

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Topics in discrete random structures

ANDERS MARTINSSON

CHALMERS



UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
AND UNIVERSITY OF GOTHENBURG
Göteborg, Sweden 2017

Topics in discrete random structures

Anders Martinsson

ISBN 978-91-7597-561-0

© Anders Martinsson, 2017

Doktorsavhandlingar vid Chalmers tekniska högskola

Ny serie nr 4242

ISSN 0346-718X

Department of Mathematical Sciences

Chalmers University of Technology

and University of Gothenburg

SE-412 96 Göteborg, Sweden

Phone: +46 (0)31-772 10 00

Cover: Artistic illustration of a jigsaw puzzle with multiple solutions, as considered in Paper IV. The motif is inspired by *The Burning Giraffe* by Salvador Dalí.

Printed in Göteborg, Sweden, 2017

Topics in discrete random structures

Anders Martinsson

*Department of Mathematical Sciences,
Chalmers University of Technology and University of Gothenburg*

Abstract

This thesis presents four papers on problems in discrete probability. A common theme of the articles is to take some class of discrete structures, impose some randomness, and then consider what happens asymptotically as the size of the structure tends to infinity.

Paper I concerns first–passage percolation on Cartesian graph powers G^d of some fixed base graph G as $d \rightarrow \infty$. We propose a natural asymptotic lower bound on the first–passage time between (v, v, \dots) and (w, w, \dots) , which we call the critical time. Our main result characterizes when this lower bound is sharp. As a consequence we are able to determine the so–called diagonal time constant of \mathbb{Z}^d as $d \rightarrow \infty$ for a large class of passage time distributions.

In Paper II we investigate a phenomenon of non–standard couplings of Markov chains, where two copies of a chain can be coupled to meet almost surely while their total variation distance stays bounded away from 0. We show that the supremum total variation distance that can be maintained in this context is $\frac{1}{2}$.

Paper III resolves affirmatively a recent conjecture by Lavrov and Loh that a uniformly chosen random edge–ordering of K_n contains a monotone Hamiltonian path with probability tending to 1 as $n \rightarrow \infty$. We further prove a partial result regarding the limiting behavior of the number of such paths, suggesting that this number, when appropriately rescaled, has log–normal distribution in the large n limit.

The topic of Paper IV is a model for a random $n \times n$ jigsaw puzzle, recently proposed by Mossel and Ross, where the shape of each edge of a piece is chosen uniformly out of q possibilities. The main question is whether this puzzle has a unique solution. We say that two solutions are similar if they only differ by permutations of duplicate pieces and rotations of pieces with rotational symmetries. We show that, with probability tending to 1 as $n \rightarrow \infty$, this puzzle has multiple non–similar solutions when $2 \leq q \leq \frac{2}{\sqrt{e}}n$, all solutions are similar when $q \geq (2 + \varepsilon)n$ for any $\varepsilon > 0$, and the solution is unique when $q = \omega(n)$.

Key words and phrases: First–passage percolation, high dimension, Cartesian power graph, non–Markovian coupling, coupling inequality, monotone paths, third moment argument, shotgun assembly, jigsaw puzzle.

Preface

The following four papers are appended to the thesis:

- I **Anders Martinsson**,
“First-passage percolation on Cartesian power graphs”,
submitted to *Annals of Probability* with current status “Minor revisions”.
- II Timo Hirscher and **Anders Martinsson**,
“Segregating Markov Chains”,
to appear in *Journal of Theoretical Probability*.
- III **Anders Martinsson**,
“Most edge-orderings of K_n have maximal altitude”,
submitted to *Random Structures & Algorithms*.
- IV **Anders Martinsson**,
“A linear threshold for uniqueness of solutions to random jigsaw puzzles”,
submitted to *Combinatorics, Probability & Computing*.

The three papers below appeared in my licentiate thesis, but are not appended here:

- A Peter Hegarty and **Anders Martinsson**,
“On the existence of accessible paths in various models of fitness landscapes”,
Annals of Applied Probability **24** (2014), no. 4, 1375–1395.
- B **Anders Martinsson**,
“Unoriented first-passage percolation on the n -cube”,
Annals of Applied Probability **26** (2016), no. 5, 2597–2625.
- C **Anders Martinsson**,
“Accessibility percolation and first-passage site percolation on the

unoriented binary hypercube”,
unpublished, available at arXiv:1501.02206

These three papers were left out of this thesis as they consider topics other than “discrete random structures”:

D Peter Hegarty and **Anders Martinsson**,
“Permutations Destroying Arithmetic Progressions in Finite Cyclic Groups”,
Electronic Journal of Combinatorics **22** (2015), no. 4, Paper #P4.39,
14pp.

E **Anders Martinsson**,
“An improved energy argument for the Hegselmann–Krause model”,
Journal of Difference Equations and Applications **22** (2016), no. 4,
630–635.

F Peter Hegarty, **Anders Martinsson**, and Edvin Wedin,
“The Hegselmann-Krause dynamics on the circle converge”,
Journal of Difference Equations and Applications **22** (2016), no. 11,
1720–1731.

Acknowledgements

First and foremost, I wish to thank my supervisor, Peter Hegarty. Throughout my five years here, you have been a great mentor, teaching me much of what I know about the academic world, as well as a constant source of thoughtful and insightful comments on the latest math problem. I doubt many other supervisors would have the dedication to help proofreading a paper on Christmas Eve. I also want to thank my co-supervisor, Johan Wästlund, for many thought-provoking and valuable discussions over the years.

My co-authors also deserve a word of thanks. To Timo Hirscher and Edvin Wedin. I am grateful for the collaborations I have had with both of you. Timo, it is remarkable how curiosity about a homework problem would turn into Paper II. Edvin, I thank you for introducing me to the Heggemann–Krause model, and I am very much looking forward to seeing progress on the freezing time.

It is probably not a surprise to anyone that completing a PhD requires much work, and can at times be very stressful because of that. In my experience, an even greater source of stress as well as frustration has been periods of e.g. making literally no progress for x months and feeling your valuable time slipping away. In this regard, the support from my fellow graduate students has been invaluable. In particular, thank you Magnus! I honestly do not know what this thesis would have been without all the times you came knocking at my office door.

I would further like to extend my gratitude to my former and current colleagues at the department. Matematiska Vetenskaper has really been a unique opportunity to interact with many awesome people, and I will fondly remember many discussions on topics ranging from heavy mathematical questions, to the behavior of an n -dimensional camera stand, to the meaning of newly made up words. To Dmitrii, Edvin, Hanna, John, Linnea, Magnus, Malin, Urban, Viktor, Åse, and many more. It has been a great pleasure to share this time here with you.

Finally, to my family and friends outside of the department. I fear the excuse “I have to work” has come out of my mouth far too often in recent years. Many thanks for your support and understanding during this time.

Anders Martinsson
Göteborg, March 26, 2017

Contents

I	Introduction	1
1	Introduction	3
2	First-passage percolation and power graphs	4
2.1	Basic definitions	5
2.2	Time constant and the shape theorem	5
2.3	Asymptotics in high dimension	8
2.4	The d -dimensional hypercube	9
2.5	Summary of Paper I	11
3	Markov chains, segregation and separation	15
3.1	The coupling argument for Markov chains	15
3.2	Meeting without mixing	18
3.3	Summary of Paper II	19
4	Altitudes of random edge-orders of the complete graph	23
4.1	The altitude of the complete graph	24
4.2	Related results	26
4.3	Random edge-orderings	27
4.4	Summary of Paper III	29
5	Jigsaw puzzles and shotgun assembly	32
5.1	Edge-matching and jigsaw puzzles	32
5.2	Summary of Paper IV	35
II	Appended papers	39
I	First-passage percolation on Cartesian power graphs	43
II	Segregating Markov chains	83
III	Most edge-orders of K_n have maximal altitude	113
IV	A linear threshold for uniqueness of solutions to random jigsaw puzzles	139

Part I

Introduction

Chapter 1

Introduction

This thesis consists of four papers on various topics in discrete probability. The common theme of the articles is to take a discrete structure, do something random to it, and then consider what happens as the size of the structure tends to infinity. In Papers I, III, and IV the main results concern questions of this form. As for Paper II, a central part of the main result, the existence of Markov chains with certain properties, is given as an asymptotic result in this sense.

Each of the following four chapters introduces the topic of one of the papers of the thesis and concludes with a summary of the corresponding paper. The chapters may be read independently of each other.

In Chapter 2 we consider the behavior of random distances in high-dimensional graphs. Random distances in graphs have a long history of being studied in mathematics under the name of first-passage percolation. The high-dimensional limit was proposed as a way to obtain quantitative results for the model. In Paper I, we consider this for a generalized class of high-dimensional graphs.

Chapter 3 concerns couplings of Markov chains, and, in particular, a counter-intuitive behavior of certain couplings where copies of a Markov chain can meet without mixing. In Paper II, we investigate this phenomenon in more detail. In Chapter 4 we describe a long-standing open problem regarding edge-orderings of graphs, and a more recently proposed randomized version which is the topic of Paper III. Lastly, Chapter 5 considers a natural model for random jigsaw puzzles. The central question is whether large jigsaw puzzles obtained in this way have a unique solution, or multiple solutions. Paper IV significantly improves earlier partial results to this end.

As an aside, for a reader who wishes to get an impression of what I have done for the last five years without bothering with too many technical details, this last chapter is probably the least technical one by far.

Chapter 2

First-passage percolation and power graphs

In Paper I, we consider a problem in first passage percolation in high dimension. First-passage percolation is a stochastic growth model on a graph, introduced by Hammersley and Welsh in 1965 [31]. A closely related growth model was introduced by Eden [26] four years earlier.

First-passage percolation is often described as a model of a porous rock submerged in water. The rock is modeled as a graph (most often \mathbb{Z}^d with nearest neighbor graph structure), where the vertices are the pores of the rock, and the edges are channels between pairs of pores through which water can flow. For each edge, we assign a random weight representing how much time it takes for water to pass through the channel. The central question is then how fast water will spread throughout the rock.

During the over 50 years since its introduction in the literature, this model has attracted much attention from mathematicians and physicists alike, and has given birth to some classical tools in mathematics, the main example being sub-additive ergodic theory. In spite of this, many natural questions about first-passage percolation remain open today.

In light of these difficulties, a number of variations of the model have been proposed. A noteworthy example is so-called last-passage percolation, for which some very interesting quantitative results are known. Another approach has been to consider underlying graphs other than lattices – for instance trees and strips. Here, we will focus on the limiting behavior of first-passage percolation in high dimension.

The aim of this chapter is to give sufficient background in first-passage percolation to put the results of Paper I into context. Considering the vast amount of research on the topic, I will not attempt to give an overview of the field as a whole. The reader is instead referred to the surveys by Howard [32] and more recently Auffinger, Damron and Hanson [6].

2.1 Basic definitions

Given a graph G and a non-negative random distribution F , first-passage percolation on G is defined as follows. For each edge $e \in E(G)$ we independently assign a weight τ_e with distribution F , called its passage time. The idea of first-passage percolation is to consider the shortest distances in G with respect to these weights. Formally, for a (self-avoiding) path γ in G we define its passage time by

$$T_G(\gamma) = \sum_{e \in \gamma} \tau_e, \quad (2.1)$$

and for any two vertices $u, v \in V(G)$, we define the first-passage time from u to v by

$$T_G(u, v) = \inf_{\gamma \text{ from } u \text{ to } v} T_G(\gamma). \quad (2.2)$$

As any paths from u to v and from v to w can be combined into a path from u to w (one might have to remove some edges in order for it to be self-avoiding), $T_G(\cdot, \cdot)$ satisfies the triangle inequality

$$T_G(u, w) \leq T_G(u, v) + T_G(v, w), \quad (2.3)$$

for all $u, v, w \in V(G)$. This is referred to as the sub-additive property of first-passage percolation.

Commonly, first-passage percolation is viewed as a growth model by defining the wet region $\{B_t\}_{t \geq 0}$ where

$$B_t = \{v \in V(G) : T_G(v_0, v) \leq t\}, \quad (2.4)$$

and v_0 is some fixed vertex. We can interpret this as that, at time $t = 0$, v_0 is connected to a water source, and the water then spreads along G according to the passage times. For integer lattices, we always assume that v_0 is the origin.

