A Walk through the Forest: the Geometry and Topology of Random Systems

Noah Alfred Maurice Halberstam

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Queens' College

August 2023

Declaration

I declare that this thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification.

Noah Halberstam
August 2023

Abstract

We prove several theorems on the geometry and topology of random walks and random forests, with analysis of the latter of these random systems often relying on analysis of the former and vice versa. The main models we consider are the static and dynamic random conductance models, the uniform spanning forest, the arboreal gas and countable Markov chains, and we will be interested in both the qualitative and quantitative behaviour of these systems over large scales. The quantitative properties of both the random system and its underlying medium are in this work and in general often encoded as a set of dimensions, or exponents, which govern how those properties scale asymptotically with distance or time. In addition to the analytical work above, we numerically investigate the relationships between the dimensions of fractal media and the random systems which sit upon them, and, in particular, provide evidence that universality should hold beyond the Euclidean setting. Material taken from a total of six papers is included. We also include an introduction explaining the background and context to these papers.

Preface

The following papers are reproduced in Chapter 2 of this thesis. All were conducted as equal collaborations.

[A] Lower Gaussian heat kernel bounds for the random conductance model in a degenerate ergodic environment
With Sebastian Andres. Published in *Stochastic Processes and their Applications* in 2021.

[B] Collisions of random walks in dynamic random environments
With Tom Hutchcroft. Published in *Electronic Journal of Probability* in 2022.

[C] What are the limits of universality?
With Tom Hutchcroft. Published in *Proceedings of the Royal Society A* in 2022.

[D] Most transient random walks have infinitely many cut times
With Tom Hutchcroft. Published in *Annals of Probability* in 2023.

[E] Logarithmic corrections to the Alexander-Orbach conjecture for the four-dimensional uniform spanning tree
With Tom Hutchcroft. Accepted pending minor revisions in *Communications in Mathematical Physics* in 2023.

[F] Uniqueness of the infinite tree in low-dimensional random forests
With Tom Hutchcroft. Accepted pending minor revisions in *Probability and Mathematical Physics* in 2023.

Acknowledgements

I would like to thank both my fantastic supervisors Sebastian Andres and Tom Hutchcroft for introducing me to the field of Probability which has yielded four fascinating, stimulating, productive and extremely enjoyable years. Thank you in particular to Tom for being an excellent mentor, guide, collaborator and friend. A special thanks also to Michelle at Caltech for looking after me in Pasadena, and rescuing me from my dire accommodation situation.

An anecdote related by George Pólya as to how he first became interested in the study of random walks:

> *At the hotel there lived also some students with whom I usually took my meals and had friendly relations. On a certain day one of them expected the visit of his fiancee, what I knew [sic], but I did not forsee that he and his fiancee would also set out for a stroll in the woods, and then suddenly I met them there. And then I met them there the same morning repeatedly. I don't remember how many times, but certainly it was too often and I felt embarrassed: It looked as if I was snooping around which was, I assure you, not the case.*

Table of contents

Chapter 1

Introduction

1.1 What is this thesis about?

This thesis is a collection of six papers about the geometry and topology of random systems, with random walks and random forests providing the primary objects of focus. Random walks and random forests are intimately connected to many other areas of statistical physics, and they also find multitudinous applications outside of mathematical physics [125], from finance [34, 138, 298] to chemistry [254, 309, 319] and biology [73, 105, 147]. As we shall see, the study of these two objects are also often intertwined with each other. All of the five analytical papers [A, B, D, E, F] will either have random walks as a primary focus or as an essential ingredient. Two of the analytical papers [E, F] will have random forests as a central object of study. The numerical paper [C] will also have some focus on a random forest model.

Of course, there are very many ways of defining random walks and random forests. We must first consider the environment or space in which the walk or forest resides. With random walks, we must then at each step specify the transition probabilities of the walker as a function of the environment. The environment can be random or deterministic, static or dynamic, and ordered or disordered. For forests, we must specify exactly how to distribute probability to the acyclic configurations: for instance, what subset of configurations is permitted? Are the permitted configurations uniformly weighted? How should we specify a distribution when the space or medium is infinite? In this thesis, we will consider random walks in a variety of different contexts with significantly varying levels of specificity, and we will examine three different types of random forest.

Once the model has been specified, we must choose an aspect of its behaviour we wish to investigate. The questions can be quantitative in nature, such as determining the exponents which govern the asymptotic relationships between various geometric quantities, or qualitative in nature, such as which almost sure topological properties these models satisfy. While, of course, these two types of questions are inextricably linked, we can say that the papers contained can be roughly divided equally between them. One particular type of question we can ask is how sensitive to the properties of a particular medium are the properties of the random system which sits upon it. This forms the subject of the numerical paper [C] in which we query the extent to which universality holds outside the Euclidean domain.

The thesis will be divided into two sections. The first, this introduction, will be used to introduce the various models which will be relevant to the latter part of this thesis. As we proceed, we will include some general discussion of the interesting and pertinent properties of these models and their relationships to each other, as well as some of the important tools which will be used later. We will also introduce the concept of *universality* in so far as it pertains to [C] and [E]. The second section is then a reproduction of the six papers (with modified formatting to fit thesis specifications).

1.2 Random Walks

In this section of the introduction we will explore the various random walk models which will be relevant to the latter part of this thesis. The majority of random walks we consider in this thesis will be Markovian in nature and so it is here that we begin.

1.2.1 Markov processes

The defining property of a Markov process is that it retains no memory of its past trajectory. In other words, if we know the present state of the process, its future is independent of its past. Formally, if $(X_n)_{n\geq 0}$, is a stochastic process taking values in some state space Ω, then (X_n) is a **Markov process** and satisfies the **Markov property** if

$$\forall n \geq 1,\ (X_i)_{i>n} \text{ is conditionally independent of } (X_i)_{i<n} \text{ given } X_n.$$

By allowing us to ignore the past, the Markov property can often greatly simplify the analysis of a random system. There is a vast literature on the general properties of Markov processes, see e.g. [284] for a broad introduction.

Aside from repeated use of the basic Markov property and its variants, there are two further crucial tools which will inform our analysis, and one builds upon the other. They involve connections between Markov chains and another fundamental probabilistic object: *the martingale*. Given a filtration $(\mathscr{F}_n)$, we recall that a **martingale** with respect to $(\mathscr{F}_n)$ is a stochastic process adapted to $(\mathscr{F}_n)$ such that

$$\forall n \geq 1,\ \mathbb{E}[M_n \mid \mathscr{F}_{n-1}] = M_{n-1}.$$

The first of these tools, many will be familiar with. We say that a Markov chain $(X_n)_{n\geq 0}$ has **transition matrix** $P = (p(x,y) : x,y \in \Omega)$ if for any $n \geq 1$ and sequence of states $u_0, \dots, u_n$, we have that

$$\mathbb{P}(X_1 = u_1, \dots, X_n = u_n \mid X_0 = u_0) = p(u_0,u_1)p(u_1,u_2)\cdots p(u_{n-1},u_n).$$

We write $p^{(n)}(x,y)$ for the Markov processes n-step transition probabilities, or in other words the entries of the matrix P^n. We will write I for the identity transition matrix $I(x,y) = \delta_{x,y}$, and for any function $f : \Omega \to \mathbf{R}$ we perform the operation Pf by treating f as a column vector. For any stochastic process $(X_n)_{n\geq 0}$, define its **natural filtration** $(\mathscr{F}_n)_{n\geq 0}$ to be filtration given by $\mathscr{F}_n = \sigma((X_i)_{i\leq n})$. We then have the following proposition.

Proposition 1. *Let $(X_n)_{n\geq 0}$ be a stochastic process with state space Ω, and for each $n \geq 0$, let $(\mathscr{F}_n)$ be the natural filtration of $(X_i)_{i\leq n}$. Then the following two statements are equivalent.*

- *$(X_n)_{n\geq 0}$ is a Markov process with transition matrix P,*
- *For all bounded functions f on E, the process $(M_n)_{n\geq 0}$ defined for each $n \geq 0$ by*

$$M_n = f(X_n) - f(X_0) - \sum_{k\leq n} (P-I)f(X_k)$$

is a martingale with respect to $(\mathscr{F}_n)_{n\geq 0}$.

In the case where f is a harmonic function, the summation on the right hand side is identically zero. This proposition allows us to control certain properties of Markov processes using the tools available for martingales, such as the optional stopping theorem [165, p. 491], and we use an extended version of this connection in [D].

The second tool which will play a role in a number of our papers is the use of *Markov type inequalities* which were introduced by Ball in 1992 [40]. We use these powerful tools to upper bounds the rate of escape of certain random walks. To introduce the relevant definitions, we will first recall the notions of stationarity and reversibility. We restrict to *irreducible* Markov processes with finite state space. We say a probability measure $\pi : \Omega \to [0,1]$ is a **stationary** distribution for the Markov process $(X_n)_{n\geq 0}$ if when we set X_0 to be distributed as π, the processes $(X_n)_{n\geq 0}$ and $(X_n)_{n\geq 1}$ have the same distribution. When this is the case, we say that the process itself is **stationary**, and observe that when (X_n) has transition matrix P, π being a stationary distribution is equivalent to $P^T\pi = \pi$. We say the Markov process $(X_n)_{n\geq 0}$

is **reversible** if for any $N > 0$, the processes $(X_0, \ldots, X_N)$ and $(X_N, \ldots, X_0)$ are identically distributed. If $(X_n)_{n\geq 0}$ is reversible, then one can show that it must also be stationary. It is also easy to check that if $(X_n)_{n\geq 0}$ has transition matrix P, then reversibility is equivalent to the existence of a probability distribution $\pi : \Omega \to [0,1]$ such that for all $x, y \in \Omega$

$$\pi(x)p(x,y) = \pi(y)p(y,x).$$

These equations are known as the **detailed-balance equations**, and when they hold, π must in fact be the stationary distribution of the process.

We say that a metric space (M,d) has **Markov type** 2 if there exists a constant $C < \infty$ such that for every stationary, reversible Markov chain $(X_n)_{n\geq 0}$ on a finite state space Ω, and every function $f : \Omega \to M$, and every $n \geq 0$, we have

$$\mathbb{E}\left[d(f(X_n), f(X_0))^2\right] \leq Cn\mathbb{E}\left[d(f(X_1), f(X_0))^2\right].$$

We say that (M,d) has **maximal Markov type** 2 if the above holds but with $d(f(X_n), f(X_0))^2$ replaced with $\max_{0\leq i\leq n} d(f(X_i), f(X_0))^2$ on the left hand side. At first glance it seems that either of these properties would be very difficult for a metric space to satisfy, as it must hold simultaneously for all Markov chains on any finite state space, and all functions from that state space to the metric space. However, it turns out that $\mathbb{R}$ is in fact maximal Markov type 2 [281].

Proposition 2. *Let $(X_n)_{n\geq 0}$ be a stationary reversible Markov process on a finite state space Ω, and let $f : \Omega \to \mathbb{R}$. Then for every $n \geq 0$, we have*

$$\mathbb{E}\left[\max_{0\leq i\leq n}(f(X_i) - f(X_0))^2\right] \leq 7n\mathbb{E}\left[\max_{0\leq i\leq n}(f(X_1) - f(X_0))^2\right].$$

The proof hinges on a particular martingale decomposition of $(f(X_n))_{n\geq 0}$ introduced in [263]. We fix $N \geq 0$ and a bounded function $f : \Omega \to \mathbb{R}$ and define $F : \Omega \to \mathbb{R}$ by $F(x) = (P-I)f(x)$. Proposition 1 gives us that

$$M_n := f(X_n) - f(X_0) - \sum_{i\leq n-1} F(X_i)$$

is a martingale with respect to the natural filtration of (X_n). However, defining $\widehat{X}_n = X_{N-n}$ for $n \leq N$ and applying Proposition 1 to $(\hat{X}_n)$ gives us that

$$\widehat{M}_n := f(\widehat{X}_n) - f(\widehat{X}_0) - \sum_{i\leq n-1} F(\widehat{X}_i)$$

is again a martingale, this time with respect to the natural filtration of $(\widehat{X}_n)$. By substituting in the relevant definitions on the right hand side and simplifying, we get that

$$f(Z_n) - f(Z_0) = \frac{1}{2}\big(M_n + \widehat{M}_{N-n} - \widehat{M}_N + F(X_0) - F(X_n)\big),$$

and so taking absolute values and maxima yields

$$2 \max_{n \le N} |f(Z_n) - f(Z_0)| \le \max_{n \le N} |M_n| + \max_{n \le N} |\widehat{M}_n| + |\widehat{M}_N| + |F(X_0)| + \max_{1 \le i \le n} |F(X_n)|.$$

While there are still some details to work out, the core of the rest of the proof is take L^2 norms and to apply Doob's L^2-maximal inequality [165, p. 497] to the martingales on the right hand side.

Of course, given that $\mathbb{R}$ is Markov type 2, so are $\mathbb{R}^d$ for $d \ge 1$, and we use this in [B, F] to prove a diffusivity result for certain random walks which embed into Euclidean space. It has been shown that trees and planar graphs also have Markov type 2 [281], and we shall see in [E] that by applying the Markov type inequality with a suitable metric space and appropriately chosen function, we can even obtain tight subdiffusive upper bounds on displacement for certain random walks.

1.2.2 Random walks on graphs and networks

Simple random walks on graphs are perhaps the most studied of all random walk models. A graph is a set of vertices, or points, together with a set of edges which specify which pairs of these points we should consider neighbours. Formally, we write $G = (V[G], E[G]) = (V, E)$, where V is some set of vertices, and $E \subseteq \{\{u, v\} : u, v \in V\}$ is the graph's set of edges. For $u, v \in V$, we write $u \sim v$ if the vertices u and v are **neighbouring** in G, i.e. $\{u, v\} \in E$, and we write $\deg(u)$ for the number of neighbours of the vertex u in G. We say the graph is **locally finite** if $\deg(u) < \infty$ for each $u \in V[G]$, and say the graph is **connected** if for any two vertices we can find a sequence of edges which connect them, i.e. for any $u, v \in V[G]$ there exists $n \ge 1$ and a sequence of vertices $u = u_0, \dots, u_n = v$ such that for each $i < n$, $u_i \sim u_{i+1}$.

Given a locally finite graph G and some starting vertex $u \in V[G]$, we define the **simple random walk** on G starting at u as the process $(X_n)_{n \ge 0}$ with $X_0 = u$ which at each time step $n \ge 0$ chooses one of the edges incident to its current vertex uniformly at random and independently from all previous time steps, and travels along this edge to the vertex at its other end. More precisely, this is the Markov process with state space $V[G]$ and transition matrix

$$p(u, v) = \frac{\mathbb{1}(u \sim v)}{\deg(u)}.$$

We can generalise this definition to allow for more flexible random walk behaviour by assigning a *conductance* $c(e)$ to each edge $e \in E[G]$. We call a triple (V, E, c) a **network**, and define the random walk $(X_n)_{n \geq 0}$ thereon to again be a Markov process, but this time with transition probabilities

$$p(u, v) = \mathbb{1}(u \sim v) \frac{c(\{u, v\})}{\sum_{w \sim u} c(\{u, w\})},$$

and so the higher an edge's conductance, the more likely the random walk will traverse it. We note that setting the conductance of each edge equal to 1, we recover the simple random walk on the graph (V, E), and that random walks on networks are in exact correspondence with countable Markov chains which satisfy the detailed balance equations for some measure π which need not be a probability measure.

1.2.3 Topological properties of random walk paths

One of the most fundamental questions we can ask about a simple random walk on a connected graph G is whether G is *recurrent* or *transient*. We say the graph is **recurrent** if a random walk on G returns to its starting vertex infinitely many times almost surely, and say it is **transient** otherwise. Implicit in this definition is the easily demonstrated fact that this property is independent of the chosen starting vertex. It is also simple to show that if the walk is transient, then the probability of infinitely many returns is zero. Graphs with a finite vertex set must of course be recurrent, but when the graph is infinite, the question of recurrence vs transience is often non-obvious. Recurrence or transience of Euclidean lattices was in fact one of the first questions asked about random walks on graphs. We define the **Euclidean lattice** with dimension $d \geq 1$ as the graph $G = (\mathbb{Z}^d, E_d)$ where the edge set E_d consists of all the nearest-neighbour edges of V, or equivalently all pairs of vertices in $\mathbb{Z}^d$ with ℓ_1 distance equal to 1. In 1921 George Pólya proved the following theorem [294].

Theorem 3. *A simple random walk on a the d-dimensional Euclidean lattice $\mathbb{Z}^d$ is recurrent in dimensions $d \leq 2$, and transient in dimensions $d \geq 3$.*

To demonstrate this, he utilized the easily proved criterion that a random walk on a connected graph is recurrent if and only if $\sum_n p^{(n)}(0,0) = \infty$ together with a computation that $p^{(n)}(0,0) \asymp n^{-d/2}$ for n even, and, trivially, $p^{(n)}(0,0) = 0$ for n odd.

One can think of the properties of transience and recurrence as statements about the *topology* of the random walk path $(X_n)_{n \geq 0}$. But recurrence/transience are by no means the only topological properties of random walk paths one could consider. Indeed, as indicated by the anecdote at the beginning of this thesis, Pólya first became interested in random walks when considering the *collisions* of random walk paths, a distinct topological property concerning a *pair* of walks. We say a graph G satisfies the **infinite collisions property** if two independent simple random walks X and Y, both started at the same vertex $u \in V[G]$ collide infinitely many times almost surely, i.e. there almost surely exist infinitely many times $n \geq 0$ such that $X_n = Y_n$. This property forms the subject of [B], and while it is in general by no means equivalent to recurrence [225], these two properties are equivalent on transitive graphs. To see this, we note the following general relation for random walks on graphs:

$$p^{(2n)}(v,v) = \sum_u p^{(n)}(v,u)^2 \frac{\deg(v)}{\deg(u)}.$$

In the transitive case, the quotient on the right hand side disappears, and so summing over $n \geq 0$, we get that the expected number of collisions between two independent random walks is equal to $\sum_{n \geq 0} p^{(2n)}(v,v)$, which is infinite for recurrent random walks and finite for transient random walks. We observe that transitivity implies that the number of collisions must be a geometric random variable and the equivalence between recurrence and the infinite collisions property becomes clear. Therefore, two independent random walks on $\mathbb{Z}^d$ have infinitely many collisions almost surely in $d \leq 2$ and only finitely many in $d \geq 3$. We can relate the number of collisions to the number of returns to an origin even more directly on $\mathbb{Z}^d$ by observing that the difference between two independent simple random walkers is itself a random walk but on a new lattice.

Another topological property which will be important to us in [E, F] is that of the *intersections* of random walks paths, where the question of interest is whether two independent random walks satisfy the **infinite intersections property**, i.e. whether their paths intersect infinitely many times almost surely, where an intersection between walks X and Y is a pair of times (n,m) such that $X_n = Y_m$. While collisions and intersections sound superficially similar, and while the former are a subset of the latter, the two properties can have very different behaviours. For instance, Erdős and Taylor [134] who initiated the study of such intersections proved that two random walks intersect infinitely often on $\mathbb{Z}^d$ for $d \leq 4$, and finitely often in $d \geq 5$, so the phase-transition occurs between dimensions 4 and 5 rather than 2 and 3 as with collisions. Unlike the infinite collisions property, recurrence is sufficient for the infinite-intersections property to hold. As we shall see, analysis of the intersections of random walks is crucial to analysis of the uniform spanning forest model (which we will introduce later).

A final topological property of random walks we will address is that of cut times, whose study was also initiated by Erdős and Taylor in 1960 [134]. We say that the random walk $(X_n)_{n \geq 0}$ has a **cut time** at time m if the sets $\{X_n\}_{n \leq m}$ and $\{X_n\}_{n > m}$ are disjoint. We will be interested in whether a random walk has infinitely or finitely many cut times almost surely.

Unlike the previous two properties, this property relates to a single random walk rather than a pair, and can be thought of as a measure of transience rather than recurrence. Indeed, transience is clearly a necessary condition for a random walk to have infinitely many cut times, and on finitely generated Cayley graphs the two properties are equivalent (although unlike for the infinite collisions property and recurrence, this equivalence is entirely non-obvious) [83, 134, 202, 238]. Moreover, when the graph is transient, the *expected* number of cut times is always infinite [64].

1.2.4 Geometric properties of random walk paths

We now shift focus from the more qualitative topological properties we have just discussed to some of the more quantitative ways to characterise the behaviour of random walks. One property we already touched upon in our discussion of Markov chains is the *rate of escape* of the random walk: what displacement of the walk from its starting location should we expect after a given amount of time. Of course there are many ways of measuring and delimiting this displacement. First of all, we must choose a metric. For an abstract graph, the graph distance metric is the obvious choice, but for graphs embedded in a substrate such as the supercritical percolation cluster of Euclidean lattices, we also have the choice of the Euclidean distance from the origin. There can be very large distortion between these two measures of distance which we call the *intrinsic* and *extrinsic* displacement respectively . We must also choose which statistic of the displacement we are interested in. With Markov type inequalities, we bounded the expected squared displacement of the walk, but we can also analyse the almost sure asymptotic behaviour. For instance, for a random walk $(X_n)_{n\geq 0}$ on a k-regular tree, we can prove by comparison of the walk to a biased walk on the integers that

$$\lim_{n\to\infty}\frac{d(X_0,X_n)}{n}=\lim_{n\to\infty}\frac{\mathbb{E}\left[d(X_0,X_n)\right]}{n}=\frac{k-2}{k}\qquad\text{almost surely.}$$

We say that a random walk is *ballistic* if its typical displacement from its starting point grows linearly and we say it is *diffusive* if its typical displacement at time n grows like $\sqrt{n}$, with *super-* and *sub-diffusive* referring to rates of growth which are faster or slower than this. Random walks on $\mathbb{Z}^d$ are diffusive in all dimensions $d\geq 1$.

In [D], we prove that slightly super-diffusive walks on networks have infinitely many cut times. This result is derived from a more general criterion which allows us to prove the existence of infinitely many cut times through control (via the Greens function) of the transition probabilities of the random walk.

The collection of transition probabilities of a random walk is also known as its *heat kernel*, and analysis of the asymptotic properties of the heat kernel is essential to understanding the behaviour of random walks. We have already seen that convergence/divergence of the sum of the return probabilities $p^{(n)}(0,0)$, which are collectively known as the *on-diagonal heat kernel*, determines recurrence or transience, and that the on-diagonal heat-kernel together with the stationary distribution determine the expected number of collisions of two independent random walks starting at the same vertex of a graph. It can easily be seen that the same holds true for the expected number of intersections of two such walks. Additionally, in [D], we prove that for irreducible countable Markov processes, if

$$p^{(n)}(x,x) = O(n^{-d/2})$$

for some $d > 2$, and for some, and therefore every element x of the state space, then the process has infinitely many cut times almost surely. When it exists, the limit

$$d_s = \lim_n \frac{-2\log p^{(n)}(x,x)}{\log n}$$

is known as the **spectral dimension** on the walk.

The transition probabilities $p^{(n)}(x,x)$ of a random walk on a graph may also decay *exponentially fast*, in which case $d_s = \infty$ and we say that the graph is **non-amenable**. Given a connected graph G, we define its **spectral radius**

$$\rho(G) = \lim_n p_{2n}(x,x)^{1/2n},$$

where we can show that this limit is well defined and independent of the choice of vertex $x \in V[G]$. The graph is then non-amenable if and only if $\rho(G) < 1$. When the graph is non-amenable and has bounded vertex degrees, then one can show that the random walk must be ballistic [259, Proposition 6.9]. We will not elucidate further here, but the spectral properties of a graph and in particular the spectral dimension and radius are strongly related to the graph's *isoperimetric profile* [277].

We can also consider the behaviour of the heat-kernel in the *off-diagonal* regime, that is $p^{(n)}(x,y)$ where $y \neq x$. In this regime we have a powerful very general upper bound, namely the Varopoulos-Carne bound [99, 316].

Theorem 4 (Varopoulos-Carne 1985)**.** *Let G be a graph, then*

$$p^{(n)}(u,v) \leq 2\sqrt{\frac{\deg(u)}{\deg(v)}} \exp\left[-\frac{d(u,v)^2}{2n}\right]$$

for every $u,v \in V[G]$ and every $n \geq 0$, where d is the graph metric on G.

See [99] for the extremely clever proof by Carne using Chebyshev polynomials. In [259] the inequality is further refined with an additional factor of $\rho(G)^n$ on the right hand side. We will use a version of this inequality in [D] to help us bound objects of the form $p_n(X_m, 0)$ under a hypothesis on the rate of escape of the random walk (X_m). Here we use a rate of escape to bound certain transition probabilities, but Varopoulos-Carne is often applied to prove results in the opposite direction. For example one can use it to show that random walks on graphs of subexponential growth cannot be ballistic, and for random walks on graphs of polynomial growth there always exists a random variable C such $d(X_0, X_n) \leq C\sqrt{n \log n}$ for all $n \geq 0$. In [A], we will be interested in having more precise control of the heat kernel, and will prove full lower Gaussian heat kernel bounds for certain classes of random walks on random networks. We leave discussion of this until our introduction of the random conductance model. Other quantitative aspects of the behaviour of random walks we will study will include the cardinality of the trace $\#\{X_i : 0 \leq i \leq N\}$ and bounds on the exit times of certain balls $\tau_r = \inf\{n \geq 0 : d(X_0, X_n) = r\}$.

1.2.5 Random walks in random environments

Up to now we have considered Markov random walks on deterministic and static graphs and networks. In this section we will look at some of the ways in which the environment can instead be random and dynamic. In this case one can think of first sampling the environment from some distribution and then sampling a random walk in this sampled environment. We can then ask questions about both the 'quenched' or 'annealed' behaviour of the random walk. Roughly these terms refer to properties which hold almost surely across all environments, and those which refer to the behaviour of the walk when averaged over the environment.

Bernoulli bond percolation. We begin by introducing one of the most commonly studied random modifications of a graph: Bernoulli bond percolation. While none of our papers study random walks on percolation, the model provides an accessible starting point for random graphs, is a precursor to the *dynamical percolation* model which is covered by one of our papers [B], and also appears as one of the two models we study numerically in [C]. We take some initial graph G and label each edge in $E[G]$ 'open' with some fixed probability $p \in [0,1]$ independently of all other edges. All edges which are not open are labelled 'closed'. The connected components of the random subgraph of G induced by the open edges are termed the percolation clusters. It is well known that for any every $d \geq 2$, there exists a critical probability $p_c = p_c(d) \in (0,1)$ such that for $p > p_c$, Bernoulli bond percolation on $\mathbb{Z}^d$ contains a unique infinite connected cluster almost surely, and for $p < p_c$ it contains no infinite connected cluster. We call $p > p_c$ the *supercritical* regime, and much work has been done characterising the behaviour of the random walk on the supercritical percolation cluster. Indeed, the random walk on the supercritical percolation cluster of $\mathbb{Z}^d$ was in fact the first *random walk in random environment* model studied, conceived by De Gennes in the 1970s [113] as the problem of 'the ant in the labyrinth'. The random walk is started at some vertex of the supercritical infinite cluster, and at each time step selects one of the *open* edges adjacent to it uniformly at random and traverses it to a new position. See [41, 74, 166, 223, 268] for a few of the highlights of the literature on random walks on percolation clusters.

The random conductance model. The random walk on the *random conductance model* follows a natural generalisation of this scheme. Instead of labelling each edge of the d-dimensional Euclidean lattice $G = (\mathbb{Z}^d, E_d)$ open or closed, we generate a random network from G by assigning a random conductance $c(e)$ to each edge $e \in E[G]$. More formally, we define the measurable space $\Omega = ([0,\infty)^{E_d}, \mathscr{B}([0,\infty])^{\otimes E_d})$ and let $\mathbb{P}$ be some measure on Ω. We now define a random walk on this random network. Rather than working directly with the discrete time random walk, we define a continuous time random walk which has the discrete time walk at its jump chain, i.e. shares the sequence of transitions. In particular, we must define how long the random walk spends at each vertex before transitioning. One natural choice leads to what is known as the *variable-speed random walk*. For each vertex $v \in \mathbb{Z}^d$, we let $c(v)$ denote the sum of the random conductances of the edges incident to v. Then each time the random walk is at vertex v, it waits an independent exponentially distributed amount of time with rate $c(v)$ before transitioning to its next vertex. We can construct this walk by attaching a Poissonian clock rate $c(e)$ to each edge of the graph. When the clock on edge e ticks, and the random walk is at one of the endpoints of e, it transitions to the other end point. Another natural choice of 'Poissonization' leads to what is known as the *constant-speed random walk*. This random walk waits again according to independently drawn exponential random variables, but this time they have a fixed constant rate. In other words, conditional on the conductances c, the walk has infinitesimal generator:

$$[\mathscr{L}f](x) = \frac{1}{c(x)} \sum_{y\sim x} c(x,y)\big(f(y) - f(x)\big),$$

where $c(x) = \sum_{y\sim x} c(x,y)$. It is this constant speed random walk which we study in [A].

Many of the questions asked about the random walk on the random conductance model of $\mathbb{Z}^d$ address under which constraints we can expect the large-scale and large-time behaviour of the walk to look like that of the simple random walk on $\mathbb{Z}^d$. For instance, there is a large collection of literature developing quenched functional central limit theorems, see e.g. [15]. Another central focus is the proof of full Gaussian heat-kernel bounds. That is, bounds of the form

$$ct^{-d/2}\exp\left[-d(x,y)^2/2t\right] \leq p^{(t)}(x,y) \leq Ct^{-d/2}\exp\left[-d(x,y)^2/2t\right],$$

for $t > 0$ and $x,y \in \mathbb{Z}^d$, where $p^{(t)}(x,y)$ are now the continuous time transition probabilities. These estimates are often derived under the assumption of *ergodicity* of the environment together with bounds, be they deterministic or probabilistic, on the conductances. When the conductances are almost surely bounded from above and below by finite and non-zero positive constants respectively, the model is known as *uniformly elliptic*. One well-studied relaxation of this constraint to allow the conductances to take values in $(0,\infty)$ under constraints on their moments and those of their reciprocals. Without these constraints, anomalous behaviour is a possibility [81].

In [18], Gaussian upper heat kernel bounds for the random conductance model are proved under moment conditions on the conductances and their reciprocals. In [A] we prove matching lower Gaussian bounds under moment conditions and one of a number of additional assumptions on the decay of correlations.

The dynamic random conductance model. A natural generalisation of the random conductance model is to allow the conductances to vary in time. In particular, on $\mathbb{Z}^d$, the state space takes the form $\Omega = [0,\infty)^{E_d} \times [0,\infty)$, where the final coordinate represents time. One simple example of the dynamic random conductance model is dynamical percolation. We choose an update rate $\Lambda > 0$, and a percolation probability $p \in [0,1]$. We then attach attach mutually independent Poisson processes to each edge each with rate Λ. At each 'arrival' of the process attached to a particular edge, we draw an independent Bernoulli random variable parameter p, and set the state of the edge according to this variable. Static Bernoulli bond percolation parameter p is the stationary distribution for this process. Random walks on dynamical percolation were first studied in [292] by Peres, Stauffer and Steif. They proved results on mixing and hitting times and displacement of the walk, and showed that the recurrence/transience criterion for simple random walks on $\mathbb{Z}^d$ extend to this model for all $\Lambda, p > 0$. See also [22, 184, 291].

For general dynamic conductance models limit theorems have also been proved under various restrictions on the conductances [10, 12, 14, 82]. In [B] we show that under stationarity, reversibility and a second moment condition, two independent random walks on a dynamic conductance model on $\mathbb{Z}^2$ will collide infinitely often almost surely. The proof utilises Markov type inequalities as well as the unimodularity of $\mathbb{Z}^2$; we briefly recap unimodularity below.

Random rooted graphs and unimodularity. We now return to static random graphs and briefly elucidate a property which will be relevant to multiple of the papers contained in this thesis: unimodularity. We begin by setting up the space of rooted graphs. We define a **rooted graph** to be a pair (G,ρ), where G is a countable locally finite connected graph, and ρ is a vertex in $V[G]$ which we call the **root**. We say two such rooted graphs are **equivalent** if there exists a graph isomorphism from one to the other which sends the root of the first to the root of the second. We let $\mathscr{G}^\bullet$ be the space of equivalence classes of rooted graphs and endow it with the Borel sigma algebra induced by the **local topology** which is roughly defined by saying that two rooted graphs are close if there exist large graph distance balls around there respective roots which admit a graph isomorphism which preserves the root. See [110] for a more precise definition. The space of doubly rooted graphs $\mathscr{G}^{\bullet\bullet}$ is defined similarly. We call a random variable taking values in $\mathscr{G}^\bullet$ a **random rooted graph**.

We say a random rooted graph (G,ρ) is **unimodular** if it satisfies the **mass-transport principle**, that for every Borel function $f:\mathscr{G}^{\bullet\bullet}\to[0,\infty)$,

$$\mathbb{E}\left[\sum_{v\in V[G]} f(G,\rho,v)\right]=\mathbb{E}\left[\sum_{v\in V[G]} f(G,v,\rho)\right].$$

We can think of this as a spatial homogeneity condition and as a generalisation of 'the root being uniformly distributed on the vertex set' from finite to infinite graphs. Indeed, finite graphs with a uniformly chosen root are trivially unimodular.

While all Cayley graphs are unimodular, there are transitive graphs which are not [259, p. 276]. We will see later that unimodular random graphs satisfy a multitude of useful properties. Some of the most elementary ones are that the unimodularity is preserved under weak limits, and under certain 'local modifications' which do not depend on a basis point, e.g. performing Bernoulli bond-percolation and then taking the cluster of the origin. Translation-invariant random subgraphs of $\mathbb{Z}^d$ will always be unimodular. In [195] it is shown that recurrent unimodular random graphs have the infinite collisions property. Inspired by this, in [E], we show that unimodular random subgraphs of $\mathbb{Z}^d$ for $d\le 4$ have the infinite intersections property. See [71] for the paper by Benjamini and Schramm in which the study of unimodular random graphs was introduced. Here they also prove that the weak (a.k.a. 'Benjamini-Schramm' or 'local') limits of finite planar random rooted graphs with uniformly bounded degree, and with the root uniformly distributed on the vertex set must be recurrent. See [8] for an in-depth look at the properties of certain stochastic systems on unimodular random graphs, and see [110] for lecture notes delivering an introduction to some of their more elementary properties.

1.3 Random Forests

In this section we introduce the three random forest models which will feature in the second part of this thesis. These are the uniform spanning forest, the arboreal gas and lattice trees. The last of these is explored numerically in [C], while the first two are explored analytically in [E] and [F] respectively. The models all have relatively simple definitions, but differ vastly in their analytical tractability and the methods which are used to study them. We begin with a discussion of the best studied and most tractable of these: the uniform spanning forest.

1.3.1 The uniform spanning forest

A **spanning tree** of a connected graph G is any acyclic connected subgraph of G which contains all of its vertices; any connected graph can be shown to have at least one such spanning tree. When such a graph G is finite then there are only finitely many spanning trees and so we can choose one of them uniformly at random. The resultant object is known as the **uniform spanning tree (UST)** of G. The **uniform spanning forests** of an infinite graph G then refers to the weak limits of the uniform spanning tree on exhaustions of G; we will define these in more detail later.

The uniform spanning forest is closely related to many other important models in statistical physics such as potential theory [68, 93], the random cluster model [164, 198], domino tilings [215], and the Abelian sandpile model [124, 207]. We shall see in [F] that the UST can also play a role in the analysis of the arboreal gas. Most important to us, however, will be its deep connection to the theory of random walks. Kirchoff, who initiated the study of UST [222], proved that the ratio of the number of spanning trees containing any particular edge to the total number of spanning tree is equal to the *effective resistance* across the edge. Rewriting this probabilistically and expressing the effective resistance in terms of random walks we have: if G is a finite connected graph, $e = \{u, v\} \in E[G]$ is some edge in G, T denotes the uniform spanning tree of G, and $(X_n)_{n\geq 0}$ is a simple random walk on G starting at u, we have

$$\mathbb{P}(e \in T) = \mathbb{P}((X_n)_{n>0} \text{ hits } v \text{ before } u).$$

It tuns out that the uniform spanning tree of a finite graph can actually be constructed from simple random walk paths via what is known as *Wilson's algorithm* [321]. To expound this we must first define the *loop-erasure* of a random walk, which was introduced by Lawler in [230]. For any $0 \le n \le \infty$ and any nearest-neighbour path $w_0, \dots, w_n$ in G which visits each vertex of G finitely many times, we recursively define the sequence of times $\ell_n(w)$ by $\ell_0(w) = 0$, and

$$\ell_{n+1}(w) = 1 + \max\{k : w_k = w_{\ell_n}\},$$

where we terminate the sequence the first time $\max\{k : w_k = w_{\ell_n}\} = m$ when $m < \infty$. The **loop-erasure** of w is then the path induced by the sequence of neighbouring vertices

$$\mathrm{LE}(w)_i = w_{\ell_i(w)}.$$

Wilson's algorithm on a finite connected graph then proceeds as follows. Fix a vertex $\rho \in V[G]$, and fix an ordering $v_0, \dots, v_{|V[G]|-2}$ of the remaining vertices. Let $(X^v)_{v \in V[G]}$ be a collection of mutually independent random walks on G, with starting locations $X_0^v = v$. Define the sets $(S_i)_{i \ge 0}$ and the stopping times $(T_i)_{i \ge 0}$ recursively as follows:

$$S_0 = \{\rho\}; \qquad T_i = \inf\{n \ge 0 : X_n^{v_i} \in S_i\} - 1; \qquad S_{i+1} = S_i \cup \mathrm{LE}(X_{T_i}^{v_i}),$$

for $0 \le i \le |V[G]| - 2$. The random variable $S_{|V[G]|-1}$ is then distributed as the uniform spanning tree of G. We can in fact choose each vertex v_i dynamically as the algorithm progresses as a function of $(S_j)_{j \le i}$. If we run a single random walk $(X_n)_{n \ge 0}$ on G and record the *first-entry edges* of each vertex $v \ne X_0$, we again obtain a sample of uniform spanning tree of G. This algorithm is known as *Aldous-Broder* [7, 87].

Infinite graphs. When the graph is infinite, the set of spanning trees usually is as well and so there is no longer necessarily an obvious way to choose a spanning tree of the graph 'uniformly' at random. There are, however, well-defined infinite volume limits of the uniform spanning tree measure. Of these, there are two canonical variations, the *wired* and *free* uniform spanning forests. We define each of these measures for a connected locally finite infinite graph G. Let $(V_n)_{n \ge 0}$ be an exhaustion of $V[G]$ by finite connected sets and for each $n \ge 0$, define the induced subgraphs $G_n = G[V_n]$, then Pemantle [288] showed that the weak limit of the sequence of measures $(UST[G_n])_{n \ge 0}$ exists and is independent of the choice of exhaustion; we call the limiting measure the **free uniform spanning forest (FUSF)** of G. Now let for each $n \ge 0$, define the graph G_n^* by contracting all of the vertices in G outside of V_n into a single vertex ρ_n, and then deleting all of the self loops of ρ_n. Again, Pemantle showed that the weak limit of the sequence of measures $(UST[G_n^*])_{n \ge 0}$ exists and is independent of the choice of exhaustion, and we call the limiting measure the **wired uniform spanning forest (WUSF)** of G. While in general the free and wired uniform spanning forests are distinct measures, Pemantle [288] showed they coincide on $\mathbb{Z}^d$ for all $d \geq 1$. When they do coincide, we simply refer to the uniform spanning forest of the graph.

In general, the wired uniform spanning forest of a graph is more easily understood than its free counterpart. This is because we can, in fact, easily extend Wilson's algorithm to infinite graphs, and the resultant measure is that of the WUSF. For recurrent infinite graphs, the algorithm is unchanged, while for transient graphs the initial set S_0 defined above as containing a single vertex is instead made to be empty, with the resultant algorithm known as *Wilson's algorithm rooted at infinity* [68]. Thus if we can control the behaviour of the loop-erasure of the random walk on an infinite graph we can gain insight into the behaviour of the WUSF.

Connectivity. While it is trivial to check that all trees in the WUSF of an infinite connected graph must be infinite almost surely, finding the number of infinite trees in the forest is not so straightforward. One important consequence of the variants of Wilson's algorithm for infinite graphs is that they allow us to control the wired uniform spanning forest's connectivity. If we assume that for any two vertices $x, y \in G$ the paths $\mathrm{LE}(X)$ and Y intersect almost surely, where X and Y are two independent random walks starting at x and y respectively, then by Wilson's algorithm, the wired uniform spanning forest of G must be connected almost surely. In [68, 261] it is shown that if the two paths have finite intersection with positive probability the WUSF is disconnected almost surely, and if they have finite intersection almost surely, the WUSF has infinitely many components almost surely. We then immediately have that for recurrent graphs, the WUSF is connected. They also show that these connectivity statements remain true even when we replace the loop erasure of X with X itself. Combining this with the theorem of Erdős and Taylor [134] who showed that two independent random walks on $\mathbb{Z}^d$ will intersect infinitely often almost surely in $d \leq 4$ and finitely often almost surely in $d \geq 5$, we can recover the following result which was first proved by Pemantle (prior to the proof of Wilson's algorithm rooted at infinity) in [288].

Theorem 5. *The uniform spanning forest of $\mathbb{Z}^d$ is a single infinite tree almost surely in $d \leq 4$, and has infinitely many infinite trees almost surely in $d \geq 5$.*

In [F], we prove that the wired uniform spanning forests of unimodular random subgraphs of $\mathbb{Z}^d$ are also connected almost surely for $d \leq 4$.

There has been much working analysing further quantitative and qualitative aspects of the behaviour of the uniform spanning forest on $\mathbb{Z}^d$ and other graphs. See e.g. [8, 68, 189, 258, 258, 314] for results concerning the number of ends of uniform spanning forests under various conditions. See [66, 196] for results regarding the adjacency structure of the component trees of the UST. In [48, 76, 190], the geometry of the uniform spanning tree and the behaviour of a random walk thereon in the high-dimensional ($d \geq 5$) mean-field regime is analysed. In [197], the geometry of the *past* in the UST is analysed at the upper-critical dimension $d = 4$, and in [E] we analyse the geometry of balls and the behaviour of a random walk on the UST, also at the upper-critical dimension. Both of these papers are concerned with calculating the exact logarithmic correction to the mean-field scaling for quantities such as the volume of balls and the intrinsic and extrinsic displacements of a random walk on the tree. See Section 1.4 for more details. See [44, 45, 305] for analysis of the properties of random walk paths on the UST in dimensions 2 and 3 and see [2, 27, 44, 44, 185, 239, 300] for results concerning the scaling limits of the uniform spanning tree and the random walk on the uniform spanning tree in dimensions 2 and 3.

1.3.2 The arboreal gas

Next we define the arboreal gas which is also known as the **weighted uniform forest model**. A **spanning forest** of a G is any subgraph of G which contains all of its vertices and is acyclic. The β-arboreal gas A is then the random subgraph of G with probability mass function

$$\mathbb{P}_\beta(A=F) = \begin{cases} (1/Z_\beta)\beta^{|F|} & F \subseteq G \text{ is a spanning forest} \\ 0 & \text{otherwise} \end{cases}, \qquad Z_\beta = \sum_{F \subseteq G \text{ a spanning forest}} \beta^{|F|}, \tag{1.1}$$

where $|F|$ denotes the cardinality of the edge set of F.

Connections to other models. The arboreal gas is connected to a number of other models in statistical physics. First, we observe that the law of A is equal to the law of Bernoulli bond percolation on G with parameter $p = \beta/(1+\beta)$ conditioned to be acyclic. It is also equal to the $q \to 0$ limit of the q-state random cluster model with p/q converging to β [199, 289]. When the graph G is connected, the law converges to that of the uniform spanning tree in the $\beta \to \infty$ limit. In the specific case $\beta = 1$ the model is known as the uniform random spanning forest of G, not to be confused with the distinct previously discussed uniform spanning tree/forests. The arboreal gas is also closely related to various supersymmetric spin systems, which has led it to receive substantial attention in the physics literature [95–97, 117].

Basic properties. Compared to the uniform spanning model, the arboreal has very few tools to work with. To begin with, there are no known random walk type sampling algorithms for the arboreal gas. Negative dependence, whilst conjectured [168] around 20 years ago, is still not known to hold, i.e. we do not know whether for any finite graph G, any $\beta > 0$ and any distinct edges $e, f \in G$,

$$\mathbb{P}_\beta(e \in A \mid f \in A) \leq \mathbb{P}_\beta(e \in A).$$

Additionally, unlike Bernoulli bond percolation there is no stochastic monotonicity with respect to the parameter i.e. it is *not* the case that for any finite graph G and $\beta \leq \beta'$, $\mathbb{P}_\beta \preceq \mathbb{P}_{\beta'}$ [Remy Poudevigne, private correspondence]. We do however have the following stochastic domination by Bernoulli bond-percolation:

$$\mathbb{P}_\beta \preceq \mathrm{B.b.p}\Big(\frac{\beta}{1+\beta}\Big).$$

The arboreal gas on $\mathbb{Z}^d$. As with the uniform spanning tree, the finite graph definition of the arboreal gas breaks down for infinite graphs. Unlike the uniform spanning tree, however, the dearth of available tools means it is not currently known whether we can take weak limits across an arbitrary exhaustion. We can, of course, take subsequential limits by compactness with arbitrary boundary conditions, and in [F] we develop a new Gibbsian reformulation for analysing these limiting measures. While we know that these measures must be supported on forests it is not immediately obvious whether these forests contain any infinite trees. For small values of β, however, stochastic domination by Bernoulli-bond percolation eliminates the possibility on any such infinite components, but this tells us nothing about the behaviour for larger values of β. Recently, however, Bauerschmidt, Crawford Helmuth and Swan proved that for subsequential limits along boxes on $\mathbb{Z}^2$, the resultant measures are supported on forests containing *no* infinite trees for *all* $\beta > 0$ [55]. Conversely, Bauerschmidt, Crawford and Helmuth prove that in $d \geq 3$, for β large enough, subsequential limits along Tori contain *at least one* infinite tree almost surely [54]. The argument they use are in large part non-probabilistic relying on re-expression of the partition and two-point functions of the arboreal gas in terms of certain supersymmetric sigma models, with the $d = 2$ result following by a Mermin-Wagner type theorem. We complement the latter of these results by proving that in dimensions $d = 3, 4$, general translation-invariant Gibbs measures for the arboreal gas (which include subsequential limits along Tori) contain *at most* one infinite tree almost surely. This means for β large enough we have existence and uniqueness. We also prove that in all dimensions and $\beta > 0$, the trees in such Gibbs measures must be one-ended almost surely.

These results follow from a new resampling property for these Gibbs measures on $\mathbb{Z}^d$ under the condition of translation-invariance, which says that the infinite component of the arboreal gas has connected trace and can be resampled on its trace as the uniform spanning forest of that trace. We use this with a random walk argument to show that translation-invariant arboreal gas Gibbs measures on $\mathbb{Z}^d$ have at most one infinite tree in $d \leq 4$.

Of course this leaves open the question of the number of infinite trees in the arboreal gas in dimensions $d \geq 5$; in [F] we conjecture that there will almost surely be infinitely many. See [56, 310] for surveys of the model and its connections to other topics.

1.3.3 Lattice trees

The final forest model we introduce is the *lattice tree* which is defined as follows. Let G be a connected infinite locally finite graph and fix some origin vertex $o \in G$. For each $n \geq 1$, we denote by S_n the set of all trees in G which contain n vertices, one of which must be the origin. The **lattice tree** of size n is then defined the random variable V_n given by selecting one of these trees uniformly at random. The lattice tree, and the related lattice animal models have been utilised as natural models of a branched polymer in a dilute solution, and their analysis was pioneered by Lubensky and Isaacson in 1979 [254].

Like the arboreal gas, lattice trees are in general far less tractable than the uniform spanning tree model as there are, again, no random walk sampling algorithms available. On the other hand, if we take $G = \mathbb{Z}^d$ with $d \gg 1$, the lattice tree model is amenable to the technique of lace expansion which has allowed significant progress to be made analytically characterising its behaviour in the mean-field regime [307]. For example, in [179], Hara and Slade partially characterise the asymptotic growth rate of $|S_n|$ as well as the asymptotic behaviour of a particular geometric features of $(V_n)_{n\geq 1}$. Convergence of $(V_n)_{n\geq 0}$ to super-Brownian motion under suitable scaling has also been proven [118, 119]. Though it has not been rigorously demonstrated, there is strong evidence [178] that the upper-critical dimension of lattice trees is 8 (See Section 1.4 for an introduction to this concept), but the results above are generally for dimensions much larger than 8.

In the low dimensional regime which will be of the most interest to us in [C], these methods fail and there has been little analytical progress (although there has been progress in the related continuum polymer model [89, 216]). There has, however, been numerical work characterising the asymptotic behaviour of lattice trees in lower dimensions, e.g. [203, 323]. The quantities we are most interested in are the *mean branch size* $B(n)$, obtained by deleting an edge of the lattice tree uniformly at random and finding the cardinality of the smaller of the two components, and the *longest path length*, measured with respect to the extrinsic $E(n)$ and intrinsic $I(n)$ metrics. We expect these quantities to scale as

$$B(n) \asymp I(n) \asymp n^{\rho}, \qquad E(n) \asymp n^{\nu},$$

for some intrinsic exponents ν and some extrinsic exponent ρ which depend on the undelying lattice, and by generating large samples of V_n via a Monte Carlo Markov chain algorithm at varying values of n, we attempt to estimate the values of ν and ρ. In [C], we are interested in the behaviour of statistical physics models on (non-Euclidean) transitive lattices below the upper-critical dimensions, and the large $d = 8$ upper-critical dimension along with numerical tractability makes lattice trees an ideal candidate for analysis.

1.4 Universality

The final concept we must introduce is *universality*, and to do so, it is convenient to first discuss *critical exponents*. If we consider some critical statistical physics model on the Euclidean lattice $\mathbb{Z}^d$, the critical exponents describe how the behaviour of the model depends on the scale at which the system is observed. The properties of these systems often obey power-law scaling relations, and the critical exponent corresponding to each of these properties is defined to be the exponent which governs this power-law. We have already seen some examples of these in Section 1.3.3, for instance the exponents ρ and ν which govern the intrinsic and extrinsic aspects of the geometry of lattice trees. In [C], we will also consider, among others, the probability that the cluster of the origin K_o has volume at least n, which is governed by the asymptotic relation

$$\mathbb{P}(|K_o| \geq n) \sim n^{2-\tau},$$

where τ is known as the Fisher critical exponent. These critical exponents are in general functions of the dimension d of the underlying Euclidean lattice, but often, at least in low dimensions, they are extremely difficult to pin down exactly, and only approximations, analytic and numerical, are known.

