

VU Research Portal

Percolation, loop soups and stochastic domination

van de Brug, T.

2015

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van de Brug, T. (2015). *Percolation, loop soups and stochastic domination*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Percolation, loop soups and stochastic domination

Tim van de Brug

Research supported by the Netherlands Organisation for Scientific Research (NWO)



Netherlands Organisation for Scientific Research

VRIJE UNIVERSITEIT

**Percolation, loop soups
and stochastic domination**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op donderdag 2 juli 2015 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Timmy van de Brug

geboren te Amstelveen

promotor: prof.dr. R.W.J. Meester
copromotoren: dr. F. Camia
dr. W. Kager

Contents

Preface	7
1 Introduction	9
1.1 Stochastic domination	9
1.2 Continuum percolation	10
1.3 Fractal percolation	12
1.4 Loop soups	13
2 Stochastic domination	17
2.1 Introduction and main results	17
2.1.1 Stochastic domination of finite vectors	18
2.1.2 Asymptotic framework	19
2.1.3 Weak convergence	20
2.1.4 Asymptotic stochastic domination	22
2.2 Stochastic domination of finite vectors	24
2.3 Weak convergence	28
2.3.1 Proof of Theorem 2.1.4	32
2.3.2 Proof of Theorem 2.1.5	35
2.3.3 Proof of Theorem 2.1.6	41
2.3.4 Law of large numbers	42
2.4 Asymptotic stochastic domination	43
2.4.1 Proof of Theorem 2.1.8	43
2.4.2 Exact conditioning	49
3 The random connection model	51
3.1 Introduction	51
3.2 Truncation and scaling	52
3.3 Proof of Theorem 3.1.1	55
3.4 Extension to larger components	60
4 Fractal percolation	65
4.1 Introduction	65
4.1.1 k -fractal percolation	65
4.1.2 Fat fractal percolation	67
4.2 Preliminaries	69
4.3 Proofs of the k -fractal results	70

4.3.1	Proof of Theorem 4.3.1	71
4.3.2	Proof of Theorem 4.3.2	74
4.3.3	Proof of Theorem 4.1.4	79
4.4	Proofs of the fat fractal results	81
4.4.1	Proof of Theorem 4.1.7	81
4.4.2	Proof of Theorem 4.1.8	82
4.4.3	Proof of Theorem 4.1.9	84
4.4.4	Proof of Theorem 4.1.11	85
5	Random walk loop soups	89
5.1	Introduction	89
5.2	Definitions and main result	90
5.3	Preliminary results	93
5.4	Finite approximation of a Brownian loop soup cluster	98
5.5	No touchings	102
5.6	Distance between Brownian loops	107
5.7	Proof of main result	111
	Summary	117
	Bibliography	119

Preface

This thesis is the result of a five-year period as a PhD student at the department of mathematics of VU University Amsterdam. I have thoroughly enjoyed this time, and I am very happy that I can stay one more year at the VU for research and teaching. Doing research was a nice way to get a deeper understanding of some topics in probability theory. It was interesting to work together with various people, each having another style of doing research and approaching mathematical problems from a different angle. It was also great to attend conferences and summer schools all over the world. This was a good opportunity to meet interesting people, obtain a broad knowledge of recent developments in the research field, and to see new places. Next to the research, I very much enjoyed teaching several courses in statistics and probability, first as a teaching assistant and last year as a lecturer. Many people, colleagues and students, made the last five years at the VU a very enjoyable time. I would like to thank a few of them in particular.

Ronald, thank you for being my advisor. It was great to come back to you some years after finishing my Master thesis. I thank you for many useful discussions, your confidence in me, and for keeping the overall view on the project. Federico, it was inspiring to work together with you. We had long and interesting discussions from which I learned a lot about loop soups and related theory, and about the international mathematical world in general. Wouter, it was very nice to work with you on the stochastic domination problem and on our recent project. You have a structured way of doing research and you always took a lot of time for discussions. Marcin, you are a smart person; it was inspiring to work with you on the loop soups. It was also nice to travel together to conferences in Brazil, Argentina, Vancouver, Marseille and Abu Dhabi. Matthijs, it was nice to work with you on the fractals, share an office and attend summer schools in Brazil and Florence. Erik, it was nice to work together on various topics.

I thank Rob van den Berg, Erik Broman, Frank den Hollander, Yves Le Jan and Anish Sarkar for being members of the reading committee. Maarten and René, thank you for being paranympths at the thesis defense, and for the pleasant and relaxed atmosphere in our shared office. Corrie and Fetsje, I have been your teaching assistant for several courses in statistics. I very much enjoyed teaching these classes and I thank you for your good advice on teaching. I thank Frank Heierman for the mathematics classes at secondary school which inspired me to start studying mathematics. Finally, I thank friends and family.

Tim van de Brug
Amsterdam, April 2015

Chapter 1

Introduction

This thesis is on probability theory, in particular on percolation, loop soups and stochastic domination. It is based on the following papers:

- (i) Erik I. Broman, Tim van de Brug, Wouter Kager, and Ronald Meester, *Stochastic domination and weak convergence of conditioned Bernoulli random vectors*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 403–434.
- (ii) Tim van de Brug and Ronald Meester, *On central limit theorems in the random connection model*, Phys. A **332** (2004), no. 1-4, 263–278.
- (iii) Erik I. Broman, Tim van de Brug, Federico Camia, Matthijs Joosten, and Ronald Meester, *Fat fractal percolation and k -fractal percolation*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 279–301.
- (iv) Tim van de Brug, Federico Camia, and Marcin Lis, *Random walk loop soups and conformal loop ensembles*, arXiv:1407.4295 [math.PR] (2014).

These four papers form the basis for Chapters 2–5, respectively. Below we give an introduction to each of the chapters, and we highlight the connections between them. We also state the main results of this thesis in an informal way.

1.1 Stochastic domination

Stochastic domination is about ordering the probability distributions of two or more random objects. A random object \mathbf{X} is said to be stochastically dominated by a similar random object \mathbf{Y} if \mathbf{X} is smaller, in a distributional sense, than \mathbf{Y} . For example, suppose that \mathbf{X} and \mathbf{Y} are random vectors on \mathbb{R}^n with laws μ and ν , respectively. Then \mathbf{X} is stochastically dominated by \mathbf{Y} if it is possible to define random vectors $\mathbf{U} = (U_1, \dots, U_n)$ and $\mathbf{V} = (V_1, \dots, V_n)$ on a common probability space such the laws of \mathbf{U} and \mathbf{V} are equal to μ and ν , respectively, and $\mathbf{U} \leq \mathbf{V}$, i.e. $U_i \leq V_i$ for all i , with probability 1. Stochastic domination and coupling techniques play an important role throughout this thesis and are used as a tool in many of the proofs. In particular, in Chapter 4 we describe a coupling of two fractal percolation

models, and we use this coupling to derive new properties of one of the models from known properties of the other model.

In Chapter 2, stochastic domination is not only used as a tool in the proofs but is our main topic of interest. We study vectors \mathbf{X} and \mathbf{Y} that each consist of n independent Bernoulli random variables with success probabilities p_1, \dots, p_n and q_1, \dots, q_n , respectively, such that $p_i \leq q_i$ for all i . Clearly, in this case \mathbf{X} is stochastically dominated by \mathbf{Y} . However, suppose that we consider the conditional law of \mathbf{X} , conditioned on the total number of successes being at least k for some integer k , and similar for \mathbf{Y} with the same integer k . Then in general the conditional law of \mathbf{X} is not necessarily stochastically dominated by the conditional law of \mathbf{Y} . We identify conditions under which we do have stochastic ordering of these conditional probability distributions.

Domination issues concerning the conditional law of Bernoulli vectors conditioned on having at least a certain number of successes have come up in the literature a number of times. In [3] and [4], the simplest case has been considered in which $p_i = p$ and $q_i = q$ for some $p < q$. In [4], the conditional domination is used as a tool in the study of random trees.

Here we study such domination issues in detail and generality. After dealing with conditioned Bernoulli random vectors as above, we consider sequences of Bernoulli vectors \mathbf{X}_n and \mathbf{Y}_n of length n that each consist of M “blocks” such that the Bernoulli random variables in block i have success probability p_i and q_i , respectively. Here M does not depend on n and the size of each block is essentially linear in n . We consider the conditional laws of \mathbf{X}_n and \mathbf{Y}_n , conditioned on the total number of successes being at least k_n , where k_n is also essentially linear in n . The main result of Chapter 2 is a complete answer to the question with what maximal probability two such conditioned Bernoulli vectors can be ordered in any coupling, when the length n of the vectors tends to infinity.

1.2 Continuum percolation

In Chapters 3–5 we study probabilistic models in which the geometry is the key feature of the model. We consider spatial models that are defined by putting random objects in two- or higher-dimensional Euclidean space. The models we study are examples of *continuum percolation* models, or related to continuum percolation models. Continuum percolation originated in the work by Gilbert [26]. We refer to Meester and Roy [39] for a rigorous introduction to continuum percolation.

In [26] Gilbert proposed the following random network. Take a Poisson point process X in the Euclidean plane and connect each pair of points x and y of X if the distance between x and y is less than R , for some constant R . This model was introduced as a model of communication networks of short range stations spread over a wide area. The points of the point process represent base stations, and two base stations at locations x and y can communicate to each other if their distance is less than R . The model also has applications in epidemiology, where the points in the network represent infected individuals or herds. If sick individuals infect all others within distance R then the disease spreads along the lines of the network.

Two natural generalizations of the model introduced in [26] are the Boolean model and the random connection model. These two models are the canonical examples of

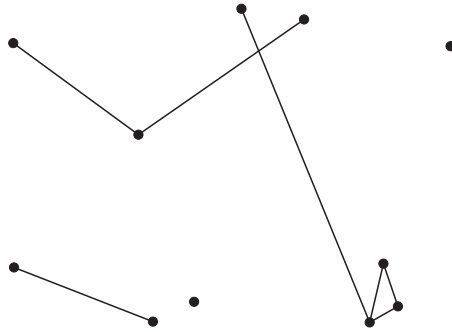


Figure 1.1: Realization of the random connection model.

models in continuum percolation. The *Boolean model* is a random collection of balls in d -dimensional Euclidean space defined as follows. Take a Poisson point process X on \mathbb{R}^d of density λ and, for each point x of X , put a ball centered at x with a random radius according to some probability distribution ρ . The radii of the balls are independent of each other and independent of the point process X . The union of the balls is called the occupied set and its complement the vacant set. Both sets consist of connected components, and connectivity properties of these components have been studied in the literature, see e.g. [39]. Note that if $\rho = R/2$ a.s. then we obtain the model from [26].

The *random connection model* is a random network defined as follows (see Figure 1.1). Take a Poisson point process X on \mathbb{R}^d of density λ , and connect each pair of points x and y of X with probability $g(|x - y|)$, independently of all other pairs of points, independently of X . Here g is a connection function, which is a non-increasing function from the positive reals to $[0, 1]$ that satisfies an integrability condition to avoid trivial cases. Thus, in the random connection model the probability of connecting two vertices x and y decreases as the distance between x and y increases. This allows for a more flexible and more realistic modeling of the networks in telecommunications and epidemiology mentioned above. Note that if $g = \mathbf{1}_R$ then we obtain the model from [26]. Percolation properties of the random connection model such as phase transitions and the existence and uniqueness of an infinite component have been studied in the literature, see e.g. [39].

In Chapter 3, we consider a sequence of random connection models X_n on \mathbb{R}^d , where X_n is a Poisson point process on \mathbb{R}^d of density λ_n , with $\lambda_n/n^d \rightarrow \lambda > 0$. The points of X_n are connected according to the connection function g_n defined by $g_n(x) = g(nx)$ for some connection function g . Let I_n be the number of isolated vertices in the random connection model X_n in some bounded set K . The main result in the paper [44] by Roy and Sarkar is a central limit theorem for I_n . Although the statement of this result is correct, the proof in [44] has errors. We explain what went wrong in the proof, and how this can be corrected. We also prove an extension to connected components larger than a single vertex in case the connection function has bounded support.

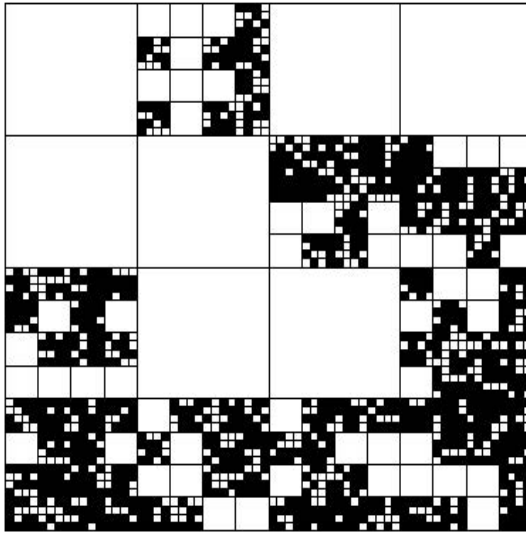


Figure 1.2: Realization of the fractal percolation process, for $d = 2$ and $N = 4$, till iteration step 3, by Matthijs Joosten. Black squares are retained; white squares are removed.

1.3 Fractal percolation

The Boolean model from Section 1.2 has a multiscale analogue called the *multiscale Boolean model* (see [39, 40]). This model is defined as follows. Take a sequence of independent Boolean models X_n on \mathbb{R}^d such that the balls in the Boolean model X_n have radius N^{-n} , for some positive constant N , and the centers of the balls are given by a Poisson point process on \mathbb{R}^d of density λN^{dn} , for some λ . Thus, as n increases, the density of the points in X_n increases and the radius of the balls decreases. The union of all balls of all Boolean models is called the occupied set and its complement the vacant set. Note that the vacant set can be obtained by removing all balls from the space. The vacant set is a *random fractal set*, i.e. a random set that is statistically self-similar in the sense that enlargements of small parts have the same law as the whole set. Indeed, if we scale the model by a factor N^n , for some integer n , then the result is statistically similar to the original model.

The original motivation for the introduction of the multiscale Boolean model was a process introduced by Mandelbrot [38] called *fractal percolation*. The fractal percolation model is a model similar in spirit to the multiscale Boolean model in the sense that random objects are removed from the space via a sequential construction. In fractal percolation the objects removed are not balls but cubes. The centers of the cubes are not formed by the points of a Poisson point process but lie on a grid.

The fractal percolation model is defined as follows (see Figure 1.2). Let $N \geq 2$, $d \geq 2$, and divide the d -dimensional unit cube in N^d subcubes of side length $1/N$. Retain each subcube with probability p and remove it with probability $1 - p$, independently of other subcubes. The closure of the union of the retained subcubes

forms a random subset D^1 of the unit cube. Next, divide each retained subcube in N^d cubes of side length $1/N^2$. Again, retain each smaller subcube with probability p and remove it with probability $1 - p$, independently of other cubes. This gives a random set $D^2 \subset D^1$. Iterating this procedure in every retained cube at every smaller scale yields a decreasing sequence of random closed subsets D^n of the unit cube. The limit set $\bigcap_n D^n$ is a random fractal set having intriguing connectivity properties which have been extensively studied in the literature, see e.g. [5, 16, 20, 24].

It is easy to extend and generalize the fractal percolation model in ways that preserve at least a certain amount of statistical self-similarity and generate random fractal sets. It is interesting to study such models to obtain a better understanding of general fractal percolation processes and explore possible new features that are not present in the model introduced by Mandelbrot [38]. In Chapter 4 we are concerned with two natural extensions which have previously appeared in the literature [17, 18, 21].

The first extension we consider is *k-fractal percolation*. In this model the d -dimensional unit cube is divided in N^d equal subcubes, k of which are retained in a uniform way while the others are removed. The procedure is then iterated inside the retained cubes at all smaller scales. We say that the model percolates if its limit set contains a connected component (in the usual topological sense) that intersects two opposite faces of the unit cube. Clearly, the probability that the model percolates is a non-decreasing function of k . We define the percolation critical value of the model as the minimal k such that the model percolates with positive probability. We show that the (properly rescaled) percolation critical value converges to the critical value of ordinary site percolation on a particular d -dimensional lattice as N tends to infinity. This is analogous to the result of Falconer and Grimmett in [24] that the critical value for Mandelbrot fractal percolation converges to the critical value of site percolation on the same d -dimensional lattice.

The second fractal percolation model we study is *fat fractal percolation*. The construction of the model is similar to Mandelbrot fractal percolation, but in fat fractal percolation the probability p_n of retaining a subcube at iteration step n depends on n . We assume that p_n is non-decreasing in n such that $\prod_n p_n > 0$. The Lebesgue measure of the limit set of the model is positive if the limit set is non-empty. We prove that either the set of connected components larger than one point has Lebesgue measure zero a.s. or its complement in the limit set has Lebesgue measure zero a.s.

1.4 Loop soups

The multiscale Boolean model mentioned in Section 1.3 is statistically self-similar in the sense that enlargements of small parts of the model have the same law as the model at the original scale. Because of the sequential construction of the model, this self-similarity only holds for enlargements by a factor N^n for integer n . There is a way to extend the multiscale Boolean model to a model that is fully scale invariant in the sense that it is self-similar to enlargements by any factor. This is the fully scale invariant Boolean model (see [6]), which is a random collection of countably many balls of different sizes.

In Chapter 5 we study a random fractal in two-dimensional Euclidean space that

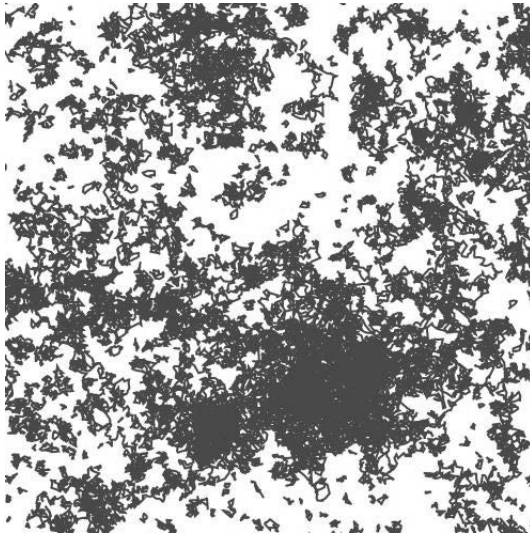


Figure 1.3: Realization of the Brownian loop soup with intensity $\lambda = 1/2$, with a cut-off on small and large time-lengths of the loops, by Matthew Kleban.

is similar to the fully scale invariant Boolean model, in the sense that random objects are removed from the space in a fully scale invariant way according to a Poisson point process. However, the objects removed from the space are not disks but the interiors of Brownian loops, i.e. two-dimensional Brownian bridges with different time-lengths. This random fractal is the *Brownian loop soup* introduced by Lawler and Werner [32] (see Figure 1.3). The Brownian loop soup is not only fully scale invariant but also conformally invariant. It has attracted much attention in the literature because of this property, and because of its relation with the Schramm-Loewner evolution (SLE) [46] and, in particular, with the *conformal loop ensembles* (CLE) (see [48]).

SLE and CLE are random objects in two-dimensional Euclidean space that appear naturally as the scaling limit of the boundaries of clusters of several interesting models of statistical mechanics such as discrete percolation and the Ising model. Indeed, in two dimensions and at the critical point, the scaling limit geometry of the boundaries of clusters of these models is known (see [13–15, 19, 49]) or conjectured (see [28, 50]) to be described by some member of the one-parameter family of Schramm-Loewner evolutions (SLE_κ with $\kappa > 0$) and related conformal loop ensembles (CLE_κ with $8/3 < \kappa < 8$). SLEs can be used to describe the scaling limit of single interfaces; CLEs are collections of loops and are therefore suitable to describe the scaling limit of the collection of all macroscopic cluster boundaries at once.

For $8/3 < \kappa \leq 4$, CLE_κ can be obtained [48] from the Brownian loop soup, as follows. A realization of the Brownian loop soup in a bounded domain D with intensity $\lambda > 0$ is the collection of loops contained in D from a Poisson realization of a conformally invariant intensity measure $\lambda\mu$ (see Chapter 5 for a definition). When $\lambda > 1/2$, there is a unique cluster [48], where a cluster is a maximal collection of loops that intersect each other. When $\lambda \leq 1/2$, the loop soup is composed of disjoint

clusters of loops [48], and the collection of outer boundaries of the outermost loop soup clusters is distributed like a conformal loop ensemble (CLE_κ) [47, 48, 55] with $8/3 < \kappa \leq 4$.

In [31] Lawler and Trujillo Ferreras introduced the *random walk loop soup* as a discrete version of the Brownian loop soup, and showed that, under Brownian scaling, it converges in an appropriate sense to the Brownian loop soup. The authors of [31] focused on individual loops, showing that, with probability going to 1 in the scaling limit, there is a one-to-one correspondence between “large” lattice loops from the random walk loop soup and “large” loops from the Brownian loop soup such that corresponding loops are close.

As explained above, the connection between the Brownian loop soup and SLE/CLE goes through its loop clusters and their boundaries. In view of this observation, it is interesting to investigate whether the random walk loop soup converges to the Brownian loop soup in terms of loop clusters and their boundaries, not just in terms of individual loops, as established in [31]. This is a natural and nontrivial question, due to the complex geometry of the loops involved and of their mutual overlaps.

In Chapter 5, we consider random walk loop soups from which the “vanishingly small” loops have been removed and establish convergence of their clusters and boundaries, in the scaling limit, to the clusters and boundaries of the corresponding Brownian loop soups. In particular, these results imply that the collection of outer boundaries of outermost clusters composed of “large” lattice loops converges to CLE.

Chapter 2

Stochastic domination and weak convergence of conditioned Bernoulli random vectors

This chapter is based on the paper [8] by Broman, Van de Brug, Kager, and Meester.

2.1 Introduction and main results

Let \mathbf{X} and \mathbf{Y} be random vectors on \mathbb{R}^n with respective laws μ and ν . We say that \mathbf{X} is stochastically dominated by \mathbf{Y} , and write $\mathbf{X} \preceq \mathbf{Y}$, if it is possible to define random vectors $\mathbf{U} = (U_1, \dots, U_n)$ and $\mathbf{V} = (V_1, \dots, V_n)$ on a common probability space such the laws of \mathbf{U} and \mathbf{V} are equal to μ and ν , respectively, and $\mathbf{U} \leq \mathbf{V}$ (that is, $U_i \leq V_i$ for all $i \in \{1, \dots, n\}$) with probability 1. In this case, we also write $\mu \preceq \nu$. For instance, when $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ are vectors of n independent Bernoulli random variables with success probabilities p_1, \dots, p_n and q_1, \dots, q_n , respectively, and $0 < p_i \leq q_i < 1$ for $i \in \{1, \dots, n\}$, we have $\mathbf{X} \preceq \mathbf{Y}$.

In this chapter, we consider the *conditional* laws of \mathbf{X} and \mathbf{Y} , conditioned on the total number of successes being at least k , or sometimes also equal to k , for an integer k . In this first section, we will state our main results and provide some intuition. All proofs are deferred to later sections.

Domination issues concerning the conditional law of Bernoulli vectors conditioned on having at least a certain number of successes have come up in the literature a number of times. In [3] and [4], a simplest case has been considered in which $p_i = p$ and $q_i = q$ for some $p < q$. In [4], the conditional domination is used as a tool in the study of random trees.

Here we study such domination issues in great detail and generality. The Bernoulli vectors we consider have the property that the p_i and q_i take only finitely many values, uniformly in the length n of the vectors. The question about stochastic ordering of the

corresponding conditional distributions gives rise to a number of intriguing questions which, as it turns out, can actually be answered. Our main result, Theorem 2.1.8, provides a complete answer to the question with what maximal probability two such conditioned Bernoulli vectors can be ordered in any coupling, when the length of the vectors tends to infinity.

In Section 2.1.1, we will first discuss domination issues for finite vectors \mathbf{X} and \mathbf{Y} as above. In order to deal with domination issues as the length n of the vectors tends to infinity, it will be necessary to first discuss weak convergence of the conditional distribution of a single vector. Section 2.1.2 introduces the framework for dealing with vectors whose lengths tend to infinity, and Section 2.1.3 discusses their weak convergence. Finally, Section 2.1.4 deals with the asymptotic domination issue when $n \rightarrow \infty$.

2.1.1 Stochastic domination of finite vectors

As above, let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be vectors of independent Bernoulli random variables with success probabilities p_1, \dots, p_n and q_1, \dots, q_n , respectively, where $0 < p_i \leq q_i < 1$ for $i \in \{1, \dots, n\}$. For an event A , we shall denote by $\mathcal{L}(\mathbf{X}|A)$ the conditional law of \mathbf{X} given A . Our first proposition states that the conditional law of the total number of successes of \mathbf{X} , conditioned on the event $\{\sum_{i=1}^n X_i \geq k\}$, is stochastically dominated by the conditional law of the total number of successes of \mathbf{Y} .

Proposition 2.1.1. *For all $k \in \{0, 1, \dots, n\}$,*

$$\mathcal{L}(\sum_{i=1}^n X_i | \sum_{i=1}^n X_i \geq k) \preceq \mathcal{L}(\sum_{i=1}^n Y_i | \sum_{i=1}^n Y_i \geq k).$$

In general, the conditional law of the full vector \mathbf{X} is not necessarily stochastically dominated by the conditional law of the vector \mathbf{Y} . For example, consider the case $n = 2$, $p_1 = p_2 = q_1 = p$ and $q_2 = 1 - p$ for some $p < \frac{1}{2}$, and $k = 1$. We then have

$$\begin{aligned} \mathbb{P}(X_1 = 1 | X_1 + X_2 \geq 1) &= \frac{1}{2 - p}, \\ \mathbb{P}(Y_1 = 1 | Y_1 + Y_2 \geq 1) &= \frac{p}{1 - (1 - p)p}. \end{aligned}$$

Hence, if p is small enough, then the conditional law of \mathbf{X} is not stochastically dominated by the conditional law of \mathbf{Y} .

We would first like to study under which conditions we do have stochastic ordering of the conditional laws of \mathbf{X} and \mathbf{Y} . For this, it turns out to be very useful to look at the conditional laws of \mathbf{X} and \mathbf{Y} , conditioned on the total number of successes being *exactly equal* to k , for an integer k . Note that if we condition on the total number of successes being exactly equal to k , then the conditional law of \mathbf{X} is stochastically dominated by the conditional law of \mathbf{Y} if and only if the two conditional laws are equal. The following proposition characterizes stochastic ordering of the conditional laws of \mathbf{X} and \mathbf{Y} in this case. First we define, for $i \in \{1, \dots, n\}$,

$$\beta_i := \frac{p_i}{1 - p_i} \frac{1 - q_i}{q_i}. \quad (2.1.1)$$

The β_i will play a crucial role in the domination issue throughout the chapter.

Proposition 2.1.2. *The following statements are equivalent:*

- (i) All β_i ($i \in \{1, \dots, n\}$) are equal;
- (ii) $\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = k) = \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i = k)$ for all $k \in \{0, 1, \dots, n\}$;
- (iii) $\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = k) = \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i = k)$ for some $k \in \{1, \dots, n-1\}$.

We will use this result to prove the next proposition, which gives a sufficient condition under which the conditional law of \mathbf{X} is stochastically dominated by the conditional law of \mathbf{Y} , in the case when we condition on the total number of successes being at least k .

Proposition 2.1.3. *If all β_i ($i \in \{1, \dots, n\}$) are equal, then for all $k \in \{0, 1, \dots, n\}$,*

$$\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i \geq k) \preceq \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i \geq k).$$

The condition in this proposition is a sufficient condition, not a necessary condition. For example, if $n = 2$, $p_1 = p_2 = \frac{1}{2}$, $q_1 = \frac{6}{10}$ and $q_2 = \frac{7}{10}$, then $\beta_1 \neq \beta_2$, but we do have stochastic ordering for all $k \in \{0, 1, 2\}$.

2.1.2 Framework for asymptotic domination

Suppose that we now extend our Bernoulli random vectors \mathbf{X} and \mathbf{Y} to infinite sequences X_1, X_2, \dots and Y_1, Y_2, \dots of independent Bernoulli random variables, which we assume to have only finitely many distinct success probabilities. It then seems natural to let \mathbf{X}_n and \mathbf{Y}_n denote the n -dimensional vectors (X_1, \dots, X_n) and (Y_1, \dots, Y_n) , respectively, and consider the domination issue as $n \rightarrow \infty$, where we condition on the total number of successes being at least $k_n = \lfloor \alpha n \rfloor$ for some fixed number $\alpha \in (0, 1)$.

More precisely, with k_n as above, let $\tilde{\mathbf{X}}_n$ be a random vector having the law $\mathcal{L}(\mathbf{X}_n | \sum_{i=1}^n X_i \geq k_n)$, and define $\tilde{\mathbf{Y}}_n$ similarly. Proposition 2.1.3 gives a sufficient condition under which $\tilde{\mathbf{X}}_n$ is stochastically dominated by $\tilde{\mathbf{Y}}_n$ for each $n \geq 1$. If this condition is not fulfilled, however, we might still be able to define random vectors \mathbf{U} and \mathbf{V} , with the same laws as $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$, on a common probability space such that the probability that $\mathbf{U} \leq \mathbf{V}$ is high (perhaps even 1). We denote by

$$\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \tag{2.1.2}$$

the supremum over all possible couplings (\mathbf{U}, \mathbf{V}) of $(\tilde{\mathbf{X}}_n, \tilde{\mathbf{Y}}_n)$ of the probability that $\mathbf{U} \leq \mathbf{V}$. We want to study the asymptotic behavior of this quantity as $n \rightarrow \infty$.

As an example (and an appetizer for what is to come), consider the following situation. For $i \geq 1$ let the random variable X_i have success probability p for some $p \in (0, \frac{1}{2})$. For $i \geq 1$ odd or even let the random variable Y_i have success probability p or $1 - p$, respectively. We will prove that $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ converges to a constant as $n \rightarrow \infty$ (Theorem 2.1.8 below). It turns out that there are three possible values of the limit, depending on the value of α :

- (i) If $\alpha < p$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 1$.
- (ii) If $\alpha = p$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow \frac{3}{4}$.

(iii) If $\alpha > p$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 0$.

In fact, to study the asymptotic domination issue, we will work in an even more general framework, which we shall describe now. For every $n \geq 1$, \mathbf{X}_n is a vector of n independent Bernoulli random variables. We assume that this vector is organized in M “blocks”, such that all Bernoulli variables in block i have the same success probability p_i , for $i \in \{1, \dots, M\}$. Similarly, \mathbf{Y}_n is a vector of n independent Bernoulli random variables with the exact same block structure as \mathbf{X}_n , but for \mathbf{Y}_n , the success probability corresponding to block i is q_i , where $0 < p_i \leq q_i < 1$ as before.

For given $n \geq 1$ and $i \in \{1, \dots, M\}$, we denote by m_{in} the size of block i , where of course $\sum_{i=1}^M m_{in} = n$. In the example above, there were two blocks, each containing (roughly) one half of the Bernoulli variables, and the size of each block was increasing with n . In the general framework, we only assume that the fractions m_{in}/n converge to some number $\alpha_i \in (0, 1)$ as $n \rightarrow \infty$, where $\sum_{i=1}^M \alpha_i = 1$. Similarly, in the example above we conditioned on the total number of successes being at least k_n , where $k_n = \lfloor \alpha n \rfloor$ for some fixed $\alpha \in (0, 1)$. In the general framework, we only assume that we are given a fixed sequence of integers k_n such that $0 \leq k_n \leq n$ for all $n \geq 1$ and $k_n/n \rightarrow \alpha \in (0, 1)$ as $n \rightarrow \infty$.

In this general framework, let $\tilde{\mathbf{X}}_n$ be a random vector having the conditional distribution of \mathbf{X}_n , conditioned on the total number of successes being at least k_n . Observe that given the number of successes in a particular block, these successes are uniformly distributed within the block. Hence, the distribution of $\tilde{\mathbf{X}}_n$ is completely determined by the distribution of the M -dimensional vector describing the numbers of successes per block. Therefore, before we proceed to study the asymptotic behavior of the quantity (2.1.2), we shall first study the asymptotic behavior of this M -dimensional vector.

2.1.3 Weak convergence

Consider the general framework introduced in the previous section. We define X_{in} as the number of successes of the vector \mathbf{X}_n in block i and write $\Sigma_n := \sum_{i=1}^M X_{in}$ for the total number of successes in \mathbf{X}_n . Then X_{in} has a binomial distribution with parameters m_{in} and p_i and, for fixed n , the X_{in} are independent. In this section, we shall study the joint convergence in distribution of the X_{in} as $n \rightarrow \infty$, conditioned on $\{\Sigma_n \geq k_n\}$, and also conditioned on $\{\Sigma_n = k_n\}$.

First we consider the case where we condition on $\{\Sigma_n = k_n\}$. We will prove (Lemma 2.3.1 below) that the X_{in} concentrate around the values $c_{in}m_{in}$, where the c_{in} are determined by the system of equations

$$\begin{cases} \frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} = \frac{1 - c_{jn}}{c_{jn}} \frac{p_j}{1 - p_j} & \forall i, j \in \{1, \dots, M\}; \\ \sum_{i=1}^M c_{in} m_{in} = k_n. \end{cases} \quad (2.1.3)$$

We will show in Section 2.3 that the system (2.1.3) has a unique solution and that

$$c_{in} \rightarrow c_i \quad \text{as } n \rightarrow \infty,$$

for some c_i strictly between 0 and 1. As we shall see, each component X_{in} is roughly normally distributed around the central value $c_{in}m_{in}$, with fluctuations around this

centre of the order \sqrt{n} . Hence, the proper scaling is obtained by looking at the M -dimensional vector

$$\boldsymbol{\mathcal{X}}_n := \left(\frac{X_{1n} - c_{1n}m_{1n}}{\sqrt{n}}, \frac{X_{2n} - c_{2n}m_{2n}}{\sqrt{n}}, \dots, \frac{X_{Mn} - c_{Mn}m_{Mn}}{\sqrt{n}} \right). \quad (2.1.4)$$

Since we condition on $\{\Sigma_n = k_n\}$, this vector is essentially an $(M-1)$ -dimensional vector, taking only values in the hyperplane

$$S_0 := \{(z_1, \dots, z_M) \in \mathbb{R}^M : z_1 + \dots + z_M = 0\}.$$

However, we want to view it as an M -dimensional vector, mainly because when we later condition on $\{\Sigma_n \geq k_n\}$, $\boldsymbol{\mathcal{X}}_n$ will no longer be restricted to a hyperplane. One expects that the laws of the $\boldsymbol{\mathcal{X}}_n$ converge weakly to a distribution which concentrates on S_0 and is, therefore, singular with respect to M -dimensional Lebesgue measure. To facilitate this, it is natural to define a measure ν_0 on the Borel sets of \mathbb{R}^M through

$$\nu_0(\cdot) := \lambda_0(\cdot \cap S_0), \quad (2.1.5)$$

where λ_0 denotes $((M-1)$ -dimensional) Lebesgue measure on S_0 , and to identify the weak limit of the $\boldsymbol{\mathcal{X}}_n$ via a density with respect to ν_0 . The density of the weak limit is given by the function $f: \mathbb{R}^M \rightarrow \mathbb{R}$ defined by

$$f(z) = \mathbb{1}_{S_0}(z) \prod_{i=1}^M \exp\left(-\frac{z_i^2}{2c_i(1-c_i)\alpha_i}\right). \quad (2.1.6)$$

Theorem 2.1.4. *The laws $\mathcal{L}(\boldsymbol{\mathcal{X}}_n | \Sigma_n = k_n)$ converge weakly to the measure which has density $f / \int f d\nu_0$ with respect to ν_0 .*

We now turn to the case where we condition on $\{\Sigma_n \geq k_n\}$. Our strategy will be to first study the case where we condition on the event $\{\Sigma_n = k_n + \ell\}$, for $\ell \geq 0$, and then sum over ℓ . We will calculate the relevant range of ℓ to sum over. In particular, we will show that for large enough ℓ the probability $\mathbb{P}(\Sigma_n = k_n + \ell)$ is so small, that these ℓ do not have a significant effect on the conditional distribution of $\boldsymbol{\mathcal{X}}_n$. For k_n sufficiently larger than $\mathbb{E}(\Sigma_n)$, only ℓ of order $o(\sqrt{n})$ are relevant, which leads to the following result:

Theorem 2.1.5. *If $\alpha > \sum_{i=1}^M p_i \alpha_i$ or, more generally, $(k_n - \mathbb{E}(\Sigma_n)) / \sqrt{n} \rightarrow \infty$, then the laws $\mathcal{L}(\boldsymbol{\mathcal{X}}_n | \Sigma_n \geq k_n)$ also converge weakly to the measure which has density $f / \int f d\nu_0$ with respect to ν_0 .*

Finally, we consider the case where we condition on $\{\Sigma_n \geq k_n\}$ with k_n below or around $\mathbb{E}(\Sigma_n)$, that is, when $(k_n - \mathbb{E}(\Sigma_n)) / \sqrt{n} \rightarrow K \in [-\infty, \infty)$. An essential difference compared to the situation in Theorem 2.1.5, is that the probabilities of the events $\{\Sigma_n \geq k_n\}$ do not converge to 0 in this case, but to a strictly positive constant. In this situation, the right vector to look at is the M -dimensional vector

$$\boldsymbol{\mathcal{X}}_n^p := \left(\frac{X_{1n} - p_1 m_{1n}}{\sqrt{n}}, \frac{X_{2n} - p_2 m_{2n}}{\sqrt{n}}, \dots, \frac{X_{Mn} - p_M m_{Mn}}{\sqrt{n}} \right).$$

It follows from standard arguments that the unconditional laws of \mathcal{X}_n^p converge weakly to a multivariate normal distribution with density $h/\int h d\lambda$ with respect to M -dimensional Lebesgue measure λ , where $h: \mathbb{R}^M \rightarrow \mathbb{R}$ is given by

$$h(z) = \prod_{i=1}^M \exp\left(-\frac{z_i^2}{2p_i(1-p_i)\alpha_i}\right). \quad (2.1.7)$$

If k_n stays sufficiently smaller than $\mathbb{E}(\Sigma_n)$, that is, when $K = -\infty$, then the effect of conditioning vanishes in the limit, and the conditional laws of \mathcal{X}_n^p given $\{\Sigma_n \geq k_n\}$ converge weakly to the same limit as the unconditional laws of \mathcal{X}_n^p . In general, if $K \in [-\infty, \infty)$, the conditional laws of \mathcal{X}_n^p given $\{\Sigma_n \geq k_n\}$ converge weakly to the measure which has, up to a normalizing constant, density h restricted to the half-space

$$H_K := \{(z_1, \dots, z_M) \in \mathbb{R}^M : z_1 + \dots + z_M \geq K\}. \quad (2.1.8)$$

Theorem 2.1.6. *If $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n} \rightarrow K$ for some $K \in [-\infty, \infty)$, then the laws $\mathcal{L}(\mathcal{X}_n^p | \Sigma_n \geq k_n)$ converge weakly to the measure which has density $h \mathbb{1}_{H_K} / \int h \mathbb{1}_{H_K} d\lambda$ with respect to λ .*

Remark 2.1.7. If $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n}$ does not converge as $n \rightarrow \infty$ and does not diverge to either ∞ or $-\infty$, then the laws $\mathcal{L}(\mathcal{X}_n^p | \Sigma_n \geq k_n)$ do not converge weakly either. This follows from our results above by considering limits along different subsequences of the k_n .

2.1.4 Asymptotic stochastic domination

Consider again the general framework for vectors \mathbf{X}_n and \mathbf{Y}_n introduced in Section 2.1.2. Recall that we write $\tilde{\mathbf{X}}_n$ for a random vector having the conditional distribution of the vector \mathbf{X}_n , given that the total number of successes is at least k_n . For $n \geq 1$ and $i \in \{1, \dots, M\}$, we let \tilde{X}_{in} denote the number of successes of $\tilde{\mathbf{X}}_n$ in block i . We define $\tilde{\mathbf{Y}}_n$ and \tilde{Y}_{in} analogously. We want to study the asymptotic behavior as $n \rightarrow \infty$ of the quantity

$$\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n),$$

where the supremum is taken over all possible couplings of $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$.

Define β_i for $i \in \{1, \dots, M\}$ as in (2.1.1). As a first observation, note that if all β_i are equal, then by Proposition 2.1.3 we have $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) = 1$ for every $n \geq 1$. Otherwise, under certain conditions on the sequence k_n , $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ will converge to a constant as $n \rightarrow \infty$, as we shall prove.

The intuitive picture behind this is as follows. Without conditioning, $\mathbf{X}_n \leq \mathbf{Y}_n$ for every $n \geq 1$. Now, as long as k_n stays significantly smaller than $\mathbb{E}(\Sigma_n)$, the effect of conditioning will vanish in the limit, and hence we can expect that $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 1$ as $n \rightarrow \infty$. Suppose now that we start making the k_n larger. This will increase the number of successes \tilde{X}_{in} of the vector $\tilde{\mathbf{X}}_n$ in each block i , but as long as k_n stays below the expected total number of successes of \mathbf{Y}_n , increasing k_n will not change the numbers of successes per block significantly for the vector $\tilde{\mathbf{Y}}_n$.

At some point, when k_n becomes large enough, there will be a block i such that \tilde{X}_{in} becomes roughly equal to \tilde{Y}_{in} . We shall see that this happens for k_n “around” the value \hat{k}_n defined by

$$\hat{k}_n := \sum_{i=1}^M \frac{p_i m_{in}}{p_i + \beta_{\max}(1 - p_i)},$$

where $\beta_{\max} := \max\{\beta_1, \dots, \beta_M\}$. Therefore, the sequence \hat{k}_n will play a key role in our main result. What will happen is that as long as k_n stays significantly smaller than \hat{k}_n , \tilde{X}_{in} stays significantly smaller than \tilde{Y}_{in} for each block i , and hence $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n) \rightarrow 1$ as $n \rightarrow \infty$. For k_n around \hat{k}_n there is a “critical window” in which interesting things occur. Namely, when $(k_n - \hat{k}_n)/\sqrt{n}$ converges to a finite constant K , $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n)$ converges to a constant P_K which is strictly between 0 and 1. Finally, when k_n is sufficiently larger than \hat{k}_n , there will always be a block i such that \tilde{X}_{in} is significantly larger than \tilde{Y}_{in} . Hence, $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n) \rightarrow 0$ in this case.

Before we state our main theorem which makes this picture precise, let us first define the non-trivial constant P_K which occurs as the limit of $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n)$ when k_n is in the critical window. To this end, let

$$I := \{i \in \{1, \dots, M\} : \beta_i = \beta_{\max}\},$$

and define positive numbers a , b and c by

$$a^2 = \sum_{i \in I} \frac{\beta_{\max} p_i (1 - p_i) \alpha_i}{(p_i + \beta_{\max}(1 - p_i))^2} = \sum_{i \in I} q_i (1 - q_i) \alpha_i; \quad (2.1.9a)$$

$$b^2 = \sum_{i \notin I} \frac{\beta_{\max} p_i (1 - p_i) \alpha_i}{(p_i + \beta_{\max}(1 - p_i))^2}; \quad (2.1.9b)$$

$$c^2 = a^2 + b^2. \quad (2.1.9c)$$

As we shall see later, these numbers will come up as variances of certain normal distributions. Let $\Phi: \mathbb{R} \rightarrow (0, 1)$ denote the distribution function of the standard normal distribution. For $K \in \mathbb{R}$, define P_K by

$$P_K = \begin{cases} 1 - \int_{-\infty}^{\frac{c-b}{ac}K} \frac{e^{-z^2/2} \Phi\left(\frac{K-az}{b}\right) - \Phi\left(\frac{K}{c}\right)}{\sqrt{2\pi} (1 - \Phi\left(\frac{K}{c}\right))} dz & \text{if } \alpha = \sum_{i=1}^M p_i \alpha_i, \\ \Phi\left(\frac{bK}{ac} - \frac{1}{a}R_K\right) + \Phi\left(-\frac{K}{a} + \frac{b}{ac}R_K\right) & \text{if } \alpha > \sum_{i=1}^M p_i \alpha_i. \end{cases} \quad (2.1.10)$$

where $R_K = \sqrt{K^2 + c^2 \log(c^2/b^2)}$. It will be made clear in Section 2.4 where these formulas for P_K come from. We will show that P_K is strictly between 0 and 1. In fact, it is possible to show that both expressions for P_K are strictly decreasing in K from 1 to 0, but we omit the (somewhat lengthy) derivation of this fact here.

Theorem 2.1.8. *If all β_i ($i \in \{1, \dots, M\}$) are equal, then we have that $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n) = 1$ for every $n \geq 1$. Otherwise, the following holds:*

(i) *If $(k_n - \hat{k}_n)/\sqrt{n} \rightarrow -\infty$, then $\sup \mathbb{P}(\tilde{X}_n \leq \tilde{Y}_n) \rightarrow 1$.*

(ii) If $(k_n - \hat{k}_n)/\sqrt{n} \rightarrow K$ for some $K \in \mathbb{R}$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow P_K$.

(iii) If $(k_n - \hat{k}_n)/\sqrt{n} \rightarrow \infty$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 0$.

Remark 2.1.9. If $\beta_i \neq \beta_j$ for some $i \neq j$, and $(k_n - \hat{k}_n)/\sqrt{n}$ does not converge as $n \rightarrow \infty$ and does not diverge to either ∞ or $-\infty$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ does not converge either. This follows from the strict monotonicity of P_K , by considering the limits along different subsequences of the k_n .

To demonstrate Theorem 2.1.8, recall the example from Section 2.1.2. Here $\beta_{\max} = 1$, $\hat{k}_n = pn$, $I = \{1\}$ and $a^2 = b^2 = \frac{1}{2}p(1-p)$. If $\alpha = p$, then we have that $(k_n - \hat{k}_n)/\sqrt{n} \rightarrow 0$ as $n \rightarrow \infty$. Hence, by Theorem 2.1.8, $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ converges to

$$P_0 = 1 - 2 \int_{-\infty}^0 \frac{e^{-z^2/2}}{\sqrt{2\pi}} (\Phi(-z) - 1/2) dz = \frac{3}{4}.$$

In fact, Theorem 2.1.8 shows that we can obtain any value between 0 and 1 for the limit by adding $[K\sqrt{n}]$ successes to k_n , for $K \in \mathbb{R}$.

Next we turn to the proofs of our results. Results in Section 2.1.1 are proved in Section 2.2, results in Section 2.1.3 are proved in Section 2.3 and finally, results in Section 2.1.4 are proved in Section 2.4.

2.2 Stochastic domination of finite vectors

Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be vectors of independent Bernoulli random variables with success probabilities p_1, \dots, p_n and q_1, \dots, q_n respectively, where $0 < p_i \leq q_i < 1$ for $i \in \{1, \dots, n\}$.

Suppose that $p_i = p$ for all i . Then $\sum_{i=1}^n X_i$ has a binomial distribution with parameters n and p . The quotient

$$\frac{\mathbb{P}(\sum_{i=1}^n X_i = k+1)}{\mathbb{P}(\sum_{i=1}^n X_i = k)} = \frac{n-k}{k+1} \frac{p}{1-p}$$

is strictly increasing in p and strictly decreasing in k , and it is also easy to see that

$$\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = k) \preceq \mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = k+1).$$

The following two lemmas show that these two properties hold for general success probabilities p_1, \dots, p_n .

Lemma 2.2.1. For $k \in \{0, 1, \dots, n-1\}$, consider the quotients

$$Q_k^n := \frac{\mathbb{P}(\sum_{i=1}^n X_i = k+1)}{\mathbb{P}(\sum_{i=1}^n X_i = k)} \quad (2.2.1)$$

and

$$\frac{\mathbb{P}(\sum_{i=1}^n X_i \geq k+1)}{\mathbb{P}(\sum_{i=1}^n X_i \geq k)}. \quad (2.2.2)$$

Both (2.2.1) and (2.2.2) are strictly increasing in p_1, \dots, p_n for fixed k , and strictly decreasing in k for fixed p_1, \dots, p_n .

Proof. We only give the proof for (2.2.1), since the proof for (2.2.2) is similar. First we will prove that Q_k^n is strictly increasing in p_1, \dots, p_n for fixed k . By symmetry, it suffices to show that Q_k^n is strictly increasing in p_1 . We show this by induction on n . The base case $n = 1$, $k = 0$ is immediate. Next note that for $n \geq 2$ and $k \in \{0, \dots, n-1\}$,

$$\begin{aligned} Q_k^n &= \frac{\mathbb{P}(\sum_{i=1}^{n-1} X_i = k)p_n + \mathbb{P}(\sum_{i=1}^{n-1} X_i = k+1)(1-p_n)}{\mathbb{P}(\sum_{i=1}^{n-1} X_i = k-1)p_n + \mathbb{P}(\sum_{i=1}^{n-1} X_i = k)(1-p_n)} \\ &= \frac{p_n + Q_k^{n-1}(1-p_n)}{p_n/Q_{k-1}^{n-1} + (1-p_n)}, \end{aligned}$$

which is strictly increasing in p_1 by the induction hypothesis (in the case $k = n-1$, use $Q_k^{n-1} = 0$, and in the case $k = 0$, use $1/Q_{k-1}^{n-1} = 0$).

