] D. Knuth. *The Art of Computer Programming, second edition, Volume 3: Sorting and Searching*. Addison-Wesley, 1998.

[60] C. Labbé and H. Lacoin. Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *The Annals of Probability*, 2016.

[61] C. Labbé and H. Lacoin. Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *The Annals of Probability*, 47:1541–1586, 2019.

[62] H. Lacoin. The cutoff profile for the simple exclusion process on the circle. *The Annals of Probability*, 44:3399–3430, 2016.

[63] H. Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *The Annals of Probability*, 44:1426–1487, 2016.

[64] H. Lacoin. The simple exclusion process on the circle has a diffusive cutoff window. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53:1402–1437, 2017.

[65] H. Lacoin and R. Leblond. Cutoff phenomenon for the simple exclusion process on the complete graph. *Latin American Journal of Probability and Mathematical Statistics*, 8:285–301, 2011.

[66] H. Lacoin, F. Simenhaus, and F. Toninelli. Zero–temperature $2d$ stochastic Ising model and anisotropic curve–shortening flow. *Journal of the European Mathematical Society*, 16(12):2557–2615, 2014.

[67] G. Lakatos and T. Chou. Totally asymmetric exclusion processes with particles of arbitrary size. *Journal of Physics A: Mathematical and General*, 36:2027, 2003.

[68] C. Landim, A. Milanés, and S. Olla. Stationary and nonequilibrium fluctuations in boundary driven exclusion processes. *Markov Processes And Related Fields*, 14:165–184, 2008.

[69] T-Y. Lee and H-T. Yau. Logarithmic Sobolev inequality for some models of random walks. *The Annals of Probability*, 26:1855–1873, 1998.

[70] D. Levin, Y. Peres, and E. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.

[71] T. Liggett. Existence theorems for infinite particle systems. *Transactions of the American Mathematical Society*, 165:471–481, 1972.

[72] T. Liggett. An Infinite Particle System with Zero Range Interactions. *The Annals of Probability*, 1:240–253, 1973.

[73] T. Liggett. A characterization of the invariant measures for an infinite particle system with interactions. *Transactions of the American Mathematical Society*, 179:433–453, 1973.

[74] T. Liggett. A characterization of the invariant measures for an infinite particle system with interactions. II. *Transactions of the American Mathematical Society*, 198:201–213, 1974.

[75] T. Liggett. Random invariant measures for Markov chains, and independent particle systems. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 45:297–313, 1978.

[76] T. Liggett. *Interacting Particle Systems*. Springer Berlin Heidelberg, 1985.

[77] T. Liggett. Stochastic models of interacting systems. *The Annals of Probability*, 25:1–29, 1997.

[78] T. Liggett. *Continuous Time Markov Processes: An Introduction*. Graduate studies in mathematics. American Mathematical Society, 2010.

[79] C. MacDonald and J. Gibbs. Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers*, 7:707–725, 1969.

[80] C. MacDonald, J. Gibbs, and A. Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, 6:1–25, 1968.

[81] R. Marchand. Strict inequalities for the time constant in first passage percolation. *The Annals of Applied Probability*, 12:1001–1038, 2002.

[82] J. Martin. Large tandem queueing networks with blocking. *Queueing Systems*, 41:45–72, 2002.

[83] J. Martin. Limiting shape for directed percolation models. *The Annals of Probability*, 32:2908–2937, 2004.

[84] J. Martin. Last-passage percolation and with general weight distribution. *Markov Processes and Related Fields*, 12:237–299, 2006.

[85] F. Martinelli, R. Morris, and C. Toninelli. Universality results for kinetically constrained spin models in two dimensions. *Communications in Mathematical Physics*, 369:761–809, 2018.

[86] Mathoverflow. How can I generate random permutations of $[n]$ with k cycles, where k is much larger than $\log n$? `https://mathoverflow.net/q/26974`. 2010.

[87] B. Morris. The mixing time for simple exclusion. *The Annals of Applied Probability*, 16:615–635, 2006.

[88] Y. Nagahata. Lower bound estimate of the spectral gap for simple exclusion process with degenerate rates. *Electronic Journal of Probability*, 17:1–19, 2012.

[89] R. Oliveira. Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk. *The Annals of Probability*, 41:871–913, 2013.

[90] J. Pitman. *Combinatorial stochastic processes*. Springer-Verlag, 2006.

[91] J. Propp and D. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223–252, 1996.

[92] M. Rossi, R. Pastor-Satorras, and A. Vespignani. Universality class of absorbing phase transitions with a conserved field. *Physical Review Letters*, 85:1803–1806, 2000.

[93] H Rost. Non–equilibrium behaviour of a many particle process: Density profile and local equilibria. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 58:41–53, 1981.

[94] J. Salez. A sharp log-Sobolev inequality for the multislice. *Annales Henri Lebesgue*, 4:1143–1161, 2021.

[95] J. Salez. Universality of cutoff for exclusion with reservoirs. *arXiv e-prints*, page arXiv:2201.03463, 2022.

[96] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996*, pages 301–413. Springer Berlin Heidelberg, 1997.

[97] T. Sasamoto and M. Wadati. Exact results for one–dimensional totally asymmetric diffusion models. *Journal of Physics A: Mathematical and General*, 31:6057, 1998.

[98] D. Schmid and A. Sly. Mixing times for the TASEP on the circle. *arXiv e-prints*, page arXiv:2203.11896, 2022.

[99] L. Shaw, R. Zia, and K. Lee. Totally asymmetric exclusion process with extended objects: A model for protein synthesis. *Physical Review E*, 68:021910, 2003.

[100] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.

[101] F. Spitzer. Interaction of Markov processes. *Advances in Mathematics*, 5:246–290, 1970.

[102] F. Spitzer. Recurrent random walk of an infinite particle system. *Transactions of the American Mathematical Society*, 198:191–199, 1974.

[103] E. Waymire. Zero–range interaction at Bose–Einstein speeds under a positive recurrent single particle law. *The Annals of Probability*, 8:441–450, 1980.

[104] H. Wilf. *Combinatorial Algorithms*. Society for Industrial and Applied Mathematics, 1989.

[105] D. Wilson. How to couple from the past using a read–once source of randomness. *Random Structures and Algorithms*, 16:85–113, 2000.

[106] D. Wilson. Mixing times of lozenge tiling and card shuffling markov chains. *The Annals of Applied Probability*, 14:274–325, 2004.

[107] H-T. Yau. Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probability and related fields*, 109:507–538, 1997.