# Geodesics in First-Passage Percolation and Random Walks on Critical Percolation Clusters

Jack Thomas Hanson

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Advisor: Professor Michael Aizenman

June 2013

# Abstract

It has been observed that statistical mechanical systems with quenched disorder can exhibit dramatically different properties than their disorder-free counterparts. The introduction of disorder can change the spatial structure of a system's ground state or the typical behavior of the system's free energy (or other quantities of interest). In many systems, the changes to spatial and quantitative behavior are closely linked.

This dissertation explores these themes in two model disordered systems: first-passage percolation and critical percolation. In first-passage percolation, a model for paths in a random potential, the energy of the ground state path of length $n$ is expected to fluctuate like $n^{\chi}$, and the ground state path is expected to deviate from a straight line by a distance of $n^{\xi}$. There are longstanding conjectures about the values of $\chi$ and $\xi$, and the relation $\chi = 2\xi - 1$ has recently been rigorously established. The prediction that $\chi < 1/2$ (in contrast to a situation where each step of unit length contributed independently to the ground state path's energy) is related to the existence of many "approximate ground states," and such techniques have proved fruitful for rigorously bounding $\chi$.

In the work presented here, we consider the related issue of wandering of infinite energy-minimizing paths–that is, infinite paths whose energy cannot be reduced by altering any finite segment. One could ask whether such paths tend to have asymptotic direction or remain in sectors. This and related questions have been studied in first-passage percolation and related models; in the setting of first-passage percolation, there are few rigorous results except under strong assumptions. The work presented here develops a framework for studying these questions, and provides the first minimal-assumption results on directional concentration of infinite ground state paths.

The critical percolation systems studied in this dissertation are two related models for the two-dimensional infinite cluster "at the critical point": the incipient infinite cluster and the invasion percolation cluster. The disorder in these systems gives them a fundamentally different geometry than an ordinary lattice, affecting their transport properties. It was previously shown rigorously that a diffusing particle on the incipient infinite cluster moves strictly slower than a diffusion on the square lattice, on average over the disorder.

The work presented here removes the average over disorder to show a quenched result: for a particular typical realization of the incipient infinite cluster, the diffusion is slow. The result is extended to diffusions on the invasion cluster. The work shows the relationship between the geometrical and transport properties of these models by deriving an upper bound for the speed in terms of critical exponents of the percolation models.

# Acknowledgements

I would first like to thank my advisor, Michael Aizenman. Without three and a half years of discussions with him, as well as his advice and support, this dissertation would not exist. Thank you, Michael!

It has been a pleasure to collaborate with and get to know Tuca Auffinger, and I look forward to our discussions next year. I owe a debt of gratitude to Charles Newman for encouragement and opportunities, and for sharing his insight. I am grateful to N. P. Ong for his guidance during my experimental project. I would also like to thank Yakov Sinai for reading this text and offering his advice.

I cannot overstate how much I have valued my time with my fellow PSPS members, Michael Damron and Phil Sosoe. The innumerable presentations of papers and sessions spent sifting through ideas have been some of my favorite memories over the last five years. Your contributions to this work and to my life are immense and undeniable.

The friends I have made here are some of the best people I have known. Thanks to Andrew & Erin, David & Allison, Doug & Kristina, Dora & Jon, Nikhil, Scott & Xin, and Zach for making these years count. My biggest regret is that I could never capture in this space what you each mean to me. A special thanks also to Sjuvon, without whose expert guidance I would currently be marooned on Fjärdlång. To all the others: thanks for being a part of my life.

Thanks to Kathleen Sullivan, whose kind heart has made the past year immeasurably better.

Last, I would like to acknowledge my family–my brother and sister, father and grandmother, mother, and aunts and uncles.

My work has been supported in part by an NSF Graduate Research Fellowship and NSF grant PHY-1104596.

# Relation to Published Work

The results presented in Chapters 3 and 4 reflect work performed as part of a research collaboration. The results in these chapters will ultimately appear in the peer-reviewed literature. The relevant preprints are:

- [10] Antonio Auffinger, Michael Damron, and Jack Hanson. Limiting geodesics for first-passage percolation on subsets of $\mathbb{Z}^2$. Submitted.

- [34] Michael Damron and Jack Hanson. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. Submitted.

The Appendix A consists of the preprint

- [35] Michael Damron, Jack Hanson, and Philippe Sosoe. Subdiffusivity of random walk on the 2d invasion percolation cluster. Accepted for publication in *Stochastic Processes and their Applications.*

To John Hanson and Roberta Hanson.

# Contents

# List of Figures

# Chapter 1

# Introduction

The work contained in this dissertation addresses questions arising from the statistical mechanics of disordered systems. The models considered here describe the behavior of particles or polymers in a lattice system with local randomness. This randomness affects the favorability of different large-scale configurations of the system: for instance, the regions which are most often visited by a diffusing particle, or the wandering of a polymer with constrained endpoints. In general, the geometrical structure of a typical state of the particle or polymer is closely related to the macroscopic distribution of regions of energetically favorable disorder. As such, there is a close relationship between the particle's typical energy or time it takes to travel a certain distance on one hand, and its typical trajectory on the other.

The two models most specifically considered here, first-passage percolation and invasion percolation, have been widely studied in both the physics and mathematics literature. This dissertation addresses questions in the model primarily through the lens of rigorous mathematical physics, with motivations in the heuristic and numerical arguments of the physics literature. These models are both especially ripe for such study in that the non-rigorous predictions for the models have far outpaced the frontier of rigorous knowledge.

In the first part of this dissertation, we study questions related to time-minimizing paths in first-passage percolation. Chapter 2 consists of an introduction to the model and the problems we will be most concerned with. In Section 2.1, we discuss some motivations behind models like first-passage percolation, and heuristics for their behavior. In Section 2.2, we define the model and describe some global or first-order properties. In this section we also present a proof of the celebrated "shape theorem" for the model. In Section 2.3, some fundamental conjectures about smaller-scale fluctuations in the model are presented alongside known rigorous results. Section 2.4 sees the introduction of infinite geodesics, which are the focus of the new results presented later in this part of the dissertation. Finally, in 2.5, we present a treatment of some results due to Hoffman which were a major impetus for these new results.

In Chapters 3 and 4, we will present our new results on infinite geodesics in first-passage percolation. In Chapter 3, we construct limiting geodesic graph structures on $\mathbb{Z}^2$ as well as a class of fractional planes. In the case of $\mathbb{Z}^2$, we prove the existence of infinite time-minimizing paths which are directionally concentrated or confined in various senses. These results provide partial answers to conjectures by C. Newman on directional properties of infinite geodesics. In the case of fractional planes, we show that the graph of finite geodesics to a point $v$ on the boundary has a limit as $v$ moves to infinity. This is a partial solution to the problem of determining whether the finite geodesics to the points $(n, 0)$ on $\mathbb{Z}^2$ have a limit in $n$–in some sense, a question of the stability of long ground-state polymers under a change of endpoint.

In Chapter 4, we present results on coalescence of infinite geodesics on $\mathbb{Z}^2$ and the half-plane. In Section 4.3, we show that the family of unigeodesics constructed as limits of geodesics to $(n, 0)$ on the half-plane must coalesce. In the following section, we show that the directionally concentrated geodesics previously constructed on $\mathbb{Z}^2$ must coalesce. We also show that the time-minimizing paths we construct are one-sided (i.e., they have one topological end). This one-sidedness shows that a particular method for constructing nonconstant ground states in a disordered ferromagnetic model will fail. These results extend previous work by Newman and Licea on coalescence, and support the conjecture in the literature that two-dimensional disordered ferromagnets have a unique ground state pair.

In Chapter 5, we provide some background on two-dimensional Bernoulli bond percolation and random walks on percolation clusters. Standard two-dimensional bond percolation is a family of models indexed by a parameter $p$. Each edge of the two-dimensional square lattice is "open" with probability $p$ and closed otherwise; the vertices linked by paths of open edges are called open clusters. At the value $p = p_c = 1/2$, there is no infinite open cluster, but the expected size of an open cluster is infinite. Two possible definitions for the "infinite open cluster at $p_c$", the incipient infinite cluster and the invasion percolation cluster, are discussed. We describe work by Kesten showing that a random walk on the incipient infinite cluster moves slower than diffusively. We then state the results of Damron, Hanson, and Sosoe [35] which provide an alternate and quenched (i.e., not averaged over possible realizations of the incipient infinite cluster) version of this result, and which extend the result to the invasion percolation cluster.

The Appendix A consists of the preprint [35] with some chiefly cosmetic modifications.

# Part I

# First-Passage Percolation: Background and Results

# Chapter 2

# First Passage Percolation

The first part of the dissertation focuses on a physical model, first-passage percolation (FPP), for systems of paths in random potentials. Related models have been proposed to describe various different physical situations; one path by which the physics community arrived at this sort of model is sketched here. The particular model of first-passage percolation is introduced, and this is followed by an overview of the heuristic and numerical predictions for such systems and overview of some rigorous results. Major conjectures for ground states ("geodesics") of first-passage models are described, and their relationship to other aspects of the model is described. This chapter lays the groundwork for the following two, in which we present results which provide a partial resolution to these conjectures.

## 2.1 Random Path Models: Physical Motivations

Systems with quenched disorder provide a broad class of interesting problems in statistical physics. In models with quenched disorder, physical degrees of freedom (for instance, spins) experience some interaction which has an essentially random part. In the case of a magnetic interaction between spins, this randomness may represent, for instance, the effect of introduced impurities [89]–the positions of impurity atoms do not thermally equilibrate on laboratory time scales. When the system is coupled to a heat bath, the degrees of freedom are allowed to equilibrate but the random terms in their Hamiltonians remain fixed for all time.

The introduction of quenched disorder to a model is capable of producing a dramatic change in behavior. The usual phase diagram may change; in one class of examples, disorder is known to destroy a previously existing ordered phase [2]. In other systems, new phases may appear which differ quantitatively from the model's previous behavior. This half of the thesis will be concerned with a model, first-passage percolation, which belongs to a universality class of random path models in which disorder plays a fundamental role by introducing a "pinned" phase.

One historical route through which these random path models were first studied was in the above-mentioned context of Ising-type ferromagnetic systems with impurities which serve to randomize the strength of local spin-spin couplings [53, 54]. Consider a system of spins in the $d$-dimensional square lattice $\mathbb{Z}^d$. If $\sigma(x) \in \{-1, +1\}$ denotes the value of the spin at site $x$, then we can write the Hamiltonian of the system as

$$H = -J \sum_{\langle xy \rangle} \sigma(x)\sigma(y) + \sum_{\langle xy \rangle} J_{x,y}\sigma(x)\sigma(y). \tag{2.1}$$

Here $J > 0$ represents the usual ferromagnetic exchange coupling of a pure Ising system; the impurities $J_{x,y}$ are random with independent and identical distribution (i.i.d.) and of small magnitude ( $|J_{x,y}/J| \ll 1$ for a typical realization of $J_{x,y}$).

The object of interest is not a complete description of the system described by (2.1) but the behavior of a typical domain wall between regions of spins with opposite sign. Assume that there is a domain wall $\Gamma = (e_i^*)$ of edges dual to those in $\mathbb{Z}^d$ which traverses the system parallel to a coordinate axis, with fixed (dual) endpoints

$(0, 0, \dots, 0)$ and $(L, 0, \dots, 0)$. The energy of $\Gamma$ is given by

$$H(\Gamma) = \sum_{e^* = \{x,y\}^*} (J - J_{x,y}), \tag{2.2}$$

a sum of (typically positive) i.i.d. terms.

Systems with Hamiltonians like (2.2) have been extensively studied in the physics literature, both numerically and heuristically (for an extensive review, see [71]). Consider a random path system with random Hamiltonian 2.2 at inverse temperature $\beta$; we will denote by $\langle \cdot \rangle$ the thermal average, and by $\mathbb{E}$ the average over disorder. Lastly, let $\Gamma$ be a path with endpoints at distance $L$ as above as above and let $D(0, L)$ denote the maximal displacement of $\Gamma$ from the first coordinate axis.

For $\beta$ large, there is predicted [29, 39, 58] to be a so-called "pinned" phase, in which disorder plays a fundamental role and the behavior of paths is not necessarily diffusive. That is, the expectation is that

$$(\mathbb{E} \left\langle D(0, L)^2 \right\rangle)^{1/2} \sim L^{\xi},$$

where $\xi$ is not in general equal to $1/2$. This behavior has been rigorously established for certain models with Hamiltonians of the form (2.2) (among others, in [86]; see also [24] for another "non-diffusivity" result). There is a wide range of further conjectures for the scaling of the expected energy $\mathbb{E}\langle H(\Gamma)\rangle \sim L^{\chi}$ in the pinned phase, including a particular relationship between $\chi$ and $\xi$. These will be discussed further later in this chapter in the context of first-passage percolation, the model for which the new results of this dissertation are derived.

It is worth noting here the expected behavior of this class of random path systems in the limit $\beta \approx 0$. In low dimensions, the pinned phase is believed to exist for all values of $\beta$ [39, 53, 54]. In large dimensions, by contrast, it is expected that there is a low-$\beta$ regime in which random paths are diffusive:

$$|D(0, L)| \sim L^{1/2}. \tag{2.3}$$

In fact, this has been rigorously established for a particular model in the case that the dimension $d > 3$ by Imbrie and Spencer [55]; see also a simplification and extension by Bolthausen [21], and an extension to a wider class of disorder distributions by Sinai [88].

The model considered in [55] makes an additional assumption of directedness of the random path system. Consider a free-boundary problem whose state space is the set of all paths $\Gamma : \{0, 1, \ldots, T\} \to \mathbb{Z}^{d-1}$ such that $\|\Gamma(t) - \Gamma(t-1)\|_1 = 1$. The Hamiltonian of the system is given by

$$H(\Gamma) = \sum_{t=0}^{T} h(t, \Gamma(t)),$$

where $\{h(t, x)\}$ is a family of independent random variables such that $h(t, x) = \pm 1$ with equal probability. A version of diffusivity proved in [55] is as follows: for any $d > 3$, there is a $\beta_0 > 0$ such that for all $\beta < \beta_0$,

$$\left|\mathbb{E}\left[\langle \|\Gamma(T)\|_2^2 \rangle\right] - T\right| \leq CT^{1-\theta}$$

for all $T > 0$, where $C$ and $\theta$ are positive (dimension-dependent) constants.

In what follows, we will consider the particular case of first-passage percolation, a model in which $\beta = \infty$–that is, the questions of first-passage percolation address the properties of ground state or energy-minimizing paths. In fact, the model has its origins in a different setting than the one discussed above: namely, the modeling of fluid flow in a random medium. As such, while the framework is very much the same, the terminology used will occasionally be different from that above: energies will be replaced by passage times, and so on. In the next section, we will give a full definition of the first-passage model and adopt the usual vocabulary of the first-passage literature.

## 2.2 Model and Global Properties

The results presented in this dissertation are proved for the particular random path model of first-passage percolation. In the following, we specialize to consideration of the conjectured behavior and established results for this model. First-passage percolation was introduced by Hammersley and Welsh [49, 50], originally as a model for fluid flow in a porous medium. In the metaphor of their original papers, paths are defined on a graph whose edges have random, spatially decorrelated “passage times” $\omega_e$, which are generally nonnegative. The sum of edge passage times along a lattice path is defined to be the path’s passage time, and two major classes of question arise:

1. What can be said about the geometry of various classes of time-minimizing paths?

2. What can be said about the passage time of the minimizing path?

The passage time functional on paths plays the role of an energy. In this sense, asking questions about various time-minimizing paths in first-passage percolation is equivalent to asking questions about ground states.

First-passage percolation can be defined and studied on a wide variety of graphs. To give more precise definitions, *we restrict ourselves to the case of first-passage percolation on the square lattice* $\mathbb{Z}^d$ and subsets thereof. Later, the discussion will be restricted further to the case that $d = 2$, for which the most detailed conjectures and results on the model are available, and for which our results are derived.

### 2.2.1 Definition of the Model

Consider the square lattice $(\mathbb{Z}^d, \mathcal{E}^d)$, where $\mathcal{E}^d$ represents the set of nearest-neighbor edges–that is, an edge $e \in \mathcal{E}^d$ is of the form $e = \{x, y\}$, where $\|x - y\|_1 = 1$. The graphs for which the model will be defined will be of the form $(V, E)$, where $V$ is a connected subset of $\mathbb{Z}^d$, and

$$E = \left\{ \{x, y\} \in \mathcal{E}^d : \, x, y \in V \right\}.$$

We will chiefly be concerned with the case $V = \mathbb{Z}^d$, we will generally refer to $(\mathbb{Z}^d, \mathcal{E}^d)$ rather than $(V, E)$ for the sake of concreteness. We will have use of general subgraphs $(V, E)$ chiefly in the presentation of our results on "fractional planes" in later chapters.

Associated with the model will be a probability space $(\Omega, \mathcal{B}, \mathbb{P})$, on which are defined the nonnegative random variables $\{\omega_e\}_{e \in E}$. To be specific, we will heretofore always make the canonical choice $\Omega = [0, \infty)^E$, with $\mathcal{B}$ the Borel sigma-algebra[1]. An element $\omega \in \Omega$ of this canonical space will then be a vector $(\omega_e)_{e \in E}$ whose coordinates are the random variables $\{\omega_e\}$. A given $\omega_e$ will be called an "edge weight", "edge passage time", or simply "passage time"; the joint distribution $\mathbb{P}$ of the passage times will be termed a "passage time distribution."

We will assume always that $\mathbb{P}$ is ergodic under translations of $\mathbb{Z}^d$ and has all of the symmetries of $\mathbb{Z}^d$. This ensures that the quantities of interest in the model have distributions which are invariant when the lattice and edge weights are jointly subjected to a transformation which leaves the lattice unchanged; for instance, the reflection $R$, acting by

$$R : \mathbb{Z}^d \times \Omega \longrightarrow \mathbb{Z}^d \times \Omega$$
$$(x, \omega) \mapsto (-x, R\omega),$$

[1] We will also use the symbol $\mathcal{B}$ to denote an object called the "limit shape" in much of what follows.

where $(R\omega)_{\{x,y\}} = \omega_{\{-x,-y\}}$.

In this chaper, unless otherwise stated, we will in fact assume that the edge weights $\{\omega_e\}$ are independent and identically distributed, or i.i.d. In this case, each edge weight is distributed on $[0,\infty)$ according to some measure $\nu$, and $\mathbb{P}$ is the product distribution. Because many of the results and conjectures for the first-passage model are specific to the i.i.d. case, the restriction to this case will often prove useful in what follows. However, the new results presented in this thesis are proved a broader class of ergodic $\mathbb{P}$, as were the results of Hoffman [51] which are presented in Section 2.5. The specific class of distributions for which Hoffman's results are valid will be defined at the beginning of Section 2.5.

A finite lattice path $\gamma$ is some alternating sequence of vertices and edges $(x_0, e_0, x_1, e_1, \ldots, e_{n-1}, x_n)$, where $x_i \in V$ and $e_i \in E$ and such that $e_i = \{x_i, x_{i+1}\}$ for all $i$.

Fix a realization $\omega$ of edge weights. Given a finite lattice path $\gamma$, we define the path passage time $\tau(\gamma)$ by the equation

$$\tau(\gamma) = \sum_{e\in\gamma} \omega_e. \tag{2.4}$$

Given two vertices $x$ and $y$ in $\mathbb{Z}^d$, we define $\tau(x,y)$, the passage time between $x$ and $y$, to be the minimal $\tau(\gamma)$ over all finite $\gamma$ connecting $x$ to $y$ (sometimes written $\gamma : x \rightsquigarrow y$). Such a minimizing $\gamma$ will be called a finite geodesic between $x$ and $y$. The question of existence of finite geodesics is somewhat delicate and will be addressed further in Subsection 2.2.2. We note here, however, that almost sure existence of finite geodesics is immediate in the case that the distribution of $\omega_e$ is bounded away from zero (i.e., in the case that $\mathbb{P}(\omega_e > \delta) = 1$ for some $\delta > 0$).

We will sometimes consider $\tau$ to be a function on (subsets of) $\mathbb{R}^d \times \mathbb{R}^d$. If $z, z' \in \mathbb{R}^d$ and there exist vertices $x, x' \in V$ with $z \in x + [-1/2, 1/2)^d$ and $z' \in x' + [-1/2, 1/2)^d$, we set $\tau(z, z') = \tau(x, x')$. This extension is especially useful in the case that $(V, E)$ is $(\mathbb{Z}^d, \mathcal{E}^d)$, where it is used in the statement of the so-called shape theorem.

### 2.2.2 Linear Order Behavior: Shape Theorem

As discussed above, the relevant questions for the first-passage model center around the study of geodesics and their passage times. A first question is the first-order behavior of the passage time–that is, the leading contribution to $\tau(0, nx)$ for $x \in \mathbb{Z}^d$ and $n$ large. It was observed by Hammersley and Welsh [49] that the passage time is subadditive; that is, $\tau(x, z) \le \tau(x, y) + \tau(y, z)$ for all vertices $x$, $y$, and $z$. This subadditivity is a manifestation of the fact that the passage time is defined by a minimization procedure, and the concatenation of two finite geodesics may be used as a convenient trial minimizer. Kingman [68] showed that this implies that the leading-order behavior of $\tau(0, nx)$ is linear and deterministic:

**Theorem 2.2.1** (Kingman). *Assume* $\mathbb{E}[\omega_e] < \infty$. *Then*

$$\frac{\tau(0, n\mathbf{e}_1)}{n} \longrightarrow \inf \frac{\mathbb{E}\tau(0, n\mathbf{e}_1)}{n} =: g(\mathbf{e}_1),$$

*almost surely and in* $L^1(\mathbb{P})$

(here $\mathbf{e}_1$ is the unit first coordinate vector). The result is a manifestation of Kingman's quite general subadditive ergodic theorem, which has seen application for similar thermodynamic limits in disordered systems–e.g., showing the existence of a deterministic free energy density for certain disordered systems in the thermodynamic limit.

The averaging result of Theorem 2.2.1 describes the growth of the minimizing passage time for point-to-point boundary conditions, as the points are translated further apart. One may ask whether this asymptotic behavior holds for a free boundary condition–that is, one endpoint of the path fixed at the origin and another allowed to wander on the boundary of a cube of linear size $n$. This question is equivalent to asking whether averaging behavior of passage times described in Theorem 2.2.1 can break down in random directions–i.e., whether there can exist exceptional directions which are abnormally favorable or unfavorable for our directed paths to travel in.

Building on the work of Richardson [84], Cox and Durrett showed [31] that this is not the case, given quite weak assumptions on the distribution $\mathbb{P}$. To describe one of their results, we define $B(t)$ to be the set of all points $x \in \mathbb{R}^d$ such that $\tau(0, x) \leq t$.

**Theorem 2.2.2** (Shape Theorem [31])**.** *Let $d \geq 2$, and define $Y = \min\{X_i\}_{i=1}^{2d}$, where the $\{X_i\}$ are distributed independently with distribution $\nu$ (i.e., the same distribution as a single $\omega_e$); assume that $\mathbb{E}Y^d < \infty$.*

- *If $g(\mathbf{e}_1) > 0$, then there exists some compact, convex $\mathcal{B} \subseteq \mathbb{R}^d$ with nonempty interior which shares the symmetries of $\mathbb{Z}^d$ such that, for arbitrary $\varepsilon > 0$,*

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

- *If $g(\mathbf{e}_1) = 0$, then for every compact $K$,*

$$\mathbb{P}\left(K \subseteq \frac{B(t)}{t} \text{ for all large } t\right) = 1.$$

Because a shape-type theorem for a different quantity will play a major role in the new results of this thesis (see Chapter 3), we will provide a proof of this result here. The proof presented here differs somewhat in perspective from the original in [31] and has been influenced by [19].

*Proof of Theorem 2.2.2.* The first step in the proof is to show that $\tau(0, x)$ has at least as many moments as $Y$.

**Claim 2.2.3.** $\mathbb{E}(\tau(0,x)^d) < \infty$ *for all $x \in \mathbb{Z}^d \setminus \{0\}$.*

To see this, note that there exist $2d$ disjoint paths $\{\gamma_i\}_{i=1}^{2d}$ from 0 to $x$; assume $\gamma_1$ is the longest (in euclidean length), and denote its length by $L_1$. Then

$$\begin{aligned}\mathbb{P}\left(\tau(0,x) > s\right) &\leq \mathbb{P}(\tau(\gamma_1) > s)^{2d} \\ &\leq \left[L_1 \mathbb{P}\left(\omega_e > s/L_1\right)\right]^{2d} \\ &= L_1^{2d}\, \mathbb{P}\left(Y > s/L_1\right).\end{aligned}$$

This proves the claim.

Using subadditive ergodic theory as mentioned above, it can be shown that there exists some function $g : \mathbb{R}^d \to \mathbb{R}$ such that, if $x \in \mathbb{Q}^d$ with $Mx \in \mathbb{Z}^d$,

$$\lim_{n\to\infty} \frac{\tau(0, nMx)}{nM} = g(x) \tag{2.5}$$

almost surely, and in fact, $\lim \tau(0, nx)/n$ exists almost surely; in particular, since $\mathbb{E}(\tau(0,x))$ is finite, $g$ is finite). The $g$ function is uniformly continuous and thus extends to a seminorm (i.e., it scales linearly and inherits the subadditivity of $\tau$) on $\mathbb{R}^d$. See [61] for a detailed discussion and proofs.

In the case that $g$ is a norm–i.e., when $g(x) \neq 0$ for all nonzero $x$–then we define $\mathcal{B}$ to be its closed unit ball. Note also that if $g(x) = 0$ for some nonzero $x$, then $g$ is identically zero. To see this, note that $g$ is symmetric about rotations and reflections, and we can build a basis for $\mathbb{R}^d$ out of reflections and rotations of $x$. This combined with subadditivity of $g$ insures that $g$ is identically zero if and only if $g(\mathbf{e}_1) = 0$.

Our proof will rely on one more tool, whose proof is taken more or less directly from the proof of (3.5) in [31]. We present it as the following lemma, whose proof we delay momentarily.

**Lemma 2.2.4.** *There is a constant $\kappa < \infty$ such that, for any $x \in \mathbb{Z}^d$,*

$$\mathbb{P}\left( \sup_{\substack{z \in \mathbb{Z}^d \\ z \neq x}} \frac{\tau(x,z)}{\|x-z\|_1} < \kappa \right) > 0. \tag{2.6}$$

We will call an $x$ in $\mathbb{Z}^d$ for which the event appearing in (2.6) occurs a "good" vertex. We can immediately leverage the information in Lemma 2.2.4 to show

**Claim 2.2.5.** *Let $\zeta \in \mathbb{Z}^d$. For a given realization of edge weights, denote by $(n_k)$ the sequence of natural numbers such that $n_k\zeta$ is a good vertex. Then with probability one, the sequence $(n_k)$ is infinite and* $\lim_{k\to\infty}(n_{k+1}/n_k) = 1$.

To see that the claim is correct, let $B_m$ denote the event that $m\zeta$ is a good vertex. Then

$$\frac{k}{n_k} = \frac{1}{n_k}\sum_{i=1}^{n_k} \mathbf{1}_{B_i};$$

the right-hand side converges to the probability in (2.6) by the ergodic theorem. Thus,

$$\frac{n_{k+1}}{n_k} = \left(\frac{n_{k+1}}{k+1}\right)\left(\frac{k}{n_k}\right)\left(\frac{k+1}{k}\right) \longrightarrow 1$$

almost surely. This proves the claim.

Let $\Xi_1$ denote the event that $\lim \tau(0,nq)/n = g(q)$ for all $q$ with rational coordinates; let $\Xi_2$ denote the event that for every $\zeta \in \mathbb{Z}^d$, the sequence $(n_k)$ defined in Claim 2.2.5 is infinite and that the ratio of successive terms tends to one. From here, the proof of Theorem 2.2.2 proceeds by contradiction. Assume the Shape Theorem does not hold. Then there exist a $\delta > 0$ and a collection of edge weight outcomes $D_\delta$ with $\mathbb{P}(D_\delta) > 0$ such that, for every outcome in $D_\delta$, there are infinitely many vertices $x \in \mathbb{Z}^d$ with

$$|\tau(0,x) - g(x)| > \delta\|x\|_1. \tag{2.7}$$

For the remainder of the proof, we will use $\|\cdot\|$ to denote the $\ell^1$ norm.

Since $\mathbb{P}(\Xi_1) = 1 = \mathbb{P}(\Xi_2)$ (by our claims about $g$ and Claim 2.2.5), the event $D_\delta \cap \Xi_1 \cap \Xi_2$ contains some outcome $\omega$; we claim $\omega$ has contradictory properties. On outcome $\omega$, there must exist a sequence $(x_i) \subseteq \mathbb{Z}^d$ satisfying the condition in (2.7). We can assume that $x_i/\|x_i\|$ converges to some $y$ with $\|y\| = 1$ by compactness of the unit sphere. Let $\delta' > 0$ be arbitrary; we will fix its value at the end of the proof. We first choose some large $N$ such that $\|x_n/\|x_n\| - y\| < \delta'$ and such that $|g(x_n) - \|x_n\|g(y)| < \delta/2$ for $n > N$. Then we have for $n > N$ (using our assumption (2.7)):

$$|\tau(0, x_n) - g(y)| > \delta\|x_n\|/2. \tag{2.8}$$

Next, we set up a sequence of approximating good vertices. We find some $z \in \mathbb{R}^d$, $\|z\| = 1$ such that $\|z - y\| < \delta'$, with the additional property that $z = x/M$ for some $x \in \mathbb{Z}^d$ and some positive integer $M$. This can be done because vectors with rational coordinates are dense in the unit sphere. On $\omega$, there must exist a sequence $(n_k)$ such that $n_k M z$ is a good vertex and such that $n_{k+1}/n_k$ tends to one. For any $n$, there exists a value of $k$ such that $n_{k+1}M > \|x_n\| > n_k M$; denote this value by $k(n)$. Finally, fix $K > 0$ such that $n_{k+1} < (1 + \delta')n_k$ and $|\tau(0, n_k M z)/n_k M - g(z)| < \delta'$ for all $k > K$. We now let $n > N$ be large enough that $k(n) > K$.

Before completing the calculation here, it is worth considering where the contradiction will arise. We have (essentially by assumption) that $\tau(0, ny) - ng(y)$ is of order $n$ for infinitely many $n$. Since $g$ is a norm, $g(y)$ and $g(z)$ are arbitrarily close; and since infinitely many of the $\{nz\}$ are good vertices, $\tau(0, ny)$ and $\tau(0, nz)$ are arbitrarily close. Thus, $|\tau(0, nz) - ng(z)|$ is large–but this is counter to the properties assumed for $\omega$.

To turn the above into a rigorous estimate, write $k$ for $k(n)$ and expand

$$\left|\frac{\tau(0,x_n)}{\|x_n\|} - g(y)\right| \leq \left|\frac{\tau(0,x_n) - \tau(0,n_kMz)}{\|x_n\|}\right| + \frac{\tau(0,n_kMz)}{n_kM}\left(1 - \frac{n_kM}{\|x_n\|}\right)$$
$$+ \left|\frac{\tau(0,n_kMz)}{n_kMz} - g(z)\right| + |g(z) - g(y)|\,.$$

There are four terms on the right-hand side of the above, which we number from left to right and bound individually in terms of $\delta'$.

Term 1. Since $n > N$ and $k > K$, we have that $n_kM < \|x_n\| \leq (1+\delta')n_{k+1}M$, that $\|y - n_kMz\| \leq \delta' n_kM$, and that $\|x_n/\|x_n\| - y < \delta'$. Therefore, $\|x_n - n_kMz\| \leq 2\delta'\|x_n\|$. Using the fact that $n_kMz$ is a good vertex yields

$$|\tau(0,x_n) - \tau(0,n_kMz)| \leq \kappa\|x_n - n_kMz\| \leq 2\kappa\delta'\|x_n\|.$$

Term 2. The relationship between $n_kM$ and $\|x_n\|$ given in the Term 1 estimates yields an upper bound for the second factor of Term 2. By the fact that $k > K$, we can bound the first factor. The overall bound is

$$[g(z) + \delta']\left(1 - (1+\delta')^{-1}\right).$$

Term 3. By the fact that $k$ is chosen greater than $K$, this term is bounded above by $\delta'$.

Term 4. If $g$ is identically zero, this term is trivially zero. If $g$ is not identically zero, it is a norm on $\mathbb{R}^d$ and is thus bounded by Euclidean norm:

$$c_L\|\cdot\| \leq g(\cdot) \leq c_U\|\cdot\|.$$

Since $\|z - y\| < \delta'$, Term 4 is bounded above by $\delta'$ times a constant depending only on $g$.

We have therefore bounded the left-hand side of (2.8) above by an expression of the form $f(\delta')\|x_n\|$, where $f$ tends to zero as $\delta' \to 0$. Since $\delta'$ was arbitrary, we can choose it such that $f(\delta')\|x_n\|$ is smaller than the right-hand side of (2.8). This contradiction proves the theorem.

□

*Proof of Lemma 2.2.4.* By translation-invariance, we may assume that $x$ is the origin. The first step is to show a weak version of the lemma for a "sparse" square lattice. We will call a pair of vertices $z, z'$ in $5\mathbb{Z}^d$ "5-adjacent" if the Euclidean distance between them is 5. If $z$ and $z'$ are 5-adjacent, there exist $2d$ disjoint paths between $z$ and $z'$ which lie entirely in the set

$$\{z, z'\} + [-5/2, 5/2)^d.$$

We will define $\hat{\tau}(z, z')$ to be the minimum of the passage times of these paths; the proof of Claim 2.2.3 shows that $\mathbb{E}(\hat{\tau}(z, z')^d)$ is finite.

Under these definitions, we can treat $5\mathbb{Z}^d$ as a renormalized lattice with 5-edges $f = \{z, z'\}$ between 5-adjacent vertices and edge passage times $\hat{\tau}(f) = \hat{\tau}(z, z')$. Then the conclusion of the preceding paragraph is that $\mathbb{E}\hat{\tau}(f)^d < \infty$; note also that $\hat{\tau}(f)$ and $\hat{\tau}(f')$ are independent if $f$ and $f'$ have no common endpoint. We extend $\hat{\tau}(\cdot, \cdot)$ to all of $5\mathbb{Z}^d$ by the usual first-passage procedure of minimizing sums of 5-edge weights over paths.

For any vertex $x \in 5\mathbb{Z}^d$, it is easy to show that there exist $2d$ disjoint paths of 5-adjacent vertices connecting 0 and $x$, with no path having more than $\|x\|_1 + 4$ vertices (see, for instance, the discussion following [61, (2.11)]); call these paths, sans their respective first and last vertices, $\{r_i\}$. Then we have (where $f$ is an arbitrary

5-edge)

$$\mathbb{P}\left(\hat{\tau}(0,x) > 10\|x\|_1\mathbb{E}\hat{\tau}(f)\right) \le \mathbb{P}\left(\hat{\tau}(r_i) > 9\|x\|_1\mathbb{E}\hat{\tau}(f), \text{for all } i\right) \\ + 4d\mathbb{P}\left(\hat{\tau}(f) > \|x\|_1\mathbb{E}\hat{\tau}(f)\right). \tag{2.9}$$

The first term on the right side of (2.9) may be bounded above by Chebyshev's inequality, using the facts that, for all $i$,

- $\mathbb{E}[\hat{\tau}(r_i)] \le 8\|x\|_1\mathbb{E}\hat{\tau}(f)$,
- $\mathrm{Var}\,\hat{\tau}(r_i) \le 8\|x\|_1\mathrm{Var}\,\hat{\tau}(f)$, and
- $\mathbb{E}\hat{\tau}(f)^d < \infty$.

Using the first two of the above implies

$$\mathbb{P}\left(\hat{\tau}(r_i) > 9\|x\|_1\mathbb{E}\hat{\tau}(f), \text{for all } i\right) = \prod_{i=1}^{2d} \mathbb{P}\left(\hat{\tau}(r_i) > 9\|x\|_1\mathbb{E}\hat{\tau}(f)\right) \\ \le C_d/\|x\|_1^{2d}, \tag{2.10}$$

where $C_d$ is some constant depending on $d$ and $\mathbb{P}$.

Applying the bound (2.10) in (2.9) and summing over all $x \in 5\mathbb{Z}^d$ shows that there exists a constant $D$ such that with probability one, $\hat{\tau}(0,y) < D\|y\|_1$ for all but finitely many $y \in 5\mathbb{Z}^d$. Since $\tau(0,y) \le \hat{\tau}(0,y)$, this proves that there exists a $\kappa_5$ such that

$$\mathbb{P}\left(\sup_{z\in 5\mathbb{Z}^d} \frac{\tau(0,z)}{\|z\|} < \kappa_5\right) > 0. \tag{2.11}$$

The above is a sparsified version of the bound claimed in (2.6); it only remains to extend this to all of $\mathbb{Z}^d$. To do this, note that if $x \in \mathbb{Z}^d$ and $z(x)$ is the nearest vertex of $5\mathbb{Z}^d$ to $x$, then defining $R_d = \sup_{x\in[-2,2]^d} \tau(0,x)$,

$$\mathbb{P}\left(\tau(x, z(x)) > \|x\|\right) \le \mathbb{P}\left(R_d > \|x\|\right). \tag{2.12}$$

Since $R_d$ has finite $d$th moment, we can sum (2.12) over all $x \in \mathbb{Z}^d$ as before to see that with probability one, all but finitely many vertices $x \in \mathbb{Z}^d$ can be reached from the nearest vertex of $5\mathbb{Z}^d$ within passage time $\|x\|$. This fact, combined with (2.11), proves the lemma. □

The case in which $g$ is identically zero is well-characterized for i.i.d. first-passage percolation. In fact, it is known that this occurs exactly when $\mathbb{P}(\omega_e = 0) \geq p_c$, the critical probability for i.i.d. bond percolation on $\mathbb{Z}^d$ (see [61] for a proof, and see Chapter 5 and [43] for more on bond percolation processes).

As noted in the proof above, assuming $g$ is not identically zero, the limit shape $\mathcal{B}$ is the closed unit ball of some norm $g$ satisfying $g(x) = \lim_n \tau(0, nx)/n$. Since the shape theorem says that $\tau(0, nx) = ng(x) + o(n)$ in a global sense, to some extent the linear order behavior of the passage time is established. However, beyond the aforementioned properties of convexity, compactness, and symmetry, very little is known about the behavior of $\mathcal{B}$ for general edge weight distributions (see [27, Section 5] for a discussion).

There are, however, conjectures about $\mathcal{B}$ based on both numerical evidence and the behavior of the related model of last-passage percolation [75, 85]. It is widely believed that for $(\omega_e)$ i.i.d. with continuous distribution and enough moments, the boundary $\partial\mathcal{B}$ of $\mathcal{B}$ is uniformly curved; as will be discussed later, the question of curvature of $\partial\mathcal{B}$ has implications for fluctuations in the model. On the other hand, uniform curvature has been rigorously ruled out in the degenerate case where the edge weight distribution has a large atom at the infimum $\lambda \neq 0$ of its support. Here the relevant condition for a "large atom" is that $\mathbb{P}(\omega_e = \lambda) \geq \vec{p}_c$, where $\vec{p}_c$ is the critical probability for directed bond percolation. In such cases, points $x$ and $y$ arbitrarily far apart will have $\tau(x, y) = \tau(\gamma)$ for some $\gamma$ consisting only of edges with weight $\lambda$, and the boundary of the limit shape has a flat edge [38].

Lastly, we will note that if the limit shape $\mathcal{B}$ is bounded (i.e., if $\mathbb{P}(\omega_e = 0) < p_c$), the almost sure existence of finite geodesics between any pair of points is immediate. In Section 3.9.2, we furnish a proof that geodesics always exist for i.i.d. measures on a broad subclass of connected subgraphs of $\mathbb{Z}^2$.

## 2.3 Fluctuations and Geodesics

The shape theorem result of the last section describes the leading-order behavior of $\tau(0, nx)$ for i.i.d. measures with $\nu(\{0\})$ small– that is, $\tau(0, nx) \sim ng(x)$. The norm $g$ is a deterministic property of the passage time distribution; the randomness of $\tau(0, nx)$ is therefore present only at order $o(n)$. In fact, $\tau(0, nx)$ is believed to exhibit fluctuations of typical order $n^\chi$ around its mean, where $\chi$ is the so-called *fluctuation exponent*. In this section we describe more fully these fluctuation conjectures, as well as the rigorous results that currently exist. The discussion of this section will again be restricted to the case of i.i.d. passage time measures $\mathbb{P}$.

A closely related set of questions in the model revolve around the typical "fluctuations" of time-minimizing paths themselves. For instance, if $\gamma$ is a finite geodesic connecting 0 and $nx$, one could ask for the typical diameter of the smallest cylinder (with axis parallel to $x$) containing $\gamma$. This diameter is expected to behave like $n^\xi$, where $\xi$ is sometimes referred to as the *wandering exponent*. We will discuss the predicted wandering behavior later in this section, followed by the conjectured relationship between $\chi$ and $\xi$, which has been the site of several recent developments in the field.

Much of work on fluctuations in the model relies on relating fluctuation properties of the passage time $\tau$ to those of the edge weights. Therefore, unless it is explicitly stated otherwise, we will assume that $0 < \mathrm{Var}\,(\omega_e) < \infty$ in the discussion that follows.

### 2.3.1 The Fluctuation Exponent $\chi$

One issue encountered in the study of either exponent is that there is no broad agreement as to their precise mathematical definitions (the recent paper [27] contains some commentary on the state of the field; for several possible definitions of the exponents, see [73, 78]). Rather than trying to give an exhaustive account of possible meanings of $\chi$, we will discuss particular properties one might expect $\chi$ to have and the results and challenges presented thereby.

For any $\chi' < \chi$, one could reasonably expect that there exists some $C > 0$ such that

$$Cn^{2\chi'} \leq \mathrm{Var}\,(\tau(0, n\mathbf{e}_1)), \quad \text{for all } n \geq 1. \tag{2.13}$$

Bounds of the type (2.13) (and corresponding upper bounds) have so far been elusive. One early result is due to Kesten [67]:

**Theorem 2.3.1.** *Consider first-passage percolation on $\mathbb{Z}^d$. Assume that $0 < \mathrm{Var}\,(\omega_e) < \infty$, and that $\mathbb{P}(\omega_e = 0 < p_c)$, where $p_c > 0$ is the critical probability for Bernoulli bond percolation on $\mathbb{Z}^d$. Then there exist constants $C, C' > 0$ such that*

$$C \leq \mathrm{Var}\,\tau(0, n\mathbf{e}_1) \leq C'n \quad \textit{for all } n \geq 1.$$

*Moreover, there exists some constants $C_1, C_2, C_3$ such that if $x \leq C_1 n$,*

$$\mathbb{P}\left(\left|\frac{\tau(0, n\mathbf{e}_1) - \mathbb{E}\tau(0, n\mathbf{e}_1)}{n^{1/2}}\right| \geq x\right) \leq C_2 \exp(-C_3 x).$$

In the article [67], the opinion is expressed that the upper bound in Theorem 2.3.1 would be in general suboptimal. That is, for i.i.d. $\mathbb{P}$ satisfying some reasonable hypotheses, there should exist, for at least some $d$, a power-law upper bound of the form

$$\mathrm{Var}\,\tau(0, n\mathbf{e}_1) \leq C'n^{1-\varepsilon} \quad \text{for all } n,$$

for some $\varepsilon > 0$. For general $d$, the exact nature of the scaling is a point of some debate; some predictions suggest, for instance, that $\operatorname{Var} \tau(0, n\mathbf{e}_1)$ should grow more slowly than any power of $n$ when $d$ is large enough (see [30] and the discussion in the introduction of [78]). In low dimensions, however, it is predicted [78, 60] that there should be a nontrivial power lower bound for $\operatorname{Var} \tau(0, n\mathbf{e}_1)$.

In the most relevant case for the results of this dissertation, the case $d = 2$, the prediction [53, 54, 59] that $\chi = 1/3$ appears to be fairly accepted in the literature (see [71] for an overview). The two-dimensional case is also where the strongest lower bound on $\operatorname{Var} \tau(0, n\mathbf{e}_1)$ has been derived. In [78], it is shown for first-passage percolation on $\mathbb{Z}^2$ that– assuming the hypotheses of Theorem 2.3.1 and that there is not a large delta mass at the infimum of the edge distribution $\nu$ – there exists a $c > 0$ such that

$$\operatorname{Var} \tau(0, n\mathbf{e}_1) \geq c \log n.$$

The result was derived independently in a special case in [81].

A similar result in the direction of upper-bounding variances was provided for all $d$ by Benjamini, Kalai, and Schramm [15] in the case that $\nu$ consists of two delta masses. The result of [15] shows that there exists a $c > 0$ such that

$$\operatorname{Var} \tau(0, n\mathbf{e}_1) \leq \frac{cn}{\log n}.$$

This result was extended to a broader class of distributions, the so-called "nearly gamma" distributions, in [14]. To date, sharper upper or lower bounds for $\chi$ remain elusive. Certain additional results on $\chi$ have been derived using properties of $\xi$ and path fluctuations; more on the relationship between fluctuations of $\tau$ and fluctuations of time-minimizing paths will be described in the next section.

In fact, methods which utilize fluctuation properties of $\tau$ have been used to study the particular path fluctuation problems which are addressed in Chapters 3 and 4.

To give an adequate description of these methods, we need to introduce results here on deviations of $\tau$ from a different mean–namely, the limiting shape of the model. Recall that if $\mathbb{P}(\omega_e = 0)$ is small enough, the limit

$$\lim_{n\to\infty} \frac{\mathbb{E}\tau(0, nx)}{n} = g(x), \tag{2.14}$$

where $g(x)$ is the norm inducing the limit shape. The rate of convergence to the limit in (2.14) has been the topic of study by Kesten [67], whose results were sharpened by Alexander [3] to show the following:

**Theorem 2.3.2.** *Assume that* $\mathbb{P}(\omega_e = 0) < p_c$ *and that* $\mathbb{E}\exp(c\omega_e) < \infty$ *for some* $c > 0$. *Then there exists* $C > 0$ *such that*

$$g(x) \leq \mathbb{E}\tau(0, x) \leq g(x) + C\|x\|_1^{1/2} \log \|x\|_1$$

*for all* $x \in \mathbb{Z}^2$ *with* $\|x\|_1 > 1$.

As will be described in what follows, the results of Kesten and Alexander have proven fruitful in relating properties of geodesics or time-minimizing paths to properties of the limit shape. A major goal of our efforts in the following chapters will be to study what can be done in this direction without relying on strong concentration results like Theorem 2.3.2.

### 2.3.2 $\xi$ and the Scaling Relation

As discussed alongside the shape theorem in Section 2.2.2, there are strange effects allowed by the presence of delta masses in $\nu$. If $\lambda$ is the infimum of the support of $\nu$, then in the case that $\lambda = 0$ and $\nu(\{\lambda\}) > p_c$ or $\lambda > 0$ and $\nu(\{\lambda\}) > \vec{p}_c$, there will be infinite clusters of edges $e$ with $\omega_e = \lambda$ whose segments will serve as geodesics between vertices arbitrarily far apart. Aside from these considerations, the presence of delta masses allows the existence of multiple distinct geodesics between pairs of vertices. Thus, in this section, we will assume that $\mathbb{P}$ is i.i.d. with continuous marginal $\nu$, ensuring that finite geodesics between pairs of vertices of $\mathbb{Z}^d$ exist and are unique (for more on the issues raised in this paragraph, see [73]).

There are numerous possible definitions of $\xi$. For concreteness, we will provide a possible choice which was studied in [73]. For $x \in \mathbb{Z}^d$, let $\mathcal{C}(x, r)$ denote the cylinder with axis $\{\alpha x : \alpha \in \mathbb{R}\}$ and radius $r$; for $x$ and $y$ in $\mathbb{Z}^d$, let $\mathcal{M}(x, y)$ be the finite geodesic between $x$ and $y$. Then a possible "point-to-point" definition for $\xi$ is given by

$$\xi := \sup\{\gamma \geq 0 : \limsup_{\|x\|_2 \to \infty} \mathbb{P}\left(\mathcal{M}(0, x) \subseteq \mathcal{C}(x, \|x\|_2^{\gamma})\right) < 1\}. \tag{2.15}$$

Various other rigorous definitions have been proposed for $\xi$, including direction-dependent versions of the point-to-point definition above [78] as well as "point-to-line" versions of the above [73]. While it appears reasonable to conjecture that these definitions would yield the same value of $\xi$, there do not appear to be results resolving this question.

While $\xi$ and $\chi$ are expected to depend on the dimension $d$, they are supposed to be the same for all $\nu$ which are sufficiently nice. Moreover, the exponents are widely believed ([39, 59, 60, 70] and others; see also [71]) to satisfy the relation

$$\chi = 2\xi - 1 \tag{2.16}$$

in a wide range of first-passage-like random path models, independent of dimension. The prediction [53, 54] that $\chi = 1/3$ and $\xi = 2/3$ in $d = 2$ is perhaps the best-known instance of (2.16). Proving (2.16) in the case of first-passage percolation has been the object of substantial efforts in the mathematical physics literature.

We will summarize some of the results in this direction here, noting that the definitions of $\xi$ and $\chi$ used in different papers are often distinct. Building on work of Wehr and Aizenman [92] in the related model of directed first-passage percolation, Newman and Piza [78] showed the inequality

$$\chi \geq \frac{1-(d-1)\xi}{2}$$

for a certain "directional" definition of $\chi$ and $\xi$. Also appearing in [78] is the rigorous lower bound

$$\chi \geq 2\xi - 1. \tag{2.17}$$

for a different "global" definition of $\chi$. Bounds on $\xi$ for various definitions (including (2.15)) appear in [73].

Recently, a complete verification of (2.16) was proved by Chatterjee [27] under certain assumptions. Building on Chatterjee's work, a simplified proof which removed an assumption on the passage-time distribution $\mathbb{P}$ was later provided by Auffinger and Damron in [9]. To state the full form of Chaterjee's result (as strengthened in [9]), we define "upper" and "lower" exponents $\xi_a, \chi_a$ and $\xi_b, \chi_b$. Then $\xi_a$ and $\chi_a$ are defined as the smallest real numbers such that, for each $\chi' > \chi_a$ and $\xi' > \xi_a$, there exists $\alpha > 0$ such that the quantities

$$\sup_{0\neq v\in\mathbb{Z}^d} \mathbb{E}\exp\left(\alpha\frac{|\tau(0,v)-\mathbb{E}\tau(0,v)|}{\|v\|_2^{\chi'}}\right),$$
$$\sup_{0\neq v\in\mathbb{Z}^d} \mathbb{E}\exp\left(\alpha\frac{R(0,v)}{\|v\|_2^{\xi'}}\right)$$

are finite, where $R(0,v)$ is the smallest $r > 0$ such that $\mathcal{M}(0,v) \subseteq \mathcal{C}(v,r)$. Similarly, $\xi_b$ and $\chi_b$ are the largest numbers such that, if $\xi' < \xi_b$ and $\chi' < \chi_b$, there exists $\alpha$

such that

$$\inf_{\substack{v \in \mathbb{Z}^d \\ \|v\|_2 > \alpha}} \frac{\operatorname{Var} \tau(0,v)}{\|v\|_2^{2\chi'}} > 0,$$

$$\inf_{\substack{v \in \mathbb{Z}^d \\ \|v\|_2 > \alpha}} \mathbb{E} \frac{R(0,v)}{\|v\|_2^{\xi'}} > 0.$$

Then the result of [9, 27] is

**Theorem 2.3.3.** *Consider i.i.d. first-passage percolation on $\mathbb{Z}^d$ such that the edge weight marginal $\nu$ does not have a large delta mass at the infimum of its support, and such that $\mathbb{E}\exp(c\,\omega_e)$ is finite for some $c > 0$. Assume that $\xi_a = \xi_b$ and $\chi_a = \chi_b$. Then* (2.16) *holds.*

The conditions under which the assumption that $\xi_a = \xi_b$, $\chi_a = \chi_b$ holds currently remain obscure.

There is another notion of geodesic wandering in the "geodesic tree", closely related to so-called *infinite geodesics*, which will play a major role in the next section (and the other two chapters of Part I). In the case that $\mathbb{P}$ is an i.i.d. measure with continuous marginal $\nu$, then if $\gamma \neq \gamma'$ are two finite paths in $\mathbb{Z}^d$, we have $\tau(\gamma) \neq \tau(\gamma')$ with $\mathbb{P}$-probability one. Therefore, (for such $\mathbb{P}$) finite geodesics must almost surely exist between each pair of vertices of $\mathbb{Z}^d$, and there will be a unique such geodesic for each pair.

Therefore, for each $x \in \mathbb{Z}^d$, we can define a graph $\mathcal{T}(x)$ with vertex set $\mathbb{Z}^d$ and whose edge set contains every $e$ which is a member of some geodesic $\mathcal{M}(x,y)$; this graph will be almost surely connected by the existence of finite geodesics.

**Claim 2.3.4.** *Every (self-avoiding) path in $\mathcal{T}(x)$ is a point-to-point geodesic.*

*Proof.* Let $\gamma = (y = v_0, e_1, v_1, \ldots, e_{n-1}, v_n = z)$ be a path as in the claim. By the definition of $\mathcal{T}(x)$, the edge $e_1$ is in some geodesic from $x$ to either $v_0$ or $v_1$. We will assume the former case; the latter is similar. Denote by $\gamma_i$ the subsequence of $\gamma$ up to vertex $v_i$; we will proceed by induction on $i$.

Now, assume that $\gamma_i$ is the terminal segment of some geodesic $g_i$ from $x$ to $v_i$. Then $e_{i+1}$ is an edge of a geodesic $\gamma'$ from $x$ to either $v_{i+1}$ or $v_i$. If $\gamma'$ is a geodesic to $v_i$, then its terminal segment must be $\gamma_i$ by uniqueness, and so $v_i$ appears multiple times in $g_i$, a contradiction since each $\omega_e > 0$ almost surely. If $\gamma'$ is a geodesic to $v_{i+1}$, then it passes through $v_i$ first; by uniqueness of geodesics, $\gamma'$ is the concatenation of $g_i$ with $e_{i+1}$ and we are done. □

By the claim above and the uniqueness of finite geodesics, we see that $\mathcal{T}(x)$ is almost surely a tree, called the *geodesic tree* or tree of infection of $x$. Regarding $\mathcal{T}(x)$ as a sort of family tree, one could ask how the tree widens spatially between generations. For instance, if $y$ is a vertex, are the descendents of $y$ in $\mathcal{T}(x)$ all in some cone with axis parallel to $y$? Questions of this type will be central in the rest of this chapter, and will be the focus of the results presented in Chapters 3 and 4.

## 2.4 Infinite Geodesics

In this section, we will largely confine ourselves to the case of i.i.d. first-passage percolation on $\mathbb{Z}^2$ such that each edge weight $\omega_e$ has continuous marginal $\nu$. This section describes issues related to infinite time-minimizing paths, or infinite geodesics, in the first-passage model. These are the questions at which the new results presented in the following chapters are aimed. In the next subsections, we will describe particular problems in infinite geodesics, some previous work on these problems, and then the contributions which will be described in full detail in Chapters 3 and 4

### 2.4.1 Number and Direction

Recall the definition of the geodesic tree $\mathcal{T}(x)$ from the previous section. An idea related to wandering of geodesic paths which appears in [76] is related to the spatial spreading of $\mathcal{T}(0)$.

Given a vertex $x \in \mathbb{Z}^2$, we will denote by $R_{\text{out}}(x)$ the set of sites $y \in \mathbb{Z}^2$ such that the geodesic from 0 to $y$ touches $x$. Let $C(x, \epsilon)$ denote the cone of $y \in \mathbb{R}^2$ such that the angle between $x$ and $y$ (considered as vectors) is smaller than $\epsilon$. Given a positive function $h$ on $\mathbb{R}$, we will say that $\mathcal{T}(0)$ is $h$-straight (for a given realization $\omega$) if for all but finitely many $x$ in $\mathbb{Z}^2$,

$$R^{\text{out}}(x) \subseteq C(x, h(\|x\|_2)).$$

Then analogously to bounding $\xi$, one could ask what for sort of functions $h$ the geodesic tree is $h$-straight with positive probability.

A result in this direction appears in [76]; the assumptions involved require some small amount of detail. We will say that $\mathcal{B}$ (or its norm $g$) is uniformly curved if for some $C > 0$ and any $z = \alpha z_1 + (1-\alpha) z_2$ with $g(z_i) = 1$ and $0 \le \alpha \le 1$, we have

$$1 - g(z) \ge C(\min\{g(z - z_1),\, g(z - z_2)\})^2.$$

**Theorem 2.4.1** ([76]). *Assume that $\mathcal{B}$ is uniformly curved and that $\mathbb{E}\exp(c\,\omega_e) < \infty$ for some $c > 0$. Then for any $\epsilon > 0$, the geodesic tree $\mathcal{T}(0)$ is almost surely $h$-straight, with $h(r) = r^{-1/4+\epsilon}$.*

As noted in [73], "it is a basic assumption behind the derivation of the relation $\chi = 2\xi - 1$" presented in [71] that (at least for certain continuous, i.i.d. $\mathbb{P}$), the limit shape should be uniformly curved.

*Sketch of proof of Theorem 2.4.1.* Let $0 < \delta < 1/4$. Letting $\|x\|_2$ be large, we examine $R^{\text{out}}(x)$ restricted to set $A(x)$ of sites $y \in C(x, g(x)^{-\delta})$ with $g(y)/g(x)$ lying between $1 - g(x)^{-2\delta}$ and 2. The boundary of $A_x$ is the set of all vertices of $A_x$ which are nearest neighbors of some vertex not in $A_x$. We will name three parts of the boundary of $A(x)$ : the "front" (the vertices adjacent to some $y$ with $g(y) > 2g(x)$), the "back" (the vertices adjacent to some $y$ with $g(y) < (1 - g(x))^{-2\delta} g(x)$), and the "side" (the rest of the boundary of $A_x$.

Assume that $R^{\text{out}}(x)$ contains a point $y$ on the side or back of $A(x)$. Then, since $x$ is on the geodesic between 0 and $y$, we have

$$\tau(0, y) = \tau(0, x) + \tau(x, y) \tag{2.18}$$

We have

$$g(y) + g(x - y) - g(x) \geq C_0 \|x\|^{1/2 - 2\delta} \tag{2.19}$$

(by curvature, or by $g(y)/g(x) \leq 1 - g(x)^{-2\delta}$, depending on the location of $y$).

Using results of Kesten and Alexander, Theorems 2.3.1 and 2.3.2, we can (for $\epsilon > 0$) replace the passage time $\tau(x, y)$ by

$$\tau(x, y) \sim g(x - y) + O(\|x - y\|_2^{1/2+\epsilon})$$

almost surely for all $y \in A(x)$ for all but a random finite collection of $x$. Combining this with (2.18) yields

$$g(y) = g(x) + g(x - y) + O(\|x\|_2^{1/2+\epsilon}).$$

This is in contradiction with (2.19).

So almost surely, for all but finitely many $x$, $R^{\text{out}}(x)$ is contained in the union of $A_x$ and the union, over all $x'$ on the front of $A_x$, of the $R^{\text{out}}(x')$. Induction now completes the proof. □

To see some consequences of this result, we make the following definitions. We will say that an infinite nearest-neighbor path $\gamma$ is an *infinite geodesic* (for a given edge weight realization) if every finite subpath of $\gamma$ is a finite geodesic. Infinite geodesics come in two varieties:

1. Indexed by $\mathbb{N}$–that is, $\gamma = (v_1, e_1, v_2, \ldots)$. These are called *unigeodesics*, *singly infinite geodesics* or (when it does not cause confusion) simply *geodesics*.

2. Indexed by $\mathbb{Z}$. These will be called *bigeodesics*.

Lastly, we will say that a self-avoiding singly infinite path $(x_1, e_1, x_2, \ldots)$ in a subgraph of $\mathbb{Z}^2$ has direction $\theta$ if $\arg x_n \to \theta$. An immediate consequence of Theorem 2.4.1 is the following:

**Corollary 2.4.2** ([76]). *Assume the hypotheses of Theorem 2.4.1. Then with probability one,*

- *Every singly infinite, self-avoiding path in $\mathcal{T}(0)$ has direction.*

- *For every $\theta \in [0, 2\pi)$, there exists a unigeodesic with direction $\theta$.*

Unfortunately, uniform curvature has not been proved for the limit shape of any i.i.d. $\mathbb{P}$. Nonetheless, the above corollary opens up a number of research questions regarding the directional properties of infinite geodesics. For instance, what is the probability that there exists a unigeodesic with direction $\theta$ for a "typical" $\mathbb{P}$? Moreover, if $\gamma$ is a unigeodesic, must it have some asymptotic direction? Finally, without making an assumption on curvature of the limit shape, is it possible to prove that unigeodesics exist in the first place?

At least we can give a quick answer to the last point: with probability one, there exists at least one unigeodesic (see [72, 76], though it is possible this fact was previously known). To see this, consider the finite geodesics $\{\gamma_n\}_n$, where $\gamma_n : 0 \rightsquigarrow n\mathbf{e}_1$. The first vertex of $\gamma_n$ is always 0 for each $n$. There are four possible choices for the second vertex of each $\gamma_n$. Therefore, there must be some edge $e$ incident to the origin such that infinitely many members of $\{\gamma_n\}$ have $e$ as their first edge. Repeating this argument on subsequences yields a singly infinite path which is easily seen to be a geodesic.

The method of the last paragraph could be called a "diagonal argument". One could next ask whether there exist multiple distinct unigeodesics with positive probability (here, distinct means sharing at most finitely many edges and vertices). In order to show the existence of more than one distinct unigeodesic, one could try to perform the diagonal argument of the preceding paragraph on two distinct sequences of finite geodesics. To do this, one would need a criterion for which different sequences of finite geodesics yield distinct unigeodesics when "diagonalized."

For i.i.d. first-passage percolation on $\mathbb{Z}^2$ with exponentially distributed edge weights, Häggström and Pemantle [48] showed that there will exist at least two distinct unigeodesics with positive probability. This result was extended to a wide range of first-passage distributions by Garet and Marchand [41] and Hoffman [52].

A 2008 paper by Hoffman [51] demonstrated that it is in fact possible to derive the existence of more than two unigeodesics from properties of the limit shape which are currently known. Given a limit shape $\mathcal{B}$, we define sides($\mathcal{B}$) to be equal to the number of sides of $\mathcal{B}$ if $\mathcal{B}$ is a polygon, and infinity otherwise. A major result of [51] is

**Theorem 2.4.3** ([51]). *Assume $\mathbb{P}$ is an i.i.d. edge-weight distribution on $\mathbb{Z}^2$ with continuous marginals and such that $\mathbb{E}\omega_e^{2+\alpha} < \infty$ for some $\alpha > 0$. Let $\varepsilon > 0$ be arbitrary, and define $G(x_1, \ldots, x_k)$ to be the event that there exist distinct unigeodesics* *beginning at vertices* $x_1, \ldots, x_k$. *Then for any* $k \le \text{sides}(\mathcal{B})$, *there exist* $x_1, \ldots, x_k$ *in* $\mathbb{Z}^2$ *such that*

$$\mathbb{P}\left(G(x_1, \ldots, x_k)\right) > 1 - \varepsilon.$$

In fact, Theorem 2.4.3 was proved not just for the i.i.d. continuous-marginal case considered in this section, but for a wide class of translation-ergodic edge weight distributions. We will devote the entirety of Section 2.5 to a presentation of a proof of Theorem 2.4.3.

Hoffman's results also include a statement that there exist locally favorable regions through which many finite geodesics will tend to pass. This line of thinking was a major impetus for the work presented in this dissertation, which in part addresses the following question:

**Q1**: Under what conditions on the passage time distribution do there exist unigeodesics with direction or directional concentration (i.e., in cones)? Is it possible to show some version of Corollary 2.4.2 without assuming uniform curvature?

Chapter 3 is addressed in part to **Q1**. There, it is shown that if $x$ is a point at which the boundary $\partial\mathcal{B}$ of $\mathcal{B}$ is differentiable with tangent line $L_x$, and if $I_x$ is the sector of angles in which $L_x$ intersects $\partial\mathcal{B}$, then there almost surely exists a unigeodesic which is concentrated in $I_x$. That is, if $I_x = [\theta_1, \theta_2]$, then there almost surely exists a unigeodesic $(x_1, e_1, x_2, \ldots)$ such that for every $\varepsilon > 0$, all but finitely many $x_i$ have $\arg x_i \in [\theta_1 - \varepsilon, \theta_2 + \varepsilon]$.