One interesting special case for the wet region is when $F = \text{Exp}(1)$. In this case, the memory-less property of the exponential distribution implies that $\{B_t\}_{t \geq 0}$ is a Markov process. This is usually described in term of the spread of an infection though G , where B_t is the set of infected vertices at time t . Initially, at $t = 0$, v_0 is the lone bearer of an infection. A healthy vertex becomes infected at rate equal to its number of infected neighbors, and once a vertex gets infected it stays infected permanently. This Markov process is known as the Richardson growth model [45]. If one considers the sequence of discrete updates $B_{t_0=0}, B_{t_1}, B_{t_2}, \dots$, then the model becomes equivalent to a version of Eden's growth model.

2.2 Time constant and the shape theorem

While first-passage percolation can be defined with any underlying graph, the classical setting is to consider \mathbb{Z}^d with nearest neighbor graph structure

for some fixed $d \geq 2$. Some immediate questions in this setting are: “How does the quantity $T_{\mathbb{Z}^d}(x, y)$ behave in the limit as $x, y \in \mathbb{Z}^d$ are far apart? How does the wet region typically look for large t ? Do we have some scaling limits of these?”

The first of these questions was partially answered already by Hammersley and Welsh. Let $e_1 = (1, 0, \dots, 0) \in \mathbb{Z}^d$. Under the assumption that $\mathbb{E}\tau < \infty$, Hammersley and Welsh observed that the limit

$$\mu = \lim_{n \rightarrow \infty} \frac{\mathbb{E}T_{\mathbb{Z}^d}(0, n e_1)}{n} \quad (2.5)$$

exists. Using the sub-additivity of $T_{\mathbb{Z}^d}(\cdot, \cdot)$ they managed to prove that

$$\frac{T_{\mathbb{Z}^d}(0, n e_1)}{n} \rightarrow \mu, \text{ in probability as } n \rightarrow \infty. \quad (2.6)$$

The constant $\mu = \mu(F, d)$ is commonly referred to as the time-constant.

Assuming $\mu > 0$, this provides a simple description of $T_{\mathbb{Z}^d}(0, n e_1)$ – it grows linearly in the distance between the specific points. In other words, the wet region will spread along the coordinate axes at an asymptotic speed of $1/\mu$. It was shown by Kesten ([34]; Thm. 6.1) that $\mu > 0$ if and only if $\mathbb{P}(\tau_e = 0) < p_c(d)$ where $p_c(d)$ denotes the critical probability in Bernoulli percolation on \mathbb{Z}^d . In particular, $\mu > 0$ if $\mathbb{P}(\tau_e = 0) = 0$.

An important advancement in the study of such time constants came with Kingman’s sub-additive ergodic theorem [35, 38]. Let us assume that

$$\mathbb{E} \min\{\tau_1, \tau_2, \dots, \tau_{2d}\} < \infty, \quad (2.7)$$

where τ_i , $1 \leq i \leq 2d$, denote independent variables of distribution F . We note that this is a natural condition as $\mathbb{E}T_{\mathbb{Z}^d}(u, v) < \infty$ for all $u, v \in \mathbb{Z}^d$ if and only if this holds (see e.g. the proof of Theorem 2.1 in [6]). Under this assumption, the sub-additive ergodic theorem implies that the limit in (2.6) holds almost surely and in L^1 . Moreover, this result holds in an arbitrary direction. For any $x \in \mathbb{R}^d$, there is a constant $\mu(x)$ such that

$$\frac{T_{\mathbb{Z}^d}(0, \lfloor n x \rfloor)}{n} \rightarrow \mu(x) \text{ a.s. and in } L^1 \text{ as } n \rightarrow \infty, \quad (2.8)$$

where $\lfloor n x \rfloor = (\lfloor n x_1 \rfloor, \lfloor n x_2 \rfloor, \dots)$. Again assuming $\mathbb{P}(\tau_e = 0) < p_c(d)$, the function $\mu(x)$ defines a norm on \mathbb{R}^d , and inherits the obvious symmetries of \mathbb{Z}^d .

Probably the most iconic result of first-passage percolation is the so-called shape theorem. This is usually attributed to Cox and Durrett [19], see also Richardson [45] and Kesten [34]. From (2.8), it follows that the wet region grows linearly in t in all directions. It is therefore natural to ask whether the rescaled wet region B_t/t has a limit in some sense as $t \rightarrow \infty$. Indeed, we can heuristically think of the above result as that $T_{\mathbb{Z}^d}(0, x)$ behaves roughly as $\mu(x)$ for x far from the origin. Hence the

rescaled version of the wet region should approximately be the unit ball with respect to the norm $\mu(x)$.

To state this formally, we introduce a smoothed version of B_t where each vertex $x \in B_t$ is replaced by the unit square (or more generally d -cube) $x + [-\frac{1}{2}, \frac{1}{2}]^d$. We let

$$\bar{B}_t = \{x + y : x \in B_t \text{ and } y \in [-\frac{1}{2}, \frac{1}{2}]^d\}. \quad (2.9)$$

Theorem 2.1. (*Cox-Durrett*) *Assume*

$$\mathbb{E} \min\{\tau_1^d, \tau_2^d, \dots, \tau_{2d}^d\} < \infty, \quad (2.10)$$

where τ_i , $1 \leq i \leq 2d$, denote independent random variables of distribution F . If $\mu > 0$, there exists a convex and compact set $\mathcal{B} \subseteq \mathbb{R}^d$, defined by

$$\mathcal{B} = \{x \in \mathbb{R}^d : \mu(x) \leq 1\} \quad (2.11)$$

with non-empty interior such that, for any $\varepsilon > 0$,

$$\mathbb{P} \left((1 - \varepsilon)\mathcal{B} \subseteq \frac{\bar{B}_t}{t} \subseteq (1 + \varepsilon)\mathcal{B} \text{ for } t \text{ sufficiently large} \right) = 1. \quad (2.12)$$

On the other hand, if $\mu = 0$, then for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$

$$\mathbb{P} \left(\mathcal{K} \subseteq \frac{\bar{B}_t}{t} \text{ for } t \text{ sufficiently large} \right) = 1. \quad (2.13)$$

Given these results, it is natural to ask, for a given non-trivial distribution F , what can be said of the time constant $\mu(e_1)$ and limit shape \mathcal{B} . For more or less trivial reasons, \mathcal{B} is always convex and symmetric with respect to permutation and changing sign of coordinates, but determining properties of these objects beyond this is considered a hard and mostly open problem.

It is believed that for continuous distributions, \mathcal{B} should be strictly convex with a differentiable boundary, but this is completely open. Early simulations for the exponential distribution [26, 45] indicated that the limit shape might be a Euclidean ball in this case. Kesten [34] showed that this is not true for a large class of distributions, including exponential distributions, if the dimension of the lattice is sufficiently high. A recent and much larger-scale simulation study by Alm and Deijfen [3] on \mathbb{Z}^2 with various distributions shows definitively that the limit shape for the exponential distribution is very close to but not exactly a Euclidean ball. Indeed the radii of \mathcal{B} along the coordinate axes and along the diagonals differ only by approximately 1%.

2.3 Asymptotics in high dimension

One drawback of the sub-additive ergodic theorem is that while it shows the existence of certain limits, it gives essentially no information about what these limits should be, and hence quantitative results in first-passage percolation are rare. One avenue to obtain quantitative results is to study the limiting behavior on \mathbb{Z}^d as $d \rightarrow \infty$. Heuristically, high dimension puts less emphasis on the geometry of the graph, and thus turns first-passage percolation into a largely combinatorial problem, which can be approached through combinatorial means.

The first high-dimensional results for the integer lattice were given by Kesten (Chapter 8 of [34]). Here we assume that the edge weights have continuous distribution with density $f(t)$, where

$$f(t) = a + o\left(\frac{1}{|\log t|}\right), \quad (2.14)$$

for some interval $t \in [0, \varepsilon)$, and that $\mathbb{E}\tau_e < \infty$. For instance, the $\text{Exp}(1)$ and $U([0, 1])$ distributions satisfy this with $a = 1$. Under these assumptions, Kesten shows that there is a constant $c > 0$ such that, for d sufficiently large, the time constant $\mu = \mu(e_1)$ satisfies

$$c \frac{\log d}{ad} \leq \mu(e_1) \leq 11 \frac{\log d}{ad}. \quad (2.15)$$

Moreover, if we let $\bar{1} = (1, 1, \dots, 1) \in \mathbb{Z}^d$, then

$$\frac{1}{6ea} \leq \liminf_{d \rightarrow \infty} \mu(\bar{1}) \leq \limsup_{d \rightarrow \infty} \mu(\bar{1}) \leq \frac{1}{2a}. \quad (2.16)$$

Relating $\mu(e_1)$ and $\mu(\bar{1})$ to the limit shape, we can observe that if \mathcal{B} is a Euclidean ball, i.e. $\mu(x)$ is proportional to the Euclidean norm, then we would get $\mu(e_1) = \mu(\bar{1})/\sqrt{d}$. Hence, from (2.15) and (2.16), Kesten could conclude that, for any non-negative distribution with finite expectation satisfying (2.14), there is a d_0 such that the limit shape of that distribution is not a Euclidean ball for $d \geq d_0$. Kesten remarks that the convergence of his estimates is rather slow. For instance, for exponential distributions you need $d \geq 1000000$.

In the case of the standard exponential distribution, more precise estimates were obtained by Dhar in [22, 23], where it was shown that, as $d \rightarrow \infty$,

$$\mu(e_1) \sim \frac{\ln d}{2d}. \quad (2.17)$$

Dhar further derived a lower bound for the diagonal time constant for this distribution. This was recently rediscovered by Couronné, Enriquez and Gerin [18]. We will use their formulation here. Let α_* denote the unique

positive solution to $\coth \alpha = \alpha$. Then for any $d \geq 1$,

$$\mu(\bar{1}) \geq \frac{1}{2} \sqrt{\alpha_*^2 - 1} \approx 0.3133 \dots \quad (2.18)$$

In 2016, Auffinger and Tang [7] generalized this result beyond the exponential distribution. We assume that $a = \lim_{t \downarrow 0} F(t)/t \in (0, \infty)$ exists, and again $\mathbb{E}\tau_e < \infty$. Under these assumptions, we have

$$\mu(e_1) \sim \frac{\ln d}{2ad}, \quad (2.19)$$

and

$$\liminf_{d \rightarrow \infty} \mu(\bar{1}) \geq \frac{1}{2a} \sqrt{\alpha_*^2 - 1}. \quad (2.20)$$

We remark that Auffinger and Tang state these estimates with the technical condition that $F(x) = a + O(1/|\log x|)$ for x near zero, but this can be overcome by stochastically sandwiching F between distributions F_1 and F_2 with constant density $a + \varepsilon$ and $a - \varepsilon$ respectively near zero.

2.4 The d -dimensional hypercube

During the development of first-passage percolation on high-dimensional integer lattices, an alternative high-dimensional graph was proposed by Aldous [1], and Fill and Pemantle [27] – the d -dimensional hypercube Q_d . This is the graph whose vertices are the binary strings of length d , and where two vertices are connected by an edge if their binary strings differ by precisely one bit. We will here let $\bar{0}$ and $\bar{1}$ denote the all zeroes and all ones vertices respectively. In this section, we always assume $\text{Exp}(1)$ passage times.

The d -dimensional hypercube is in many ways similar to \mathbb{Z}^d . They are both in a sense d -dimensional structures, with similar geometries. We have the obvious embeddings of Q_d into \mathbb{Z}^d . They also have similar group structures – \mathbb{Z}^d is a Cayley graph under coordinate-wise addition, whereas Q_d is a Cayley graph under coordinate-wise addition modulo 2. On the other hand, the hypercube is combinatorially a simpler structure. Fill and Pemantle thus argued that understanding first-passage percolation on Q_d may lead to a better understanding of the high-dimensional integer lattice.