For any given critical statistical physics model, there is often a dimension d_c known as the *upper-critical dimension* above which the model is considered *mean-field* and these functions plateau. For Bernoulli bond percolation this dimension is 6, for the UST it is 4 and for lattice trees it is (strongly expected to be) 8. At the upper-critical dimension itself, the power law scaling relations are expected to admit a logarithmic correction to their mean-field scaling. For instance, in [E], we show that the volume $V(n)$ of the ball of intrinsic radius n for the UST on $\mathbb{Z}^4$ satisfies

$$\mathbb{E}[|V(n)|] = \frac{n^2}{(\log n)^{1/3-o(1)}},$$

where 2 is the mean field exponent for this quantity, as derived in [48].

The term *universality* refers to the surprising phenomenon that these critical exponents are *invariant* under local perturbations of the geometry of the Euclidean lattice $\mathbb{Z}^d$: as long as the large scale geometry remains the same, we should expect the same critical exponents. For instance, we expect identical critical exponents for percolation on square, triangular and hexagonal lattices. This phenomenon is traditionally explained with the heuristic that these local details disappear under the *renormalization group flow* as these geometries all share the same $\mathbb{R}^d$ scaling limit.

In [C], we seek to understand whether this phenomenon extends *beyond* the Euclidean setting. While there is strong evidence that these exponents are invariant to changes in local geometry, is it the case that they are invariant under changes in large-scale geometry conditional on fixing the volume growth dimension? While for $d \leq 3$ it has been shown that the only transitive geometries with polynomial growth dimension d are Euclidean, in four dimensions and above, there are in fact multiple possible transitive geometries; for instance, the Cayley graph of the Heisenburg group is a four dimensional transitive lattice. These disparate transitive geometries are not rough isometric to Euclidean lattices and have distinct scaling limits, despite having the same volume growth dimension. We can therefore ask, for each $d \geq 4$, whether the critical exponents of any statistical physics model varies according to this choice of large-scale geometry. We provide strong numerical evidence that, perhaps surprisingly, the critical exponents of Bernoulli bond-percolation and lattice trees are shared across disparate geometries of the same polynomial growth dimension.

Transitive lattices are only one special type of fractal geometry and the second half of [C] concerns more general fractal geometries. In general, fractal geometries have multiple dimensions associated to them, such as their Hausdorff, spectral and topological dimensions, and while these dimensions all agree for Euclidean lattices, they may each take a different value on other fractal geometries. We show that no universality should be expected to hold for general fractals, even if we allow the critical exponents to be a function of a large number of these dimensions, thus refuting a conjecture of Balankin et al [37]. Indeed, we provide two fractals which share Hausdorff, topological, topological Hausdorff [39] and spectral dimension, but for which percolation exhibits differing critical exponents. See [137] for general background on fractals and their properties.

Another intriguing phenomena also related to the renormalisation group is that seemingly disparate statistical physics models may share critical exponents. We say that any two such models belong to the same *universality class*. One particularly important such class is the Alexander-Orbach class which has been proven to contain the *incipient infinite cluster* of critical high dimensional percolation [223], high dimensional oriented percolation [49], and the high dimensional UST [190]. In [E] we prove logarithmic corrections to mean-field behaviour for the last of these three models, which constitutes the first time such corrections have been rigorously computed for a random walk on a random fractal in this universality class.

Chapter 2

[A] Lower Gaussian heat kernel bounds for the Random Conductance Model in a degenerate ergodic environment

Abstract. We study the random conductance model on $\mathbb{Z}^d$ with ergodic, unbounded conductances. We prove a Gaussian lower bound on the heat kernel given a polynomial moment condition and some additional assumptions on the correlations of the conductances. The proof is based on the well-established chaining technique. We also obtain bounds on the Green's function.

2.1 Introduction

2.1.1 The Model

We let $G = (\mathbb{Z}^d, E_d)$, where $E_d = \{\{x,y\} \in \mathbb{Z}^d \times \mathbb{Z}^d : |x-y| = 1\}$, be the d-dimensional lattice for a fixed dimension $d \geq 2$. We write $x \sim y$ if $(x,y) \in E_d$. We consider the space of positive weightings on the edges of the graph, $\Omega = (0,\infty)^{E_d}$, and for $\omega \in \Omega$, we access the weight at a particular edge $e \in E_d$ by $\omega(e)$, which we will also refer to as the *conductance* on an edge e. For $x,y \in \mathbb{Z}^d$ and $\omega \in \Omega$ we set $\omega(x,y) = \omega(y,x) = \omega(\{x,y\})$ if $\{x,y\} \in E_d$, else $\omega(x,y) = 0$. For any fixed ω, we define measures μ^ω and ν^ω on $\mathbb{Z}^d$ by

$$\mu^\omega(x) := \sum_{y \sim x} \omega(x,y) \qquad \text{and} \qquad \nu^\omega(x) := \sum_{y \sim x} \frac{1}{\omega(x,y)}.$$

For any $z \in \mathbb{Z}^d$ we denote by $\tau_z\colon \Omega \to \Omega$ the space shift by z defined by

$$(\tau_z \omega)(\{x,y\}) := \omega(\{x+z,y+z\}), \qquad \forall \{x,y\} \in E_d.$$

We equip Ω with a σ-algebra $\mathscr{F}$. Further, we will denote by $\mathbb{P}$ a probability measure on $(\Omega,\mathscr{F})$, and we write $\mathbb{E}$ for the expectation with respect to $\mathbb{P}$. Throughout the paper we will assume that the conductances are stationary and ergodic.

Assumption 6 (Stationarity and ergodicity). $\mathbb{P}$ is stationary and ergodic with respect to translations of $\mathbb{Z}^d$, i.e. $\mathbb{P}\circ\tau_x^{-1} = \mathbb{P}$ for all $x \in \mathbb{Z}^d$ and $\mathbb{P}[A] \in \{0,1\}$ for any $A \in \mathscr{F}$ such that $\tau_x(A) = A$ for all $x \in \mathbb{Z}^d$.

We now introduce the *random conductance model (RCM)*. For a given $\omega \in \Omega$, we consider the continuous time Markov chain $X = \{X_t : t \geq 0\}$ on $\mathbb{Z}^d$ with generator

$$(\mathscr{L}^\omega f)(x) = \frac{1}{\mu^\omega(x)} \sum_{y\sim x} \omega(x,y)\big(f(y) - f(x)\big).$$

This stochastic process, also known as the *constant speed random walk* (CSRW), waits at x for an exponential time with mean 1, and then chooses the next position $y \sim x$ with probability $\omega(x,y)/\mu^\omega(x)$. We also recall that the Markov chain X is reversible with respect to μ^ω. We denote by P_x^ω the law of the walk starting at the vertex $x \in \mathbb{Z}^d$, and by E_x^ω the expectation with respect to this law. For $x,y \in \mathbb{Z}^d$ and $t > 0$, we let $p^\omega(t,x,y)$ be the transition density (or the heat kernel associated with $\mathscr{L}^\omega$) with respect to the measure μ^ω, i.e.

$$p^\omega(t,x,y) = \frac{\mathrm{P}_x^\omega[X_t = y]}{\mu^\omega(y)}.$$

2.1.2 Main Results

The random conductance model has been the subject of extensive research for more than a decade, see [77, 226] for surveys of the model and references therein. More recent results include the derivation of quenched functional central limit theorems [16, 57, 80, 121] and local limit theorems [14, 17, 24, 47, 58] for the RCM with unbounded ergodic conductances under moment conditions. In this paper we will focus on heat kernel estimates, see e.g. [18, 19, 41, 46, 47, 75, 79, 115, 141] for previous results. In particular, we will obtain Gaussian type lower bounds on the heat kernel in the case of ergodic unbounded conductances.

It is known that Gaussian bounds do not hold in general: for example, under i.i.d. conductances with fat tails at zero, the heat kernel decay may be sub-diffusive due to a trapping phenomenon – see [75, 79]. Moreover, in [17, Theorem 5.4], it is proved that in the general ergodic setting, moment bounds on the conductances and their reciprocals are a necessary condition for upper and lower near-diagonal Gaussian bounds to hold. In [18], this necessary condition is shown to be sufficient for full upper Gaussian heat kernel bounds.

Gaussian lower bounds have been shown on i.i.d. percolation clusters in [41], and for variable speed random walks under i.i.d. conductances bounded away from zero in [46]. However, in the general ergodic setting, as of yet, Gaussian lower bounds have only been proved under the stronger condition of uniformly elliptic conductances [115], i.e. $c^{-1} \leq \omega(e) \leq c$, $e \in E_d$, for some $c \geq 1$. In this paper we relax the uniform ellipticity assumption, substituting it for the combination of a polynomial moment condition together with an assumption concerning the correlations of the conductances, see Assumption 8. It is unknown whether moment conditions by themselves should be sufficient for the lower bound to hold. The main available technique for proving lower bounds, the *chaining method* (see [136]), fails at present in this generality (see Section 2.1.3 below for a more in-depth discussion), while our assumptions are sufficient to ensure the functionality of this method. However, it seems that other techniques would be required in order to weaken these assumptions. One possible approach would be to use techniques from quantitative stochastic homogenization that lead to much stronger quantitative homogenization results for heat kernels and Green functions, see [29, Chapters 8–9] for details. This technique has been adapted to Bernoulli bond percolation clusters in [112], and it is expected that it also applies to other degenerate models.

We will begin by recalling the already established Gaussian upper bound in [18], for which we will need some more notation. For $A \subset \mathbb{Z}^d$ non-empty and finite, and $p \in [1,\infty)$, we introduce space-averaged l^p norms on functions $\phi : A \to \mathbb{R}$ by

$$\|\phi\|_{p,A} := \left(\frac{1}{|A|} \sum_{x \in A} \left|\phi(x)\right|^p \right)^{\frac{1}{p}} \qquad \text{and} \qquad \|\phi\|_{\infty,A} := \max_{x \in A} |\phi(x)|,$$

where $|A|$ denotes the cardinality of the set A. For $x \in \mathbb{Z}^d$ we denote by $B(x,r) := \{y \in \mathbb{Z}^d : |x-y| < r\}$ balls in $\mathbb{Z}^d$ centered at x with respect to the graph distance, where $|x| := \sum_{i=1}^d |x_i|$ for $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$. Suppose now that $\omega(e) \in L^p(\mathbb{P})$ and $\omega(e)^{-1} \in L^q(\mathbb{P})$ for any $p, q \geq 1$. Then, under Assumption 6, the spatial ergodic theorem gives that, $\mathbb{P}$-a.s., for any $x \in \mathbb{Z}^d$,

$$\bar{\mu}_p := \mathbb{E}[\mu^\omega(0)^p] = \lim_{n \to \infty} \|\mu^\omega\|^p_{p,B(x,n)}, \qquad \bar{\nu}_q := \mathbb{E}[\nu^\omega(0)^q] = \lim_{n \to \infty} \|\nu^\omega\|^q_{q,B(x,n)}.$$

In particular, for $\mathbb{P}$-a.e. ω and each $x \in \mathbb{Z}^d$, there exists $N_1(x) = N_1(\omega, x, p, q) \in \mathbb{N}$ such that

$$\sup_{n \geq N_1(x)} \|\mu^\omega\|_{p,B(x,n)}^p \leq 2\bar{\mu}_p, \qquad \sup_{n \geq N_1(x)} \|\nu^\omega\|_{q,B(x,n)}^q \leq 2\bar{\nu}_q. \tag{2.1}$$

We will choose $N_1(x)$ to be the minimal such random variable. The Gaussian upper heat kernel bound is as follows:

Theorem 7. *Suppose that Assumption 6 holds and suppose there exist $p, q \in (1, \infty]$ with $1/p + 1/q < 2/d$ such that $\omega(e) \in L^p(\mathbb{P})$ and $\omega(e)^{-1} \in L^q(\mathbb{P})$ for any $e \in E_d$. Then, there exist constants $c_i = c_i(d, p, q, \bar{\mu}_p, \bar{\nu}_q)$ such that, for $\mathbb{P}$-a.e. ω, for any given t and x with $\sqrt{t} \geq N_1(x)$ and all $y \in \mathbb{Z}^d$ the following hold.*

(i) If $|x - y| \leq c_1 t$ then

$$p^\omega(t, x, y) \;\leq\; c_2 t^{-d/2} \exp\big(-c_3 |x-y|^2/t\big).$$

(ii) If $|x - y| \geq c_1 t$ then

$$p^\omega(t, x, y) \;\leq\; c_2 t^{-d/2} \exp\big(-c_4 |x-y|(1 \vee \log(|x-y|/t))\big).$$

Proof. See [18, Theorem 1.6] and a more general version with a streamlined proof in [19, Theorem 3.2]. □

We now state the additional assumptions we require, followed by our main results. We will then discuss why these additional assumptions are needed and how they interact with the strategy of the proof.

For $\omega, \omega' \in \Omega$ we write $\omega \leq \omega'$ if $\omega(e) \leq \omega'(e)$ for all $e \in E_d$. We say that a function $f : \Omega \to \mathbb{R}$ is non-decreasing if $f(\omega) \leq f(\omega')$ whenever $\omega \leq \omega'$.

Assumption 8. At least one of the following four conditions holds.

(A1) (i) *FKG inequality.* For any finite set of edges $A \subset E_d$, and any non-decreasing functions $f, g : \Omega \to \mathbb{R}$ depending only on $\{\omega(e) : e \in A\}$, we have

$$\mathbb{C}\mathrm{ov}(f, g) \geq 0, \tag{2.2}$$

whenever the covariance exists.

(ii) *Polynomial mixing.* There exist constants $\gamma > d^2$ and $C_{\mathrm{mix}} \in (0, \infty)$ such that for any non-decreasing function $f \in L^2(\mathbb{P})$ depending only on $\{\omega(0, y), |y| = 1\}$, and any

$x \in \mathbb{Z}^d \setminus \{0\}$,

$$\mathbb{C}\mathrm{ov}\left(f, f \circ \tau_x\right) \leq C_{\mathrm{mix}} \|f\|_{L^2(\mathbb{P})}^2 |x|^{-\gamma}.$$

(A2) *Spectral gap.* There exists $C_{\mathrm{sg}} \in (0, \infty)$ such that

$$\mathbb{E}\left[(f - \mathbb{E}[f])^2\right] \leq C_{\mathrm{sg}} \sum_{e \in E_d} \mathbb{E}\left[\left(\partial_e f\right)^2\right], \tag{2.3}$$

for any $f \in L^2(\mathbb{P})$. Here, the 'vertical derivative' $\partial_e f$ is defined as

$$\partial_e f(\omega) := \limsup_{h \to 0} \frac{f(\omega + h\delta_e) - f(\omega)}{h},$$

where $\delta_e : E_d \to \{0, 1\}$ stands for the Dirac function satisfying $\delta_e(e) = 1$ and $\delta_e(e') = 0$ if $e' \neq e$.

(A3) *Finite range dependence.* There exists a positive constant $\mathfrak{R} \in \mathbb{N}$, such that for any $x \in \mathbb{Z}^d$, the collection of random variables $\big(\omega(\{x, x+e\}) : |e| = 1\big)$ is independent of $\big(\omega(\{z, z+e\}) : |z - x| \geq \mathfrak{R}, |e| = 1\big)$.

(A4) *Negative association.* For any two disjoint finite sets of edges $A, B \subset E_d$, and any non-decreasing functions $f, g : \Omega \to \mathbb{R}$ depending only on $\{\omega(e) : e \in A\}$ and $\{\omega(e) : e \in B\}$ respectively, we have

$$\mathbb{C}\mathrm{ov}(f, g) \leq 0,$$

whenever the covariance exists.

The FKG inequality was first (formally) investigated by Fortuin, Kastelyn and Ginibre in [143] in connection with correlation properties of Ising spin systems. The inequality is in fact a natural property of a very wide range of statistical mechanics models, including the random cluster model (with $q \geq 1$) [164], Yukawa quantum field theory models [282], Gaussian free fields [78, Proposition 5.22] and interlacement percolation [189]. We note that by [135, Theorem 3.3], it is sufficient to check that (2.2) holds for bounded continuous non-decreasing functions. On the other hand, the opposite assumption of negative association (A4) also holds for some prominent models, including the uniform spanning tree, the random cluster model (with $q \leq 1$), and simple exclusion models, we refer to [289] also for more motivation and background for this condition. We note that in the case of Gaussian fields, pairwise positive and negative correlation are enough to imply the FKG inequality [212] and negative associativity [213], respectively.

The spectral gap condition in (A2) and the finite range dependence in (A3) also appear as decorrelation assumptions in the context of quantitative stochastic homogenization, see for instance [161] for (A2), and [29] for (A3). In a sense, the spectral gap condition in (A2), introduced in [161], can be interpreted as a quantified version of ergodicity, as it implies an optimal variance decay for the semigroup associated with the "process of the environment as seen from the particle" induced by the simple random walk on $\mathbb{Z}^d$, cf. [160, Proposition 1 and Remark 5].

Throughout the paper we write c to denote a positive constant which may change on each appearance, while constants denoted c_i will be the same through the paper. The constants will depend only on d, p, q, the moments of $\mu^\omega(0)$ and $\nu^\omega(0)$, and the parameters $\gamma, C_{\text{mix}}, C_{\text{sg}}, \mathfrak{R}$ in Assumption 8 as appropriate, unless the dependencies are specified in the particular context.

Theorem 9. *Let $d \geq 2$ and suppose that Assumptions 6 and 8 hold. Then there exist constants $c_5, c_6, c_7 \in (0, \infty)$ and $p_0, q_0 \in [1, \infty)$ such that if $\omega(e) \in L^{p_0}(\mathbb{P})$ and $\omega(e)^{-1} \in L^{q_0}(\mathbb{P})$ the following holds. For $\mathbb{P}$-a.e. ω and any $x \in \mathbb{Z}^d$, there exists a random constant $N(x) = N(\omega, x)$ satisfying*

$$\mathbb{P}(N(x) > r) \leq c_5 r^{-\alpha}, \qquad \forall r > 0, \tag{2.4}$$

for some $\alpha > d(d-1) - 2$, such that for all $y \in \mathbb{Z}^d$ and $t \geq N(x)(1 \vee |x-y|)$,

$$p^\omega(t, x, y) \geq c_6 t^{-d/2} \exp\big(-c_7 |x-y|^2/t\big). \tag{2.5}$$

Remark 1. (i) Minimal choices for p_0 and q_0 are $p_0 > p\kappa\chi$ and $q_0 > q\kappa\chi$ with $\chi := d^2(1 + \frac{d^2-2}{\gamma - d^2})$ under (A1), and $p_0 = 2p\kappa d$ and $q_0 = 2q\kappa d$ under (A2), (A3) or (A4), for any $p, q > 1$ satisfying $1/p + 1/q < 2/d$, and with $\kappa = \kappa(p, q, d)$ as in Proposition 17 below. More precisely, the quantity κ originally appears in the random constant of the parabolic Harnack inequality in [17], which serves as one main ingredient in the proof of Theorem 9.

(ii) Given the two-sided heat kernel bounds provided by Theorems 7 and 9, the law of iterated logarithm (LIL) for the sample paths of the random walk can be established, see [130, 228]. However, it is expected that the LIL can be derived more easily under much weaker assumptions by exploiting the decomposition of the random walk into a martingale part and a corrector function, used in many proofs of a quenched functional central limit theorem, together with the sublinearity of the corrector (see e.g. [11, 16, 58, 80]) and an LIL for the martingale part.

In $d \geq 3$, we can use Theorems 9 and 7 to derive the following bound on the Green kernel, $g^\omega(x,y)$, defined by

$$g^\omega(x,y) := \int_0^\infty p_t^\omega(x,y)\,\mathrm{d}t, \qquad x,y \in \mathbb{Z}^d.$$

We refer to [11, Theorem 1.2] for precise estimates and asymptotics in the case of general non-negative i.i.d. conductances, to [17, Theorem 1.14] for a local limit theorem for g^ω in the case of ergodic conductances satisfying a moment condition, and to [20] for recent results on the Green kernel in dimension $d = 2$.

Theorem 10. *Let $d \geq 3$ and suppose that Assumption 6 holds.*

(i) Suppose there exist $p, q \in (1, \infty]$ with $1/p + 1/q < 2/d$ such that $\omega(e) \in L^p(\mathbb{P})$ and $\omega(e)^{-1} \in L^q(\mathbb{P})$ for any $e \in E_d$. For $\mathbb{P}$-a.e. ω, there exist $c_8 \in (0, \infty)$ and a random constant $N_2(x) = N_2(\omega, x)$ such that for all $x, y \in \mathbb{Z}^d$ with $|x - y| \geq N_2(x)$,

$$g^{\omega}(x,y) \leq c_8 \, |x-y|^{2-d} . \tag{2.6}$$

Additionally, suppose that Assumption 8 is satisfied. Then there exist $c_9, c_{10}, c_{11} \in (0, \infty)$ and $p_0, q_0 \in [1, \infty)$ such that if $\omega(e) \in L^{p_0}(\mathbb{P})$ and $\omega(e)^{-1} \in L^{q_0}(\mathbb{P})$, then the following hold.

(ii) For all $x, y \in \mathbb{Z}^d$ with $|x - y| > N(x)$,

$$g^{\omega}(x,y) \geq c_9 \, |x-y|^{2-d}. \tag{2.7}$$

(iii) For any $x, y \in \mathbb{Z}^d$ with $x \neq y$,

$$c_{10} \, |x-y|^{2-d} \leq \mathbb{E}\left[g^{\omega}(x,y)\right] \leq c_{11} \, |x-y|^{2-d} . \tag{2.8}$$

Example 11 (RCMs defined by Ginzburg-Landau $\nabla\phi$ interface models)**.** One class of conductances satisfying the assumptions of Theorem 9 can be constructed from the Ginzburg-Landau $\nabla\phi$-interface model (see [146]), a well established model for an interface separating two pure thermodynamical phases. The interface is described by a random field of height variables $\phi = \{\phi(x); x \in \mathbb{Z}^d\}$ sampled from a Gibbs measure formally given by $Z^{-1} \exp(-H(\varphi)) \prod_{x \in \mathbb{Z}^d} d\varphi(x)$ with formal Hamiltonian $H(\varphi) = \sum_{e \in E_d} V(\nabla\varphi(e))$ and potential $V \in C^2(\mathbb{R}; \mathbb{R}_+)$, which we suppose to be even and strictly convex. Note that in the special case $V(x) = \frac{1}{2}x^2$, the field ϕ becomes a discrete Gaussian free field. In $d \geq 3$ this can be made rigorous by taking the thermodynamical limit, while in dimension $d \geq 1$ one considers the gradient process instead. Then, thanks to the strict convexity we have the Brascamp-Lieb inequality, which allows one to show that any environment with random conductances of the form $\{\omega(x,y) = \lambda(\nabla\phi(e)), e \in E_d\}$ for any positive, even, globally Lipschitz function $\lambda \in C^1(\mathbb{R})$ satisfies the spectral gap condition in Assumption 8-(A2), see [23, Section 7] for details. The Brascamp-Lieb inequality also implies that exponential moments for gradient fields under the Gibbs measure exist (cf. [146, 280]). Thus, the environment $\{\omega(e), e \in E_d\}$ as chosen above also satisfies the required moment condition in Theorem 9. The assumption of a strictly convex potential can be relaxed, see [24].

The $\nabla\phi$ interface model also satisfies the FKG inequality, see again e.g. [146, 280], and for models with *massive* Hamiltonians formally given by

$$H(\varphi) = \sum_{e \in E_d} V(\nabla\varphi(e)) + \frac{m^2}{2} \sum_{x \in \mathbb{Z}^d} \phi(x)^2, \qquad m > 0,$$

we have exponential correlation decay, see [280, Theorem B]. In particular, Assumption 8-(A1) holds, and Theorem 9 applies, for instance, to conductances of the form $\omega(x,y) = \exp(\phi(x) + \phi(y))$, $\{x,y\} \in E_d$.

2.1.3 The Method

It is well known that Gaussian lower and upper bounds on the heat kernel are equivalent in many situations to a parabolic Harnack inequality (PHI), e.g. in the case of uniformly elliptic conductances, see [115]. Indeed, the PHI implies near-diagonal bounds which are then converted into off-diagonal bounds via the established *chaining method* (see e.g. [41, 115, 136]).

In our context, a PHI has been obtained in [17]. Unfortunately, due to the special structure of the constant in the PHI in the case of unbounded conductances (see (2.14) below), in particular its dependence on $\|\mu^\omega\|_{p,B(x,n)}$ and $\|\nu^\omega\|_{q,B(x,n)}$, we cannot directly deduce off-diagonal Gaussian lower bounds from it. In order to get effective Gaussian off-diagonal bounds using the chaining argument, one needs to apply the Harnack inequality on a number of balls with radius n over a distance of order n^2. In general, however, the ergodic theorem does not give the required uniform control on the convergence of space-averages of stationary random variables over such balls (see [3]). Therefore, in order to obtain lower Gaussian bounds we will need to make use of one of the additional conditions on the correlations stated in Assumption 8. Specifically, in Proposition 12 we employ any one of these conditions to derive a certain concentration estimate. Then, in Proposition 18 (and Corollary 19), we manipulate these to give us the desired uniform control on the space-averages of the conductances over the aforementioned chain of balls of radius n. Finally, we utilize this uniform control within the chaining argument to yield the desired Gaussian off-diagonal lower bound.

Near-diagonal heat kernel bounds can also be deduced from a local limit theorem, cf. [17, Lemma 5.3]. Recently, such local limit theorems have been derived for a more general class of RCMs in [14, 24] via De Giorgi's iteration technique, circumventing the need for a PHI. However, the bounds obtained from arguments in [14, 24] involve random constants which are implicit functions of the averages $\|\mu^\omega\|_{p,B(x,n)}$ and $\|\nu^\omega\|_{q,B(x,n)}$, while the chaining argument requires the more explicit dependence on the averages in the PHI in [17]. Note that in [17] the PHI has only been derived for the CSRW, so we obtain the lower heat kernel bounds in Theorem 9 for the CSRW only, while the upper bounds in [19] have been established for a general class of speed measures.

The rest of the paper is organised as follows. In Section 2.2, we first deduce some concentration estimates from the correlation decay conditions in Assumption 8, which are then used in Section 2.3 to prove the lower Gaussian bounds in Theorem 9. Finally, in Section 2.4 we show Theorem 10.

2.2 Concentration estimates under decorrelation assumptions

Recall that $\bar{\mu}_p := \mathbb{E}[\mu^\omega(0)^p]$ and $\bar{\nu}_q := \mathbb{E}[\nu^\omega(0)^q]$ for any $p,q \in [1,\infty)$. In this section we will derive some moment estimates on the deviations of $\mu^\omega(x)$ and $\nu^\omega(x)$ from their means under Assumption 8. For that purpose, we define the centred random variables

$$\Delta\mu_p^\omega(x) := \mu^\omega(x)^p - \bar{\mu}_p, \qquad \Delta\nu_q^\omega(x) := \nu^\omega(x)^q - \bar{\nu}_q, \qquad x \in \mathbb{Z}^d,$$

for any $p,q \in [1,\infty)$ such that $\bar{\mu}_p$ and $\bar{\nu}_q$ are finite. Our moment bounds on $\Delta\mu_p^\omega$ and $\Delta\nu_q^\omega$ will take the form given in the following definition.

Definition 1. *For any $p,q \in [1,\infty)$ and $1 \le \theta < \eta < \infty$ we say that $\mathbb{P}$ satisfies a (p,q,η,θ)-moment bound, if there exists $c \in (0,\infty)$ such that*

$$\mathbb{E}\Bigg[\Big|\sum_{x\in R} \Delta\mu_p^\omega(x)\Big|^\eta\Bigg] \le c\,|R|^\theta \qquad \textit{and} \qquad \mathbb{E}\Bigg[\Big|\sum_{x\in R} \Delta\nu_q^\omega(x)\Big|^\eta\Bigg] \le c\,|R|^\theta \tag{2.9}$$

for all hyper-rectangles $R \subset \mathbb{Z}^d$.

In the next proposition, which is the main result in this section, we gather and derive the relations between Assumption 8 and (p,q,η,θ)-moment bounds.

Proposition 12. *Let $\zeta, p, q \in [1,\infty)$ and let $R \subset \mathbb{Z}^d$ be a hyper-rectangle. Suppose that Assumptions 6 and 8 hold, and that $\zeta < \gamma/d$ if under (A1). There exist constants $p_0, q_0, \eta, \theta \in [1,\infty)$ with $\eta - \theta \ge \zeta$ such that if $\omega(e) \in L^{p_0}(\mathbb{P})$, $\omega(e)^{-1} \in L^{q_0}(\mathbb{P})$ for any $e \in E_d$, then the (p,q,η,θ)-moment bound holds.*

We will prove Proposition 12 under each of the assumptions separately, referencing the necessary materials before incorporating them into the proof. The following lemma is easily implied by [36, Corollary 1].

Lemma 13. *Let $\{Y(x) : x \in \mathbb{Z}^d\}$ be a random field satisfying the FKG inequality, and which is stationary with respect to translation, and suppose that*

$$\sum_{|x|\geq n} \mathbb{C}\mathrm{ov}\big(Y(0), Y(x)\big) = O(n^{-\nu}) \qquad \textit{and} \qquad \mathbb{E}\Big[|Y(0)|^{\eta+\delta}\Big] < \infty,$$

for some $\delta, \nu > 0$ and $\eta > 2$. Then, for any hyper-rectangle $R \subset \mathbb{Z}^d$,

$$\mathbb{E}\left[\Big|\sum_{x\in R} Y(x)\Big|^{\eta}\right] = O(n^{\theta}),$$

for $\theta > \max\{\eta/2, \chi(1-d^{-1}\min\{1, \nu\delta/\chi\})/(\eta+\delta-2)\}$, where $\chi = \delta + (\eta+\delta)(\eta-2)$.

Proof. We apply stationarity and the positivity of covariances due to the FKG inequality to [36, Corollary 1] to give the result. Indeed, by stationarity any hyper-rectangle can be shifted into $\mathbb{N}^d$, and positivity of the covariances allows us to bound the summation of covariances over $\mathbb{N}^d$ by the summation of covariances over $\mathbb{Z}^d$. $\square$

Proof of Proposition 12 under (A1): We will deal with the moment bound on the summation of the $\Delta\mu_p^{\omega}(x)$. The argument for the $\Delta\nu_q^{\omega}(x)$ follows identically. We will apply Lemma 13 to the field $Y(x) = \mu^{\omega}(x)^p$, $x \in \mathbb{Z}^d$. Then,

$$\sum_{|x|\geq n} \mathbb{C}\mathrm{ov}\big(Y(0), Y(x)\big) \leq c \sum_{|x|\geq n} |x|^{-\gamma} \leq cn^{-(\gamma-d)},$$

where in the first inequality we used the polynomial mixing condition in (A1). We can therefore take $\nu = \gamma - d$ in Lemma 13.

We now let $\eta = d\zeta$ and $p_0 = p\alpha$ with $\alpha > \frac{d\zeta(d\zeta-2)}{\gamma-d\zeta} + d\zeta$. Then in Lemma 13 we take $\delta = \alpha - d\zeta$ and note that $\nu > d\zeta - d$, to give that (2.9) holds with any $\theta > \chi(1 - d^{-1}\min\{\nu\delta/\chi, 1\})/(\eta+\delta-2)$. A computation then yields $\eta - \theta > \zeta$ for θ chosen close enough to the lower bound above. $\square$

We now turn to Proposition 12 under Assumption (A2). First, we recall that under the spectral gap condition, we have the following p-version of the spectral gap estimate. For

$p \geq 1$ and any $f \in L^{2p}(\Omega, \mathbb{P})$ with $\mathbb{E}[f] = 0$,

$$\mathbb{E}\left[|f|^{2p}\right] \leq c(p, C_{\text{sg}})\,\mathbb{E}\left[\Big(\sum_{e\in E_d}\big(\partial_e f\big)^2\Big)^p\right], \tag{2.10}$$

which basically follows by applying (2.3) to the function $|u|^p$, see [161, Lemma 2].

Proof of Proposition 12 under (A2): We will follow a similar argument given in [23, Lemma 2.10]. Again, we will only show the moment estimate for $\Delta\mu_p^\omega$. Take $p_0 = 2\zeta p$. Noting that $f := \sum_{y\in R}\Delta\mu_p^\omega(y)$ has mean zero, we use the spectral gap estimate in the form (2.10) which yields

$$\mathbb{E}\left[\Big|\sum_{y\in R}\Delta\mu_p^\omega(y)\Big|^{2\zeta}\right] \leq c\,\mathbb{E}\left[\Big(\sum_{e\in E_d}\big|\partial_e u\big|^2\Big)^\zeta\right].$$

Now we observe that, for any $e = \{\bar{e}, \underline{e}\} \in E_d$,

$$\partial_e\big[\Delta\mu_p^\omega(y)\big] = \partial_e\big[\mu^\omega(y)^p\big] = p\,\mu^\omega(y)^{p-1}\,\mathbb{1}_{\{\bar{e},\underline{e}\}}(y),$$

so that

$$\partial_e u \leq \begin{cases} p\big(\mu^\omega(\underline{e})^{p-1} + \mu^\omega(\bar{e})^{p-1}\big) & \text{if } \underline{e}\in R \text{ or } \bar{e}\in R,\\ 0 & \text{else.}\end{cases}$$

Hence,

$$\mathbb{E}\left[\Big|\sum_{y\in R}\Delta\mu_p^\omega(y)\Big|^{2\zeta}\right] \leq c\,|R|^\zeta\,\mathbb{E}\left[\mu^\omega(0)^{2\zeta(p-1)}\right],$$

and so we have obtained the requisite moment bounds with $\eta/2 = \theta = \zeta$. □

Lemma 14. *Let $p \in (2,\infty)$. There exists a constant $c_{12} = c_{12}(p)$ such that if $Y_1, \dots Y_n \in L^p(\mathbb{P})$ are independent random variables satisfying $\mathbb{E}\left[Y_j\right] = 0$ for all $j \in \{1,\dots,n\}$, then*

$$\mathbb{E}\left[\Big|\sum_{i=1}^n Y_j\Big|^p\right]^{1/p} \leq c_{12}\max\left\{\Big(\sum_{j=1}^n\mathbb{E}\left[|Y_j|^p\right]\Big)^{1/p}, \Big(\sum_{j=1}^n\mathbb{E}\left[|Y_j|^2\right]\Big)^{1/2}\right\}.$$

Proof. This can be extracted from [299, Theorem 3]. □

Proof of Proposition 12 under (A3): Take $p_0 = 2p\zeta$, again considering only the moment bound on the sum of the $\Delta\mu_p^\omega(x)$ as the argument for $\Delta\nu_q^\omega(x)$ is the same. Let $(e_i)_{1\leq i\leq d}$

denote the standard unit vectors. We call two vertices $x, y \in R$ equivalent if $x - y = \pm \Re e$ for some $e \in \{e_1, \dots, e_d\}$. Write the equivalence classes as $E_1, \dots, E_m$, and observe that we must have $m \leq \Re^d$. Note that the size of each equivalence class is trivially bounded above by $|R|$. We apply the finite range assumption to give that for each fixed i, the $(\mu^\omega(x))_{x \in E_i}$ are mutually independent, and therefore

$$
\begin{aligned}
\mathbb{E}\left[\Big|\sum_{y \in R(x)} \Delta\mu_p^\omega(y)\Big|^{2\zeta}\right] &= \mathbb{E}\left[\Big|\sum_{i \leq m} \sum_{y \in E_i} \Delta\mu_p^\omega(y)\Big|^{2\zeta}\right] \\
&\leq c \sum_{i \leq m} \mathbb{E}\left[\Big|\sum_{y \in E_i} \Delta\mu_p^\omega(y)\Big|^{2\zeta}\right] \leq c\, |R|^\zeta,
\end{aligned}
$$

where in the final step we apply Lemma 14 for each i in the summation, with (Y_j) an enumeration of $(\Delta\mu_p^\omega(y))_{y \in E_i}$. Thus (2.9) holds with $\eta/2 = \theta = \zeta$. $\square$

Lemma 15. *Let $\{Y_i, 1 \leqslant i \leqslant n\}$ be a negatively associated sequence. Further, let $\{Y_i^*, 1 \leqslant i \leqslant n\}$ be a sequence of independent random variables such that Y_i^* and Y_i have the same distribution for each $i = 1, 2, \dots, n$. Then*

$$
\mathbb{E}\left[\phi\left(\sum_{i=1}^n Y_i\right)\right] \leq \mathbb{E}\left[\phi\left(\sum_{i=1}^n Y_i^*\right)\right]
$$

for any convex function ϕ on $\mathbb{R}$, whenever the expectation on the right hand side exists.

Proof. This follows from [302, Theorem 1]. $\square$

Proof of Proposition 12 under (A4): Let $p_0 = 2p\zeta$ and $q_0 = 2q\zeta$. Then, we apply Lemma 15 and Lemma 14, with (Y_i) an enumeration of $(\Delta\mu_p^\omega(y)^*)_{y \in R}$ (and $(\Delta\nu_q^\omega(y)^*)_{y \in R}$, respectively) to give (2.9) with $\eta/2 = \theta = \zeta$. $\square$

As a first consequence of the concentration estimate in Proposition 12 we record the following tail estimate on the random variables $N_1(x)$, $x \in \mathbb{Z}^d$, defined via (2.1).

Lemma 16. *Suppose that Assumption 6 holds and that $\mathbb{P}$ satisfies a (p, q, η, θ)-moment bound, with $\zeta := \eta - \theta > 0$. Then there exists $c_{13} \in (0, \infty)$ such that*

$$
\mathbb{P}\big(N_1(x) > n\big) \leq c_{13}\, n^{1 - d\zeta}, \qquad \forall n \in \mathbb{N}. \tag{2.11}
$$

Proof. Note that, for any $n \in \mathbb{N}$, we have by a union bound

$$
\mathbb{P}(N_1(x) > n) \leq \sum_{m \geq n} \Big(\mathbb{P}\left[\|\mu^\omega\|_{p, B(x,m)}^p > 2\bar{\mu}_p\right] + \mathbb{P}\left[\|\nu^\omega\|_{q, B(x,m)}^q > 2\bar{\nu}_q\right] \Big). \tag{2.12}
$$

For the first term we get by Proposition 12 and Markov's inequality,

$$\sum_{m\geq n} \mathbb{P}\left[\|\mu^\omega\|^p_{p,B(x,m)} > 2\bar{\mu}_p\right] = \sum_{m\geq n} \mathbb{P}\left[\Big|\sum_{y\in B(x,m)} \Delta\mu^\omega_p(y)\Big|^\eta > \bar{\mu}_p^\eta \left|B(x,m)\right|^\eta\right]$$
$$\leq c\sum_{m\geq n} m^{-d\zeta} \leq c n^{1-d\zeta}.$$

Repeating the same argument with the second term in (2.12) gives the claim. □

2.3 Heat kernel lower bounds

We first recall the near-diagonal heat kernel bound in [17, Proposition 4.7], which will be a key ingredient in the proof of the main theorem.

Proposition 17. *Suppose that Assumption 6 holds, and suppose there exist $p,q \in (1,\infty]$ with $1/p+1/q < 2/d$ such that $\omega(e) \in L^p(\mathbb{P})$ and $\omega(e)^{-1} \in L^q(\mathbb{P})$ for any $e \in E_d$. Then there exists $c_{14} = c_{14}(d)$ such that for any $t \geq 1$, $x_1 \in \mathbb{Z}^d$ and $x_2 \in B(x_1, \frac{1}{2}\sqrt{t})$,*

$$p^\omega(t,x_1,x_2) \geq \frac{c_{14}}{C_{\mathrm{PH}}} t^{-\frac{d}{2}}, \tag{2.13}$$

where $C_{\mathrm{PH}} = C_{\mathrm{PH}}\big(\|\mu^\omega\|_{p,B(x_1,\sqrt{t})}, \|\nu^\omega\|_{q,B(x_1,\sqrt{t})}\big)$ is the constant appearing in the parabolic Harnack inequality in [17, Theorem 1.4], more explicitly given by

$$C_{\mathrm{PH}}\big(\|\mu^\omega\|_{p,B}, \|\nu^\omega\|_{q,B}\big) = c\exp\Big(c\left(1\vee\|\mu^\omega\|_{p,B}\right)^\kappa \left(1\vee\|\nu^\omega\|_{q,B}\right)^\kappa\Big) \tag{2.14}$$

for some positive $c = c(d,p,q)$ and $\kappa = \kappa(d,p,q) \geq 1$.

Theorem 9 will be proven by the well-established chaining technique. More precisely, we will apply Proposition 17 on a certain sequence of balls. Given a vertex $x = (x_1,\dots,x_d) \in \mathbb{Z}^d$ and $0 < r \leq 4|x|$, we specify a nearest-neighbour path $P[x]$ of length $D := |x|$ from 0 to x. Setting $p_0(x) := 0$ and $p_i(x) := (x_1,\dots,x_i,0,\dots,0) \in \mathbb{Z}^d$, $1 \leq i \leq d$, we define $P[x]$ to be the path that consists of d consecutive straight line segments connecting $p_0(x), p_1(x),\dots,p_d(x)$. Next, for any $k \in \mathbb{N}$ with $\frac{12D}{r} \leq k \leq \frac{16D}{r}$, we choose a subset $\{z_0,\dots z_k\} \subset P[x]$ such that $z_0 = 0$, $z_k = x$, $d(z_j,z_{j-1}) \leq \frac{r}{12}$ for $1 \leq j \leq k$ and such that, for each $j \leq k$, $\big|B(z_j,r)\cap\{z_0,\dots z_k\}\big| \leq c$ for some $c = c(d)$. Set $B_j := B(z_j, r/48)$. Finally, we let $s := Dr/k$, then $\frac{1}{16}r^2 \leq s \leq \frac{1}{12}r^2$.

Proposition 18. *Suppose that Assumption 6 holds. Further, for fixed $p,q \in [1,\infty)$ assume that $\mathbb{P}$ satisfies a (p,q,η,θ)-moment bound with $\zeta := \eta - \theta > d$. Then there exist constants*

$c_{15}, c_{16} \in (0,\infty)$ *and a random variable* $N_3 = N_3(\omega)$ *satisfying*

$$\mathbb{P}(N_3 > \rho) \le c_{15} \rho^{-d(\zeta-1)+2}, \qquad \forall \rho > 0, \tag{2.15}$$

such that, $\mathbb{P}$*-a.s., for all* $r \ge N_3$, $x \in \mathbb{Z}^d$ *and* $(B_j)_{1\le j\le k}$ *defined as right above, the following holds. If* $r \le 4|x|$, *for any collection of vertices* $y_0, \dots, y_k$ *with* $y_0 = 0$, $y_j \in B_j$ *for* $1 \le j \le k-1$ *and* $y_k = x$ *we have*

$$\sum_{j=0}^{k-1} \Big(1 \vee \|\mu^\omega\|_{p,B(y_j,\sqrt{s})}\Big) \Big(1 \vee \|\nu^\omega\|_{q,B(y_j,\sqrt{s})}\Big) \le c_{16} k. \tag{2.16}$$

Proof. Set $B_{y_j} := B(y_j, \sqrt{s})$ for abbreviation. Then note that there exists $c = c(d) \in (0,\infty)$ such that

$$\big|\big\{ j \in \{1, \dots k\} : z \in B_{y_j} \big\}\big| \le c, \qquad \forall z \in \bigcup_{j \le k} B_{y_j}. \tag{2.17}$$

We divide the rest of the proof into several steps.

Step 1. For $x \in \mathbb{Z}^d$ and r as in the statement, we will define a collection $(\mathscr{R}_i^{x,r})_{0 \le i \le d}$ of $d+1$ hyper-rectangles in $\mathbb{Z}^d$ which covers $\bigcup_{j \le k} B_{y_j}$ for any selection $y_j \in B_j$. For simplicity, we will only give the definition for $x \in \mathbb{Z}^d \cap [0,\infty)^d$ – it can be easily adjusted to the other regions of $\mathbb{Z}^d$. For any $m, l \in \mathbb{N}$, $u \in \mathbb{Z}^d$ and $i = 1, \dots d$ we write

$$R_i(u,m,l) := u + \big\{ v \in \mathbb{Z}^d : 0 \le v_i \le l, \, |v_j| \le m, \text{ for all } j \neq i \big\}$$

for the d-dimensional hyper-rectangle with base point u and dimension l along the e_i axis and m along the remaining coordinate axes. Now define

$$\mathscr{R}_0^{x,r} := \big([0,r] \times [-r,0]^{d-1}\big) \cap \mathbb{Z}^d \quad \text{and} \quad \mathscr{R}_i^{x,r} := R_i\big(p_{i-1}(x), r, x_i + r\big), \quad 1 \le i \le d.$$

Then note that $\bigcup_{0 \le i \le d} \mathscr{R}_i^{x,r} \supseteq \bigcup_{j \le k} B_{y_j}$.

Step 2. In this step we will show that there exists a random N_3 satisfying (2.15) such that for all $x \in \mathbb{Z}^d$ and all $r \ge N_3$,

$$\sum_{j \le k} \|\mu^\omega\|_{p,B_{y_j}}^p \le ck, \qquad \text{and} \qquad \sum_{j \le k} \|\nu^\omega\|_{q,B_{y_j}}^q \le ck. \tag{2.18}$$

We will only discuss the first inequality as the arguments for the second are identical. By (2.17) and the fact that the hyper-rectangles $(\mathscr{R}_i^{x,r})_{0\le i\le d}$ cover $\bigcup_{j\le k} B_{y_j}$, we have that

$$\begin{aligned}\sum_{j\le k}\|\mu^\omega\|^p_{p,B_{y_j}} &\le c\,r^{-d}\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\mu^\omega(y)^p\\ &\le c\,r^{-d}\left(\big(|x|+r\big)r^{d-1}\bar{\mu}_p+\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\Delta\mu_p^\omega(y)\right)\\ &\le c\,k+c\,r^{-d}\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\Delta\mu_p^\omega(y),\end{aligned}\tag{2.19}$$

where we have used that $|x|\le ckr$. Now we apply the moment-bound hypothesis with Proposition 12 and Markov's inequality to give

$$\mathbb{P}\left(\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\Delta\mu_p^\omega(y)>kr^d\right)\le c\left(|x|\,r^{d-1}\right)^{-\zeta},$$

where we have used that $kr^d\ge c\,|x|\,r^{d-1}$. Now fix $\rho,l\in\mathbb{N}$ with $l\ge\rho$. By applying a union bound, and summing over $\partial B(l):=\{x\in\mathbb{Z}^d:|x|=l\}$ and $r\ge\rho$, we get

$$\mathbb{P}\left(\exists x\in\partial B(l),r\in[\rho,4l]\cap\mathbb{N}:\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\Delta\mu_p^\omega(y)>kr^d\right)\le c\,l^{d-1-\zeta}\rho^{-\zeta(d-1)+1}.$$

Set

$$A_\rho:=\left\{\exists x\in\mathbb{Z}^d,r\in\mathbb{N}:|x|\ge\rho,\,r\in\big[\rho,4|x|\big],\sum_{0\le i\le d}\sum_{y\in\mathscr{R}_i^{x,r}}\Delta\mu_p^\omega(y)>kr^d\right\}.$$

Since $\zeta>d$, we can apply another union bound over $l\ge\rho$ to obtain

$$\mathbb{P}(A_\rho)\le c\,\rho^{d-\zeta}\rho^{-\zeta(d-1)+1}=c\,\rho^{d(1-\zeta)+1}.\tag{2.20}$$

However, $d(1-\zeta)+1<-1$, so by Borel-Cantelli, for $\mathbb{P}$-a.e. ω, there exists $N_3=N_3(\omega)$ such that A_ρ does not occur for $\rho\ge N_3$. Substituting into (5.29) completes the proof of (2.18). Moreover, via a union bound, (2.20) implies that N_3 can be constructed such that (2.15) holds.

Step 3. By Hölder's inequality,

$$\sum_{j\le k}\left(1\vee\|\mu^\omega\|_{p,B_{y_j}}\right)\left(1\vee\|\nu^\omega\|_{q,B_{y_j}}\right)$$
$$\le k^{1-\frac{1}{p}-\frac{1}{q}}\left(\sum_{j\le k}\left(1\vee\|\mu^\omega\|^p_{p,B_{y_j}}\right)\right)^{1/p}\left(\sum_{j\le k}\left(1\vee\|\nu^\omega\|^q_{q,B_{y_j}}\right)\right)^{1/q},$$

so that the statement follows from Step 2. □

Remark 2. With a more convoluted covering argument, replacing union bounds with bounds on maxima, the requirement of $\zeta > d$ in Proposition 18 can be decreased to $\zeta > d-1$, and thus we only need $\gamma > d(d-1)$, and the minimal moment conditions of Remark 1 can be reduced to $p_0 > p\kappa\chi$ and $q_0 > q\kappa\chi$ with $\chi = d(d-1)\left[1+\frac{d(d-1)-2}{\gamma-d(d-1)}\right]$ under (A1), and $p_0 = 2p\kappa(d-1)$ and $q_0 = 2q\kappa(d-1)$ under (A2), (A3) or (A4). We do not include this argument as it brings greatly increased complication for very limited improvement.

Corollary 19. *In the setting of Proposition 18, assume that $\mathbb{P}$ satisfies a $(\kappa p,\kappa q,\eta,\theta)$-moment bound. Then there exists a constant $c_{17}\in(0,\infty)$ and a random variable $N_4 = N_4(\omega)$ satisfying* (2.15) *such that, $\mathbb{P}$-a.s., for all $r\ge N_4$, $x\in\mathbb{Z}^d$ with $r\le 4|x|$,*

$$\sum_{j=0}^{k-1}\left(1\vee\|\mu^\omega\|_{p,B(y_j,\sqrt{s})}\right)^\kappa\left(1\vee\|\nu^\omega\|_{q,B(y_j,\sqrt{s})}\right)^\kappa \le c_{17}k. \tag{2.21}$$

Proof. This follows exactly as Proposition 18 after applying Jensen's inequality to $\|\mu^\omega\|_{p,B_{y_j}}$ and $\|\nu^\omega\|_{q,B_{y_j}}$, and then replacing μ^ω by $(\mu^\omega)^\kappa$ and ν^ω by $(\nu^\omega)^\kappa$. □

Proof of Theorem 9. By translation invariance of the measure it suffices to show a lower bound on $p^\omega(t,0,x)$. To begin with, we must establish the necessary moment conditions to deploy the tools developed in the previous section.