Another question is provoked by the diagonal argument for the existence of unigeodesics. If $(\gamma_n)_n$ is a sequence of finite geodesics, we say that $\gamma_n \to \gamma$ for some path $\gamma$ if for all $N > 0$, the first $N$ steps of $\gamma_n$ equal those of $\gamma$ for all $n$ large. Then the diagonal argument presented above works by finding subsequential limits of geodesics.

**Q2**: Under what circumstances do geodesic limits exist? If $\gamma_n$ is the geodesic between 0 and $n\mathbf{e}_1$, does there exist a $\gamma$ such that $\gamma_n \to \gamma$?

This question has particular relevance because a natural way to try to construct geodesics with direction $\theta$ would be to find a limit of finite geodesics to points $x_n$ with $\arg x_n \to \theta$. Moreover, one of the major obstacles to further progress in [34] arises because limits of geodesics are not known to exist, and taking subsequences breaks ergodicity of a certain measure constructed. Some of the results of Chapter 3 are directed toward **Q2**. It is shown there that the a version of the question can be answered on a subclass of infinite subgraphs of $\mathbb{Z}^2$ with boundaries.

### 2.4.2 Merging of Geodesics

Assume for the moment uniform curvature, or any other assumption which would establish a result like Corollary 2.4.2. Knowing that there almost surely exists a geodesic with direction $\theta$ for every $\theta$, one could ask how many typically exist, where again we say that two geodesics are distinct if they have finite symmetric difference. Is it possible that there are more than one such unigeodesic, and does the answer change if we mandate that the geodesics have the same initial vertex?

A major result in this direction is due to Licea and Newman [72], who proved the following result:

**Theorem 2.4.4.** *Let $\mathbb{P}$ be an i.i.d. edge weight distribution on $\mathbb{Z}^2$, and assume that the single-edge marginal $\nu$ has no atoms. Then there exists some set $D \subseteq [0, 2\pi)$ of full Lebesgue measure such that if $\theta \in D$, then there is zero probability that there exist distinct geodesics with direction $\theta$.*

It is not known whether the set $D$ is all of $[0, 2\pi)$, or whether any particular angle lies in $D$. However, $D$ was later shown to have at most a countably infinite complement by Zerner (see [77]).

While Theorem 2.4.4 is rather general, there remains the issue of extending the size of $D$. Also, without a curvature assumption, it is not clear how to find directional geodesics in the first place.

**Q3**: Do any natural methods for constructing infinite geodesics with direction produce families of non-distinct unigeodesics–that is, a family of unigeodesics $\{\gamma_x\}_{x\in\mathbb{Z}^2}$ with $\gamma_x \setminus \gamma_y$ finite for all $x$ and $y$?

Several results presented in Chapter 4 address **Q3**. For first-passage on $\mathbb{Z}^2$, it is shown that merging families of unigeodesics can be constructed as subsequential limits of geodesics to lines, and that these families are directed in the sense described below **Q1**. It is also shown that limiting (not subsequential) unigeodesics exist on certain subgraphs of $\mathbb{Z}^2$ with boundaries, and that these limiting geodesics are a nondistinct family as above.

### Busemann Functions and "Surface View"

A major goal of showing merging or nondistinctness of geodesics, as described in [76], is the study of the microstructure of the surface of the growing region

$$B(t) = \{x : \tau(0,x) \le t\}.$$

One way to study the surface of $B(t)$ for $t$ large would be to define some sort of "point at infinity" and examine the first-passage times from this point to the vertices of $\mathbb{Z}^2$. Using his and his collaborators' results on existence and uniqueness of unigeodesics with direction, Newman [76] showed that for certain sequences $x_n$ with $\arg x_n \to \theta$, the limit

$$\lim_{n\to\infty} \tau(y, x_n) - \tau(z, x_n)$$

exists almost surely. Going in the opposite direction, this type of limit (which could be called a Busemann function) was used to great effect in [51] to establish existence of geodesics. Busemann-type functions are a major tool in the analysis of [10, 34].

### 2.4.3 Bigeodesics

A final major question is

**Q4**: Under what circumstances do bigeodesics exist?

It is clear that bigeodesics must exist for certain $\mathbb{P}$. For instance, if $\mathbb{P}$ is a product measure and $\mathbb{P}(\omega_e = 1) = 1$, then any doubly infinite path which moves only up and to the right is a bigeodesic almost surely. However, if $\mathbb{P}$ has continuous marginals and obeys some moment assumptions, there are some plausible heuristic arguments against bigeodesics [77].

A step toward ruling out bigeodesics appeared in [72] and showed that "directed" bigeodesics almost surely could not exist, at least for a single fixed direction. That is, calling a bigeodesic $(\theta, \theta')$ directed if its two ends have direction $\theta$ and $\theta'$, the result of [72] is that under the assumptions of Theorem 2.4.4 and for fixed $\theta, \theta' \in D$ (the same $D$ as previously), there almost surely exists no $(\theta, \theta')$-directed bigeodesic.

This result leaves the question open in three ways. First is the familiar issue of whether particular directions may be absent from $D$. Second, bigeodesics are only ruled out almost surely for a fixed pair of directions; as there are uncountably many directions, it is not clear a priori that bigeodesics with random direction can exist. Finally, barring a result like uniform curvature of the limit shape, it is not clear that bigeodesics must be directed.

The work of Chapter 4 addresses the problem of bigeodesics on $\mathbb{Z}^2$ from a different point of view. A major effort of this work goes to describing and characterizing a method of constructing unigeodesics via subsequential limits of finite geodesics to lines. It is shown in Section 4.5.4 that the unigeodesics so constructed will not be doubly infinite, ruling out the possibility that this method produces bigeodesics.

It is worth mentioning that the nonexistence of bigeodesics in the half-plane has been established by Wehr and Woo [93]. Moreover, Wehr has shown [91] that on $\mathbb{Z}^d$, if there can (with positive probability) exist at least one bigeodesic, then the number of bigeodesics is almost surely infinite.

**Relation to disordered spin models**

Questions about geodesics have implications for the ground states of disordered ferromagnetic spin models. Examples of such systems include the disordered Ising ferromagnet, a variant of the usual Ising model in which nearest-neighbor couplings take values according to some (positive) distribution. Consider the lattice dual to $\mathbb{Z}^2$, defined by

$$(\mathbb{Z}^2_*, \mathcal{E}^2_*) = (\mathbb{Z}^2, \mathcal{E}^2) + \frac{1}{2}(\mathbf{e}_1 + \mathbf{e}_2) \ ,$$

and define a "spin configuration" $\sigma = (\sigma_x)_{x \in \mathbb{Z}^2_*} \in \{+1, -1\}^{\mathbb{Z}^2_*}$. Let $(J_{x,y})_{\langle x,y \rangle \in \mathcal{E}^2_*}$ have joint distribution $\mu$ which is ergodic and such that $\mu(J_{x,y} > 0) = 1$. For any configuration $\sigma$ and any finite $S \subseteq \mathbb{Z}^2_*$ define the (random) energy functional

$$H_S(\sigma) = - \sum_{\substack{\langle x,y \rangle \in \mathcal{E}^2_* \\ x \in S}} J_{x,y} \sigma_x \sigma_y \ .$$

We will call $\sigma$ a *ground state* for couplings $(J_{x,y})$ if, for each configuration $\tilde{\sigma}$ such that $\tilde{\sigma}_x = \sigma_x$ for all $x$ outside of some finite set, we have

$$H_S(\sigma) \leq H_S(\tilde{\sigma}) \text{ for all finite } S \subseteq \mathbb{Z}^2_* \ .$$

It is an open problem to describe the set of ground states for this ferromagnetic model. In particular it is not known how many ground states there are for a given $(J_{x,y})$, although it is conjectured (see, for instance, [77]) that if $\mu$ is continuous there should be only two almost surely. These two are the constant configurations $\sigma = \pm 1$, which are clearly $\mu$-almost surely ground states. If any nonconstant ground states $\sigma$ exist, they cannot have finite regions of disagreement; that is, there can be no finite $S$ such that $\sigma_x = +1$ for all $x \in S$ and $\sigma_x = -1$ for all $y \in \partial S$ or vice-versa (here, $\partial S$ is the set of sites at $\ell^1$ distance one from $S$). Therefore, any nonconstant ground state must have a two-sided (and circuitless) infinite (original lattice) path of edges dual to bonds $\langle x, y\rangle$ such that $\sigma_x = -\sigma_y$.

We can push forward $\mu$ to a first-passage edge distribution $\mathbb{P}$ on edge-weight configurations of $\mathbb{Z}^2$ by defining $\omega_e = J_{x,y}$, where $\langle x, y\rangle$ is the edge dual to $e$. If such a first-passage configuration had a bigeodesic, then the configuration $\sigma$ which takes the value $+1$ on one side of the bigeodesic and $-1$ on the other would be a nonconstant ground state for the associated spin model.

In addition to the conjectures and partial proofs against the existence of bigeodesics in the first-passage model, there are arguments in the physics literature [40] against the existence of nonconstant ground states. In particular, it is believed that for distributions $\mu$ satisfying some weak conditions (for example, i.i.d. $\mu$ with continuous edge-weight distribution and finite second moment), there should be almost surely no nonconstant ground states.

If one were to argue against the existence of non-constant ground states, one would try to rule out the possibility of constructing such states by standard means. From the point of view of first-passage percolation, it is natural to try to construct bigeodesics by taking limits of finite geodesics to points or lines. In this light, the results presented in subsequent chapters that certain constructions do not produce bigeodesics can be taken as evidence against the existence of nonconstant ground states in the disordered ferromagnet.

## 2.5 Multiple Geodesics Under General Assumptions

In this section, we present an abbreviated form of Hoffman's [51] result (Theorem 2.4.3) showing the existence of four unigeodesics under general assumptions. The perspective here differs somewhat from the original (having been influenced by the work which will appear in [34]), and it also incorporates simplifications originally suggested by M. Damron [33].

We note that the results of [51] actually hold for any $\mathbb{P}$ which is translation-ergodic, has unique passage times (i.e., the passage times of different finite paths are different a.s.), has the symmetries of $\mathbb{Z}^2$, satisfies $\mathbb{E}\omega_e^{2+\alpha} < \infty$ for some $\alpha > 0$, and such that the limit shape $\mathcal{B}$ is bounded. Note that the shape theorem has been proved for this class of measures [19], though there is no simple characterization of the conditions under which boundedness of $\mathcal{B}$ holds [47].

### 2.5.1 Notation

The proofs will require the introduction of some new notation. Because modified versions of some definitions appear in the following chapters, we specify that any notation introduced here expires at the conclusion of this section.

For any point $v \in \partial B$ (that is, such that $g(v) = 1$), let $L_v$ denote the unique tangent line to $\partial B$ at $v$ (if it exists), and define $w(v)$ to some choice of $g$-unit vector parallel to $L_v$. For any $S \subseteq \mathbb{R}^2$, we will define the function $B_S$ on $\mathbb{R}^2 \times \mathbb{R}^2$ by

$$B_S(x, y) = \tau(x, S) - \tau(y, S).$$

We note the relations

$$|B_S(x,y)| \le \tau(x,y); \tag{2.20}$$

$$B_S(x,y) + B_S(y,z) = B_S(x,z). \tag{2.21}$$

The first is a consequence of subadditivity of $\tau$; the second is immediate from the definition of $B_S$.

When approximating passage times, we will need a uniform bound for small increments; to this end, we define

$$\beta = 4 \sup_{\substack{x \in \mathbb{R}^2 \\ \|x\|_\infty \le 4}} \mathbb{E}\tau(0,x) < \infty.$$

We define the lower density (or simply "density" for short) of a sequence $\{a_n\}$ of natural numbers by

$$\liminf_{N\to\infty} \frac{1}{N} \# \left( \{a_n\} \cap \{1, \ldots, N\} \right).$$

### 2.5.2 Translation Invariance

One issue which must be dealt with is the fact that the passage time distribution is not precisely translation-invariant because $\tau(\cdot,\cdot)$ is a function on $\mathbb{R}^2$. That is, while $\tau(x,y)$ has the same distribution as $\tau(x+z, y+z)$ for $z \in \mathbb{Z}^2$, this is not necessarily the case for $z \in \mathbb{R}^2$. The following bound, which says that $\tau$ is "almost translation-invariant", allows us to largely ignore this fact in what follows.

**Lemma 2.5.1.** *For all $v, w, z \in \mathbb{R}^2$, we have*

$$|\mathbb{E}\tau(v,w) - \mathbb{E}\tau(v+z, w+z)| \le 4\beta. \tag{2.22}$$

*Proof.* We will write $v'$ for the closest vertex of $\mathbb{Z}^2$ to $v$ (breaking ties by some deterministic rule) and use analogous definitions for $w'$ and $z'$. Then, by the invariance of the first-passage model under translations by vectors in $\mathbb{Z}^2$,

$$\mathbb{E}\tau(v', w') - \mathbb{E}\tau(v' + z', w' + z') = 0.$$

Using this fact and the triangle inequality for $\tau$, we see that the left-hand side of (2.22) is bounded above by

$$\mathbb{E}\left[\tau(v, v') + \tau(w, w') + \tau(v' - z', v - z) + \tau(w' - z', w - z)\right].$$

Each term in the above is the passage time between two points of $\mathbb{R}^2$ which are at Euclidean distance at most four. Therefore, each term is bounded above by $\beta$. This proves the bound (2.22). $\square$

### 2.5.3 Core of the proof

Given vertices $\{x_i\}_{i=1}^N \subseteq \mathbb{Z}^2$, we define $G_i$ to be the (random) set of points which are reached first from $x_i$; that is,

$$G_i(\omega) = \{z \in \mathbb{Z}^2 : B_z(x_i, x_j) < 0 \text{ for all } j \neq i\}. \tag{2.23}$$

The sets $G_i$ and $G_j$ are disjoint if $i \neq j$. Note that if $z \in G_i$ and $\gamma$ is the geodesic from $x_i$ to $z$, then $\gamma$ contains only vertices of $G_i$. Indeed, if $\gamma$ contained some $y$ such that $B_y(x_i, x_j) \geq 0$, then we would have

$$\tau(x_i, z) = \tau(x_i, y) + \tau(y, z) \geq \tau(x_j, y) + \tau(y, z) \geq \tau(x_j, z)$$

and so $z \notin G_i$.

If there exists an infinite sequence $\{z_m\} \subseteq G_i$ , consider the sequence of geodesics $\{\gamma_i(m)\}_m$ which connect $x_i$ to the members of $\{z_m\}$. As before, we can use a "diagonal argument" to find some infinite geodesic $\gamma_i = \lim_k \gamma_i(m_k)$. By the arguments of the preceding paragraph, $\gamma_i \subseteq G_i$. We have therefore proved the following:

**Lemma 2.5.2.** *Let $\{x_i\}_{i=1}^N$ denote $N$ distinct vertices of $\mathbb{Z}^2$, with associated random sets $G_i$ defined in* (2.23). *If for a configuration $\omega$ of edge weights, $|G_i(\omega)| = \infty$ for $i = 1, \ldots, N$, then there exist at least $N$ disjoint unigeodesics in the configuration $\omega$.*

Hoffman's strategy is to control the growth of $B_{nx_i}(x_i, x_j)$ as $n$ becomes large to show that every $G_i$ can be simultaneously infinite with positive probability. For $n$ large, one would imagine that $B_{n\mathbf{e}_1}(-M\mathbf{e}_1, M\mathbf{e}_1)$ would typically be nonzero of order $M$. In fact a weak version of this is true, and will be proved in the next section. A similar argument holds for $B_{-n\mathbf{e}_1}(M\mathbf{e}_1, -M\mathbf{e}_1)$ This allows the choice $x_1 = M\mathbf{e}_1$, $x_2 = -M\mathbf{e}_1$, and shows that with positive probability both $G_1$ and $G_2$ are infinite and so there exist at least two geodesics.

To find a set $\{x_i\}$ with more than two elements such that each $G_i$, one needs to control $B_{nv}(Mv, Mv')$, where $v'$ is not a multiple of $v$. Hoffman does this by choosing $v$ as a point of differentiability of $\partial B$ and $w$ as a unit tangent vector at $v$, and showing that $B_{nv}(0, Mw)$ grows sublinearly in $M$ because $\tau(Mw, nv)$ does. This is described in Theorem 2.5.7 and the preceding discussion.

### 2.5.4 Growth of $B_{n,v}$

A crucial idea in Hoffman's work, whose explication has already been years in development [41, 48, 52], is that if $n \gg M > 0$, the point $nv$ should be closer to $Mv$ than the origin in the first-passage metric. That is, $\tau(0, nv) - \tau(Mv, nv) \sim M$ in some sense. While this relation appears very natural, its proof requires an averaging "trick," which gives only a weaker averaged result. We begin with a lemma in terms of expectation.

**Lemma 2.5.3.** *For any $v \in \partial B$ and $\varepsilon > 0$, there exists an $M_0 = M_0(\varepsilon, v)$ such that for all $r \in \mathbb{R}^2$ and for each $M > M_0$, the density of $n$ such that*

$$(1+\varepsilon)M \geq \mathbb{E}B_{nv}(r, Mv) \geq (1-\varepsilon)M \tag{2.24}$$

*is at least $1 - \varepsilon$.*

*Proof.* We will restrict ourselves to $r = 0$ for the first portion of the proof; we will at the end describe the extension to general $r$. To show the first inequality, we can use (2.20) to see that

$$B_{nv}(0, Mv) \leq \tau(0, Mv)$$

for all $n$. By the shape theorem and the fact that $g(v) = 1$, there exists an $M_0$ such that if $M > M_0$, $\mathbb{E}\tau(0, Mv) \leq (1+\varepsilon)M$, proving the first inequality.

We will now show the second bound. Choose some $k$ depending on $n$ and $M$ such that

$$kM \leq n < (k+1)M.$$

By the shape theorem, there exists an $M_1$ such that if $M > M_1$ and $n \geq 2M(1-\varepsilon)\varepsilon^{-1}$,

$$(1-\varepsilon)(k+1)M \leq (1-\varepsilon/2)n \leq \mathbb{E}\tau(0, nv). \tag{2.25}$$

We now decompose the passage time, using the invariance of $\mathbb{E}\tau(\cdot,\cdot)$ under shifts by elements of $\mathbb{Z}^2$:

$$\begin{aligned}
\mathbb{E}\tau(0, nv) &= \sum_{j=0}^{k-1} \left[\tau(jMv, nv) - \tau((j+1)Mv, nv)\right] + \mathbb{E}\tau(kMv, nv) \\
&\leq 2k\beta + \sum_{j=0}^{k-1} \mathbb{E}\left[\tau(0, (n-jM)v) - \tau(Mv, (n-jM)v)\right] + \mathbb{E}\tau(kMv, nv) \\
&\leq Ck\beta + \sum_{j=0}^{k-1} \mathbb{E}B_{(n-kM+jM)v}(0, Mv).
\end{aligned} \tag{2.26}$$

The constant $C$ is the result of both the $2k$ of the previous line and the fact that $\mathbb{E}\tau(kMv, nv) \leq \mathbb{E}\tau(0, (n-kM)v) + \beta$, where $\mathbb{E}\tau(0, (n-kM)v)$ can be bounded by a multiple of $k\beta$ since $(n-kM)v/k$ is uniformly bounded in Euclidean norm.

We can now find $M_2 > M_1$ such that if $M > M_2$, then $Ck\beta < \varepsilon(k+1)M$. Putting this together with (2.26) gives

$$(1-2\varepsilon)(k+1)M \leq \sum_{j=0}^{k-1} \mathbb{E}B_{(n-kM+jM)v}(0, Mv). \tag{2.27}$$

As in the proof of the first inequality of the lemma, we can find (using the shape theorem) an $M_3 > M_2$ such that for $M > M_3$ we have $\mathbb{E}\tau(0, Mv) \leq (1+\varepsilon)M$. Using this and the bound $B_S(0, Mv) \leq \tau(0, Mv)$, we see that if $M > M_3$, each term in (2.27) is bounded above by $(1+\varepsilon)M$. From here, algebraic manipulation implies that the number of terms in (2.27) which are at least as large as $(1-\sqrt{\varepsilon})M$ is at least $(1-4\sqrt{\varepsilon})k$.

Because $\varepsilon$ was arbitrary, we can apply the preceding with $\varepsilon' = \varepsilon^2/16$ to see that the number of terms of the sequence

$$\mathbb{E}B_{(n-kM+M)v}(0, Mv), \mathbb{E}B_{(n-kM+2M)v}(0, Mv), \ \ldots\ , \mathbb{E}B_{nv}(0, Mv) \tag{2.28}$$

which are at least $(1-\varepsilon)M$ is at least $(1-\varepsilon)k$. From here the lemma is proved once we take $n \to \infty$ in the appropriate way; we now will see how to do this. Setting $j(M, n) = n - kM$, we can restate the result about (2.28) as saying that the number of terms of the sequence

$$\mathbb{E}B_{(j+M)v}(0, Mv), \mathbb{E}B_{(j+2M)v}(0, Mv), \ \ldots\ , \mathbb{E}B_{(j+kM)v}(0, Mv)$$

which are at least $(1-\varepsilon)M$ is bounded below by $(1-\varepsilon)k$. Taking $n$ to infinity along a sequence $\{n_i\}$ such that $n_i \mod M = j$, we see that the density of terms of the sequence $\{\mathbb{E}B_{(j+lM)v}(0,Mv)\}_l$ which are at least $(1-\varepsilon)M$ is bounded below by $(1-\varepsilon)$. Applying the result for $j = 0, \ldots, M-1$ completes the proof in the case $r = 0$.

Note that by the translation-invariant properties of the model, the result of the lemma for $r \in \mathbb{Z}^2$ follows immediately from the case $r = 0$. For general $r \in \mathbb{R}^2$, we write $r = r' + \delta_r$, where $r' \in \mathbb{Z}^2$ and $\delta_r \in [0,1)^2$. Using the fact that the lemma holds for $r'$, in conjunction with the bound

$$|\mathbb{E}(B_{nv}(r, Mv) - B_{nv}(r', Mv))| \leq \mathbb{E}\tau(r, r')$$

and (2.22) proves the result. □

We will extend Lemma 2.5.3 from a statement about expectations to a statement that the bound (2.24) holds with high probability. To do this, we will need a lemma about uniform integrability of $\tau$.

**Lemma 2.5.4.** *For all $z \in \mathbb{R}^2$ and all $\varepsilon > 0$, there exist $M_0$ and $\delta$ such that*

$$\mathbb{E}\left(\tau(0, Mz)\mathbf{1}_A\right) \leq \varepsilon M$$

*For all $M > M_0$ and all events $A$ with $\mathbb{P}(A) < \delta$.*

*Proof.* By the bound in (2.22), it suffices to consider integer $M$. As in the preceding, the major difficulty will be with $z \notin \mathbb{Z}^2$. To deal with this, we consider the space

$$\Omega^* = [0,\infty)^{\mathcal{E}^2} \times [0,1)^2$$

under the probability measure $\mathbb{P}^* = \mathbb{P} \times \lambda$, where $\lambda$ is the uniform (Lebesgue) measure on $[0,1)^2$. All of the random variables and events (for instance, $\tau(x,y)$) defined on the first-passage space $\Omega = [0,\infty)^{\mathcal{E}^2}$ may be considered as random variables or events on $\Omega^*$ which do not depend on the coordinates in $[0,1)^2$.

For an element $(\omega, a_1, b_1) \in \Omega^*$, let $w = (a_1, b_1) + z$, and decompose $w$ uniquely as $w = w' + (a_2, b_2)$ where $w' \in \mathbb{Z}^2$ and $(a_2, b_2) \in [0, 1)^2$. We will define the translation operator $T_z$ on $\Omega^*$ by $T_z(\omega, a_1, b_1) = (T_{w'}\omega, a_2, b_2)$, where $(T_{w'}\omega)_e = \omega_{e-w'}$. The measure $\mathbb{P}^*$ is stationary under the action of $T_z$. Therefore, if we fix some starting point $(a_1, b_1)$ and write

$$\begin{aligned}\frac{1}{M}\tau((a_1, b_1), Mz + (a_1, b_1)) &\le \frac{1}{M}\sum_{j=0}^{M-1}\tau(jz + (a_1, b_1), (j+1)z + (a_1, b_1)) \\ &= \frac{1}{M}\sum_{j=0}^{M-1} h(T_z^{(j)}(\omega, a_1, b_1)) \qquad (2.29)\end{aligned}$$

(where $h(\omega, a, b) = \tau((a, b), (a, b) + z)$ evaluated in configuration $\omega$), then (2.29) converges almost surely and in $L^1(\mathbb{P}^*)$ to some random variable $\overline{h} : \Omega^* \to \mathbb{R}$ by the ergodic theorem. Let $\mathbb{E}^*$ denote expectation under $\mathbb{P}^*$.

Recalling that $A \subseteq \Omega$ may be regarded as the subset $A \times [0, 1)^2 \subseteq \Omega^*$, we write

$$\int_{(a,b)\in[0,1)^2} \mathbb{E}\left[\tau((a, b), (a, b) + Mz)\mathbf{1}_A\right] \mathrm{d}\lambda \le \frac{1}{M}\mathbb{E}^*\left(\sum_{j=0}^{M-1} h(T_z^{(j)}(\omega, a, b)\mathbf{1}_A\right). \qquad (2.30)$$

If $\varepsilon > 0$, we can use the fact that $\overline{h}$ has finite expected value to find some $\delta > 0$ such that

$$\mathbb{E}^*(\overline{h}\mathbf{1}_B) < \varepsilon/4$$

for all $B \subseteq \Omega^*$ satisfying $\mathbb{P}^*(B) < \delta$. Then, assuming that $\mathbb{P}^*(A) < \delta$, we can bound above the right-hand side of (2.30) by

$$\begin{aligned}\frac{1}{M}\mathbb{E}^*\left(\sum_{j=0}^{M-1} h(T_z^{(j)}(\omega, a, b)\mathbf{1}_A\right) &\le \left|\mathbb{E}^*\left(\frac{1}{M}\sum_{j=0}^{M-1}(h(T_z^{(j)}(\omega, a, b) - \overline{h})\mathbf{1}_A\right)\right| + \left|\mathbb{E}^*(\overline{h}\mathbf{1}_A)\right| \\ &\le \left|\mathbb{E}^*\left(\frac{1}{M}\sum_{j=0}^{M-1}(h(T_z^{(j)}(\omega, a, b) - \overline{h}\right)\right| + \varepsilon/4 \\ &< \varepsilon/2,\end{aligned}$$

for $M$ large enough, by the $L^1$ convergence to $\overline{h}$.

Now we may note that with $\mathbb{P}$-probability one, for all $(a, b) \in [0, 1)^2$,

$$|\tau(0, Mz) - \tau((a, b), (a, b) + Mz)| \le \tau(0, (a, b)) + \tau(Mz, Mz + (a, b)).$$

In particular, we see using Lemma 2.22 that

$$\frac{1}{M} |\mathbb{E}\tau(0, Mz)\mathbf{1}_A - \mathbb{E}\tau((a, b), (a, b) + Mz)\mathbf{1}_A| \le 2\beta/M < \varepsilon/2,$$

for $M$ large enough. Putting this together with the upper bound of $\varepsilon/2$ for the left-hand side of (2.30) proves the claim.

**Theorem 2.5.5.** *For any $v \in \partial B$ and $\varepsilon > 0$, there exists an $M_0 = M_0(\varepsilon, v)$ such that for all $r \in \mathbb{R}^2$ and for each $M > M_0$, the density of $n$ such that*

$$\mathbb{P}\left(M(1-\varepsilon) < B_{nv}(r, r + Mv) < M(1+\varepsilon)\right) > 1 - \varepsilon \tag{2.31}$$

*is at least $1 - \varepsilon$.*

We will let $\varepsilon' > 0$ be fixed (to be chosen at the end of the proof). Let the event

$$E_n(M) := \{B_{nv}(0, Mv) > (1 + \varepsilon')M\}.$$

We note that, since $B_{nv}(\cdot, \cdot) \le \tau(\cdot, \cdot)$, we have

$$\mathbb{P}\left(E_n(M)\right) \le \mathbb{P}\left(\tau(0, Mv) > (1 + \varepsilon')M\right) \to 0 \tag{2.32}$$

as $M \to \infty$. Use (2.32) to find some $M_1$ such that, if $M > M_1$, $\mathbb{P}(E_n(M)) < \delta$ for all $n$. At this point, we have proved the second inequality of (2.31); it remains to prove the first. Using Lemma 2.5.4, we can find some $\delta \in (0, \varepsilon')$ such that $\mathbb{E}\left(\tau(0, Mv)\mathbf{1_A}\right) < \varepsilon' M$ whenever $\mathbb{P}(A) < \delta$. We will henceforth suppress the argument of $E_n$.

Find an $M_0 > M_1$ such that (2.24) holds (but with $\varepsilon$ replaced with $\varepsilon'$) for all $n$ in some set $\Xi(\varepsilon', M)$ of density at least $1 - \varepsilon'$. For $M > M_0$ and $n \in \Xi(\varepsilon', M)$, we may write

$$\begin{aligned}(1 - \varepsilon')M \leq \mathbb{E}B_{nv}(0, Mv) &= \mathbb{E}[B_{nv}(0, Mv)(\mathbf{1}_{E_n} + 1 - \mathbf{1}_{E_n})] \\ &\leq \varepsilon' M + \mathbb{E}[B_{nv}(0, Mv)\mathbf{1}_{E_n^c}],\end{aligned}$$

where the superscript $c$ denotes the usual set complement. In particular, $\mathbb{E}[B_{nv}(0, Mv)\mathbf{1}_{E_n^c}] > (1 - 2\varepsilon')M$. The following claim can be verified with a simple computation.

**Claim 2.5.6.** *Let $X$ be a random variable which is a.s. bounded above by $(1 + \varepsilon')$, and assume that $\mathbb{E}(X) > (1 - 2\varepsilon')$. Then if $\varepsilon'$ is sufficiently small,*

$$\mathbb{P}\left(X \geq (1 - \sqrt{\varepsilon'})\right) \geq 1 - 3\sqrt{\varepsilon'}.$$

As has been seen, the random variable $X := M^{-1}B_{nv}(0, Mv)\mathbf{1}_{E_n^c}$ satisfies the hypotheses of Claim 2.5.6 for $M > M_0$ and $n \in \Xi(\varepsilon', M)$. We now choose $\varepsilon' = \varepsilon^2/9$; noting that $X \geq (1 - \varepsilon)$ implies $B_{nv}(0, Mv) \geq M(1 - \varepsilon)$ completes the proof. □

### 2.5.5 $B_{nv}$ in direction $w$

Having found bounds on the growth of $B_{nv}(0, Mv)$ as $M$ becomes large, we now will study the growth of $B_{nv}(0, Mw)$. Recall the fact that $w$ is a unit tangent vector at the point $v \in \partial B$, which is a point of differentiability of $\partial B$. One immediate consequence of this is that

$$\lim_{a \to 0} \frac{1 - g(v - aw)}{a} = 0. \tag{2.33}$$

Now, $\tau(0, nv) \sim ng(v) = n$ for $n$ large. Similarly,

$$\tau(Mw, nv) \sim ng(v - Mw/n) = n(1 + o(M/n)),$$

by (2.33). In particular, we expect that in the regime $n \gg M$,

$$B_{nv}(0, Mw) = \tau(0, nv) - \tau(Mw, nv) \sim o(M).$$

Such a sublinear growth bound will be derived rigorously, though as in the previous section, we are limited to proving bounds for values of $n$ in sets of high density.

**Theorem 2.5.7.** *Let $\varepsilon > 0$. There exists some $M_0$ such that, for all $r \in \mathbb{R}^2$ and $M > M_0$, the density of $n$ such that*

$$\mathbb{P}\left(|B_{nv}(r, r + Mw)| < \varepsilon M\right) > 1 - \varepsilon$$

*is at least $1 - \varepsilon$.*

*Proof.* Note that $B_{nv}(r, r + Mw) = -B_{nv}(r + Mw, r)$, and note that $-w$ is also a unit tangent vector of $\partial B$ at $v$. Therefore, we need only prove the bounds of the form $B_{nv}(r, r + Mw) < \varepsilon M$.

Similarly to the other proofs of this section, we may assume $r = 0$. Let $\varepsilon > 0$. We have, for $a > 0$ and $M'$ arbitrary,

$$\begin{aligned} B_{nv}(0, aM'w) &= B_{nv}(-M'v, aM'w) - B_{nv}(-M'v, 0) \\ &\leq \tau(-M'v, aM'w) - B_{nv}(-M'v, 0). \end{aligned} \tag{2.34}$$

By the shape theorem, there is some $M_1(a, \varepsilon)$ such that $M' > M_1$ implies that

$$\mathbb{P}\left(\tau(-M'v, aM'w) < (1 + \varepsilon a/4)M'g(v + aw)\right) \geq 1 - \varepsilon/2. \tag{2.35}$$

By (2.33), we have $|g(v + aw) - 1| < \varepsilon a/4$ for all $a < a_0(\varepsilon) < 1$; we will fix such a value of $a$. Using this in (2.35) yields, for $M' > M_1$,

$$\mathbb{P}\left(\tau(-M'v, aM'w) < (1 + \varepsilon a/2)M'\right) \geq 1 - \varepsilon/2. \tag{2.36}$$

Using Theorem 2.5.5, we may also find an $M_0 > M_1$ such that if $M' > M_0$, then for all $n$ in some set of density at least $1 - \varepsilon$,

$$\mathbb{P}\left((1 - \varepsilon a/2)M' < B_{nv}(-M'v, 0)\right) > 1 - \varepsilon/2.$$

This fact and (2.36) applied to (2.34) imply that for $M' > M_0$, the set of $n$ such that

$$B_{nv}(0, aM'w) < \varepsilon aM'$$

has density at least $1 - \varepsilon$. Setting $M = aM'$ completes the proof. □

The result of Theorem 2.5.5, which is essentially a "one-dimensional bound," would suffice to show the existence of at least two geodesics. To establish that at least four geodesics exist, the following theorem–which incorporates the additional information of Theorem 2.5.7–will prove useful.

**Theorem 2.5.8.** *Fix $\varepsilon > 0$. Let $v \in \partial B$ be a point at which $\partial B$ is differentiable, and let $w$ denote the unit tangent vector at $v$. Let $y \in \mathbb{R}^2$ be written (uniquely) as $y = (1-t)v + sw$, for $s, t \in \mathbb{R}$. Then there exists some $M_0$ such that, for all $M > M_0$, the density of $n$ such that*

$$\mathbb{P}\left(B_{nv}(My, Mv) > M(t - \varepsilon)\right) > (1 - \varepsilon)$$

*is at least $1 - \varepsilon$.*

*Proof.* Begin by writing

$$B_{nv}(My, Mv) = B_{nv}(My, My + tv) + B_{nv}(My + tv, Mv). \tag{2.37}$$

Using Theorem 2.5.5, we can find some $M_0'$ such that $M > M_0'$ implies the density of $n$ such that

$$\mathbb{P}\left(B_{nv}(My, My + tv) > Mt(1 - \varepsilon/2)\right) > 1 - \varepsilon/2$$

is at least $1 - \varepsilon/2$. Similarly, by Theorem 2.5.7, we can find $M_0 > M_0'$ such that $M > M_0$ implies that the density of $n$ such that

$$\mathbb{P}\left(B_{nv}(My + tv, Mv) > -M\varepsilon/2)\right) > 1 - \varepsilon/2$$

is at least $1 - \varepsilon/2$. Putting these two bounds together with (2.37) completes the proof. □

### 2.5.6 Proof of Main Theorems

We will demonstrate Theorem 2.4.3 by showing that the condition of Lemma 2.5.2 is satisfied with high probability.

*Proof.* Let $k \geq 0$ be an integer which is at most the number of sides of $B$. There exists a set of points $x_1 \ldots, x_k \in \partial B$ such that there is a unique tangent line $L_i$ of $\partial B$ at $x_i$, and such that $L_i \neq L_j$ for $i \neq j$. Let $x_i \neq x_j$ be arbitrary members of this sequence, and decompose $x_j = (1-t)x_i + sw(x_i)$ as in the statement of Theorem 2.5.8. Because $L_j \neq L_i$, we have $x_j \notin L_i$ and so $t \neq 1$.

Applying Theorem 2.5.8, there exists some $M_0 > 0$ and $c > 0$ such that if $M > M_0$, the density of $n$ such that

$$\mathbb{P}\left(B_{nx_i}(Mx_j, Mx_i) > cM \text{ for all } i \neq j\right) > 1 - \varepsilon \tag{2.38}$$

is at least $1 - \varepsilon$. The condition of Lemma 2.5.2 now follows for the collection $\{Mx_i\}$.

□

In particular, since the limit shape has at least four sides by symmetry, there can be at least four unigeodesics. Hoffman is also able to show a form of directional concentration of long finite geodesics, insofar as there are favorable regions through which a large density of them tend to pass through. Hoffman's methods are a major impetus behind the work presented in the following chapters.

# Chapter 3

# Geodesic Graphs and Busemann Distributions

It is worth recalling the main role of the Busemann function in Hoffman's [51] results. Given a set $S$ and two vertices $x$ and $y$, let $\gamma$ denote the geodesic from $x$ to $S$ (if it exists). The Busemann function $B_S(x, y)$ has the property that

$$y \in \gamma \Longrightarrow B_S(x, y) = \tau(x, y).$$

If one can show that $B_S(x, y) \neq \tau(x, y)$, then one rules out the existence of a geodesic from $x$ to $S$ which passes through $y$.

So far, this technique has been applied in a local or "one geodesic at a time" manner to show the existence of geodesics. In what follows, we will take a more global picture by looking at the interplay between Busemann functions and "geodesic graphs" which are in some sense global collections of geodesics. One motivating example for such a geodesic graph is the tree $\mathcal{T}(x)$ defined in the preceding chapter. However, as we will see, our techniques will be most powerful when the graph's distribution inherits some translation invariance from $(V, E)$ (note $\mathcal{T}(x)$, being rooted at $x$, has no translation symmetry).

## 3.1 Assumptions and Definitions

In what follows, we will consider only first-passage percolation on a connected subgraph of $\mathbb{Z}^2$, which we will denote by $(V, E)$ (in the case that $(V, E) = (\mathbb{Z}^2, \mathcal{E}^2)$, we will often simply refer to the graph as $\mathbb{Z}^2$). Recall the definition of the model from Section 2.2.1. If $S$ is a subset of $\mathbb{R}^2$, we define the Busemann function

$$B_S(x, y) = \tau(x, S) - \tau(y, S). \tag{3.1}$$

We will call a graph of the form

$$(\mathbb{Z}^2, \mathcal{E}^2) \cap \{(x_1, x_2) : x_2 \geq m\}$$

for $m \in \mathbb{Z}$ an "upper half plane."

We denote the standard orthonormal basis vectors for $\mathbb{R}^2$ by $\mathbf{e}_1$ and $\mathbf{e}_2$. If our graph $(V, E)$ is either $(\mathbb{Z}^2, \mathcal{E}^2)$ or an upper half-plane, we can define translations by unit vectors. The translation operators $T_{\mathbf{e}_i}$, $i = 1, 2$ act on a configuration $\omega$ as follows: $(T_{\mathbf{e}_i}(\omega))_{e'} = \omega_{e'+\mathbf{e}_i}$. If $(V, E) = (\mathbb{Z}^2, \mathcal{E}^2)$ and $\mathbb{P}$ is invariant under these translations, the passage times have a certain translation-covariance: for $i = 1, 2$,

$$\tau(x, S)(T_{\mathbf{e}_i}\omega) = \tau(x + \mathbf{e}_i, S + \mathbf{e}_i)(\omega) \ , \tag{3.2}$$

where $S + \mathbf{e}_i = \{x + \mathbf{e}_i : x \in S\}$. In the case of the half-plane, (3.2) holds for $i = 1$ (assuming that $\mathbb{P}$ is translation-invariant). On $\mathbb{Z}^2$, we analogously define for $x \in \mathbb{Z}^2$, the operators

$$(T_x\omega)_e = \omega_{e+x};$$

this definition is also in effect for half-planes, but with $x = (x_1, x_2)$, where $x_2 \geq 0$.

There are several different classes of assumptions under which our theorems will be valid. These will be assumptions on both the edge weight distribution $\mathbb{P}$ and the geometry of $(V, E)$. In order to save ourselves from having to describe these in detail multiple times, we will fix a few common classes of assumptions which will be useful.

### 3.1.1 Assumptions on the full plane

Two classes of assumptions we commonly consider will be called **A1** and **A2**. In both cases, $(V, E) = (\mathbb{Z}^2, \mathcal{E}^2)$, with certain other requirements depending on whether $\mathbb{P}$ is i.i.d. or merely ergodic. In the first case, $\mathbb{P}$ is taken i.i.d.:

**A1** First-passage percolation on $\mathbb{Z}^2$. Here $\mathbb{P}$ is a product measure whose common distribution satisfies the criterion of Cox and Durrett [31]: if $e_1, \ldots, e_4$ are the four edges touching the origin,

$$\mathbb{E}\left[\min_{i=1,\ldots,4} \omega_{e_i}\right]^2 < \infty \ . \tag{3.3}$$

Furthermore we assume $\mathbb{P}(\omega_e = 0) < p_c = 1/2$, the bond percolation threshold for $\mathbb{Z}^2$.

Condition (3.3) is implied by, for example, the assumption $\mathbb{E}\omega_e < \infty$.

The other assumption is on distributions that are only translation-invariant.

**A2** First-passage percolation on $\mathbb{Z}^2$. $\mathbb{P}$ is a measure satisfying the conditions of Hoffman [51]:

(a) $\mathbb{P}$ is ergodic with respect to translations of $\mathbb{Z}^2$;

(b) $\mathbb{P}$ has all the symmetries of $\mathbb{Z}^2$;

(c) $\mathbb{E}\omega_e^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$;

(d) the limit shape $\mathcal{B}$ for $\mathbb{P}$ is bounded (see the next paragraph).

We have seen that **A1** suffice to show the shape theorem, Theorem 2.2.2, with a limiting norm $g$ which is not identically zero. Recall also that, in proving the shape theorem, we showed that

$$\mathbb{E}(\tau(x,y))^2 < \infty \quad \text{for all } x, y. \tag{3.4}$$

Under **A2**, $\tau$ has a finite second moment as a consequence of our moment assumption on $\omega_e$.

Under **A2**, the shape theorem has also been proved [19]; the final assumption of **A2** is tantamount to requiring that $g$ not be identically zero. The statement that $\mathcal{B}$ has nonempty interior is not explicitly proved in [19] but follows from the maximal lemma stated there.

Since the limit shape is bounded and has nonempty interior, there are constants $0 < C_1, C_2 < \infty$ such that

$$C_1\|x\|_2 \le g(x) \le C_2\|x\|_2 \text{ for all } x \in \mathbb{R}^2 . \tag{3.5}$$

We say a measure $\mathbb{P}$ on $\Omega$ *admits geodesics* if

$$\mathbb{P}(\exists \text{ a geodesic } \gamma : x \rightsquigarrow y) = 1 \text{ for all } x, y \in V .$$

As discussed in Section 2.2.2, any $\mathbb{P}$ which satisfies the shape theorem with $g$ not identically zero admits geodesics. In particular, measures satisfying **A1** and **A2** admit geodesics.

For some of our results, we will need a slight strengthening of **A1** and **A2** which ensures that there is at most one finite geodesic between a pair of vertices.

**A1'** $\mathbb{P}$ satisfies **A1** and the common distribution of $\omega_e$ is continuous.

**A2'** $\mathbb{P}$ satisfies **A2** and $\mathbb{P}$ has unique passage times.

The phrase "unique passage times" means that for any pair of edge-nonempty distinct paths $\gamma$ and $\gamma'$, $\mathbb{P}(\tau(\gamma) = \tau(\gamma')) = 0$.

### 3.1.2 Assumptions on fractional planes

The graphs $(V, E)$ admitted by these assumptions will be infinite connected domains (of $\mathbb{Z}^2$) with infinite complement. That is, $V \subseteq \mathbb{Z}^2$ is infinite and connected (in $(\mathbb{Z}^2, \mathcal{E}^2)$), and its complement $V^c$ is also infinite and connected. As shorthand, we will call such a graph $(V, E)$ a *fractional plane*.

We give an alternate characterization of the class of fractional planes. We will need the graph dual to the square lattice, the vertex set of which is $(\mathbb{Z}^2)^* = \mathbb{Z}^2 + (1/2, 1/2)$ and the edge set of which is $(\mathcal{E}^2)^* = \mathcal{E}^2 + (1/2, 1/2)$. The edge $e^*$ is said to be dual to $e \in \mathcal{E}^2$ if it bisects $e$. We prove in Section 3.9.1 that there exists some path of dual edges $\Gamma = (e_i^*)_{i \in \mathbb{Z}}$ which does not (vertex) self-intersect and such that $(V, E)$ is one of the two components of the graph formed from $(\mathbb{Z}^2, \mathcal{E}^2)$ by removing the edges $(e_i)$ dual to those edges $(e_i^*)$. Let $v_i$ be the endpoint of $e_i$ that lies in $V$. Note that while $\Gamma$ is not self-intersecting, a particular $v_i$ may appear multiple times (at most 3 times).

Recall the definition of the phrase "admits geodesics" from the assumptions on full-plane measures. Some of our results on fractional planes $(V, E)$ make the assumption that the edge weight distribution $\mathbb{P}$ admits geodesics. We show that i.i.d. $\mathbb{P}$ admit geodesics in Section 3.9.2 for general fractional planes $(V, E)$.

## 3.2 Geodesic Graphs to $S$

As stated previously, we would ultimately like to construct a geodesic graph which has some sort of translation-invariance (assuming $(V, E)$ does). The first step in this program is to construct a graph of geodesics to some set $S$, which we will later "translate to infinity." In what follows, we will assume without further comment that $S \subseteq \mathbb{R}^2$ contains at least one point $z$ such that, for some $x \in V$, we have $z \in x + [-1/2, 1/2)^2$. This will ensure that the passage time $\tau(\cdot, S)$ is always defined. We will also define the set of directed edges

$$\vec{E} = \{\langle x, y\rangle : \quad \{x, y\} \in E\}.$$

Assume that $\mathbb{P}$ admits geodesics. For any $S \subseteq V$ and configuration $\omega$, we denote the set of edges in all geodesics from a point $v \in V$ to $S$ as $G_S(v)$. We regard each geodesic in $G_S(v)$ as a directed path, giving orientation $\langle x, y\rangle$ to an edge if $\tau(x, S) \geq \tau(y, S)$ (the direction in which the edge is crossed), and set $\vec{G}_S(v)$ to be the union of these directed edges. Let $\mathbb{G}_S(\omega)$ be the directed graph induced by the edges in $\cup_v \vec{G}_S(v)$. Last, define the configuration $\eta_S(\omega)$ of directed edges by

$$\eta_S(\omega)(\langle x, y\rangle) = \begin{cases} 1 & \text{if } \langle x, y\rangle \in \vec{G}_S(v) \text{ for some } v \\ 0 & \text{otherwise} \end{cases}.$$

For $S \subseteq \mathbb{R}^2$ we define $\eta_S(\omega)$ and $\mathbb{G}_S(\omega)$ using the geodesics to the set $\hat{S}$, where

$$\hat{S} = \{y \in V : y + [-1/2, 1/2)^2 \cap S \neq \varnothing\}. \tag{3.6}$$

**Proposition 3.2.1.** *Let $S \subseteq \mathbb{R}^2$. The graph $\mathbb{G}_S$ and the collection $(\eta_S)$ satisfy the following properties $\mathbb{P}$-almost surely.*

1. *Every finite directed path is a geodesic. It is a subpath of a geodesic ending in $S$.*

2. *If there is a directed path from $x$ to $y$ in $\mathbb{G}_S$ then $B_S(x, y) = \tau(x, y)$.*

*3. Assume that* $(V,E) = (\mathbb{Z}^2, \mathcal{E}^2)$. *If* $\mathbb{P}$ *is translation-invariant, then for* $i = 1, 2$,

$$\eta_S(e)(T_{\mathbf{e}_i}\omega) = \eta_{S+\mathbf{e}_i}(e + \mathbf{e}_i)(\omega) \ . \tag{3.7}$$

*Therefore the finite dimensional distributions of* $\eta_S$ *obey a translation invariance:*

$$(\eta_S(e)) \underset{d}{=} (\eta_{S+\mathbf{e}_i}(e + \mathbf{e}_i)) \ .$$

*In the case that* $(V,E)$ *is an upper half plane, the above holds for* $i = 1$ *under the assumption that* $\mathbb{P}$ *is invariant under* $T_{\mathbf{e}_1}$.

*Proof.* The third property follows from translation covariance of passage times (3.2). The second property follows from the first and Proposition 3.2.4.

To prove the first, let $\gamma$ be a directed path in $\mathbb{G}_S$ and write the edges of $\gamma$ in order as $e_1, \ldots, e_n$. Write $J \subseteq \{1, \ldots, n\}$ for the set of $k$ such that the path $\gamma_k$ induced by $e_1, \ldots, e_k$ is a subpath of a geodesic from some vertex to $S$. We will show that $n \in J$. By construction of $\mathbb{G}_S$, the edge $e_1$ is in a geodesic from some point to $S$, so $1 \in J$. Now suppose that $k \in J$ for some $k < n$; we will show that $k+1 \in J$. Take $\sigma$ to be a geodesic from a point $z$ to $S$ which contains $\gamma_k$ as a subpath. Write $\sigma'$ for the portion of the path from $z$ to the far endpoint $v_k$ of $e_k$ (the vertex to which $e_k$ points). The edge $e_{k+1}$ is also in $\mathbb{G}_S$ so it is in a geodesic from some point to $S$. If we write $\hat{\sigma}$ for the piece of this geodesic from $v_k$ of $e_k$ to $S$, we claim that the concatenation of $\sigma'$ with $\hat{\sigma}$ is a geodesic from $z$ to $S$. To see this, write $\tau_{\tilde{\gamma}}$ for the passage time along a path $\tilde{\gamma}$:

$$\tau(z,S) = \tau_\sigma(z, v_k) + \tau_\sigma(v_k, S) = \tau_{\sigma'}(z, v_k) + \tau_{\hat{\sigma}}(v_k, S) \ .$$

The last equality holds since both the segment of $\hat{\sigma}$ from $v_k$ to $S$ and the segment of $\sigma$ from $v_k$ to $S$ are geodesics, so they have equal passage time. Hence $k+1 \in J$ and we are done. □

Note that each vertex $x \notin \hat{S}$ has out-degree at least 1 in $\mathbb{G}_S$. Furthermore it is possible to argue using part 1 of the previous proposition and the shape theorem that there are no infinite directed paths in $\mathbb{G}_S$. Since we will not use this result later, we omit the proof. Once we take limits of measures on such graphs later, infinite paths will appear.

If $\mathbb{P}$ has unique passage times, we can say more about the structure of $\mathbb{G}_S$.

**Proposition 3.2.2.** *Assume $\mathbb{P}$ admits geodesics and has unique passage times. The following properties hold $\mathbb{P}$-almost surely.*

1. *Each vertex $x \notin \hat{S}$ has out-degree 1. Here $\hat{S}$ is defined as in* (3.6).

2. *Viewed as an undirected graph, $\mathbb{G}_S$ has no circuits.*

*Proof.* For the first property note that every vertex $x \notin \hat{S}$ has out-degree at least 1 because there is a geodesic from the vertex to $S$ and the first edge is directed away from $x$. Assuming $x$ has out-degree at least 2 then we write $e_1$ and $e_2$ for two such directed edges. By the previous proposition, there are two geodesics $\gamma_1$ and $\gamma_2$ from $x$ to $S$ such that $e_i \in \gamma_i$ for $i = 1, 2$. If either of these paths returned to $x$ then there would exist a finite path with passage time equal to 0. There would then be two distinct paths with passage time 0 (the concatenation of a zero passage time path with its reversed path has zero passage time), contradicting unique passage times. This implies that $\gamma_1$ and $\gamma_2$ have distinct edge sets. However, they have the same passage time, again contradicting unique passage times.

For the second property suppose that there is a circuit in the undirected version of $\mathbb{G}_S$. Each vertex has out-degree 1, so this is actually a directed circuit and thus a geodesic. But then it has passage time zero, giving a contradiction as above. □

Property 2 implies that $\mathbb{G}_S$, viewed as an undirected graph, is a forest. It has more than one component if and only if $\hat{S}$ has size at least 2. We will see later that under certain assumptions, after taking limits of measures on these graphs, the number of components will reduce to 1.

### 3.2.1 Busemann functions: Properties

In order to analyze the structure of the geodesic graph, we will often rely on properties of Busemann functions. We list below some basic properties of Busemann functions. One of the most interesting is the additivity property 1. It is the reason that the asymptotic shape for the Busemann function is a half space whereas the asymptotic shape for $\tau$ is a compact set.

**Proposition 3.2.3.** *Let $S \subseteq \mathbb{R}^2$. The Busemann function $B_S$ satisfies the following properties $\mathbb{P}$-almost surely for $x, y, z \in V$:*

1. *(Additivity)*

$$B_S(x,y) = B_S(x,z) + B_S(z,y) \ . \tag{3.8}$$

2. *As in 3.7, if $(V,E)$ is either $\mathbb{Z}^2$ or an upper half-plane and $\mathbb{P}$ is translation-invariant, then*

$$B_S(x,y)(T_{\mathbf{e}_i}\omega) = B_{S+\mathbf{e}_i}(x+\mathbf{e}_i, y+\mathbf{e}_i)(\omega) \tag{3.9}$$

*(recall that in a half-plane this holds only for $i = 1$). Therefore the finite-dimensional distributions of $B_S$ obey a translation invariance:*

$$(B_S(x,y)) \underset{d}{=} (B_{S+\mathbf{e}_i}(x+\mathbf{e}_i, y+\mathbf{e}_i)) \ .$$

3. 

$$|B_S(x,y)| \le \tau(x,y) \ . \tag{3.10}$$

*Proof.* The first property follows from the definition. The third is a consequence of subadditivity of $\tau(y, S)$. The second item follows from the statement (3.2) for passage times. $\square$

The last property we need regards the relation between geodesics and Busemann functions. Though it is simple, it will prove to be important later.

**Proposition 3.2.4.** *Let $S \subseteq \mathbb{R}^2$ and $x \in \mathbb{Z}^2$. If $\gamma$ is a geodesic from $x$ to $S$ and $y$ is a vertex of $\gamma$ then $B_S(x, y) = \tau(x, y)$.*

*Proof.* Write $\tau_\gamma(x, y)$ for the passage time along $\gamma$ between $x$ and $y$. Since every segment of a geodesic itself a geodesic, $\tau(x, S) - \tau(y, S) = \tau_\gamma(x, S) - \tau_\gamma(y, S) = \tau_\gamma(x, y) = \tau(x, y)$. $\square$

Using this proposition and additivity of the Busemann function we can relate $B_S(x, y)$ to coalescence. If $\gamma_x$ and $\gamma_y$ are geodesics from $x$ and $y$ to $S$ (respectively) and they meet at a vertex $z$ then $B_S(x, y) = \tau(x, z) - \tau(y, z)$. This is a main reason why Busemann functions are useful for studying geodesics.

## 3.3 Results

The results of this chapter are directed at the issues **Q1** and **Q2** raised in Section 2.4.1. The ultimate goal is to take limits of geodesic graphs $\mathbb{G}_S$ defined in the last section. Using these limits, various properties of infinite geodesics in the model can be derived.

### 3.3.1 Results on fractional planes

In fractional planes, we can actually show that limits of $\mathbb{G}_{S_n}$ exist for an appropriate sequence of sets $S_n$. Moreover, the limiting behavior can be established under very weak assumptions. Moreover, there exists a limiting Busemann function to a point "at infinity" constructed by taking limits of Busemann functions. This will prove useful in the analysis of the geodesic graphs.

We consider geodesics to the sequence $\{v_n\}_n$ of boundary vertices of $(V, E)$. As shorthand, we will set

$$B_j(x, y) := B_{\{v_j\}}(x, y),$$
$$\mathbb{G}_j := \mathbb{G}_{\{v_j\}}$$

for each $j \in \mathbb{Z}$.

The first result shows that asymptotic limits of the $(B_n)$ exist under no assumptions on $\omega$. That is, it holds for all passage time configurations.

**Theorem 3.3.1.** *Consider first-passage percolation on a fractional plane $(V, E)$. For any $x, y \in V$ and $\omega \in \Omega$,*

$$B(x, y) := \lim_{n \to \infty} B_n(x, y) \text{ exists .} \tag{3.11}$$

For the second result we consider a measure $\mathbb{P}$ on $\Omega$ (with the product Borel sigma algebra) that admits geodesics; that is,

$$\mathbb{P}(\exists \text{ a geodesic } \gamma : x \rightsquigarrow y) = 1 \text{ for all } x, y \in V \text{ .}$$

We will define a notion of convergence for geodesic graphs. We say that $\eta_n \to \eta \in \{0, 1\}^{\vec{E}}$ if for each $e \in \vec{E}$, $\eta_n(e) \to \eta(e)$. In this case we write $\mathbb{G}_n \to \mathbb{G}$, where $\mathbb{G}$ is the directed graph corresponding to $\eta$.

**Theorem 3.3.2.** *Suppose that $\mathbb{P}$ is a distribution on a fractional plane which admits geodesics. Then with probability one, $(\mathbb{G}_n)$ converges to a graph $\mathbb{G}$. Each directed path in $\mathbb{G}$ is a geodesic.*

This result can be taken as an answer to **Q2** of Section 2.4.1 for a particular family of finite geodesics.

### 3.3.2 Results on $(\mathbb{Z}^2, \mathcal{E}^2)$

In the case of the full plane, we no longer have a proof of the existence of the existence of geodesic limits. However, by taking subsequential limits of geodesic graphs, we can build a translation-invariant structure which gives us information about the directional properties of unigeodesics. We can regard the results here as partial answers to **Q1** of Section 2.4.1.

#### Directional results

Below we will show that under **A1** or **A2** there are geodesics that are asymptotically directed in sectors of aperture no bigger than $\pi/2$. Under a certain directional condition on the boundary of the limit shape (see Corollary 3.3.4) we show existence of geodesics with asymptotic direction. To our knowledge, the only work of this type so far [76, Theorem 2.1] requires a global curvature assumption to show the existence of geodesics in even one direction.

To describe the results, we endow $[0, 2\pi)$ with the distance of $S^1$: say that $dist(\theta_1, \theta_2) < r$ if there exists an integer $m$ such that $|\theta_1 - \theta_2 - 2\pi m| < r$. For $\Theta \subseteq [0, 2\pi)$ we say that a path $\gamma = x_0, x_1, \dots$ is *asymptotically directed in* $\Theta$ if for each $\varepsilon > 0$, $\arg x_k \in \Theta_\varepsilon$ for all large $k$, where $\Theta_\varepsilon = \{\theta : dist(\theta, \phi) < \varepsilon \text{ for some } \phi \in \Theta\}$. For $\theta \in [0, 2\pi)$, write $v_\theta$ for the unique point of $\partial\mathcal{B}$ with argument $\theta$. Recall that a supporting line $L$ for $\mathcal{B}$ at $v_\theta$ is one that touches $\mathcal{B}$ at $v_\theta$ such that $\mathcal{B}$ lies on one side of $L$. If $\theta$ is an angle such that $\partial\mathcal{B}$ is differentiable at $v_\theta$ (and therefore has a unique supporting line $L_\theta$ (the tangent line) at this point), we define an interval of angles $I_\theta$:

$$I_\theta = \{\theta' : v_{\theta'} \in L_\theta\} \ . \tag{3.12}$$

**Theorem 3.3.3.** *Assume either* **A1** *or* **A2**. *If $\partial\mathcal{B}$ is differentiable at $v_\theta$, then with probability one there is an infinite geodesic containing the origin which is asymptotically directed in $I_\theta$.*

The meaning of the theorem is that there is a measurable set $\mathcal{A}$ with $\mathbb{P}(\mathcal{A}) = 1$ such that if $\omega \in \mathcal{A}$, there is an infinite geodesic containing the origin in $\omega$ which is asymptotically directed in $I_\theta$. This also applies to any result we state with the phrases "with probability one there is an infinite geodesic" or "with probability one there is a collection of geodesics."

We now state two corollaries. A point $x \in \partial\mathcal{B}$ is *exposed* if there is a supporting line for $\mathcal{B}$ that touches $\mathcal{B}$ only at $x$.

**Corollary 3.3.4.** *Assume either* **A1** *or* **A2**. *Suppose that $v_\theta$ is an exposed point of differentiability of $\partial\mathcal{B}$. With probability one there exists an infinite geodesic containing the origin with asymptotic direction $\theta$.*

*Proof.* Apply Theorem 3.3.3, noting that $I_\theta = \{\theta\}$. □

In the next corollary we show that there are infinite geodesics asymptotically directed in certain sectors. Because the limit shape is convex and compact, it has at least 4 extreme points. Angles corresponding to the arcs connecting these points can serve as the sectors.

**Corollary 3.3.5.** *Assume either* **A1** *or* **A2**. *Let $\theta_1 \neq \theta_2$ be such that $v_{\theta_1}$ and $v_{\theta_2}$ are extreme points of $\mathcal{B}$. If $\Theta$ is the set of angles corresponding to some arc of $\partial\mathcal{B}$ connecting $v_{\theta_1}$ to $v_{\theta_2}$, then with probability one there exists an infinite geodesic containing the origin which is asymptotically directed in $\Theta$.*

*Proof.* Choose $\theta_3 \in \Theta$ such that $\theta_1 \neq \theta_3 \neq \theta_2$ and $\mathcal{B}$ has a unique supporting line $L_{\theta_3}$ at $v_{\theta_3}$ (this is possible since the boundary is differentiable almost everywhere). Let $C$ be the closed arc of $\partial\mathcal{B}$ from $v_{\theta_1}$ to $v_{\theta_2}$ that contains $v_{\theta_3}$ and write $D$ for its open complementary arc. We claim $D \subseteq I_{\theta_3}^c$. This will prove the corollary after applying Theorem 3.3.3 with $\theta = \theta_3$.

For a contradiction, suppose that $L_{\theta_3}$ intersects $D$ at some point $v_\phi$ and write $S$ for the segment of $L_{\theta_3}$ between $v_{\theta_3}$ and $v_\phi$. Since $L_{\theta_3}$ is a supporting line, the set $\mathcal{B}$ lies entirely on one side of it. On the other hand, since $\mathcal{B}$ is convex and $v_{\theta_3}, v_\phi \in \mathcal{B}$, $S \subseteq \mathcal{B}$. Therefore $S \subseteq \partial\mathcal{B}$ and must be an arc of the boundary. It follows that one of $v_{\theta_1}$ or $v_{\theta_2}$ is in the interior of $S$, contradicting the fact that these are extreme points of $\mathcal{B}$. □

**Remark 3.3.6.** *If $\mathbb{P}$ is a product measure with $\mathbb{P}(\omega_e = 1) = \vec{p}_c$ and $\mathbb{P}(\omega_e < 1) = 0$, where $\vec{p}_c$ is the critical value for directed percolation, [8, Theorem 1] implies that $(1/2, 1/2)$ is an exposed point of differentiability of $\mathcal{B}$. Corollary 3.3.4 then gives a geodesic in the direction $\pi/4$. Though all points of $\partial\mathcal{B}$ (for all measures not in the class of Durrett-Liggett [38]) should be exposed points of differentiability, this is the only proven example.*

**Remark 3.3.7.** *From [47, Theorem 1.3], for any compact convex set $\mathcal{C}$ which is symmetric about the axes with nonempty interior, there is a measure $\mathbb{P}$ satisfying* **A2** *(in fact, with bounded passage times) which has $\mathcal{C}$ as a limit shape. Taking $\mathcal{C}$ to be a Euclidean disk shows that there exist measures for which the corresponding model obeys the statement of Corollary 3.3.4 in any deterministic direction $\theta$.*

### Global results

In this section we use the terminology of Newman [76]. Call $\theta$ a *direction of curvature* if there is a Euclidean ball $B_\theta$ with some center and radius such that $\mathcal{B} \subseteq B_\theta$ and $\partial B_\theta \cap \mathcal{B} = \{v_\theta\}$. We say that $\mathcal{B}$ has *uniformly positive curvature* if each direction is a direction of curvature and there exists $M < \infty$ such that the radius of $B_\theta$ is bounded by $M$ for all $\theta$.