To prove that Q_k^n is strictly decreasing in k for fixed p_1, \dots, p_n , note that since Q_k^n is strictly increasing in p_n for fixed $k \in \{1, \dots, n-2\}$, we have

$$0 < \frac{\partial}{\partial p_n} Q_k^n = \frac{\partial}{\partial p_n} \frac{p_n + Q_k^{n-1}(1-p_n)}{p_n/Q_{k-1}^{n-1} + (1-p_n)} = \frac{1 - Q_k^{n-1}/Q_{k-1}^{n-1}}{(p_n/Q_{k-1}^{n-1} + (1-p_n))^2}.$$

Hence, $Q_k^{n-1} < Q_{k-1}^{n-1}$. This argument applies for any $n \geq 2$. \square

Let $\mathbf{X}^k = (X_1^k, \dots, X_n^k)$ have the conditional law of \mathbf{X} , conditioned on the event $\{\sum_{i=1}^n X_i = k\}$. Our next lemma gives an explicit coupling of the \mathbf{X}^k in which they are ordered. The existence of such a coupling was already proved in [27, Proposition 6.2], but our explicit construction is new and of independent value. In our construction, we freely regard \mathbf{X}^k as a random subset of $\{1, \dots, n\}$ by identifying \mathbf{X}^k with $\{i \in \{1, \dots, n\} : X_i^k = 1\}$. For any $K \subset \{1, \dots, n\}$, let $\{\mathbf{X}_K = \mathbf{1}\}$ denote the event $\{X_i = 1 \forall i \in K\}$, and for any $I \subset \{1, \dots, n\}$ and $j \in \{1, \dots, n\}$, define

$$\gamma_{j,I} := \sum_{L \subset \{1, \dots, n\} : |L|=|I|+1} \frac{\mathbb{1}(j \in L)}{|L \setminus I|} \mathbb{P}(\mathbf{X}_L = \mathbf{1} \mid \sum_{i=1}^n X_i = |I| + 1).$$

Lemma 2.2.2. *For any $I \subset \{1, \dots, n\}$, the collection $\{\gamma_{j,I}\}_{j \in \{1, \dots, n\} \setminus I}$ is a probability vector. Moreover, if I is picked according to \mathbf{X}^k and then j is picked according to $\{\gamma_{j,I}\}_{j \in \{1, \dots, n\} \setminus I}$, the resulting set $J = \{I, j\}$ has the same distribution as if it was picked according to \mathbf{X}^{k+1} . Therefore, we can couple the sequence $\{\mathbf{X}^k\}_{k=1}^n$ such that $\mathbb{P}(\mathbf{X}^1 \leq \mathbf{X}^2 \leq \dots \leq \mathbf{X}^{n-1} \leq \mathbf{X}^n) = 1$.*

Proof. Throughout the proof, I, J, K and L denote subsets of $\{1, \dots, n\}$, and we simplify notation by writing $\Sigma_n := \sum_{i=1}^n X_i$. First observe that

$$\sum_{j \notin I} \gamma_{j,I} = \sum_{L : |L|=|I|+1} \mathbb{P}(\mathbf{X}_L = \mathbf{1} \mid \Sigma_n = |I| + 1) = 1,$$

which proves that the $\{\gamma_{j,I}\}_{j \notin I}$ form a probability vector, since $\gamma_{j,I} \geq 0$.

Next note that for any K containing j ,

$$\frac{\mathbb{P}(\mathbf{X}_K = \mathbf{1} \mid \Sigma_n = |K|)}{\mathbb{P}(\mathbf{X}_{K \setminus \{j\}} = \mathbf{1} \mid \Sigma_n = |K| - 1)} = \frac{\mathbb{P}(X_j = 1) \mathbb{P}(\Sigma_n = |K| - 1)}{\mathbb{P}(X_j = 0) \mathbb{P}(\Sigma_n = |K|)}. \quad (2.2.3)$$

Now fix J , and for $j \in J$, let $I = I(j, J) = J \setminus \{j\}$. Then for $j \in J$, by (2.2.3),

$$\begin{aligned} \gamma_{j,I} &= \frac{\mathbb{P}(\mathbf{X}_J = \mathbf{1} \mid \Sigma_n = |J|)}{\mathbb{P}(\mathbf{X}_I = \mathbf{1} \mid \Sigma_n = |I|)} \sum_{L: |L|=|J|} \frac{\mathbf{1}(j \in L)}{|L \setminus I|} \mathbb{P}(\mathbf{X}_{L \setminus \{j\}} = \mathbf{1} \mid \Sigma_n = |I|) \\ &= \frac{\mathbb{P}(\mathbf{X}_J = \mathbf{1} \mid \Sigma_n = |J|)}{\mathbb{P}(\mathbf{X}_I = \mathbf{1} \mid \Sigma_n = |I|)} \sum_{K: |K|=|I|} \frac{\mathbf{1}(j \notin K)}{|J \setminus K|} \mathbb{P}(\mathbf{X}_K = \mathbf{1} \mid \Sigma_n = |I|), \end{aligned}$$

where the second equality follows upon writing $K = L \setminus \{j\}$, and using $|L \setminus I| = |L \setminus J| + 1 = |K \setminus J| + 1 = |J \setminus K|$ in the sum. Hence, by summing first over j and then over K , we obtain

$$\sum_{j \in J} \gamma_{j,I} \mathbb{P}(\mathbf{X}_I = \mathbf{1} \mid \Sigma_n = |I|) = \mathbb{P}(\mathbf{X}_J = \mathbf{1} \mid \Sigma_n = |J|). \quad \square$$

Corollary 2.2.3. *For $k \in \{0, 1, \dots, n-1\}$ we have*

$$\mathcal{L}(\mathbf{X} \mid \sum_{i=1}^n X_i \geq k) \leq \mathcal{L}(\mathbf{X} \mid \sum_{i=1}^n X_i \geq k+1).$$

Proof. Using Lemma 2.2.2, we will construct random vectors \mathbf{U} and \mathbf{V} on a common probability space such that \mathbf{U} and \mathbf{V} have the conditional distributions of \mathbf{X} given $\{\sum_{i=1}^n X_i \geq k\}$ and \mathbf{X} given $\{\sum_{i=1}^n X_i \geq k+1\}$, respectively, and $\mathbf{U} \leq \mathbf{V}$ with probability 1.

First pick an integer m according to the conditional law of $\sum_{i=1}^n X_i$ given $\{\sum_{i=1}^n X_i \geq k\}$. If $m \geq k+1$, then pick \mathbf{U} according to the conditional law of \mathbf{X} given $\{\sum_{i=1}^n X_i = m\}$, and set $\mathbf{V} = \mathbf{U}$. If $m = k$, then first pick an integer $m+\ell$ according to the conditional law of $\sum_{i=1}^n X_i$ given $\{\sum_{i=1}^n X_i \geq k+1\}$. Next, pick \mathbf{U} and \mathbf{V} such that \mathbf{U} and \mathbf{V} have the conditional laws of \mathbf{X} given $\{\sum_{i=1}^n X_i = m\}$ and \mathbf{X} given $\{\sum_{i=1}^n X_i = m+\ell\}$, respectively, and $\mathbf{U} \leq \mathbf{V}$. This is possible by Lemma 2.2.2. By construction, $\mathbf{U} \leq \mathbf{V}$ with probability 1, and a little computation shows that \mathbf{U} and \mathbf{V} have the desired marginal distributions. \square

Now we are in a position to prove Propositions 2.1.1, 2.1.2 and 2.1.3.

Proof of Proposition 2.1.1. By Lemma 2.2.1 we have that for $\ell \in \{1, \dots, n-k\}$,

$$\frac{\mathbb{P}(\sum_{i=1}^n X_i \geq k+\ell)}{\mathbb{P}(\sum_{i=1}^n X_i \geq k)} = \prod_{j=0}^{\ell-1} \frac{\mathbb{P}(\sum_{i=1}^n X_i \geq k+j+1)}{\mathbb{P}(\sum_{i=1}^n X_i \geq k+j)}$$

is strictly increasing in p_1, \dots, p_n . This implies that for $\ell \in \{1, \dots, n-k\}$,

$$\mathbb{P}(\sum_{i=1}^n X_i \geq k+\ell \mid \sum_{i=1}^n X_i \geq k) \leq \mathbb{P}(\sum_{i=1}^n Y_i \geq k+\ell \mid \sum_{i=1}^n Y_i \geq k). \quad \square$$

Proof of Proposition 2.1.2. Let $x, y \in \{0, 1\}^n$ be such that $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and let $k = \sum_{i=1}^n x_i$. Write $I = \{i \in \{1, \dots, n\} : x_i = 1\}$ and, likewise, $J = \{i \in \{1, \dots, n\} : y_i = 1\}$, and recall the definition (2.1.1) of β_i . We have

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{X} = x \mid \sum_{i=1}^n X_i = k)}{\mathbb{P}(\mathbf{X} = y \mid \sum_{i=1}^n X_i = k)} &= \frac{\prod_{i \in I} p_i \prod_{i \notin I} (1-p_i)}{\prod_{i \in J} p_i \prod_{i \notin J} (1-p_i)} \\ &= \prod_{i \in I \setminus J} \frac{p_i}{1-p_i} \prod_{i \in J \setminus I} \frac{1-p_i}{p_i} = \frac{\prod_{i \in I \setminus J} \beta_i \mathbb{P}(\mathbf{Y} = x \mid \sum_{i=1}^n Y_i = k)}{\prod_{i \in J \setminus I} \beta_i \mathbb{P}(\mathbf{Y} = y \mid \sum_{i=1}^n Y_i = k)}. \end{aligned} \quad (2.2.4)$$

Since $|I| = |J| = k$, we have $|I \setminus J| = |J \setminus I|$. Hence, (i) implies (ii), and (ii) trivially implies (iii). To show that (iii) implies (i), suppose that $\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = k) = \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i = k)$ for a given $k \in \{1, \dots, n-1\}$. Let $i \in \{2, \dots, n\}$ and let K be a subset of $\{2, \dots, n\} \setminus \{i\}$ with exactly $k-1$ elements. Choosing $I = \{1\} \cup K$ and $J = K \cup \{i\}$ in (2.2.4) yields $\beta_i = \beta_1$. \square

Proof of Proposition 2.1.3. By Proposition 2.1.2 and Lemma 2.2.2, we have for $m \in \{0, 1, \dots, n\}$ and $\ell \in \{0, 1, \dots, n-m\}$

$$\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i = m) \preceq \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i = m + \ell).$$

Using this result and Proposition 2.1.1, we will construct random vectors \mathbf{U} and \mathbf{V} on a common probability space such that \mathbf{U} and \mathbf{V} have the conditional distributions of \mathbf{X} given $\{\sum_{i=1}^n X_i \geq k\}$ and \mathbf{Y} given $\{\sum_{i=1}^n Y_i \geq k\}$, respectively, and $\mathbf{U} \leq \mathbf{V}$ with probability 1.

First, pick integers m and $m + \ell$ such that they have the conditional laws of $\sum_{i=1}^n X_i$ given $\{\sum_{i=1}^n X_i \geq k\}$ and $\sum_{i=1}^n Y_i$ given $\{\sum_{i=1}^n Y_i \geq k\}$, respectively, and $m \leq m + \ell$ with probability 1. Secondly, pick \mathbf{U} and \mathbf{V} such that they have the conditional laws of \mathbf{X} given $\{\sum_{i=1}^n X_i = m\}$ and \mathbf{Y} given $\{\sum_{i=1}^n Y_i = m + \ell\}$, respectively, and $\mathbf{U} \leq \mathbf{V}$ with probability 1. A little computation shows that the vectors \mathbf{U} and \mathbf{V} have the desired marginal distributions. \square

We close this section with a minor result, which gives a condition under which we do not have stochastic ordering.

Proposition 2.2.4. *If $p_i = q_i$ for some $i \in \{1, \dots, n\}$ but not for all i , then for $k \in \{1, \dots, n-1\}$,*

$$\mathcal{L}(\mathbf{X} | \sum_{i=1}^n X_i \geq k) \not\preceq \mathcal{L}(\mathbf{Y} | \sum_{i=1}^n Y_i \geq k).$$

Proof. Without loss of generality, assume that $p_n = q_n$. We have

$$\begin{aligned} & \mathbb{P}(X_n = 1 | \sum_{i=1}^n X_i \geq k) \\ &= \frac{p_n \mathbb{P}(\sum_{i=1}^{n-1} X_i \geq k-1)}{p_n \mathbb{P}(\sum_{i=1}^{n-1} X_i \geq k-1) + (1-p_n) \mathbb{P}(\sum_{i=1}^{n-1} X_i \geq k)} \\ &= \frac{p_n}{p_n + (1-p_n) \mathbb{P}(\sum_{i=1}^{n-1} X_i \geq k) / \mathbb{P}(\sum_{i=1}^{n-1} X_i \geq k-1)} \\ &> \frac{q_n}{q_n + (1-q_n) \mathbb{P}(\sum_{i=1}^{n-1} Y_i \geq k) / \mathbb{P}(\sum_{i=1}^{n-1} Y_i \geq k-1)} \\ &= \mathbb{P}(Y_n = 1 | \sum_{i=1}^n Y_i \geq k), \end{aligned}$$

where the strict inequality follows from Lemma 2.2.1. \square

2.3 Weak convergence

We now turn to the framework for asymptotic domination described in Section 2.1.2 and to the setting of Section 2.1.3. Recall that X_{in} is the number of successes of the vector \mathbf{X}_n in block i . We want to study the joint convergence in distribution of the X_{in} as $n \rightarrow \infty$, conditioned on $\{\Sigma_n \geq k_n\}$, and also conditioned on $\{\Sigma_n = k_n\}$. Since we are interested in the limit $n \rightarrow \infty$, we may assume from the outset that the values of n we consider are so large that k_n and all m_{in} are strictly between 0 and n , to avoid degenerate situations.

We will first consider the case where we condition on the event $\{\Sigma_n = k_n\}$. Lemma 2.3.1 below states that the X_{in} will then concentrate around the values $c_{in}m_{in}$, where the c_{in} are determined by the system of equations (2.1.3), which we repeat here for the convenience of the reader:

$$\begin{cases} \frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} = \frac{1 - c_{jn}}{c_{jn}} \frac{p_j}{1 - p_j} & \forall i, j \in \{1, \dots, M\}; \\ \sum_{i=1}^M c_{in} m_{in} = k_n. \end{cases} \quad (2.1.3)$$

Before we turn to the proof of this concentration result, let us first look at the system (2.1.3) in more detail. If we write

$$A_n = \frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} \quad (2.3.1)$$

for the desired common value for all i , then

$$c_{in} = \frac{p_i}{p_i + A_n(1 - p_i)}.$$

Note that this is equal to 1 for $A_n = 0$ and to p_i for $A_n = 1$, and strictly decreasing to 0 as $A_n \rightarrow \infty$, so that there is a unique $A_n > 0$ such that

$$\sum_{i=1}^M c_{in} m_{in} = \sum_{i=1}^M \frac{p_i m_{in}}{p_i + A_n(1 - p_i)} = k_n. \quad (2.3.2)$$

It follows that the system (2.1.3) does have a unique solution, characterized by this value of A_n . Moreover, it follows from (2.3.2) that if $k_n > \mathbb{E}(\Sigma_n) = \sum_{i=1}^M p_i m_{in}$, then $A_n < 1$. Furthermore, $k_n/n \rightarrow \alpha$ and $m_{in}/n \rightarrow \alpha_i$. Hence, by dividing both sides in (2.3.2) by n , and taking the limit $n \rightarrow \infty$, we see that the A_n converge to the unique positive number A such that

$$\sum_{i=1}^M \frac{p_i \alpha_i}{p_i + A(1 - p_i)} = \alpha,$$

where $A = 1$ if $\alpha = \sum_{i=1}^M p_i \alpha_i$. As a consequence, we also have that

$$c_{in} \rightarrow c_i = \frac{p_i}{p_i + A(1 - p_i)} \quad \text{as } n \rightarrow \infty.$$

Note that the c_i are the unique solution to the system of equations

$$\begin{cases} \frac{1-c_i}{c_i} \frac{p_i}{1-p_i} = \frac{1-c_j}{c_j} \frac{p_j}{1-p_j} & \forall i, j \in \{1, \dots, M\}; \\ \sum_{i=1}^M c_i \alpha_i = \alpha. \end{cases}$$

Observe also that $c_i = p_i$ in case $A = 1$, or equivalently $\sum_{i=1}^M p_i \alpha_i = \alpha$, which is the case when the total number of successes k_n is within $o(n)$ of the mean $\mathbb{E}(\Sigma_n)$. The concentration result:

Lemma 2.3.1. *Let c_{1n}, \dots, c_{Mn} satisfy (2.1.3). Then for each i and all positive integers r , we have that*

$$\mathbb{P}(|X_{in} - c_{in}m_{in}| \geq Mr \mid \Sigma_n = k_n) \leq 2Me^{-(M-1)r^2/n}.$$

Proof. The idea of the proof is as follows. Condition on $\{\Sigma_n = k_n\}$, and consider the event that for some $i \neq j$ we have that $X_{in} = c_{in}m_{in} + s$, and $X_{jn} = c_{jn}m_{jn} - t$, for some positive numbers s and t . We will show that if the c_{in} satisfy (2.1.3), the event obtained by increasing X_{in} by 1 and decreasing X_{jn} by 1 has smaller probability. This establishes that the conditional distribution of the X_{in} is maximal at the central values $c_{in}m_{in}$ identified by the system (2.1.3). The precise bound in Lemma 2.3.1 also follows from the argument.

Now for the details. Let s and t be nonnegative real numbers such that $c_{in}m_{in} + s$ and $c_{jn}m_{jn} - t$ are integers. By the binomial distributions of X_{in} and X_{jn} and their independence, if it is the case that $0 \leq c_{in}m_{in} + s < m_{in}$ and $0 < c_{jn}m_{jn} - t \leq m_{jn}$, then

$$\begin{aligned} & \frac{\mathbb{P}(X_{in} = c_{in}m_{in} + s + 1, X_{jn} = c_{jn}m_{jn} - t - 1)}{\mathbb{P}(X_{in} = c_{in}m_{in} + s, X_{jn} = c_{jn}m_{jn} - t)} \\ &= \left(\frac{m_{in} - c_{in}m_{in} - s}{c_{in}m_{in} + s + 1} \frac{p_i}{1-p_i} \right) \left(\frac{c_{jn}m_{jn} - t}{m_{jn} - c_{jn}m_{jn} + t + 1} \frac{1-p_j}{p_j} \right) \\ &\leq \left(\frac{m_{in} - c_{in}m_{in} - s}{c_{in}m_{in}} \frac{p_i}{1-p_i} \right) \left(\frac{c_{jn}m_{jn} - t}{m_{jn} - c_{jn}m_{jn}} \frac{1-p_j}{p_j} \right). \end{aligned}$$

Hence, if the c_{in} satisfy (2.1.3), then using $1 - z \leq \exp(-z)$ we obtain

$$\begin{aligned} & \frac{\mathbb{P}(X_{in} = c_{in}m_{in} + s + 1, X_{jn} = c_{jn}m_{jn} - t - 1)}{\mathbb{P}(X_{in} = c_{in}m_{in} + s, X_{jn} = c_{jn}m_{jn} - t)} \\ &\leq \left(1 - \frac{s}{m_{in} - c_{in}m_{in}} \right) \left(1 - \frac{t}{c_{jn}m_{jn}} \right) \leq \exp\left(-\frac{s+t}{n}\right). \end{aligned}$$

It follows by iteration of this inequality, that for *all* real $s, t \geq 0$ and all integers $u \geq 0$,

$$\begin{aligned} & \mathbb{P}(X_{in} = c_{in}m_{in} + s + u, X_{jn} = c_{jn}m_{jn} - t - u) \\ &\leq \exp\left(-\frac{(s+t)u}{n}\right) \mathbb{P}(X_{in} = c_{in}m_{in} + s, X_{jn} = c_{jn}m_{jn} - t). \end{aligned} \quad (2.3.3)$$

Now fix i , and observe that for all integers $r > 0$,

$$\begin{aligned} \mathbb{P}(X_{in} \geq c_{in}m_{in} + Mr, \Sigma_n = k_n) \\ = \sum_{\substack{\ell_1, \dots, \ell_M \in \mathbb{N}_0: \\ \ell_1 + \dots + \ell_M = k_n}} \mathbb{1}(\ell_i \geq c_{in}m_{in} + Mr) \mathbb{P}(X_{kn} = \ell_k \ \forall k). \end{aligned}$$

But if $\ell_1 + \dots + \ell_M = k_n$ and $\ell_i \geq c_{in}m_{in} + Mr$, then there must be some $j \neq i$ such that $\ell_j \leq c_{jn}m_{jn} - r$. Therefore,

$$\begin{aligned} \mathbb{P}(X_{in} \geq c_{in}m_{in} + Mr, \Sigma_n = k_n) \\ \leq \sum_{j=1}^M \sum_{\substack{\ell_1, \dots, \ell_M \in \mathbb{N}_0: \\ \ell_1 + \dots + \ell_M = k_n}} \mathbb{1} \left(\begin{array}{l} \ell_i \geq c_{in}m_{in} + Mr \\ \ell_j \leq c_{jn}m_{jn} - r \end{array} \right) \mathbb{P}(X_{kn} = \ell_k \ \forall k). \end{aligned}$$

By independence of the X_{in} and using (2.3.3) with $s = (M-1)r$, $t = 0$ and $u = r$, we now obtain

$$\begin{aligned} \mathbb{P}(X_{in} \geq c_{in}m_{in} + Mr, \Sigma_n = k_n) \\ \leq e^{-(M-1)r^2/n} \sum_{j=1}^M \sum_{\substack{\ell_1, \dots, \ell_M \in \mathbb{N}_0: \\ \ell_1 + \dots + \ell_M = k_n}} \mathbb{1} \left(\begin{array}{l} \ell_i \geq c_{in}m_{in} + Mr - r \\ \ell_j \leq c_{jn}m_{jn} \end{array} \right) \mathbb{P}(X_{kn} = \ell_k \ \forall k) \\ \leq M e^{-(M-1)r^2/n} \mathbb{P}(\Sigma_n = k_n). \end{aligned}$$

This proves that

$$\mathbb{P}(X_{in} \geq c_{in}m_{in} + Mr \mid \Sigma_n = k_n) \leq M e^{-(M-1)r^2/n}.$$

Similarly, one can prove that

$$\mathbb{P}(X_{in} \leq c_{in}m_{in} - Mr \mid \Sigma_n = k_n) \leq M e^{-(M-1)r^2/n}. \quad \square$$

As we have already mentioned, we expect that the X_{in} have fluctuations around their centres of the order \sqrt{n} . It is therefore natural to look at the M -dimensional vector

$$\boldsymbol{\mathcal{X}}_n := \left(\frac{X_{1n} - x_{1n}}{\sqrt{n}}, \frac{X_{2n} - x_{2n}}{\sqrt{n}}, \dots, \frac{X_{Mn} - x_{Mn}}{\sqrt{n}} \right), \quad (2.3.4)$$

where the vector $x_n = (x_{1n}, \dots, x_{Mn})$ represents the centre around which the X_{in} concentrate. To prove weak convergence of $\boldsymbol{\mathcal{X}}_n$, we will not set x_{in} equal to $c_{in}m_{in}$, because the latter numbers are not necessarily integer, and it will be more convenient if the x_{in} are integers. So instead, for each fixed n , we choose the x_{in} to be nonnegative integers such that $|x_{in} - c_{in}m_{in}| < 1$ for all i , and $\sum_{i=1}^M x_{in} = k_n$. Of course, the vector $\boldsymbol{\mathcal{X}}_n$ as it is defined in (2.3.4), and the vector defined in (2.1.4) have the same weak limit. In our proofs of Theorems 2.1.4 and 2.1.5, $\boldsymbol{\mathcal{X}}_n$ will refer to the vector defined in (2.3.4).

If we condition on $\{\Sigma_n = k_n\}$, then the vector $\boldsymbol{\mathcal{X}}_n$ will only take values in the hyperplane

$$S_0 := \{(z_1, \dots, z_M) \in \mathbb{R}^M : z_1 + \dots + z_M = 0\}.$$

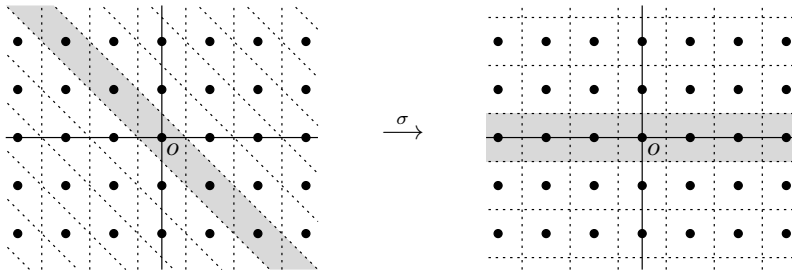


Figure 2.1: The shear transformation σ (illustrated here for $M = 2$) maps sheared cubes to cubes. The dots are the sites of the integer lattice \mathbb{Z}^2 . The gray band on the left encompasses those sheared cubes that intersect S_0 .

However, as we have already explained in the introduction, we still regard \mathcal{X}_n as an M -dimensional vector, because we will also condition on $\{\Sigma_n \geq k_n\}$, in which case \mathcal{X}_n is not restricted to a hyperplane. To deal with this, it turns out that for technical reasons which will become clear later, it is useful to introduce the projection $\pi: (z_1, \dots, z_M) \mapsto (z_1, \dots, z_{M-1})$ and the shear transformation $\sigma: (z_1, \dots, z_M) \mapsto (z_1, \dots, z_{M-1}, z_1 + \dots + z_M)$. We can then define a metric ρ on \mathbb{R}^M by setting $\rho(x, y) := |\sigma x - \sigma y|$, where $|\cdot|$ denotes Euclidean distance. See Figure 2.1 for an illustration.

Using the projection π , we now define a new measure μ_0 on the Borel subsets of \mathbb{R}^M , which is concentrated on S_0 , by

$$\mu_0(\cdot) := \lambda^{M-1}(\pi(\cdot \cap S_0)),$$

where λ^{M-1} is the ordinary Lebesgue measure on \mathbb{R}^{M-1} . Note that up to a multiplicative constant, μ_0 is equal to the measure ν_0 defined in Section 2.3, so we could have stated Theorems 2.1.4 and 2.1.5 equally well with μ_0 instead of ν_0 . In the proofs it turns out to be more convenient to work with μ_0 , however, so that is what we shall do.

Our proofs of Theorems 2.1.4 and 2.1.5 resemble classical arguments to prove weak convergence of random vectors living on a lattice via a local limit theorem and Scheffé's theorem, see for instance [1, Theorem 3.3]. However, we cannot use these classic results here, for two reasons. First of all, in Theorem 2.1.5 our random vectors live on an M -dimensional lattice, but in the limit all the mass collapses onto a lower-dimensional hyperplane, leading to a weak limit which is singular with respect to M -dimensional Lebesgue measure. The classic arguments do not cover this case of a singular limit.

Secondly, we are considering *conditioned* random vectors, for which it is not so obvious how to obtain a local limit theorem directly. Our solution is to get rid of the conditioning by considering *ratios* of conditioned probabilities, and prove a local limit theorem for these ratios. An extra argument will then be needed to prove weak convergence. Since we cannot resort to classic arguments here, we have to go through the proofs in considerable detail.

2.3.1 Proof of Theorem 2.1.4

As we have explained above, the key idea in the proof of Theorem 2.1.4 is that we can get rid of the awkward conditioning by considering *ratios* of conditional probabilities, rather than the conditional probabilities themselves. Thus, we will be dealing with ratios of binomial probabilities, and the following lemma addresses the key properties of these ratios needed in the proof. The lemma resembles standard bounds on binomial probabilities, but we point out that here we are considering *ratios* of binomial probabilities which centre around $c_{in}m_{in}$ rather than around the mean p_im_{in} . We also note that actually, the lemma is stronger than required to prove Theorem 2.1.4, but we will need this stronger result to prove Theorem 2.1.5 later.

Lemma 2.3.2. *Recall the definition (2.3.1) of A_n . Fix $i \in \{1, 2, \dots, M\}$ and let b_1, b_2, \dots be a sequence of positive integers such that $b_n/\sqrt{n} \rightarrow 0$ as $n \rightarrow \infty$. Then, for every $z \in \mathbb{R}$,*

$$\sup_{\substack{x: |x-x_{in}| < b_n \\ r: |r-z\sqrt{n}| < b_n}} \left| \frac{1}{A_n^r} \frac{\mathbb{P}(X_{in} = x+r)}{\mathbb{P}(X_{in} = x)} - \exp\left(-\frac{z^2}{2c_i(1-c_i)\alpha_i}\right) \right| \rightarrow 0.$$

Furthermore, there exist constants $B_i^1, B_i^2 < \infty$ such that for all n and r ,

$$\sup_{x: |x-x_{in}| < b_n} \frac{1}{A_n^r} \frac{\mathbb{P}(X_{in} = x+r)}{\mathbb{P}(X_{in} = x)} \leq B_i^1 \left(1 + \frac{r^4}{n^2}\right) \exp\left(B_i^2 \frac{|r|}{\sqrt{n}} - \frac{1}{2} \frac{r^2}{n}\right).$$

Proof. Robbins' note on Stirling's formula [43] states that for all $m = 1, 2, \dots$,

$$\sqrt{2\pi} m^{m+1/2} e^{-m+1/(12m+1)} < m! < \sqrt{2\pi} m^{m+1/2} e^{-m+1/(12m)},$$

from which it is straightforward to show that for all $m = 0, 1, 2, \dots$ (so including $m = 0$), there exists an η_m satisfying $1/7 < \eta_m < 1/5$ such that

$$m! = \sqrt{2\pi(m+\eta_m)} m^m e^{-m} = \sqrt{2\pi\llbracket m \rrbracket} m^m e^{-m}, \quad (2.3.5)$$

where we have introduced the notation $\llbracket m \rrbracket := m + \eta_m$.

Since X_{in} has the binomial distribution with parameters m_{in} and p_i ,

$$\frac{1}{A_n^r} \frac{\mathbb{P}(X_{in} = x+r)}{\mathbb{P}(X_{in} = x)} = \frac{x!}{(x+r)!} \frac{(m_{in}-x)!}{(m_{in}-x-r)!} \left(\frac{c_{in}}{1-c_{in}}\right)^r.$$

Using (2.3.5), we can write this as the product of the three factors

$$\begin{aligned} P_{in}^1(x, r) &= \left(\frac{\llbracket x \rrbracket}{\llbracket x+r \rrbracket} \frac{\llbracket m_{in}-x \rrbracket}{\llbracket m_{in}-x-r \rrbracket} \right)^{1/2} \\ P_{in}^2(x, r) &= \left(\frac{c_{in}m_{in}}{x} \frac{m_{in}-x}{m_{in}-c_{in}m_{in}} \right)^r \\ P_{in}^3(x, r) &= \left(\frac{x}{x+r} \right)^{x+r} \left(\frac{m_{in}-x}{m_{in}-x-r} \right)^{m_{in}-x-r} \end{aligned}$$

for all x and r such that $0 < x < m_{in}$ and $0 \leq x + r \leq m_{in}$.

To study the convergence of $P_{in}^3(x, r)$, first write

$$P_{in}^3(x, r) = \left(1 - \frac{r}{x+r}\right)^{x+r} \left(1 + \frac{r}{m_{in} - x - r}\right)^{m_{in} - x - r}.$$

Using the fact that for all $u > -1$, $(1+u)$ lies between $\exp(u - \frac{1}{2}u^2)$ and $\exp(u - \frac{1}{2}u^2/(1+u))$, a little computation now shows that $P_{in}^3(x, r)$ is wedged in between

$$\exp\left(-\frac{1}{2} \frac{(m_{in} - r)r^2}{x(m_{in} - x - r)}\right) \quad \text{and} \quad \exp\left(-\frac{1}{2} \frac{(m_{in} + r)r^2}{(x+r)(m_{in} - x)}\right).$$

From this fact, it follows that for fixed $z \in \mathbb{R}$,

$$\sup_{\substack{x: |x-x_{in}| < b_n \\ r: |r-z\sqrt{n}| < b_n}} \left| P_{in}^3(x, r) - \exp\left(-\frac{z^2}{2c_i(1-c_i)\alpha_i}\right) \right| \rightarrow 0,$$

because $x_{in}/m_{in} \rightarrow c_i$, hence $x = c_i m_{in} + o(n)$ and $r = z\sqrt{n} + o(\sqrt{n})$ under the supremum, and $m_{in}/n \rightarrow \alpha_i$. Since $|x_{in} - c_{in}m_{in}| < 1$, we also have that

$$\sup_{\substack{x: |x-x_{in}| < b_n \\ r: |r-z\sqrt{n}| < b_n}} |P_{in}^1(x, r) - 1| \rightarrow 0 \quad \text{and} \quad \sup_{\substack{x: |x-x_{in}| < b_n \\ r: |r-z\sqrt{n}| < b_n}} |P_{in}^2(x, r) - 1| \rightarrow 0.$$

Together with the uniform convergence of $P_{in}^3(x, r)$, this establishes the first part of Lemma 2.3.2.

We now turn to the second part of the lemma. If x and r are such that $0 < x < m_{in}$ and $0 \leq x + r \leq m_{in}$, then $m_{in} - r \geq x > 0$ and $m_{in} + r \geq m_{in} - x > 0$, hence from the bounds on $P_{in}^3(x, r)$ given in the previous paragraph we can conclude that

$$P_{in}^3(x, r) \leq \exp\left(-\frac{1}{2} \frac{r^2}{m_{in}}\right) \leq \exp\left(-\frac{1}{2} \frac{r^2}{n}\right).$$

Next observe that if x is such that $|x - x_{in}| < b_n$, then $|x - c_{in}m_{in}| < 1 + b_n$, from which it follows that uniformly in n , for all x and r such that $0 < x < m_{in}$, $0 \leq x + r \leq m_{in}$ and $|x - x_{in}| < b_n$,

$$P_{in}^2(x, r) \leq \left(1 + \text{const.} \times \frac{b_n}{n}\right)^{|r|} \leq \exp\left(\text{const.} \times \frac{|r|}{\sqrt{n}}\right).$$

To finish the proof, it remains to bound $P_{in}^1(x, r)$. To this end, observe first that uniformly in n , for all x and r such that $|x - x_{in}| < b_n$ and $|r| < n^{3/4}$, $P_{in}^1(x, r)$ is bounded by a constant. On the other hand, uniformly for all x and r such that $0 < x < m_{in}$ and $0 \leq x + r \leq m_{in}$, $P_{in}^1(x, r)$ is bounded by a constant times n , and $n \leq r^4/n^2$ if $|r| \geq n^{3/4}$. Combining these observations, we see that uniformly in n , for all x and r satisfying $|x - x_{in}| < b_n$ and $0 \leq x + r \leq m_{in}$,

$$P_{in}^1(x, r) \leq \text{const.} \times \left(1 + \frac{r^4}{n^2}\right). \quad \square$$

Proof of Theorem 2.1.4. For a point z in \mathbb{R}^M , let $\lceil z \rceil$ be the point in \mathbb{Z}^M ρ -closest to z (take the lexicographically smallest one if there is a choice). Graphically, this means that the collection of those points z for which $\lceil z \rceil = a$ comprises the sheared cube $a + \sigma^{-1}(-1/2, 1/2]^M$, see Figure 2.1. Now, for each fixed $z \in \mathbb{R}^M$, set $r_n^z = (r_{1n}^z, \dots, r_{Mn}^z) := \lceil z\sqrt{n} \rceil$. Observe that because (for fixed n) the x_{in} sum to k_n , if $r_n^z \in S_0$ we have that

$$\frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z \mid \Sigma_n = k_n)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n = 0 \mid \Sigma_n = k_n)} = \frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n = 0)} = \prod_{i=1}^M \frac{\mathbb{P}(X_{in} = x_{in} + r_{in}^z)}{\mathbb{P}(X_{in} = x_{in})}, \quad (2.3.6)$$

where we have used the independence of the components X_{in} . If $r_n^z \notin S_0$, on the other hand, this ratio obviously vanishes.

We now apply Lemma 2.3.2 to (2.3.6), taking $b_n = M$ for every $n \geq 1$. Since $\sum_{i=1}^M r_{in}^z = 0$ if $r_n^z \in S_0$ and hence $\prod_{i=1}^M A_n^{r_{in}^z} = 1$, the first part of Lemma 2.3.2 immediately implies that for all $z \in \mathbb{R}^M$,

$$\frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z \mid \Sigma_n = k_n)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n = 0 \mid \Sigma_n = k_n)} \rightarrow \mathbb{1}_{S_0}(z) \prod_{i=1}^M \exp\left(-\frac{z_i^2}{2c_i(1-c_i)\alpha_i}\right) = f(z)$$

as $n \rightarrow \infty$. To see how this will lead to Theorem 2.1.4, define $f_n: \mathbb{R}^M \rightarrow \mathbb{R}$ by

$$f_n(z) := (\sqrt{n})^M \mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z \mid \Sigma_n = k_n).$$

Then f_n is a probability density function with respect to M -dimensional Lebesgue measure λ . Moreover, if \mathcal{Z}_n is a random vector with this density, then the vector $\mathcal{Z}'_n = \lceil \mathcal{Z}_n \sqrt{n} \rceil / \sqrt{n}$ has the same distribution as the vector \mathcal{X}_n , conditioned on $\{\Sigma_n = k_n\}$. Since clearly \mathcal{Z}_n and \mathcal{Z}'_n must have the same weak limit, it is therefore sufficient to show that the weak limit of \mathcal{Z}_n has density $f / \int f d\mu_0$ with respect to μ_0 .

Now, by what we have established above, we already know that

$$\frac{f_n(z)}{f_n(0)} = \frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z \mid \Sigma_n = k_n)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n = 0 \mid \Sigma_n = k_n)} \rightarrow f(z) \quad \text{for every } z \in \mathbb{R}^M.$$

Moreover, the second part of Lemma 2.3.2 applied to (2.3.6) shows that the ratios $f_n(z)/f_n(0)$ are uniformly bounded by some μ_0 -integrable function $g(z)$. Thus it follows by dominated convergence that for every Borel set $A \subset \mathbb{R}^M$,

$$\int_A \frac{f_n(z)}{f_n(0)} d\mu_0(z) \rightarrow \int_A f(z) d\mu_0(z).$$

Next observe that $1 = \int f_n d\lambda = \int n^{-1/2} f_n d\mu_0$, because by the conditioning, f_n is nonzero only on the sheared cubes which intersect S_0 . Therefore, taking $A = \mathbb{R}^M$ in the previous equation yields $n^{-1/2} f_n(0) \rightarrow (\int f d\mu_0)^{-1}$, which in turn implies that for every Borel set A ,

$$\int_A n^{-1/2} f_n(z) d\mu_0(z) \rightarrow \frac{\int_A f(z) d\mu_0(z)}{\int f d\mu_0}.$$

In general, $\int_F f_n d\lambda \neq \int_F n^{-1/2} f_n d\mu_0$ for an arbitrary Borel set F , but we have equality here for sufficiently large n if F is a finite union of sheared cubes. Hence,

if A is open, we can approximate A from the inside by unions of sheared cubes contained in A to conclude that

$$\liminf_{n \rightarrow \infty} \int_A f_n(z) d\lambda(z) \geq \frac{\int_A f(z) d\mu_0(z)}{\int f d\mu_0}. \quad \square$$

2.3.2 Proof of Theorem 2.1.5

We now turn to the case where we condition on $\{\Sigma_n \geq k_n\}$, for the same fixed sequence $k_n \rightarrow \infty$ as before. To treat this case, we are going to consider what happens when we condition on the event that $\Sigma_n = k_n + \ell$ for some $\ell \geq 0$, and later sum over ℓ . It will be important for us to know the relevant range of ℓ to sum over. In particular, for large enough ℓ we expect that the probability $\mathbb{P}(\Sigma_n = k_n + \ell)$ will be so small, that these ℓ will not influence the conditional distribution of the vector \mathcal{X}_n in an essential way. The relevant range of ℓ can be determined from the following lemma:

Lemma 2.3.3. *For all positive integers s ,*

$$\mathbb{P}(\Sigma_n \geq k_n + 2Ms) \leq M \exp\left(-\frac{(k_n - \mathbb{E}(\Sigma_n) + Ms)s}{Mn}\right) \mathbb{P}(\Sigma_n \geq k_n).$$

Proof. Let u be such that $0 < u < (1 - p_i)m_{in}$. Observe that then, for all integers m such that $p_i m_{in} + u \leq m \leq m_{in}$,

$$\frac{\mathbb{P}(X_{in} = m + 1)}{\mathbb{P}(X_{in} = m)} = \frac{m_{in} - m}{m + 1} \frac{p_i}{1 - p_i} \leq \frac{p_i m_{in} - u \frac{p_i}{1 - p_i}}{p_i m_{in} + u},$$

hence

$$\frac{\mathbb{P}(X_{in} = m + 1)}{\mathbb{P}(X_{in} = m)} \leq 1 - \frac{u}{p_i m_{in} + u} \left(1 + \frac{p_i}{1 - p_i}\right) \leq 1 - \frac{u}{m_{in}} \leq 1 - \frac{u}{n}.$$

Since $1 - z \leq \exp(-z)$, by repeated application of this inequality it follows that for all $u > 0$ and all positive integers t , if m is an integer such that $m \geq p_i m_{in} + u$, then

$$\mathbb{P}(X_{in} = m + t) \leq \exp\left(-\frac{ut}{n}\right) \mathbb{P}(X_{in} = m). \quad (2.3.7)$$

Now observe that if $\Sigma_n \geq \mathbb{E}(\Sigma_n) + Mr + 2Ms$, where s is a positive integer, and r a real number such that $r + s > 0$, then for some k it must be the case that $X_{kn} \geq p_k m_{kn} + r + 2s$. Therefore,

$$\begin{aligned} & \mathbb{P}(\Sigma_n \geq \mathbb{E}(\Sigma_n) + Mr + 2Ms) \\ & \leq \sum_{\substack{\ell_1, \dots, \ell_M \in \mathbb{N}_0: \\ \ell_1 + \dots + \ell_M \geq \mathbb{E}(\Sigma_n) + Mr + 2Ms}} \sum_{k=1}^M \mathbb{1}(\ell_k \geq p_k m_{kn} + r + 2s) \mathbb{P}(X_{in} = \ell_i \ \forall i). \end{aligned}$$

But by (2.3.7), taking $u = r + s$ and $t = s$,

$$\begin{aligned} & \mathbb{1}(\ell_k \geq p_k m_{kn} + r + 2s) \mathbb{P}(X_{in} = \ell_i \ \forall i) \\ & \leq \exp\left(-\frac{(r + s)s}{n}\right) \mathbb{P}(X_{kn} = \ell_k - s, X_{in} = \ell_i \ \forall i \neq k), \end{aligned}$$

and therefore

$$\begin{aligned} \mathbb{P}(\Sigma_n \geq \mathbb{E}(\Sigma_n) + Mr + 2Ms) & \\ & \leq M \exp\left(-\frac{(r+s)s}{n}\right) \mathbb{P}(\Sigma_n \geq \mathbb{E}(\Sigma_n) + Mr + 2Ms - s) \\ & \leq M \exp\left(-\frac{(r+s)s}{n}\right) \mathbb{P}(\Sigma_n \geq \mathbb{E}(\Sigma_n) + Mr). \end{aligned}$$

Choosing r such that $k_n \equiv \mathbb{E}(\Sigma_n) + Mr$ yields Lemma 2.3.3 (observe that the bound holds trivially if $r + s \leq 0$). \square

Lemma 2.3.3 shows that if $\alpha > \sum_{i=1}^M p_i \alpha_i$, then for sufficiently large n , $\mathbb{P}(\Sigma_n \geq k_n + \ell)$ will already be much smaller than $\mathbb{P}(\Sigma_n \geq k_n)$ when ℓ is of order $\log n$. However, when $\alpha = \sum_{i=1}^M p_i \alpha_i$, we need to consider ℓ of bigger order than \sqrt{n} for $\mathbb{P}(\Sigma_n \geq k_n + \ell)$ to become much smaller than $\mathbb{P}(\Sigma_n \geq k_n)$. In either case, Lemma 2.3.3 shows that ℓ of larger order than \sqrt{n} become irrelevant.

Keeping this in mind, we will now look at the conditional distribution of the vector \mathbf{X}_n , conditioned on $\{\Sigma_n = k_n + \ell\}$. The first thing to observe is that for $\ell > 0$, the locations of the centres around which the components X_{in} concentrate will be shifted to larger values. Indeed, these centres are located at $c_{in}^\ell m_{in}$, where the c_{in}^ℓ are of course determined by the system of equations

$$\begin{cases} \frac{1 - c_{in}^\ell}{c_{in}^\ell} \frac{p_i}{1 - p_i} = \frac{1 - c_{jn}^\ell}{c_{jn}^\ell} \frac{p_j}{1 - p_j} & \forall i, j \in \{1, \dots, M\}; \\ \sum_{i=1}^M c_{in}^\ell m_{in} = k_n + \ell. \end{cases} \quad (2.3.8)$$

To find an explicit expression for the size of the shift $c_{in}^\ell - c_{in}$, we can substitute $c_{in}^\ell = c_{in} + \delta_{in}$ into (2.3.8), and then perform an expansion in powers of the correction δ_{in} to guess this correction to first order. This procedure leads us to believe that c_{in}^ℓ must be of the form

$$c_{in}^\ell = c_{in} + c_{in}(1 - c_{in})d_n^\ell + e_{in}^\ell, \quad (2.3.9)$$

where

$$d_n^\ell := \frac{\ell}{\sum_{j=1}^M c_{jn}(1 - c_{jn})m_{jn}},$$

and e_{in}^ℓ should be a higher-order correction. The following lemma shows that the error terms e_{in}^ℓ are indeed of second order in d_n^ℓ , so that the effective shift in c_{in} by adding ℓ extra successes to our Bernoulli variables is given by $c_{in}(1 - c_{in})d_n^\ell$. For convenience, we assume in the lemma that $|d_n^\ell| \leq 1/2$, which means that $|\ell|$ cannot be too large, but by Lemma 2.3.3, this does not put too severe a restriction on the range of ℓ we can consider later.

Lemma 2.3.4. *For all ℓ (positive or negative) such that $|d_n^\ell| \leq 1/2$, we have that $|e_{in}^\ell| \leq (d_n^\ell)^2$ for all $i = 1, \dots, M$.*

Proof. For ease of notation, write $\sigma_{in} := c_{in}(1 - c_{in})$. As before, we write

$$A_n^\ell = \frac{1 - c_{in}^\ell}{c_{in}^\ell} \frac{p_i}{1 - p_i} = \frac{1 - c_{in} - \sigma_{in}d_n^\ell - e_{in}^\ell}{c_{in} + \sigma_{in}d_n^\ell + e_{in}^\ell} \frac{p_i}{1 - p_i}$$

for the desired common value for all i , so

$$e_{in}^\ell = \frac{p_i(1 - c_{in} - \sigma_{in}d_n^\ell) - A_n^\ell(1 - p_i)(c_{in} + \sigma_{in}d_n^\ell)}{A_n^\ell(1 - p_i) + p_i}. \quad (2.3.10)$$

As before, the value of A_n^ℓ is uniquely determined by the requirement $\sum_{i=1}^M c_{in}^\ell m_{in} = k_n + \ell$. Since $\sum_{i=1}^M c_{in} m_{in} = k_n$ and $\sum_{i=1}^M \sigma_{in} d_n^\ell m_{in} = \ell$, this requirement says that

$$\sum_{i=1}^M e_{in}^\ell m_{in} = 0.$$

In particular, the e_{in}^ℓ cannot be all positive or all negative, from which we derive, using (2.3.10), that A_n^ℓ must satisfy the double inequalities

$$\min_{i=1, \dots, M} \left\{ \frac{p_i(1 - c_{in} - \sigma_{in}d_n^\ell)}{(1 - p_i)(c_{in} + \sigma_{in}d_n^\ell)} \right\} \leq A_n^\ell \leq \max_{i=1, \dots, M} \left\{ \frac{p_i(1 - c_{in} - \sigma_{in}d_n^\ell)}{(1 - p_i)(c_{in} + \sigma_{in}d_n^\ell)} \right\}.$$

A simple calculation establishes that

$$\frac{p_i(1 - c_{in} - \sigma_{in}d_n^\ell)}{(1 - p_i)(c_{in} + \sigma_{in}d_n^\ell)} = \frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} \left(1 + \sum_{k=1}^{\infty} \frac{-(1 - c_{in})d_n^{\ell k}}{1 - c_{in}} \right),$$

from which (using $|d_n^\ell| \leq 1/2$) we can conclude that

$$\frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} (1 - d_n^\ell) \leq A_n^\ell \leq \frac{1 - c_{in}}{c_{in}} \frac{p_i}{1 - p_i} (1 - d_n^\ell + 2(d_n^\ell)^2),$$

since by (2.1.3), neither the lower bound nor the upper bound here depends on i .

Inserting the lower bound on A_n^ℓ into (2.3.10) gives

$$e_{in}^\ell \leq \frac{\sigma_{in}(1 - c_{in})(d_n^\ell)^2}{1 - (1 - c_{in})d_n^\ell} \leq \frac{1}{2}(d_n^\ell)^2,$$

where in the last step we used that $|d_n^\ell| \leq 1/2$ and $\sigma_{in} \leq 1/4$. Likewise, substituting the upper bound on A_n^ℓ into (2.3.10) yields

$$e_{in}^\ell \geq -\frac{\sigma_{in}(1 + c_{in})(d_n^\ell)^2 + 2\sigma_{in}(1 - c_{in})(d_n^\ell)^3}{1 - (1 - c_{in})d_n^\ell + 2(1 - c_{in})(d_n^\ell)^2} \geq -\frac{2\sigma_{in}(d_n^\ell)^2}{1 - 1/2} \geq -(d_n^\ell)^2. \quad \square$$

For future use, we state the following corollary:

Corollary 2.3.5. *If $(k_n - \sum_{i=1}^M c_i m_{in})/\sqrt{n} \rightarrow K$ for some $K \in [-\infty, \infty]$, then for $i \in \{1, \dots, M\}$,*

$$\frac{(c_{in} - c_i)m_{in}}{\sqrt{n}} \rightarrow \frac{c_i(1 - c_i)\alpha_i}{\sum_{j=1}^M c_j(1 - c_j)\alpha_j} K.$$

Remark 2.3.6. If $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n} \rightarrow K \in \mathbb{R}$, then $\alpha = \sum_{i=1}^M p_i \alpha_i$ and we have $c_i = p_i$ for all $i \in \{1, \dots, M\}$. In this situation, Corollary 2.3.5 states that the vectors $\mathcal{X}_n^p - \mathcal{X}_n$, and hence also the same vectors conditioned on $\{\Sigma_n \geq k_n\}$, converge pointwise to the vector whose i -th component is

$$\frac{p_i(1 - p_i)\alpha_i}{\sum_{j=1}^M p_j(1 - p_j)\alpha_j} K.$$

Proof of Corollary 2.3.5. First, suppose that $K \in \mathbb{R}$. If $\ell = \sum_{i=1}^M c_i m_{in} - k_n$ and the c_{in}^ℓ satisfy (2.3.8), then $c_{in}^\ell = c_i$. Hence, by Lemma 2.3.4,

$$c_i - c_{in} = c_{in}(1 - c_{in})d_n^\ell + O((d_n^\ell)^2),$$

where

$$d_n^\ell = \frac{\sum_{i=1}^M c_i m_{in} - k_n}{\sum_{j=1}^M c_{jn}(1 - c_{jn})m_{jn}} = O(n^{-1/2}).$$

This implies

$$\frac{(c_i - c_{in})m_{in}}{\sqrt{n}} = \frac{c_{in}(1 - c_{in})m_{in}}{\sum_{j=1}^M c_{jn}(1 - c_{jn})m_{jn}} \frac{\sum_{i=1}^M c_i m_{in} - k_n}{\sqrt{n}} + O(n^{-1/2}),$$

from which the result follows.