We assume that there exist some $p,q\in(1,\infty)$ with $1/p+1/q<2/d$ such that $\omega(e)\in L^p(\mathbb{P})$ and $\omega(e)^{-1}\in L^q(\mathbb{P})$. This will allow us to apply Proposition 17 involving the constant $\kappa=\kappa(p,q,d)$. If working under assumption (A1), recall that $\gamma>d^2$ and fix $d<\zeta<\gamma/d$; otherwise, just fix $\zeta>d$. Then Proposition 12 provides us with $p_0,q_0\ge 1$ such that if $\omega(e)\in L^{p_0}(\mathbb{P})$ and $\omega(e)^{-1}\in L^{q_0}(\mathbb{P})$, then $(1,1,\eta,\theta)$ and $(\kappa p,\kappa q,\eta,\theta)$-moment bounds hold with $\eta-\theta\ge\zeta$. This will allow us to apply Lemma 16, Proposition 18, and Corollary 19 as required. We then set $N := N_1(0)\vee N_3\vee N_4$ and combine the tail bounds in Lemma 16 and (2.15) to obtain that N satisfies the tail bound in (2.4).

Set again $D := |x|$, and assume now that $t\ge N(D\vee 1)$. We will split the proof into two cases, $D^2/t\le 1/4$ and $D^2/t>1/4$.

Case 1: $D^2/t \leq 1/4$. Then $x \in B(0, \frac{1}{2}\sqrt{t})$, so by Proposition 17,

$$p^{\omega}(t,0,x) \geq \frac{c_{14}}{C_{\mathrm{PH}}} t^{-\frac{d}{2}}$$

with $C_{\mathrm{PH}} = C_{\mathrm{PH}}\big(\|\mu^{\omega}\|_{p,B(0,\sqrt{t})}, \|\nu^{\omega}\|_{q,B(0,\sqrt{t})}\big)$. Since $C_{\mathrm{PH}}\big(\|\mu^{\omega}\|_{p,B}, \|\nu^{\omega}\|_{q,B}\big)$ is increasing in $\|\mu^{\omega}\|_{p,B}$ and $\|\nu^{\omega}\|_{q,B}$ (cf. (2.14) above) and $t \geq N_1(0)$,

$$C_{\mathrm{PH}}\big(\|\mu^{\omega}\|_{p,B(0,\sqrt{t})}, \|\nu^{\omega}\|_{q,B(0,\sqrt{t})}\big) \leq C_{\mathrm{PH}}\Big(\big(2\bar{\mu}_p\big)^{1/p}, \big(2\bar{\nu}_q\big)^{1/q}\Big),$$

and therefore $p^{\omega}(t,0,x) \geq ct^{-d/2}$.

Case 2: $D^2/t > 1/4$. Set $r := t/D \geq 1 \vee N_3 \vee N_4$. We deploy the chaining setup as introduced right below Proposition 17. Recall that $s := Dr/k = t/k$ with $\frac{12D}{r} \leq k \leq \frac{16D}{r}$ so that $1 \leq \frac{1}{16}r^2 \leq s \leq \frac{1}{12}r^2$, and note that $k \geq 3$. Then, for any collection of vertices $y_0, \dots y_k$ with $y_0 = 0$, $y_j \in B_j$ for $1 \leq j \leq k-1$ and $y_k = x$, we have $d(y_i, y_{i+1}) \leq r/8 \leq \sqrt{s}/2$ so that by Proposition 17,

$$p^{\omega}(s,y_i,y_{i+1}) \geq \frac{c_{14}}{C_{\mathrm{PH}}\big(\|\mu^{\omega}\|_{p,B_{y_j}}, \|\nu^{\omega}\|_{q,B_{y_j}}\big)} s^{-\frac{d}{2}} \geq \frac{c}{C_{\mathrm{PH}}\big(\|\mu^{\omega}\|_{p,B_{y_j}}, \|\nu^{\omega}\|_{q,B_{y_j}}\big)} r^{-d},$$

with $B_{y_j} := B(y_j, \sqrt{s})$. Further, recall the representation of C_{PH} in (2.14) and that $\mathrm{P}^{\omega}_{y_j}[X_s = y_{j+1}] = p^{\omega}(s,y_j,y_{j+1})\mu^{\omega}(y_{j+1})$. Hence, by the Markov property,

$$\begin{aligned}
\mathrm{P}^{\omega}_0[X_t = x] &= \mathrm{P}^{\omega}_0[X_{ks} = x] \geq \mathrm{P}^{\omega}_0\big[X_{js} \in B_j, 1 \leq j \leq k-1, X_{ks} = x\big] \\
&\geq \sum_{y_1 \in B_1, \dots, y_{k-1} \in B_{k-1}} \frac{c^k \Big(\prod_{j=1}^{k-1} r^{-d}\mu^{\omega}(y_j)\Big) s^{-d/2}\mu^{\omega}(x)}{\exp\Big(c\sum_{j=0}^{k-1}\big(1 \vee \|\mu^{\omega}\|_{p,B_{y_j}}\big)^{\kappa}\big(1 \vee \|\nu^{\omega}\|_{q,B_{y_j}}\big)^{\kappa}\Big)} \\
&\geq c^k \left(\prod_{j=1}^{k-1} \|\mu^{\omega}\|_{1,B_j}\right) s^{-d/2}\mu^{\omega}(x),
\end{aligned}$$

where we used Corollary 19 in the last step. In particular,

$$p^{\omega}(t,0,x) \geq c^k \left(\prod_{j=1}^{k-1} \|\mu^{\omega}\|_{1,B_j}\right) t^{-d/2}. \tag{2.22}$$

Now, by the harmonic-geometric mean inequality and Jensen's inequality, we have

$$\left(\prod_{j=1}^{k-1} \|\mu^{\omega}\|_{1,B_j}\right)^{\frac{1}{k-1}} \geq \frac{k-1}{\sum_{j=1}^{k-1} \|\mu^{\omega}\|_{1,B_j}^{-1}} \geq \frac{c\,(k-1)}{\sum_{j=1}^{k-1} \|\nu^{\omega}\|_{1,B_j}}.$$

We use Proposition 18 (setting $p = q = 1$, $y_j = z_j$, and replacing B_{y_j} with B_j) to obtain

$$\sum_{j=1}^{k-1} \|\nu^\omega\|_{1,B_j} \le c\,(k-1),$$

so that

$$\left(\prod_{j=1}^{k-1} \|\mu^\omega\|_{1,B_j}\right) \ge c^{k-1}.$$

Combining this with (2.22) yields $p^\omega(t,0,x) \ge c\,c_{18}^k t^{-d/2}$ for some $c_{18} \in (0,1)$, which gives the bound (2.5) by the choice of k. □

2.4 Green kernel estimates

In this final section we utilize Theorems 7 and 9 to establish Theorem 10. We refer to [47, Section 6] for similar arguments.

Proof of Theorem 10. (i) First we deduce the upper bound (2.6) on the Green kernel. For any distinct $x, y \in \mathbb{Z}^d$, we decompose the integral as

$$g^\omega(x,y) = \frac{1}{\mu^\omega(x)} \int_0^{N_1(x)^2} \mathrm{P}_y^\omega(X_t = x)\,\mathrm{d}t + \int_{N_1(x)^2}^{N_{x,y}} p_t^\omega(x,y)\,\mathrm{d}t + \int_{N_{x,y}}^\infty p_t^\omega(x,y)\,\mathrm{d}t \qquad (2.23)$$

with $N_{x,y} := N_1(x)^2 \vee (|x-y|/c_1)$, where we used that by the symmetry of the heat kernel $p_t^\omega(x,y) = p_t^\omega(y,x)$. By Theorem 7 we can bound the last two terms of (2.23) by

$$\int_{N_1(x)^2}^{N_{x,y}} p_t^\omega(x,y)\,\mathrm{d}t \le c_2 e^{-c_4|x-y|} \int_1^\infty t^{-d/2}\,\mathrm{d}t \le c\,|x-y|^{2-d},$$

and

$$\int_{N_{x,y}}^\infty p_t^\omega(x,y)\,\mathrm{d}t \le c_2 \int_0^\infty t^{-d/2} e^{-c_3|x-y|^2/t}\,\mathrm{d}t \le c\,|x-y|^{2-d}.$$

It is left to bound the first term in the right hand side of (2.23). Recall that the random walk X spends i.i.d. $\mathrm{Exp}(1)$-distributed waiting times between its jumps. Set $\lambda := N_1(x)^2$ and $r := |x-y| \ge 1$. In particular, the random walk starting at y needs to perform at least r

jumps to get to x. Thus,

$$\begin{aligned}\frac{1}{\mu^\omega(x)}\int_0^\lambda \mathrm{P}_y^\omega(X_t=x)\,\mathrm{d}t &\le \frac{\lambda}{\mu^\omega(x)}\mathrm{P}_y^\omega\big(X_t=x \text{ for any } t\in[0,\lambda]\big)\\ &\le \frac{\lambda}{\mu^\omega(x)}\mathrm{Pois}(\lambda)\big([r,\infty)\big).\end{aligned} \tag{2.24}$$

Here $\mathrm{Pois}(\lambda)$ denotes the Poisson distribution with parameter λ, which we recall to have exponential tails (see e.g. [206, Remark 2.6]). So there exists $N_2=N_2(\omega,x)$ such that for each $y\in\mathbb{Z}^d$ with $|x-y|\ge N_2(\omega,x)$ the first term in (2.23) is bounded from above by $c\,|x-y|^{2-d}$, which completes the proof of (2.6).

(ii) This follows directly from Theorem 9, which gives for $x,y\in\mathbb{Z}^d$ with $|x-y|>N(x)$,

$$\begin{aligned}\int_0^\infty p_t^\omega(x,y)\,\mathrm{d}t &\ge \int_{N(x)|x-y|}^\infty c_6\,t^{-d/2}e^{-c_7|x-y|^2/t}\,\mathrm{d}t \ge \int_{|x-y|^2}^\infty c_6\,t^{-d/2}e^{c_7|x-y|^2/t}\,\mathrm{d}t\\ &= |x-y|^{2-d}\int_1^\infty c_6\,t^{-d/2}\,e^{-c_7/t}\,\mathrm{d}t = c\,|x-y|^{2-d}.\end{aligned}$$

(iii) First, we carry out some preparation for the proof of the upper bound. In particular, we show the Green kernel has finite second moments. By the symmetry of the heat kernel and the on-diagonal part of the upper bound in Theorem 7, note that

$$\begin{aligned}g^\omega(x,y) &= \int_0^{N_1(x)^2} p_t^\omega(y,x)\,\mathrm{d}t + \int_{N_1(x)^2}^\infty p_t^\omega(x,y)\,\mathrm{d}t \le \frac{N_1(x)^2}{\mu^\omega(x)} + cN_1(x)^{2-d}\\ &\le cN_1(x)^2\nu^\omega(x),\end{aligned}$$

where we used Jensen's inequality in the last step. Assuming that $\omega(e)\in L^{p_0}(\mathbb{P})$ and $\omega(e)^{-1}\in L^{q_0}(\mathbb{P})$ for suitable $p_0,q_0\in(1,\infty)$, we apply Proposition 12 together with Lemma 16 (with $\zeta=d$) to obtain that $N_1\in L^\beta(\mathbb{P})$ for any $\beta<d^2-1$. Thus, by Hölder's inequality,

$$\mathbb{E}\left[g^\omega(x,y)^\beta\right]<\infty, \tag{2.25}$$

for any $\beta<(d^2-1)(q_0-1)/(2q_0)$. Then, as $d\ge 3$, assuming $q_0>2$ ensures that the second moment of the Green kernel exists.

We can now prove the upper bound of (2.8). To do so, we show that the random variable N_2 introduced in (i) satisfies the tail bound

$$\mathbb{P}\left[N_2>u\right]\le c\,u^{2-d}, \qquad \forall u\ge 1. \tag{2.26}$$

Indeed, if (2.26) holds true, then we obtain

$$\mathbb{E}\left[g^{\omega}(x,y)\right] \le c\,|x-y|^{2-d} + \mathbb{E}\left[g^{\omega}(x,y)^2\right]^{1/2}\,\mathbb{P}\left[N_2 > |x-y|\right] \le c\,|x-y|^{2-d},$$

where we used (2.6) and the Cauchy-Schwarz inequality in the first step, and (2.26) and (2.25) in the second step.

In order to show (2.26), recall that N_2 has been chosen as a value of r such that $\frac{\lambda}{\mu^{\omega}(x)}\mathrm{Pois}(\lambda)\big([r,\infty)\big) \le cr^{2-d}$ with $\lambda := N_1(x)^2$. By Chernoff's inequality (cf. e.g. [206, Corollary 2.4 and Remark 2.6]), $\mathrm{Pois}(\lambda)\big([r,\infty)\big) \le e^{-r+7\lambda}$ for $r > 7\lambda$. Hence, N_2 can be chosen as a constant times $\lambda = N_1^2$, so the tail bound on N_2 can be dominated by a tail bound on N_1^2, which is provided by Proposition 12 and Lemma 16 (with the choice $\zeta = d$) under suitable moment conditions on $\omega(e)$ and $\omega(e)^{-1}$. More precisely,

$$\mathbb{P}(N_2 > u) \le \mathbb{P}(N_1^2 > cu) \le c\,u^{\frac{1-d^2}{2}} \le c\,u^{2-d}, \qquad u \ge 1,$$

since $(1-d^2)/2 < 2-d$ for $d \ge 3$, which completes the proof of (2.26).

Finally, we prove the lower bound of (2.8), which follows again from Theorem 9. Choose $K \in (0,\infty)$ such that $\mathbb{P}(N(x) \le K) = \mathbb{P}(N(0) \le K) \ge 1/2$, then

$$\mathbb{E}\left[g^{\omega}(x,y)\right] \ge \mathbb{E}\left[\int_{N(x)}^{\infty} p_t^{\omega}(x,y)\,\mathrm{d}t\right] \ge \frac{1}{2}\int_K^{\infty} c_6\,t^{-d/2}\,e^{-c_7|x-y|^2/t}\,\mathrm{d}t.$$

If $|x-y|^2 \ge K$ then we can bound the integral on the right hand side as in the proof of (2.7) to give (2.8). On the other hand, there are only finitely many vertices $z \in B(0,\sqrt{K})$, and for each such z we have $\mathbb{E}\,[g^{\omega}(0,z)] > 0$. Therefore, $\inf_{y\in B(x,\sqrt{K})}\mathbb{E}\,[g^{\omega}(x,y)] = \inf_{z\in B(0,\sqrt{K})}\mathbb{E}\,[g^{\omega}(0,z)] > 0$. Thus, we can adjust the constant c_{10} such that (2.8) also holds for $|x-y|^2 \le K$. $\square$

Chapter 3

[B] Collisions of Random Walks in Dynamic Random Environments

Abstract. We study dynamic random conductance models on $\mathbb{Z}^2$ in which the environment evolves as a reversible Markov process that is stationary under space-time shifts. We prove under a second moment assumption that two conditionally independent random walks in the same environment collide infinitely often almost surely. These results apply in particular to random walks on dynamical percolation.

3.1 Introduction

A graph is said to have the **infinite collisions property** if two independent random walks started at the same location collide (occupy the same location at the same time) infinitely often almost surely. For Euclidean lattices, Polya [295] observed that the study of collisions can be reduced to the study of returns on an auxiliary lattice, and hence that the infinite collisions property holds if and only if the dimension is at most two. In fact, *for transitive graphs*, the infinite collisions property is always equivalent to recurrence: The number of collisions and the number of returns are geometric random variables with the same mean. For bounded degree graphs that are *not* transitive, the infinite collisions property is strictly stronger than recurrence. Indeed, while it is easy to see that bounded degree transient graphs cannot have infinite collisions, Krishnapur and Peres [225] showed that there exist bounded degree graphs, including the *infinite comb graph*, that are recurrent but which do not have the infinite collisions property. See e.g. [104] for further examples.

Despite the existence of these counterexamples, it is natural to expect that the infinite collisions property is equivalent to recurrence for most graphs and networks arising in applications. Indeed, it is now known that the two properties are equivalent for many random walks in random environments that are spatially homogeneous in some distributional sense [52, 103]. The most general such result is due to Hutchcroft and Peres [195], who proved that every recurrent *reversible random rooted network* has the infinite collisions property. An important class of examples to which these result apply are the translation-invariant random conductance models on $\mathbb{Z}^d$; see [77] for background. Note that while earlier results such as those of [52] had relied on a fine analysis of the random walk in specific examples, the method of [195] is entirely qualitative and does not rely on heat-kernel estimates. Further results on collisions of random walks in random environments include [102, 122, 123, 148, 151].

In this paper we study collisions of random walks on *dynamic* random conductance models (dynamic RCMs), in which the environment itself is permitted to vary over time. Such models have recently been of burgeoning interest, with works establishing, for example, quenched invariance principles [10, 12, 82], quenched and annealed local limit theorems [13, 25], heat kernel estimates [116, 278], and Green kernel asymptotics [21]. We restrict attention to the class of dynamic RCMs in which the conductances themselves form a *strongly reversible* Markov process whose law is invariant under space-time shifts. We will refer to such environments as **stationary, strongly reversible Markovian environments**; see Section 3.2 for detailed definitions. This class includes many of the most natural and interesting examples of dynamic RCMs appearing in the literature, including dynamical percolation [184, 290, 291, 291, 292], the simple symmetric exclusion process [31, 297, 308], and dynamic RCMs in which the conductances evolve according to an SDE such as those arising in the Helffer-Sjöstrand representation of gradient fields, see e.g. [116, 183]. Previous works studying random walks in general (reversible and non-reversible) Markovian environments include [32, 33, 128].

We now state our main theorem. We write E_d for the edge set of $\mathbb{Z}^d$, and consider our random environments to be random locally integrable functions from $\mathbb{R} \times E_d$ to $[0,\infty)$. The walk in the environment η is in continuous time, and is defined formally in Section 3.2, with generator given in (3.2). We say that a stationary Markovian random environment $\eta : \mathbb{R} \times E_d \to [0,\infty)$ is **strongly reversible** if the conditional distributions of η and its reversal given the instantaneous sigma-algebra $\mathscr{F}_0$ are almost surely equal, where $\mathscr{F}_{[s,t]}$ is the sigma-algebra generated by the restriction of η to $[s,t]$ and $\mathscr{F}_0 := \bigcap\{\mathscr{F}_{[s,t]} : s \leq 0 \leq t, s < t\}$; see Section 3.2 for more detailed definitions.

Theorem 20. *Let $\eta : \mathbb{R} \times E_2 \to [0,\infty)$ be a stationary random environment on $\mathbb{Z}^2$ and let $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ be two doubly-infinite random walks on η, both started from the origin at time zero, that are conditionally independent given the environment η. Suppose that at least one of the following conditions holds:*

(A1): *The environment η is Markovian, strongly reversible, and satisfies the second moment condition $\|\eta\|_2^2 := \sup_{a<b} \frac{1}{|b-a|^2}\mathbb{E}[(\int_a^b \sum_{x\sim 0} \eta_s(\{0,x\})\,\mathrm{d}s)^2] < \infty$.*

(A2): *The backwards walk $(X_{-t})_{t\geq 0}$ satisfies a (quenched or annealed) invariance principle under Brownian scaling with Brownian motion on $\mathbb{R}^2$ as the limiting distribution.*

Then X and Y collide infinitely often almost surely: the set $\{n \in \mathbb{N} : X_n = Y_n\}$ has infinite cardinality almost surely and the set $\{t \in [0,\infty) : X_t = Y_t\}$ has infinite Lebesgue measure almost surely.

Remark 3. Having infinitely many integer collision times implies very generally that the Lebesgue measure of the collision times is infinite by a standard application of Tonelli's Theorem as shown in Lemma 31.

Invariance principles are known in the ergodic setting in the non-elliptic case with rates bounded from above (and 0 only on intervals with lengths of finite expectation) [82], and with elliptic rates under moment conditions on the conductances and their reciprocals [25]. Such environments need not be reversible, so there exist examples that satisfy (A2) but not (A1). On the other hand, most examples arising in applications do satisfy the simpler condition (A1), for which our proof is self-contained and relies on a simpler and more general analysis than that required to establish an invariance principle. Indeed, (A1) applies to highly non-elliptic environments for which invariance principles do not hold, such as the random walk on the uniform spanning tree of $\mathbb{Z}^2$ which has a non-Brownian scaling limit [44]. Dynamical percolation and the simple symmetric exclusion process are covered by either hypothesis (A1) or (A2).

Both results will be deduced from the following more general theorem. Note that the hypotheses of this theorem hold trivially under the assumption (A2) of Theorem 20; in Section 3.2.2 we use the theory of *Markov-type inequalities* to prove that they also hold under the assumption (A1).

Theorem 21 (A weak diffusive estimate suffices)**.** *Let $\eta : \mathbb{R} \times E_2 \to \mathbb{R}_{\geq 0}$ be a stationary random environment on $\mathbb{Z}^2$ and let $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ be two doubly-infinite random walks on η, both started from the origin at time zero, that are conditionally independent given the environment η. Suppose that for every $\varepsilon > 0$ there exists $K < \infty$ and $\delta > 0$ such that*

$$\mathbb{P}\left(\limsup_{n\to\infty} \min_{0\leq m\leq n} \mathbb{P}^\eta\left(\|X_{-m}\|_2 \leq K\sqrt{n}\right) \geq \delta\right) \geq 1-\varepsilon. \tag{3.1}$$

Then X and Y collide infinitely often almost surely: the set $\{n \in \mathbb{N} : X_n = Y_n\}$ has infinite cardinality and the set $\{t \in \mathbb{R}_{\geq 0} : X_t = Y_t\}$ has infinite Lebesgue measure almost surely.

Under some additional non-degeneracy assumptions, we are able to prove similar infinite-collision theorems in which the two walks X and Y are not required to start at the same location. We say a random environment η is **irreducible** if for each two vertices x and y there exist times $s < t$ such that the conditional transition probability $P^{\eta}_{s,t}(x,y)$ is positive with positive probability. We say that a stationary environment η is **time-ergodic** if it has probability either zero or one to belong to any time-shift-invariant measurable subset of Ω. (Note that being time-ergodic is a stronger condition than being space-time ergodic.)

Corollary 22. *Let $\eta : \mathbb{R} \times E_2 \to [0,\infty)$ be a irreducible, time-ergodic, stationary random environment on $\mathbb{Z}^2$ and let $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ be two doubly-infinite random walks on η, started at two vertices x and y at time zero, that are conditionally independent given the environment η. If η satisfies the hypotheses of either Theorem 20 or Theorem 21 then X and Y collide infinitely often almost surely: the set $\{n \in \mathbb{N} : X_n = Y_n\}$ has infinite cardinality and the set $\{t \in [0,\infty) : X_t = Y_t\}$ has infinite Lebesgue measure almost surely.*

Corollaries for the voter model. Let us now briefly describe a corollary of our results for the *voter model* in two-dimensional dynamic random environments. Roughly speaking, the voter model in the environment $\eta : \mathbb{R} \times E_d \to \mathbb{R}$ is the interacting particle system on $\mathbb{Z}^d$ in which each vertex has an opinion belonging to $[0,1]$ and the opinion of x changes to match the opinion of y at rate $\eta_t(\{x,y\})$. Since this model is tangential to the main results of this paper, we omit the precise definition of the model and refer the reader to [249] for background. The following is an immediate consequence of Corollary 22 and the standard duality between the voter model and coalescing random walk described in [249, §5] and [Aldous and Fill, §14], which readily generalises to the dynamic case.

Corollary 23. *Let $\eta : \mathbb{R} \times E_2 \to \mathbb{R}_{\geq 0}$ be a stationary random environment on $\mathbb{Z}^2$. If the* reversal *of η satisfies the hypotheses of Corollary 22, then the only ergodic stationary measures for the voter model in η are the constant (a.k.a. consensus) measures.*

One-dimensional models. Our methods can also be used to prove that one-dimensional stationary random environments have the infinite collision property under a *first moment* condition. This is much simpler than the two-dimensional case. Once this proposition is proven, one can also formulate and prove one-dimensional analogues of Corollaries 22 and 23 similarly to the two-dimensional case; we omit the details.

Proposition 24. *Let η be a stationary random environment on $\mathbb{Z}$ with $\|\eta\|_1 < \infty$. Then η has the infinite collisions property almost surely: If X and Y are two random walks on η, both started from the origin at time zero, that are conditionally independent given η, then the set $\{n \in \mathbb{N} : X_n = Y_n\}$ has infinite cardinality and the set $\{t \in [0,\infty) : X_t = Y_t\}$ has infinite Lebesgue measure almost surely.*

About the proof and organisation. This remainder of this paper will be divided into two sections. In Section 2 we define necessary terminology, before establishing moment bounds on the number of jumps the random walk takes in a given interval, as well as non-explosivity in Proposition 25. Then, in Corollary 29, we use the Markov-Type inequality, along with the previously derived moment bounds, to prove a diffusive upper bound on the displacement of the random walk on the environment.

In Section 3, we will use these results to complete the proof of the theorem. In Proposition 107, we extend to the time-varying setting an argument of Hutchcroft and Peres [195] to give a sufficient condition for dynamic environments to satisfy the infinite collisions property. Namely, we prove, utilizing the Mass-transport Principle, that if the expected number of collisions *of the backwards walks* conditioned on the environment is infinite almost surely, then the number of collisions is infinite almost surely. Then, in Theorems 20 and 21, we complete the proof by demonstrating that in two dimensions, the diffusive bound on displacement implies the previously derived sufficient condition on the conditional expectations. We finish by proving Corollary 22.

3.2 Stationary Random Environments

Fix $d \geq 1$. We work on the d-dimensional Euclidean lattice $(\mathbb{Z}^d, E_d)$, where $E_d = \{\{x,y\} \in \mathbb{Z}^d \times \mathbb{Z}^d : \|x-y\|_1 = 1\}$. We write $x \sim y$ if $\{x,y\} \in E_d$, and $B(x,r)$ for the l^1 ball centred at x with radius r. For each $e = \{x,y\} \in E_d$ and $z \in \mathbb{Z}^d$, we write $e - z$ for the edge $\{x-z, y-z\}$. We define an **environment** to be a non-negative element of the space $L^1_{\mathrm{loc}}(E_d \times \mathbb{R})$ of locally integrable, measurable functions $E_d \times \mathbb{R} \to \mathbb{R}$ modulo a.e. equivalence, where we recall that $f : E_d \times \mathbb{R} \to \mathbb{R}$ is said to be **locally integrable** if $\int_a^b |f_t(e)|\,\mathrm{d}t < \infty$ for every $a < b$ and every edge $e \in E_d$. (Here and elsewhere we follow the usual convention of writing the time variable as a subscript.) We recall that $L^1_{\mathrm{loc}}(E_d \times \mathbb{R})$ can be endowed with a unique topology, called the *local* L^1 *topology*, with the property that f^n converges to f if and only if $\int_a^b |f^n_t(e) - f_t(e)|\,\mathrm{d}t \to 0$ as $n \to \infty$ for every $a < b$ and $e \in E_d$. We write $\Omega = \{\eta \in L^1_{\mathrm{loc}}(E_d \times \mathbb{R}) : \eta_t(e) \geq 0$ for every $e \in E_d$ and a.e. $t \in \mathbb{R}\}$ for the space of environments, which we equip with the associated subspace topology and Borel σ-algebra. For each environment $\eta \in \Omega$ and $x \in \mathbb{Z}^d$ we write $\eta_t(x) = \sum_{y \sim x} \eta_t(\{x,y\})$.

We refer to a random variable taking values in Ω as a **random environment**. For each $x \in \mathbb{Z}^d$ and $t \in \mathbb{R}$ we write $\tau_{x,t} : \Omega \to \Omega$ for the space-time shift defined by $\tau_{x,t}\eta_s(e) = \eta_{s-t}(e-x)$ and say that a random environment η is **stationary** if $\tau_{x,t}(\eta)$ has the same distribution as η for every $x \in \mathbb{Z}^d$ and $t \in \mathbb{R}$. Similarly, we define the time-reversal map $R : \Omega \to \Omega$ by $R(\eta)_t(e) = \eta_{-t}(e)$ and say that a random environment η is **reversible** if $R(\eta)$ has the same distribution as η. For each $a<b$, let $\mathscr{F}_{[a,b]}$ be the σ-algebra generated by the restriction of η to $[a,b]$. We say that η is a **Markovian random environment** if $\mathscr{F}_{[a_1,a_2]}$ and $\mathscr{F}_{[c_1,c_2]}$ are conditionally independent given $\mathscr{F}_{[b_1,b_2]}$ whenever $a_2<b_2$ and $c_1>b_1$ (that is, if the past and the future are conditionally independent given the present). For each $t\in\mathbb{R}$, we define the **instantaneous sigma-algebra** $\mathscr{F}_t=\bigcap\{\mathscr{F}_{[a,b]}:a<t<b\}$, and say that η is **strongly reversible** if the conditional distributions of η and $R(\eta)$ given $\mathscr{F}_0$ are the same almost surely. For example, if θ is a uniform random element of $[0,2\pi]$, then the environment η defined by $\eta_t(e)=(\sin(t+\theta))_{t\in\mathbb{R}}$ for every $e\in E_d$ and $t\in\mathbb{R}$ is a stationary reversible Markovian environment that is not strongly reversible.

Let $\mathbb{Z}^d_\infty=\mathbb{Z}^d\cup\{\infty\}$ be the one-point compactification of $\mathbb{Z}^d$ and let $D(\mathbb{R},\mathbb{Z}^d_\infty)$ be the space of $\mathbb{Z}^d_\infty$-valued càdlàg functions on $\mathbb{R}$, which we equip with the Skorohod topology and associated Borel σ-algebra. The point at infinity is included to deal with the possibility of an explosion. For each starting space-time location $(u,s)\in\mathbb{Z}^d\times\mathbb{R}$ and environment $\eta\in\Omega$, there exists a unique probability measure $\mathbb{P}^\eta_{u,s}$ on $D(\mathbb{R},\mathbb{Z}^d_\infty)$ under which the coordinate process $(X_t)_{t\in\mathbb{R}}$ is an inhomogeneous continuous time Markov Chain on $\mathbb{Z}^d$ starting at u at time s and with self-adjoint time-dependent generator $(\mathscr{L}^\eta_t)_{t\in\mathbb{R}}$ defined by

$$\mathscr{L}^\eta_t f(x)=\sum_{y\sim x}\eta_t(\{x,y\})(f(y)-f(x)). \tag{3.2}$$

We denote the transition probabilities of this Markov chain by $P^\eta_{t_1,t_2}(u,v)=\mathbb{P}^\eta_{u,t_1}(X_{t_2}=v)$ for each t_1,t_2 and $u,v\in\mathbb{Z}^d$. We say that an environment η is **non-explosive** if $\mathbb{P}^\eta_{u,s}$ is supported on paths that make at most finitely many jumps in any bounded interval of time for every $u\in V$ and $s\in\mathbb{R}$.

A Poissonian reformulation. As usual, one can equivalently define the random walk in the environment η using Poisson processes rather than generators. We first briefly recall how point processes in $E_d\times\mathbb{R}$ can be used to define walks. Let $\mathscr{D}$ be the set of subsets $U\subset\mathbb{R}\times E_d$ that are discrete (i.e. consist only of isolated points), and for which $U\cap(E_d\times\{t\})$ contains at most one point for each $t\in\mathbb{R}$. For each $U\in\mathscr{D}$, let $J=J(U)$ be the set of space-time points $(u,t)\in\mathbb{Z}^d\times\mathbb{R}$ such that $(\{u,v\},t)\in U$ for some neighbour v of u. Given $U\in\mathscr{D}$ and a space-time coordinate $(u,t)\notin J(U)$, we define the induced cádlág path $F_{u,t}(U)=(F_{u,t}(U)_s)_{s\in\mathbb{R}}\in D(\mathbb{R},\mathbb{Z}^d)$ which starts with $F_{u,t}(U)_t=u$ and follows the points of U forwards and backwards in time, traversing an edge $e=\{x,y\}$ at time $s\geq t$ if $\lim_{\varepsilon\downarrow 0}F_{u,t}(U)_{s-\varepsilon}\in\{x,y\}$ and $(e,s)\in U$ and, similarly, traversing an edge $e=\{x,y\}$ at time $s\leq t$ if $\lim_{\varepsilon\downarrow 0}F_{u,t}(U)_{s+\varepsilon}\in\{x,y\}$ and $(e,s)\in U$. We define $T_\infty+$ and T^-_∞ to be the forward and backward explosion times of $F_{u,t}(U)$, and set $F_{u,t}(U)_s=\infty$ for all $s\geq T^+_\infty$ and $s\leq T^-_\infty$.

Translation and reflection equivariance. An important property of this construction is that for any $U \in \mathscr{D}$ and any two space-time points $(u,s),(v,t) \in (\mathbb{Z}^d \times \mathbb{R}) \setminus J(U)$ we have that

$$F_{u,s}(U)_t = v \iff F_{v,t}(U)_s = u \iff F_{u,s}(U) = F_{v,t}(U), \tag{3.3}$$

where the final equality is an equality of functions. Indeed, if we start a particle at (u,s) then follow the points of U forwards in time until we hit v at time $t \geq s$, then if we instead start at v at time t and follow the points of U backwards in time until time s, we will end up at u. A further important property of the map $F : \mathscr{D} \times \mathbb{Z}^d \times \mathbb{R}$ is that it is equivariant with respect to space-time shifts and time-reversals. That is, if we define the space-time shifts

$$\begin{aligned} \tau_{x,t} : \mathscr{D} &\longrightarrow \mathscr{D} \\ U &\longmapsto \big\{(e-x, s-t) : (e,s) \in U\big\} \end{aligned} \qquad \begin{aligned} \tau_{x,t} : D(\mathbb{R}, \mathbb{Z}^d_\infty) &\longrightarrow D(\mathbb{R}, \mathbb{Z}^d_\infty) \\ (\zeta_s)_{s\in\mathbb{R}} &\longmapsto (\zeta_{s-t} - x)_{s\in\mathbb{R}} \end{aligned}$$

for each $x \in \mathbb{Z}^d$ and $t \in \mathbb{R}$ and the time-reversal maps

$$\begin{aligned} R : \mathscr{D} &\longrightarrow \mathscr{D} \\ U &\longmapsto \big\{(e, -s) : (e,s) \in U\big\} \end{aligned} \qquad \begin{aligned} R : D(\mathbb{R}, \mathbb{Z}^d_\infty) &\longrightarrow D(\mathbb{R}, \mathbb{Z}^d_\infty) \\ (\zeta_s)_{s\in\mathbb{R}} &\longmapsto \Big(\lim_{\varepsilon\downarrow 0} \zeta_{-s+\varepsilon}\Big)_{s\in\mathbb{R}} \end{aligned}$$

then we have that

$$\tau_{x,t}(F_{u,s}(U)) = F_{u-x,s-t}(\tau_{x,t}(U)) \qquad \text{and} \qquad R(F_{u,s}(U)) = F_{u,-s}(R(U))$$

for every $(x,t) \in \mathbb{Z}^d \times \mathbb{R}$, $U \in \mathscr{D}$, and $(u,s) \in (\mathbb{Z}^d \times \mathbb{R}) \setminus J(U)$.

Given an environment η, we may take U to be the inhomogeneous Poisson process on $E_d \times \mathbb{R}$ with intensity η, which belongs to $\mathscr{D}$ almost surely since η is locally integrable. It is a standard and easily verified fact that the resulting process $F_{u,t}(U)$ then has law $\mathbb{P}^\eta_{u,t}$ for each $u \in \mathbb{Z}^d$ and $t \in \mathbb{R}$. Fixing η and taking expectations over U in (3.3) therefore yield the detailed-balance equations

$$P^\eta_{s,t}(u,v) = P^\eta_{t,s}(v,u), \tag{3.4}$$

which also follow directly by self-adjointness of the generators. Moreover, if U is a Poisson process with intensity η, then $R(U)$ is a Poisson process with intensity $R(\eta)$, and it follows that if $X = (X_t)_{t\in\mathbb{R}}$ has law $\mathbb{P}^\eta_{u,s}$, then $R(X)$ has law $\mathbb{P}^{R(\eta)}_{u,-s}$. It follows in particular that if η is a stationary reversible random environment and $X = (X_t)_{t\in\mathbb{R}}$ is the associated random walk started at (u,s), then X and $R(X)$ have the same marginal distribution (the *conditional* distributions of these processes given η need not be the same).

3.2.1 Moment conditions

Let $d \geq 1$ and let $\eta \in \Omega$ be a stationary random environment on $\mathbb{Z}^d$. Recall that we write $\eta_t(x) := \sum_{y \sim x} \eta_t(\{x,y\})$ for the total conductance of all edges incident to x at time t. For each $p \geq 1$ we define the **infinitesimal p-norm** $\|\eta\|_p$ of η to be

$$\|\eta\|_p := \sup_{[a,b] \subset \mathbb{R}} \frac{1}{b-a} \mathbb{E}\left[\left(\int_a^b \eta_s(0)\, \mathrm{d}s\right)^p\right]^{1/p} = \limsup_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \mathbb{E}\left[\left(\int_{[0,\varepsilon]} \eta_s(0)\, \mathrm{d}s\right)^p\right]^{1/p},$$

where the equivalence of these two quantities follows by stationarity and Minkowski's inequality. Note that $\|\eta\|_p$ is increasing in $p \geq 1$ and that if η is, say, bounded and a.s. cádlág, so that $\eta_t(x)$ is well-defined pointwise, then $\|\eta\|_p = \|\eta_t(x)\|_p$ for every $x \in \mathbb{Z}^d$ and $t \in \mathbb{R}$.

The next proposition shows that first and second moment bounds on the total conductance at a *fixed* vertex imply first and second moment bounds on the number of times the walk jumps. We will deduce in particular that $\|\eta\|_1 < \infty$ is a sufficient condition for non-explosivity, recovering [12, Lemma 4.1]. For each two integers $p \geq 1$ and $1 \leq \ell \leq p$, we write $\left\{ {p \atop \ell} \right\}$ for the Stirling numbers of the second kind, which are defined to be the unique non-negative integers such that $x^p = \sum_{\ell=1}^p \left\{ {p \atop \ell} \right\} \ell! \binom{x}{\ell}$ for every $x \in \mathbb{R}$. (Equivalently, $\left\{ {p \atop \ell} \right\}$ is the number of ways to partition a set of size p into ℓ non-empty subsets.)

Proposition 25. *Let $d \geq 1$, let η be a stationary random environment on $\mathbb{Z}^d$, let $(u,s) \in \mathbb{Z}^d \times \mathbb{R}$ be a space-time location, and let $X = (X_t)_{t \in \mathbb{R}}$ be the associated random walk started at the origin at time zero. For each $0 \leq a < b$ let $N[a,b]$ denote the cardinality of the set of jump times $\{t \in [a,b] : X_{t^-} \neq X_t\}$. Then*

$$\mathbb{E}\left[N[a,b]^p\right] \leq \sum_{\ell=1}^{p} \left\{ {p \atop \ell} \right\} \ell! |a-b|^\ell \|\eta\|_\ell^\ell$$

for every integer $p \geq 1$. In particular, if $\|\eta\|_1 < \infty$, then η is non-explosive almost surely.

The most important consequence of this theorem is the statement that if $\|\eta\|_p < \infty$ for some integer $p \geq 1$, then $\mathbb{E}\left[N[a,b]^p\right] < \infty$ for every $a < b$. We will only use the cases $p = 1,2$ of this proposition, but prove the general case for possible future applications since it is not much more work.

The proof of Proposition 25 will rely on the construction of the *censored random walk* in finite volume, which we now introduce. Let η be a stationary random environment on $\mathbb{Z}^d$, let U be a Poisson process with intensity η, and let $X = F_{0,0}(U)$ be the associated random walk in η started at $(0,0)$. Consider the sequence of l_1 boxes $B_k = B(0,k) \cap \mathbb{Z}^d$ for $k \geq 1$, and let

$E_{d,k}$ be the set of edges of $\mathbb{Z}^d$ with both endpoints in B_k. For each $k \geq 1$, let S_k be a uniform random element of B_k independent of η and U, and let $X^k = F_{S_k,0}(U)$ be a random walk in η started at $(S_k, 0)$. Stationarity of η implies that $X^k - S_k = (X_t^k - S_k)_{t\in\mathbb{R}}$ and X have the same distribution for every $k \geq 1$.

For each $k \geq 1$, let $U^k = U \cap (E_{d,k} \times \mathbb{R})$, and define the **censored random walk** $Z^k = (Z_t^k)_{t\in\mathbb{R}} = F_{S_k,0}(U^k)$. In other words, the censored random walk Z^k is coupled with the random walk X^k by setting $Z_0^k = X_0^k$, and then letting Z^k follow the same Poisson point process U as X^k, forwards and backwards in time, but ignoring the edges which lead out of B_k. Thus, Z^k is guaranteed to equal to X^k up until the first time X^k leaves the ball B_k. Observe that censored random walks cannot explode since the rate of transition of the walk at any time is bounded above by the total conductance of all the edges contained within the box, which is finite by assumption.

Note that if η is a stationary Markovian random environment and $k \geq 1$, then both (η, X) and (η, Z^k) are Markov processes in the sense that the future and the past are conditionally independent given the present; see Section 3.2.2 for details. However, the censored random walk has the advantage that the associated Markov process admits a stationary *probability* measure. Indeed, we will argue more generally that if η is a stationary random environment then (η, Z^k) is time-stationary in the sense $\tau_{0,t}(\eta, Z^k) := (\tau_{0,t}(\eta), \tau_{0,t}(Z^k))$ has the same distribution as (η, Z^k) for every $k \geq 1$ and $t \in \mathbb{R}$.

Lemma 26. *Let $d \geq 1$ and let η be a stationary random environment. Then the processes $(\eta_t, Z_t^k)_{t\in\mathbb{R}}$ are stationary for each $k \geq 1$.*

Proof. Fix $k \geq 1$. Let U be a Poisson process with intensity η and let U^k be defined as above. We have by (3.3) that if $(u,s),(v,t) \notin J(U)$, then

$$F_{u,s}(U^k)_t = v \iff F_{v,t}(U^k)_s = u \iff F_{u,s}(U^k) = F_{v,t}(U^k). \tag{3.5}$$

One implication of this is that for any $s,t \in \mathbb{R}$, the function $\sigma_{s,t} : B_k \to B_k$ given by $\sigma_{s,t}(u) = [F_{s,u}(U^k)]_t$ is almost surely a bijection with the property that

$$F_{s,u}(U^k) = F_{t,\sigma_{s,t}(u)}(U^k)$$

for every $u \in B_k$. Letting S_k be a uniform random element of B_k independent of η and U, we deduce that S_k and $\sigma_{s,t}(S_k)$ have the same conditional distribution given η and U and hence that

$$\tau_{0,t}\Big(\eta, F_{0,S_k}(U^k)\Big) = \tau_{0,t}\Big(\eta, F_{t,\sigma_{0,t}(S_k)}(U^k)\Big) \sim \tau_{0,t}\Big(\eta, F_{t,S_k}(U^k)\Big) \sim \Big(\eta, F_{0,S_k}(U^k)\Big)$$

for every $t \in \mathbb{R}$, where we used stationarity of η and shift-equivariance of F in the final equality in distribution. This completes the proof of stationarity. $\square$

We will deduce Proposition 25 from the following analogous statement for the censored random walk.

Lemma 27. *Let $d \geq 1$, let η be a stationary random environment on $\mathbb{Z}^d$, let $k \geq 1$, and let Z^k be the censored random walk in η. For each $0 \leq a < b$ let $N_k[a,b]$ denote the cardinality of the set of jump times $\{t \in [a,b] : Z^k_{t^-} \neq Z^k_t\}$. Then*

$$\mathbb{E}\left[N_k[a,b]^p\right] \leq \sum_{\ell=1}^{p} \left\{ {p \atop \ell} \right\} \ell! |a-b|^\ell \|\eta\|_\ell^\ell$$

for every integer $p \geq 1$.

Proof. By stationarity, we can without loss of generality assume that $a = 0$. We fix $b \geq 0$ and $k \geq 1$, and write $N = N_k$, and $Z = Z^k$. For each $n \in \mathbb{N}$ and $i \in \mathbb{Z}$, define

$$A_{i,n} = \mathbb{1}\left(N\left[\frac{(i-1)b}{n}, \frac{ib}{n}\right] > 0\right) \qquad \text{and} \qquad \Sigma_n = \sum_{i=1}^{n} A_{i,n}.$$

Since $N = \lim_{n\to\infty} \Sigma_n$ almost surely and (Σ_{2^n}) is a monotone increasing sequence, the monotone convergence theorem implies that $\mathbb{E}\left[N^p\right] = \lim_{n\to\infty} \mathbb{E}\left[\Sigma_{2^n}^p\right]$ for every $p \geq 0$. Since $\mathbb{E}\left[\Sigma_{2^n}^p\right] = \sum_{\ell=1}^{p} \left\{ {p \atop \ell} \right\} \ell! \mathbb{E}\binom{\Sigma_{2^n}}{\ell}$ it therefore suffices to prove that

$$\mathbb{E}\binom{\Sigma_n}{\ell} = \sum_{1 \leq i_1 < \cdots < i_\ell \leq n} \mathbb{E}\left[\prod_{j=1}^{\ell} A_{i_j,n}\right] \leq b^\ell \|\eta\|_\ell^\ell$$

for every $\ell \geq 1$. Writing $\mathbf{E}^\eta$ for expectations conditional on the environment η and the uniform starting point $S_k = Z_0^k \in B_k$, we will prove by induction on ℓ that the stronger inequality

$$\mathbb{E}\left[\mathbf{E}^\eta\left[\prod_{j=1}^{\ell} A_{i_j,n}\right]^q\right] \leq \left(\frac{b\|\eta\|_{q\ell}}{n}\right)^{q\ell} \tag{3.6}$$

holds for every $n \geq 1$, $\ell \geq 0$, $q \geq 1$, and every increasing sequence $i_1 < i_2 < \ldots < i_\ell$ in $\mathbb{R}$, where we take the empty product to be 1. (Note that we do not assume that q is an integer.)

The $\ell = 0$ case holds vacuously. Assume that the claim holds for some $\ell \geq 0$ and let $i_0 < \ldots < i_\ell$ be an increasing sequence of times. Then we have by stationarity (Lemma 26) and the fact that $(Z^k_t)_{t\leq 0}$ and $(Z^k_t)_{t\geq 0}$ are conditionally independent given η and S_k that

$$\mathbb{E}\left[\mathbf{E}^\eta\left[\prod_{j=0}^{\ell} A_{i_j,n}\right]^q\right] = \mathbb{E}\left[\mathbf{E}^\eta\left[\prod_{j=0}^{\ell} A_{i_j-i_0,n}\right]^q\right] \leq \mathbb{E}\left[\mathbf{E}^\eta\left[A_{0,n}\right]^q \cdot \mathbf{E}^\eta\left[\prod_{j=1}^{\ell} A_{i_j,n}\right]^q\right].$$

Applying Hölder's inequality and the induction hypothesis yields that

$$\begin{aligned}\mathbb{E}\left[\mathbf{E}^\eta\left[\prod_{j=0}^{\ell} A_{i_j,n}\right]^q\right] &\leq \mathbb{E}\Big[\mathbf{E}^\eta\left[A_{0,n}\right]^{q(\ell+1)}\Big]^{1/(\ell+1)}\mathbb{E}\left[\mathbf{E}^\eta\left[\prod_{j=1}^{\ell} A_{i_j,n}\right]^{q(\ell+1)/\ell}\right]^{\ell/(\ell+1)} \\ &\leq \left(\frac{b\|\eta\|_{q(\ell+1)}}{n}\right)^{q\ell}\mathbb{E}\Big[\mathbf{E}^\eta\left[A_{0,n}\right]^{q(\ell+1)}\Big]^{1/(\ell+1)}. \quad (3.7)\end{aligned}$$

Conditioned on η and $Z_0 = S_k$, the indicator random variable $A_{0,n}$ is equal to 1 if and only if at least one of the Poisson clocks attached to an edge incident to Z_0 rings in the interval $[-b/n, 0]$, so that

$$\mathbf{E}^\eta\left[A_{0,n}\right] = 1 - \exp\left[-\int_{-b/n}^{0} \eta_t(Z_0)\,\mathrm{d}t\right] \leq \int_{-b/n}^{0} \eta_t(Z_0)\,\mathrm{d}t$$

and hence by stationarity of η that

$$\mathbb{E}\Big[\mathbf{E}^\eta\left[A_{0,n}\right]^{q(\ell+1)}\Big] \leq \mathbb{E}\left[\left(\int_{-b/n}^{0} \eta_t(Z_0)\,\mathrm{d}t\right)^{q(\ell+1)}\right] \leq \left(\frac{b\|\eta\|_{q(\ell+1)}}{n}\right)^{q(\ell+1)}.$$

Substituting this estimate into (3.7) completes the induction step and hence the proof of (3.6). □

Proof of Proposition 25. Fix $b > 0$. Lemma 27 implies that the first moment of $\max_{0\leq t\leq b} d(Z^k_t, S_k) \leq N_k[0,b]$ is bounded above uniformly in k. We also note that for any fixed distance $l > 0$ the probability that the distance between S_k and the boundary of B_k is less than l decreases to zero as k tends to infinity. Combining these two observations, the probability that $Z^k - S_k$ hits the boundary of $B_k - S_k$ before time b tends to zero as $k \to \infty$. Since Z^k and X^k are equal up to the first time the boundary is hit, and, by stationarity, the law of $(X^k_t - S_k)_{0\leq t\leq b}$ is equal to the law of $(X_t)_{0\leq t\leq b}$, it follows that $(Z^k_t - S_k)_{0\leq t\leq b}$ converges in distribution to $(X_t)_{0\leq t\leq b}$ as $k \to \infty$. It follows that the law of $(N_k[a,b])_{0\leq a\leq b}$ converges weakly to the law of $(N[a,b])_{0\leq a\leq b}$, and hence by Fatou's lemma that

$$\mathbb{E}\left[N[a,b]^p\right] \leq \liminf_{k\to\infty} \mathbb{E}\left[N_k[a,b]^p\right] \leq \sum_{\ell=1}^{p} \left\{ {p \atop \ell} \right\} \ell! |a-b|^\ell \|\eta\|_\ell^\ell,$$

where the second inequality follows by Lemma 27. □

3.2.2 Diffusive upper bounds via Markov-type inequalities

In this section we use *Markov-type inequalities* to establish diffusive upper bounds on the displacement of random walks in stationary reversible Markovian environments, generalising an argument of Peres, Stauffer, Steif [292, Theorem 1.9] from the setting of dynamical percolation. To do this, we will need a version of the Markov-type inequality that applies to Markov processes defined on uncountable state spaces and that need not be well-defined pointwise. The proof of this inequality is in fact very similar to the usual discrete-time proof of Naor, Peres, and Sheffield [281] as presented in [259, Lemma 13.15]. Markov-type inequalities were first studied by Keith Ball in his work on the Lipschitz extension problem [40], and have recently found many important applications in probability theory including e.g. [149, 173, 244, 292, 293].