Recall that in [76, Theorem 2.1], Newman has shown that under the assumptions (a) $\mathbb{P}$ is a product measure with $\mathbb{E}e^{\beta\omega_e} < \infty$ for some $\beta > 0$, (b) the limit shape $\mathcal{B}$ has uniformly positive curvature and (c) $\omega_e$ is a continuous variable, two things are true with probability one.

1. For each $\theta \in [0, 2\pi)$, there is an infinite geodesic with asymptotic direction $\theta$.

2. Every infinite geodesic has an asymptotic direction.

As far as we know, there has been no weakening of these assumptions.

Below we improve on Newman's theorem. We first reduce the moment assumption on $\mathbb{P}$ to that of **A1**. Next we extend the theorem to non-i.i.d. measures. Newman's proof uses concentration inequalities of Kesten [67] and Alexander [3], which require exponential moments on the distribution (and certainly independence). So to weaken the moment assumptions we need to use a completely different method, involving Busemann functions instead.

For this theorem, we need slightly stronger hypotheses. Recall the definitions of **A1'** and **A2'**.

**Theorem 3.3.8.** *Assume either* **A1'** *or* **A2'** *and that $\mathcal{B}$ has uniformly positive curvature.*

1. *With $\mathbb{P}$-probability one, for each $\theta$ there is an infinite geodesic with direction $\theta$.*

2. *With $\mathbb{P}$-probability one, every infinite geodesic has a direction.*

The same method of proof shows the following.

**Corollary 3.3.9.** *Assume either* **A1'** *or* **A2'** *and suppose $v_\theta$ is an exposed point of differentiability of $\partial\mathcal{B}$ for all $\theta$. Then the conclusions of Theorem 3.3.8 hold.*

**Remark 3.3.10.** *The proofs of the above two results only require that the set of extreme points of $\mathcal{B}$ is dense in $\partial\mathcal{B}$. In fact, a similar result holds for a sector in which extreme points of $\mathcal{B}$ are dense in the arc corresponding to this sector.*

## 3.4 Graph Limits on Fractional Planes

The work in this section will be devoted to proving Theorems 3.3.1 and 3.3.2. Our method is motivated by the "paths crossing" trick of Alm and Wierman [4].

### 3.4.1 Existence of Busemann Limits

The main goal of this subsection is prove Theorem 3.3.1. We begin with $x, y \in \{v_i\}_{i \in \mathbb{Z}}$.

**Proposition 3.4.1.** *For any $x, y \in \{v_i\}_i$ and $\omega \in \Omega$, the limit in (3.11) exists. Moreover, the convergence is monotone.*

*Proof.* We assume that $x = v_i$ and $y = v_j$ for $i < j$ and we let $\epsilon > 0$. Fix any $n_2 > n_1 > j$ such that $v_{n_1} \neq v_{n_2}$. We can now choose vertex self-avoiding paths $\gamma : x \rightsquigarrow v_{n_1}$ and $\gamma' : y \rightsquigarrow v_{n_2}$ to satisfy

$$\tau(\gamma) \leq \tau(x, v_{n_1}) + \epsilon \text{ and } \tau(\gamma') \leq \tau(y, v_{n_2}) + \epsilon \ .$$

Form a continuous path $\beta$ (in $\mathbb{R}^2$) by taking $\gamma$, adjoining half of the edge $e_{n_1}$, adjoining the segment of $\Gamma$ between $e^*_{n_1}$ and $e^*_i$, and then finally appending half of the edge $e_i$, to form a continuous circuit based at $x$. Since this circuit is a Jordan curve, it separates $\mathbb{R}^2$ into an interior and an exterior. See Figure 3.1 for an illustration of $\beta$.

Our first observation is that either $y \in \beta$ or $y$ is in the interior of $\beta$ (and in fact, $y \in \beta$ only if $y \in \gamma$). The reason is that $y$ is an endpoint of one of the $e_i$'s, which must cross $\beta$. Since the other endpoint of this edge is in $V^c$, it cannot be in the interior of $\beta$ (or on $\beta$). The Jordan curve theorem implies that these endpoints are in different components, and thus if $y \notin \beta$, it must be in the interior of $\beta$. We make the following claim:

**Claim 3.4.2.** *$\gamma' \cap \gamma$ contains a vertex of $\mathbb{Z}^2$.*

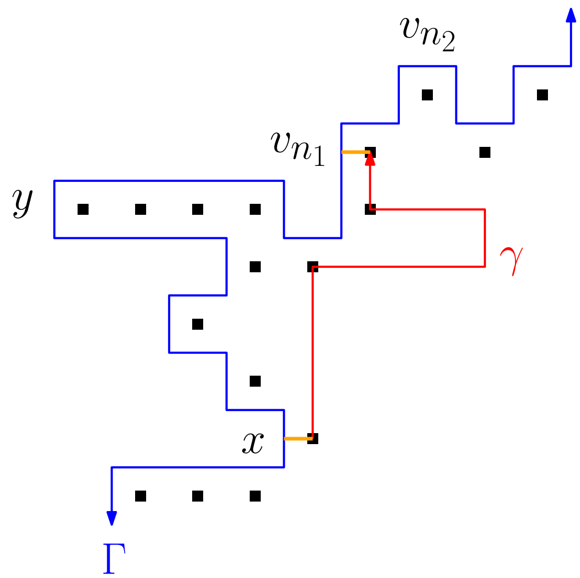



Figure 3.1: Construction of the Jordan curve $\beta$. It consists of the red path $\gamma$, two half dual edges in yellow and the blue segment of $\Gamma$ between $v_{n_1}$ and $x$.

To show the claim, we first prove that $v_{n_2}$ is either on $\beta$ or in the exterior of $\beta$. Accordingly, assume $v_{n_2}$ is not on $\beta$. Notice that neither endpoint of $e_{n_2}$ can touch $\beta$. Furthermore the edge $e_{n_2}$ cannot intersect $\beta$ because $e^*_{n_2}$ is not contained in $\beta$. Therefore both endpoints are in the same component of the complement of $\beta$ and since the other one is in $V^c$, they must be in the exterior of $\beta$.

Now, considering $\gamma'$ as a continuous plane curve, we note that $\gamma'$ must intersect $\beta$ (since it has to reach $v_{n_2}$, which is not in the interior of $\beta$), but it cannot intersect $\Gamma$. Therefore, it must intersect $\gamma$; this intersection must happen at a vertex, though it may of course also happen at one or more edges. This proves the claim.

We will complete the existence proof for the limit in (3.11) by showing that $B_n(x, y)$ is monotone in $n$ for fixed $x$ and $y$. Let $n_1$ and $n_2$ be as above. For any path $\sigma : a \rightsquigarrow b$ and $c \in \sigma$ write $\sigma\ |_c$ for the segment of $\sigma$ from the first meeting of $c$ onward and $\sigma\ |^c$ for the segment of $\sigma$ to the first meeting of $c$. Then letting $w$ be a point in

$\gamma' \cap \gamma$,

$$\begin{aligned}\tau(x, v_{n_2}) + \tau(y, v_{n_1}) &\le [\tau(\gamma \mid^{w}) + \tau(\gamma' \mid_{w})] + [\tau(\gamma' \mid^{w}) + \tau(\gamma \mid_{w})] \\ &= [\tau(\gamma \mid^{w}) + \tau(\gamma \mid_{w})] + [\tau(\gamma' \mid^{w}) + \tau(\gamma' \mid_{w})] \\ &= \tau(\gamma) + \tau(\gamma') \le \tau(x, v_{n_1}) + \tau(y, v_{n_2}) + 2\epsilon \ .\end{aligned}$$

Taking $\epsilon \to 0$,

$$\tau(x, v_{n_2}) + \tau(y, v_{n_1}) \le \tau(x, v_{n_1}) + \tau(y, v_{n_2}). \tag{3.13}$$

We can rearrange the terms in (3.13) to find that

$$B_{n_2}(x, y) \le B_{n_1}(x, y).$$

Since $B_n(x, y)$ is a sequence bounded below by $-\tau(x, y)$, $\lim B_n(x, y)$ exists. □

We now move on to general $x, y \in V$ and prove the limit in (3.11) exists. We will need a few geometric notions. Let $\alpha$ denote the vertex set of a finite, connected subgraph of $(V, E)$ which contains some $v_i$. Denote by $(V', E')$ the graph formed by setting $V' = V \setminus \alpha$ and letting $E'$ be formed from $E$ by removing every edge with an endpoint in $\alpha$. The graph $(V', E')$ may have multiple components, but the following claim allows us to find a single component defining the Busemann function.

**Claim 3.4.3.** *There exists a component $(\overline{V}, \overline{E})$ of $(V', E')$ and an $M < \infty$ such that, for all $n > M$, $v_n \in \overline{V}$. Moreover, $(\overline{V}, \overline{E})$ is formed from $(\mathbb{Z}^2, \mathcal{E}^2)$ by the removal of edges dual to a doubly infinite, self-avoiding path $\overline{\Gamma}$ in the dual lattice.*

*Proof.* Note that if $v_n \neq v_{n+1}$, then there exists a path in $(V, E)$ between $v_n$ and $v_{n+1}$ of Euclidean length at most two. Since $\|v_n\|_1 \to \infty$, we can choose $M$ such that

$$\mathrm{dist}(\{v_n\}_{n > M}, \alpha) \ge 2,$$

where $\text{dist}(\cdot,\cdot)$ is the $(V,E)$ graph distance. Then $\{v_n\}_{n>M}$ must all lie in one component of $(V',E')$, which we denote by $(\overline{V},\overline{E})$.

It remains to show that $(\overline{V},\overline{E})$ can be formed from $(\mathbb{Z}^2,\mathcal{E}^2)$ by cutting along a doubly infinite, loop-free dual path $\overline{\Gamma}$. By Proposition 3.9.1 in Section 3.9.1, it suffices to show that both $\overline{V}$ and $\mathbb{Z}^2\setminus\overline{V}$ are infinite and connected (as subsets of $\mathbb{Z}^2$). Both claims are true for $\overline{V}$. Moreover, $\mathbb{Z}^2\setminus\overline{V}$ is infinite, since it contains $V^c$. Because $\alpha$ is connected and contains a point of $\{v_i\}_i$, we see that $\mathbb{Z}^2\setminus\overline{V}$ is connected; it consists of the union of $\alpha$, $V^c$, and the sites of $V$ which were only reachable from the large $v_n$'s via sites of $\alpha$ (see Figure 3.2). Therefore, by the above, the boundary between $(\overline{V},\overline{E})$ and $G$ is a doubly infinite self-avoiding dual path, proving the claim. □

We note that, by Proposition 3.4.1 and the linearity of the Busemann function, we need only prove the existence of the limit in (3.11) when $y\notin\{v_i\}_i$ but $x$ is some $v_m$ (which can be chosen as a function of $y$). Fix $y$, and denote by $\alpha$ the vertex set of some (vertex self-avoiding, finite) path in $(V,E)$ which starts at a vertex adjacent to $y$ and ends at a vertex $w\in\{v_i\}_i$. Form the graph $(\overline{V},\overline{E})$ as in Claim 3.4.3; denote by $\overline{\Gamma}$ the doubly-infinite dual path whose existence is established in the claim, and define $\{\bar{v}_i\}_i$ analogously to $\{v_i\}_i$. We may choose an orientation of $\{\bar{v}_i\}_i$ such that the following holds. There exists $\kappa\in\mathbb{Z}$ such that for all large $n$, $v_n=\bar{v}_{n+\kappa}$.

If $\bar{\tau}$ and $\overline{B}_n$ are the passage times and Busemann functions in $(\overline{V},\overline{E})$ (defined in the obvious way), then

$$\overline{B}(\bar{v}_i,\bar{v}_j)=\lim_{n\to\infty}\overline{B}_n(\bar{v}_i,\bar{v}_j) \tag{3.14}$$

exists for all $i$ and $j$ by Proposition 3.4.1. Fix $m$ large enough that $v_n=\bar{v}_{n+\kappa}$ for $n\ge m$, and choose any $n>m$.

Note that $y$ is adjacent to some vertex of $\alpha$; therefore, if $y\in\overline{V}$, then $y=\bar{v}_l$ for some $l$. We will want to apply the following lemma to both $z=y$ and $z=v_m$:

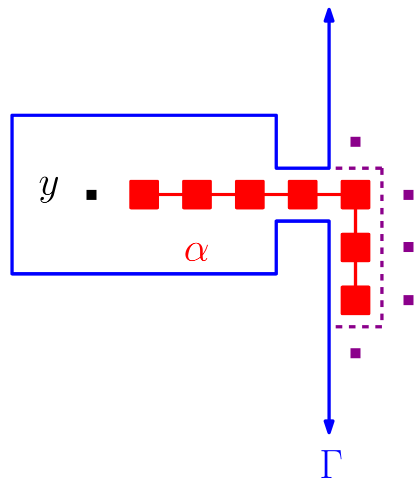



Figure 3.2: Removal of the vertex set $\alpha$ from $V$. The enlarged red squares represent $\alpha$ and the dotted purple path is the segment of $\overline{\Gamma}$ that does not lie in $\Gamma$. The vertices $\bar{v}_j$ for $j \in J$ are depicted in purple.

**Lemma 3.4.4.** *Let $z \in V$ be such that either $z \in \{\bar{v}_i\}_i$ or $z \notin \overline{V}$. Denote by $J \subseteq \mathbb{Z}$ the finite set of indices such that $\bar{v}_j$ is at Euclidean distance one from $\alpha$. Then*

$$\tau(z, v_n) = \min_{j \in J}\{\tau(z, \bar{v}_j) + \overline{\tau}(\bar{v}_j, v_n)\} \ . \tag{3.15}$$

*Proof.* Let $\epsilon > 0$ and $j \in J$. Then find paths $\gamma : z \rightsquigarrow \bar{v}_j$ in $(V, E)$ and $\bar{\gamma} : \bar{v}_j \rightsquigarrow v_n$ in $(\overline{V}, \overline{E})$ such that $\tau(\gamma) \leq \tau(z, \bar{v}_j) + \epsilon$ and $\overline{\tau}(\bar{\gamma}) \leq \overline{\tau}(\bar{v}_j, v_n) + \epsilon$. Build a path $\sigma : z \rightsquigarrow v_n$ in $(V, E)$ by concatenating $\gamma$ with $\bar{\gamma}$. Then

$$\tau(z, v_n) \leq \tau(\sigma) = \tau(\gamma) + \overline{\tau}(\bar{\gamma}) \leq \tau(z, \bar{v}_j) + \overline{\tau}(\bar{v}_j, v_n) + 2\epsilon \ .$$

Taking $\epsilon \to 0$ and a minimum over $j \in J$ gives the inequality $\leq$ in (3.15).

To prove the other inequality, let $\sigma : z \rightsquigarrow v_n$ in $(V, E)$ be a path such that $\tau(\sigma) \leq \tau(z, v_n) + \epsilon$. The path $\sigma$ must have a terminal segment $\bar{\gamma}$ which lies in $(\overline{V}, \overline{E})$ from some $\bar{v}_{j_0}$ to $v_n$ – this terminal segment may be equal to the singleton $\{v_n\}$. Write

$\gamma$ for the segment of $\sigma$ from $z$ to the last meeting of $\bar{v}_{j_0}$. Then

$$\begin{aligned}\min_{j \in J} \{\tau(z, \bar{v}_j) + \overline{\tau}(\bar{v}_j, v_n)\} &\leq \tau(z, \bar{v}_{j_0}) + \overline{\tau}(\bar{v}_{j_0}, v_n) \\ &\leq \tau(\gamma) + \overline{\tau}(\bar{\gamma}) = \tau(\sigma) \leq \tau(z, v_n) + \epsilon \ .\end{aligned}$$

Taking $\epsilon \to 0$ proves (3.15). □

So, defining

$$\varphi_j(z, n) := \tau(z, \bar{v}_j) + \overline{\tau}(\bar{v}_j, v_n) - \overline{\tau}(\bar{v}_1, v_n) \ ,$$

we see that $\tau(z, v_n) = \overline{\tau}(\bar{v}_1, v_n) + \min_{j \in J} \varphi_j(z, n)$. Moreover,

$$\lim_{n \to \infty} \varphi_j(z, n) =: \varphi_j(z)$$

exists by (3.14), and therefore so does

$$\lim_{n \to \infty} [\tau(z, v_n) - \overline{\tau}(\bar{v}_1, v_n)] \ . \tag{3.16}$$

Finally, we can use the above to show convergence of $B_n(y, v_m)$ as $n \to \infty$. Write

$$\begin{aligned}\lim_{n \to \infty} B_n(y, v_m) &= \lim_{n \to \infty} [\tau(y, v_n) - \tau(v_m, v_n)] \\ &= \lim_{n \to \infty} [\tau(y, v_n) - \overline{\tau}(\bar{v}_1, v_n) + \overline{\tau}(\bar{v}_1, v_n) - \tau(v_m, v_n)] \\ &= \lim_{n \to \infty} [\tau(y, v_n) - \overline{\tau}(\bar{v}_1, v_n)] - \lim_{n \to \infty} [\tau(v_m, v_n) - \overline{\tau}(\bar{v}_1, v_n)] \, ;\end{aligned}$$

Using (3.16) with $z = y$ and $z = v_m$ completes the proof.

### 3.4.2 Geodesic Limits

Our aim in this subsection is to prove Theorem 3.3.2.

The second statement of the theorem follows directly from the Section 3.2: each directed path in $\mathbb{G}_n$ is a geodesic. So we prove the first statement and show that for each directed edge $(x,y)$ in $\vec{E}$, with probability one the value of $\eta_n((x,y))$ is eventually constant. Fix $x \in V$ and choose $m \in \mathbb{N}$ such that, defining (with $d(\cdot,\cdot)$ the graph distance in $(V,E)$)

$$S_m = \{w \in V : \, d(x,w) \leq m\}$$
$$\partial S_m = \{w \in V : d(x,w) = m+1\} \ ,$$

we have $S_m \cap \{v_i\}_i \neq \emptyset$. Setting $\alpha = S_m$, we may apply Claim 3.4.3 to find $(\overline{V}, \overline{E})$, a component of the graph generated by removing $\alpha$ from $(V,E)$ containing $v_n$ for all large $n$. As before, it can be alternatively created by cutting $(\mathbb{Z}^2, \mathcal{E}^2)$ along a doubly infinite self-avoiding dual path $\overline{\Gamma}$. We will decorate expressions with an overline when they are meant for the model in $(\overline{V}, \overline{E})$ (for instance, $\overline{\tau}$). For the remainder, we also fix $\omega \in \Omega$ such that for each $x, y \in V$, there is a geodesic from $x$ to $y$.

For each $\zeta \in T_m := \partial S_m \cap \overline{V}$, and $n$ such that $v_n \in \overline{V}$, we define the quantity

$$f_n(\zeta) = \tau(x,\zeta) + \overline{\tau}(\zeta, v_n) \ . \tag{3.17}$$

Let $\mathfrak{m}_n$ be the set of minimizers of $f_n$.

**Lemma 3.4.5.** *There exists* $\mathfrak{m} \subset T_m$ *such that* $\mathfrak{m}_n = \mathfrak{m}$ *for all large* $n$.

*Proof.* First, note that by Proposition 3.4.1, for $\zeta, \zeta' \in T_m$,

$$\begin{aligned} f_n(\zeta) - f_n(\zeta') &= \tau(x,\zeta) + \overline{\tau}(\zeta, v_n) - \tau(x,\zeta') - \overline{\tau}(\zeta', v_n) \\ &= \tau(x,\zeta) - \tau(x,\zeta') + \overline{B}_n(\zeta,\zeta') \end{aligned}$$

is eventually monotone. Suppose that $\zeta \in T_m$ satisfies $\zeta \notin \mathfrak{m}_n$ for infinitely many $n$. Then we can find $\zeta'$ such that $f_n(\zeta) - f_n(\zeta') > 0$ for infinitely many $n$. By monotonicity this means that actually $f_n(\zeta) - f_n(\zeta') > 0$ for all large $n$ and thus $\zeta \notin \mathfrak{m}_n$ for all large $n$. This also implies that if $\zeta \in \mathfrak{m}_n$ for infinitely many $n$ then $\zeta \in \mathfrak{m}_n$ for all large $n$, completing the proof. □

Given this lemma, the theorem will follow once we show that $\eta_n((x,y)) = 1$ if and only if $\{x,y\}$ is in a geodesic from $x$ to a vertex of $\mathfrak{m}_n$. Note that $T_m$ is equal to the set of vertices in $\overline{V}$ at Euclidean distance one from $S_m$. Applying Lemma 3.4.4 with $z = x$, any $\zeta \in T_m$ satisfies

$$\zeta \in \mathfrak{m}_n \text{ if and only if } f_n(\zeta) = \tau(x, v_n) \ .$$

So suppose first that $\eta_n((x,y)) = 1$; then $\{x,y\}$ is in a geodesic $\gamma$ from $x$ to $v_n$. $\gamma$ has a last intersection $\zeta$ with $T_m$. Then the segment $\bar{\gamma}$ of $\gamma$ from this intersection to $v_n$ has

$$\tau(\zeta, v_n) = \tau(\bar{\gamma}) \geq \overline{\tau}(\zeta, v_n) \ .$$

But $\overline{\tau}(\zeta, v_n) \geq \tau(\zeta, v_n)$, so $\tau(\bar{\gamma}) = \overline{\tau}(\zeta, v_n)$. Therefore

$$\tau(x, v_n) = \tau(\gamma) = \tau(x,\zeta) + \tau(\bar{\gamma}) = \tau(x,\zeta) + \overline{\tau}(\zeta, v_n) = f_n(\zeta) \ ,$$

giving $\zeta \in \mathfrak{m}_n$. Furthermore the segment of $\gamma$ up to the last intersection with $\zeta$ is a geodesic from $x$ to $\zeta$ that contains $\{x,y\}$.

Conversely, suppose that $\{x,y\}$ is in a geodesic $\gamma_1$ from $x$ to a vertex $\zeta$ of $\mathfrak{m}_n$; we will show that $\eta_n((x,y)) = 1$. Choose $\gamma_2$ as any geodesic from $\zeta$ to $v_n$. Concatenate them to form a path $\gamma$ from $x$ to $v_n$. We compute

$$\tau(\gamma) = \tau(\gamma_1) + \tau(\gamma_2) = \tau(x,\zeta) + \tau(\zeta, v_n) \leq \tau(x,\zeta) + \overline{\tau}(\zeta, v_n) = f_n(\zeta) \ .$$

However since $\zeta \in \mathfrak{m}_n$, $f_n(\zeta) = \tau(x, v_n)$, so $\tau(\gamma) \leq \tau(x, v_n)$. The opposite inequality holds because $\gamma : x \rightsquigarrow v_n$, so $\gamma$ is a geodesic from $x$ to $v_n$. It remains to show that $\tau(x, v_n) \geq \tau(y, v_n)$. But this holds because $y$ appears in $\gamma$ after the first appearance of $x$. Therefore if we write $\sigma$ for the segment of $\gamma$ from the first intersection with $y$ to $v_n$, then

$$\tau(x, v_n) = \tau(\gamma) \geq \tau(\sigma) = \tau(y, v_n) \ .$$

## 3.5 Subsequential limits on $\mathbb{Z}^2$: Construction

We would like to replicate the success of the limit construction on fractional planes, now in the case that $(V, E) = (\mathbb{Z}^2, \mathcal{E}^2)$. Unfortunately, without the "paths crossing" tool provided by the boundary, we are unaware of any way to show that such a limit exists. We instead are forced to construct a limiting graph by use of subsequences. However, because $(V, E)$ is now highly symmetric, our limiting graph will inherit these symmetry properties. From these, much can be deduced about limiting geodesics in the graph.

Our ultimate goal is to prove Theorems 3.3.3 and 3.3.8. These will be derived over the next sections after a period of tool-building, in which we construct a general framework for dealing with subsequential limits of geodesic graphs and Busemann functions. In this section, we construct the aforementioned framework and prove some of its basic properties.

For the remainder of this section, we assume that $(V, E)$ is $\mathbb{Z}^2$ and $\mathbb{P}$ is a measure satisfying **A1** or **A2**.

We will choose a one-parameter family of lines $L_\alpha = L + \alpha \mathbf{v}$ for $\mathbf{v}$ a normal vector to $L$ and consider the Busemann functions $B_{L_\alpha}(x, y)$. The main question is whether or not the limit

$$\lim_{\alpha \to \infty} B_{L_\alpha}(x, y) \tag{3.18}$$

exists for $x, y \in \mathbb{Z}^2$. If one could show this, then one could prove many results about FPP, for instance, that infinite geodesics with an asymptotic direction always exist. Under an assumption of uniformly positive curvature of the limit shape $\mathcal{B}$ and exponential moments for the common distribution of the $\omega_e$'s (in the case that $\mathbb{P}$ is a product measure) Newman [76] has shown the existence of this limit for Lebesgue-almost every unit vector $\mathbf{v}$.

We will try to overcome the difficulty of existence of limits (3.18) by enlarging the space to work with subsequential limits in a systematic way. This technique is inspired by work [6, 7] on ground states of short-range spin glasses.

### 3.5.1 Definition of $\mu$

We begin by assigning a space for our passage times. Let $\Omega_1 = \mathbb{R}^{\mathbb{Z}^2}$ be a copy of $\Omega$. A sample point in $\Omega_1$ we call $\omega$ as before. Our goal is to enhance this space to keep track of Busemann functions and geodesic graphs. We will take limits in a fixed direction, so for the remainder of this section, let $\varpi \in \partial\mathcal{B}$ and let $g_\varpi$ be any linear functional on $\mathbb{R}^2$ that takes its maximum on $\mathcal{B}$ at $\varpi$ with $g_\varpi(\varpi) = 1$. The nullspace of $g_\varpi$ is then a translate of a supporting line for $\mathcal{B}$ at $\varpi$. For $\alpha \in \mathbb{R}$, define

$$L_\alpha = \left\{x \in \mathbb{R}^2 : g_\varpi(x) = \alpha\right\} \ .$$

For future reference, we note the inequality

$$\text{for all } x \in \mathbb{R}^2,\ g_\varpi(x) \leq g(x) \ . \tag{3.19}$$

It clearly holds if $x \neq 0$. Otherwise since $x/g(x) \in \mathcal{B}$, $1 \geq g_\varpi(x/g(x)) = g_\varpi(x)/g(x)$.

Given $\alpha \in \mathbb{R}$ and $\omega \in \Omega_1$, write $B_\alpha(x, y)(\omega) = B_{L_\alpha}(x, y)(\omega)$. Define the space $\Omega_2 = (\mathbb{R}^2)^{\mathbb{Z}^2}$ with the product topology and Borel sigma-algebra and the *Busemann* *increment configuration* $B_\alpha(\omega) \in \Omega_2$ as

$$B_\alpha(\omega) = \big( B_\alpha(v, v+\mathbf{e}_1),\, B_\alpha(v, v+\mathbf{e}_2) \big)_{v \in \mathbb{Z}^2} \,.$$

We also consider directed graphs of geodesics. These are points in a directed graph space $\Omega_3 = \{0,1\}^{\vec{\mathcal{E}}^2}$, where $\vec{\mathcal{E}}^2$ is the set of oriented edges $\langle x, y \rangle$ of $\mathbb{Z}^2$, and we use the product topology and Borel sigma-algebra. For $\eta \in \Omega_3$, write $\mathbb{G} = \mathbb{G}(\eta)$ for the directed graph induced by the edges $e$ such that $\eta(e) = 1$. Using the definition from Section 3.2

$$\eta_\alpha(\omega) = \eta_{L_\alpha}(\omega) \in \Omega_3 \text{ and } \mathbb{G}_\alpha(\omega) = \mathbb{G}(\eta_\alpha(\omega)) \text{ for } \alpha \in \mathbb{R} \,.$$

Set $\widetilde{\Omega} = \Omega_1 \times \Omega_2 \times \Omega_3$, equipped with the product topology and Borel sigma-algebra;

$$(\omega, \Theta, \eta) = (\omega(e), \theta_1(x), \theta_2(x), \eta(f) : e \in \mathcal{E}^2, x \in \mathbb{Z}^2,\ f \in \vec{\mathcal{E}}^2)$$

denotes a generic element of the space $\widetilde{\Omega}$. Define the map

$$\Phi_\alpha : \Omega_1 \longrightarrow \widetilde{\Omega} \text{ by } \omega \mapsto (\omega, B_\alpha(\omega), \eta_\alpha(\omega)) \,. \tag{3.20}$$

Because $\Phi_\alpha$ is measurable, we can use it to push forward the distribution $\mathbb{P}$ to a probability measure $\mu_\alpha$ on $\widetilde{\Omega}$. Given the family $(\mu_\alpha)$ and $n \in \mathbb{N}$, we define the empirical average

$$\mu_n^*(\cdot) := \frac{1}{n} \int_0^n \mu_\alpha(\cdot)\, \mathrm{d}\alpha. \tag{3.21}$$

To prove that this defines a probability measure, one must show that for each measurable $A \subseteq \widetilde{\Omega}$, the map $\alpha \mapsto \mu_\alpha(A)$ is Lebesgue-measurable. The proof is deferred to Section 3.9.3.

From $B_\alpha(x,y) \leq \tau(x,y)$, the sequence $(\mu_n^*)_{n=1}^\infty$ is seen to be tight and thus has a subsequential weak limit $\mu$. We will call the marginal of $\mu$ on $\Omega_2$ a *Busemann increment distribution* and the marginal on $\Omega_3$ a *geodesic graph distribution*. It will be important to recall the Portmanteau theorem, a basic result about weak convergence. The following are equivalent if $(\nu_k)$ is a sequence of Borel probability measures on a metric space $X$:

$$\lim_{k\to\infty} \nu_k \to \nu \text{ weakly}$$

$$\limsup_{k\to\infty} \nu_k(A) \leq \nu(A) \text{ if } A \text{ is closed} \tag{3.22}$$

$$\liminf_{k\to\infty} \nu_k(A) \geq \nu(A) \text{ if } A \text{ is open .} \tag{3.23}$$

(See, for example, [57, Theorem 3.25].) Because $\widetilde{\Omega}$ is metrizable, these statements apply.

In this section and the next, we prove general properties about the measure $\mu$ and focus on the marginal on $\Omega_2$. In Sections 3.7 and 4.5 we study the marginal on $\Omega_3$ and in Section 3.8 relate results back to the original FPP model. It is important to remember that $\mu$ depends among other things not only on $\varpi$, but on the choice of the linear functional $g_\varpi$. We will suppress mention of $\varpi$ in the notation. Furthermore we will use $\mu$ to represent the measure and also its marginals. For instance, if we write $\mu(A)$ for an event $A \subseteq \Omega_2$ we mean $\mu(\Omega_1 \times A \times \Omega_3)$.

### 3.5.2 Translation invariance of $\mu$.

We will show that $\mu$ inherits translation invariance from $\mathbb{P}$. The natural translations $\tilde{T}_m,\ m = 1, 2$ act on $\widetilde{\Omega}$ as follows:

$$\left[\tilde{T}_m(\omega, \Theta, \eta)\right](e, x, f) = (\omega_{e-\mathbf{e}_m}, \theta_1(x-\mathbf{e}_m), \theta_2(x-\mathbf{e}_m), \eta(f-\mathbf{e}_m)) \ .$$

Here, for example, we interpret $e - \mathbf{e}_m$ for the edge $e = (y, z)$ as $(y - \mathbf{e}_m, z - \mathbf{e}_m)$.

**Lemma 3.5.1.** *For any $\alpha \in \mathbb{R}$ and $m = 1, 2$, $\mu_\alpha \circ \tilde{T}_m = \mu_{\alpha + g_\varpi(\mathbf{e}_m)}$.*

*Proof.* Let $A$ be a cylinder event for the space $\widetilde{\Omega}$ of the form

$$A = \left\{\omega_{e_i} \in \mathbf{B}_i, \theta_{r_j}(x_j) \in \mathbf{C}_j, \eta(f_k) = a_k : i = 1, \ldots, l,\ j = 1, \ldots, m,\ k = 1, \ldots, n\right\} ,$$

where each $\mathbf{B}_i, \mathbf{C}_j$ is a (real) Borel set with $a_k \in \{0, 1\}$, each $r_j \in \{1, 2\}$, and each $e_i \in \mathcal{E}^2, x_j \in \mathbb{Z}^2$ and $f_k \in \vec{\mathcal{E}}^2$. We will show that for $m = 1, 2$,

$$\mu_\alpha \left(\tilde{T}_m^{-1} A\right) = \mu_{\alpha + g_\varpi(\mathbf{e}_m)}(A) . \tag{3.24}$$

Such $A$ generate the sigma-algebra so this will imply the lemma. For $m \in \{1, 2\}$,

$$\tilde{T}_m^{-1}(A) = \left\{\omega_{e_i - \mathbf{e}_m} \in \mathbf{B}_i, \theta_{r_j}(x_j - \mathbf{e}_m) \in \mathbf{C}_j, \eta(f_k - \mathbf{e}_m) = a_k\right\} .$$

Rewriting $\mu_\alpha(\cdot) = \mathbb{P}(\Phi_\alpha^{-1}(\cdot))$ and using the definition of $\Phi_\alpha$ (3.20),

$$\mu_\alpha(\tilde{T}_m^{-1}(A)) = \mathbb{P}\left(\omega_{e_i - \mathbf{e}_m} \in \mathbf{B}_i, B_\alpha(x_j - \mathbf{e}_m, x_j - \mathbf{e}_m + \mathbf{e}_{r_j}) \in \mathbf{C}_j, \eta_\alpha(f_k - \mathbf{e}_m)(\omega) = a_k\right) .$$

Note that translation invariance of $\mathbb{P}$ allows to shift the translation by $\mathbf{e}_m$ from the arguments of $\omega$, $B_\alpha$ and $\eta_\alpha$ to the position of the line $L_\alpha$. We have equality in distribution:

$$\omega_{e - \mathbf{e}_m} \underset{d}{=} \omega_e,\ B_\alpha(x - \mathbf{e}_m, y - \mathbf{e}_m) \underset{d}{=} B_\beta(x, y) \text{ and } \eta_\alpha(e - \mathbf{e}_m) \underset{d}{=} \eta_\beta(e) ,$$

where $\beta = \alpha + g_\varpi(\mathbf{e}_m)$. In fact, using the translation covariance statements (3.2), (3.9) and (3.7), equality of the above sort holds for the joint distribution of the $\omega$'s, Busemann increments and graph variables appearing in the event $A$. This proves (3.24). ☐

**Proposition 3.5.2.** *$\mu$ is invariant under the translations $\tilde{T}_m$, $m = 1, 2$.*

*Proof.* Let $f$ be a continuous function (bounded by $D \geq 0$) on the space $\widetilde{\Omega}$, and fix $\epsilon > 0$. Choose an increasing sequence $(n_k)$ such that $\mu^*_{n_k} \to \mu$ weakly as $k \to \infty$. We can then find $k_0$ such that $|\mu(f) - \mu^*_{n_k}(f)| < \epsilon/3$ for $k > k_0$. By Lemma 3.5.1, $\mu_\alpha \circ \tilde{T}_m = \mu_{\alpha + g_\varpi(\mathbf{e}_m)}$ for $m = 1, 2$. Therefore

$$\left[\mu^*_{n_k} \circ \tilde{T}_m\right](f) = \frac{1}{n_k} \int_{g_\varpi(\mathbf{e}_m)}^{n_k + g_\varpi(\mathbf{e}_m)} \mu_\alpha(f)\, \mathrm{d}\alpha$$

$$\Rightarrow \left|\left[\mu^*_{n_k} \circ \tilde{T}_m\right](f) - \mu^*_{n_k}(f)\right| \leq \frac{1}{n_k}\left|\int_0^{g_\varpi(\mathbf{e}_m)} \mu_\alpha(f)\, \mathrm{d}\alpha\right| + \frac{1}{n_k}\left|\int_{n_k}^{n_k + g_\varpi(\mathbf{e}_m)} \mu_\alpha(f)\, \mathrm{d}\alpha\right|$$

$$\leq \frac{2 g_\varpi(\mathbf{e}_m) D}{n_k} \to 0 \text{ as } k \to \infty\ .$$

As $\tilde{T}_m$ is a continuous on $\widetilde{\Omega}$, $(\mu^*_{n_k} \circ \tilde{T}_m)$ converges weakly to $\mu \circ \tilde{T}_m$, so there exists $k_1 > k_0$ such that $|\mu \circ \tilde{T}_m(f) - \mu^*_{n_k} \circ \tilde{T}_m(f)| < \epsilon/3$ for all $k > k_1$, and $k_2 > k_1$ with $2g_\varpi(\mathbf{e}_m)D/n_{k_2} < \epsilon/3$. So $|\mu(f) - \mu \circ \tilde{T}_m(f)| < \epsilon$ for all $\varepsilon > 0$, giving $\mu = \mu \circ \tilde{T}_m$. ☐

### 3.5.3 Reconstructed Busemann functions

We wish to reconstruct an "asymptotic Busemann function" $f : \mathbb{Z}^2 \to \mathbb{R}$ by summing the Busemann increments of $\Theta \in \Omega_2$. That $\Theta$ is almost surely curl-free allows the construction to proceed independent of the path we sum over. For this we need some definitions.

Given $\Theta \in \Omega_2$, $x \in \mathbb{Z}^2$ and $z \in \mathbb{Z}^2$ with $\|z\|_1 = 1$ we set $\theta(x,z) = \theta(x,z)(\Theta)$ equal to

$$\theta(x,z) = \begin{cases} \theta_1(x) & z = \mathbf{e}_1 \\ \theta_2(x) & z = \mathbf{e}_2 \\ -\theta_1(x-\mathbf{e}_1) & z = -\mathbf{e}_1 \\ -\theta_2(x-\mathbf{e}_2) & z = -\mathbf{e}_2 \end{cases} .$$

For any finite lattice path $\gamma$ we write its vertices in order as $x_1, \ldots, x_n$ and set

$$f(\gamma) = f(\gamma)(\Theta) = \sum_{i=1}^{n-1} \theta(x_i, x_{i+1} - x_i) .$$

**Lemma 3.5.3.** *With $\mu$-probability one, $f$ vanishes on all circuits:*

$$\mu\left(f(\gamma) = 0 \textit{ for all circuits } \gamma\right) = 1 .$$

*Proof.* Pick a circuit $\gamma$ and let $A \subseteq \widetilde{\Omega}_2$ denote the event $\{\Theta : f(\gamma) = 0\}$. Choose an increasing sequence $(n_k)$ such that $\mu^*_{n_k} \to \mu$ weakly. For fixed $\gamma$, $f(\gamma)$ is a continuous function on $\widetilde{\Omega}$, so the event $A$ is closed, giving $\mu(A) \geq \limsup_k \mu^*_{n_k}(A)$ by (3.22) . However, for each $\alpha$, by additivity of $B_\alpha(\cdot,\cdot)$ (see (3.8)),

$$\mu_\alpha(A) = \mathbb{P}\left(\sum_{i=1}^{n} B_\alpha(x_i, x_{i+1}) = 0\right) = 1 .$$

Thus $\mu^*_n(A) = 1$ for all $n$ and $\mu(A) = 1$. There are countably many $\gamma$'s so we are done. □

Using the lemma we may define the reconstructed Busemann function. Fix a deterministic family of finite paths $\{\gamma_{x,y}\}$, one for each pair $(x,y) \in \mathbb{Z}^2$ and define

$$f(x,y) = f(x,y)(\Theta) := f(\gamma_{x,y}) .$$

Although we use fixed paths $\gamma_{x,y}$, this is only to ensure that $f$ is a continuous function on $\widetilde{\Omega}$. Actually, for any $\Theta$ in the $\mu$-probability one set of Lemma 3.5.3 and vertices $x, y \in \mathbb{Z}^2$ we could equivalently define $f(x, y) = f(\gamma)$, where $\gamma$ is any finite lattice path from $x$ to $y$. To see that it would then be well-defined (that is, only a function of $x, y$ and the configuration $\Theta$) is a standard argument. If we suppose that $\gamma_1$ and $\gamma_2$ are finite lattice paths from $x$ to $y$ and $\Theta$ is given as above, the concatenation of $\gamma_1$ with $\gamma_2$ (traversed in the opposite direction) is a circuit and thus has $f$-value zero. However, by definition, this is the difference of $f(\gamma_1)$ and $f(\gamma_2)$ and proves the claim.

We now give some properties about asymptotic Busemann functions that come over from the original model. The third says that $f$ retains translation covariance. This will allow us to prove the existence of almost-sure limits using the ergodic theorem in the next section.

**Proposition 3.5.4.** *The reconstructed Busemann function satisfies the following properties for $x, y, z \in \mathbb{Z}^2$.*

1.
$$f(x, y) + f(y, z) = f(x, z) \ \mu\text{-almost surely} . \tag{3.25}$$

2. *For* $m = 1, 2$
$$f(x, y)(\tilde{T}_m\Theta) = f(x - \mathbf{e}_m, y - \mathbf{e}_m)(\Theta) \ \mu\text{-almost surely} . \tag{3.26}$$

3.
$$f(x, y) : \widetilde{\Omega} \to \mathbb{R} \text{ is continuous} . \tag{3.27}$$

4. *$f$ is bounded by $\tau$:*
$$|f(x, y)| \leq \tau(x, y)) \ \mu\text{-almost surely} . \tag{3.28}$$

*Proof.* The first two properties follow from path-independence of $f$ and the third holds because $f$ is a sum of finitely many Busemann increments, each of which is a continuous function. We show the fourth property. For $x, y \in \mathbb{Z}^2$, the event

$$\{(\omega, \Theta) : |f(x,y)(\Theta)| - \tau(x,y)(\omega) \leq 0\}$$

is closed because $|f(x,y)| - \tau(x,y)$ is continuous. For every $\alpha$, (3.10) gives $|B_\alpha(x,y)| \leq \tau(x,y)$ with $\mathbb{P}$-probability one, so the above event has $\mu_\alpha$-probability one. Taking limits and using (3.22), $\mu(|f(x,y)(\Theta)| \leq \tau(x,y)(\omega)) = 1$. □

### 3.5.4 Expected value of $f$

In this section we compute $\mathbb{E}_\mu f(0,x)$ for all $x \in \mathbb{Z}^2$. The core of our proof is a argument from Hoffman [51], which was developed using an averaging argument due to Garet-Marchand [41]. The presentation we give below is inspired by that of Gouéré [42, Lemma 2.6]. In fact, the proof shows a stronger statement. Without need for a subsequence,

$$\mathbb{E}_{\mu_n^*} f(0,x) \to g_\varpi(x) \ .$$

**Theorem 3.5.5.** *For each $x \in \mathbb{Z}^2$, $\mathbb{E}_\mu f(0,x) = g_\varpi(x)$.*

*Proof.* We will use an elementary lemma that follows from the shape theorem.

**Lemma 3.5.6.** *The following convergence takes place almost surely and in $L^1(\mathbb{P})$:*

$$\frac{\tau(0, L_\alpha)}{\alpha} \to 1 \ \textit{as}\ \alpha \to \infty \ .$$

*Proof.* Since $\alpha\varpi \in L_\alpha$,

$$\limsup_{\alpha \to \infty} \frac{\tau(0, L_\alpha)}{\alpha} \leq \lim_{\alpha \to \infty} \frac{\tau(0, \alpha\varpi)}{\alpha} = 1 \ .$$

On the other hand, given $\varepsilon > 0$ and any $\omega$ for which the shape theorem holds, we can find $K$ such that for all $x \in \mathbb{R}^2$ with $\|x\|_1 \geq K$, $\tau(0,x) \geq g(x)(1-\varepsilon)$. So if $\alpha$ is large enough that all $x \in L_\alpha$ have $\|x\|_1 \geq K$, then we can use (3.19):

$$\tau(0, L_\alpha) = \min_{x \in L_\alpha} \tau(0,x) \geq (1-\varepsilon) \min_{x \in L_\alpha} g(x) \geq (1-\varepsilon)\alpha \ .$$

Consequently, $\liminf_{\alpha \to \infty} \tau(0, L_\alpha)/\alpha \geq 1$, giving almost sure convergence in the lemma.

For $L^1$ convergence, note $0 \leq \tau(0, L_\alpha)/\alpha \leq \tau(0, \alpha\varpi)/\alpha$, so the dominated convergence theorem and $L^1$ convergence of point to point passage times completes the proof. □

For any $x \in \mathbb{Z}^2$ and integer $n \geq 1$, use the definition of $\mu_n^*$ to write

$$\mathbb{E}_{\mu_n^*}(f(-x,0)) = \frac{1}{n}\left[\int_0^n \mathbb{E}\tau(-x, L_\alpha)\ \mathrm{d}\alpha - \int_0^n \mathbb{E}\tau(0, L_\alpha)\ \mathrm{d}\alpha\right] \ .$$

Using translation covariance of passage times,

$$\int_0^n \mathbb{E}\tau(-x, L_\alpha)\ \mathrm{d}\alpha = \int_0^n \mathbb{E}\tau(0, L_{\alpha + g_\varpi(x)})\ \mathrm{d}\alpha = \int_{g_\varpi(x)}^{n+g_\varpi(x)} \mathbb{E}\tau(0, L_\alpha)\ \mathrm{d}\alpha \ .$$

Therefore

$$\mathbb{E}_{\mu_n^*}(f(-x,0)) = \frac{1}{n}\left[\int_n^{n+g_\varpi(x)} \mathbb{E}\tau(0, L_\alpha)\ \mathrm{d}\alpha - \int_0^{g_\varpi(x)} \mathbb{E}\tau(0, L_\alpha)\ \mathrm{d}\alpha\right] \ . \tag{3.29}$$

Choose $(n_k)$ to be an increasing sequence such that $\mu_{n_k}^* \to \mu$ weakly. We claim that

$$\mathbb{E}_{\mu_{n_k}^*} f(-x,0) \to \mathbb{E}_\mu f(-x,0) \ . \tag{3.30}$$

To prove this, note that for any $R > 0$, if we define the truncated variable

$$f_R(-x,0) = \operatorname{sgn} f(-x,0) \min\{R, |f(-x,0)|\} ,$$

then continuity of $f$ on $\widetilde{\Omega}$ gives $\mathbb{E}_{\mu^*_{n_k}} f_R(-x,0) \to \mathbb{E}_\mu f_R(-x,0)$. To extend this to (3.30), it suffices to prove that for each $\varepsilon > 0$, there exists $R > 0$ such that

$$\limsup_{k\to\infty} \mathbb{E}_{\mu^*_{n_k}} |f(-x,0)| I(|f(-x,0)| \geq R) < \varepsilon , \tag{3.31}$$

where $I(A)$ is the indicator of the event $A$. Because $\mathbb{E}_{\mu^*_{n_k}} f(-x,0)^2 \leq \mathbb{E}\tau(-x,0)^2 < \infty$ for all $k$ by (3.4), condition (3.31) follows from the Cauchy-Schwarz inequality. This proves (3.30).

Combining (3.29) and (3.30), we obtain the formula

$$\mathbb{E}_\mu f(-x,0) = \lim_{k\to\infty} \frac{1}{n_k} \int_{n_k}^{n_k+g_\varpi(x)} \mathbb{E}\tau(0, L_\alpha)\ \mathrm{d}\alpha = \lim_{k\to\infty} \int_0^{g_\varpi(x)} \frac{\mathbb{E}\tau(0, L_{\alpha+n_k})}{n_k}\ \mathrm{d}\alpha . \tag{3.32}$$

By Lemma 3.5.6, for each $\alpha$ between 0 and $g_\varpi(x)$,

$$\lim_{k\to\infty} \frac{\mathbb{E}\tau(0, L_{\alpha+n_k})}{n_k} = \lim_{k\to\infty} \frac{\mathbb{E}\tau(0, L_{\alpha+n_k})}{\alpha + n_k} \cdot \frac{\alpha + n_k}{n_k} = 1 .$$

So using $\mathbb{E}\tau(0, L_{\alpha+n_k}) \leq \mathbb{E}\tau(0, L_{2n_k})$ for large $k$, we can pass the limit under the integral in (3.32) to get $\mathbb{E}_\mu f(0,x) = \mathbb{E}_\mu f(-x,0) = g_\varpi(x)$. □

## 3.6 Subsequential limits on $\mathbb{Z}^2$: Asymptotics

We continue the analysis of the subsequential limiting distribution $\mu$ constructed in the preceding section; we will continue to assume **A1** or **A2** throughout. In this section we study the asymptotic behavior of the reconstructed Busemann function $f$. We will see that $f$ is asymptotically a projection onto a line and if the boundary of the limit shape is differentiable at $\varpi$, we give the explicit form of the hyperplane. Without this assumption we show that the line is a translate of a supporting line for $\mathcal{B}$ at $\varpi$.

One of the advantages of constructing $f$ from our measure $\mu$ is that we can use the ergodic theorem and translation invariance to show the existence of limits. This gives us almost as much control on the Busemann function as we would have if we could show existence of the limit in (3.18). If we knew this, we would not need differentiability at $\varpi$ to deduce the form of the random hyperplane for $f$; we could derive it from ergodicity and symmetry.

### 3.6.1 Radial limits

In this section we will prove the existence of radial limits for $f$. This is the first step to deduce a shape theorem, which we will do in the next section. We extend the definition of $f$ to all of $\mathbb{R}^2 \times \mathbb{R}^2$ in the usual way: $f(x, y)$ is defined as $f(\tilde{x}, \tilde{y})$ where $\tilde{x}$ and $\tilde{y}$ are the unique points in $\mathbb{Z}^2$ such that $x \in \tilde{x} + [-1/2, 1/2)^2$ and $y \in \tilde{y} + [-1/2, 1/2)^2$.

**Proposition 3.6.1.** *Let $q \in \mathbb{Q}^2$. Then*

$$\rho_q := \lim_{n\to\infty} \frac{1}{n} f(0, nq) \text{ exists } \mu\text{-almost surely .}$$

*Proof.* Choose $M \in \mathbb{N}$ such that $Mq \in \mathbb{Z}^2$. We will first show that

$$\lim_{n\to\infty} \frac{1}{Mn} f(0, nMq) \text{ exists } \mu\text{-almost surely .} \tag{3.33}$$

To do this, we note that since $\tau(0, Mq) \in L^2(\mu)$ (from (3.4)), it is also in $L^1$. Using (3.28), $f(0, Mq) \in L^1(\mu)$ as well. Define the map $\tilde{T}_q$ on $\Omega_2$ as

$$\left[\tilde{T}_q \Theta\right](x) = (\theta_1(x - Mq), \theta_2(x - Mq)) \; .$$

This is a composition of maps $\tilde{T}_m$, $m = 1, 2$, so it is measure-preserving. By (3.25) and (3.26),

$$f(0, nMq)(\Theta) = \sum_{i=1}^{n} f((i-1)Mq, iMq)(\Theta) = \sum_{i=0}^{n-1} f(0, Mq)(\tilde{T}_q^{-i}(\Theta)) \; .$$

Applying the ergodic theorem finishes the proof of (3.33).

To transform (3.33) into the statement of the proposition we need to "fill in the gaps." Choose $M$ as above and for any $n$ pick $a_n \in \mathbb{Z}$ such that $a_n M \leq n < (a_n+1)M$. Then

$$\left| \frac{f(0,nq)}{n} - \frac{f(0,a_nMq)}{a_nM} \right| \leq \left| \frac{f(0,a_nMq)}{a_nM} \right| \left| 1 - \frac{a_nM}{n} \right| + \frac{1}{n} \left| f(0,a_nMq) - f(0,nq) \right| \; .$$

The first term on the right converges to 0. To show the same for the second term we use the fact that $f(x,y) \in L^1(\mu, \Omega_2)$ for all $x, y \in \mathbb{R}^2$. Indeed, the difference $f(0, a_nMq) - f(0, nq)$ is equal to $f(nq, a_nMq)$, which has the same distribution as $f(0, (a_nM - n)q)$. For each $\varepsilon > 0$,

$$\sum_{n \geq 1} \mu(|f(0,(n-a_nM)q)| \geq \varepsilon n) \leq \frac{1}{\varepsilon} \sum_{i=1}^{M} \|f(0,-iq)\|_{L^1(\mu)} < \infty \; .$$

So only finitely many of the events $\{|f(0,a_nMq) - f(0,nq)| \geq \varepsilon n\}$ occur and we are done. □

The last proposition says that for each $q$ there exists a random variable $\rho_q = \rho(q, \Theta)$ such that $\mu$-almost surely, the above limit equals $\rho_q$. Assume now that we fix $\Theta$ such that this limit exists for all $q \in \mathbb{Q}^2$. We will consider $\rho_q$ as a function of $q$. The next theorem states that $\rho_q$ represents a random projection onto a vector $\varrho$.

**Theorem 3.6.2.** *There exists a random vector $\varrho = \varrho(\Theta)$ such that*

$$\mu\left(\rho_q = \varrho \cdot q \text{ for all } q \in \mathbb{Q}^2\right) = 1 \; .$$

*Furthermore $\varrho$ is translation invariant:*

$$\varrho(\tilde{T}_m\Theta) = \varrho(\Theta) \text{ for } m = 1, 2 \ .$$

*Proof.* We will show that $q \mapsto \rho_q$ is a (random) linear map on $\mathbb{Q}^2$. Specifically, writing an arbitrary $q \in \mathbb{Q}^2$ as $(q_1, q_2)$, we will show that

$$\mu\left(\rho_q = q_1\rho_{\mathbf{e}_1} + q_2\rho_{\mathbf{e}_2} \text{ for all } q \in \mathbb{Q}^2\right) = 1 \ . \tag{3.34}$$

Then, setting $\varrho = (\rho_{\mathbf{e}_1}, \rho_{\mathbf{e}_2})$, we will have proved the theorem.

The first step is to show translation invariance of $\rho_q$. Given $q \in \mathbb{Q}^2$, let $M \in \mathbb{N}$ be such that $Mq \in \mathbb{Z}^2$. For $m = 1, 2$, translation covariance implies

$$\begin{aligned} |f(0, nMq)(\tilde{T}_m\Theta) - f(0, nMq)(\Theta)| &= |f(-\mathbf{e}_m, nMq - \mathbf{e}_m)(\Theta) - f(0, nMq)(\Theta)| \\ &\leq |f(-\mathbf{e}_m, 0)(\Theta)| + |f(nMq - \mathbf{e}_m, nMq)(\Theta)| \ . \end{aligned}$$

Furthermore, given $\delta > 0$,

$$\sum_n \mu\left(|f(nMq - \mathbf{e}_m, nMq)| > \delta n\right) \leq \sum_n \mu\left(|f(0, \mathbf{e}_m)| > \delta n\right) \leq \frac{1}{\delta}\|f(0, \mathbf{e}_m)\|_{L^1(\mu)} < \infty \ .$$

Therefore only finitely many of the events $\{|f(nMq - \mathbf{e}_m, nMq)| > \delta n\}$ occur and

$$\rho_q(\tilde{T}_m\Theta) = \lim_{n\to\infty} \frac{f(0, nMq)(\tilde{T}_m\Theta)}{nM} = \lim_{n\to\infty} \frac{f(0, nMq)(\Theta)}{nM} = \rho_q(\Theta) \text{ almost surely} \ .$$

To complete the proof we show that $q \mapsto \rho_q$ is almost surely additive. Over $\mathbb{Q}$, this suffices to show linearity and thus (3.34). Let $q_1, q_2 \in \mathbb{Q}^2$ and choose $M \in \mathbb{N}$ with $Mq_1, Mq_2 \in \mathbb{Z}^2$. By Proposition 3.6.1, for $\varepsilon > 0$, we can pick $N$ such that if $n \geq N$ then the following hold:

1. $\mu\left(|(1/nM)f(0,nMq_1) - \rho_{q_1}| > \varepsilon/2\right) < \varepsilon/2$ and

2. $\mu\left(|(1/nM)f(0,nMq_2) - \rho_{q_1}| > \varepsilon/2\right) < \varepsilon/2$.

Writing $\tilde{T}_{-q}(\Theta)(x) = \Theta(x + Mq)$ and using translation invariance of $\rho_{q_2}$,

$$\begin{aligned}
& f(0,nM(q_1+q_2))(\Theta) - nM\rho_{q_1}(\Theta) - nM\rho_{q_2}(\Theta) \\
= \quad & f(0,nMq_1)(\Theta) - nM\rho_{q_1}(\Theta) + f(0,nMq_2)(\tilde{T}^n_{-q_1}\Theta) - nM\rho_{q_2}(\tilde{T}^n_{-q_1}\Theta) \ .
\end{aligned}$$

So by translation invariance of $\mu$ and items 1 and 2 above,

$$\begin{aligned}
& \mu(|(1/nM)f(0,nM(q_1+q_2)) - (\rho_{q_1} + \rho_{q_2})| > \varepsilon) \\
\leq \quad & \mu(|(1/nM)f(0,nMq_1) - \rho_{q_1}| > \varepsilon/2) + \mu(|(1/nM)f(0,nMq_2) - \rho_{q_2}| > \varepsilon/2) < \varepsilon \ .
\end{aligned}$$

Thus $(1/nM)f(0,nM(q_1+q_2))$ converges in probability to $\rho_{q_1} + \rho_{q_2}$. By Proposition 3.6.1, this equals $\rho_{q_1+q_2}$. □

### 3.6.2 A shape theorem

We will now upgrade the almost-sure convergence in each rational direction, from Proposition 3.6.1, to a sort of shape theorem for the Busemann function $f$. The major difference is that, unlike in the usual shape theorem of first-passage percolation, the limiting shape of $f$ is allowed to be random.

**Theorem 3.6.3.** *For each $\delta > 0$,*

$$\mu\left(|f(0,x) - x \cdot \varrho| < \delta\|x\|_1 \textit{ for all } x \textit{ with } \|x\|_1 \geq M \textit{ and all large } M\right) = 1. \quad (3.35)$$

As in the proofs of the usual shape theorems, we will need a lemma which allows us to compare $f$ in different directions. A result showing that with positive probability, $f(0,x)$ grows at most linearly in $\|x\|$ will be sufficient for our purposes. The fourth item of Proposition 3.5.4 allows us to derive such a bound by comparison with the usual passage time $\tau(0,x)$.

**Lemma 3.6.4.** *There exist deterministic $K < \infty$ and $p_g > 0$ depending only on the passage time distribution such that*

$$\mathbb{P}\left(\sup_{\substack{x\in\mathbb{Z}^2\\ x\neq 0}} \frac{\tau(0,x)}{\|x\|_1} \leq K\right) = p_g > 0.$$

*Proof.* By the first-passage shape theorem, there exists $\lambda < \infty$ and $T, p_g > 0$ such that

$$\mathbb{P}\left(\forall t \geq T,\ B(t)/t \supseteq [-\lambda,\lambda]^2\right) = p_g\ .$$

(Here we are using (3.5).) Choosing $K = T + 2/\lambda$ completes the proof.

□

The development of the shape theorem from this point is similar to that of the usual first-passage shape theorem for ergodic passage time distributions.

We will say that $z \in \mathbb{Z}^2$ is "good" for a given outcome if

$$\sup_{\substack{x\in\mathbb{Z}^2\\ x\neq z}} \frac{\tau(z,x)}{\|x-z\|_1} \leq K\ .$$

Note that $\mathbb{P}(z \text{ is good}) = p_g > 0$ for all $z \in \mathbb{Z}^2$.

**Lemma 3.6.5.** *Let $\zeta$ be a nonzero vector with integer coordinates, and let $z_n = n\zeta$. Let $(n_k)$ denote the increasing sequence of integers such that $z_{n_k}$ is good. $\mathbb{P}$-almost surely, $(n_k)$ is infinite and* $\lim_{k\to\infty}(n_{k+1}/n_k) = 1$.

*Proof.* The ergodic theorem shows that $(n_k)$ is a.s. infinite. Let $B_i$ denote the event that $z_i$ is good. By another application of the ergodic theorem,

$$\frac{k}{n_k} = \frac{1}{n_k}\sum_{i=1}^{n_k} \mathbf{1}_{B_i} \longrightarrow p_g \quad \text{a.s.} \tag{3.36}$$

Thus,

$$\frac{n_{k+1}}{n_k} = \left(\frac{n_{k+1}}{k+1}\right)\left(\frac{k}{n_k}\right)\left(\frac{k+1}{k}\right) \longrightarrow 1 \quad \text{a.s.},$$

since the first and second factors converge to $p_g$ and $p_g^{-1}$ by (3.36).

□

In what follows, we will use the fact that there is a positive density of good sites to show convergence of $f(0,z)/\|z\|_1$ in all directions. Given the convergence of $f(0,nq)/n$ for each rational $q$, we will find enough good sites along lines close to $nq$ to let us to bound the difference $|f(0,nq) - f(0,z)|$. To describe this procedure, we need to make several definitions. Call a vector $\zeta$ satisfying the a.s. event of Lemma 3.6.5 a good direction. We will extend this definition to $\zeta \in \mathbb{Q}^2$: such a $\zeta$ will be called a good direction if $m\zeta$ is, where $m$ is the smallest natural number such that $m\zeta \in \mathbb{Z}^2$.

By countability, there exists a probability one event $\Omega''$ on which each $\zeta \in \mathbb{Q}^2$ is a good direction. For each integer $M \geq 1$, let $V_M = \{x/M : x \in \mathbb{Z}^2\}$, and let $V = \cup_{M\geq 1} V_M$. Set $B = \{z \in \mathbb{R}^2 : z \in V, \|z\|_1 = 1\}$ and note that $B$ is dense in the unit sphere of $\mathbb{R}^2$ (with norm $\|\cdot\|_1$). By Theorem 3.6.2, we can find a set $\hat{\Omega} \subseteq \Omega_2$ with $\mu(\hat{\Omega}) = 1$ such that, for all $\Theta \in \hat{\Omega}$,

$$\lim_{n\to\infty} \frac{1}{n} f(nz_0)(\Theta) = z_0 \cdot \varrho(\Theta) \text{ for all } z_0 \in B \ . \tag{3.37}$$

*Proof of Theorem 3.6.3.* Assume that there exist $\delta > 0$ and an event $D_\delta$ with $\mu(D_\delta) > 0$ such that, for every outcome in $D_\delta$, there are infinitely many vertices $x \in \mathbb{Z}^2$ with $|f(x) - x \cdot \varrho| \geq \delta \|x\|_1$. Then $D_\delta \cap \hat{\Omega} \cap \Omega''$ is nonempty and so it contains some outcome $(\omega, \Theta, \eta)$. We will derive a contradiction by showing that $(\omega, \Theta, \eta)$, by way of its membership in these three sets, has contradictory properties.