Next, suppose that $K = \infty$. Since c_{in} is increasing as a function of k_n , we have by the first part of the proof

$$\liminf_{n \rightarrow \infty} \frac{(c_{in} - c_i)m_{in}}{\sqrt{n}} \geq \frac{c_i(1 - c_i)\alpha_i}{\sum_{j=1}^M c_j(1 - c_j)\alpha_j} L$$

for all $L \in \mathbb{R}$. Hence, the left-hand side is equal to ∞ . The proof for the case $K = -\infty$ is similar. \square

When we condition on $\{\Sigma_n = k_n + \ell\}$, then in analogy with what we have done before, the natural scaled vector to consider would be the vector

$$\boldsymbol{\mathcal{X}}_n^\ell := \left(\frac{X_{1n} - x_{1n}^\ell}{\sqrt{n}}, \frac{X_{2n} - x_{2n}^\ell}{\sqrt{n}}, \dots, \frac{X_{Mn} - x_{Mn}^\ell}{\sqrt{n}} \right),$$

where the components of the vector $x_n^\ell = (x_{1n}^\ell, \dots, x_{Mn}^\ell)$ identify the centres around which the X_{in} concentrate. Here, the x_{in}^ℓ are nonnegative integers chosen such that $|x_{in}^\ell - c_{in}^\ell m_{in}| < 1$ for all i , and $\sum_{i=1}^M x_{in}^\ell = k_n + \ell$. Note that the vector $\boldsymbol{\mathcal{X}}_n^\ell$ is simply a translation of $\boldsymbol{\mathcal{X}}_n$ by $(x_n^\ell - x_n)/\sqrt{n}$. Since Lemma 2.3.3 shows that if k_n is sufficiently larger than $\mathbb{E}(\Sigma_n)$, only values of ℓ up to small order in n are relevant, the statement of Theorem 2.1.5 should not come as a surprise. To prove it, we need to refine the arguments we used to prove Theorem 2.1.4.

Proof of Theorem 2.1.5. Assume that $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n} \rightarrow \infty$, and let

$$a_n := 2M \left[\sqrt{n} \left(\frac{\sqrt{n}}{k_n - \mathbb{E}(\Sigma_n)} \right)^{1/2} \right].$$

Note that then $a_n \rightarrow \infty$ but $a_n/\sqrt{n} \rightarrow 0$. Furthermore, Lemma 2.3.3 and a short computation show that

$$\frac{\mathbb{P}(\Sigma_n > k_n + a_n)}{\mathbb{P}(\Sigma_n \geq k_n)} \rightarrow 0.$$

It is easy to see that from this last fact it follows that

$$\sup_A \left| \mathbb{P}(\boldsymbol{\mathcal{X}}_n \in A \mid \Sigma_n \geq k_n) - \mathbb{P}(\boldsymbol{\mathcal{X}}_n \in A \mid k_n \leq \Sigma_n \leq k_n + a_n) \right| \rightarrow 0,$$

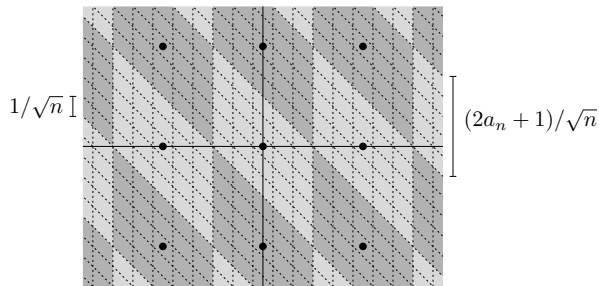


Figure 2.2: We coarse-grain our densities by combining $(2a_n + 1)^M$ sheared cubes into larger sheared cubes. Here, we show this coarse-graining for $M = 2$ and $a_n = 2$. The dots are the points in $((2a_n + 1)\mathbb{Z})^M / \sqrt{n}$. The combined sheared cubes have been colored in a chessboard fashion as a visual aid.

where the supremum is over all Borel subsets A of \mathbb{R}^M . It is therefore sufficient to consider the limiting distribution of the vector \mathcal{X}_n conditioned on the event $\{k_n \leq \Sigma_n \leq k_n + a_n\}$, rather than on the event $\{\Sigma_n \geq k_n\}$.

As in the proof of Theorem 2.1.4, for $z \in \mathbb{R}^M$ we let $r_n^z = \lceil z\sqrt{n} \rceil$, and we define the functions $f_n: \mathbb{R}^M \rightarrow \mathbb{R}$ by setting

$$f_n(z) := (\sqrt{n})^M \mathbb{P}(\sqrt{n} \mathcal{X}_n = r_n^z \mid k_n \leq \Sigma_n \leq k_n + a_n).$$

As before, this is a probability density function with respect to Lebesgue measure λ on \mathbb{R}^M , and if \mathcal{Z}_n is a random vector with this density, then the vector $\mathcal{Z}'_n = \lceil \mathcal{Z}_n \sqrt{n} \rceil / \sqrt{n}$ has the same distribution as the vector \mathcal{X}_n conditioned on the event $\{k_n \leq \Sigma_n \leq k_n + a_n\}$. Hence, it is enough to show that the weak limit of \mathcal{Z}_n has density $f / \int f d\mu_0$ with respect to μ_0 .

An essential difference compared to the situation in Theorem 2.1.4, however, is that the densities f_n are no longer supported by the collection of points z for which r_n^z is in the hyperplane S_0 (i.e. the union of those sheared cubes that intersect S_0). Rather, the support now encompasses all the points z for which r_n^z is in any of the hyperplanes

$$S_\ell := \{(z_1, \dots, z_M) \in \mathbb{R}^M : z_1 + \dots + z_M = \ell\}, \quad \ell = 0, 1, \dots, a_n,$$

because if $r_n^z \in S_\ell$, then the event $\{\sqrt{n} \mathcal{X}_n = r_n^z\}$ is contained in the event $\{\Sigma_n = k_n + \ell\}$. For this reason, the densities f_n are not so convenient to work with here. Instead, it is more convenient to “coarse-grain” our densities by spreading the mass over sheared cubes of volume $((2a_n + 1)/\sqrt{n})^M$ rather than volume $(1/\sqrt{n})^M$, to the effect that all the mass is again contained in the collection of sheared (coarse-grained) cubes intersecting S_0 .

To this end, for given n we partition \mathbb{R}^M into the collection of sets

$$\left\{ \frac{1}{\sqrt{n}} (a + \sigma^{-1}(-a_n - 1/2, a_n + 1/2]^M) : a \in ((2a_n + 1)\mathbb{Z})^M \right\}. \quad (2.3.11)$$

See Figure 2.2. For a given point $z \in \mathbb{R}^M$, we denote by Q_n^z the sheared cube in this partition containing z . Now we can define the coarse-grained densities

$$\begin{aligned} g_n(z) &:= \left(\frac{\sqrt{n}}{2a_n + 1} \right)^M \mathbb{P}(\mathcal{X}_n \in Q_n^z \mid k_n \leq \Sigma_n \leq k_n + a_n) \\ &= \left(\frac{\sqrt{n}}{2a_n + 1} \right)^M \int_{Q_n^z} f_n(y) d\lambda(y). \end{aligned}$$

By construction, these are again probability density functions with respect to M -dimensional Lebesgue measure λ . Moreover, each of these densities is supported on the collection of sheared cubes in (2.3.11) that intersect S_0 , and is constant on each sheared cube Q_n^z . In particular, for any given point $z \in \mathbb{R}^M$ we have

$$\int_{Q_n^z} g_n(y) d\lambda(y) = \frac{2a_n + 1}{\sqrt{n}} \int_{Q_n^z} g_n(y) d\mu_0(y).$$

Finally, because $a_n/\sqrt{n} \rightarrow 0$ it is clear that if \mathcal{Z}_n'' has density g_n , then its weak limit will coincide with that of \mathcal{Z}_n , and hence also with that of the vector \mathcal{X}_n conditioned on the event $\{k_n \leq \Sigma_n \leq k_n + a_n\}$.

Suppose now that we could prove that

$$\frac{2a_n + 1}{\sqrt{n}} g_n(z) \rightarrow \frac{f(z)}{\int f d\mu_0} \quad \text{for every } z \in \mathbb{R}^M. \quad (2.3.12)$$

Then it would follow from Fatou's lemma that for every open set $A \subset \mathbb{R}^M$,

$$\liminf_{n \rightarrow \infty} \int_A \frac{2a_n + 1}{\sqrt{n}} g_n(z) d\mu_0(z) \geq \frac{\int_A f(z) d\mu_0(z)}{\int f d\mu_0}.$$

By approximating the open set A by unions of sheared cubes contained in A , as in the proof of Theorem 2.1.4, it is then clear that this would imply that

$$\liminf_{n \rightarrow \infty} \int_A g_n(z) d\lambda(z) \geq \frac{\int_A f(z) d\mu_0(z)}{\int f d\mu_0}.$$

It therefore only remains to establish (2.3.12).

Since (2.3.12) holds by construction for $z \notin S_0$, we only need to consider the case $z \in S_0$. So let us fix $z \in S_0$, and look at $g_n(z)$. By definition, this is just the rescaled conditional probability that the vector \mathcal{X}_n lies in the sheared cube Q_n^z , given that $k_n \leq \Sigma_n \leq k_n + a_n$. In other words, if we define $C_n^z := \sqrt{n}Q_n^z \cap \mathbb{Z}^M$ and $C_{\ell_n}^z := C_n^z \cap S_\ell$, then we have

$$\begin{aligned} g_n(z) &= \left(\frac{\sqrt{n}}{2a_n + 1} \right)^M \sum_{r \in C_n^z} \mathbb{P}(\sqrt{n}\mathcal{X}_n = r \mid k_n \leq \Sigma_n \leq k_n + a_n) \\ &= \left(\frac{\sqrt{n}}{2a_n + 1} \right)^M \sum_{\ell=0}^{a_n} \sum_{r \in C_{\ell_n}^z} \frac{\mathbb{P}(\sqrt{n}\mathcal{X}_n = r \mid \Sigma_n = k_n + \ell) \mathbb{P}(\Sigma_n = k_n + \ell)}{\mathbb{P}(k_n \leq \Sigma_n \leq k_n + a_n)}. \end{aligned}$$

Since $C_{\ell_n}^z$ contains exactly $(2a_n + 1)^{M-1}$ points, from this equality we conclude that to prove (2.3.12), it is sufficient to show that

$$\sup_{0 \leq \ell \leq a_n} \sup_{r \in C_{\ell_n}^z} \left| (\sqrt{n})^{M-1} \mathbb{P}(\sqrt{n} \mathcal{X}_n = r \mid \Sigma_n = k_n + \ell) - \frac{f(z)}{\int f d\mu_0} \right| \rightarrow 0. \quad (2.3.13)$$

The proof of (2.3.13) proceeds along the same line as the proof of pointwise convergence in Theorem 2.1.4, based on Lemma 2.3.2. However, there is a catch: because we are now conditioning on $\Sigma_n = k_n + \ell$, the X_{in} are no longer centred around x_{in} , but around x_{in}^ℓ . We therefore first write the conditional probabilities in a form analogous to what we had before, by using that

$$\mathbb{P}(\sqrt{n} \mathcal{X}_n = r \mid \Sigma_n = k_n + \ell) = \mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = r + x_n - x_n^\ell \mid \Sigma_n = k_n + \ell).$$

Writing $r^\ell := r + x_n - x_n^\ell$ for convenience, we now want to study the ratios

$$\frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = r^\ell \mid \Sigma_n = k_n + \ell)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = 0 \mid \Sigma_n = k_n + \ell)} = \frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = r^\ell)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = 0)} = \prod_{i=1}^M \frac{\mathbb{P}(X_{in} = x_{in}^\ell + r_i^\ell)}{\mathbb{P}(X_{in} = x_{in}^\ell)}$$

for ℓ and r satisfying $0 \leq \ell \leq a_n$ and $r \in C_{\ell_n}^z$.

By equation (2.3.9) and Lemma 2.3.4 we have that $\sup_\ell |x_{in}^\ell - x_{in}| = o(\sqrt{n})$, from which it follows that also $\sup_{\ell, r} |r^\ell - z\sqrt{n}| = o(\sqrt{n})$, where the suprema are over all $\ell \in \{0, \dots, a_n\}$ and $r \in C_{\ell_n}^z$. Thus, by the first part of Lemma 2.3.2,

$$\sup_{0 \leq \ell \leq a_n} \sup_{r \in C_{\ell_n}^z} \left| \frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = r^\ell \mid \Sigma_n = k_n + \ell)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = 0 \mid \Sigma_n = k_n + \ell)} - f(z) \right| \rightarrow 0,$$

where we have used that for all terms concerned, $\prod_{i=1}^M A_n^{r_i^\ell} = 1$ because $r^\ell \in S_0$. Furthermore, from the second part of Lemma 2.3.2 it follows that the functions

$$z \mapsto \frac{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = \lceil z\sqrt{n} \rceil \mid \Sigma_n = k_n + \ell)}{\mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = 0 \mid \Sigma_n = k_n + \ell)}$$

are bounded uniformly in n and in all $\ell \in \{0, \dots, a_n\}$ by a μ_0 -integrable function. In the same way as in the proof of Theorem 2.1.4, it follows from these facts (with the addition that we have uniform bounds) that

$$\sup_{0 \leq \ell \leq a_n} \left| (\sqrt{n})^{M-1} \mathbb{P}(\sqrt{n} \mathcal{X}_n^\ell = 0 \mid \Sigma_n = k_n + \ell) - \frac{1}{\int f d\mu_0} \right| \rightarrow 0.$$

From this we conclude that (2.3.13) does hold, which completes the proof of Theorem 2.1.5. \square

2.3.3 Proof of Theorem 2.1.6

Proof of Theorem 2.1.6. Suppose that $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n} \rightarrow K$ for some $K \in [-\infty, \infty)$. Let \mathcal{X} be a random vector having a multivariate normal distribution with density

$h/\int h d\lambda$ with respect to λ . By standard arguments, \mathcal{X}_n^p converges weakly to \mathcal{X} . Therefore, for a rectangle $A \subset \mathbb{R}^M$ we have

$$\mathbb{P}(\mathcal{X}_n^p \in A, \Sigma_n \geq k_n) = \mathbb{P}(\mathcal{X}_n^p \in A \cap H_{\frac{k_n - \mathbb{E}(\Sigma_n)}{\sqrt{n}}}) \rightarrow \mathbb{P}(\mathcal{X} \in A \cap H_K),$$

since $A \cap H_{K+\varepsilon}$ is a λ -continuity set for all $\varepsilon \in \mathbb{R}$. Taking $A = \mathbb{R}^M$ gives

$$\mathbb{P}(\Sigma_n \geq k_n) \rightarrow \mathbb{P}(\mathcal{X} \in H_K).$$

Hence, for all rectangles $A \subset \mathbb{R}^M$

$$\mathbb{P}(\mathcal{X}_n^p \in A \mid \Sigma_n \geq k_n) \rightarrow \frac{\mathbb{P}(\mathcal{X} \in A \cap H_K)}{\mathbb{P}(\mathcal{X} \in H_K)}. \quad \square$$

2.3.4 Law of large numbers

Finally, we prove a law of large numbers, which we will need in Section 2.4. Let \tilde{X}_{in} denote a random variable with the conditional law of X_{in} , conditioned on the event $\{\Sigma_n \geq k_n\}$. If $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n} \rightarrow K$ for some $K \in [-\infty, \infty]$, then an immediate consequence of Theorems 2.1.5 and 2.1.6 is that \tilde{X}_{in}/n converges in probability to either $p_i\alpha_i$ or $c_i\alpha_i$. The following theorem shows that such a law of large numbers holds for a general sequence k_n such that $k_n/n \rightarrow \alpha$.

Theorem 2.3.7. *For $i \in \{1, \dots, M\}$, the random variable \tilde{X}_{in}/n converges in probability to $p_i\alpha_i$ if $\alpha \leq \sum_{i=1}^M p_i\alpha_i$, or to $c_i\alpha_i$ if $\alpha \geq \sum_{i=1}^M p_i\alpha_i$.*

Proof. If $\alpha \neq \sum_{i=1}^M p_i\alpha_i$, then $(k_n - \mathbb{E}(\Sigma_n))/\sqrt{n}$ goes to $-\infty$ or ∞ as $n \rightarrow \infty$, and the result immediately follows from Theorem 2.1.5 and Theorem 2.1.6.

Now suppose that $\alpha = \sum_{i=1}^M p_i\alpha_i$. Then $c_i = p_i$ for all $i \in \{1, \dots, M\}$. Recall that in general the c_i and A are determined by the equations

$$c_i = \frac{p_i}{p_i + A(1 - p_i)} \quad \text{and} \quad \sum_{i=1}^M \frac{p_i\alpha_i}{p_i + A(1 - p_i)} = \alpha.$$

The constant A is continuous as a function of α , hence $c_i = c_i[\alpha]$ is also continuous as a function of α . Therefore, if $\alpha = \sum_{i=1}^M p_i\alpha_i$, then for each $\varepsilon > 0$ we can choose $\delta > 0$ such that $c_i[\alpha + \delta]\alpha_i \leq p_i\alpha_i + \frac{1}{2}\varepsilon$. By Corollary 2.2.3 we have, for large enough n ,

$$\begin{aligned} \mathbb{P}(X_{in} \geq (p_i\alpha_i + \varepsilon)n \mid \Sigma_n \geq k_n) \\ &\leq \mathbb{P}(X_{in} \geq (p_i\alpha_i + \varepsilon)n \mid \Sigma_n \geq (\alpha + \delta)n) \\ &\leq \mathbb{P}(X_{in} \geq (c_i[\alpha + \delta]\alpha_i + \frac{1}{2}\varepsilon)n \mid \Sigma_n \geq (\alpha + \delta)n), \end{aligned}$$

which tends to 0 as $n \rightarrow \infty$ by Theorem 2.1.5. Similarly, using Corollary 2.2.3 and Theorem 2.1.6 instead of Theorem 2.1.5, we obtain

$$\mathbb{P}(X_{in} \leq (p_i\alpha_i - \varepsilon)n \mid \Sigma_n \geq k_n) \rightarrow 0.$$

We conclude that \tilde{X}_{in}/n converges in probability to $p_i\alpha_i = c_i\alpha_i$. □

2.4 Asymptotic stochastic domination

2.4.1 Proof of Theorem 2.1.8

Consider the general framework for vectors \mathbf{X}_n and \mathbf{Y}_n of Section 2.1.2 in the setting of Section 2.1.4. We will split the proof of Theorem 2.1.8 into four lemmas. In the statements of these lemmas, we will need the constant $\hat{\alpha}$, which is defined as the limit as $n \rightarrow \infty$ of \hat{k}_n/n :

$$\hat{k}_n = \sum_{i=1}^M \frac{p_i m_{in}}{p_i + \beta_{\max}(1 - p_i)}, \quad \text{hence} \quad \hat{\alpha} = \sum_{i=1}^M \frac{p_i \alpha_i}{p_i + \beta_{\max}(1 - p_i)}.$$

Let us first look at the definition of $\hat{\alpha}$ in more detail. In Section 2.1.4, we informally introduced the sequence \hat{k}_n as a critical sequence such that if k_n is around \hat{k}_n , then there exists a block i such that the number of successes \tilde{X}_{in} of the vector $\tilde{\mathbf{X}}_n$ in block i is roughly the same as \tilde{Y}_{in} . We will now make this precise. Recall that the c_i and the constant A are determined by

$$c_i = \frac{p_i}{p_i + A(1 - p_i)} \quad \text{and} \quad \sum_{i=1}^M \frac{p_i \alpha_i}{p_i + A(1 - p_i)} = \alpha.$$

Furthermore, note that

$$\frac{p_i}{p_i + \beta_i(1 - p_i)} = q_i,$$

and recall that we defined $I = \{i \in \{1, \dots, M\} : \beta_i = \beta_{\max}\}$. The ordering of α and $\hat{\alpha}$ gives information about the ordering of the c_i and q_i . This is stated in the following remark, which follows from the equations above.

Remark 2.4.1. We have the following:

- (i) If $\alpha < \hat{\alpha}$, then $A > \beta_{\max}$ and $c_i < q_i$ for all $i \in \{1, \dots, M\}$.
- (ii) If $\alpha = \hat{\alpha}$, then $A = \beta_{\max}$ and $c_i = q_i$ for $i \in I$, while $c_i < q_i$ for $i \notin I$.
- (iii) If $\alpha > \hat{\alpha}$, then $A < \beta_{\max}$ and $c_i > q_i$ for some $i \in \{1, \dots, M\}$.
- (iv) $\sum_{i=1}^M p_i \alpha_i \leq \hat{\alpha} \leq \sum_{i=1}^M q_i \alpha_i$, with $\hat{\alpha} = \sum_{i=1}^M p_i \alpha_i$ if and only if $\beta_{\max} = 1$, and $\hat{\alpha} = \sum_{i=1}^M q_i \alpha_i$ if and only if all β_i ($i \in \{1, \dots, M\}$) are equal.

Our law of large numbers, Theorem 2.3.7, states that \tilde{X}_{in}/n converges in probability to $p_i \alpha_i$ if $\alpha \leq \sum_{i=1}^M p_i \alpha_i$, and to $c_i \alpha_i$ if $\alpha \geq \sum_{i=1}^M p_i \alpha_i$. This law of large numbers applies analogously to the vector $\tilde{\mathbf{Y}}_n$. If we define d_1, \dots, d_M as the unique solution of the system

$$\begin{cases} \frac{1 - d_i}{d_i} \frac{q_i}{1 - q_i} = \frac{1 - d_j}{d_j} \frac{q_j}{1 - q_j} & \forall i, j \in \{1, \dots, M\}, \\ \sum_{i=1}^M d_i \alpha_i = \alpha, \end{cases}$$

then \tilde{Y}_{in}/n converges in probability to $q_i\alpha_i$ if $\alpha \leq \sum_{i=1}^M q_i\alpha_i$, and to $d_i\alpha_i$ if $\alpha \geq \sum_{i=1}^M q_i\alpha_i$. These laws of large numbers and the observations in Remark 2.4.1 will play a crucial role in the proofs in this section.

Now we define one-dimensional (possibly degenerate) distribution functions $F_K: \mathbb{R} \rightarrow [0, 1]$ for $K \in [-\infty, \infty]$, which will come up in the proofs as the distribution functions of the limit of a certain function of the vectors $\tilde{\mathbf{X}}_n$. Recall from Section 2.1.3 the definitions (2.1.5), (2.1.6), (2.1.7) and (2.1.8) of the measure ν_0 , the functions f and h and the half-space H_K . Write $u = (u_1, \dots, u_M)$. Then

$$F_K(z) = \begin{cases} \frac{\int_{H_K \cap \{\sum_{i \in I} u_i \leq z\}} h(u) d\lambda(u)}{\int_{H_K} h d\lambda} & \text{if } K < \infty, \alpha = \sum_{i=1}^M p_i\alpha_i, \\ \frac{\int_{\{\sum_{i \in I} u_i \leq z - z_K\}} f(u) d\nu_0(u)}{\int f d\nu_0} & \text{if } K < \infty, \alpha > \sum_{i=1}^M p_i\alpha_i, \\ 0 & \text{if } K = \infty, \end{cases} \quad (2.4.1)$$

where

$$z_K = \frac{\sum_{i \in I} c_i(1 - c_i)\alpha_i}{\sum_{i=1}^M c_i(1 - c_i)\alpha_i} K. \quad (2.4.2)$$

The following lemmas, together with Proposition 2.1.3, imply Theorem 2.1.8.

Lemma 2.4.2. *If $\alpha < \hat{\alpha}$, then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 1$.*

Lemma 2.4.3. *Suppose that $\alpha > \hat{\alpha}$ and $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$. Then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 0$.*

Lemma 2.4.4. *Suppose that $\alpha = \hat{\alpha}$ and $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$. Suppose furthermore that $(k_n - \hat{k}_n)/\sqrt{n} \rightarrow K$ for some $K \in [-\infty, \infty]$. Then $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow \inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1$.*

Lemma 2.4.5. *If $\alpha = \hat{\alpha}$ and $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$, then*

$$\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = \begin{cases} 1 & \text{if } K = -\infty, \\ P_K & \text{if } K \in \mathbb{R}, \\ 0 & \text{if } K = \infty. \end{cases} \quad \text{where } 0 < P_K < 1,$$

The constant a in Lemma 2.4.4 is the constant defined in (2.1.9a). The infimum in Lemma 2.4.4 can actually be computed, as Lemma 2.4.5 states, and attains the values stated in Theorem 2.1.8, with P_K as defined in (2.1.10).

We will prove Theorem 2.1.8 by proving each of the Lemmas 2.4.2–2.4.5 in turn. The idea behind the proof of Lemma 2.4.2 is as follows. If we do not condition at all, then $\mathbf{X}_n \preceq \mathbf{Y}_n$ for every $n \geq 1$. If $\alpha < \sum_{i=1}^M p_i\alpha_i$, then the effect of conditioning vanishes in the limit and $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 1$ as $n \rightarrow \infty$. If $\sum_{i=1}^M p_i\alpha_i \leq \alpha < \hat{\alpha}$, then $c_i < q_i$ for all $i \in \{1, \dots, M\}$. Hence, for large n we have that \tilde{X}_{in} is significantly smaller than \tilde{Y}_{in} for all $i \in \{1, \dots, M\}$, from which it will again follow that $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 1$.

Proof of Lemma 2.4.2. First, suppose that $\alpha < \sum_{i=1}^M p_i \alpha_i$. Let \mathbf{X}_n and \mathbf{Y}_n be defined on a common probability space (Ω, \mathcal{F}, P) such that $\mathbf{X}_n \leq \mathbf{Y}_n$ on all of Ω . Pick $\omega_1 \in \Omega$ according to the measure $P(\cdot \mid \sum_{i=1}^M X_{in} \geq k_n)$ and pick $\omega_2 \in \Omega$ independently according to the measure $P(\cdot \mid \sum_{i=1}^M Y_{in} \geq k_n)$. If ω_2 is in the event $\{\sum_{i=1}^M X_{in} \geq k_n\} \in \mathcal{F}$, set $\tilde{\mathbf{Y}}_n(\omega_1, \omega_2) := \mathbf{Y}_n(\omega_1)$, otherwise set $\tilde{\mathbf{Y}}_n(\omega_1, \omega_2) := \mathbf{Y}_n(\omega_2)$. Set $\tilde{\mathbf{X}}_n(\omega_1, \omega_2) := \mathbf{X}_n(\omega_1)$ regardless of the value of ω_2 . It is easy to see that this defines a coupling of $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$ on the space $(\Omega \times \Omega, \mathcal{F} \times \mathcal{F})$ with the correct marginals for $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$. Moreover, in this coupling we have $\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n$ at least if $\omega_2 \in \{\sum_{i=1}^M X_{in} \geq k_n\}$. Hence

$$\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \geq \frac{\mathbb{P}(\sum_{i=1}^M X_{in} \geq k_n)}{\mathbb{P}(\sum_{i=1}^M Y_{in} \geq k_n)},$$

which tends to 1 as $n \rightarrow \infty$ (e.g. by Chebyshev's inequality).

Secondly, suppose that $\sum_{i=1}^M p_i \alpha_i \leq \alpha < \hat{\alpha}$. By Remark 2.4.1(i), $c_i < q_i$ for all $i \in \{1, \dots, M\}$. For each coupling of $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$ we have

$$\mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \geq \mathbb{P}(\tilde{X}_{in} \leq (c_i + q_i)\alpha_i n/2 \leq \tilde{Y}_{in} \ \forall i \in \{1, \dots, M\}),$$

which tends to 1 as $n \rightarrow \infty$ by Theorem 2.3.7 and Remark 2.4.1(iv). \square

The next lemma, Lemma 2.4.3, treats the case $\alpha > \hat{\alpha}$. In this case, we have that for large n , \tilde{X}_{in} is significantly larger than \tilde{Y}_{in} for some $i \in \{1, \dots, M\}$, from which it follows that $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \rightarrow 0$.

Proof of Lemma 2.4.3. First, suppose that $\hat{\alpha} < \alpha < \sum_{i=1}^M q_i \alpha_i$. Then $c_i > q_i$ for some $i \in \{1, \dots, M\}$ by Remark 2.4.1(iii). Hence, by Theorem 2.3.7 and Remark 2.4.1(iv),

$$\begin{aligned} \mathbb{P}(\tilde{X}_{in} \geq (c_i + q_i)\alpha_i n/2) &\rightarrow 1, \\ \mathbb{P}(\tilde{Y}_{in} \geq (c_i + q_i)\alpha_i n/2) &\rightarrow 0. \end{aligned}$$

It follows that $\mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ tends to 0 uniformly over all couplings.

Next, suppose that $\alpha \geq \sum_{i=1}^M q_i \alpha_i$ and $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$. Then there exists $i \in \{1, \dots, M\}$ such that $c_i \neq d_i$, since

$$\frac{1 - d_i}{d_i} \frac{d_j}{1 - d_j} \beta_j = \frac{1 - q_i}{q_i} \frac{p_j}{1 - p_j} = \beta_i \frac{1 - c_i}{c_i} \frac{c_j}{1 - c_j}.$$

In fact, we must have $c_i > d_i$ for some $i \in \{1, \dots, M\}$, because $\sum_{i=1}^M c_i \alpha_i = \sum_{i=1}^M d_i \alpha_i$. By Theorem 2.3.7, it follows that

$$\begin{aligned} \mathbb{P}(\tilde{X}_{in} \geq (c_i + d_i)\alpha_i n/2) &\rightarrow 1, \\ \mathbb{P}(\tilde{Y}_{in} \geq (c_i + d_i)\alpha_i n/2) &\rightarrow 0. \end{aligned}$$

Again, $\mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$ tends to 0 uniformly over all couplings. \square

We now turn to the proof of Lemma 2.4.4. Under the assumptions of this lemma, $c_i = q_i$ for $i \in I$ and $c_i < q_i$ for $i \notin I$. The proof proceeds in four steps. In step 1, we show that the blocks $i \notin I$ do not influence the asymptotic behavior of $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n)$, because for these blocks, \tilde{X}_{in} is significantly smaller than \tilde{Y}_{in} for large n . In step 2, we show that the parts of the vectors $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$ that correspond to the blocks $i \in I$ are stochastically ordered, if and only if the total numbers of successes in these parts of the vectors are stochastically ordered. At this stage, the original problem of stochastic ordering of random vectors has been reduced to a problem of stochastic ordering of random variables. In step 3, we use our central limit theorems to deduce the asymptotic behavior of the total numbers of successes in the blocks $i \in I$. In step 4, we apply the following lemma, which follows from [45, Proposition 1], to these total numbers of successes:

Lemma 2.4.6. *Let X and Y be random variables with distribution functions F and G respectively. Then we have*

$$\sup \mathbb{P}(X \leq Y) = \inf_{z \in \mathbb{R}} F(z) - G(z) + 1,$$

where the supremum is taken over all possible couplings of X and Y .

Proof of Lemma 2.4.4. Write $m_{In} := \sum_{i \in I} m_{in}$. Let \mathbf{X}_{In} and $\tilde{\mathbf{X}}_{In}$ denote the m_{In} -dimensional subvectors of \mathbf{X}_n and $\tilde{\mathbf{X}}_n$, respectively, consisting of the components that belong to the blocks $i \in I$. Define \mathbf{Y}_{In} and $\tilde{\mathbf{Y}}_{In}$ analogously.

Step 1. Note that for each coupling of $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$,

$$\begin{aligned} \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) &\geq \mathbb{P}(\tilde{\mathbf{X}}_{In} \leq \tilde{\mathbf{Y}}_{In}, \tilde{X}_{in} \leq (c_i + q_i)\alpha_i n / 2 \leq \tilde{Y}_{in} \ \forall i \notin I) \\ &\geq \mathbb{P}(\tilde{\mathbf{X}}_{In} \leq \tilde{\mathbf{Y}}_{In}) - \\ &\quad \sum_{i \notin I} \left\{ \mathbb{P}\left(\tilde{X}_{in} > \frac{c_i + q_i}{2} \alpha_i n\right) + \mathbb{P}\left(\tilde{Y}_{in} < \frac{c_i + q_i}{2} \alpha_i n\right) \right\}. \end{aligned} \quad (2.4.3)$$

By Remark 2.4.1(ii), $c_i < q_i$ for $i \notin I$. Hence, it follows from Remark 2.4.1(iv) and Theorem 2.3.7 that the sum in (2.4.3) tends to 0 as $n \rightarrow \infty$, uniformly over all couplings. Since clearly $\sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) \leq \sup \mathbb{P}(\tilde{\mathbf{X}}_{In} \leq \tilde{\mathbf{Y}}_{In})$,

$$\left| \sup \mathbb{P}(\tilde{\mathbf{X}}_n \leq \tilde{\mathbf{Y}}_n) - \sup \mathbb{P}(\tilde{\mathbf{X}}_{In} \leq \tilde{\mathbf{Y}}_{In}) \right| \rightarrow 0,$$

where the suprema are taken over all possible couplings of $(\tilde{\mathbf{X}}_n, \tilde{\mathbf{Y}}_n)$ and $(\tilde{\mathbf{X}}_{In}, \tilde{\mathbf{Y}}_{In})$, respectively.

Step 2. The β_i for $i \in I$ are all equal. Hence, by Proposition 2.1.2 and Lemma 2.2.2 we have for $m \in \{0, 1, \dots, m_{In}\}$ and $\ell \in \{0, 1, \dots, m_{In} - m\}$

$$\mathcal{L}(\mathbf{X}_{In} | \sum_{i \in I} X_{in} = m) \preceq \mathcal{L}(\mathbf{Y}_{In} | \sum_{i \in I} Y_{in} = m + \ell). \quad (2.4.4)$$

Now let B be any collection of vectors of length m_{In} with exactly m components equal to 1 and $m_{In} - m$ components equal to 0. Then

$$\begin{aligned} \mathbb{P}(\tilde{\mathbf{X}}_{In} \in B) &= \mathbb{P}(\mathbf{X}_{In} \in B | \sum_{i=1}^M X_{in} \geq k_n) \\ &= \frac{\mathbb{P}(\mathbf{X}_{In} \in B) \mathbb{P}(\sum_{i \notin I} X_{in} \geq k_n - m)}{\mathbb{P}(\sum_{i=1}^M X_{in} \geq k_n)}. \end{aligned}$$

Taking C to be the collection of all vectors in $\{0, 1\}^{m_{I_n}}$ with exactly m components equal to 1, we obtain

$$\mathbb{P}(\tilde{\mathbf{X}}_{I_n} \in B \mid \sum_{i \in I} \tilde{X}_{in} = m) = \frac{\mathbb{P}(\tilde{\mathbf{X}}_{I_n} \in B)}{\mathbb{P}(\tilde{\mathbf{X}}_{I_n} \in C)} = \mathbb{P}(\mathbf{X}_{I_n} \in B \mid \sum_{i \in I} X_{in} = m),$$

and likewise for \mathbf{Y}_{I_n} and $\tilde{\mathbf{Y}}_{I_n}$. Hence, (2.4.4) is equivalent to

$$\mathcal{L}(\tilde{\mathbf{X}}_{I_n} \mid \sum_{i \in I} \tilde{X}_{in} = m) \preceq \mathcal{L}(\tilde{\mathbf{Y}}_{I_n} \mid \sum_{i \in I} \tilde{Y}_{in} = m + \ell).$$

With a similar argument as in the proof of Proposition 2.1.3, it follows that

$$\sup \mathbb{P}(\tilde{\mathbf{X}}_{I_n} \leq \tilde{\mathbf{Y}}_{I_n}) = \sup \mathbb{P}(\sum_{i \in I} \tilde{X}_{in} \leq \sum_{i \in I} \tilde{Y}_{in}).$$

Step 3. First observe that by Remark 2.4.1(iv), $\alpha < \sum_{i=1}^M q_i \alpha_i$. Hence, by Theorem 2.1.6 (note that $(k_n - \mathbb{E}(\sum_{i=1}^M Y_{in}))/\sqrt{n} \rightarrow -\infty$) and the continuous mapping theorem,

$$\mathbb{P}(\sum_{i \in I} (\tilde{Y}_{in} - q_i m_{in})/\sqrt{n} \leq z) \rightarrow \Phi(z/a) \quad \text{for every } z \in \mathbb{R}. \quad (2.4.5)$$

Next observe that by Remark 2.4.1(ii), $c_i = q_i$ for $i \in I$ and $A = \beta_{\max}$, from which it follows that $\hat{k}_n = \sum_{i=1}^M c_i m_{in}$. Hence, Corollary 2.3.5 gives

$$\sum_{i \in I} (c_{in} - q_i) m_{in}/\sqrt{n} \rightarrow z_K, \quad (2.4.6)$$

with z_K as defined in (2.4.2). In the case $\alpha > \sum_{i=1}^M p_i \alpha_i$, Theorem 2.1.5, (2.4.6) and the continuous mapping theorem now immediately imply

$$\mathbb{P}(\sum_{i \in I} (\tilde{X}_{in} - q_i m_{in})/\sqrt{n} \leq z) \rightarrow F_K(z) \quad \text{for every } z \in \mathbb{R}. \quad (2.4.7)$$

Note that if $K = \pm\infty$, F_K is degenerate in this case: we have $F_K(z) = 1$ for all $z \in \mathbb{R}$ if $K = -\infty$ and $F_K(z) = 0$ for all $z \in \mathbb{R}$ if $K = \infty$.

Now consider the case $\alpha = \sum_{i=1}^M p_i \alpha_i$. By Remark 2.4.1(iv), in this case we have $\beta_{\max} = 1$, which implies that $\hat{k}_n = \sum_{i=1}^M p_i m_{in} = \mathbb{E}(\Sigma_n)$ and $p_i = q_i$ for all $i \in \{1, \dots, M\}$. Hence, if $K = \infty$, then (2.4.6) and Theorem 2.1.5 again imply (2.4.7) with $F_K(z) = 0$ everywhere. If $K \in [-\infty, \infty)$, then we obtain (2.4.7) directly from Theorem 2.1.6; F_K is non-degenerate in this case (also for $K = -\infty$).

Step 4. The distribution functions on the left-hand sides of (2.4.5) and (2.4.7) are non-decreasing and bounded between 0 and 1, hence they converge uniformly on compact sets. It follows by Lemma 2.4.6 that

$$\sup \mathbb{P}(\sum_{i \in I} \tilde{X}_{in} \leq \sum_{i \in I} \tilde{Y}_{in}) \rightarrow \inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1. \quad \square$$

Finally, we turn to the proof of Lemma 2.4.5. The key to computing the infimum of $F_K(z) - \Phi(z/a) + 1$ is to first express the distribution function F_K , defined in (2.4.1), in a simpler form.

Proof of Lemma 2.4.5. In the case $\alpha > \sum_{i=1}^M p_i \alpha_i$ and $K = -\infty$, F_K is 1 everywhere, hence $\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = 1$. In the case $K = \infty$, F_K is 0 everywhere, hence $\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = 0$. We will now study the remaining cases.

Consider the case $\alpha = \hat{\alpha} = \sum_{i=1}^M p_i \alpha_i$ and $K \in [-\infty, \infty)$. Let $\mathbf{Z} = (Z_1, \dots, Z_M)$ be a random vector which has the multivariate normal distribution with density $h/\int h d\lambda$. By Remark 2.4.1(iv) we have $\beta_{\max} = 1$. Note that therefore, $\frac{1}{a} \sum_{i \in I} Z_i$, $\frac{1}{b} \sum_{i \notin I} Z_i$ and $\frac{1}{c} \sum_{i=1}^M Z_i$, with a, b and c as defined in (2.1.9), all have the standard normal distribution. Moreover, $\sum_{i \in I} Z_i$ and $\sum_{i \notin I} Z_i$ are independent.

For $K = -\infty$, it follows that $F_K(z) = \Phi(z/a)$, hence $\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = 1$. For $K \in \mathbb{R}$, observe that $\mathbf{Z} \in H_K$ is equivalent with $\frac{1}{c} \sum_{i=1}^M Z_i \geq K/c$. Likewise, $\mathbf{Z} \in H_K \cap \{u \in \mathbb{R}^M : \sum_{i \in I} u_i \leq z\}$ is equivalent with $\frac{1}{a} \sum_{i \in I} Z_i \leq z/a$ and $\frac{1}{b} \sum_{i \notin I} Z_i \geq (K - \sum_{i \in I} Z_i)/b$. It follows that

$$\begin{aligned} F_K(z) &= \frac{\int_{H_K} h d\lambda}{\int h d\lambda} \frac{\int_{H_K \cap \{\sum_{i \in I} u_i \leq z\}} h(u) d\lambda(u)}{\int h d\lambda} \\ &= \frac{1}{1 - \Phi(K/c)} \int_{-\infty}^{z/a} \int_{\frac{K-au}{b}}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} \frac{e^{-v^2/2}}{\sqrt{2\pi}} dv du \\ &= \int_{-\infty}^{z/a} \frac{e^{-u^2/2}}{\sqrt{2\pi}} \frac{1 - \Phi\left(\frac{K-au}{b}\right)}{1 - \Phi\left(\frac{K}{c}\right)} du, \end{aligned}$$

hence

$$F_K(z) - \Phi(z/a) = \int_{-\infty}^{z/a} \frac{e^{-u^2/2}}{\sqrt{2\pi}} \frac{\Phi\left(\frac{K}{c}\right) - \Phi\left(\frac{K-au}{b}\right)}{1 - \Phi\left(\frac{K}{c}\right)} du. \quad (2.4.8)$$

Clearly, the derivative of this expression with respect to z is 0 if and only if $(K-z)/b = K/c$, that is, $z = z_{\min} = K - bK/c$. Plugging this value for z into (2.4.8) shows that $\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = P_K$, with P_K as defined in (2.1.10). Moreover, $P_K > 0$ because $F_K(z_{\min}) > 0$, and $P_K < 1$ because the integrand in (2.4.8) is negative for $u < z_{\min}/a$.

Finally, consider the case $\alpha > \sum_{i=1}^M p_i \alpha_i$ and $K \in \mathbb{R}$. This time, let $\mathbf{Z} = (Z_1, \dots, Z_M)$ be a random vector which has the singular multivariate normal distribution with density $f/\int f d\nu_0$ with respect to ν_0 . Then a little computation shows that (Z_1, \dots, Z_{M-1}) has a multivariate normal distribution with mean 0 and a covariance matrix Σ given by

$$\left\{ \begin{array}{l} \Sigma_{ii} = \frac{\sigma_i^2 \sum_{k=1, k \neq i}^M \sigma_k^2}{\sum_{k=1}^M \sigma_k^2} \quad \text{for } i \in \{1, \dots, M-1\}, \\ \Sigma_{ij} = \frac{-\sigma_i^2 \sigma_j^2}{\sum_{k=1}^M \sigma_k^2} \quad \text{for } i, j \in \{1, \dots, M-1\} \text{ with } i \neq j, \end{array} \right.$$

where $\sigma_i^2 = c_i(1 - c_i)\alpha_i$ for $i \in \{1, \dots, M\}$. Similarly, every subvector of \mathbf{Z} of dimension less than M has a multivariate normal distribution.

By the definition (2.4.1) of F_K , $z_K + \sum_{i \in I} Z_i$ has distribution function F_K . Since $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$, we have $|I| \leq M - 1$. It follows that $\sum_{i \in I} Z_i$ has a normal distribution with mean 0 and variance

$$\sum_{i \in I} \frac{\sigma_i^2 \sum_{k=1, k \neq i}^M \sigma_k^2}{\sum_{k=1}^M \sigma_k^2} + \sum_{i \in I} \sum_{j \in I \setminus \{i\}} \frac{-\sigma_i^2 \sigma_j^2}{\sum_{k=1}^M \sigma_k^2} = \frac{(\sum_{i \in I} \sigma_i^2)(\sum_{i \notin I} \sigma_i^2)}{\sum_{i=1}^M \sigma_i^2}. \quad (2.4.9)$$

By Remark 2.4.1(ii), $A = \beta_{\max}$ and hence for $i \in \{1, \dots, M\}$,

$$\sigma_i^2 = c_i(1 - c_i)\alpha_i = \frac{\beta_{\max}p_i(1 - p_i)\alpha_i}{(p_i + \beta_{\max}(1 - p_i))^2}.$$

It follows that the variance (2.4.9) is equal to a^2b^2/c^2 , with a , b , and c as defined in (2.1.9). Furthermore, $z_K = a^2K/c^2$. We conclude that F_K is the distribution function of a normally distributed random variable with mean a^2K/c^2 and variance a^2b^2/c^2 , so that $F_K(z) = \Phi\left(\frac{c}{ab}(z - a^2K/c^2)\right)$. Since $a^2b^2/c^2 < a^2$, we see that $F_K(z) < \Phi(z/a)$ for small enough z . Hence $F_K(z) - \Phi(z/a)$ attains a minimum value which is strictly smaller than 0. This minimum is strictly larger than -1 because $F_K(z) > 0$ for all $z \in \mathbb{R}$.

To find the minimum, we compute the derivative of $F_K(z) - \Phi(z/a)$ with respect to z . It is not difficult to verify that the minimum is attained for

$$z = z_{\min} = K - \frac{b}{c}\sqrt{K^2 + c^2 \log(c^2/b^2)},$$

from which it follows that $\inf_{z \in \mathbb{R}} F_K(z) - \Phi(z/a) + 1 = P_K$, with P_K as defined in (2.1.10). From the remarks above we know that $0 < P_K < 1$. \square

2.4.2 Conditioning on exactly k_n successes

For the sake of completeness, we finally treat the case of conditioning on the total number of successes being equal to k_n . The situation is not very interesting here.

Theorem 2.4.7. *Let $\hat{\mathbf{X}}_n$ be a random vector having the conditional distribution of \mathbf{X}_n , conditioned on the event $\{\Sigma_n = k_n\}$. Define $\hat{\mathbf{Y}}_n$ similarly. If all β_i ($i \in \{1, \dots, M\}$) are equal, then $\hat{\mathbf{X}}_n$ and $\hat{\mathbf{Y}}_n$ have the same distribution for every $n \geq 1$. Otherwise, $\sup \mathbb{P}(\hat{\mathbf{X}}_n = \hat{\mathbf{Y}}_n) \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. If all β_i ($i \in \{1, \dots, M\}$) are equal, then by Proposition 2.1.2 we have that $\hat{\mathbf{X}}_n$ and $\hat{\mathbf{Y}}_n$ have the same distribution for every $n \geq 1$. If $\beta_i \neq \beta_j$ for some $i, j \in \{1, \dots, M\}$, then it can be shown that $\sup \mathbb{P}(\hat{\mathbf{X}}_n \leq \hat{\mathbf{Y}}_n) \rightarrow 0$ as $n \rightarrow \infty$, by a similar argument as in the proof of Lemma 2.4.3; instead of Theorem 2.3.7 use Lemma 2.3.1. \square

Chapter 3

On central limit theorems in the random connection model

This chapter is based on the paper [53] by Van de Brug and Meester.

3.1 Introduction

Let (X, λ, g) denote a Poisson random connection model, where X is the underlying Poisson point process on \mathbb{R}^d with density $\lambda > 0$, and where g is a *connection function* which we assume to be a non-increasing and which satisfies $0 < \int_{\mathbb{R}^d} g(|x|) dx < \infty$. In words, this amounts to saying that any two points x and y of X are connected with probability $g(|x - y|)$, independently of all other pairs, independently of X . The random connection model plays an important role in many areas, for instance in telecommunications and epidemiology. In telecommunications, the points of the point process can represent base stations, and the connection function then tells us that two base stations at locations x and y respectively, can communicate to each other with probability $g(|x - y|)$. In epidemiology, the connection function can for instance represent the probability that an infected herd at location x infects another herd at location y .

Let K be a bounded Borel subset of \mathbb{R}^d with non-empty interior and boundary of Lebesgue measure zero. Consider a sequence of positive real numbers λ_n with $\lambda_n/n^d \rightarrow \lambda$, let X_n be a Poisson process on \mathbb{R}^d with density λ_n and let g_n be the connection function defined by $g_n(x) = g(nx)$. Consider the sequence of Poisson random connection models (X_n, λ_n, g_n) on \mathbb{R}^d . Let $I_n(g)$ be the number of isolated vertices of (X_n, λ_n, g_n) in K . Roy and Sarkar [44] claim to prove the following result.

Theorem 3.1.1.

$$\frac{I_n(g) - \mathbb{E} I_n(g)}{\sqrt{\text{Var} I_n(g)}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty, \quad (3.1.1)$$

where \rightsquigarrow denotes convergence in distribution.

Although the statement of this result is correct, the proof in [44] is not. In this chapter, we explain what went wrong in their proof, and how this can be corrected. In

addition, we prove an extension to larger components in case the connection function has bounded support.