We now introduce the relevant definitions. Let $\mathbb{X}$ be a Polish space, and let $\mathscr{X} = \mathscr{X}(\mathbb{R},\mathbb{X})$ be the set of Borel-measurable functions from $\mathbb{R}$ to $\mathbb{X}$ modulo almost-everywhere equivalence. For each $s \in \mathbb{R}$ we define the time-shift $\tau_s : \mathscr{X} \to \mathscr{X}$ by $\tau_s\zeta(t) = \zeta(t-s)$ for every $\zeta \in \mathscr{X}$ and $t \in \mathbb{R}$, and define the reversal $R : \mathscr{X} \to \mathscr{X}$ by $R(\zeta)(t) = \zeta(-t)$ for every $\zeta \in \mathscr{X}$ and $t \in \mathbb{R}$. Let Z be a random variable taking values in $\mathscr{X}$, and for each $a < b$ let $\mathscr{F}_{[a,b]}$ be the σ-algebra generated by the restriction of Z to $[a,b]$. We say that Z is a **Markov process** if $\mathscr{F}_{[a_1,a_2]}$ and $\mathscr{F}_{[c_1,c_2]}$ are conditionally independent given $\mathscr{F}_{[b_1,b_2]}$ whenever $a_2 < b_2$ and $c_1 > b_1$ (that is, if the past and the future are conditionally independent given the present). We say that Z is **stationary** if $\tau_s\mathscr{X}$ has the same distribution as Z for every $s \in \mathbb{R}$, and that Z is **reversible** if $R(Z)$ and Z have the same distribution. For each $t \in \mathbb{R}$, we define the **instantaneous sigma-algebra** $\mathscr{F}_t = \bigcap\{\mathscr{F}_{[a,b]} : a < t < b\}$, and say that Z is **strongly reversible** if the conditional distributions of Z and $R(Z)$ given $\mathscr{F}_0$ are the same almost surely.

Proposition 28. *Let $\mathbb{X}$ be a Polish space, and let $Z \in \mathscr{X}(\mathbb{R},\mathbb{X})$ be a stationary, strongly reversible Markov process. Let $d \geq 1$ and let $f : \mathscr{X} \to \mathbb{R}^d$ be measurable with respect to the instantaneous sigma-algebra $\mathscr{F}_0$ and reversible in the sense that $f(Z) = f(R(Z))$ almost surely. Then we have that*

$$\mathbb{E}\left[\max_{0\leq m\leq n} \|f(\tau_{2mt}Z) - f(Z)\|_2^2\right] \leq 25n\mathbb{E}\Big[\|f(\tau_t Z) - f(Z)\|_2^2\Big]. \tag{3.8}$$

for every $n \geq 1$ and $t > 0$ and hence that

$$\mathbb{E}\left[\operatorname*{ess\,sup}_{0\leq s\leq t} \|f(\tau_s Z) - f(Z)\|_2^2\right] \leq \frac{25t}{2}\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}\left[\|f(\tau_\varepsilon Z) - f(Z)\|_2^2\right] \tag{3.9}$$

for every $t > 0$.

Remark 4. If θ is a uniform random element of $[0,2\pi]$ then $(X_t)_{t\in\mathbb{R}} = (\sin(t+\theta))_{t\in\mathbb{R}}$ is a stationary reversible Markov process $X : \mathbb{R} \to \mathbb{R}$ that is not strongly reversible and does not satisfy the conclusions of the Markov-type inequality. Indeed, if we consider the identity function $f : \mathbb{R} \to \mathbb{R}$ then

$$\mathbb{E}\left[\|f(X_t) - f(X_0)\|_2^2\right] = \int_0^{2\pi} [\sin(t+\theta) - \sin(\theta)]^2 \, d\theta = 2\pi\,(1-\cos(t)) = \Theta(t^2) \qquad \text{as } t \downarrow 0,$$

so that $\mathbb{E}\left[\|f(X_{nt}) - f(X_0)\|_2^2\right] \gg n\mathbb{E}\left[\|f(X_t) - f(X_0)\|_2^2\right]$ when t is small and n is large. Further processes with similar properties include e.g. *piecewise deterministic Markov processes* and the *integrated Ornstein-Uhlenbeck process mod* 1.

Proof of Proposition 28. Without loss of generality we may take $d = 1$, the higher-dimensional cases following by summing the inequalities (3.8) and (3.9) over the coordinates of f. We may also assume that f is bounded, truncating f to $[-r,r]$ and using monotone convergence to take the limit as $r \to \infty$ otherwise. Note that if θ is a uniform random number in $[1/2,1]$ and $N = N(\theta,n) = \lceil nt/2\theta \rceil$ for each $n \geq 1$ then $\max_{0\leq m\leq N} \|f(\tau_{2m\theta t/n}Z) - f(Z)\|_2^2$ converges in probability to $\operatorname{ess\,sup}_{0\leq s\leq t} \|f(\tau_s Z) - f(Z)\|_2^2$ as $n \to \infty$ (this follows by e.g. the Lebesgue differentiation theorem), so that (3.9) follows from (3.8) and Fatou's lemma.

The main idea, taken from [281], is to write the maximum we are interested in terms of two martingales, one going forwards in time and the other backwards in time, and then use Doob's L^2 maximal inequality. For each $t \in \mathbb{R}$, let $\mathscr{G}_t^{\rightarrow} = \bigcap_{s>t} \mathscr{F}_{(-\infty,s]}$ and let $\mathscr{G}_t^{\leftarrow} = \bigcap_{s<t} \mathscr{F}_{[s,\infty)}$, so that $\mathscr{F}_t \subseteq \mathscr{G}_t^{\rightarrow} \cap \mathscr{G}_t^{\leftarrow}$ for each $r \in \mathbb{R}$. Since Z is a Markov process, $\mathscr{F}_s$ and $\mathscr{G}_t^{\rightarrow}$ are conditionally independent given $\mathscr{F}_t$ when $s > t$, while $\mathscr{F}_s$ and $\mathscr{G}_t^{\leftarrow}$ are conditionally independent given $\mathscr{F}_t$ when $s < t$. Fix $t > 0$ and $n \in \mathbb{N}$, and for each $1 \leq m \leq 2n$ let

$$D_m^{\rightarrow} = f(\tau_{mt}Z) - \mathbb{E}\left[f(\tau_{mt}Z) \mid \mathscr{G}_{(m-1)t}^{\rightarrow}\right] = f(\tau_{mt}Z) - \mathbb{E}\left[f(\tau_{mt}Z) \mid \mathscr{F}_{(m-1)t}\right],$$

where the almost-sure equivalence of these two quantities follows from the assumption that Z is a Markov-process and that f is $\mathscr{F}_0$-measurable. In particular, the process $(D_i^{\rightarrow})_{m=1}^{2n}$ is a martingale difference sequence with respect to the filtration $(\mathscr{G}_{mt}^{\rightarrow})_{m=0}^{n}$. Similarly, for each $1 \leq m \leq 2n$ we define

$$\begin{aligned} D_m^{\leftarrow} = f(\tau_{(2n-m)t}Z) - \mathbb{E}\left[f(\tau_{(2n-m)t}Z) \mid \mathscr{G}_{(2n-m+1)t}^{\leftarrow}\right] \\ = f(\tau_{(2n-m)t}Z) - \mathbb{E}\left[f(\tau_{(2n-m)t}Z) \mid \mathscr{F}_{(2n-m+1)t}\right]. \end{aligned}$$

As before, the almost-sure equivalence of these quantities follows from the assumption that Z is a Markov-process and that f is $\mathscr{F}_0$-measurable. In particular, the process $(D_m^{\leftarrow})_{m=1}^{2n}$ is a martingale difference sequence with respect to the filtration $(\mathscr{G}_{(2n-m)t}^{\leftarrow})_{m=0}^{n}$. Moreover, for each $2 \leq m \leq 2n$ we have that

$$D_m^{\rightarrow} - D_{2n-m+2}^{\leftarrow} = f(\tau_{mt}Z) - f(\tau_{(m-2)t}Z) - \mathbb{E}\left[f(\tau_{mt}Z) - f(\tau_{(m-2)t}Z) \mid \mathscr{F}_{(m-1)t}\right] \\ = f(\tau_{mt}Z) - f(\tau_{(m-2)t}Z) \quad (3.10)$$

almost surely, where we used stationarity and strong reversibility to deduce that $f(\tau_{mt}Z)$ and $f(\tau_{(m-2)t}Z)$ have the same conditional distribution given $\mathscr{F}_{(m-1)t}$ almost surely and hence that the central conditional expectation is almost surely zero. We obtain by algebra that

$$f(\tau_{2kt}Z) - f(Z) = \sum_{m=1}^{k} D_{2m}^{\rightarrow} - \sum_{m=1}^{k} D_{n-2m+2}^{\leftarrow}$$

for every $1 \leq k \leq n$. It follows that

$$\max_{0\leq k\leq n} |f(\tau_{2kt}Z) - f(Z)| \leq \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2m}^{\rightarrow}\right| + \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2n-2m+2}^{\leftarrow}\right| \\ \leq \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2m}^{\rightarrow}\right| + \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2m}^{\leftarrow}\right| + \left|\sum_{m=1}^{n} D_{2m}^{\leftarrow}\right|$$

and hence by Cauchy-Schwarz that

$$\max_{0\leq k\leq n} |f(\tau_{2kt}Z) - f(Z)|^2 \leq \frac{5}{2} \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2m}^{\rightarrow}\right|^2 + \frac{5}{2} \max_{0\leq k\leq n} \left|\sum_{m=1}^{k} D_{2m}^{\leftarrow}\right|^2 + 5\left|\sum_{m=1}^{n} D_{2m}^{\leftarrow}\right|^2.$$

Applying Doob's L^2 maximal inequality and the orthogonality of martingale differences, we obtain that

$$\mathbb{E}\left[\max_{0\leq k\leq n} |f(\tau_{2kt}Z) - f(Z)|^2\right] \leq 10 \sum_{m=1}^{n} \mathbb{E}\left[(D_{2m}^{\rightarrow})^2\right] + 15 \sum_{m=1}^{n} \mathbb{E}\left[(D_{2m}^{\leftarrow})^2\right].$$

Using stationarity and reversibility once more, we obtain that

$$\begin{aligned}\mathbb{E}\left[\max_{0\le k\le n}|f(\tau_{2kt}Z)-f(Z)|^2\right] &\le 25n\mathbb{E}\left[\left(f(\tau_t Z)-\mathbb{E}\left[f(\tau_t Z)\mid \mathscr{F}_0\right]\right)^2\right]\\ &=25n\mathbb{E}\left[\left(f(\tau_t Z)-f(Z)-\mathbb{E}\left[f(\tau_t Z)-f(Z)\mid \mathscr{F}_0\right]\right)^2\right]\\ &=25n\mathbb{E}\left[\mathrm{Var}(f(\tau_t Z)-f(Z)\mid \mathscr{F}_0)\right]\le 25n\,\mathrm{Var}(f(\tau_t Z)-f(Z)),\end{aligned}$$

which implies the claim. □

Proposition 28 has the following corollary for random walks in reversible random environments.

Corollary 29. *Let $d\ge 1$, let η be a stationary, strongly reversible Markovian random environment on $\mathbb{Z}^d$ and let $X=(X_t)_{t\in\mathbb{R}}$ be the associated random walk started at the origin at time zero. If $\|\eta\|_2<\infty$, then*

$$\mathbb{E}\left[\max_{-t\le s\le t}\|X_s-X_0\|_2^2\right]\le 25t\,\|\eta\|_1$$

for every $t\ge 0$.

Proof. Let $k\ge 1$ and let $(Z_t^k)_{t\in\mathbb{R}}$ be the censored random walk started at a uniform random element S_k of B_k as in Section 3.2.1. By Lemma 26, (η_t,Z_t^k) is a stationary Markov process. Moreover, if we consider this process to take values in the space of measurable functions $\mathscr{X}=\mathscr{X}(\mathbb{R},\mathbb{R}^{E_d}\times\mathbb{Z}^d)$ then it is strongly reversible: this follows by time-reversal equivariance of F and the fact that, given η, the reversed Poisson process $R(U)$ has the same conditional distribution as a Poisson process with intensity $R(\eta)$. Thus, we may apply Proposition 28 to the function $f:\mathscr{X}\to\mathbb{R}^d$ given by $f(\omega,\zeta)=\zeta_0$, to obtain that

$$\mathbb{E}\left[\max_{0\le s\le t}\|Z_t^k-Z_0^k\|_2^2\right]\le\frac{25t}{2}\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}\left[\|Z_\varepsilon^k-Z_0^k\|_2^2\right]$$

for every $t>0$ and $k\ge 1$. Since the Euclidean displacement is trivially bounded by the total number of jumps, we obtain that

$$\begin{aligned}\mathbb{E}\left[\max_{0\le s\le t}\|Z_t^k-Z_0^k\|_2^2\right]&\le\frac{25t}{2}\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{E}\left[N_k[0,\varepsilon]^2\right]\\ &\le\frac{25t}{2}\limsup_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\left(\varepsilon\|\eta\|_1+2\varepsilon^2\|\eta\|_2^2\right)=\frac{25t}{2}\|\eta\|_1.\end{aligned}$$

Taking the limit as $k \to \infty$, it follows by a similar weak convergence and Fatou argument to that used in the proof of 25 that

$$\mathbb{E}\left[\max_{0\le s\le t}\|X_t - X_0\|_2^2\right] \le \frac{25t}{2}\|\eta\|_1$$

for every $t \ge 0$ also. The claimed two-sided version of this inequality follows by reversibility. □

Remark 5. It has been pointed out to us by a referee that it should be possible to prove diffusive upper bounds using the Kipnis-Varadahn method of martingale approximation of additive functionals of reversible Markov processes [31, 220]. We prefer to use Markov-type inequalities as they are self-contained and more familiar to us, and hope that we will help popularize the use of these powerful inequalities within the RWRE community. Let us remark further that Markov-type inequalities are particularly useful for obtaining sharp quantitative control of the limiting variance of RWRE processes, and indeed will be used in forthcoming work of the second author to study the asymptotic diffusivity of random walks on slightly supercritical percolation clusters.

3.3 Proof of the main theorem

In this section will will prove Theorem 20 and its corollaries. We begin with the following general criterion for infinite collisions at integer times, from which our main theorems will be deduced. Recall that we write $\mathbb{E}^\eta$ for conditional expectations given the environment η.

Proposition 30. *Let $d \ge 1$, let $\eta : \mathbb{R}\times E_d \to [0,\infty)$ be a stationary, non-explosive random environment on $\mathbb{Z}^d$ and let $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ be random walks in η, both started at the origin at time zero, that are conditionally independent given η. Then we have the implication*

$$\left(\mathbb{E}^\eta \sum_{n\ge 0}\mathbb{1}_{\{X_{-n}=Y_{-n}\}} = \infty \textit{ almost surely}\right) \Rightarrow \left(\sum_{n\ge 0}\mathbb{1}_{\{X_n=Y_n\}} = \infty \textit{ almost surely}\right). \tag{3.11}$$

The proof of this proposition is adapted from the methods of [195], and relies on the *mass-transport principle* for $\mathbb{Z}^d$. Recall that a function $f : \mathbb{Z}^d \times \mathbb{Z}^d \to [0,\infty]$ is said to be a *transport function* if it is diagonally invariant in the sense that $f(x,y) = f(x+z,y+z)$ for every $x,y,z \in \mathbb{Z}^d$. The mass-transport principle for $\mathbb{Z}^d$ states that

$$\sum_{x\in\mathbb{Z}^d} f(0,x) = \sum_{x\in\mathbb{Z}^d} f(x,0).$$

for every transport function f.

Proof. Suppose that $\mathbb{E}^\eta \sum_{n\geq 0} \mathbb{1}_{\{X_{-n}=Y_{-n}\}} = \infty$ almost surely. Recall that $P^\eta_{t_1,t_2}(\cdot,\cdot)$ denotes the transition probabilities of the random walk conditional on the environment η. For each $u \in \mathbb{Z}^d$ and $n \in \mathbb{Z}$ we let $q^\eta_{\mathrm{fin}}(u,n)$ denote the conditional probability given η that two conditionally independent random walks started at the space-time location (u,n) occupy the same position for only finitely many positive integer times $m \geq n$, and let $q^\eta_0(u,n)$ denote the conditional probability that the two walks started at (u,n) do not occupy the same position at any integer time strictly greater than n. Decomposing according to the last integer time at which the two walks occupy the same position, and where they do so, we get that

$$q^\eta_{\mathrm{fin}}(u,n) = \sum_{v\in\mathbb{Z}^d} \sum_{m\geq n} P^\eta_{n,m}(u,v)^2 q^\eta_0(v,m).$$

By space-shift invariance, $f(u,v) = \sum_{m\geq 0} \mathbb{E}\left[P^\eta_{0,m}(u,v)^2 q^\eta_0(v,m)\right]$ is a transport function and we can apply the mass-transport principle to get that

$$\mathbb{E}\left[q^\eta_{\mathrm{fin}}(0,0)\right] = \mathbb{E}\left[\sum_{v\in\mathbb{Z}^d} \sum_{m\geq 0} P^\eta_{0,m}(0,v)^2 q^\eta_0(v,m)\right] = \mathbb{E}\left[\sum_{v\in\mathbb{Z}^d} \sum_{m\geq 0} P^\eta_{0,m}(v,0)^2 q^\eta_0(0,m)\right],$$

and hence by time-shift invariance applied to each term that

$$\begin{aligned}\mathbb{E}\left[q^\eta_{\mathrm{fin}}(0,0)\right] &= \mathbb{E}\left[\sum_{v\in\mathbb{Z}^d} \sum_{m\geq 0} P^\eta_{-m,0}(v,0)^2 q^\eta_0(0,0)\right] \\ &= \mathbb{E}\left[q^\eta_0(0,0) \sum_{v\in\mathbb{Z}^d} \sum_{m\geq 0} P^\eta_{0,-m}(0,v)^2\right] = \mathbb{E}\left[q^\eta_0(0,0)\mathbb{E}^\eta\left[\sum_{n\geq 0} \mathbb{1}_{\{X_{-n}=Y_{-n}\}}\right]\right].\end{aligned} \tag{3.12}$$

Since $q^\eta_{\mathrm{fin}}(0,0)$ is at most one and $\mathbb{E}^\eta \sum_{n\geq 0} \mathbb{1}_{\{X_{-n}=Y_{-n}\}} = \infty$ a.s. by assumption, we must have that $q^\eta_0(0,0) = 0$ a.s. and hence that $q^\eta_{\mathrm{fin}}(0,0) = 0$ a.s. also. This implies the claim. $\square$

Next, we note that infinite collisions at infinite times quite generally implies that the Lebesgue measure of the set of all positive collision times is infinite almost surely.

Lemma 31. *Let $d \geq 1$, let $\eta : \mathbb{R} \times E_d \to [0,\infty)$ be a stationary, non-explosive random environment on $\mathbb{Z}^d$ and let $(X_t)_{t\in\mathbb{R}}$ and $(Y_t)_{t\in\mathbb{R}}$ be random walks in η, started at x and y at time zero, that are conditionally independent given η. If the set $\{n \in \mathbb{N} : X_n = Y_n\}$ has infinite cardinality almost surely, then the set $\{t \in [0,\infty) : X_t = Y_t\}$ has infinite Lebesgue measure almost surely.*

Proof. Let U_1 and U_2 be two conditionally independent Poisson processes with intensity η and let $X^s = F_{0,s}(U_1)$ and $Y^s = F_{0,s}(U_2)$ for each $s \in \mathbb{R}$. It follows by stationarity of η that the law of (X^s, Y^s) does not depend on s. Let T be the infimal positive time at which either of the walks X^0 or Y^0 takes a jump, so that $0 < T \leq \infty$ almost surely and $(X^s, Y^s) = (X^0, Y^0)$ for all $0 \leq s < T$. Then we have that

$$\begin{aligned}\mathfrak{Leb}\{t \in [0,\infty) : X_t^0 = Y_t^0\} &= \int_0^1 \left|\{n \in \mathbb{N} : X_{n+s}^0 = Y_{n+s}^0\}\right| \mathrm{d}s \\ &\geq \int_0^{T\wedge 1} \left|\{n \in \mathbb{N} : X_{n+s}^0 = Y_{n+s}^0\}\right| \mathrm{d}s = \int_0^{T\wedge 1} \left|\{n \in \mathbb{N} : X_{n+s}^s = Y_{n+s}^s\}\right| \mathrm{d}s.\end{aligned}$$

Since $T > 0$ almost surely and the integrand $\left|\{n \in \mathbb{N} : X_{n+s}^s = Y_{n+s}^s\}\right|$ is almost surely infinite for each $s \geq 0$, it follows by Tonelli's theorem that both sides are almost surely infinite, completing the proof. $\square$

We now apply Proposition 107 and Lemma 31 to prove Theorems 20 and 21.

Proof of Theorem 21. For each $K < \infty$ and $\delta > 0$, let $A_{K,\delta} \subseteq \Omega$ be the set of environments η such that

$$\limsup_{n\to\infty} \min_{0\leq m\leq n} \mathbb{P}^\eta\left(\|X_{-m}\|_2^2 \leq Kn\right) \geq \delta.$$

By assumption, for every $\varepsilon > 0$ there exists K and δ such that $\mathbb{P}(\eta \in A_{K,\delta}) \geq 1-\varepsilon$. Thus, in view of Proposition 107 and Lemma 31, it suffices to prove that if $K < \infty$ and $\delta > 0$ then $\sum_{m=1}^\infty \mathbb{P}^\eta(X_{-m} = Y_{-m}) = \infty$ for *every* environment $\eta \in A_{K,\delta}$.

Fix $K < \infty$ and $\delta > 0$ and suppose that $\eta \in A_{K,\delta}$ holds. We can recursively define a sequence of positive integer times $n_1, n_2, \ldots$, depending on η, such that $n_{i+1} \geq 2(n_i + 1)$ for each $i \geq 1$ and

$$\min_{0\leq m\leq n_i} \mathbb{P}^\eta\left(\|X_{-m}\|_2^2 \leq Kn_i\right) \geq \frac{\delta}{2}$$

for every $i \geq 1$. For each $r \geq 1$, let $\Lambda_r \subseteq \mathbb{Z}^2$ be the set of lattice points with Euclidean norm at most r. Then there exists a constant c such that

$$\mathbb{P}^\eta(X_{-m} = Y_{-m}) \geq \sum_{x\in\Lambda_r} P_{0,-m}^\eta(0,x)^2 \geq \frac{1}{|\Lambda_r|}\left[\sum_{x\in\Lambda_r} P_{0,-m}^\eta(0,x)\right]^2 \geq \frac{c}{r^2}\mathbb{P}^\eta(X_{-m} \in \Lambda_r)^2$$

for every $m, r \geq 1$ and hence that

$$\sum_{m=n_i+1}^{n_{i+1}} \mathbb{P}^\eta(X_{-m} = Y_{-m}) \geq \frac{c}{Kn_{i+1}} \sum_{m=n_i+1}^{n_{i+1}} \mathbb{P}^\eta(\|X_{-m}\|_2^2 \leq Kn_{i+1})^2$$

$$\geq \frac{c}{2K} \min_{1 \leq m \leq n_{i+1}} \mathbb{P}^\eta\left(\|X_{-m}\|_2^2 \leq Kn_{i+1}\right)^2 \geq \frac{c\delta^2}{8K}$$

for every $i \geq 1$. Summing over $i \geq 1$, it follows that $\sum_{m=1}^{\infty} \mathbb{P}^\eta(X_{-m} = Y_{-m}) = \infty$ as claimed. □

Proof of Theorem 20. It suffices to prove that the conditions (A1) and (A2) each imply the weak diffusive estimate on the backwards process (3.1) needed to apply Theorem 21. This is obvious in the case (A2) that the backwards process satisfies a (quenched or annealed) invariance principle with Brownian scaling. (It is not a problem if the limiting covariance is random.) In the case (A1) that the environment is strongly reversible and Markovian, we have by Markov's inequality and Corollary 29 that

$$\mathbb{P}\left(\min_{m \leq n} \mathbb{P}^\eta\left(\|X_{-m}\|_2^2 \leq Kn\right) \leq \delta\right) \leq \mathbb{P}\left(\mathbb{P}^\eta\left(\max_{m \leq n} \|X_{-m}\|_2^2 > Kn\right) \geq 1 - \delta\right)$$

$$\leq \mathbb{P}\left(\mathbb{E}^\eta\left[\max_{m \leq n} \|X_{-m}\|_2^2\right] \geq K(1-\delta)n\right) \leq \frac{25}{K(1-\delta)} \|\eta\|_1$$

for every $K < \infty$, $\delta > 0$, and $n \geq 1$, and hence by Fatou's lemma that

$$\mathbb{P}\left(\limsup_{n \to \infty} \min_{m \leq n} \mathbb{P}^\eta\left(\|X_{-m}\|_2^2 \leq Kn\right) \leq \delta\right) \leq \frac{25}{K(1-\delta)} \|\eta\|_1$$

for every $K < \infty$ and $\delta > 0$. This implies the claim. □

We next prove Proposition 24, which concerns the one-dimensional case.

Proof of Proposition 24. Bounding the total displacement by the number of jumps, Proposition 25 implies that $\mathbb{E} \max_{0 \leq m \leq n} \|X_{-m}\| \leq \mathbb{E}N[-n, 0] \leq n\|\eta_1\|$ for every $n \geq 1$. In the one dimensional case, this linear bound is sufficient to guarantee that $\mathbb{E}^\eta \sum_{n \geq 0} \mathbb{1}(X_{-n} = Y_{-n}) = 0$ almost surely; the details are very similar to the proof of Theorem 21 and are omitted. □

It remains only to prove Corollary 22, which concerns the case that the two walks do not start at the same vertex, and will be deduced from Theorems 20 and 21 together with the following general lemma.

Lemma 32. *Let $d \geq 1$ and let $\eta : \mathbb{R} \times E_d \to [0,\infty)$ be an irreducible, time-ergodic, stationary random environment on $\mathbb{Z}^d$. Let $(X_t)_{t\in\mathbb{R}}$, $(X'_t)_{t\in\mathbb{R}}$, $(Y_t)_{t\in\mathbb{R}}$, and $(Z_t)_{t\in\mathbb{R}}$ be random walks in η, started at some vertices x, x, y, and z at time zero respectively, that are conditionally independent given η. If $\{n \in \mathbb{N} : X_n = X'_n\}$ is infinite almost surely, then $\{n \in \mathbb{N} : Y_n = Z_n\}$ is infinite almost surely.*

Proof of Lemma 32. By stationarity, we can without loss of generality assume that $x = y = 0$. For each $z \in \mathbb{Z}^d$ and $t \in \mathbb{R}$ we define $A_{z,t}$ to be the set of environments η for which $P_t^\eta(0,z) > 0$. We will first use irreducibility and time-ergodicity of η to prove that $\mathbb{P}(A_{z,t}) \to 1$ as $t \to \infty$ for each fixed $z \in \mathbb{Z}^d$. Irreducibility give us that there exists some $t_0 > 0$ such that $\eta \in A_{z,t_0}$ with positive probability. We deduce by stationarity and time-ergodicity that $\tau_n \eta \in A_{z,t_0}$ for infinitely many positive integers almost surely, and hence that $\mathbb{P}$(there exists $m \leq t$ such that $\tau_m \eta \in A_{z,t_0}) \to 1$ as $t \to \infty$. Since the walk always has a positive conditional probability not to move in any given time interval, we have that

$$\tau_t \eta \in A_{z,t_0} \iff P^\eta_{t,t+t_0}(0,z) > 0 \Rightarrow P^\eta_{0,t+t_0}(0,z) > 0 \iff \eta \in A_{z,t+t_0}$$

for every $t \geq 0$, and hence that

$$\mathbb{P}(\eta \in A_{z,t+t_0}) \geq \mathbb{P}(\text{there exists } m \leq t \text{ such that } \tau_m \eta \in A_{z,t_0}) \to 1$$

as $n \to \infty$ as claimed.

For each $n \in \mathbb{N}$ and $\eta \in A_{z,n}$, the event $B_{u,n} = \{X_n = 0, X'_n = z\}$ has positive conditional probability. Let Y' and Z' be random walks on η, started at $(0,n)$ and (z,n), that are conditionally independent of each other and of (X,X') given η, so that $(\tau_n Y', \tau_n Z')$ has the same marginal distribution as (Y,Z). We have by the Markov property that

$$\mathbb{P}^\eta\left(\sum_{m\geq 0} \mathbb{1}_{\{Y'_m = Z'_m\}} = \infty\right) = \mathbb{P}^\eta\left(\sum_{m\geq n} \mathbb{1}_{\{X_m = Y_m\}} = \infty \middle| B_{u,n}\right) = 1$$

almost surely on the event $A_{z,n}$, and hence by stationarity that

$$\mathbb{P}\left(\sum_{m\geq 0} \mathbb{1}_{\{Y_m = Z_m\}} = \infty\right) = \mathbb{P}\left(\sum_{m\geq 0} \mathbb{1}_{\{Y'_m = Z'_m\}} = \infty\right) \geq \mathbb{P}(A_{z,n})$$

for every $n \geq 1$. The claim follows since the right hand side tends to 1 as $n \to \infty$. $\square$

Chapter 4

[C] What Are the Limits of Universality?

Abstract. It is a central prediction of renormalisation group theory that the critical behaviours of many statistical mechanics models on Euclidean lattices depend only on the dimension and not on the specific choice of lattice. We investigate the extent to which this universality continues to hold beyond the Euclidean setting, taking as case studies Bernoulli bond percolation and lattice trees. We present strong numerical evidence that the critical exponents governing these models on transitive graphs of polynomial volume growth depend only on the volume-growth dimension of the graph and not on any other large-scale features of the geometry. For example, our results strongly suggest that percolation, which has upper-critical dimension six, has the same critical exponents on $\mathbb{Z}^4$ and the Heisenberg group despite the distinct large-scale geometries of these two lattices preventing the relevant percolation models from sharing a common scaling limit. On the other hand, we also show that no such universality should be expected to hold on fractals, even if one allows the exponents to depend on a large number of standard fractal dimensions. Indeed, we give natural examples of two fractals which share Hausdorff, spectral, topological, and topological Hausdorff dimensions but exhibit distinct numerical values of the percolation Fisher exponent τ. This gives strong evidence against a conjecture of Balankin et al. [Phys. Lett. A 2018].

4.1 Introduction

For many models of statistical physics, the critical behaviour of the system is believed to be dependent solely on the large-scale geometry of the substrate, independently of the microscopic details of its geometry. The behaviour at criticality is encoded in a set of **critical exponents** which describe how properties of the model are dependent on the length scale at which the system is observed. These critical exponents are often summarised as a function of the dimension of the substrate under consideration, and, fascinatingly, apparently unrelated models are often found to share the same critical exponents. This phenomenon is known as **universality**, and systems with identical exponents are grouped together into **universality classes**. For background on the universality phenomenon and its renormalization group interpretations, see e.g. [98, 187, 264].

Underlying the phenomenon of universality is the fact that Euclidean lattices have a single well-defined dimension which determines all their large-scale geometric features via their common scaling limit $\mathbb{R}^d$. In contrast, it is possible in more general settings to have many potentially inequivalent notions of dimension, and even to have multiple substrates for which all these notions of dimension agree but which nevertheless have highly distinct large-scale geometries. This raises several interesting questions: can we characterise the set of geometric features of the substrate on which the critical exponents depend? It is possible that they depend only on the dimensions? To what extent do the answers to these questions depend on the model under consideration? In other words, how universal is universality?

In this paper, we study these questions in two classes of geometric setting: transitive (possibly non-Euclidean) lattices with polynomial volume growth and self-similar fractals. Our results in these two cases push in opposite directions. For transitive lattices we present clear numerical evidence that the critical exponents depend only on the dimension, suggesting that a very strong form of universality should hold in this setting. In stark contrast, we construct two self-similar fractals for which a large number of standard dimensions coincide but which do not appear to have the same critical exponents for Bernoulli bond percolation, showing that no such universality should be expected to hold in this case.

4.1.1 Transitive graphs of polynomial growth

We now introduce the class of transitive graphs that we will study. Recall that a graph is said to be **transitive** if any vertex can be mapped to any other vertex by a symmetry of the graph. A transitive graph has **polynomial volume growth** if there exists a constant C such that $|B(v,r)| \leq Cr^C$ for every $r \geq 1$, where $B(v,r)$ is the graph distance ball of radius r around the vertex v. While the hypercubic lattices $\mathbb{Z}^d$ are trivially seen to be transitive graphs of polynomial volume growth, there are also many examples of highly non-Euclidean transitive graphs of polynomial volume growth. Indeed, the possible large scale geometries of these graphs are classified by famous theorems of Gromov [169] and Trofimov [313] which imply that every transitive graph of polynomial volume growth is quasi-isometric to a Cayley graph of a torsion-free nilpotent group. A theorem of Bass [53] and Guivarc'h [171] then implies that every transitive graph of polynomial growth has a well-defined **integer** dimension d such that

$$C^{-1}r^d \leq |B(v,r)| \leq Cr^d$$

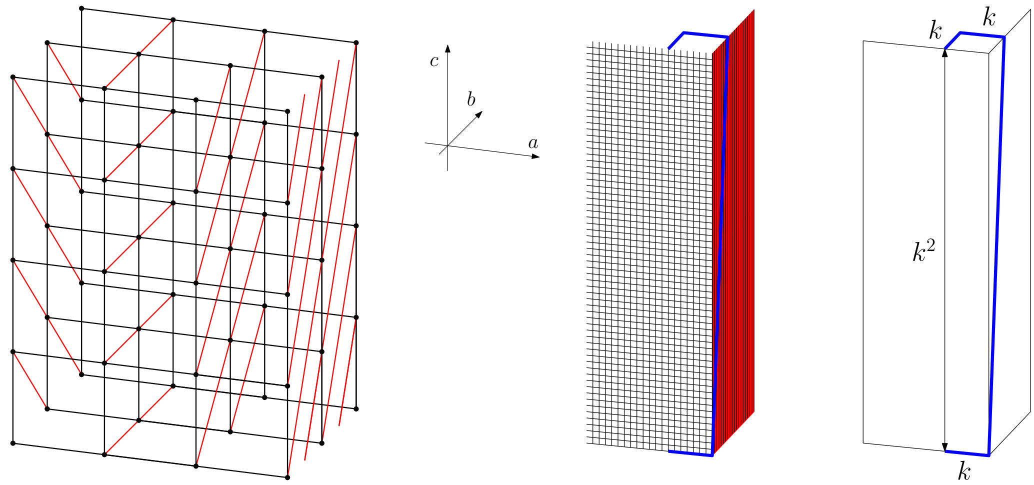


Fig. 4.1 The non-Euclidean geometry of the Heisenberg group. Left: A section of a Cayley graph of the Heisenberg group with generators a, b, and $c = [a,b]$. This graph may be obtained from the cubic lattice $\mathbb{Z}^3$ by applying a vertical shear of coefficient n to each of the hyperplanes $\{(a,b,c) : a = n\}$. (Note that the $a \leftrightarrow b$ asymmetry of this picture arises from our choice to take the right Cayley graph rather than the left Cayley graph.) Right: One may reach $(0,0,k^2)$ from $(0,0,0)$ in $4k$ steps by first going k steps in the a direction, then k steps in the direction $(0,1,k)$, then k steps in the negative a direction, then finally coming back k steps in the negative b direction. This leads to the Heisenberg group having volume-growth dimension 4 rather than 3. In fact, the graph metric on the Heisenberg group is comparable to the quasi-norm $\|(a,b,c)\| = |a| + |b| + |c|^{1/2}$. To illustrate just how alien the geometry of this space is, let us mention a theorem of Monti and Rickley [276] which states that any three non-collinear points in the continuum Heisenberg group have the entire space as their convex hull.

for some constant C and every $r \geq 1$. This same dimension d also arises as the spectral and isoperimetric dimensions of the graph by a theorem of Coulhon and Saloff-Coste [107].

In the low-dimensional cases $1 \leq d \leq 3$, it is a consequence of the Bass-Guivarc'h formula and the classification of low-dimensional nilpotent Lie algebras [114] that there is only one possible large-scale geometry, namely that of $\mathbb{Z}^d \approx \mathbb{R}^d$. For $d = 4$, there are exactly two possible large-scale geometries exemplified by the abelian group $\mathbb{Z}^4$ and the **Heisenberg group** (Figure 4.1), i.e. the 3×3 matrix group

$$\mathscr{H} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in \mathbb{Z} \right\}.$$

The fact that Heisenberg group has distinct large-scale geometry from $\mathbb{Z}^4$ is evidenced by the fact that its scaling limit is not $\mathbb{R}^4$ but is instead the **continuum Heisenberg group** equipped with its **Carnot–Carathéodory metric** – a self-similar **sub-Riemannian** manifold that is homeomorphic to $\mathbb{R}^3$ but has Hausdorff dimension 4 [240]. For $d=5$ there are again exactly two quasi-isometry classes, namely those of $\mathbb{Z}^5$ and $\mathscr{H}\times\mathbb{Z}$. In higher dimensions the number of possibilities is much larger, and indeed the classification of possible geometries is not completely understood [106, Section 19.7]. As with the Heisenberg and continuum Heisenberg groups above, each finitely generated torsion-free nilpotent group has an associated nilpotent Lie group, known as its **Mal'cev completion**, which contains the group as a lattice and which carries a self-similar sub-Riemannian metric arising as the scaling limit of its Cayley graphs by a theorem of Pansu [285]. Further background on these topics can be found in the surveys [106, 129, 170, 240].

In this paper we simulate critical Bernoulli bond percolation on $\mathscr{H}$ and $\mathscr{H}\times\mathbb{Z}$, and uniform lattice trees on $\mathscr{H}$, $\mathscr{H}\times\mathbb{Z}$, and two non-Euclidean seven-dimensional geometries known as $G_{4,3}$ and $G_{5,8}$. Here, a **uniform lattice tree** is simply a finite subtree of the lattice chosen uniformly at random among those subtrees that contain the origin and have some fixed number of vertices n; detailed definitions of both models and of the graphs we work with are given in Section 4.2 and Section 4.3. As summarised in Table 4.1, the numerical values of the critical exponents we obtain are in good agreement with previous results for Euclidean lattices, providing strong evidence in favour of the following conjecture:

Conjecture 33. *The critical exponents describing Bernoulli percolation and lattice trees on transitive graphs of polynomial growth are each determined by the volume-growth dimension of the graph.*

We also expect similar conjectures to hold for many other models; see Section 4.5 for further discussion. Note that the exponent estimates reported in Table 4.1 are only one facet of the evidence we provide in favour of Conjecture 33, with a more nuanced perspective on the data presented in Section 4.3.

Conjecture 33 is uncontroversial in high dimensional settings: The critical exponents describing percolation and lattice trees are strongly believed to take their mean-field values above the upper-critical dimensions of $d_c=6$ and $d_c=8$ respectively [177, 178] and the heuristic arguments in support of this do not rely on the Euclidean geometry of $\mathbb{Z}^d$ in any way. (Hara and Slade's rigorous derivation of mean-field behaviour for these models in high dimensions via the lace expansion [177, 178] does however rely on specific features of Euclidean geometry, and it is an open problem to extend their analysis to the non-Euclidean case.) For models with $d_c=4$ such as the Ising model, φ^4 field theory, the self-avoiding walk, and the uniform spanning tree, the dearth of possible low-dimensional geometries causes the analogous conjecture to reduce the standard universality principle for Euclidean lattices. Indeed, we chose to study lattice trees in part because their high upper-critical dimension

(a) CRITICAL EXPONENT ESTIMATES FOR PERCOLATION.

		$\mathscr{H}$	$\mathbb{Z}^4$		$\mathscr{H} \times \mathbb{Z}$	$\mathbb{Z}^5$		$d \geq 6$
τ	cluster-size distribution	2.315	2.313 2.314 2.311 2.314 2.312	[287] [322] [272] [324] [162]	2.420	2.412 2.422 2.418 2.417	[287] [272] [324] [162]	2.5
σ	size of large clusters	0.476	0.480 0.474	[324] [162]	0.499	0.494 0.493	[324] [162]	0.5

(b) CRITICAL EXPONENT ESTIMATES FOR LATTICE TREES.

		$\mathscr{H}$	$\mathbb{Z}^4$		$\mathscr{H} \times \mathbb{Z}$	$\mathbb{Z}^5$		$G_{4,3}$	$G_{5,8}$	$\mathbb{Z}^7$		$d \geq 8$
ρ	intrinsic radius	0.595	0.609 0.607	(new) [323]	0.570	0.576 0.578	(new) [323]	0.526	0.524	0.530	[323]	0.5
ν	extrinsic radius	0.420	0.417 0.415 0.416	(new) [323] [186]	0.358	0.358 0.359 0.359	(new) [323] [186]	0.286	0.283	0.291 0.283	[323] [186]	0.25

Table 4.1 A summary of our results for transitive graphs of polynomial volume growth. All estimates are presented to three decimal places for ease of comparison. For percolation, the exponents τ and σ heuristically describe the distribution of the size of the cluster of the origin at and near criticality via the ansatz $\mathbb{P}_p(|K| = s) \approx s^{1-\tau} g(|p - p_c|^{1/\sigma} \cdot s)$ for some rapidly decaying function g. These exponents are equivalent to those known as δ and Δ by the relations $\tau = 2 + 1/\delta$ and $\sigma = 1/\Delta$. For lattice trees, the exponents ρ and ν are defined so that a typical n-vertex lattice tree will have intrinsic and extrinsic radii of order n^ρ and n^ν respectively. Note that Gracey's estimates [162] are obtained using (non-rigorous) renormalization group methods rather than numerically, and that the percolation estimates credited to Zhang et al. [324] were computed from their estimates of the exponents ν and d_f using the scaling relations $\tau - 1 = d/d_f$ and $\sigma = 1/\nu d_f$. In each case, our results are consistent with those obtained for the Euclidean lattices of corresponding dimension, with the small differences in numerical values reasonably attributed to finite-size effects and noise.

allowed for the analysis of a larger number of interesting examples. (One could however make a similar universality conjecture concerning the **logarithmic corrections to scaling** at the upper-critical dimension, so that our conjectures would remain interesting for, say, the 4d Ising model. We are inclined to believe such a conjecture but have not tested it numerically.)

PERCOLATION EXPONENTS ON SELF-SIMILAR FRACTALS.

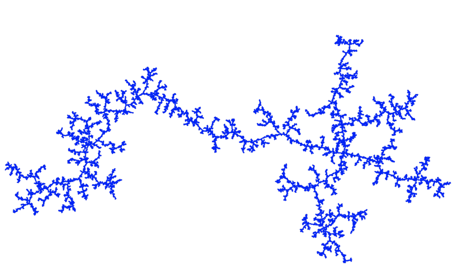
	Dimension				Exponent	
	Hausdorff	Spectral	Topological	Top. Hausdorff	τ	σ
H_1	3	7/3	2	3	2.195(5)	?
H_2	3	7/3	2	3	2.151(1)	0.385(5)
H_3	8	5	4	8	2.66(1)	0.41(1)
H_4	8	5	4	8	?	0.41(1)

Table 4.2 Summary of results of percolation on self-similar fractals. Note the unambiguous separation in numerical values of the Fisher exponent τ between the two equidimensional fractals H_1 and H_2 and the coincidence in the numerical values of σ for the two fractals H_3 and H_4. We found the finite-size effects to be **much** larger on these graphs than on the transitive graphs we considered and – despite our considering clusters of up to 10^9 vertices – for some graphs the relevant log-log plots were too far from linear to reliably extract any exponent value at all. Again, the more detailed data presented in Section 4.4 give a much more complete picture of the situation than the raw exponent estimates presented here. In particular, we find the data presented in Figure 4.11 to demonstrate very convincingly that H_1 and H_2 have distinct values of τ.

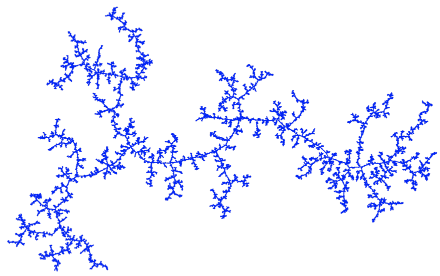
We note that for percolation our simulations on $\mathscr{H}$ and $\mathscr{H} \times \mathbb{Z}$ exhaust all available non-Euclidean geometries below the upper-critical dimension $d_c = 6$, so that our results lend particularly strong support to the conjecture in this case.

If the conjecture is true, it may be difficult to explain using existing methodology. Indeed, the equality of critical exponents on different Euclidean lattices of the same dimension is often explained as a consequence of the stronger statement that the two models have the same **scaling limit**. In our setting, however, it is certainly **not** the case that e.g. percolation on $\mathbb{Z}^4$ and $\mathscr{H}$ have a common scaling limit, since one scaling limit would be defined on $\mathbb{R}^4$ while the other would be defined on the continuum Heisenberg group. Again, we stress that the continuum Heisenberg group is self-similar and non-Riemannian, so that it is **not** approximated by Euclidean space on any scale. In light of these difficulties, we are optimistic that further investigation into Conjecture 33 may also significantly deepen our understanding of the original Euclidean models.

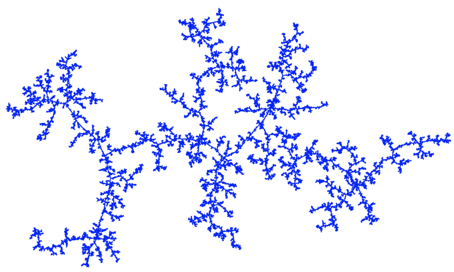
One very interesting possibility is that the **intrinsic geometries** of the models share a common scaling limit across different geometries of the same dimension, even though the extrinsic scaling limits must be different. For example, it may be that large uniform lattice trees on $\mathbb{Z}^4$ and the Heisenberg group have a common scaling limit **when considered as abstract metric trees**. The simulations presented in Figure 4.2 show that such a conjecture is at least plausible and is worthy of further investigation in future work. Still, such a conjecture



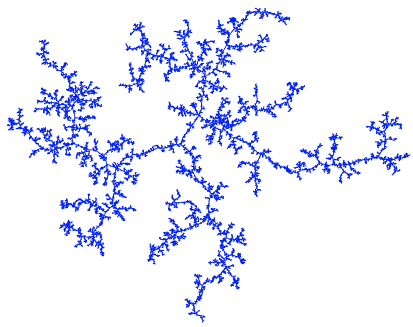
(a) THE FOUR-DIMENSIONAL HYPERCUBIC LATTICE $\mathbb{Z}^4$.



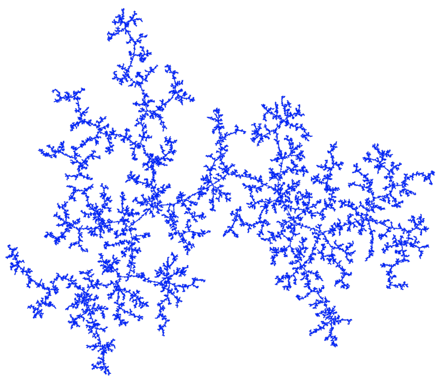
(b) THE FOUR-DIMENSIONAL HEISENBERG GROUP $\mathscr{H}$.



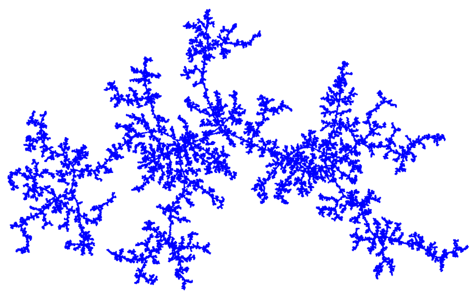
(c) THE FIVE-DIMENSIONAL HYPERCUBIC LATTICE $\mathbb{Z}^5$.



(d) THE FIVE-DIMENSIONAL PRODUCT SPACE $\mathscr{H} \times \mathbb{Z}$.



(e) THE SEVEN-DIMENSIONAL GEOMETRY $G_{4,3}$.



(f) THE SEVEN-DIMENSIONAL GEOMETRY $G_{5,8}$.

Fig. 4.2 A visual cross-comparison of large, approximately uniform lattice trees in six different geometries, each with $60,000$ vertices. These trees were sampled via the MCMC method described in Section 4.2.2 and drawn in the plane using Mathematica's SpringElectricalEmbedding algorithm with parameter RepulsiveForcePower $= -3$. Note that this is **not** an isometric embedding, and tends to distort distances rather severely. The difference in exponents between the low-dimensional and high-dimensional cases manifests itself in the seven-dimensional lattice trees looking much "bushier" than their four-dimensional lattice tree counterparts. The reader may like to compare these figures to the simulations of Aldous's continuum random tree [4] that are available on e.g. Igor Kortchemski's webpage https://igor-kortchemski.perso.math.cnrs.fr/images.html, noting that the continuum random tree arises as the scaling limit of large uniform lattice trees in dimensions eight and above [118].

would be difficult to confirm in light of the distinct extrinsic scaling limits and would not obviously explain e.g. the coincidence of exponents describing the **extrinsic** geometry of lattice trees.