As Q_d is a finite graph it does not make sense to ask for the limiting behavior of $T_{Q_d}(x, y)$ for x and y far apart, but we can instead consider the limiting behavior as $d \rightarrow \infty$ directly. A natural question in this setting is: “How does the first-passage time between two points depend on their Hamming distance? In particular, how does the first-passage time between two opposite corners of the cube behave, say for instance $\bar{0}$ and $\bar{1}$?”

In considering first-passage percolation on the hypercube, a simplification that appears in the literature is to introduce an orientation of its

edges. We define the oriented d -dimensional hypercube $Q_d^{\mathcal{O}}$ as the oriented version of Q_d where each edge is directed towards the vertex with the extra '1'.

As it turns out, for both $Q_d^{\mathcal{O}}$ and Q_d , there are simple explicit lower bounds on the first-passage times. Let us restrict attention to the first-passage times between $\bar{0}$ and $\bar{1}$. In this case, we have

$$\mathbb{P}(T_{Q_d^{\mathcal{O}}}(\bar{0}, \bar{1}) \leq t) \leq t^d, \quad (2.21)$$

and

$$\mathbb{P}(T_{Q_d}(\bar{0}, \bar{1}) \leq t) \leq (\sinh t)^d, \quad (2.22)$$

for all $t \geq 0$. As a consequence of this, the first-passage time from $\bar{0}$ to $\bar{1}$ is, with probability tending to one as $d \rightarrow \infty$, at least $1 - \varepsilon$ in the oriented case, and at least $\sinh^{-1}(1) - \varepsilon = \ln(1 + \sqrt{2}) - \varepsilon$ in the unoriented case, for any $\varepsilon > 0$.

The lower bound for $Q_d^{\mathcal{O}}$ was first observed by Aldous in 1989 (see Example G7 of [1]). It was derived by considering the expected number of paths from $\bar{0}$ to $\bar{1}$ with passage time at most t . Aldous conjectured that this bound tells the truth in the sense that

$$T_{Q_d^{\mathcal{O}}}(\bar{0}, \bar{1}) \xrightarrow{P} 1 \text{ as } d \rightarrow \infty. \quad (2.23)$$

This was proven correct four years later by Fill and Pemantle using essentially a complementary second moment analysis, but with a twist to reduce the variance.

The bound in the unoriented case, (2.22), was first shown by Fill and Pemantle using an argument due to Durrett. Despite its similarity to (2.21), it is derived using a very different approach. The central idea is to consider what they call the branching translation process on Q_d : Initially the process consists of one particle placed at $\bar{0}$. Each existing particle generates offspring at rate d . Whenever some new offspring is born, it is placed at a uniformly chosen neighbor to its parent's location. Comparing this process to Richardson's model, one can show that the probability that $x \in B_t$, i.e. $T_{Q_d}(\bar{0}, x) \leq t$, is at most the expected number of particles at x at time t , which can be computed exactly.

Due to the result for the oriented hypercube, Fill and Pemantle conclude that

$$\ln(1 + \sqrt{2}) - o(1) \leq T_{Q_d}(\bar{0}, \bar{1}) \leq 1 + o(1) \quad (2.24)$$

with probability tending to 1 as $d \rightarrow \infty$. The problem of determining $T_{Q_d}(\bar{0}, \bar{1})$ further remained open until a recent paper by me [39] not included in this thesis, where it was shown that

$$T_{Q_d}(\bar{0}, \bar{1}) \rightarrow \ln(1 + \sqrt{2}) \quad (2.25)$$

in probability and L^p for any $p \geq 1$. In particular, I show that $T_{Q_d}(\bar{0}, \bar{1})$ has mean $\ln(1 + \sqrt{2}) + O(\frac{1}{d})$ and standard deviation $\Theta(\frac{1}{d})$ and that the minimizing path from $\bar{0}$ to $\bar{1}$ can be characterized by a certain random walk between these points.

Another quantity related to first-passage percolation on Q_d proposed by Fill and Pemantle is the so-called covering time. This is the smallest time at which the wet region covers the entire graph, or, equivalently, the maximum value of $T_{Q_d}(\bar{0}, x)$ over all $x \in V(Q_d)$. Fill and Pemantle showed that, with probability tending to 1 as $d \rightarrow \infty$, the covering time is at least $\frac{1}{2} \ln(2 + \sqrt{5}) + \ln 2 - o(1) \approx 1.415$. The currently best known upper bound on the covering time is $\ln(1 + \sqrt{2}) + \ln 2 + o(1) \approx 1.574$, which follows by combining results by Bollobás and Kohayakawa [10] and me [39].

2.5 Summary of Paper I

The aim of Paper I is to investigate first-passage percolation on a large class of “high-dimensional” graphs, including the hypercube and the integer lattice, to which many ideas from the hypercube extend. This analysis allows us to derive many new quantitative results. As the approach is quite general, I believe it may lead to a more unified approach to first-passage percolation in high dimension.

For any pair of graphs $H_1 = (V_1, E_1)$ and $H_2 = (V_2, E_2)$, their Cartesian graph product, denoted by $H_1 \square H_2$, is a graph whose vertices are the pairs (v_1, v_2) of vertices $v_1 \in V_1$ and $v_2 \in V_2$. We define the edges of the product graph such that taking a step in this graph corresponds to choosing one coordinate, and taking one step in the corresponding factor. Formally, the edge set of the product is the disjoint union $(E_1 \times V_2) \cup (V_1 \times E_2)$, where an edge of the form (e, v) or (v, e) is interpreted as an edge between (w_1, v) and (w_2, v) or between (v, w_1) and (v, w_2) respectively, where w_1 and w_2 denote the end-points of e . Given any base graph G and any integer $d \geq 1$ we define the d :th Cartesian power graph G^d as the d -fold Cartesian graph product $G \square \dots \square G$. For any vertex $v \in G$, we let $\bar{v} = (v, v, \dots, v) \in G^d$.

It is natural to think of G^d as a d -dimensional graph. Indeed, its vertices can be represented by d -dimensional vectors with coordinates in G . Moreover, this generalizes the d -dimensional graphs mentioned in earlier sections:

- \mathbb{Z}^d is the d :th Cartesian power of \mathbb{Z} with edges $\{i, i + 1\}$ for $i \in \mathbb{Z}$.
- Q_d^O is the d :th Cartesian power of the graph consisting of vertices 0 and 1 with a directed graph from 0 to 1.
- Q_d is the d :th Cartesian power of K_2 . To match earlier notation, we denote the vertices of K_2 by 0 and 1.

It turns out that, for any graph of the form G^d , there is a natural lower bound on the first-passage time between two vertices, which generalizes the lower bounds (2.21) and (2.22) for the hypercube. For any graph H and any pair of vertices $v, w \in H$, let $\Gamma_H(v, w)$ denote the set of trails from v to w in H . In the case where $v = w$, this should include the trivial trail of length 0. We define the exponential generating function

$$m_H(v, w, t) = \sum_{\gamma \in \Gamma(v, w)} \frac{t^{|\gamma|}}{|\gamma|!} \quad (2.26)$$

where $|\gamma|$ denote the length of γ , that is, the number of edges counted with multiplicity. For $t = 0$, we interpret any term of the form $\frac{0^0}{0!}$ as 1. It is not too hard (see Proposition 1.1 in Paper I and the subsequent discussion) to show that, for any such graph, first-passage percolation with $\text{Exp}(1)$ passage times has the lower bound

$$\mathbb{P}(T_H(v, w) < t) \leq m_H(v, w, t), \quad (2.27)$$

for all $t \geq 0$.

A valuable observation (Proposition 1.2 in Paper I) is that the exponential generating function is multiplicative with respect to Cartesian products in the sense that

$$m_{H_1 \square H_2}((v_1, v_2), (w_1, w_2), t) = m_{H_1}(v_1, w_1, t) m_{H_2}(v_2, w_2, t). \quad (2.28)$$

In the case of $H = G^d$, we can rewrite (2.27) as

$$\mathbb{P}(T_{G^d}((v_1, \dots, v_d), (w_1, \dots, w_d)) \leq t) \leq \prod_{i=1}^d m_G(v_i, w_i, t). \quad (2.29)$$

In particular, if we focus on the first-passage time between diagonal vertices \bar{v} and \bar{w} , we obtain the simple bound

$$\mathbb{P}(T_{G^d}(\bar{v}, \bar{w}) < t) \leq (m_G(v, w, t))^d. \quad (2.30)$$

As a consequence of this, the critical value of t , $t^* = t_G(v, w)$, given by the solution to $m_G(v, w, t) = 1$ is an asymptotic lower bound on $T_{G^d}(\bar{v}, \bar{w})$ in the sense that

$$\mathbb{P}(T_{G^d}(\bar{v}, \bar{w}) < t^* - \varepsilon) \rightarrow 0 \text{ as } d \rightarrow \infty, \quad (2.31)$$

for any $\varepsilon > 0$.

Given this lower bound on high-dimensional first-passage percolation, it is natural to ask when the bound is sharp. That is, for which choices of G , v , and w is it true that $T_{G^d}(\bar{v}, \bar{w}) \rightarrow t_G^*(v, w)$ in some sense as $d \rightarrow \infty$? The main result of Paper I is a necessary and sufficient condition for this to hold.

Let G, v and w be given. Assume there is a path from v to w in G . For any $s, t \geq 0$ such that $s + t \leq t^*$, we define

$$f_G^{vw}(s, t) = \sum_{x, y \in V(G)} m_G(v, x, s) m_G(x, y, t) \cdot \ln(m_G(x, y, t)) m_G(y, w, t^* - s - t). \quad (2.32)$$

As it turns out, sharpness of t^* is completely determined by the maximum value of this function.

Theorem 2.2. *Assume G has bounded degree, and let F be a non-negative distribution such that $a = \lim_{t \downarrow 0} F(t)/t \in (0, \infty)$ exists. Consider first-passage percolation on G^d with passage times of distribution F . If $f_G^{vw} \leq 0$ for all $s, t \geq 0$ such that $s + t \leq t^*$, then*

$$T_{G^d}(\bar{v}, \bar{w}) \rightarrow \frac{t^*}{a} \text{ in probability as } d \rightarrow \infty. \quad (2.33)$$

Moreover, if F has finite expectation, then convergence also holds in L^1 . On the other hand, if $f_G^{vw} > 0$ for some such s, t , then there exists some constant $c = c(G, v, w) > 0$ such that

$$\mathbb{P}\left(T_{G^d}(\bar{v}, \bar{w}) > \frac{t^* + c}{a}\right) \rightarrow 1 \text{ as } d \rightarrow \infty. \quad (2.34)$$

Given an explicit graph G and vertices v and w , we can at least in principle compute f_G^{vw} to check its maximum. Unfortunately, these calculations can be quite complicated. To remedy this, the following result gives a simple combinatorial condition to ensure convergence as in (2.33).

Corollary 2.3. *Let F and a be as above, and let G be a connected graph with bounded degree. Each of the following conditions are sufficient for $T_{G^d}(\bar{v}, \bar{w}) \rightarrow t^*/a$ in probability (and L^1 assuming F has finite expectation) as $d \rightarrow \infty$ for all $v, w \in G$:*

1. G is a Cayley graph of a group \mathcal{G} generated by a finite normal set S , that is $|S|$ and $gSg^{-1} = S \forall g \in \mathcal{G}$. In particular, this always holds for Cayley graphs of finitely generated abelian groups.
2. For all pairs of vertices $x, y \in G$ there is a graph automorphism $\phi = \phi_{xy}$ such that $\phi(x) = y$ and $\phi(y) = x$.

We further give two applications of this result. First, we consider the diagonal time constant $\mu(\bar{1})$ in \mathbb{Z}^d in the limit as $d \rightarrow \infty$. In particular, concerning the first-passage time from $\bar{0}$ to \bar{k} for some integer $k \geq 1$, we have

$$m_{\mathbb{Z}}(0, k, t) = \sum_{i=0}^{\infty} \frac{t^{k+2i}}{i!(k+i)!}. \quad (2.35)$$

Furthermore, one can show that $t_{\mathbb{Z}}^*(0, k)/k \rightarrow \frac{1}{2}\sqrt{\alpha_*^2 - 1}$ when $k \rightarrow \infty$, where α is defined as before in (2.18). As a consequence of our main result, we show that

$$\mu_d(\bar{1}) \rightarrow \frac{1}{2a}\sqrt{\alpha_*^2 - 1} \text{ as } d \rightarrow \infty, \quad (2.36)$$

for any distribution F as above with finite expectation. Hence, the lower bounds (2.18) and (2.20) observed in [7, 18, 22] are asymptotically sharp in high dimension.