3.2 Truncation and scaling

The central limit theorem (3.1.1) is relatively easy to show when g has bounded support, see [44]. Hence, the strategy adopted in [44] is to truncate the relevant connection functions, and let the truncation go to infinity. This means that there are two operations involved: scaling and truncation. The root of the problem lies in the fact that these two operations do not commute.

Following [44] we define for $R > 0$ and $n \in \mathbb{N}$ connection functions $g_R, g^R, g_{n,R}, g_n^R : [0, \infty) \rightarrow [0, 1]$ by

$$\begin{aligned} g_R(x) &= \mathbb{1}_{\{x \leq R\}} g(x), & g^R(x) &= \mathbb{1}_{\{x > R\}} g(x), \\ g_{n,R}(x) &= \mathbb{1}_{\{x \leq R\}} g(nx), & g_n^R(x) &= \mathbb{1}_{\{x > R\}} g(nx), \end{aligned}$$

where the *indicator function* $\mathbb{1}_{x \leq R}$ is by definition equal to 1 when $x \leq R$ and equal to 0 when $x > R$, and similarly for the other indicator functions. Note that the notation g_R can formally not be used to denote $\mathbb{1}_{\{x \leq R\}} g(\cdot)$, since g_n has already been defined as $g(n\cdot)$. Nevertheless we will adopt this notation, because we think that this will not cause any confusion. Henceforth g_R will always denote $\mathbb{1}_{\{x \leq R\}} g(\cdot)$ and g_n will always denote $g(n\cdot)$. Let $L_R(g)$ be the number of isolated vertices of (X, λ, g_R) in K that are not isolated in (X, λ, g) . Let $J_{n,R}(g)$ be the number of isolated vertices of $(X_n, \lambda_n, g_{n,R})$ in K and let $L_{n,R}(g) = J_{n,R}(g) - I_n(g)$ be the number of isolated vertices of $(X_n, \lambda_n, g_{n,R})$ in K that are not isolated in (X_n, λ_n, g_n) .

The authors of [44] claim the following (without proof).

Statement 3.2.1. *If (3.1.1) is true when the connection function g has bounded support, then it is the case that*

$$\frac{J_{n,R}(g) - \mathbb{E} J_{n,R}(g)}{\sqrt{\text{Var} J_{n,R}(g)}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty, \quad (3.2.1)$$

for any connection function g .

They then proceed, via a number of moment estimates involving $J_{n,R}(g)$ and $L_{n,R}(g)$, to show that the truth of (3.2.1) for any connection function g , implies the full central limit theorem in (3.1.1).

One problem with their argument is that Statement 3.2.1 is not true, as it would imply that we would be able to write $J_{n,R}(g) = I_n(h)$ for some connection function h with bounded support. This would mean that $g_{n,R}$ can be seen as a scaling of h , that is,

$$\mathbb{1}_{\{x \leq R\}} g(nx) = h(nx),$$

but this leads to $h(x) = \mathbb{1}_{\{x \leq nR\}} g(x)$, which clearly does not make any sense in general.

It seems then that the authors of [44] interchange truncation and scaling, but these two operations do not commute. This mixing up becomes already apparent when we

look at their Lemma 5 which states (without proof) that

$$\lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \mathbb{E} L_{n,R}(g) = p(\lambda, g_R)(1 - p(\lambda, g^R)); \quad (3.2.2)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \text{Var} L_{n,R}(g) &= p(\lambda, g_R)(1 - p(\lambda, g^R)) + \lambda \int_{\mathbb{R}^d} (1 - g(|x|)) \\ &\left[p(\lambda, g_R)^2 p_\lambda^{g_R, g_R}(x, 0) - 2p(\lambda, g_R)^2 p(\lambda, g^R) p_\lambda^{g_R, g}(x, 0) + \right. \\ &\left. p(\lambda, g)^2 p_\lambda^{g, g}(x, 0) \right] - p(\lambda, g_R)^2 (1 - p(\lambda, g^R))^2 dx; \end{aligned} \quad (3.2.3)$$

where ℓ denotes Lebesgue measure on \mathbb{R}^d and

$$p(\mu, h) = e^{-\mu \int_{\mathbb{R}^d} h(|y|) dy} \quad \text{and} \quad p_\mu^{h_1, h_2}(x_1, x_2) = e^{\mu \int_{\mathbb{R}^d} h_1(|y-x_1|) h_2(|y-x_2|) dy}.$$

However, the following proposition shows that (3.2.2) and (3.2.3) are not correct; see the forthcoming Lemma 3.3.3 for a corresponding correct (and useful) statement.

Proposition 3.2.2. *For $R > \sup\{|x_1 - x_2| : x_1, x_2 \in K\}$ we have*

$$\lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \mathbb{E} L_{n,R}(g) = 0; \quad (3.2.4)$$

$$\lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \text{Var} L_{n,R}(g) = 0. \quad (3.2.5)$$

Proof. For $R > 0$ and $n \in \mathbb{N}$ define $k_{nR}, k^{nR} : [0, \infty) \rightarrow [0, 1]$ by

$$k_{nR}(x) = \mathbb{1}_{\{x \leq nR\}} g(x), \quad k^{nR}(x) = \mathbb{1}_{\{x > nR\}} g(x), \quad x \in [0, \infty).$$

We have as $n \rightarrow \infty$,

$$\begin{aligned} p(\lambda_n, g_{n,R}) &= p(\lambda_n/n^d, k_{nR}) \rightarrow p(\lambda, g); \\ p(\lambda_n, g_n^R) &= p(\lambda_n/n^d, k^{nR}) \rightarrow 1. \end{aligned}$$

According to [44] Lemma 4 we have for $R > \sup\{|x_1 - x_2| : x_1, x_2 \in K\}$,

$$\mathbb{E} L_R(g) = \lambda \ell(K) p(\lambda, g_R)(1 - p(\lambda, g^R)), \quad (3.2.6)$$

and therefore,

$$(\lambda_n \ell(K))^{-1} \mathbb{E} L_{n,R}(g) = p(\lambda_n, g_{n,R})(1 - p(\lambda_n, g_n^R)) \rightarrow 0, \quad n \rightarrow \infty,$$

which proves (3.2.4).

To prove (3.2.5), we use Lemma 4 in [44] which says that for $R > \sup\{|x_1 - x_2| : x_1, x_2 \in K\}$, we have

$$\begin{aligned} \text{Var} L_R(g) &= \lambda \ell(K) p(\lambda, g_R)(1 - p(\lambda, g^R)) + \lambda^2 \int_K \int_K (1 - g(|x_1 - x_2|)) \\ &\left[p(\lambda, g_R)^2 p_\lambda^{g_R, g_R}(x_1, x_2) - 2p(\lambda, g_R)^2 p(\lambda, g^R) p_\lambda^{g_R, g}(x_1, x_2) + \right. \\ &\left. p(\lambda, g)^2 p_\lambda^{g, g}(x_1, x_2) \right] - p(\lambda, g_R)^2 (1 - p(\lambda, g^R))^2 dx_2 dx_1. \end{aligned} \quad (3.2.7)$$

We use (3.2.7) with $\lambda = \lambda_n$ and $g = g_n$. Note that as $n \rightarrow \infty$

$$\begin{aligned} p_{\lambda_n}^{g_n, R, g_n, R}(x/n, 0) &= p_{\lambda_n/n^d}^{k_n R, k_n R}(x, 0) \rightarrow p_{\lambda}^{g, g}(x, 0); \\ p_{\lambda_n}^{g_n, R, g_n}(x/n, 0) &= p_{\lambda_n/n^d}^{k_n R, g}(x, 0) \rightarrow p_{\lambda}^{g, g}(x, 0); \\ p_{\lambda_n}^{g_n, g_n}(x/n, 0) &= p_{\lambda_n/n^d}^{g, g}(x, 0) \rightarrow p_{\lambda}^{g, g}(x, 0). \end{aligned}$$

We have

$$\begin{aligned} &\frac{\lambda_n}{\ell(K)} \int_K \int_K (1 - g_n(|x_1 - x_2|)) \left[p(\lambda_n, g_{n, R})^2 p_{\lambda_n}^{g_n, R, g_n, R}(x_1, x_2) - \right. \\ &2p(\lambda_n, g_{n, R})^2 p(\lambda_n, g_n^R) p_{\lambda_n}^{g_n, R, g_n}(x_1, x_2) + p(\lambda_n, g_n)^2 p_{\lambda_n}^{g_n, g_n}(x_1, x_2) \left. \right] - \\ &p(\lambda_n, g_{n, R})^2 (1 - p(\lambda_n, g_n^R))^2 dx_2 dx_1 \\ &= \frac{\lambda_n}{n^d \ell(K)} \int_K \int_{n(K-x_1)} (1 - g(|x_2|)) \left[p(\lambda_n/n^d, k_n R)^2 p_{\lambda_n/n^d}^{k_n R, k_n R}(0, x_2) - \right. \\ &2p(\lambda_n/n^d, k_n R)^2 p(\lambda_n/n^d, k_n^R) p_{\lambda_n/n^d}^{k_n R, g}(0, x_2) + p(\lambda_n/n^d, g)^2 p_{\lambda_n/n^d}^{g, g}(0, x_2) \left. \right] - \\ &p(\lambda_n/n^d, k_n R)^2 (1 - p(\lambda_n/n^d, k_n^R))^2 dx_2 dx_1. \end{aligned}$$

By Lemma 3.3.1 below with $x = -x_2$ we can apply the dominated convergence theorem. Combining the result with (3.2.4) yields (3.2.5). \square

In what follows, we proceed along the way that we believe the authors of [44] had in mind.

For this, we introduce for $R > 0$ and $n \in \mathbb{N}$ connection functions $g_{R, n}, g^{R, n} : [0, \infty) \rightarrow [0, 1]$ as follows:

$$g_{R, n}(x) = \mathbf{1}_{\{x \leq R/n\}} g(nx), \quad g^{R, n}(x) = \mathbf{1}_{\{x > R/n\}} g(nx).$$

Note the difference between $g_{R, n}$ and $g_{n, R}$ and between $g^{R, n}$ and g_n^R . Let $J_{R, n}(g)$ be the number of isolated vertices of $(X_n, \lambda_n, g_{R, n})$ in K and let $L_{R, n}(g) = J_{R, n}(g) - I_n(g)$ be the number of isolated vertices of $(X_n, \lambda_n, g_{R, n})$ in K that are not isolated in (X_n, λ_n, g_n) . Note that the notations $g_{R, n}$, $J_{R, n}(g)$ and $L_{R, n}(g)$ can formally not be used here, since $g_{n, R}$, $J_{n, R}(g)$ and $L_{n, R}(g)$ have already been defined. Nevertheless we will adopt these notations, because henceforth we will use the function $g_{n, R}$ and the random variables $J_{n, R}(g)$ and $L_{n, R}(g)$ no more. We now claim that the following is true (compare the incorrect Statement 3.2.1 above)

Statement 3.2.3. *If (3.1.1) is true when the connection function g has bounded support, then it is the case that*

$$\frac{J_{R, n}(g) - \mathbb{E} J_{R, n}(g)}{\sqrt{\text{Var} J_{R, n}(g)}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty, \quad (3.2.8)$$

for any connection function g .

To see this, observe that

$$J_{R, n}(g) = I_n(g_R),$$

as can be seen by direct computation. Since g_R has bounded support, Statement 3.2.3 follows. The moral of this is, that we should base the proof on $J_{R,n}(g)$ and $L_{R,n}(g)$ instead of $J_{n,R}(g)$ and $L_{n,R}(g)$. In the next section we show that the proof idea of [44] can still be carried out, although the computations involved are a little more complicated now.

3.3 Proof of Theorem 3.1.1

We start with a technical lemma, needed for applications of dominated convergence.

Lemma 3.3.1. *There exists N such that for $R > 0$, $n \geq N$ and $x \in \mathbb{R}^d$*

$$\left| (1 - g(|x|)) \left[p(\lambda_n/n^d, g_R)^2 p_{\lambda_n/n^d}^{g_R, g_R}(x, 0) - 2p(\lambda_n/n^d, g_R)^2 p(\lambda_n/n^d, g^R) p_{\lambda_n/n^d}^{g_R, g}(x, 0) + p(\lambda_n/n^d, g)^2 p_{\lambda_n/n^d}^{g, g}(x, 0) \right] - p(\lambda_n/n^d, g_R)^2 (1 - p(\lambda_n/n^d, g^R))^2 \right| \leq Cg(|x|/2), \quad (3.3.1)$$

where C is a constant not depending on x , n or R .

Proof. Since $p(\lambda_n/n^d, g_R)p(\lambda_n/n^d, g^R) = p(\lambda_n/n^d, g)$, the expression between the absolute value signs in (3.3.1) is equal to

$$\begin{aligned} & -g(|x|) \left[p(\lambda_n/n^d, g_R)^2 p_{\lambda_n/n^d}^{g_R, g_R}(x, 0) - 2p(\lambda_n/n^d, g_R)p(\lambda_n/n^d, g)p_{\lambda_n/n^d}^{g_R, g}(x, 0) + \right. \\ & \left. p(\lambda_n/n^d, g)^2 p_{\lambda_n/n^d}^{g, g}(x, 0) \right] + p(\lambda_n/n^d, g_R)^2 (p_{\lambda_n/n^d}^{g_R, g_R}(x, 0) - 1) - \\ & 2p(\lambda_n/n^d, g_R)p(\lambda_n/n^d, g)(p_{\lambda_n/n^d}^{g_R, g}(x, 0) - 1) + p(\lambda_n/n^d, g)^2 (p_{\lambda_n/n^d}^{g, g}(x, 0) - 1). \end{aligned} \quad (3.3.2)$$

Let N be such that $\frac{3}{4}\lambda \leq \lambda_n/n^d \leq \frac{3}{2}\lambda$, $n \geq N$. Then since

$$\int_{\mathbb{R}^d} g_R(|y|) + g(|y|) dy \geq 2 \int_{\mathbb{R}^d} g_R(|y|) dy \geq 2 \int_{\mathbb{R}^d} g_R(|y|)g(|y+x|) dy,$$

we have for $n \geq N$

$$\begin{aligned} & p(\lambda_n/n^d, g_R)p(\lambda_n/n^d, g)p_{\lambda_n/n^d}^{g_R, g}(x, 0) \\ & \leq e^{-\frac{3}{4}\lambda \int_{\mathbb{R}^d} g_R(|y|) + g(|y|) dy + \frac{3}{2}\lambda \int_{\mathbb{R}^d} g_R(|y-x|)g(|y|) dy} \leq 1. \end{aligned} \quad (3.3.3)$$

Also,

$$p(\lambda_n/n^d, g_R)^2 p_{\lambda_n/n^d}^{g_R, g_R}(x, 0) \leq 1, \quad p(\lambda_n/n^d, g)^2 p_{\lambda_n/n^d}^{g, g}(x, 0) \leq 1,$$

which follows from (3.3.3) by taking $g = g_R$ or letting $R \rightarrow \infty$ respectively. Hence for $n \geq N$ the absolute value of (3.3.2) is bounded by

$$4g(|x|) + 4(p_{2\lambda}^{g, g}(x, 0) - 1). \quad (3.3.4)$$

To give an upper bound for the second term in this expression, note that for

$y \in \mathbb{R}^d$ either $|y| \geq |x|/2$ or $|y - x| \geq |x|/2$, so

$$\begin{aligned} & \int_{\mathbb{R}^d} g(|y - x|)g(|y|) dy \\ & \leq \int_{|y| < |x|/2} g(|y - x|)g(|y|) dy + \int_{|y| \geq |x|/2} g(|y - x|)g(|y|) dy \\ & \leq g(|x|/2) \int_{|y| < |x|/2} g(|y|) dy + g(|x|/2) \int_{|y| \geq |x|/2} g(|y - x|) dy \\ & \leq 2g(|x|/2) \int_{\mathbb{R}^d} g(|y|) dy. \end{aligned}$$

Choose M such that $4\lambda g(M/2) \int_{\mathbb{R}^d} g(|y|) dy \leq 1$. Then since $e^t \leq 1 + et$, $t \leq 1$, we have for $|x| \geq M$

$$e^{4\lambda g(|x|/2) \int_{\mathbb{R}^d} g(|y|) dy} \leq 1 + 4e\lambda g(|x|/2) \int_{\mathbb{R}^d} g(|y|) dy.$$

For $|x| < M$ we have

$$e^{4\lambda g(|x|/2) \int_{\mathbb{R}^d} g(|y|) dy} \leq e^{4\lambda \int_{\mathbb{R}^d} g(|y|) dy} \leq 1 + g(|x|/2)g(M/2)^{-1}[e^{4\lambda \int_{\mathbb{R}^d} g(|y|) dy} - 1].$$

Combining the above inequalities yields

$$p_{2\lambda}^{g, g}(x, 0) - 1 \leq Cg(|x|/2), \quad (3.3.5)$$

where C is a constant not depending on x , n or R . We conclude that (3.3.4) is bounded by $4(1 + C)g(|x|/2)$. \square

Lemma 3.3.2.

$$\mathbb{E} L_{R,n}(g) = \lambda_n \ell(K) p(\lambda_n, g_{R,n})(1 - p(\lambda_n, g^{R,n})); \quad (3.3.6)$$

$$\begin{aligned} \text{Var} L_{R,n}(g) &= \lambda_n \ell(K) p(\lambda_n, g_{R,n})(1 - p(\lambda_n, g^{R,n})) + \lambda_n^2 \int_K \int_K (1 - g_n(|x_1 - x_2|)) \\ & \left[p(\lambda_n, g_{R,n})^2 p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) - 2p(\lambda_n, g_{R,n})^2 p(\lambda_n, g^{R,n}) p_{\lambda_n}^{g_{R,n}, g_n}(x_1, x_2) + \right. \\ & \left. p(\lambda_n, g_n)^2 p_{\lambda_n}^{g_n, g_n}(x_1, x_2) \right] - p(\lambda_n, g_{R,n})^2 (1 - p(\lambda_n, g^{R,n}))^2 dx_2 dx_1 + \\ & \lambda_n^2 p(\lambda_n, g_{R,n})^2 \int_K \int_K g^{R,n}(|x_1 - x_2|) p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) dx_2 dx_1. \quad (3.3.7) \end{aligned}$$

Proof. The first statement (3.3.6) is proved as in [44] Lemma 4.

For a Borel subset B of \mathbb{R}^d let $X_n(B)$ be the number of points in $X_n \cap B$. For $t > 0$ denote $K^t = K + \{x \in \mathbb{R}^d : |x| < t\}$. In the model (X_n, λ_n, g_n) let $L_{R,n,t}(g)$ be the number of points ξ in $X_n \cap K$ such that ξ is not connected to any point in $X_n \cap K^t$ at a distance R/n or less from ξ but ξ is connected to some point in $X_n \cap K^t$ at a distance greater than R/n from ξ . Since $L_{R,n,t}(g) \rightarrow L_{R,n}(g)$, $t \rightarrow \infty$, and $L_{R,n,t}(g) \leq X_n(K)$, $t > 0$, and $\mathbb{E} X_n(K)^2 < \infty$, the dominated convergence theorem gives

$$\mathbb{E} L_{R,n,t}(g) \rightarrow \mathbb{E} L_{R,n}(g), \quad \text{Var} L_{R,n,t}(g) \rightarrow \text{Var} L_{R,n}(g), \quad t \rightarrow \infty.$$

In order to compute the moments of $L_{R,n,t}(g)$, note that

$$L_{R,n,t}(g) \sim \sum_{i=1}^{X_n(K^t)} \mathbb{1}_{F_i},$$

where \sim denotes equality in distribution, ξ_i , $i \in \mathbb{N}$ are independent random variables, independent of $X_n(K^t)$, uniformly distributed on K^t and connected to each other according to g_n , and $F_i = \{\xi_i \in K; \xi_i \text{ is not connected to any } \xi_j, j \leq X_n(K^t), \text{ at a distance } R/n \text{ or less from } \xi_i; \xi_i \text{ is connected to some } \xi_j, j \leq X_n(K^t), \text{ at a distance greater than } R/n \text{ from } \xi_i\}$.

Since

$$L_{R,n,t}(g)^2 \sim \sum_{i=1}^{X_n(K^t)} \mathbb{1}_{F_i} + \sum_{i=1}^{X_n(K^t)} \sum_{\substack{j=1 \\ j \neq i}}^{X_n(K^t)} \mathbb{1}_{F_i} \mathbb{1}_{F_j},$$

the variance of $L_{R,n,t}(g)$ can be written as

$$\begin{aligned} \text{Var} L_{R,n,t}(g) &= \mathbb{E} L_{R,n,t}(g) + \\ &\sum_{m=2}^{\infty} m(m-1) \mathbb{P}(F_1 \cap F_2 \mid X_n(K^t) = m) \mathbb{P}(X_n(K^t) = m) - (\mathbb{E} L_{R,n,t}(g))^2. \end{aligned} \quad (3.3.8)$$

We have

$$\begin{aligned} &\mathbb{P}(F_1 \cap F_2 \mid \{\xi_1 \text{ is connected to } \xi_2\} \mid X_n(K^t) = m) \\ &= \frac{1}{\ell(K^t)^m} \int_K \int_K g^{R,n}(|x_1 - x_2|) \int_{K^t} \dots \int_{K^t} \\ &\quad \prod_{i=3}^m (1 - g_{R,n}(|x_i - x_1|))(1 - g_{R,n}(|x_i - x_2|)) dx_m \dots dx_3 dx_2 dx_1 \\ &= \frac{1}{\ell(K^t)^m} \int_K \int_K g^{R,n}(|x_1 - x_2|) \\ &\quad \left[\int_{K^t} (1 - g_{R,n}(|y - x_1|))(1 - g_{R,n}(|y - x_2|)) dy \right]^{m-2} dx_2 dx_1, \end{aligned}$$

whence

$$\begin{aligned} &\sum_{m=2}^{\infty} m(m-1) \mathbb{P}(F_1 \cap F_2 \cap \{\xi_1 \text{ is connected to } \xi_2\} \mid X_n(K^t) = m) \mathbb{P}(X_n(K^t) = m) \\ &= \lambda_n^2 \int_K \int_K g^{R,n}(|x_1 - x_2|) \sum_{m=0}^{\infty} \frac{e^{-\lambda_n \ell(K^t)} \lambda_n^m}{m!} \\ &\quad \left[\int_{K^t} (1 - g_{R,n}(|y - x_1|))(1 - g_{R,n}(|y - x_2|)) dy \right]^m dx_2 dx_1 \\ &= \lambda_n^2 \int_K \int_K g^{R,n}(|x_1 - x_2|) \\ &\quad e^{\lambda_n \int_{K^t} -g_{R,n}(|y-x_1|) - g_{R,n}(|y-x_2|) + g_{R,n}(|y-x_1|)g_{R,n}(|y-x_2|) dy} dx_2 dx_1 \\ &\rightarrow \lambda_n^2 p(\lambda_n, g_{R,n})^2 \int_K \int_K g^{R,n}(|x_1 - x_2|) p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) dx_2 dx_1, \end{aligned} \quad (3.3.9)$$

as $t \rightarrow \infty$, where we use the dominated convergence theorem.

Furthermore,

$$\begin{aligned}
& \mathbb{P}(F_1 \cap F_2 \cap \{\xi_1 \text{ is not connected to } \xi_2\} \mid X_n(K^t) = m) \\
&= \frac{1}{\ell(K^t)^m} \int_K \int_K (1 - g_n(|x_1 - x_2|)) \int_{K^t} \dots \int_{K^t} \\
& \quad \left[1 - \prod_{i=3}^m (1 - g^{R,n}(|x_i - x_1|)) \right] \prod_{i=3}^m (1 - g_{R,n}(|x_i - x_1|)) \\
& \quad \left[1 - \prod_{i=3}^m (1 - g^{R,n}(|x_i - x_2|)) \right] \prod_{i=3}^m (1 - g_{R,n}(|x_i - x_2|)) dx_m \dots dx_3 dx_2 dx_1.
\end{aligned} \tag{3.3.10}$$

Exactly as in [44] Lemma 4, one can now show that

$$\begin{aligned}
& \sum_{m=2}^{\infty} m(m-1) \mathbb{P}(F_1 \cap F_2 \cap \{\xi_1 \text{ is not connected to } \xi_2\} \mid X_n(K^t) = m) \mathbb{P}(X_n(K^t) = m) \\
& \rightarrow \lambda_n^2 \int_K \int_K (1 - g_n(|x_1 - x_2|)) \left[p(\lambda_n, g_{R,n})^2 p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) - \right. \\
& \quad \left. 2p(\lambda_n, g_{R,n})^2 p(\lambda_n, g^{R,n}) p_{\lambda_n}^{g_{R,n}, g^n}(x_1, x_2) + p(\lambda_n, g_n)^2 p_{\lambda_n}^{g_n, g^n}(x_1, x_2) \right] dx_2 dx_1,
\end{aligned} \tag{3.3.11}$$

as $t \rightarrow \infty$, where we use the dominated convergence theorem.

Combining (3.3.8), (3.3.6), (3.3.9) and (3.3.11) yields (3.3.7). \square

The following lemma replaces the incorrect Lemma 5 (equation (3.2.2) and (3.2.3) in our current chapter) of [44].

Lemma 3.3.3.

$$\lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \mathbb{E} L_{R,n}(g) = p(\lambda, g_R)(1 - p(\lambda, g^R)); \tag{3.3.12}$$

$$\begin{aligned}
& \lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \mathbb{V}ar L_{R,n}(g) = p(\lambda, g_R)(1 - p(\lambda, g^R)) + \lambda \int_{\mathbb{R}^d} (1 - g(|x|)) \\
& \quad \left[p(\lambda, g_R)^2 p_{\lambda}^{g_R, g^R}(x, 0) - 2p(\lambda, g_R)^2 p(\lambda, g^R) p_{\lambda}^{g^R, g}(x, 0) + \right. \\
& \quad \left. p(\lambda, g)^2 p_{\lambda}^{g, g}(x, 0) \right] - p(\lambda, g_R)^2 (1 - p(\lambda, g^R))^2 dx + \\
& \quad \lambda p(\lambda, g_R)^2 \int_{\mathbb{R}^d} g^R(|x|) p_{\lambda}^{g^R, g^R}(x, 0) dx.
\end{aligned} \tag{3.3.13}$$

Proof. Assertion (3.3.12) follows from (3.3.6) by direct computation. We will deduce (3.3.13) from (3.3.7). By the dominated convergence theorem

$$\begin{aligned}
& \frac{\lambda_n}{\ell(K)} p(\lambda_n, g_{R,n})^2 \int_K \int_K g^{R,n}(|x_1 - x_2|) p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) dx_2 dx_1 \\
&= \frac{\lambda_n}{n^d \ell(K)} p(\lambda_n, g_{R,n})^2 \int_K \int_{n(K-x_1)} g^R(|x_2|) p_{\lambda_n/n^d}^{g^R, g^R}(0, x_2) dx_2 dx_1 \\
&\rightarrow \lambda p(\lambda, g_R)^2 \int_{\mathbb{R}^d} g^R(|x|) p_{\lambda}^{g^R, g^R}(x, 0) dx, \quad n \rightarrow \infty.
\end{aligned} \tag{3.3.14}$$

Furthermore,

$$\begin{aligned}
& \frac{\lambda_n}{\ell(K)} \int_K \int_K (1 - g_n(|x_1 - x_2|)) \left[p(\lambda_n, g_{R,n})^2 p_{\lambda_n}^{g_{R,n}, g_{R,n}}(x_1, x_2) - \right. \\
& 2p(\lambda_n, g_{R,n})^2 p(\lambda_n, g^{R,n}) p_{\lambda_n}^{g_{R,n}, g^n}(x_1, x_2) + p(\lambda_n, g_n)^2 p_{\lambda_n}^{g_n, g^n}(x_1, x_2) \left. \right] - \\
& p(\lambda_n, g_{R,n})^2 (1 - p(\lambda_n, g^{R,n}))^2 dx_2 dx_1 \\
& = \frac{\lambda_n}{n^d \ell(K)} \int_K \int_{n(K-x_1)} (1 - g(|x_2|)) \left[p(\lambda_n/n^d, g_R)^2 p_{\lambda_n/n^d}^{g_R, g_R}(0, x_2) - \right. \\
& 2p(\lambda_n/n^d, g_R)^2 p(\lambda_n/n^d, g^R) p_{\lambda_n/n^d}^{g_R, g}(0, x_2) + p(\lambda_n/n^d, g)^2 p_{\lambda_n/n^d}^{g, g}(0, x_2) \left. \right] - \\
& p(\lambda_n/n^d, g_R)^2 (1 - p(\lambda_n/n^d, g^R))^2 dx_2 dx_1.
\end{aligned}$$

By Lemma 3.3.1 with $x = -x_2$, we can apply the dominated convergence theorem. Combining the result with (3.3.12) and (3.3.14) yields (3.3.13). \square

Corollary 3.3.4.

$$\lim_{R \rightarrow \infty} \lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \mathbb{E} L_{R,n}(g) = 0; \quad (3.3.15)$$

$$\lim_{R \rightarrow \infty} \lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \text{Var} L_{R,n}(g) = 0. \quad (3.3.16)$$

Proof. The dominated convergence theorem gives

$$\begin{aligned}
p(\lambda, g_R) &\rightarrow p(\lambda, g), & p(\lambda, g^R) &\rightarrow 1, \\
p_{\lambda}^{g_R, g_R}(x, 0) &\rightarrow p_{\lambda}^{g, g}(x, 0), & p_{\lambda}^{g_R, g}(x, 0) &\rightarrow p_{\lambda}^{g, g}(x, 0),
\end{aligned}$$

as $R \rightarrow \infty$. Now (3.3.15) follows from (3.3.12). Another application of the dominated convergence theorem yields

$$\int_{\mathbb{R}^d} g^R(|x|) p_{\lambda}^{g_R, g_R}(x, 0) dx \rightarrow 0, \quad R \rightarrow \infty.$$

Finally, the integrand in the first integral on the right hand side of (3.3.13) tends to 0 as $R \rightarrow \infty$. By Lemma 3.3.1 with $\lambda_n = \lambda n^d$, we can apply the dominated convergence theorem to conclude (3.3.16). \square

Finally, we can prove the main result:

Theorem 3.3.5. *If for $R > 0$*

$$\frac{J_{R,n}(g) - \mathbb{E} J_{R,n}(g)}{\sqrt{\text{Var} J_{R,n}(g)}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty, \quad (3.3.17)$$

then (3.1.1) holds.

Proof. Lemma 3 of [44] shows that

$$\lim_{n \rightarrow \infty} (\lambda_n \ell(K))^{-1} \text{Var} I_n(g) = p(\lambda, g) + \lambda p(\lambda, g)^2 \int_{\mathbb{R}^d} (1 - g(|x|)) p_{\lambda}^{g, g}(x, 0) - 1 dx. \quad (3.3.18)$$

It follows from (3.3.18), Corollary 3.3.4 and Chebyshev's inequality that

$$\lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{L_{R,n}(g) - \mathbb{E} L_{R,n}(g)}{\sqrt{\text{Var} I_n(g)}} \right| \geq \varepsilon \right) \leq \lim_{R \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\text{Var} L_{R,n}(g)}{\varepsilon^2 \text{Var} I_n(g)} = 0, \quad \varepsilon > 0.$$

Moreover, applying (3.3.18) also with g replaced by g_R gives

$$\lim_{n \rightarrow \infty} \frac{\text{Var} J_{R,n}(g)}{\text{Var} I_n(g)} = \delta_R,$$

where δ_R is a constant. (This was incorrectly claimed in [44] with $L_{n,R}(g)$ instead of $L_{R,n}(g)$.) Because

$$(1 - g_R(|x|)) p_\lambda^{g_R, g_R}(x, 0) - 1 \geq (1 - g_R(|x|)) \cdot 1 - 1 \geq -g(|x|)$$

and by (3.3.5)

$$(1 - g_R(|x|)) p_\lambda^{g_R, g_R}(x, 0) - 1 \leq 1 \cdot p_\lambda^{g, g}(x, 0) - 1 \leq Cg(|x|/2),$$

where C is a constant not depending on x or R , we have by the dominated convergence theorem $\lim_{R \rightarrow \infty} \delta_R = 1$. Now if (3.3.17) holds, then for $x \in \mathbb{R}$

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{P} \left(\frac{I_n(g) - \mathbb{E} I_n(g)}{\sqrt{\text{Var} I_n(g)}} \leq x \right) \\ & \leq \lim_{\varepsilon \downarrow 0} \lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\frac{J_{R,n}(g) - \mathbb{E} J_{R,n}(g)}{\sqrt{\text{Var} I_n(g)}} \leq x + \varepsilon \right) + \\ & \quad \mathbb{P} \left(\left| \frac{L_{R,n}(g) - \mathbb{E} L_{R,n}(g)}{\sqrt{\text{Var} I_n(g)}} \right| \geq \varepsilon \right) \\ & = \Phi(x). \end{aligned}$$

A similar argument yields

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\frac{I_n(g) - \mathbb{E} I_n(g)}{\sqrt{\text{Var} I_n(g)}} \leq x \right) \geq \Phi(x),$$

which completes the proof of the theorem. \square

3.4 Extension to larger components

In this section, we discuss larger components. A central limit theorem for larger components needs another approach, even when the connection function has bounded support. The reason for this is that the exact moment computations of the preceding sections no longer seem possible. At this point, we can only prove a central limit theorem when the connection function g has bounded support. For this, we use a result of [2], from which it follows that in order to prove a central limit theorem, certain mixing conditions suffice. For convenience, the central limit theorem in this section is stated a little different from the earlier ones, in the sense that we do not scale

the connection function and the density, but instead take larger and larger subsets of the space. This is equivalent to the case where $\lambda_n = \lambda n^d$ in the original setup.

For a subset Λ of \mathbb{Z}^d , let the inner boundary of Λ be denoted by $\partial\Lambda$, and its cardinality by $|\Lambda|$. Let the random variable $I^r(\Lambda) = I^r(\Lambda, g)$ be defined as $1/r$ times the number of vertices of (X, λ, g) in $\Lambda + (0, 1]^d$ that are contained in a component of size r . For $z \in \mathbb{Z}^d$ write $I^r(z) = I^r(\{z\})$. We will prove the following central limit theorem.

Theorem 3.4.1. *Consider a random connection model with connection function g of bounded support. Then for any increasing sequence $(\Lambda_n)_{n \in \mathbb{N}}$ of finite subsets of \mathbb{Z}^d with $\bigcup_{n \in \mathbb{N}} \Lambda_n = \mathbb{Z}^d$ and $|\partial\Lambda_n|/|\Lambda_n| \rightarrow 0$, $n \rightarrow \infty$, we have*

$$\frac{I^r(\Lambda_n) - \mathbb{E} I^r(\Lambda_n)}{\sqrt{\text{Var} I^r(\Lambda_n)}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty. \quad (3.4.1)$$

In order to prove this result, we use the main theorem in [2]. The conditions of his theorem involve three mixing conditions which are trivially satisfied when g has bounded support, and which we do not repeat here. Under these three mixing conditions, it is shown in [2] that if in addition

$$\sum_{z \in \mathbb{Z}^d} \text{Cov}(I^r(0), I^r(z)) > 0, \quad (3.4.2)$$

then it is the case that

$$\frac{I^r(\Lambda_n) - \mathbb{E} I^r(\Lambda_n)}{\sqrt{|\Lambda_n| \sum_{z \in \mathbb{Z}^d} \text{Cov}(I^r(0), I^r(z))}} \rightsquigarrow N(0, 1), \quad n \rightarrow \infty. \quad (3.4.3)$$

Because of the following elementary lemma, which we give without proof, (3.4.2) and (3.4.3) imply our Theorem 3.4.1.

Lemma 3.4.2. *Let $(Y_z)_{z \in \mathbb{Z}^d}$ be a stationary random field with $\mathbb{E} Y_0^2 < \infty$. Let $(\Lambda_n)_{n \in \mathbb{N}}$ be a sequence of finite non-empty subsets of \mathbb{Z}^d with $|\partial\Lambda_n|/|\Lambda_n| \rightarrow 0$, $n \rightarrow \infty$. If*

$$\sum_{z \in \mathbb{Z}^d} |\text{Cov}(Y_0, Y_z)| < \infty, \quad (3.4.4)$$

then

$$\frac{1}{|\Lambda_n|} \text{Var} \sum_{z \in \Lambda_n} Y_z \rightarrow \sum_{z \in \mathbb{Z}^d} \text{Cov}(Y_0, Y_z), \quad n \rightarrow \infty.$$

Note that (3.4.4) is satisfied because g has bounded support. It remains to prove (3.4.2). We give the proof in the two-dimensional case, but the method clearly generalizes to other dimensions.

With a slight abuse of notation, for a Borel subset B of \mathbb{R}^2 let $I^r(B)$ henceforth be defined as $1/r$ times the number of vertices of (X, λ, g) in B that are contained in a component of size r . According to Lemma 3.4.2, it suffices to show that there exists $M \in \mathbb{N}$ and $\gamma > 0$ such that for all n ,

$$\text{Var} I^r((0, nM]^2) \geq \gamma n^2. \quad (3.4.5)$$

We estimate the variance in (3.4.5) with the following general abstract trick, which we learned from Rob van den Berg (personal communication).

16	15	14	13					
9	8	7	12					
4	3	6	11					
1	2	5	10	17				

Figure 3.1: The enumeration of cubes in the first quadrant.

Lemma 3.4.3. *Let Y be a random variable with finite second moment, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $n \in \mathbb{N}$ and let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n$ be sub- σ -algebras of \mathcal{A} with $\mathbb{E}(Y|\mathcal{F}_0) = \mathbb{E}Y$ and $\mathbb{E}(Y|\mathcal{F}_n) = Y$ a.s. Then we have*

$$\text{Var}Y = \sum_{i=1}^n \mathbb{E} [\mathbb{E}(Y|\mathcal{F}_i) - \mathbb{E}(Y|\mathcal{F}_{i-1})]^2.$$

Proof. For $1 \leq i \leq n$, denote $\Delta_i = \mathbb{E}(Y|\mathcal{F}_i) - \mathbb{E}(Y|\mathcal{F}_{i-1})$. We write the variance of Y with a telescoping sum as $\text{Var}Y = \mathbb{E}(\sum_{i=1}^n \Delta_i)^2$. For $1 \leq i < j \leq n$ we have $\mathbb{E} \Delta_i \Delta_j = \mathbb{E} \mathbb{E}(\Delta_i Y|\mathcal{F}_j) - \mathbb{E} \mathbb{E}(\Delta_i Y|\mathcal{F}_{j-1}) = 0$. Hence $\text{Var}Y = \sum_{i=1}^n \mathbb{E} \Delta_i^2$, as required. \square

Let R be such that $g(x) = 0$, $x \geq R$. Define $\mu = \mathbb{E}I^r((0, 1]^2) > 0$. Choose an integer $M > 3\lambda R/\mu$. We will show that (3.4.5) holds for this M , and this is sufficient to prove Theorem 3.4.1.

Partition the first quadrant of \mathbb{R}^2 into cubes of side length M , and denote these cubes by B_k , $k \in \mathbb{N}$, where the indices run as indicated in Figure 3.1. For $n \in \mathbb{N}$ let K_n be the set of indices $k \in \{1, \dots, (n-1)^2\}$ that are shaded in Figure 3.1.

For $k \in \bigcup_{n \in \mathbb{N}} K_n$, we define the following sets:

$$\begin{aligned} C_k &= (rR, rR) + B_k; \\ D_k &= B_k + (-rR, rR]^2; \\ L_k &= D_k \cap \bigcup_{i=1}^{k-1} B_i; \\ U_k &= D_k \setminus \bigcup_{i=1}^{k-1} B_i; \end{aligned}$$

see Figure 3.2.

For $k \in \mathbb{N}$, let \mathcal{F}_k be the σ -algebra generated by the points of X in $\bigcup_{i=1}^k B_i$. We will first show that for $n \in \mathbb{N}$ and $k \in K_n$ the difference $\mathbb{E}(I^r((0, nM]^2)|\mathcal{F}_{k-1}) - \mathbb{E}(I^r((0, nM]^2)|\mathcal{F}_k)$ is bounded below by a positive uniform constant, with positive probability which is also uniform in k and n .

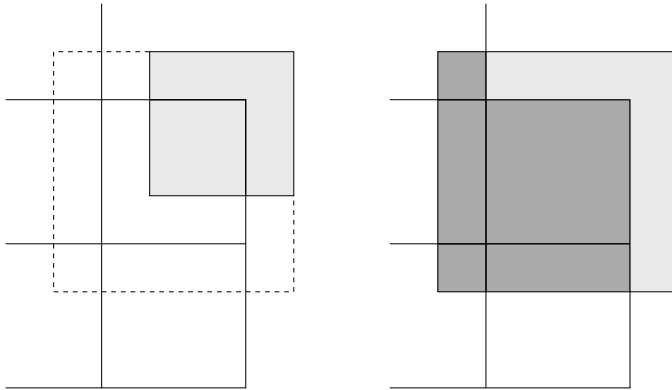


Figure 3.2: The shaded region on the left is C_k . The dark shaded region on the right is L_k and the light shaded region on the right is U_k .

On the one hand, we have

$$\begin{aligned} \mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_{k-1}) &\geq \mathbb{E}(I^r(C_k) | \mathcal{F}_{k-1}) + \mathbb{E}(I^r((0, nM]^2 \setminus D_k) | \mathcal{F}_{k-1}) \\ &= \mu M^2 + \mathbb{E}(I^r((0, nM]^2 \setminus D_k) | \mathcal{F}_k), \end{aligned} \quad (3.4.6)$$

since $I^r(C_k)$ is independent of \mathcal{F}_{k-1} and since the σ -algebra generated by $I^r((0, nM]^2 \setminus D_k)$ and the points of X in $\bigcup_{i=1}^{k-1} B_i$, is independent of the points of X in B_k .

On the other hand, we also have

$$\begin{aligned} \mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_k) &\leq (1/r) \mathbb{E}(X(L_k) | \mathcal{F}_k) + (1/r) \mathbb{E}(X(U_k) | \mathcal{F}_k) + \mathbb{E}(I^r((0, nM]^2 \setminus D_k) | \mathcal{F}_k) \\ &= 0 + 2\lambda R(M + rR) + \mathbb{E}(I^r((0, nM]^2 \setminus D_k) | \mathcal{F}_k), \end{aligned} \quad (3.4.7)$$

with probability at least $e^{-\lambda(M+2rR)^2}$, since $X(L_k)$ is \mathcal{F}_k -measurable and $X(U_k)$ is independent of \mathcal{F}_k .

Combining (3.4.6) and (3.4.7) yields for $n \in \mathbb{N}$ and $k \in K_n$,

$$\begin{aligned} \mathbb{P}(\mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_{k-1}) - \mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_k) \geq \mu M^2 - 2\lambda R(M + rR)) \\ \geq e^{-\lambda(M+2rR)^2}). \end{aligned}$$

Now observe that the box $(0, nM]^2$ contains at least αn^2 boxes indexed by an element of K_n , for some $\alpha > 0$. Hence, since $\mu M^2 - 2\lambda R(M + rR) > 0$, we have by Lemma 3.4.3

$$\begin{aligned} \text{Var} I^r((0, nM]^2) &\geq \sum_{k \in K_n} \mathbb{E}[\mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_k) - \mathbb{E}(I^r((0, nM]^2) | \mathcal{F}_{k-1})]^2 \\ &\geq \alpha n^2 (\mu M^2 - 2\lambda R(M + rR))^2 e^{-\lambda(M+2rR)^2}, \end{aligned}$$

proving the result.

Chapter 4

Fat fractal percolation and k -fractal percolation

This chapter is based on the paper [7] by Broman, Van de Brug, Camia, Joosten, and Meester.

4.1 Introduction

In [38] Mandelbrot introduced the following fractal percolation model. Let $N \geq 2$, $d \geq 2$ be integers and consider the unit cube $[0, 1]^d$. Divide the unit cube into N^d subcubes of side length $1/N$. Each subcube is retained with probability p and discarded with probability $1 - p$, independently of other subcubes. The closure of the union of the retained subcubes forms a random subset D_p^1 of $[0, 1]^d$. Next, each retained subcube in D_p^1 is divided into N^d cubes of side length $1/N^2$. Again, each smaller subcube is retained with probability p and discarded with probability $1 - p$, independently of other cubes. We obtain a new random set $D_p^2 \subset D_p^1$. Iterating this procedure in every retained cube at every smaller scale yields an infinite decreasing sequence of random subsets $D_p^1 \supset D_p^2 \supset D_p^3 \supset \dots$ of $[0, 1]^d$. We define the limit set $D_p := \bigcap_{n=1}^{\infty} D_p^n$. We will refer to this model as the Mandelbrot fractal percolation (MFP) model with parameter p .

It is easy to extend and generalize the classical Mandelbrot model in ways that preserve at least a certain amount of statistical self-similarity and generate random fractal sets. It is interesting to study such models to obtain a better understanding of general fractal percolation processes and explore possible new features that are not present in the MFP model. In this chapter we are concerned with two natural extensions which have previously appeared in the literature, as we mention below. We will next introduce the models and state our main results.

4.1.1 k -fractal percolation

Let $N \geq 2$ be an integer and divide the unit cube $[0, 1]^d$, $d \geq 2$, into N^d subcubes of side length $1/N$. Fix an integer $0 < k \leq N^d$ and retain k subcubes in a uniform

way, that is, all configurations where k cubes are retained have equal probability, other configurations have probability 0. Let D_k^1 denote the random set which is obtained by taking the closure of the union of all retained cubes. Iterating the described procedure in retained cubes and on all smaller scales yields a decreasing sequence of random sets $D_k^1 \supset D_k^2 \supset D_k^3 \supset \dots$. We are mainly interested in the connectivity properties of the limiting set $D_k := \bigcap_{n=1}^{\infty} D_k^n$. This model was called the *micro-canonical fractal percolation process* by Lincoln Chayes in [18] and both *correlated fractal percolation* and *k out of N^d fractal percolation* by Dekking and Don [21]. We will adopt the terms *k -fractal percolation* and *k -model*.

For $F \subset [0, 1]^d$, we say that the unit cube is *crossed by F* if there exists a connected component of F which intersects both $\{0\} \times [0, 1]^{d-1}$ and $\{1\} \times [0, 1]^{d-1}$. Define $\theta(k, N, d)$ as the probability that $[0, 1]^d$ is crossed by D_k . Similarly, $\sigma(p, N, d)$ denotes the probability that $[0, 1]^d$ is crossed by D_p . Let us define the critical probability $p_c(N, d)$ for the MFP model and the critical threshold value $k_c(N, d)$ for the k -model by

$$p_c(N, d) := \inf\{p : \sigma(p, N, d) > 0\}, \quad k_c(N, d) := \min\{k : \theta(k, N, d) > 0\}.$$

Let \mathbb{L}^d be the d -dimensional lattice with vertex set \mathbb{Z}^d and with edge set given by the adjacency relation: $(x_1, \dots, x_d) = x \sim y = (y_1, \dots, y_d)$ if and only if $x \neq y$, $|x_i - y_i| \leq 1$ for all i and $x_i = y_i$ for at least one value of i . Let $p_c(d)$ denote the critical probability for site percolation on \mathbb{L}^d . It is known (see [24]) that $p_c(N, d) \rightarrow p_c(d)$ as $N \rightarrow \infty$. We have the following analogous result for the k -model.

Theorem 4.1.1. *For all $d \geq 2$, we have that*

$$\lim_{N \rightarrow \infty} \frac{k_c(N, d)}{N^d} = p_c(d).$$

Remark 4.1.2. Note that the choice for the unit cube in the definitions of $\theta(k, N, d)$ and $\sigma(p, N, d)$ (and thus implicitly also in the definitions of $k_c(N, d)$ and $p_c(N, d)$) is rather arbitrary: We could define them in terms of crossings of other shapes such as annuli, for example, and obtain the same conclusion, i.e. $k_c(N, d)/N^d \rightarrow p_c(d)$ as $N \rightarrow \infty$, where $\theta(k, N, d)$ and $k_c(N, d)$ are defined using the probability that D_k crosses an annulus. One advantage of using annuli is that the percolation function $\sigma(p, N, d)$ is known to have a discontinuity at $p_c(N, d)$ for all N, d and any choice of annulus [6, Corollary 2.6]. (This is known to be the case also when $p_c(N, d)$ is defined using the unit cube if $d = 2$ [16, 20], but for $d \geq 3$ it is proven only for N sufficiently large [5].) In the present chapter we stick to the “traditional” choice of the unit cube.

Remark 4.1.3. For the MFP model it is the case that, for $p > p_c(d)$,

$$\sigma(p, N, d) \rightarrow 1, \tag{4.1.1}$$

as $N \rightarrow \infty$. This is part (b) of Theorem 2 in [24]. During the course of the proof of Theorem 4.1.1 we will prove a similar result for the k -model, see Theorem 4.3.2.

Next, consider the following generalization of both the k -model and the MFP model. Let $d \geq 2, N \geq 2$ be integers and let $Y = Y(N, d)$ be a random variable taking values in $\{0, \dots, N^d\}$. Divide the unit cube into N^d smaller cubes of side

length $1/N$. Draw a realization y according to Y and retain y cubes uniformly. Let D_Y^1 denote the closure of the union of the retained cubes. Next, every retained cube is divided into N^d smaller subcubes of side length $1/N^2$. Then, for every subcube C in D_Y^1 (where we slightly abuse notation by viewing D_Y^1 as the set of retained cubes in the first iteration step) draw a new (independent) realization $y(C)$ of Y and retain $y(C)$ subcubes in C uniformly, independently of all other subcubes. Denote the closure of the union of retained subcubes by D_Y^2 . Repeat this procedure in every retained subcube at every smaller scale and define the limit set $D_Y := \bigcap_{n=1}^{\infty} D_Y^n$. We will call this model the *generalized fractal percolation model* (GFP model) with generator Y . Define $\phi(Y, N, d)$ as the probability of the event that $[0, 1]^d$ is crossed by D_Y .