We remark that there is an extensive literature investigating critical behaviour on **hyperbolic** lattices including e.g. [35, 62, 158, 191, 224, 265, 271]. These lattices are very different from the non-Euclidean lattices we consider in this paper. Indeed, hyperbolic lattices are of infinite-dimensional volume growth and are therefore expected to exhibit mean-field behaviour for both models; this has been proven rigorously for percolation on arbitrary hyperbolic lattices in [191] and for lattice trees on certain hyperbolic lattices by Madras and Wu [265]. We believe our paper is the first to systematically investigate critical exponents on transitive non-Euclidean lattices **below** the upper-critical dimension.

4.1.2 Self-similar fractals

The self-similar Carnot groups arising as scaling limits of transitive graphs of polynomial growth can be thought of as very special examples of **fractal** spaces. As such, it is natural to wonder to what extent the phenomena discussed above extend to more general fractals. The situation here is more complicated. We will restrict our attention in the fractal case to Bernoulli percolation, where previous works investigating the effect of fractal geometry on percolation critical probabilities and critical exponents include [37, 38, 61, 86, 152–157, 181, 182, 250, 275, 315]. In these works, percolation on families of fractals with varying fractal and spectral dimensions is investigated, with the focus often on Sierpinski-type fractals. (Of course one does not sample percolation directly on the continuum fractals but rather on appropriately chosen 'prefractal' graphical approximants; we discuss this further in Section 4.4.) These results demonstrate that, in contrast to our Conjecture 33, percolation critical exponents on fractals cannot depend solely on the Hausdorff dimension (which is the most popular continuum analogue of the volume-growth dimension).

Once it is known that universality does not hold across fractals with identical Hausdorff dimensions, a next natural hypothesis is that critical exponents are instead a function of some **set** of properties or dimensions which better capture the geometry of the fractal. A specific proposal to this effect was made by Balankin et al. [37], who suggested that the critical exponents should be determined by a set of three fractal dimensions, namely the Hausdorff, spectral, and topological Hausdorff dimensions. The more general view that the spectral dimension is important to the determination of critical behaviour has been advocated by many authors; see the introduction of [273] for an overview.

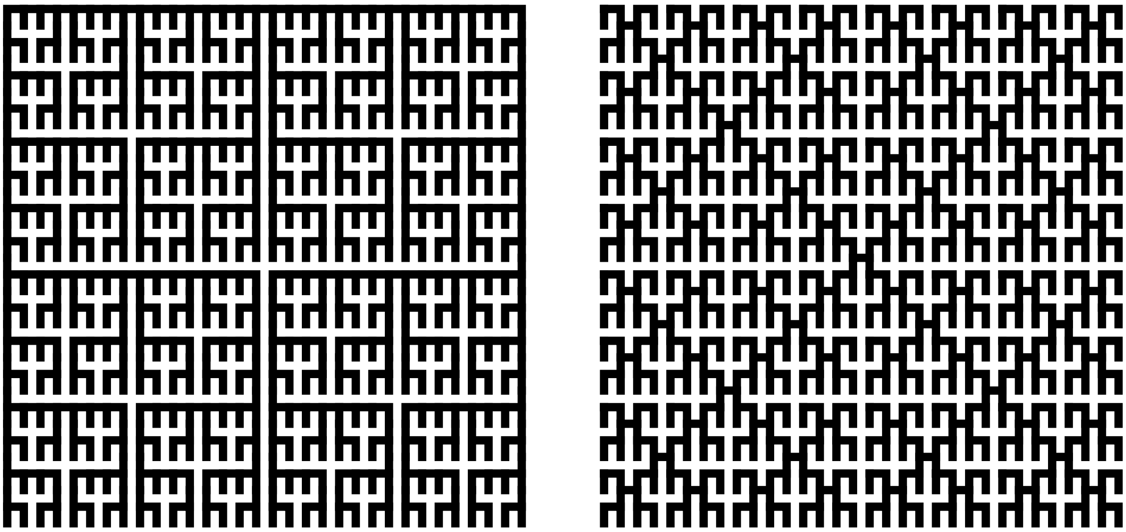


Fig. 4.3 Discrete approximations of the self-similar fractal trees T_O (left) and T_I (right). Each tree is constructed as a scaling limit of a recursively defined, self-similar spanning tree of the square lattice. In the 'outer' tree T_O, the tree associated to the $(n+1)$th dyadic scale is formed by connecting four copies of the scale n tree 'around the outside' by adding edges on the centre left, centre right, and top of the square. In the 'inner' tree T_I, the tree associated to the $(n+1)$th dyadic scale is formed by connecting four copies of the scale n tree 'in the middle' by adding three edges to the centre of the square.

In this paper we make a novel contribution to this problem by cross-comparing percolation critical exponents between pairs of fractals which are distinct but for which many important notions of dimension coincide. The specific examples we consider are constructed as products of various recursively-defined self-similar fractal trees. We use these examples due to the flexibility of their construction and the ease of computation of their associated fractal dimensions. The dimensions of the four different fractal products we consider and our numerical estimates of their percolation critical exponents are summarised in Table 4.2.

We begin by constructing two fractal trees T_I and T_O, each with Hausdorff dimension 2, such that the two fractal products $H_1 = T_O \times [0,1]$ and $H_2 = T_I \times [0,1]$ share the same Hausdorff, spectral, topological and topological Hausdorff dimensions. Both trees T_I and T_O are defined as scaling limits of self-similar spanning trees of the square lattice $\mathbb{Z}^2$ as depicted in Figure 4.3. We present strong numerical evidence that the two fractals H_1 and H_2 have distinct values of the percolation Fisher exponent τ which characterises the cluster size-distribution at criticality. This provides strong evidence against the aforementioned conjecture of Balankin et al. [37]. Moreover, since the two fractals we consider are very similar in a large number of ways beyond these dimensions, our results suggest that any universality principle applying to fractals must be much weaker than in the transitive case.

On the other hand, a more positive picture emerges when one considers the critical exponent σ which characterises the size of the percolation scaling window. Indeed, for H_1 and H_2 our results were inconclusive but consistent with the hypothesis that the two values of σ coincide. In order to investigate this potential phenomenon further, we constructed and analysed two further fractal tree products which we call H_3 and H_4. As with H_1 and H_2, the two fractals H_3 and H_4 share many notions of dimension despite having distinct geometries in other regards, but are 'higher-dimensional' overall than H_1 and H_2. We present strong numerical evidence that H_3 and H_4 have a shared value of σ. This may be related to the phenomenon of **weak universality** as discussed in [274], and weakly suggests that the exponent σ may indeed be a function of some small set of parameters associated to the fractal.

Organisation: The rest of the paper is structured as follows: In Section 4.2, we recall the definitions of the two models we will study and the exponents we wish to compute, and describe the methodologies used in our simulations. Further details of an improvement to the invasion percolation methodology are given in Appendix A. In Section 4.3 we give background on the four geometries $\mathscr{H}$, $\mathscr{H} \times \mathbb{Z}$, $G_{4,3}$, and $G_{5,8}$ and present our numerical results regarding percolation and lattice trees in these geometries. In Section 4.4, we describe the four fractals H_1, H_2, H_3, and H_4, and present the relevant numerics. Finally, we summarise our findings and discuss possible directions for future work in Section 4.5.

4.2 Models and algorithms

In this section, we give relevant background on percolation and lattice trees, and review the methodology we use to compute the critical exponents describing these models.

4.2.1 Bernoulli bond percolation

Fix $p \in [0,1]$. Given a graph $G = (V,E)$, we attach independent and identically distributed (i.i.d.) Bernoulli random variables $(\omega_e)_{e\in E}$ of parameter p to the edges of the graph and say that an edge e is **open** if $\omega_e = 1$ and **closed** if $\omega_e = 0$. We denote the associated product probability measure by $\mathbb{P}_p$. Given any vertex $v \in V$, we define the **cluster** K_v of v to be the set of vertices that are accessible from o by paths consisting only of open edges. Given an infinite graph G, we define the **critical probability** p_c to be the infimal value of p for which infinite clusters exist with positive probability. Note that the value of p_c depends strongly on the microscopic details of the graph and is not universal.

We now introduce the exponents we consider and some relevant (non-rigorous) scaling theory, referring the reader to e.g. [163, Chapter 9] for further background. Let o be a fixed vertex of G, which we regard as the origin. Assuming they are well-defined, the exponents τ and ρ describe the distribution of the volume and (extrinsic) radius of the cluster of the origin at criticality by

$$\mathbb{P}_{p_c}(|K_o| \geq s) \approx s^{2-\tau} \qquad \text{as } s \uparrow \infty \text{ and}$$
$$\mathbb{P}_{p_c}(\mathrm{rad}(K_o) \geq r) \approx r^{-1/\rho} \qquad \text{as } r \uparrow \infty,$$

where $\mathrm{rad}(K_o)$ is the maximum distance in G between o and another point of K_o. (We keep the meaning of the symbol $\approx$ intentionally vague; it should not be read as corresponding to any specific or consistent notion of asymptotic equivalence.) Below the upper-critical dimension, these exponents are expected to determine each other via the hyperscaling relation $\tau = (2d\rho - 1)/(d\rho - 1)$ [163, Chapter 9]. It is a standard assumption of scaling theory that there exists a further exponent σ such that

$$\mathbb{P}_p(|K_o| = s) \approx s^{1-\tau} g_{\pm}\big(|p - p_c|^{1/\sigma} \cdot s\big)$$

for some rapidly decaying functions g_- and g_+ describing the cases $p \leq p_c$ and $p \geq p_c$ respectively. In particular, this ansatz predicts that the probability $\mathbb{P}_p(|K_o| = n)$ is of the same order as its critical value when $n \ll |p - p_c|^{-1/\sigma}$ and is very small when $n \gg |p - p_c|^{-1/\sigma}$, and we think of $|p - p_c|^{-1/\sigma}$ as describing the "typical size of a large finite cluster".

The calculations we perform in this paper will utilise a slightly different approach to scaling theory, adapted from the presentation of [253], which we now overview. When G is transitive, we fix an origin vertex o as above and write $P_{\geq s} = P_{\geq s, p_c}$ for the cluster size distribution at criticality, where $P_{\geq s,p} = \mathbb{P}_p(|K_o| \geq s)$. For non-transitive fractals, we take the origin o to be a vertex selected uniformly at random (in a sense which will be made precise later), and then define $P_{\geq s,p} = \mathbf{E}[\mathbb{P}_p(|K_o| \geq s)]$, where $\mathbf{E}[\cdot]$ denotes the expectation with respect to the random origin o. We will assume as a basis for calculations that the critical cluster size distribution is described by the ansatz

$$P_{\geq s, p_c} = A_0 s^{2-\tau}(1 + A_1 s^{-\Omega} + \cdots) \tag{4.1}$$

for some $\tau, \Omega > 0$ and $A_0, A_1 \in \mathbb{R}$. The exponent τ is known as the Fisher exponent, and Ω is the leading correction-to-scaling component whose impact becomes negligible for large s. Both these exponents are expected to be universal in the sense that they only depend on the large scale geometry of the graph, whereas the constants A_i are non-universal and will also depend on the microscopic geometry of the graph. In order to compute σ, we will assume similarly that

$$P_{\geq s,p} = C_0 s^{2-\tau}(1 + C_1(p - p_c)s^{\sigma} + \cdots) \tag{4.2}$$

for $p \neq p_c$ and values of s that are not too large or small.

We use an algebraic manipulation, as in [253], to derive a more convenient scaling relation for σ which only relies on properties of the percolation clusters at criticality. Taking derivatives of Eq. (4.2) with respect to p, we get

$$\frac{dP_{\geq s,p}}{dp} = C_0 C_1 s^{2-\tau+\sigma} + \cdots \tag{4.3}$$

If we let $g_{n,t}$ be the number of possible cluster configurations containing the origin and exactly n open edges, and with t closed edges adjacent to the cluster, then

$$P_{\geq s,p} = \sum_{n \geq s} \sum_{t} g_{n,t} p^n (1-p)^t.$$

Taking derivatives with respect to p gives

$$\frac{dP_{\geq s,p}}{dp} = \sum_{n \geq s} \sum_{t} g_{n,t} p^n (1-p)^t \big(\frac{n}{p} - \frac{t}{1-p}\big) = \frac{\mathbb{E}[n\mathbf{1}_{n \geq s}]}{p} - \frac{\mathbb{E}[t\mathbf{1}_{n \geq s}]}{1-p}$$

so that

$$\frac{\mathbb{E}_{p_c}[n\mathbf{1}_{n \geq s}]}{p_c} - \frac{\mathbb{E}_{p_c}[t\mathbf{1}_{n \geq s}]}{1-p_c} = C_0 C_1 s^{2-\tau+\sigma} + \cdots,$$

and hence

$$\frac{\mathbb{E}_{p_c}[n \mid n \geq s]}{p_c} - \frac{\mathbb{E}_{p_c}[t \mid n \geq s]}{1-p_c} = C_1 s^{\sigma} + \cdots. \tag{4.4}$$

As in [253], this will be used as an assumed formula to compute σ using only information at criticality. Let us note however that the expectation on the left hand side has poor numerical properties since the associated random variable $n/p - t/(1-p)$ is heavy tailed at criticality. Indeed, Eq. (4.4) should really be interpreted as a statement about the $p \uparrow p_c$ limit since the left hand side is not well-defined at criticality.

Lastly, we define $Q_{\geq r} = \mathbb{P}_{p_c}(\mathrm{rad}(K_o) \geq r)$, and assume the form

$$Q_{\geq r} = F_0 r^{-1/\rho}(1 + F_1 r^{-\zeta} + \ldots),$$

for some $\rho, \zeta > 0$ and $F_0, F_1 \in \mathbb{R}$, where F_0 and F_1 are not expected to be universal.

Methodology. We now describe the methods we used to compute critical exponents for percolation. Our first step was to estimate the value of the critical probability p_c. To do this, we began by employing the **invasion percolation** methodology of [271, 272], using the bulk-to-boundary ratio developed by Leath in [242] as an estimator of the critical probability, and then using the extrapolation hypothesis developed in [272] to further refine the resulting estimate. In fact, we implemented a simple improvement to this methodology that resulted in substantial run-time reductions and which we describe in detail in Appendix A.

Invasion percolation is a stochastic model for the transport of fluid through porous media [100, 247, 320]. It operates by assigning i.i.d. uniform random variables U_e taking values in $[0,1]$ to the edges of some graph G with root vertex o. We then define the sequences $(e_n)_{n\geq 1}$, $(V_n)_{n\geq 0}$, $(E_n)_{n\geq 0}$, and $(F_n)_{n\geq 0}$ recursively as follows:

1. Start with $V_0 = \{o\}$, $E_0 = \varnothing$, and $F_0 = \{\{x,o\} : x \sim o\}$.
2. At each step $n \geq 1$, let e_n be the element of F_{n-1} minimizing U_e, let $E_n = E_{n-1} \cup \{e\}$, let V_n be the set of vertices adjacent to at least one edge of E_n, and let F_n be the set of edges that have at least one endpoint in V_n but do not belong to E_n.

We call V_n the **invasion cluster** up to time n, and F_n the **frontier** at time n. The **bulk-to-boundary ratios** are given by the random sequence

$$a_n = \frac{|E_n|}{|E_n| + |F_n|} = \frac{n}{n + |F_n|}.$$

It is proven in [259, Chapter 11] that $\limsup_{n\to\infty} U_{e_n} = p_c$ almost surely, and it is believed that

$$a_n \to p_c \qquad \text{as } n \to \infty;$$

this has been proven [101, 167] for $\mathbb{Z}^d$, but it is expected to hold for a much wider range of graphs. Assuming that this limiting relation holds, one may estimate the critical probability p_c by running invasion percolation for a long time and computing the resulting bulk-to-boundary ratio.

In [271] this method was improved via the following extrapolation argument. For invasion percolation on the binary tree, the bulk to boundary ratio can be shown to satisfy the asymptotics

$$\mathbb{E}a_n \approx \frac{p_c}{1 + An^{-1}}$$

for some constant A. The authors of [271] conjecture and verify numerically that for Euclidean lattices one has the analogous formula

$$\mathbb{E}a_n \approx \frac{p_c}{1 + An^{-\delta}} \tag{4.5}$$

with a high degree of accuracy for some positive constants A and δ. We will assume that such a formula also holds in our settings and, following [271], use the ansatz

$$\mathbb{E}a_n = \frac{p_c}{1 + An^{-\delta}(1 + Bn^{-\delta'} + Cn^{-\delta''} + \cdots)}$$

as a basis from which to calculate p_c; this allows us to gather data for a relatively small number of time-steps and then use curve fitting to give an estimate of p_c.

The use of a heap or sorted list for storing/extracting values on the frontier lets us compute S_n in time $O(n \log n)$ and memory $O(n)$. The simple improvement we outline in Appendix A reduces the size of the sorted list used for the frontier by a power and thus significantly reduces the running time.

The major advantages of using invasion percolation for computing p_c are as follows:

1. It does not require us to store large blocks of the relatively high dimensional lattices on which we simulate percolation - instead we need only store a number of edges or vertices which is linear in the number of steps of the algorithm thus far.
2. It does not require us to assume **a priori** values of any critical exponents, unlike the methods of [322].
3. It does not require detailed understanding of the geometry of the graph under consideration, unlike the wrapping method used in [283] or the multi-scale analysis used in [37].

Invasion percolation also allows us to narrow in on a relatively precise value of p_c with far smaller memory and time requirements than by starting from scratch utilizing a logarithmic search with the Leath algorithm as in [253].

Having estimated p_c via invasion percolation as described above, we then used the **Leath algorithm** [243] to generate a population of samples for the cluster at the origin. For each sample, we recorded both the cardinality of the cluster and the number of boundary edges. For the lattices, we used least-mean-squares to fit the parameters in the following equations, approximately valid at $p \approx p_c$, to give estimates of τ and σ:

1. $\log_2 P_{\geq s} = (2-\tau)\log_2(s) + Bs^{-\Omega} + C$
2. $\log_2\left(\frac{\mathbb{E}[n|n\geq s]}{p} - \frac{\mathbb{E}[t|n\geq s]}{1-p}\right) = \sigma \log_2(s) + D.$

For the fractals, we plotted

(i) $\log_2 P_{\geq s}$ against $\log_2(s)$ (ii) $\log_2\left(\frac{\mathbb{E}[n \mid n \geq s]}{p} - \frac{\mathbb{E}[t \mid n \geq s]}{1-p}\right)$ against $\log_2(s)$, at $p \approx p_c$ and calculated the gradient of the approximately linear sections at large s to give estimates of τ and σ.

In the case of the non-Euclidean lattices, the values of τ and σ we obtained were very close to the values in the literature for the corresponding Euclidean lattices. Having obtained these estimates, we then utilised the methodology of [253, 322] to improve our value for the critical probability and give further credence to our conclusion. To this end, we sampled the cluster at the origin using the Leath algorithm at a range of values of p near our initial estimate of p_c and plotted the following graphs:

$$\text{(i) } s^{\tau_E - 2} P_{\geq s} \text{ against } s^{-\Omega_E} \qquad \text{(ii) } s^{\tau_E - 2} P_{\geq s} \text{ against } s^{\sigma_E},$$

where $\tau_E, \sigma_E, \Omega_E$ are the corresponding estimates of the Euclidean exponents calculated in previous literature. For the first graph, looking at relation Eq. (4.1), we expect that the curve does not deviate from its linear trajectory for small $s^{-\Omega}$ when p is close to p_c, and for the second graph, looking at relation Eq. (4.2), we expect a plateau for large s^{σ} when p is close to p_c. We observed that this was indeed the case, lending further credibility to the accuracy of our estimates.

We calculated estimates of the extrinsic exponent for each transitive lattice by recording the maximum extrinsic distance of any vertex visited in runs of the Leath algorithm. We plotted the following curve and calculated the gradient of its approximately linear final segment to give $-1/\rho$:

- $\log_2 Q_{\geq r}$ against $\log_2(r)$.

For the percolation on products of fractal trees, where we did not have reference values of τ and σ to compare with, we instead refined our values of the critical exponents and p_c by plotting $\log P_{\geq s,p}$ against $\log s$ over very large ranges of s for a selection of probabilities p, and finding the value of p which gave the smallest deviation from linearity for medium and large s as in [180, 253].

4.2.2 Lattice Trees

Given a transitive connected graph G and a fixed vertex o of G, a **lattice tree** is a finite connected subgraph of G that contains o and is a tree, i.e. does not contain any cycles. We let T_n be the set of n-vertex lattice trees; the uniform lattice tree of size n is then just the random variable given by selecting one of these tree uniformly at random. We will study how the following quantities depend on the tree size n:

- **The mean branch size,** $B(n)$: If we take an edge from the lattice tree and delete it, the branch size is the cardinality of the smaller of the two resultant subtrees. The mean branch size is the expectation of this quantity over the lattice tree and over an edge picked uniformly at random from the lattice tree.
- **The intrinsic longest path,** $I(n)$: This is the expected length of the longest path in the lattice tree, where the length of the path is given by the **intrinsic metric**, i.e. the graph metric of the tree.
- **The extrinsic displacement of the longest path,** $E(n)$: This is the expected extrinsic distance between the two endpoints of some maximal-length intrinsic path in the tree. (The method used to pick a particular such path when it is non-unique is described below; the details of this should not be important.)

Here we are using the extrinsic displacement of the longest path as an easier-to-compute substitute for the true extrinsic diameter of the tree, which we expect to be of the same order. Assuming they are well-defined, the exponents ρ and ν describe the asymptotics of $B(n)$, $I(n)$, and $E(n)$ via

$$B(n) \approx I(n) \approx n^{\rho} \qquad \text{and} \qquad E(n) \approx n^{\nu}.$$

We remark as a point of general interest that the exact equality $\nu = 0.5$ is believed to hold for **three-dimensional** lattice trees. This equality has been proven rigorously for **branched polymers** [89, 217], which are believed to be in the same universality class as lattice trees.

Methodology. In order to sample approximately uniform lattice trees, we employed a combination of two Markov chain Monte Carlo (MCMC) algorithms. The MCMC algorithms involve evolving some arbitrary initial tree t_0 by applying a sequence of randomly chosen operations to form a process $(t_i)_{i\geq 1}$ on the space of lattice trees. The possible operations and their probabilities are chosen such that, when restricted to the set of n-vertex lattice trees T_n, the resultant process is Markovian, irreducible and aperiodic, and has the uniform distribution as its invariant distribution. Standard Markov chain theory then implies that the process will converge to the uniform measure on T_n as $i \to \infty$. After an initial mixing period, the process is sampled at regular intervals, and measurements of interest are calculated and recorded. In the absence of bounds on the mixing/relaxation time, we fixed the number of steps of the algorithm between samples in such a way that the autocorrelation of the measured quantities across the samples was found to be negligible. The two algorithms we combined were the **cut-and-paste** algorithm (CP) developed in [203] and the **cycle-breaking** algorithm (BC) described in [144]. The resultant algorithm, which we term the **Cut-and-Paste Break-Cycle algorithm** (CPBC), involves alternating between (CP)- and (BC)-type operations according to a probabilistic criterion. For details of the algorithm, see Appendix B.

Once we have sampled the lattice tree, we must calculate the exponents. We computed the critical exponents ν and ρ by the same method described in [203, 323] where they were calculated for Euclidean lattices of dimensions 2 through 7. We used breadth-first search and dynamic programming techniques to calculate the mean branch size. To find an intrinsic longest path, we use the method introduced by Dijkstra around 1960: We choose a vertex v in the tree (at random, the choice being immaterial), and then find a vertex u in the tree with maximum intrinsic distance from v using breadth-first search. We then find a vertex u' in the tree with maximum intrinsic distance from u, and record the intrinsic distance between u and u'. The fact that this produces a pair of vertices at maximal intrinsic distance from each other is proven formally in [90]. Once this is done, we compute the extrinsic distance between u and u' either exactly or using an approximating quasi-norm as discussed in the next section, with the details being context-dependent. In each case, we averaged the outputs of these computations over a large number of runs to estimate $I(n)$, $E(n)$, and $B(n)$, plotted log-log plots of these quantities against n, and calculated estimates of ν and ρ by measuring the gradients of the final sections of the resulting curves.

4.3 Transitive lattices

In this section we define the various Cayley graphs we consider and report the outcomes of our simulations on these Cayley graphs. Given a finitely generated group Γ and a finite set S which generates Γ, the (right) **Cayley graph** $\mathrm{Cay}(\Gamma,S)$ is defined to be the undirected graph with vertex set Γ and edge set $\{\{\gamma,\gamma s\} : \gamma \in \Gamma, s \in S \cup S^{-1}\}$. Cayley graphs are always transitive since each element γ of Γ defines an automorphism of $\mathrm{Cay}(\Gamma,S)$ by left multiplication. The graph metric on $\mathrm{Cay}(\Gamma,S)$ is also known as the **word metric** and can be expressed as

$$d_S(\gamma_1,\gamma_2) = \min\left\{n \geq 0 : \exists s_1,\ldots,s_n \in S \cup S^{-1} \text{ such that } \gamma_2 = \gamma_1 s_1 \cdots s_n\right\},$$

and observe that this coincides with the graph metric. For each of the groups we consider, the word metric is comparable to a **quasi-norm** that is much easier to compute. We will use these quasi-norms in place of the word metric when computing distances on $G_{4,3}$ and $G_{5,8}$.

Recall that two metric spaces (X,d_X) and (Y,d_Y) are said to be **quasi-isometric** if there exist positive constants α and β and a function $\phi : X \to Y$ such that $\alpha^{-1} d_X(x,y) - \beta \leq d_Y(\phi(x),\phi(y)) \leq \alpha d_X(x,y) + \beta$ for every $x,y \in X$ and for every $y \in Y$ there exists $x \in X$ with $d_Y(y,\phi(x)) \leq \beta$. It is easily seen that different Cayley graphs of the same finitely generated group are quasi-isometric to each other and that e.g. $\mathbb{Z}^d$ is quasi-isometric to $\mathbb{R}^d$ for each $d \geq 1$.

When Γ is a group, the **lower central series** of Γ is defined recursively by $\Gamma_1 = \Gamma$ and $\Gamma_{i+1} = [\Gamma_i, \Gamma] = \langle \{[a,b] : a \in \Gamma_i, b \in \Gamma\} \rangle$. The group Γ is said to be **nilpotent** if there exists $s \geq 1$, known as the **step** of Γ, so that Γ_s is abelian and hence that $\Gamma_i = \{\mathrm{id}\}$ for every $i > s$. The Bass-Guivarc'h formula [53, 171] states that if Γ is a torsion-free nilpotent group then Γ has volume growth dimension $\sum_{i=1}^{s} i r_i$ where r_i is the rank of the abelian group Γ_i / Γ_{i+1}. The quantity $\sum_{i=1}^{s} i r_i$ is also known as the **homogeneous dimension** of the group. It is a consequence of Pansu's theorem [285] that both the step s and the sequence $(r_1, \ldots, r_s)$ are quasi-isometry invariants of nilpotent groups.

We chose four non-Euclidean groups to study, namely $\mathscr{H}$, $\mathscr{H} \times \mathbb{Z}$, $G_{4,3}$ and $G_{5,8}$. In some cases we also carried out simulations on $\mathbb{Z}^4$ and $\mathbb{Z}^5$ so that we could directly compare our results to the Euclidean case. The upper critical dimension of percolation is 6, so we limited our study to the most interesting dimensions of four and five where mean-field behaviour does not hold but there is more than one quasi-isometry class of geometries to consider. The upper critical dimension for lattice trees is 8, meaning that more interesting possibilities are available. We chose to study the two seven-dimensional groups $G_{4,3}$ and $G_{5,8}$ since they were highly distinct from the other examples we considered, being neither abelian, generalised Heisenberg, nor products thereof. These groups are defined as lattices in the nilpotent Lie groups corresponding to the nilpotent Lie algebras notated in [114] and [251, Table 1] as $\mathscr{L}_{4,3}$ and $\mathscr{L}_{5,8}$. The multiplication rules and generating sets of these groups were computed using Maple. A complete taxonomy of possible low-dimensional geometries can be found in [251, Tables 1-4 and Figure 5].

We briefly introduce each of the groups we consider, with most of the relevant information succinctly summarised in Appendix C.

The Heisenberg Group $\mathscr{H}$. The discrete Heisenberg group can be defined as the set of integer-valued upper-triangular 3×3 matrices under matrix multiplication. We identify each matrix $M \in \mathscr{H}$ with an element of $\mathbb{Z}^3$ via the bijection $\phi : \mathbb{Z}^3 \to \mathscr{H}$ given by

$$\phi\big((a,b,c)\big) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

and use these coordinates to represent elements of the group. These are known as the **Mal'cev** coordinates. Multiplication of two elements is therefore given by:

$$(a_1,b_1,c_1)\cdot(a_2,b_2,c_2) = (a_1+a_2, b_1+b_2, c_1+c_2+a_1b_2).$$

The Heisenberg group is generated by the elements $a = (1,0,0)$ and $b = (0,1,0)$ as witnessed by the identity

$$(x,y,z) = b^y[a,b]^z a^x,$$

where $[a,b]$ is the commutator $aba^{-1}b^{-1}$. We will work with the right-Cayley graph $\Gamma_{\mathscr{H}} = \mathrm{Cay}(\mathscr{H},\{a,b\})$. (Note that this is not the Cayley graph depicted in Figure 4.1, which has generating set $\{a,b,c\}$.) The graph metric on this Cayley graph is equivalent [140, 3.1.6] to the quasi-norm

$$\|(a,b,c)\| = |a|+|b|+|c|^{1/2}.$$

We will also make use of a formula for computing graph distances in this Cayley graph that is described in [84]; this formula is too long to reproduce here but is easily implemented on a computer. The Heisenberg group has step 2 and $(r_1,r_2) = (2,1)$.

All of the above mentioned facts have obvious consequences for the product space $\mathscr{H}\times\mathbb{Z}$, for which we will consider the Cayley graph generated by $a = (1,0,0,0)$, $b = (0,1,0,0)$, and $d = (0,0,0,1)$. This group has step 2 and $(r_1,r_2) = (3,1)$.

The seven-dimensional geometry $G_{4,3}$. The group $G_{4,3}$ is defined as a lattice in the nilpotent Lie group corresponding to the Lie algebra notated in [114] as $\mathscr{L}_{4,3}$. Concretely, the group is defined as the set $4\mathbb{Z}\times 2\mathbb{Z}\times 2\mathbb{Z}\times 2\mathbb{Z}$ equipped with the multiplication operation

$$\begin{aligned}(a_1,b_1,c_1,d_1)\times(a_2,b_2,c_2,d_2) &= \\ &(b_1d_2+\frac{1}{2}d_2^2c_1+a_2+a_1, c_1d_2+b_1+b_2, c_1+c_2, d_1+d_2),\end{aligned}$$

which has identity element $(0,0,0,0)$. The group is generated by the elements $2b = (0,2,0,0)$, $2c = (0,0,2,0)$, and $2d = (0,0,0,2)$ as witnessed by the formula

$$(4x,2y,2z,2w) = [2b,2d]^x(2d)^w(2b)^y(2c)^z.$$

We work with the Cayley graph $\mathrm{Cay}(G_{4,3},\{2b,2c,2d\})$, whose word metric is comparable to the quasi-norm

$$\|(a,b,c,d)\| = |a|^{1/3}+|b|^{1/2}+|c|+|d|.$$

The group $G_{4,3}$ has step 3 and $(r_1,r_2,r_3) = (2,1,1)$.

The seven-dimensional geometry $G_{5,8}$. The group $G_{5,8}$ is defined as a lattice in the nilpotent Lie group corresponding to the Lie algebra notated in [114] as $\mathscr{L}_{5,8}$. Concretely, the group is defined as the set $\mathbb{Z}^5$ equipped with the multiplication operation

$$(a_1,b_1,c_1,d_1,e_1)\times(a_2,b_2,c_2,d_2,e_2) = \\ (a_1+a_2+b_1d_2, b_1+b_2, -d_1e_2+c_1+c_2, d_1+d_2, e_1+e_2),$$

which has identity element $(0,0,0,0,0)$. The group is generated by the elements $b = (0,1,0,0,0)$, $d = (0,0,0,1,0)$, and $e = (0,0,0,0,1)$ as witnessed by the formula

$$(x,y,z,w,v) = [e,d]^z e^v d^w [b,d]^x b^y.$$

We work with the Cayley graph $\Gamma_{G_{5,8}} = \mathrm{Cay}(G_{5,8},\{b,d,e\})$, whose word metric is comparable to the quasi-norm

$$\|(a,b,c,d,e)\| = |a|^{1/2}+|b|+|c|^{1/2}+|d|+|e|.$$

The group $G_{5,8}$ has step 2 and $(r_1,r_2) = (3,2)$.

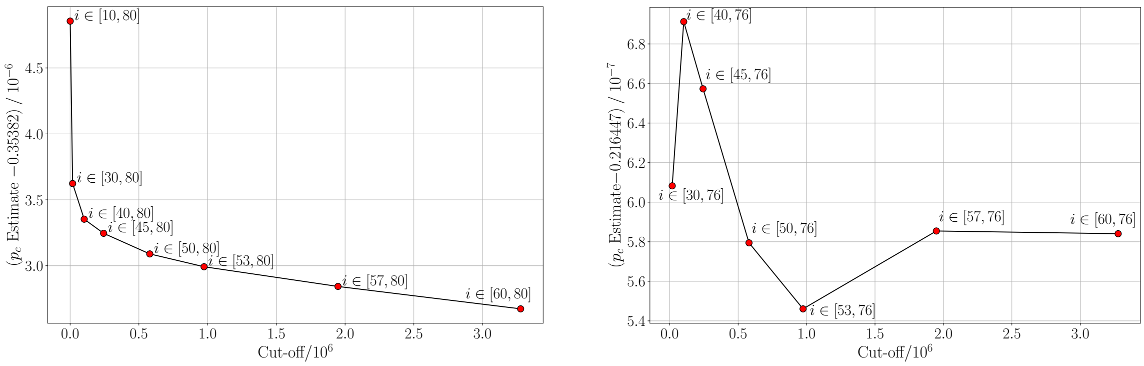
4.3.1 Results for percolation

We now describe the simulations we carried out for percolation on $\mathscr{H}$ and $\mathscr{H}\times\mathbb{Z}$ and the results that we obtained.

Estimating $\mathbf{p_c}$. In each case, we began with a small number of initial runs of invasion percolation, as described in Appendix A, in order to approximate the constants F and z in Appendix A Eq. (1) to achieve a speed up for further runs.

For the Heisenberg group $\mathscr{H}$ we then generated approximately 9×10^5 samples, each with a total number of $\lfloor 100\times 2^{67/4}\rfloor = 11{,}021{,}797$ steps, and a further $100{,}000$ samples each with a total number of $\lfloor 100\times 2^{80/4}\rfloor = 104{,}857{,}600$ steps. We recorded and averaged the sampled bulk-to-boundary ratios a_n at $n = \lfloor 100\times 2^{i/4}\rfloor$ for $5\le i\le 80$, for a total of 76 points. We then used weighted least mean squares to fit the parameters p_c, A, δ in 4.5 to the data. We noticed that removing the points at small n (at the beginning of the runs) shifted our estimate of p_c, lessening the effect of finite-size effects. We plotted the effect of removing small values of n in Figure 4.4a and extrapolated from the resulting data to obtain the estimate $p_c\approx 0.3538225(10)$.

We then repeated this procedure for $\mathscr{H}\times\mathbb{Z}$. This time we generated approximately 6×10^5 samples each with a total number of $\lfloor 100\times 2^{67/4}\rfloor = 11{,}021{,}797$ steps, and a further



(a) The Heisenberg group

(b) The product space $\mathscr{H} \times \mathbb{Z}$

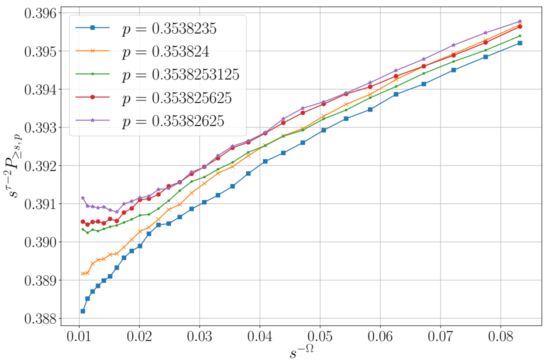
Fig. 4.4 Estimated percolation thresholds for $\mathscr{H}$ and $\mathscr{H} \times \mathbb{Z}$ with varying numbers of excluded initial points. A label of the form $i \in [a,b]$ indicates that the fit was calculated with $a_{\lfloor 100 \times 2^{i/4} \rfloor}, i = a \dots b$.

approximately 200,000 samples each with a total number of $\lfloor 100 \times 2^{76/4} \rfloor = 52,428,800$ steps. We recorded and averaged the sampled bulk-to-boundary ratios a_n at $n = \lfloor 100 \times 2^{i/4} \rfloor$ for $5 \le i \le 76$, for a total of 72 points. Again, plotting the effect of removing small values of n in Figure 4.4b and extrapolating yielded the estimate $p_c \approx 0.2164476(1)$.

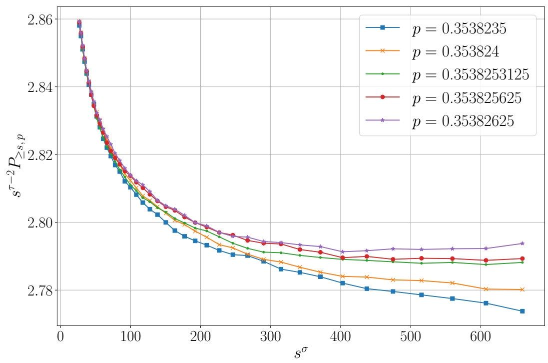
Estimating intrinsic exponents. Having obtained these estimates for the critical probability, we sampled percolation at $p = 0.3538225$ for $\mathscr{H}$ and $p = 0.2164476$ for $\mathscr{H} \times \mathbb{Z}$ using the Leath algorithm. In each case we collected approximately 10^8 samples each with 2^{20}=1,048,576 time steps. We calculated $P_{\ge s}$ and $\mathbb{E}[n/p - t/(1-p) | n \ge s]$ empirically from these samples, fitted the data to the ansatz equations presented in Section 4.2.1, and obtained the estimates $\tau = 2.315$ and $\sigma = 0.4758$ for $\mathscr{H}$ and $\tau = 2.420$ and $\sigma = 0.4988$ for $\mathscr{H} \times \mathbb{Z}$. All these results were in close agreement with previously derived values for $\mathbb{Z}^4$ and $\mathbb{Z}^5$ (see Table 4.1), giving weight to the claim of Conjecture 33.

Refinement and confirmation. Next, we refined our values for the critical probability, and simultaneously added extra weight to the claim that the critical exponents are shared by the Euclidean and non-Euclidean lattices, employing the methods outlined in [253, 322].

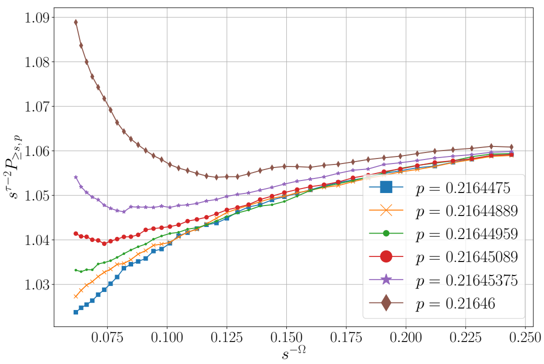
We ran the Leath algorithm at multiple values of p, with between 10^8 and 10^9 samples per value of p, and with each run having $2^{20} = 1,048,576$ steps. As presented in Figure 4.5, we then plotted graphs of $s^{\tau-2} P_{\ge s}$ against $s^{-\Omega}$ and against s^{σ}, where we used the values of τ, σ, and Ω for $\mathbb{Z}^4$ and $\mathbb{Z}^5$ as computed in [287], [162], and [322] respectively. If Conjecture 33 is true, then as explained in [253], the plots against $s^{-\Omega}$ should look approximately linear when $p = p_c$ while the plots against s^{σ} should plateau for large s when $p = p_c$. As such, the figures indicate that the critical probability for $\mathscr{H}$ lies between $p = 0.353824$ and



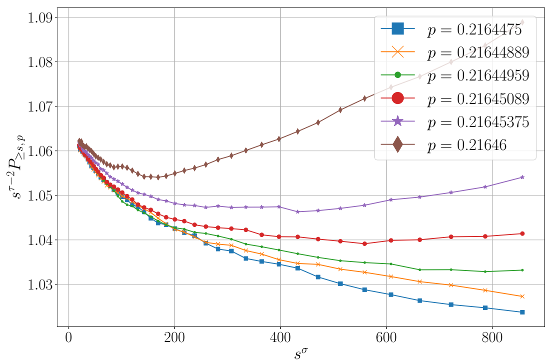
(a) Plot of $s^{\tau-2}P_{\geq s}$ against $s^{-\Omega}$ for $\mathscr{H}$. Smaller deviations from linearity for small $s^{-\Omega}$ indicates that p is closer to p_c.



(b) Plot of $s^{\tau-2}P_{\geq s}$ against s^{σ} for $\mathscr{H}$. A plateau at large s^{σ} indicates that p is close to p_c.



(c) Plot of $s^{\tau-2}P_{\geq s}$ against $s^{-\Omega}$ for $\mathscr{H}\times\mathbb{Z}$. Smaller deviations from linearity for small $s^{-\Omega}$ indicates that p is closer to p_c.



(d) Plot of $s^{\tau-2}P_{\geq s}$ against s^{σ} for $\mathscr{H}\times\mathbb{Z}$. A plateau at large s^{σ} indicates that p is close to p_c.

Fig. 4.5 Runs of the Leath algorithm on $\mathscr{H}$ and $\mathscr{H}\times\mathbb{Z}$ for different values of the percolation probability p.

$p = 0.3538253125$ while the critical probability for $\mathscr{H}\times\mathbb{Z}$ lies between $p = 0.21644889$ and $p = 0.21644959$. In each case, the fact that we do indeed see approximately linear behaviour in the plots against $s^{-\Omega}$ and a large-s plateau in the plots against s^{σ} strongly suggests that Conjecture 33 is true and e.g. the values of τ, σ, and Ω are the same for $\mathscr{H}$ and $\mathbb{Z}^4$.

Estimating extrinsic exponents. Finally, we ran the Leath algorithm on the two non-Euclidean graphs $\mathscr{H}$, $\mathscr{H}\times\mathbb{Z}$ and the two Euclidean graphs $\mathbb{Z}^4$ and $\mathbb{Z}^5$. This time, however, instead of running for a fixed number of steps, we halted the algorithm when it first visited a vertex with extrinsic distance $2^{10} = 1024$ away from the origin for the four-dimensional lattices, and $2^9 = 512$ away from the origin for the five-dimensional lattices. For the non-

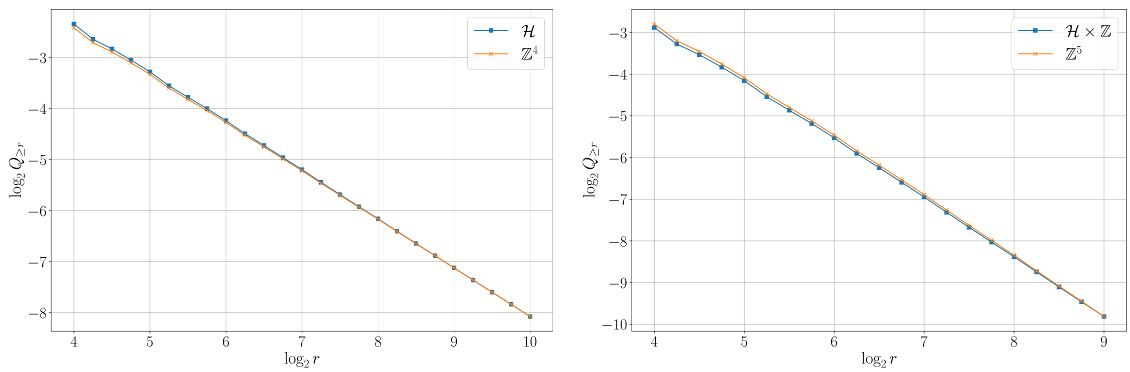
	$\mathscr{H}$	$\mathscr{H}\times\mathbb{Z}$
p_c	0.3538247(7)	0.21644925(36)

Table 4.3 Critical probability estimates.

Euclidean lattices, we ran the algorithm at the previously calculated critical percolation estimates displayed in Table 4.3, and used $p_c = 0.1601312$ for $\mathbb{Z}^4$, extracted from [272, 322], and $p_c = 0.11817145$ for $\mathbb{Z}^5$, extracted from [272]. For each of these graphs, we then plotted $\log_2 Q_{\geq s}$ against $\log_2 s$ for $s = 2^{i/4}$ with $16 \leq i \leq 40$ for the four-dimensional lattices, and $16 \leq i \leq 36$ for the five-dimensional lattices. We calculated the gradients of the final sections of the curves to give the estimates $\rho = 1.047$ for $\mathscr{H}$, $\rho = 1.049$ for $\mathbb{Z}^4$, $\rho = 0.701$ for $\mathscr{H}\times\mathbb{Z}$, and $\rho = 0.683$ for $\mathbb{Z}^5$. The large finite-size effects, especially in the five-dimensional case, meant that the computational resources available to us were insufficient to compute ρ to a high level of precision. Using the scaling relation $\tau = 1 + d/(d - 1/\rho)$, we computed secondary estimates $\tau = 2.314$ for $\mathscr{H}$, $\tau = 2.313$ for $\mathbb{Z}^4$, and $\tau = 2.400$ for $\mathscr{H}\times\mathbb{Z}$, $\tau = 2.414$ for $\mathbb{Z}^5$.

4.3.2 Results for lattice trees

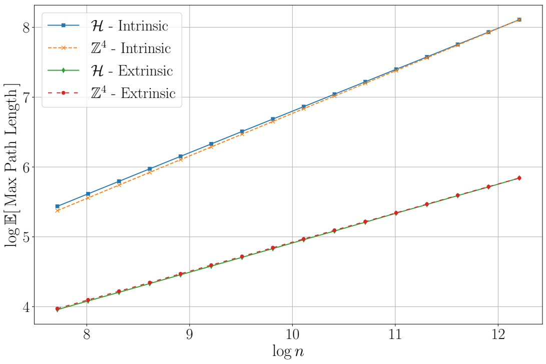
We now describe the simulations we carried out of lattice trees on $\mathscr{H}$, $\mathbb{Z}^4$, $\mathscr{H}\times\mathbb{Z}$, $\mathbb{Z}^5$, $G_{4,3}$, and $G_{5,8}$ and the results that we obtained. We ran our own simulations on $\mathbb{Z}^4$ and $\mathbb{Z}^5$ for better comparability with our non-Euclidean simulations since the simulations of [323] used much smaller tree sizes, and [186] did not estimate the intrinsic exponent.



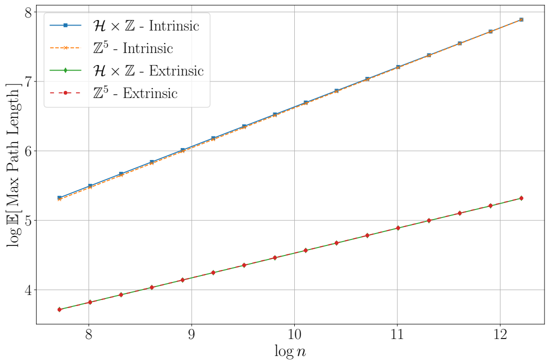
(a) Four-dimensional lattices.

(b) Five-dimensional lattices.

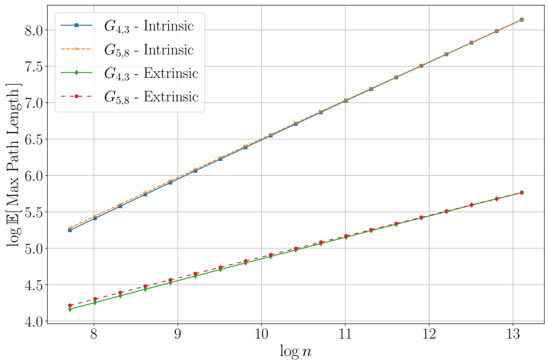
Fig. 4.6 Log-log plots of the radius tail $Q_{\geq r}$ against r for the transitive lattices $\mathscr{H}$, $\mathbb{Z}^4$, $\mathscr{H}\times\mathbb{Z}$, and $\mathbb{Z}^5$, produced using runs of the Leath algorithm at their respective critical probabilities. The curves are vertically translated to allow for easy comparison of their final gradients, which are seen to be in close agreement in both cases.



(a) $\mathscr{H}$ and $\mathbb{Z}^4$.



(b) $\mathscr{H} \times \mathbb{Z}$ and $\mathbb{Z}^5$.



(c) $G_{4,3}$ and $G_{5,8}$.

Fig. 4.7 Log-log plots of mean intrinsic and extrinsic distances between the end-points of maximum (intrinsic) length paths in the lattice tree as functions of tree size. In each case, the vertical positioning of the curves has been adjusted for ease of comparison of the gradients of the final segments.

For each of the graphs that we considered, we initialised the CPBC MCMC algorithm with tree sizes $s = \lfloor 10000 \times 20^{i/10} \rfloor$ from $i = -5$ to $i = 10$ for the four and five dimensional lattices, and up to $i = 13$ for the seven dimensional lattices. We collected between 100,000 samples and 500,000 samples for each tree size. For a tree of size s, we evolved the algorithm for an initial $4s$ steps, and then collected a sample every $2s$ steps thereafter. The initial trees of size s were taken to be paths with $s/2$ vertices lying along a suitable coordinate axis with additional edges coming off each vertex in another fixed coordinate direction. An estimate of the extrinsic exponent ν was calculated by finding the gradients of the final section of the relevant log-log curve. An estimate of the intrinsic exponent ρ was found by first averaging the two log-log curves for branch-size and intrinsic radius, before taking the gradient.