By compactness of the $\ell^1$ unit ball, we can find a sequence $\{x_n\}$ in $\mathbb{Z}^2$ with $\|x_n\| \to \infty$ and $y \in \mathbb{R}^2$ with $\|y\|_1 = 1$ such that $x_n/\|x_n\|_1 \to y$ and

$$\left| \frac{f(x_n)[\Theta]}{\|x_n\|_1} - y \cdot \varrho[\Theta] \right| > \frac{\delta}{2} \text{ for all } n \ . \tag{3.38}$$

Let $\delta' > 0$ be arbitrary (we will ultimately take it to be small). Our first goal is the approximation of $x_n$ by multiples of some element of $B$. Choose $z \in B$ such that $\|z - y\|_1 < \delta'$ and let $\{n_k\}$ denote the increasing sequence of integers such that $n_k z$ is good. (Here if $z \notin \mathbb{Z}^2$, then $z$ being good means that $Mz$ is good, where $Mz$ was chosen after Lemma 3.6.5 to be $\mathbb{Z}^2$. Therefore $(n_k)$ would then be of the form $(Ml_k)$ for some increasing sequence $l_k$.) Note that $n_{k+1}/n_k \to 1$ by Lemma 3.6.5 so we are able to choose a $K > 0$ such that

$$n_{k+1} < (1 + \delta')n_k \text{ and } \left| \frac{f(0, n_k z)}{n_k} - \varrho \cdot z \right| \leq \delta' \text{ for all } k > K \ . \tag{3.39}$$

By the triangle inequality, the left-hand side of (3.38) is bounded above by

$$\left| \frac{f(0, x_n)}{\|x_n\|_1} - \frac{f(0, n_k z)}{\|x_n\|_1} \right| + \left| \frac{f(0, n_k z)}{\|x_n\|_1} - \frac{f(0, n_k z)}{n_k} \right| + \left| \frac{f(0, n_k z)}{n_k} - \varrho \cdot z \right| + |\varrho \cdot z - \varrho \cdot y| \tag{3.40}$$

for arbitrary $n$ and $n_k$. Choose some $N_0$ such that $\|x_n - \|x_n\|_1\, y\|_1 \leq \delta'\|x_n\|_1$ for all $n > N_0$, and note that

$$\|x_n - \|x_n\|_1 z\|_1 \leq \|x_n - \|x_n\|_1 y\|_1 + \|x_n\|_1 \, \|y - z\|_1 \leq 2\|x_n\|_1 \delta' \text{ for } n > N_0 \ . \tag{3.41}$$

For any $n$, let $k = k(n)$ be the index such that $n_{k+1} \geq \|x_n\|_1 > n_k$. If $n$ is so large that $k(n) > K$, then $\| \, \|x_n\|_1 z - n_k z\|_1 < \delta'\|x_n\|_1$. Combining this observation with

(3.41) gives

$$\|x_n - n_k z\|_1 \le 3\delta' \|x_n\|_1 \text{ for } \|x_n\|_1 \in (n_k, n_{k+1}] \text{ when } k = k(n) > K \ . \tag{3.42}$$

For the remainder of the proof, fix any $n > N_0$ such that $k = k(n) > K$, so that (3.42) holds. We will now control the terms in (3.40), working our way from right to left. The rightmost term may be bounded by noting

$$|\varrho \cdot z - \varrho \cdot y| = |\varrho \cdot (z - y)| \le \|z - y\|_2 \|\varrho\|_2 \le \delta' \|\varrho\|_2 \ .$$

The second term from the right is bounded above by $\delta'$ by (3.39). To bound the third term from the right, note that $n_k < \|x_n\|_1 \le n_{k+1}$, so by (3.39),

$$\begin{aligned} \left| \frac{f(0, n_k z)}{\|x_n\|_1} - \frac{f(0, n_k z)}{n_k} \right| &= \left| \frac{f(0, n_k z)}{n_k} \right| \left( 1 - \frac{n_k}{\|x_n\|_1} \right) \\ &\le [|\varrho \cdot z| + \delta'] \left( 1 - \frac{1}{1 + \delta'} \right) \ . \end{aligned}$$

It remains to bound the first term of (3.40). To do this, note that by (3.42),

$$|f(0, x_n) - f(0, n_k z)| = |f(n_k z, x_n)| \le \tau(n_k z, x_n) \le K \|x - n_k z\|_1 \le 3K\delta' \|x_n\|_1 \ .$$

So

$$\left| \frac{f(0, x_n)}{\|x_n\|_1} - \frac{f(0, n_k z)}{\|x_n\|_1} \right| \le 3K\delta' \ .$$

Applying our estimates for each term in (3.40) to the left side of (3.38) gives

$$\frac{\delta}{2} \le 3K\delta' + (|\varrho \cdot z| + \delta') \left( 1 - \frac{1}{1 + \delta'} \right) + \delta' + \delta' \|\varrho\|_2 \ .$$

Because this holds for all $\delta' > 0$, and because the right-hand side goes to zero as $\delta' \to 0$, we have derived a contradiction and proved the theorem. □

### 3.6.3 General properties of $\varrho$

In this short section we study the random vector $\varrho$. In the case that $\partial\mathcal{B}$ is differentiable at $\varpi$, the vector $\varrho$ is deterministic and we give the explicit form.

The main theorem of the section is below. It says that the line

$$L_\varrho := \{x \in \mathbb{R}^2 : \varrho \cdot x = 1\}$$

is $\mu$-almost surely a supporting line for $\mathcal{B}$ at $\varpi$.

**Theorem 3.6.6.** *With $\mu$-probability one, $\varrho \cdot \varpi = 1$ and $\varrho \cdot x \leq 1$ for all $x \in \mathcal{B}$. Thus $L_\varrho$ is a supporting line for $\mathcal{B}$ at $\varpi$.*

This theorem has an important corollary. It follows directly from the fact that there is a unique supporting line for $\mathcal{B}$ at points of differentiability of $\partial\mathcal{B}$.

**Corollary 3.6.7.** *If $\partial\mathcal{B}$ is differentiable at $\varpi$ then*

$$\mu\big(\varrho = (g_\varpi(\mathbf{e}_1), g_\varpi(\mathbf{e}_2))\big) = 1 \ .$$

*Proof of Theorem 3.6.6.* Using Theorem 3.5.5, we first find the expected value of $\varrho \cdot y$ for $y \in \mathbb{R}^2$. We simply apply the dominated convergence theorem with the bound $|f(0, my)| \leq \tau(0, my)$. Letting $y_m \in \mathbb{Z}^2$ be such that $my \in y_m + [-1/2, 1/2)^2$,

$$\mathbb{E}_\mu(\varrho \cdot y) = \lim_{m \to \infty} \frac{1}{m} \mathbb{E}_\mu f(0, my) = \lim_{m \to \infty} g_\varpi(y_m/m) = g_\varpi(y) \ .$$

The theorem follows from this statement and

$$\mu\left(x \cdot \varrho \leq g(x) \text{ for all } x \in \mathcal{B}\right) = 1 \ . \tag{3.43}$$

Indeed, assuming this, we have

$$\mu(\varrho \cdot \varpi \leq 1) = 1 \text{ and } \mathbb{E}_\mu(\varrho \cdot \varpi) = g_\varpi(\varpi) = 1 \ ,$$

giving $\varrho \cdot \varpi = 1$ with $\mu$-probability one. To prove (3.43), first take $x \in \mathbb{Q}^2 \cap \mathcal{B}$. Then by (3.28), for all $n$, $f(nx) \leq \tau(nx)$ with $\mu$-probability one. Dividing by $n$ and taking limits with Proposition 3.6.1 and the shape theorem we get $x \cdot \varrho \leq g(x)$. For non-rational $x \in \mathcal{B}$ we extend the inequality by almost sure continuity of both sides in $x$. □

## 3.7 Subsequential limits on $\mathbb{Z}^2$: Geodesic graphs

In this section we study the behavior of $\mu$ on $\Omega_3$. We will continue to assume **A1** or **A2** throughout this section. Given $\eta \in \Omega_3$ recall from Section 3.5.1 the definition of the geodesic graph $\mathbb{G}$ of $\eta$ as the directed graph induced by the edges $e$ for which $\eta(e) = 1$. In this section we prove a fundamental property about infinite directed paths in this graph which relates them to the asymptotic Busemann function constructed from $\Theta$.

### 3.7.1 Basic properties

We begin by showing that properties of $\eta_\alpha$ from Section 3.2 carry over to $\eta$. We use some new notation. We say that $y \in \mathbb{Z}^2$ is connected to $z \in \mathbb{Z}^2$ in $\mathbb{G}$ (written $y \to z$) if there exists a sequence of vertices $y = y_0, y_1, \ldots, y_n = z$ such that $\eta(\langle y_k, y_{k+1} \rangle) = 1$ for all $k = 0, \ldots, n-1$. We say that a path in $\mathbb{G}$ is a geodesic (for the configuration $(\omega, \Theta, \eta)$) if it is a geodesic in $\omega$.

**Proposition 3.7.1.** *With $\mu$-probability one, the following statements hold for $x, y, z \in \mathbb{Z}^2$.*

1. *Each directed path in $\mathbb{G}$ is a geodesic.*

2. *If $x \to y$ in $\mathbb{G}$ then $f(x,y) = \tau(x,y)$.*

3. *If $x \to z$ and $y \to z$ in $\mathbb{G}$ then $f(x,y) = \tau(x,z) - \tau(y,z)$.*

4. *There exists an infinite self-avoiding directed path starting at $x$ in $\mathbb{G}$.*

*Proof.* The third item follows directly from the second and additivity of $f$ (from (3.25)). For the first item, if $\gamma$ is a deterministic finite directed path, write $A_\gamma$ for the event that all edges of $\gamma$ are edges of $\mathbb{G}$ and

$$B_\gamma = A_\gamma^c \cup (A_\gamma \cap \{\gamma \text{ is a geodesic}\}) \ .$$

The event in question equals the intersection over all finite $\gamma$'s of $B_\gamma$, so it suffices to show that for each $\gamma$, $\mu(B_\gamma) = 1$.

By part 1 of Proposition 3.2.1, for all $\alpha \in \mathbb{R}$ the $\mathbb{P}$-probability that all directed paths in $\mathbb{G}_\alpha(\omega)$ are geodesics is 1. By pushing forward to $\widetilde{\Omega}$, for each $\alpha$, $\mu_\alpha(B_\gamma) = 1$ and thus $\mu_n^*(B_\gamma) = 1$ for all $n$. Once we show that $B_\gamma$ is a closed event, we will be done, as we can then apply (3.22). To show this we note that the event that a given finite path is a geodesic is a closed event. Indeed, letting $\gamma_1$ and $\gamma_2$ be finite paths, the function $\tau(\gamma_1) - \tau(\gamma_2)$ is continuous on $\widetilde{\Omega}$. Therefore the event $\{\omega \in \Omega_1 : \tau(\gamma_1) \leq \tau(\gamma_2)\}$ is closed. We then write

$$\{\gamma_1 \text{ is a geodesic}\} = \bigcap_{\gamma_2} \{\tau(\gamma_1) \leq \tau(\gamma_2)\} \ ,$$

where the intersection is over all finite paths $\gamma_2$ with the same endpoints as those of $\gamma_1$. Thus $\{\gamma_1 \text{ is a geodesic}\}$ is closed. Since $A_\gamma$ depends on finitely many edge variables $\eta(e)$, it is closed and its complement is closed. Therefore $B_\gamma$ is closed and we are done.

For item 2, we write $\gamma_{xy}$, any path from $x$ to $y$ in $\mathbb{G}$, in order as $x = x_0, x_1, \ldots, x_n = y$ and use additivity of $f$:

$$f(x,y) = \sum_{i=0}^{n-1} f(x_i, x_{i+1}) \ .$$

For each $i$, $x_i \to x_{i+1}$, and by item 1, $\gamma_{xy}$ is a geodesic. This means that we only need to show that if $x$ and $y$ are neighbors such that $\eta(\langle x,y\rangle) = 1$ then $f(x,y) = \omega_{\langle x,y\rangle}$, the passage time of the edge between $x$ and $y$. By part 2 of Proposition 3.2.1, for each $\alpha$, with $\mathbb{P}$-probability one, if $\eta_\alpha(\langle x,y\rangle) = 1$ then $B_\alpha(x,y) = \omega_{\langle x,y\rangle}$. By similar reasoning to that in the last item,

$$\{\eta(\langle x,y\rangle) = 0\} \cup \left(\{\eta(\langle x,y\rangle) = 1\} \cap \{f(x,y) = \omega_{\langle x,y\rangle}\}\right)$$

is closed and since it has $\mu_\alpha$-probability 1 for all $\alpha$, it also has $\mu$-probability one.

We now argue for item 4. By translation-invariance we can just prove it for $x = 0$. For $n \geq 1$ let $A_n \subseteq \Omega_3$ be the event that there is a self-avoiding directed path starting at 0 in $\mathbb{G}$ that leaves $[-n,n]^2$. We claim that $\mu(A_n) = 1$ for all $n$. Taking $n \to \infty$ will prove item 4.

For each $\alpha > 0$ so large that $[-n,n]^2$ is contained on one side of $L_\alpha$, let $\gamma$ be a geodesic from 0 to $L_\alpha$. This path is contained in $\mathbb{G}_\alpha$. We may remove loops from $\gamma$ so that it is self-avoiding, and still a geodesic. It will also be directed in the correct way: as we traverse the path from 0, each edge will be directed in the direction we are traveling. So for all large $\alpha > 0$, with $\mathbb{P}$-probability one, there is a self-avoiding directed path starting at 0 in $\mathbb{G}_\alpha$ that leaves $[-n,n]^2$. Thus $\mu_\alpha(A_n) = 1$ for all large $\alpha$ and $\mu^*_{n_k}(A_n) \to 1$ as $k \to \infty$. The indicator of $A_n$ is continuous on $\widetilde{\Omega}$, as $A_n$ depends on $\eta(f)$ for finitely many edges $f$, so $\mu(A_n) = 1$. $\square$

**Proposition 3.7.2.** *Assume* **A1'** *or* **A2'**. *With $\mu$-probability one, the following statements hold.*

1. *Each vertex in* $\mathbb{Z}^2$ *has out-degree 1 in* $\mathbb{G}$. *Consequently from each vertex* $x$ *emanates exactly one infinite directed path* $\Gamma_x$.

2. *Viewed as an undirected graph,* $\mathbb{G}$ *has no circuits.*

*Proof.* For $x \in \mathbb{Z}^2$, let $A_x \subseteq \widetilde{\Omega}$ be the event that $\eta(\langle x, y\rangle) = 1$ for only one neighbor $y$ of $x$. Note that the indicator of $A_x$ is a bounded continuous function, so since $\mu_\alpha(A_x) = 1$ for all $\alpha$ such that $x$ is not within Euclidean distance 1 of $L_\alpha$ (from part 1 of Proposition 3.2.2 – here $\hat{S}$ is contained in the set of vertices within distance 1 of $L_\alpha$) it follows that $\mu(A_x) = 1$. For each $z$ that is not a neighbor of $x$, $\eta(\langle x, z\rangle) = 0$ with $\mu_\alpha$-probability one for all $\alpha$. This similarly implies that in $\mathbb{G}$ with $\mu$-probability one, there is no edge between $x$ and such a $z$.

To prove the second statement, fix any circuit $\mathcal{C}$ in $\mathbb{Z}^2$ and let $A_\mathcal{C}$ be the event that each edge of $\mathcal{C}$ is in $\mathbb{G}$. Because there are no circuits in $\mathbb{G}_\alpha$ with $\mathbb{P}$-probability one, we have $\mu_n^*(A_\mathcal{C}) = 0$ for all $n$. The indicator of $A_\mathcal{C}$ is a continuous function on $\widetilde{\Omega}$, so we may take limits and deduce $\mu(A_\mathcal{C}) = 0$. There are a countable number of circuits, so we are done. □

### 3.7.2 Asymptotic directions

Recall the definition $L_\varrho = \{x \in \mathbb{R}^2 : x \cdot \varrho = 1\}$ for the vector $\varrho = \varrho(\Theta)$ of Theorem 3.6.2. Set

$$J_\varrho = \{\theta : L_\varrho \text{ touches } \mathcal{B} \text{ in direction } \theta\} . \tag{3.44}$$

The main theorem of this subsection is as follows.

**Theorem 3.7.3.** *With* $\mu$*-probability one, for all* $x \in \mathbb{Z}^2$*, the following holds. Each directed infinite self-avoiding path in* $\mathbb{G}$ *which starts at* $x$ *is asymptotically directed in* $J_\varrho$.

*Proof.* We will prove the theorem for $x = 0$. Assuming we do this, then using translation invariance of $\mu$ and $\varrho$ it will follow for all $x$.

Let $\varepsilon_k = 1/k$ for $k \geq 1$ and $\delta > 0$. We will show that if $S_0 = \{x \in \mathbb{Z}^2 : 0 \to x$ in $\mathbb{G}\}$ then

$$\text{for each } k \geq 1,\ \mu(\arg x \in (J_\varrho)_{\varepsilon_k} \text{ for all but finitely many } x \in S_0) > 1 - \delta\ . \quad (3.45)$$

Here we write $(J_\varrho)_{\varepsilon_k}$ for all angles $\theta$ with $dist(\theta, \theta') < \varepsilon_k$ for some $\theta' \in J_\varrho$. The line $L_\varrho$ only touches $\mathcal{B}$ in directions in $J_\varrho$ so by convexity, $v_\theta \cdot \varrho < 1$ for all $\theta \notin J_\varrho$. Since the set of angles not in $(J_\varrho)_{\varepsilon_k}$ is compact in $[0, 2\pi)$ (using the metric $dist$), we can find a random $a \in (0,1)$ with $v_\theta \cdot \varrho < 1 - a$ for all $\theta \notin (J_\varrho)_{\varepsilon_k}$. We can then choose $a$ to be deterministic such that

$$\mu\left(v_\theta \cdot \varrho < 1 - a \text{ for all } \theta \notin (J_\varrho)_{\varepsilon_k}\right) > 1 - \delta/3\ . \quad (3.46)$$

By the shape theorem there exists $M_0$ such that $M \geq M_0$ implies

$$\mathbb{P}(\tau(0,x) \geq g(x)(1 - a/2) \text{ for all } x \text{ with } \|x\|_1 \geq M) > 1 - \delta/3\ .$$

The marginal of $\mu$ on $\Omega_1$ is $\mathbb{P}$ so this holds with $\mu$ in place of $\mathbb{P}$. By part 2 of Proposition 3.7.1,

$$\mu(f(x) \geq g(x)(1 - a/2) \text{ for all } x \text{ with } \|x\|_1 \geq M \text{ and } 0 \to x) > 1 - \delta/3\ . \quad (3.47)$$

Choose $C > 0$ such that $\|x\|_1 \leq Cg(x)$ for all $x \in \mathbb{R}^2$. This is possible by (3.5). By Theorem 3.6.3, there exists $M_1 \geq M_0$ such that $M \geq M_1$ implies

$$\mu\left(|f(x) - x \cdot \varrho| < \frac{a}{2C}\|x\|_1 \text{ for all } x \text{ with } \|x\|_1 \geq M\right) > 1 - \delta/3\ .$$

This implies that for $M \geq M_1$,

$$\mu\left(|f(x) - x \cdot \varrho| < \frac{a}{2} g(x) \text{ for all } x \text{ with } \|x\|_1 \geq M\right) > 1 - \delta/3 \ . \tag{3.48}$$

We claim that the intersection of the events in (3.46), (3.47) and (3.48) implies the event in (3.45). Indeed, take a configuration in the intersection of the three events for some $M \geq M_1$. For a contradiction, assume there is an $x \in S_0$ with $\arg x \notin (J_\varrho)_{\varepsilon_k}$ and $\|x\|_1 \geq M$. Then

$$(x/g(x)) \cdot \varrho < 1 - a \text{ by } (3.46) \ .$$

However, since the event in (3.47) occurs and $\|x\|_1 \geq M$,

$$f(0, x) \geq g(x)(1 - a/2) \ .$$

Last, as the event in (3.48) occurs,

$$f(0, x) < x \cdot \varrho + \frac{a}{2} g(x) \ .$$

Combining these three inequalities,

$$g(x)(1 - a/2) \leq x \cdot \varrho + (a/2)g(x) < g(x)(1 - a) + (a/2)g(x) \ ,$$

or $g(x)(1 - a/2) < g(x)(1 - a/2)$, a contradiction. This completes the proof.

□

## 3.8 Proof of results on $\mathbb{Z}^2$

In this section, we apply the subsequential limit construction of the past sections to prove the claimed results about directional properties of geodesics on $\mathbb{Z}^2$.

### 3.8.1 Proof of Theorem 3.3.3

Suppose that $\partial\mathcal{B}$ is differentiable at $v_\theta = \varpi$ and construct the measure $\mu$ as in Section 3.5.1. Using the notation of Theorem 3.7.3, we set

$$L_\varrho = \{x \in \mathbb{R}^2 : x \cdot \varrho = 1\} \ .$$

From the theorem, we deduce that with $\mu$-probability 1, $\Gamma_0$ is asymptotically directed in $J_\varrho$. But by the assumption of differentiability, $J_\varrho = I_\theta$ with $\mu$-probability 1 and thus

$$\mu\left(\Gamma_0 \text{ is asymptotically directed in } I_\theta\right) = 1 \ . \tag{3.49}$$

By Proposition 3.7.1, each finite piece of $\Gamma_0$ is a geodesic, so $\Gamma_0$ is an infinite geodesic. Define $\hat{\Omega} \subseteq \Omega_1$ as the set

$$\hat{\Omega} = \{\omega \in \Omega_1 : \mu(\Gamma_0 \text{ is asymptotically directed in } I_\theta \mid \omega) = 1\} \ .$$

The inner probability measure is the regular conditional probability measure. The set $\hat{\Omega}$ is measurable and because the marginal of $\mu$ on $\Omega_1$ is $\mathbb{P}$, it satisfies $\mathbb{P}(\hat{\Omega}) = 1$. Further, for each $\omega \in \hat{\Omega}$ there is an infinite geodesic from 0 which is asymptotically directed in $I_\theta$.

### 3.8.2 Proof of Theorem 3.3.8

In this section we assume either **A1'** or **A2'**. Assume that the limit shape $\mathcal{B}$ has uniformly positive curvature. Then the boundary $\partial\mathcal{B}$ cannot contain any straight line segments. This implies that the extreme points $ext(\mathcal{B})$ are dense in $\partial\mathcal{B}$. Choose some countable set $D \subseteq ext(\mathcal{B})$ that is dense in $\partial\mathcal{B}$. For any $\theta_1$ and $\theta_2$ with $0 < dist(\theta_1, \theta_2) < \pi$, let $I(\theta_1, \theta_2)$ be the set of angles corresponding to the shorter closed arc of $\partial\mathcal{B}$ from $v_{\theta_1}$ to $v_{\theta_2}$. By Corollary 3.3.5, for each $\theta_1, \theta_2 \in D$ with $0 < dist(\theta_1, \theta_2) < \pi$, with probability one there is an infinite geodesic from 0 asymptotically directed in $I(\theta_1, \theta_2)$. The collection of such sets of angles is countable, so there exists an event $\Omega' \subseteq \Omega_1$ such that $\mathbb{P}(\Omega') = 1$ and for each $\omega \in \Omega'$,

1. for each $\theta_1, \theta_2 \in D$ such that $dist(\theta_1, \theta_2) < \pi$, there exists an infinite geodesic containing 0 and asymptotically directed in $I(\theta_1, \theta_2)$ and

2. for each $x, y \in \mathbb{Z}^2$ there is exactly one geodesic from $x$ to $y$.

We claim that for each $\omega \in \Omega'$, both statements of the theorem hold: for each $\theta$ there is an infinite geodesic with asymptotic direction $\theta$ and each infinite geodesic has a direction.

To prove the first statement, let $\omega \in \Omega'$ and $\theta \in [0, 2\pi)$. For distinct angles $\theta_1$ and $\theta_2$ such that $0 < dist(\theta_i, \theta) < \pi$ we write $\theta_1 >_\theta \theta_2$ if $I(\theta_1, \theta)$ contains $\theta_2$. Because $D$ is dense in $\partial\mathcal{B}$, we can find two sequences $(\theta_n^1)$ and $(\theta_n^2)$ such that (a) $0 < dist(\theta_n^i, \theta) < \pi$ for all $n$ and $i$, (b) for $i = 1, 2$, $dist(\theta_n^i, \theta) \to 0$ as $n \to \infty$ and (c) for each $i = 1, 2$ and $n$, $\theta_n^j >_\theta \theta_{n+1}^j$. Let $v_n$ be the point $nv_\theta$ and let $\gamma_n$ be the geodesic from 0 to $v_n$. Define $\gamma$ as any subsequential limit of $(\gamma_n)$. By this we mean a path $\gamma$ such that for each finite subset $E$ of $\mathbb{R}^2$, the intersection $\gamma_n \cap E$ equals $\gamma \cap E$ for all large $n$. We claim that $\gamma$ has asymptotic direction $\theta$.

Let $\varepsilon > 0$ and choose $N$ such that $dist(\theta, \theta_N^j) < \varepsilon$ for $j = 1, 2$. Because $\omega \in \Omega'$, for $j = 1, 2$, we can choose an infinite geodesic $\gamma_N^j$ containing 0 with asymptotic direction in $I(\theta_N^j, \theta_{N+1}^j)$. Write $P$ for the union of $\gamma_N^1$ and $\gamma_N^2$. This complement of $P$ in $\mathbb{R}^2$ consists of two open connected components (as $P$ cannot contain a circuit). Because both paths are directed away from $\theta$, exactly one of these two components contains all but finitely many of the $nv_\theta$'s. Let $C_1$ be the union of $P$ with this component and let $C_2$ be the other component.

Choose $N_0$ so that $nv_\theta \in C_1$ for all $n \geq N_0$. We claim now that each finite geodesic $\gamma_n$ for $n \geq N_0$ is contained entirely in $C_1$. If this were not true, $\gamma_n$ would contain a vertex $z$ in $C_2$ and therefore it would cross $P$ to get from $z$ to $v_n$. Then if $w$ is any vertex on $\gamma_n \cap P$ visited by $\gamma_n$ after $z$, then there would be two different geodesics from 0 to $w$ and this would contradict unique passage times. Therefore, as $\gamma_n$ is contained in $C_1$ for all large $n$, so must $\gamma$. This implies that $\gamma$ is asymptotically directed in the set of angles within distance $\varepsilon$ of $\theta$ (for each $\varepsilon > 0$) and therefore has asymptotic direction $\theta$.

To prove the second statement choose $\omega \in \Omega'$ and let $\gamma$ be an infinite geodesic. If $\gamma$ does not have an asymptotic direction then, writing $x_n$ for the $n$-th vertex of $\gamma$, we can find an angle $\phi \in [0, 2\pi)$ such that $\phi$ is a limit point of $\{\arg x_n : n \geq 1\}$ (under the metric *dist*) but $(\arg x_n)$ does not converge to $\phi$. So there exists a number $\varepsilon$ with $0 < \varepsilon < \pi$ and a subsequence $(x_{n_k})$ of $(x_n)$ such that for each $m$, $dist(\arg x_{n_{2m}}, \phi) < \varepsilon/2$ but $dist(\arg x_{n_{2m+1}}, \phi) > \varepsilon$. By the first part of the theorem we can find infinite geodesics $\gamma_1$ and $\gamma_2$ from 0 such that $\gamma_1$ has asymptotic direction $\phi + 3\varepsilon/4$ and $\gamma_2$ has asymptotic direction $\phi - 3\varepsilon/4$. Now it is clear that if we write $P$ for the union of $\gamma_1$ and $\gamma_2$ then $\gamma$ must both contain infinitely many vertices of $P$ and infinitely many vertices of $P^c$. This again contradicts unique passage times.

*Proof of Corollary 3.3.9.* If $\theta$ is an exposed point of differentiability then by Corollary 3.3.4, with probability one there exists an infinite geodesic from 0 in each rational direction. Then the proof above goes through with minor modifications. □

## 3.9 Three technical results

In this section, we prove three results whose proofs had previously been delayed. In the first subsection, we prove that fractional planes have a boundary which is a single doubly infinite dual path. In the second, we show that geodesics always exist on fractional planes if $\mathbb{P}$ is i.i.d. In the third, we prove that on the full plane, the measures $\mu_\alpha$ are measurable in $\alpha$ (in the sense that $\mu_\alpha(A)$ is measurable in $\alpha$).

### 3.9.1 Dual edge boundary of $V$

Consider first-passage percolation on a fractional plane $(V, E)$.

For any set $V_1 \subseteq \mathbb{Z}^2$, let $F$ be the edge boundary of $V_1$:

$$F = F(V_1) = \{\{x, y\} : x \in V_1, y \in V_1^c\} .$$

**Proposition 3.9.1.** *Let $V_1 \subseteq \mathbb{Z}^2$ be infinite, connected and such that $V_1^c$ is infinite and connected. The dual edge set $F^*$ consists of a single doubly infinite dual path which is non-self intersecting. That is, it is connected, infinite and each dual vertex $v^*$ in $W^*$, the set of endpoints of dual edges in $F^*$, has degree exactly 2 in the connected infinite graph $G^* = (W^*, F^*)$.*

*Proof.* Assume first that $G^*$ has a cycle. We can then extract from this cycle a self-avoiding one, whose parametrization yields a Jordan curve. This curve must contain a vertex of $\mathbb{Z}^2$ in its interior, showing that either $V_1$ or $V_1^c$ must be finite, a contradiction.

Next we prove that each dual vertex $v^* \in W^*$ has degree 2 in $G^*$. If $v^*$ has degree 1, then it has one incident dual edge $e^* \in F^*$ and this is dual to an edge $e \in F$. One endpoint of $e$ is in $V_1$ and one is in $V_1^c$, but they can be connected outside of $F$ using the 3 other edges dual to those which have $v^*$ as an endpoint, a contradiction. This means each $v^* \in W^*$ has degree at least 2 in $G^*$. However if $v^*$ has degree at least 3 in $G^*$ then three such dual edges $e_1^*, e_2^*$ and $e_3^*$ incident to $v^*$ are the first edges of disjoint self-avoiding infinite dual paths $P_1, P_2, P_3$. These paths split $\mathbb{Z}^2$ into at least 3 components, violating the fact that $(\mathbb{Z}^2, \mathcal{E}^2) \setminus F$ has two components.

Last we must show that $G^*$ is connected. Indeed, if $G^*$ were not connected, it would have two components $G_1^*, G_2^*$ (and possibly others). Since each dual vertex of $G_i^*$ must have degree two, and since there can be no cycles, $G_1^*$ and $G_2^*$ must be disjoint, self-avoiding, doubly infinite dual paths. But this breaks $\mathbb{Z}^2$ into at least three components, a contradiction. $\square$

### 3.9.2 Existence of Geodesics

Consider first-passage percolation on a fractional plane $(V, E)$. In this section, we prove that if $\mathbb{P}$ is a product measure and $x$ and $y$ are arbitrary vertices of $V$, then there almost surely exists a (finite) geodesic between $x$ and $y$. For $V = \mathbb{Z}^2$ this was proved by Wierman and Reh [95]; for general $d$, this appears to be open (see the remark under Theorem 8.1.8 in [98]). The proof will rely on the following "partial shape theorem."

**Lemma 3.9.2.** *Assume that $\mathbb{P}(\omega_e = 0) < 1/2$. Then, with probability one,*

$$\liminf_{\|x\|_1 \to \infty} \frac{\tau(0,x)}{\|x\|_1} > 0 \ .$$

*Proof.* Because $(V, E)$ is a subgraph of $(\mathbb{Z}^2, \mathcal{E}^2)$, it suffices to show the lemma in the first-passage model on $\mathbb{Z}^2$. So let $(\omega_e)$ be a passage time realization on $\mathcal{E}^2$ and define the truncated $\hat{\omega}_e = \min\{\omega_e, 1\}$, with $\hat{\tau}$ the passage time in the environment $(\hat{\omega}_e)$. Then by the shape theorem (see [78, Theorem 1] and the references therein), the lemma holds for $\hat{\tau}$. However $\tau \geq \hat{\tau}$ so we are done. □

**Theorem 3.9.3.** *Let $x$ and $y$ be elements of $V$. Then, almost surely, there exists a geodesic $\gamma : x \rightsquigarrow y$.*

*Proof.* The proof will be broken up into two cases, depending on the probability that $\omega_e = 0$. In both cases, we will show that if we write for $N \in \mathbb{N}$,

$$\tau_N(x,y) = \min_{\substack{\gamma : x \rightsquigarrow y \\ \gamma \subseteq \left(x + [-N,N]^2\right) \cap V}} \tau(\gamma) \ ,$$

then

$$\mathbb{P}(\tau_N(x,y) = \tau(x,y) \text{ for all large } N) = 1 \ . \tag{3.50}$$

This suffices to prove the theorem, as a function on a finite set attains its minimum.

**Case I:** $\mathbb{P}(\omega_e = 0) < 1/2$. In this case, we fix some deterministic path $\gamma_0$ in $V$ connecting $x$ and $y$ and define $N = N(\tau(\gamma_0))$ to be the smallest number such that

$$\min_{z \in V \setminus (x + [-N,N]^2)} \tau(x,z) > \tau(\gamma_0) \ .$$

Note that $N$ is almost surely finite by Lemma 3.9.2. Then no path containing a vertex of $V \setminus (x + [-N,N]^2)$ can have passage time less than or equal to $\tau(x,y)$. In particular, (3.50) holds.

**Case II:** $\mathbb{P}(\omega_e = 0) \geq 1/2$. Choose a deterministic $N_0 > 1$ such that there exists a path connecting $x$ and $y$ lying entirely in $[-N_0, N_0]^2 \cap V$. We will consider $\mathbb{P}$ to actually be defined on $\mathbb{R}^{\mathcal{E}^2}$, though of course the weights of edges outside of $E$ will have no bearing on the first-passage model in $(V, E)$.

Consider a sequence of annuli $A_n \subseteq \mathbb{R}^2$ of the form

$$A_n = [-N_0^{n+1}, N_0^{n+1}]^2 \setminus (-N_0^n, N_0^n)^2;$$

denote by $G_n$ the event that there is a (vertex) self-avoiding circuit $\alpha$ in $A_n$ of edges $e$ such that $\omega_e = 0$. By the RSW theorem for independent percolation (see [20, Section 3.1]), we have

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} G_n\right) = 1 \ .$$

For any $N \in \mathbb{N}$ write $L_N = N_0^{N+1}$. For a given $\omega$ such that $G_N$ occurs, choose $\alpha$ as above and consider it as a continuous plane curve. Further, let $\gamma$ be any vertex self-avoiding path in $(V, E)$ from $x$ to $y$. We will show that there exists another path $\gamma'$ in $[-L_N, L_N]^2$ from $x$ to $y$ such that $\tau(\gamma') \leq \tau(\gamma)$. This suffices to complete the proof. To do so, we use the following construction. Let $\beta$ be any path from $x$ to $y$ in $(V, E)$ lying entirely in $[-N_0, N_0]^2$ . Since $\gamma$ intersects $\beta$ at $x$ and $y$ we may list their common vertices in order (along $\gamma$) as $x = x_1, \ldots, x_k = y$. We proceed along

$\gamma$ from each $x_i$ to $x_{i+1}$, calling this subpath $\gamma_i$. If $\gamma_i$ is not just one edge of $\beta$, we create a Jordan curve $C$ by concatenating the portion of $\beta$ from $x_i$ to $x_{i+1}$ with $\gamma_i$. If $\alpha$ intersects the interior of $C$ then we choose any common point $p$ and proceed in both directions along $\alpha$ from it. In each direction we must meet $C$ again; otherwise $\alpha$ was in the interior of $C$, which is false. Furthermore we meet $C$ before we meet $\Gamma$, since $\Gamma$ is in the exterior of $C$. Therefore the component of $\alpha \cap \text{int } C$ containing $p$ is a segment of $\alpha$ from some vertex $a$ to another $b$. Since $a$ and $b$ are in $C$ they must be in $\gamma_i$ and we can replace the segment of $\gamma_i$ from $a$ to $b$ with this segment of $\alpha$. In this way we obtain a new path we call $\tilde{\gamma}_i$ and corresponding Jordan curve $\tilde{C}$. Note that $\tau(\tilde{\gamma}_i) \leq \tau(\gamma_i)$. See Figure 3.3 for a depiction of this procedure.

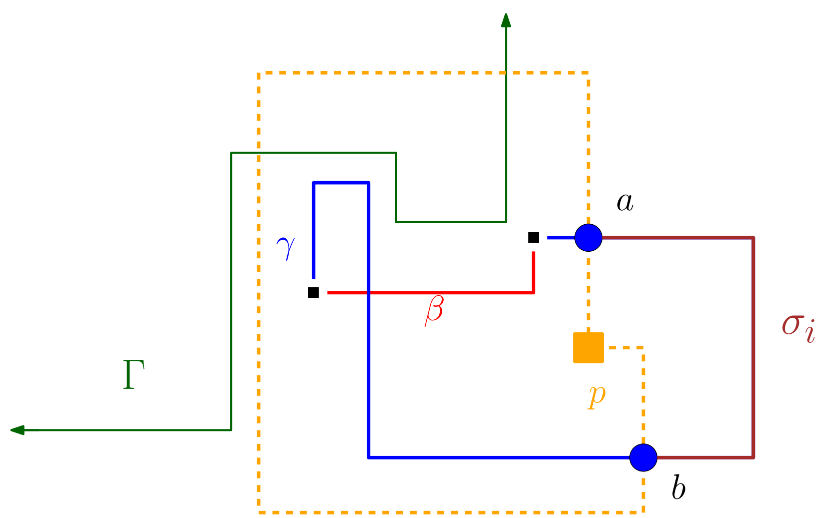



Figure 3.3: Modifying the path $\gamma$ by replacing a segment $\sigma_i$ of $\gamma$ with a segment of $\alpha$. In the figure, $\alpha$ is the dotted path in yellow and $p$ is a point on $\alpha$ in the interior of $C$, the Jordan curve formed by the union of $\gamma_i$ with $\beta$.

It remains to show that the procedure defined above eventually terminates in some path $\hat{\gamma}_i$ and Jordan curve $\hat{C}$. At this point $\alpha$ will not intersect the interior of $\hat{C}$, implying that $\hat{\gamma}_i$ does not leave $[-L_N, L_N]^2$. To prove this, assume that $p \in \alpha \cap \text{int } C$ and define $a$ and $b$ as above. Let $\sigma_i$ be the segment of $\gamma_i$ from $a$ to $b$. If $\sigma_i$ does not leave $\alpha$ then it must be the complementary segment of $\alpha$ from $a$ to $b$, implying that $\alpha \subset (C \cup \text{int } C)$. Then int $\alpha \subset$ int $C$, a contradiction, since $\beta$ is in the interior of

$\alpha$. Therefore we can find some edge adjacent to $\alpha$ in $\sigma_i$. When we construct $\hat{\gamma}_i$, we remove this edge from $\gamma_i$ and only add edges of $\alpha$. Since there are only finitely many edges adjacent to $\alpha$, the process terminates. □

### 3.9.3 Measurability of $\alpha \mapsto \mu_\alpha(A)$.

Consider first-passage percolation on $\mathbb{Z}^2$ with a distribution $\mathbb{P}$ satisfying **A1** or **A2**. In this section we show that for all Borel measurable $A \subseteq \widetilde{\Omega}$, $\alpha \mapsto \mu_\alpha(A)$ is Lebesgue measurable. By the monotone class theorem, it suffices to consider the case that $A$ is a cylinder event; that is, that there exists $M > 0$ such that $A$ depends only on passage times $\omega_e$, Busemann increments $(\theta_1(v), \theta_2(v))$ and graph variables $\eta(f)$ for vertices $v$ in $[-M, M]^2$, and edges $e$ and directed edges $f$ with both endpoints in $[-M, M]^2$. Recall that for $\alpha \in \mathbb{R}$,

$$\hat{L}_\alpha = \{x \in \mathbb{Z}^2 : x + [-1/2, 1/2)^2 \cap L_\alpha \neq \varnothing\}$$

and that passage times to $L_\alpha$ are actually defined to $\hat{L}_\alpha$. We are interested in how this set changes near $[-M, M]^2$ as we vary $\alpha$. For this reason, define for each $v \in \mathbb{Z}^2$

$$C_v^- = \inf\{\alpha : v \in \hat{L}_\alpha\} \text{ and } C_v^+ = \sup\{\alpha : v \in \hat{L}_\alpha\}\ .$$

It follows that for all $v$, $C_v^- < C_v^+$ and

$$v \in \begin{cases} \hat{L}_\alpha & \text{if } \alpha \in (C_v^-, C_v^+) \\ \hat{L}_\alpha^c & \text{if } \alpha \in \mathbb{R} \setminus [C_v^-, C_v^+] \end{cases} .$$

Define the set

$$X = \cup_{v \in \mathbb{Z}^2} \{C_v^-, C_v^+\}$$

and note that $X$ is countable. To prove Lebesgue measurability of $\alpha \mapsto \mu_\alpha(A)$, we show that

$$f(\alpha) := \mu_\alpha(A) \text{ is continuous except at } \alpha \in X \; . \tag{3.51}$$

Let $\alpha \in [0,n] \setminus X$ and let $\varepsilon > 0$. For any integer $N \geq M$ such that $[-N,N]^2$ intersects $\hat{L}_\alpha$ let $\mathcal{P}_N$ be the collection of all lattices paths whose vertices are in $[-N,N]^2$. Last define the approximate passage times for $x \in [-N,N]^2$

$$\tau_N(x, L_\alpha) = \min_{\substack{x \in \gamma \in \mathcal{P}_N \\ \gamma \cap \hat{L}_\alpha \neq \varnothing}} \tau(\gamma)$$

and geodesics $G_N(x, L_\alpha)$ to be the minimizing paths. Let $G(x, L_\alpha)$ be the original geodesic from $x$ to $L_\alpha$. Using the shape theorem, we can choose $N$ large enough that

$$\mathbb{P}\left( \min_{\substack{v \in [-M,M]^2 \\ w \notin (-N,N)^2}} \tau(v,w) > \max_{v \in [-M,M]^2} \tau(v, L_\alpha) \right) \geq 1 - \varepsilon \; . \tag{3.52}$$

For $N$ fixed as above, the condition that $\alpha \notin X$ implies that we can choose $\delta > 0$ such that the interval $(\alpha - \delta, \alpha + \delta)$ is contained in the complement of the finite set

$$X_N = \cup_{v \in [-N,N]^2} \{C_v^-, C_v^+\} \; .$$

It follows that

$$\text{for all } \beta \text{ with } |\alpha - \beta| < \delta, \; \hat{L}_\alpha \cap [-N,N]^2 = \hat{L}_\beta \cap [-N,N]^2 \; . \tag{3.53}$$

Having fixed $\delta$ above we now prove that if $|\beta - \alpha| < \delta$ then $|\mu_\alpha(A) - \mu_\beta(A)| < \varepsilon$. Using the definition of $\Phi_\alpha$ we can first give an upper bound

$$|\mu_\alpha(A) - \mu_\beta(A)| \leq \mathbb{P}(\Phi_\alpha^{-1}(A) \Delta \Phi_\beta^{-1}(A)) \; , \tag{3.54}$$

where $\Delta$ is the symmetric difference operator. Note that the events on the right side are determined by (a) $\omega_e$ for $e$ with both endpoints in $[-M, M]^2$, (b) the geodesics $G(x, L_\alpha)$ and $G(x, L_\beta)$ from all points $x \in [-M, M]^2$ to the lines $L_\alpha$ and $L_\beta$ and (c) the passage times of these geodesics. Therefore the right side of (3.54) is bounded above by

$$\mathbb{P}(\exists\ x \in [-M, M]^2 \text{ such that } G(x, L_\alpha) \neq G(x, L_\beta))\ .$$

However if such an $x$ exists then by (3.53), one of two geodesics must exit the box $[-N, N]^2$. A subpath of this geodesic must cross from $[-M, M]^2$ to the complement of $(-N, N)^2$, so the event $E(M, N)$ in (3.52) cannot occur. Thus

$$|\mu_\alpha(A) - \mu_\beta(A)| \leq \mathbb{P}(E(M, N)^c) < \varepsilon \text{ if } |\beta - \alpha| < \delta\ ,$$

so $f$ is continuous at $\alpha$, giving measurability of $f$.

# Chapter 4

# Coalescence in geodesic graphs

This chapter is devoted to presenting new results on the coalescence of unigeodesics constructed using the methods of Chapter 3. The theorems are differ from those of [72] by being "in construction." Rather than assuming the existence of directed unigeodesics and showing that these unigeodesics coalesce, the theorems below show coalescence for a constructed family of unigeodesics on $(V, E)$. As a biproduct of these results, it will be shown that our construction does not produce bigeodesics on $\mathbb{Z}^2$. One can regard our theorems as partial answers to questions **Q3**, and **Q4** of Sections 2.4.2 and 2.4.3.

The theorems are proven for the families of unigeodesics we have constructed as limits of point-to-point geodesics on fractional planes and point-to-plane geodesics on $\mathbb{Z}^2$. The coalescence results below require that our measure $\mathbb{P}$ be translation-invariant, at least under $T_{\mathbf{e}_1}$. As such, we are forced to specialize our fractional plane results to the case of upper half-planes.

Recall the formal definition of coalescence: we say that two infinite directed paths $\Gamma$ and $\Gamma'$ coalesce if their (edge) symmetric difference is finite.

## 4.1 Finite Energy and Half-Planes

For the main theorems on coalescence we need an extra assumption in the case that $\mathbb{P}$ is not a product measure. It allows us to apply "edge modification" arguments. Write $\omega = (\omega_e, \check{\omega})$, where $\check{\omega}_f = (\omega)_{f \neq e}$.

**Definition 4.1.1.** *We say that $\mathbb{P}$ has the upward finite energy property if for each $\lambda > 0$ such that $\mathbb{P}(\omega_e \geq \lambda) > 0$,*

$$\mathbb{P}\left(\omega_e \geq \lambda \,\middle|\, \check{\omega}\right) > 0 \quad \textit{almost surely} . \tag{4.1}$$

Note that if $\mathbb{P}$ is a product measure, it has the upward finite energy property.

### 4.1.1 The Half-Plane

The previous results on fractional planes $(V, E)$ required very few assumptions on $\mathbb{P}$. In our specialization to the half-plane, we will need a slight strengthening of our assumptions, which we detail here.

Considering the vertex set $V = V_H = \{(x_1, x_2) \in \mathbb{Z}^2 : x_2 \geq 0\}$ and $E_H$ the induced set of edges, we can analyze first-passage percolation more closely on $\mathbb{H} = (V_H, E_H)$, taking advantage of translation invariance of standard measures. The relevant space is $\Omega_H = [0, \infty)^{E_H}$. Note that the family of translation operators $\{T_x : x \in V_H\}$ on $\Omega_H$ (see Section 3.1) are well-defined on $\mathbb{H}$.

For the results on $\mathbb{H}$ we will consider a probability measure $\mathbb{P}$ satisfying one of two assumptions, labeled **B1** and **B2** below. Assumption **B2** includes the upward finite energy property defined above.

The assumptions we need are:

**B1**. $\mathbb{P}$ is a product measure with continuous marginals OR

**B2**. $\mathbb{P}$ is the restriction to $[0,\infty)^{E_H}$ of a Borel probability measure $\hat{\mathbb{P}}$ on $[0,\infty)^{\mathcal{E}^2}$ that satisfies the upward finite energy property and the assumptions of Hoffman [51]:

(a) $\hat{\mathbb{P}}$ is ergodic relative to the translations $T_x$ for $x \in \mathbb{Z}^2$,

(b) $\hat{\mathbb{P}}$ has all the symmetries of $\mathbb{Z}^2$,

(c) $\int \omega_e^{2+\alpha}\, \mathrm{d}\hat{\mathbb{P}} < \infty$ for some $\alpha > 0$,

(d) $\hat{\mathbb{P}}$ has unique passage times: with probability one, no two (edge) nonempty distinct paths have the same passage time and

(e) the limiting shape for $\hat{\mathbb{P}}$ is bounded.

Under parts (a)-(c) of assumption **B2**, Kingman's theorem implies that if we write $\tau'$ for the passage time in $\mathbb{Z}^2$ then for each $y \in \mathbb{Z}^2$, the limit $g(y) = \lim_{n\to\infty} \tau'(0,ny)/n$ exists almost surely and in $L^1$. Part (e) of assumption **B2** is then the statement that $\inf_{y\neq 0} \frac{g(y)}{\|y\|_1} > 0$.

Under either of these assumptions, one can show that $\mathbb{P}$ admits geodesics. Under **B1**, we show it in Section 3.9.2. Under **B2** it follows from the shape theorem proved by Boivin [19] and boundedness of the limit shape. This means we can use the results from the previous chapter.

## 4.2 Results

Recall the two methods of producing geodesic graphs from the previous chapter.

### 4.2.1 Result on $\mathbb{H}$

Our first result is a general theorem about the coalescence structure of the limiting geodesic graph on $\mathbb{H}$. For the statement of the main theorem, we use the shorthand $x \to y$ for vertices $x, y$ in a directed graph $\vec{G}$ if there is a directed path from $x$ to $y$ in $\vec{G}$.

**Theorem 4.2.1.** *Assume* **B1** *or* **B2**. *Writing $x_n = (n, 0)$, the geodesic graphs $(\mathbb{G}_n)$ converge almost surely to a directed graph $\mathbb{G}$ with the following properties:*

1. *each vertex in $V_H$ has out-degree 1,*
2. *viewed as an undirected graph, $\mathbb{G}$ has no circuits,*
3. *for each $x \in V_H$, the backward cluster $B_x = \{y \in V_H : y \to x\}$ is finite and*
4. *writing $\Gamma_x$ for the unique self-avoiding infinite directed path in $\mathbb{G}$ starting from $x$, for all $x, y \in V_H$, $\Gamma_x$ and $\Gamma_y$ coalesce. That is, their edge symmetric difference is finite.*

Note that the finiteness of $B_x$ has been proven under stronger assumptions by Wehr-Woo [93]. We therefore relegate the proof of this fact to Section 4.4. We modify their arguments to account for the fact that, without their moment assumption on $\omega_e$, their large deviations estimate no longer holds.

### 4.2.2 Results on $\mathbb{Z}^2$

We return to the setting of the full plane $(\mathbb{Z}^2, \mathcal{E}^2)$. As our graph measure $\mu$ was constructed as a subsequential limit, we lack a statement analogous to the convergence of finite geodesics on the half-plane. However, our results still allow us to construct some limiting family of geodesics, and we can prove similar coalescence results.

**Theorem 4.2.2.** *Assume either* **A1'** *or both* **A2'** *and the upward finite energy property. Let $v \in \mathbb{R}^2$ be any nonzero vector and for $\beta \in \mathbb{R}$ define*

$$L_\beta(v) = \{y \in \mathbb{R}^2 : y \cdot v = \beta\} \ .$$

*There exists an event $\mathcal{A}$ with $\mathbb{P}(\mathcal{A}) = 1$ such that for each $\omega \in \mathcal{A}$, the following holds. There exists an ($\omega$-dependent) increasing sequence $(\alpha_k)$ of real numbers with $\alpha_k \to \infty$ such that $\mathbb{G}_{L_{\alpha_k}(v)}(\omega) \to G(\omega)$, a directed graph with the following properties.*

1. *Viewed as an undirected graph, $G$ has no circuits.*

2. *Each $x \in \mathbb{Z}^2$ has out-degree 1 in $G$.*

3. *(All geodesics coalesce.) Write $\Gamma_x$ for the unique infinite path in $G$ from $x$. If $x, y \in \mathbb{Z}^2$ then $\Gamma_x$ and $\Gamma_y$ coalesce.*

4. *(Backward clusters are finite.) For all $x \in \mathbb{Z}^2$, the set $\{y \in \mathbb{Z}^2 : y \to x \text{ in } G\}$ is finite.*

Our last theorem deals with coalescence and asymptotic directions. Before stating it, we note its relation to the results of Licea and Newman [72]. The result reduces the complement of the set $D$ (see Section 2.4.2) to be empty for existence of coalescing geodesics (item 1 above). It however does not address uniqueness; in principle, different subsequences can produce families of geodesics which do not coalesce. We reduce the finite exponential moment condition of [72], extend to non-i.i.d. measures and replace the global curvature assumption with a directional condition. Without this condition, part 3 gives the existence of coalescing geodesics directed in sectors. For the statement, recall the definition of $I_\theta$ in (3.12).

**Theorem 4.2.3.** *Assume either* **A1'** *or both* **A2'** *and the upward finite energy property. Let $\theta \in [0, 2\pi)$.*

1. *If $\partial\mathcal{B}$ is differentiable at $v_\theta$ then with probability one there exists a collection $\{\gamma_x : x \in \mathbb{Z}^2\}$ of infinite geodesics in $\omega$ such that*

   (a) *each $x$ is a vertex of $\gamma_x$;*

   (b) *each $\gamma_x$ is asymptotically directed in $I_\theta$;*

   (c) *for all $x, y \in \mathbb{Z}^2$, $\gamma_x$ and $\gamma_y$ coalesce and*

   (d) *each $x$ is on $\gamma_y$ for only finitely many $y$.*

2. *If $v_\theta$ is an exposed point of differentiability of $\mathcal{B}$ then the above geodesics all have asymptotic direction $\theta$.*

3. *Suppose $\theta_1 \neq \theta_2$ are such that $v_{\theta_1}$ and $v_{\theta_2}$ are extreme points of $\mathcal{B}$. If $\Theta$ is the set of angles corresponding to some arc of $\partial\mathcal{B}$ connecting $v_{\theta_1}$ to $v_{\theta_2}$ then the above geodesics can be taken to be asymptotically directed in $\Theta$.*

Theorems 4.2.2 and 4.2.3 follow from a stronger result. In Sections 3.7 and 4.5, we prove that any subsequential limit $\mu$ defined as in Section 3.5.1 is supported on geodesic graphs with properties 1-4 of Theorem 4.2.2.

**Remark 4.2.4.** *The finiteness of backward clusters in the graphs produced in the previous two theorems (see item 4 of the first and item 1(d) of the second) and in Theorem 4.2.1 is the aforementioned connection to nonexistence of bigeodesics. It shows that when constructing infinite geodesics using our limiting procedure, it is impossible for doubly infinite paths to arise.*

## 4.3 Geodesics graphs on $\mathbb{H}$

In this section we prove Theorem 4.2.1. In what follows, we will consider first-passage percolation under $\mathbb{H}$ with a measure $\mathbb{P}$ satisfying **B1** or **B2**.

Because $\mathbb{P}$ admits geodesics, Theorem 3.3.2 implies that the sequence of graphs $(\mathbb{G}_n)$ converge almost surely to a directed graph $\mathbb{G}$, each of whose directed paths is a geodesic. As $\mathbb{P}$ also has unique passage times, Proposition 3.2.2 states that each vertex of $\mathbb{G}_n$ has out-degree one and there are no undirected circuits, so these same properties survive in the limit for $\mathbb{G}$. The finiteness of backward clusters is a consequence of non-existence of bigeodesics in the half-plane, proved by Wehr and Woo [93]. Unfortunately this result was only proved under **B1** with the additional assumption $\mathbb{E}\omega_e < \infty$, so we provide a proof in Section 4.4 under either **B1** or **B2**.

This section is devoted to showing coalescence of directed paths in $\mathbb{G}$. Because each vertex in $\mathbb{G}_H$ has out-degree one, it suffices to show that each $\Gamma_v$ and $\Gamma_w$ share a vertex. The main difficulty will be proving this statement for all $v, w$ on the first coordinate axis; that is, the set $L_0$, where

$$\text{for } k \in \mathbb{N} \cup \{0\},\ L_k := \{(x, k) : x \in \mathbb{Z}\}\ .$$

To see why this implies coalescence for all paths, assume we have proved this statement and note that it suffices then to show that for all $v, w \in V_H$ with $w \in L_0$, the geodesics $\Gamma_v$ and $\Gamma_w$ coalesce. Write $v = (v_1, v_2)$ and consider the set

$$\tilde{L}_v = \{(v_1, y) \in V_H : 0 \le y \le v_2\}\ .$$

With probability one, for each $v' \in \tilde{L}_v$, the backward cluster $B_{v'}$ is finite. Thus we can find $m, n \in \mathbb{Z}$ with $m < v_1 < n$ such that for all $v' \in \tilde{L}_v$, both points $(m, 0)$ and $(n, 0)$ are not in $B_{v'}$. This means in particular that $\Gamma_{(m,0)}$ and $\Gamma_{(n,0)}$ cannot intersect $\tilde{L}_v$ and, since they coalesce, they must meet "above" $v$. In other words, $v$ is in the bounded component of $V_H \setminus (\Gamma_{(m,0)} \cup \Gamma_{(n,0)})$ (viewing these paths only as their vertex sets). By planarity, $\Gamma_v$ must intersect $\Gamma_{(m,0)}$. Because $\Gamma_{(m,0)}$ coalesces with $\Gamma_w$, this completes the proof.

So we move to proving coalescence starting from the first coordinate axis. We will prove by contradiction, so assume either **B1** or **B2** but that

$$\text{with positive probability, there are vertices } v, w \in L_0 \text{ with } \Gamma_v \cap \Gamma_w = \emptyset\ . \tag{4.2}$$

### 4.3.1 Estimates on density of disjoint geodesics

#### Definitions

For each $k \in \mathbb{N} \cup \{0\}$ and $m, n \in \mathbb{Z}$ with $m < n$ define $N_{m,n}^{(k)}$ as the largest number $N$ such that we can find vertices $v_1, \ldots, v_N \in [m, n] \times \{k\}$ such that

(a) $\Gamma_{v_1}, \ldots, \Gamma_{v_N}$ are pairwise disjoint and

(b) for all $i$, $\Gamma_{v_i} \cap [L_0 \cup \cdots \cup L_k] = \{v_i\}$.

Similarly, for $k \in \mathbb{N}$ let $M_{m,n}^{(k)}$ be the largest $M$ such that we can find $v_1, \ldots, v_M \in [m, n] \times \{k\}$ such that (a) and (b) above hold but also (c) for all $i = 1, \ldots, M$, no $v \in L_0$ has $\Gamma_v \cap \Gamma_{v_i} \neq \emptyset$.

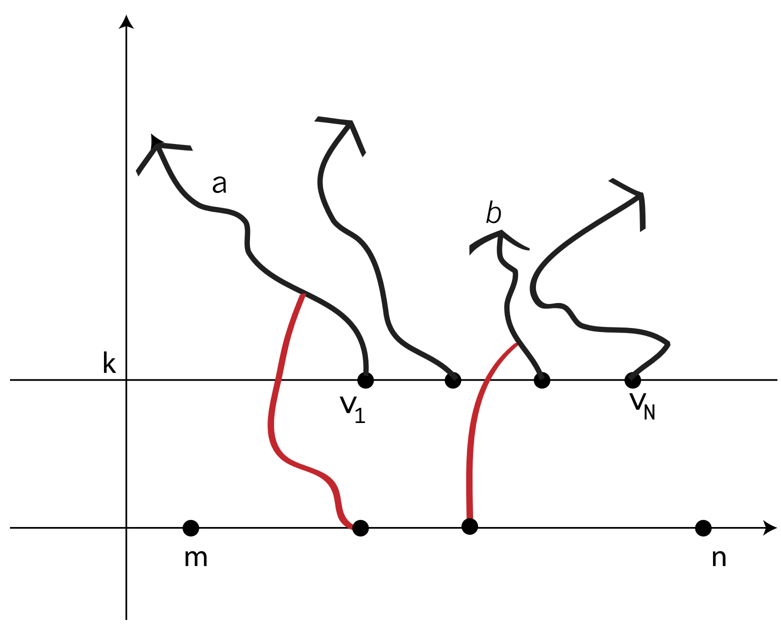



Figure 4.1: In this example $N_{m,n}^{(k)}$ is at least 4. The black paths are geodesics emanating from vertices on the line $L_k$. They do not intersect each other and they intersect $L_k$ only at their initial points. The red paths are segments of geodesics starting from $L_0$. Note that in this example the paths a and $b$ do not contribute to the random variable $M_{m,n}^{(k)}$.

**Lemma 4.3.1.** *For each $k_1 \in \mathbb{N} \cup \{0\}$ and $k_2 \in \mathbb{N}$, there exist deterministic $\alpha_{k_1}, \beta_{k_2} \geq 0$ such that*

$$\lim_{n\to\infty} \frac{N_{0,n}^{(k_1)}}{n} = \alpha_{k_1} \text{ and } \lim_{n\to\infty} \frac{M_{0,n}^{(k_2)}}{n} = \beta_{k_2} \text{ almost surely and in } L^1(\mathbb{P}) \ .$$

*We have the characterization*

$$\alpha_{k_1} = \inf_{n\in\mathbb{N}} \frac{\mathbb{E}N_{0,n}^{(k_1)}}{n} \text{ and } \beta_{k_2} = \inf_{n\in\mathbb{N}} \frac{\mathbb{E}M_{0,n}^{(k_2)}}{n} \ .$$

*Furthermore, assuming* (4.2), $\alpha_0 > 0$.

*Proof.* Note that for all $m < n < p$ in $\mathbb{Z}$ and $k_1 \in \mathbb{N} \cup \{0\}$, $k_2 \in \mathbb{N}$, we have

$$N_{m,p}^{(k_1)} \leq N_{m,n}^{(k_1)} + N_{n,p}^{(k_1)} \text{ and } M_{m,p}^{(k_2)} \leq M_{m,n}^{(k_2)} + M_{n,p}^{(k_2)} \ .$$

Further $\max\{N_{m,n}^{(k_1)}, M_{m,n}^{(k_2)}\} \leq n - m + 1$ surely, so they have finite mean and $(N_{m,n}^{(k_1)}, M_{m,n}^{(k_2)})$ has the same distribution as $(N_{0,n-m}^{(k_1)}, M_{0,n-m}^{(k_2)})$. Therefore we can apply Kingman's subadditive ergodic theorem to find deterministic $\alpha_{k_1}, \beta_{k_2} \geq 0$ such that

$$\frac{1}{n} N_{0,n}^{(k_1)} \to \alpha_{k_1} \text{ and } \frac{1}{n} M_{0,n}^{(k_2)} \to \beta_{k_2} \text{ almost surely and in } L^1(\mathbb{P}) \ .$$

Furthermore, $\alpha_{k_1} = \inf_{n\in\mathbb{N}} \mathbb{E}N_{0,n}^{(k_1)}/n$ and $\beta_{k_2} = \inf_{n\in\mathbb{N}} \mathbb{E}M_{0,n}^{(k_2)}/n$.

We claim now that under assumption (4.2), $\alpha_0 > 0$. By countability and invariance of $\mathbb{P}$ under $T_{(1,0)}$, we can find $i_0 \in \mathbb{N}$ such that $\mathbb{P}(A(1, i_0)) > 0$, where $A(1, i_0)$ is the event that $\Gamma_{(1,0)}$ and $\Gamma_{(i_0,0)}$ do not intersect. Note that if $i_1 < i_2 < i_3 < i_4$ are integers such that $\Gamma_{(i_l,0)}$ and $\Gamma_{(i_{l+1},0)}$ are disjoint for $l = 1, 3$, then by planarity, at least three of them must be disjoint. So the ergodic theorem implies that with probability one, $A(1, i_0) \circ T_{(j,0)}$ occurs for infinitely many $j$ and therefore we can find 4 geodesics starting from $L_0$ that are all disjoint. Clearly at least two of these must intersect $L_0$

only finitely often. This implies that for some $j_0 \in \mathbb{N}$, $\mathbb{P}(B(1,j_0)) > 0$, where $B(1,j_0)$ is the event that $\Gamma_{(1,0)}$ and $\Gamma_{(j_0,0)}$ do not intersect and only touch $L_0$ at their initial points.

Again, by the ergodic theorem,

$$\frac{1}{N}\sum_{l=0}^{N} T^l_{(j_0,0)} 1_{B(1,j_0)} \to \mathbb{P}(B(1,j_0)) \text{ almost surely and in } L^1(\mathbb{P}) \ .$$

The reasoning given above, but applied to sets $\{j_1, j_2, \ldots\}$ of size bigger than 4, implies that for $n \in \mathbb{N}$,

$$N^{(0)}_{0,j_0 n} - 1 \geq \sum_{l=0}^{n} T^l_{(j_0,0)} 1_{B(1,j_0)} \ .$$

Dividing by $j_0 n$ and taking $n \to \infty$, we find $\alpha_0 \geq \mathbb{P}(B(1,j_0))/j_0 > 0$. □

**Lower bound on $\alpha_k$**

**Proposition 4.3.2.** *For each $k \in \mathbb{N}$, $\alpha_k \geq \beta_k + \alpha_0$.*

*Proof.* For the proof we need a lemma stating that any geodesic starting at $L_0$ intersects $L_k$ only finitely often.