By taking Y equal to an integer k , resp. to a binomially distributed random variable with parameters N^d and p , we obtain the k -model, resp. the MFP model with parameter p . If Y is stochastically dominated by a binomial random variable with parameters N^d and p , where $p < p_c(N, d)$, then by standard coupling techniques it follows that $\phi(Y, N, d) = 0$. Likewise, if $Y(N, d)$ dominates a binomial random variable with parameters N^d and p , where $p > p_c(d)$, then $\phi(Y(N, d), N, d) \geq \sigma(p, N, d) \rightarrow 1$ as $N \rightarrow \infty$, as mentioned in Remark 4.1.3. The following theorem, which generalizes (4.1.1), shows that the latter conclusion still holds if for some $p > p_c(d)$, $\mathbb{P}(Y(N, d) \geq pN^d) \rightarrow 1$ as $N \rightarrow \infty$.

Theorem 4.1.4. *Consider the GFP model with generator $Y(N, d)$. Let $p > p_c(d)$. Suppose that $\mathbb{P}(Y(N, d) \geq pN^d) \rightarrow 1$ as $N \rightarrow \infty$. Then*

$$\lim_{N \rightarrow \infty} \phi(Y(N, d), N, d) = 1.$$

Remark 4.1.5. Observe that by Chebyshev's inequality the condition of Theorem 4.1.4 is satisfied if, for some $p > p_c(d)$, $\mathbb{E}Y(N, d) \geq pN^d$ for all $N \geq 2$ and $\text{Var}(Y(N, d))/N^{2d} \rightarrow 0$ as $N \rightarrow \infty$.

Open problem 4.1.6. *It is a natural question to ask whether a “symmetric version” of Theorem 4.1.4 is true. That is, if e.g. $\mathbb{P}(Y(N, d) \leq pN^d) \rightarrow 1$ as $N \rightarrow \infty$, for some $p < p_c(d)$, implies $\phi(Y(N, d), N, d) \rightarrow 0$ as $N \rightarrow \infty$. The proof of Theorem 4.1.4 can not be adapted to this situation.*

4.1.2 Fat fractal percolation

Let $(p_n)_{n \geq 1}$ be a non-decreasing sequence in $(0, 1]$ such that $\prod_{n=1}^{\infty} p_n > 0$. We call *fat fractal percolation* a model analogous to the MFP model, but where at every iteration step n a subcube is retained with probability p_n and discarded with probability $1 - p_n$, independently of other subcubes. Iterating this procedure yields a decreasing sequence of random subsets $D_{\text{fat}}^1 \supset D_{\text{fat}}^2 \supset D_{\text{fat}}^3 \supset \dots$ and we will mainly study connectivity properties of the limit set $D_{\text{fat}} := \bigcap_{n=1}^{\infty} D_{\text{fat}}^n$. In [17] it is shown that if $p_n \rightarrow 1$ and $\prod_{n=1}^{\infty} p_n = 0$, then the limit set does not contain a directed crossing from left to right.

For a point $x \in D_{\text{fat}}$, let C_{fat}^x denote its *connected component*:

$$C_{\text{fat}}^x := \{y \in D_{\text{fat}} : y \text{ connected to } x \text{ in } D_{\text{fat}}\}.$$

We define the set of “dust” points by $D_{\text{fat}}^d := \{x \in D_{\text{fat}} : C_{\text{fat}}^x = \{x\}\}$. Define $D_{\text{fat}}^c := D_{\text{fat}} \setminus D_{\text{fat}}^d$, which is the union of connected components larger than one

point. Let λ denote the d -dimensional Lebesgue measure. It is easy to prove that $\lambda(D_{\text{fat}}) > 0$ with positive probability, see Proposition 4.4.1. Moreover, we can show that the Lebesgue measure of the limit set is positive a.s. given non-extinction, i.e. $D_{\text{fat}} \neq \emptyset$.

Theorem 4.1.7. *We have that $\lambda(D_{\text{fat}}) > 0$ a.s. given non-extinction.*

It is a natural question to ask whether both D_{fat}^c and D_{fat}^d have positive Lebesgue measure. The following theorem shows that they cannot simultaneously have positive Lebesgue measure.

Theorem 4.1.8. *Given non-extinction of the fat fractal process, it is the case that either*

$$\lambda(D_{\text{fat}}^d) = 0 \text{ and } \lambda(D_{\text{fat}}^c) > 0 \text{ a.s.} \quad (4.1.2)$$

or

$$\lambda(D_{\text{fat}}^d) > 0 \text{ and } \lambda(D_{\text{fat}}^c) = 0 \text{ a.s.} \quad (4.1.3)$$

Part (ii) of the following theorem gives a sufficient condition under which (4.1.2) holds. Furthermore, the theorem shows that the limit set either has an empty interior a.s. or can be written as the union of finitely many cubes a.s.

Theorem 4.1.9. *We have that*

- (i) *If $\prod_{n=1}^{\infty} p_n^{N^{dn}} = 0$, then D_{fat} has an empty interior a.s.;*
- (ii) *If $\prod_{n=1}^{\infty} p_n^{N^n} > 0$, then $\lambda(D_{\text{fat}}^d) = 0$ a.s.;*
- (iii) *If $\prod_{n=1}^{\infty} p_n^{N^{dn}} > 0$, then D_{fat} can be written as the union of finitely many cubes a.s.*

Open problem 4.1.10. *Part (ii) of Theorem 4.1.9 shows that if $\prod_{n=1}^{\infty} p_n^{N^n} > 0$, then (4.1.2) holds. However, we do not have an example for which (4.1.3) holds, and we do not know whether (4.1.3) is possible at all.*

In two dimensions, we have the following characterizations of $\lambda(D_{\text{fat}}^c)$ being positive a.s. given non-extinction of the fat fractal process.

Theorem 4.1.11. *Let $d = 2$. The following statements are equivalent.*

- (i) *$\lambda(D_{\text{fat}}^c) > 0$ a.s., given non-extinction of the fat fractal process;*
- (ii) *There exists a set $U \subset [0, 1]^2$ with $\lambda(U) > 0$ such that for all $x, y \in U$ it is the case that $\mathbb{P}(x \text{ is in the same connected component as } y) > 0$;*
- (iii) *There exists a set $U \subset [0, 1]^2$ with $\lambda(U) = 1$ such that for all $x, y \in U$ it is the case that $\mathbb{P}(x \text{ is in the same connected component as } y) > 0$.*

Let us now outline the rest of the chapter. The next section will be devoted to a formal introduction of the fractal percolation processes in the unit cube. We also define an ordering on the subcubes which will facilitate the proofs of Theorems 4.1.1 and 4.1.4 in Section 4.3. In Section 4.4 we prove our results concerning fat fractal percolation.

4.2 Preliminaries

In this section we set up an ordering for the subcubes of the fractal processes in the unit cube which will turn out to be very useful during the course of the proofs. We also give a formal probabilistic definition of the different fractal percolation models. We follow [24] almost verbatim in this section; a simple reference to [24] would however not be very useful for the reader, so we repeat some definitions here.

Order $J^d := \{0, 1, \dots, N-1\}^d$ in some way, say lexicographically by coordinates. For a positive integer n , write $J^{d,n} := \{(\mathbf{i}_1, \dots, \mathbf{i}_n) : \mathbf{i}_j \in J^d, 1 \leq j \leq n\}$ for the set of n -vectors with entries in J^d . Set $J^{d,0} := \{\emptyset\}$. With $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_n) = ((i_{1,1}, \dots, i_{1,d}), \dots, (i_{n,1}, \dots, i_{n,d}))$ we associate the subcube of $[0, 1]^d$ given by

$$C(\mathbf{I}) = c(\mathbf{I}) + [0, N^{-n}]^d,$$

where

$$c(\mathbf{I}) = \left(\sum_{j=1}^n N^{-j} i_{j,1}, \dots, \sum_{j=1}^n N^{-j} i_{j,d} \right)$$

and $c(\emptyset)$ is defined to be the origin. Such a cube $C(\mathbf{I})$ is called a *level- n cube* and we write $|\mathbf{I}| = n$. A concatenation of $\mathbf{I} \in J^{d,n}$ and $\mathbf{j} \in J^d$ is denoted by (\mathbf{I}, \mathbf{j}) , which is in $J^{d,n+1}$. We define the set of indices for all cubes until (inclusive) level- n as $\mathcal{J}^{(n)} := J^{d,0} \cup J^{d,1} \cup \dots \cup J^{d,n}$ and we order them in the following way. We declare $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_a) < \mathbf{I}' = (\mathbf{i}'_1, \dots, \mathbf{i}'_b)$ if and only if

- either $\mathbf{i}_r < \mathbf{i}'_r$ (according to the order on J^d) where $r \leq \min\{a, b\}$ is the smallest index so that $\mathbf{i}_r \neq \mathbf{i}'_r$ holds;
- or $a > b$ and $\mathbf{i}_r = \mathbf{i}'_r$ for $r = 1, \dots, b$.

To clarify this ordering we give a short example, see Figure 4.1. Suppose $N = 2$, $d = 2$ and J^2 is ordered by $(1, 1) > (1, 0) > (0, 1) > (0, 0)$, then the ordering of $\mathcal{J}^{(2)}$ starts with

$$\begin{aligned} \emptyset &> ((1, 1)) > ((1, 1), (1, 1)) > ((1, 1), (1, 0)) \\ &> ((1, 1), (0, 1)) > ((1, 1), (0, 0)) > ((1, 0)) > \dots \end{aligned}$$

We introduce the following formal probabilistic definition of the fractal percolation models. As noted before, the k -model and MFP model can be obtained from the GFP model with generator Y by setting $Y \equiv k$, resp. Y binomially distributed with parameters N^d and $p \in [0, 1]$. Therefore, we only provide a formal probabilistic definition of the GFP model and the fat fractal percolation model. Define the index set $\mathcal{J} := \bigcup_{n=0}^{\infty} J^{d,n}$. We define a family of random variables $\{Z_{\text{model}}(\mathbf{I})\}$, where $\mathbf{I} \in \mathcal{J}$ and – here as well as in the rest of the section – “model” stands for either p , fat, k or Y .

1. GFP model with generator Y : For every $\mathbf{I} \in \mathcal{J}$, let $y(\mathbf{I})$ denote a realization of Y , independently of other \mathbf{I}' . We define $J(\mathbf{I})$ as a uniform choice of $y(\mathbf{I})$ different indices of J^d , independently of other $J(\mathbf{I}')$. For $\mathbf{j} \in J^d$ define

$$Z_Y(\mathbf{I}, \mathbf{j}) = \begin{cases} 1, & \mathbf{j} \in J(\mathbf{I}), \\ 0, & \text{otherwise.} \end{cases}$$

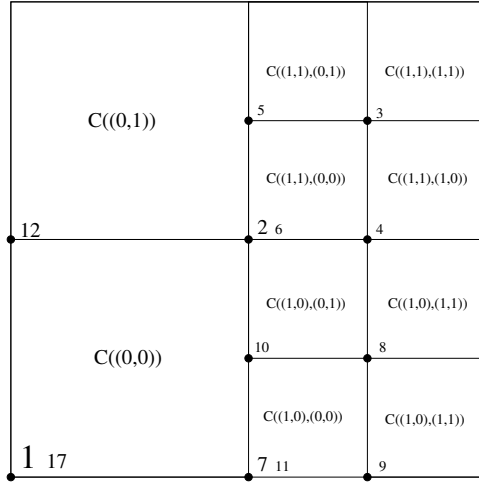


Figure 4.1: Illustration of the ordering of subcubes in $\mathcal{J}^{(2)}$, for $N = 2$ and $d = 2$. A black dot denotes the corner point $c(\mathbf{I})$ of a subcube $C(\mathbf{I})$. The number in the lower left corner of a subcube indicates the rank of the subcube in the ordering: e.g. the unit cube, i.e. $C(\emptyset)$, has rank 1 and $C((0,0))$ has rank 17.

2. Fat fractal percolation with parameters $(p_n)_{n \geq 1}$: For every $\mathbf{I} \in \mathcal{J}$ and $\mathbf{j} \in J^d$, let $n = |\mathbf{I}|$ and define

$$Z_{\text{fat}}(\mathbf{I}, \mathbf{j}) = \begin{cases} 1, & \text{with probability } p_{n+1}, \\ 0, & \text{with probability } 1 - p_{n+1}, \end{cases}$$

independently of all other $Z_{\text{fat}}(\mathbf{I}')$.

For each $\mathbf{I} \in \mathcal{J}$ we define the indicator function $\mathbf{1}_{\text{model}}(\mathbf{I})$ by

$$\mathbf{1}_{\text{model}}(\emptyset) = 1, \quad \mathbf{1}_{\text{model}}(\mathbf{I}) = Z_{\text{model}}(\mathbf{i}_1)Z_{\text{model}}(\mathbf{i}_1, \mathbf{i}_2) \cdots Z_{\text{model}}(\mathbf{I}),$$

where $\mathbf{I} = (\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_n) \in J^{d,n}$. We retain the subcube $C(\mathbf{I})$ if $\mathbf{1}_{\text{model}}(\mathbf{I}) = 1$ and we write D_{model}^n for the set of retained level- n cubes. Note that $D_{\text{model}}^1, D_{\text{model}}^2, \dots$ correspond to the sets informally constructed in the introduction. We denote by $\mathbb{P}_{\text{model}}$ the distribution of the corresponding model on $\Omega = \{0,1\}^{\mathcal{C}}$, where $\mathcal{C} := \{C(\mathbf{I}) : \mathbf{I} \in \mathcal{J}\}$ denotes the collection of all subcubes, endowed with the usual sigma algebra generated by the cylinder events. To simplify the notation, we will drop the subscripts fat, k, p, Y when there is no danger of confusion.

4.3 Proofs of the k -fractal results

In this section we prove Theorem 4.1.1 and Theorem 4.1.4. The proof of Theorem 4.1.1 is divided in two parts. First we treat the subcritical case and show that $\liminf_{N \rightarrow \infty} k_c(N, d)/N^d \geq p_c(d)$.

Theorem 4.3.1. *Consider the k -model. We have*

$$\liminf_{N \rightarrow \infty} k_c(N, d)/N^d \geq p_c(d).$$

In the supercritical case, we prove that the crossing probability converges to 1 as $N \rightarrow \infty$. Again, for future reference we state this as a theorem.

Theorem 4.3.2. *Let $p > p_c(d)$ and let $(k(N))_{N \geq 2}$ be a sequence of integers such that $k(N)/N^d \geq p$, for all $N \geq 2$. We have*

$$\lim_{N \rightarrow \infty} \theta(k(N), N, d) = 1.$$

Theorem 4.1.1 follows immediately from these two theorems.

We prove Theorems 4.3.1 and 4.3.2 in Sections 4.3.1 and 4.3.2, respectively. In Section 4.3.3 we prove Theorem 4.1.4, using the idea of the proof of Theorem 4.3.1 and the result of Theorem 4.3.2.

4.3.1 Proof of Theorem 4.3.1

Let $p < p_c(d)$ and consider a sequence $(k(N))_{N \geq 2}$ such that $k(N)/N^d \leq p$, for all $N \geq 2$, and $k(N)/N^d \rightarrow p$ as $N \rightarrow \infty$. Our goal is to show that the probability that the unit cube is crossed by $D_{k(N)}$, is equal to zero for all N large enough. Let $N \geq 2$ and let D_{p_0} be the limit set of an MFP process with parameters p_0 and N , where $p < p_0 < p_c(d)$. First, part (a) of Theorem 2 in [24] states that

$$p_c(d) \leq p_c(N, d), \tag{4.3.1}$$

for all N . Hence, the MFP process with parameter $p_0 < p_c(d)$ is subcritical. Therefore, a natural approach to prove that the probability that $D_{k(N)}$ crosses the unit cube equals zero for N large enough would be to couple the limit set $D_{k(N)}$ to the limit set D_{p_0} in such a way that $D_{k(N)} \subset D_{p_0}$. However, a “direct” coupling between the limit sets $D_{k(N)}$ and D_{p_0} is not possible, since with fixed positive probability at each iteration of the MFP process the number of retained subcubes is less than $k(N)$. We therefore need to find a more refined coupling.

The following is an informal strategy of the proof. We will define an event E on which the MFP process contains an infinite tree of retained subcubes, such that each subcube in this tree contains at least $k(N)$ retained subcubes in the tree. Next, we perform a construction of two auxiliary random subsets of the unit cube, from which it will follow that the law of $D_{k(N)}$ is stochastically dominated by the conditional law of D_{p_0} , conditioned on the event E . In particular, the probability that $D_{k(N)}$ crosses $[0, 1]^d$ is less than or equal to the conditional probability that D_{p_0} crosses the unit cube, given E . The latter probability is zero for N large enough, since the event E has positive probability for N large enough and the MFP process is subcritical.

Let us start by defining the event E . Consider an MFP process with parameters p_0 and N . For notational convenience we call the unit cube the level-0 cube. A level- n cube, $n \geq 0$, is declared θ -good if it is retained and contains at least $k(N)$ retained level- $(n+1)$ subcubes. (We adopt the convention that $[0, 1]^d$ is automatically retained.) Recursively, we define the notion m -good, for $m \geq 0$. A level- n cube, for

$n \geq 0$, is $(m+1)$ -good if it is retained and contains at least $k(N)$ m -good subcubes. We say that the unit cube is ∞ -good if it is m -good for every $m \geq 0$. Define the following events

$$\begin{aligned} E_m &:= \{[0, 1]^d \text{ is } m\text{-good}\}, \\ E &:= \{[0, 1]^d \text{ is } \infty\text{-good}\}. \end{aligned} \quad (4.3.2)$$

The following lemma states that we can make the probability of E arbitrary close to 1, for N large enough. In particular, E has positive probability for large enough N , which will be sufficient for the proof of Theorem 4.3.1.

Lemma 4.3.3. *Let $p_0 < p_c(d)$. Let $(k(N))_{N \geq 2}$ be a sequence of integers satisfying $\limsup_{N \rightarrow \infty} k(N)/N^d < p_0$. Consider an MFP model with parameters p_0 and N . For all $\varepsilon > 0$ there exists N_0 such that $\mathbb{P}_{p_0}(E) > 1 - \varepsilon$ for all $N \geq N_0$.*

Proof. Let $\delta > 0$ and N_0 be such that $k(N)/N^d \leq p_0 - 2\delta =: p$ for all $N \geq N_0$. Choose $N_1 \geq N_0$ so large that $p_0/(4\delta^2 N^d) < \delta$ for $N \geq N_1$. We will show that

$$\mathbb{P}_{p_0}(E_m) \geq 1 - \frac{1}{4\delta^2 N^d}, \quad (4.3.3)$$

for all $m \geq 0$ and $N \geq N_1$. Since E_m decreases to E as $m \rightarrow \infty$, it follows that

$$\mathbb{P}_{p_0}(E) = \lim_{m \rightarrow \infty} \mathbb{P}_{p_0}(E_m) \geq 1 - \frac{1}{4\delta^2 N^d},$$

for $N \geq N_1$. Now take $N_2 \geq N_1$ so large that $1 - \frac{1}{4\delta^2 N^d} > 1 - \varepsilon$ for all $N \geq N_2$. It remains to show (4.3.3).

We prove (4.3.3) by induction on m . Consider the event E_0 , i.e. the event that the unit cube contains at least $k(N)$ retained level-1 subcubes. Let $X(n, p)$ denote a binomially distributed random variable with parameters $n \in \mathbb{N}$ and $p \in [0, 1]$. Since the number of retained level-1 cubes has a binomial distribution with parameters N^d and p_0 , it follows from Chebyshev's inequality that, for every $N \geq N_1$, we have (writing \mathbb{P} for the probability measure governing the binomially distributed random variables)

$$\begin{aligned} \mathbb{P}_{p_0}(E_0) &= \mathbb{P}(X(N^d, p_0) \geq k(N)) \\ &\geq \mathbb{P}(X(N^d, p_0) \geq pN^d) \\ &\geq 1 - \frac{\text{Var}X(N^d, p_0)}{4\delta^2 N^{2d}} \\ &= 1 - \frac{p_0(1-p_0)N^d}{4\delta^2 N^{2d}} \\ &\geq 1 - \frac{1}{4\delta^2 N^d}. \end{aligned}$$

Next, let $m \geq 0$ and $N \geq N_1$ and suppose that (4.3.3) holds for this m and N . Recall that E_{m+1} is the event that the unit cube contains at least $k(N)$ m -good level-1 cubes. The probability that a level-1 cube is m -good, given that it is retained, is equal to $\mathbb{P}_{p_0}(E_m)$. Using the induction hypothesis, we get

$$\begin{aligned} \mathbb{P}_{p_0}(E_{m+1}) &= \mathbb{P}(X(N^d, p_0 \mathbb{P}_{p_0}(E_m)) \geq k(N)) \\ &\geq \mathbb{P}(X(N^d, p_0(1 - \frac{1}{4\delta^2 N^d})) \geq k(N)). \end{aligned}$$

By our choices for δ and N it follows that $p_0(1 - \frac{1}{4\delta^2 N^d}) > p + \delta$. Hence, using Chebyshev's inequality, we get

$$\begin{aligned} \mathbb{P}(X(N^d, p_0(1 - \frac{1}{4\delta^2 N^d})) \geq k(N)) &\geq \mathbb{P}(X(N^d, p + \delta) \geq k(N)) \\ &\geq \mathbb{P}(X(N^d, p + \delta) \geq pN^d) \\ &\geq 1 - \frac{\text{Var}X(N^d, p + \delta)}{\delta^2 N^{2d}} \\ &\geq 1 - \frac{1}{4\delta^2 N^d}. \end{aligned}$$

Therefore, the induction step is valid and we have proved (4.3.3). \square

Proof of Theorem 4.3.1. Let p, p_0 be such that $p < p_0 < p_c(d)$. Let $(k(N))_{N \geq 2}$ be a sequence such that $k(N)/N^d \leq p$, for all $N \geq 2$, and $k(N)/N^d \rightarrow p$ as $N \rightarrow \infty$. Consider an MFP model with parameters p_0 and N and define the event E as in (4.3.2). Henceforth, we assume that N is so large that $\mathbb{P}_{p_0}(E) > 0$, which is possible by Lemma 4.3.3. In order to prove Theorem 4.3.1 we will use E to construct two random subsets, \tilde{D}_{p_0} and $\tilde{D}_{k(N)}$, of the unit cube, on a common probability space and with the following properties:

- (i) $\tilde{D}_{k(N)} \subset \tilde{D}_{p_0}$;
- (ii) the law of \tilde{D}_{p_0} is stochastically dominated by the conditional law of D_{p_0} , conditioned on the event E ;
- (iii) the law of $\tilde{D}_{k(N)}$ is the same as the law of $D_{k(N)}$.

It follows that the law of $D_{k(N)}$ is stochastically dominated by the conditional law of D_{p_0} , conditioned on the event E . Hence, the probability that the unit cube is crossed by $D_{k(N)}$ is at most the conditional probability that D_{p_0} crosses the unit cube, conditioned on the event E . By (4.3.1) the MFP process with parameter p_0 is subcritical, thus the latter probability equals zero. Using the fact that $k(N)/N^d \rightarrow p$ as $N \rightarrow \infty$, we conclude that

$$\liminf_{N \rightarrow \infty} \frac{k_c(N, d)}{N^d} \geq p.$$

Since $p < p_c(d)$ was arbitrary, we get

$$\liminf_{N \rightarrow \infty} \frac{k_c(N, d)}{N^d} \geq p_c(d).$$

It remains to construct random sets $\tilde{D}_{p_0}, \tilde{D}_{k(N)}$ with the properties (i)-(iii). First we construct two sequences $(\tilde{D}_{p_0}^n)_{n \geq 1}, (\tilde{D}_{k(N)}^n)_{n \geq 1}$ of decreasing random subsets. Let \mathcal{L} be the conditional law of the number of ∞ -good level-1 cubes of the MFP process, conditioned on the event E . Note that the support of \mathcal{L} is $\{k(N), k(N) + 1, \dots, N^d\}$. Furthermore, for a fixed level- n cube $C(\mathbf{I})$, \mathcal{L} is also equal to the conditional law of the number of ∞ -good level- $(n+1)$ subcubes in $C(\mathbf{I})$, conditioned on $C(\mathbf{I})$ being ∞ -good.

Choose an integer l according to \mathcal{L} and choose l level-1 cubes uniformly. Define $\tilde{D}_{p_0}^1$ as the closure of the union of these l level-1 cubes. Choose $k(N)$ out of these l cubes in a uniform way and define $\tilde{D}_{k(N)}^1$ as the closure of the union of these $k(N)$ cubes. For each level-1 cube $C(\mathbf{I}) \subset \tilde{D}_{p_0}^1$, pick an integer $l(\mathbf{I})$ according to \mathcal{L} , independently of other cubes, and choose $l(\mathbf{I})$ level-2 subcubes of $C(\mathbf{I})$ in a uniform way. Define $\tilde{D}_{p_0}^2$ as the closure of the union of all selected level-2 cubes. For each level-1 cube $C(\mathbf{I}) \subset \tilde{D}_{k(N)}^1$, uniformly choose $k(N)$ out of the $l(\mathbf{I})$ selected level-2 subcubes. Define $\tilde{D}_{k(N)}^2$ as the closure of the union of the $k(N)^2$ selected level-2 cubes of $C(\mathbf{I})$. Iterating this procedure yields two infinite decreasing sequences of random subsets $(\tilde{D}_{p_0}^n)_{n \geq 1}, (\tilde{D}_{k(N)}^n)_{n \geq 1}$.

Now define

$$\tilde{D}_{p_0} := \bigcap_{n=1}^{\infty} \tilde{D}_{p_0}^n, \quad \tilde{D}_{k(N)} := \bigcap_{n=1}^{\infty} \tilde{D}_{k(N)}^n.$$

By construction, for each $n \geq 1$, we have that (1) $\tilde{D}_{k(N)}^n \subset \tilde{D}_{p_0}^n$, (2) the law of $\tilde{D}_{p_0}^n$ is stochastically dominated by the conditional law of $D_{p_0}^n$ given E and (3) the law of $\tilde{D}_{k(N)}^n$ is equal to the law of $D_{k(N)}^n$. It follows that the limit sets $\tilde{D}_{p_0}, \tilde{D}_{k(N)}$ satisfy properties (i)-(iii). \square

4.3.2 Proof of Theorem 4.3.2

Let us start by outlining the proof. The first part consists mainly of setting up the framework, where we use the notation of Falconer and Grimmett [24], which will enable us in the second part to prove that the subcubes of the fractal process satisfy certain “good” properties with probability arbitrarily close to 1 as $N \rightarrow \infty$. Informally, a subcube is good when there exist many connections inside the cube between its faces and when it is also connected to other good subcubes. Therefore, the probability of crossing the unit cube converges to 1 as $N \rightarrow \infty$.

Although we will partly follow [24], it does not seem possible to use Theorem 2.2 of [24] directly. First, we state (a slightly adapted version of) Lemma 2 of [24], which concerns site percolation with parameter π on \mathbb{L}^d . We let every vertex of \mathbb{L}^d be colored *black* with probability π and *white* otherwise, independently of other vertices. We write P_π for the ensuing product measure with density $\pi \in [0, 1]$. We call a subset C of \mathbb{L}^d a *black cluster* if it is a maximal connected subset (with respect to the adjacency relation on \mathbb{L}^d) of black vertices. Denote the cube with vertex set $\{1, 2, \dots, N\}^d$ by B_N . Let \mathcal{L} be the set of edges of the unit cube $[0, 1]^d$, that is \mathcal{L} contains all sets of the form

$$L_r(\mathbf{a}) = \{a_1\} \times \{a_2\} \times \cdots \times \{a_{r-1}\} \times [0, 1] \times \{a_{r+1}\} \times \cdots \times \{a_d\}$$

as r ranges over $\{1, \dots, d\}$ and $\mathbf{a} = (a_1, a_2, \dots, a_d)$ ranges over $\{0, 1\}^d$. For each $L = L_r(\mathbf{a}) \in \mathcal{L}$ we write

$$L_N = \{\mathbf{x} \in B_N : x_i = \max\{1, a_i N\} \text{ for } 1 \leq i \leq d, i \neq r\}$$

for the corresponding edge of B_N .

Lemma 4.3.4. *Suppose $\pi > p_c(d), \varepsilon > 0$ and let q be a positive integer. There exist positive integers u and N_1 such that the following holds for all $N \geq N_1$. Let $U(1), \dots, U(q)$ be subsets of vertices of B_N such that for each $r \in \{1, \dots, q\}$, (i) $|U(r)| \geq u$ and (ii) there exists $L \in \mathcal{L}$ such that $U(r) \subset L_N$. Then,*

$$P_\pi \left(\begin{array}{l} \text{there exists a black cluster } C_N \text{ such that } |C_N \cap L_N| \geq u \\ \text{for all } L \in \mathcal{L}, \text{ and } |C_N \cap U(r)| \geq 1, \text{ for all } r \in \{1, \dots, q\} \end{array} \right) \geq 1 - \frac{\varepsilon}{2}. \quad (4.3.4)$$

Our goal is to show that the following holds uniformly in n : With probability arbitrarily close to 1 as $N \rightarrow \infty$, there is a sequence of cubes in $D_{k(N)}^n$, each with at least one edge in common with the next, which crosses the unit cube. In order to prove this we examine the cubes $C(\mathbf{I})$, for $\mathbf{I} \in \mathcal{J}^{(n)}$, in turn according to the ordering on $\mathcal{J}^{(n)}$, and declare some of them to be good according to the rule given below. Since the probabilistic bounds on the goodness of cubes will hold uniformly in n , the desired conclusion follows.

Fix integers $n, u, k \geq 1$ until Lemma 4.3.7. For $m \geq 1$, identify a level- m cube with a vertex in $B_{N^m} \subset \mathbb{L}^d$ in the canonical way. A set of level- m cubes $\{C(\mathbf{I}_1), \dots, C(\mathbf{I}_l)\}$ is called *edge-connected* if they form a connected set with respect to the adjacency relation of \mathbb{L}^d . Whether a cube $C(\mathbf{I})$, for $\mathbf{I} \in \mathcal{J}^{(n)}$, is called (n, u) -good or not, is determined by the following inductive procedure. Let $\mathbf{I} \in \mathcal{J}^{(n)}$, and assume that the goodness of $C(\mathbf{I}')$ has been decided for all $\mathbf{I}' < \mathbf{I}$. We have the following possibilities:

- (a) $|\mathbf{I}| = n$. Then $C(\mathbf{I})$ is always declared (n, u) -good.
- (b) $0 \leq |\mathbf{I}| = m < n$.

In the latter case we act as follows. Note that the subcubes $C(\mathbf{I}, \mathbf{j})$ with $\mathbf{j} \in J^d$ have already been examined, since $(\mathbf{I}, \mathbf{j}) < \mathbf{I}$. Define the following set of level- $(m+1)$ subcubes of $C(\mathbf{I})$,

$$\mathcal{D}(\mathbf{I}) := \{C(\mathbf{I}, \mathbf{j}) : \mathbf{j} \in J^d \text{ with } C(\mathbf{I}, \mathbf{j}) \text{ } (n, u)\text{-good and } Z_k(\mathbf{I}, \mathbf{j}) = 1\}. \quad (4.3.5)$$

We declare $C(\mathbf{I})$ to be (n, u) -good if there exists an edge-connected set $\mathcal{H}(\mathbf{I}) \subset \mathcal{D}(\mathbf{I})$ such that

- (i) Each edge of $C(\mathbf{I})$ intersects at least u cubes of $\mathcal{H}(\mathbf{I})$;
- (ii) For every (n, u) -good level- m cube $C(\mathbf{I}')$ with $\mathbf{I}' < \mathbf{I}$ that has (at least) one edge in common with $C(\mathbf{I})$, there are a cube of $\mathcal{H}(\mathbf{I}')$ and a cube of $\mathcal{H}(\mathbf{I})$ with a common edge.

(If there is more than one candidate for $\mathcal{H}(\mathbf{I})$ we use some deterministic rule to choose one of them.) This procedure determines whether $C(\mathbf{I})$ is (n, u) -good for each \mathbf{I} in turn. Note that it is easier for higher level cubes to be (n, u) -good than for lower level cubes. In particular, for the unit cube, i.e. $C(\emptyset)$, it is the hardest to be (n, u) -good.

The next lemma shows that if the unit cube is (n, u) -good then there is a sequence of cubes in D_k^n , each with at least one edge in common with the next, which connects the “left-hand side” of $[0, 1]^d$ with its “right-hand side”. If such a sequence of cubes exists in D_k^n we say that *percolation occurs in D_k^n* .

Lemma 4.3.5. *Suppose $[0, 1]^d$ is (n, u) -good, then percolation occurs in D_k^n .*

Proof. Assume that the unit cube, i.e. $C(\emptyset)$, is (n, u) -good. We will show, with a recursive argument, that for $1 \leq m \leq n$ there exists an edge-connected chain of retained (n, u) -good level- m cubes which joins $\{0\} \times [0, 1]^{d-1}$ and $\{1\} \times [0, 1]^{d-1}$. In particular, this holds for $m = n$ and hence percolation occurs in D_k^n .

Since the unit cube is assumed to be (n, u) -good, $\mathcal{D}(\emptyset)$ contains by definition an edge-connected subset $\mathcal{H}(\emptyset)$ of retained (n, u) -good level-1 subcubes, such that each edge of $C(\emptyset)$ intersects at least u cubes of $\mathcal{H}(\emptyset)$. In particular, there is a sequence of retained (n, u) -good edge-connected level-1 cubes that connects the left-hand side of $[0, 1]^d$ with its right-hand side.

Let $1 \leq m < n$ and assume that there exists an edge-connected chain $C(\mathbf{I}_1), \dots, C(\mathbf{I}_l)$ of retained (n, u) -good level- m cubes which connects the left-hand side of $[0, 1]^d$ with its right-hand side. For each i , $1 \leq i \leq l$, either $\mathbf{I}_i < \mathbf{I}_{i+1}$ or $\mathbf{I}_{i+1} < \mathbf{I}_i$. By condition (ii), there exist level- $(m+1)$ cubes of $\mathcal{H}(\mathbf{I}_{i+1})$ which are edge-connected to level- $(m+1)$ cubes of $\mathcal{H}(\mathbf{I}_i)$. These level- $(m+1)$ cubes $C(\mathbf{J})$ are all (n, u) -good and have $Z_k(\mathbf{J}) = 1$, by (4.3.5) and the definition of the $\mathcal{H}(\mathbf{I})$. It follows that there is an edge-connected chain of retained (n, u) -good level- $(m+1)$ cubes $C(\mathbf{J})$ which joins $\{0\} \times [0, 1]^{d-1}$ and $\{1\} \times [0, 1]^{d-1}$. \square

For $\mathbf{I} \in \mathcal{J}^{(n)}$, define the index $\mathbf{I}^- \in \mathcal{J}^{(n)}$ by

$$\mathbf{I}^- = \max\{\mathbf{I}' : \mathbf{I}' < \mathbf{I} \text{ and } |\mathbf{I}'| \leq |\mathbf{I}|\}.$$

If there is no such index, \mathbf{I}^- is left undefined. For each $\mathbf{I} \in \mathcal{J}^{(n)}$ we let $\mathcal{F}(\mathbf{I})$ denote the σ -field

$$\mathcal{F}(\mathbf{I}) = \sigma(Z_k(\mathbf{I}', \mathbf{j}) : |\mathbf{I}'| \leq n-1, \mathbf{I}' \leq \mathbf{I}, \mathbf{j} \in J^d).$$

If \mathbf{I}^- is undefined, we take $\mathcal{F}(\mathbf{I}^-)$ to be the trivial σ -field. Note that $\mathcal{F}(\mathbf{I})$ is generated by those Z_k that have been examined prior to deciding whether $C(\mathbf{I})$ is (n, u) -good. In particular, by virtue of the ordering on the cubes as introduced in Section 4.2, $\mathcal{F}(\mathbf{I}^-)$ does *not* contain any information about subcubes of \mathbf{I} .

Let $p > p_c(d)$ and let $(k(N))_{N \geq 2}$ be a sequence such that $k(N)/N^d \geq p$, for all $N \geq 2$. We want to prove that, for every $\varepsilon > 0$, the probability that $[0, 1]^d$ is (n, u) -good in the $k(N)$ -model is at least $1 - \varepsilon$, for $N \geq N_0$, where N_0 is an integer which has to be taken sufficiently large to satisfy certain probabilistic bounds but is independent of n .

Let us first give a sketch of the proof. Fix $N \geq N_0$ and consider the $k(N)$ -model. We use a recursive argument. The smallest level- n cube according to the ordering on $\mathcal{J}^{(n)}$ is by definition (n, u) -good. Let $\mathbf{I} \in \mathcal{J}^{(n)}$ and assume that

$$\mathbb{P}_{k(N)}(C(\mathbf{I}') \text{ is } (n, u)\text{-good} \mid \mathcal{F}(\mathbf{I}'^-)) \geq 1 - \varepsilon$$

for all $\mathbf{I}' < \mathbf{I}$. We prove that, given $\mathcal{F}(\mathbf{I}^-)$, $C(\mathbf{I})$ is (n, u) -good with probability at least $1 - \varepsilon$. The proof of this consists of a coupling between a product measure with density $\pi \in (p_c(d), (1 - \varepsilon)p)$ in the box B_N and the law of the set of subcubes $C(\mathbf{I}, \mathbf{j})$ of $C(\mathbf{I})$ which are (n, u) -good and satisfy $Z_{k(N)}(\mathbf{I}, \mathbf{j}) = 1$. Applying Lemma 4.3.4 to the product measure combined with the coupling yields that the subcubes satisfy properties (i) and (ii) with probability at least $1 - \varepsilon$. Therefore, given $\mathcal{F}(\mathbf{I}^-)$, $C(\mathbf{I})$ is (n, u) -good with probability at least $1 - \varepsilon$. Iterating this argument then yields that the unit cube is (n, u) -good with probability at least $1 - \varepsilon$, for $N \geq N_0$.

The proof in [24] of the analogous result that $\sigma(p, N, d) \rightarrow 1$ as $N \rightarrow \infty$ for $p > p_c(d)$ is considerably less involved. In the context of [24], subcubes are retained with probability p independently of other cubes, which is not the case in k -fractal percolation. Therefore, they can directly show that there exists $\pi > p_c(d)$ such that, for $\mathbf{I} \in \mathcal{J}^{(n)}$, the law of the set of subcubes $C(\mathbf{I}, \mathbf{j})$ of $C(\mathbf{I})$ which are good and satisfy $Z_p(\mathbf{I}, \mathbf{j}) = 1$, dominates an i.i.d. process on the box B_N with density π .

We need the following result for binomially distributed random variables, which we state as a lemma for future reference. Since the result follows easily from Chebyshev's inequality, we omit the proof.

Lemma 4.3.6. *Let $p > p_c(d)$ and let $(k(N))_{N \geq 2}$ be a sequence of integers such that $k(N)/N^d \geq p$ for all $N \geq 2$. Let $\varepsilon > 0$ be such that $(1 - \varepsilon)p > p_c(d)$, let $\pi \in (p_c(d), (1 - \varepsilon)p)$ and define $M := ((1 - \varepsilon)p + \pi)N^d/2$. There exists N_2 such that*

$$\mathbb{P}(\{X(k(N), 1 - \varepsilon) \geq M\} \cap \{X'(N^d, \pi) \leq M\}) \geq 1 - \varepsilon/2,$$

for $N \geq N_2$, where X and X' are independent, binomially distributed random variables with the indicated parameters.

We now prove that, for any $\varepsilon > 0$, the unit cube is (n, u) -good with probability at least $1 - \varepsilon$, for N large enough but independent of n .

Lemma 4.3.7. *Let $p > p_c(d)$ and let $(k(N))_{N \geq 2}$ be a sequence of integers such that $k(N)/N^d \geq p$, for all $N \geq 2$. Let $\varepsilon > 0$ be such that $(1 - \varepsilon)p > p_c(d)$. Take $\pi \in (p_c(d), (1 - \varepsilon)p)$ and set $q = 3^d$. Let u and N_1 be given by Lemma 4.3.4. Let N_2 be given by Lemma 4.3.6. Set $N_0 = \max\{N_1, N_2\}$. Then, for all $n \geq 1$,*

$$\mathbb{P}_{k(N)}([0, 1]^d \text{ is } (n, u)\text{-good}) \geq 1 - \varepsilon, \quad (4.3.6)$$

for all $N \geq N_0$.

Proof. Fix $N \geq N_0$ and $n \geq 1$ and consider the $k(N)$ -fractal model. Our aim is to show that

$$\mathbb{P}_{k(N)}(C(\mathbf{I}) \text{ is } (n, u)\text{-good} \mid \mathcal{F}(\mathbf{I}^-)) \geq 1 - \varepsilon \quad (4.3.7)$$

holds for all $\mathbf{I} \in \mathcal{J}^{(n)}$. Taking $\mathbf{I} = \emptyset$ then yields (4.3.6). We prove this with a recursive argument. Let \mathbf{I}_0 be the smallest index in $\mathcal{J}^{(n)}$, according to the ordering on $\mathcal{J}^{(n)}$. By virtue of the ordering, we have $|\mathbf{I}_0| = n$. Hence, by definition, $C(\mathbf{I}_0)$ is (n, u) -good. In particular, (4.3.7) holds for \mathbf{I}_0 .

The recursive step is as follows. Take an index $\mathbf{I} \in \mathcal{J}^{(n)}$ and assume that

$$\mathbb{P}_{k(N)}(C(\mathbf{I}') \text{ is } (n, u)\text{-good} \mid \mathcal{F}(\mathbf{I}'^-)) \geq 1 - \varepsilon, \quad (4.3.8)$$

has been established for all indices \mathbf{I}' in $\mathcal{J}^{(n)}$ less than \mathbf{I} . We have to show that (4.3.7) holds for \mathbf{I} given this assumption. We have two cases:

(a) $|\mathbf{I}| = n$; then $\mathbb{P}_{k(N)}(C(\mathbf{I}) \text{ is } (n, u)\text{-good}) = 1$ and (4.3.7) is true.

(b) $0 \leq |\mathbf{I}| = m < n$.

For case (b), given $\mathcal{F}(\mathbf{I}^-)$, the goodness of $C(\mathbf{I}')$ is determined (in particular) for all $\mathbf{I}' < \mathbf{I}$ with $|\mathbf{I}| = m$. Let

$$\mathcal{Q} = \left\{ \mathbf{I}' : \begin{array}{l} \mathbf{I}' < \mathbf{I} \text{ and } C(\mathbf{I}') \text{ is an } (n, u)\text{-good level-}m \\ \text{cube with an edge in common with } C(\mathbf{I}) \end{array} \right\}.$$

For each $\mathbf{I}' \in \mathcal{Q}$, let $E(\mathbf{I}')$ be some common edge of $C(\mathbf{I})$ and $C(\mathbf{I}')$. Since $C(\mathbf{I}')$ is (n, u) -good, there are at least u level- $(m+1)$ subcubes in $\mathcal{H}(\mathbf{I}')$ which intersect $E(\mathbf{I}')$; call this set of subcubes $\mathcal{U}(\mathbf{I}')$. To see whether $C(\mathbf{I})$ is (n, u) -good, we look at $C(\mathbf{I}, \mathbf{j}(l))$ where $\mathbf{j}(l), 1 \leq l \leq N^d$, are the vectors of J^d arranged in order. We have $(\mathbf{I}, \mathbf{j}(l)) < \mathbf{I}$, so by the induction hypothesis (4.3.8) we have

$$\mathbb{P}_{k(N)}(C(\mathbf{I}, \mathbf{j}(l)) \text{ is } (n, u)\text{-good} \mid \mathcal{F}((\mathbf{I}, \mathbf{j}(l))^-)) \geq 1 - \varepsilon, \quad (4.3.9)$$

for all l . Note that $\mathcal{F}((\mathbf{I}, \mathbf{j}(1))^-) = \mathcal{F}(\mathbf{I}^-)$.

We identify each subcube of $C(\mathbf{I})$ in the canonical way with a vertex in B_N . We will construct three random subsets G_1, G_2, G_3 of B_N on a common probability space with the following properties:

- (I) the law of G_1 equals the law of the set of subcubes $C(\mathbf{I}, \mathbf{j})$ of $C(\mathbf{I})$ which are (n, u) -good and satisfy $Z_{k(N)}(\mathbf{I}, \mathbf{j}) = 1$;
- (II) G_2 is obtained by first selecting $k(N)$ vertices of B_N uniformly and then retaining each selected vertex with probability $1 - \varepsilon$, independently of other vertices;
- (III) the law of G_3 is the Bernoulli product measure with density π on B_N ;
- (IV) $G_1 \supset G_2$;
- (V) $\mathbb{P}(G_2 \supset G_3) \geq 1 - \varepsilon/2$.

From (4.3.9) and a standard coupling technique, sometimes referred to as sequential coupling (see e.g. [35]), the construction of G_1 and G_2 with properties (I), (II) and (IV) is straightforward. The construction of G_3 such that properties (III) and (V) hold is given below. Let $|G_2|$ denote the cardinality of the set G_2 . Define $M = ((1 - \varepsilon)p + \pi)N^d/2$ and let R be a number drawn from a binomial distribution with parameters N^d and π , independently of G_1 and G_2 . If $|G_2| \geq M$ and $M \geq R$ we select R vertices uniformly out of the $|G_2|$ retained vertices of G_2 and call this set G_3 . Otherwise, we select, independently of G_1 and G_2 , R vertices of B_N in a uniform way and call this set G_3 . From the construction (note that also G_2 was obtained in a uniform way) it is clear that G_3 satisfies property (III). Observe that $|G_2|$ has a binomial distribution with parameters $k(N)$ and $1 - \varepsilon$. From Lemma 4.3.6 it follows that

$$\mathbb{P}(\{|G_2| \geq M\} \cap \{R \leq M\}) \geq 1 - \varepsilon/2.$$

Hence, property (V) also holds.

Let us now return to the goodness of $C(\mathbf{I})$. As before, we identify the random subsets G_1, G_2, G_3 of B_N with the corresponding sets of subcubes of $C(\mathbf{I})$ in the canonical way. It then follows from property (III) and Lemma 4.3.4 (note that \mathcal{Q} has cardinality at most $3^d = q$) that G_3 has an edge-connected subset which satisfies the following properties with probability at least $1 - \varepsilon/2$:

- (i) intersects every edge of $C(\mathbf{I})$ with at least u cubes;
- (ii) contains a cube that is edge-connected to a cube of $\mathcal{U}(\mathbf{I}')$, for all $\mathbf{I}' \in \mathcal{Q}$.

Combining properties (IV), (V) and the previous paragraph we obtain

$$\begin{aligned} & \mathbb{P}_{k(N)}(C(\mathbf{I}) \text{ is } (n, u)\text{-good} \mid \mathcal{F}(\mathbf{I}^-)) \\ & \geq \mathbb{P}(\{G_1 \supset G_3\} \cap \{G_3 \text{ satisfies properties (i) and (ii)}\}) \\ & \geq 1 - \varepsilon. \end{aligned}$$

Therefore, (4.3.7) holds for the index \mathbf{I} given that (4.3.8) holds for all indices $\mathbf{I}' < \mathbf{I}$. A recursive use of this argument – recall that (4.3.7) is valid for \mathbf{I}_0 (the smallest index according to the ordering) – yields that (4.3.7) holds for all \mathbf{I} . Taking $\mathbf{I} = \emptyset$ in (4.3.7) proves the lemma. \square

We are now able to conclude the proof of Theorem 4.3.2.

Proof of Theorem 4.3.2. Let $p > p_c(d)$ and consider a sequence $(k(N))_{N \geq 2}$ such that $k(N)/N^d \geq p$, for all $N \geq 2$. We get, using both Lemma 4.3.7 and Lemma 4.3.5, that for any $\varepsilon > 0$ such that $(1 - \varepsilon)p > p_c(d)$, there exists N_0 , depending on ε , such that

$$\mathbb{P}_{k(N)}(\text{percolation in } D_{k(N)}^n) \geq \mathbb{P}_{k(N)}([0, 1]^d \text{ is } (n, u)\text{-good}) \geq 1 - \varepsilon, \quad (4.3.10)$$

for $N \geq N_0$. It is well known (see e.g. [24]) that

$$\{[0, 1]^d \text{ is crossed by } D_{k(N)}\} = \bigcap_{n=1}^{\infty} \{\text{percolation in } D_{k(N)}^n\}.$$

Hence, taking the limit $n \rightarrow \infty$ in (4.3.10) yields that for $\varepsilon > 0$ small enough

$$\mathbb{P}_{k(N)}([0, 1]^d \text{ is crossed by } D_{k(N)}) \geq 1 - \varepsilon, \quad (4.3.11)$$

for $N \geq N_0$. Therefore,

$$\theta(k(N), N, d) \rightarrow 1,$$

as $N \rightarrow \infty$. \square

4.3.3 Proof of Theorem 4.1.4

Proof of Theorem 4.1.4. We use the idea of the proof of Theorem 4.3.1 and the result of Theorem 4.3.2. Fix some p_0 such that $p_c(d) < p_0 < p$ and set $k(N) := \lfloor p_0 N^d \rfloor$. Consider the event F that in the GFP model with generator $Y = Y(N, d)$ there exists an infinite tree of retained subcubes such that each subcube in the tree contains at least $k(N)$ retained subcubes in the tree. Similar to the proof of Lemma 4.3.3, we prove that $\mathbb{P}(F) \rightarrow 1$ as $N \rightarrow \infty$. We then show that the law of $D_{k(N)}$ is stochastically dominated by the conditional law of D_Y , conditioned on the event F . By Theorem 4.3.2 we can then conclude that $\phi(Y(N, d), N, d) \rightarrow 1$ as $N \rightarrow \infty$.