As a second application we consider the first-passage time between two general points in Q_d . Let F be as above, and let $S(d, k)$ denote the first-passage time between two points at Hamming distance k in the d -dimensional hypercube. By symmetry of the cube, it does not matter which pair of such points one chooses. For any $x \in [0, 1]$, let $\vartheta(x)$ be the non-negative solution to

$$(\sinh \vartheta)^x (\cosh \vartheta)^{1-x} = 1. \quad (2.37)$$

We show that this function characterizes the limiting behavior of $S(d, k)$ in the sense that

$$\left| S(d, k) - \frac{1}{a}\vartheta\left(\frac{k}{d}\right) \right| \rightarrow 0 \quad (2.38)$$

in probability (and L^1 -norm assuming finite expectation) uniformly over $0 \leq k \leq d$ as $d \rightarrow \infty$.

Chapter 3

Markov chains, segregation and separation

Coupling is one of the main tools in comparing random distributions. The idea is to define a pair of random variables (X, Y) on a joint probability space such that they, viewed individually, have certain distributions of interest, and jointly depend on each other in some way that allows one to relate their distributions.

In particular, a common approach to investigate the long-term behavior of Markov chains is to consider a coupling of two copies of the chain with different initial states. Formally, this is a joint process $\{(X_n, Y_n)\}_{n=0}^{\infty}$, where $\{X_n\}_{n=0}^{\infty}$ and $\{Y_n\}_{n=0}^{\infty}$ individually evolve according to the transition probabilities of the Markov chain. The idea is to make the copies depend on each other in some clever way such that they meet as fast as possible and then stick together. This relates to the so-called mixing time of the chain. Probably the most well-known example of mixing times of Markov chains is the mathematical investigation of how long it takes to shuffle a deck of cards, but questions of this type span a wide range of areas of probability theory.

The topic of Paper II is a phenomenon of certain non-standard couplings of Markov chains, first observed by Häggström [30], in which two copies of a Markov chain can be coupled to meet a.s. without mixing.

3.1 The coupling argument for Markov chains

Given two random distributions μ and ν on a finite or countable set S , interpreted as probability measures, we define their total variation distance

by

$$\|\mu - \nu\|_{TV} = \sup_{A \subseteq S} |\mu(A) - \nu(A)|. \quad (3.1)$$

In estimating this distance, it is often useful to interpret it in terms of couplings of the distributions. We have that

$$\|\mu - \nu\|_{TV} = \inf\{\mathbb{P}(X \neq Y) : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}, \quad (3.2)$$

see for instance Proposition 4.7 in [37]. In particular, for any such coupling, $\mathbb{P}(X \neq Y)$ gives an upper bound on the total variation distance between their distributions.

Let $\{X_n\}_{n=0}^\infty$ denote a Markov chain on a finite state space S with transition probabilities

$$\mathbb{P}(X_{n+1} = y | X_n = x) = P(x, y) \quad (3.3)$$

for all $x, y \in S$. We let $P^n(x, y)$ denote the probability of $X_n = y$, and $P^n(x, \cdot)$ the distribution of X_n respectively, conditioned on $X_0 = x$. It is a well-known fact that, under some mild assumptions, a Markov chain “mixes” in the sense that it converges to a distribution as $n \rightarrow \infty$, not depending on its starting position. We say that the Markov chain is irreducible if for all $x, y \in S$ there exists an $n \geq 0$ such that $P^n(x, y) > 0$. Moreover, the chain is aperiodic if for any $x \in S$, $P^n(x, x) > 0$ for all sufficiently large n .

With these definitions, the standard convergence result for finite Markov chains can be stated as follows, see for instance Theorem 4.9 in [37].

Theorem 3.1. *Suppose $\{X_n\}_{n=0}^\infty$ is an irreducible and aperiodic Markov chain on a finite state space S . Then there exists a unique distribution π on S , called the stationary distribution of the chain, and constants $\alpha \in (0, 1)$ and $C > 0$ such that*

$$\|P^n(x, \cdot) - \pi\|_{TV} \leq C\alpha^n, \quad (3.4)$$

for all $x \in S$ and $n \geq 0$.

We remark that convergence in this sense does not hold for all Markov chains. In particular, the chains of interest in Paper II do not mix in this way. Nevertheless, mixing Markov chains are central to introduce some important concepts to the paper.

Given a Markov chain that converges to a unique stationary distribution π , a central question is how fast this mixing occurs. For irreducible aperiodic chains, Theorem 3.1 technically gives us a lower bound on this rate, but this tends to be extremely far from its true value. In order to quantify this convergence, one can define

$$d(n) = \sup_{x \in S} \|P^n(x, \cdot) - \pi\|_{TV}. \quad (3.5)$$

Sometimes, in particular when dealing with coupling arguments, it is more convenient to use the related function

$$\bar{d}(n) = \sup_{x, y \in S} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV}. \quad (3.6)$$

One can show that $d(n) \leq \bar{d}(n) \leq 2d(n)$, so these are indeed closely related. Moreover, both functions are non-increasing, and $2d$ and \bar{d} are sub-multiplicative in the sense that $2d(m+n) \leq 2d(m) \cdot 2d(n)$ and $\bar{d}(m+n) \leq \bar{d}(m) \cdot \bar{d}(n)$. See Lemma 2.20 in [2] for a proof of this.

The standard coupling approach for estimating these quantities can be described as follows. Suppose that we have a coupling $\{(X_n, Y_n)\}_{n=0}^{\infty}$ of two copies of the Markov chain started at states x and y respectively, where the chains depend on each other in a clever way such that they eventually meet up and then stick together, that is, there exists a random time τ such that $X_n = Y_n$ for all $n \geq \tau$. Then by (3.2), we have

$$\|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} \leq \mathbb{P}(X_n \neq Y_n) \leq 1 - \mathbb{P}(\tau \leq n). \quad (3.7)$$

Hence, in order to obtain good upper bounds on $\bar{d}(n)$, it suffices to find good ways to couple the chains such that τ is as small as possible for all $x, y \in S$.

Often, how this type of argument is done in practice is that one defines some rule for how the chains should be coupled until they meet for the first time, one takes τ to be this first meeting time, and then modifies the coupling such that $X_n = Y_n$ for all $n \geq \tau$. We can think of this modification as “gluing” the chains together. If this can be done, then (3.7) holds with τ as the first meeting time. For most couplings of this type in the literature, the act of gluing is completely undramatic and trivially valid. Indeed, this issue has required so little attention that an earlier version of the standard reference [37] erroneously claimed that any coupling of two copies of a Markov chain can be glued together in this way.

The question of when couplings can be glued was addressed by Rosenthal [48]. His paper gives an example of a coupling of two copies of a Markov chain that meet but where gluing cannot be done. More specifically, attempting to glue the chains will change the marginal distributions of the coupling, hence it is no longer a coupling of the same processes. Rosenthal further proposes a natural sufficient condition on a coupling $\{(X_n, Y_n)\}_{n=0}^{\infty}$ in order for gluing to be possible. For all $n \geq 0$ and all $x_0, \dots, x_n, y_0, \dots, y_n, z \in S$, we require

$$\mathbb{P}(X_{n+1} = z | X_k = x_k, Y_k = y_k \forall 0 \leq k \leq n) = P(x_n, z) \quad (3.8)$$

and

$$\mathbb{P}(Y_{n+1} = z | X_k = x_k, Y_k = y_k \forall 0 \leq k \leq n) = P(y_n, z). \quad (3.9)$$

A coupling satisfying this is said to be faithful. This is sometimes also called Markovian in the literature, but this term may be misleading as it might be interpreted as that the joint process $\{(X_n, Y_n)\}_{n=0}^\infty$ should have the Markov property, which is a strictly weaker condition. For readers familiar with more formal treatment of stochastic processes, it can be noted that faithfulness is equivalent to saying that $\{X_n\}_{n=0}^\infty$ and $\{Y_n\}_{n=0}^\infty$ should have the Markov property with respect to a common filtration.

3.2 Meeting without mixing

The general philosophy of coupling arguments for Markov chains is that meeting implies mixing. In particular, we can note that, by (3.7), if there exists a faithful coupling $\{(X_n, Y_n)\}_{n=0}^\infty$ with the property that the chains will almost surely meet at some point, then, in the limit as $n \rightarrow \infty$, X_n has the same distribution as Y_n . Indeed, this makes sense as the ability to couple the chains to meet almost surely means that they in some sense have to live on the same set of states, and as they have the same transition probabilities on these states, one would expect them to have similar distributions.

In light of this intuition, it might be surprising that there are cases where two copies of a Markov chain, started in distinct states x and y , can be coupled to meet almost surely while their total variation distance stays bounded away from 0. In Paper II, we call this phenomenon segregation of x and y .

The first observation of a segregating Markov chain was made by Häggström in [30]. His example is illustrated in Figure 3.1. The Markov chain consists of six states, and a chain with initial state either x or y will end up in one of the absorbing states a or b after exactly two steps. A direct calculation shows that

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} = |P^2(x, a) - P^2(y, a)| = (1 - 2p)^2, \quad (3.10)$$

which is non-zero unless $p = \frac{1}{2}$.

Häggström notes that it is possible to explicitly construct a coupling of two copies of this chain, started in x and y respectively, such that the copies meet a.s. whenever $p \in [1 - \frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}]$. Let us consider $p = 1 - \frac{1}{2}\sqrt{2}$ so that (3.10) is maximized. In this case, it is straight-forward to check that

$$\mathbb{P} \left(\begin{array}{c} \diagdown \\ \downarrow \\ \diagup \end{array} \right) = \mathbb{P} \left(\begin{array}{c} \diagup \\ \downarrow \\ \diagdown \end{array} \right) = \frac{3 - 2\sqrt{2}}{2}, \quad (3.11)$$

$$\mathbb{P} \left(\begin{array}{c} \diagup \\ \downarrow \\ \diagdown \end{array} \right) = \mathbb{P} \left(\begin{array}{c} \downarrow \\ \diagdown \\ \diagup \end{array} \right) = \mathbb{P} \left(\begin{array}{c} \diagdown \\ \downarrow \\ \diagup \end{array} \right) = \mathbb{P} \left(\begin{array}{c} \downarrow \\ \diagup \\ \diagdown \end{array} \right) = \frac{\sqrt{2} - 1}{2}, \quad (3.12)$$

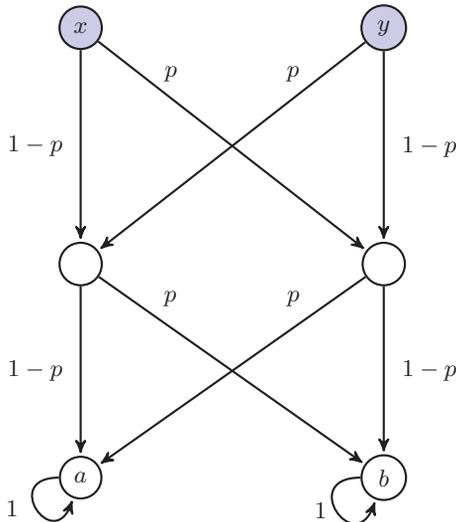


Figure 3.1: An example of segregating Markov chain proposed by Hägström.

defines a coupling of copies of the Markov chain started in x and y respectively in which the chains meet a.s. On the other hand, $\|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} \rightarrow (\sqrt{2} - 1)^2 = 3 - 2\sqrt{2} \approx 0.17153$ as $n \rightarrow \infty$.

As the example above illustrates, general couplings of copies of Markov chains can behave both quantitatively and qualitatively very different from faithful couplings, and perhaps also from what one would intuitively expect. This naturally leads us to ask two related questions. First, what is the relation between meeting probabilities and distributional distance for general couplings of two copies of a Markov chain? Second, how can we understand the phenomenon of segregation?