**Lemma 4.3.3.** *Assume* (4.2). *For each $v \in L_0$ and $k \in \mathbb{N}$, with probability one, the set $\Gamma_v \cap L_k$ is finite.*

*Proof.* Assume that there exists $k \in \mathbb{N}$ such that with positive probability, there exists $v \in L_0$ with $\Gamma_v \cap L_k$ infinite. By countability and invariance of $\mathbb{P}$ under $T_{(1,0)}$,

$$\mathbb{P}(B) > 0 \ , \text{ where } B = \left\{\#(\Gamma_{(0,0)} \cap L_k) = \infty\right\} \ .$$

By Lemma 4.3.1, we can find $N_0 \in \mathbb{N}$ such that

$$\mathbb{P}(N^{(0)}_{1,N_0+1} > k+2) > 1 - \mathbb{P}(B)/2$$

and then by translation invariance, with positive $\mathbb{P}$-probability, the event $B \cap \{N^{(0)}_{1,N_0+1} > k+2\} \cap \{N^{(0)}_{-1-N_0,-1} > k+2\}$ occurs. However any outcome in this event must have contradictory properties, as we now explain. Since $B$ occurs, $\Gamma_{(0,0)}$ must intersect infinitely many vertices of either $L_k \cap \{(x,y) : x \geq 0\}$ or $L_k \cap \{(x,y) : x \leq 0\}$. Let us assume the first; the subsequent argument is similar in the other case. Then $\Gamma_{(0,0)}$ must be disjoint from at least $k+1$ different geodesics $\Gamma_{v_1}, \ldots, \Gamma_{v_{k+1}}$ with $v_i \in L_0 \cap [1, N_0+1]$ for all $i$, but it must intersect some vertex $(x,k)$ for $x > N_0$. By planarity, the geodesics $\Gamma_{v_i}$ must all intersect the set $\{(x,j) : 0 \leq j \leq k\}$, but then they cannot be disjoint. This is a contradiction. $\square$

Returning to the proof of Proposition 4.3.2, fix $k \in \mathbb{N}$. For each $m \in \mathbb{Z}$, define $d_k(m)$ as the first coordinate of the last vertex (by the natural ordering) on $\Gamma_{(m,0)}$ in the line $L_k$. This quantity exists almost surely by Lemma 4.3.3. For any $a, b \in \mathbb{Z}$ with $a < b$, define the set

$$X_{a,b} = \{j \in \mathbb{Z} : d_k(j) \in [a,b]\} \ .$$

We claim that for some fixed $N_0 \in \mathbb{N}$,

$$\mathbb{P}\left(X_{-N_0,n+N_0} \text{ contains } [0,n] \text{ for infinitely many } n \in \mathbb{N}\right) \geq 1/2 \ . \tag{4.3}$$

To show this, first choose $N_0 \in \mathbb{N}$ such that $\mathbb{P}(|d_k(0)| \leq N_0) \geq 3/4$. Next note that by invariance of $\mathbb{P}$ under $T_{(1,0)}$, $\mathbb{P}(d_k(n) \leq n + N_0) \geq 3/4$ for all $n \in \mathbb{N}$. These two events occur simultaneously with probability at least 1/2, so

$$\mathbb{P}\left(d_k(0) \geq -N_0 \text{ and } d_k(n) \leq n + N_0 \text{ for infinitely many } n \in \mathbb{N}\right) \geq 1/2 \ .$$

Last, observe that by planarity, the function $m \mapsto d_k(m)$ is monotonic. This implies that if $d_k(0) \geq -N_0$ and $d_k(n) \leq n + N_0$ then the set $X_{-N_0,n+N_0}$ contains $[0, n]$.

The second step is to prove that

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{N_{0,n}^{(k)} - M_{0,n}^{(k)}}{n} \geq \alpha_0\right) \geq 1/4 \ . \tag{4.4}$$

Because $(N_{0,n}^{(k)} - M_{0,n}^{(k)})/n$ converges almost surely to $\alpha_k - \beta_k$, this suffices to complete the proof of the proposition. First, given $\epsilon > 0$, by Lemma 4.3.1, pick $N_1$ such that

$$\mathbb{P}\left(N_{0,n}^{(0)}/n \geq \alpha_0 - \epsilon \text{ for all } n \geq N_1\right) \geq 3/4 \ .$$

On this event, for $n \geq N_1$, setting $a_n = \lfloor n(\alpha_0 - \epsilon)\rfloor$, we may find $x_1^{(n)}, \ldots, x_{a_n}^{(n)} \in [0, n]$ such that the geodesics $\Gamma_{(x_1^{(n)},0)}, \ldots, \Gamma_{(x_{a_n}^{(n)},0)}$ are pairwise disjoint. If, in addition, the event in (4.3) occurs, then for infinitely many $n$, all of $d_k(x_1^{(n)}), \ldots, d_k(x_{a_n}^{(n)})$ are in $[-N_0, n+N_0]$. Note that the geodesics emanating from each of the points $(d_k(x_i^{(n)}), k)$ are disjoint and do not intersect $L_0 \cup \cdots \cup L_k$ except for their initial vertices. Next, choose a maximal set $\hat{\Gamma}_1^{(n)}, \ldots, \hat{\Gamma}_{M_{-N_0,n+N_0}^{(k)}}^{(n)}$ of geodesics starting in $[-N_0, n+N_0] \times \{k\}$ which are disjoint, intersect $L_0 \cup \cdots \cup L_k$ only at their initial vertices, and such that no $v \in L_0$ has $\Gamma_v \cap \hat{\Gamma}_i^{(n)} \neq \emptyset$ for $i = 1, \ldots, M_{-N_0,n+N_0}^{(k)}$. Note that these $\hat{\Gamma}$'s are disjoint from the geodesics starting from the points $(d_k(x_i^{(n)}), k)$. Therefore for each $n \geq N_1$, with probability at least 1/4 we have $N_{-N_0,n+N_0}^{(k)} \geq a_n + M_{-N_0,n+N_0}^{(k)}$. Thus

$$\mathbb{P}\left(N_{-N_0,n+N_0}^{(k)} \geq a_n + M_{-N_0,n+N_0}^{(k)} \text{ for infinitely many } n\right) \geq 1/4 \ .$$

By invariance of $\mathbb{P}$ under $T_{(1,0)}$,

$$\mathbb{P}\left(N^{(k)}_{0,n+2N_0} - M^{(k)}_{0,n+2N_0} \geq a_n \text{ for infinitely many } n\right) \geq 1/4 \ .$$

Finally, as $(n+2N_0)/n \to 1$ as $n \to \infty$ and $\epsilon$ is arbitrary, (4.4) holds. □

**Upper bound on $\alpha_k$**

In this section we combine the lower bound from last section with an upper bound to conclude that $\beta_k = 0$. In what follows, we will denote by $G(x,y)$ the unique geodesic between $x$ and $y$.

**Proposition 4.3.4.** *For $k \in \mathbb{N}$, $\alpha_k \leq \alpha_0$. Therefore $\beta_k = 0$.*

We will couple together the upper half-plane with shifted half-planes. For any $k \in \mathbb{N}$ we consider the shifted configuration $T_{(0,k)}\omega$ and the unique geodesics $G(v,(n,0))$ in this configuration. Specifically, for any $\omega \in \Omega_H$ and $v \in V_H^k = \{(x,y) \in V_H : y \geq k\}$, we set

$$G_n^{(k)}(v) = T_{(0,-k)}\left[G(v-(0,k),(n,0))(T_{(0,k)}\omega)\right] \ , \tag{4.5}$$

where for a path $\gamma$ in $\mathbb{H}$ we denote by $T_{(0,-k)}\gamma$ the path $\gamma$ shifted up by $k$ units. By Theorem 3.3.2, there is an almost sure limit $G^{(k)}(v) = \lim_{n\to\infty} G_n^{(k)}(v)$.

**Lemma 4.3.5.** *Let $k \in \mathbb{N}$. With probability one, for all $v \in L_k$, if $\Gamma_v \cap [L_0 \cup \cdots \cup L_{k-1}] = \emptyset$ then*

$$\Gamma_v = G^{(k)}(v) \ .$$

*Proof.* Let $v \in L_k$ such that $\Gamma_v \cap [L_0 \cup \cdots \cup L_{k-1}] = \emptyset$ and write it as $v = (v_1, v_2)$. Let $\sigma$ be the non-self intersecting continuous curve obtained by concatenating (a) the edges of $\Gamma_v$, (b) the vertical line segment connecting $(v_1, -1/2)$ and $v$ and (c) the ray $\{(x,-1/2) \in \mathbb{R}^2 : x \geq v_1\}$. One component of the complement of $\sigma$ contains all vertices of $L_{k-1}$ to the right of $v-(0,1)$ and the other contains all vertices of $L_{k-1}$

to the left of $v-(0,1)$; call the first $C_1$ and the second $C_2$. Because the sequence $G(v,(n,0))$ converges to $\Gamma_v$ as $n\to\infty$, there exists $N_0$ such that if $n\geq N_0$ then $G(v,(n,0))$ does not contain any vertices of the form $(v_1,y)$ for $y<v_2$. For $n\geq N_0$ the geodesic $G(v,(n,0))$ cannot contain any vertices in $C_2$. For if it did, it would start at $v$, go through a vertex in $C_2$, and then touch $(n,0)$, a vertex in $C_1$. Because this geodesic cannot cross $\{(v_1,y):y<v_2)\}$, it must cross $\Gamma_v$ and violate unique passage times.

For $n\geq N_0$, let $w_n$ denote the first intersection of $G(v,(n,0))$ with $L_{k-1}$. The vertex $v_n$ directly before this must be in $L_k$ and the segment $\gamma_n$ of $G(v,(n,0))$ from $v$ to $v_n$ has all vertices in $V_H^k$. Therefore writing $v_n=(a_n,k)$, we have $\gamma_n=G^{(k)}_{a_n}(v)$. Because $\Gamma_v$ does not intersect $L_0\cup\cdots\cup L_{k-1}$, $\|w_n\|_1\to\infty$. However $w_n$ is in $C_1$, so $a_n\to+\infty$. Taking $n$ to infinity, these segments converge to $G^{(k)}(v)$. However they converge to $\Gamma_v$. □

For $n\in\mathbb{N}$, choose $r=N^{(k)}_{0,n}$ pairwise disjoint geodesics $\Gamma_{v_1},\ldots,\Gamma_{v_r}$ for $v_1,\ldots,v_r\in[0,n]\times\{k\}$ such that for each $i=1,\ldots,r$, $\Gamma_{v_i}\cap[L_0\cup\cdots\cup L_k]=\{v_i\}$. By Lemma 4.3.5, $r\leq N^{(0)}_{0,n}(T_{(0,k)}(\omega))$. Therefore

$$\frac{N^{(k)}_{0,n}(\omega)}{n}\leq\frac{N^{(0)}_{0,n}(T_{(0,k)}(\omega))}{n}\text{ for all }n\in\mathbb{N}\ .$$

Taking $n\to\infty$ and using invariance of $\mathbb{P}$ under $T_{(0,k)}$, we find $\alpha_k\leq\alpha_0$.

### 4.3.2 Deriving a contradiction

In this section we will show that assuming (4.2), there exists $k\geq 1$ such that $\beta_k>0$. This will contradict Proposition 4.3.4 and complete the proof of coalescence starting from the first-coordinate axis.

#### Lemmas for edge modification

The first lemma will let us apply an edge modification argument. For a typical element $\omega$ and edge $e \in E_H$ we write $\omega = (\omega_e, \check{\omega})$. We say an event $A \subset \Omega_H$ is *e-increasing* if, for all $(\omega_e, \check{\omega}) \in A$ and $r > 0$, $(\omega_e + r, \check{\omega}) \in A$. The following lemma is a consequence of the finite energy property.

**Lemma 4.3.6.** *Let $\lambda > 0$ be such that $\mathbb{P}(\omega_e \geq \lambda) > 0$. If $A \subset \Omega_H$ is e-increasing with $\mathbb{P}(A) > 0$ then*

$$\mathbb{P}(A,\ \omega_e \geq \lambda) > 0 \ .$$

*Proof.* We estimate

$$\begin{aligned}\mathbb{P}(A,\ \omega_e \geq \lambda) &= \mathbb{E}\left[\mathbb{E}[1_A(\omega_e, \check{\omega}) 1_{\{\omega_e \geq \lambda\}} \mid \check{\omega}]\right] \\ &\geq \mathbb{E}\left[1_A(\lambda, \check{\omega}) \mathbb{P}(\omega_e \geq \lambda \mid \check{\omega})\right] \ .\end{aligned}$$

Because $A$ is $e$-increasing, the variable $1_A 1_{\{\omega_e \leq \lambda\}}$ is less than or equal to the random variable $1_A(\lambda, \check{\omega})$. Therefore if the statement of the lemma is false then $1_A(\lambda, \check{\omega})$ is positive on a set of positive probability. By the upward finite energy property, $\mathbb{P}(\omega_e \geq \lambda \mid \check{\omega})$ is positive almost surely, so the above estimates give $\mathbb{P}(A,\ \omega_e \geq \lambda) > 0$, a contradiction. □

The second lemma is a shape theorem-type upper bound. For it, we define

$$\lambda_0^+ = \sup\{\lambda \geq 0 : \mathbb{P}(\omega_e \geq \lambda) > 0\} \ . \tag{4.6}$$

**Lemma 4.3.7.** *Suppose that $\lambda_0^+ < \infty$. There exists $c^+ < \lambda_0^+$ such that*

$$\mathbb{P}\left(\tau(0, x) \leq c^+ \|x\|_1 \text{ for all but finitely many } x \in V_H\right) = 1 \ .$$

*Proof.* Because $\mathbb{P}$ has unique passage times, the marginal of $\omega_e$ is not concentrated at a point and therefore $\mathbb{E}\omega_e < \lambda_0^+$. For any $x \in V_H$ choose a deterministic path $\gamma_x : 0 \rightsquigarrow x$ in $\mathbb{H}$ with $\|x\|_1$ number of edges. Then

$$\mathbb{E}\tau(0,x) \leq \mathbb{E}\tau(\gamma_x) = \|x\|_1 \mathbb{E}\omega_e \ .$$

We now set $c^+ = \frac{\mathbb{E}\omega_e + \lambda_0^+}{2}$ and argue that this value satisfies the condition of the lemma. The argument will be similar to the proof of the shape theorem in the full space.

For any $z \in \mathbb{Q}^2$ with second coordinate non-negative, let $N$ be any natural number such that $Nz \in V_H$. Then for $n \in \mathbb{N}$, write $n = \lfloor \frac{n}{N} \rfloor + r$, where $0 \leq r < N$ and estimate

$$\tau(0,nz) \leq N\lambda_0^+ \|z\|_1 + \sum_{i=0}^{\lfloor \frac{n}{N} \rfloor - 1} \tau(0,Nz)(T_{Nz}^i \omega) \ .$$

Divide by $n$ and use the ergodic theorem to find

$$\limsup_{n\to\infty} \frac{\tau(0,nz)}{n} \leq \frac{\mathbb{E}\tau(0,Nz)}{N} \leq \|z\|_1 \mathbb{E}\omega_e \ . \tag{4.7}$$

Let $\Omega_H'$ be the full-probability event on which (4.7) holds for all $z \in \mathbb{Q}^2$ with second coordinate non-negative. Assume by way of contradiction that on some positive probability event $A$, the lemma does not hold for the $c^+$ fixed above. Then we can find $\omega \in A \cap \Omega_H'$; we will show that this $\omega$ has contradictory properties.

Let $(z_n)$ be a sequence of vertices in $V_H$ such that $\|z_n\|_1 \to \infty$ and

$$\tau(0,z_n) > c^+ \|z_n\|_1 \text{ for all } n \in \mathbb{N} \ .$$

By compactness (and by restricting to a subsequence), given a positive $a$ such that $a\lambda_0^+ < c^+ - \mathbb{E}\omega_e$, we can find some $z \in \mathbb{Q}^2$ with second coordinate non-negative and

$$\|z\|_1 = 1 \text{ such that } \left\| \frac{z_n}{\|z_n\|_1} - z \right\|_1 < a \text{ for all } n \in \mathbb{N} .$$

Then we can estimate

$$\tau(0, z_n) \leq \tau(0, \|z_n\|_1 z) + \tau(\|z_n\|_1 z, z_n) \leq \tau(0, \|z_n\|_1 z) + \left\| \|z_n\|_1 z - z_n \right\|_1 \lambda_0^+ .$$

Therefore

$$c^+ < \frac{\tau(0, z_n)}{\|z_n\|_1} \leq \frac{\tau(0, \|z_n\|_1 z)}{\|z_n\|_1} + \left\| z - \frac{z_n}{\|z_n\|_1} \right\|_1 \lambda_0^+ .$$

Taking limsup on the right side gives $c^+ \leq \mathbb{E}\omega_e + a\lambda_0^+$, a contradiction. □

The final lemma deals with spatial concentration of geodesics emanating from the first coordinate axis. For $v_1, v_2, v_3 \in L_0$ let $B(v_1, v_2, v_3)$ be the event that

1. the geodesics $\Gamma_{v_1}, \Gamma_{v_2}$ and $\Gamma_{v_3}$ are disjoint,

2. they intersect $L_0$ only at their initial points, and

3. they intersect each $L_k$ in finitely many vertices.

**Lemma 4.3.8.** *Suppose $v_1 = (x_1, 0), v_2 = (x_2, 0)$ and $v_3 = (x_3, 0)$ with $x_1 < x_2 < x_3$. Let $B^G(v_1, v_2, v_3)$ be the subevent of $B(v_1, v_2, v_3)$ on which for each $\epsilon > 0$, there are infinitely many $k \in \mathbb{N}$ such that the last intersections $\zeta_k$ and $\zeta_k'$ of $\Gamma_{v_1}$ and $\Gamma_{v_3}$ with $L_k$ satisfy $\|\zeta_k - \zeta_k'\|_1 < \epsilon k$. Then*

$$\mathbb{P}(B^G(v_1, v_2, v_3) \mid B(v_1, v_2, v_3)) = 1 .$$

*Proof.* For $z \in L_0$ and $k \in \mathbb{N}$, denote by $\zeta_k(z)$ the last point of intersection of $\Gamma_z$ with $L_k$, which exists almost surely on $B(v_1, v_2, v_3)$ by Lemma 4.3.3. By translation invariance we will take $v = v_3 - v_1$ and consider

$$C_k(v) = \{\|\zeta_k(v) - \zeta_k(0)\|_1 \geq \epsilon k\} \ .$$

For $k, n \in \mathbb{N}$, define $X_n^{(k)} = \sum_{j=0}^{n-1} 1_{C_k(v)}(T_{(jd,0)}(\omega))$, where $d = \|v\|_1 + 1$. By the ergodic theorem, putting $p_k = \mathbb{P}(C_k(v))$,

$$X_n^{(k)}/n \to p_k \text{ almost surely} \ . \tag{4.8}$$

As previously in the paper, for $l \in \mathbb{Z}$ and $k \in \mathbb{N}$, define $d_k(l)$ as the first coordinate of $\zeta_k(l)$ and note that by planarity, $d_k(l)$ is monotone in $l$. Therefore for $n \in \mathbb{N}$, the difference $d_k(nd) - d_k(0)$ is at least equal to $\epsilon k X_n^{(k)}$, so

$$\frac{d_k(nd) - nd - d_k(0)}{n} \geq \frac{\epsilon k X_n^{(k)} - nd}{n} = \epsilon k X_n^{(k)}/n - d \ .$$

Combining with (4.8), almost surely,

$$\liminf_{n \to \infty} \frac{d_k(nd) - nd - d_k(0)}{n} \geq \epsilon k p_k - d \ .$$

Because $d_k(nd) - nd$ and $d_k(0)$ have the same distribution, $(d_k(nd) - nd - d_k(0))/n \to 0$ in probability. Therefore

$$p_k \leq d/(\epsilon k) \ ,$$

giving $p_k \to 0$. In particular, with probability one, $C_k(v)^c$ occurs for infinitely many $k$. □

**Main argument**

We will first assume that $\lambda_0^+ < \infty$ and that (4.2) holds. By Proposition 4.3.2, $\alpha_0 > 0$ and so we can find $v_1, v_2, v_3$ and $p > 0$ such that $\mathbb{P}(B(v_1, v_2, v_3)) \geq p$, where this event was defined before Lemma 4.3.8. Fix any positive

$$\epsilon < \frac{\lambda_0^+ - c^+}{8\lambda_0^+} \ . \tag{4.9}$$

We first define a modified event which combines conditions from the previous section. Specifically, for $k \in \mathbb{N}$ we set $B'(k) = B'(v_1, v_3; k)$ as the event that

1. the geodesics $\Gamma_{v_1}$ and $\Gamma_{v_3}$ are disjoint and intersect $L_j$ in a finite set for all $j \in \mathbb{N} \cup \{0\}$,

2. writing $w_1 = w_1(k)$ and $w_3 = w_3(k)$ for the last intersections of $\Gamma_{v_1}$ and $\Gamma_{v_3}$ with $L_k$, there is a vertex $x^*$ in $L_k$ between $w_1$ and $w_3$ such that $\Gamma_{x^*}$ is disjoint from $\Gamma_{v_1}$ and $\Gamma_{v_3}$ and $\Gamma_{x^*}$ intersects $L_k$ only at $x^*$,

3. the finite geodesics $r_1(k)$ and $r_3(k)$, defined as the segments of $\Gamma_{v_1}, \Gamma_{v_3}$ from $L_0$ to each of $w_1$ and $w_3$ satisfy $\tau(r_i(k)) \leq c^+\|v_i - w_i\|_1$ for $i = 1, 3$ and

4. $\|w_1 - w_3\|_1 < \epsilon k$.

The first two conditions hold together for all $k$ simultaneously with probability at least $p$. This is because whenever $B(v_1, v_2, v_3)$ occurs, almost surely each $\Gamma_{v_i}$ intersects each $L_k$ in a finite set, so we can let $x^*$ be the last intersection point of $\Gamma_{v_2}$ with $L_k$. Next, by Lemma 4.3.7 we can find $k_0$ such that

$$\mathbb{P}(\tau(v_i, w) \leq c^+\|v_i - w\|_1 \text{ for all } i = 1, 3 \text{ and } w \in \cup_{k=k_0}^{\infty} L_k) > 1 - p/2 \ .$$

This implies that the first three conditions hold for all $k \geq k_0$ with probability at least $p/2$. Using Lemma 4.3.8,

$$\mathbb{P}(B'(k)) > 0 \text{ for infinitely many } k \geq k_0 \ . \tag{4.10}$$

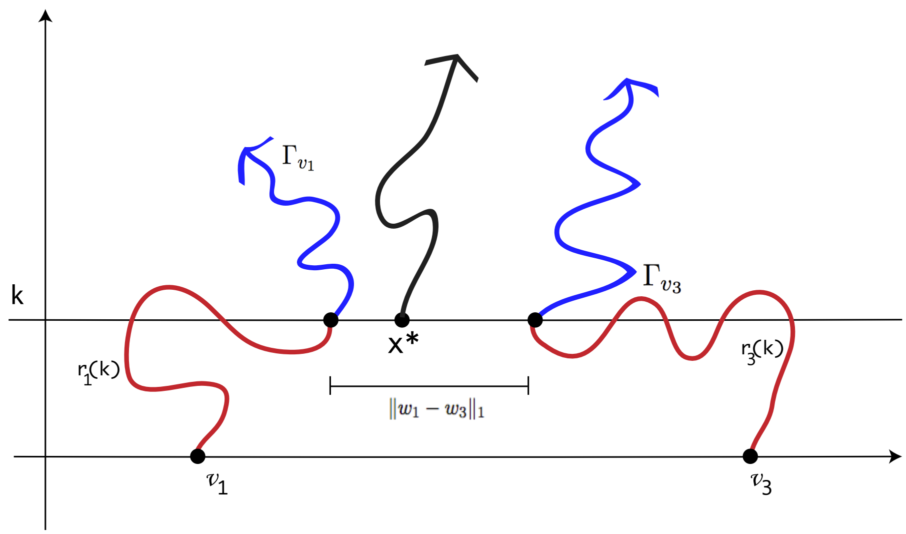



Figure 4.2: The event $B'(k)$. The geodesics $\Gamma_{v_i}, i = 1, 3$, are the concatenation of blue and red paths. The black geodesic $\Gamma_{x^*}$ does not intersect either $\Gamma_{v_1}$ or $\Gamma_{v_3}$ and intersects $L_k$ only at $x^*$. The red paths satisfy $\tau(r_i(k)) \leq c^+\|v_i - w_i\|_1$ while $\|w_1 - w_3\|_1 < \epsilon k$.

We then fix any such $k \geq k_0$ with

$$4\|v_3 - v_1\|_1 \lambda_0^+ < \frac{\lambda_0^+ - c^+}{2} k \ . \tag{4.11}$$

Next we modify the edge-weights for a set of edges between the geodesics $\Gamma_{v_1}$ and $\Gamma_{v_3}$. For any configuration $\omega$ in $B'(k)$ write $X_1$ for the closed subset of $\mathbb{R}^2$ with boundary curves $\Gamma_{v_1}, \Gamma_{v_3}$ and the segment of the first coordinate axis between $v_1$ and $v_3$. Let $X_2$ be the component of $X_1 \cap \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq k\}$ containing $v_1$. Last, define the set $X \subset E_H$ consisting of all edges not in $\Gamma_{v_1}$ or $\Gamma_{v_3}$ but such that both endpoints are in $X_2$. Because there are only countably many choices, (4.10) implies there is a deterministic choice $X'$ and a vertex $y \in L_k$ such that

$$\mathbb{P}(B'(k),\ X = X',\ x^* = y) > 0 \ . \tag{4.12}$$

Here the notation $x^* = y$ means that the (deterministic) vertex $y$ satisfies condition 2 of the definition of $B'(k)$.

We next show that

$$\mathbb{P}\left(B'(k),\ X = X',\ x^* = y,\ \bigcap_{e \in X'} \left\{\omega_e \geq \frac{c^+ + \lambda_0^+}{2}\right\}\right) > 0\ . \tag{4.13}$$

To prove this we enumerate the edges $e_1, \ldots, e_r$ of $X'$ and repeatedly apply Lemma 4.3.6. By (4.12), we simply need to verify that

$$B'(k) \cap \{X = X',\ x^* = y\} \cap \bigcap_{i=1}^{j-1} \left\{\omega_{e_i} \geq \frac{c^+ + \lambda_0^+}{2}\right\} \text{ is } e_j\text{-increasing for all } j = 2, \ldots, r\ .$$

So take $\omega$ in the event on the left for some $j = 2, \ldots, r$ with $\omega'$ such that $\omega'_f = \omega_f$ for $f \neq e_j$ and $\omega'_{e_j} \geq \omega_{e_j}$. First we claim that $\Gamma_{v_1}$, $\Gamma_y$ and $\Gamma_{v_3}$ are unchanged from $\omega$ to $\omega'$. To see this, note that since $e_j$ is not in $\Gamma_{v_1}$, $\Gamma_y$ or $\Gamma_{v_3}$ we can find $n_1 = n_1(\omega)$ such that if $n \geq n_1$ then $e_j$ is also not in any of the geodesics $G(v_1, (n, 0))$, $G(y, (n, 0))$ or $G(v_3, (n, 0))$ in $\omega$. Therefore these remain geodesics in $\omega'$; taking the limit as $n \to \infty$ proves the claim. Now it is clear that $X = X'$ in $\omega'$ and conditions 1 - 4 of $B'(k)$ hold in $\omega'$. Obviously if $\omega_{e_i} \geq (1/2)(c^+ + \lambda_0^+)$ for $i = 1, \ldots, j-1$ in $\omega$ then this is still true in $\omega'$. This proves (4.13).

On the event in (4.13), no point $v \in L_0$ can have $\Gamma_v \cap \Gamma_y \neq \emptyset$. We will now argue for this fact and explain why it leads to a contradiction. If such a $v$ exists it must be on the segment of $L_0$ strictly between $v_1$ and $v_3$; this is a direct consequence of planarity and the fact that each vertex in $\mathbb{G}_H$ has out degree one. Therefore $\Gamma_v$ must start at $L_0$ and use only edges in $X'$ until its exit from $L_0 \cup \ldots \cup L_k$. Writing $w$ for the first vertex of $\Gamma_v$ in $L_k$, we must then have

$$\tau(v, w) \geq \frac{c^+ + \lambda_0^+}{2} \|v - w\|_1\ . \tag{4.14}$$

On the other hand, we can give an upper bound for the passage time from $v$ to $w$ by taking the path obtained by concatenating (a) the segment of $L_0$ from $v$ to $v_1$, (b) the geodesic $r_1$ and (c) the segment of $L_k$ from $w_1$ to $w$. We get the bound

$$\begin{aligned}\tau(v,w) &\leq \left[\|v_3 - v_1\|_1 + \epsilon k\right]\lambda_0^+ + c^+\|v_1 - w_1\|_1 \\ &\leq 2\left[\|v_3 - v_1\|_1 + \epsilon k\right]\lambda_0^+ + c^+\|v - w\|_1 \ .\end{aligned}$$

Combining this with (4.14), we find

$$(\lambda_0^+ - c^+)k \leq 4\left[\|v_3 - v_1\|_1 + \epsilon k\right]\lambda_0^+ \ .$$

This contradicts (4.9) and (4.11).

To summarize, we have now shown that for some fixed $w_1, w_2, w_3 \in L_k$ such that the segment of $L_k$ between $w_1$ and $w_3$ contains $w_2$, $C = C(w_1, w_2, w_3)$ has positive probability, where this event is defined by the conditions

1. $\Gamma_{w_1}, \Gamma_{w_2}$ and $\Gamma_{w_3}$ are disjoint and intersect $L_0 \cup \cdots \cup L_k$ only in $w_1$, $w_2$ and $w_3$ respectively and

2. no $v \in L_0$ has $\Gamma_{w_2} \cap \Gamma_v \neq \emptyset$.

Fix any $m, n \in \mathbb{Z}$ with $m < n$ and $w_1, w_3 \in [m,n] \times \{k\}$. Let $l \in \mathbb{N}$ be bigger than $\|w_3 - w_1\|_1$ and recall the notation $M_{m,n}^{(k)}$ from Section 4.3.1. Note that if $C \cap T_{(l,0)}C$ occurs then $M_{m,n+l}^{(k)} \geq 2$. Iterating this reasoning, for any $j \in \mathbb{N}$,

$$M_{m,n+jl}^{(k)}(\omega) \geq \sum_{i=0}^{j-1} 1_C(T_{(l,0)}^i \omega) \ .$$

Diving by $j$ and using the ergodic theorem gives $\beta_k > 0$, a contradiction. This proves that assumption (4.2) is false in the case $\lambda_0^+ < \infty$ and thus all geodesics starting from $L_0$ coalesce.

In the case that $\lambda_0^+ = \infty$ the argument is much easier and we will just explain the idea. If (4.2) holds then we still find $v_1, v_2, v_3$ in $L_0$ with $v_2$ in the segment of $L_0$

between $v_1$ and $v_3$ and such that the $\Gamma_{v_i}$'s are disjoint and intersect $L_0$ in only $v_1, v_2$ and $v_3$. Again pick $y$ as the last intersection point of $\Gamma_{v_2}$ with $L_1$. Letting $S$ be the set of edges touching any vertex of $L_0$ between $v_1$ and $v_3$ (and therefore not in $\Gamma_{v_1}$ or $\Gamma_{v_3}$), we then modify the edge-weights for edges in $S$ to be larger than some $C_{big} > 0$. Using Lemma 4.3.6 we can find $C_{big}$ large enough so that on this event, no vertex $v$ of $L_0$ can have $\Gamma_v \cap \Gamma_y \neq \emptyset$. As before, this implies $\beta_1 > 0$, a contradiction.

## 4.4 Absence of bigeodesics in $\mathbb{H}$

In this section we outline the modifications needed to carry over the proof of the main theorem of [93] to our setting. An infinite geodesic indexed by $\mathbb{Z}$ is called a bigeodesic. When we assume unique passage times, such a path is (vertex) self-avoiding.

### 4.4.1 Lemmas from Wehr-Woo

Assume either **B1** or **B2** and let $K^*$ be the event

$$K^* = \{\text{there exists a bigeodesic}\} \ .$$

Note that for all $x$, $\mathbb{P}(\#B_x = \infty, (K^*)^c) = 0$, where $B_x$ was defined in Theorem 4.2.1. By horizontal translation ergodicity, $\mathbb{P}(K^*)$ is zero or one; let us assume for a contradiction that $\mathbb{P}(K^*) = 1$.

Any bigeodesic $\gamma$ divides $\mathbb{R}^2 \setminus \gamma$ into two components, say $R^+ = R^+(\gamma)$ and $R^- = R^-(\gamma)$; that is,

$$R^+(\gamma) \cap R^-(\gamma) = \emptyset \ ,$$

$$R^+(\gamma) \cup R^-(\gamma) = \mathbb{R}^2 \setminus \gamma \ ,$$

$$\partial R^+ = \partial R^- = \gamma \ ,$$

where $R^-$ is a region that contains $(0,-1)$ and where $\partial A$ denotes the usual boundary of a set $A \subset \mathbb{R}^2$. Hence by unique passage times, for any points $x, y \in R^-(\gamma)$, no bond $b$ belonging to the finite geodesic $G(x,y)$ can be an element of $R^+(\gamma)$. The following is [93, Proposition 4].

**Proposition 4.4.1.** *Consider the sequence* $G((-n,0),(n,0))$ *for* $n \in \mathbb{N}$. *With probability 1, this sequence has a limit:*

$$\gamma_0 = \lim_{n\to\infty} G((-n,0),(n,0)) \ .$$

*Moreover,* $\gamma_0$ *is a bigeodesic and for any bigeodesic* $\gamma$*,*

$$\gamma_0 \subset \left[R^-(\gamma) \cup \gamma\right] \ .$$

*Proof.* The same proof as in [93] works here. The only assumption needed is that of unique passage times. □

The next is [93, Lemma 5].

**Lemma 4.4.2.** *Let* $n \in \mathbb{N}$ *and* $\mathbb{H}' = \{(x_1,x_2) \in \mathbb{R}^2 : x_2 \le n\}$. *With probability 1, for any bigeodesic* $\gamma$ *intersecting* $z = (z_1, z_2)$ *with* $z_2 < n$*,*

$$\mathbb{H}' \cap R^+(\gamma) \neq \emptyset \textit{ and all its components are bounded} \ .$$

*The boundary of each component is a self-avoiding loop, which is a bond-disjoint union of segments of* $\gamma$ *and segments of the boundary of* $\mathbb{H}'$.

*Proof.* Because we do not assume independence of the variables $(\omega_e)$, we must modify the proof of [93], replacing independence with the upward finite energy property.

In order to prove the boundedness of each component of $\mathbb{H}' \cap R^+(\gamma)$ it is sufficient to prove that

$$\mathbb{P}(\text{there is a bigeodesic with an infinite connected part in } \mathbb{H}') = 0 \ . \tag{4.15}$$

For each $k \in \mathbb{Z}$ consider a rectangular box

$$C_k = C_k(m,n) = \{(x_1,x_2) : 2km \leq x_1 \leq (2k+1)m,\ 0 \leq x_2 \leq n\} \ .$$

Let $T_k$ be the minimum passage time of all paths in $C_k$ which start at a vertex in the left boundary of $C_k$ and end at a vertex in the right boundary of $C_k$, without intersecting the top boundary. Let $\hat{C}_k$ for the set of edges in $\partial C_k$ that do not lie on the first coordinate axis; then set

$$E_k = \left\{ \sum_{e \in \hat{C}_k} \tau_e < T_k \right\} \ .$$

We claim that for some $m$ large enough, $\mathbb{P}(E_k) > 0$ for all $k$. To prove this, we consider two cases. Assume first that $\lambda_0^+$, defined in (4.6), is finite. Then by the ergodic theorem, writing $e_k = \{(k,0),(k+1,0)\}$, $(1/m)\sum_{k=0}^{m-1} \omega_{e_k} \to \mathbb{E}\omega_e$. Therefore, using the bound $\omega_e \leq \lambda_0^+$,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{e \in \hat{C}_0} \omega_e = \mathbb{E}\omega_e \ .$$

As $\mathbb{P}$ has unique passage times, $\mathbb{E}\omega_e < \lambda_0^+$, so choose $m$ such that

$$\mathbb{P}\left( \sum_{e \in \hat{C}_0} \omega_e < \frac{\mathbb{E}\omega_e + \lambda_0^+}{2} m \right) > 0 \ .$$

Writing $C_k^0$ for the set of edges with an endpoint in $C_k \setminus \hat{C}_k$, we see that the above event is $e$-increasing for all $e \in C_0^0$. So by Lemma 4.3.6,

$$\mathbb{P}\left(\sum_{e \in \hat{C}_0} \omega_e < \frac{\mathbb{E}\omega_e + \lambda_0^+}{2} m,\ \omega_f \geq \frac{\mathbb{E}\omega_e + \lambda_0^+}{2} \text{ for all } f \in C_0^0\right) > 0 \ .$$

On this event, each path which passes from the left to the right side of $C_0$, taking only edges in $C_0^0$, must have passage time at least $\frac{\mathbb{E}\omega_e + \lambda_0^+}{2} m$. So for such $m$, horizontal translation invariance gives $\mathbb{P}(E_k) > 0$.

In the case that $\lambda_0^+ = \infty$, the proof of $\mathbb{P}(E_k) > 0$ is easier. We simply modify the edge-weights for edges in $C_0^0$ to be larger than the sum of the boundary edge-weights with positive probability. In either case, the ergodic theorem shows that

$$\mathbb{P}(E_k \text{ occurs for infinitely many } k > 0 \text{ and } k < 0) = 1 \ .$$

For any $k$ such that $E_k$ occurs, no geodesic can pass from the left to the right side of $C_k$ taking only edges in $C_k^0$, because we can replace the segment between the left and right sides by a portion of the boundary $\partial C_0$. This shows (4.15). The rest of the lemma follows immediately. □

We now move to [93, Proposition 6], the main observation showing that unique passage times implies that $\gamma_0$ must intersect any large box with probability bounded below uniformly of the position of the box. For $l \in \mathbb{N}$, let us write $B = B(l) = [-l, l] \times [0, 2l]$ and let $K$ be the event that at least one bigeodesic intersects $B$. Define for $L \in \mathbb{N}$, translations of $B$ by

$$B_{i,j} = B_{i,j}(l, L) = B + (iL, jL) \text{ for } (i, j) \in V_H \ .$$

For $L > 2l$, the $B_{i,j}$ are mutually disjoint.

**Proposition 4.4.3.** *Let* $\delta = 1 - \mathbb{P}(K)$. *Then*

$$\mathbb{P}(B_{i,j} \subset R^+(\gamma_0)) \leq \delta \ ,$$

$$\mathbb{P}(B_{i,j} \subset R^-(\gamma_0)) \leq \delta \ .$$

*Proof.* The proof is the same as that in [93]. □

## 4.5 Coalescence on $\mathbb{Z}^2$

In this section we consider first-passage percolation on $\mathbb{Z}^2$ with a $\mathbb{P}$ satisfying **A1'** or **A2'** and the finite energy property. At the end of the section, we will prove the theorems of Section 4.2.2. As in the preceding chapter, the results will follow from properties of the subsequential limit measure $\mu$ and the graph $\mathbb{G}$ which is sampled from this measure. As such, the majority of this section will be devoted to an analysis of paths in $\mathbb{G}$.

We will first prove that all directed infinite paths coalesce in $\mathbb{G}$. Recall that under either **A1'** or **A2'**, for $x \in \mathbb{Z}^2$, $\Gamma_x$ is the unique infinite directed path in $\mathbb{G}$ starting at $x$.

**Theorem 4.5.1.** *Assume either* **A1'** *or both* **A2'** *and the upward finite energy property. With* $\mu$*-probability one, for each* $x, y \in \mathbb{Z}^2$*, the paths* $\Gamma_x$ *and* $\Gamma_y$ *coalesce.*

The proof will be long, so we first explain the main ideas. We apply the technique of Licea-Newman [72], whose central tool is a Burton-Keane type argument [23]. We proceed by contradiction, so suppose there are vertices $x, y$ such that $\Gamma_x$ and $\Gamma_y$ do not coalesce. By results of the last section, they cannot even intersect. We show in Sections 4.5.1 and 4.5.2 that there are many triples of non-intersecting paths $\Gamma_{x_1}, \Gamma_{x_2}$ and $\Gamma_{x_3}$ such that $\Gamma_{x_2}$ is "shielded" from all other infinite paths in $\mathbb{G}$. To do this, we must use the information in Theorem 3.7.3 about asymptotic directions.

A contradiction comes in Section 4.5.3 from translation invariance because when $\Gamma_{x_2}$ is shielded, the component of $x_2$ in $\mathbb{G}$ has a unique least element in a certain lexicographic-like ordering of $\mathbb{Z}^2$. This is a different concluding argument than that given in [72], where these shielded paths are used for a Burton-Keane "lack of space" proof.

We now give the proof. For the entirety we will assume either **A1'** or both **A2'** and the upward finite energy property.

### 4.5.1 Constructing "building blocks"

Assume for the sake of contradiction that there are disjoint $\Gamma_x$'s in $\mathbb{G}$. Then for some vertex $z_0$, the event $A_0(z_0) \subseteq \widetilde{\Omega}$ has positive $\mu$-probability, where

$$A_0(z_0) = \{\Gamma_{z_0} \text{ and } \Gamma_0 \text{ share no vertices}\} \ .$$

We begin with a geometric lemma. It provides a (random) line such that with probability one, any path that is asymptotically directed in $J_\varrho$ (from (3.44)) intersects this line finitely often. We will need some notation which is used in the rest of the proof.

Let $\varpi'$ be a vector with

$$\arg \varpi' \in \{j\pi/4, \, j = 0, \ldots, 7\} \text{ and } \|\varpi'\|_\infty = 1 \ , \tag{4.16}$$

where $\| \cdot \|_\infty$ is the $\ell^\infty$ norm. (A precise value of $j$ will be fixed shortly.) Define (for $N \in \mathbb{N}$) $L'_N = \{z \in \mathbb{R}^2 : \varpi' \cdot z = N\}$. For such an $N$ and for $x \in \mathbb{Z}^2$, write $x \prec L'_N$ if $\varpi' \cdot x < N$ and $x \succ L'_N$ if $\varpi' \cdot x > N$. The symbols $\preceq$ and $\succeq$ are interpreted in the obvious way. We use the terms "far side of $L'_N$" and "near side of $L'_N$" for the sets of $x \in \mathbb{R}^2$ with $x \succ L'_N$ and $x \prec L'_N$, respectively. Note that any lattice path $\gamma$ intersecting both sides of $L'_N$ contains a vertex $z \in L'_N$.

**Lemma 4.5.2.** *There is a measurable choice of $\varpi'$ as in (4.16) such that with $\mu$-probability one, the following holds. For each vertex $x$ and each integer $N$,*

$$\Gamma_x \cap \{z \in \mathbb{Z}^2 : z \preceq L'_N\} \text{ is finite .}$$

*In other words, $\Gamma_x$ eventually lies on the far side of $L'_N$ for all $x$ and $N$.*

*Proof.* The limit shape $\mathcal{B}$ is convex and compact, so it has an extreme point $p$. Because it is symmetric with respect to the rotation $R$ of $\mathbb{R}^2$ by angle $\pi/2$, the points $p_i = R^i p$, $i = 1, \ldots, 3$ are all extreme points of $\mathcal{B}$. $J_\varrho$ is an interval of angles corresponding to points of contact between $\mathcal{B}$ and one of its supporting lines, so it is connected (in the topology induced by $dist$) and must lie between (inclusively) $\arg p_i$ and $\arg p_{i+1}$ for some $i = 0, \ldots, 3$ (here we identify $p_4 = p_0$). Therefore $\operatorname{diam} J_\varrho \leq \pi/2$ almost surely and contains at most three elements of the set $\{j\pi/4 : j = 0, \ldots, 7\}$ (and they must be consecutive). Choose five of the remaining elements to be consecutive and label them $j_1\pi/4, \ldots, (j_1+4)\pi/4$. The interval $[j_1\pi/4, (j_1+4)\pi/4]$ defines a half-plane $H$ in $\mathbb{R}^2$ and since the distance between this interval and $J_\varrho$ is positive (measured with $dist$), for all sufficiently small $\varepsilon > 0$, the sector

$$\{x \in \mathbb{R}^2 : x \neq 0 \text{ and } dist(\arg x, \phi) < \varepsilon \text{ for some } \phi \in J_\varrho\}$$

is contained in $H^c$. This implies the statement of the lemma for a (random) $\varpi'$ equal to the normal to $H$. Since $\varpi'$ can be chosen as a measurable function of $\varrho$ (which is clearly Borel measurable on $\widetilde{\Omega}$), we are done. □

For the rest of the proof, fix a deterministic $\varpi'$ as in (4.16) that satisfies Lemma 4.5.2 with positive probability on the event $A_0(z_0)$. (This is possible because there are only eight choices for $\varpi'$.) Let $A'_0(0, z_0)$ be the intersection of $A_0(z_0)$ and the event in the lemma. On $A'_0(0, z_0)$, $\Gamma_0$ and $\Gamma_{z_0}$ eventually cease to intersect $L'_0$. In particular, they each have a last intersection with $L_0'$. Since there are only countably many possible pairs of such last intersections, we see that some pair $(y, y')$ in $L_0'$ occurs with positive probability; that is, $\mu(A(y,y')) > 0$, where $A(y,y')$ is defined by the conditions

I. $\Gamma_y \cap \Gamma_{y'} = \varnothing$;

II. $\Gamma_y$ intersects $L_0'$ only at $y$; $\Gamma_{y'}$ intersects $L_0'$ only at $y'$ and

III. $\Gamma_u \cap L_N'$ is nonempty and bounded for $u = y, y'$ and all integers $N \geq 0$.

(Note that condition III follows directly from the preceding lemma because $\Gamma_u$ contains infinitely many vertices.) By translation invariance, there exists $z \in L_0'$ with $\mu(A(0,z)) > 0$.

Fix

$$\varsigma = \text{ a nonzero vector with the smallest integer coordinates normal to } \varpi' \tag{4.17}$$

(it will be a rotation of either (0,1) or (1,1) by a multiple of $\pi/2$). Defining $\widetilde{T}_\varsigma : \widetilde{\Omega} \to \widetilde{\Omega}$ as the translation by $\varsigma$ (that is, $\widetilde{T}_1^{a_1} \circ \widetilde{T}_2^{a_2}$, where $\varsigma = a_1\mathbf{e}_1 + a_2\mathbf{e}_2$),

$$\mathbf{1}_{A(0,z)}\left((\omega, \Theta, \eta)\right) = \mathbf{1}_{A(\varsigma, z+\varsigma)}\left(\widetilde{T}_\varsigma(\omega, \Theta, \eta)\right) .$$

Since $\mu$ is invariant under the action of $\widetilde{T}_\varsigma$, the ergodic theorem implies

$$\frac{1}{N}\sum_{j=0}^{N-1} \mathbf{1}_{A(j\varsigma, z+j\varsigma)}\left((\omega, \Theta, \eta)\right) = \frac{1}{N}\sum_{j=0}^{N-1} \mathbf{1}_{A(0,z)}\left(\widetilde{T}_\varsigma^j(\omega, \Theta, \eta)\right) \to g(\omega, \Theta, \eta), \tag{4.18}$$

where $g$ is a function in $L^1(\mu)$; the convergence is both $\mu$-almost sure and in $L^1(\mu)$, so $\int g \, \mathrm{d}\mu = \mu(A(0,z)) > 0$. Using this in (4.18) gives infinitely many $j$ with

$$\mu\left(A(0,z) \cap A(j\varsigma, z+j\varsigma)\right) > 0. \tag{4.19}$$

We fix $j > \|z\|_1$ to ensure $\Gamma_{j\varsigma}$ and $\Gamma_{z+j\varsigma}$ are outside the region bounded by $L_0'$, $\Gamma_0$, and $\Gamma_z$.

What is the significance of the event in (4.19)? When it occurs, we are guaranteed that there is a line $L_0'$ and four directed paths remaining on its far side apart from their initial vertices. We claim that at least three of them never intersect. Indeed, ordering the paths using the direction of $\varsigma$, we are guaranteed that the "first two" paths do not intersect each other, nor do the "last two." But if the middle two paths ever intersect, they would merge beyond that point and the three remaining paths could not touch.

For $x_1, x_2 \in L_0'$, let $B(0, x_1, x_2)$ be the event that $\Gamma_0, \Gamma_{x_1}$ and $\Gamma_{x_2}$ (a) never intersect, (b) stay on the far side of $L_0'$ except for their initial vertices and (c) intersect $L_N'$ in a bounded set for each $N \geq 1$. Then the above implies

$$B(0, z, j\varsigma) \cup B(0, z, z + j\varsigma) \supseteq A(0, z) \cap A(j\varsigma, z + j\varsigma) \ .$$

Therefore we may choose $x_1, x_2 \in L_0'$ such that the portion of $L_0'$ from 0 to $x_2$ contains $x_1$ and so that $\mu(B(0, x_1, x_2)) > 0$. The vertices $x_1$ and $x_2$ are fixed for the rest of the proof.

### 4.5.2 Constructing $B'$

Our next step is to refine $B(0, x_1, x_2)$ to a positive probability subevent $B'(x^*; N, R)$ on which no paths $\Gamma_z$ with $z \preceq L_N'$ (outside of some large polygon) merge with $\Gamma_{x_1}$. We will need to pull events back from $\widetilde{\Omega}$ to $\Omega_1$ to do an edge modification and this will present a considerable difficulty. Our strategy is reminiscent of that in [6]. In the first subsection we give several lemmas that we will need. In the next subsection we will define $B'$ and show it has positive probability.

**Lemmas for $B'$**

We wish to construct a barrier of high-weight edges on the near side of some $L'_N$. Set

$$\lambda_0^+ = \sup\left\{\lambda > 0 : \mathbb{P}\left(\omega_e \in [\lambda, \infty)\right) > 0\right\} \ .$$

Because we do not wish to assume $\lambda_0^+ = \infty$, our barrier will occupy some wide polygon (in the case that $\lambda_0^+ = \infty$, many of the complications which we address below can be neglected; we direct the interested reader to [72]). To control the exit of our directed paths from the polygon, we will need a lemma about weak angular concentration of paths:

**Lemma 4.5.3.** *For $x \in L'_0$ such that $\Gamma_x \cap L'_N \neq \varnothing$, define $\zeta_N(x)$ to be the $\varsigma$-coordinate of the first intersection of $\Gamma_x$ with $L'_N$. That is, this first intersection may be written uniquely as $\zeta_N(x)\varsigma + b\varpi'$ for some number $b$. Denote the event*

$$B_G(x_a, x_b, x_c) := B(x_a, x_b, x_c) \cap \Big\{\text{for every } \varepsilon > 0, |\zeta_{N_i}(x_a) - \zeta_{N_i}(x_c)| < \varepsilon N \\ \text{for infinitely many } N\Big\}.$$

*Then $\mu(B_G(0, x_1, x_2) \mid B(0, x_1, x_2)) = 1$.*

*Proof.* Let $B_B(x_a, x_b, x_c, N, \varepsilon)$ denote the event that $B(x_a, x_b, x_c)$ occurs but that $|\zeta_N(x_a) - \zeta_N(x_c)| \geq \varepsilon N$. Fix some $\ell$ greater than the absolute value of the $\varsigma$-coordinate of $x_2$.

Note that, if $\liminf_N \mu(B_B(0, x_1, x_2, N, \varepsilon)) = 0$ for every $\varepsilon > 0$, then the lemma holds. To see this, for every $n > 0$ take a sequence $(N_n)$ such that $\mu(B_B(0, x_1, x_2, N_n, 1/n) \leq n^{-2}$ and apply the Borel-Cantelli theorem.

So assume for the sake of contradiction that there are some $\varepsilon, p_B > 0$ such that $\mu(B_B(0, x_1, x_2, N, \varepsilon)) > p_B > 0$ for all large $N$; fix some such $N > 6\ell/\varepsilon p_B$ for the remainder of the proof, and define

$$\widehat{B}(n) = B_B(n\ell\varsigma, n\ell\varsigma + x_1, n\ell\varsigma + x_2, N, \varepsilon).$$

We will look at shifted versions of $\widehat{B}$ and define a function $f$ to count the number of $\zeta_N(y)$ arising from some shifted version of $\widehat{B}$.

Define $f_N : \mathbb{Z} \to \{0,1\}$, where $f_N(m) = 1$ if there is some $n \in \mathbb{Z}$ such that $\widehat{B}(n)$ occurs and either $\zeta(n\ell\varsigma) = m$ or $\zeta(n\ell\varsigma + x_2) = m$. That is,

$$f_N(m) = \sup_{n\in\mathbb{Z}} \left\{ \mathbf{1}_{\widehat{B}(n)} \max_{y=n\ell\varsigma, x_2+n\ell\varsigma} \mathbf{1}_{\zeta_N(y)=m} \right\}.$$

This form makes it clear that $f_N(m)$ is a measurable random variable for each $N$ and $m$. Finally, set

$$\bar{f}_N(L) = \frac{1}{\ell L + 1} \sum_{m=0}^{\ell L} f_N(m).$$

We claim that $\bar{f}_N(L)$ must satisfy contradictory inequalities for $L$ large. Denote by $\mathcal{C}(L)$ the box of vertices points lying between $L'_0$ and $L'_N$ with $\varsigma$-coordinate between 0 and $L\ell$.

We first note that, on the event $\hat{B}(n) \cap \hat{B}(n')$, if $\zeta(n\ell\varsigma) = m$ and $\zeta(n\ell\varsigma + x_2) = m'$, then $\zeta(n'\varsigma), \zeta(n'\varsigma + x_2) \notin (m, m')$ because $n\varsigma$ and $n\varsigma + x_2$ start outside of the region bounded by $L'_0, \Gamma_{n\varsigma}, \Gamma_{n\varsigma+x_2}$. Moreover, $|m - m'| \geq \varepsilon N$.Thus, by the definition of $f$, if $f(m) = 1$ then $f(m_1) = 0$ either for all $m_1 \in (m, m + N\varepsilon)$ or $m_1 \in (m - N\varepsilon, m)$. In particular, we have the almost sure bound

$$\begin{aligned} &\sum_{m=0}^{\ell L} f(m) \leq 3 + \frac{\ell L \|\varsigma\|_1}{N\varepsilon} \\ &\Longrightarrow \bar{f}_N(L) \leq \frac{3}{N\varepsilon} \quad \mu - \text{a.s.} \end{aligned} \tag{4.20}$$

for all $L$ larger than some $L_{\min}$.

On the other side, note that if $\cap_{i=1}^{2\lambda}\hat{B}(n_i)$ occurs, then the non-merging of each pair $\Gamma_{n\varsigma}, \Gamma_{n\varsigma+x_2}$ and planarity ensure that we can find a subset $\Lambda$ of the set of paths

$$\{\Gamma_{n_i\varsigma}, \Gamma_{n_i\varsigma+x_2}\}_{i=1}^{2\lambda}$$

such that $|\Lambda| = \lambda$ and such that the paths in $\Lambda$ are pairwise disjoint. Now, note that if all the paths of $\Lambda$ have their starting points in $\mathcal{C}(L)$, then if $\Gamma \in \Lambda$ either its first intersection with $L'_N$ occurs within $\mathcal{C}(L)$ or outside.

If the latter occurs, then $\Gamma$ must intersect a point of $\mathcal{C}(L)$ with $\varsigma$-coordinate 0 or $L$. There are at most $c_N$ such points, where $c_N$ is an $L$-independent constant. In particular, we have

$$\begin{aligned} \sum_{m=0}^{\ell L} f_N(m) &\geq \left[\sum_{n=0}^{L} \mathbf{1}_{\hat{B}(n)}\right] - c_N \\ \Longrightarrow \mathbb{E}_\mu \bar{f}_N(L) &\geq \frac{p_B}{2\ell} \end{aligned} \tag{4.21}$$

for all $L > L'_{\min} > L_{\min}$. Combining (4.20) and (4.21) for $L$ large enough, we have

$$\frac{p_B}{2\ell} \leq \frac{3}{N\varepsilon} < \frac{p_B}{2\ell},$$

a contradiction. □

The next lemma is a modification of the usual first-passage shape theorem.

**Lemma 4.5.4.** *There exists a deterministic $c^+ < \lambda_0^+$ such that, $\mathbb{P}$-a.s.,*

$$\lim_{M\to\infty} \sup_{\|x\|_1 \geq M} \tau(0,x)/\|x\|_1 < c^+ .$$

*Proof.* Because either **A1'** or **A2'** hold, $\mathbb{E}(\tau_e) < \lambda_0^+$. For any $z \in \mathbb{Z}^2$, choose a deterministic path $\gamma_z$ with number of edges equal to $\|z\|_1$. For $x \in \mathbb{Q}^2$ and $n \geq 1$

with $nx \in \mathbb{Z}^2$,

$$\mathbb{E}\tau(0, nx) \leq \mathbb{E}\tau(\gamma_{nx}) = n\|x\|_1 \mathbb{E}\tau_e \ , \text{ so } g(x) \leq \|x\|_1 \mathbb{E}\tau_e \ .$$

This extends to all $x \in \mathbb{R}^2$ by continuity, so the shape theorem gives the result. □

We need a lemma to pull events back from $\widetilde{\Omega}$ to $\Omega_1$. Fix an increasing sequence $(n_k)$ such that $\mu_{n_k}^* \to \mu$ weakly.

**Lemma 4.5.5.** *Let $E \subseteq \widetilde{\Omega}$ be open with $\mu(E) > \beta$. There exists $C_\beta > 0$ and $K_0$ such that for $k \geq K_0$, the Lebesgue measure of the set $\{\alpha \in [0, n_k] : \mu_\alpha(E) > \beta/2\}$ is at least $C_\beta\, n_k$.*

*Proof.* Call the Lebesgue measure of the above set $\lambda$. Since $E$ is open, (3.23) allows us to pick $K_0$ such that if $k \geq K_0$ then $\mu_{n_k}^*(E) > \beta$. For such $k$, we can write

$$\frac{1}{n_k}\left(\lambda + (n_k - \lambda)\beta/2\right) \geq \mu_{n_k}^*(E) > \beta, \text{ giving } \lambda > \frac{n_k \beta}{2(1-\beta/2)} \ .$$

Setting $C_\beta := \beta(2-\beta)^{-1}$ completes the proof. □

The last lemma is based on [6, Lemma 3.4] and will be used in the edge-modification argument. To push the upward finite energy property forward from $\Omega_1$ to $\widetilde{\Omega}$ we need concrete lower bounds for probabilities of modified events. We write a typical element of $\Omega_1$ as $\omega = (\omega_e, \check{\omega})$, where $\check{\omega} = (\omega_f)_{f \neq e}$. We say an event $A \subseteq \Omega_1$ is *$e$-increasing* if, for all $(\omega_e, \check{\omega}) = \omega \in A$ and $r > 0$, $(\omega_e + r, \check{\omega}) \in A$.

**Lemma 4.5.6.** *Let $\lambda > 0$ be such that $\mathbb{P}\left(\omega_e \geq \lambda\right) > 0$. For each $\vartheta > 0$ there exists $C = C(\vartheta, \lambda) > 0$ such that for all edges $e$ and all $e$-increasing events $A$ with $\mathbb{P}(A) \geq \vartheta$,*

$$\mathbb{P}\left(A,\ \omega_e \geq \lambda\right) \geq C\ \mathbb{P}\left(A\right)\ .$$

*Proof.* If $\mathbb{P}(A,\ \omega_e < \lambda) \le (1/2)\mathbb{P}(A)$ then

$$\mathbb{P}(A,\ \omega_e \ge \lambda) \ge (1/2)\mathbb{P}(A)\ . \tag{4.22}$$

Otherwise, we assume that

$$\mathbb{P}(A,\ \omega_e < \lambda) \ge (1/2)\mathbb{P}(A)\ . \tag{4.23}$$

We then need to define an extra random variable. Let $\omega_e'$ be a variable such that, given $\check{\omega}$ from $\omega \in \Omega_1$, it is an independent copy of the variable $\omega_e$. In other words, letting $\mathbb{Q}$ be the joint distribution of $(\omega, \omega_e')$ on the space $\Omega_1 \times \mathbb{R}$, for $\mathbb{Q}$-almost every $\check{\omega}$,

- $\omega_e'$ and $\omega_e$ are conditionally independent given $\check{\omega}$ and
- the distributions $\mathbb{Q}(\omega_e \in \cdot \mid \check{\omega})$ and $\mathbb{Q}(\omega_e' \in \cdot \mid \check{\omega})$ are equal.

(This can be defined, for instance, by setting $\mathbb{Q}(A \times B) = \int_A \mathbb{P}(\omega_e \in B \mid \check{\omega})\, \mathrm{d}\mathbb{P}(\omega)$ for Borel sets $A \subseteq \Omega_1$ and $B \subseteq \mathbb{R}$.)

We now write $\mathbb{P}(A, \omega_e \ge \lambda)$ as

$$\begin{aligned}
\mathbb{Q}[(\omega_e, \check{\omega}) \in A,\ \omega_e \in [\lambda, \infty)] &\ge \mathbb{Q}\left[(\omega_e, \check{\omega}) \in A,\ \omega_e \in [\lambda, \infty),\ \omega_e' \in [0, \lambda)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{(\omega_e, \check{\omega}) \in A}\, \mathbf{1}_{\omega_e \in [\lambda, \infty)}\, \mathbf{1}_{\omega_e' \in [0, \lambda)}\right] \\
&\ge \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{(\omega_e', \check{\omega}) \in A}\, \mathbf{1}_{\omega_e \in [\lambda, \infty)}\, \mathbf{1}_{\omega_e' \in [0, \lambda)}\right] && (4.24) \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{(\omega_e', \check{\omega}) \in A}\, \mathbf{1}_{\omega_e' \in [0, \lambda)}\, \mathbb{E}_{\mathbb{Q}}\left(\mathbf{1}_{\omega_e \in [\lambda, \infty)}\ \mid \check{\omega}, \omega_e'\right)\right]. && (4.25)
\end{aligned}$$

In (4.24), we have used that $A$ is $e$-increasing. Using conditional independence in (4.25),

$$\mathbb{P}(A,\ \omega_e \ge \lambda) \ge \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{(\omega_e', \check{\omega}) \in A}\, \mathbf{1}_{\omega_e' \in [0, \lambda)}\, \mathbb{E}_{\mathbb{Q}}\left(\mathbf{1}_{\omega_e \in [\lambda, \infty)}\ \mid \check{\omega}\right)\right]\ . \tag{4.26}$$

By the upward finite energy property,

$$\mathbb{E}_{\mathbb{Q}}(\mathbf{1}_{\omega_e \in [\lambda,\infty)} \mid \check{\omega}) = \mathbb{E}(1_{\omega_e \in [\lambda,\infty)} \mid \check{\omega}) > 0 \ \ \mathbb{Q}\text{-almost surely} \ ,$$

so choose $c > 0$ such that

$$\mathbb{Q}\left[\mathbb{E}_{\mathbb{Q}}(\mathbf{1}_{\omega_e \in [\lambda,\infty)} \mid \check{\omega}) \geq c\right] \geq 1 - (\vartheta/4) \ .$$

Note that this choice of $c$ depends only on $\lambda$ and $\vartheta$. By (4.23) and the assumption $\mathbb{P}(A) \geq \vartheta$, the right side is at least $1 - (1/2)\mathbb{P}(A,\ \omega_e < \lambda)$, implying

$$\mathbb{Q}\left[(\omega_e', \check{\omega}) \in A,\ \omega_e' \in [0,\lambda),\ \mathbb{E}_{\mathbb{Q}}(\mathbf{1}_{\omega_e \in [\lambda,\infty)} \mid \check{\omega}) \geq c\right] \geq (1/2)\mathbb{P}(A,\ \omega_e < \lambda) \ .$$

Combining with (4.26), we find $\mathbb{P}(A,\ \omega_e \geq \lambda) \geq (c/2)\mathbb{P}(A,\ \omega_e < \lambda)$. We finish the proof by writing

$$\mathbb{P}(A) = \mathbb{P}(A, \omega_e < \lambda) + \mathbb{P}(A, \omega_e \geq \lambda) \leq \left[\frac{2}{c} + 1\right] \mathbb{P}(A, \omega_e \geq \lambda) \ .$$

Observing this inequality and (4.22), we set $C = \min\{1/2, c/(2+c))\}$.

□

**Defining $B'$**

We begin with the definition of the "barrier event" $B'$. For an integer $R > N$, let

$$S(R,N) = \{y \in \mathbb{Z}^2 : 0 \leq y \cdot \varpi' \leq N,\ |y \cdot \varsigma| \leq R\} \ .$$

For any vertex $x^* \in S(R,N) \cap L_N'$, define $B'(x^*; R, N)$ by the condition

$$\text{for all } z \in \mathbb{Z}^2 \setminus S(R,N) \text{ with } z \preceq L_N',\ \Gamma_z \cap \Gamma_{x^*} = \varnothing \ . \tag{4.27}$$

**Proposition 4.5.7.** *There exist values of $R, N$ and $x^*$ such that $\mu(B'(x^*; R, N)) > 0$.*

Our strategy is to pull back cylinder approximations of $B(0, x_1, x_2)$ to $\Omega_1$ to find events that depend on $\mathbb{G}$ in the vicinity of $0, x_1$ and $x_2$. We will find a subevent which is monotone increasing in the weights of edges lying in $S(R, N)$ between the pulled-back versions of $\Gamma_0$ and $\Gamma_{x_2}$. When we look at the subevent on which all of these weights are large ("edge modification"), the pullback of $\Gamma_{x_1}$ will be unchanged (past $S(R, N)$), and no pullback of any $\Gamma_z$ can intersect it if $z \preceq L'_N$ and $z \notin S(R, N)$. We will then choose $x^*$ to be a certain point on $\Gamma_{x_1} \cap L'_N$. The constants $N$ and $R$ will be chosen to guarantee that the pullback of $\Gamma_{x_1}$ is so isolated. Pushing forward the subevent to $\widetilde{\Omega}$ will complete the proof.