Consider the construction of D_Y . We will use the same definition of m -good as in Section 4.3.1, that is, if a level- n cube is retained and contains at least $k(N)$ retained subcubes, we call this level- n cube 0-good. Recursively, we say that a level- n cube is

$(m+1)$ -good if it is retained and contains at least $k(N)$ m -good level- $(n+1)$ subcubes. We call the unit cube ∞ -good if it is m -good for every $m \geq 0$. Define the following events

$$\begin{aligned} F_m &:= \{[0, 1]^d \text{ is } m\text{-good}\}, \\ F &:= \{[0, 1]^d \text{ is } \infty\text{-good}\}. \end{aligned}$$

We will show that for every $\varepsilon > 0$ such that $(1 - \varepsilon)p > p_0$ there exists $N_0 = N_0(\varepsilon)$ such that, for all $m \geq 0$,

$$\mathbb{P}(F_m) > 1 - \varepsilon, \quad \text{for all } N \geq N_0. \quad (4.3.12)$$

The proof of (4.3.12) is similar to the proof of Lemma 4.3.3. Let $\varepsilon > 0$ be such that $(1 - \varepsilon)p > p_0$. Take $\delta > 0$ such that $(1 - \varepsilon)p > p_0 + \delta$. Then, take N_0 so large that

$$1 - \frac{1}{4\delta^2 N} > 1 - \varepsilon/2 \quad \text{and} \quad (4.3.13)$$

$$\mathbb{P}(Y \geq pN^d) > 1 - \varepsilon/2, \quad (4.3.14)$$

for all $N \geq N_0$. We prove that (4.3.12) holds for this N_0 and all $m \geq 0$, by induction on m . Since $k(N) = \lfloor p_0 N^d \rfloor \leq pN^d$ it follows from (4.3.14) that $\mathbb{P}(F_0) > 1 - \varepsilon$, for all $N \geq N_0$.

Next, assume that (4.3.12) holds for some $m \geq 0$. The probability that a level-1 cube is m -good, given that it is retained, is equal to $\mathbb{P}(F_m)$. It follows that, given that the number of retained level-1 cubes equals y , the number of m -good level-1 cubes has a binomial distribution with parameters y and $\mathbb{P}(F_m)$. By our choices for N_0 and δ we get

$$\begin{aligned} \mathbb{P}(F_{m+1}) &= \sum_{y \geq k(N)} \mathbb{P}(X(y, \mathbb{P}(F_m)) \geq k(N)) \mathbb{P}(Y = y) \\ &\geq \mathbb{P}(X(\lfloor pN^d \rfloor, \mathbb{P}(F_m)) \geq p_0 N^d) \mathbb{P}(Y \geq \lfloor pN^d \rfloor) \\ &\geq \mathbb{P}(X(\lfloor pN^d \rfloor, 1 - \varepsilon) \geq p_0 N^d) (1 - \varepsilon/2) \\ &\geq \left(1 - \frac{\text{Var}X(\lfloor pN^d \rfloor, 1 - \varepsilon)}{(p_0 - (1 - \varepsilon)p)^2 N^{2d}}\right) (1 - \varepsilon/2) \\ &\geq \left(1 - \frac{(1 - \varepsilon)\varepsilon p N^d}{\delta^2 N^{2d}}\right) (1 - \varepsilon/2) \\ &\geq \left(1 - \frac{1}{4\delta^2 N^d}\right) (1 - \varepsilon/2) \\ &\geq (1 - \varepsilon/2)(1 - \varepsilon/2) > 1 - \varepsilon, \end{aligned}$$

for all $N \geq N_0$. Hence, the induction step is valid.

Analogously to the proof of Theorem 4.3.1 we use the event $F = \bigcap_{m=1}^{\infty} F_m$ to construct two random subsets $\tilde{D}_{k(N)}$ and \tilde{D}_Y on a common probability space, with the following properties:

- (i) $\tilde{D}_{k(N)} \subset \tilde{D}_Y$;

- (ii) the law of \tilde{D}_Y is stochastically dominated by the conditional law of D_Y , conditioned on the event F ;
- (iii) the law of $\tilde{D}_{k(N)}$ is equal to the law of $D_{k(N)}$.

This construction is the same (modulo replacing the binomial distribution with Y) as in the proof of Theorem 4.3.1 and is therefore omitted.

From properties (i)-(iii) and Theorem 4.3.2 we get

$$\begin{aligned}
 & \mathbb{P}([0, 1]^d \text{ is crossed by } D_{Y(N,d)} | F) \\
 & \geq \mathbb{P}([0, 1]^d \text{ is crossed by } \tilde{D}_{Y(N,d)}) \\
 & \geq \mathbb{P}([0, 1]^d \text{ is crossed by } \tilde{D}_{k(N)}) \\
 & = \mathbb{P}([0, 1]^d \text{ is crossed by } D_{k(N)}) \rightarrow 1,
 \end{aligned}$$

as $N \rightarrow \infty$. Since (4.3.12) implies that $\mathbb{P}(F) \rightarrow 1$ as $N \rightarrow \infty$, we obtain

$$\mathbb{P}([0, 1]^d \text{ is crossed by } D_{Y(N,d)}) \rightarrow 1,$$

as $N \rightarrow \infty$. □

4.4 Proofs of the fat fractal results

In this section we prove our results concerning fat fractal percolation. First, we state an elementary property of the fat fractal percolation model; it follows immediately from Fubini's theorem and we omit the proof.

Proposition 4.4.1. *The expected Lebesgue measure of the limit set of fat fractal percolation is given by*

$$\mathbb{E} \lambda(D_{\text{fat}}) = \prod_{n=1}^{\infty} p_n.$$

4.4.1 Proof of Theorem 4.1.7

Since $\prod_{n=1}^{\infty} p_n > 0$ it follows from Proposition 4.4.1 that with positive probability the limit set has positive Lebesgue measure given $D_{\text{fat}} \neq \emptyset$. Theorem 4.1.7 states that the latter holds with probability 1.

Proof of Theorem 4.1.7. Let Z_n denote the number of retained level- n cubes after iteration step n and set $Z_0 := 1$. Since the retention probabilities p_n vary with n , the process $(Z_n)_{n \geq 1}$ is a so-called branching process in a time-varying environment. Following the notation of Lyons in [37] let L_n be a random variable, having the distribution of Z_n given that $Z_{n-1} = 1$. Note that L_n has a binomial distribution with parameters N^d and p_n .

Define the process $(W_n)_{n \geq 1}$ by

$$W_n := \frac{Z_n}{\prod_{i=1}^n p_i N^d}.$$

It is straightforward to show that $(W_n)_{n \geq 1}$ is a martingale:

$$\begin{aligned} \mathbb{E}[W_n | W_{n-1}] &= \frac{\mathbb{E}[Z_n | Z_{n-1}]}{\prod_{i=1}^n p_i N^d} = \frac{Z_{n-1}}{\prod_{i=1}^n p_i N^d} \mathbb{E}[Z_n | Z_{n-1} = 1] \\ &= \frac{Z_{n-1} p_n N^d}{\prod_{i=1}^n p_i N^d} = W_{n-1}. \end{aligned}$$

The Martingale Convergence Theorem tells us that W_n converges almost surely to a random variable W . Theorem 4.14 of [37] states that if

$$A := \sup_n \|L_n\|_\infty < \infty,$$

then $W > 0$ a.s. given non-extinction. It is clearly the case that $A < \infty$, because L_n can take at most the value N^d . Therefore, W_n converges to a random variable W which is strictly positive a.s. given non-extinction.

The Lebesgue measure of the retained cubes at each iteration step n is equal to Z_n/N^{dn} . We have

$$\lambda(D_{\text{fat}}^n) = \frac{Z_n}{N^{dn}} = \frac{(\prod_{i=1}^n p_i N^d) W_n}{N^{dn}} = \left(\prod_{i=1}^n p_i \right) W_n. \quad (4.4.1)$$

Letting $n \rightarrow \infty$ in (4.4.1) yields $\lambda(D_{\text{fat}}) = (\prod_{i=1}^\infty p_i) W$. Since $\prod_{i=1}^\infty p_i > 0$ and $W > 0$ a.s. given non-extinction, we get the desired result. \square

4.4.2 Proof of Theorem 4.1.8

We start with a heuristic strategy for the proof. For a fixed configuration $\omega \in \Omega$, let us call a point x in the unit cube *conditionally connected* if the following property holds: If we change ω by retaining all cubes that contain x , then x is contained in a connected component larger than one point. We show that for almost all points x it is the case that x is conditionally connected with probability 0 or 1. We define an ergodic transformation T on the unit cube. The transformation T enables us to prove that the probability for a point x to be conditionally connected has the same value for λ -almost all x . From this we can then conclude that either the set of dust points or the set of connected components contains all Lebesgue measure.

Proof of Theorem 4.1.8. First, we have to introduce some notation. Let U be the collection of points in $[0, 1]^d$ not on the boundary of a subcube. For each $x \in U$ there exists a unique sequence $(C(\mathbf{x}_1, \dots, \mathbf{x}_n))_{n \geq 1}$ of cubes of the fractal process, where $\mathbf{x}_j \in J^d$ for all j , such that $\bigcap_{n \geq 1} C(\mathbf{x}_1, \dots, \mathbf{x}_n) = \{x\}$. Therefore, we can define an invertible transformation $\phi : U \rightarrow (J^d)^\mathbb{N}$ by $\phi(x) = (\mathbf{x}_1, \mathbf{x}_2, \dots)$. For each $n \in \mathbb{N}$ let μ_n be the uniform measure on (X_n, \mathcal{F}_n) , where $X_n = J^d$ and \mathcal{F}_n is the power set of X_n . Let $(X, \mathcal{F}, \mu) = \bigotimes_{n=1}^\infty (X_n, \mathcal{F}_n, \mu_n)$ be the product space. Since $\phi : (U, \mathcal{B}(U), \lambda) \rightarrow (X, \mathcal{F}, \mu)$ is an invertible measure-preserving transformation, we have that (X, \mathcal{F}, μ) is by definition isomorphic to $(U, \mathcal{B}(U), \lambda)$. Here $\mathcal{B}(U)$ denotes the Borel σ -algebra.

Next, we define the transformation $T : U \rightarrow U$, which will play a crucial role in the rest of the proof. Define the auxiliary shift transformation $T^* : X \rightarrow X$ by

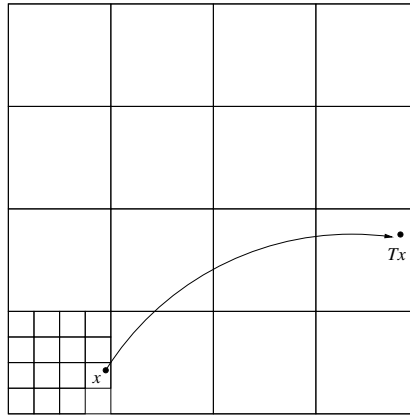


Figure 4.2: Illustration of the transformation T . Note that the relative position of x in the level-1 cube is the same as the relative position of Tx in the unit cube.

$T^*((\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots)) = (\mathbf{x}_2, \mathbf{x}_3, \dots)$, for $(\mathbf{x}_1, \mathbf{x}_2, \dots) \in X$. The transformation T^* is measure preserving with respect to the measure μ and also ergodic, see for instance [54]. Let $T := \phi^{-1} \circ T^* \circ \phi$ be the induced transformation on U and note that T is isomorphic to T^* and hence also ergodic. Informally, T sends a point $x \in U$ to the point Tx , in such a way that the relative position of Tx in the unit cube is the same as the relative position of x in its level-1 cube $C(\mathbf{x}_1)$; see Figure 4.2.

Recall that $\omega \in \Omega$ denotes a particular realization of the fat fractal percolation process. For $x \in U$, we define the following event.

$$A^x := \{\omega : \text{if we set } \omega(C(\mathbf{x}_1, \dots, \mathbf{x}_n)) = 1 \text{ for all } n \geq 1, \text{ then } C_{\text{fat}}^x \neq \{x\}\}.$$

In other words, A^x consists of those configurations ω such that when we change the configuration by retaining all $C(\mathbf{x}_1, \dots, \mathbf{x}_n)$, then in this new configuration, x is in the same connected component as some $y \neq x$. Observe that

$$A^x \cap \{x \in D_{\text{fat}}\} = \{x \in D_{\text{fat}}^c\}. \quad (4.4.2)$$

It is easy to see that A^x is a tail event. Hence, by Kolmogorov's 0-1 law we get $\mathbb{P}(A^x) \in \{0, 1\}$ for all $x \in U$.

However, a priori it is not clear that for almost all x in the unit cube $\mathbb{P}(A^x)$ has the same value. To this end, define the set $V := \{x \in U : \mathbb{P}(A^x) = 0\}$. We will show that $\lambda(V) \in \{0, 1\}$. Recall that the relative position of Tx in the unit cube is the same as the relative position of x in the level-1 cube $C(\mathbf{x}_1)$. It is possible to construct a coupling between the fractal process in the unit cube and the fractal process in $C(\mathbf{x}_1)$, given that $C(\mathbf{x}_1)$ is retained, with the following property: For every cube $C(\mathbf{I})$ in $C(\mathbf{x}_1)$, it is the case that if $TC(\mathbf{I})$ is retained in the fractal process in the unit cube, then $C(\mathbf{I})$ is also retained in the fractal process in $C(\mathbf{x}_1)$, given that $C(\mathbf{x}_1)$ is retained. It is straightforward that such a coupling exists since the retention probabilities p_n are non-decreasing in n . Hence,

$$\mathbb{P}(A^{Tx}) \leq \mathbb{P}(A^x | C(\mathbf{x}_1) \text{ is retained}). \quad (4.4.3)$$

Furthermore, since A^x is a tail event, we have

$$\mathbb{P}(A^x) = \mathbb{P}(A^x | C(\mathbf{x}_1) \text{ is retained}). \quad (4.4.4)$$

It follows from (4.4.3) and (4.4.4) that $\mathbb{P}(A^{T^x}) \leq \mathbb{P}(A^x)$ for all x . This implies that $V \subset T^{-1}V$. Because T is measure preserving it follows that

$$\lambda(V \Delta T^{-1}V) = \lambda(V \setminus T^{-1}V) + \lambda(T^{-1}V \setminus V) = 0 + \lambda(T^{-1}V) - \lambda(V) = 0.$$

Ergodicity of T now yields that $\lambda(V) \in \{0, 1\}$.

Suppose $\lambda(V) = 0$. Then $\mathbb{P}(x \in D_{\text{fat}}^d) = \mathbb{P}(\{x \in D_{\text{fat}}\} \setminus A^x) = 0$ for almost all $x \in [0, 1]^d$, by (4.4.2). Applying Fubini's theorem gives

$$\begin{aligned} \mathbb{E} \lambda(D_{\text{fat}}^d) &= \int_{\Omega} \int_{[0,1]^d} \mathbb{1}_{D_{\text{fat}}^d}(x, \omega) d\lambda d\mathbb{P} \\ &= \int_{[0,1]^d} \int_{\Omega} \mathbb{1}_{D_{\text{fat}}^d}(x, \omega) d\mathbb{P} d\lambda \\ &= \int_{[0,1]^d} \mathbb{P}(x \in D_{\text{fat}}^d) d\lambda = 0. \end{aligned}$$

Therefore $\lambda(D_{\text{fat}}^d) = 0$ a.s. By Theorem 4.1.7 we have $\lambda(D_{\text{fat}}^c) > 0$ a.s. given non-extinction.

Next suppose that $\lambda(V) = 1$. Then with a similar argument we can show that $\lambda(D_{\text{fat}}^c) = 0$ and $\lambda(D_{\text{fat}}^d) > 0$ a.s. given non-extinction. \square

4.4.3 Proof of Theorem 4.1.9

Proof of Theorem 4.1.9. (i) Suppose that D_{fat} has a non-empty interior with positive probability. Then we have

$$\begin{aligned} 0 &< \mathbb{P}(D_{\text{fat}} \text{ has non-empty interior}) \\ &= \mathbb{P}(\exists n, \exists \mathbf{i}_1, \dots, \mathbf{i}_n : C(\mathbf{i}_1, \dots, \mathbf{i}_n) \subset D_{\text{fat}}) \\ &\leq \sum_{n, \mathbf{i}_1, \dots, \mathbf{i}_n} \mathbb{P}(C(\mathbf{i}_1, \dots, \mathbf{i}_n) \subset D_{\text{fat}}). \end{aligned}$$

Since we sum over countably many cubes, there must exist n and $\mathbf{i}_1, \dots, \mathbf{i}_n$ such that $\mathbb{P}(C(\mathbf{i}_1, \dots, \mathbf{i}_n) \subset D_{\text{fat}}) > 0$. Hence, by translation invariance, $\mathbb{P}(C(\mathbf{i}_1, \dots, \mathbf{i}_n) \subset D_{\text{fat}}) > 0$ for this specific n and all $\mathbf{i}_1, \dots, \mathbf{i}_n$. We can apply the FKG inequality to obtain $\mathbb{P}(D_{\text{fat}} = [0, 1]^d) = \mathbb{P}(C(\mathbf{i}_1, \dots, \mathbf{i}_n) \subset D_{\text{fat}} \forall \mathbf{i}_1, \dots, \mathbf{i}_n) > 0$. Since $\mathbb{P}(D_{\text{fat}} = [0, 1]^d) = \prod_{n=1}^{\infty} p_n^{N^{dn}}$, this proves the first part of the theorem.

(ii) Suppose $\prod_{n=1}^{\infty} p_n^{N^{dn}} > 0$. Then for each $x \in [0, 1]^{d-1}$ we have $\mathbb{P}(\{x\} \times [0, 1] \subset D_{\text{fat}}) \geq \prod_{n=1}^{\infty} p_n^{N^{dn}} > 0$. Let λ_{d-1} denote $(d-1)$ -dimensional Lebesgue measure.

Applying Fubini's theorem gives

$$\begin{aligned}
& \mathbb{E} \lambda_{d-1}(\{x \in [0, 1]^{d-1} : \{x\} \times [0, 1] \subset D_{\text{fat}}\}) \\
&= \int_{\Omega} \int_{[0, 1]^{d-1}} \mathbb{1}_{\{x\} \times [0, 1] \subset D_{\text{fat}}} d\lambda_{d-1} d\mathbb{P} \\
&= \int_{[0, 1]^{d-1}} \int_{\Omega} \mathbb{1}_{\{x\} \times [0, 1] \subset D_{\text{fat}}} d\mathbb{P} d\lambda_{d-1} \\
&= \int_{[0, 1]^{d-1}} \mathbb{P}(\{x\} \times [0, 1] \subset D_{\text{fat}}) d\lambda_{d-1} > 0.
\end{aligned}$$

Hence,

$$\lambda_{d-1}(\{x \in [0, 1]^{d-1} : \{x\} \times [0, 1] \subset D_{\text{fat}}\}) > 0 \quad (4.4.5)$$

with positive probability. Observe that

$$D_{\text{fat}}^c \supset \bigcup_{x \in [0, 1]^{d-1} : \{x\} \times [0, 1] \subset D_{\text{fat}}} \{x\} \times [0, 1].$$

In particular,

$$\lambda(D_{\text{fat}}^c) \geq \lambda_{d-1}(\{x \in [0, 1]^{d-1} : \{x\} \times [0, 1] \subset D_{\text{fat}}\}).$$

From (4.4.5) we conclude that $\lambda(D_{\text{fat}}^c) > 0$ with positive probability. It now follows from Theorem 4.1.8 that the Lebesgue measure of the dust set is 0 a.s.

(iii) Next assume that $\prod_{n=1}^{\infty} p_n^{N^{dn}} > 0$. For each level n , we have $\mathbb{P}(D_{\text{fat}}^n = D_{\text{fat}}^{n-1}) \geq p_n^{N^{dn}}$. Since $\prod_{n=1}^{\infty} p_n^{N^{dn}} > 0$ is equivalent to $\sum_{n=1}^{\infty} (1 - p_n^{N^{dn}}) < \infty$, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(D_{\text{fat}}^n \neq D_{\text{fat}}^{n-1}) \leq \sum_{n=1}^{\infty} (1 - p_n^{N^{dn}}) < \infty.$$

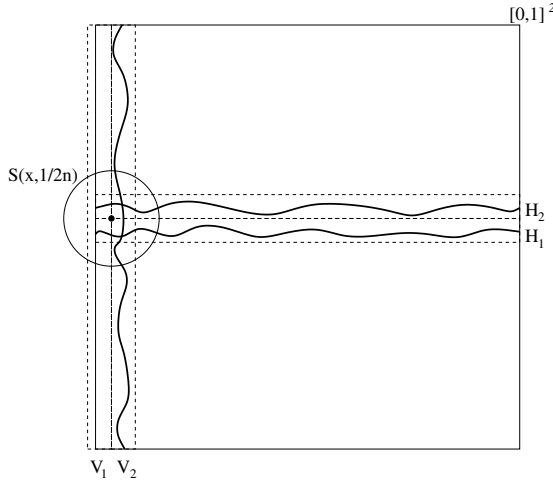
Applying the Borel-Cantelli lemma gives that, with probability 1, $\{D_{\text{fat}}^n \neq D_{\text{fat}}^{n-1}\}$ occurs for only finitely many n . Hence, with probability 1 there exists an n such that D_{fat} can be written as the union of level- n cubes. \square

4.4.4 Proof of Theorem 4.1.11

Proof of Theorem 4.1.11. (iii) \Rightarrow (ii). Trivial.

(ii) \Rightarrow (i). Suppose $\mathbb{P}(x \text{ connected to } y) > 0$ for all $x, y \in U$, for some set $U \subset [0, 1]^2$ with $\lambda(U) > 0$. Fix $y \in U$. By Fubini's theorem

$$\begin{aligned}
\mathbb{E} \lambda(D_{\text{fat}}^c) &= \int_{\Omega} \int_{[0, 1]^2} \mathbb{1}_{D_{\text{fat}}^c}(x, \omega) d\lambda(x) d\mathbb{P}(\omega) \\
&= \int_{[0, 1]^2} \int_{\Omega} \mathbb{1}_{D_{\text{fat}}^c}(x, \omega) d\mathbb{P}(\omega) d\lambda(x) \\
&= \int_{[0, 1]^2} \mathbb{P}(x \in D_{\text{fat}}^c) d\lambda(x) \\
&\geq \int_{U \setminus \{y\}} \mathbb{P}(x \text{ connected to } y) d\lambda(x) > 0.
\end{aligned}$$

Figure 4.3: Realization of the event Γ_x .

Hence $\lambda(D_{\text{fat}}^c) > 0$ with positive probability. By Theorem 4.1.8 it follows that $\lambda(D_{\text{fat}}^c) > 0$ a.s. given non-extinction of the fat fractal process.

(i) \Rightarrow (iii). Next suppose that $\lambda(D_{\text{fat}}^c) > 0$ a.s. given non-extinction of the fat fractal process. For points $x \in [0, 1]^2$ not on the boundary of a subcube, define the event A^x as in the proof of Theorem 4.1.8. It follows from the proof of Theorem 4.1.8 that $\mathbb{P}(A^x) = 1$ for all $x \in V$, for some set $V \subset [0, 1]^2$ with $\lambda(V) = 1$. By (4.4.2) we have for all $x \in V$

$$\mathbb{P}(x \in D_{\text{fat}}^c) = \mathbb{P}(x \in D_{\text{fat}}) > 0.$$

Let $x \in V$. Then

$$0 < \mathbb{P}(x \in D_{\text{fat}}^c) \leq \sum_{n=1}^{\infty} \mathbb{P}(\text{diam}(C_{\text{fat}}^x) > \frac{1}{n}),$$

where $\text{diam}(C_{\text{fat}}^x)$ denotes the diameter of the set C_{fat}^x . So there exists a natural number n_x such that $\mathbb{P}(\text{diam}(C_{\text{fat}}^x) > \frac{1}{n_x}) > 0$. Hence

$$\mathbb{P}(x \text{ connected to } S(x, \frac{1}{2n_x})) > 0,$$

where $S(x, \frac{1}{2n_x})$ is a circle centered at x with radius $\frac{1}{2n_x}$. Write $x = (x_1, x_2)$ and define the following subsets of \mathbb{R}^2

$$\begin{aligned} H_1 &= [0, 1] \times [x_2 - \frac{1}{4n_x}, x_2], \\ H_2 &= [0, 1] \times [x_2, x_2 + \frac{1}{4n_x}], \\ V_1 &= [x_1 - \frac{1}{4n_x}, x_1] \times [0, 1], \\ V_2 &= [x_1, x_1 + \frac{1}{4n_x}] \times [0, 1]. \end{aligned}$$

Note that for every $x \in [0, 1]^2$ it is the case that at least one horizontal strip H_i and at least one vertical strip V_j is entirely contained in $[0, 1]^2$. Define the event Γ_x by

$$\begin{aligned} \Gamma_x &= \bigcap_{i \in \{1, 2\}: H_i \subset [0, 1]^2} \{\text{horizontal crossing in } H_i\} \\ &\cap \bigcap_{j \in \{1, 2\}: V_j \subset [0, 1]^2} \{\text{vertical crossing in } V_j\}. \end{aligned}$$

See Figure 4.3 for an illustration of the event Γ_x . From Theorem 2 in [16] it follows that in the MFP model with parameter $p \geq p_c(N, 2)$, the limit set D_p connects the left-hand side of $[0, 1]^2$ with its right-hand side with positive probability. It then follows from the RSW lemma (e.g. Lemma 5.1 in [20]) and the FKG inequality that $\mathbb{P}_p(\Gamma_x) > 0$. Let A_n denote the event of complete retention until level n , i.e. $\omega(C(\mathbf{I})) = 1$ for all $\mathbf{I} \in \mathcal{J}^{(n-1)}$. Since $\prod_{n=1}^{\infty} p_n > 0$ there exists an integer n_0 such that $p_n \geq p_c(N, 2)$ for all $n \geq n_0$. Hence, the probability measure $\mathbb{P}_{\text{fat}}(\cdot | A_{n_0})$ dominates $\mathbb{P}_{p_c(N, 2)}(\cdot)$. Since $\mathbb{P}_{\text{fat}}(A_{n_0}) > 0$ it follows that $\mathbb{P}_{\text{fat}}(\Gamma_x) > 0$.

Observe that for $x, y \in V$

$$\begin{aligned} &\{x \text{ connected to } y\} \\ &\supset \{x \text{ connected to } S(x, \frac{1}{2n_x})\} \cap \Gamma_x \cap \{y \text{ connected to } S(y, \frac{1}{2n_y})\} \cap \Gamma_y. \end{aligned}$$

Since all four events on the right-hand side are increasing and have positive probability, we can apply the FKG inequality to conclude that for all $x, y \in V$ we have $\mathbb{P}(x \text{ connected to } y) > 0$. \square

Chapter 5

Random walk loop soups and conformal loop ensembles

This chapter is based on the paper [52] by Van de Brug, Camia, and Lis.

5.1 Introduction

Several interesting models of statistical mechanics, such as percolation and the Ising and Potts models, can be described in terms of clusters. In two dimensions and at the critical point, the scaling limit geometry of the boundaries of such clusters is known (see [13–15, 19, 49]) or conjectured (see [28, 50]) to be described by some member of the one-parameter family of Schramm-Loewner evolutions (SLE_κ with $\kappa > 0$) and related conformal loop ensembles (CLE_κ with $8/3 < \kappa < 8$). What makes SLEs and CLEs natural candidates is their conformal invariance, a property expected of the scaling limit of two-dimensional statistical mechanical models at the critical point. SLEs can be used to describe the scaling limit of single interfaces; CLEs are collections of loops and are therefore suitable to describe the scaling limit of the collection of all macroscopic boundaries at once. For example, the scaling limit of the critical percolation exploration path is SLE_6 [14, 49], and the scaling limit of the collection of all critical percolation interfaces in a bounded domain is CLE_6 [13, 15].

For $8/3 < \kappa \leq 4$, CLE_κ can be obtained [48] from the Brownian loop soup, introduced by Lawler and Werner [32] (see Section 5.2 for a definition), as we explain below. A sample of the Brownian loop soup in a bounded domain D with intensity $\lambda > 0$ is the collection of loops contained in D from a Poisson realization of a conformally invariant intensity measure $\lambda\mu$. When $\lambda \leq 1/2$, the loop soup is composed of disjoint clusters of loops [48] (where a cluster is a maximal collection of loops that intersect each other). When $\lambda > 1/2$, there is a unique cluster [48] and the set of points not surrounded by a loop is totally disconnected (see [6]). Furthermore, when $\lambda \leq 1/2$, the outer boundaries of the outermost loop soup clusters are distributed like conformal loop ensembles (CLE_κ) [47, 48, 55] with $8/3 < \kappa \leq 4$. More precisely, if $8/3 < \kappa \leq 4$, then $0 < (3\kappa - 8)(6 - \kappa)/4\kappa \leq 1/2$ and the collection of all outer boundaries of the outermost clusters of the Brownian loop soup with intensity $\lambda = (3\kappa - 8)(6 - \kappa)/4\kappa$ is

distributed like CLE_κ [48]. For example, the continuum scaling limit of the collection of all macroscopic outer boundaries of critical Ising spin clusters is conjectured to correspond to CLE_3 and to a Brownian loop soup with $\lambda = 1/4$. Note that the relation between λ and κ mentioned above is not exactly as stated in [48]; it has become apparent recently (see e.g. the discussion in the introduction of [36]) that there is a mistake in [48, 55].

In [31] Lawler and Trujillo Ferreras introduced the random walk loop soup as a discrete version of the Brownian loop soup, and showed that, under Brownian scaling, it converges in an appropriate sense to the Brownian loop soup. The authors of [31] focused on individual loops, showing that, with probability going to 1 in the scaling limit, there is a one-to-one correspondence between “large” lattice loops from the random walk loop soup and “large” loops from the Brownian loop soup such that corresponding loops are close.

In [33] Le Jan showed that the random walk loop soup has remarkable connections with the discrete Gaussian free field, analogous to Dynkin’s isomorphism [22, 23] (see also [9]). Such considerations have prompted an extensive analysis of more general versions of the random walk loop soup (see e.g. [34, 51]).

As explained above, the connection between the Brownian loop soup and SLE/CLE goes through its loop clusters and their boundaries. In view of this observation, it is interesting to investigate whether the random walk loop soup converges to the Brownian loop soup in terms of loop clusters and their boundaries, not just in terms of individual loops, as established by Lawler and Trujillo Ferreras [31]. This is a natural and nontrivial question, due to the complex geometry of the loops involved and of their mutual overlaps.

In this chapter, we consider random walk loop soups from which the “vanishingly small” loops have been removed and establish convergence of their clusters and boundaries, in the scaling limit, to the clusters and boundaries of the corresponding Brownian loop soups (see Figure 5.1). We work in the same set-up as [31], which in particular means that the number of loops of the random walk loop soup after cut-off diverges in the scaling limit. We use tools ranging from classical Brownian motion techniques to recent loop soup results. Indeed, properties of planar Brownian motion as well as properties of CLEs play an important role in the proofs of our results.

5.2 Definitions and main result

We recall the definitions of the Brownian loop soup and the random walk loop soup. A *curve* γ is a continuous function $\gamma : [0, t_\gamma] \rightarrow \mathbb{C}$, where $t_\gamma < \infty$ is the time length of γ . A *loop* is a curve with $\gamma(0) = \gamma(t_\gamma)$. A planar Brownian loop of time length t_0 started at z is the process $z + B_t - (t/t_0)B_{t_0}$, $0 \leq t \leq t_0$, where B is a planar Brownian motion started at 0. The Brownian bridge measure μ_{z, t_0}^\sharp is a probability measure on loops, induced by a planar Brownian loop of time length t_0 started at z . The (rooted) Brownian loop measure μ is a measure on loops, given by

$$\mu(C) = \int_{\mathbb{C}} \int_0^\infty \frac{1}{2\pi t_0^2} \mu_{z, t_0}^\sharp(C) dt_0 dA(z),$$

where C is a collection of loops and A denotes two-dimensional Lebesgue measure, see Remark 5.28 of [29]. For a domain D let μ_D be μ restricted to loops which stay

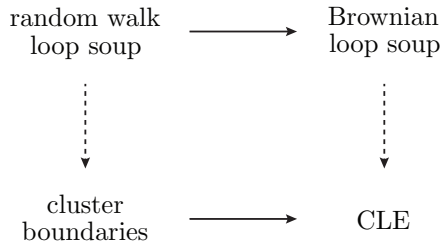


Figure 5.1: Schematic diagram of relations between discrete and continuous loop soups and their cluster boundaries. Horizontal arrows indicate a scaling limit. In this chapter we show the convergence corresponding to the bottom horizontal arrow.

in D .

The (rooted) *Brownian loop soup* with intensity $\lambda \in (0, \infty)$ in D is a Poissonian realization from the measure $\lambda\mu_D$. The Brownian loop soup introduced by Lawler and Werner [32] is obtained by forgetting the starting points (roots) of the loops. The geometric properties we study in this chapter are the same for both the rooted and the unrooted version of the Brownian loop soup. Let \mathcal{L} be a Brownian loop soup with intensity λ in a domain D , and let \mathcal{L}^{t_0} be the collection of loops in \mathcal{L} with time length at least t_0 .

The (rooted) random walk loop measure $\tilde{\mu}$ is a measure on nearest neighbor loops in \mathbb{Z}^2 , which we identify with loops in the complex plane by linear interpolation. For a loop $\tilde{\gamma}$ in \mathbb{Z}^2 , we define

$$\tilde{\mu}(\tilde{\gamma}) = \frac{1}{t_{\tilde{\gamma}}} 4^{-t_{\tilde{\gamma}}},$$

where $t_{\tilde{\gamma}}$ is the time length of $\tilde{\gamma}$, i.e. its number of steps. The (rooted) *random walk loop soup* with intensity λ is a Poissonian realization from the measure $\lambda\tilde{\mu}$. For a domain D and positive integer N , let $\tilde{\mathcal{L}}_N$ be the collection of loops $\tilde{\gamma}_N$ defined by $\tilde{\gamma}_N(t) = N^{-1}\tilde{\gamma}(2N^2t)$, $0 \leq t \leq t_{\tilde{\gamma}}/(2N^2)$, where $\tilde{\gamma}$ are the loops in a random walk loop soup with intensity λ which stay in ND . Note that the time length of $\tilde{\gamma}_N$ is $t_{\tilde{\gamma}}/(2N^2)$. Let $\tilde{\mathcal{L}}_N^{t_0}$ be the collection of loops in $\tilde{\mathcal{L}}_N$ with time length at least t_0 .

We will often identify curves and processes with their range in the complex plane, and a collection of curves C with the set in the plane $\bigcup_{\gamma \in C} \gamma$. For a bounded set A , we write $\text{Ext}A$ for the *exterior* of A , i.e. the unique unbounded connected component of $\mathbb{C} \setminus \bar{A}$. By $\text{Hull}A$, we denote the *hull* of A , which is the complement of $\text{Ext}A$. We write $\partial_o A$ for the topological boundary of $\text{Ext}A$, called the *outer boundary* of A . Note that $\partial A \supset \partial_o A = \partial \text{Ext}A = \partial \text{Hull}A$. For sets A, A' , the *Hausdorff distance* between A and A' is given by

$$d_H(A, A') = \inf\{\delta > 0 : A \subset (A')^\delta \text{ and } A' \subset A^\delta\},$$

where $A^\delta = \bigcup_{x \in A} B(x; \delta)$ with $B(x; \delta) = \{y : |x - y| < \delta\}$.

Let \mathcal{A} be a collection of loops in a domain D . A *chain of loops* is a sequence of loops, where each loop intersects the loop which follows it in the sequence. We call $C \subset \mathcal{A}$ a *subcluster* of \mathcal{A} if each pair of loops in C is connected via a finite chain

of loops from C . We say that C is a finite subcluster if it contains a finite number of loops. A subcluster which is maximal in terms of inclusion is called a *cluster*. A cluster C of \mathcal{A} is called *outermost* if there exists no cluster C' of \mathcal{A} such that $C' \neq C$ and $\text{Hull}C \subset \text{Hull}C'$. The *carpet* of \mathcal{A} is the set $D \setminus \bigcup_C (\text{Hull}C \setminus \partial_o C)$, where the union is over all outermost clusters C of \mathcal{A} . For collections of subsets of the plane $\mathcal{A}, \mathcal{A}'$, the *induced Hausdorff distance* is given by

$$d_H^*(\mathcal{A}, \mathcal{A}') = \inf\{\delta > 0 : \forall A \in \mathcal{A} \exists A' \in \mathcal{A}' \text{ such that } d_H(A, A') < \delta, \\ \text{and } \forall A' \in \mathcal{A}' \exists A \in \mathcal{A} \text{ such that } d_H(A, A') < \delta\}.$$

The main result of this chapter is the following theorem:

Theorem 5.2.1. *Let D be a bounded, simply connected domain, take $\lambda \in (0, 1/2]$ and $16/9 < \theta < 2$. As $N \rightarrow \infty$,*

- (i) *the collection of hulls of all outermost clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ converges in distribution to the collection of hulls of all outermost clusters of \mathcal{L} , with respect to d_H^* ,*
- (ii) *the collection of outer boundaries of all outermost clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ converges in distribution to the collection of outer boundaries of all outermost clusters of \mathcal{L} , with respect to d_H^* ,*
- (iii) *the carpet of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ converges in distribution to the carpet of \mathcal{L} , with respect to d_H .*

Note that since $\theta < 2$, $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ contains loops of time length, and hence also diameter, arbitrarily small as $N \rightarrow \infty$, so the number of loops in $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ diverges as $N \rightarrow \infty$. Theorem 5.2.1 has an analogue for the random walk loop soup with killing and the massive Brownian loop soup as defined in [12]; our proof extends to that case.

As an immediate consequence of Theorem 5.2.1 and the loop soup construction of conformal loop ensembles by Sheffield and Werner [48], we have the following corollary:

Corollary 5.2.2. *Let D be a bounded, simply connected domain, take $\lambda \in (0, 1/2]$ and $16/9 < \theta < 2$. Let $\kappa \in (8/3, 4]$ be such that $\lambda = (3\kappa - 8)(6 - \kappa)/4\kappa$. As $N \rightarrow \infty$, the collection of outer boundaries of all outermost clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ converges in distribution to CLE_κ , with respect to d_H^* .*

Note that the relation between λ and κ in Corollary 5.2.2 is not exactly as stated in [48]. It has become apparent recently (see e.g. the discussion in the introduction of [36]) that in [48, 55] a factor $1/2$ is missing in the relation between the loop soup intensity and the CLE parameter. In particular, this implies that the critical intensity of the Brownian loop soup is not $\lambda = 1$ but $\lambda = 1/2$.

We conclude this section by giving an outline of the chapter and explaining the structure of the proof of Theorem 5.2.1. The largest part of the proof is to show that, for large N , with high probability, for each large cluster C of \mathcal{L} there exists a cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ such that $d_H(\text{Ext}C, \text{Ext}\tilde{C}_N)$ is small. We will prove this fact in three steps.

First, let C be a large cluster of \mathcal{L} . We choose a finite subcluster C' of C such that $d_H(\text{Ext}C, \text{Ext}C')$ is small. A priori, it is not clear that such a finite subcluster exists

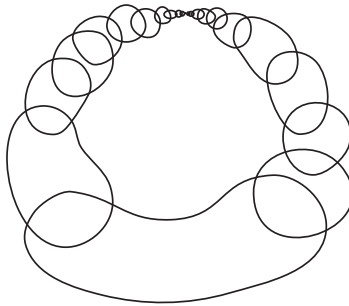


Figure 5.2: A cluster whose exterior is not well-approximated by the exterior of any finite subcluster.

– see, e.g., Figure 5.2 which depicts a cluster containing two disjoint infinite chains of loops at Euclidean distance zero from each other. A proof that, almost surely, a finite subcluster with the desired property exists is given in Section 5.4, using results from Section 5.3. The latter section contains a number of definitions and preliminary results used in the rest of the chapter.

Second, we approximate the finite subcluster C' by a finite subcluster \tilde{C}'_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$. Here we use Corollary 5.4 of Lawler and Trujillo Ferreras [31], which gives that, with probability tending to 1, there is a one-to-one correspondence between loops in $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ and loops in $\mathcal{L}^{N^{\theta-2}}$ such that corresponding loops are close. To prove that $d_H(\text{Ext}C', \text{Ext}\tilde{C}'_N)$ is small, we need results from Section 5.3 and the fact that a planar Brownian loop has no “touchings” in the sense of Definition 5.3.1 below. The latter result is proved in Section 5.5.

Third, we let \tilde{C}_N be the full cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ that contains \tilde{C}'_N . In Section 5.6 we prove an estimate which implies that, with high probability, for non-intersecting loops in $\mathcal{L}^{N^{\theta-2}}$ the corresponding loops in $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ do not intersect. We deduce from this that, for distinct subclusters $\tilde{C}'_{1,N}$ and $\tilde{C}'_{2,N}$, the corresponding clusters $\tilde{C}_{1,N}$ and $\tilde{C}_{2,N}$ are distinct. We use this property to conclude that $d_H(\text{Ext}C, \text{Ext}\tilde{C}_N)$ is small.

5.3 Preliminary results

In this section we give precise definitions and rigorous proofs of deterministic results which are important tools in the proof of our main result. Let γ_N be a sequence of curves converging uniformly to a curve γ , i.e. $d_\infty(\gamma_N, \gamma) \rightarrow 0$ as $N \rightarrow \infty$, where

$$d_\infty(\gamma, \gamma') = \sup_{s \in [0,1]} |\gamma(st_\gamma) - \gamma'(st_{\gamma'})| + |t_\gamma - t_{\gamma'}|.$$

The distance d_∞ is a natural distance on the space of curves mentioned in Section 5.1 of [29]. We will identify topological conditions that, imposed on γ (and γ_N), will yield convergence in the Hausdorff distance of the exteriors, outer boundaries and hulls of γ_N to the corresponding sets defined for γ . Note that, in general, uniform convergence of the curves does not imply convergence of any of these sets. We define

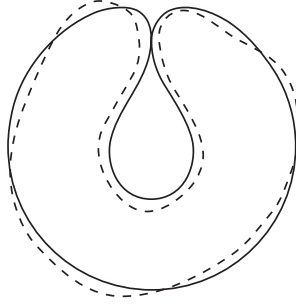


Figure 5.3: A curve with a touching, and an approximating curve (dashed).

a notion of touching (see Figure 5.3) and prove that if γ has no touchings then the desired convergence follows:

Definition 5.3.1. We say that a curve γ has a *touching* (s, t) if $0 \leq s < t \leq t_\gamma$, $\gamma(s) = \gamma(t)$ and there exists $\delta > 0$ such that for all $\varepsilon \in (0, \delta)$, there exists a curve γ' with $t_\gamma = t_{\gamma'}$, such that $d_\infty(\gamma, \gamma') < \varepsilon$ and $\gamma'[s^-, s^+] \cap \gamma'[t^-, t^+] = \emptyset$, where (s^-, s^+) is the largest subinterval of $[0, t_\gamma]$ such that $s^- \leq s \leq s^+$ and $\gamma'(s^-, s^+) \subset B(\gamma(s); \delta)$, and t^-, t^+ are defined similarly using t instead of s .

Theorem 5.3.2. Let γ_N, γ be curves such that $d_\infty(\gamma_N, \gamma) \rightarrow 0$ as $N \rightarrow \infty$, and γ has no touchings. Then,

$$d_H(\text{Ext}\gamma_N, \text{Ext}\gamma) \rightarrow 0, \quad d_H(\partial_o\gamma_N, \partial_o\gamma) \rightarrow 0, \quad \text{and} \quad d_H(\text{Hull}\gamma_N, \text{Hull}\gamma) \rightarrow 0.$$

To prove the main result of this chapter, we will also need to deal with similar convergence issues for sets defined by collections of curves. For two collections of curves C, C' let

$$d_\infty^*(C, C') = \inf\{\delta > 0 : \forall \gamma \in C \exists \gamma' \in C' \text{ such that } d_\infty(\gamma, \gamma') < \delta, \\ \text{and } \forall \gamma' \in C' \exists \gamma \in C \text{ such that } d_\infty(\gamma, \gamma') < \delta\}.$$

We will also need a modification of the notion of touching:

Definition 5.3.3. Let γ_1 and γ_2 be curves. We say that the pair γ_1, γ_2 has a *mutual touching* (s, t) if $0 \leq s \leq t_{\gamma_1}$, $0 \leq t \leq t_{\gamma_2}$, $\gamma_1(s) = \gamma_2(t)$ and there exists $\delta > 0$ such that for all $\varepsilon \in (0, \delta)$, there exist curves γ'_1, γ'_2 with $t_{\gamma_1} = t_{\gamma'_1}$, $t_{\gamma_2} = t_{\gamma'_2}$, such that $d_\infty(\gamma_1, \gamma'_1) < \varepsilon$, $d_\infty(\gamma_2, \gamma'_2) < \varepsilon$ and $\gamma'_1[s^-, s^+] \cap \gamma'_2[t^-, t^+] = \emptyset$, where (s^-, s^+) is the largest subinterval of $[0, t_{\gamma_1}]$ such that $s^- \leq s \leq s^+$ and $\gamma'_1(s^-, s^+) \subset B(\gamma_1(s); \delta)$, and t^-, t^+ are defined similarly using γ_2 and t , instead of γ_1 and s .

Definition 5.3.4. We say that a collection of curves has a *touching* if it contains a curve that has a touching or it contains a pair of distinct curves that have a mutual touching.

The next result is an analog of Theorem 5.3.2.

Theorem 5.3.5. *Let C_N, C be collections of curves such that $d_\infty^*(C_N, C) \rightarrow 0$ as $N \rightarrow \infty$, and C contains finitely many curves and C has no touchings. Then,*

$$d_H(\text{Ext}C_N, \text{Ext}C) \rightarrow 0, \quad d_H(\partial_o C_N, \partial_o C) \rightarrow 0, \quad \text{and} \quad d_H(\text{Hull}C_N, \text{Hull}C) \rightarrow 0.$$

The remainder of this section is devoted to proving Theorems 5.3.2 and 5.3.5. We will first identify a general condition for the convergence of exteriors, outer boundaries and hulls in the setting of arbitrary bounded subsets of the plane. We will prove that if a curve does not have any touchings, then this condition is satisfied and hence Theorem 5.3.2 follows. At the end of the section, we will show how to obtain Theorem 5.3.5 using similar arguments.

Proposition 5.3.6. *Let A_N, A be bounded subsets of the plane such that $d_H(A_N, A) \rightarrow 0$ as $N \rightarrow \infty$. Suppose that for every $\delta > 0$ there exists N_0 such that, for all $N > N_0$, $\text{Ext}A_N \subset (\text{Ext}A)^\delta$. Then,*

$$d_H(\text{Ext}A_N, \text{Ext}A) \rightarrow 0, \quad d_H(\partial_o A_N, \partial_o A) \rightarrow 0, \quad \text{and} \quad d_H(\text{Hull}A_N, \text{Hull}A) \rightarrow 0.$$

To prove Proposition 5.3.6, we will first prove that one of the inclusions required for the convergence of exteriors is always satisfied under the assumption that $d_H(A_N, A) \rightarrow 0$. For sets A, A' let $d_\mathcal{E}(A, A')$ be the *Euclidean distance* between A and A' .

Lemma 5.3.7. *Let A_N, A be bounded sets such that $d_H(A_N, A) \rightarrow 0$ as $N \rightarrow \infty$. Then, for every $\delta > 0$, there exists N_0 such that for all $N > N_0$, $\text{Ext}A \subset (\text{Ext}A_N)^\delta$.*

Proof. Suppose that the desired inclusion does not hold. This means that there exists $\delta > 0$ such that, after passing to a subsequence, $\text{Ext}A \not\subset (\text{Ext}A_N)^\delta$ for all N . This is equivalent to the existence of $x_N \in \text{Ext}A$, such that $d_\mathcal{E}(x_N, \text{Ext}A_N) \geq \delta$. Since $d_H(A_N, A) \rightarrow 0$ and the sets are bounded, the sequence x_N is bounded and we can assume that $x_N \rightarrow x \in \overline{\text{Ext}A}$ when $N \rightarrow \infty$. It follows that for N large enough, $d_\mathcal{E}(x, \text{Ext}A_N) > \delta/2$ and hence $B(x; \delta/2)$ does not intersect $\text{Ext}A_N$. We will show that this leads to a contradiction. To this end, note that since $x \in \overline{\text{Ext}A}$, there exists $y \in \text{Ext}A$ such that $|x - y| < \delta/4$. Furthermore, $\text{Ext}A$ is an open connected subset of \mathbb{C} , and hence it is path connected. This means that there exists a continuous path connecting y with x which stays within $\text{Ext}A$. We denote by φ its range in the complex plane. Note that $d_\mathcal{E}(\varphi, \overline{A}) > 0$. For N sufficiently large, $d_H(\overline{A}_N, \overline{A}) < d_\mathcal{E}(\varphi, \overline{A})$ and so \overline{A}_N does not intersect φ . This implies that \overline{A}_N does not disconnect y from x . Hence, $y \in \text{Ext}A_N$ and $B(x; \delta/2)$ intersects $\text{Ext}A_N$ for N large enough, which is a contradiction. This completes the proof. \square

Lemma 5.3.8. *Let A, A' be bounded sets and let $\delta > 0$. If $d_H(A, A') < \delta$ and $\text{Ext}A \subset (\text{Ext}A')^\delta$, then $\partial_o A \subset (\partial_o A')^{2\delta}$ and $\text{Hull}A' \subset (\text{Hull}A)^{2\delta}$.*

Proof. We start with the first inclusion. From the assumption, it follows that $\overline{A} \subset (A')^\delta$ and $\overline{\text{Ext}A} \subset (\text{Ext}A')^{2\delta}$. Take $x \in \partial_o A$. Since $\partial_o A \subset \overline{A} \subset (A')^\delta \subset (\text{Hull}A')^\delta$, we have that $B(x; \delta) \cap \text{Hull}A' \neq \emptyset$. Since $\partial_o A \subset \overline{\text{Ext}A} \subset (\text{Ext}A')^{2\delta}$, we have that $B(x; 2\delta) \cap \text{Ext}A' \neq \emptyset$. The ball $B(x; 2\delta)$ is connected and intersects both $\text{Ext}A'$ and its complement $\text{Hull}A'$. This implies that $B(x; 2\delta) \cap \partial_o A' \neq \emptyset$. The choice of x was arbitrary, and hence $\partial_o A \subset (\partial_o A')^{2\delta}$.