3.3 Summary of Paper II

Paper II, which is joint work with Timo Hirscher, studies the connection between meeting times and total variation distance for general couplings of Markov chains and specifically segregating Markov chains in further detail. We in particular ask for the constant κ given by

$$\kappa = \sup \lim_{n \rightarrow \infty} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV}, \quad (3.13)$$

where the supremum is taken over all finite Markov chains and states x and y , such that two copies of the chain, one started in x and the other

in y , can be coupled to meet a.s.

The first result of our paper is a non-trivial upper bound on the total variation distance between two copies of a Markov chain in terms of their meeting probability with respect to a not necessarily faithful coupling.

Proposition 3.2. *Let $\{(X_n, Y_n)\}_{n=0}^\infty$ be a coupling of two copies of a countable state Markov chain with initial states x and y , where $\tau = \inf\{n : X_n = Y_n\}$. Then*

$$\|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} \leq 1 - \frac{1}{2}\mathbb{P}(\tau \leq n). \quad (3.14)$$

We note that this bound is very similar to the usual coupling inequality (3.7). In particular, assuming $\tau < \infty$ a.s., it follows that

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} \leq \frac{1}{2}, \quad (3.15)$$

which implies the upper bound $\kappa \leq \frac{1}{2}$. In fact, a slightly more intricate argument, see Proposition 6.9 of Paper II, shows that the inequality in (3.15) is always strict even for countable state Markov chains. It can further be remarked that the bound in Proposition 3.2 holds in much greater generality, see the discussion in Section 4 of Paper II.

In Section 5 of Paper II, we describe a class of finite state Markov chains indexed by a real parameter $p \in (0, 1)$, see Figure 3.2, that segregates two states x and y , while

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} = p^{\frac{1}{1-p}}. \quad (3.16)$$

In particular, letting $p \rightarrow 1$, this tends to e^{-1} , which proves that $\kappa \geq e^{-1}$. In this case, the segregating coupling is described by the simple rule that, whenever $\{Y_n\}_{n=0}^\infty$ moves to the side for the first time, $\{X_n\}_{n=0}^\infty$ should move to the side in the next time step.

Lastly, in Section 6 of Paper II we prove that for any $\varepsilon > 0$ there exists a finite state Markov chain that segregates two states x and y while

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - P^n(y, \cdot)\|_{TV} \geq \frac{1}{2} - \varepsilon. \quad (3.17)$$

As a consequence, we can conclude that $\kappa = \frac{1}{2}$. More precisely, we show that for any $\varepsilon > 0$ there exist a finite state Markov chain, states x and y , and a positive integer T such that copies of the chain started in x and y respectively can be coupled such that their first meeting time is at most T a.s., while

$$\|P^T(x, \cdot) - P^T(y, \cdot)\|_{TV} \geq \frac{1}{2} - \varepsilon. \quad (3.18)$$

Given such a chain, it is straight-forward to construct a chain that satisfies (3.17).

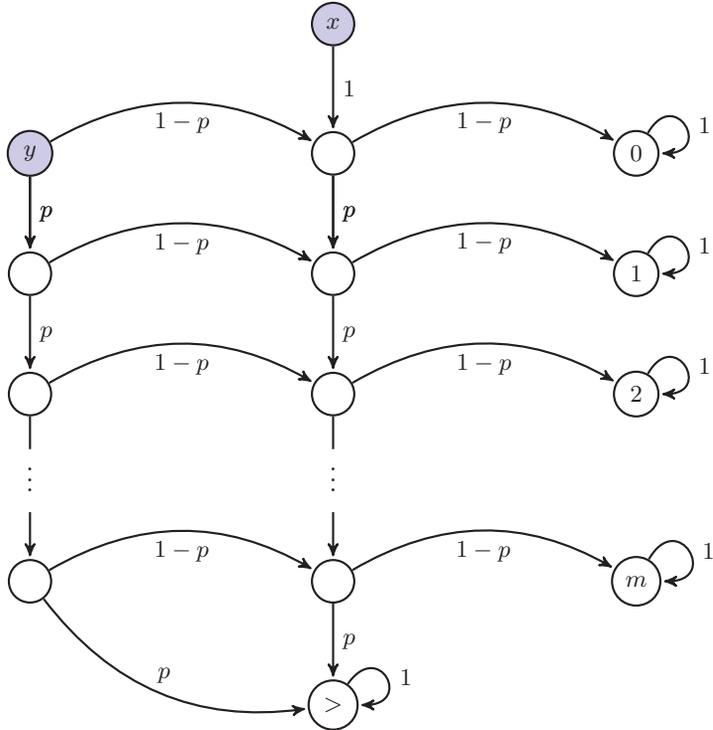


Figure 3.2: The “two drunkards on a cliff” Markov chain described in Section 5 of Paper II to obtain segregation with total variation distance $e^{-1} - \varepsilon$ for any $\varepsilon > 0$. This chain can be described as follows. Two drunkards, X and Y with initial positions x and y , walk along a cliffside, represented by the absorbing states on the right. For each step, they stumble to the side with probability $1-p$, and fall down the cliff after one and two stumbles respectively.

The Markov chains considered in this result are so-called finite birth-death chains with almost absorbing end points x and y . For these chains, it is relatively easy to estimate the total variation distance as in (3.18). The hard part is to prove the existence of the corresponding couplings. In order to do this, the key idea is to consider two related optimization problems.

For a given Markov chain, two initial states x and y , and a positive time T , we define the optimal meeting probability as

$$\mathcal{C}_T(x, y) = \max \mathbb{P}(X_n = Y_n \text{ for some } 0 \leq n \leq T), \quad (3.19)$$

where the maximum is taken over all couplings with initial states x and y . We can interpret optimal coupling in this sense as a linear optimization problem, which means that it has a corresponding dual problem. This dual turns out to have a natural interpretation, which we call separation.

A sequence $\mathcal{A} = \{A_n\}_{n=0}^T$ of subsets of the state space of a Markov chain is called a separating sequence. The separation of \mathcal{A} is given by

$$\begin{aligned} \mathcal{S}_T^{\mathcal{A}}(x, y) &= \mathbb{P}(X_n \in A_n \text{ for all } 0 \leq n \leq T | X_0 = x) \\ &\quad + \mathbb{P}(X_n \in A_n \text{ for all } 0 \leq n \leq T | X_0 = y). \end{aligned} \quad (3.20)$$

We say that the separating sequence is non-trivial if both terms in the right-hand side are non-zero. Moreover, the optimal separation $\mathcal{S}_T(x, y)$ is the maximum separation over all separating sequences. Note that the trivial separating sequences obtained by taking A_n constantly equal to S or to \emptyset have a separation of 1, hence the optimal separation is always at least 1.

We show that optimal meeting probability and optimal separation are strongly dual to each other in the sense that

$$\mathcal{C}_T(x, y) = 2 - \mathcal{S}_T(x, y), \quad (3.21)$$

see Theorem 6.3 of Paper II. In particular, in order to show that the optimal meeting probability is 1, it suffices to show that no non-trivial separating sequence yields a separation of more than 1.

In applying this to our finite birth-and-death chains, the idea is that, at least roughly, the best non-trivial way to separate the end-points is to cut the chain in the middle. By tweaking parameters of the chain and T , we can make this less than 1 while maintaining a total variation distance of approximately $\frac{1}{2}$.

Chapter 4

Altitudes of random edge-orders of the complete graph

The topic of Paper III is a problem concerning random orderings of the edges of the complete graph. While the precise formulation of this problem is quite new, it is motivated by a long-standing open problem in combinatorics posed by Chvátal and Komlós [15].

Speaking very broadly, we can describe a large class of problems within extremal combinatorics by the following philosophy. We are given a discrete structure, on which an ordering is imposed. Then, regardless of the choice of ordering, what is the maximum size of a “monotone substructure” that we are guaranteed to find?

One well-known example of this type of problem is the so-called Erdős-Szekeres theorem. Let a_1, a_2, \dots, a_n be a sequence of n distinct elements from a totally ordered set, say for instance \mathbb{R} , or $\{1, 2, \dots, n\}$. The Erdős-Szekeres theorem states that any such sequence has a subsequence of length $\lceil \sqrt{n} \rceil$ which is monotone. Moreover, this is sharp in the sense that for any $n \geq 1$, there are such length n sequences where no subsequence of length at least $\lceil \sqrt{n} \rceil + 1$ is monotone.

In 1971, Chvátal and Komlós proposed another problem of this type, this time concerning monotone (self-avoiding) paths in edge-ordered graphs. For a graph $G = (V, E)$, an edge-ordering on G is a bijective map $\phi : E \rightarrow \{1, 2, \dots, |E|\}$, where $\phi(e_1) < \phi(e_2)$ is interpreted as that e_1 comes before e_2 in the ordering. A pair (G, ϕ) as above is called an edge-ordered graph. A trail v_0, v_1, \dots, v_l in an edge-ordered graph is said to be monotone if either

$$\phi(\{v_0, v_1\}) < \phi(\{v_1, v_2\}) < \dots < \phi(\{v_{l-1}, v_l\}) \quad (4.1)$$

or

$$\phi(\{v_0, v_1\}) > \phi(\{v_1, v_2\}) > \cdots > \phi(\{v_{l-1}, v_l\}). \quad (4.2)$$

The altitude of an edge-ordering ϕ on a graph G , denoted by $f(G, \phi)$, is the maximal length of a ϕ -monotone (self-avoiding) path in G . The altitude of a graph G is defined as the minimum altitude $f(G) = \min_{\phi} f(G, \phi)$ over all edge-orderings of G . We similarly define $f^*(G, \phi)$ as the length of the longest ϕ -monotone trail in G , and $f^*(G) = \min_{\phi} f^*(G, \phi)$.

4.1 The altitude of the complete graph

Slightly rephrased to fit better together with the notation of later results, the question proposed by Chvátal and Komlós was to determine the values of $f^*(K_n)$ and $f(K_n)$ for all n . The former quantity was completely determined by Graham and Kleitman in a paper [29] published in 1973: $f^*(K_n) = n - 1$ unless $n = 3$ or 5 , $f^*(K_3) = 3$ and $f^*(K_5) = 5$. Winkler [49] described an elegant argument for the lower bound, due to Ehud Friedgut, which is known as the pedestrian argument: We initially place one “pedestrian” at each vertex of K_n . We then go through the edges in increasing order, and for each edge we swap the two pedestrians at its end-points. It follows that the movement of each pedestrian corresponds to an increasing trail, and on average a pedestrian moves $2\binom{n}{2}/n = n - 1$ steps.

In comparison to $f^*(K_n)$, the value of $f(K_n)$ has proven to be much more difficult to determine, and is still largely an open problem. Graham and Kleitman showed that

$$\sqrt{n - 3/4} - 1/2 \leq f(K_n) \leq \frac{3}{4}n + O(1). \quad (4.3)$$

The lower bound is obtained by modifying their argument for $f^*(K_n)$. To obtain the upper bound, they partition the vertices of K_n into four sets S_1, S_2, S_3, S_4 of roughly the same size, and consider an edge-ordering that first orders the edges in each part, then edges between S_1 and S_2 or S_3 and S_4 , then edges between S_1 and S_3 or S_2 and S_4 , then finally between S_1 and S_4 or S_2 and S_3 .

In his Master’s thesis, also from 1973, Rödl [47] gave an elegant argument that $f(G) \geq (1 - o(1))\sqrt{d}$ where d is the average degree of G . More precisely, for any integer $k \geq 0$ it holds that if $d \geq 2(1 + 2 + \cdots + k) = k(k + 1)$, then $f(G) \geq k$. Let G be an edge-ordered graph with average degree at least $2(1 + 2 + \cdots + k)$. We construct the subgraph G' of G obtained by removing the k largest edges connected to each vertex. As the average degree of G' is at least $d - 2k \geq 2(1 + 2 + \cdots + (k - 1))$, it follows by induction that it contains a monotone path of length $k - 1$. By adjoining a suitable edge from $G \setminus G'$, this can be extended to a monotone path of length k in G .