4.4 Self-similar fractals

In this section we give a brief introduction to the self-similar fractals we consider, and describe our results concerning critical percolation on them. The fractals we consider will be

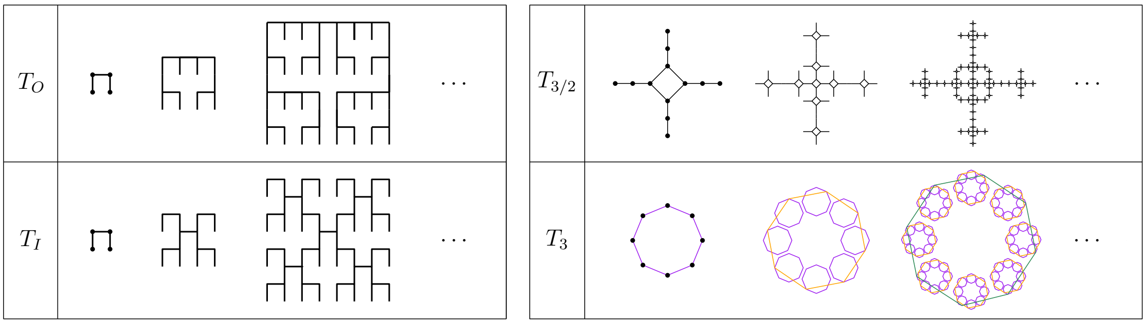


Fig. 4.8 Illustration of the recursive construction of graphical approximants to the fractal trees we consider. The graphs used to approximate $T_{3/2}$ and T_3 are not trees but **quasi-trees**, with bounded-length cycles that disappear in the continuum limit. Note also that we have drawn the edges of the graphs approximating T_3 with different lengths in order to represent them cleanly in the plane; as a result, these drawings do not accurately represent the intrinsic geometry of the graphs in question. The colours are included to aid visualisation since the drawing is not planar. The tree T_4 is similar but is constructed from 16-gons rather than octagons.

defined as the scaling limits of sequences of 'prefractal' graphs generated by an initial seed graph and a recursive rule describing how the generation $n+1$ prefractal is constructed from copies of the generation n prefractal. In addition to the continuum fractal scaling limit, we can also take the **Benjamini-Schramm limit** of this growing sequence of prefractal graphs, which describes how the graph looks in the vicinity of a uniform random vertex. In each of the cases we consider, the Benjamini-Schramm limit exists and is an infinite, locally finite random rooted graph, so that we can define the critical probability p_c and critical exponents τ and σ with respect to this infinite limit graph.

Recursive rules. Each of the fractal **trees** we consider will be constructed using a hierarchical coordinate system in the following way, which makes their Benjamini-Schramm limit easy to describe. Let $N \in \mathbb{N}$ and define $\mathbb{X} = \mathbb{Z}_N^\infty$, where $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. We call $\mathbb{X}$ the **coordinate space** of the fractal and write points in $\mathbb{X}$ as $x = (\ldots, x_1, x_0)$. The number N will represent the number of 'marked points' that are used to specify specify how to construct the prefractal in one generation from the prefractal at the previous generation. A **seed graph** is defined to be a connected, undirected graph with vertex set $\mathbb{Z}_N$. Given an undirected graph G_g on $\mathbb{Z}_N \times \mathbb{Z}_N$, we define the **contraction** $C[G_g]$ of G_g to be the graph with vertex set $\mathbb{Z}_N$, and with two vertices $i, j \in \mathbb{Z}_N$ connected if an only if there exist $a, b \in \mathbb{Z}_N$ such that $(i,a) \sim (j,b)$ in G_g. We say that an undirected graph G_g on $\mathbb{Z}_N$ is a **generator graph** if its contraction $C[G_g]$ is connected.

	N	Seed edges $E[G_s]$	Generator edges $E[G_g]$
T_3	8	$\{\{i,i+1\}: 0 \leq i \leq 7\}$	$\{\{(i,i+1),(j,j+1)\}: 0 \leq i,j \leq 7, i-j = \pm 1\}$
T_4	16	$\{\{i,i+1\}: 0 \leq i \leq 15\}$	$\{\{(i,i+1),(j,j+1)\}: 0 \leq i,j \leq 15, i-j = \pm 1\}$
$T_{3/2}$	8	$\{\{2i,2i+2\}: 0 \leq i \leq 3\}$	$\{\{(2i,2i+5),(2j,2j+5)\}: 0 \leq i,j \leq 3, i-j = \pm 1\}$

Table 4.4 Formal encodings of the fractal trees T_3, T_4, and $T_{3/2}$ used in our explicit recursive scheme for constructing fractals. All addition is computed modulo N.

Given a seed graph G_s and a generator G_g, we define the fractal graph $\mathbb{G} = \mathbb{G}(G_g, G_s)$ to be the graph with vertex set $\mathbb{X}$ and where two distinct points $x = (\ldots, x_1, x_0)$ and $y = (\ldots, y_1, y_0)$ in $\mathbb{X}$ are connected by an edge if one of the following two conditions hold:

- $x_i = y_i$ for every $i \geq 1$ and $x_0 \sim y_0$ in G_s, or
- $m = \inf\{i \geq 1 : x_j = y_j$ for every $j \geq i\}$ is finite and strictly larger than one, $x_i = x_{m-1}$ and $y_i = y_{m-1}$ for every $0 \leq i \leq m-1$, and (x_m, x_{m-1}) is adjacent to (y_m, y_{m-1}) in G_g.

Note that $x, y \in \mathbb{X}$ belong to the same connected component of $\mathbb{G}$ if and only if $x_i = y_i$ for all sufficiently large i, so that $\mathbb{G}$ has uncountably many connected components. For each n the finite subgraphs of $\mathbb{G}$ induced by the sets $\Lambda_n(y) = \{x : x_i = y_i$ for every $i \geq n\}$ have isomorphism class that does not depend on the choice of $y \in \mathbb{X}$, and we define G_n to be a graph with this isomorphism class. This ensures that G_1 is equal to the seed graph G_s, while for each $n \geq 1$ we can form G_{n+1} by attaching edges between N copies of G_n according to the combinatorics of the generator G_g. The Benjamini-Schramm limit of the graph sequence $(G_n)_{n\geq 1}$ is equal to the rooted graph (G_∞, o) defined by taking $o \in \mathbb{X}$ to have i.i.d. uniform coordinates in $\mathbb{Z}/N\mathbb{Z}$ and taking G_∞ to be the connected component of o in the uncountably infinite graph $\mathbb{G}$.

Algorithmically, this representation of the infinite-volume prefractal (G_∞, o) has the advantage that the initial sequence of coordinates $(o_1, \ldots, o_k)$ typically determines the isomorphism class of a large neighbourhood around o, and we can sample more terms of this sequence on an as-needed basis as we explore the percolation cluster of o or run invasion percolation from o.

The fractal trees T_O and T_I presented in Figure 4.3 are both easily represented via this recursive scheme with $N = 4$: In both cases we take the seed graph to have edge set $\{\{0,1\},\{1,2\},\{2,3\}\}$. For the 'outer' tree T_O we take the generator graph to have edge set $\{\{(0,1),(1,0)\},\{(1,2),(2,1)\},\{(2,3),(3,2)\}\}$, while for the 'inner' tree T_I we take the generator graph to have edge set $\{\{(0,2),(1,3)\},\{(1,3),(2,0)\},\{(2,0),(3,1)\}\}$. We encourage the reader to work through this simple example to see how our fractal encoding scheme works in practice. The reader may also find it enlightening to consider how the infinite line graph $\mathbb{Z}$ can be expressed as a Benjamini-Schramm limit of graphs defined through a similar recursive scheme.

Besides T_O and T_I we will also consider three further fractal trees which we call T_3, T_4, and $T_{3/2}$. In fact, it will be convenient to consider graphical approximants of these trees that are not themselves trees, but are quasi-trees in the sense that they include cycles of bounded length which disappear in the continuum limit. The formal definitions of these fractal trees in terms of our recursive scheme are stated in Table 4.4 with graphical representations of the first three generations given in Figure 4.8.

4.4.1 Fractal dimensions

Let us now briefly review the definitions and background on the dimensions we consider, referring the reader to [137] for further background.

We begin with the Hausdorff dimension and topological dimension, which are both classical. Given a non-empty metric space X, the d-dimensional Hausdorff outer measure of a set $S \subset X$ is defined as

$$\mathscr{H}^d(S) = \liminf_{r\to 0} \Big\{ \sum_I r_i^d : \text{ there is a cover of } S \text{ by balls of radii } 0 < r_i < r \Big\}.$$

The **Hausdorff dimension** of X is then

$$\dim_H X = \inf\{d \geq 0 : \mathscr{H}^d(X) < \infty\}.$$

In non-pathological examples, one typically has that a continuum fractal has Hausdorff dimension d if and only if the Benjamini-Schramm limit of its prefractal approximants has volume-growth dimension d in the sense that $|B(o,r)| \approx r^d$ as $r \to \infty$. Moreover, in non-pathological examples one also has that the Hausdorff dimension is additive in the sense that $\dim_H X \times Y = \dim_H X + \dim_H Y$; see [137, Chapter 7] for precise theorems to this effect. All the examples we consider will have the very strong property of being **Ahlfors regular**, which ensures that the Hausdorff dimension is indeed additive for these examples.

Suppose that we construct a fractal via a recursive rule as discussed at the beginning of this section, and let $(G_n)_{n\geq 1}$ be the associated sequence of prefractal graphical approximants so that G_n has N^n vertices for some $N \geq 2$. If the ratio of diameters of G_{n+1} and G_n tends to N^α as $n \to \infty$, then the associated continuum fractal will typically have Hausdorff dimension $\log N/\log N^{\alpha} = 1/\alpha$. Again, all the examples we consider are sufficiently well-behaved that these heuristics can easily be turned into rigorous proofs with a little work. See [137, Section 9.2] for detailed justifications of various related formulae. It follows from these considerations that both fractal trees T_O and T_I have Hausdorff dimension 2: at each successive scale of approximation the number of vertices is multiplied by four while the diameter roughly doubles. Similarly, in T_3 and T_4 the diameter roughly doubles in each generation while the volume increases by a factor of 8 or 16 as appropriate, so that these trees have Hausdorff dimensions $\log 8/\log 2 = 3$ and $\log 16/\log 2 = 4$ respectively. Finally, in $T_{3/2}$ the diameter roughly quadruples at each scale while the volume increases by a factor of 8, so that $T_{3/2}$ has Hausdorff dimension $\log 8/\log 4 = 3/2$.

The **topological dimension** (a.k.a. lower inductive dimension) $\dim_t X$ of a separable metric space X is defined inductively by $\dim_t \varnothing = -1$ and

$$\dim_t X = \inf\{d : X \text{ has a basis } U \text{ such that } \dim_t \partial U \leq d-1 \text{ for every } U \in U\}.$$

Note that **real trees** such as $\mathbb{R}$, $[0,1]$, and the fractal trees we consider always have topological dimension 1. The topological dimension is not additive in general but always satisfies the inequality $\dim_t X \times Y \leq \dim_t X + \dim_t Y$ [133, Theorem 1.5.16]. Moreover, if Y is a subspace of X then $\dim_t Y \leq \dim_t X$, a fact referred to as the **subspace theorem** [133, Theorem 1.1.2]. If $X = T_1 \times T_2 \times \cdots \times T_k$ is a product of real trees then it follows that $\dim_t X \leq k$, and since X contains a copy of the space $[0,\varepsilon]^k$ for some $\varepsilon > 0$ it follows from the subspace theorem that $\dim_t X = k$. This equality determines the topological dimension for all the examples we consider.

The **topological Hausdorff dimension** is a much more recent notion of dimension that was introduced by Balka, Buczolich and Elekes [39]. The topological Hausdorff dimension $\dim_{tH} X$ of a non-empty metric space X is defined to be

$$\dim_{tH} X = \inf\{d : X \text{ has a basis } \mathscr{U} \text{ such that } \dim_H \partial U \leq d-1 \text{ for every } U \in \mathscr{H}\},$$

where $\dim_H \varnothing$ is defined to be -1. It is proven in [39, Theorem 4.21] that

$$\dim_{tH}(X \times [0,1]) = 1 + \dim_H(X), \tag{4.6}$$

for every non-empty and separable metric space X. This allows us to always reduce the computation of the topological Hausdorff dimension to that of the Hausdorff dimension by working only with products with $[0,1]$; this corresponds to taking products with $\mathbb{Z}$ for the relevant Benjamini-Schramm limits.

It remains to introduce the **spectral dimension**, which is most easily defined for the infinite Benjamini-Schramm limit (G_∞, o) associated to the fractal. Indeed, an infinite connected graph G is said to have spectral dimension $d_s = \dim_s G$ if the simple random walk return probabilities $p_n(v,v)$ satisfy

$$p_{2n}(v,v) = n^{-d_2/2+o(1)}$$

as $n \to \infty$ for each vertex v of G. (In principle, one can define the spectral dimension of a continuum fractal directly by first defining Brownian motion on that fractal, but this is a very delicate matter in general.) It is easily seen from the definition that the spectral dimension is additive with respect to products in the sense that if G and H are two infinite, connected graphs then $\dim_s G \times H = \dim_s G + \dim_s H$. Most fractal **trees** T have spectral and Hausdorff dimensions related by the formula

$$\dim_s T = \frac{2\dim_H T}{\dim_H T + 1},$$

and it is not difficult to justify that this equality does indeed hold for all the fractal trees we consider. (Indeed, this equality should hold whenever the **effective resistance** between a vertex and the boundary of the ball of radius r grows like $r^{1-o(1)}$, and for trees this holds whenever subsequential limits do not have vertices of infinite degree; this can be deduced from the same methods used in [193, Section 8].) Thus, T_O, T_I, $T_{3/2}$, T_3, and T_4 have spectral dimensions $4/3$, $4/3$, $6/5$, $3/2$, and $8/5$ respectively.

4.4.2 Equidimensional fractal products

We now define the two pairs of equidimensional fractal products on which we will study percolation. The first pair is given by

$$H_1 = T_O \times [0,1] \quad \text{and} \quad H_2 = T_I \times [0,1].$$

It follows from the above discussion that these two fractals both have Hausdorff dimension 3, topological dimension 2, spectral dimension $7/3$ and topological Hausdorff dimension 3. The second pair is given by

$$H_3 = T_{3/2} \times T_{3/2} \times T_4 \times [0,1] \text{ and } H_4 = T_3 \times T_3 \times [0,1] \times [0,1].$$

These two fractals both have Hausdorff dimension

$$\frac{3}{2}+\frac{3}{2}+4+1=3+3+1+1=8,$$

topological dimension 4, spectral dimension

$$\frac{6}{5}+\frac{6}{5}+\frac{8}{5}+1=\frac{3}{2}+\frac{3}{2}+1+1=5,$$

and topological Hausdorff dimension 8. This pair of examples is interesting to study in part because the two fractals H_3 and H_4 seem to have 'the same dimensions for different reasons', with different components of their defining products making up different proportions of their shared Hausdorff and spectral dimensions. This would seem to make them a prime candidate for a failure of universality, although in the end the large finite-size effects made it difficult for us to compare the Fisher exponents in the two cases.

Again, we do not work directly with continuum fractals, but instead consider the Benjamini-Schramm limits defined via the recursive schemes specifying the trees T_O, T_I, T_3, T_4, and $T_{3/2}$ above. Thus, for example, when we simulate percolation on H_3 we are really simulating percolation on the product of two independent copies of the Benjamini-Schramm limit associated to $T_{3/2}$, a further independent copy of the Benjamini-Schramm limit associated to T_4, and one copy of $\mathbb{Z}$. Since we will always use these same graphical approximations, for clarity of exposition we will abuse the terminology by speaking simply of 'percolation on H_1' and so on.

4.4.3 Results

We now discuss the results of our simulations of percolation on the self-similar fractals, beginning with the equidimensional pair H_1 and H_2. As with the transitive lattices, we began by running invasion percolation with approximately 10^6 samples, each time recording the bulk-to-boundary ratios a_n at $n=\lfloor 100\times 2^{i/4}\rfloor$, $0\leq i\leq 67$. The outcome of these simulations is recorded in Figure 4.9.

In order to estimate p_c from this data, we carried out a similar analysis to the transitive case, varying the amount we cut-off at the beginning before curve-fitting. This gave the initial estimates $p_c=0.4249$ for H_1 and $p_c=0.4232$ for H_2.

We emphasise that the oscillations in Figure 4.9 are in fact a feature rather than noise. This was a consistent appearance throughout our simulations for fractals, both for invasion percolation and the Leath algorithm. For the former the oscillations decayed, while for the latter they grew. For the former it meant more initial data had to be discarded, and for the

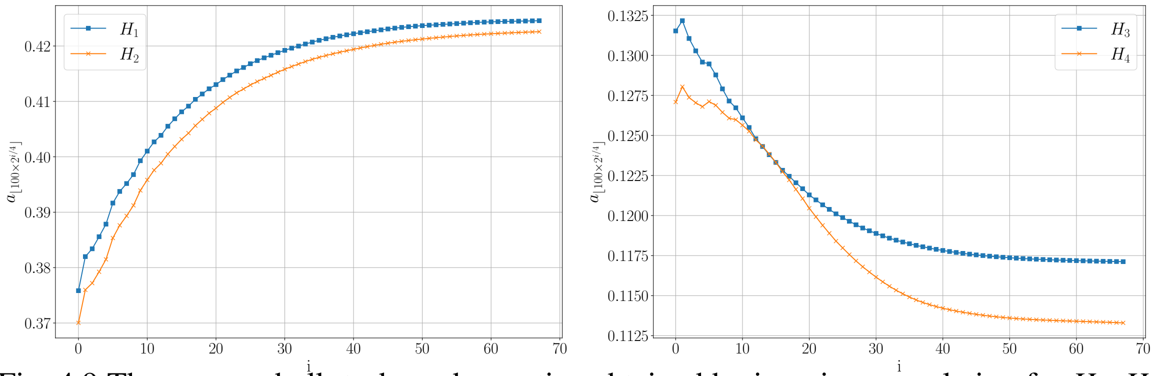


Fig. 4.9 The average bulk-to-boundary ratios obtained by invasion percolation for H_1, H_2 (left) and H_3, H_4 (right). Note the very pronounced finite-size effects for H_4.

latter it made estimating linearity more difficult. This became more of a problem for some of the higher dimensional fractals, in particular in relation to runs of the Leath algorithm.

Having obtained an initial estimate for p_c, we then sampled the percolation cluster for H_1 and H_2 at differing values of p, and produced log-log plots of the average volume tail $P_{\geq s,p} = \mathbf{E}\left[\mathbb{P}_p(|K_o| \geq s)\right]$ against s at different values of p, taking $s = \lfloor 2^{i/4} \rfloor$ with maximal i ranging up to 27 and with between 10^6 and 10^8 samples for each value of p. The outcomes of this investigation are recorded in Figure 4.10 below. Drawing tangents along the curves of Figure 4.10 reveals that $p_c = 0.42545(5)$ for H_1 and $p_c = 0.423225(25)$ for H_2. It is interesting to note that invasion percolation gave a far more accurate reading for H_2 than H_1.

Finding the gradient of the these critical log-log plots gave $\tau = 2.195(5)$ for H_1 and a more precise reading of $2.151(1)$ for H_2, where uncertainties were estimated by varying the portions of the midsections of the curves over which the gradients were calculated and calculating over multiple curves corresponding to probabilities within the aforementioned ranges for p_c. The more prominent presence of the oscillations for H_1 made the reading for

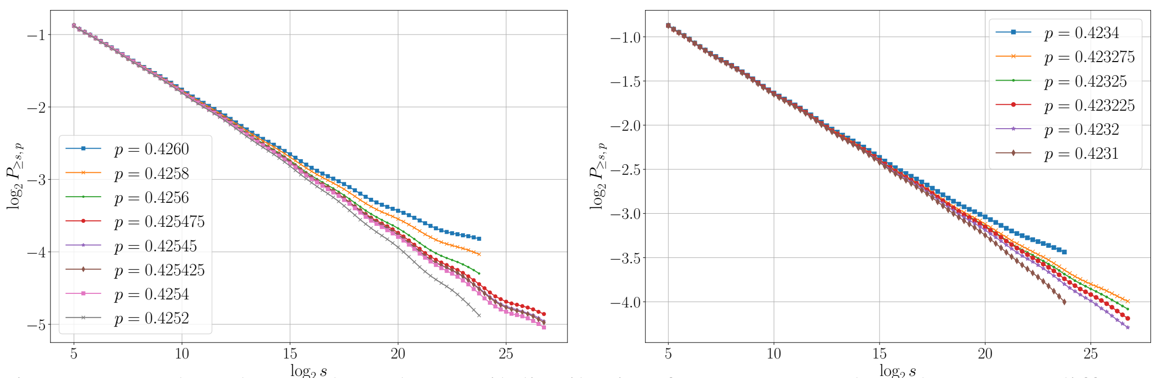


Fig. 4.10 Log-log plots of the volume tail distribution for H_1 (top) and H_2 (bottom) at different values of the percolation probability p. Smaller deviations from linearity indicate that p is closer to the critical probability p_c.

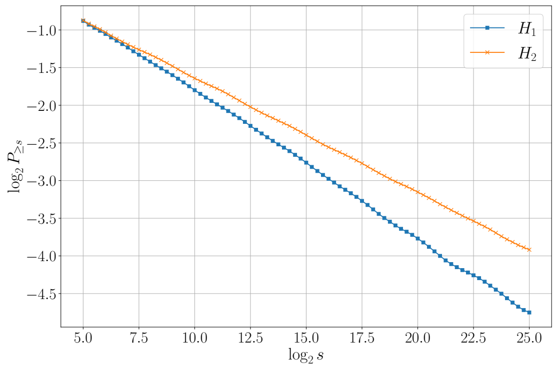


Fig. 4.11 Log-log plots of the volume tail distribution for H_1 and H_2 at the estimated critical probabilities of $p = 0.42545$ for H_1 and $p = 0.423225$ for H_2. The two lines clearly have distinct slopes, lending strong evidence to the claim that H_1 and H_2 have distinct values of the critical exponent τ.

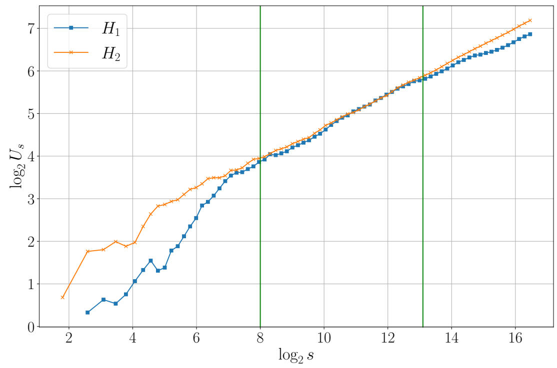


Fig. 4.12 Plot of $\log_2 U_s$ against $\log_2 s$ with U_s calculated by averaging $U_{s,p}$ at $p = 0.4242, 0.4246, 0.4248$ for H_1 and $p = 0.4231, 0.4234$ for H_2.

τ less precise. A direct visual comparison of the two critical log-log volume tail plots is provided in Figure 4.11.

We now turn to estimating the exponent σ, for which our results are less clear. Unfortunately, we found the method based on the ansatz Eq. (4.4) to work very poorly for these graphs, with the resulting expectation requiring a prohibitively large number of samples to stabilise. As such, we resorted to a more ad-hoc analysis to estimate σ. First, we rearranged the ansatz formula Eq. (4.2) to obtain that

$$u_{s,p} := \frac{\log P_{\geq s,p} - \log P_{\geq s,p_c}}{p - p_c} = C_1 s^{\sigma} + \dots. \tag{4.7}$$

Then, for each fractal, we took the average U_s of $u_{s,p}$ over a selection of near-critical p and used curve-fitting over s to output a value of σ. The outcome of this investigation is recorded in Figure 4.12. As can be seen from this figure, the results of this investigation are inconclusive at best, with large non-linearities in the curve for H_1 preventing us from getting a reliable estimate of σ in this case. There seems to be a section of alignment, but not enough to confirm or disconfirm that the σ critical exponents are the same. The deviation at the end may indicate a different σ exponent, or it could be due to the the imprecision of our estimate of p_c, or it could indicate that $p - p_c$ is large enough that the approximations in the derivation of Eq. (4.7) are not valid.

H_3 **and** H_4**.** We now turn to our results for the equidimensional fractal products H_3 and H_4. Running invasion percolation and extrapolating as above (see Figure 4.9) gave the estimates $p_c = 0.11705$ for H_3 and $p_c = 0.11326$ for H_4. Having obtained this estimate we then sampled the percolation cluster at a variety of nearby values of p and produced log-log plots of both the volume tail distribution and the quantity $\mathbb{E}[n/p - t/(1-p)|n \geq s]$, as described in Section 4.2.1, where we used between 5×10^8 and 3×10^9 samples to estimate each of the relevant quantities. The outcomes of these investigations are recorded in Figures 4.13 and 4.14.

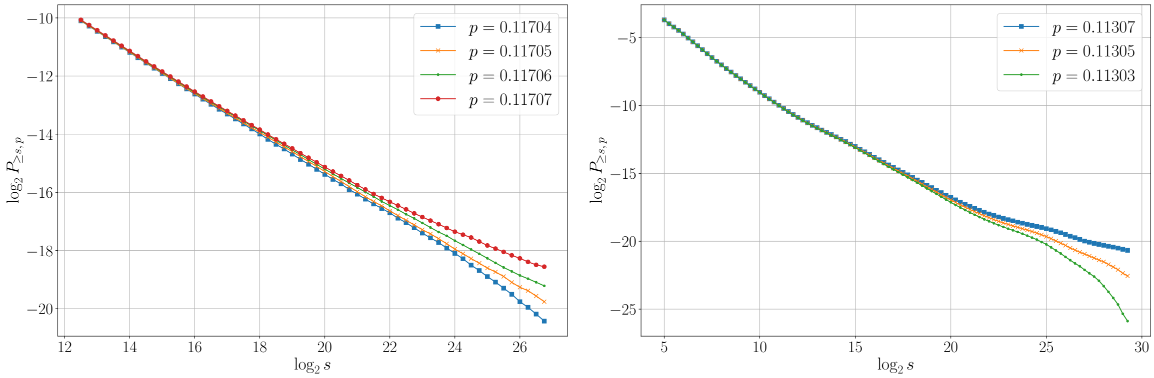


Fig. 4.13 Log-log plots of the volume tail distribution for H_3 (left) and H_4 (right) with different values of the percolation probability p. Smaller deviations from linearity indicate that p is closer to p_c. The large deviations in linearity present in all the curves plotted for H_4 make them difficult to compare to the critical volume tail distribution curve for H_3, or to reliably estimate the relevant value of τ in this case.

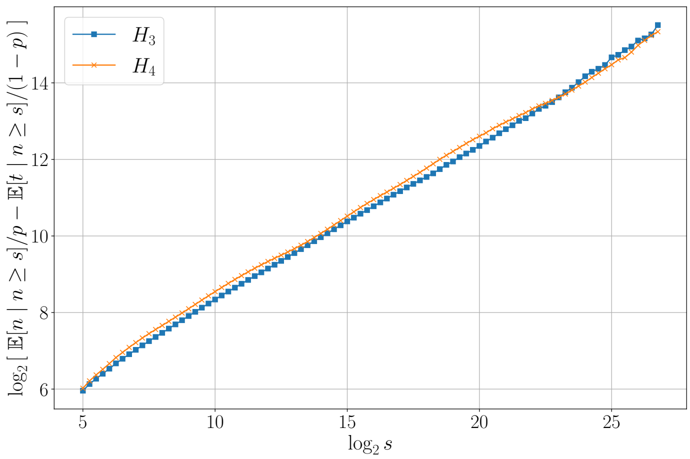


Fig. 4.14 Graphs to estimate σ, with $p = 0.11705$ for H_3 and $p = 0.11305$ for H_4.

Plotting tangents to the final segments of each of the curves in Figure 4.13 and finding the closest linear fit with the midsection of the curve gave $p_c = 0.11705(1)$ for H_3 and $p_c = 0.11305(2)$ for H_4. Invasion percolation therefore gave an extremely accurate estimate for H_3 but a much less accurate estimate for H_4. This was due to the prominent oscillatory behaviour in the bulk-to-boundary ratios, most likely due to the presence of the T_4 tree in H_4.

By considering tangents, we estimated $\tau = 2.66(1)$ for H_3 and $\sigma = 0.41(1)$ for both H_3 and H_4, suggesting that these two fractals share the same value of the exponent σ. Unfortunately, the large deviations from linearity in the volume tail distribution plots for H_4 prevented us from obtaining an estimate on τ for this fractal to any reasonable level of accuracy, and it is unclear whether one should expect H_3 and H_4 to share a common value of this exponent.

4.5 Discussion and open questions

Summary. In this paper we presented strong numerical evidence in support of our conjecture that the critical exponents governing critical percolation and lattice trees on transitive lattices of polynomial volume growth depend only on the dimension and not on any other features of the large-scale geometry. For self-similar fractals, we showed that the situation is more complicated: we presented examples of two fractals having the same Hausdorff, spectral, topological, and topological Hausdorff dimensions, but which have distinct numerical values of the percolation Fisher exponent τ. On the other hand, we do not rule out that the exponent σ is determined by these dimensions. This may be related to the phenomenon of **weak universality** as discussed in [274] and deserves closer investigation in future work.

Open Questions. We now present a collection of open problems and directions for future research:

1. Provide theoretical reasoning either in support of or against Conjecture 33. Is there a reason these exponents might be extremely close without being exactly the same?
2. Does the conjecture hold for other models with an upper-critical dimension $d_c > 4$ such as the minimal spanning tree and invasion percolation? It may be interesting to consider the $|\varphi|^c$ spin model, which has upper-critical dimension $2c/(c-2)$ when $c > 2$, and so can be made arbitrarily large.
3. Are the exponents describing logarithmic corrections at the upper-critical dimension d_c independent of the choice of d_c-dimensional transitive graph? This question is also interesting for models with $d_c = 4$ such as the Ising model, self-avoiding walk, and the uniform spanning tree.
4. Further investigate the extent to which the exponent σ is constrained by the dimensions we consider. Do H_1 and H_2 have the same value of σ? Is there a reason why σ would

be less sensitive to the geometry than τ? Is this related to the phenomenon of weak universality?

Of course, there are many other directions that one might pursue in relation to our work. In addition to the endless variety of fractals, there are also many other transitive graphs of polynomial growth for which the problems studied in this paper are interesting [114, 251]. There are also many other exponents associated with the models that one could seek to estimate - in particular, the exponents characterising the intrinsic radii of critical percolation clusters, and the exponent characterising the sub-exponential correction to growth of the number of lattice trees of size n [89, 204, 217], although we note that it has been argued in [286] that a scaling relation between the growth correction exponent and the extrinsic exponent holds.

Finally, let us remark that our focus in this paper has been to investigate a large number of different examples rather than devoting too much computing time to a very in-depth analysis of any particular example. It may be worthwhile in the future to subject one or two of the quantities we investigated to more intensive study.

Simulation code:
https://gitfront.io/r/user-5838678/a2bfdf5d13aa3ad8f3d2cc247b5100aed0c51e83/LoU/.

Chapter 5

[D] Most transient random walks have infinitely many cut times

Abstract. We prove that if $(X_n)_{n\geq 0}$ is a random walk on a transient graph such that the Green's function decays at least polynomially along the random walk, then $(X_n)_{n\geq 0}$ has infinitely many cut times almost surely. This condition applies in particular to any graph of spectral dimension strictly larger than 2. In fact, our proof applies to general (possibly nonreversible) Markov chains satisfying a similar decay condition for the Green's function that is sharp for birth-death chains. We deduce that a conjecture of Diaconis and Freedman (Ann. Probab. 1980) holds for the same class of Markov chains, and resolve a conjecture of Benjamini, Gurel-Gurevich, and Schramm (Ann. Probab. 2011) on the existence of infinitely many cut times for random walks of positive speed.

5.1 Introduction

Let $(x_n)_{n\geq 0}$ be a sequence taking values in some set Ω. A **cut time** of $(x_n)_{n\geq 1}$ is a time $n \in \mathbb{Z}_{\geq 0}$ for which the sets $\{x_i : i \leq n\}$ and $\{x_i : i > n\}$ are disjoint. The study of cut times of random walks was initiated by Erdős and Taylor in 1960 [134], who proved lower bounds on the densities of cut times for simple random walks on the integer lattices $\mathbb{Z}^d$ for $d \geq 5$, showing that in this case the *doubly infinite* random walk has a positive density of cut times. The lower dimensional cases $d = 3,4$ are more complicated, with the singly infinite random walk having an infinite, density zero set of cut times and the doubly infinite random walk having no cut times almost surely; see [91, 92, 108, 131, 234, 236, 237] for highlights of the literature and [238] for an overview. Extending these results beyond the simple random walk on $\mathbb{Z}^d$, James and Peres [202] and Blanchere [83] proved that every centered, finite-range random walk on a transient Cayley graph has infinitely many cut times almost surely; see also the recent work [252] for a more robust analysis. The proofs of these results rely on delicate estimates on the gradient of the Green's function that are not available in more general settings, with the works [83, 202] also employing a case analysis of the different possible transitive low-dimensional geometries.

Indeed, while transience is of course a necessary condition for a random walk to have infinitely many cut times, the converse implication quickly breaks down once we leave the transitive setting: James, Lyons and Peres [201] constructed an example of a birth-death chain that is transient but has finitely many cut times almost surely (see also [109]), and Benjamini, Gurel-Gurevich, and Schramm [64] showed that the same behaviour is possible for random walks on bounded degree graphs. On the other hand, Benjamini, Gurel-Gurevich, and Schramm [64] also prove that a graph is transient if and only if the *expected* number of cut times of the random walk is infinite, which suggests that most 'non-pathological' transient random walks should indeed have infinitely many cut times. It is also known that the set of edges crossed by a random walk always spans a recurrent graph almost surely [63, 65], a property that holds trivially when there are infinitely many cut times.

In this paper, we prove a new, very easily satisfied criterion for a transient Markov chain to have infinitely many cut times almost surely, applying in particular to any Markov chain in which the Green's function decays at least polynomially along a trajectory of the chain. Our result demonstrates that most transient Markov chains arising in examples will have infinitely many cut times almost surely, and, in particular, provides a simple and unified treatment of the transitive locally finite case.

We now state our main theorem. Let $M = (\Omega, P)$ be an irreducible Markov chain consisting of countable state space Ω and transition kernel P. For each $x \in \Omega$, we write $\mathbb{P}_x$ and $\mathbb{E}_x$ for probabilities and expectations taken with respect to the law of the Markov chain trajectory $(X_n)_{n\geq 0}$ started at x, and write $\mathbb{G}(x,y)$ for the Green's function $\mathbb{G}(x,y) = \sum_{n\geq 0} P^n(x,y) = \mathbb{E}_x \sum_{n\geq 0} \mathbb{1}(X_n = y)$. We say that a sequence of non-negative numbers $(a_n)_{n\geq 0}$ **decays at least polynomially** as $n \to \infty$ if there exists a constant $c > 0$ and an integer N such that $a_n \leq n^{-c}$ for every $n \geq N$.

Theorem 34. *Let $M = (\Omega, P)$ be a countable Markov chain and let $(X_n)_{n\geq 0}$ be a trajectory of M started at some state $x \in \Omega$. If there exists a decreasing bijection $\Phi : [0,\infty) \to (0,1]$ such that*

$$\int_0^1 \frac{1}{u(1 \vee \log \Phi^{-1}(u))}\, \mathrm{d}u = \infty \qquad \textit{and} \qquad \limsup_{n\to\infty} \frac{\mathbb{G}(X_n,y)}{\Phi(n)} < \infty \quad \textit{a.s. } \forall y \in \Omega, \tag{5.1}$$

then the trajectory $(X_n)_{n\geq 0}$ has infinitely many cut times almost surely. In particular, the same conclusion holds if $\mathbb{G}(X_n,y)$ decays at least polynomially as $n\to\infty$ for each fixed $y\in\Omega$ almost surely.

We stress that the $\Phi^{-1}(u)$ term appearing in (5.1) denotes the *inverse* of Φ rather than its reciprocal. Note that if the Markov chain is irreducible we can replace the decay condition appearing here with the condition that $\limsup_{n\to\infty}\Phi(n)^{-1}\mathbb{G}(X_n,X_0)<\infty$ a.s.; In general it suffices that $\limsup_{n\to\infty}\Phi(n)^{-1}\mathbb{G}(X_n,X_m)<\infty$ for each $m\geq 0$ a.s.

Remark 6. Theorem 34 applies to some decay rates that are slightly slower than polynomial, such as that given by $\Phi(n)=\exp(-\frac{\log n}{\log\log n})$. In Section 5.4, we discuss how the results of Csáki, Földes, and Révész [109] imply that the integral condition of Theorem 34 is sharp for birth-death chains and hence cannot be improved in general.

Theorem 34 easily implies various sufficient conditions for a Markov chain trajectory to have infinitely many cut times almost surely. One particularly simple such condition is as follows.

Corollary 35. *Let $M=(\Omega,P)$ be a countable Markov chain and let $X=(X_n)_{n\geq 0}$ be a trajectory of M started at some state $x\in\Omega$. If for each $y\in\Omega$ there exist constants $C=C_{xy}<\infty$ and $d=d_{xy}>2$ such that $P^n(x,y)\leq Cn^{-d/2}$ for every $n\geq 1$, then X has infinitely many cut times almost surely.*

Note that if M is irreducible then the hypothesis of this corollary is equivalent to the on-diagonal heat kernel estimate $P^n(x,x)=O(n^{-d/2})$ holding for some $d>2$; for graphs, this is (by definition) equivalent to the *spectral dimension* of the graph being strictly larger than 2. As such, Corollary 35 is already sufficient to treat most natural examples of transient graphs arising in examples. For instance, graphs satisfying an isoperimetric inequality of dimension strictly greater than 2 satisfy this hypothesis [226, Corollary 3.2.10].

Proof of Corollary 35 given Theorem 34. Fix $x,y\in\Omega$ and suppose that $C<\infty$ and $d>2$ are such that $P^n(x,y)\leq Cn^{-d/2}$ for every $n\geq 1$. We have by the Markov property that

$$\begin{aligned}\mathbb{E}_x\left[\mathbb{G}(X_n,y)\right] &= \mathbb{E}_x\#\{\text{visits to } y \text{ after time } n\}\\ &= \sum_{m=n}^{\infty}P^m(x,y)\leq C\sum_{m=n}^{\infty}m^{-d/2}\leq\frac{2C}{d-2}n^{-(d-2)/2}\end{aligned}$$

for every $n\geq 1$, and hence by Borel-Cantelli that

$$\mathbb{G}(X_{2^k},y)\leq k^2 2^{-(d-2)k/2}\qquad\text{for all sufficiently large } k \text{ almost surely.}$$

If τ denotes the first time after time 2^k that X hits y then the stopped process $(\mathbb{G}(X_m,y))_{m=2^k}^{\tau}$ is a non-negative martingale, and it follows by the optional stopping theorem that

$$\mathbb{P}_x\left(\exists m \geq 2^k \text{ such that } \mathbb{G}(X_m,y) \geq k^4 2^{-(d-2)k/2} \mid \mathbb{G}(X_{2^k},y) \leq k^2 2^{-(d-2)k/2}\right) \leq \frac{1}{k^2}$$

for all sufficiently large k. Thus, a further application of Borel-Cantelli yields that

$$\mathbb{G}(X_n,y) \leq (\log_2 n)^4 \left(\frac{n}{2}\right)^{-(d-2)/2} \tag{5.2}$$

for all sufficiently large n almost surely. Since y was arbitrary, the hypotheses of Theorem 34 are satisfied and X has infinitely many cut times almost surely. □

As mentioned above, earlier results concerning random walks on groups relied on relatively fine control of the Green's function and its gradient, which was used to prove the existence of infinitely many cut times via a second moment argument. The far weaker and more distributed nature of our decay hypothesis causes this second moment argument to break down. Instead, we compare expectations and conditional expectations of certain special types of cut times as the process $(\mathbb{G}(X_n,X_0))_{n\geq 0}$ crosses a small exponential scale $[e^{-k-1},e^{-k}]$. Roughly speaking, this allows us to integrate all of the available information across time, compensating for the looser information. See Section 5.2 for details.

Superdiffusive random walks have infinitely many cut times. As an application of Theorem 34, we also prove that walks on graphs and networks (i.e. reversible Markov chains) satisfying a weak superdiffusivity condition have infinitely many cut times almost surely. Given a network $N=(V,E,c)$ with underlying graph (V,E) and conductances $c : E \to (0,\infty)$, we define the conductance $c(v)$ of a vertex v to be the total conductance of all edges emanating from v.

Theorem 36. *Let $N=(V,E,c)$ be a locally finite, connected network with $\inf_v c(v) > 0$ and let X be a random walk on N. If there exists $r > 3/2$ such that*

$$\liminf_{n\to\infty} \frac{d(X_0,X_n)}{n^{1/2}(\log n)^r} > 0 \qquad \textit{almost surely}, \tag{5.3}$$

then X has infinitely many cut times almost surely.

This result resolves a conjecture of Benjamini, Gurel-Gurevich, and Schramm [64], who asked whether random walks on graphs with positive linear liminf speed have infinitely many cut times almost surely. For *bounded degree* graphs where the walk has positive speed, our proof yields that the walk has a *positive density* of cut times a.s., yielding a very strong version of their conjecture.

Theorem 37. *Let G be a bounded degree graph and let X be a random walk on G.*

$$\textit{If} \quad \liminf_{n\to\infty} \frac{1}{n} d(X_0, X_n) > 0 \quad \textit{a.s.} \quad \textit{then}$$

$$\liminf_{n\to\infty} \frac{1}{n} \#\{0 \le m \le n : m \textit{ is a cut time for } X\} > 0 \quad \textit{a.s.}$$

Note that Theorem 36 is not an immediate consequence of Theorem 34, as we are not aware of any general result allowing us to deduce Green's function decay estimates from distance estimates without further assumptions on the graph: the Varopoulos-Carne inequality [99, 317] tells us that $p_m(X_n, X_0)$ is small when $d(X_0, X_n)$ is much larger than $m^{1/2}$, but does not give any control whatsoever of the large-time contribution to the Green's function $\sum_{m \ge n^2} p_m(X_n, X_0)$. To circumvent this obstacle, we consider adding a spatially-dependent killing to our network. We tune the rate of killing to be weak enough that the walk has a positive chance to live forever when superdiffusive, and strong enough that we can control the decay of the killed Green's function along the walk. We prove that this killed walk has infinitely many cut times almost surely on the event that it survives forever, from which Theorem 36 easily follows.

The Diaconis-Freedman conjecture. Let $M = (\Omega, P)$ be a transient Markov chain, and let $X = (X_n)_{n \ge 0}$ be a trajectory of M. The **partially exchangeable σ-algebra** of X is defined to be the exchangeable σ-algebra generated by the sequence of increments $((X_n, X_{n+1}))_{n \ge 0}$, that is, the set of events that are determined by the sequence of increments and that are invariant under permutations of this sequence that fix all but finitely many terms. This σ-algebra arises naturally in the work of Diaconis and Freedman [126], who proved that every partially exchangeable sequence of random variables can be expressed as a Markov process in a random environment. This can be thought of as a partially-exchangeable version of de Finetti's theorem and plays an important role in the theory of reinforced random walks [26, 270]. Their study of the partially exchangeable σ-algebra led Diaconis and Freedman to make the following conjecture. Given a trajectory $X = (X_n)_{n \ge 0}$, we define the **crossing number** of an ordered pair of states (x, y) to be the number of integers n such that $(X_n, X_{n+1}) = (x, y)$.

Conjecture 38 (Diaconis-Freedman 1980)**.** *Let X be a trajectory of a transient Markov chain. Then the partially exchangeable σ-algebra of X is generated by the crossing numbers of X.*

An equivalent statement of this conjecture is that if we condition on the crossing numbers then the resulting process has trivial exchangeable σ-algebra almost surely. Note that there is a close analogy between this conjecture and the problem of computing the Poisson boundary for lamplighter random walks [214, 260]. As observed in [202], it is easily seen that the Diaconis-Freedman conjecture holds whenever X has infinitely many cut times almost surely. As such, our main results imply that the Diaconis-Freedman conjecture holds for most transient Markov chains arising in examples.

Corollary 39. *Let $M = (\Omega, P)$ be an irreducible transient Markov chain with trajectory $(X_n)_{n\geq 0}$. If (M,x) satisfies the hypotheses of either Theorem 34 or Corollary 35 then the partially exchangeable σ-algebra of X is generated by its crossing numbers.*

Organisation. Section 5.2 contains the proof of our main theorem, Theorem 34. First, in Section 5.2.1, we describe the overarching strategy behind the proof of Theorem 34 and give a proof in the much simpler special case in which the Green's function decays exponentially along the random walk. We then introduce relevant technical preliminaries in Sections 5.2.2 and 5.2.3 before proving Theorem 34 in Section 5.2.4. Finally, we prove our results concerning superdiffusive walks in Section 5.3 and prove that Theorem 34 is sharp for birth-death chains in Section 5.4.

Notation. Given a sequence of real numbers $(z_n)_{n\geq 0}$, we will often write $(z_n^*)_{n\geq 0}$ for the associated sequence of running minima $z_n^* = \min_{0\leq m\leq n} z_m$.

5.2 Proof of the main theorem

In this section we prove our main theorem, Theorem 34. We will work mostly under the additional assumption that M is irreducible, locally finite (i.e. that there are finitely many possible transitions from each state), and has $P(x,x) = 0$ except possibly for one absorbing state $\dagger$, before showing that the general case follows from this case at the end of the proof. It will be convenient to work throughout with the hitting probabilities

$$\mathbb{H}(x,y) = \mathbb{P}_x(\text{hit } y) = \frac{\mathbb{G}(x,y)}{\mathbb{G}(y,y)}$$

rather than the Green's function. This can be done with minimal changes to each of the other statements since $\mathbb{H}(X_n, y)$ decays at the same rate as $\mathbb{G}(X_n, y)$ for each fixed y.

Let us now give some relevant definitions. We define a **Markov chain with killing** to be a tuple $M = (\Omega, P, \dagger)$ where Ω is a countable state space, $P : \Omega \times \Omega \to [0,1]$ is the transition kernel and $\dagger \in \Omega$ is a distinguished **graveyard** state satisfying $p(\dagger,\dagger) = 1$. We say that a Markov chain with killing is **locally finite** if the set $\{v : p(u,v) > 0\}$ is finite for every $u \in \Omega$

and say that a Markov chain with killing is **irreducible** if for every $u,v \in \Omega \setminus \{\dagger\}$ there exists $n \in \mathbb{N}$ such that $P^n(u,v) > 0$. We say the chain is **transient** if every state other than † is visited at most finitely many times almost surely. Given a trajectory X of a Markov chain with killing, we define for each $x \in \Omega$ the hitting time $\tau_x = \inf\{n \geq 0 : X_n = x\}$, and say that a trajectory of the chain is *killed* if $\tau_\dagger < \infty$.

Theorem 40. *Let $M = (\Omega, P, \dagger)$ be a transient, locally finite, irreducible Markov chain with killing such that $P(x,x) = 0$ for every $x \neq \dagger$, let $X = (X_n)_{n\geq 0}$ be a trajectory of M, and let $\phi : [0,\infty) \to [0,\infty)$ be an increasing bijection such that*

$$\sum_{n=1}^{\infty} \frac{1}{1 \vee \log(\phi^{-1}(n))} = \infty. \tag{5.4}$$

If the event $\mathscr{G} = \{\limsup_{n\to\infty} e^{\phi(n)} \mathbb{H}(X_n, X_0) < \infty\}$ has positive probability, then conditional on $\mathscr{G}$, X is either killed or has infinitely many cut times almost surely.

Note that (5.4) becomes equivalent to (5.1) when $\Phi(x) = e^{-\phi(x)}$ as established in the following lemma; we found the condition in terms of Φ given in Theorem 34 to be easier to think about in examples, while the condition in terms of ϕ given in Theorem 40 is better suited to the proof.

Lemma 41. *Let $\Phi : [0,\infty) \to (0,1]$ be a decreasing bijection and let $\phi = -\log \Phi$. Then*

$$\int_0^1 \frac{1}{u(1 \vee \log \Phi^{-1}(u))} \mathrm{d}u = \infty \qquad \textit{if and only if} \qquad \sum_{n=1}^{\infty} \frac{1}{(1 \vee \log \phi^{-1}(n))} = \infty.$$

Proof of Lemma 41. We will prove that if the integral involving Φ diverges then the sum involving ϕ diverges, this being the only direction of the lemma that we need. The reverse direction is proved similarly. Since Φ is decreasing, we have that

$$\begin{aligned}\int_0^1 \frac{1}{u(1 \vee \log \Phi^{-1}(u))} \mathrm{d}u &= \sum_{k=1}^{\infty} \int_{e^{-k}}^{e^{-k+1}} \frac{1}{u \log(1 \vee \Phi^{-1}(u))} \mathrm{d}u \\ &\leq \sum_{k=1}^{\infty} \frac{e^{-k+1}}{e^{-k}(1 \vee \log \Phi^{-1}(e^{-k+1}))} = \sum_{k=1}^{\infty} \frac{e}{(1 \vee \log \Phi^{-1}(e^{-k+1}))},\end{aligned} \tag{5.5}$$

and the claim follows since $\Phi^{-1}(e^{-k+1}) = \phi^{-1}(k-1)$. □

5.2.1 The overarching strategy and the special case of exponential decay

In this section we describe the high-level strategy underlying Theorem 34 and present a proof in the much simpler case of an exponentially decaying hitting probability process

$\mathbb{H}(X_n)$. We then document the issues that arise when attempting to extend this method to the subexponential case and outline how we overcome them.

The high-level idea is to construct a function $F : \Omega \to [0,\infty)$ such that there are infinitely many times n when the trajectory (X_m) of the irreducible Markov chain $M = (\Omega, P)$ satisfies

$$\max_{m \geq n} F(X_m) < \min_{m<n} F(X_m).$$

In other words, at each of these times n, the process $F(X_i)$ must *drop* lower than it has previously, and this drop must be a *permadrop*, i.e. $F(X_i)$ must not *recover* to any level achieved prior to the drop. Indeed, if this condition holds then the walk cannot return to any vertex it has previously visited and therefore has a cut time at n. For the 'drop' part of this condition to hold infinitely often, it is sufficient that the process $(F(X_n))_{n\geq 0}$ converges to zero, and given transience of the Markov chain, a candidate such as $F(x) = d(o,x)^{-1}$ would suffice. Indeed, studying graph distances appears to be a particularly natural choice in the superdiffusive regime. Unfortunately, there seem to be very limited tools available to prove that this function yields infinitely many permadrops, even when the random walk has positive speed.