We are left with proving the second inclusion. From the assumption, it follows that $\overline{A'} \subset A^\delta$ and $\text{Ext}A \subset (\text{Ext}A')^\delta$. Since $\partial_o A' \subset \overline{A'} \subset A^\delta \subset (\text{Hull}A)^\delta$, we have that $(\partial_o A')^\delta \subset (\text{Hull}A)^{2\delta}$. Since $\text{Ext}A \subset (\text{Ext}A')^\delta = \text{Ext}A' \cup (\partial_o A')^\delta$, by taking complements we have that $\text{Hull}A' \setminus (\partial_o A')^\delta \subset \text{Hull}A \subset (\text{Hull}A)^{2\delta}$. By taking the union with $(\partial_o A')^\delta$, we obtain that $\text{Hull}A' \subset (\text{Hull}A)^{2\delta}$. \square

Proof of Proposition 5.3.6. It follows from Lemmas 5.3.7 and 5.3.8. \square

Remark 5.3.9. In the proof of Theorem 5.2.1, we will use equivalent formulations of Theorem 5.3.5 and Lemma 5.3.7 in terms of metric rather than sequential convergence. The equivalent formulation of Lemma 5.3.7 is as follows: For any bounded set A and $\delta > 0$, there exists $\varepsilon > 0$ such that if $d_H(A, A') < \varepsilon$, then $\text{Ext}A \subset (\text{Ext}A')^\delta$. The equivalent formulation of Theorem 5.3.5 is similar.

Without loss of generality, from now till the end of this section, we assume that all curves have time length 1 (this can always be achieved by a linear time change).

Definition 5.3.10. We say that $s, t \in [0, 1]$ are δ -connected in a curve γ if there exists an open ball B of diameter δ such that $\gamma(s)$ and $\gamma(t)$ are connected in $\gamma \cap B$.

Lemma 5.3.11. *Let γ_N, γ be curves such that $d_\infty(\gamma_N, \gamma) \rightarrow 0$ as $N \rightarrow \infty$, and γ has no touchings. Then for any $\delta > 0$ and s, t which are δ -connected in γ , there exists N_0 such that s, t are 4δ -connected in γ_N for all $N > N_0$.*

Proof. Fix $\delta > 0$. If the diameter of γ is at most δ , then it is enough to take N_0 , such that $d_\infty(\gamma_N, \gamma) < \delta$ for $N > N_0$.

Otherwise, let $s, t \in [0, 1]$ be δ -connected in γ and let x be such that $\gamma(s)$ and $\gamma(t)$ are in the same connected component of $\gamma \cap B(x; \delta/2)$. We say that $I = [a, b] \subset [0, 1]$ defines an *excursion* of γ from $\partial B(x; \delta)$ to $B(x; \delta/2)$ if I is a maximal interval satisfying

$$\gamma(a, b) \subset B(x; \delta) \quad \text{and} \quad \gamma(a, b) \cap B(x; \delta/2) \neq \emptyset.$$

Note that if $[a, b]$ defines an excursion, then the diameter of $\gamma[a, b]$ is at least $\delta/2$. Since γ is uniformly continuous, it follows that there are only finitely many excursions. Let $I_i = [a_i, b_i]$, $i = 1, 2, \dots, k$, be the intervals which define them.

It follows that $\gamma \cap B(x; \delta/2) \subset \bigcup_{i=1}^k \gamma[I_i]$, and hence $\gamma(s)$ and $\gamma(t)$ are in the same connected component of $\bigcup_{i=1}^k \gamma[I_i]$. If $s, t \in I_i$ for some i , then it is enough to take N_0 such that $d_\infty(\gamma_N, \gamma) < \delta$ for $N > N_0$, and the claim of the lemma follows. Otherwise, using the fact that $\gamma[I_i]$ are closed connected sets, one can reorder the intervals in such a way that $s \in I_1$, $t \in I_l$, and $\gamma[I_i] \cap \gamma[I_{i+1}] \neq \emptyset$ for $i = 1, \dots, l-1$. Let (s_i, t_i) be such that $s_i \in I_i$, $t_i \in I_{i+1}$, and $\gamma(s_i) = \gamma(t_i) = z_i$. Since (s_i, t_i) is not a touching, we can find $\varepsilon_i \in (0, \delta)$ such that $\gamma'(s_i)$ is connected to $\gamma'(t_i)$ in $\gamma' \cap B(z_i; \delta)$ for all γ' with $d(\gamma, \gamma') < \varepsilon_i$. Hence, if N_0 is such that $d(\gamma_N, \gamma) < \min\{\varepsilon, \delta\}$ for $N > N_0$, where $\varepsilon = \min_i \varepsilon_i$, then $\gamma_N(s)$ and $\gamma_N(t)$ are connected in $\bigcup_{i=1}^l \gamma_N[I_i] \cup (\gamma_N \cap \bigcup_{i=1}^{l-1} B(z_i; \delta))$, and therefore also in $\gamma_N \cap B(x; 2\delta)$. \square

Lemma 5.3.12. *If γ is a curve, then there exists a loop whose range is $\partial_o \gamma$ and whose winding around each point of $\text{Hull}\gamma \setminus \partial_o \gamma$ is equal to 2π .*

Proof. Let $D' = \{x \in \mathbb{C} : |x| > 1\}$. By the proof of Theorem 1.5(ii) of [10], there exists a one-to-one conformal map φ from D' onto $\text{Ext}\gamma$ which extends to a continuous function $\bar{\varphi} : \overline{D'} \rightarrow \overline{\text{Ext}\gamma}$, and such that $\bar{\varphi}[\partial D'] = \partial_o\gamma$. Let $\gamma_r(t) = \bar{\varphi}(e^{it2\pi}(1+r))$ for $t \in [0, 1]$ and $r \geq 0$. It follows that the range of γ_0 is $\partial_o\gamma$. Moreover, since φ is one-to-one, γ_r is a simple curve for $r > 0$ and hence its winding around every point of $\text{Hull}\gamma \setminus \partial_o\gamma$ is equal to 2π . Since $d_\infty(\gamma_0, \gamma_r) \rightarrow 0$ when $r \rightarrow 0$, the winding of γ_0 around every point of $\text{Hull}\gamma \setminus \partial_o\gamma$ is also equal to 2π . \square

Lemma 5.3.13. *Let γ_N, γ be curves such that $d_\infty(\gamma_N, \gamma) \rightarrow 0$ as $N \rightarrow \infty$. Suppose that for any $\delta > 0$ and s, t which are δ -connected in γ , there exists N_0 such that s, t are 4δ -connected in γ_N for all $N > N_0$. Then, for every $\delta > 0$, there exists N_0 such that for all $N > N_0$, $\text{Ext}\gamma_N \subset (\text{Ext}\gamma)^\delta$.*

Proof. Fix $\delta > 0$. By Lemma 5.3.12, let γ_0 be a loop whose range is $\partial_o\gamma$ and whose winding around each point of $\text{Hull}\gamma \setminus \partial_o\gamma$ equals 2π . Let

$$0 = t_0 < t_1 < \dots < t_l = 1,$$

be a sequence of times satisfying

$$t_{i+1} = \inf\{t \in [t_i, 1] : |\gamma_0(t) - \gamma_0(t_i)| = \delta/32\} \quad \text{for } i = 0, \dots, l-2,$$

and $|\gamma_0(t) - \gamma_0(t_{l-1})| < \delta/32$ for all $t \in [t_{l-1}, 1)$. This is well defined, i.e. $l < \infty$, since γ_0 is uniformly continuous. Note that t_i and t_{i+1} are $\delta/8$ -connected in γ_0 . For each t_i , we choose a time τ_i , such that $\gamma(\tau_i) = \gamma_0(t_i)$ and $\tau_l = \tau_0$. It follows that τ_i and τ_{i+1} are $\delta/8$ -connected in γ . Let N_i be so large that τ_i and τ_{i+1} are $\delta/2$ -connected in γ_N for all $N > N_i$, and let $M = \max_i N_i$. The existence of such N_i is guaranteed by the assumption of the lemma.

Let $M' > M$ be such that $d_\infty(\gamma_N, \gamma) < \delta/16$ for all $N > M'$. Take $N > M'$. We will show that $\text{Ext}\gamma_N \subset (\text{Ext}\gamma)^\delta$. Suppose by contradiction, that $x \in \text{Ext}\gamma_N \cap (\mathbb{C} \setminus (\text{Ext}\gamma)^\delta) = \text{Ext}\gamma_N \cap (\text{Hull}\gamma \setminus (\partial_o\gamma)^\delta)$. Since $\text{Ext}\gamma_N$ is open and connected, it is path connected and there exists a continuous path \wp connecting x with ∞ and such that $\wp \subset \text{Ext}\gamma_N$.

We will construct a loop γ^* which is contained in $\mathbb{C} \setminus \wp$, and which disconnects x from ∞ . This will yield a contradiction. By the definition of M , for $i = 0, \dots, l-1$, there exists an open ball B_i of diameter $\delta/2$, such that $\gamma_N(\tau_i)$ and $\gamma_N(\tau_{i+1})$ are connected in $\gamma_N \cap B_i$, and hence also in $B_i \setminus \wp$. Since the connected components of $B_i \setminus \wp$ are open, they are path connected and there exists a curve γ_i^* which starts at $\gamma_N(\tau_i)$, ends at $\gamma_N(\tau_{i+1})$, and is contained in $B_i \setminus \wp$. By concatenating these curves, we construct the loop γ^* , i.e.

$$\gamma^*(t) = \gamma_i^*\left(\frac{t - t_i}{t_{i+1} - t_i}\right) \quad \text{for } t \in [t_i, t_{i+1}], \quad i = 0, \dots, l-1.$$

By construction, $\gamma^* \subset \mathbb{C} \setminus \wp$. We will now show that γ^* disconnects x from ∞ by proving that its winding around x equals 2π . By the definition of t_{i+1} , $\gamma_0(t_i, t_{i+1}) \subset B(\gamma_0(t_i); \delta/16)$. Since $d_\infty(\gamma_N, \gamma) < \delta/16$ and $\gamma_0(t_i) = \gamma(\tau_i)$, it follows that $\gamma_0(t_i, t_{i+1}) \subset B(\gamma_N(\tau_i); \delta/8)$. By the definition of γ_i^* , $\gamma_i^* \subset B_i \subset B(\gamma_N(\tau_i); \delta/2)$. Combining these two facts, we conclude that $d_\infty(\gamma_0, \gamma^*) < 5\delta/8$. Since the winding

of γ_0 around every point of $\text{Hull}\gamma \setminus \partial_o\gamma$ is equal to 2π , and since $x \in \text{Hull}\gamma$ and $d_{\mathcal{E}}(x, \gamma_0) \geq \delta$, the winding of γ^* around x is also equal to 2π . This means that γ^* disconnects x from ∞ , and hence $\wp \cap \gamma^* \neq \emptyset$, which is a contradiction. \square

Proof of Theorem 5.3.2. It is enough to use Proposition 5.3.6, Lemma 5.3.11 and Lemma 5.3.13. \square

Proof of Theorem 5.3.5. The proof follows similar steps as the proof of Theorem 5.3.2. To adapt Lemma 5.3.11 to the setting of collections of curves, it is enough to notice that a finite collection of nontrivial curves, when intersected with a ball of sufficiently small radius, looks like a single curve intersected with the ball. To generalize Lemma 5.3.13, it suffices to notice that the outer boundary of each connected component of C is given by a curve as in Lemma 5.3.12. \square

5.4 Finite approximation of a Brownian loop soup cluster

Let \mathcal{L} be a Brownian loop soup with intensity $\lambda \in (0, 1/2]$ in a bounded, simply connected domain D . The following theorem is the main result of this section.

Theorem 5.4.1. *Almost surely, for any cluster C of \mathcal{L} , there exists a sequence of finite subclusters C_N of C such that as $N \rightarrow \infty$,*

$$d_H(\text{Ext}C_N, \text{Ext}C) \rightarrow 0, \quad d_H(\partial_o C_N, \partial_o C) \rightarrow 0, \quad \text{and} \quad d_H(\text{Hull}C_N, \text{Hull}C) \rightarrow 0.$$

We will need the following result.

Lemma 5.4.2. *Almost surely, for each cluster C of \mathcal{L} , there exists a sequence of finite subclusters C_N increasing to C (i.e. $C_N \subset C_{N+1}$ for all N and $\bigcup_N C_N = C$), and a sequence of loops $\ell_N : [0, 1] \rightarrow \mathbb{C}$ converging uniformly to a loop $\ell : [0, 1] \rightarrow \mathbb{C}$, such that the range of ℓ_N is equal to C_N , and hence the range of ℓ is equal to \overline{C} .*

Proof. This follows from the proof of Lemma 9.7 in [48]. Note that in [48], a cluster C is replaced by the collection of simple loops η given by the outer boundaries of $\gamma \in C$. However, the same argument works also for C and the loops γ . \square

To prove Theorem 5.4.1, we will show that the loops ℓ_N, ℓ from Lemma 5.4.2 satisfy the conditions of Lemma 5.3.13. Then, using Proposition 5.3.6 and Lemma 5.3.13, we obtain Theorem 5.4.1. We will first prove some necessary lemmas.

Lemma 5.4.3. *Almost surely, for all $\gamma \in \mathcal{L}$ and all subclusters C of \mathcal{L} such that γ does not intersect C , it holds that $d_{\mathcal{E}}(\gamma, C) > 0$.*

Proof. Fix k and let γ_k be the loop in \mathcal{L} with k -th largest diameter. Using an argument similar to that in Lemma 9.2 of [48], one can prove that, conditionally on γ_k , the loops in \mathcal{L} which do not intersect γ_k are distributed like $\mathcal{L}(D \setminus \gamma_k)$, i.e. a Brownian loop soup in $D \setminus \gamma_k$. Moreover, $\mathcal{L}(D \setminus \gamma_k)$ consists of a countable collection of disjoint loop soups, one for each connected component of $D \setminus \gamma_k$. By conformal invariance, each of these loop soups is distributed like a conformal image of a copy of \mathcal{L} . Hence, by Lemma 9.4

of [48], almost surely, each cluster of $\mathcal{L}(D \setminus \gamma_k)$ is at positive distance from γ_k . This implies that the unconditional probability that there exists a subcluster C such that $d_{\mathcal{E}}(\gamma_k, C) = 0$ and γ_k does not intersect C is zero. Since k was arbitrary and there are countably many loops in \mathcal{L} , the claim of the lemma follows. \square

Lemma 5.4.4. *Almost surely, for all x with rational coordinates and all rational $\delta > 0$, no two clusters of the loop soup obtained by restricting \mathcal{L} to $B(x; \delta)$ are at Euclidean distance zero from each other.*

Proof. This follows from Lemma 9.4 of [48], the restriction property of the Brownian loop soup, conformal invariance and the fact that we consider a countable number of balls. \square

Lemma 5.4.5. *Almost surely, for every $\delta > 0$ there exists $t_0 > 0$ such that every subcluster of \mathcal{L} with diameter larger than δ contains a loop of time length larger than t_0 .*

Proof. Let $\delta > 0$ and suppose that for all $t_0 > 0$ there exists a subcluster of diameter larger than δ containing only loops of time length less than t_0 .

Let $t_1 = 1$ and let C_1 be a subcluster of diameter larger than δ containing only loops of time length less than t_1 . By the definition of a subcluster there exists a finite chain of loops C'_1 which is a subcluster of C_1 and has diameter larger than δ . Let $t_2 = \min\{t_\gamma : \gamma \in C'_1\}$, where t_γ is the time length of γ . Let C_2 be a subcluster of diameter larger than δ containing only loops of time length less than t_2 . By the definition of a subcluster there exists a finite chain of loops C'_2 which is a subcluster of C_2 and has diameter larger than δ . Note that by the construction $\gamma_1 \neq \gamma_2$ for all $\gamma_1 \in C'_1, \gamma_2 \in C'_2$, i.e. the chains of loops C'_1 and C'_2 are disjoint as collections of loops, i.e. $\gamma_1 \neq \gamma_2$ for all $\gamma_1 \in C'_1, \gamma_2 \in C'_2$. Iterating the construction gives infinitely many chains of loops C'_i which are disjoint as collections of loops and which have diameter larger than δ .

For each chain of loops C'_i take a point $z_i \in C'_i$, where C'_i is viewed as a subset of the complex plane. Since the domain is bounded, the sequence z_i has an accumulation point, say z . Let z' have rational coordinates and δ' be a rational number such that $|z - z'| < \delta/8$ and $|\delta - \delta'| < \delta/8$. The annulus centered at z' with inner radius $\delta'/4$ and outer radius $\delta'/2$ is crossed by infinitely many chains of loops which are disjoint as collections of loops. However, the latter event has probability 0 by Lemma 9.6 of [48] and its consequence, leading to a contradiction. \square

Proof of Theorem 5.4.1. We restrict our attention to the event of probability 1 such that the claims of Lemmas 5.4.2, 5.4.3, 5.4.4 and 5.4.5 hold true, and such that there are only finitely many loops of diameter or time length larger than any positive threshold. Fix a realization of \mathcal{L} and a cluster C of \mathcal{L} . Take C_N, ℓ_N and ℓ defined for C as in Lemma 5.4.2. By Proposition 5.3.6 and Lemma 5.3.13, it is enough to prove that the sequence ℓ_N satisfies the condition that for all $\delta > 0$ and $s, t \in [0, 1]$ which are δ -connected in ℓ , there exists N_0 such that s, t are 4δ -connected in ℓ_N for all $N > N_0$.

To this end, take $\delta > 0$ and s, t such that $\ell(s)$ is connected to $\ell(t)$ in $\ell \cap B(x, \delta/2)$ for some x . Take x' with rational coordinates and δ' rational such that $B(x; \delta/2) \subset$

$B(x'; \delta'/2)$ and $\overline{B(x'; \delta')} \subset B(x; 2\delta)$. If $C \subset B(x'; \delta')$, then $\ell_N(s)$ is connected to $\ell_N(t)$ in $\ell_N \cap B(x; 2\delta)$ for all N and we are done. Hence, we can assume that

$$C \cap \partial B(x'; \delta') \neq \emptyset. \quad (5.4.1)$$

When intersected with $\overline{B(x'; \delta')}$, each loop $\gamma \in C$ may split into multiple connected components. We call each such component of $\gamma \cap \overline{B(x'; \delta')}$ a *piece* of γ . In particular if $\gamma \subset \overline{B(x'; \delta')}$, then the only piece of γ is the full loop γ . The collection of all pieces we consider is given by $\{\varphi : \varphi \text{ is a piece of } \gamma \text{ for some } \gamma \in C\}$. A *chain of pieces* is a sequence of pieces such that each piece intersects the next piece in the sequence. Two pieces are in the same *cluster of pieces* if they are connected via a finite chain of pieces. We identify a collection of pieces with the set in the plane given by the union of the pieces. Note that there are only finitely many pieces of diameter larger than any positive threshold, since the number of loops of diameter larger than any positive threshold is finite and each loop is uniformly continuous.

Let C_1^*, C_2^*, \dots be the clusters of pieces such that

$$C_i^* \cap B(x'; \delta'/2) \neq \emptyset \text{ and } C_i^* \cap \partial B(x'; \delta') \neq \emptyset. \quad (5.4.2)$$

We will see later in the proof that the number of such clusters of pieces is finite, but we do not need this fact yet. We now prove that

$$\overline{C_i^*} \cap \overline{C_j^*} \cap B(x'; \delta'/2) = \emptyset \text{ for all } i \neq j. \quad (5.4.3)$$

To this end, suppose that (5.4.3) is false and let $z \in \overline{C_i^*} \cap \overline{C_j^*} \cap B(x'; \delta'/2)$ for some $i \neq j$.

First assume that $z \in C_i^*$. Then, by the definition of clusters of pieces, $z \notin C_j^*$. It follows that C_j^* contains a chain of infinitely many different pieces which has z as an accumulation point. Since there are only finitely many pieces of diameter larger than any positive threshold, the diameters of the pieces in this chain approach 0. Since $d_{\mathcal{E}}(z, \partial B(x'; \delta')) > \delta'/2$, the pieces become full loops at some point in the chain. Let $\gamma \in C$ be such that $z \in \gamma$. It follows that there exists a subcluster of loops of C , which does not contain γ and has z as an accumulation point. This contradicts the claim of Lemma 5.4.3 and therefore it cannot be the case that $z \in C_i^*$.

Second assume that $z \notin C_i^*$ and $z \notin C_j^*$. By the same argument as in the previous paragraph, there exist two chains of loops of C which are disjoint, contained in $B(x'; \delta')$ and both of which have z as an accumulation point. These two chains belong to two different clusters of \mathcal{L} restricted to $B(x'; \delta')$. Since x' and δ' are rational, this contradicts the claim of Lemma 5.4.4, and hence it cannot be the case that $z \notin C_i^*$ and $z \notin C_j^*$. This completes the proof of (5.4.3).

We now define a particular collection of pieces P . By Lemma 5.4.5, let $t_0 > 0$ be such that every subcluster of \mathcal{L} of diameter larger than $\delta'/4$ contains a loop of time length larger than t_0 . Let P be the collection of pieces which have diameter larger than $\delta'/4$ or are full loops of time length larger than t_0 . Note that P is finite. Each chain of pieces which intersects both $B(x'; \delta'/2)$ and $\partial B(x'; \delta')$, contains a piece of diameter larger than $\delta'/4$ intersecting $\partial B(x'; \delta')$ or contains a chain of full loops which intersects both $B(x'; \delta'/2)$ and $\partial B(x'; 3\delta'/4)$. In the latter case it contains a subcluster of \mathcal{L} of diameter larger than $\delta'/4$ and therefore a full loop of time length

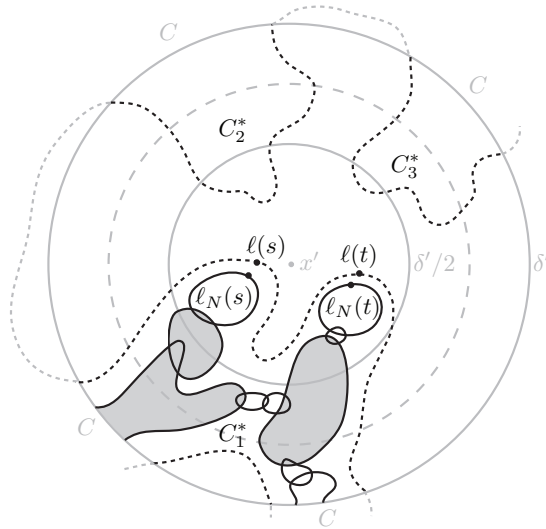


Figure 5.4: Illustration of the last part of the proof of Theorem 5.4.1 with $C_i^* = C_1^*$. The pieces drawn with solid lines form the set $C_i^* \cap \ell_N$. The shaded pieces represent the set $C_i^* \cap P$.

larger than t_0 . Hence, each chain of pieces which intersects both $B(x'; \delta'/2)$ and $\partial B(x'; \delta')$ contains an element of P . Since P is finite, it follows that the number of clusters of pieces C_i^* satisfying (5.4.2) is finite.

Since the range of ℓ is \overline{C} and the number of clusters of pieces C_i^* is finite,

$$\begin{aligned} \ell \cap B(x'; \delta'/2) &= \overline{C} \cap B(x'; \delta'/2) \\ &= \overline{\bigcup_i C_i^*} \cap B(x'; \delta'/2) = \bigcup_i \overline{C_i^*} \cap B(x'; \delta'/2). \end{aligned} \quad (5.4.4)$$

By (5.4.3), (5.4.4) and the fact that $\ell(s)$ is connected to $\ell(t)$ in $\ell \cap B(x'; \delta'/2)$,

$$\ell(s), \ell(t) \in \overline{C_i^*} \cap B(x'; \delta'/2), \quad (5.4.5)$$

for some i . From now on see also Figure 5.4.

Let ε be the Euclidean distance between $\{\ell(s), \ell(t)\}$ and $\partial B(x'; \delta'/2) \cup \bigcup_{j \neq i} \overline{C_j^*}$. By (5.4.3) and (5.4.5), $\varepsilon > 0$. Let M be such that $d_\infty(\ell_N, \ell) < \varepsilon$ and $\ell_N \cap \partial B(x'; \delta') \neq \emptyset$ for $N > M$. The latter can be achieved by (5.4.1). Let $N > M$. By the definitions of ε and M , we have that $\ell_N(s), \ell_N(t) \in B(x'; \delta'/2)$ and $\ell_N(s), \ell_N(t) \notin C_j^*$ for $j \neq i$. It follows that

$$\ell_N(s), \ell_N(t) \in C_i^* \cap B(x'; \delta'/2).$$

Since ℓ_N is a finite subcluster of C , it also follows that there are finite chains of pieces $G_N^*(s), G_N^*(t) \subset C_i^* \cap \ell_N$ (not necessarily distinct) which connect $\ell_N(s), \ell_N(t)$, respectively, to $\partial B(x'; \delta')$.

Since $G_N^*(s), G_N^*(t)$ intersect both $B(x'; \delta'/2)$ and $\partial B(x'; \delta')$, we have that $G_N^*(s), G_N^*(t)$ both contain an element of P . Moreover, P is finite, any two elements of C_i^* are

connected via a finite chain of pieces and $\ell_N (= C_N)$ increases to the full cluster C . Hence, all elements of $C_i^* \cap P$ are connected to each other in $C_i^* \cap \ell_N$ for N sufficiently large. It follows that $G_N^*(s)$ is connected to $G_N^*(t)$ in $C_i^* \cap \ell_N$ for N sufficiently large. Hence, $\ell_N(s)$ is connected to $\ell_N(t)$ in $\ell_N \cap \overline{B(x'; \delta')}$ for N sufficiently large. This implies that s, t are 4δ -connected in ℓ_N for N sufficiently large. \square

5.5 No touchings

Recall the definitions of touching, Definitions 5.3.1, 5.3.3 and 5.3.4. In this section we prove the following:

Theorem 5.5.1. *Let B_t be a planar Brownian motion. Almost surely, B_t , $0 \leq t \leq 1$, has no touchings.*

Corollary 5.5.2.

- (i) *Let B_t^{loop} be a planar Brownian loop with time length 1. Almost surely, B_t^{loop} , $0 \leq t \leq 1$, has no touchings.*
- (ii) *Let \mathcal{L} be a Brownian loop soup with intensity $\lambda \in (0, \infty)$ in a bounded, simply connected domain D . Almost surely, \mathcal{L} has no touchings.*

We start by giving a sketch of the proof of Theorem 5.5.1. Note that ruling out isolated touchings can be done using the fact that the intersection exponent $\zeta(2, 2)$ is larger than 2 (see [30]). However, also more complicated situations like accumulations of touchings can occur. Therefore, we proceed as follows. We define excursions of the planar Brownian motion B from the boundary of a disk which stay in the disk. Each of these excursions has, up to a rescaling in space and time, the same law as a process W which we define below. We show that the process W possesses a particular property, see Lemma 5.5.6 below. If B had a touching, it would follow that the excursions of B would have a behavior that is incompatible with this particular property of the process W .

As a corollary to Theorem 5.5.1, Corollary 5.5.2 and Theorem 5.3.2, we obtain the following result. It is a natural result, but we could not find a version of this result in the literature and therefore we include it here.

Corollary 5.5.3. *Let S_t , $t \in \{0, 1, 2, \dots\}$, be a simple random walk on the square lattice \mathbb{Z}^2 , with $S_0 = 0$, and define S_t for non-integer times t by linear interpolation.*

- (i) *Let B_t be a planar Brownian motion started at 0. As $N \rightarrow \infty$, the outer boundary of $(N^{-1}S_{2N^2t}, 0 \leq t \leq 1)$ converges in distribution to the outer boundary of $(B_t, 0 \leq t \leq 1)$, with respect to d_H .*
- (ii) *Let B_t^{loop} be a planar Brownian loop of time length 1 started at 0. As $N \rightarrow \infty$, the outer boundary of $(N^{-1}S_{2N^2t}, 0 \leq t \leq 1)$, conditional on $\{S_{2N^2} = 0\}$, converges in distribution to the outer boundary of $(B_t^{\text{loop}}, 0 \leq t \leq 1)$, with respect to d_H .*

To define the process W mentioned above, we recall some facts about the three-dimensional Bessel process and its relation with Brownian motion, see e.g. Lemma 1 of [11] and the references therein. The three-dimensional Bessel process can be defined as the modulus of a three-dimensional Brownian motion.

Lemma 5.5.4. *Let X_t be a one-dimensional Brownian motion starting at 0 and Y_t a three-dimensional Bessel process starting at 0. Let $0 < a < a'$ and define $\tau = T_a(X) = \inf\{t \geq 0 : |X_t| = a\}$, $\tau' = T_{a'}(X)$, $\sigma = \sup\{t < \tau : X_t = 0\}$, $\rho = T_a(Y)$ and $\rho' = T_{a'}(Y)$. Then,*

- (i) *the two processes $(X_{\sigma+u}, 0 \leq u \leq \tau - \sigma)$ and $(Y_u, 0 \leq u \leq \rho)$ have the same law,*
- (ii) *the process $(Y_{\rho+u}, 0 \leq u \leq \rho' - \rho)$ has the same law as the process $(X_{\tau+u}, 0 \leq u \leq \tau' - \tau)$ conditional on $\{\forall u \in [0, \tau' - \tau], X_{\tau+u} \neq 0\}$.*

Next we recall the skew-product representation of planar Brownian motion, see e.g. Theorem 7.26 of [41]: For a planar Brownian motion B_t starting at 1, there exist two independent one-dimensional Brownian motions X_t^1 and X_t^2 starting at 0 such that

$$B_t = \exp(X_{H(t)}^1 + iX_{H(t)}^2),$$

where

$$H(t) = \inf \left\{ h \geq 0 : \int_0^h \exp(2X_u^1) du > t \right\} = \int_0^t \frac{1}{|B_u|^2} du.$$

We define the process W_t as follows. Let X_t be a one-dimensional Brownian motion starting according to some distribution on $[0, 2\pi)$. Let Y_t be a three-dimensional Bessel process starting at 0, independent of X_t . Define

$$V_t = \exp(-Y_{H(t)} + iX_{H(t)}),$$

where

$$H(t) = \inf \left\{ h \geq 0 : \int_0^h \exp(-2Y_u) du > t \right\}.$$

Let B_t be a planar Brownian motion starting at 0, independent of X_t and Y_t , and define

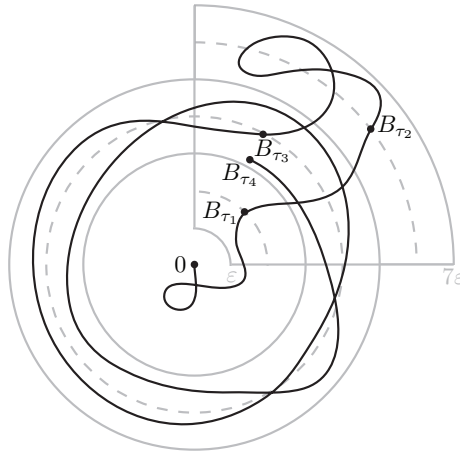
$$W_t = \begin{cases} V_t, & 0 \leq t \leq \tau_{\frac{1}{2}}, \\ V_{\tau_{\frac{1}{2}}} + B_{t-\tau_{\frac{1}{2}}}, & \tau_{\frac{1}{2}} < t \leq \tau, \end{cases}$$

with

$$\begin{aligned} \tau_{\frac{1}{2}} &= \inf\{t > 0 : |V_t| = \frac{1}{2}\}, \\ \tau &= \inf\{t > \tau_{\frac{1}{2}} : |V_{\tau_{\frac{1}{2}}} + B_{t-\tau_{\frac{1}{2}}}| = 1\}. \end{aligned}$$

Note that W_t starts on the unit circle, stays in the unit disk and is stopped when it hits the unit circle again.

Next we derive the property of W which we will use in the proof of Theorem 5.5.1. For this, we need the following property of planar Brownian motion:

Figure 5.5: The event E_ε .

Lemma 5.5.5. *Let B be a planar Brownian motion started at 0 and stopped when it hits the unit circle. Almost surely, there exists $\varepsilon > 0$ such that for all curves γ with $d_\infty(\gamma, B) < \varepsilon$ we have that γ disconnects $\partial B(0; \varepsilon)$ from $\partial B(0; 1)$.*

Proof. We construct the event E_ε , for $0 < \varepsilon \leq 1/7$, illustrated in Figure 5.5. Loosely speaking, E_ε is the event that B disconnects 0 from the unit circle in a strong sense, by crossing an annulus centered at 0 and winding around twice in this annulus. Let

$$\begin{aligned} \tau_1 &= \inf\{t \geq 0 : |B_t| = 2\varepsilon\}, & \tau_2 &= \inf\{t \geq 0 : |B_t| = 6\varepsilon\}, \\ \tau_3 &= \inf\{t > \tau_2 : |B_t| = 4\varepsilon\}, & \tau_4 &= \inf\{t > \tau_3 : |\arg(B_t/B_{\tau_3})| = 4\pi\}, \end{aligned}$$

where \arg is the continuous determination of the angle. Let

$$\begin{aligned} A_1 &= \{z \in \mathbb{C} : \varepsilon < |z| < 7\varepsilon, |\arg(z/B_{\tau_1})| < \pi/4\}, \\ A_2 &= \{z \in \mathbb{C} : 3\varepsilon < |z| < 5\varepsilon\}. \end{aligned}$$

Define the event E_ε by

$$E_\varepsilon = \{\tau_4 < \infty, B[\tau_1, \tau_3] \subset A_1, B[\tau_3, \tau_4] \subset A_2\}.$$

By construction, if E_ε occurs then for all curves γ with $d_\infty(\gamma, B) < \varepsilon$ we have that γ disconnects $\partial B(0; 2\varepsilon)$ from $\partial B(0; 6\varepsilon)$. It remains to prove that almost surely E_ε occurs for some ε . By scale invariance of Brownian motion, $\mathbb{P}(E_\varepsilon)$ does not depend on ε , and it is obvious that $\mathbb{P}(E_\varepsilon) > 0$. Furthermore, the events $E_{1/7^n}$, $n \in \mathbb{N}$, are independent. Hence almost surely E_ε occurs for some ε . \square

Lemma 5.5.6. *Let $\gamma : [0, 1] \rightarrow \mathbb{C}$ be a curve with $|\gamma(0)| = |\gamma(1)| = 1$ and $|\gamma(t)| < 1$ for all $t \in (0, 1)$. Let W denote the process defined above Lemma 5.5.5 and assume that $W_0 \notin \{\gamma(0), \gamma(1)\}$ a.s. Then the intersection of the following two events has probability 0:*

(i) $\gamma \cap W \neq \emptyset$,

(ii) for all $\varepsilon > 0$ there exist curves γ', γ'' such that $d_\infty(\gamma, \gamma') < \varepsilon$, $d_\infty(W, \gamma'') < \varepsilon$ and $\gamma' \cap \gamma'' = \emptyset$.

Proof. The idea of the proof is as follows. We run the process W_t till it hits $\partial B(0; a)$, where $a < 1$ is close to 1. From that point the process is distributed as a conditional Brownian motion. We run the Brownian motion till it hits the trace of the curve γ . From that point the Brownian motion winds around such that the event (ii) cannot occur, by Lemma 5.5.5.

Let $T_a(W) = \inf\{t \geq 0 : |W_t| = a\}$ and let P be the law of $W_{T_a(W)}$. Let B_t be a planar Brownian motion with starting point distributed according to the law P and stopped when it hits the unit circle. Let $\tau = \inf\{t > 0 : |W_t| = 1\}$. By Lemma 5.5.4 and the skew-product representation, if $a \in (\frac{1}{2}, 1)$, the process

$$(W_t, T_a(W) \leq t \leq \tau)$$

has the same law as

$$(B_t, 0 \leq t \leq T_1(B)) \text{ conditional on } \{T_{1/2}(B) < T_1(B)\},$$

where $T_1(B) = \inf\{t \geq 0 : |B_t| = 1\}$. Let E_1, E_2 be similar to the events (i) and (ii), respectively, from the statement of the lemma, but with B instead of W , i.e.

$$E_1 = \{\gamma \cap B \neq \emptyset\},$$

$$E_2 = \{\text{for all } \varepsilon > 0 \text{ there exist curves } \gamma', \gamma'' \text{ such that } d_\infty(\gamma, \gamma') < \varepsilon, \\ d_\infty(B, \gamma'') < \varepsilon, \gamma' \cap \gamma'' = \emptyset\}.$$

Let $T_\gamma(W) = \inf\{t \geq 0 : W_t \in \gamma\}$ be the first time W_t hits the trace of the curve γ .

The probability of the intersection of the events (i) and (ii) from the statement of the lemma is bounded above by

$$\begin{aligned} & \mathbb{P}(E_1 \cap E_2 \mid T_{1/2}(B) < T_1(B)) + \mathbb{P}(T_\gamma(W) \leq T_a(W)) \\ & \leq \frac{\mathbb{P}(E_2 \mid E_1) \mathbb{P}(E_1)}{\mathbb{P}(T_{1/2}(B) < T_1(B))} + \mathbb{P}(T_\gamma(W) \leq T_a(W)). \end{aligned} \quad (5.5.1)$$

The second term in (5.5.1) converges to 0 as $a \rightarrow 1$, by the assumption that $W_0 \notin \{\gamma(0), \gamma(1)\}$ a.s. The first term in (5.5.1) is equal to 0. This follows from the fact that

$$\mathbb{P}(E_2 \mid E_1) = 0, \quad (5.5.2)$$

which we prove below, using Lemma 5.5.5.

To prove (5.5.2) note that $E_1 = \{T_\gamma(B) \leq T_1(B)\}$, where $T_\gamma(B) = \inf\{t \geq 0 : B_t \in \gamma\}$. Define $\delta = 1 - |B_{T_\gamma(B)}|$ and note that $\delta > 0$ a.s. The time $T_\gamma(B)$ is a stopping time and hence, by the strong Markov property, $B_t, t \geq T_\gamma(B)$, is a Brownian motion. Therefore, by translation and scale invariance, we can apply Lemma 5.5.5 to B_t started at time $T_\gamma(B)$ and stopped when it hits the boundary of the ball centered at $B_{T_\gamma(B)}$ with radius δ . It follows that (5.5.2) holds. \square

Proof of Theorem 5.5.1. For $\delta_0 > 0$ we say that a curve $\gamma : [0, 1] \rightarrow \mathbb{C}$ has a δ_0 -touching (s, t) if (s, t) is a touching and we can take $\delta = \delta_0$ in Definition 5.3.1, and moreover $A \cap \partial B(\gamma(s); \delta_0) \neq \emptyset$ for all $A \in \{\gamma[0, s], \gamma(s, t), \gamma(t, 1]\}$. The last condition ensures that if (s, t) is a δ_0 -touching then γ makes excursions from $\partial B(\gamma(s); \delta_0)$ which visit $\gamma(s)$.

Since $B_t \neq B_0$ for all $t \in (0, 1]$ a.s., we have that $(0, t)$ is not a touching for all $t \in (0, 1]$ a.s. By time inversion, $B_1 - B_{1-u}$, $0 \leq u \leq 1$, is a planar Brownian motion and hence $(s, 1)$ is not a touching for all $s \in [0, 1)$ a.s. For every touching (s, t) with $0 < s < t < 1$ there exists $\delta' > 0$ such that for all $\delta \leq \delta'$ we have that (s, t) is a δ -touching a.s. (A touching (s, t) that is not a δ -touching for any $\delta > 0$ could only exist if $B_u = B_0$ for all $u \in [0, s]$ or $B_u = B_1$ for all $u \in [t, 1]$.) We prove that for every $\delta > 0$ we have almost surely,

$$B \text{ has no } \delta\text{-touchings } (s, t) \text{ with } 0 < s < t < 1. \quad (5.5.3)$$

By letting $\delta \rightarrow 0$ it follows that B has no touchings a.s.

To prove (5.5.3), fix $\delta > 0$ and let $z \in \mathbb{C}$. We define excursions W^n , for $n \in \mathbb{N}$, of the Brownian motion B as follows. Let

$$\tau_0 = \inf\{u \geq 0 : |B_u - z| = 2\delta/3\},$$

and define for $n \geq 1$,

$$\sigma_n = \inf\{u > \tau_{n-1} : |B_u - z| = \delta/3\},$$

$$\rho_n = \sup\{u < \sigma_n : |B_u - z| = 2\delta/3\},$$

$$\tau_n = \inf\{u > \sigma_n : |B_u - z| = 2\delta/3\}.$$

Note that $\rho_n < \sigma_n < \tau_n < \rho_{n+1}$ and that ρ_n, σ_n, τ_n may be infinite. The reason that we take $2\delta/3$ instead of δ is that we will consider δ -touchings (s, t) not only with $B_s = z$ but also with $|B_s - z| < \delta/3$. We define the excursion W^n by

$$W_u^n = B_u, \quad \rho_n \leq u \leq \tau_n.$$

Observe that W^n has, up to a rescaling in space and time and a translation, the same law as the process W defined above Lemma 5.5.5. This follows from Lemma 5.5.4, the skew-product representation and Brownian scaling.

If B has a δ -touching (s, t) with $|B_s - z| < \delta/3$, then there exist $m \neq n$ such that

- (i) $W^m \cap W^n \neq \emptyset$,
- (ii) for all $\varepsilon > 0$ there exist curves γ^m, γ^n such that $d_\infty(\gamma^m, W^m) < \varepsilon$, $d_\infty(\gamma^n, W^n) < \varepsilon$ and $\gamma^m \cap \gamma^n = \emptyset$.

By Lemma 5.5.6, with W^m playing the role of W and W^n of γ , for each m, n such that $m \neq n$ the intersection of the events (i) and (ii) has probability 0. Here we use the fact that $W_{\rho_m}^m \notin \{W_{\rho_n}^n, W_{\tau_n}^n\}$ a.s. Hence B has no δ -touchings (s, t) with $|B_s - z| < \delta/3$ a.s. We can cover the plane with a countable number of balls of radius $\delta/3$ and hence B has no δ -touchings a.s. \square

Proof of Corollary 5.5.2. First we prove part (i). For any $u_0 \in (0, 1)$, the laws of the processes $B_u^{\text{loop}}, 0 \leq u \leq u_0$, and $B_u, 0 \leq u \leq u_0$, are mutually absolutely continuous, see e.g. Exercise 1.5(b) of [41]. Hence by Theorem 5.5.1 the process $B_u^{\text{loop}}, 0 \leq u \leq 1$, has no touchings (s, t) with $0 \leq s < t \leq u_0$ a.s., for any $u_0 \in (0, 1)$. Taking a sequence of u_0 converging to 1, we have that $B_u^{\text{loop}}, 0 \leq u \leq 1$, has no touchings (s, t) with $0 \leq s < t < 1$ a.s. By time reversal, $B_1^{\text{loop}} - B_{1-u}^{\text{loop}}, 0 \leq u \leq 1$, is a planar Brownian loop. It follows that $B_u^{\text{loop}}, 0 \leq u \leq 1$, has no touchings $(s, 1)$ with $s \in (0, 1)$ a.s. By Lemma 5.5.5, the time pair $(0, 1)$ is not a touching a.s.

Second we prove part (ii). By Corollary 5.5.2 and the fact that there are countably many loops in \mathcal{L} , we have that every loop in \mathcal{L} has no touchings a.s. We prove that each pair of loops in \mathcal{L} has no mutual touchings a.s. To this end, we discover the loops in \mathcal{L} one by one in decreasing order of their diameter, similarly to the construction in Section 4.3 of [42]. Given a set of discovered loops $\gamma_1, \dots, \gamma_{k-1}$, we prove that the next loop γ_k and the already discovered loop γ_i have no mutual touchings a.s., for each $i \in \{1, \dots, k-1\}$ separately. Note that, conditional on $\gamma_1, \dots, \gamma_{k-1}$, we can treat γ_i as a deterministic loop, while γ_k is a (random) planar Brownian loop. Therefore, to prove that γ_k and γ_i have no mutual touchings a.s., we can define excursions of γ_i and γ_k and apply Lemma 5.5.6 in a similar way as in the proof of Theorem 5.5.1. We omit the details. \square

5.6 Distance between Brownian loops

In this section we give two estimates, on the Euclidean distance between non-intersecting loops in the Brownian loop soup and on the overlap between intersecting loops in the Brownian loop soup. We will only use the first estimate in the proof of Theorem 5.2.1. As a corollary to the two estimates, we obtain a one-to-one correspondence between clusters composed of “large” loops from the random walk loop soup and clusters composed of “large” loops from the Brownian loop soup. This is an extension of Corollary 5.4 of [31]. For intersecting loops γ_1, γ_2 we define their *overlap* by

$$\text{overlap}(\gamma_1, \gamma_2) = 2 \sup\{\varepsilon \geq 0 : \text{for all loops } \gamma'_1, \gamma'_2 \text{ such that } d_\infty(\gamma_1, \gamma'_1) \leq \varepsilon, \\ d_\infty(\gamma_2, \gamma'_2) \leq \varepsilon, \text{ we have that } \gamma'_1 \cap \gamma'_2 \neq \emptyset\}.$$

Proposition 5.6.1. *Let \mathcal{L} be a Brownian loop soup with intensity $\lambda \in (0, \infty)$ in a bounded, simply connected domain D . Let $c > 0$ and $16/9 < \theta < 2$. For all non-intersecting loops $\gamma, \gamma' \in \mathcal{L}$ of time length at least $N^{\theta-2}$ we have that $d_\varepsilon(\gamma, \gamma') \geq cN^{-1} \log N$, with probability tending to 1 as $N \rightarrow \infty$.*

Proposition 5.6.2. *Let \mathcal{L} be a Brownian loop soup with intensity $\lambda \in (0, \infty)$ in a bounded, simply connected domain D . Let $c > 0$ and $\theta < 2$ sufficiently close to 2. For all intersecting loops $\gamma, \gamma' \in \mathcal{L}$ of time length at least $N^{\theta-2}$ we have that $\text{overlap}(\gamma, \gamma') \geq cN^{-1} \log N$, with probability tending to 1 as $N \rightarrow \infty$.*

Corollary 5.6.3. *Let D be a bounded, simply connected domain, take $\lambda \in (0, \infty)$ and $\theta < 2$ sufficiently close to 2. Let $\mathcal{L}, \mathcal{L}^{N^{\theta-2}}, \tilde{\mathcal{L}}_N, \tilde{\mathcal{L}}_N^{N^{\theta-2}}$ be defined as in Section 5.2. For every N we can define $\tilde{\mathcal{L}}_N$ and \mathcal{L} on the same probability space in such a way that the following holds with probability tending to 1 as $N \rightarrow \infty$. There is a one-to-one correspondence between the clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ and the clusters of $\mathcal{L}^{N^{\theta-2}}$ such*

that for corresponding clusters, $\tilde{C} \subset \tilde{\mathcal{L}}_N^{N^{\theta-2}}$ and $C \subset \mathcal{L}^{N^{\theta-2}}$, there is a one-to-one correspondence between the loops in \tilde{C} and the loops in C such that for corresponding loops, $\tilde{\gamma} \in \tilde{C}$ and $\gamma \in C$, we have that $d_\infty(\gamma, \tilde{\gamma}) \leq cN^{-1} \log N$, for some constant c which does not depend on N .

Proof. Let c be two times the constant in Corollary 5.4 of [31]. Combine this corollary and Propositions 5.6.1 and 5.6.2 with the c in Propositions 5.6.1 and 5.6.2 equal to six times the constant in Corollary 5.4 of [31]. \square

In Propositions 5.6.1 and 5.6.2 and Corollary 5.6.3, the probability tends to 1 as a power of N . This can be seen from the proofs. We will use Proposition 5.6.1, but we will not use Proposition 5.6.2 in the proof of Theorem 5.2.1. Because of this, and because the proofs of Propositions 5.6.1 and 5.6.2 are based on similar techniques, we omit the proof of Proposition 5.6.2. To prove Proposition 5.6.1, we first prove two lemmas.

Lemma 5.6.4. *Let B be a planar Brownian motion and let B^{loop, t_0} be a planar Brownian loop with time length t_0 . There exist $c_1, c_2 > 0$ such that, for all $0 < \delta < \delta'$ and all $N \geq 1$,*

$$\mathbb{P}(\text{diam}B[0, N^{-\delta}] \leq N^{-\delta'/2}) \leq c_1 \exp(-c_2 N^{\delta'-\delta}), \quad (5.6.1)$$

$$\mathbb{P}(\text{diam}B^{\text{loop}, N^{-\delta}} \leq N^{-\delta'/2}) \leq c_1 \exp(-c_2 N^{\delta'-\delta}). \quad (5.6.2)$$

Proof. First we prove (5.6.2). By Brownian scaling,

$$\begin{aligned} & \mathbb{P}(\text{diam}B^{\text{loop}, N^{-\delta}} \leq N^{-\delta'/2}) \\ &= \mathbb{P}(\text{diam}B^{\text{loop}, 1} \leq N^{-(\delta'-\delta)/2}) \\ &\leq \mathbb{P}(\sup_{t \in [0,1]} |X_t^{\text{loop}}| \leq N^{-(\delta'-\delta)/2})^2, \end{aligned}$$

where X_t^{loop} is a one-dimensional Brownian bridge starting at 0 with time length 1. The distribution of $\sup_{t \in [0,1]} |X_t^{\text{loop}}|$ is the asymptotic distribution of the (scaled) Kolmogorov-Smirnov statistic, and we can write, see e.g. Theorem 1 of [25],

$$\begin{aligned} & \mathbb{P}(\sup_{t \in [0,1]} |X_t^{\text{loop}}| \leq N^{-(\delta'-\delta)/2}) \\ &= \sqrt{2\pi} N^{(\delta'-\delta)/2} \sum_{k=1}^{\infty} e^{-(2k-1)^2 \pi^2 8^{-1} N^{\delta'-\delta}} \\ &\leq \sqrt{2\pi} N^{(\delta'-\delta)/2} \sum_{k=1}^{\infty} e^{-(2k-1)\pi^2 8^{-1} N^{\delta'-\delta}} \\ &= \sqrt{2\pi} N^{(\delta'-\delta)/2} e^{\pi^2 8^{-1} N^{\delta'-\delta}} \sum_{k=1}^{\infty} (e^{-2\pi^2 8^{-1} N^{\delta'-\delta}})^k \\ &= \sqrt{2\pi} N^{(\delta'-\delta)/2} e^{-\pi^2 8^{-1} N^{\delta'-\delta}} (1 - e^{-2\pi^2 8^{-1} N^{\delta'-\delta}})^{-1} \\ &\leq c e^{-N^{\delta'-\delta}}, \end{aligned} \quad (5.6.3)$$

for some constant c and all $0 < \delta < \delta'$ and all $N \geq 1$. This proves (5.6.2).