*Proof.* We will first fix some parameters to prepare for the main argument. Recall the definition of $c^+$ from Lemma 4.5.4 and let

$$\lambda^+ := \min\{\lambda_0^+, 2c^+\} ,$$

and put $\delta^+ := \lambda^+ - c^+ > 0$ (giving $\lambda^+ = 2c^+$ when $\lambda_0^+ = \infty$). Choose once and for all some

$$\varepsilon < \frac{\delta^+}{16\lambda^+}, \tag{4.28}$$

such that also

$$\limsup_{\|x\|_1 \to \infty} \sup_{y : \|y - x\|_1 \le \varepsilon \|x\|_1} \frac{\tau(0, y)}{\|x\|_1} < \lambda^+ - \frac{7\delta^+}{8} \quad \mu\text{-a.s.} \tag{4.29}$$

This follows from Lemma 4.5.4 because if $\|y\|_1$ is large, $\|y - x\|_1 \le \varepsilon \|x\|_1$ gives $\tau(0, y)/\|x\|_1 \le (\tau(0, y)/\|y\|_1)(1 + \varepsilon) < c^+(1 + \varepsilon)$. Fix $\beta > 0$ with $\mu(B(0, x_1, x_2)) > \beta$.

The majority of the proof will consist of defining a few events in sequence, the second of which we will pull back to the space $\Omega_1$ to do the edge modification. We will need to choose further parameters to ensure that each of these events has positive probability. For an arbitrary outcome in $\widetilde{\Omega}$ and $N \geq 0$, denote by $r_0(N)$ and $r_2(N)$ the segments of $\Gamma_0$ and $\Gamma_{x_2}$ up to their first intersections with $L'_N$ (if they exist) and let $w_N$ denote the midpoint of the segment of $L'_N$ lying between these first intersections. The first event $B^\circ(R, N, \varepsilon)$ is defined by the conditions (for $R, N \geq 1$)

1. $\Gamma_0, \Gamma_{x_1}$ and $\Gamma_{x_2}$ never intersect,

2. they stay on the far side of $L'_0$ except for their initial vertices,

3. $\Gamma_0$ and $\Gamma_{x_2}$ intersect $L'_N$ and their first intersection points are within $\ell^1$ distance $\varepsilon N$ of each other,

4. for $i = 0, 2$, $\tau(r_i(N)) < (\lambda^+ - 7\delta^+/8)\|w_N\|_1$ and

5. $\Gamma_0$ and $\Gamma_{x_2}$ do not touch any $x \preceq L'_N$ with $x \notin S(R, N)$.

See Figure 4.3 for a depiction of the event $B^\circ(R, N, \varepsilon)$.

We claim that there exists $N_0$ and $R_0$ such that

$$\mu(B^\circ(R_0, N_0, \varepsilon)) > 0 \ . \tag{4.30}$$

We also need $N_0$ to satisfy a technical requirement. It will be used at the end of the proof:

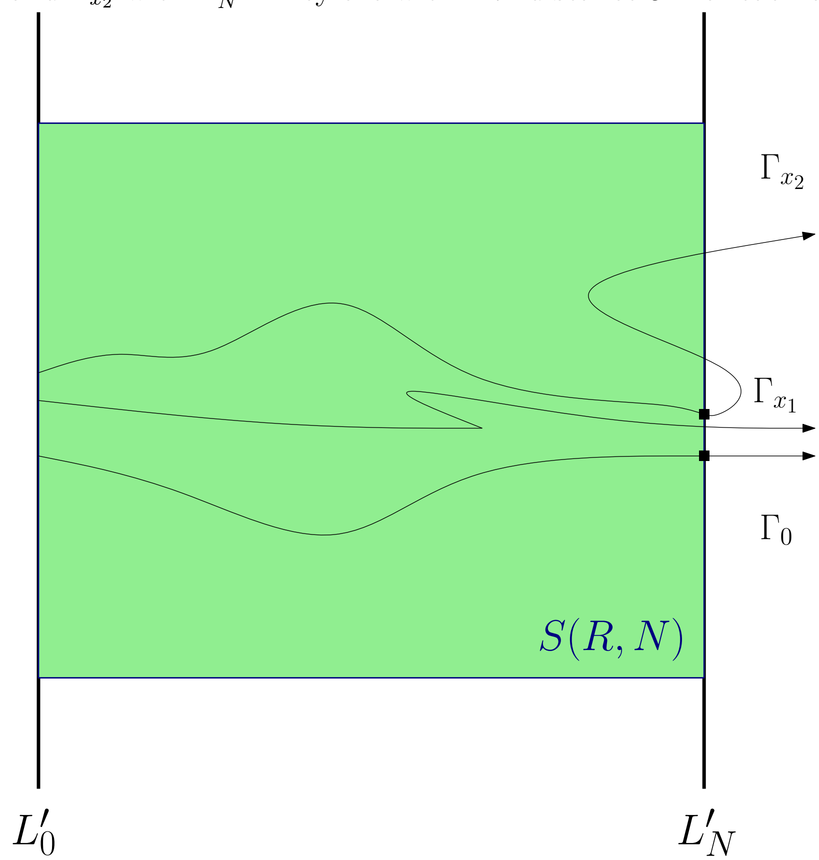
$$\|x_2\|_1 \leq \varepsilon N_0 \ . \tag{4.31}$$

To pick $N_0$, first choose $N_1 > 0$ so large that if $N \geq N_1$ then

$$\mathbb{P}\left(\forall z, z' \text{ with } \|z\|_1 \geq N, \text{ and } \frac{\|z - z'\|_1}{\|z\|_1} \leq \varepsilon, \ \frac{\tau(0, z')}{\|z\|_1} < (\lambda^+ - \frac{7\delta^+}{8})\right) > 1 - \beta/4 \ , \tag{4.32}$$

and $\|x_2\|_1 \leq \varepsilon N$. This is possible by (4.29). Write $E_0(N)$ for the event in (4.32) and $E_{x_2}(N)$ for $E_0(N)$ translated so that 0 is mapped to $x_2$. Then $\mathbb{P}(B(0, x_1, x_2) \cap$

Figure 4.3: The event $B^\circ(R, N, \varepsilon)$. The solid dots represent the first intersection points of $\Gamma_0$ and $\Gamma_{x_2}$ with $L'_N$. They are within $\ell^1$ distance $\varepsilon N$ of each other.




$E_0(N) \cap E_{x_2}(N)) > \beta/2$. By Lemma 4.5.3, we can then choose $N_0 \geq N_1$ such that

$$\mu(B(0, x_1, x_2) \cap E_0(N_0) \cap E_{x_2}(N_0) \cap C(0, x_2; N_0)) > 0 \ , \tag{4.33}$$

where $C(0, x_2; N_0)$ is the event that $\Gamma_0$ and $\Gamma_{x_2}$ intersect $L'_{N_0}$ and their first intersection points are within $\ell^1$ distance $\varepsilon N_0$ of each other. On the event in (4.33), the endpoints of the $r_i(N_0)$'s are within distance $\varepsilon N_0$ of $w_{N_0}$ and since they are on $L'_{N_0}$, their $\ell^1$ distance from 0 or $x_2$ is at least $N_0$. Therefore $\tau(r_i(N_0)) < (\lambda^+ - 7\delta^+/8)\|w_{N_0}\|_1$ for $i = 0, 2$. This shows that the intersection of four of the five events in the definition of $B^\circ(R, N_0, \varepsilon)$ occurs with positive probability. For the fifth, recall that on $B(0, x_1, x_2)$, the paths $\Gamma_0$, $\Gamma_{x_1}$ and $\Gamma_{x_2}$ contain only finitely many vertices $z \preceq L'_{N_0}$. Thus we can choose $R_0$ large enough (depending on $N_0$) to satisfy condition 5 and complete the proof of (4.30).

Fix these $R = R_0$ and $N = N_0$ from now on. The next event we define is a cylinder approximation of the first event. It will be needed to pull back to $\Omega_1$. For $M > 0$ and $x \in \mathbb{Z}^2$, let $\Gamma_x^M$ be the finite path formed by starting at $x$ and then passing along out-edges of $\mathbb{G}$ until we first reach a vertex of $\mathbb{R}^2 \setminus (-M,M)^2$. (Note that by this definition, $\Gamma_x^M = \{x\}$ whenever $x \notin (-M,M)^2$.) We define $B_M^\circ(R,N,\varepsilon)$ with the same conditions as $B^\circ(R,N,\varepsilon)$, except replacing the paths $\Gamma_{(\cdot)}$ by the segments $\Gamma_{(\cdot)}^M$. In addition, however, we impose the restriction that, writing

$$\partial M = [-M,M]^2 \setminus (-M,M)^2 \ ,$$

we have

$$\Gamma_y^M \cap \partial M \subseteq \{z \in \mathbb{R}^2 : z \succ L_N'\}, \ y = 0, x_2 \ . \tag{4.34}$$

Of course, if $\Gamma_0^M$ (etc.) does not intersect $L_N'$, then $B_M^\circ$ does not occur. Then $B_M^\circ(R,N,\varepsilon)$ is open for all $M$ and we claim that

$$B^\circ(R,N,\varepsilon) = \cup_{M_0=1}^\infty \cap_{M=M_0}^\infty B_M^\circ(R,N,\varepsilon) \ . \tag{4.35}$$

Assuming we show this, then there exists some $M_0$ such that $\mu(\cap_{M=M_0}^\infty B_M^\circ(R,N,\varepsilon)) > 0$ and so there is some $\beta'$ with

$$\mu(B_M^\circ(R,N,\varepsilon)) > \beta' \text{ for all } M \geq M_0 \ . \tag{4.36}$$

To prove (4.35), note that the right side is the event that $B_M^\circ(R,N,\varepsilon)$ occurs for all $M$ bigger than some random $M_0$. Suppose that an outcome is in the left side. Then the paths $\Gamma_0$, $\Gamma_{x_1}$ and $\Gamma_{x_2}$ are disjoint and remain on the far side of $L_0'$ (except for their first vertices), so the same is true for each $\Gamma_{(\cdot)}^M$ for all $M \geq 1$. Also $\Gamma_0^M$ and $\Gamma_{x_2}^M$ do not touch any $x \preceq L_N'$ with $x \notin S(R,N)$ for all $M \geq 1$. Because $\Gamma_0$ and $\Gamma_{x_2}$ intersect $L_N'$, so do $\Gamma_0^M$ and $\Gamma_{x_2}^M$ for all $M$ bigger than some random $M_1$. Their first intersection points are the same as those of $\Gamma_0$ and $\Gamma_{x_2}$, so for $M \geq M_1$, their first intersection points with $L_N'$ are within $\ell^1$ distance $\varepsilon N$ of each other. Further, the passage times of the segments up to $L_N'$ are strictly bounded above by $(\lambda^+ - 7\delta^+/8)\|w_N\|_1$. Last, because $\Gamma_0$ and $\Gamma_{x_2}$ do not touch any $x \preceq L_N'$ with $x \notin S(R,N)$, they share only finitely many vertices with $\{z \in \mathbb{Z}^2 : z \preceq L_N'\}$ and so must eventually lie on the far side of $L_N'$. This allows us to further increase $M_1$ to an $M_0$ such that if $M \geq M_0$ then in addition (4.34) holds.

Suppose conversely that the right side of (4.35) occurs. Then for all $M$ bigger than some random $M_0$, the six events comprising $B_M^\circ(R,N,\varepsilon)$ occur. In particular, the paths $\Gamma_0$, $\Gamma_{x_1}$ and $\Gamma_{x_2}$ are disjoint and stay on the far side of $L_0'$ except for their first vertices (parts 1 and 2 of $B^\circ(R,N,\varepsilon)$). Furthermore $\Gamma_0$ and $\Gamma_{x_2}$ cannot touch any $x \preceq L_N'$ with $x \notin S(R,N)$ (part 5). For $M \geq M_0$, the paths $\Gamma_0^M$ and $\Gamma_{x_2}^M$ intersect $L_N'$, with their first intersection points within distance $\varepsilon N$ of each other (with passage time strictly bounded above by $(\lambda^+ - 7\delta^+/8)\|w\|_1$). These are the same first intersection points as $\Gamma_0$ and $\Gamma_{x_2}$, so parts 3 and 4 of $B^\circ(R,N,\varepsilon)$ occur.

We now pull the cylinder approximation $B_M^\circ(R,N,\varepsilon)$ back to $\Omega_1$ using Lemma 4.5.5. Because this is an open event and satisfies (4.36) for $M \geq M_0$, we can find an $M$-dependent number $K_0$ such that if $k \geq K_0$, then there is a set $\Lambda_{M,k}$ of values of $\alpha \in [0, n_k]$ which has Lebesgue measure at least $C_{\beta'} n_k$, on which $\mu_\alpha(B_M^\circ(R,N,\varepsilon)) > \beta'/2$. Pull back to $\Omega_1$, setting $B_M^\alpha := \Phi_\alpha^{-1}(B_M^\circ(R,N,\varepsilon))$, where $\Phi_\alpha$ was defined in (3.20). (Here we have suppressed mention of $R, N, \varepsilon$ in the notation, as they are fixed for the remainder of the proof.) Then

$$\mathbb{P}(B_M^\alpha) > \beta'/2 \text{ for all } \alpha \in \Lambda_{M,k} \text{ if } M \geq M_0 \text{ and } k \geq K_0(M) \ . \tag{4.37}$$

We henceforth restrict to values of $M$, $\alpha$ and $k$ such that (4.37) holds. In the end of the proof we will take $k \to \infty$ and then $M \to \infty$. In particular then we will be thinking of

$$\alpha \gg M \gg N \ ,$$

the latter of which is fixed. Some of the remaining definitions will only make sense for such $\alpha$, $M$ and $N$ but this does not affect the argument.

Next we define the third of our four events, now working on $\Omega_1$. Let $s_y^\alpha$ be the geodesic from $y \in \mathbb{Z}^2$ to $L_\alpha$ (recall this was defined for $\varpi$ and not $\varpi'$), and $s_y^\alpha(M)$ the path $s_y^\alpha$ up to its first intersection with $\mathbb{R}^2 \setminus (-M, M)^2$. If $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ intersect $L'_N$ then write $r_i^\alpha(M)$, $i = 0, 2$ for the portions up to the first intersection point. As before, let $w_N^\alpha$ be the midpoint of the segment of $L'_N$ between these two intersection points. Let $\mathcal{R}_1^\alpha(M)$ be the closed connected subset (in $\mathbb{R}^2$) of $\{x \in \mathbb{R}^2 : x \succeq L'_0\}$ with boundary curves $s_0^\alpha(M)$, $s_{x_2}^\alpha(M)$, $L'_0$ and $\partial M$. Similarly let $\mathcal{R}_2^\alpha(M)$ be the closed connected subset of $\mathcal{R}_1^\alpha(M)$ with the following boundary curves: the portions of $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ after their last intersections with $L'_N$, the segment of $L'_N$ between these intersections and last, $\partial M$. Note that when (4.34) holds, $\mathcal{R}_2^\alpha(M)$ is contained in $\{z \in \mathbb{R}^2 : z \succeq L'_N\}$. See Fig. 4.4 for an illustration of these definitions.

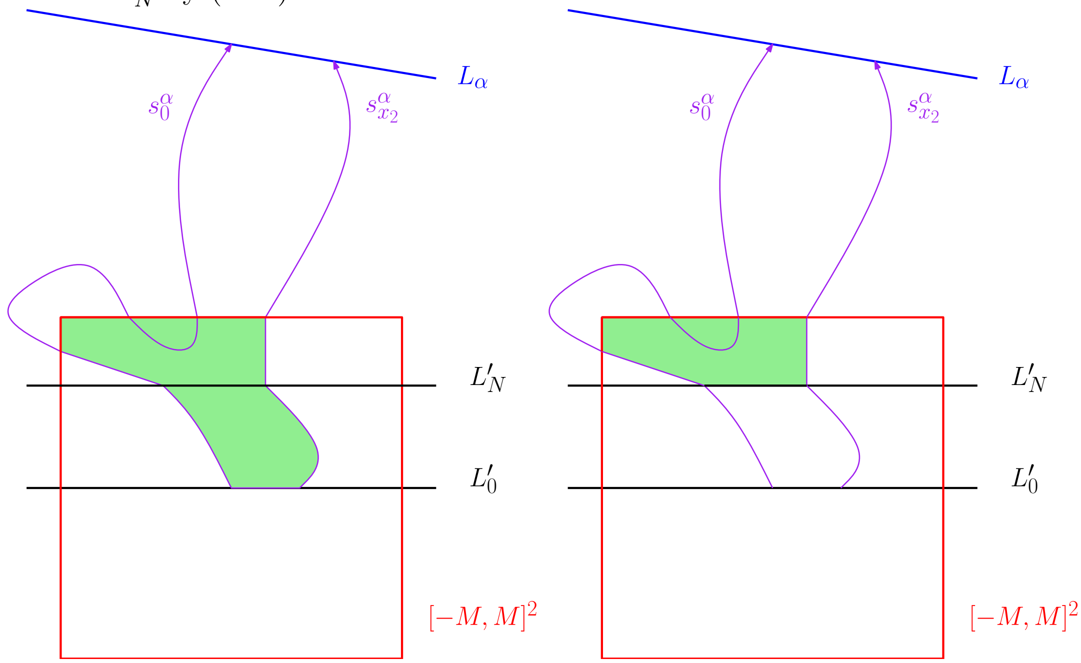
The event $\hat{B}_M^\alpha \subseteq \Omega_1$ is then defined by the following conditions:

- $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ intersect $L'_0$ only once, are disjoint, and do not touch any $y \preceq L'_N$ with $y \notin S(R, N)$.

- $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ intersect $L'_N$ and their first intersection points are within $\ell^1$ distance $\varepsilon N$ of each other; the paths $r_i^\alpha(M)$ satisfy $\tau(r_i^\alpha(M)) < (\lambda^+ - 7\delta^+/8)\|w_N^\alpha\|_1$, for $i = 0, 2$.

- $s_y^\alpha(M) \cap \partial M \subseteq \{z \in \mathbb{R}^2 : z \succ L'_N\}$ for $y = 0, x_2$,

Figure 4.4: The regions $\mathcal{R}_1^\alpha(M)$ and $\mathcal{R}_2^\alpha(M)$. The left figure shows $\mathcal{R}_1^\alpha(M)$ in green. It has boundary curves $L_0'$, $\partial M$, $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$. The right figure shows $\mathcal{R}_2^\alpha(M) \subseteq \mathcal{R}_1^\alpha(M)$ in green. It has boundary curves $L_N'$, $\partial M$, and the pieces of $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ from their last intersections with $L_N'$. Note that $\mathcal{R}_2^\alpha(M)$ is contained in the far side of $L_N'$ by (4.34).




- there is a vertex $X^* \in L_N' \cap S(R, N)$ such that $s_{X^*}^\alpha(M)$ is disjoint from $s_0^\alpha(M)$ and $s_{x_2}^\alpha(M)$ but is contained in $\mathcal{R}_2^\alpha(M)$, and
- the portions of $s_0^\alpha$, $s_{X^*}^\alpha$ and $s_{x_2}^\alpha$ beyond $[-M, M]^2$ do not contain a vertex of $S(R, N)$;

We claim there is an $M_0' \geq M_0$ such that

$$\mathbb{P}(\hat{B}_M^\alpha) > \beta'/4 \text{ for all } M \geq M_0' \,. \tag{4.38}$$

Verifying this requires us to define an auxiliary event. Let $H_M \subseteq \Omega_1$ denote the event that no geodesic from any point in $S(R, N)$ returns to $S(R, N)$ after its first intersection with $\partial M$. Then $\mathbb{P}(H_M) \to 1$ as $M \to \infty$. So for any $M$ larger than some

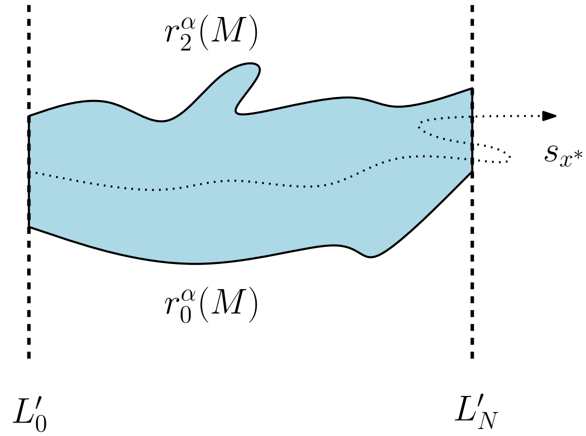
$M_0' \geq M_0$, $\mathbb{P}(H_M) > 1 - \beta'/4$, giving

$$\mathbb{P}(B_M^\alpha \cap H_M) > \beta'/4 \text{ for all } M \geq M_0' \ .$$

To finish the proof of (4.38) we show that $B_M^\alpha \cap H_M \subseteq \hat{B}_M^\alpha$. Note that the first three conditions of $\hat{B}_M^\alpha$ are immediately implied by $B_M^\alpha$; they are the analogues on $\Omega_1$ of the conditions that make up $B_M^\circ(N, R, \varepsilon)$ (each $\Gamma_{(\cdot)}^M$ is replaced by $s_{(\cdot)}^\alpha(M)$). For the fourth condition, note that when $B_M^\alpha$ occurs, $s_0^\alpha(M)$, $s_{x_1}^\alpha(M)$ and $s_{x_2}^\alpha(M)$ stay on the far side of $L_0'$ (aside from their initial vertices) and stop when they touch $\partial M$. Therefore by planarity, $s_{x_1}^\alpha(M)$ is contained in $\mathcal{R}_1^\alpha(M)$. In particular, if we choose $X^*$ to be the last intersection point of $s_{x_1}^\alpha(M)$ with $L_N'$, then $s_{X^*}^\alpha(M)$ is trapped in $\mathcal{R}_2^\alpha(M)$. We can see this as follows. The last vertex of $s_{X^*}^\alpha(M)$ is clearly in this region because it must be in $\mathcal{R}_1^\alpha(M) \cap \partial M$ and this equals $\mathcal{R}_2^\alpha(M) \cap \partial M$. Proceeding backward along $s_{X^*}^\alpha(M)$ from this final vertex, the path can only leave $\mathcal{R}_2^\alpha(M)$ if it (a) leaves $[-M, M]^2$ (b) crosses $s_0^\alpha(M)$ or $s_{x_2}^\alpha(M)$ or (c) crosses $L_N'$. Because none of these can happen, the fourth condition holds. As for the fifth, it is implied by $H_M$, so we have proved (4.38).

Our fourth and final event will fix some random objects to be deterministic so that we can apply the edge modification lemma. On the event $\hat{B}_M^\alpha$, let $U$ denote the (random) closed connected subset of $[-M, M]^2$ with boundary curves $L_0'$, $L_N'$, $r_0^\alpha(M)$ and $r_2^\alpha(M)$. Note that $U \subseteq S(R, N)$. Furthermore we note that on $\hat{B}_M^\alpha$, $U \cap \mathcal{R}_2^\alpha(M)$ is contained in $L_N'$. This is because $\mathcal{R}_2^\alpha(M) \subseteq \{z : z \succeq L_N'\}$, whereas $U \subseteq \{z : z \preceq L_N'\}$. Last, define $U_{\mathcal{E}}$ to be the random set of edges with both endpoints in $U$ and which are not edges in $s_0^\alpha(M), s_{x_2}^\alpha(M), L_0'$ or $L_N'$. See Figure 4.5 for an illustration of these definitions.

Figure 4.5: Illustration of definitions on $\hat{B}_M^\alpha$. The region $U$ is in blue and is contained in $S(R, N)$ (not pictured). It is bounded by curves $L'_0$, $L'_N$, $r_0^\alpha(M)$ and $r_2^\alpha(M)$. The path $s_{x^*}$ begins at the final intersection point of the dotted path with $L'_N$.




On $\hat{B}_M^\alpha$, there are at most $2^{64NR}$ possibilities for $U$ and $U_\mathcal{E}$ and at most $2R$ choices for $X^*$. So there exist some deterministic $U'$, $U'_\mathcal{E}$, and $x^*$ such that, if we define

$$\tilde{B}_M^\alpha := \hat{B}_M^\alpha \cap \{U = U',\, U_\mathcal{E} = U'_\mathcal{E}\} \cap \{X^* = x^*\}\ ,$$

then

$$\mathbb{P}(\tilde{B}_M^\alpha) > 2^{-2-64NR}\beta'/2R \text{ for } M \geq M'_0 \text{ and } \alpha \in \Lambda_{M,k}\ . \tag{4.39}$$

The meaning of the event $\{X^* = x^*\}$ is that the deterministic point $x^*$ satisfies the conditions in the fourth and fifth items of the description of $\hat{B}_M^\alpha$.

In the rest of the proof we perform the edge modification and push forward to $\widetilde{\Omega}$. To apply Lemma 4.5.6 we need to verify that $\tilde{B}_M^\alpha$ is $e$-increasing for all $e \in U'_\mathcal{E}$. For this purpose, suppose that $\omega \in \tilde{B}_M^\alpha$ and that $\omega'$ is another configuration such that $\omega'_e \geq \omega_e$ for some fixed $e \in U'_\mathcal{E}$ but $\omega'_f = \omega_f$ for all other $f \neq e$. By construction, $e$ is not an edge of $s_0^\alpha(M)$, $s_{x^*}^\alpha(M)$ or $s_{x_2}^\alpha(M)$ ($e \notin s_{x^*}^\alpha(M)$ since $e$ is contained in $U_\mathcal{E}$, which does not meet $L'_N$, so is not in $\mathcal{R}_2^\alpha(M) \supseteq s_{x^*}^\alpha(M)$). Furthermore because $s_0^\alpha$, $s_{x^*}^\alpha$ and $s_{x_2}^\alpha$ do not re-enter $S(R, N)$ after leaving $[-M, M]^2$ and all edges of $U'_\mathcal{E}$ have both endpoints in $S(R,N)$, $e$ cannot be on these paths either. This means that

$$s_y^\alpha(\omega) = s_y^\alpha(\omega') \text{ for } y = 0, x^*, x_2 \text{ and } U(\omega) = U(\omega'),\ U_{\mathcal{E}}(\omega) = U_{\mathcal{E}}(\omega')\ .$$

So the fifth condition of $\hat{B}_M^\alpha$ occurs in $\omega'$. The paths $s_y^\alpha(M)$ are then equal in $\omega$ and $\omega'$, so conditions 1, the first part of 2, and 3 and 4 hold in $\omega'$. As $e$ is not on any of these paths, their passage times are the same in $\omega'$. This gives the second part of condition 2 of $\hat{B}_M^\alpha$ and shows that $\tilde{B}_M^\alpha$ is $e$-increasing.

Now we conclude the proof in a slightly different manner depending on whether or not $\lambda_0^+$ is finite; we focus first on the case that $\lambda_0^+ < \infty$. We will use Lemma 6.6, but several times in sequence, appending events onto $\hat{B}_M^\alpha$. Precisely we note for reference that if $e_1, \ldots, e_j$ are edges and $a_1, \ldots, a_j \in \mathbb{R}$ then

$$\hat{B}_\alpha^M \cap \left[\cap_{i=1}^j \{\omega_{e_i} \geq a_i\}\right] \text{ is } e\text{-increasing for } e \in U'_{\mathcal{E}}\ .$$

Using Lemma 4.5.6 once for each edge $e \in U'_{\mathcal{E}}$ and the upper bound $|U'_{\mathcal{E}}| \leq 32NR$, we can find some constant $C_{N,R}$ such that, defining

$$B_M'^\alpha := \tilde{B}_M^\alpha \cap \left\{\forall e \in U'_{\mathcal{E}},\, \omega_e \geq \lambda^+ - \delta^+/4\right\}\ ,$$

we have

$$\mathbb{P}\left(B_M'^\alpha\right) > C_{N,R} > 0 \text{ for all } M \geq M'_0 \text{ and } \alpha \in \Lambda_{M,k} \text{ when } k \geq K_0(M)\ .$$

(For the first application of the lemma we use $\vartheta = 2^{-2-64NR}\beta'/2R$, for the second, a smaller $\vartheta$, and so on.)

We claim that on $B_M'^\alpha$, no $z \in \mathbb{Z}^2 \cap [-M,M]^2$ with $z \preceq L'_N$ and $z \notin S(R,N)$ has $s_z^\alpha(M) \cap s_{x^*}^\alpha(M) \neq \varnothing$. We argue by first estimating the passage time between vertices from $L'_0$ to $L'_N$ in $U'$. For any outcome in $B'^{\alpha}_M$, given vertices $x \in U' \cap L'_0$ and $y \in U' \cap L'_N$, there is a path from $x$ to $y$ formed by moving along $L'_0$ to 0, taking $r^{\alpha}_0$ to $L'_N$, and moving similarly along $L'_N$ to $y$. This gives

$$\tau(x, y) < (\lambda^+ - 7\delta^+/8)\|w^{\alpha}_N\|_1 + (N\varepsilon + \|x_2\|_1)\lambda^+. \tag{4.40}$$

Using the choice of $\varepsilon$ from (4.28) and condition (4.31) to bound the right side of (4.40),

$$\tau(x, y) \leq (\lambda^+ - 3\delta^+/4)\|w^{\alpha}_N\|_1. \tag{4.41}$$

Suppose now that a point $z$ exists as in the claim. Since $s^{\alpha}_0(M)$ and $s^{\alpha}_{x_2}(M)$ do not touch any $y \notin S(R, N)$ with $y \preceq L'_N$ (see item 1 in the definition of $\hat{B}^{\alpha}_M$),

$$\mathcal{R}^{\alpha}_1(M) \cap \{y : y \preceq L'_N\} \subseteq S(R, N) \ .$$

This implies $z \notin \mathcal{R}^{\alpha}_1(M)$, whereas $x^* \in \mathcal{R}^{\alpha}_1(M)$. As $s^{\alpha}_z(M)$ cannot touch $s^{\alpha}_0(M)$ or $s^{\alpha}_{x_2}(M)$ (else it would merge with one of them) it would have to enter $\mathcal{R}^{\alpha}_1(M)$ through $L'_0$ and pass through all of $U'$ from $L'_0$ to $L'_N$, thus taking only edges of $U'_{\mathcal{E}}$. The portion $\gamma'$ of $\gamma$ from its first intersection with $L'_0$ to its first intersection with $L'_N$ would then satisfy

$$\begin{aligned}\tau(\gamma') &\geq \left(\lambda^+ - \delta^+/4\right)\left[\|w^{\alpha}_N\|_1 - \|x_2\|_1 - N\varepsilon\right] \\ &\geq (\lambda^+ - \delta^+/4)\|w^{\alpha}_N\|_1 - 2\|w^{\alpha}_N\|_1\varepsilon\lambda^+ \\ &\geq (\lambda^+ - 3\delta^+/8)\|w^{\alpha}_N\|_1,\end{aligned}$$

in contradiction with the estimate of (4.41). This establishes the claim.

For the final step in the case that $\lambda^+_0 < \infty$, note that by the previous claim, the pushforward, $\Phi_\alpha(B'^{\alpha}_M)$, is a sub-event of $B'_M = B'_M(x^*; R, N)$, defined exactly as the event $B' = B'(x^*; R, N)$ in (4.27) except with $\Gamma_{x^*}$ and $\Gamma_z$ replaced by the truncated paths $\Gamma_{x^*}^M$ and $\Gamma_z^M$ and considering only $z \in [-M, M]^2$. Thus

$$\mu_\alpha(B'_M) \geq C_{N,R} \text{ for all } M \geq M'_0,\ k \geq K_0(M) \text{ and } \alpha \in \Lambda_{M,k}\ ,$$

with $\Lambda_{M,k} \subseteq [0, n_k]$ of Lebesgue measure at least $C_{\beta'} n_k$. As the indicator of $B'_M$ is continuous,

$$\mu(B'_M) = \lim_{k \to \infty} \mu^*_{n_k}(B'_M) \geq C_{N,R} C_{\beta'}\ .$$

Last,

$$\mu(B') = \mu(B'_M \text{ for infinitely many } M) \geq C_{N,R} C_{\beta'} > 0\ ,$$

completing the proof in the case $\lambda_0^+ < \infty$.

If $\lambda_0^+ = \infty$, we are no longer guaranteed the estimate (4.41), since the passage time of a path taking $N\varepsilon$ steps along $L'_N$ is not necessarily bounded above by $N\varepsilon\lambda^+$. However, writing $\tilde{E}$ for the set of edges with an endpoint within $\ell^1$ distance 1 of $U'$ but not in $U'_{\mathcal{E}}$ and noting

$$A_C := \{\text{for all } e \in \tilde{E},\ \tau_e \leq C\}$$

satisfies $\mathbb{P}(A_C) \to 1$ as $C \to \infty$ independently of $k$ and $M$, we can choose $C_{\text{big}}$ such that

$$\mathbb{P}(\tilde{B}^\alpha_M \cap A_{C_{\text{big}}}) > 0$$

independently of $k$ and $M$. This event is still monotone increasing in the appropriate edge variables. In particular, we can modify the edges in $U'_{\mathcal{E}}$ to be each larger than $2C_{\text{big}}|\tilde{E}|$ and the rest of the proof follows as in the case $\lambda_0^+ < \infty$. □

### 4.5.3 Deriving a contradiction

Given that the event $B'(x^*; R, N)$ of the preceding section has positive probability, we now derive a contradiction, proving that all paths in $\mathbb{G}$ must merge. The next lemma is an example of a mass-transport principle. (See [16, 45, 46] for a more comprehensive treatment.)

**Lemma 4.5.8.** *Let $m : \mathbb{Z}^2 \times \mathbb{Z}^2 \to [0, \infty)$ be such that $m(x, y) = m(x + z, y + z)$ for all $x, y, z \in \mathbb{Z}^2$. Then*

$$\forall x \in \mathbb{Z}^2, \quad \sum_{y \in \mathbb{Z}^2} m(x, y) = \sum_{y \in \mathbb{Z}^2} m(y, x) \ .$$

*Proof.* Write

$$\sum_{y \in \mathbb{Z}^2} m(x, y) = \sum_{z \in \mathbb{Z}^2} m(x, x + z) = \sum_{z \in \mathbb{Z}^2} m(x - z, x) = \sum_{y \in \mathbb{Z}^2} m(y, x) \ .$$

□

Given a realization of $\mathbb{G}$ and $x \in \mathbb{Z}^2$, order the set

$$C_x = \{y \in \mathbb{Z}^2 : y \to x \text{ in } \mathbb{G}\} \tag{4.42}$$

using a dictionary-type ordering where $y$ precedes $y'$ if either $\varpi' \cdot y < \varpi' \cdot y'$ or if both $\varpi' \cdot y = \varpi' \cdot y'$ and $y \cdot \varsigma < y' \cdot \varsigma$ (where $\varsigma$ was fixed in (4.17)); clearly this defines a total ordering. If there is a least element $y$ under this ordering, we will call $y$ the progenitor of $x$ (relative to $\mathbb{G}$). We define the $\mathbb{G}$-dependent function $m_{\mathbb{G}}$ on pairs of vertices $x, y$ by

$$m_{\mathbb{G}}(x, y) = \begin{cases} 1 & \text{if } y \text{ is the progenitor of } x \\ 0 & \text{otherwise,} \end{cases}$$

and let $m(x,y) := \mathbb{E}_\mu(m_\mathbb{G}(x,y))$. Note that $m(x,y) = m(x+z, y+z)$ by the fact that $\mathbb{G}$ has a translation-invariant distribution.

Since each $x$ can have at most one progenitor,

$$\sum_{y \in \mathbb{Z}^2} m(x,y) \leq 1 \text{ for all } x \in \mathbb{Z}^2 \ . \tag{4.43}$$

On the other hand, if $B'(x^*; R, N)$ occurs, then $\Gamma_z$ cannot intersect $\Gamma_{x^*}$ if $z \preceq L'_N$ and $z \notin S(R,N)$. Therefore, on this event, there is some vertex $y \in S(R,N)$ which is the progenitor of infinitely many vertices of $\Gamma_{x^*}$. In particular,

$$\sum_{y \in \mathbb{Z}^2} m(y,x) = \infty. \tag{4.44}$$

The contradiction implied by (4.43), (4.44) and Lemma 4.5.8 gives $\mu(B'(x^*; R, N)) = 0$. However this contradicts the previous section and completes the proof of Theorem 4.5.1.

### 4.5.4 Absence of backward infinite paths

In this section, we move on from Theorem 4.5.1 to show that because all paths in $\mathbb{G}$ coalesce, all paths in the "reverse" direction terminate. That is, recalling the definition of $C_x$ in (4.42),

**Theorem 4.5.9.** *For each $x \in \mathbb{Z}^2$, $|C_x| < \infty$ with $\mu$-probability one.*

**Remark 4.5.10.** *The proof below applies to the following general setting. Suppose $\nu$ is a translation-invariant probability measure on directed subgraphs of $\mathbb{Z}^2$ and there is a line $L \subseteq \mathbb{R}^2$ such that $\nu$-almost surely (a) each $x$ has exactly one forward path and it is infinite (b) all forward paths coalesce and (c) each forward infinite path emanating from a vertex on $L$ intersects it finitely often. Then all backward clusters are finite $\nu$-almost surely.*

We assume that, contrary to the theorem, there exists $x \in \mathbb{Z}^2$ with $\mu(|C_x| = \infty) > 0$ for the remainder of this section to derive a contradiction. Using Lemma 4.5.2, choose a deterministic $\varpi'$ with argument in $\{j\pi/4 : j = 0, \ldots, 7\}$ such that with positive $\mu$-probability on $\{|C_x| = \infty\}$, each $\Gamma_z$ eventually lies on the far side of each $L'_N$. Note that this event is translation-invariant, so by conditioning on it, we may assume that it occurs with probability 1 (and $\mu$ is still translation-invariant).

**Claim 4.5.11.** *There exist vertices $z \neq z'$ in $L'_0$ such that*

$$\mu\left(|C_z| = \infty,\, |C_{z'}| = \infty,\, \Gamma_z \cap L'_0 = \{z\},\, \Gamma_{z'} \cap L'_0 = \{z'\}\right) > 0 \ . \tag{4.45}$$

*Proof.* By translation-invariance, we may assume that the $x$ with $\mu(|C_x| = \infty) > 0$ satisfies $x \prec L'_0$. $\mu$-almost surely, $\Gamma_x$ has a last intersection with $L'_0$. There are countably many choices for such a last intersection, so there exists a vertex $z \in L'_0$ such that

$$\mu\left(|C_z| = \infty,\, \Gamma_z \cap L'_0 = \{z\}\right) > 0 \ .$$

Translating by $\varsigma$ (chosen from (4.17)), the ergodic theorem gives $z, z'$ satisfying (4.45). □

*Proof of Theorem 4.5.9.* Given an outcome in the event in (4.45), $\Gamma_z$ and $\Gamma_{z'}$ almost surely merge. So there is some random $z_{\mathbb{G}} \in \mathbb{Z}^2$ which is the first intersection point of $\Gamma_z$ and $\Gamma_{z'}$ ("first" in the sense of both the ordering in $\Gamma_z$ and in the ordering of $\Gamma_{z'}$). Again $z_{\mathbb{G}}$ can take only countably many values, and so there is a $z_0$ which occurs with positive probability; call the intersection of the event in (4.45) with the event $\{z_{\mathbb{G}} = z_0\}$ by the name $B$.

We now consider the graph $\mathbb{G}$ as an undirected graph, in which vertices $x$ and $y$ are adjacent if $\langle x, y\rangle$ or $\langle y, x\rangle$ are in $\mathbb{G}$ (we abuse notation by using the same symbol for both the directed and undirected versions of $\mathbb{G}$). We define an encounter point of the undirected $\mathbb{G}$ to be a vertex whose removal splits $\mathbb{G}$ into at least three infinite components. Note that $B \subseteq \{z_0 \text{ is an encounter point}\}$; by translation invariance, we see that there is a uniform $c_t > 0$ such that the probability of any fixed vertex to be an encounter point is at least $c_t$.

We are now in the setting of Burton-Keane [23]. To briefly synopsize, the number of points on the boundary of $[-M, M]^2$ must be at least the number of encounter points within. In particular, the number of encounter points is surely bounded above by $8M$. But since each point within has probability at least $c_t$ to be an encounter point, the expected number of encounter points within $[-M, M]^2$ is at least $c_t M^2$. This is a contradiction for large $M$.

□

### 4.5.5 Proof of Theorem 4.2.2

Assume either **A1'** or both **A2'** and the upward finite energy property. Let $v \in \mathbb{R}^2$ be nonzero and $\varepsilon > 0$. We will prove that the statement of the theorem holds with probability at least $1-\varepsilon$. Choose $\varpi \in \partial\mathcal{B}$ to be parallel to $v$ and construct a measure $\mu$ as in Section 3.5.1. Let $(n_k)$ be an increasing sequence such that $\mu^*_{n_k} \to \mu$ weakly.

We will define a double sequence of cylinder events that approximate the events in the theorem. For $m \leq n$, a configuration $\eta \in \Omega_3$ and $x, y \in [-m, m]^2 \cap \mathbb{Z}^2$, we say that $x$ is $n$-connected to $y$ ($x \to_n y$) if there exists a directed path from $x$ to $y$ whose vertices stay in $[-n, n]^2$. We say that $x$ and $y$ are $n$-connected ($x \leftrightarrow_n y$) if there is an undirected path connecting $x$ and $y$ in $[-n, n]^2$. For $m \leq n$ write $A_{m,n} \subseteq \Omega_3$ for the event that

1. all vertices $v \in [-m, m]^2$ have exactly one forward neighbor in $\mathbb{G} \cap [-n, n]^2$,

2. there is no undirected circuit contained in $[-m, m]^2$,

3. for all vertices $v, w \in [-m, m]^2$, there exists $z \in [-n, n]^2$ such that $v \to_n z$ and $w \to_n z$ and

4. for all vertices $v \in [-m,m]^2$ there is no $z \in [-n,n]^2 \backslash (-n,n)^2$ such that $z \to_n v$.

We claim that for any $m$ there exists $n(m) \geq m$ such that $\mu(A_{m,n(m)}) > 1-\varepsilon/4^{m+2}$. To prove this, let $\hat{\Omega} \subseteq \widetilde{\Omega}$ be the event that (a) all vertices have one forward neighbor in $\mathbb{G}$, (b) $\mathbb{G}$ has no undirected circuits, (c) for all $x, y \in \mathbb{Z}^2$, $\Gamma_x$ and $\Gamma_y$ coalesce and (d) $|C_x| < \infty$ for all $x \in \mathbb{Z}^2$. By Proposition 3.7.2, Theorem 4.5.1 and Theorem 4.5.9, the $\mu$-probability of $\hat{\Omega}$ is 1. Therefore conditions 1 and 2 above have probability 1 for all $m$ and $n$. For any configuration in $\hat{\Omega}$ and $m \geq 1$ we can then choose a random and finite $N(m) \geq m$ to be minimal so that conditions 3 and 4 hold for all $n \geq N(m)$. Taking $n(m)$ so large that $\mu(N(m) \geq n(m)) \leq \varepsilon/4^{m+1}$ completes the proof of the claim.

We now pull $A_{m,n(m)}$ back to $\Omega_1$, using the fact that it is a cylinder event in $\Omega_3$ and thus its indicator function is continuous. There is an $m$-dependent number $K_0(m)$ such that if $k \geq K_0(m)$ then $\mu^*_{n_k}(A_{m,n(m)}) > 1 - \varepsilon/4^{m+2}$. By definition of $\mu^*_{n_k}$ in (3.21) and $\Phi_\alpha$ in (3.20), the set $\Lambda_{m,k}$ of values of $\alpha \in [0, n_k]$ such that $\mathbb{P}(\Phi_\alpha^{-1}(A_{m,n(m)})) > 1 - \varepsilon/2^{m+2}$ has Lebesgue measure at least $n_k(1 - 2^{-(m+2)})$.

The next step is to construct a deterministic sequence $(a_m)_{m\geq 1}$ of real numbers such that

$$a_m \to \infty \text{ and } \mathbb{P}\left(\cap_{j=1}^m \Phi_{a_m}^{-1}(A_{j,n(j)})\right) \geq 1 - \varepsilon/2 \text{ for all } m \ . \tag{4.46}$$

We do this by induction on $m$. For $m = 1$, let $a_1$ be any number in the set $\Lambda_{1,K_0(1)}$. By definition then $\mathbb{P}(\Phi_{a_1}^{-1}(A_{1,n(1)})) \geq 1-\varepsilon/2$. Assuming that we have fixed $a_1, \ldots, a_m$, we now define $a_{m+1}$. Let $k$ be such that $k \geq \max\{K_0(1), \ldots, K_0(m+1)\}$ and $n_k \geq 3a_m$ and consider $\Lambda_{1,k}, \ldots, \Lambda_{m+1,k}$ as above. The intersection of these sets has Lebesgue measure at least $3n_k/4$ so choose $a_{m+1}$ as any element of the nonempty set $(3a_m/2, n_k] \cap \left[\cap_{i=1}^{m+1} \Lambda_{i,k}\right]$. For this choice,

$$1 - \mathbb{P}\left(\cap_{j=1}^{m+1} \Phi_{a_{m+1}}^{-1}(A_{j,n(j)})\right) \leq \sum_{j=1}^{\infty} \varepsilon/2^{j+2} = \varepsilon/4 \ .$$

As $a_{m+1} \geq 3a_m/2$, the condition $a_m \to \infty$ holds and we are done proving (4.46).

From (4.46), we deduce $\mathbb{P}(A) \geq 1 - \varepsilon/2$, where

$$A = \{\cap_{j=1}^{m} \Phi_{a_m}^{-1}(A_{j,n(j)}) \text{ occurs for infinitely many } m\} \ .$$

We complete the proof by showing that the statement of the theorem holds for any $\omega \in A$. Fix such an $\omega$ and a random subsequence $(a_{m_k})$ of $(a_m)$ such that $\omega \in \cap_{j=1}^{m_k} \Phi_{a_{m_k}}^{-1}(A_{j,n(j)})$ for all $k$. By extracting a further subsequence, we may assume that $\mathbb{G}_{L_{a_{m_k}}(\varpi)}$ converges to some graph $G$. The event $\Phi_\alpha^{-1}(A_{j,n(j)})$ is exactly that the graph $\mathbb{G}_{L_\alpha(\varpi)}$ satisfies the conditions of $A_{j,n(j)}$ above, so in particular, it has no undirected circuits in $[-j,j]^2$, all directed paths starting in $[-j,j]^2$ coalesce before leaving $[-n(j),n(j)]^2$, no directed paths connect $[-n(j),n(j)]^2 \setminus (-n(j),n(j))^2$ to $[-j,j]^2$, and all vertices in $[-j,j]^2$ have one forward neighbor in $[-n(j),n(j)]^2$. On the subsequence $(a_{m_k})$, the events $\Phi_{a_{m_k}}^{-1}(A_{1,n(1)})$ occur for all $k$, so $G$ must satisfy the conditions of $A_{1,n(1)}$ as well. The same is true for $A_{j,n(j)}$ for all $j$, so $G$ satisfies the conditions of the theorem.

### 4.5.6 Proof of Theorem 4.2.3

This theorem follows directly from results of the previous sections. Assume either **A1'** or both **A2'** and the upward finite energy property. For the first part of the theorem, suppose that $\partial\mathcal{B}$ is differentiable at $v_\theta$. Choose $\varpi = v_\theta$ and construct the measure $\mu$ as in Section 3.5.1. Given $(\omega, \Theta, \eta) \in \widetilde{\Omega}$, let $\mathbb{G}(\eta)$ be the geodesic graph associated to $\eta$. By Theorems 3.7.3, 4.5.1 and 4.5.9, with $\mu$-probability one, all directed paths in $\mathbb{G}$ are asymptotically directed in $I_\theta$, they coalesce, and no vertex $x$ has $|C_x|$ infinite. Call this event $A$ and define

$$\hat{\Omega} = \{\omega \in \Omega_1 : \mu(A \mid \omega) = 1\} \ .$$

$\mu(\cdot \mid \omega)$ is the regular conditional probability measure. $\hat{\Omega}$ is a measurable set and satisfies $\mathbb{P}(\hat{\Omega}) = 1$ since the marginal of $\mu$ on $\Omega_1$ is $\mathbb{P}$. Further, for each $\omega \in \hat{\Omega}$, the theorem holds.

For the other two parts of the theorem we simply argue as in the proof of Corollaries 3.3.4 and 3.3.5. In the former case we just notice that if $v_\theta$ is also exposed, then $I_\theta = \{\theta\}$. In the latter case, we find a point $v_\theta$ on the arc joining $v_{\theta_1}$ to $v_{\theta_2}$ at which $\partial\mathcal{B}$ is differentiable. The set $I_\theta$ contains only angles associated to points on the arc and we are done.

# Part II

# Random Walks on the Invasion Percolation Cluster

# Chapter 5

# Critical Percolation Models

In this chapter, we introduce some fundamental facts about percolation models – especially critical percolation – and motivate the results of [35] in preparation for the final chapter of this dissertation. The account here is far from exhaustive and does not track the historical development of the subject. We direct the interested reader to the volumes [20, 43] and the recent paper [44].

## 5.1 Bernoulli Percolation

Percolation theory was introduced by Broadbent and Hammersley [22] as a prototypical model for random media. We will be concerned here chiefly with the two-dimensional case of independent Bernoulli nearest-neighbor bond percolation, which we will take license to abbreviate as Bernoulli percolation or simply percolation (the definition of the model in higher dimensions will be clear; for results in higher dimensions, see one of the above references). This is a model on the usual two-dimensional square lattice $(\mathbb{Z}^2, \mathcal{E}^2)$. The model lives on a family of probability spaces parametrized by some $p \in [0, 1]$. These spaces take the form $(\Omega, \mathcal{B}, \mathbb{P}_p)$; here $\Omega = \{0, 1\}^{\mathcal{E}^2}$, $\mathcal{B}$ is the usual Borel sigma-algebra, and $\mathbb{P}$ is the i.i.d. measure with $\mathbb{P}(\omega_e = 1) = p$.

A given realization $\omega \in \Omega$ can be taken to define a random graph structure on $\mathbb{Z}^2$. An edge $e$ will be said to be "open" for $\omega$ if $\omega_e = 1$; otherwise, $e$ is said to be "closed." Then two vertices $x, y \in \mathbb{Z}^2$ are said to be connected by an open path for $\omega$ if there exists a path $(x = v_1, e_1, v_2, \ldots e_{n-1}, v_n = y)$, where $v_i \in \mathbb{Z}^2$, $e_i = \{v_i, v_{i+1}\} \in \mathcal{E}^2$, and $\omega_{e_i} = 1$ for all $i$. The open cluster $C(x)$ of a site $x$ is the set of all sites connected to $x$ by open paths. Analogous notions exist for closed paths and closed clusters. Many of the important questions in the model relate to the properties of typical open and closed clusters for different values of the parameter $p$.

A well-known property of open-connectedness is that there is a critical probability $p_c = 1/2$ [62] with the following properties (among others):

1. For $p \leq p_c$,
$$\mathbb{P}_p\left(|C(x)| = \infty\right) = 0;$$

2. For $p > p_c$,
$$\mathbb{P}_p\left(|C(x)| = \infty\right) > 0.$$

If for a realization $\omega$ there is some $x$ with $|C(x)| = \infty$, we say there "is percolation for $\omega$". The regime $p < p_c$ corresponds to exponential decay of the radius of $C$:

$$\mathbb{P}_p\left(C(0) \text{ contains vertices at Euclidean distance } n \text{ from } 0\right) \leq 2\exp(-Kn)$$

for some $K > 0$ depending on $p$.

By contrast, if $p = p_c$, open clusters are finite but typically large:

$$\mathbb{P}_p\left(C(0) \text{ contains vertices at Euclidean distance } n \text{ from } 0\right) \geq \frac{1}{8n}.$$

At the critical point, connections tend to be long but sparse. For instance, by duality the probability that there exists an open connection between a point on the left and a point on the right side of a square of side length $n$ (an "open crossing") is 1/2 independently of $n$. However, for $n$ large, the typical number of disjoint open crossings remains of order one (see [17]).

### 5.1.1 Random Walks on Open Clusters.

A topic which has seen some study (see the introduction to the next chapter for a longer account of the literature) is the behavior of a random walk on an open percolation cluster. This is described formally by introducing a new probability space, the space of trajectories $(X_n)_n$ of a particle in discrete time. For a given realization $\omega$ of the percolation cluster, the Markov chain of the walk $(X_n)_n$ is defined as follows: $X_0 = 0$ almost surely; after time zero, conditioned on $X_{n-1}$, $X_n$ is chosen uniformly from the neighbors $y$ of $X_{n-1}$ such that the edge $\{y, X_{n-1}\}$ is open for $\omega$.

Of most interest are the long-time asymptotics of the random walk on an infinite percolation cluster. If $p > p_c$, this is achieved by drawing from the percolation measure $\mathbb{P}$ conditioned on the event $|C(0)| = \infty$ and then running the above random walk $(X_n)_n$. In this case, the infinite open cluster of the origin has density [23] in the plane, and one might expect the random walk on $C(0)$ to behave much like a random walk on $(\mathbb{Z}^2, \mathcal{E}^2)$. This has been shown in a strong sense through the work of numerous authors over a number of years. We direct the reader to the recent paper [18], in which a version of convergence of the walk $(X_n)_n$ to isotropic Brownian motion is shown; the introduction and references of [18] contain some review of the history of these problems.

On the other hand, the exponential decay of the radius of $C(0)$ for $p < p_c$ implies that the system has a finite correlation length in some sense, and the typical open cluster will be an isolated speck in any sort of scaling limit. The expected cluster size at $p = p_c$ is infinite, however, and one could ask the question of whether there is some meaningful way of asking about the long-time behavior of a random walk on an "arbitrarily large critical cluster." There has been some work on ascribing meaning to the idea of an infinite open cluster at criticality, which we describe in what follows.

### 5.1.2 The Incipient Infinite Cluster

One method by which one could try to describe a large cluster "at criticality" is to condition the critical percolation measure on the existence of a connection to infinity. That is, to define the limit

$$\nu(\cdot) = \lim_{n\to\infty} \mathbb{P}_{p_c}\left(\cdot \mid 0 \mapsto \partial B(n)\right), \tag{5.1}$$

where $B(n) = [-n, n]^2$ and $\partial B(n)$ is its boundary, and where $0 \mapsto \partial S$ means that there exists an open path connecting 0 and some vertex of the set $S$.

The existence of the limit of measures (5.1) was proved by Kesten [64]. The typical cluster $C(0)$ sampled from $\nu$ is infinite but has zero density (in fact, it occupies a fraction of $B(n)$ which goes to zero like a power of $n$). This signals that the geometry of such a cluster is quite different from that in the case $p > p_c$, and signals that a random walk on $C(0)$ would no longer be guaranteed to look like a diffusion on $(\mathbb{Z}^2, \mathcal{E}^2)$.

The measure $\nu$ defined in (5.1) is frequently referred to as the distribution of the "incipient infinite cluster" or IIC. The behavior of the random walk $(X_n)_n$ on the IIC was studied in [65], and a major point of difference between the IIC and $\mathbb{Z}^2$ random walks was shown. Recall that a random walk in $(\mathbb{Z}^2, \mathcal{E}^2)$ started at 0 typically moves a distance of order $n^{1/2}$ after $n$ steps. For the purpose of stating the theorem, consider $(X_n)_n$ to be a family of random variables on the enlarged probability space of IIC configurations and walk trajectories.

**Theorem 5.1.1** ([65]). *There exists some $\varepsilon > 0$ such that the family $\{n^{-1/2+\varepsilon} X_n\}_n$ is tight.*

The above result is described as a "subdiffusivity" result–the walk moves more slowly than a diffusion on the square lattice in an "averaged sense". The corresponding result in a quenched setting is proved in Chapter A:

**Theorem 5.1.2.** *Let $\tau(n)$ be the smallest $N$ such that $X_N \notin B(n)$. There exists $\varepsilon > 0$ such that, for $\nu$-almost every realization of the IIC, for almost every realization of the random walk,*

$$\tau(n) \geq n^{2+\varepsilon}$$

*for all $n$ larger than some random $n_0$.*

The methods used in Chapter A simplify those of [65] using recent results on the tortuousity of open paths (specifically, those of [1, 82]).

### 5.1.3 Invasion Percolation

Another possible description of an infinite critical cluster is the so-called "Invasion Percolation Cluster" or IPC. This model was introduced and numerically studied in [26, 96]. The construction of the IPC measure on $\Omega$ proceeds by steps. Let $\{t(e)\}$ denote a family of independent uniform random variables on $[0, 1]$ indexed by $e \in \mathcal{E}^2$. Note that here we describe only the two-dimensional IPC, the case of most interest in what follows.

Given a realization of the family $\{t(e)\}$, define the graph $G_0 = (V_0, E_0)$, where the initial vertex set $V_0 = \{0\}$ and $E_0 = \varnothing$. Given $G_{i-1}$, consider the edges incident to $V_{i-1}$ (i.e., the edges of $\mathcal{E}^2$ which touch a vertex of $V_{i-1}$ but do not lie in $E_{i-1}$); denote by $e_i$ the incident edge such that $t(e_i)$ is minimal. Then $E_i$ is defined by

$$E_i = E_{i-1} \cup \{e_i\}$$

and $V_i$ is the union of $V_{i-1}$ and the vertices touched by $e_i$. The graph $G_i = (V_i, E_i)$; finally, the IPC is defined by the limit

$$S = \lim_{n \to \infty} G_n.$$

The distribution of the IPC will here be denoted by $\mu$.

Much past work has been devoted to the characterization of the IPC and its relation to the IIC. It is known that in certain senses, the IPC asymptotically "looks like" the IIC [56]. A natural question that arises is whether the subdiffusivity of the random walk on the IIC means that the corresponding behavior holds on the IPC.

The second contribution of [35], as will be presented in Chapter A, is to show that this subdiffusivity holds using adaptations of techniques from [65]. Moreover, a lower bound is given for the subdiffusivity exponent $\varepsilon$ in terms of the "arm exponents" for critical percolation in two dimensions.

### 5.1.4 Idea of proof

Because [35] has to deal with a number of technical issues, we describe here the major ideas of the proofs in the case of the IPC. Recall that the main goal of the paper is to demonstrate a subdiffusive upper bound for the speed of a random walk on a typical realization of the IPC. As we will see in the Appendix, we can prove this result in two versions with different techniques. One form of the theorem, which parallels the presentation of the theorem of Kesten, states

**Theorem 5.1.3.** *Let $\tau(n)$ be the smallest $N$ such that $X_N \notin B(n)$. There exists $\varepsilon > 0$ such that, for $\nu$-almost every realization of the IPC, for almost every realization of the random walk,*

$$\tau(n) \geq n^{2+\varepsilon}$$

*for all $n$ larger than some random $n_0$.*

The proof of Theorem 5.1.3 is based on results on the intrinsic or chemical distance in the IPC which did not exist at the time of [65]. Let $\mathrm{dist}_{\mathrm{IPC}}(x, y)$, the intrinsic distance, denote the length of the shortest path in the IPC between vertices $x$ and $y$ (we define $\mathrm{dist}_{\mathrm{IPC}}(A, B)$ for sets $A, B$ in the usual fashion). Building on techniques of Aizenman and Burchard [1], a result of Pisztora [82] says that the IPC distance between the origin and the boundary of $[-n, n]^2$ is typically very large, in the following sense. There exist constants $C, C'$ and some $s > 1$ such that

$$\mathbb{P}\left(\mathrm{dist}_{\mathrm{IPC}}(0, \partial B(n)) \leq Cn^s\right) \leq C'n^{-2}.$$

Therefore, asymptotically, a random walk must go on the order of $n^s$ steps to reach the boundary of a box of size $n$.

However, there exist general bounds for reversible Markov chains [25, 90] which imply that the random walk on the IPC is at most diffusive in the intrinsic distance in the sense that

$$\mathrm{dist}_{\mathrm{IPC}}(0, X_k)^2 \sim k.$$

These two facts imply Theorem 5.1.3. This approach has the advantage of providing a relatively rapid proof of subdiffusivity using recent percolation technology.

On the other hand, using methods parallel to Kesten's original proof, a stronger result can be derived. This takes the form

**Theorem 5.1.4.** *Let $\tau(n)$ be as before. There exists $\varepsilon > 0$ such that, for $\nu$-almost every realization of the IPC, for almost every realization of the random walk,*

$$\tau(n) \geq n^{2+\kappa+\varepsilon}$$

*for all $n$ larger than some random $n_0$. Here $\kappa$ is a constant which can be bounded away from zero in terms of physically relevant quantities (see the discussion below).*

To discuss $\kappa$ and the techniques used to prove the theorem, we briefly describe "arm exponents" in critical percolation (for more discussion, see the introduction to the Appendix). A fact of life in critical percolation is that a given point is connected to the boundary by an open path (or two disjoint open paths) by a power law. In the following equations, we use "connected" to mean "connected by an open path".

$$\mathbb{P}_{p_c}(0 \text{ is connected to } [n, \infty) \times \mathbb{R}) \leq Cn^{-\eta_1} \tag{5.2}$$

$$\mathbb{P}_{p_c}(0 \text{ has two disjoint connections to } \partial B(n)) \leq Cn^{-\eta_2}. \tag{5.3}$$

We call the event in (5.2) a "one-arm event" and $\eta_1$ a "one-arm exponent"; two-arm events are defined analogously.

Kesten's insight was to note that in some sense, the IIC in the box $B(n)$ looks like a collection of points at which a one-arm event occurs, but to travel to the boundary $\partial B(n)$ requires one move along the backbone–the collection of sites with disjoint open connections to 0 and $\partial B(n)$–which are sites at which a two-arm event occurs. Moreover, the random walk should spend an amount of time on the backbone roughly equal to the volume fraction of the IIC it occupies. Because two-arm points are rarer than one-arm points, this implies that the random walk "wastes time" going down dead-ends which offer no hope of reaching $\partial B(n)$.

Using an argument based on Kesten's, Theorem 5.1.4 is shown with a $\kappa$ bounded below by $\eta_1\eta_2/2$.

# Appendix A

# Subdiffusivity of Random Walk on the 2d Invasion Percolation Cluster

Michael Damron, Jack Hanson, and Philippe Sosoe

arXiv:1210.7267

## A.1 Abstract

We derive quenched subdiffusive lower bounds for the exit time $\tau(n)$ from a box of size $n$ for the simple random walk on the planar invasion percolation cluster. The first part of the paper is devoted to proving an almost sure analog of H. Kesten's subdiffusivity theorem for the random walk on the incipient infinite cluster and the invasion percolation cluster using ideas of M. Aizenman, A. Burchard and A. Pisztora. The proof combines lower bounds on the instrinsic distance in these graphs and general inequalities for reversible Markov chains. In the second part of the paper, we present a sharpening of Kesten's original argument, leading to an explicit almost sure lower bound for $\tau(n)$ in terms of percolation arm exponents. The methods give $\tau(n) \geq n^{2+\epsilon_0+\kappa}$, where $\epsilon_0 > 0$ depends on the instrinsic distance and $\kappa$ can be taken to be $\frac{5}{384}$ on the hexagonal lattice.

## A.2 Introduction

H. Kesten has proved [65] that the simple random walk $\{X(n)\}_{n\geq 0}$ started at $\mathbf{0}$ on the incipient infinite cluster (IIC) [64] in two-dimensional Bernoulli bond percolation is subdiffusive in the sense that there exists $\epsilon > 0$ such that the family

$$\{n^{-1/2+\epsilon}X(n)\}_{n\geq 0} \tag{A.1}$$

is tight. The purpose of the current work is to explain how a "quenched" version of this result can be obtained and extended to the random walk in an environment generated by a related two-dimensional model, invasion percolation. (The model is defined in the next section). We present some refinements of Kesten's method, which provides a general framework for proving subdiffusivity of random walks in stochastic geometric models. In the case of two-dimensional invasion percolation (as well as the incipient infinite cluster), the ideas in [65] can be used to give explicit bounds on $\epsilon$ from (A.1) in terms of known critical exponents (see (A.2) and (A.3) below).