These considerations make it natural to instead study the decay of hitting probabilities along the random walk: when the chain is transient the **hitting probability process** $(\mathbb{H}(X_n,X_0))_{n\geq 0}$ automatically tends to zero, and we can use the fact that the process is a martingale to attempt to analyse the number of permadrops. Indeed, Benjamini, Gurel-Gurevich and Schramm [64] used martingale techniques to show that the *expected* number of permadrops of this process is always infinite when the chain is transient and hence that every transient chain has infinitely many cut times in expectation. Thus, a natural approach to the cut times problem is to find sufficient conditions for the number of permadrops of this process to be infinite almost surely.

Let us first consider the special case in which M is irreducible and $\mathbb{H}(X_n,X_0)$ decays exponentially. Note that this case is already sufficient to resolve the conjecture of Benjamini, Gurel-Gurevich, and Schramm [64] in conjunction with the spatially-dependent-killing argument of Section 5.3.

Proposition 42. *If M is irreducible and the hitting probability process $Z_n := \mathbb{H}(X_n,X_0)$ decays exponentially in the sense that $\liminf_{n\to\infty} \frac{1}{n} \log 1/Z_n > 0$ a.s. then X has infinitely many cut times a.s.*

The proof of this proposition will rely on Lévy's zero-one law [248], which is a special case of the martingale convergence theorem.

Lemma 43 (Lévy's zero-one law). *Let $(\Omega, F, \mathbb{P})$ be a probability space, and let $\mathbb{E}$ denote expectation with respect to $\mathbb{P}$. Let $(F_n)_{n\geq 0}$ be a filtration and let A be an $F_\infty = \sigma(\cup_n F_n)$ measurable event. Then*

$$\lim_{k\to\infty} \mathbb{P}[A \mid F_k] = \mathbb{1}_A \qquad \textit{almost surely.}$$

For application later in the paper, we will deduce Proposition 42 from the following lemma.

Lemma 44. *Let $\alpha < 1$ and let A be the event that $Z_n^* \leq \alpha Z_{n-1}^*$ infinitely often. If A has positive probability then X has infinitely many cut times a.s. on the event A.*

Proof. Define the sequence of times $(t_n)_{n\geq 0}$ recursively by setting $t_0 = 0$ and for each $n \geq 1$ setting

$$t_n = \inf\{m > t_{n-1} : Z_m \leq \alpha Z_{m-1}^*\}$$

with the convention that $\inf \emptyset = \infty$. We observe that the event A is equal to $\{t_n < \infty \;\forall n \geq 0\}$ and for each $n \geq 1$ consider the *permadrop* event $\mathscr{A}_n = \{t_n < \infty$ and $Z_m < \alpha^{-1} Z_{t_n}$ for every $m \geq t_n\}$, so that $Z_m < Z_{t_n-1}^*$ for every $m \geq t_n$ on the event A_n. Let $\mathscr{A}$ be the event that infinitely many of the events $\mathscr{A}_n$ hold, so that $\mathscr{A} \subseteq A$ and X has infinitely many cut times whenever $\mathscr{A}$ holds. Since the filtration $(\mathscr{F}_{t_n})$ has the σ-algebra generated by the entire random walk as its union, Lévy's zero-one law implies that

$$\lim_{n\to\infty} \mathbb{P}(t_n < \infty \text{ and } \exists m \geq n \text{ s.t. } A_m \text{ occurs} \mid \mathscr{F}_{t_n}) = \mathbb{1}(\mathscr{A})$$

almost surely. On the other hand, since Z is a supermartingale, we have by optional stopping that

$$\mathbb{P}(t_n < \infty \text{ and } \exists m \geq n \text{ s.t. } \mathscr{A}_m \text{ occurs} \mid \mathscr{F}_{t_n}) \geq \mathbb{P}(\mathscr{A}_n \mid \mathscr{F}_{t_n})\mathbb{1}(t_n < \infty) \geq (1-\alpha)\mathbb{1}(t_n < \infty)$$

almost surely for each $n \geq 1$. Since the latter estimate is bounded away from zero as $n \to \infty$ on the event A, we deduce that $\mathscr{A}$ holds almost surely conditional on A and hence that X has infinitely many cut times almost surely conditional on A. $\square$

Proof of Proposition 42. For each $\alpha \in (0,1)$, let E_α be the event that $Z_n = Z_n^* \leq \alpha Z_{n-1}^*$ for infinitely many $n \geq 1$. Since M is irreducible, Z_n is positive for every $n \geq 0$. Since Z also decays exponentially almost surely, the sequence $Z_n^* = \inf_{m\leq n} Z_m$ is also positive and decays exponentially almost surely. In particular, there exists a $[0,1]$-valued random variable α

satisfying $\alpha < 1$ almost surely such that

$$Z_n = Z_n^* \leq \alpha Z_{n-1}^* \tag{5.6}$$

for infinitely many $n \geq 1$. This implies that the countable union $\bigcup_{k\geq 1} E_{(k-1)/k}$ has probability 1 and so the result follows from Lemma 44. □

Problems in the subexponential case. As we have just seen, it is straightforward to show that Z_n has infinitely many permadrops whenever it decays exponentially: the large decay rate guarantees an infinite supply of drops of a constant relative size, and the optional stopping theorem bounds the probability of each of these drops being a permadrop below by a constant. This constant lower bound means we can rely on soft techniques like Lévy's zero-one law to deduce that permadrop events occur infinitely often without having to worry about their dependencies. However, even if we did have to think about dependencies, we could choose the drops far enough away from each other such that we could easily control the correlations between their recovery events. (The exact argument is somewhat subtle: it is not necessarily true that the correlations are small, but the conditional probability of there being a permadrop on one scale given what has happened on previous scales is bounded away from 0.)

When we move to the subexponential case, this argument quickly begins to break down. Indeed, the best we were able to do by optimizing the above approach was to handle the case of stretched-exponential decay $Z_n = e^{-\Theta(n^z)}$ for $z > 1/2$. Let us now overview the problems that arise when attempting to perform such an optimization. First, without access to Lévy's zero-one law, we now have to consider correlations between recovery events. Perhaps more significantly, however, subexponential decay gives us only very loose information about the local behaviour of the hitting probability process. We know the extent to which it must decrease over long periods of time, but have relatively little structural information about how this decrease occurs or about the positions and sizes of the drops: the overall fall in value of the process could be made up of frequent small drops, rare large drops, or any combination thereof. Consider for instance the case of stretched exponential decay $Z_n = e^{-\Theta(n^z)}$ for $z \in (0,1)$. This decay could be achieved by drops by a factor of size $1 - n^{z-1}$ at a positive density of times, or, say, by halving at each time of the form $n^{1/z}$. The only restriction is that we cannot have too much of the decay made up of very small drops, as this would contradict the assumed decay of the process.

In an attempt to adapt the arguments used in the exponential decay case, a natural starting place would be to attempt to extract a sparse sequence of roughly independent drops of guaranteed size. For instance, in the stretched exponential decay case $Z_n = e^{-\Theta(n^z)}$, we can set up the infinite sequence of stopping times

$$t_n = \inf\left\{m > t_{n-1} : Z_m < a_n Z_{t_{n-1}} \text{ and } Z_m < (1-n^{z'-1})Z_m^*\right\}$$

for some decreasing sequence (a_n) very slowly converging to 0, and $z' \in (0,z)$, where the sequence a_n should be chosen to allow us to safely ignore dependencies between successive steps. It turns out that this works well for $z > 1/2$: a deterministic argument proves that if (Z_m) has only finitely many drops of any constant relative size, then for n large enough, the drop at time t_n must approximately have size at least $1 - n^{(z'-1)/z}$, and optional stopping allows us to control the dependencies between the recovery events. Optional stopping then gives a $n^{(z'-1)/z}$ probability of the drop at time t_n being a permadrop, and a simple generalisation of Borel-Cantelli then implies that there are infinitely many permadrops almost surely. For $z \leq \frac{1}{2}$, however, the sequence $n^{(z'-1)/z}$ has a convergent sum and the argument breaks down. At this stage we are very far from handling polynomial decay!

Addressing the problems. To get results when the process decays slower than $e^{-n^{1/2}}$, we can no longer just extract sparse sequences and must begin to consider neighbouring drops and the interactions between their recovery events. We attempted to employ a second moment method, bounding each $\mathbb{P}(A_i \cap A_j)$ from above where A_i is the event that the ith drop is a permadrop. Unfortunately, due to the looseness of the information that we have regarding the locations and sizes of the drops, this method proved difficult to implement and did not seem capable of producing optimal results. To overcome the outlined issues, we instead analyse the path of the hitting probability process as it traverses a series of spatial scales. At each scale we upper bound the expected number of *large* permadrops *conditional on there being at least one*, and simultaneously *lower bound* the *unconditional* expected number of large permadrops. We modulate the definition of “large" across scales to ensure that the former quantity is not too large and the latter is not too small: we need the threshold for the drop sizes we consider to be small enough that we get an adequate supply of drops to lower bound the unconditional expectation while being large enough to prevent an accumulation of drops amplifying the conditional expectation. Once we have done this with a well-chosen choice of thresholding function, comparing these two quantities allows us to lower bound the probability that there is a permadrop on each scale; considering a whole scale simultaneously, rather than individual pairs of drops, allowed us to tackle the flexibility present in the structure of the decay. The Borel-Cantelli counterpart then has a natural application demonstrating that there are infinitely many permadrops when the decay of the hitting probability is strong enough.

Rather than working directly with the hitting probability process of the Markov chain, we work with an augmented continuous time process which we call the *drawbridge process*. This makes the hitting probability process a continuous martingale away from 1 and lets us use optional stopping to get exact expressions for permadrop probabilities rather than one-sided inequalities: this is important since we need to prove both upper and lower bounds on relevant expectations. As mentioned above, we will work primarily in the setting of locally finite Markov chains that are irreducible bar the presence of a graveyard state, before deducing a result for general Markov chains via a simple reduction argument.

5.2.2 The drawbridge process

At several points in our analysis we will want to apply the optional stopping theorem to get equalities rather than one-sided inequalities, making it convenient to work with continuous rather than discrete martingales. For the random walk on a graph, it is well-known that one can embed the discrete-time random walk inside a continuous-time continuous process by considering Brownian motion on an appropriately constructed metric graph known as the *cable graph* [142, 256]. We now construct a similar way of embedding a *non-reversible* locally finite Markov chain inside a continuous Markov process, which we call the *drawbridge process*, and hence of embedding the discrete-time hitting probability process inside a continuous martingale. While there are precedents for considering similar processes [246], it appears to be much less well known than the cable process, and we give a fairly detailed introduction to keep the paper self-contained.

Before giving a precise definition let us first give the intuition behind the name. Let $M = (\Omega, P, \dagger)$ be a locally finite Markov chain with killing and suppose that $P(x,x) = 0$ for every $x \neq \dagger$. Consider the corresponding directed graph G with vertex set Ω and with a directed edge from a vertex u to a vertex v if $u \neq v$ and $P(u,v) > 0$. We can make this abstract graph physical, in some sense, by assigning the positive real length $1/P(u,v)$ to each directed edge (u,v). While it is nonsensical to think of a Brownian motion which can only travel in one direction, we can recover restrictions in motion through the use of “drawbridges". More specifically, we envision Brownian motion on a modified version of the metric graph, in which one places a “drawbridge" along each directed edge of the metric graph. Each drawbridge has two states, **raised** and **lowered**. When the drawbridge at (u,v) is raised, the connection between the part of the edge near v and the vertex v itself is severed, and the Brownian motion cannot cross from v onto the edge (u,v). Conversely, when (u,v) is lowered it is possible for the Brownian motion to enter the edge from either u or v. For each vertex u, we call the drawbridges across the edges emanating from u in the corresponding directed graph the **outgoing drawbridges** from u. The drawbridge process will be defined by taking the Brownian motion on this metric graph and raising and lowering drawbridges as the Brownian motion moves so that, at each time, the outgoing drawbridges from the last vertex it visited are lowered and all other drawbridges are raised.

We now make this precise. Let $M = (\Omega, P, \dagger)$ be a locally finite Markov chain with killing such that $P(x,x) = 0$ for every $x \neq \dagger$. For each state $x \in \Omega$, we define the set of outgoing states $x^{\rightarrow}$ to be $\{y \in \Omega \setminus \{x\} : P(x,y) > 0\}$. We define the *star graph* $S[x]$ to be the metric graph with vertex set $\{x\} \cup x^{\rightarrow}$, with edge set $\{\{x,y\} : y \in x^{\rightarrow}\}$, and with edge lengths $1/P(x,y)$, so that $S[\dagger]$ is the metric graph consisting of the single vertex $\{\dagger\}$ and no edges. In an abuse of notation, we will identify vertices in the star graph with their corresponding states; the precise meaning will be clear from context. We construct the metric space $\mathscr{S}$ from the disjoint union $\mathscr{S}^{\sqcup} = \sqcup_{x \in \Omega} S[x] = \{(x,y) : x \in \Omega, y \in S[x]\}$ by gluing together $(x,y) \in \{x\} \times S[x]$ and $(y,y) \in \{y\} \times S[y]$ for every $x \in \Omega$ and $y \in x^{\rightarrow}$. Note that every point in $\mathscr{S}$ has a unique representation of the form (x,y) where $x \in \Omega$ and $y \in S[x] \setminus x^{\rightarrow}$.

Let $(x_0, y) \in \mathscr{S}$ be such that $x_0 \in \Omega$ and $y \in S[x_0] \setminus x_0^{\rightarrow}$. We construct the drawbridge process on $\mathscr{S}$ starting at (x_0, y) as follows. First we start a Brownian motion $(B_t^0)_{t \geq 0}$ on $\mathscr{S}[x_0]$ starting at (x_0, y) at time $\mathscr{T}(0) = 0$ and run until the stopping time $\mathscr{T}(1) = \inf\{t > \mathscr{T}(0) : B_t^0 \in \{x_0\} \times x_0^{\rightarrow}\}$, so that if $\mathscr{T}(1) < \infty$ then $B^0_{\mathscr{T}(1)} = (x_0, x_1)$ for some $x_1 \in x_0^{\rightarrow}$. If $\mathscr{T}(1)$ is finite, we then run a Brownian motion $(B_t^1)_{t=\mathscr{T}(1)}^{\mathscr{T}(2)}$ on $\mathscr{S}[x_1]$, started from (x_1, x_1) at time $\mathscr{T}(1)$ and run until the stopping time $\mathscr{T}(2) = \inf\{t > \mathscr{T}(1) : B_t^1 \in \{x_1\} \times x_1^{\rightarrow}\}$. We iterate this construction to generate a possibly infinite sequence $((B_t^i)_{\mathscr{T}(i) \leq t \leq \mathscr{T}(i+1)})_i$, noting that $\mathscr{T}(i)$ is almost surely finite whenever $x_i \neq \dagger$. If the sequence terminates because $\mathscr{T}(i) = \infty$ for some $i \in N$, which almost surely happens exactly when the process first visits the graveyard state $\dagger$, then we define $\mathscr{T}(j) = \infty$ for $j > i$, set $\tau_\dagger = i - 1$ and $\mathscr{T}_\dagger = \mathscr{T}(\tau_\dagger)$. If the sequence does not terminate, we set $\tau_\dagger = \mathscr{T}_\dagger = \infty$. Finally, we construct the *drawbridge process* $(\mathscr{X}_t)_{t \geq 0}$ by concatenating, in order, the images of the paths of the Brownian motions $(B^i)_{0 \leq i \leq \tau_\dagger}$ in $\mathscr{S}$ under the gluing map $\mathscr{S}^{\sqcup} \to \mathscr{S}$.

We let $\mathbf{P}_{x,y}$, $\mathbf{E}_{x,y}$ denote probability and expectation with respect to the law of $\mathscr{X}$ started at (x,y) and write $\mathbf{P}_x = \mathbf{P}_{x,x}$ and $\mathbf{E}_x = \mathbf{E}_{x,x}$. Observe that if $\mathscr{X}$ is started at (x,x) for some $x \in \Omega$ then the discrete-time process $X = (X_n)_{n=0}^{\tau_\dagger}$ defined by $(\mathscr{X}_{\mathscr{T}(n)})_{n=0}^{\tau_\dagger} = (X_n, X_n)_{n=0}^{\tau_\dagger}$ has the distribution of a trajectory of the Markov chain stopped when it hits the graveyard state $\dagger$. Thus, if we fix an arbitrary 'origin' state $o \neq \dagger$ and let $\mathscr{T}_o = \inf\{t \geq 0 : \mathscr{X}_t = (o,o)\}$ be the hitting time of o, setting $\mathscr{T}_o = \infty$ if o is never hit, then the hitting probability $\mathscr{H}(x,y) = \mathbf{P}_{x,y}(\mathscr{T}_o < \infty)$ satisfies $\mathscr{H}(x,x) = \mathbb{H}(x) = \mathbb{H}(x,o)$ for all $x \in \Omega$. We define $(\mathscr{Z}_t)_{t \geq 0} = (\mathscr{H}(\mathscr{X}_t))_{t \geq 0}$ and $(Z_n)_{n \geq 0} = (\mathbb{H}(X_n))_{n \geq 0} = (\mathscr{Z}_{\mathscr{T}(n)})_{n \geq 0}$, noting that if $\sigma_1 \leq \sigma_2$ are stopping times for $\mathscr{X}$ such that $\mathscr{X}_t \neq o$ almost surely for every $\sigma_1 < t < \sigma_2$ then $(\mathscr{Z}_t)_{t=\sigma_1}^{\sigma_2}$ is a continuous time martingale with respect to its natural filtration.

5.2.3 Deterministic preliminaries

We now set up the notation to record the behaviour of the running minima of a non-negative sequence as it converges to zero across a series of exponential scales. Recall that for each sequence $(z_n)_{n\geq 1}$ we write $(z_n^*)_{n\geq 0} = (\min_{m\leq n} z_m)_{n\geq 0}$ for the associated sequence of running minima. Given a non-negative sequence $(z_n)_{n\geq 0}$ converging to 0 with $z_0 > 0$, we construct a sequence of logarithmic scales over which to analyse and control its behaviour. We record the associated notation in the following definition.

Definition 2 (Notation for drops at scale k)**.** *Fix a non-negative sequence $z = (z_n)_{n\geq 0}$ with $z_n > 0$ and $z_n \to 0$ as $n \to \infty$. Let $k_0 = k_0(z) = \lceil 2\log z_0^{-1} \rceil$, and for each $k \geq 1$ define the kth scale interval $I_k = [e^{-k-1}, e^{-k}]$. For each $k \geq k_0$ we define the set D_k by adjoining the set of running minima on the kth scale to the endpoints of the corresponding interval I_k so that*

$$D_k = D_k(z) = (\{z_m^* : m \geq 1\} \cap I_k) \cup \{e^{-k}, e^{-k-1}\}$$

for each $k \geq k_0$. We define $N_k = N_k(z) = |D_k(z)| - 1$ and label the elements of $D_k(z)$ in decreasing order as $(d_{i,k})_{1\leq i\leq N_k} = (d_{i,k}(z))_{0\leq i\leq N_k}$ so that

$$e^{-k} = d_{0,k} > d_{1,k} > \cdots > d_{N_k,k} = e^{-k-1} \qquad \text{and} \qquad \prod_{i=0}^{N_k-1} \frac{d_{i+1,k}}{d_i, k} = e^{-1}.$$

We call the pairs $(d_{i,k}, d_{i+1,k})$ for $0 \leq i \leq N_k - 1$ the ***drops*** *of z on scale k. When context makes clear which sequence z we are referring to, we will drop it from our notation. Similarly, when it is clear that we are referring to a particular scale k we will drop the second subscript on the $d_{i,k}$ by writing $d_i = d_{i,k}$.*

We begin the proof by proving the following deterministic lemma. Roughly speaking, this lemma says that for sequences which decay to zero sufficiently quickly, we can define a threshold between large and small drops in such a way that the following hold:

1. For a good proportion of scales, a good proportion of the decay is made up of large drops.
2. The large drops are large enough that there cannot be too many such drops at any particular scale.

The actual threshold we will use will be a simple function of the ψ which is outputted by this lemma. Later, we will apply this lemma to the hitting probability process. Given an increasing bijection $\psi : [0,\infty) \to [0,\infty)$ and a sequence of non-negative numbers $(z_n)_{n\geq 0}$, we say that $(z_n)_{n\geq 0}$ is **ψ-good on scale** k if

$$\prod\left\{\frac{d_{i+1,k}}{d_{i,k}}:0\leq i<N_k \text{ is such that } \frac{d_{i+1,k}}{d_{i,k}}\leq \exp\left[-\frac{k}{2\psi^{-1}(k)}\right]\right\} \\ \leq \prod\left\{\frac{d_{i+1,k}}{d_{i,k}}:0\leq i<N_k \text{ is such that } \frac{d_{i+1,k}}{d_{i,k}}> \exp\left[-\frac{k}{2\psi^{-1}(k)}\right]\right\},$$

or equivalently if

$$\prod\left\{\frac{d_{i+1,k}}{d_{i,k}}:0\leq i<N_k \text{ is such that } \frac{d_{i+1,k}}{d_{i,k}}\leq \exp\left[-\frac{k}{2\psi^{-1}(k)}\right]\right\}\leq e^{-1/2},$$

where we write $\prod\{A_i : i\in I\} = \prod_{i\in I} A_i$ for reasons of legibility. That is, $(z_n)_{n\geq 0}$ is ψ-good on scale k if at least half of the total decay across the scale comes from drops of size at least $\Psi(k) := e^{-k/2\psi^{-1}(k)}$ in a geometric sense. Note that ψ^{-1} denotes the inverse of ψ; we will typically think of ψ as being a slowly growing function so that ψ^{-1} satisfies $\psi^{-1}(x)\gg x$.

Lemma 45 (Good, well-separated scales)**.** *Let $\phi : [0,\infty)\to[0,\infty)$ be an increasing bijection such that*

$$\sum_{k=1}^{\infty}\frac{1}{1\vee\log\left(\phi^{-1}(k)\right)}=\infty. \tag{5.7}$$

Then there exist an increasing bijection $\psi : [0,\infty)\to[0,\infty)$ satisfying $\psi(x)\leq\phi(x)$ and $\psi(x)\leq\sqrt{x}$ for every $x\geq 0$, and a strictly increasing function $a:\mathbb{N}\to\mathbb{N}$ satisfying

$$\lim_{n\to\infty} a(k)-a(k-1)=+\infty \qquad \text{and} \qquad \sum_{k=1}^{\infty}\frac{1}{1\vee\log(\psi^{-1}(a(k)))}=\infty, \tag{5.8}$$

such that if $(z_n)_{n\geq 0}$ is a sequence of positive reals with $z_0>0$ satisfying $\limsup_{n\to\infty} e^{\phi(n)}z_n<\infty$, then

$$\liminf_{k\to\infty}\frac{1}{k}\#\{1\leq r\leq k : z \text{ is } \psi\text{-good on scale } a(r)\}\geq\frac{1}{2}. \tag{5.9}$$

As mentioned above, the function ψ, which we think of as “ϕ with some room”, is used to define the threshold for a drop on scale k to be “large”, with z being ψ-good on scale k precisely when a good proportion of the total decay on this scale comes from drops that are larger than this threshold. Meanwhile, the sequence a is used to take a sparse sequence of spatial scales so that we can safely ignore dependencies between scales while keeping various series divergent so that we can still hope to conclude via Borel-Cantelli.

The proof of Lemma 45 will rely on the following elementary analytic facts.

Lemma 46. *Suppose that* $f : \mathbb{N} \to [0,\infty)$ *is a decreasing function satisfying* $\sum_{n=1}^{\infty} f(n) = \infty$.

1. *If* $A \subseteq \mathbb{N}$ *has positive density in the sense that* $\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}(n \in A) > 0$ *then*

$$\sum_{n\in A} f(n) = \infty \qquad \textit{and} \qquad \liminf_{N\to\infty} \frac{\sum_{n=1}^{N} \mathbb{1}(n\in A) f(n)}{\sum_{n=1}^{N} f(n)} > 0.$$

2. *There exists a convex, strictly increasing function* $a : \mathbb{N} \to \mathbb{N}$ *with* $\lim_{n\to\infty} a(n) - a(n-1) = \infty$ *such that* $\sum_{n=1}^{\infty} f(a(n)) = \infty$.

Proof of Lemma 46. Fix f as in the statement of the lemma. We begin with the first statement. Let $A \subseteq \mathbb{N}$ be such that $\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}(n \in A) > 0$ and let k be such that $\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}(n \in A) > 2/k$, so that there exists ℓ_0 such that $\sum_{n=k^\ell+1}^{k^{\ell+1}} \mathbb{1}(n\in A) \geq 2k^\ell - k^\ell = k^\ell$ for every $\ell \geq \ell_0$. Letting $N \geq k^{\ell_0+1}$ and setting $\ell_1 = \lfloor \log_k N \rfloor$, we have that

$$\sum_{n=k^{\ell_0}+1}^{N} \mathbb{1}(n\in A) f(n) \geq \sum_{\ell=\ell_0}^{\ell_1-1} f(k^{\ell+1}) \sum_{n=k^\ell+1}^{k^{\ell+1}} \mathbb{1}(n\in A) \geq \sum_{\ell=\ell_0}^{\ell_1-1} k^\ell f(k^{\ell+1})$$

and that

$$\sum_{n=k^{\ell_0}+1}^{N} f(n) \leq \sum_{\ell=\ell_0}^{\ell_1} k^{\ell+1} f(k^\ell) \leq k^{\ell_0+1} f(k^{\ell_0}) + k^2 \sum_{\ell=\ell_0}^{\ell_1-1} k^\ell f(k^{\ell+1}).$$

Since f was assumed to be divergent, it follows that

$$\liminf_{N\to\infty} \frac{\sum_{n=1}^{N} \mathbb{1}(n\in A) f(n)}{\sum_{n=1}^{N} f(n)} \geq \frac{1}{k^2} > 0,$$

and hence that $\sum_{n\in A} f(n) = \infty$ as claimed.

We now prove the second statement. The first statement implies that for any $a,b \geq 1$ there exists $m = m(a,b) \geq 1$ such that $\sum_{n=1}^{m} f(a+bn) \geq 1$. This fact allows us to recursively construct a pair of integer sequences $(b_i)_{i\geq 0}$ and $(d_i)_{i\geq 0}$ by setting $b_0 = 0$ and recursively defining

$$d_i = \min\left\{m : \sum_{n=1}^{m} f(b_i + 2^i n) \geq 1\right\} \qquad \text{and} \qquad b_{i+1} = b_i + 2^i d_i \qquad \text{for each } i \geq 0,$$

and we observe that both d_i and b_i must be finite for $i \geq 0$. We must then have

$$\sum_{i=1}^{\infty} \sum_{n=1}^{d_i} f(b_i + 2^i n) = \sum_{n=1}^{\infty} f(a(n)) = \infty,$$

where $a(n)$ is the convex, strictly increasing sequence defined by

$$(a(1),a(2),\dots) = (b_1+2, b_1+2\cdot 2,\dots,b_1+2\cdot d_1, b_2+2^2, b_2+2^2\cdot 2,\dots,b_2+2^2\cdot d_2,\\ b_3+2^3, b_3+2^3\cdot 2,\dots).$$

This sequence has increasing increments tending to infinity by construction, completing the proof. $\square$

We now apply Lemma 46 to prove Lemma 45.

Proof of Lemma 45. We may assume without loss of generality that $\phi(x) \le \sqrt{x}$ for every $x \ge 0$, replacing ϕ with $\tilde{\phi} = \min\{\phi, \sqrt{x}\}$ otherwise. Indeed, since ϕ is increasing, we have that $\tilde{\phi}^{-1}(y) \le \max\{\phi^{-1}(y), y^2\}$, and we have by the Cauchy condensation test that

$$\sum_{k=1}^{\infty} \frac{1}{1 \vee \log \max\{\phi^{-1}(k), k^2\}} = \infty \quad \text{iff} \quad \sum_{k=1}^{\infty} \frac{2^k}{1 \vee \max\{\log \phi^{-1}(2^k), k \log 4\}} = \infty.$$

If there are infinitely many k such that $4^k \ge \phi^{-1}(2^k)$ then the right hand series trivially diverges, while if not then it diverges as a consequence of the Cauchy condensation test applied to $\sum_{k=1}^{\infty} \frac{1}{1 \vee \log \phi^{-1}(k)}$.

We begin by applying Lemma 46 to the function $f(k) = 1/1 \vee \log \phi^{-1}(k)$ to give a strictly increasing function $a(k) : \mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} a(k) - a(k-1) = \infty$ and

$$\sum_{k=1}^{\infty} \frac{1}{1 \vee \log \phi^{-1}(a(k))} = \infty. \tag{5.10}$$

Extend a arbitrarily to an increasing bijection $a : [0,\infty) \to [0,\infty)$ and define $\psi : [0,\infty) \to [0,\infty)$ to be the inverse of the increasing bijection

$$\psi^{-1}(x) = 8\phi^{-1}\Big(a\left(8a^{-1}(x)\right)\Big),$$

so that ψ is strictly increasing, bounded above by ϕ, and satisfies

$$\sum_{k=1}^{\infty} \frac{1}{1 \vee \log(\psi^{-1}(a(k)))} = \sum_{k=1}^{\infty} \frac{1}{1 \vee (\log 8 + \log(\phi^{-1}(a(8k))))} = \infty$$

by Lemma 46. Since ϕ and a are increasing and $\phi(x) \le \sqrt{x}$ for every $x \ge 0$ we also have that $\psi^{-1}(x) \ge 8x^2$ and $\psi(x) \le \sqrt{x}$ for every $x \ge 0$.

Let $(z_n)_{n\ge 0}$ be a sequence of positive numbers and let C be such that $z_n \le Ce^{-\phi(n)}$ for every $n \ge 1$. We will use the notation of Definition 2. Observe that if $k \ge k_0$ is such that z is *not* ψ-good on scale k then we must have that

$$\#\left\{i : \frac{d_{i+1,k}}{d_{i,k}} > \exp\left[-\frac{1}{2\psi^{-1}(k)}\right]\right\} > \frac{\log e^{-1/2}}{\log \exp(-k/(2\psi^{-1}(k))} = \frac{1}{k}\psi^{-1}(k).$$

Discounting the endpoints of the interval, it follows that

$$\#\big(\{z_m^* : m \geq 0\} \cap (e^{-k-1}, e^{-k})\big) > \frac{1}{k}\psi^{-1}(k) - 2 \geq \frac{1}{2k}\psi^{-1}(k)$$

whenever z is not ψ-good on scale k, where we used that $\psi^{-1}(k) \geq 8k^2$ in the final inequality. Let B be the set of positive integers $k \geq k_0$ such that z is not ψ-good on scale $a(k)$ and let

$$A = \left\{k \geq k_0 : \frac{1}{k}|B \cap \{k_0, \ldots, k\}| \geq \frac{1}{2}\right\}.$$

We wish to prove that A is finite, and observe that A is finite if and only if $A \cap B$ is finite. For each $k \in A \cap B$ with $k \geq 4k_0$, we have that $|B \cap \{\lfloor k/4 \rfloor, \ldots k\}| \geq k/4$ and hence, since ψ^{-1} is increasing, that

$$\begin{aligned} |\{n : z_n \geq e^{-a(k)-1}\}| &\geq \sum_{i \in B \cap \{k_0, \ldots, k\}} \left|\{z_m^* : m \geq 0\} \cap (e^{-a(i)-1}, e^{-a(i)})\right| \\ &\geq \frac{1}{2} \sum_{i \in B \cap \{k_0, \ldots, k\}} \frac{1}{k}\psi^{-1}(a(i)) \geq \frac{1}{8}\psi^{-1}(a(\lfloor k/4 \rfloor)). \end{aligned}$$

As such, for each $k \in A \cap B$ with $k \geq 4k_0 \geq 4$, there must exist $n \geq \frac{1}{8}\psi^{-1}(a(\lfloor k/4 \rfloor))$ such that $z_n \geq e^{-a(k)-1}$. On the other hand, for such n, we also have that

$$\phi(n) \geq \phi\left(\frac{1}{8}\psi^{-1}(a(\lfloor k/4 \rfloor))\right) = a(8\lfloor k/4 \rfloor) \geq a(2k),$$

and since $a(2k) - a(k) \to \infty$ and $\limsup_{n\to\infty} e^{\phi(n)} z_n < \infty$, we deduce that $A \cap B$ is finite as claimed. □

5.2.4 Proof of Theorem 40

We now apply the deterministic tools from Section 5.2.3 to prove Theorem 40. Let us fix for the remainder of this subsection a transient Markov chain with killing $M = (\Omega, P, \dagger)$ with distinguished origin vertex o such that M is irreducible, locally finite, and satisfies $P(x,x) = 0$ for every $x \neq \dagger$. Fix $X_0 \in \Omega \setminus \{\dagger\}$, let $(\mathscr{X}_t)_{t\geq 0}$ be the drawbridge process started at (X_0, X_0), let $(X_n)_{n\geq 0}$ be the associated discrete-time trajectory of the Markov process, and let $(\mathscr{Z}_t)_{t\geq 0}$

and $(Z_n)_{n\geq 0}$ be the associated continuous- and discrete-time hitting probability processes as defined in Subsection 5.2.2. We write **P** and **E** for probabilities and expectations taken with respect to the joint law of these processes.

Fix an increasing bijection $\phi : [0,\infty) \to [0,\infty)$ as in the statement of the theorem, and let ψ be as in Lemma 45. For each $k \geq 1$ and $0 \leq i \leq N_k - 1$ we say that the drop $(d_{i,k}, d_{i+1,k})$ is a **large drop** if

$$\frac{d_{i+1,k}}{d_{i,k}} \leq \Psi(k) := \exp\left[-\frac{k}{2\psi^{-1}(k)}\right]$$

and there exists n such that $Z_n^* = d_{i+1,k}$. (The latter condition can fail if $d_{i+1,k} = e^{-k-1}$.) If $(d_{i,k}, d_{i+1,k})$ is a large drop and Z first hits $d_{i+1,k}$ at some time n, we say that $(d_{i,k}, d_{i+1,k})$ is a **large permadrop** if we have additionally that

$$\mathscr{Z}_t < d_{i,k} \text{ for every } t \geq \mathscr{T}(n).$$

We say that an arbitrary pair of values (a,b) in $[e^{-k-1}, e^{-k}]$ with $a > b$ is a large drop or large permadrop if $(a,b) = (d_{i,k}, d_{i+1,k})$ for some large drop or large permadrop $(d_{i,k}, d_{i+1,k})$ as appropriate.

Given $k \geq k_0$ and $i \geq 1$, we write $\tau_i = \tau_{i,k}$ for the ith time the discrete process Z reaches a new running minimum smaller than e^{-k}, and write $\mathscr{T}_i = \mathscr{T}(\tau_{i,k})$ for the corresponding time for the continuous process $\mathscr{Z}$, noting that $\mathscr{T}_{i,k}$ is a stopping time for $\mathscr{X}$ for each $i,k \geq 1$. Note that when $1 \leq i \leq N_k$, $\tau_{i,k}$ can be defined equivalently as the first time that $Z_n \leq d_{i,k}$. Recall that if ρ is a stopping time for $\mathscr{X}$ then $\mathscr{F}_\rho$ denotes the σ-algebra generated by $(\mathscr{X}_t)_{t=0}^{\rho}$. Given such a stopping time, we lighten notation by writing $\mathbf{E}_\rho[\,\cdot\,] = \mathbf{E}[\,\cdot \mid \mathscr{F}_\rho]$ and $\mathbf{P}_\rho[\,\cdot\,] = \mathbf{P}[\,\cdot \mid \mathscr{F}_\rho]$.

The following two estimates on the distribution of the random variable

$$R_k := \#\{0 \leq i \leq N_k - 1 : (d_{i,k}, d_{i+1,k}) \text{ is a large permadrop}\}$$

lie at the heart of the paper. We will use these propositions to bound the probability $\mathbf{P}_{\mathscr{T}_{1,k}}(R_k \geq 1)$ in Proposition 48 in terms on the probability appearing in Proposition 47.

Proposition 47. *The estimate*

$$\mathbf{E}_{\mathscr{T}_{1,k}}[R_k] \geq \frac{1}{4}\mathbf{P}_{\mathscr{T}_{1,k}}\Big(Z \text{ is } \psi\text{-good on scale } k \text{ and } \exists n \geq 0 \text{ such that } e^{-k-1} < Z_n^* \leq e^{-k-3/4}\Big) \tag{5.11}$$

holds almost surely for every $k \geq 1$.

Proposition 48. *There exists a universal constant C such that the estimate*

$$\mathbf{E}_{\mathscr{T}_{1,k}}[R_k] \leq \left(2 + C\log\frac{2\psi^{-1}(k)}{k}\right)\mathbf{P}_{\mathscr{T}_{1,k}}(R_k \geq 1) \tag{5.12}$$

holds almost surely for every $k \geq 1$.

Proof of Proposition 47. To lighten notation, we drop ks from subscripts wherever possible. We can use the optional stopping theorem to compute

$$\begin{aligned}\mathbf{E}_{\mathscr{T}_1}[R] &\geq \mathbf{E}_{\mathscr{T}_1}\left[\sum_{i\geq 0}\mathbb{1}(Z_{\tau_{i+1}} > e^{-k-1})\mathbf{P}_{\mathscr{T}_{i+1}}((d_i, d_{i+1}) \text{ is a large permadrop})\right] \\ &= \mathbf{E}_{\mathscr{T}_1}\left[\sum_{i=0}^{N_k-2}\left(1-\frac{d_{i+1}}{d_i}\right)\mathbb{1}\left(\frac{d_{i+1}}{d_i} \leq \Psi\right)\right]\end{aligned} \tag{5.13}$$

and applying the inequality $1 - x \geq \log x$ yields that

$$\mathbf{E}_{\mathscr{T}_1}[R] \geq \mathbf{E}_{\mathscr{T}_1}\left[\log\prod\left\{\frac{d_i}{d_{i+1}} : 0 \leq i \leq N_k - 2, \frac{d_{i+1}}{d_i} \leq \Psi\right\}\right]. \tag{5.14}$$

When Z is ψ-good on scale k and there exists n such that $e^{-k-1} < Z_n^* \leq e^{-k-3/4}$ we have that

$$\prod\left\{\frac{d_i}{d_{i+1}} : 0 \leq i \leq N_k - 1, \frac{d_{i+1}}{d_i} \leq \Psi\right\} \geq e^{1/2} \qquad \text{and} \qquad \frac{d_{N_k-1}}{d_{N_k}} \leq e^{1/4},$$

so that the claim follows from (5.14). □

Proof of Proposition 48. We fix a scale $I_k = [e^{-k-1}, e^{-k}]$ and calculate the conditional expectation of the number of large permadrops on that scale given there is at least one. Since $k \geq 1$ is fixed throughout, we will drop it from notation when possible. By countability of the state space Ω, there are only countably many possible permadrops. We condition on the the "first large permadrop" being $(a,b) \subset I_k$ as follows. Let R' be the number of large permadrops in the scale excluding possibly the last drop, so that

$$R' = \sum_{0\leq i\leq N_k-2}\mathbb{1}((d_i, d_{i+1}) \text{ is a large permadrop}),$$

and let $R'' = (R'-1)\vee 0$ be the amount that R' exceeds 1. Given an arbitrary pair $e^{-k-1} \leq b < a \leq e^{-k}$, we say that (a,b) is the **first large permadrop** if $(a,b) = (d_i, d_{i+1})$ for some $0 \leq i \leq N_k - 1$ such that the pair (d_i, d_{i+1}) is a large permadrop and (d_j, d_{j+1}) is not a large permadrop for any $j < i$. We write τ_b for the first time Z_n hits b (letting $\tau_b = \infty$ if this never occurs), write $\mathscr{T}_b = \mathscr{T}(\tau_b)$, and seek to upper bound the conditional expectation

$$\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first large permadrop})R''\right].$$

If $b = e^{-k-1}$ then this conditional expectation is zero, so assume $b > e^{-k-1}$.

Note that if a large drop (d_i, d_{i+1}) is *not* a permadrop then there must exist a *recovery time* at which $\mathscr{Z}$ hits d_i for the first time after $\mathscr{T}_{i+1}$. Let

$$L = L(a,b) = \{(d_i, d_{i+1}) : 1 \leq i \leq N_k - 2,\, d_i \leq b, \text{ and } d_{i+1}/d_i \leq \Psi(k)\}$$

be the set of large drops on the scale k after (a,b) and possibly excluding the last drop, define the event $K = K(a,b)$ by

$$K = \{(a,b) \text{ is a drop and every previous drop in the scale recovers before time } \mathscr{T}_b\},$$

and observe that

$$\begin{aligned}
&\mathbb{1}((a,b) \text{ is the first large permadrop})R'' \\
&\quad = \mathbb{1}(K)\mathbb{1}((a,b) \text{ is a large permadrop}) \sum_{(d_i,d_{i+1})\in L} \mathbb{1}((d_i,d_{i+1}) \text{ is a large permadrop}).
\end{aligned}$$

Since $\mathscr{T}_i$ is a stopping time for each $i \geq 1$, we can use the optional stopping theorem to compute that if $b/a \leq \Psi(k)$ then

$$\begin{aligned}
&\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first large permadrop})R''\right] \\
&\quad = \mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) \sum_{(d_i,d_{i+1})\in L} \mathbb{1}((a,b) \text{ and } (d_i,d_{i+1}) \text{ are permadrops})\right] \\
&\quad = \mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) \sum_{(d_i,d_{i+1})\in L} \mathbb{1}(\mathscr{Z}_t < a \text{ for } t \in (\mathscr{T}_b, \mathscr{T}_{i+1}) \text{ and } \mathscr{Z}_t < d_i \text{ for } t > \mathscr{T}_{i+1})\right] \\
&\quad = \mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) \sum_{(d_i,d_{i+1})\in L} \big(1 - \frac{d_{i+1}}{d_i}\big)\mathbb{1}\big(\mathscr{Z}_t < a \text{ for } t \in (\mathscr{T}_b, \mathscr{T}_{i+1})\big)\right] \\
&\quad = \mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) \sum_{(d_i,d_{i+1})\in L} \big(1 - \frac{d_{i+1}}{d_i}\big)\mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b} \text{ hits } d_{i+1} \text{ before } a)\right],
\end{aligned} \tag{5.15}$$

where the third equality follows by taking conditional expectations with respect to $\mathscr{F}_{\mathscr{T}_{i+1}}$ for the term in the summation corresponding to (d_i, d_{i+1}) and then applying optional stopping.

Write $\Psi = \Psi(k)$ and let $\mathscr{L} = \mathscr{L}(b)$ be the set of all finite sets S of ordered pairs of numbers in $[e^{-k-1}, e^{-k}]$ satisfying the following conditions:

1. If $(x,y) \in S$ then $x \leq b$ and $y/x \leq \Psi$.
2. If (x,y) and (z,w) are distinct elements of S then the open intervals (y,x) and (w,z) are disjoint.

If we consider the (random) function $F : \mathscr{L} \to [0,\infty)$ defined by

$$F(S) = \sum_{(x,y)\in S} \left(1 - \frac{y}{x}\right) \mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b} \text{ hits } y \text{ before } a),$$

then we can rewrite (5.15) as

$$\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first permadrop}) R''_k\right] \leq \mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) F(L)\right], \tag{5.16}$$

and we claim that the inequality

$$F(S) \leq (1-\Psi^2) \sum_{i=1}^{\lceil -1/\log\Psi \rceil} \mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b} \text{ hits } \Psi^i b \text{ before } a), \tag{5.17}$$

is satisfied deterministically for every $S \in \mathscr{L}$.

Before proving the claimed inequality (5.17), let us first see how it implies (5.12). Substituting (5.17) into (5.16) yields that

$$\begin{aligned}
&\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first permadrop}) R''\right] \\
&\qquad \leq (1-\Psi^2)\mathbf{E}_{\mathscr{T}_b}\mathbb{1}(K) \sum_{i=1}^{\lceil -1/\log\Psi \rceil} \mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b} \text{ hits } \Psi^i b \text{ before } a) \\
&\qquad = (1-\Psi^2)\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}(K) \sum_{i=1}^{\lceil -1/\log\Psi \rceil} \frac{a-b}{a-\Psi^i b}\right] \\
&\qquad = (1-\Psi^2)\frac{a-b}{a}\mathbb{1}(K) \sum_{i=1}^{\lceil -1/\log\Psi \rceil} \frac{1}{1-(b/a)\Psi^i},
\end{aligned}$$

where we have applied optional stopping in the first equality. Since $1/(1-x\Psi^i)$ is an increasing function of x it follows that if $b \leq \Psi a$ then

$$\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first permadrop}) R''\right] \leq \frac{a-b}{b}\mathbb{1}(K)(1-\Psi^2) \sum_{i=1}^{-\lceil 1/\log\Psi \rceil} \frac{1}{1-\Psi^{i+1}}.$$

Now, we have by calculus that $1-\Psi^{i+1} \geq (i+1)(1-\Psi)$ for every $1 \leq i \leq -\lceil 1/\log \Psi \rceil$ and hence that there exists a universal constant C such that

$$\begin{aligned}\mathbf{E}_{\mathscr{T}_b}\left[\mathbb{1}((a,b) \text{ is the first permadrop})R''\right] &\leq \frac{a-b}{b}\mathbb{1}(K)\frac{1-\Psi^2}{1-\Psi}\sum_{i=1}^{-\lceil 1/\log \Psi \rceil}\frac{1}{i+1}\\ &\leq C\frac{a-b}{b}\mathbb{1}(K)\log\frac{2\psi^{-1}(k)}{k},\end{aligned}$$

where we used the definition of $\Psi = \Psi(k)$ in the final inequality. Since we also have by optional stopping that

$$\mathbf{P}_{\mathscr{T}_b}\left((a,b) \text{ is the first permadrop}\right) = \frac{a-b}{b}\mathbb{1}(K), \tag{5.18}$$

we can take expectations over $\mathscr{F}_{\mathscr{T}_b}$ conditional on $\mathscr{F}_{\mathscr{T}_1}$ to deduce that

$$\begin{aligned}&\mathbf{E}_{\mathscr{T}_1}\left[\mathbb{1}((a,b) \text{ is the first large permadrop})R_k''\right]\\ &\qquad\leq C\log\left[\frac{2\psi^{-1}(k)}{k}\right]\cdot \mathbf{P}_{\mathscr{T}_1}\left((a,b) \text{ is the first large permadrop}\right).\end{aligned} \tag{5.19}$$

We stress that this holds for *every* pair of real numbers $a > b$ in $[e^{-k-1}, e^{-k}]$, but that both sides will be equal to zero for all but countably many such pairs (since there are only countably many possible permadrops). Summing over the countable set of pairs giving a non-zero contribution yields that

$$\mathbf{E}_{\mathscr{T}_1}[R] \leq 2\mathbf{P}_{\mathscr{T}_1}(R \geq 1) + \mathbf{E}_{\mathscr{T}_1}[R''] \leq \left(2 + C\log\frac{2\psi^{-1}(k)}{k}\right)\mathbf{P}_{\mathscr{T}_1}(R \geq 1)$$

as claimed.

It remains to prove (5.17). Given a set $S \in \mathscr{L}$, we say S is **slack** if there exists an element $(x,y) \in S$ such that $y/x < \Psi^2$ and **taut** otherwise. Observe that if $S \in \mathscr{L}$ is slack and $(x,y) \in S$ satisfies $y/x < \Psi^2$ then the set $S' = S \cup \{(x,\Psi x),(\Psi x, y)\} \setminus \{(x,y)\}$ also belongs to $\mathscr{L}$ and satisfies $F(S) \leq F(S')$. Indeed, the latter inequality follows from the pointwise inequality

$$\begin{aligned}&\mathbb{1}((\mathscr{Z}_t)_{t > \mathscr{T}_b} \text{ hits } y_i \text{ before } a)\left(1 - \frac{y_i}{x_i}\right)\\ &\leq \mathbb{1}((\mathscr{Z}_t)_{t > \mathscr{T}_b} \text{ hits } \Psi x_i \text{ before } a)\left(1 - \frac{\Psi x_i}{x_i}\right) + \mathbb{1}((\mathscr{Z}_t)_{t > \mathscr{T}_b} \text{ hits } y_i \text{ before } a)\left(1 - \frac{y_i}{\Psi x_i}\right).\end{aligned}$$

To verify this inequality, note that if the indicator on the left is one, then so are both indicators on the right, and when all three indicators are equal to one, the inequality is equivalent to the elementary inequality

$$1-\frac{\Psi x_i}{x_i}+1-\frac{y_i}{\Psi x_i}-\left(1-\frac{y_i}{x_i}\right)=\frac{\Psi(1-\Psi)x_i-(1-\Psi)y_i}{\Psi x_i}$$
$$\geq \frac{\Psi(1-\Psi)x_i-(1-\Psi)\Psi^2 x_i}{\Psi x_i}=(1-\Psi)^2\geq 0,$$

which holds since $y_i<\Psi^2 x_i$. Given a slack set $S\in\mathscr{L}$, we can therefore iterate this operation until we obtain a taut set $S^\bullet$ with $F(S)\leq F(S^\bullet)$; this iterative process must terminate after finitely many steps since $|S'|=|S|+1$ and every set in $\mathscr{L}$ contains at most $\lceil -1/\log\Psi\rceil$ pairs of points. Enumerate the pairs of points of $S^\bullet$ in decreasing order as (x_1,y_1), (x_2,y_2), ..., (x_ℓ,y_ℓ). Since every pair $(x,y)\in S^\bullet$ satisfies $y/x\leq\Psi^2$ and every two distinct pairs of points in $S^\bullet$ span disjoint open intervals of $[e^{-k-1},b]$ we must have that $y_i\leq\Psi^i b$ for every $1\leq i\leq\ell$ and hence that $\ell\leq\lceil -1/\log\Psi\rceil$ as previously mentioned. It follows that

$$F(S)\leq F(S^\bullet)=\sum_{i=1}^{\ell}\mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b}\text{ hits }y_i\text{ before }a)\big(1-\frac{y_i}{x_i}\big)$$
$$\leq\sum_{i=1}^{\lceil -1/\log\Psi\rceil}\mathbb{1}((\mathscr{Z}_t)_{t>\mathscr{T}_b}\text{ hits }\Psi^i b\text{ before }a)(1-\Psi^2),$$

as claimed, where we used that $S^\bullet$ is taut in the second inequality. □

With these bounds in hand, we can now prove Theorem 40. The proof will apply the Borel-Cantelli counterpart [88] which is an extension of the second Borel-Cantelli lemma to dependent events.