Next we prove (5.6.1). We can write $X_t^{\text{loop}} = X_t - tX_1$, where X_t is a one-dimensional Brownian motion starting at 0. Hence

$$\sup_{t \in [0,1]} |X_t^{\text{loop}}| \leq \sup_{t \in [0,1]} |X_t| + |X_1| \leq 2 \sup_{t \in [0,1]} |X_t|. \quad (5.6.4)$$

By Brownian scaling, (5.6.4) and (5.6.3),

$$\begin{aligned}
& \mathbb{P}(\text{diam}B[0, N^{-\delta}] \leq N^{-\delta'/2}) \\
&= \mathbb{P}(\text{diam}B[0, 1] \leq N^{-(\delta'-\delta)/2}) \\
&\leq \mathbb{P}(\sup_{t \in [0,1]} |X_t| \leq N^{-(\delta'-\delta)/2})^2 \\
&\leq \mathbb{P}(\sup_{t \in [0,1]} |X_t^{\text{loop}}| \leq 2N^{-(\delta'-\delta)/2})^2 \\
&\leq c^2 e^{-\frac{1}{2}N^{\delta'-\delta}}.
\end{aligned}$$

This proves (5.6.1). \square

Lemma 5.6.5. *There exist $c_1, c_2 > 0$ such that the following holds. Let $c > 0$ and $0 < \delta < \delta' < 2$. Let γ be a (deterministic) loop with $\text{diam}\gamma \geq N^{-\delta'/2}$. Let B^{loop, t_0} be a planar Brownian loop starting at 0 of time length $t_0 \geq N^{-\delta}$. Then for all $N > 1$,*

$$\begin{aligned}
& \mathbb{P}(0 < d_{\mathcal{E}}(B^{\text{loop}, t_0}, \gamma) \leq cN^{-1} \log N) \\
&\leq c_1 N^{-1/2 + \delta'/4} (c \log N)^{1/2} + c_1 \exp(-c_2 N^{\delta'-\delta})
\end{aligned}$$

Proof. We use some ideas from the proof of Proposition 5.1 of [31]. By time reversal, we have

$$\begin{aligned}
& \mathbb{P}(0 < d_{\mathcal{E}}(B^{\text{loop}, t_0}, \gamma) \leq cN^{-1} \log N) \\
&\leq 2 \mathbb{P}(0 < d_{\mathcal{E}}(B^{\text{loop}, t_0}[0, \frac{1}{2}t_0], \gamma) \leq cN^{-1} \log N, B^{\text{loop}, t_0}[0, \frac{3}{4}t_0] \cap \gamma = \emptyset) \\
&= 4t_0 \lim_{\varepsilon \downarrow 0} \varepsilon^{-2} \mathbb{P}(0 < d_{\mathcal{E}}(B[0, \frac{1}{2}t_0], \gamma) \leq cN^{-1} \log N, B[0, \frac{3}{4}t_0] \cap \gamma = \emptyset, \\
&\quad |B_{t_0}| \leq \varepsilon), \tag{5.6.5}
\end{aligned}$$

where B is a planar Brownian motion starting at 0. The equality (5.6.5) follows from the following relation between the law $\mu_{0, t_0}^{\#}$ of $(B_t^{\text{loop}, t_0}, 0 \leq t \leq t_0)$ and the law μ_{0, t_0} of $(B_t, 0 \leq t \leq t_0)$:

$$\mu_{0, t_0}^{\#} = 2t_0 \lim_{\varepsilon \downarrow 0} \varepsilon^{-2} \mu_{0, t_0} \mathbb{1}_{\{|\gamma(t_0)| \leq \varepsilon\}},$$

see Section 5.2 of [29] and Section 3.1.1 of [32].

Next we bound the probability

$$\mathbb{P}(0 < d_{\mathcal{E}}(B[0, \frac{1}{2}t_0], \gamma) \leq cN^{-1} \log N, B[0, \frac{3}{4}t_0] \cap \gamma = \emptyset). \tag{5.6.6}$$

If the event in (5.6.6) occurs, then B_t hits the $cN^{-1} \log N$ neighborhood of γ before time $\frac{1}{2}t_0$, say at the point x . From that moment, in the next $\frac{1}{4}t_0$ time span, B_t either stays within a ball containing x (to be defined below) or exits this ball without touching γ . Hence, using the strong Markov property, (5.6.6) is bounded above by

$$\sup_{\substack{x \in \mathbb{C}, y \in \gamma \\ |x-y| \leq cN^{-1} \log N}} \mathbb{P}(\tau_y^x > \frac{1}{4}t_0) + \mathbb{P}(B^x[0, \tau_y^x] \cap \gamma = \emptyset), \tag{5.6.7}$$

where B^x is a planar Brownian motion starting at x and τ_y^x is the exit time of B^x from the ball $B(y; \frac{1}{4}N^{-\delta'/2})$.

To bound the second term in (5.6.7), recall that $\text{diam}\gamma \geq N^{-\delta'/2}$, so γ intersects both the center and the boundary of the ball $B(y; \frac{1}{4}N^{-\delta'/2})$. Hence we can apply the Beurling estimate (see e.g. Theorem 3.76 of [29]) to obtain the following upper bound for the second term in (5.6.7),

$$c_1(4cN^{\delta'/2}N^{-1}\log N)^{1/2}, \quad (5.6.8)$$

for some constant $c_1 > 1$ which in particular does not depend on the curve γ . The above reasoning to obtain the bound (5.6.8) holds if $cN^{-1}\log N < \frac{1}{4}N^{-\delta'/2}$ and hence for large enough N . If N is small then the bound (5.6.8) is larger than 1 and holds trivially. To bound the first term in (5.6.7) we use Lemma 5.6.4,

$$\begin{aligned} \mathbb{P}(\tau_y^x > \tfrac{1}{4}t_0) &\leq \mathbb{P}(\tau_y^x > \tfrac{1}{4}N^{-\delta}) \leq \mathbb{P}(\text{diam}B[0, \tfrac{1}{4}N^{-\delta}] \leq \tfrac{1}{2}N^{-\delta'/2}) \\ &\leq c_2 \exp(-c_3N^{\delta'-\delta}), \end{aligned}$$

for some constants $c_2, c_3 > 0$.

We have that

$$\begin{aligned} \mathbb{P}(|B_{t_0}| \leq \varepsilon \mid 0 < d_{\mathcal{E}}(B[0, \tfrac{1}{2}t_0], \gamma) \leq cN^{-1}\log N, B[0, \tfrac{3}{4}t_0] \cap \gamma = \emptyset) \\ \leq \sup_{x \in \mathbb{C}} \mathbb{P}(|B_{\tfrac{1}{4}t_0}^x| \leq \varepsilon) = \mathbb{P}(|B_{\tfrac{1}{4}t_0}| \leq \varepsilon) \leq \frac{8}{\pi} \varepsilon^2 t_0^{-1}. \end{aligned} \quad (5.6.9)$$

The first inequality in (5.6.9) follows from the Markov property of Brownian motion. The equality in (5.6.9) follows from the fact that $B_{\frac{1}{4}t_0}^x$ is a two-dimensional Gaussian random vector centered at x . By combining (5.6.5), the bound on (5.6.6), and (5.6.9), we conclude that

$$\begin{aligned} \mathbb{P}(0 < d_{\mathcal{E}}(B^{\text{loop}, t_0}, \gamma) \leq cN^{-1}\log N) \\ \leq \frac{32}{\pi} [c_1(4cN^{\delta'/2}N^{-1}\log N)^{1/2} + c_2 \exp(-c_3N^{\delta'-\delta})]. \quad \square \end{aligned}$$

Proof of Proposition 5.6.1. Let $2 - \theta =: \delta < \delta' < 2$ and let X_N be the number of loops in \mathcal{L} of time length at least $N^{-\delta}$. First, we give an upper bound on X_N . Note that X_N is stochastically less than the number of loops γ in a Brownian loop soup in the full plane \mathbb{C} with $t_\gamma \geq N^{-\delta}$ and $\gamma(0) \in D$. The latter random variable has the Poisson distribution with mean

$$\lambda \int_D \int_{N^{-\delta}}^{\infty} \frac{1}{2\pi t_0^2} dt_0 dA(z) = \lambda A(D) \frac{1}{2\pi} N^\delta,$$

where A denotes two-dimensional Lebesgue measure. By Chebyshev's inequality, $X_N \leq N^\delta \log N$ with probability tending to 1 as $N \rightarrow \infty$.

Second, we bound the probability that \mathcal{L} contains loops of large time length with small diameter. By Lemma 5.6.4,

$$\begin{aligned} \mathbb{P}(\exists \gamma \in \mathcal{L}, t_\gamma \geq N^{-\delta}, \text{diam}\gamma < N^{-\delta'/2}) \\ \leq N^\delta \log N c_1 \exp(-c_2N^{\delta'-\delta}) + \mathbb{P}(X_N > N^\delta \log N), \end{aligned} \quad (5.6.10)$$

for some constants $c_1, c_2 > 0$. The expression (5.6.10) converges to 0 as $N \rightarrow \infty$.

Third, we prove the proposition. To this end, we discover the loops in \mathcal{L} one by one in decreasing order of their time length, similarly to the construction in Section 4.3 of [42]. This exploration can be done in the following way. Let $\mathcal{L}_1, \mathcal{L}_2, \dots$ be a sequence of independent Brownian loop soups with intensity λ in D . From \mathcal{L}_1 take the loop γ_1 with the largest time length. From \mathcal{L}_2 take the loop γ_2 with the largest time length smaller than t_{γ_1} . Iterating this procedure yields a random collection of loops $\{\gamma_1, \gamma_2, \dots\}$, which is such that $t_{\gamma_1} > t_{\gamma_2} > \dots$ a.s. By properties of Poisson point processes, $\{\gamma_1, \gamma_2, \dots\}$ is a Brownian loop soup with intensity λ in D .

Given a set of discovered loops $\gamma_1, \dots, \gamma_{k-1}$, we bound the probability that the next loop γ_k comes close to γ_i but does not intersect γ_i , for each $i \in \{1, \dots, k-1\}$ separately. Note that, because of the conditioning, we can treat γ_i as a deterministic loop, while γ_k is random. Therefore, to obtain such a bound, we can use Lemma 5.6.5 on the event that $t_{\gamma_k} \geq N^{-\delta}$ and $\text{diam}\gamma_i \geq N^{-\delta'/2}$. We use the first and second steps of this proof to bound the probability that \mathcal{L} contains more than $N^\delta \log N$ loops of large time length, or loops of large time length with small diameter. Thus,

$$\begin{aligned} & \mathbb{P}(\exists \gamma, \gamma' \in \mathcal{L}, t_\gamma, t_{\gamma'} \geq N^{-\delta}, 0 < d_{\mathcal{E}}(\gamma, \gamma') \leq cN^{-1} \log N) \\ & \leq (N^\delta \log N)^2 [c_3 N^{-1/2+\delta'/4} (c \log N)^{1/2} + c_3 \exp(-c_4 N^{\delta'-\delta})] + \\ & \quad \mathbb{P}(\{X_N > N^\delta \log N\} \cup \{\exists \gamma \in \mathcal{L}, t_\gamma \geq N^{-\delta}, \text{diam}\gamma < N^{-\delta'/2}\}), \end{aligned} \quad (5.6.11)$$

for some constants $c_3, c_4 > 0$. If $\delta' < 2/9$, then (5.6.11) converges to 0 as $N \rightarrow \infty$. \square

5.7 Proof of main result

Proof of Theorem 5.2.1. By Corollary 5.4 of [31], for every N we can define on the same probability space $\tilde{\mathcal{L}}_N$ and \mathcal{L} such that the following holds with probability tending to 1 as $N \rightarrow \infty$: There is a one-to-one correspondence between the loops in $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ and the loops in $\mathcal{L}^{N^{\theta-2}}$ such that, if $\tilde{\gamma} \in \tilde{\mathcal{L}}_N^{N^{\theta-2}}$ and $\gamma \in \mathcal{L}^{N^{\theta-2}}$ are paired in this correspondence, then $d_\infty(\tilde{\gamma}, \gamma) < cN^{-1} \log N$, where c is a constant which does not depend on N .

We prove that in the above coupling, for all $\delta, \alpha > 0$ there exists N_0 such that for all $N \geq N_0$ the following holds with probability at least $1 - \alpha$: For every outermost cluster C of \mathcal{L} there exists an outermost cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ such that

$$d_H(C, \tilde{C}_N) < \delta, \quad d_H(\text{Ext}C, \text{Ext}\tilde{C}_N) < \delta, \quad (5.7.1)$$

and for every outermost cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ there exists an outermost cluster C of \mathcal{L} such that (5.7.1) holds. By Lemma 5.3.8, (5.7.1) implies that $d_H(\partial_o C, \partial_o \tilde{C}_N) < 2\delta$ and $d_H(\text{Hull}C, \text{Hull}\tilde{C}_N) < 2\delta$. Also, (5.7.1) implies that the Hausdorff distance between the carpet of \mathcal{L} and the carpet of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ is less than or equal to δ . Hence this proves the theorem.

Fix $\delta, \alpha > 0$. To simplify the presentation of the proof of (5.7.1), we will often use the phrase “with high probability”, by which we mean with probability larger than a certain lower bound which is uniform in N . It is not difficult to check that we can choose these lower bounds in such a way that (5.7.1) holds with probability at least $1 - \alpha$.

First we define some constants. By Lemma 9.7 of [48], a.s. there are only finitely many clusters of \mathcal{L} with diameter larger than any positive threshold; moreover they are all at positive distance from each other. Let $\rho \in (0, \delta/2)$ be such that, with high probability, for every $z \in D$ we have that $z \in \text{Hull}C$ for some outermost cluster C of \mathcal{L} with $\text{diam}C \geq \delta/2$, or $d_{\mathcal{E}}(z, C) < \delta/4$ for some outermost cluster C of \mathcal{L} with $\rho < \text{diam}C < \delta/2$. The existence of such a ρ follows from the fact that a.s. \mathcal{L} is dense in D and that there are only finitely many clusters of \mathcal{L} with diameter at least $\delta/2$. We call a cluster or subcluster *large* (*small*) if its diameter is larger than (less than or equal to) ρ .

Let $\varepsilon_1 > 0$ be such that, with high probability,

$$|\text{diam}C - \rho| > \varepsilon_1$$

for all clusters C of \mathcal{L} . Let $\varepsilon_2 > 0$ be such that, with high probability,

$$d_{\mathcal{E}}(C_1, C_2) > \varepsilon_2$$

for all distinct large clusters C_1, C_2 of \mathcal{L} . For every large cluster C_1 of \mathcal{L} , let $\wp(C_1)$ be a path connecting $\text{Hull}C_1$ with ∞ such that, for all large clusters C_2 of \mathcal{L} such that $\text{Hull}C_1 \not\subset \text{Hull}C_2$, we have that $\wp(C_1) \cap \text{Hull}C_2 = \emptyset$. Let $\varepsilon_3 > 0$ be such that, with high probability,

$$d_{\mathcal{E}}(\wp(C_1), \text{Hull}C_2) > \varepsilon_3$$

for all large clusters C_1, C_2 of \mathcal{L} such that $\text{Hull}C_1 \not\subset \text{Hull}C_2$. By Lemma 5.3.7 (and Remark 5.3.9) we can choose $\varepsilon_4 > 0$ such that, with high probability, for every large cluster C of \mathcal{L} ,

$$\text{if } d_H(C, \tilde{C}) < \varepsilon_4, \text{ then } \text{Ext}C \subset (\text{Ext}\tilde{C})^{\min\{\delta, \varepsilon_2\}/8}$$

for any collection of loops \tilde{C} .

Let $t_0 > 0$ be such that, with high probability, every subcluster C of \mathcal{L} with $\text{diam}C > \rho - \varepsilon_1$ contains a loop of time length larger than t_0 . Such a t_0 exists by Lemma 5.4.5. In particular, every large subcluster of \mathcal{L} contains a loop of time length larger than t_0 . Note that the number of loops with time length larger than t_0 is a.s. finite.

From now on the proof is in six steps, and we start by giving a sketch of these steps (see Figure 5.6). First, we treat the large clusters. For every large cluster C of \mathcal{L} , we choose a finite subcluster C' of C such that $d_H(C, C')$ and $d_H(\text{Ext}C, \text{Ext}C')$ are small, using Theorem 5.4.1. Second, we approximate C' by a subcluster \tilde{C}'_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ such that $d_H(\text{Ext}C', \text{Ext}\tilde{C}'_N)$ is small, using the one-to-one correspondence between random walk loops and Brownian loops, Theorem 5.3.5 and Corollary 5.5.2. Third, we let \tilde{C}_N be the cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ that contains \tilde{C}'_N . Here we make sure, using Proposition 5.6.1, that for distinct subclusters $\tilde{C}'_{1,N}, \tilde{C}'_{2,N}$, the corresponding clusters $\tilde{C}_{1,N}, \tilde{C}_{2,N}$ are distinct. It follows that $d_H(C, \tilde{C}_N)$ and $d_H(\text{Ext}C, \text{Ext}\tilde{C}_N)$ are small. Fourth, we show that the obtained clusters \tilde{C}_N are large. We also show that we obtain in fact all large clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ in this way. Fifth, we prove that a large cluster C of \mathcal{L} is outermost if and only if the corresponding large cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ is outermost. Sixth, we deal with the small outermost clusters.

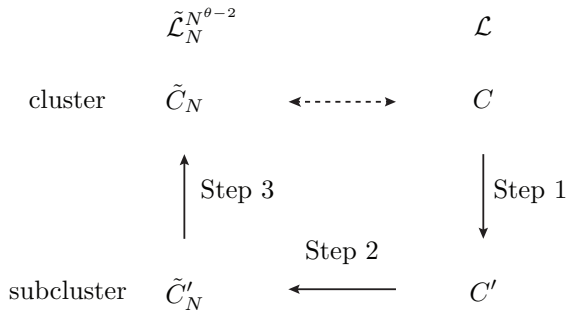


Figure 5.6: Schematic diagram of the proof of Theorem 5.2.1. We start with a cluster C of \mathcal{L} and, following the arrows, we construct a cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$. The dashed arrow indicates that C and \tilde{C}_N satisfy (5.7.1).

Step 1. Let \mathcal{C} be the collection of large clusters of \mathcal{L} . By Lemma 9.7 of [48], the collection \mathcal{C} is finite a.s. For every $C \in \mathcal{C}$ let C' be a finite subcluster of C such that C' contains all loops in C which have time length larger than t_0 and

$$d_H(C, C') < \min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/16, \tag{5.7.2}$$

$$d_H(\text{Ext}C, \text{Ext}C') < \min\{\delta, \varepsilon_2\}/16, \tag{5.7.3}$$

a.s. This is possible by Theorem 5.4.1. Let \mathcal{C}' be the collection of these finite subclusters C' .

Step 2. For every $C' \in \mathcal{C}'$ let $\tilde{C}'_N \subset \tilde{\mathcal{L}}_N^{N^{\theta-2}}$ be the set of random walk loops which correspond to the Brownian loops in C' , in the one-to-one correspondence from the first paragraph of this proof. This is possible for large N , with high probability, since then

$$\bigcup C' \subset \mathcal{L}^{N^{\theta-2}},$$

where $\bigcup C' = \bigcup_{C' \in \mathcal{C}'} C'$. Let $\tilde{\mathcal{C}}'_N$ be the collection of these sets of random walk loops \tilde{C}'_N .

Now we prove some properties of the elements of $\tilde{\mathcal{C}}'_N$. By Corollary 5.5.2, C' has no touchings a.s. Hence, by Theorem 5.3.5 (and Remark 5.3.9), for large N , with high probability,

$$d_H(\text{Ext}C', \text{Ext}\tilde{C}'_N) < \min\{\delta, \varepsilon_2\}/16. \tag{5.7.4}$$

Next note that almost surely, $d_{\mathcal{E}}(\gamma, \gamma') > 0$ for all non-intersecting loops $\gamma, \gamma' \in \mathcal{L}$, and $\text{overlap}(\gamma, \gamma') > 0$ for all intersecting loops $\gamma, \gamma' \in \mathcal{L}$. Since the number of loops in $\bigcup C'$ is finite, we can choose $\eta > 0$ such that, with high probability, $d_{\mathcal{E}}(\gamma, \gamma') > \eta$ for all non-intersecting loops $\gamma, \gamma' \in \bigcup C'$, and $\text{overlap}(\gamma, \gamma') > \eta$ for all intersecting loops $\gamma, \gamma' \in \bigcup C'$. For large N , $cN^{-1} \log N < \eta/2$ and hence with high probability,

$$\gamma_1 \cap \gamma_2 = \emptyset \text{ if and only if } \tilde{\gamma}_1 \cap \tilde{\gamma}_2 = \emptyset, \text{ for all } \gamma_1, \gamma_2 \in \bigcup C', \tag{5.7.5}$$

where $\tilde{\gamma}_1, \tilde{\gamma}_2$ are the random walk loops which correspond to the Brownian loops γ_1, γ_2 , respectively. By (5.7.5), every $\tilde{C}'_N \in \tilde{\mathcal{C}}'_N$ is connected and hence a subcluster of

$\tilde{\mathcal{L}}_N^{N^{\theta-2}}$. Also by (5.7.5), for distinct $C'_1, C'_2 \in \mathcal{C}'$, the corresponding $\tilde{C}'_{1,N}, \tilde{C}'_{2,N} \in \tilde{\mathcal{C}}'_N$ do not intersect each other when viewed as subsets of the plane.

Step 3. For every $\tilde{C}'_N \in \tilde{\mathcal{C}}'_N$ let \tilde{C}_N be the cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ which contains \tilde{C}'_N . Let $\tilde{\mathcal{C}}_N$ be the collection of these clusters \tilde{C}_N . We claim that for distinct $\tilde{C}'_{1,N}, \tilde{C}'_{2,N} \in \tilde{\mathcal{C}}'_N$, the corresponding $\tilde{C}_{1,N}, \tilde{C}_{2,N} \in \tilde{\mathcal{C}}_N$ are distinct, for large N , with high probability. This implies that there is one-to-one correspondence between elements of $\tilde{\mathcal{C}}'_N$ and elements of $\tilde{\mathcal{C}}_N$, and hence between elements of $\mathcal{C}, \mathcal{C}', \tilde{\mathcal{C}}'_N$ and $\tilde{\mathcal{C}}_N$.

To prove the claim, we combine Proposition 5.6.1 and the one-to-one correspondence between random walk loops and Brownian loops to obtain that, for large N , with high probability,

$$\text{if } \gamma_1 \cap \gamma_2 = \emptyset \text{ then } \tilde{\gamma}_1 \cap \tilde{\gamma}_2 = \emptyset, \text{ for all } \gamma_1, \gamma_2 \in \mathcal{L}^{N^{\theta-2}}, \quad (5.7.6)$$

where $\tilde{\gamma}_1, \tilde{\gamma}_2$ are the random walk loops which correspond to the Brownian loops γ_1, γ_2 , respectively. Let $\tilde{C}'_{1,N}, \tilde{C}'_{2,N} \in \tilde{\mathcal{C}}'_N$ be distinct. Let $C'_1, C'_2 \in \mathcal{C}'$ be the finite subclusters of Brownian loops which correspond to $\tilde{C}'_{1,N}, \tilde{C}'_{2,N}$, respectively. By construction, C'_1, C'_2 are contained in clusters of $\mathcal{L}^{N^{\theta-2}}$ which are distinct. Hence by (5.7.6), $\tilde{C}_{1,N}, \tilde{C}_{2,N}$ are distinct.

Next we prove that, for large N , with high probability,

$$d_H(C, \tilde{C}_N) < \min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/4, \quad (5.7.7)$$

$$d_H(\text{Ext}C, \text{Ext}\tilde{C}_N) < \min\{\delta, \varepsilon_2\}/4, \quad (5.7.8)$$

which implies that C and \tilde{C}_N satisfy (5.7.1). To prove (5.7.7), let N be sufficiently large, so that in particular $cN^{-1} \log N < \min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/16$. By (5.7.2), with high probability,

$$C \subset (C')^{\min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/16} \subset (\tilde{C}'_N)^{\min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/8} \subset (\tilde{C}_N)^{\min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/8}.$$

By (5.7.6), $\tilde{C}_N \subset C^{\min\{\delta, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}/16}$. This proves (5.7.7). To prove (5.7.8), note that by (5.7.7) and the definition of ε_4 , $\text{Ext}C \subset (\text{Ext}\tilde{C}_N)^{\min\{\delta, \varepsilon_2\}/8}$. By (5.7.3) and (5.7.4),

$$\text{Ext}\tilde{C}_N \subset \text{Ext}\tilde{C}'_N \subset (\text{Ext}C)^{\min\{\delta, \varepsilon_2\}/8}.$$

This proves (5.7.8).

Step 4. We prove that, for large N , with high probability, all $\tilde{C}_N \in \tilde{\mathcal{C}}_N$ are large, and that all large clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ are elements of $\tilde{\mathcal{C}}_N$. This gives that, for large N , with high probability, there is a one-to-one correspondence between large clusters C of \mathcal{L} and large clusters \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ such that (5.7.7) and (5.7.8) hold, and hence such that (5.7.1) holds.

First we show that, for large N , with high probability, all $\tilde{C}_N \in \tilde{\mathcal{C}}_N$ are large. By (5.7.7) and the definition of ε_1 , for large N , with high probability, $\text{diam}\tilde{C}_N > \text{diam}C - \varepsilon_1 > \rho$, i.e. \tilde{C}_N is large.

Next we prove that, for large N , with high probability, all large clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ are elements of $\tilde{\mathcal{C}}_N$. Let \tilde{G}_N be a large cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$. Let $G_N \subset \mathcal{L}^{N^{\theta-2}}$ be the set of Brownian loops which correspond to the random walk loops in \tilde{G}_N . By (5.7.6), G_N is connected and hence a subcluster of \mathcal{L} . If $cN^{-1} \log N < \varepsilon_1/2$, then $\text{diam}G_N > \rho - \varepsilon_1$.

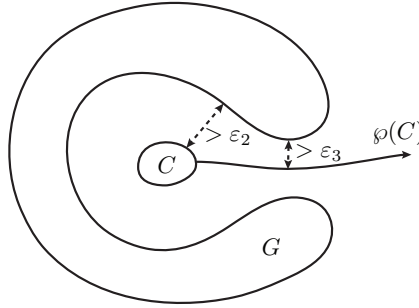


Figure 5.7: The case $\text{Hull}C \cap \text{Hull}G = \emptyset$ in Step 5.

Let G be the cluster of \mathcal{L} which contains G_N . We have that $\text{diam}G > \rho - \varepsilon_1$ and hence by the definition of ε_1 , with high probability, G is large, i.e. $G \in \mathcal{C}$. Let \tilde{G}_N^* be the element of $\tilde{\mathcal{C}}_N$ which corresponds to G . We claim that

$$\tilde{G}_N^* = \tilde{G}_N, \tag{5.7.9}$$

which implies that $\tilde{G}_N \in \tilde{\mathcal{C}}_N$.

To prove (5.7.9), let G' be the element of \mathcal{C}' which corresponds to G . Since G_N is a subcluster of \mathcal{L} with $\text{diam}G_N > \rho - \varepsilon_1$, G_N contains a loop γ of time length larger than t_0 . Since $\gamma \in G$ and $t_\gamma > t_0$, by the construction of G' , we have that $\gamma \in G'$. Hence $\tilde{\gamma} \in \tilde{G}_N^*$, where $\tilde{\gamma}$ is the random walk loop corresponding to the Brownian loop γ . Since $\gamma \in G_N$, by the definition of G_N , we have that $\tilde{\gamma} \in \tilde{G}_N$. It follows that $\tilde{\gamma} \in \tilde{G}_N \cap \tilde{G}_N^*$, which implies that (5.7.9) holds.

Step 5. Let C, G be distinct large clusters of \mathcal{L} , and let \tilde{C}_N, \tilde{G}_N be the large clusters of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ which correspond to C, G , respectively. We prove that, for large N , with high probability,

$$\text{Hull}C \subset \text{Hull}G \text{ if and only if } \text{Hull}\tilde{C}_N \subset \text{Hull}\tilde{G}_N. \tag{5.7.10}$$

It follows from (5.7.10) that a large cluster C of \mathcal{L} is outermost if and only if the corresponding large cluster \tilde{C}_N of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ is outermost.

To prove (5.7.10), suppose that $\text{Hull}C \subset \text{Hull}G$. By the definition of ε_2 , $(\text{Hull}C)^{\varepsilon_2/2} \subset \mathbb{C} \setminus (\text{Ext}G)^{\varepsilon_2/2}$. By (5.7.8), $\text{Ext}\tilde{G}_N \subset (\text{Ext}G)^{\varepsilon_2/2}$. By (5.7.7), $\tilde{C}_N \subset \mathbb{C}^{\varepsilon_2/2} \subset (\text{Hull}C)^{\varepsilon_2/2}$. Hence

$$\tilde{C}_N \subset (\text{Hull}C)^{\varepsilon_2/2} \subset \mathbb{C} \setminus (\text{Ext}G)^{\varepsilon_2/2} \subset \mathbb{C} \setminus \text{Ext}\tilde{G}_N = \text{Hull}\tilde{G}_N.$$

It follows that $\text{Hull}\tilde{C}_N \subset \text{Hull}\tilde{G}_N$.

To prove the reverse implication of (5.7.10), suppose that $\text{Hull}\tilde{C}_N \subset \text{Hull}\tilde{G}_N$. There are three cases: $\text{Hull}C \subset \text{Hull}G$, $\text{Hull}G \subset \text{Hull}C$ and $\text{Hull}C \cap \text{Hull}G = \emptyset$. We will show that the second and third case lead to a contradiction, which implies that $\text{Hull}C \subset \text{Hull}G$. For the second case, suppose that $\text{Hull}G \subset \text{Hull}C$. Then, by the previous paragraph, $\text{Hull}\tilde{G}_N \subset \text{Hull}\tilde{C}_N$. This contradicts the fact that $\text{Hull}\tilde{C}_N \subset \text{Hull}\tilde{G}_N$ and $\tilde{C}_N \cap \tilde{G}_N = \emptyset$.

For the third case, suppose that $\text{Hull}C \cap \text{Hull}G = \emptyset$. Let $\wp(C)$ be the path from the definition of ε_3 , which connects $\text{Hull}C$ with ∞ such that $\wp(C) \cap \text{Hull}G = \emptyset$ (see Figure 5.7). By the definition of $\varepsilon_2, \varepsilon_3$, with high probability,

$$((\text{Hull}C)^{\min\{\varepsilon_2, \varepsilon_3\}/2} \cup \wp(C)) \cap (\text{Hull}G)^{\min\{\varepsilon_2, \varepsilon_3\}/2} = \emptyset.$$

By (5.7.7), for large N , with high probability,

$$\tilde{C}_N \subset C^{\min\{\varepsilon_2, \varepsilon_3\}/2} \subset (\text{Hull}C)^{\min\{\varepsilon_2, \varepsilon_3\}/2}.$$

Similarly, $\tilde{G}_N \subset (\text{Hull}G)^{\min\{\varepsilon_2, \varepsilon_3\}/2}$. It follows that there exists a path from \tilde{C}_N to ∞ that avoids \tilde{G}_N . This contradicts the assumption that $\text{Hull}\tilde{C}_N \subset \text{Hull}\tilde{G}_N$.

Step 6. Finally we treat the small outermost clusters. Let G be a small outermost cluster of \mathcal{L} . By the definition of ρ , with high probability, there exists an outermost cluster C of \mathcal{L} with $\rho < \text{diam}C < \delta/2$ such that $d_{\mathcal{E}}(C, G) < \delta/4$. It follows that

$$\begin{aligned} d_H(C, G) &\leq d_{\mathcal{E}}(C, G) + \max\{\text{diam}C, \text{diam}G\} < 3\delta/4, \\ d_H(\text{Ext}C, \text{Ext}G) &\leq \frac{1}{2} \max\{\text{diam}C, \text{diam}G\} < \delta/4. \end{aligned}$$

Note that C is large, and let \tilde{C}_N be the large outermost cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$ which corresponds to C . Since C and \tilde{C}_N satisfy (5.7.7) and (5.7.8), we obtain that $d_H(G, \tilde{C}_N) < \delta$ and $d_H(\text{Ext}G, \text{Ext}\tilde{C}_N) < \delta/2$.

Next, by the one-to-one correspondence between elements of \mathcal{C} and $\tilde{\mathcal{C}}_N$ satisfying (5.7.7) and (5.7.8), for large N , with high probability,

$$d_H\left(\bigcap_{C \in \mathcal{C}} \text{Ext}C, \bigcap_{\tilde{C}_N \in \tilde{\mathcal{C}}_N} \text{Ext}\tilde{C}_N\right) < \delta/4. \quad (5.7.11)$$

Let \tilde{G}_N be a small outermost cluster of $\tilde{\mathcal{L}}_N^{N^{\theta-2}}$, then we have $\tilde{G}_N \subset \bigcap_{\tilde{C}_N \in \tilde{\mathcal{C}}_N} \text{Ext}\tilde{C}_N$. By (5.7.11) and the fact that \mathcal{L} is dense in D , a.s. there exists an outermost cluster C of \mathcal{L} with $\text{diam}C < \delta/2$ such that $d_{\mathcal{E}}(C, \tilde{G}_N) < \delta/2$. It follows that

$$\begin{aligned} d_H(C, \tilde{G}_N) &\leq d_{\mathcal{E}}(C, \tilde{G}_N) + \max\{\text{diam}C, \text{diam}\tilde{G}_N\} < \delta, \\ d_H(\text{Ext}C, \text{Ext}\tilde{G}_N) &\leq \frac{1}{2} \max\{\text{diam}C, \text{diam}\tilde{G}_N\} < \delta/4. \end{aligned}$$

This completes the proof. □

Summary

This thesis is on probability theory, in particular on percolation, loop soups and stochastic domination. It is based on the papers [8], [53], [7] and [52], which form the basis for Chapters 2–5, respectively. Chapter 1 contains an introduction.

In Chapter 2 we study *stochastic domination* of conditioned Bernoulli random vectors. We consider sequences of vectors \mathbf{X}_n and \mathbf{Y}_n that each consist of n independent Bernoulli random variables. We assume that \mathbf{X}_n and \mathbf{Y}_n each consist of M “blocks” such that the Bernoulli random variables in block i have success probability p_i and q_i , respectively, with $p_i \leq q_i$ for all i . Here M does not depend on n and the size of each block is essentially linear in n . We consider the conditional laws of \mathbf{X}_n and \mathbf{Y}_n , conditioned on the total number of successes being at least k_n , where k_n is also essentially linear in n . In general, the conditional law of \mathbf{X}_n is not necessarily stochastically dominated by the conditional law of \mathbf{Y}_n . We give a complete answer to the question with what maximal probability two such conditioned Bernoulli random vectors can be ordered in any coupling, when the length n of the vectors tends to infinity.

In Chapter 3 we study the *random connection model*, which is a model in *continuum percolation* (see [39]) defined as follows. Take a Poisson point process X on \mathbb{R}^d of density λ and connect each pair of points x and y in X with probability $g(|x - y|)$, independently of other pairs of points, independently of the point process X . Here g is a connection function, which is a non-increasing function from the positive reals to $[0, 1]$. We consider a sequence of random connection models X_n , where X_n is a Poisson point process on \mathbb{R}^d of density λ_n such that $\lambda_n/n^d \rightarrow \lambda > 0$. The points of X_n are connected according to the connection function g_n defined by $g_n(x) = g(nx)$, for some connection function g . Let I_n be the number of isolated points in the random connection model X_n in some bounded set K . The main result in [44] by Roy and Sarkar is a central limit theorem for I_n . Although the statement of this result is correct, the proof in [44] has errors. We explain what went wrong in the proof, and how this can be corrected. We also prove an extension to components larger than a single point in case the connection function has bounded support.

In Chapter 4 we study two variations on the fractal percolation model introduced by Mandelbrot [38]. The first variation is *k-fractal percolation*, defined as follows. Divide the d -dimensional unit cube in N^d equal subcubes and retain k of them in a uniform way while the others are removed. Then iterate the procedure inside the retained subcubes at all smaller scales. We prove that the (properly rescaled) percolation critical value of the model converges to the critical value of ordinary site percolation on a particular d -dimensional lattice as N tends to infinity. This is analo-

gous to the result of Falconer and Grimmett [24] that the critical value of Mandelbrot fractal percolation converges to the critical value of site percolation on the same d -dimensional lattice. The second model we study is *fat fractal percolation*. In this model subcubes are retained with probability p_n at iteration step n of the construction, where p_n is non-decreasing in n such that $\prod_n p_n > 0$. The Lebesgue measure of the limit set is positive a.s. given non-extinction. We prove that either the set of connected components larger than one point has Lebesgue measure zero a.s. or its complement in the limit set has Lebesgue measure zero a.s.

In Chapter 5 we study the *random walk loop soup*, which is a Poissonian collection of lattice loops. It has been extensively studied because of its connections to the discrete Gaussian free field [33], but was originally introduced by Lawler and Trujillo Ferreras [31] as a discrete version of the *Brownian loop soup* of Lawler and Werner [32], a conformally invariant Poissonian collection of planar loops with deep connections to *conformal loop ensembles* (CLE) [48] and the Schramm-Loewner evolution (SLE). Lawler and Trujillo Ferreras [31] showed that, roughly speaking, in the continuum scaling limit, “large” lattice loops from the random walk loop soup converge to “large” loops from the Brownian loop soup. Their results, however, do not extend to clusters of loops, which are interesting because the connection between Brownian loop soup and CLE goes via cluster boundaries. We study the scaling limit of clusters of “large” lattice loops, showing that they converge to Brownian loop soup clusters. In particular, our results imply that the collection of outer boundaries of outermost clusters composed of “large” lattice loops converges to CLE.

Bibliography

- [1] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749 (2000e:60008)
- [2] E. Bolthausen, *On the central limit theorem for stationary mixing random fields*, Ann. Probab. **10** (1982), no. 4, 1047–1050. MR 672305 (84c:60035)
- [3] Erik Broman, Olle Häggström, and Jeffrey Steif, *Refinements of stochastic domination*, Probab. Theory Related Fields **136** (2006), no. 4, 587–603, doi:10.1007/s00440-006-0496-1. MR 2257137 (2008k:60040)
- [4] Erik Broman and Ronald Meester, *Survival of inhomogeneous Galton-Watson processes*, Adv. in Appl. Probab. **40** (2008), no. 3, 798–814, doi:10.1239/aap/1222868186. MR 2454033 (2009k:60220)
- [5] Erik I. Broman and Federico Camia, *Large- N limit of crossing probabilities, discontinuity, and asymptotic behavior of threshold values in Mandelbrot’s fractal percolation process*, Electron. J. Probab. **13** (2008), no. 33, 980–999. MR 2413292 (2009g:60130)
- [6] ———, *Universal behavior of connectivity properties in fractal percolation models*, Electron. J. Probab. **15** (2010), 1394–1414. MR 2721051 (2011m:60028)
- [7] Erik I. Broman, Tim van de Brug, Federico Camia, Matthijs Joosten, and Ronald Meester, *Fat fractal percolation and k -fractal percolation*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 279–301. MR 3069366
- [8] Erik I. Broman, Tim van de Brug, Wouter Kager, and Ronald Meester, *Stochastic domination and weak convergence of conditioned Bernoulli random vectors*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 403–434. MR 3069371
- [9] David Brydges, Jürg Fröhlich, and Thomas Spencer, *The random walk representation of classical spin systems and correlation inequalities*, Comm. Math. Phys. **83** (1982), no. 1, 123–150. MR 648362 (83i:82032)
- [10] Krzysztof Burdzy and Gregory F. Lawler, *Nonintersection exponents for Brownian paths. II. Estimates and applications to a random fractal*, Ann. Probab. **18** (1990), no. 3, 981–1009. MR 1062056 (91g:60097)

- [11] Krzysztof Burdzy and Wendelin Werner, *No triple point of planar Brownian motion is accessible*, Ann. Probab. **24** (1996), no. 1, 125–147. MR 1387629 (97j:60147)
- [12] Federico Camia, *Off-criticality and the massive Brownian loop soup*, arXiv:1309.6068 [math.PR] (2013).
- [13] Federico Camia and Charles M. Newman, *Two-dimensional critical percolation: the full scaling limit*, Comm. Math. Phys. **268** (2006), no. 1, 1–38. MR 2249794 (2007m:82032)
- [14] ———, *Critical percolation exploration path and SLE_6 : a proof of convergence*, Probab. Theory Related Fields **139** (2007), no. 3–4, 473–519. MR 2322705 (2008k:82040)
- [15] ———, *SLE_6 and CLE_6 from critical percolation*, Probability, geometry and integrable systems, Math. Sci. Res. Inst. Publ., vol. 55, Cambridge Univ. Press, Cambridge, 2008, pp. 103–130. MR 2407594 (2009h:82032)
- [16] J. T. Chayes, L. Chayes, and R. Durrett, *Connectivity properties of Mandelbrot’s percolation process*, Probab. Theory Related Fields **77** (1988), no. 3, 307–324. MR 931500 (89d:60193)
- [17] L. Chayes, Robin Pemantle, and Yuval Peres, *No directed fractal percolation in zero area*, J. Statist. Phys. **88** (1997), no. 5–6, 1353–1362. MR 1478072 (99d:82029)
- [18] Lincoln Chayes, *Aspects of the fractal percolation process*, Fractal geometry and stochastics (Finsterbergen, 1994), Progr. Probab., vol. 37, Birkhäuser, Basel, 1995, pp. 113–143. MR 1391973 (97g:60131)
- [19] Dmitry Chelkak, Hugo Duminil-Copin, Clément Hongler, Antti Kemppainen, and Stanislav Smirnov, *Convergence of Ising interfaces to Schramm’s SLE curves*, C. R. Math. Acad. Sci. Paris **352** (2014), no. 2, 157–161. MR 3151886
- [20] F. M. Dekking and R. W. J. Meester, *On the structure of Mandelbrot’s percolation process and other random Cantor sets*, J. Statist. Phys. **58** (1990), no. 5–6, 1109–1126. MR 1049059 (91c:60140)
- [21] Michel Dekking and Henk Don, *Correlated fractal percolation and the Palis conjecture*, J. Stat. Phys. **139** (2010), no. 2, 307–325. MR 2602964 (2011i:60175)
- [22] E. B. Dynkin, *Gaussian and non-Gaussian random fields associated with Markov processes*, J. Funct. Anal. **55** (1984), no. 3, 344–376. MR 734803 (86h:60085a)
- [23] ———, *Local times and quantum fields*, Seminar on stochastic processes, 1983 (Gainesville, Fla., 1983), Progr. Probab. Statist., vol. 7, Birkhäuser Boston, Boston, MA, 1984, pp. 69–83. MR 902412 (88i:60080)
- [24] K. J. Falconer and G. R. Grimmett, *On the geometry of random Cantor sets and fractal percolation*, J. Theoret. Probab. **5** (1992), no. 3, 465–485. MR 1176432 (94b:60115)

- [25] W. Feller, *On the Kolmogorov-Smirnov limit theorems for empirical distributions*, Ann. Math. Statistics **19** (1948), 177–189. MR 0025108 (9,599i)
- [26] E. N. Gilbert, *Random plane networks*, J. Soc. Indust. Appl. Math. **9** (1961), 533–543. MR 0132566 (24 #A2406)
- [27] Johan Jonasson and Olle Nerman, *On maximum entropy π ps-sampling with fixed sample size*, Report 13, Chalmers University of Technology and Gothenburg University, available at <http://www.math.chalmers.se/Math/Research/Preprints/1996/13.pdf>, 1996.
- [28] Wouter Kager and Bernard Nienhuis, *A guide to stochastic Löwner evolution and its applications*, J. Statist. Phys. **115** (2004), no. 5-6, 1149–1229. MR 2065722 (2005f:82037)
- [29] Gregory F. Lawler, *Conformally invariant processes in the plane*, Mathematical Surveys and Monographs, vol. 114, American Mathematical Society, Providence, RI, 2005. MR 2129588 (2006i:60003)
- [30] Gregory F. Lawler, Oded Schramm, and Wendelin Werner, *Values of Brownian intersection exponents. II. Plane exponents*, Acta Math. **187** (2001), no. 2, 275–308. MR 1879851 (2002m:60159b)
- [31] Gregory F. Lawler and José A. Trujillo Ferreras, *Random walk loop soup*, Trans. Amer. Math. Soc. **359** (2007), no. 2, 767–787 (electronic). MR 2255196 (2008k:60084)
- [32] Gregory F. Lawler and Wendelin Werner, *The Brownian loop soup*, Probab. Theory Related Fields **128** (2004), no. 4, 565–588. MR 2045953 (2005f:60176)
- [33] Yves Le Jan, *Markov loops and renormalization*, Ann. Probab. **38** (2010), no. 3, 1280–1319. MR 2675000 (2011e:60179)
- [34] ———, *Markov paths, loops and fields*, Lecture Notes in Mathematics, vol. 2026, Springer, Heidelberg, 2011, Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR 2815763 (2012i:60158)
- [35] Thomas M. Liggett and Jeffrey E. Steif, *Stochastic domination: the contact process, Ising models and FKG measures*, Ann. Inst. H. Poincaré Probab. Statist. **42** (2006), no. 2, 223–243. MR 2199800 (2007i:60131)
- [36] Titus Lupu, *Convergence of the two-dimensional random walk loop soup clusters to CLE*, arXiv:1502.06827 [math.PR] (2015).
- [37] Russell Lyons, *Random walks, capacity and percolation on trees*, Ann. Probab. **20** (1992), no. 4, 2043–2088. MR 1188053 (93k:60175)
- [38] Benoit B. Mandelbrot, *The fractal geometry of nature*, W. H. Freeman and Co., San Francisco, Calif., 1982, Schriftenreihe für den Referenten. [Series for the Referee]. MR 665254 (84h:00021)

- [39] Ronald Meester and Rahul Roy, *Continuum percolation*, Cambridge Tracts in Mathematics, vol. 119, Cambridge University Press, Cambridge, 1996. MR 1409145 (98d:60193)
- [40] M. V. Menshikov, S. Yu. Popov, and M. Vachkovskaia, *On the connectivity properties of the complementary set in fractal percolation models*, Probab. Theory Related Fields **119** (2001), no. 2, 176–186. MR 1818245 (2002d:60085)
- [41] Peter Mörters and Yuval Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010, With an appendix by Oded Schramm and Wendelin Werner. MR 2604525 (2011i:60152)
- [42] Şerban Nacu and Wendelin Werner, *Random soups, carpets and fractal dimensions*, J. Lond. Math. Soc. (2) **83** (2011), no. 3, 789–809. MR 2802511 (2012k:28006)
- [43] Herbert Robbins, *A remark on Stirling’s formula*, Amer. Math. Monthly **62** (1955), 26–29. MR 0069328 (16,1020e)
- [44] Rahul Roy and Anish Sarkar, *High density asymptotics of the poisson random connection model*, Phys. A **318** (2003), no. 1-2, 230–242.
- [45] Ludger Rüschendorf, *Random variables with maximum sums*, Adv. in Appl. Probab. **14** (1982), no. 3, 623–632, doi:10.2307/1426677. MR 665297 (83j:60021)
- [46] Oded Schramm, *Scaling limits of loop-erased random walks and uniform spanning trees*, Israel J. Math. **118** (2000), 221–288. MR 1776084 (2001m:60227)
- [47] Scott Sheffield, *Exploration trees and conformal loop ensembles*, Duke Math. J. **147** (2009), no. 1, 79–129. MR 2494457 (2010g:60184)
- [48] Scott Sheffield and Wendelin Werner, *Conformal loop ensembles: the Markovian characterization and the loop-soup construction*, Ann. of Math. (2) **176** (2012), no. 3, 1827–1917. MR 2979861
- [49] Stanislav Smirnov, *Critical percolation in the plane: conformal invariance, Cardy’s formula, scaling limits*, C. R. Acad. Sci. Paris Sér. I Math. **333** (2001), no. 3, 239–244. MR 1851632 (2002f:60193)
- [50] ———, *Conformal invariance in random cluster models. I. Holomorphic fermions in the Ising model*, Ann. of Math. (2) **172** (2010), no. 2, 1435–1467. MR 2680496 (2011m:60302)
- [51] Alain-Sol Sznitman, *Topics in occupation times and Gaussian free fields*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2012. MR 2932978
- [52] Tim van de Brug, Federico Camia, and Marcin Lis, *Random walk loop soups and conformal loop ensembles*, arXiv:1407.4295 [math.PR] (2014).

- [53] Tim van de Brug and Ronald Meester, *On central limit theorems in the random connection model*, Phys. A **332** (2004), no. 1-4, 263–278. MR 2048371 (2004m:82045)
- [54] Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982. MR 648108 (84e:28017)
- [55] Wendelin Werner, *SLEs as boundaries of clusters of Brownian loops*, C. R. Math. Acad. Sci. Paris **337** (2003), no. 7, 481–486. MR 2023758 (2005b:60221)