Our main result is the following:

**Theorem A.2.1.** *Let $\{X(k)\}_{k\geq 0}$ be a simple random walk on the invasion percolation cluster (IPC), and $\tau(n)$ the first time $X(k)$ exits the box $S(n) = [-n, n]^2$:*

$$\tau(n) = \inf\{k \geq 0 : |X(k)|_\infty = n\}.$$

*There exists $\epsilon_0 > 0$ such that, for almost every realization of the random walk, and almost every realization of the IPC, there is a (random) $n_0$ such that*

$$\tau(n) \geq n^{2+\kappa+\epsilon_0} \text{ for } n \geq n_0.$$

*$\kappa$ is a constant that can be estimated in terms of the behaviour of the one-arm and two-arm probabilities in critical percolation (with measure $\mathbb{P}_{p_c}$):*

$$\kappa \geq \frac{1}{2}\eta_1\eta_2,$$

*where $\eta_1, \eta_2 > 0$ are exponents such that*

$$\mathbb{P}_{p_c}(\mathbf{0} \text{ is connected to } [n,\infty)\times\mathbb{R}) \leq \mathbf{C}_1 n^{-\eta_1} \tag{A.2}$$

*and*

$$\mathbb{P}_{p_c}\begin{pmatrix}\mathbf{0} \text{ has two disjoint open} \\ \text{connections to } S(n)^c\end{pmatrix} \leq \mathbf{C}_2 n^{-\eta_2}, \tag{A.3}$$

*for some constants $\mathbf{C}_1$ and $\mathbf{C}_2$.*

**Remark A.2.2.** *If one repeats the arguments of this paper in the setting of random walk on the IIC or IPC of the hexagonal lattice, one can use the exact values of* *the one-arm and two-arm exponents to give a stronger bound on $\kappa$. Indeed, it is not necessary in that case to use the van den Berg-Kesten inequality [17] in* (A.22)*, therefore giving*

$$\kappa \geq (1/2)\eta_2(\eta_2 - \eta_1).$$

*Using the conjectured value* $\eta_2 = \frac{5}{48} + \frac{1}{4}$ *[13], we get a lower bound* $\frac{17}{384}$*. Without using this value, but using* $\eta_2 \geq 1/4$ *[13], we get* $\kappa \geq \eta_1/8 = \frac{5}{384}$*.*

**Remark A.2.3.** *This result is stronger than the corresponding theorem for the random walk on the IIC stated in [65, Theorem 1.27], but it is derived by a modification of the strategy used there. In particular, Kesten proves that*

$$\mathbb{P}(\tau(n) \geq n^{2+\epsilon}) \to 1,$$

*for the "averaged" measure* $\mathbb{P}$*, which incorporates averaging with respect to the IIC measure constructed in [64]. Closer examination of his proof reveals that one can take* $\epsilon = \eta_1^2/4$*, and that the estimates in [65] are sufficient to establish a "quenched" result by a simple application of the Borel-Cantelli lemma. A substantial part of the present paper is concerned with presenting arguments to overcome the difficulties in adapting Kesten's proof to the invasion percolation cluster.*

The second result of the paper concerns a simple derivation of subdiffusivity of random walk on the IPC using results in [1] and [82] concerning the length of the shortest path from the origin to $\partial S(n)$ (the *chemical distance*) in near-critical percolation. The work of these authors implies that for large $n$ this length is of order at least $n^s$, where $s > 1$. Although Theorem A.2.4 is contained in Theorem A.2.1, it is of interest because its proof represents a significant reduction in complexity from the original argument of Kesten.

**Theorem A.2.4** (Quenched Kesten theorem for the IPC)**.** *Let* $\tau(n)$ *be the time for a random walker on the invasion percolation cluster to exit* $S(n)$*. There exists* $\epsilon > 0$

*such that, for $\mathbb{P}_{\mathrm{IPC}}$-almost every $\omega$ and almost-every realization of the random walk,*

$$\tau(n) \geq n^{2+\epsilon}$$

*for $n$ greater than some random $n_0$.*

**Remark A.2.5.** *A similar, but simpler, argument applies to the incipient infinite cluster and gives an alternative proof that the random walk on the IIC is almost surely subdiffusive. See the Appendix for details.*

**Remark A.2.6.** *$\epsilon > 0$ in the statement of Theorem A.2.4 depends on the value of $s$ obtained by the methods of Aizenman-Burchard and Pisztora. $s$ is both very small and difficult to calculate explicitly. Kesten's comparison argument (explained in Section A.5) yields an improvement of the estimate for $\tau(n)$ in the previous theorem by a factor of the form $n^\kappa$, which leads to Theorem A.2.1. We note that any explicit bound on $s$ would be directly reflected in that theorem. Indeed, if one has upper and lower bounds (with high enough probability)*

$$Cn^{s_1} \leq dist_{IPC}(x, y) \leq Cn^{s_2}, \;\; x, y \in IPC$$

*then one can get the lower bound $\tau(n) \geq Cn^a$ for any $a$ satisfying*

$$a < 2s_1 + \eta_1 \left( 2\frac{s_1}{s_2} - \frac{2-\eta_2}{s_2} \right).$$

*On the hexagonal lattice, this can be improved as above to*

$$a < 2s_1 + (\eta_2 - \eta_1) \left( 2\frac{s_1}{s_2} - \frac{2-\eta_2}{s_2} \right).$$

*One can actually show $s_2$ can be taken to be $2 - \eta_2$, which yields the improved bound (assuming again the exact value of $\eta_2$)*

$$\kappa \geq \frac{\eta_2(\eta_2 - \eta_1)}{2 - \eta_2} = \frac{17}{316}.$$

**Remark A.2.7.** *The improvement due to $s_1$ and $s_2$ in the previous remark comes from choosing $q$ larger in* (A.16)*. It is actually a common misconception that Kesten's original "lost in the bushes" argument gives a lower bound for $\tau(n)$ proportional to the ratio of volume of the IIC to the volume of its backbone. The reason this is false is that it is not clear how to increase $q$ to order $n$. The parameter $q$ gives the scale at which volume estimates can be applied.*

There has been little success with rigorous results for random walks on low-dimensional critical models (for instance, the IIC and IPC). One notable example is the work of D. Shiraishi [87] on random walk on non-intersecting two-sided random walk trace. For results in high dimensions, we mention the recent work of G. Kozma and A. Nachmias [69] on the IIC in dimensions $d \geq 19$ and of M. Barlow, A. Járai, T. Kumagai and G. Slade on the IIC for oriented percolation [11]. On a critical Galton-Watson tree, Kesten [65] found the asymptotics of $\tau(n)$ and constructed a scaling limit for random walk on the IIC (see also [32] and [12]). Later, O. Angel, J. Goodman, F. den Hollander and G. Slade [5] found similar results for random walk on the IPC on a regular tree.

After setting some notation below, we give the definition of the invasion percolation model in Section A.3, and recall some useful properties of the IPC derived in previous literature. We then prove Theorem A.2.4 in Section A.4, and explain how Kesten's volume comparison argument is used to obtain Theorem A.2.1 in Section A.5. Section A.6 contains the derivation of estimates used in the proof of Theorem 1.

For convenience, we work on the square lattice $\mathbb{Z}^2$, but our results extend to planar lattices for which the Russo-Seymour-Welsh estimates hold true.

### A.2.1 Notation

In this section, we give notation used throughout the paper for future reference. For any vertex (lattice point) $v = (v_1, v_2) \in \mathbb{Z}^2$, $S(n, v)$ is the box

$$\begin{aligned} S(n,v) &= ([v_1 - n, v_1 + n] \times [v_2 - n, v_2 + n]) \cap \mathbb{Z}^2 \\ &= \{x \in \mathbb{Z}^2 : |x - v|_\infty \le n\}. \\ |x|_\infty &= \max(|x_1|, |x_2|). \end{aligned}$$

$S(n)$ is the box $S(n, \mathbf{0})$, centred at the origin. $\partial S(n, v)$ refers to the internal vertex boundary of $S(n, v)$:

$$\partial S(n,v) = \{x \in S(n) : \exists y \in S(n)^c, |x - y|_\infty = 1\}$$

We also define

$$\Lambda(n) = S(n) \cap \mathrm{IPC},$$

where IPC is defined in the next section. For a graph $G$, the set of edges is denoted by $E(G)$.

For each $p \in [0,1]$, the independent bond percolation measure $\mathbb{P}_p$ is an infinite product of Bernoulli measures with parameter $p$ indexed by the edges of $\mathbb{Z}^2$. For a finite set $I \subset E(\mathbb{Z}^2)$, and a vector $v \in \{0,1\}^I$ we have

$$\mathbb{P}_p(\sigma \in \{0,1\}^{E(\mathbb{Z}^2)} : \sigma(e) = v(e) \text{ for } e \in I) = p^{\sharp\{e : v(e) = 1\}}(1-p)^{\sharp\{e : v(e) = 0\}}.$$

A *configuration* $\sigma$ is an element of $\{0,1\}^{E(\mathbb{Z}^2)}$. An edge $e$ is said to be *open* in the configuration $\sigma$ if $\sigma(e) = 1$, and *closed* otherwise.

If $A$ and $B$ are subsets of $\mathbb{Z}^2$ we denote by

$$\mathbb{P}_p(A \to B)$$

the probability of the event that $A$ and $B$ are connected by a path of open edges. The notation $\mathbb{P}_{\text{IPC}}(A \to B)$ is defined analogously. We write

$$\mathbb{P}(A \xrightarrow{p} B)$$

to denote the probability that $A$ and $B$ are connected by $p$-open edges.

We will use the connection probabilities $\pi$ and $\rho$ defined as

$$\pi(p,n) = \mathbb{P}_p(\mathbf{0} \to [n,\infty) \times \mathbb{R})$$
$$\rho(p,n) = \mathbb{P}_p(\mathbf{0} \to \partial S(n) \text{ by two disjoint open paths}).$$

These probabilities refer to independent bond percolation with parameter $p$. When no parameter is specified, it is understood that $p = p_c = 1/2$; that is,

$$\pi(n) = \pi(p_c, n), \;\; \rho(n) = \rho(p_c, n).$$

We denote by $\mathbb{P}_{\text{IPC}}$ the invasion percolation measure on bond configurations in $\mathbb{Z}^2$. Throughout, $\omega$ will denote a realization of the IPC; that is, a subgraph of $\mathbb{Z}^2$ sampled from $\mathbb{P}_{\text{IPC}}$. For each such $\omega$, we denote by $\mathbf{P}^\omega$ the probability measure associated with the simple random walk on the invasion cluster in the realization $\omega$ (which by definition contains the origin $\mathbf{0} = (0,0) \in \mathbb{Z}^2$).

For $x > 0$ , we denote by $\log x = \log_2 x$ the logarithm of $x$ in base 2.

Throughout the paper, $\mathbf{C}_i$ will denote constants chosen independent of $n$. We use the notation $A \lesssim B$ if there exists a constant $C$ such that

$$A \leq CB.$$

This notation is only used if the implicit constant $C$ is deterministic; that is, it does not depend on the realization of the IPC or of the random walk. The notation

$$A \lesssim_c B$$

is used to emphasize that the implicit constant depends on the parameter $c$.

The notation $A \asymp B$ denotes the existence of two positive constants $D_1$ and $D_2$ such that

$$D_1 B \leq A \leq D_2 B.$$

If $f(n)$ and $g(n)$ are two positive sequences, we use the notation $f(n) \gg g(n)$ to mean

$$\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$$

## A.3 Invasion percolation

### A.3.1 Definition of the model

The planar invasion percolation cluster is a random subgraph of the lattice $\mathbb{Z}^2$ which can be constructed from the familiar coupling of the independent bond percolation measures $\mathbb{P}_p$, $0 < p < 1$. To every edge $e$ of $\mathbb{Z}^2$, viewed as a graph, associate a random variable $w(e)$, uniformly distributed in $[0, 1]$, $\{w(e) : e \in E\}$ and $\{w(e') : e' \in E'\}$ being independent for $E \cap E' = \emptyset$. An edge is called *$p$-open* if $w(e) \leq p$ and is *$p$-closed* otherwise. The distribution of the set of $p$-open edges is that of a Bernoulli bond-percolation process at density $p$. The distribution of $(w(e))_{e \in E(\mathbb{Z}^2)}$ is a product measure which will be denoted by $\mathbb{P}$.

The IPC consists of a union of subgraphs of $\mathbb{Z}^2$ constructed by an iterative process: we start at the origin $\{0, \emptyset\} \equiv G_0$. At every stage, we form $G_{i+1}$ by adding to the current (finite) graph $G_i$ the edge $e$ with the least weight $w(e)$ among

$$\Delta G_i = \{e = (u, v) \in E(\mathbb{Z}^2), e \notin E(G_i) \text{ but } u \in G_i\}$$

as well as the endpoints of $e$. The IPC is defined to be the union $\bigcup_{i \geq 0} G_i$.

Since the percolation probability $\theta(p) = \mathbb{P}_p(\mathbf{0} \to \infty)$ at $p_c = 1/2$ is zero [62], the IPC contains infinitely many edges $e$ with $w(e) > 1/2$. On the other hand, for any $p > p_c$, by the Russo-Seymour-Welsh theorem, the IPC will intersect the (unique) $p$-open infinite cluster almost surely (see [28] for general $d$). By construction, once an edge $e$ in the $p$-open infinite cluster has been added, all edges added to the IPC after $e$ have weight no bigger than $p$.

### A.3.2 Correlation length

We will later require bounds for the probability that the IPC intersects the $p$-open infinite cluster, for some fixed $p$, by the time it reaches an annulus of size $n$. Such estimates can be found in [97], [56]. An important notion in this context is the *finite-size scaling length* $L(p, \epsilon)$. To define it, consider for $p > p_c$ the probability

$$\sigma(n, m, p) = \mathbb{P}_p(\exists \text{ a } p\text{-open horizontal crossing of } [0, n] \times [0, m]).$$

Then $L(p, \epsilon)$ is defined to be

$$L(p, \epsilon) = \min_{n \geq 0}\{\sigma(n, n, p) \geq 1 - \epsilon\}.$$

From [66], it is known that $L(p, \epsilon_1) \asymp L(p, \epsilon_2)$ for $0 < \epsilon_1, \epsilon_2 \leq \epsilon_0$, so we shall fix $\epsilon = \epsilon_0$ and henceforth simply refer to $L(p) = L(p, \epsilon)$. We note the following properties of $L$:

1. $L(p)$ is right-continuous, non-increasing in $(p_c, 1)$ and $L(p) \to \infty$ as $p \downarrow p_c$.

2. Taking $\epsilon_0$ small enough, there exists $K > 0$ such that [56, (2.8)]:

$$\sigma(2mL(p), mL(p), p) \geq 1 - \exp(-Km), \; m \geq 1.$$

3. Again from [56, Eq. 2.10], there exists $D$ independent of $p$ such that

$$\lim_{\delta \downarrow 0} \frac{L(p-\delta)}{L(p)} \leq D.$$

Let $\log^{(j)}$ be the $j$-th iterate of log, and

$$\log^* n = \min\{j > 0 : \log^{(j)} n \leq 16\}, \; n \geq 16.$$

Define, for $n \geq 16$ and $j = 1, \ldots, \log^* n$

$$p_n(j) = \min_{p > p_c} \left\{ L(p) \leq \frac{n}{M \log^{(j)} n} \right\}.$$

$M > 0$ is a constant to be determined later. Note that if $m \leq n$, then $p_n(1) \geq p_m(1)$ when $m$ is sufficiently large.

By (3) above, there exists $D > 0$ such that

$$M \log^{(j)} n \leq \frac{n}{L(p_n(j))} \leq DM \log^{(j)} n. \tag{A.4}$$

Item (2) in the list above implies [56, (2.21)]

$$\mathbb{P}(H_n(j)) \equiv \mathbb{P}\begin{pmatrix} \exists\, p_n(j)\text{-open circuit } \mathcal{C} \text{ around } 0 \text{ in } S(n/2)\setminus S(n/4) \\ \text{and } \mathcal{C} \text{ is connected to } \infty \text{ by a } p_n(j)\text{-open path} \end{pmatrix} \tag{A.5}$$
$$\geq 1 - \mathbf{C}_3 \exp(-\mathbf{C}_0 M \log^{(j)} n).$$

The measure $\mathbb{P}$ refers to the coupling of the $p$-Bernoulli measures described earlier. If the event $H_n(j)$ occurs, the IPC intersects the $p_n(j)$-open infinite cluster by the time it reaches $S(n)^c$. The bound (A.5) plays a role in estimates derived in Section A.6.

## A.4 Proof of Theorem A.2.4

We begin by giving a brief sketch of the main idea. The first step is to consider a restriction of the random walk to a certain subset of the IPC, the backbone. The exit time for this walk from a box of size $n$ is controlled using the Varopolous-Carne inequality. This inequality implies that the exit time is at least of order $d^2$, where $d$ is the chemical (instrinsic) distance to the boundary of the box of size $n$ through the IPC. In Lemma A.4.1, we outline an argument of A. Pisztora that proves that $d$ grows superlinearly with $n$. All of these estimates are tight enough to apply Borel-Cantelli and close the proof of subdiffusivity.

### A.4.1 Random walk on the backbone

The simple random walk started at $\mathbf{0}$ on the IPC is the Markov chain $\{X(k)\}_{k\geq 0}$ with the set of sites in the IPC as its state space, such that $X(0) = \mathbf{0}$, and with transition probabilities given by

$$\mathbf{P}^{\omega}(X(k+1) = y \mid X(k) = x) = \frac{\mathbf{1}[(x,y) \in E(\text{IPC})]}{\deg(x, \text{IPC})}.$$

The random variable $\deg(x, \mathrm{IPC})$ denotes the number of sites $y$ such that the edge $(x, y)$ belongs to the IPC.

Below, it will be convenient to work with a modification of $X$ that is reversible on $\Lambda(n)$. Thus, we let $\{X^n(k)\}_{k\geq 0}$ be the Markov chain started at the origin and defined by the transition probabilities

$$\mathbf{P}^\omega(X^n(k+1) = y \mid X^n(k) = x) = \frac{\mathbf{1}[(x, y) \in E(\Lambda(n))]}{\deg(x, \Lambda(n))}.$$

Note that the distribution of $X^n(k)$ coincides with that of $X(k)$ for $k \leq \tau^*(n)$, where

$$\tau^*(n) = \inf\{k \geq 0 : |X^n(k)|_\infty = n\}.$$

Moreover, the distribution of $\tau^*(n)$ is equal to the distribution of the exit time $\tau(n) = \inf\{k \geq 0 : |X(k)|_\infty = n\}$ defined in terms of the "full" random walk $X$ on $\Lambda(n)$. Thus, it will suffice to obtain bounds on $\tau^*(n)$.

The "backbone" $B(n)$ of $\Lambda(n)$ is the set of sites in $\Lambda(n)$ connected in the invasion cluster to $\mathbf{0}$ and to $\partial S(n)$ by two disjoint paths. A simple argument (see [65, Lemma 3.13]) shows that whenever $X^n$ leaves the backbone, it must return at the site where it left before it reaches $\partial S(n)$. Thus the random walk $X^n$ on $\Lambda(n)$ induces a random walk $X^{n,B}$ on $B(n)$ which moves only when $X^n$ is in $B(n)$. That is, if we define

$$\begin{aligned} \sigma_0 &= 0 \\ \sigma_m &= \inf\{k > \sigma_{m-1} : X^n(k) \in B(n)\} \\ X^{n,B}(k) &\equiv X^n(\sigma_k), \end{aligned}$$

then $X^{n,B}$ is a random walk on the backbone $B(n)$, with transition probabilities given by

$$\mathbf{P}^\omega(X^{n,B}(k+1) = y \mid X^{n,B}(k) = x) = \begin{cases} \frac{\mathbf{1}[x,y \in B(n),(x,y)\in E(\mathbb{Z}^2))]}{\deg(x,\Lambda(n))}, & y \neq x \\ \frac{\deg(x,\Lambda(n)) - \deg(x,B(n))}{\deg(x,\Lambda(n))}, & x = y. \end{cases} \tag{A.6}$$

Here, $\deg(x, B(n))$ is defined as the number of edges $(x, y)$ in $\Lambda(n)$ such that $x, y \in B(n)$.

### A.4.2 Estimate on the speed of the walk

Irrespective of the geometry of $B(n)$, $X^{n,B}$ must travel at least $n$ steps in $B(n)$ to reach $\partial S(n)$, because the distance between any two points in $B(n) \subset \mathbb{Z}^2$ is no less than the corresponding chemical distance in $\mathbb{Z}^2$. This fact was used by Kesten to conclude that the time spent by the walker on the backbone is of order at least $n^2/\log n$ with high probability. The Carne-Varopoulos bound ([25], [90]; see also [74, Theorem 13.4]) allows us to obtain a better estimate by considering the chemical distance on $B(n)$. It implies that the reversible Markov chain $X^{n,B}$ has at most diffusive speed in the intrinsic metric of the backbone. If $\mu$ is the stationary measure for the walk $X^{n,B}$ ($\mu$ depends on $\omega$), then

$$\mathbf{P}^\omega(X^{n,B}(k) = y \mid X^{n,B}(0) = \mathbf{0}) \leq 2\sqrt{\mu(y)/\mu(0)}\,\exp(-\operatorname{dist}_{B(n)}(\mathbf{0}, y)^2/(2k)).$$

The right side of this expression refers to the chemical distance in the backbone $B(n)$. The ratio appearing on the right can be bounded independently of the realization $\omega$ of the invasion percolation, since the stationary measure $\mu$ satisfies

$$1/4 \leq \frac{\mu(x)}{\mu(y)} \leq 4,$$

for any $x, y \in B(n)$. Since $B(n) \subset \mathbb{Z}^2$, we have the inequality of graph distances:

$$\mathrm{dist}_{B(n)} \geq \mathrm{dist}_{\mathbb{Z}^2} = \mathrm{dist}_1 .$$

Summing this bound over $\lambda\sqrt{k} \leq |y|_\infty \leq k$, we find

$$\mathbf{P}^\omega(\mathrm{dist}_{B(n)}(\mathbf{0}, X^{n,B}(k)) \geq \lambda\sqrt{k}) \lesssim k^2 \exp(-\lambda^2/\mathbf{C}_4).$$

Suppose we restrict our attention to realizations $\omega$ of the environment such that the chemical distance in $B(n)$ satisfies

$$\mathrm{dist}_{B(n)}(\mathbf{0}, \partial S(k)) \geq \mathbf{C}_5 k^s, \quad k \geq n_0(\omega)$$

for some $n_0(\omega)$ and some deterministic constants $s > 1$, $\mathbf{C}_5 > 0$. For such $\omega$, $\lambda \geq 1$ and $n$ sufficiently large, we have:

$$\begin{aligned} \mathbf{P}^\omega(|X^{n,B}(k)|_\infty \geq \lambda k^{1/(2s)}) &\leq \mathbf{P}^\omega(\mathrm{dist}_{B(n)}(X^{n,B}(k), \mathbf{0}) \geq \mathbf{C}_5 \lambda k^{1/2}) \\ &\lesssim k^2 \exp(-\lambda^2/\mathbf{C}_6). \end{aligned} \tag{A.7}$$

### A.4.3 Chemical distance in the IPC

It follows from work of Aizenman and Burchard [1] that the chemical distance inside a large box in independent bond percolation with parameter $p_c = 1/2$ is bounded below by a power $s > 1$ of the Euclidean distance in $\mathbb{Z}^2$ with high probability. Pisztora [82] showed how to extend this result to $p > p_c$ suitably close to $1/2$, and to the invasion percolation cluster. We reproduce the argument leading to his result, in a form that suits our needs, in the lemma below. Theorem A.2.4 follows from these results and the considerations above.

**Lemma A.4.1** ([82], Theorem 1.3)**.** *There exist* $\mathbf{C}_7$ *and* $s > 1$ *such that*

$$\mathbb{P}_{\text{IPC}}\left(\text{dist}_{\Lambda(n)}(\mathbf{0}, \partial S(n)) \le \mathbf{C}_7 n^s\right) \lesssim n^{-2}. \tag{A.8}$$

*Proof.* The models considered in [1] are defined by families $\{\mathbf{P}_\ell\}_{\ell>0}$ of probability measures on collections of curves in a compact region $\mathcal{R}$. For each $\ell$, $\mathbf{P}_\ell$ is supported on unions of polygonal curves with step size $\ell$. The realizations in the support of $\mathbf{P}_\ell$ are denoted by $\mathcal{F}_\ell$.

A truncated version of capacity is used to obtain lower bounds on the minimal number $N(A, \ell)$ of sets of diameter $\ell$ required to cover a given set $A \subset \mathcal{R}$:

$$N(A, \ell) \ge \text{cap}_{s,\ell}(A) \cdot \ell^{-s} \tag{A.9}$$

where

$$\text{cap}_{s,\ell}^{-1}(A) = \inf_{\mu(A)=1} \iint \frac{1}{\max(|x-y|^s, \ell^{-s})}\ \mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

The infimum is over Borel probability measures supported on $A$.

Under the assumption "Hypothesis H2," the authors of [1] obtain uniform bounds for $\text{cap}_{s,\ell}$: if there exist some $K, \sigma > 0$, and $0 < \rho < 1$ such that for every $k$ and collection of $k$ rectangles $A_1, \ldots, A_k$ of lengths $l_1, \ldots, l_k \ge \ell$ and cross-section[1] $\sigma$, and satisfying

$$\text{dist}(A_j, \cup_{i \ne j} A_i) \ge \text{diam}\, A_j$$

for all $j$, we have

$$\mathbf{P}_\ell(\text{all } A_i \text{ are traversed by segments of a curve in } \mathcal{F}_\ell) \le K\rho^k,$$

[1]The cross-section of a rectangle is the ratio of its short side to its long side.

then the capacity $\text{cap}_{s,\ell}$ of macroscopic curves is bounded below for some $s > 1$ [1, Theorem 1.3]: all curves $\mathcal{C}$ in $\mathcal{F}_\ell$ with $\text{diam}(\mathcal{C}) \geq 1/10$ satisfy

$$\text{cap}_{s,\ell}(\mathcal{C}) \geq C(s, \omega, \ell).$$

$C(s, \omega, \ell)$ is a random variable which is *stochastically bounded below* in the sense that

$$\mathbf{P}_\ell(C(s, \omega, \ell) \leq u) \to 0 \tag{A.10}$$

uniformly in $\ell$ as $u \to 0$.

We will apply the results in [1], with $\ell = n^{-1}$ to bond percolation on the rescaled lattice

$$\mathcal{R}_n = (1/n)\mathbb{Z}^2 \cap [-1, 1]^2.$$

For $p \in [0, 1]$, let $\mathbb{P}_p^n$ denote the independent bond percolation measure with parameter $p$ on the edges of $\mathcal{R}_n$. $\mathbb{P}_p^n$ induces a probability measure on configurations $\mathcal{F}_{1/n}$ of curves in $\mathcal{R} = [-1, 1]^2$: the percolation configuration is a union of connected paths of $p$-open edges, each edge being identified with a line segment of length $1/n$.

In the case of independent percolation, Hypothesis H2 reduces to the existence of a cross-section $\sigma$ and $\rho < 1$ such that the probability that there exists an open-crossing of a rectangle of cross-section $\sigma$ is less than $\rho$. By the Russo-Seymour-Welsh estimates, Hypothesis H2 is satisfied for $\{\mathbb{P}_{p_c}^n\}_{n \geq 1}$.

The lower bound (A.9) gives an estimate for the chemical distance in $\mathcal{F}_{1/n}$ between any two sets in $[-1, 1]^2$. Any $p_c$-open path in $\mathcal{R}_n$ connecting subsets $A$ and $B$ of $\mathcal{R}_n$ at Euclidean distance

$$\text{dist}(A, B) \geq 1/10$$

contains at least $C(s, \omega, 1/n) \cdot n^s$ bonds. Denote by $\text{dist}_{\mathcal{F}_{1/n}}(A, B)$ the (random) number of bonds in the shortest $p_c$-open path connecting $A$ and $B$ in $\mathcal{R}_n$. By (A.10), given any $\epsilon > 0$, we can choose $C(\epsilon)$ such that for all $n$,

$$\mathbb{P}^n_{p_c}(\mathrm{dist}_{\mathcal{F}_{1/n}}(A, B) \leq C(\epsilon) \cdot n^s) \leq \epsilon.$$

The scaling

$$x \mapsto nx$$

defines a measure-preserving bijection between $(E(\mathcal{R}_n), \mathbb{P}^n_{p_c})$ and $(E(S(n)), \mathbb{P}_{p_c})$. It follows that for each $\epsilon > 0$, there exists a constant $C(\epsilon)$ such that for all subsets $A$, $B$ of $S(n)$ at Euclidean distance $n/10$ from each other,

$$\mathbb{P}_{p_c}\left(\begin{array}{c}\text{there exists an open path connecting } A \text{ and } B\\ \text{in } S(n) \text{ with no more than } C(\epsilon)n^s \text{ bonds}\end{array}\right) \leq \epsilon.$$

Note that if $B$ cuts $A$ from $S(n)^c$ in $\mathbb{Z}^2$, the restriction that the path be contained in $S(n)$ is superfluous. This point will be relevant below.

The observation in [82] is that the Aizenman-Burchard bounds remain valid for $p > p_c$ as long as $n$ is smaller than the correlation length $L(p)$. The estimate used to obtain (A.10) depends only on $\sigma$ and $\rho$ [1, p. 446]. It follows from the definition of $L(p)$ and the Russo-Seymour-Welsh estimates that there exists $\rho < 1$ such that for rectangles of cross-section ratio 1/3, say, with long side $n \leq 3L(p)$,

$$\mathbb{P}_p(\exists \text{ an open crossing of } [0, n] \times [0, n/3]) \leq \rho.$$

Thus (A.10) remains true uniformly for $\ell^{-1} \leq 3L(p)$. Repeating the argument above, we see that we can choose $C(\epsilon)$ independent of $p \in (p_c, 1)$ to make the probability

$$\mathbb{P}_p(\mathrm{dist}(\partial S(L(p)), \partial S(3L(p)) \leq C(\epsilon)L(p)^s)$$

smaller than an arbitrary $\epsilon > 0$. The distance refers to the chemical distance in the union of all percolation clusters in the box $S(3L(p))$. Since $L(p) \to \infty$ as $p \downarrow p_c$, for any fixed $\epsilon$, $L(p)$ is much greater than $C(\epsilon)$, and so the estimate on the distance is not vacuous. More precisely, we find

$$\limsup_{p \downarrow p_c} \mathbb{P}_p(\mathrm{dist}(\partial S(L(p)), \partial S(3L(p)) \leq C(\epsilon) L(p)^s) \leq \epsilon. \tag{A.11}$$

A block argument with blocks of size $3L(p)$ converts the initial estimate (A.11) into an exponential bound for the macroscopic chemical distance in near-critical percolation (see the proof of [82, Theorem 1.3, pp. 12-14]). There exist constants $\mathbf{C}_8$, $\mathbf{C}_9$ such that if $p$ is sufficiently close to $p_c$:

$$\mathbb{P}_p(\mathrm{dist}(\mathbf{0}, \partial S(n)) \leq \mathbf{C}_8 n (L(p))^{s-1}) \lesssim \exp\left(-\mathbf{C}_9 \frac{n}{L(p)}\right). \tag{A.12}$$

With this in hand, (A.8) follows from the construction described in Section A.3. We outline the argument. The occurrence of the event

$$\tilde{H}_n(1) = \left\{ \begin{array}{c} \exists\ p_{(n/4)-1}(1)\text{-open circuit } \mathcal{C} \text{ around } \mathbf{0} \text{ in } S(n/2) \setminus S(n/4) \\ \text{and } \mathcal{C} \text{ is connected to } \infty \text{ by a } p_{(n/4)-1}(1)\text{-open path} \end{array} \right\}$$

implies that all edges of the IPC in $\Lambda(n) \setminus S(n/2)$ are $p_{(n/4)-1}(1)$-open.

$$\begin{aligned} &\mathbb{P}_{\mathrm{IPC}}(\mathrm{dist}_{\Lambda(n)}(\mathbf{0}, \partial S(n)) \leq (\mathbf{C}_8/5) \cdot n (L(p_{(n/4)-1}(1)))^{s-1}) && \text{(A.13)} \\ &\leq \mathbb{P}\left( \begin{array}{c} \exists\ x \in \partial S(3n/4) : \mathrm{dist}_{\Lambda(n)}(x, \partial S(x, n/4-1)) \\ \leq (\mathbf{C}_8/5) \cdot n (L(p_n(1)))^{s-1}; \tilde{H}_n(1) \end{array} \right) + \mathbb{P}(\tilde{H}_n(1)^c) \\ &\lesssim n \mathbb{P}_{p_{(n/4)-1}(1)}(\mathrm{dist}(\mathbf{0}, \partial S(n/4-1)) \leq (\mathbf{C}_8/5) \cdot n (L(p_{(n/4)-1}(1)))^{s-1}) + n^{-M\mathbf{C}_{10}} \\ &\lesssim n \exp\left(-\mathbf{C}_{11} \frac{n}{L(p_{(n/4)-1}(1))}\right) + n^{-M\mathbf{C}_{10}}. \end{aligned}$$

The final inequality follows from (A.12) and (A.5). Recalling (A.4):

$$\frac{n}{L(p_n(1))} \geq M \log n,$$

and choosing $M$ suitably large in the definition of $p_n(1)$, we find:

$$n \exp\left(-\mathbf{C}_{11} \frac{n}{L(p_{(n/4)-1}(1))}\right) + n^{-M\mathbf{C}_{10}} \lesssim n^{-2}.$$

By slightly lowering $s$ to absorb the logarithm, the probability on the left of (A.13) can be made to match the form of the left side of (A.8). □

**Remark A.4.2.** *The final part of the proof of Lemma A.4.1 shows that for any $0 < R_1 < R_2 < R_3$ and any $k > 0$, one can find constants (depending on $R_i$ and $k$) such that*

$$\mathbb{P}_{\mathrm{IPC}}(\mathrm{dist}_{\Lambda(R_3 n)}(\partial S(R_2 n), \partial S(R_1 n) \cup \partial S(R_3 n)) \leq C n^s) \lesssim n^{-k}.$$

*Here $s > 1$ is the constant appearing in (A.8). Such a statement will be used in Section A.5 below.*

*Proof of Theorem A.2.4.* For $s > 1$, let $L_n$ be the event

$$\{\mathrm{dist}_{\Lambda(n)}(\mathbf{0}, \partial S(n)) \leq \mathbf{C}_7 n^s\}.$$

By (A.8) in the previous lemma, we have

$$\sum_{n \geq 1} \mathbb{P}_{\mathrm{IPC}}(L_n) < \infty \tag{A.14}$$

for some $s > 1$. Applying the Borel-Cantelli lemma and choosing

$$\omega \in \{L_n \text{ occurs infinitely often}\}^c,$$

we can use (A.7) with $\lambda = 4^s \cdot \mathbf{C}_6^{s/2} (\log n)^{s/2}$ and $\mathbf{C}_5 = \mathbf{C}_7$; for some $N(\omega)$ we have:

$$\sum_{n \geq N(\omega)} \mathbf{P}^\omega(\tau^*(n) \leq n^{2s}/\lambda^2) \leq \sum_{n \geq N(\omega)} \sum_{k \leq n^{2s}/\lambda^2} \mathbf{P}^\omega(|X^{n,B}(k)|_\infty \geq n) < \infty.$$

A second application of the Borel-Cantelli lemma leads to Theorem A.2.4. □

Note that for the argument above it was not necessary to consider $X^{n,B}$. However, the decomposition of the IPC into a backbone and "dangling ends" will be central in the derivation of Theorem A.2.1 below. The proof of Theorem A.2.4 shows that $X^{n,B}$ alone already contributes at least $n^{2+\epsilon}$ steps to $\tau(n)$.

## A.5 Kesten's comparison argument

Our modification of Kesten's argument compares the volume of sites in the invasion percolation cluster (IPC) to the volume of sites on the backbone to conclude that the walk must be subdiffusive.

### A.5.1 Preliminaries and a key lemma

We assume for simplicity of notation that $n = 3m$, $m \in \mathbb{Z}^+$. We introduce two stopping times:

$$\tau(2m) = \inf\{k \geq 0 : X(k) \in \partial S(2m)\}$$
$$\sigma^+(m) = \inf\{k \geq \tau(2m) : X(k) \in \partial S(m) \cup \partial S(n)\}.$$

By definition, we clearly have:

$$\tau(n) = \tau(3m) \geq \sigma^+(m) - \tau(2m).$$

Hence, it will suffice to obtain a lower bound on the right side of the previous expression.

$$Y(k) = X(\tau(2m) + k),\ k \geq 0$$

is a simple random walk on IPC; now define $Y^n$ to be the simple random walk on (the possibly disconnected)

$$\Gamma(n) = \mathrm{IPC} \cap (S(n) \setminus S(m))$$

with initial point $Y(0)$. Letting $\sigma^*(n)$ be the hitting time of $\partial S(m) \cup \partial S(n)$ by the walk $Y^n$, we note that $\sigma^*(n)$ has the same distribution as $\sigma^+(m) - \tau(n)$.

A key tool in Kesten's argument is the following result from [65], expressing the spatial "smoothness" of the local times for a reversible Markov chain.

**Lemma A.5.1** ([65], Lemma 3.18)**.** *Let $x$, $y$ be two sites in $\Gamma(n)$, and let*

$$L(x, k) = \sharp\{l : 0 \leq l \leq k,\ Y^n(l) = x\}$$

*be the local time at a site $x$ of the walk $Y^n$. Then, for some $L_0 > 0$ and any $\lambda > 1$:*

$$\mathbf{P}^\omega\left(\exists k,\ L(y,k) \geq \lambda\, \mathrm{dist}_{\Gamma(n)}(x,y) \text{ and } L(x,k) \leq \frac{1}{2}\frac{\deg_{\Gamma(n)}(x)}{\deg_{\Gamma(n)}(y)} L(y,k)\right) \lesssim \mathrm{dist}_{\Gamma(n)}(x,y)\exp(-\lambda/L_0). \quad \text{(A.15)}$$

In [65], Lemma A.5.1 is stated in terms of the intrinsic distance on the incipient infinite cluster. Replacing $\|x - y\|_{m,w}$, $d(x)$ and $d(y)$ in the proof of Lemma 3.18 in [65] by $\mathrm{dist}_{\Gamma(n)}$, $\deg_{\Gamma(n)}(x)$ and $\deg_{\Gamma(n)}(y)$, respectively, we obtain Lemma A.5.1 above.

We also modify our definition of the backbone. $\tilde{B}(n)$ is defined to be the set of sites in $\Gamma(n)$ connected by two disjoint paths (in $\Gamma(n)$) to $\partial S(n)$ and $\partial S(m)$. $Y^{n,\tilde{B}}$ is the induced walk on $\tilde{B}$, defined analogously to $X^{n,B}$ in Section A.4. We let $b(n)$ be the number of steps $Y^{n,\tilde{B}}$ takes between 0 and $\sigma^*(n)$; $b(n)$ is the time spent by $Y^n$ on $\tilde{B}(n)$.

### A.5.2 Sketch of the proof of Theorem A.2.1

Kesten's comparison argument will be applied to $Y^n$. The idea is to consider a "thickening" of the backbone of size $q$. By Lemma A.5.1, if a box $S(v,q)$ of size $q$ contains a site $x \in \tilde{B}(n)$ with $L(x,\sigma^*(n)) \gg q^2 L_0$, the random walk visits all accessible sites of $\Gamma(n)$ inside $S(v,q)$ at least $CL(x,\sigma^*(n))$ times, with high probability. If it is traversed by a portion of the random walk, the box $S(v,q)$ typically contains $q^2\pi(q)$ sites of $\Gamma(n)$, and at most $q^2\rho(q)$ sites of $\tilde{B}(n)$. Thus the time spent by $Y^n$ in $S(q,v)$ up to $\sigma^*(n)$ is larger than the time $Y^{n,\tilde{B}}$ spends there by a factor of at least $\pi(q)/\rho(q)$. By choosing $q$ appropriately, the set of sites $y$ on the backbone which do not satisfy the lower bound of order $q^2 L_0$ on $L(y,\sigma^*(n))$ will make a contribution bounded by a fraction of the total time spent on the backbone.

### A.5.3 Proof of Theorem A.2.1

To realize the strategy just described, we tile $S(n) \setminus S(m)$ by squares of size

$$q = Q \cdot \frac{n^{\eta_2/2}}{(\log n)^{3/2}} \tag{A.16}$$

for a constant $Q$ to be determined. Here $\eta_2$ is the exponent appearing in (A.3). We note for future reference that

$$\eta_2 \leq 1 \text{ so that } q = o(\sqrt{n}) \ . \tag{A.17}$$

This bound on $\eta_2$ can be proved using the method of [17, Cor. 3.15]. For the details, the reader can see a standard sketch of a similar inequality (for crossings of an annulus) under equation (A.31) later in the paper.

For $\mathbf{j} = (j_1, j_2) \in \mathbb{Z}^2$, define

$$D(\mathbf{j}, q) = [qj_1, q(j_1+1)) \times [qj_2, q(j_2+1)),$$
$$F(\mathbf{j}, q) = [q(j_1-1), q(j_1+2)] \times [q(j_2-1), q(j_2+2)].$$

Given a realization $\omega$ of IPC and a realization of the walk, we follow the path of $Y^n$ until $\sigma^*(n)$ by introducing two sequences $\{l_i\}$ and $\{\mathbf{j}_i\}$, first setting $l_0 = 0$ and $\mathbf{j}_0$ to be the index such that $Y^n(l_0) \in D(\mathbf{j}_0, q)$ and then defining $l_i$ by

$$l_{i+1} = \min\{l > l_i, Y(l) \notin F(\mathbf{j}_i, q)\}$$
$$Y^n(l_{i+1}) \in D(\mathbf{j}_{i+1}, q).$$

$Y^n$ may reach $\partial S(m) \cup \partial S(n)$ before leaving $F(\mathbf{j}_i, q)$, in which case $l_i$ ends the sequence. We let $\mathcal{C}(i)$ denote the component of $F(\mathbf{j}_i, q) \cap \Gamma(n)$ containing $Y^n(l_i)$. $Y^n$ may return several times to the same square, so $\mathcal{C}(j)$ may be equal to $\mathcal{C}(i)$ for $i \neq j$. Enumerating the $\mathcal{C}(i)$ without repetition as $\iota_0, \iota_1, \ldots, \iota_\lambda$, with $\mathcal{C}(\iota_\lambda)$ the component of $F(\mathbf{j}_i, q)$ where $i$ is such that

$$Y^n(\sigma^*(n)) \in F(\mathbf{j}_i, q),$$

we define

$$\Lambda(\iota) = \sum_{x \in \mathcal{C}(\iota)} L(x, \sigma^*(n))$$
$$\Theta(\iota) = \sum_{x \in \tilde{B}(n) \cap \mathcal{C}(\iota)} L(x, \sigma^*(n)).$$

Since any $x$ belongs to at most 16 different $F$ squares, we have:

$$\frac{1+\sigma^*(n)}{1+b(n)} = \frac{\sum_x L(x,\sigma^*(n))}{\sum_{x\in\tilde{B}(n)} L(x,\sigma^*(n))} \geq \frac{1}{16}\frac{\sum_\iota \Lambda(\iota)}{\sum_\iota \Theta(\iota)}.$$

We now state volume estimates analogous to those obtained in [65] for the incipient infinite cluster; they will be derived in the next section. We will only be concerned with those indices in the set

$$\mathcal{J} = \{\mathbf{j} \in \mathbb{Z}^2 : F(\mathbf{j},q) \cap (S(n)\setminus S(m)^\circ) \neq \emptyset\}.$$

The first estimate is for the number of backbone sites in any $F$ square; for any $\mathbf{j}\in\mathcal{J}$, we have:

$$\mathbb{P}_{\mathrm{IPC}}\left(\sharp(\tilde{B}(n)\cap F(\mathbf{j},q)) \geq \frac{c}{\mathbf{C}_{12}} q^2\rho(q)\log q\right) \lesssim q^{-c}. \tag{A.18}$$

The second provides, with high probability, a lower bound for the number of sites of the IPC in a box $F(\mathbf{j},q)$, $\mathbf{j}\in\mathcal{J}$, given that there is a crossing of $F(\mathbf{j},q)\setminus D(\mathbf{j},q)$:

$$\mathbb{P}_{\mathrm{IPC}}\begin{pmatrix}\text{there exists a crossing } r\subseteq\Gamma(n) \text{ of } F(\mathbf{j},q)\setminus D(\mathbf{j},q)\\ \text{with } \sharp\{x\in F(\mathbf{j},q) : x \text{ connected to } r \text{ in}\\ \Gamma(n)\cap F(\mathbf{j},q)\} \leq q^2\pi(q)/(\log q)^4\end{pmatrix} \lesssim_c q^{-c} \tag{A.19}$$

Here $c$ is arbitrary but the implicit constant depends on the choice of $c$.

### A.5.4 The events $E_i(n)$ and $W_i(n)$

We now define the events $E_i(n)$, $1\leq i\leq 4$ and $W_i$, $1\leq i\leq 3$. The ratio $\sum\Lambda(\iota)/\sum\Theta(\iota)$ will be bounded below by $\pi(q)/\rho(q)$ on the event $(\cap_i E_i)\cap(\cap_i W_i)$.

1.

$$E_1(n) = \left\{\omega : \mathrm{dist}_{\Gamma(n)}(\partial S(2m), \partial S(n)\cup\partial S(m)) \geq \mathbf{C}_{13} n^s\right\}.$$

2.

$$E_2(n) = \left\{ \begin{matrix} \omega : \sharp(\tilde{B}(n) \cap F(\mathbf{j}, q)) \leq \mathbf{C}_{14} q^2 \rho(q) \log q \\ \text{for all } \mathbf{j} \in \mathcal{J} \end{matrix} \right\}.$$

3.

$$E_3(n) = \left\{ \begin{matrix} \omega : \sharp\{x \in F(\mathbf{j}, q) : x \text{ connected to } r \text{ in} \\ \Gamma(n) \cap F(\mathbf{j}, q)\} \geq q^2 \pi(q)/(\log q)^4 \text{ for all} \\ \mathbf{j} \in \mathcal{J} \text{ and any crossing } r \subseteq \Gamma(n) \text{ of } F(\mathbf{j}, q) \setminus D(\mathbf{j}, q) \end{matrix} \right\}.$$

4.

$$E_4(n) = \left\{ \omega : \sharp\tilde{B}(n) \leq \frac{4}{\mathbf{C}_{15}} n^2 \rho(n) (\log n)^2 \right\}.$$

5.

$$W_1(n) = \{b(n) \geq n^{2s'} / \log n\}.$$

6.

$$W_2(n) = \left\{ \begin{matrix} 1/8 \leq \dfrac{L(x, \sigma^*(n))}{L(y, \sigma^*(n))} \leq 8 \text{ for each pair } x, y \in S(n) \setminus S^\circ(m) \\ \text{such that } x, y \text{ belong to the union of two clusters} \\ \mathcal{C}(i), \mathcal{C}(i+1) \text{ traversed consecutively by } Y^n \text{ and} \\ L(x, \sigma^*(n)) \geq 320 L_0 q^2 \log n \end{matrix} \right\}.$$

7.

$$W_3(n) = \left\{ \begin{matrix} L(x, \sigma^*(n)) \leq 2560 L_0 q^2 \log n \\ \text{for any } x \text{ in a cluster } \mathcal{C}(i) \text{ such that} \\ L(y, \sigma^*(n)) \leq 320 L_0 q^2 \log n \\ \text{for some } y \in \mathcal{C}(i) \end{matrix} \right\}.$$

By the remark following the proof of Lemma A.4.1, there exists $\mathbf{C}_{13}$ such that

$$\mathbb{P}_{\mathrm{IPC}}(E_1(n)^c) \lesssim n^{-2},$$

with the same constant $s$ as in (A.8). For any $1 < s' < s$ and $\omega \in E_1(n)$, we use the Carne-Varopoulos estimate (A.7), applied to the symmetric chain $Y^{n,\tilde{B}}$, to show as in the proof of Theorem A.2.4:

$$\mathbf{P}^{\omega}(W_1(n)^c) \lesssim n^{-2}, \quad \omega \in E_1(n), \tag{A.20}$$

giving the bound (A.2.4).

Recall the definition of $q$ in (A.16). We have

$$q^{-1} = o(n^{-\eta_2/4}).$$

Noting that there are, up to a constant, at most $n^2/q^2$ indices $\mathbf{j}$ in $\mathcal{J}$, and choosing $\mathbf{C}_{14} \geq 16/(\mathbf{C}_{12}\eta_2)$ in (A.18), and accordingly in the definition of $E_2(n)$, we find

$$\mathbb{P}_{\mathrm{IPC}}(E_2(n)^c) \lesssim n^{-2}.$$

By the estimate (A.25) in Section A.6:

$$\mathbb{P}_{\mathrm{IPC}}(E_4(n)^c) \lesssim n^{-2}.$$

By (A.19) (for some $c$ large enough), we have

$$\mathbb{P}_{\mathrm{IPC}}(E_3(n)^c) \lesssim n^{-2}.$$

Finally, we have

$$\mathbf{P}^\omega(W_2(n)^c), \mathbf{P}^\omega(W_3(n)^c) \lesssim n^{-2}$$

uniformly in $\omega$. Indeed, suppose $x$ and $y$ are two sites as in the description of $W_2(n)$, then, for $n$ sufficiently large,

$$\mathrm{dist}_{\Gamma(n)}(x,y) \leq (6q+1)^2 \leq 40q^2$$

for any $\omega$. Using

$$1 \leq \deg_{\Gamma(n)}(x), \deg_{\Gamma(n)}(y) \leq 4$$

for any $x, y$ in the IPC, we find that on $W_2(n)^c$, for some pair $x, y$ either

$$L(x,\sigma^*(n)) \leq \frac{1}{8}L(y,\sigma^*(n)) \leq \frac{1}{2}\frac{\deg_{\Gamma(n)}(y)}{\deg_{\Gamma(n)}(x)}L(y,\sigma^*(n))$$

and

$$L(y,\sigma^*(n)) \geq 8 \cdot 320 L_0 q^2 \log n \geq 64 L_0 \log n \,\mathrm{dist}_{\Gamma(n)}(x,y)$$

or

$$L(y,\sigma^*(n)) \leq \frac{1}{2}\frac{\deg_{\Gamma(n)}(y)}{\deg_{\Gamma(n)}(x)}L(x,\sigma^*(n))$$

and

$$L(x,\sigma^*(n))) \geq 8L(y,\sigma^*(n)) \geq 8L_0 \log n \,\mathrm{dist}_{\Gamma(n)}(x,y).$$

The first case is contained in the event appearing in (A.15). So is the second case, after reversing the roles of $x$ and $y$ in that event. Using

$$\mathrm{dist}_{\Gamma(n)}(x,y) \lesssim n^2,$$

applying Lemma A.5.1, and taking the union over all pairs, $x, y \in S(3m) \setminus S(m)$, we find that, whatever $\omega$ in the support of $\mathbb{P}_{\mathrm{IPC}}$:

$$\mathbf{P}^\omega(W_2(n)^c) \lesssim \sum_{x,y \in S(3m)\setminus S(m)} \sharp(S(3m) \setminus S(m)) \cdot \exp(-8 \log n) \lesssim n^6 n^{-8}.$$

A similar argument applies to $W_3(n)$.

### A.5.5 End of the proof

Applying the Borel-Cantelli lemma to

$$E_1(n)^c \cup E_2(n)^c \cup E_3(n)^c \cup E_4(n)^c,$$

we find that for $\mathbb{P}_{\mathrm{IPC}}$-almost every $\omega$, there exists $N(\omega)$ such that $\cap_i E_i(n)$ holds when $n \geq N(\omega)$. For any such $\omega$, a further application of the Borel-Cantelli lemma shows that, $\mathbf{P}^\omega$-almost surely, $\cap_i W_i(n)$ holds for $n$ large enough.

It remains to show that whenever all the events above hold, we have the subdiffusive bound of Theorem A.2.1. First, on $E_4(n) \cap W_3(n)$, if we denote by $\sum^*$ the sum over indices $\iota$ such that $\mathcal{C}(\iota)$ contains a site $x_\iota \in \tilde{B}(n)$ with

$$L(x_\iota, \sigma^*(n)) \geq 320 L_0 q^2 \log n,$$

then, assuming $W_1(n)$ also occurs, adjusting the constant $Q$ in the definition of $q$ (see (A.16)):

$$\begin{aligned}\Big(\sum_\iota - \sum_\iota{}^*\Big)\ \Theta(\iota) &\leq 16 \cdot 320 L_0 q^2 \log n \cdot \sharp\tilde{B}(n) \\ &\leq \frac{1}{2} b(n).\end{aligned}$$

It follows that

$$\frac{1+\sigma^*(n)}{1+b(n)} \geq \frac{1}{32}\frac{\sum^* \Lambda(\iota)}{\sum^* \Theta(\iota)}. \tag{A.21}$$

On $W_2(n)$, letting $y_\iota$ be the lexicographically earliest point of $\tilde{B}(n)$ in $\mathcal{C}(\iota)$, we have for those indices $\iota \leq \lambda$ occurring in $\sum^*$:

$$\Lambda(\iota) \geq \frac{1}{8}L(y_\iota, \sigma^*(n)) \cdot \sharp\mathcal{C}(\iota)$$
$$\Theta(\iota) \leq 8L(y_\iota, \sigma^*(n)) \cdot \sharp\tilde{B}(n) \cap \mathcal{C}(\iota).$$

For each $\iota \leq \lambda$, $F(\mathbf{j}_\iota, q) \setminus D(\mathbf{j}_\iota, q)$ contains an invaded crossing in $\mathcal{C}(\iota)$. Thus, on $E_2(n) \cap E_3(n)$, we can write:

$$\frac{\Lambda(\iota)}{\Theta(\iota)} \geq \frac{1}{8^2\mathbf{C}_{14}}\frac{\pi(q)}{\rho(q)}\frac{1}{(\log q)^5}, \quad \iota \leq \lambda$$

Bounding every $\Lambda(\iota)$ term below individually in (A.21) and using the BK inequality, we find:

$$\frac{1+\sigma^*(n)}{1+b(n)} \gtrsim \frac{\pi(q)}{\rho(q)}\frac{1}{(\log q)^6} \gtrsim \frac{1}{\pi(q)(\log q)^5}. \tag{A.22}$$

On $W_1(n)$, we have $b(n) \geq n^{2s'}/\log n$. Recalling the definition of $q$ from (A.16), we have the following bound for $\pi(q)$.

$$\pi(q) \lesssim q^{-\eta_1} \lesssim n^{-(1/2)\eta_1\eta_2}(\log n)^{3\eta_1/2}$$

Choosing $\epsilon_0 > 0$ such that $2 < 2+\epsilon_0 < 2s'$, we obtain:

$$\tau(n) \geq \sigma^*(n) \gtrsim \frac{n^{2s'}}{\log n} \cdot \frac{1}{\pi(q)(\log q)^5} \gg n^{2+\frac{1}{2}\eta_1\eta_2+\epsilon_0},$$

the desired result.

## A.6 Derivation of the volume estimates

In this section we prove the volume estimates (A.18) and (A.19).

### A.6.1 Estimates for the size of the backbone

We show that the following moment bounds hold for $\mathbf{j} \in \mathcal{J}$:

$$\mathbb{E}_{\mathrm{IPC}}(\sharp F(\mathbf{j}, q) \cap \tilde{B}(n))^k \leq k! \cdot (\mathbf{C}_{16} q^2 \rho(q))^k \tag{A.23}$$

$$\mathbb{E}_{\mathrm{IPC}}(\sharp \tilde{B}(n))^k \leq k! \cdot k^k \cdot (\mathbf{C}_{17} n^2 \rho(n))^k, \tag{A.24}$$

for $k = 1, 2...$ and constants $\mathbf{C}_{16}$, $\mathbf{C}_{17}$.

The estimate (A.23) implies the existence, for $\lambda > 0$ small enough, of the exponential moment:

$$\mathbb{E}_{\mathrm{IPC}} \exp(\lambda \sharp F(\mathbf{j}, q) \cap \tilde{B}(n)) < \infty.$$

Applying Chebyshev's inequality with $\lambda = 1/(2\mathbf{C}_{16} q^2 \rho(q))$ yields $(A.18)$.

From (A.24), we obtain the finiteness, for sufficiently small $\lambda$, of the stretched exponential moment:

$$\mathbb{E}_{\mathrm{IPC}} \exp(\lambda \sharp (\tilde{B}(n))^{1/2}) < \infty.$$

Using Chebyshev's inequality with $\lambda = 1/(2(e\mathbf{C}_{17} n^2 \rho(n))^{1/2})$, we obtain, for each $c > 0$:

$$\mathbb{P}_{\mathrm{IPC}} \left( \sharp \tilde{B}(n) \geq \frac{c}{\mathbf{C}_{15}} n^2 \rho(n) (\log n)^2 \right) \lesssim n^{-c^{1/2}}. \tag{A.25}$$

To derive (A.23) and (A.24), we follow the method introduced by Járai [56] to estimate the moments of the volume $|\Lambda|$ of the IPC in a box. We will instead apply this argument to the volume of a backbone, and then combine it with an inductive argument of Nguyen [79].

We begin with the first moment ($k = 1$) in (A.23). If $F(\mathbf{j}, q) \subset S(n) \setminus (S(m))^\circ$, the number of sites of $F(\mathbf{j}, q)$ with two disjoint connections in the IPC to $\partial F(\mathbf{j}, q)$ provides an upper bound for the volume of $F(\mathbf{j}, q) \cap \tilde{B}(n)$. Let $Z_q(\mathbf{j}, j)$ denote the set of sites in $F(\mathbf{j}, q)$ with two $p_{2m}(j)$-open connections to $\partial F(\mathbf{j}, q)$. Note that

$$\tilde{B}(n) \subset S(m)^c.$$

On $H_{2m}(j)$ (defined in (A.5)), every edge of the IPC in $S(m)^c$ is $p_{2m}(j)$-open, as noted at the end of Section A.3, and thus:

$$\begin{aligned} \mathbb{E}(\sharp(\tilde{B}(n) \cap F(\mathbf{j}, q))) \leq\ & \mathbb{E}(\sharp(\tilde{B}(n) \cap F(\mathbf{j}, q)); H_{2m}(1)^c) \\ & + \sum_{j=2}^{\log^* 2m} \mathbb{E}(\sharp Z_q(\mathbf{j}, j-1); H_{2m}(j-1) \cap H_{2m}(j)^c) \\ & + \mathbb{E}(\sharp Z_q(\mathbf{j}, \log^* 2m)). \end{aligned} \tag{A.26}$$

The first term is bounded up to a constant factor by:

$$(3q)^2 \cdot \mathbb{P}(H_{2m}(1)^c) \lesssim q^2 (2m)^{-\mathbf{C}_0 M}.$$

The terms of the sum are estimated using the Harris-FKG inequality:

$$\begin{aligned} \mathbb{E}(\sharp Z_q(\mathbf{j}, j-1); H_{2m}(j-1) \cap H_{2m}(j)^c) &\leq \mathbb{E}(\sharp Z_q(\mathbf{j}, j-1)) \cdot \mathbb{P}(H_{2m}(j)^c) \\ &\lesssim \mathbb{E}(\sharp Z_q(\mathbf{j}, j-1)) \exp(-M \mathbf{C}_0 \log^{(j)} 2m) \end{aligned} \tag{A.27}$$

By decomposing $F(\mathbf{j}, q)$ according to the distance $l$ to $\partial F(\mathbf{j}, q)$, we find:

$$\mathbb{E}\sharp(Z_q(\mathbf{j}, j-1)) \lesssim \sum_{l=1}^{3q/2} q\rho(p_{2m}(j-1), l) \tag{A.28}$$

$$\lesssim \sum_{l \leq \lfloor L(p_{2m}(j-1)) \rfloor} q\rho(p_{2m}(j-1), l) \tag{A.29}$$

$$+ \ q^2 \rho(p_{2m}(j-1), L(p_{2m}(j-1))) \cdot \mathbf{1}[\ 3q/2 > \lfloor L(p_{2m}(j)) \rfloor\ ].$$

By the same argument used for (7) in [64] (see Remark (37) there), the sum up to $L(p_{2m}(j-1))$ in (A.29) is bounded up to a constant by

$$q^2 \rho(p_{2m}(j-1), L(p_{2m}(j-1))).$$

The proof in [64] is carried out for $p = p_c$, but the implicit constants that appear are due to applications of RSW theory and thus are uniformly bounded in $p > p_c$. By comparability of the arm exponents below $L(p)$ [66] (see also [80, Theorem 26]), we have

$$\rho(p_{2m}(j-1), L(p_{2m}(j-1)) \lesssim \rho(p_c, L(p_{2m}(j-1))).$$

Thus, finally, in (A.28), we have (since $q \leq 2m$)

$$\mathbb{E}\sharp(Z_q(\mathbf{j}, j-1)) \lesssim q^2 \rho(p_c, 2m) M \log^{(j-1)} 2m \tag{A.30}$$

$$\lesssim q^2 \rho(p_c, q) M \log^{(j-1)} 2m,$$

where in the first step we have used the inequality

$$\frac{\rho(p_c, r)}{\rho(p_c, s)} \gtrsim \frac{s}{r} \tag{A.31}$$

for $r \leq s$. A similar inequality for $\pi(p_c, n)$ was used in [56], where the author indicates that it can be proved by the argument in [17, Corollary 3.15]. The proof of (A.31) follows the same general strategy, but does not use the van den Berg-Kesten inequality: let $k = 0, 1, 2, \ldots, \lceil r/s \rceil$, $v_k = \mathbf{0} \pm 2ks$ and consider the annuli $S(v_k, r) \setminus S^\circ(v_k, s)$. The inner squares of these $2\lceil r/s \rceil + 1$ annuli are adjacent. The event that there exists a $p_c$-open left-right crossing of $S(r)$ has probability bounded below uniformly in $r$, and implies that one of the annuli is crossed by two disjoint $p_c$-open paths. By quasi-multiplicativity, this probability is comparable to $\rho(r)/\rho(s)$, and (A.31) follows by a union bound.

Inserting (A.30) into (A.27) and then into (A.26), we find

$$\mathbb{E}(\sharp(\tilde{B}(n) \cap F(q, \mathbf{j}))) \lesssim q^2 \rho(p_c, q) \times \left( \frac{\exp(-M\mathbf{C}_0 \log 2m)}{\rho(p_c, q)} + M \sum_{j=2}^{\log^* 2m} (\log^{(j-1)} 2m)^{-M\mathbf{C}_0} \log^{(j-1)} 2m + M \right)$$

The final term corresponds to $\mathbb{E}(\sharp Z_q(\mathbf{j}, \log^* 2m))$, which is $O(q^2\rho(p_c, q))$ by (A.30). Using (A.31), we may choose $M$ large enough to make the first term $O(1)$. An important point made in [56] is that choosing $M$ possibly larger, we may bound the contribution from the sum in the parentheses by a constant. Indeed, we have

$$\sup_{n \geq 1} \sum_{j=1}^{\log^* n} \left( \log^{(j)} n \right)^{-1} \lesssim 1. \tag{A.32}$$

This establishes (A.23) for $k = 1$. To deal with the higher moments, we use the following general lemma:

**Lemma A.6.1.** *Let $p_c \leq p \leq 1$, $n \geq 1$, and $C_n(p)$ be the set of sites of $S(n)$ with two disjoint $p$-open connections to $\partial S(n)$. There exists a constant $\mathbf{C}_{18}$ independent of $n$ and $p$ such that, for any $k \geq 1$, the following inductive bound holds:*

$$\mathbb{E}(\sharp C_n(p))^{k+1} \leq \mathbf{C}_{18}(k+1) n^2 \rho(p, n) \mathbb{E}(\sharp C_n(p))^k.$$

*Proof.* The result is essentially due to Nguyen [79], who proved that for $p \geq p_c$ and $L \leq L(p)$,

$$\mathbb{E}(\sharp W_L)^{k+1} \leq \mathbf{C}_{19}(k+1)L^2\pi(p,L)\mathbb{E}(\sharp W_L)^k, \quad k \geq 1,$$

where $W_L$ is the set of sites in $S(L)$ connected to $\partial S(L)$ by a $p$-open path. $\mathbf{C}_{19}$ is a constant uniform in $k, L$ and $p$.