It is claimed that Alspach, Heinrich and Graham proved that $f(K_n) \leq \frac{7}{12}n$ (unpublished, see [14]), though as noted in [13] this false for small n . Currently, the best known upper bound is $f(K_n) \leq (\frac{1}{2} + o(1))n$, which was shown in a paper by Calderbank, Chung and Sturtevant in 1984 [14]. When $n = 2^k$ for some $k \geq 1$, their construction can be described as follows. Enumerate the vertices of K_n from 0 to $2^k - 1$. We think of these as k -bit binary strings. We then label each edge of K_n according to

$$l(\{v, w\}) = v \text{ XOR } w, \quad (4.4)$$

where XOR denotes bitwise exclusive or, and finally order the edges according to their labels, with ties broken arbitrarily. Using an elaborate argument, the authors show that any monotone path in this ordering has length at most $(\frac{1}{2} + o(1))n$. Essentially the same construction can be used for any n .

Until recently, the above results comprised everything known about $f(K_n)$. However, in early 2015, De Silva, Molla, Pfender, Retter and Tait uploaded a preprint to arxiv.org [21] (published 2016), where it was shown that, for any positive sequence $\{\omega_n\}_{n=1}^\infty$ tending to infinity as $n \rightarrow \infty$, we have

$$f(G(n, p)) \geq (1 + o(1)) \min(\sqrt{n}, \frac{np}{\omega_n \ln n}), \quad (4.5)$$

with probability tending to 1 uniformly in p , where $G(n, p)$ denotes the usual Erdős–Rényi graph. In particular, with probability tending to 1, $G(n, n^{-1/2+\varepsilon})$ has altitude at least $(1 - o(1))\sqrt{n}$. As altitude is non-decreasing when adding edges, this strongly indicates that the altitude of K_n should be much larger.

Later the same year, Milans uploaded a preprint to arxiv.org [40], showing that for any graph G on n vertices with average degree $d = \Omega(n^{2/3} (\ln n)^{4/3})$, we have

$$f(G) = \Omega\left(\frac{d}{n^{1/3} (\ln n)^{2/3}}\right). \quad (4.6)$$

In particular,

$$f(K_n) \geq \left(\frac{1}{20} - o(1)\right) \left(\frac{n}{\ln n}\right)^{2/3}, \quad (4.7)$$

finally making a significant improvement upon Graham and Kleitman's lower bound after over four decades.

Beyond these results, the value of $f(K_n)$ remains an open question. It is conjectured, first by Graham and Kleitman, and then repeated by Milans, that $f(K_n)$ should lie closer to the upper than to the lower bounds. Burger, Cockayne and Mynhardt [13] compute the values of $f(K_n)$ for $3 \leq n \leq 8$. It can be noted that these all satisfy $f(K_n) \geq \frac{1}{2}n$, though $n \leq 8$ is probably too small to infer a pattern.

4.2 Related results

While any progress on the altitude of the complete graph has been scarce for the last three decades, a number of studies have appeared during this time which consider the altitude of other classes of graphs.

It is clear from the definition of altitude that only empty graphs have altitude zero, and the only graphs with altitude one are non-empty graphs where each component contains at most two vertices. In 1987, Bialostocki and Roditty [8] further characterized all graphs of altitude two by proving that $f(G) \geq 3$ if and only if G contains, as a subgraph, either one of six fixed graphs, or an odd cycle of length at least 5.

In computing upper bounds for the altitude of a graph, multiple authors have observed that f is sub-additive in the sense that

$$f\left(\bigcup_{i=1}^k G_i\right) \leq \sum_{i=1}^k f(G_i), \quad (4.8)$$

where G_1, G_2, \dots, G_k denote graphs on the same vertex set. Let $\Delta = \Delta(G)$ denote the maximal degree of any vertex in G . By Vizing's theorem (see e.g. Theorem 1.13 in [9]), any graph G can be partitioned into at most $\Delta(G) + 1$ matchings. Combining this with the sub-additivity of f , it follows that

$$f(G) \leq \Delta(G) + 1. \quad (4.9)$$

Given this bound, it is natural to ask how large one can make $f(G)$ for a fixed value of $\Delta(G)$. To this end, in 2001 Yuster [50] proved that there exists a sequence $\{G_k\}_{k=2}^{\infty}$ of graphs such that $\Delta(G_k) = k$ and $f(G_k) \geq k - o(k)$. Later the same year, Alon [4] improved this result using the following observation: It is a well-known fact that for any $k \geq 2$ there are k -regular graphs with girth more than k . By the pedestrian argument above, any edge-ordering of any such graph contains a monotone trail of length k . As the girth of the graph is larger than k , this trail must be a path. Hence for any $\Delta \geq 2$ there exist graphs G with $\Delta(G) = \Delta$ and $f(G) \geq \Delta$.

An in-depth study of the altitude of 3-regular graphs by Mynhardt, Burger, Clark, Falvai and Henderson [43] showed that there are 3-regular graphs with altitude 4, e.g. the so-called flower snarks. Hence the upper bound in (4.9) can be obtained exactly for $\Delta = 3$, but for $\Delta \geq 4$ it remains an open problem whether the maximally obtainable altitude is Δ or $\Delta + 1$.

Roditty, Shoham, and Yuster [46] asked for the maximal altitude of any planar graph. They showed that the planar graph obtained by taking a long cycle and adding two vertices which are connected to everything in the cycle has altitude at least 5. On the other hand, by considering partitions of planar graphs into forests and using (4.8), they showed that any planar graph has altitude at most 9. Combining their argument with

a later result for the so-called caterpillar arboricity of planar graphs by Gonçalves [28], this upper bound can be improved to 8.

A recent article by De Silva, Molla, Pfender, Retter and Tait [21] considered the altitude of the d -dimensional hypercube Q_d . It is not too hard to partition Q_d into d matchings, hence by (4.8), $f(Q_d) \leq d$. This bound is believed to be sharp for all $d \geq 1$, but already $d = 4$ is hard to check by computer. De Silva et al. showed as a partial result in this direction that $f(Q_d) \geq d/\ln d$ for all $d \geq 2$.

Concerning computational aspects of altitude, Katrenič and Semanišin [33] showed that computing the altitude of an edge-ordered graph is NP -hard. In fact, even the simpler problems of (i) determining whether a certain edge-ordered graph admits a monotone Hamiltonian path or (ii) estimating the altitude of an edge-ordered graph within a constant factor are known to have no efficient solutions unless $P = NP$. Although it seems likely to be true, it remains an open question whether the problem of computing the altitude of a non edge-ordered graph is also NP -hard.

The study of altitudes of graphs has further led to the related concept of the depression of a graph. Given a graph G , its depression $\varepsilon(G)$ is defined as the smallest non-negative integer k such that any edge-ordering of G contains a maximal monotone path of length at most k , where a monotone path is said to be maximal if it is not contained in a longer such path. This has been characterized for various graphs, see for instance [16, 17, 42].

4.3 Random edge-orderings

In a recent paper, Lavrov and Loh [36] proposed a randomized version of Chvátal and Komlós' original question. Suppose we choose an edge-ordering of K_n uniformly at random. Then, what can we say about the altitude of this ordering? That is, what is the maximum length of a monotone path in a random edge-ordering?

It is natural to begin by considering the performance of the “obvious” greedy algorithm to explicitly construct long monotone paths – start in an arbitrary vertex and then successively traverse the minimally labeled edge which maintains the increasing property and does not lead to a previously visited vertex. Lavrov and Loh show that the monotone path obtained in this way has length concentrated around $(1 - \frac{1}{e})n \approx 0.632n$.

The behavior of the greedy path can intuitively be understood as follows. Let us assume the edge-ordering is obtained by assigning independent $U(0, 1)$ edge weights to K_n . The first edge in the path is the minimum out of $n - 1$ edges, and so the weight of this edge is approximately $\text{Exp}(n - 1)$ -distributed. After the first step, there are $n - 2$ edges that lead to new vertices, and so the weight of the second edge is approximately an $\text{Exp}(n - 2)$ random variable higher than the weight of the first

edge. By continuing in this way, the k :th edge in this path (if it exists) will have a weight about

$$\frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-k} \approx \ln n - \ln(n-k). \quad (4.10)$$

This continues until weights become close to 1, which occurs at $k \approx (1 - \frac{1}{e})n$.

Clearly, this greedy approach is not optimal as it only makes use of a very limited amount of information each time it chooses a new edge. To improve upon this, Lavrov and Loh propose a more clever way to successively choose the next edge, which they call the k -greedy algorithm. The precise formulation is a bit technical. They show that for each k , the path obtained in this way has length concentrated around some value $(1 - e^{-\alpha_k})n$. The value of α_k can be estimated numerically. In particular, letting $k \rightarrow \infty$ we get that $1 - e^{-\alpha_k}$ converges to a value ≈ 0.853 , and so, with probability tending to 1 as $n \rightarrow \infty$, a random edge-ordering of K_n contains a monotone path of length at least, say, $0.85n$.

It should probably be noted that already the first greedy algorithm shows that typical random edge-orderings contain monotone paths longer than $\frac{1}{2}n$. Hence, these will not help us to improve the upper bound on $f(K_n)$. This should not be very surprising – the central idea of the upper bounds in [14, 29] is to order the edges according to some global structure on K_n , which is more or less the opposite of what a random ordering does.

Besides providing explicit algorithms for finding long monotone paths, Lavrov and Loh further gave a very interesting partial result regarding the length of the longest monotone path in a random edge ordering. First, we have the trivial upper bound of $n - 1$: the length of a Hamiltonian path. It is natural to ask whether typically any of these are monotone.

To investigate this, let X_n be the number of increasing Hamiltonian paths in a random edge-ordering of K_n . For $n \geq 3$ the number of monotone such paths is then twice this value. There are $n!$ Hamiltonian paths in K_n , and the probability that each of these is increasing is $\frac{1}{(n-1)!}$. Hence, by linearity of expectation,

$$\mathbb{E}X_n = \frac{n!}{(n-1)!} = n. \quad (4.11)$$

Heuristically, the fact that the expectation of X_n tends to infinity would seem to indicate that X_n is typically non-zero. Using an elegant combinatorial argument, Lavrov and Loh showed that

$$\mathbb{E}X_n^2 = (1 + o(1))en^2. \quad (4.12)$$

This shows that the standard deviation of X is $\sim \sqrt{e-1} \cdot n$, which is too large to show concentration around its mean using Chebyshev's inequality.

Instead, by applying the Paley–Zygmund inequality, they obtained the weaker estimate

$$\mathbb{P}(X_n > 0) \geq \frac{\mathbb{E}[X_n]^2}{\mathbb{E}[X_n^2]} \rightarrow \frac{1}{e} \text{ as } n \rightarrow \infty, \quad (4.13)$$

and so, with probability at least $\frac{1}{e} - o(1)$, a random edge-ordering of K_n contains a monotone Hamiltonian path.

In conclusion of these results, the authors state that empirical simulations seem to indicate that a stronger result is true, and give the following conjecture.

Conjecture 4.1. If an edge-ordering of K_n is chosen uniformly at random, then it admits to a monotone Hamiltonian path with probability tending to 1 as $n \rightarrow \infty$.

Indeed, given (4.13), the negation of this conjecture would mean that there are arbitrarily large values of n where a random edge-ordering is sometimes organized to block all Hamiltonian paths, and sometimes allows many to be monotone. As each individual Hamiltonian path contains relatively few of the edges of K_n , this would at least be very peculiar.

4.4 Summary of Paper III

In Paper III, we consider the problem of random edge-orderings of K_n as described in the previous section. In particular, we investigate the behavior of the random variable X_n in greater detail. Our approach is reminiscent of Stein’s method in the sense that it relates X_n to its size-biased version X_n^* , but as far as I know, the precise approach is novel to the paper.

The main result resolves the conjecture by Lavrov and Loh affirmatively – a uniformly chosen random edge-ordering of K_n admits a monotone Hamiltonian path with probability $1 - o(1)$. This is a consequence of a more technical result below. Before delving into that, let us note some more properties of X_n that follows from this. First, for any $x > 0$ and any positive integer $k \geq 0$ we have the lower tail estimate

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \leq x \cdot n) \leq e^{k(k+1)} x^k. \quad (4.14)$$

As a consequence, with probability $1 - o(1)$, the number of monotone Hamiltonian paths is of order n . Second, for a suitable truncation \hat{X}_n of X_n such that $\hat{X}_n \leq X_n$ and $\mathbb{P}(X_n = \hat{X}_n) \rightarrow 1$ as $n \rightarrow \infty$, it holds that

$$\lim_{n \rightarrow \infty} \frac{1}{n^k} \mathbb{E} \hat{X}_n^k = e^{k(k-1)/2}, \quad (4.15)$$

for any $k = 1, 2, \dots$. Third, for any large constant $M > 0$, the probabilities that $X_n/n \geq M$ and $X_n/n \leq \frac{1}{M}$ respectively are both bounded away

from 0 as $n \rightarrow \infty$, so the sequence X_n/n does not have bounded support in the large n limit.

It seems very reasonable to expect that the sequence $\{X_n/n\}_{n=1}^\infty$ should converge in distribution, but it is not clear how to prove this. We note however that $X_n/n \geq 0$ and $\mathbb{E}X_n/n = 1$ for all $n \geq 1$, and hence the sequence is tight, meaning that no probability mass “escapes to infinity” as $n \rightarrow \infty$. This implies that any subsequence of these random variables contains a subsubsequence which converges in distribution. We note that if all such converging sequences have the same limit, then X_n/n converges to this distribution as $n \rightarrow \infty$.

The three properties of X_n listed above are all direct consequences of the following characterization of limit points of X_n/n .

Proposition 4.2. *Let $F(x)$ denote the cumulative distribution function of the limit of any weakly converging subsequence $\{X_{n_i}/n_i\}_{i=1}^\infty$. We then have*

$$\int_0^\infty x^k dF(x) = e^{k(k-1)/2} \quad (4.16)$$

for any $k \in \mathbb{Z}$. In particular, F has the same moments as a log-normal random variable with $\mu = -\frac{1}{2}$ and $\sigma = 1$. Moreover, if we let $G(t) = F(e^t)$, equivalently $G(t)$ is the CDF of the limit of $\ln(X_{n_i}/n_i)$, then

$$G(t) = \int_{-\infty}^t e^{-(s+\frac{1}{2})^2/2} d\nu(s) \quad (4.17)$$

for some 1-periodic positive measure $\nu(s)$ on \mathbb{R} .

Recall that the log-normal distribution with parameters μ and σ is the distribution of $Y = e^Z$, where Z has normal distribution with mean μ and standard deviation σ .

An important caveat concerning this result is that the log-normal distribution is M -indeterminate, meaning that there exist other random variables that have the same moment sequence. Hence, even though we uniquely determine the moment sequence of limit points of X_n/n , this does not uniquely determine the limit distribution. In fact, any distribution that satisfies (4.17) for some 1-periodic measure has this property. Nevertheless, it seems fairly clear that the only natural distribution that can be obtained in this way is the log-normal one.

Conjecture 4.3. As $n \rightarrow \infty$, X_n/n converges in distribution to a $\log \mathcal{N}(-\frac{1}{2}, 1)$ random variable.

Let us briefly mention some of the proof ideas of this paper. Given a non-negative random variable ξ with finite positive expectation μ , we say that ξ^* has ξ -size biased distribution if

$$\mathbb{E}[\xi f(\xi)] = \mu \mathbb{E}f(\xi^*), \quad (4.18)$$

for all functions f for which these expectations exist. In the case of X_n , its size-biased version X_n^* then satisfies

$$\mathbb{P}(X_n^* = k) = \frac{k}{n} \mathbb{P}(X_n = k), \quad (4.19)$$

for $k = 0, 1, \dots$.

As it turns out, there is a natural coupling of X_n and X_n^* defined as follows. Uniformly choose an edge-ordering of K_n and define X_n according to this. Thereafter, fix some Hamiltonian path P and rearrange the indices of edges along P such that it becomes increasing. We can now let X_n^* be the number of increasing Hamiltonian paths in the modified edge-ordering.

The key observation in this analysis is the following simple relation between X_n and X_n^* .

Proposition 4.4. *In the coupling above, we have*

$$\mathbb{E} \left[(X_n^* - e X_n)^2 \right] = o(n^2). \quad (4.20)$$

It can be mentioned, that despite the simplicity of this statement, its proof consists of some fairly technical third-moment estimates for X_n , which take up the majority of the paper.

Intuitively, this proposition states that in the large n limit, size-biasing scales up X_n by a factor of e but does not otherwise affect its distribution. In particular, the proposition implies that any limit point of X_n/n has this property exactly. Random variables where size-biasing scales up the distribution by a constant factor are known in the literature [5], and in particular when the constant is e , these have the properties described in Proposition 4.2.

Chapter 5

Jigsaw puzzles and shotgun assembly

Suppose a random jigsaw puzzle is constructed by cutting out an $n \times n$ grid of jigsaw pieces from a large board where the borders between adjacent pieces are chosen uniformly out of q possible shapes. Then, what is the probability that this puzzle has a unique solution? This is the topic of Paper IV. Despite its natural formulation, this problem seems to have been first considered very recently in a paper by Mossel and Ross [41].

In the aforementioned paper, Mossel and Ross proposed a class of problems going under the following philosophy. There exists some, possibly random, graph G which is assigned either a random edge or vertex labeling. For each vertex $v \in G$, its labeled r -neighborhood $\mathcal{N}_r(v)$ is the induced labeled subgraph of all vertices at distance at most r from v . You are told the collection of all such r -neighborhoods for some r . Given this information, is it possible to reconstruct G ?

Probably the most well-known problem of this type is to reconstruct a random string given a collection of overlapping substrings, see for instance [25]. In particular, this is an important problem in genetic sequencing, where it goes under the name of shotgun assembly. Mossel and Ross hence called their class of problems shotgun assembly of labeled graphs.

As a more recreational version of shotgun assembly, the authors proposed the random jigsaw puzzle problem. They in particular asked: “How large should q be in order for the puzzle to have a unique solution? Moreover, how can this solution be found efficiently?”

5.1 Edge-matching and jigsaw puzzles

A jigsaw puzzle is a collection of square pieces where each of the four edges of the piece is assigned a label, called its shape, out of q options.

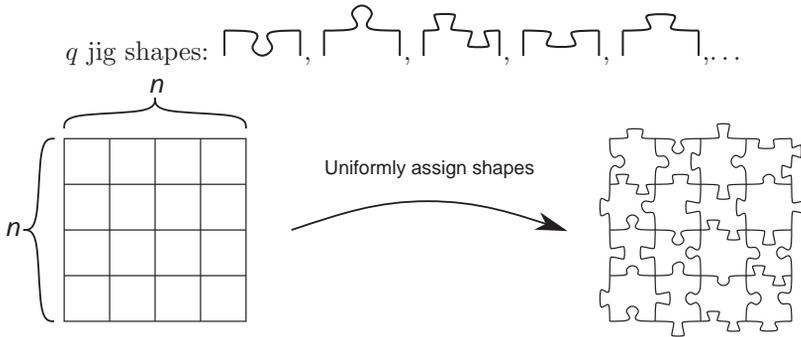


Figure 5.1: Illustration of the random jigsaw puzzle model proposed by Mossel and Ross.

For each possible shape i we assume that there exists a unique inverted shape j (this may be i itself) such that two edges with respective shapes i and j “fit together”. An edge-matching puzzle is a collection of square pieces where each side is instead given one of q possible color. Here, two edges fit together if they have the same color.

A solution to the respective puzzles is an arrangement of the pieces into a certain form, here it will always be an $n \times n$ square, such that neighboring pieces fit together. It is sometimes assumed that the pieces have a fixed orientation, in which case a solution must orient all pieces the same way. Otherwise, we assume that the pieces are allowed to be rotated any multiple of 90° but not flipped upside down.

We shall here be concerned with the following models for random edge-matching and jigsaw puzzles, as defined in [11, 41]. Arrange n^2 identical squares into an $n \times n$ pattern. For each edge of each square, we choose its shape/color uniformly from q options under the restriction that any pairs of touching edges should have inverted shapes/the same color. This is illustrated in Figure 5.1. Note that this means that, unlike most real jigsaw puzzles, also edges along the boundary are assigned shapes/colors. This assumption should not affect the behavior of the model very much, but simplifies the analysis. An assignment of shapes to a puzzle is called a carving, and an assignment of colors a coloring.

A natural question regarding these models is how likely it is for such puzzles to have a unique solution. Equivalently, how likely is it that the original solution of the puzzle can be uniquely recovered from the collection of shuffled pieces. One would intuitively expect this probability to be increasing in q – more shapes/colors gives us more information about the puzzle and generally decreases the probability that the pieces fit together in some other arrangement.

In order to treat these questions, a few caveats need to be mentioned.

First, if we allow the pieces to be rotated, then any global rotation of a solution also solves the puzzle. To simplify terminology, we will consider two solutions as identical if they only differ by a global rotation. Second, if the puzzle contains duplicate pieces, or, in the case where rotation is allowed, pieces with rotational symmetry, then any appropriate rearrangement of such pieces in a solution yields a new solution. We say that two solutions are similar if they only differ by global rotation and rearrangement of duplicate and rotationally symmetric pieces.

Mossel and Ross showed that for $2 \leq q \leq o(n^{2/3})$, with probability tending to 1 as $n \rightarrow \infty$, a random edge-matching puzzle contains a pair of disjoint 2×1 blocks with identical colors around their borders but different in their centers. Hence by exchanging these in the original solution, we see that the puzzle has at least two non-similar solutions. Conversely, if $q = \omega(n^2)$ then the solution is unique with probability tending to one. It should be mentioned that it is not entirely clear from their presentation whether they allow pieces to be rotated, but the fact is their argument goes through either way.

This result was later improved in two independent works by Bordenave, Feige and Mossel [11] and Nenadov, Pfister, and Steger [44]. Incidentally, both were published on arxiv.org on May 11:th 2016, and both prove essentially the same result: When $q \geq n^{1+\varepsilon}$, a random edge-matching puzzle has a unique solution with probability tending to 1, for any fixed $\varepsilon > 0$. On the other hand, when $q = o(n)$, with probability tending to 1, the puzzle contains duplicate pieces and hence multiple, but possibly all similar, solutions. It should be mentioned that Nenadov et al. assume that rotations are allowed. Bordenave et al. on the other hand state that they, for simplicity, assume that the pieces are not allowed to be rotated, but argue in the last section how this assumption can be removed and also how the same result can be extended to the model of random jigsaw puzzles above.

As part of their proof in the case of $q \geq n^{1+\varepsilon}$, Bordenave et al. present an algorithm that reconstructs the original solution with probability tending to 1 in time $n^{O(1/\varepsilon)}$. Slightly paraphrased, their idea is that in order to check whether two pieces with matching edges were adjacent in the original solution, it suffices to check whether it is possible to construct a $k \times k$ local solution centered around them for some fixed $k = k(\varepsilon)$.

By comparison, the general algorithmic problems of finding a solution to a given edge-matching or jigsaw puzzle are known to be NP -complete [20], with a related hardness of approximation result given in [12]. The problem also seems to be hard in practice. In the summer of 2007, Christopher Monckton announced a \$2 million prize for the first complete solution to the (slightly oxymoronically named) Eternity II puzzle [51]. This is an edge-matching puzzle consisting of 256 square pieces with 22 distinct colors that should be assembled into a 16×16 grid. The competition ended on December 31:st 2010 with no complete solution being found, and at

the time of writing the puzzle is claimed to remain unsolved.

5.2 Summary of Paper IV

In Paper IV we consider the transition between multiple and unique solutions to random edge-matching and jigsaw puzzles in further detail. We here assume that the pieces are allowed to be rotated. The result is summarized in the following theorem.

Theorem 5.1. *As $n \rightarrow \infty$, the following holds with probability tending to 1 for a random jigsaw puzzle with q shapes or random edge-matching puzzle with q colors.*

1. *For $2 \leq q \leq \frac{2}{\sqrt{e}}n$, there are at least two non-similar solutions.*
2. *For $q \geq (2 + \varepsilon)n$, for any fixed $\varepsilon > 0$, all solutions are similar.*
3. *For $q = \omega(n)$, the solution is unique (up to global rotation).*

I further conjecture that the event that all solutions are similar has a sharp threshold at $q = \frac{2}{\sqrt{e}}n$ in the sense that, in addition to point 1 above, all solutions should be similar with probability tending to 1 for $q \geq (\frac{2}{\sqrt{e}} + \varepsilon)n$, for all $\varepsilon > 0$.