When $n \leq L(p)$, the proof in [79] is easily adapted to the variables $\sharp C_n(p)$. We define the event

$$A(x) = \{x \text{ has two disjoint open connections to } \partial S(n)\}.$$

The idea is to write

$$\begin{aligned}\mathbb{E}(\sharp C_n(p))^{k+1} &= \sum_{x_1,\ldots,x_{k+1}\in S(n)} \mathbb{P}_p(\cap_{i=1}^k A(x_i), A(x_{k+1})) \\ &= \sum_{l=1}^{n/2} \sum_{x_1,\ldots,x_k\in S(n)} \sum_{x_{k+1}\in R_l\cap S(n)} \mathbb{P}_p\left(\cap_{i=1}^k A(x_i), A(x_{k+1})\right), \end{aligned} \tag{A.33}$$

where we have set

$$R_l = R_l(x_1, \ldots x_k) = \{x : \mathrm{dist}_\infty(x, \{x_1, \ldots, x_k\} \cup \partial S(n)) = l\}.$$

Letting

$$\mathrm{Circ}_{k,l} = \left\{ \begin{array}{c} \text{there exists an open circuit} \\ \text{around } x_{k+1} \text{ in } S(l, x_{k+1}) \setminus S(l/2, x_{k+1}) \end{array} \right\},$$

we have $\mathbb{P}(\mathrm{Circ}_{k,l}) \geq \mathbf{C}_{20} > 0$ uniformly in $l$ and $n$ (even for $l > L(p)$). By the FKG inequality:

$$\begin{aligned}\mathbb{P}_p\left(\cap_{i=1}^k A(x_i),\ A(x_{k+1})\right) \leq & \frac{1}{\mathbf{C}_{20}}\mathbb{P}_p\left(\cap_{i=1}^k A(x_i), A(x_{k+1}), \mathrm{Circ}_{k,l}\right)\\ \leq & \frac{1}{\mathbf{C}_{20}}\mathbb{P}_p\begin{pmatrix}\cap_{i=1}^k \tilde{A}(x_i,l), x_{k+1} \text{ has two disjoint}\\ \text{connections to } \partial S(l/2, x_{k+1})\end{pmatrix}.\end{aligned}$$

$\tilde{A}(x_i, l)$ is the event that $x_i$ is connected to $\partial S(n)$ by two disjoint open paths outside of $S(l/2, x_{k+1})$. By independence, the last quantity on the right is bounded, up to a constant, by

$$\mathbb{P}_p\left(\cap_{i=1}^k A(x_i)\right)\rho(p, l/2).$$

For any $l$, we have:

$$\sharp R_l \lesssim (k+1)\cdot n.$$

Returning to (A.33), we find

$$\mathbb{E}(\sharp C_n(p))^{k+1} \lesssim (k+1)\cdot \mathbb{E}(\sharp C_n(p))^k \cdot n\sum_{l=1}^{n/2}\rho(p, l/2). \tag{A.34}$$

If $n \leq L(p)$, we have

$$\rho(p, l/2) \asymp \rho(p_c, l/2) \asymp \rho(p_c, l),$$

and the estimate (see the remark concerning (A.29) above)

$$\sum_{l=1}^{n/2}\rho(p_c, l) \lesssim n\rho(p_c, n) \lesssim n\rho(p, n)$$

leads to the inductive estimate claimed above.

If $n \geq L(p)$, we split the sum as we did in the treatment of the first moment of $\sharp Z_q$:

$$\begin{aligned}\sum_{l=1}^{n/2} \rho(p,l/2) &= \left( \sum_{l=1}^{\lfloor L(p) \rfloor} + \sum_{l=\lfloor L(p) \rfloor +1}^{n/2} \right) \rho(p,l/2) \\ &\lesssim \sum_{l=1}^{\lfloor L(p) \rfloor} \rho(p,l/2) + n\rho(p,L(p)/2) \\ &\lesssim n\rho(p,L(p)) + n\rho(p,L(p)/2) \\ &\lesssim n\rho(p,n).\end{aligned} \tag{A.35}$$

Here we have used that for $L \geq L(p)$, $\rho(p,L) \asymp \rho(p,L(p))$. This follows by a variant of the argument presented in [80, Section 7.4] . This establishes the lemma. □

Using Lemma A.6.1, induction and the fact that $q \leq 2m/(DM \log 2m)$ for large $m$, we obtain

$$\mathbb{E}(\sharp Z_q(\mathbf{j}, j))^k \leq k!(\mathbf{C}_{18} q^2 \rho(p_{2m}(j), q))^k \leq k!(\mathbf{C}_{18} q^2 \rho(p_c, q))^k. \tag{A.36}$$

Thus, arguing as for (A.26):

$$\begin{aligned}\mathbb{E}((\sharp(\tilde{B}(n) \cap F(q,\mathbf{j}))^k) &\lesssim k!\ q^{2k} (\rho(p_c,q))^k (\mathbf{C}_{18})^k \\ &\quad \times \left( \frac{\exp(-M\mathbf{C}_0 \log 2m)}{k!(\mathbf{C}_{18}\rho(p_c,q))^k} + \sum_{j=2}^{\log^* 2m} (\log^{(j-1)} 2m)^{-M\mathbf{C}_0} + 1 \right).\end{aligned}$$

Using the value of $q$ from (A.16) and choosing $M = \eta_2 k / 2\mathbf{C}_0$, we use (A.32) to get (A.23) in the case where $\mathbf{j}$ is such that $F(q,\mathbf{j}) \subset S(3m) \setminus S(m)^\circ$. For a general $\mathbf{j} \in \mathcal{J}$, the intersection

$$F(q,\mathbf{j}) \cap (S(3m) \setminus S(m)^\circ)$$

is a union of at most two rectangles with side lengths $r_1$ and $r_2$, $r_i \leq 3q/2$. Repeating the arguments above, we see that the size of the intersection of each of these rectangles with the backbone $\tilde{B}(n)$ enjoys the moment bounds (A.23), with $q^2\rho(p_c, q)$ replaced by $r_1 r_2 \rho(p_c, r)$, with $r = \max\{r_1, r_2\}$. Using (A.31), we obtain the upper bound $r^2\rho(p_c, r)$. For the higher moments, moment bounds of the form (A.23) with a larger (but still uniform in $\max(r_1, r_2) \leq 3q/2$) constant $\mathbf{C}_{16}$ are valid. (A.18) now follows by a union bound.

The proof of (A.24) follows a very similar pattern to the above. Instead of $Z_q(\mathbf{j}, j)$, we consider the sets

$$Z_n(j), \quad 1 \leq j \leq \log^* 2m$$

of points of $S(3m) \setminus S(m)$ with two disjoint $p_{2m}(j)$-open connections to $\partial S(m) \cup \partial S(3m)$. Repeating the steps for the case $k = 1$ gives (A.24) for the first moment. In the case of higher moments, we need to modify inequality (A.36), replacing it with

$$\mathbb{E}(\sharp Z_n(j))^k \leq k!(\mathbf{C}_{18} n^2 \rho(p_{2m}(j), n))^k \leq k!(M\mathbf{C}_{21} n^2 \rho(p_c, n) \log^{(j)} 2m)^k$$

for some $\mathbf{C}_{21}$. Decomposing as before over the events $(H_{2m}(j))$ and choosing $M \geq (k+1)/\mathbf{C}_0$ leads to (A.24).

### A.6.2 Proof of (A.19)

In its general outline, the proof is similar to that of (3.24) in [65], with some parameters chosen differently because we wish to bound only logarithmic deviations from the mean. However, the estimates in [65] are carried out for critical percolation ($p = p_c$), and the proof of the initial estimate (A.37) below in the supercritical case introduces an additional technical difficulty.

As explained in Section A.3, the entire (finite) $p_c$-open cluster of any site in the IPC also belongs to the IPC. Thus, for any crossing $r$ (in the IPC) of $F(\mathbf{j}, q) \setminus D(\mathbf{j}, q)$, the number of sites in $F(\mathbf{j}, q) \setminus D(\mathbf{j}, q)$ connected to $r$ by $p_c$-open paths provides a lower bound for the quantity in (A.19).

The starting point is the following:

**Lemma A.6.2** ([64])**.** *Let $r$ be a deterministic path crossing $S(27n) \setminus (S(n))^\circ$. Let $Z(n)$ be the set of sites $p_c$-connected to $r$ inside this annulus. We have the lower bound:*

$$\mathbb{P}_{p_c}(\sharp Z(n) \geq \mathbf{C}_{22} n^2 \pi(n)) \geq \mathbf{C}_{23},$$

*for some constants $\mathbf{C}_{22}, \mathbf{C}_{23} > 0$ independent of $n$.*

The proof is essentially that of (56) in [64]. Kesten's idea is to compute the first and second moments of the number $Y(n)$ of sites in $S(9n) \setminus (S(3n))^\circ$ connected to open circuits in $S(3n) \setminus (S(n))^\circ$ and $S(27n) \setminus (S(9n))^\circ$ (and thus to $r$) and use the Harris-FKG inequality and the second moment method.

Fix some $\delta > 0$ (to be chosen later). We first show that, for any $0 < t < q$ (entailing in particular $t < L(p_n(1))$ by (A.17)), and any coordinates $\mathbf{v} = (v_1, v_2)$ such that

$$T(\mathbf{v}) = [-t + v_1, t + v_1] \times [-t + v_2, 4t + v_2] \subset S(3m) \setminus (S(m))^\circ,$$

we have, for some constants $\mathbf{C}_{24}$, $\mathbf{C}_{25} > 0$,

$$\mathbb{P}\left(\begin{array}{c} \exists \text{ a } p_n(1)\text{-open crossing } r \text{ of } J(\mathbf{v}) = [v_1, t + v_1] \times [v_2, 3t + v_2] \\ \text{such that } \sharp\{x \in T(\mathbf{v}) : x \xrightarrow{p_c} r \text{ in } T(\mathbf{v})\} \leq \mathbf{C}_{24} t^2 \pi(t)/(\log t)^\delta \end{array}\right) \lesssim \frac{1}{(\log t)^{\mathbf{C}_{25}}}. \quad \text{(A.37)}$$

### A.6.3 Proof of (A.37)

For any crossing $r$ of $J(\mathbf{v})$, let

$$Z(T(\mathbf{v}), r) = \{x \in T(\mathbf{v}) : x \text{ is connected to } r \text{ in } T(\mathbf{v}) \text{ by a } p_c\text{-open path}\}.$$

The probability on the left of (A.37) equals

$$\begin{aligned}\mathbb{P}\left(\exists \text{ a } p_n(1)\text{-open crossing } r :\ \sharp Z(T(\mathbf{v}), r) \le \frac{\mathbf{C}_{24} t^2 \pi(t)}{(\log t)^\delta}\right) \\ \le \mathbb{P}\left(\exists \text{ a } p_c\text{-open crossing } r' :\ \sharp Z(T(\mathbf{v}), r') \le \frac{\mathbf{C}_{24} t^2 \pi(t)}{(\log t)^\delta}\right) \\ + \mathbb{P}\left(\begin{array}{c}\exists \text{ a } p_n(1)\text{-open crossing } r \text{ such that } r \\ \text{intersects no } p_c\text{-open crossing of } J(\mathbf{v})\end{array}\right). \quad \text{(A.38)}\end{aligned}$$

The precise meaning of "$r$ intersects no $p_c$-open crossing of $J(\mathbf{v})$" is that no site in $J(\mathbf{v})$ is a common endpoint of an edge in $r$ and an edge in some horizontal $p_c$-open crossing of $J(\mathbf{v})$. In particular, $r$ is edge-disjoint from all $p_c$-open crossings.

Both terms on the right in (A.38) will be bounded, up to a constant factor, by $(\log t)^{-\delta \mathbf{C}_{25}}$. We begin by estimating the first term in (A.38). For any crossing lattice path $r'$ of $J(\mathbf{v})$, let $J^-(r')$ be the set of edges with an endpoint that can be connected to $[v_1, v_1 + t] \times \{v_2\}$ by a path in $J(\mathbf{v})$ that does not touch $r'$ (below $r'$). Note that $J^-(r')$ may include edges not entirely contained in $J(\mathbf{v})$. The lowest $p_c$-open crossing $R_1$ of $J = J(\mathbf{v})$ is defined as the horizontal crossing of the rectangle by $p_c$-open edges such that the component $J^-(R_1)$, is minimal. $R_k$ is defined inductively as the lowest crossing of $J \setminus (J^-(R_{k-1}) \cup R_{k-1})$ (defined analogously – see [63, Prop. 2.3] for the existence of $R_k$ and precise definitions). For a given (lattice path) crossing $r'$ of $J(\mathbf{v})$, write $\Sigma_{r'}$ for the sigma algebra generated by the status of edges in $r' \cup J^-(r')$. We define $K$ to be the maximal $k$ such that $R_k$ exists. The veracity of the following string of inequalities is then evident:

$$\begin{aligned}&\mathbb{P}\left(\exists \text{ a } \ p_c\text{-open crossing } r' : \ \sharp Z(T(\mathbf{v}), r') \leq \frac{\mathbf{C}_{24} t^2 \pi(t)}{(\log t)^{\delta}}\right)\\ &\leq \sum_{k\geq 1} \mathbb{P}\left( \ \sharp Z(T(\mathbf{v}), R_k) \leq \frac{\mathbf{C}_{24} t^2 \pi(t)}{(\log t)^{\delta}}; K \geq k\right)\\ &\leq \sum_{k\geq 1}\sum_{r''} \mathbb{E}\left(\mathbb{P}( \ \sharp Z(T(\mathbf{v}), R_k) \leq \mathbf{C}_{24} t^2 \pi(t)/(\log t)^{\delta} \mid \Sigma_{r''}); R_k = r'', K \geq k)\right). \quad (A.39)\end{aligned}$$

On $\{R_k = r'', K \geq k\}$, we have the following uniform estimate for the conditional probability given $\Sigma_{r''}$:

$$\mathbb{P}(\sharp Z(T(\mathbf{v}), r'') \leq \mathbf{C}_{24} t^2 \pi(t)/(\log t)^{\delta} \mid \Sigma_{r''}) \lesssim \frac{1}{(\log t)^{\mathbf{C}_{25}}}. \tag{A.40}$$

To see this, consider the left endpoint $l_{r''}$ of the crossing $r''$, and annuli

$$A(l_{r''}, 3^k) = S(l_{r''}, 3 \cdot 3^k) \setminus S(l_{r''}, 3^k), \quad \frac{t}{(\log t)^{\delta/2}} \leq 3^k \leq t.$$

For $3^{3j} \leq t/27$, the existence of circuits $C'_j$ around $l_{r''}$ in $A(l_{r''}, 3^{3j})$ and $C''_j$ in $A(l_{r''}, 3^{3j+2})$, all of whose edges outside $J^-(r'')$ are $p_c$-open implies that any site in $A(l_{r''}, 3^{3j+1}) \cap ([-t, 0) \times \mathbb{R})$ connected to

$$\partial S(l_{r''}, 3^{3j}) \cup \partial S(l_{r''}, 3^{3j+3})$$

is $p_c$-connected to the crossing $r''$. Thus, using the Harris-FKG inequality, independence of the edge configurations in $J^-(r'')$ and $[-t, 0) \times \mathbb{R}$ and the second moment method as in the discussion preceding (A.37), there exist constants $\mathbf{C}_{24}$, $\mathbf{C}_{26}$, such that for each $j$ with $t/(\log t)^{\delta/2} \leq 3^{3j} \leq t/27$:

$$\begin{aligned}
&\mathbb{P}\left(\sharp Z(T(\mathbf{v}), r'') \leq \mathbf{C}_{24} t^2 \pi(t)/(\log t)^\delta \mid \Sigma_{r''}\right) \\
&\leq \mathbb{P}\left(\sharp\{x \in A(l_{r''}, 3^{3j+1}) \cap ([-t,0) \times \mathbb{R}) : x \xrightarrow{p_c} r''\} \leq \mathbf{C}_{24} 3^{6j} \pi(3^{3j}) \mid \Sigma_{r''}\right) \\
&\leq 1 - \mathbf{C}_{26}.
\end{aligned}$$

There are $(\delta/(6\log 3))\log\log t + O(1)$ admissible indices $j$, and so by independence of the configuration in the different annuli, we find

$$\mathbb{P}\left(\sharp Z(T(\mathbf{v}), r'') \leq \mathbf{C}_{24} t^2/(\log t)^\delta \mid \Sigma_{r''}\right) \lesssim (1 - \mathbf{C}_{26})^{(\delta/(6\log 3))\log\log t},$$

which is the same as (A.40).

Returning to the double sum of (A.39):

$$\begin{aligned}
&\mathbb{P}\left(\exists \text{ a } p_c\text{-open crossing } r :\ \sharp Z(T(\mathbf{v}), r) \leq \frac{\mathbf{C}_{22} t^2 \pi(t)}{(\log t)^\delta}\right) \\
&\lesssim \sum_{k \geq 1} \sum_r \frac{1}{(\log t)^{\mathbf{C}_{25}}} \mathbb{P}\left(R_k = r, K \geq k\right) \\
&= \frac{1}{(\log t)^{\mathbf{C}_{25}}} \sum_{k \geq 1} \mathbb{P}\left(K \geq k\right) \qquad \text{(A.41)}
\end{aligned}$$

By the Russo-Seymour-Welsh method, the $\mathbb{P}_{p_c}$ probability of a dual vertical crossing of $J(\mathbf{v})$ is bounded below by some $\epsilon > 0$. Thus, by disjointness of the $R_k$'s and the BK inequality,

$$\mathbb{P}\left(K \geq k\right) \leq \mathbb{P}(\exists\, k \text{ disjoint } p_c\text{-open crossings of } J(\mathbf{v})) \leq (1-\epsilon)^k.$$

This allows us to bound the sum in (A.41) by $C/\epsilon$.

We now estimate the second term on the right in (A.38). Denote by $\Xi$ the event that there exists a $p_n(1)$-open crossing $r$ of $J(\mathbf{v})$ such that $r$ intersects no $p_c$-open crossing of $J(\mathbf{v})$. For any $K_0 > 0$, we have

$$\mathbb{P}(\Xi) \le \mathbb{P}(\Xi, K \le K_0) + \mathbb{P}(K > K_0).$$

As previously, $K$ denotes the maximal number of disjoint $p_c$-open crossings of $J(\mathbf{v})$. We will choose $K_0 = c \log\log n$, so as to give the following bound for the second term above:

$$\mathbb{P}(K > K_0) \lesssim \exp(-\mathbf{C}_{27} \log\log n) = (\log n)^{-\mathbf{C}_{27}},$$

where the constant $\mathbf{C}_{27}$ is a constant such that $\mathbf{C}_{27} \ge \mathbf{C}_{25}$.

For the first term, we have the union bound

$$\mathbb{P}(\Xi, K \le K_0) \le \sum_{k=0}^{\lceil c \log\log n \rceil} \mathbb{P}(\Xi, K = k).$$

It will be shown below (see Lemma A.6.3) that there is a constant $\mathbf{C}_{28}$ such that, for any $\mathbf{v}$ with $T(\mathbf{v}) \subset S(3m) \setminus (S(m))^\circ$,

$$\mathbb{P}(\Xi, K = k) \le (\mathbf{C}_{28} \log t)^{2k} (p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t), \tag{A.42}$$

where $\pi_4(t) = \pi_4(t, p_c)$ is the "alternating 4-arm probability," associated to the event that $\langle \mathbf{0}, \mathbf{e}_1 \rangle$ is connected to $\partial S(t)$ by two disjoint $p_c$-open paths and its dual edge is connected to $\partial S(t)$ by two disjoint $p_c$-closed dual paths. Thus

$$\begin{aligned} \mathbb{P}(\Xi, K \le K_0) &\lesssim (\log\log n) \exp(2K_0 \log(\mathbf{C}_{28} \log t))(p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t) \\ &\le \exp\left(\mathbf{C}_{29} (\log\log n)^2\right) \cdot (p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t), \end{aligned} \tag{A.43}$$

for a constant $\mathbf{C}_{29}$. The factor $(p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t)$ is $O(n^{-c})$. Indeed, it was shown in [66] that, uniformly for $p > p_c$ sufficiently close to $p_c$:

$$L(p)^2 \pi_4(L(p), p_c)(p - p_c) \asymp 1. \tag{A.44}$$

Applying this to $p = p_t(1)$, and using (A.4) and $\pi_4(t/(M \log t)) \asymp \pi_4(t/\log t)$ [80, Proposition 12], we find, for $t$ large enough:

$$\frac{t^2}{(\log t)^2} \pi_4(t/\log t) \cdot (p_t(1) - p_c) \asymp 1.$$

Thus we have

$$(p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t) \lesssim \frac{p_n(1) - p_c}{p_t(1) - p_c} \cdot (\log t)^2.$$

Here we have used $\pi_4(t) \leq \pi_4(t/\log t)$. Using (A.44) again, we have:

$$\frac{p_n(1) - p_c}{p_t(1) - p_c} \asymp \frac{L(p_t(1))^2}{L(p_n(1))^2} \cdot \frac{\pi_4(L(p_t(1))}{\pi_4(L(p_n(1)))}.$$

By quasimultiplicativity [80, Proposition 12]:

$$\frac{\pi_4(L(p_t(1)))}{\pi_4(L(p_n(1)))} \asymp \frac{1}{\pi_4\left(L(p_t(1)), L(p_n(1))\right)},$$

where $\pi_4(n, N) = \pi_4(n, N; p_c)$ is the probability that there are four arms of alternating occupation status connecting $\partial S(n)$ to $\partial S(N)$ in $S(N) \setminus S(n)^\circ$. Using Reimer's inequality [83] and the (exact) scaling of the 5-arm exponent (see [80, Theorem 23] or [94]), we have:

$$\begin{aligned}\left(\frac{L(p_t(1))}{L(p_n(1))}\right)^2 &\asymp \pi_5\left(L(p_t(1)), L(p_n(1))\right) \\ &\lesssim \pi\left(L(p_t(1)), L(p_n(1))\right) \cdot \pi_4\left(L(p_t(1)), L(p_n(1))\right).\end{aligned}$$

Here, $\pi(n, N)$ is the one-arm probability, that $\partial S(n)$ is connected to $\partial S(N)$ by an open path. Since the one-arm probability satisfies the power-type upper bound $\pi(n) \lesssim n^{-\eta_1}$ for some $\eta_1 \leq 1/2$ (apply the BK inequality to the bound on $\eta_2$ in (A.17)), we find that $(p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t)$ is bounded, up to a constant, by

$$(\log t)^2 \left( \frac{L(p_t(1))}{L(p_n(1))} \right)^{\eta_1} \lesssim (\log n)^2 \left( \frac{t}{n} \right)^{\eta_1}. \tag{A.45}$$

Since we assume $t < q$, and $q = o(n^{1/2})$ (see (A.17)), we find

$$(p_n(1) - p_c) \cdot t^2 \cdot (\log t)^2 \cdot \pi_4(t) = O(n^{-c}),$$

for some $c > 0$. Returning to (A.43), we have the bound:

$$\mathbb{P}(\Xi, K \leq K_0) \lesssim n^{-c/2}.$$

It remains to prove (A.42). This is done in Lemma A.6.4 below. Before proceeding, let us introduce a definition: A *$p_c$-closed arm with $k$ defects* is a path of dual edges, all of which except for $k$ of them are $p_c$-closed. The proof of Lemma A.6.4 depends on the following:

**Lemma A.6.3.** *Let $\Xi$ be the event that there exists a $p_n(1)$-open crossing $r$ of $J(\mathbf{v})$ such that $r$ intersects no $p_c$-open crossing of $J(\mathbf{v})$, and $K$ be the maximal number of horizontal $p_c$-open crossings of $J(\mathbf{v})$. Suppose $K = k$; then there exists an edge $e \in J(\mathbf{v})$ such that*

1. *$e$ has two disjoint $p_n(1)$-open arms to $\{v_1\} \times [v_2, v_2 + 3t]$ (the left side of $J(\mathbf{v})$) and $\{v_1 + t\} \times [v_2, v_2 + 3t]$ (the right side of $J(\mathbf{v})$), respectively.*

2. *$e^*$, the dual edge to $e$, has two disjoint $p_c$-closed arms, each with at most $k$ defects to $[v_1, v_1+t] \times \{v_2\}$ (the bottom side of $J(\mathbf{v})$) and to $[v_1, v_1+t] \times \{v_2+3t\}$ (the top side of $J(\mathbf{v})$), respectively.*

3. *$w(e) \in [p_c, p_n(1)]$.*

*Proof.* On the event $\{K = k\}$, Menger's theorem [37, Section 3.3] implies that there is a dual path $s$ joining the top of $J(\mathbf{v})$ to the bottom, all of whose edges, with exactly $k$ exceptions, are closed and which moreover does not intersect itself. This path must intersect the horizontal $p_n(1)$-open crossing $r$ [63, Prop. 2.2] along a $p_n(1)$-open edge $e$. This edge must then be $p_c$-closed. The dual edge $e^*$, being part of the non-self-intersecting $s$ with $k$ defects, has two dual arms joining it to the top and bottom of $J(\mathbf{v})$. (See Figure A.1.) Moreover, the total number of defects on these two arms is $k$. This establishes the lemma. □

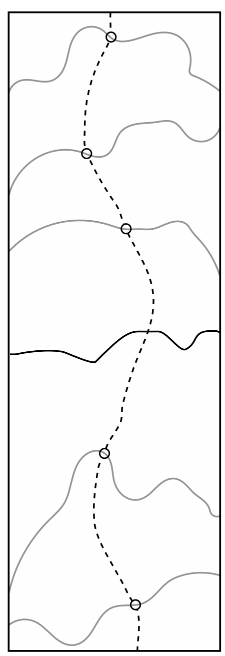


Figure A.1: Depiction of the application of Menger's theorem in the proof of Lemma A.6.3. The dotted path has $k$ defects, shown as empty circles. The solid black path represents a $p_n(1)$-open crossing and the grey paths represent disjoint $p_c$-open crossings.

The proof of (A.37) is concluded by the following lemma, which establishes the estimate (A.42):

**Lemma A.6.4.** *There is a constant* $\mathbf{C}_{28}$ *such that, for each* $k \geq 1$*, the following bound holds:*

$$\mathbb{P}(\Xi, K = k) \leq (\mathbf{C}_{28} \log t)^{2k}(p_n(1) - p_c) \cdot t^2 \cdot \pi_4(t). \tag{A.46}$$

It suffices to estimate the probability that there is an edge in $J(\mathbf{v})$ satisfying the two conditions in Lemma A.6.3. To that end, we will show that the expected number of such edges in $J(\mathbf{v})$ is bounded by the quantity on the right side of equation (A.46). For $e \in J(\mathbf{v})$, let $A_e^k$ be the event that $e$ satisfies the conditions of Lemma A.6.3.

The key step is the existence of a constant $\mathbf{C}_{29}$ such that

$$\mathbb{P}\left(A_e^k\right) \leq (\mathbf{C}_{29} \log t)^{2k}(p_n(1) - p_c)\mathbb{P}(B_e), \tag{A.47}$$

where $B_e$ is the event that $e$ has two disjoint $p_c$-open arms joining it to the left and right sides of $J(\mathbf{v})$ respectively, and $e^*$ has two disjoint $p_c$-closed dual arms to the top and bottom sides of $J(\mathbf{v})$. The effect of the arms with defects is to produce the logarithmic factor indicated in the equations above:

$$\mathbb{P}\left(A_e^k\right) \leq (\mathbf{C}_{29} \log t)^{2k}(p_n(1) - p_c)\mathbb{P}(A_e), \tag{A.48}$$

where $A_e$ is defined analogously to $B_e$ above, except that the open connections are required to be $p_n(1)$-open rather than $p$-open. This follows from the argument in [80, Prop. 17], where it is shown that if $A_{j,\sigma}^d(n)$ denotes the probability that the origin is connected to $\partial S(n)$ by $j$ paths, with $d$ defects in total, whose occupation status is specified by the sequence $\sigma \in \{\text{open}, \text{closed}\}^j$, then

$$\mathbb{P}(A_{j,\sigma}^d(n)) \lesssim_d (1 + \log n)^d \mathbb{P}(A_{j,\sigma}(n)). \tag{A.49}$$

$A_{j,\sigma}(n)$ is the event that there are $j$ arms (without defects) to $\partial S(n)$ (with occupation status as in $\sigma$). Inspection of the proof in [80] reveals that the constant implicit in (A.49) is of the form $(\mathbf{C}_{29})^d$. Separating the four arms as in [66] verifies that (A.48) holds.

It now remains to show that $\mathbb{P}(A_e) \lesssim \mathbb{P}(B_e)$; that is, that for $n$ sufficiently large, we can change the $p_n(1)$-open arms in the definition of $A_e$ to be $p_c$-open at the cost of a constant probability factor. For edges $e$ at distance $t/2$ from the boundary, this follows immediately from [36, Lemma 6.3]. We briefly sketch how the proof given there can be adapted to the case where $e$ is close to the boundary. We write $\mathbb{P}(A_e)$ as $\mathbb{P}(A_e(p_n(1), p_c))$, where for $p, q \in [p_c, 1)$, $A_e(p, q)$ denotes the event that $e$ has two disjoint $p$-open arms to opposite vertical sides of $J(\mathbf{v})$ and $e^*$ has two disjoint $q$-closed dual arms to the top and bottom of $J(\mathbf{v})$. Using Russo's formula as in [36, (39)] , we find

$$\frac{\mathrm{d}}{\mathrm{d}p}\mathbb{P}(A_e(p, p_c)) = \sum_{e' \neq e} \mathbb{P}(A_e(\cdot, p_c), A_{e'}(\cdot, p), D_{e,e'}(p)). \tag{A.50}$$

$A_e(\cdot, p_c)$ is the event that $e^*$ has two disjoint $p_c$-closed dual connections to the top and bottom of $J(\mathbf{v})$, and $D_{e,e'}(p)$ is the event that there exist three disjoint $p$-open paths joining, respectively, one vertical side of $J(\mathbf{v})$ to one endpoint of $e$, the other endpoint of $e$ to an endpoint of $e'$, and the other endpoint of $e'$ to the remaining vertical side of $J(\mathbf{v})$. Note that our notation differs somewhat from the one in [36]. For the purposes of illustration, we will henceforth suppose that $e = \langle (v_1 + l, v_2 + \lfloor 3t/2 \rfloor), (v_1 + l + 1, v_2 + \lfloor 3t/2 \rfloor) \rangle$ for some $l < t/4$; that is, $e$ is close to the left side of $J(\mathbf{v})$. The sum on the right of (A.50) can be rewritten as:

$$\left( \sum_{j=1}^{\lfloor l/2 \rfloor} + \sum_{j=\lfloor l/2 \rfloor + 1}^{l} + \sum_{j=l+1}^{3t} \right) \sum_{e' : |e'_x - e_x| = j} \mathbb{P}(A_e(\cdot, p_c), A_{e'}(\cdot, p), D_{e,e'}(p)). \tag{A.51}$$

$e_x$ denotes the left endpoint of the edge $e$, if $e$ is a horizontal edge, and its bottom endpoint if $e$ is a vertical edge. The first sum is bounded by

$$\begin{aligned}\sum_{j=1}^{\lfloor l/2 \rfloor} \sum_{e':|e'_x - e_x|=j} & \mathbb{P}(A_e(\lfloor j/2 \rfloor; p, p_c))\mathbb{P}(A_{e'}(\lfloor j/2 \rfloor; p, p)) \\ & \times \mathbb{P}(A(\lfloor 3j/2 \rfloor, l; p, p_c))\mathbb{P}(T(l,t;p,p_c)) \\ \lesssim & \sum_{j=1}^{\lfloor l/2 \rfloor} j\mathbb{P}(A_e(\lfloor j/2 \rfloor; p, p_c))\mathbb{P}(A(1, \lfloor j/2 \rfloor; p, p)) \\ & \times \mathbb{P}(A(\lfloor 3j/2 \rfloor, l; p, p_c))\mathbb{P}(T(l,t;p,p_c))\end{aligned} \tag{A.52}$$

$A(n, N; p, p_c)$ denotes the probability that there are four arms of alternating occupation status joining $\partial S(n)$ to $\partial S(N)$; $T(l, t; p, p_c)$ is the event that there are two $p_c$-closed arms, as well as a $p$-open arm connecting $\partial S(l)$ to $\partial S(t)$. Using gluing constructions similar to those in proofs of quasi-multiplicativity, and the fact that we may change the length of any connections involved by constant factors at the cost of constant factors in the probabilities, we have:

$$\mathbb{P}(A_e(\lfloor j/2 \rfloor; p, p_c))\mathbb{P}(A(\lfloor 3j/2 \rfloor, l; p, p_c))\mathbb{P}(T(l,t;p,p_c)) \asymp \mathbb{P}(A_e(p, p_c)).$$

For $p \le p_n(1) < p_t(1)$, we can use [80, Theorem 27] to assert

$$\mathbb{P}(A(1, \lfloor j/2 \rfloor; p, p)) \asymp \mathbb{P}(A(1, \lfloor j/2 \rfloor; p_c, p_c)).$$

We can now follow [36] exactly (see equations (42) and (43) and the surrounding discussion) to show that the sum in (A.52) is bounded by:

$$\mathbb{P}(A_e(p, p_c)) \cdot l^2 \pi_4(t) \le \mathbb{P}(A_e(p, p_c)) \cdot t^2 \pi_4(t).$$

To deal with the second sum in (A.51), we note that when

$$|e_x - e'_x| \geq \lfloor l/2 \rfloor + 1,$$

the conjunction of the events $A_e(\cdot, p_c), A_{e'}(\cdot, p)$ and $D_{e,e'}(p)$ appearing in the probability on the right of the equation implies that $e$ has 2 $p$-open, and $e^*$ two $p_c$-closed arms to distance $\lfloor l/4 \rfloor$, that $e'$ has four alternating arms with parameter $p$ to the boundary of the intersection of $S(e'_x, \lfloor l/4 \rfloor)$ with $J(\mathbf{v})$, three of which reach to distance $\lfloor l/4 \rfloor$, and finally that $\partial S(e_x, \lfloor 5l/4 \rfloor)$ has two $p_c$-closed arms to the top and bottom of $J(\mathbf{v})$ and a $p$-open arm to the right side of $J(\mathbf{v})$, and all these connections occur inside $J(\mathbf{v})$. Using these observations, an argument similar to the previous case and a summation analogous to that in the proof of [94, Lemma 6.2], shows that we can estimate (in addition, using the remarks in [80, Section 4.6] to change the $p$-open and closed arms in a half-plane to $p_c$-open and closed arms)

$$\sum_{j=\lfloor l/2 \rfloor+1}^{l} \sum_{e':|e'_x - e_x|=j} \mathbb{P}(A_e(\cdot, p_c), A_{e'}(\cdot, p), D_{e,e'}(p)) \lesssim \mathbb{P}(A_e(p, p_c)) \cdot l^2 \pi_4(l).$$

Turning to the final sum on the left in (A.51), we can again closely follow [36] to bound this term by

$$t^2 \pi_4(t) \cdot \mathbb{P}(A_e(p, p_c)).$$

The estimates outlined above for the left side of (A.50) imply

$$\frac{\mathrm{d}}{\mathrm{d}p} \log \mathbb{P}(A_e(p, p_c)) \leq C(l^2 \pi_4(l) + t^2 \pi_4(t)).$$

Integrating this from $p_c$ to $p_n(1)$ and using (A.44), we find

$$\mathbb{P}(A_e(p_n(1), p_c)) \lesssim \mathbb{P}(A_e(p_c, p_c)),$$

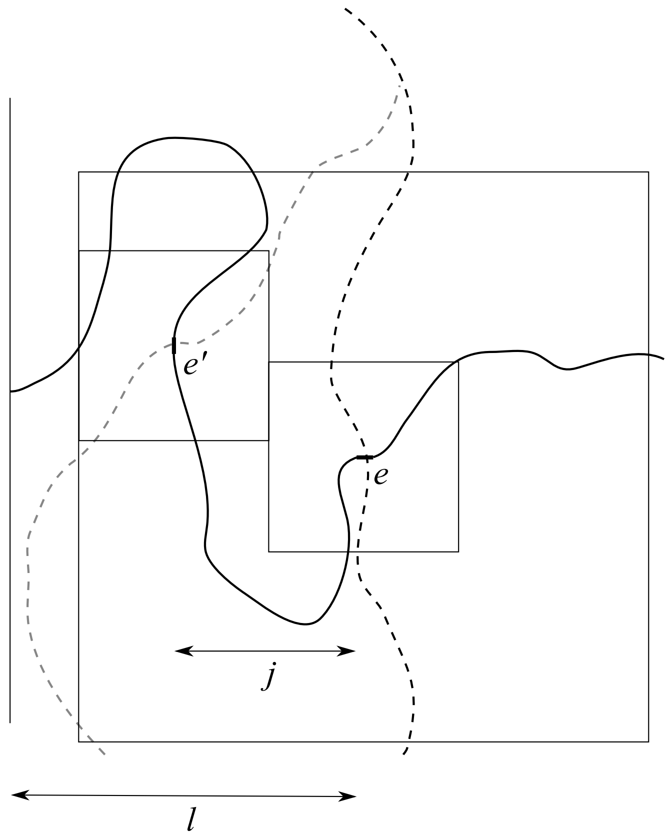



Figure A.2: Illustration of the setup for management of the second sum in (A.51). The edge $e$ is at distance $l$ from the left boundary of $J(\mathbf{v})$ and the distance between $e'$ and $e$ is $j$, a number between $l/2$ and $l$. The dark dotted curve represents a $p_c$-closed dual path (given by Menger's theorem) and the dark solid curve represents a $p_n(1)$-open path, connecting $e'$ and $e$ to each other and to the left and right sides of $J(\mathbf{v})$. The grey dotted curve represents a $p_n(1)$-closed dual path connecting the edge dual to $e'$ with the top and bottom of $J(\mathbf{v})$.

which is what we wanted to prove. We have thus established (A.47); that is, we have shown

$$\mathbb{P}(\Xi, K = k) \leq \mathbb{E}[\sharp N_k] \lesssim (\mathbf{C}_{30} \log t)^{2k} \cdot (p_n(1) - p_c) \sum_{e \in J(\mathbf{v})} \mathbb{P}(B_e),$$

where $N_k$ is the number of edges satisfying the conditions in Lemma A.6.3. Note that $B_e$ is equal to the event that the edge $e$ is *pivotal* for the existence of a left-right $p_c$-open crossing of $J(\mathbf{v})$. Following [94, Lemma 6.2], we can show

$$\sum_{e \in J(\mathbf{v})} \mathbb{P}(B_e) \lesssim t^2 \cdot \pi_4(t).$$

This concludes the proof of the lemma.

### A.6.4 Final Peierls argument

We use a block argument and a Peierls argument to upgrade (A.37) to (A.19). The annulus $F(\mathbf{j},q)\setminus D(\mathbf{j},q)$, centred at

$$v = q\left(j_1+\frac{1}{2}, j_2+\frac{1}{2}\right)$$

is tiled with smaller squares of side length

$$t=\frac{q}{\log q}.$$

The existence of a $p_n(1)$-open crossing of $F(\mathbf{j},q)\setminus D(\mathbf{j},q)$ implies that of a crossing

$$\bar{r}=(x(0),x(1),\ldots,x(\xi)),$$

of $S(v,3q/2-5t)\setminus S(v,q/2+5t)$ along edges of $\mathbb{Z}^2$, with $x(0)\in S(v,q/2+5t)$ and $x(\xi)\in S(v,3q/2)\setminus(S(v,3q/2-5t))^\circ$. The reason for considering this smaller annulus will become clear below. We can now introduce sequences $\mathbf{j}_0,\ldots,\mathbf{j}_\lambda$, and $l_0=0,\ldots,l_\lambda$ relative to the sequence $x(i)$ and squares of size $t$; that is,

$$x(l_i)\in D(\mathbf{j}_i,t)$$

$$l_{i+1}=\min\{l>l_i : x(l)\notin F(\mathbf{j}_i,t)\}.$$

The first observation is that we have a lower bound on $\lambda$ due to the difference in scales:

$$q-10t\le |x(\xi)-x(0)|\le \sum_{l=0}^{\lambda-1}|x(l_{i+1})-x(l_i)|+|x(\xi)-x(l_\lambda)|\le 2\sqrt{2}t(\lambda+1),$$

implying

$$\lambda \geq \mathbf{C}_{31} \log q.$$

The second observation is that

$$|\mathbf{j}_{i+1}(k) - \mathbf{j}_i(k)| \leq 2, \; k = 1, 2,$$

where $\mathbf{j}_i(k)$ denotes the $k$-th coordinate of the vector $\mathbf{j}$. From this, for each fixed $\lambda$, given $\mathbf{j}_i$, there are at most 16 choices for $\mathbf{j}_{i+1}$ and so at most

$$4\left(\frac{q}{t} + 11\right) 16^{\lambda}$$

choices for the sequence $\mathbf{j}_0, \ldots, \mathbf{j}_\lambda$. The first factor is an estimate for the number of choices of squares $D(\mathbf{j}, t)$ with $x(0) \in D(\mathbf{j}, t)$.

The third observation is that $\bar{r}$ must contain, between $x(\mathbf{j}_i)$ and $x(\mathbf{j}_{i+1})$, a "short" crossing $r_i$ of a $t \times 3t$ or $3t \times t$ rectangle $R_i$ (that is, the crossing is between the long sides).

Denote by $\tilde{R}_i$ the $2t \times 5t$ or $5t \times 2t$ rectangle around $R_i$, as in (A.37). Then

$$\tilde{R}_i \subset S(v, 3q/2) \setminus S(v, q/2),$$

and so

$$\hat{Z}(v, q/2, \bar{r}) = \{x \in S(v, 3q/2) : x \xrightarrow{p_c} \bar{r} \text{ in } S(v, 3q/2) \setminus S(v, q/2)\}$$
$$\geq \max_{0 \leq i \leq \lambda} Z(\tilde{R}_i, r_i),$$

where $Z(\tilde{R}_i, r_i)$ is the number of points in $\tilde{R}_i$ connected to $r_i$ by a $p_c$-open path in $\tilde{R}_i$. It follows that

$$\mathbb{P}\left(\begin{array}{c}\exists p_{2m}(1)\text{-open crossing of } F(\mathbf{j},q)\setminus D(\mathbf{j},q) \\ \text{with } \hat{Z}(v,q/2,r)\leq q^2\pi(q)/(\log q)^4\end{array}\right)$$
$$\leq \sum_{\tilde{R}_0,\ldots,\tilde{R}_\lambda} \mathbb{P}\left(\begin{array}{c}\text{for all } i\leq\lambda, \exists p_{2m}(1)\text{-open crossing } r_i \text{ in } R_i \\ \text{with } Z(\tilde{R}_i, r_i)\leq q^2\pi(q)/(\log q)^4\end{array}\right). \quad \text{(A.53)}$$

The sum is over all possible finite sequences of squares $\{\tilde{R}_i\}_{i\leq\lambda}$, for all $\lambda \geq \mathbf{C}_{31}\log q$. This quantity is controlled by choosing a subsequence of $\mathbf{C}_{32}\lambda$ disjoint $\tilde{R}_i$: each rectangle intersects a fixed number of other such rectangles. The events appearing in the last probability are independent for disjoint $\tilde{R}_i$'s. Their probability can be bounded using (A.37) (with $\delta = 1$), since our choice of $t$ implies

$$\frac{q^2}{(\log q)^4}\pi(q) \leq \mathbf{C}_{33}\frac{t^2}{\log t}\pi(t),$$

for large $q$. Moreover, one can use the bound on the number of sequences of $\mathbf{j}$'s (there are at most 4 choices of $R_i$ for a given $\mathbf{j}_i$) to control the entire sum: the last line in (A.53) is bounded up to a constant factor by:

$$\sum_{\lambda\geq\mathbf{C}_{31}\log q} q64^\lambda(\mathbf{C}_{33}(\log t)^{-\mathbf{C}_{34}})^{\mathbf{C}_{32}\lambda}.$$

For $q$ large enough, this sum is bounded (again up to a constant) by:

$$\exp(-\mathbf{C}_{35}\log q\cdot\log\log t) \ll q^{-c}$$

for any $c > 0$.

On $H_{2m}(1)$, any crossing $r$ in the portion of the IPC $\Gamma(n)$ consists of $p_{2m}(1)$-open edges. Since any site $p_c$-connected to a site in the IPC also belongs to the IPC, we find that the probability in (A.19) is bounded by:

$$\mathbb{P}(H_{2m}(1)^c) + \mathbb{P}\begin{pmatrix} \exists \text{ a } p_{2m}(1)\text{-open crossing of } F(\mathbf{j},q) \setminus D(\mathbf{j},q) \\ \text{with } \hat{Z}(v,q/2,r) \le q^2\pi(q)/(\log q)^4 \end{pmatrix}$$
$$\lesssim (2m)^{-M\mathbf{C}_0} + \exp(-\mathbf{C}_{35}\log q(\log\log q - \log\log\log q)). \quad \text{(A.54)}$$

Choosing $M$ appropriately in the definition of $p_n(1)$ (depending on the parameter $c$ in (A.19)) establishes the claim.

## A.7 Quenched subdiffusivity on the Incipient Infinite Cluster

In this section, we justify Remark 3 above and outline the derivation of a result analogous to Theorem A.2.4 for the random walk on H. Kesten's *Incipient Infinite Cluster* (IIC). For cylinder events $A$, the IIC measure is defined by

$$\mathbb{P}_{\mathrm{IIC}}(A) = \lim_{l\to\infty} \mathbb{P}_{p_c}(A \mid \mathbf{0} \to \partial S(l)). \quad \text{(A.55)}$$

It was shown in [64] that the limit (A.55) exists and that the resulting set function extends to a measure. Note that the connected cluster of the origin, $C(\mathbf{0})$, is $\mathbb{P}_{\mathrm{IIC}}$-almost surely unbounded. We will refer to this cluster as the IIC. We have the following result:

**Theorem A.7.1** (Quenched Kesten theorem for the IIC)**.** *Let $\{X_k\}_{k\ge 0}$ denote a simple random walk on the incipient infinite cluster started at $\mathbf{0}$. Let $\tau(n)$ denote the first exit time of $X_k$ from $S(n)$. There exists $\epsilon > 0$ such that, for $\mathbb{P}_{\mathrm{IIC}}$-almost every $\omega$*

*and almost-every realization of* $\{X_k\}$*, there is a (random)* $n_0$ *such that*

$$\tau(n) \geq n^{2+\epsilon}$$

*for* $n$ *greater than* $n_0$*.*

We can proceed along the lines of the proof of estimate (A.20), and consider a suitable modification of the random walk whose distribution coincides with that of $X$ from the first hitting time $\tau(2m)$ of $\partial S(2m)$ to the first hitting time of $\partial S(m) \cup \partial S(3m)$ after time $\tau(2m)$, $\sigma^+(m)$. To use the argument leading to (A.20) in our case, we merely need to show that we can prove an estimate equivalent to the one obtained for $\mathbb{P}_{\mathrm{IPC}}(E_1(n)^c)$ in Section A.5.

We will show that there are constants $C > 0$ and $s > 1$ such that

$$\mathbb{P}_{\mathrm{IIC}}(\{\mathrm{dist}_{C(\mathbf{0})}(\partial S(2m), \partial S(n) \cup \partial S(m)) \leq Cn^s\}) \lesssim n^{-2}, \tag{A.56}$$

By the argument given in the proof of Lemma A.4.1, there exists $C > 0$ and $s > 1$ such that

$$\mathbb{P}_{p_c}\left(\begin{array}{c}\exists \text{ an open path in } S(3m) \setminus S(m)^\circ \text{ connecting } \partial S(2m) \\ \text{to } \partial S(m) \text{ or } \partial S(3m) \text{ with less than } Cn^s \text{ edges}\end{array}\right) \lesssim n^{-2}. \tag{A.57}$$

Let us denote the event on the left by $G(n)$. Clearly

$$\{\mathrm{dist}_{C(\mathbf{0})}(\partial S(2m), \partial S(n) \cup \partial S(m)) \leq Cn^s\} \subset G(n).$$

$G(n)$ depends only on the status of edges inside $S(3m) \setminus S(m)^\circ$. Write the conditional probability in the definition of $\mathbb{P}_{\mathrm{IIC}}$ as a ratio:

$$\mathbb{P}_{p_c}(G(n) \mid \mathbf{0} \to \partial S(l)) = \frac{\mathbb{P}_{p_c}(G(n), \mathbf{0} \to \partial S(l))}{\mathbb{P}_{p_c}(\mathbf{0} \to \partial S(l))}.$$

For $l > 3m$, we have, by independence and monotonicity

$$\mathbb{P}_{p_c}(G(n), \mathbf{0} \to \partial S(l)) \leq \mathbb{P}_{p_c}(G(n))\mathbb{P}_{p_c}(\mathbf{0} \to \partial S(m))\mathbb{P}_{p_c}(\partial S(3m) \to \partial S(l)). \quad \text{(A.58)}$$

Now

$$\mathbb{P}_{p_c}(\mathbf{0} \to \partial S(m)) \asymp \mathbb{P}_{p_c}(\mathbf{0} \to \partial S(3m)),$$

and by quasi-multiplicativity

$$\mathbb{P}_{p_c}(\mathbf{0} \to \partial S(3m)) \cdot \mathbb{P}_{p_c}(\partial S(3m) \to \partial S(l)) \asymp \mathbb{P}_{p_c}(\mathbf{0} \to \partial S(l)).$$

Using this in (A.58), we have, by (A.57):

$$\mathbb{P}_{p_c}(G(n) \mid \mathbf{0} \to \partial S(l)) \lesssim n^{-2},$$

from which (A.56) follows at once.

# Bibliography


[1] Michael Aizenman and Almut Burchard. Hölder regularity and dimension bounds for random curves. *Duke Mathematical Journal*, 99(3):419–453, 1999.

[2] Michael Aizenman and Jan Wehr. Rounding of first-order phase transitions in systems with quenched disorder. *Physical Review Letters*, 62(21):2503–2506, 1989.

[3] Kenneth S. Alexander. Approximation of subadditive functions and convergence rates in limiting-shape results. *The Annals of Probability*, 25(1):30–55, 1997.

[4] Sven Erick Alm and John C. Wierman. Inequalities for means of restricted first-passage times in percolation theory. *Combinatorics Probability and Computing*, 8(4):307–315, 1999.

[5] Omer Angel, Jesse Goodman, Frank den Hollander, and Gordon Slade. Invasion percolation on regular trees. *The Annals of Probability*, 36(2):420–466, 2008.

[6] Louis-Pierre Arguin and Michael Damron. On the number of ground states of the Edwards-Anderson spin glass model. *arXiv:1110.6913*, 2011.

[7] Louis-Pierre Arguin, Michael Damron, Charles M. Newman, and Daniel L. Stein. Uniqueness of ground states for short-range spin glasses in the half-plane. *Communications in Mathematical Physics*, 300(3):641–657, 2010.

[8] Antonio Auffinger and Michael Damron. Differentiability at the edge of the percolation cone and related results in first-passage percolation. *arXiv:1105.4172*, 2011.

[9] Antonio Auffinger and Michael Damron. A simplified proof of the relation between scaling exponents in first-passage percolation. *arXiv:1109.0523*, 2011.

[10] Antonio Auffinger, Michael Damron, and Jack Hanson. Limiting geodesics for first-passage percolation on subsets of Zˆ2. *arXiv:1302.5413*, 2013.

[11] Martin T. Barlow, Antal A. Járai, Takashi Kumagai, and Gordon Slade. Random walk on the incipient infinite cluster for oriented percolation in high dimensions. *Communications in Mathematical Physics*, 278(2):385–431, 2008.

[12] Martin T. Barlow and Takashi Kumagai. Random walk on the incipient infinite cluster on trees. *Illinois Journal of Mathematics*, 50(1):33–65, 2006.

[13] Vincent Beffara and Pierre Nolin. On monochromatic arm exponents for 2D critical percolation. *The Annals of Probability*, 39(4):1286–1304, 2011.

[14] Michel Benaïm and Raphaël Rossignol. Exponential concentration for first passage percolation through modified poincar inequalities. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44(3):544–573, 2008.

[15] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sublinear distance variance. *The Annals of Probability*, 31(4):1970–1978, 2003.

[16] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm. Uniform spanning forests. *The Annals of Probability*, 29(1):1–65, 2001.

[17] Jacob van den Berg and Harry Kesten. Inequalities with applications to percolation and reliability. *Journal of Applied Probability*, 22(3):556–569, 1985.

[18] Noam Berger and Marek Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probability Theory and Related Fields*, 137(1-2):83–120, January 2007.

[19] Daniel Boivin. First passage percolation: The stationary case. *Probability Theory and Related Fields*, 86(4).

[20] Bela Bollobás and Oliver Riordan. *Percolation*. Cambridge University Press, 2006.

[21] Erwin Bolthausen. A note on the diffusion of directed polymers in a random environment. *Communications in Mathematical Physics*, 123(4):529–534, 1989.

[22] Simon Broadbent and John Hammersley. Percolation processes I. crystals and mazes. *Proceedings of the Cambridge Philosophical Society*, 53:629–641, 1957.

[23] Robert M. Burton and Michael Keane. Density and uniqueness in percolation. *Communications in Mathematical Physics*, 121(3):501–505, 1989.

[24] Philippe Carmona and Yueyun Hu. On the partition function of a directed polymer in a gaussian random environment. *Probability Theory and Related Fields*, 124(3):431–457, 2002.

[25] Thomas K. Carne. A transmutation formula for Markov chains. *Bulletin des Sciences Mathématiques*, 109(4):399–405, 1985.

[26] Richard Chandler, Joel Koplik, Kenneth Lerman, and Jorge F. Willemsen. Capillary displacement and percolation in porous media. *Journal of Fluid Mechanics*, 119:249–267, 1982.

[27] Sourav Chatterjee. The universal relation between scaling exponents in first-passage percolation. *arXiv:1105.4566*.

[28] Jennifer T. Chayes, Lincoln Chayes, and Charles M. Newman. The stochastic geometry of invasion percolation. *Communications in Mathematical Physics*, 101(3):383–407, 1985.

[29] J. Cook and Bernard Derrida. Polymers on disordered hierarchical lattices: A nonlinear combination of random variables. *Journal of Statistical Physics*, 57(1-2):89–139, 1989.

[30] J. Cook and Bernard Derrida. Directed polymers in a random medium: 1/d expansion and the n-tree approximation. *Journal of Physics A: Mathematical and General*, 23(9):1523–1554, 1990.

[31] J. Theodore Cox and Richard Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *The Annals of Probability*, 9(4):583–603, 1981.

[32] David Croydon and Takashi Kumagai. Random walks on galton-watson trees with infinite variance offspring distribution conditioned to survive. *Electronic Journal of Probability*, 13:1419–1441, 2008.

[33] Michael Damron. Private communication.

[34] Michael Damron and Jack Hanson. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. *arXiv:1209.3036*.

[35] Michael Damron, Jack Hanson, and Philippe Sosoe. Subdiffusivity of random walk on the 2d invasion percolation cluster. *arXiv:1210.7267*.

[36] Michael Damron, Artëm Sapozhnikov, and Bálint Vágvölgyi. Relations between invasion percolation and critical percolation in two dimensions. *The Annals of Probability*, 37(6):2297–2331, 2009.

[37] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics 173. Springer, fourth edition, 2010.

[38] Richard Durrett and Thomas M. Liggett. The shape of the limit set in richardson's growth model. *The Annals of Probability*, 9(2):186–193, 1981.

[39] Daniel S. Fisher and David A. Huse. Directed paths in a random potential. *Physical Review B*, 43(13):10728–10742, 1991.

[40] Gabor Forgacs, Reinhard Lipowsky, and Theo M. Nieuwenhuizen. The behavior of interfaces in ordered and disordered systems. In C. Domb and J. Lebowitz, editors, *Phase Transitions and Critical Phenomena*, volume 14, pages 135–363. Academic Press, 1991.

[41] Olivier Garet and Régine Marchand. Coexistence in two-type first-passage percolation models. *The Annals of Applied Probability*, 15(1):298–330, 2005.

[42] Jean-Baptiste Gouéré. Shape of territories in some competing growth models. *The Annals of Applied Probability*, pages 1273–1305, 2007.

[43] Geoffrey Grimmett. *Percolation*. Springer, 1999.

[44] Geoffrey Grimmett and Harry Kesten. Percolation since Saint-Flour. *arXiv:1207.0373*, 2012.

[45] Olle Häggström. Infinite clusters in dependent automorphism invariant percolation on trees. *The Annals of Probability*, pages 1423–1436, 1997.

[46] Olle Häggström. Invariant percolation on trees and the mass-transport method. In *Bulletin of the International Statistical Institute*, volume 52nd Session Proceedings, Tome LVIII, Book 1, pages 1511–1522. 1999.

[47] Olle Häggström and Ronald Meester. Asymptotic shapes for stationary first passage percolation. *The Annals of Probability*, 23(4):1511–1522, 1995.

[48] Olle Häggström and Robin Pemantle. First passage percolation and a model for competing spatial growth. *Journal of Applied Probability*, 35(3):683–692, 1998.

[49] John M. Hammersley. First-passage percolation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 28(3):491–496, 1966.

[50] John M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley*. Springer-Verlag.

[51] Christopher Hoffman. Geodesics in first passage percolation. *The Annals of Applied Probability*, 18(5).

[52] Christopher Hoffman. Coexistence for Richardson type competing spatial growth models. *The Annals of Applied Probability*, 15(1):739–747, 2005.

[53] David A. Huse and Christopher L. Henley. Pinning and roughening of domain walls in ising systems due to random impurities. *Physical Review Letters*, 54(25):2708–2711, 1985.

[54] David A. Huse, Christopher L. Henley, and Daniel S. Fisher. Huse, henley, and fisher respond. *Physical Review Letters*, 55(26):2924–2924, 1985.

[55] John Z. Imbrie and Thomas Spencer. Diffusion of directed polymers in a random environment. *Journal of Statistical Physics*, 52(3-4):609–626, 1988.

[56] Antal A. Járai. Invasion percolation and the incipient infinite cluster in 2D. *Communications in Mathematical Physics*, 236(2):311–334, 2003.

[57] Olav Kallenberg. *Foundations of modern probability*. Springer Verlag, second edition, 2002.

[58] Mehran Kardar and David R. Nelson. Commensurate-incommensurate transitions with quenched random impurities. *Physical Review Letters*, 55(11):1157–1160, 1985.

[59] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889–892, 1986.

[60] Mehran Kardar and Yi-Cheng Zhang. Scaling of directed polymers in random media. *Physical Review Letters*, 58(20):2087–2090, 1987.

[61] Harry Kesten. Aspects of first passage percolation. In P. L. Hennequin, editor, *École d'été de Probabilités de Saint Flour XIV - 1984*, number 1180 in Lecture Notes in Mathematics. Springer Berlin Heidelberg.

[62] Harry Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in Mathematical Physics*, 74(1):41–59, 1980.

[63] Harry Kesten. *Percolation theory for mathematicians*. Birkhäuser Boston, 1982.

[64] Harry Kesten. The incipient infinite cluster in two-dimensional percolation. *Probability Theory and Related Fields*, 73(3):369–394, 1986.

[65] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 22(4):425–487, 1986.

[66] Harry Kesten. Scaling relations for 2D-percolation. *Communications in Mathematical Physics*, 109(1):109–156, 1987.

[67] Harry Kesten. On the speed of convergence in first-passage percolation. *The Annals of Applied Probability*, 3(2):296–338, 1993.

[68] John F. C. Kingman. The ergodic theory of subadditive stochastic processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(3):499–510, 1968.

[69] Gady Kozma and Asaf Nachmias. The Alexander-Orbach conjecture holds in high dimensions. *Inventiones Mathematicae*, 178(3):635–654, 2009.

[70] Joachim Krug. Scaling relation for a growing interface. *Physical Review A*, 36(11):5465–5466, 1987.

[71] Joachim Krug and Herbert Spohn. Kinetic roughening of growing surfaces. In *Solids Far from Equilibrium: Growth, Morphology and Defects*.

[72] Cristina Licea and Charles M. Newman. Geodesics in two-dimensional first-passage percolation. *The Annals of Probability*, 24(1):399–410, 1996.

[73] Cristina Licea, Charles M. Newman, and Marcelo S. T. Piza. Superdiffusivity in first-passage percolation. *Probability Theory and Related Fields*, 106(4):559–591, 1996.

[74] Russell Lyons and Yuval Peres. *Probability on trees and networks.* Cambridge University Press, in preparation, 2012.

[75] James B. Martin. Limiting shape for directed percolation models. *The Annals of Probability*, 32(4):2908–2937, 2004.

[76] Charles M. Newman. A surface view of first-passage percolation. In *Proceedings of the International Congress of Mathematicians*, page 10171023, 1995.

[77] Charles M. Newman. *Topics in disordered systems.* Springer, 1997.

[78] Charles M. Newman and Marcelo S. T. Piza. Divergence of shape fluctuations in two dimensions. *The Annals of Probability*, 23(3):977–1005, 1995.

[79] Bao Gia Nguyen. Typical cluster size for two-dimensional percolation processes. *Journal of Statistical Physics*, 50(3-4):715–726, 1988.

[80] Pierre Nolin. Near-critical percolation in two dimensions. *Electronic Journal of Probability*, 13:1562–1623, 2008.

[81] Robin Pemantle and Yuval Peres. Planar first-passage percolation times are not tight. In *NATO ASI Series C Mathematical and Physical Sciences-Advanced Study Institute*, volume 420, pages 261–264. Kluwer Acad.Publ. Dordrecht, 1994.

[82] Agoston Pisztora. Scaling inequalities for shortest paths in regular and invasion percolation. *Carnegie-Mellon CNA preprint*, 2000.

[83] David Reimer. Proof of the van den Berg - Kesten conjecture. *Combinatorics, Probability, and Computing*, 9(1):27–32, 2000.

[84] Daniel Richardson. Random growth in a tessellation. In *Proc. Cambridge Philos. Soc*, volume 74, pages 515–528, 1973.

[85] Hermann Rost. Non-equilibrium behaviour of a many particle process: Density profile and local equilibria. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 58(1):41–53, 1981.

[86] Timo Seppäläinen. Scaling for a one-dimensional directed polymer with boundary conditions. *The Annals of Probability*, 40(1):19–73, 2012.

[87] Daisuke Shiraishi. Two-sided random walks conditioned to have no intersections. *arXiv:1106.5295.*

[88] Yakov Sinai. A remark concerning random walks with random potentials. *Fundamenta Mathematicae*, 147(2):173–180, 1995.

[89] Daniel L. Stein. Spin glasses: Old and new complexity. *AIP Conference Proceedings*, 1389(1).

[90] Nicholas Varopoulos. Long range estimates for markov chains. *Bulletin des sciences mathématiques*, 109(3):225–252, 1985.

[91] Jan Wehr. On the number of infinite geodesics and ground states in disordered systems. *Journal of Statistical Physics*, 87:439–447, 1997.

[92] Jan Wehr and Michael Aizenman. Fluctuations of extensive functions of quenched random couplings. *Journal of Statistical Physics*, 60(3-4):287–306, 1990.

[93] Jan Wehr and Jung Woo. Absence of geodesics in first-passage percolation on a half-plane. *The Annals of Probability*, 26(1):358–367, 1998.

[94] Wendelin Werner. Lectures on two-dimensional critical percolation. *IAS-Park City Summer School notes*, 2007.

[95] John C. Wierman and Wolfgang Reh. On conjectures in first passage percolation theory. *The Annals of Probability*, pages 388–397, 1978.

[96] David Wilkinson and Jorge F. Willemsen. Invasion percolation: a new form of percolation theory. *Journal of Physics A: Mathematical and General*, 16(14):3365–3376, 1983.

[97] Yu Zhang. The fractal volume of the two-dimensional invasion percolation cluster. *Communications in Mathematical Physics*, 167(2):237–254, 1995.

[98] Yu Zhang. Double behavior of critical first-passage percolation. In *Perplexing problems in probability*, pages 143–158. Springer, 1999.