The first part of the theorem is a consequence of the following observation. The statement “all solutions are similar” is equivalent to saying that the unordered collection of jigsaw pieces uniquely determines the carving of the original solution, up to global rotation. Hence there are at most four times as many carvings where all solutions are similar as there are unordered collections of n^2 jigsaw pieces. When $2 \leq q \leq \frac{2}{\sqrt{e}}n$, the number of the latter is much smaller than the total number of carvings.

One can further note that the event that the solution is unique is the intersection of the event that all solutions are similar and the event that the puzzle contains no duplicate or rotationally symmetric pieces. In particular, the expected numbers of both rotationally symmetric pieces and pairs of duplicate pieces are $o(1)$ if and only if $q = \omega(n)$. Hence, part 2 of the theorem implies part 3.

The majority of Paper IV is dedicated to proving part 2 of the theorem. The general approach is similar to those of Bordenave et al. and Nenadov et al. – if there exist some alternative solution to the puzzle, then there exists local solutions where a relatively large number of pieces have neighbors different from what they have in the original configuration. In order to prove that this cannot happen, one wants to bound the expected number of such local solutions. One main obstacle to this approach is that, for a given arrangement of pieces, the events that pairs of new neighboring pieces fit together are not generally independent. A novelty of this paper is to circumvent this by controlling these dependencies in terms of the number of shapes that appear locally more than once.

Bibliography

- [1] David Aldous, *Probability approximations via the Poisson clumping heuristic*, Applied Mathematical Sciences, vol. 77, Springer-Verlag, New York, 1989. MR969362 (90k:60004)
- [2] David Aldous and James Allen Fill, *Reversible Markov Chains and Random Walks on Graphs*. Unfinished monograph, recompiled 2014, available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [3] Sven Erick Alm and Maria Deijfen, *First passage percolation on \mathbb{Z}^2 : a simulation study*, J. Stat. Phys. **161** (2015), no. 3, 657–678, DOI 10.1007/s10955-015-1356-0. MR3406703
- [4] N. Alon, *Problems and results in extremal combinatorics. I*, Discrete Math. **273** (2003), no. 1-3, 31–53, DOI 10.1016/S0012-365X(03)00227-9. EuroComb'01 (Barcelona). MR2025940
- [5] R. Arratia, L. Goldstein, and F. Kochman, *Size bias for one and all* (2015+). available at <http://arxiv.org/abs/1308.2729>.
- [6] Antonio Auffinger, M Damron, and J Hanson, *50 years of first passage percolation* (2016). arxiv preprint arXiv:1511.03262.
- [7] Antonio Auffinger and Si Tang, *On the time constant of high dimensional first passage percolation*, Electron. J. Probab. **21** (2016), Paper No. 24, 23, DOI 10.1214/16-EJP1. MR3485366
- [8] A. Bialostocki and Y. Roditty, *A monotone path in an edge-ordered graph*, Internat. J. Math. Math. Sci. **10** (1987), no. 2, 315–320, DOI 10.1155/S0161171287000383. MR886388
- [9] Béla Bollobás, *Extremal graph theory*, London Mathematical Society Monographs, vol. 11, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978. MR506522
- [10] Béla Bollobás and Yoshiharu Kohayakawa, *On Richardson's model on the hypercube*, Combinatorics, geometry and probability (Cambridge, 1993), Cambridge Univ. Press, Cambridge, 1997, pp. 129–137. MR1476439 (98j:05122)
- [11] C. Bordenave, U. Feige, and E. Mossel, *Shotgun assembly of random jigsaw puzzles* (2016+). available at <http://arxiv.org/abs/1605.03086>.
- [12] J. Bosboom, E.D. Demaine, M.L. Demaine, A. Hesterberg, P. Manurangsi, and A. Yodpinyanee, *Even $1 \times n$ Edge-Matching and Jigsaw Puzzles are Really Hard* (2016+). available at <http://arxiv.org/abs/1701.00146>.
- [13] A. P. Burger, E. J. Cockayne, and C. M. Mynhardt, *Altitude of small complete and complete bipartite graphs*, Australas. J. Combin. **31** (2005), 167–177. MR2113705
- [14] A. R. Calderbank, F. R. K. Chung, and D. G. Sturtevant, *Increasing sequences with nonzero block sums and increasing paths in edge-ordered graphs*, Discrete Math. **50** (1984), no. 1, 15–28, DOI 10.1016/0012-365X(84)90031-1. MR747708

- [15] V. Chvátal and J. Komlós, *Some combinatorial theorems on monotonicity*, *Canad. Math. Bull.* **14** (1971), 151–157. MR0337676
- [16] E. J. Cockayne, G. Geldenhuys, P. J. P. Grobler, C. M. Mynhardt, and J. H. van Vuuren, *The depression of a graph*, *Util. Math.* **69** (2006), 143–160. MR2212805
- [17] E. J. Cockayne and C. M. Mynhardt, *A lower bound for the depression of trees*, *Australas. J. Combin.* **35** (2006), 319–328. MR2239325
- [18] Olivier Couvonné, Nathanaël Enriquez, and Lucas Gerin, *Construction of a short path in high-dimensional first passage percolation*, *Electron. Commun. Probab.* **16** (2011), 22–28, DOI 10.1214/ECP.v16-1595. MR2753301
- [19] J. Theodore Cox and Richard Durrett, *Some limit theorems for percolation processes with necessary and sufficient conditions*, *Ann. Probab.* **9** (1981), no. 4, 583–603. MR624685 (82k:60208)
- [20] Erik D. Demaine and Martin L. Demaine, *Jigsaw puzzles, edge matching, and polyomino packing: connections and complexity*, *Graphs Combin.* **23** (2007), no. suppl. 1, 195–208, DOI 10.1007/s00373-007-0713-4. MR2320628
- [21] Jessica De Silva, Theodore Molla, Florian Pfender, Troy Retter, and Michael Tait, *Increasing paths in edge-ordered graphs: the hypercube and random graph*, *Electron. J. Combin.* **23** (2016), no. 2, Paper 2.15, 9. MR3512637
- [22] D. Dhar, *Asymptotic shape of Eden clusters*. B, *On Growth and Form: Fractal and Non-Fractal Patterns in Physics* (H. Eugene and Ostrowsky Stanley Nicole, ed.), NATO ASI Series, vol. 100, Springer, Netherlands, 1986, pp. 288–292, DOI 10.1007/978-94-009-5165-5_31.
- [23] ———, *First passage percolation in many dimensions*, *Phys. Lett. A* **130** (1988), no. 4-5, 308–310, DOI 10.1016/0375-9601(88)90616-0. MR949856
- [24] Richard Durrett, *Lecture notes on particle systems and percolation*, *The Wadsworth & Brooks/Cole Statistics/Probability Series*, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1988. MR940469 (89k:60157)
- [25] M. Dyer, A. Frieze, and S. Suen, *The Probability of Unique Solutions of Sequencing by Hybridization*, *Journal of Computational Biology* **1** (1994), no. 2, 105–110.
- [26] Murray Eden, *A two-dimensional growth process*, *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, Vol. IV, Univ. California Press, Berkeley, Calif., 1961, pp. 223–239. MR0136460
- [27] James Allen Fill and Robin Pemantle, *Percolation, first-passage percolation and covering times for Richardson’s model on the n -cube*, *Ann. Appl. Probab.* **3** (1993), no. 2, 593–629. MR1221168 (94h:60145)
- [28] D. Gonçalves, *Caterpillar arboricity of planar graphs*, *Discrete Math.* **307** (2007), no. 16, 2112–2121, DOI 10.1016/j.disc.2005.12.055. MR2326172
- [29] R. L. Graham and D. J. Kleitman, *Increasing paths in edge ordered graphs*, *Period. Math. Hungar.* **3** (1973), 141–148. Collection of articles dedicated to the memory of Alfréd Rényi, II. MR0327568
- [30] Olle Häggström, *A note on disagreement percolation*, *Random Structures Algorithms* **18** (2001), no. 3, 267–278, DOI 10.1002/rsa.1008. MR1824276
- [31] J. M. Hammersley and D. J. A. Welsh, *First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory*, *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif.*, Springer-Verlag, New York, 1965, pp. 61–110. MR0198576
- [32] C. Douglas Howard, *Models of first-passage percolation*, *Probability on discrete structures*, *Encyclopaedia Math. Sci.*, vol. 110, Springer, Berlin, 2004, pp. 125–173. MR2023652 (2005b:60258)

- [33] J. Katrenič and G. Semanišin, *Finding monotone paths in edge-ordered graphs*, Discrete Appl. Math. **158** (2010), no. 15, 1624–1632, DOI 10.1016/j.dam.2010.05.018. MR2671540
- [34] Harry Kesten, *Aspects of first passage percolation*, École d’été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 125–264, DOI 10.1007/BFb0074919. MR876084 (88h:60201)
- [35] J. F. C. Kingman, *Subadditive ergodic theory*, Ann. Probability **1** (1973), 883–909. With discussion by D. L. Burkholder, Daryl Daley, H. Kesten, P. Ney, Frank Spitzer and J. M. Hammersley, and a reply by the author. MR0356192
- [36] Mikhail Lavrov and Po-Shen Loh, *Increasing Hamiltonian paths in random edge orderings*, Random Structures Algorithms **48** (2016), no. 3, 588–611, DOI 10.1002/rsa.20592. MR3481275
- [37] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. MR2466937
- [38] Thomas M. Liggett, *An improved subadditive ergodic theorem*, Ann. Probab. **13** (1985), no. 4, 1279–1285. MR806224
- [39] Anders Martinsson, *Unoriented first-passage percolation on the n -cube*, Ann. Appl. Probab. **26** (2016), no. 5, 2597–2625, DOI 10.1214/15-AAP1155. MR3563188
- [40] K Milans, *Monotone Paths in Dense Edge-Ordered Graphs* (2015+). available at <http://arxiv.org/abs/1509.02143>.
- [41] E. Mossel and N. Ross, *Shotgun assembly of labeled graphs* (2015+). available at <http://arxiv.org/abs/1504.07682>.
- [42] C. M. Mynhardt, *Trees with depression three*, Discrete Math. **308** (2008), no. 5-6, 855–864, DOI 10.1016/j.disc.2007.07.039. MR2378919
- [43] C. M. Mynhardt, A. P. Burger, T. C. Clark, B. Falvai, and N. D. R. Henderson, *Altitude of regular graphs with girth at least five*, Discrete Math. **294** (2005), no. 3, 241–257, DOI 10.1016/j.disc.2005.02.007. MR2137566
- [44] R. Nenadov, P. Pfister, and A. Steger, *Unique reconstruction threshold for random jigsaw puzzles* (2016+). available at <http://arxiv.org/abs/1605.03043>.
- [45] Daniel Richardson, *Random growth in a tessellation*, Proc. Cambridge Philos. Soc. **74** (1973), 515–528, DOI 10.1017/S0305004100077288.
- [46] Y. Roditty, B. Shoham, and R. Yuster, *Monotone paths in edge-ordered sparse graphs*, Discrete Math. **226** (2001), no. 1-3, 411–417, DOI 10.1016/S0012-365X(00)00174-6. MR1802612
- [47] V. Rödl (1973). Master’s thesis, Charles University (1973).
- [48] Jeffrey S. Rosenthal, *Faithful couplings of Markov chains: now equals forever*, Adv. in Appl. Math. **18** (1997), no. 3, 372–381, DOI 10.1006/aama.1996.0515. MR1436487
- [49] Peter Winkler, *Puzzled: Solutions and Sources*, Commun. ACM **51** (September 2008), no. 9, 103–103, DOI 10.1145/1378727.1389570.
- [50] Raphael Yuster, *Large monotone paths in graphs with bounded degree*, Graphs Combin. **17** (2001), no. 3, 579–587, DOI 10.1007/s003730170031. MR1859935
- [51] <http://www.prnewswire.co.uk/news-releases/eternity-puzzle-is-back-this-summer-with-a-us2-million-prize-for-the-first-person-to-find-a-solution-153539245.html>.