Lemma 49 (Borel-Cantelli Counterpart)**.** *If $(E_n)_{n\geq 0}$ is an increasing sequence of events satisfying the divergence condition $\sum_{n\geq 1}\mathbb{P}(E_n\mid E_{n-1}^c)=\infty$, then $\mathbb{P}(\bigcup_{n\geq 1}E_n)=1$.*

Setting $E_n=\cup_{1\leq i\leq n}A_i$ for $n\geq 1$ where $(A_i)_{i\geq 1}$ is an arbitrary sequence of events, the Borel-Cantelli counterpart implies in particular that

$$\text{if}\qquad\sum_{n\geq 1}\mathbb{P}\left(A_n\mid(\cup_{i=1}^{n-1}A_i)^c\right)=\infty\qquad\text{then}\qquad\mathbb{P}\left(\bigcup_{n\geq 1}A_n\right)=1.\tag{5.20}$$

Proof of Theorem 40. We assume that $\tau_\dagger=\infty$ with positive probability, the claim holding vacuously otherwise. We recall that we have applied Lemma 45 to ϕ from the statement of the theorem which yields the functions a and ψ. We continue to use the notation $k_0=k_0[Z]$ as in Definition 2. We let $\mathscr{G}$ be the event that $Z_n\leq e^{-\phi(n)}$ for all sufficiently large n. Since each permadrop gives rise to a distinct cut time, it suffices to prove that $\sum_k R_k = \infty$ almost surely on the event that Z is not killed. For each $k \geq 1$, define the events

$$A_k = \{R_{a(k)} \geq 1\} = \{\text{there is at least one large permadrop on the } a(k)\text{th scale}\},$$
$$B_k = \{\text{there exists } n \text{ such that } e^{-a(k)-1} < Z_n^* \leq e^{-a(k)-3/4}\},$$

and the event

$$C_k = \{\text{there exists } n \text{ such that } \sqrt{e^{-a(k-1)-1}e^{-a(k)}} \leq Z_n^* < e^{-a(k-1)-1}\}.$$

If $(B_k \cap C_k)^c$ holds for infinitely many k then Z is either killed or satisfies $Z_{n+1}^* \leq e^{-1/4} Z_n^*$ for infinitely many n, in which case the claim follows from Lemma 44. As such, it suffices to prove that

$$\mathbf{P}\left(\bigcup_{k=k_1}^{\infty} A_k \mid \mathscr{G} \text{ holds, } \tau_\dagger = \infty, \text{ and } B_k \cap C_k \text{ holds for all sufficiently large } k\right) = 1$$

for all $k_1 \geq k_0$, whenever the event being conditioned on has positive probability. We will assume for contradiction that there exists $k_1 \geq k_0$ such that this does not hold, and fix such a k_1 for the remainder of the proof.

For each $k \geq k_0$, define the event

$$G_k = \{Z \text{ is } \psi\text{-good on scale } a(k)\}.$$

It follows from Propositions 47 and 48 that there exists a universal constant $c > 0$ such that

$$\mathbf{P}_{\mathscr{T}_{1,a(k)}}(A_k) \geq \frac{c \cdot \mathbf{P}_{\mathscr{T}_{1,a(k)}}(G_k \cap B_k)}{1 + \log \psi^{-1}(a(k))} \tag{5.21}$$

for every $k \geq 1$, where we used that $\psi(x) \leq \sqrt{x}$ to bound $\log(\psi^{-1}(a(k))/a(k)) \geq \frac{1}{2} \log \psi^{-1}(a(k))$. For each $k \geq k_0$, let F_k be the event that for each $\ell < k$ and $0 \leq i \leq N_{a(\ell)-1}$ such that the drop $(d_{i,a(\ell)}, d_{i+1,a(\ell)})$ is *not* a permadrop, the process $\mathscr{Z}$ hits $d_{i,a(\ell)}$ at a time between $\mathscr{T}_{i+1,a(\ell)}$ and $\mathscr{T}_{1,a(k)}$. The event F_k is constructed precisely to force decorrelation between A_k and $(\bigcup_{i=k_1}^{k-1} A_i)^c$. Indeed, the intersection $F_k \cap C_k \setminus \cup_{i=k_1}^{k-1} A_i$

is measurable with respect to $\mathscr{F}_{\mathscr{T}_{1,a(k)}}$ and we can apply (5.21) to deduce that

$$\begin{aligned}\mathbf{P}(A_k \cap F_k \cap C_k \setminus \cup_{i=k_1}^{k-1} A_i) &= \mathbf{E}\left[\mathbb{1}(F_k \cap C_k \setminus \cup_{i=k_1}^{k-1} A_i)\mathbf{P}_{\mathscr{T}_{1,a(k)}}(A_k)\right] \\ &\geq \frac{c}{1+\log\psi^{-1}(a(k))}\mathbf{E}\left[\mathbb{1}(F_k \cap C_k \setminus \cup_{i=k_1}^{k-1} A_i)\mathbf{P}_{\mathscr{T}_{1,a(k)}}(G_k \cap B_k)\right] \\ &= \frac{c}{1+\log\psi^{-1}(a(k))}\mathbf{P}(F_k \cap C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{k-1} A_i)\end{aligned}$$

for every $k \geq k_0$. On the other hand, on $F_k^c \cap C_k$, the process $\mathscr{Z}$ has to recover after time $\mathscr{T}_{1,a(k)}$ by a multiplicative factor of at least $e^{-a(k)}/\sqrt{e^{-a(k)-a(k-1)-1}} = \sqrt{e^{-a(k)+a(k-1)+1}}$, and so we can apply optional stopping to $\mathscr{Z}$ at this time to upper bound

$$\mathbf{P}(F_k^c \cap C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{k} A_i) \leq \mathbf{P}(F_k^c \cap C_k) \leq \sqrt{e^{-a(k)+a(k-1)+1}}.$$

Hence, we have that

$$\mathbf{P}(A_k \setminus \cup_{i=k_1}^{k-1} A_i) \geq \frac{c}{1+\log\psi^{-1}(a(k))}\left(\mathbf{P}(C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{k-1} A_i) - \sqrt{e^{-a(k)+a(k-1)+1}}\right) \vee 0.$$

We deduce by linearity of expectation that

$$\sum_{k=k_1}^{\ell} \mathbf{P}(A_k \setminus \cup_{i=k_1}^{k-1} A_i) \geq c \cdot \mathbf{E}\left[\sum_{k=k_1}^{\ell} \frac{(\mathbb{1}(C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{k-1} A_i) - \sqrt{e^{-a(k)+a(k-1)+1}}) \vee 0}{1+\log\psi^{-1}(a(k))}\right]. \tag{5.22}$$

On the event that $\mathscr{G} \setminus \cup_{i=k_1}^{\infty} A_i$ holds and $B_k \cap C_k$ holds for all sufficiently large k (which has positive probability by assumption), we have by choice of ψ in Lemma 45 that $\liminf_{k\to\infty} \frac{1}{k}\sum_{\ell=k_1}^{k} \mathbb{1}(C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{\infty} A_i) > 0$ and hence by Lemma 46 that there exists an almost surely positive $\eta > 0$ and almost surely finite k_2 such that

$$\sum_{k=k_1}^{\ell} \frac{(\mathbb{1}(C_k \cap G_k \cap B_k \setminus \cup_{i=k_1}^{k-1} A_i)}{1+\log\psi^{-1}(a(k))} \geq \eta \sum_{k=k_1}^{\ell} \frac{1}{1+\log\psi^{-1}(a(k))}$$

for every $\ell \geq k_2$. Since $a(k) - a(k-1) \to \infty$ as $k \to \infty$, the other term in (5.22) is of lower order than this and we deduce that

$$\sum_{k=k_1}^{\infty} \mathbf{P}(A_k \setminus \cup_{i=k_1}^{k-1} A_i) = \infty.$$

It follows from the Borel-Cantelli counterpart that $\mathbf{P}(\cup_{k=k_1}^{\infty} A_k) = 1$, contradicting the definition of k_1. □

5.2.5 Completing the proof of Theorem 34

In this section we deduce Theorem 34 from Theorem 40. Note that the proof establishes a slightly stronger claim giving the almost sure existence of infinitely many cut times on the event that hitting probabilities decay quickly, without needing to assume that the latter occurs almost surely.

Proof of Theorem 34. Let $M = (\Omega, P, \dagger)$ be a transient Markov chain with killing, let $X = (X_n)_{n\geq 0}$ be a trajectory of M, and let $\phi : [0,\infty) \to [0,\infty)$ be a strictly increasing function such that

$$\sum_{n=1}^{\infty} \frac{1}{1 \vee \log(\phi^{-1}(n))} = \infty. \tag{5.23}$$

It suffices by Lemma 41 to prove that if the event $\mathscr{G} = \{\limsup_{n\to\infty} e^{\phi(n)} \mathbb{H}(X_n, X_m) < \infty$ for every $m \geq 0$ such that $X_m \neq \dagger\}$ has positive probability, then X is either killed or has infinitely many cut times almost surely conditional on $\mathscr{G}$. Compared to this statement, Theorem 40 has three additional hypotheses: that M is locally finite, that M is irreducible, and that $P(x,x) = 0$ for every $x \neq \dagger$. We will show that these assumptions can each be removed via a simple reduction argument.

Removing the condition that $P(x,x) = 0$ for every $x \neq \dagger$: First suppose that M is irreducible and locally finite but does not necessarily satisfy $P(x,x) = 0$ for every $x \neq \dagger$. Since M is irreducible and transient, $P(x,x) \neq 1$ for every $x \neq \dagger$. Consider the Markov chain $M' = (\Omega, P', \dagger)$ where P' is the transition matrix defined by $P'(x,x) = 0$ for every $x \neq \dagger$ and

$$P'(x,y) = \frac{P(x,y)}{1 - P(x,x)} \qquad \text{for every } x \in \Omega \setminus \{\dagger\} \text{ and } y \in \Omega \setminus \{x\}.$$

We can couple trajectories X and Y of M and M' so that X visits the same states as Y in the same order but possibly includes additional steps where it stays at the same non-graveyard vertex for more than one consecutive step. In particular, if Y has infinitely many cut times then X does also. Since the hitting probabilities for M and M' are equal and $Y_n \in \{X_m : m \geq n\}$ for every $n \geq 1$, if the event $\mathscr{G}$ holds for X then the analogous event holds for Y also, and the claim follows from Theorem 40.

Removing the condition that M is irreducible: Now suppose that M is locally finite but not necessarily irreducible, and does not necessarily satisfy $P(x,x) = 0$ for every $x \neq \dagger$. For each communicating class $C \neq \{\dagger\}$ of M we can define a Markov chain with killing $M_C = (C \cup \{\dagger\}, P_C, \dagger)$, where $P_C(u,v) = P(u,v)$ for each $u,v \in C$ and $P_C(u,\dagger) = \sum_{v \notin C} P(u,v)$ for each $u \in C$. When a trajectory X of the original Markov chain M enters a communicating class $C \neq \{\dagger\}$, it can be coupled with a trajectory of M_C up to the first time that it leaves C, at which time the coupled trajectory of M_C is killed. Observe that a trajectory of M must either pass though infinitely many communicating classes or enter some final communicating class C_f. If $C_f = \{\dagger\}$, the trajectory is killed and there is nothing to prove. Each time the trajectory (X_n) enters a new communicating class $C \neq \{\dagger\}$, the coupling with a trajectory of M_C together with the previous part of the proof implies that, conditional on $\mathscr{G}$, the walk will almost surely either stay in C forever and have infinitely many cut times or leave C. Thus, if $\mathscr{G}$ holds and X eventually stays in a single communicating class, then it is either killed or has infinitely many cut times almost surely. On the other hand, if X visits infinitely many communicating classes then the set of times at which it enters a new communicating class constitute an infinite set of cut times, so that the claim also holds in this case.

Removing the condition that M is locally finite: We now let M be arbitrary; it remains only to remove the restriction that it is locally finite. We assume that trajectory X starts at a non-recurrent state $X_0 \in \Omega$, the claim holding vacuously otherwise. We merge all the recurrent communicating classes of M into the single state $\dagger$ to give a Markov chain with killing $M' = (\Omega', P', \dagger)$, noting that we can couple trajectories of M and M' such that they are identical up to the first time the two trajectories enter a recurrent communicating class (which corresponds to be killed in M'). We enumerate the states in $\Omega' \setminus \{\dagger\}$ as $(y_i)_{i \geq 1}$ and for each state $y \in \Omega$ define $y^{\rightarrow} = \{z \in \Omega : P(y,z) > 0\}$. Fix $\varepsilon > 0$. Since every state in $\Omega' \setminus \{\dagger\}$ is transient, we can select for each $i \geq 0$ a subset L_i of the states in $y_i^{\rightarrow}$ such that $y_i^{\rightarrow} \setminus L_i$ is finite and the trajectory (X_n) on M' starting at X_0 satisfies

$$\mathbb{P}(\exists j \in \mathbb{N} \text{ such that } X_j = y_i \text{ and } X_{j+1} \in L_i) < \varepsilon 2^{-i}.$$

It follows by a union bound that the event $\mathscr{L} = \{\exists i, j \in \mathbb{N} \text{ such that } X_j = y_i \text{ and } X_{j+1} \in L_i\}$ that the trajectory ever makes a transition of this type has probability at most ε. We construct a new Markov chain with killing $M'' = (\Omega', P'', \dagger)$ where, for each $i \geq 1$, transitions from y_i to L_i are redirected to the graveyard state. That is, for each $i \geq 1$, we set $P''(y_i, v) = 0$ for every $v \in L_i$, set $P''(y_i, v) = P'(y_i, v)$ for each $v \notin L_i \cup \{\dagger\}$, and set $P''(y_i, \dagger) = P'(y_i, \dagger) + \sum_{v \in L_i} P'(y_i, v)$. This construction ensures that M'' is locally finite. We can couple trajectories X on M' and Y on M'' to be identical up until the time that X makes a transition from y_i to L_i for some $i \geq 1$, after which Y is killed. It follows from this coupling that

$$\mathbb{H}^M(x,y) \geq \mathbb{H}^{M''}(x,y) \qquad \text{for every } x, y \in \Omega' \setminus \{\dagger\}, \tag{5.24}$$

and hence under this coupling that $\mathbb{H}^{M''}(Y_n, Y_m) \leq \mathbb{H}^{M''}(X_n, X_m)$ whenever $n \geq m$ is such that $Y_n \neq \dagger$. Since M'' is locally finite it follows that, under this coupling, Y is either killed or has infinitely many cut times on the event $\mathscr{G} = \{\limsup_{n \to \infty} e^{\phi(n)} \mathbb{H}(X_n, X_m) < \infty$ for every $m \geq 0$ such that $X_m \neq \dagger\}$. The claim follows since X and Y coincide forever with probability at least $1 - \varepsilon$ and $\varepsilon > 0$ was arbitrary. $\square$

5.3 Superdiffusive walks

In this section we prove Theorem 36, which states that random walks on networks satisfying a mild superdiffusivity condition have infinitely many cut times almost surely. It will once again be convenient to work within a more general framework that allows for random walks to be killed. We define a **network with killing** to be a tuple $N = (V, E, c, K)$ where (V, E, c) is a network and $K : V \to [0, \infty)$ is a **killing function**. Given a network with killing $N = (V, E, c, K)$, the random walk on N is the Markov chain with state space $V \cup \{\dagger\}$ and with transition matrix defined by

$$P(u,v) = \frac{c(u,v)}{c(u) + K(u)} \quad \text{for } u, v \in V \qquad\qquad P(u,\dagger) = \frac{K(u)}{c(u) + K(u)} \quad \text{for } u \in V,$$

and $P(\dagger, \dagger) = 1$, where $c(u)$ denotes the total conductance of all edges emanating from u, and $c(u, v)$ denotes the conductance of the edge between u and v if one exists, and is zero otherwise. We will follow the standard practice of writing $p_n(u, v) = P^n(u, v)$ for transition probabilities.

The starting point of our analysis is the following well-known theorem of Varopoulos and Carne [99, 317] (see also [259]). While usually stated without allowing for killing, the same proof[1] applies equally well to networks with killing; the important thing is that P satisfies

[1] As pointed out to us by the referee, this version of the inequality can also be proved from the standard version by appending a loop of weight $K(u)$ to each vertex $u \in V$ and coupling a second random walk to the original which is killed when the original random walk first uses one of these loops, and which is identical to the original random walk up to that time.

the self-adjointness relation $(c(u)+K(u))P(u,v)=(c(v)+K(v))P(v,u)$ for every $u,v\in V$ and that the restriction of P to V is substochastic.

Theorem 50 (Varopoulos-Carne Inequality)**.** *The transition probabilities $p_n(x,y)$ of a random walk on a network with killing $N=(V,E,c,K)$ satisfy*

$$p_n(x,y)\leq\sqrt{\frac{c(y)+K(y)}{c(x)+K(x)}}\exp\left[-\frac{d(x,y)^2}{2n}\right]\rho^n. \tag{5.25}$$

for every $x,y\in V$ and $n\geq 1$, where ρ is the spectral radius of the restriction of P to V.

We will bound the spectral radius term trivially by 1 in all our applications of this inequality.

While we would naively like to use Varoupoulos-Carne together with our superdiffusivity hypothesis to obtain bounds on the decay of the Green's function along the random walk, and conclude by applying Theorem 34, unfortunately, this does not seem to be possible in general. Indeed, while it is possible to obtain bounds on the small-time and medium-time transition probabilities of the walk using the Varopoulos-Carne inequality, this inequality gives us no control of the large-time contributions to the Green's function. In our efforts to circumvent this issue, we will establish some rather general conditions under which we can compare the decay of $\mathbb{G}(X_n,X_0)$ and $p_n(X_n,X_0)$ that may be of independent interest.

5.3.1 Comparing $p_n(X_n,X_0)$ and $\mathbb{G}(X_n,X_0)$ assuming superpolynomial decay

The first step of our proof is to give conditions under which the a.s. rates of decay of $p_n(X_n,X_0)$ and $\mathbb{G}(X_n,X_0)$ can be compared. Given a connected network with killing N, we say that N satisfies the **superpolynomial decay condition** if

$$\lim_{n\to\infty}\frac{\log\sup_{u\in V}p_n(u,v)}{\log n}=-\infty \qquad \text{for some (and hence every) } v\in V. \tag{SPD}$$

Proposition 51. *Let N be a network with killing and let $X=(X_n)_{n\geq 0}$ be a random walk started at o. If the transition probabilities to o satisfy the superpolynomial decay condition* (SPD) *then*

$$\lim_{n\to\infty}\frac{\log p_n(X_n,X_0)}{\log\mathbb{G}(X_n,X_0)}=1$$

almost surely, with the convention that this ratio is equal to 1 when $X_n=\dagger$.

This proposition is not really needed for Theorem 36, since the superpolynomial decay hypothesis (SPD) would already suffice to deduce the claim from Theorem 34. It will, however, be used more seriously in the proof of Theorem 37. For random walks on finitely generated groups with positive speed, which always satisfy (SPD) by [259, Corollary 6.32], Proposition 51 implies that the Avéz entropy and exponential decay rate of the Green's function coincide, recovering a result of Benjamini and Peres [70, Proposition 6.2]. Similar results for groups that are *not* finitely generated have been obtained in [85].

Proposition 51 will be deduced from the following elementary observation.

Lemma 52. *The transition probabilities $p_n(x,y)$ of a random walk $(X_n)_{n\geq 0}$ on a network with killing $N=(V,E,c,K)$ satisfy*

$$\mathbb{E}\left[\frac{p_m(X_n,X_0)}{p_n(X_n,X_0)}\right] \leq \mathbb{P}(X_m \neq \dagger)+\mathbb{P}(X_n=\dagger) \leq 2, \tag{5.26}$$

for every $x\in V$ and $m,n\geq 0$, with the convention that the ratio is 1 *when $X_n=\dagger$.*

Proof of Lemma 52. Let A be the set of vertices $x\in V$ such that $p_n(X_0,x)>0$. Then we have that

$$\begin{aligned}\mathbb{E}\left[\frac{p_m(X_n,X_0)}{p_n(X_n,X_0)}\mathbb{1}(X_n\neq\dagger)\right] &= \mathbb{E}\left[\frac{p_m(X_0,X_n)}{p_n(X_0,X_n)}\mathbb{1}(X_n\neq\dagger)\right]\\ &= \sum_{x\in A} p_n(X_0,x)\frac{p_m(X_0,x)}{p_n(X_0,x)} = \sum_{x\in A} p_m(X_0,x) \leq \mathbb{P}(X_m\neq\dagger),\end{aligned}$$

which is easily seen to imply the claim. $\square$

Proof of Proposition 51. Using Lemma 52, an application of the Borel-Cantelli Lemma implies that there exists an almost surely finite random variable γ such that

$$p_m(X_n,X_0) \leq \gamma(m+1)^2(n+1)^2 p_n(X_n,X_0) \tag{5.27}$$

for every $n,m\geq 0$. Fix $\varepsilon>0$, let $n\geq 1$ and let $N=\lceil p_n(X_n,X_0)^{-\varepsilon}\rceil$. We deduce by summing (5.27) over $0\leq m\leq N$ that

$$\mathbb{G}(X_n,X_0)=\sum_{m=0}^{\infty} p_m(X_n,X_0) \leq \gamma(N+1)^3(n+1)^2 p_n(X_n,X_0)+\sum_{m=N+1}^{\infty} p_m(X_n,X_0).$$

Since $p_m(X_n,X_0) \leq \sup_v p_m(v,X_0)$ decays superpolynomially in m by (SPD) we can write this estimate in asymptotic notation as

$$\mathbb{G}(X_n,X_0) \leq p_n(X_n,X_0)^{1-3\varepsilon-o(1)} + p_n(X_n,X_0)^{\omega(1)} \qquad \text{a.s. as } n \to \infty \text{ for each fixed } \varepsilon > 0,$$

where $o(1)$ and $\omega(1)$ denote quantities tending to 0 and $+\infty$ respectively. The claim follows since $\varepsilon > 0$ was arbitrary and the inequality $\mathbb{G}(X_n,X_0) \geq p_n(X_n,X_0)$ holds trivially. □

Since $p_n(X_n,X_0)$ decays superpolynomially under the superdiffusivity assumption (5.3) by Varopoulos-Carne and we only require polynomial decay of $\mathbb{G}(X_n,X_0)$ to apply Theorem 34, to prove Theorem 36 it would suffice for us to have a much weaker comparison of the two quantities than that provided by Proposition 51. Such comparison inequalities can be provided by the proof of Proposition 51 under much weaker assumptions on transition probabilities that are only barely stronger than transience. For example, this argument is able to handle the ballistic case under the mild additional assumption that there exists $c > 0$ such that

$$\sup_u p_n(u,v) \leq \frac{C_v}{n(\log n)^{1+c}} \qquad \text{for every } v \in V \text{ and } n \geq 2, \tag{5.28}$$

where C_v is a finite constant depending on the choice of v. Unfortunately, we believe that such transition probability estimates need not hold in general, even when the random walk has positive speed. Indeed, identifying the origin of $\mathbb{Z}^2$ with the root of a binary tree gives an example where the random walk has positive liminf speed almost surely but where $p_n(0,0)$ is at least the probability that the walk makes an excursion of length n from the origin to itself in $\mathbb{Z}^2$, which is of order $n^{-1}(\log n)^{-2}$. Replacing $\mathbb{Z}^2$ in this example by a tree of slightly superquadratic growth should allow one to construct examples where the random walk has positive speed but where (5.28) does not hold for any $c > 0$; we do not pursue this further here. We believe that there exist examples where the random walk has positive speed but where $\mathbb{G}(X_n,X_0)$ decays very slowly, but this seems to require a more involved construction.

5.3.2 Spatially-dependent killing and the proof of Theorem 36

We now describe how we circumvent the issue discussed at the end of the previous subsection by introducing *spatially dependent killing* to our network, where we will take $K(x)$ to be a function of the distance of x from some fixed origin vertex o. We will show under the hypotheses of Theorem 36 that this killing function can be chosen to decay sufficiently quickly that the random walk has a positive probability never to be killed, but decay sufficiently slowly that the resulting network with killing satisfies (SPD).

We begin by finding the marginal rate of decay under which the resulting network with killing automatically satisfies (SPD). Given a network $N = (V,E,c)$ and a fixed origin vertex o, we write $\langle x \rangle = 2 \vee d(o,x)$ for each $x \in V$ to avoid division by zero.

Lemma 53. *Let $N = (V,E,c)$ be a network with $c_{\min} = \inf_{x\in V} c(x) > 0$, fix a vertex $o \in V$, let $\gamma \in \mathbb{R}$ and let $K : V \to [0,\infty)$ be the killing function defined by $K(x) = c(x)\min\{1, \langle x\rangle^{-2}(\log\langle x\rangle)^{\gamma}\}$. Then there exists a positive constant $c = c(\gamma)$ such that*

$$p_n(x,o) \leq \sqrt{\frac{8c(o)}{c_{\min}}} \exp\left[-c\,(\log n)^{\gamma/2}\right]$$

for every $x \in V$ and $n \geq 2$. In particular, if $\gamma > 2$ then (V,E,c,K) satisfies (SPD).

The rough idea behind this lemma is as follows: Suppose we run a random walk for time n started at some vertex x. If $d(o,X_m) \gg \sqrt{n}$ for some $0 \leq m \leq n$ then the probability of hitting the origin at time n is small as a consequence of Varopoulos-Carne. On the other hand, if this never happens, the higher rate of killing ensures that the walk is killed before time n with high probability and is therefore unlikely to hit the origin at time n.

Proof of Lemma 53. Let $\mathbb{P}_x$ denote the law of the random walk $X = (X_n)_{n\geq 0}$ on the network with killing (V,E,c,K) started at some fixed vertex $x \in V$, and let $\tau_\dagger$ denote the time the walk is killed (i.e. first visits the graveyard state $\dagger$). We define $d(o,\dagger) = \infty$ and decompose

$$\begin{aligned}\mathbb{P}_x(X_n = o) = \mathbb{P}_x(&X_n = o \text{ and } d(o,X_m) > r \text{ for some } 0 \leq m \leq n)\\ &+ \mathbb{P}_x(X_n = o \text{ and } d(o,X_m) \leq r \text{ for every } 0 \leq m \leq n) \quad (5.29)\end{aligned}$$

for each $n, r \geq 2$, where r is a parameter we will optimize over at the end of the proof. We begin by analysing the first term on the right hand side of (5.29). Let κ be the stopping time $\kappa := \inf\{m \geq 0 : d(o,X_m) > r\}$. We apply the strong Markov property at κ together with Varopoulos-Carne to give that

$$\begin{aligned}\mathbb{P}_x(X_n = o \text{ and } d(o,X_m) > r &\text{ for some } 0 \leq m \leq n)\\ &\leq \sum_{m=0}^{n} \sum_{z\in V} \mathbb{P}_x(\kappa = m, X_\kappa = z)\mathbb{P}_z(X_{n-m} = o)\\ &\leq \sqrt{\frac{c(o)+K(o)}{c(z)+K(z)}} \exp\left[-\frac{r^2}{2n}\right] \leq \sqrt{\frac{2c(o)}{c_{\min}}} \exp\left[-\frac{r^2}{2n}\right],\end{aligned}$$

where $c_{\min} = \inf_{z\in V} c(z)$, and where the final inequality follows by definition of K. We now turn our attention to the second term on the right hand side of (5.29). Each time the walk makes a step at distance at most r it is killed with probability at least $\frac{1}{2}(1 \wedge r^{-2}(\log r)^{\gamma})$. Letting $c_1 = c_2(\gamma)$ be a positive constant such that this probability is at least $c_1 r^{-2}(\log r)^{\gamma}$, we deduce that

$$\begin{aligned}\mathbb{P}_x\big(X_n = o \text{ and } d(o,X_m) \le r \text{ for every } 0 \le m \le n\big)\\ &\le \mathbb{P}_x\big(\tau_\dagger > n \text{ and } d(o,X_m) \le r \text{ for every } 0 \le m \le n\big)\\ &\le \left(1 - \frac{c_1(\log r)^{\gamma}}{r^2}\right)^n \le \exp\left[-\frac{c_1(\log r)^{\gamma} n}{r^2}\right],\end{aligned}$$

where we used the inequality $1-t \le e^{-t}$ in the final inequality. Substituting these two estimates into (5.29) yields that

$$\mathbb{P}_x(X_n = o) \le \sqrt{\frac{2c(o)}{c_{\min}}}\left(\exp\left[-\frac{r^2}{2n}\right] + \exp\left[-\frac{c_1(\log r)^{\gamma} n}{r^2}\right]\right),$$

and the claim follows by taking $r = \lceil n^{1/2}(\log n)^{\gamma/4}\rceil$. □

Let $N = (V,E,c)$ be a network, let o be a vertex of N, and let $X = (X_n)_{n\ge 0}$ be the random walk on N. Let $r > 0$ and let $\mathscr{S}_r$ be the event that

$$\liminf_{n\to\infty} \frac{d(o,X_n)}{n^{1/2}(\log n)^r} > 0.$$

We next wish to show that for any choice of r, we can choose the killing function K as in Lemma 53 such that if $\mathscr{S}_r$ holds, the walk does not "feel" the effects of the killing. More precisely, we can ensure the killing function decays quickly enough such that conditional on the path of the walk, the walk almost surely has a positive probability of never getting killed. To formulate this lemma, let us first note that we can couple the random walks on (V,E,c) and (V,E,c,K) so that they coincide up until the killing time $\tau_\dagger$. Writing X for the unkilled walk and writing $\mathbf{P}_x$ for the joint law of X and $\tau_\dagger$ when X is started at $x \in V$, this coupling is determined by the equality

$$\mathbf{P}_x(\tau_\dagger = n \mid X) = K(X_{n-1})\prod_{i=0}^{n-2}(1 - K(X_i)).$$

Lemma 54. *Let $N = (V,E,c)$ be a network with $c_{\min} = \inf_{x\in V} c(x) > 0$, fix a vertex $o \in V$, let $\gamma \in \mathbb{R}$, and let $K : V \to [0,\infty)$ be the killing function defined by $K(x) = c(x)\min\{1, \langle x\rangle^{-2}(\log\langle x\rangle)^{\gamma}\}$. If X is a random walk on N and $\gamma + 1 < 2r$, then $\mathbf{P}_x(\tau_\dagger = \infty \mid X) > 0$ almost surely on the event $\mathscr{S}_r$.*

Proof of Lemma 54. We can write the conditional probability $\mathbf{P}_x(\tau_\dagger = \infty \mid X)$ as an infinite product

$$\mathbf{P}_x(\tau_\dagger = \infty \mid X) = \prod_{i=0}^{\infty}(1-K(X_i)), \qquad \text{which is positive if and only if} \qquad \sum_{i=0}^{\infty} K(X_i) < \infty.$$

We have by calculus that there exists a random variable α taking values in $[1,\infty]$ that is finite on the event $\mathscr{S}_r$ and satisfies $K(X_n) \leq \alpha(\log n)^{\gamma-2r}n^{-1}$ for every $n \geq 1$, and it follows that if $2r > 1+\gamma$ then $\sum_{i=0}^{\infty} K(X_i) < \infty$ on the event $\mathscr{S}_r$ as required. □

We are now ready to complete the proofs of Theorems 36 and 37.

Proof of Theorem 36. Let $r > 3/2$ and $2 < \gamma < 2r-1$. Let $N = (V,E,c)$ be a network with $c_{\min} = \inf_{x\in V} c(x) > 0$, fix a vertex $o \in V$, and let $K : V \to [0,\infty)$ be the killing function defined by $K(x) = c(x)\min\{1, \langle x\rangle^{-2}(\log\langle x\rangle)^{\gamma}\}$. Couple the random walk X on N with the killing time $\tau_\dagger$ as above, write $X^\dagger$ for the killed walk, and assume that the superdiffusivity event $\mathscr{S}_r$ has positive probability. Let $p_n^\dagger$ and $\mathbb{G}_\dagger$ denote transition probabilities and the Green's function with respect to the killed network $N_\dagger = (V,E,c,K)$. Lemma 53 implies that $N_\dagger$ satisfies the superpolynomial decay condition (SPD), and we deduce from Proposition 51 that

$$\lim_{n\to\infty} \frac{\log p_n^\dagger(X_n^\dagger, X_0^\dagger)}{\log \mathbb{G}_\dagger(X_n^\dagger, X_0^\dagger)} = 1 \tag{5.30}$$

almost surely, where the ratio is considered to be equal to 1 when $X_n^\dagger = \dagger$. Varopoulos-Carne yields that

$$p_n^\dagger(X_n^\dagger, X_0^\dagger) = \exp\left[-\Omega\left((\log n)^{2r}\right)\right] \qquad \text{as } n\to\infty$$

when $\mathscr{S}_r$ holds, and hence by (5.30) that

$$\mathbb{G}_\dagger(X_n^\dagger, X_0^\dagger) = \exp\left[-\Omega\left((\log n)^{2r}\right)\right] \qquad \text{as } n\to\infty$$

almost surely on the event $\mathscr{S}_r$. (Here we recall that $\Omega(f(n))$ denotes a quantity that is lower bonded by a (possibly random) positive multiple of $f(n)$ for large values of n.) Since $r > 3/2 > 1/2$, this decay is superpolynomial, and it follows from Theorem 34 that $X^\dagger$ is either killed or has infinitely many cut times almost surely on the event $\mathscr{S}_r$. Since we also have that the conditional probability $\mathbf{P}_x(\tau_\dagger = \infty \mid X)$ is almost surely positive on the event $\mathscr{S}_r$, we deduce that X has infinitely many cut times almost surely on the event $\mathscr{S}_r$ as claimed. □

Proof of Theorem 37. Since G has bounded degrees and the walk has positive liminf speed almost surely, it follows as above that we can take a bounded killing function K so that the walk has a.s. positive conditional probability not to be killed and the killed Green's function $\mathbb{G}_\dagger(X_n^\dagger, X_o^\dagger)$ decays exponentially. On the other hand, since the degrees and the killing function are both bounded, there exists a positive constant c such that $\mathbb{G}_\dagger(X_{n+1}^\dagger, X_o^\dagger) \geq c \cdot \mathbb{G}_\dagger(X_n^\dagger, X_o^\dagger)$ for every $n \geq 0$ such that $X_{n+1}^\dagger \neq \dagger$. Combined with exponential decay this implies that if we define $A_a = \{n : Z_{n+1}^\dagger \leq a \min_{m \leq n} Z_m^\dagger\}$, where $Z_n^\dagger = \mathbb{G}_\dagger(X_n^\dagger, X_o^\dagger)$, and define $\mathscr{A}_a$ to be the event that A_a has positive liminf density then $\bigcup_{k=1}^\infty \mathscr{A}_{(k-1)/k}$ has probability 1. On the other hand, by optional stopping, for each $n \geq 1$ the conditional probability that n is a cut time given everything the walk has done up to time n is bounded below by $1-a$ whenever $n \in A_a$. From here the claim follows easily by standard arguments and we omit the details. □

5.4 Sharpness for birth-death chains

In this final section, we demonstrate that the integral condition given in Theorem 34 is sharp by comparing our results to those of Csáki, Földes, and Révész [109] on the cut times of birth-death chains. Throughout this section, $(X_n)_{n\geq 0}$ will denote a random walk on $\mathbb{Z}_{\geq 0}$ with transition probabilities of the form

$$E_i := \mathbb{P}(X_{n+1} = i+1 \mid X_n = i) = 1 - \mathbb{P}(X_{n+1} = i-1 \mid X_n = i) = \begin{cases} 1 & \text{if } i = 0 \\ 1/2 + p_i & \text{otherwise} \end{cases},$$

where $-1/2 < p_i < 1/2$ for each $i \geq 1$. For each $m \geq 0$, define

$$D(m) = 1 + \sum_{j=1}^{\infty} \prod_{i=1}^{j} \left(\frac{1}{E_{m+i}} - 1 \right).$$

The aforementioned work [109] establishes the following dichotomy. (Here we rephrase their theorem in terms of cut times and omit the strengthened conclusion concerning strong cut points.)

Theorem 55 ([109, Theorem 1.1])**.** *Let $(X_n)_{n\geq 0}$ be a transient birth-death chain as defined above with $0 \leq p_i < 1/2$ for each $i \geq 1$.*

- *If $\sum_{n=2}^\infty (D(n)\log n)^{-1} < \infty$, then (X_n) has finitely many cut times a.s.*
- *If $D(n) \leq n(\log n)^{1/2}$ and $\sum_{n=2}^\infty (D(n)\log n)^{-1} = \infty$, then (X_n) has infinitely many cut times a.s.*

We use this Theorem to prove the following partial converse of Theorem 34. We let $\mathbb{G}(n)=\mathbb{G}(n,0)$ denote the Green's function associated with (X_n) and say a function $F:[0,\infty)\to(0,\infty)$ is **eventually log-convex** if there exists $r\geq 0$ such that the restriction of F to the interval $[r,\infty)$ is log-convex.

Proposition 56. *Given any decreasing differentiable bijection $\Phi:[0,\infty)\to(0,1]$ that is eventually log-convex and satisfies*

$$\int_0^1 \frac{1}{u(1\vee\log\Phi^{-1}(u))}\,\mathrm{d}u<\infty,$$

there exists a nearest-neighbour random walk $(X_n)_{n\geq 0}$ on $\mathbb{Z}_{\geq 0}$, with Green's function $\mathbb{G}(n)=\mathbb{G}(n,0)$, such that $\limsup_{n\to\infty}\Phi(n)^{-1}\mathbb{G}(X_n)<\infty$ and $(X_n)_{n\geq 0}$ has finitely many cut times almost surely.

In the proof of this proposition, we will utilize the following two elementary identities relating the quantities p_n, $\mathbb{G}(n)$, and $D(n)$ for each $n\geq 1$:

$$D(n-1)=\frac{\mathbb{G}(n-1)}{\mathbb{G}(n-1)-\mathbb{G}(n)}, \tag{5.31}$$

$$p_n=\frac{1}{2}\frac{\mathbb{G}(n-1)+\mathbb{G}(n+1)-2\mathbb{G}(n)}{\mathbb{G}(n-1)-\mathbb{G}(n+1)}. \tag{5.32}$$

The first identity follows from [109, (2.1)] and the elementary identity $\mathbb{H}(n+1,n)=\mathbb{H}(n+1)/\mathbb{H}(n)$, where $\mathbb{H}(n)$ is the probability that (X_m) will hit 0 when $X_0=n$, and the second identity follows from [109, (2.2)] together with the first.

Plugging (5.31) into $\sum_{n=2}^{\infty}(D(n)\log n)^{-1}$, we observe that their summation criterion is roughly related to our integral condition by a change of variables. We prove Proposition 56 by formalising this relationship for a walk whose Green's function is an appropriate transformation of the input function Φ. We then conclude by proving a very weak lower bound on the displacement of the walk from 0.

Proof of Proposition 56. Let f be the decreasing, log-convex function $f(x):=e^{-\sqrt{\log(x+2)}}$, and let $M\geq 2$ be the smallest integer such that the restriction of Φ to $[M,\infty)$ is log-convex. We begin by defining the function $\widetilde{\Phi}:[0,\infty)\to(0,\infty)$ by

$$\widetilde{\Phi}(x)=\Phi((x+M)^4)f(x),$$

and noting some its properties. First, observe that $\widetilde{\Phi}\leq\Phi$ is strictly positive, strictly decreasing, log-convex and differentiable. Moreover, since $\widetilde{\Phi}(x)\leq\Phi((x+M)^4)\wedge f(x)$, we also have that $\widetilde{\Phi}^{-1}(x) \geq (\Phi^{-1}(x)^{1/4} - M) \vee f^{-1}(x)$, and hence that there exists a $C < \infty$ finite such that

$$\begin{aligned}
\int_0^1 \frac{1}{u(1 \vee \log \widetilde{\Phi}^{-1}(u))}\,\mathrm{d}u &\leq \int_0^1 \frac{1}{u(1 \vee \log[(\Phi^{-1}(x)^{1/4} - M) \vee f^{-1}(x)])}\,\mathrm{d}u \\
&= \int_0^1 \min\left\{ \frac{1}{u(1 \vee \log f^{-1}(x))}, \frac{1}{u(1 \vee \log[(\Phi^{-1}(x)^{1/4} - M)])} \right\} \mathrm{d}u \\
&\leq C + C\int_0^1 \frac{1}{u(1 \vee \log \Phi^{-1}(x))}\mathrm{d}u < \infty, \qquad (5.33)
\end{aligned}$$

where for functions $F \in \{\widetilde{\Phi}, f\}$, we use the convention that $1 \vee \log F^{-1}(u) = 1$ when $F^{-1}(u)$ is not defined. We also note that the logarithmic derivative $(\log \widetilde{\Phi})'$ of $\widetilde{\Phi}$, which is increasing by log-convexity of $\widetilde{\Phi}$, satisfies the inequality

$$-\frac{\mathrm{d}}{\mathrm{d}x}\log \widetilde{\Phi}(x) \geq -\frac{\mathrm{d}}{\mathrm{d}x}\log f(x) = \frac{1}{2(x+2)\sqrt{\log(x+2)}}. \qquad (5.34)$$

We now use the function $\widetilde{\Phi}$ to define a Markov chain satisfying the desired properties. For $i \geq 1$, we define

$$p_i = \frac{1}{2}\frac{\widetilde{\Phi}(n-1) + \widetilde{\Phi}(n+1) - 2\widetilde{\Phi}(n)}{\widetilde{\Phi}(n-1) - \widetilde{\Phi}(n+1)}, \qquad (5.35)$$

which is non-negative since $\widetilde{\Phi}$ is convex and strictly less than $1/2$ since $\widetilde{\Phi}$ is strictly decreasing. We can therefore define a nearest-neighbour random walk $(X_n)_{n\geq 0}$ on the integers with $X_0 = 0$ and with transition probabilities

$$\mathbb{P}(X_{n+1} = i+1 \mid X_n = i) = 1 - \mathbb{P}(X_{n+1} = i-1 \mid X_n = i) = \frac{1}{2} + p_i \qquad \text{for } i \geq 1,$$

and $\mathbb{P}(X_{n+1} = 1 \mid X_n = 0) = 1$. Comparing (5.35) and (5.32), it follows by induction on n that the Green's function of this Markov chain is given by

$$\mathbb{G}(n) = C\widetilde{\Phi}(n) \qquad \text{for } n \geq 0,$$

for some constant $C = \mathbb{G}(0)/\widetilde{\Phi}(0)$ independent of n. Therefore, to complete the proof, it suffices to show that

$$\limsup_{n\to\infty} \frac{\widetilde{\Phi}(X_n)}{\Phi(n)} \leq \limsup_{n\to\infty} \frac{\Phi((X_n+M)^4)}{\Phi(n)} < \infty \quad \text{a.s.} \qquad (5.36)$$

and that X has at most finitely many cut times almost surely.

We first apply Theorem 55 to prove that X has finitely many cut times almost surely. Let $N \geq 3$ be large enough such that $\Phi(N-1) < f(2)$. We can calculate

$$\sum_{n=N}^{\infty} \frac{1}{D(n)\log n} = \sum_{n=N}^{\infty} \frac{\widetilde{\Phi}(n) - \widetilde{\Phi}(n+1)}{\widetilde{\Phi}(n)\log n}$$
$$\leq \sum_{n=N}^{\infty} \frac{-\widetilde{\Phi}'(n)}{\widetilde{\Phi}(n)\log n} \leq \int_{N-1}^{\infty} \frac{-\widetilde{\Phi}'(x)}{\widetilde{\Phi}(x)\log x}\,\mathrm{d}x \leq \int_{0}^{\Phi(N-1)} \frac{\mathrm{d}u}{u \log \widetilde{\Phi}^{-1}(u)} < \infty,$$

where the first equality follows from (5.31), the first inequality is by convexity, the second follows by integral comparison as $(\mathrm{d}/\mathrm{d}x)[\log \widetilde{\Phi}(x)]$ is increasing, the third follows by the substitution $u = \widetilde{\Phi}(x)$ and the inequality $\widetilde{\Phi} \leq \Phi$, and the fourth follows from (5.33). The claim then follows from Theorem 55.

Finally, we prove (5.36). As $\widetilde{\Phi}$ is decreasing, it suffices to show that

$$\inf_{m \geq n} X_m \geq n^{1/2-o(1)} \quad \text{a.s. as } n \to \infty \qquad \text{and hence that} \qquad \liminf_{n\to\infty} \frac{X_n}{n^{1/4}} > 1 \quad \text{a.s.} \tag{5.37}$$

Since $\widetilde{\Phi}$ is log-convex, $\widetilde{\Phi}(m+1) \geq \widetilde{\Phi}(m)\widetilde{\Phi}(1)/\widetilde{\Phi}(0)$ for every $m \geq 1$. For each $m \geq 0$, define $H_m = |\{n \geq 0 : X_n = m\}|$. By [109, Lemma B], for each $m \geq 0$, H_m is a geometric random variable with parameter

$$1/\mu_m := \frac{1+2p_m}{2} \cdot \frac{\widetilde{\Phi}(m) - \widetilde{\Phi}(m+1)}{\widetilde{\Phi}(m)} \geq \frac{1}{2} \frac{\widetilde{\Phi}(m) - \widetilde{\Phi}(m+1)}{\widetilde{\Phi}(m)}$$
$$\geq \frac{-\widetilde{\Phi}'(m+1)}{2\widetilde{\Phi}(m)} \geq -\frac{\widetilde{\Phi}(1)\widetilde{\Phi}'(m+1)}{2\widetilde{\Phi}(0)\widetilde{\Phi}(m+1)} \geq \frac{\widetilde{\Phi}(1)}{4\widetilde{\Phi}(0)(m+2)\sqrt{\log(m+2)}},$$

where the second inequality follows by convexity of $\widetilde{\Phi}$ and the final inequality follows from (5.34). Since each H_m is a geometric random variable with mean of order $m^{1+o(1)}$, it follows by an elementary application of Borel-Cantelli that $\max_{m \leq n} H_m = n^{1+o(1)}$ almost surely as $n \to \infty$, and hence that $\max_{m \leq n} X_m \geq n^{1/2-o(1)}$ almost surely as $n \to \infty$. On the other hand, letting τ_n be the hitting time of n for each $n \geq 1$, we have by optional stopping that

$$\mathbb{P}(X \text{ visits } m \text{ after } \tau_n) \leq \frac{\mathbb{G}(n)}{\mathbb{G}(m)} = \frac{\widetilde{\Phi}(n)}{\widetilde{\Phi}(m)} \leq \frac{f(n)}{f(m)}$$

for each $0 \leq m \leq n$. Since $f(2^k)/f(\lfloor 2^{(1-\varepsilon)k} \rfloor)$ is superpolynomially small in k for each fixed $\varepsilon > 0$, we deduce by a further simple Borel-Cantelli argument that $\inf_{m \geq n} X_m = (\max_{m \leq n} X_m)^{1-o(1)} \geq n^{1/2-o(1)}$ almost surely as $n \to \infty$, completing the proof. $\square$

Chapter 6

[E] Logarithmic corrections to the Alexander–Orbach conjecture for the four-dimensional uniform spanning tree

Abstract. We compute the precise logarithmic corrections to Alexander–Orbach behaviour for various quantities describing the geometric and spectral properties of the four-dimensional uniform spanning tree. In particular, we prove that the volume of an intrinsic n-ball in the tree is $n^2(\log n)^{-1/3+o(1)}$, that the typical intrinsic displacement of an n-step random walk is $n^{1/3}(\log n)^{1/9-o(1)}$, and that the n-step return probability of the walk decays as $n^{-2/3}(\log n)^{1/9-o(1)}$.

6.1 Introduction

The behaviour of random walks on random fractals has been the subject of intense study since the 1970s [113], and a sophisticated and widely applicable theory has now developed on the topic [49, 219, 226]. In particular, it is now well established that the asymptotic behaviour of spectral quantities such as exit times, return probabilities, and walk displacement are determined under mild conditions by geometric properties such as volume growth and resistance growth [60, 226], with very general results to this effect established in the recent work of Lee [244]. This theory has led to a fairly complete understanding of several notable motivating examples including random planar maps [111, 172, 174, 175], high-dimensional percolation and branching random walks [49, 59, 223], and uniform spanning trees in two dimensions [44], three dimensions [27], and high dimensions ($d > 4$) [193]. The analysis of other important examples such as two-dimensional critical percolation remain largely open despite significant partial progress [149, 218, 219].

As suggested by this list of examples, many of the most interesting random fractals arise from critical statistical mechanics models, and for many such models the geometric and spectral properties of the associated random fractal depends heavily on the dimension in which the model is considered. Indeed, for many random fractals arising in statistical mechanics, a dichotomy emerges around an *upper-critical dimension* [325], denoted d_c, which is equal to 4 for the uniform spanning tree and 6 for percolation: below this dimension, the behaviour of the fractal is highly dependent on the geometry of the underlying space, while above this dimension the fractal displays *mean-field behaviour*, meaning that its large-scale behaviour is the same as it would be in a 'geometrically trivial' setting such as the complete graph or the binary tree. For many models the mean-field regime is described by *Alexander-Orbach* behaviour [9, 50, 219], in which the relevant random fractal has quadratic volume growth, spectral dimension $4/3$, and typical n-step walk displacement of order $n^{1/3}$. Indeed, Alexander-Orbach behaviour has been proven to hold for high-dimensional *oriented* percolation by Barlow, Jarai, Kumagai, and Slade [49], high-dimensional percolation by Kozma and Nachmias [223], and for the high-dimensional uniform spanning tree by the second author [193]. (An interesting example that is *not* expected to exhibit Alexander-Orbach behaviour in high dimensions is the minimal spanning forest, mean-field models of which have cubic volume growth and spectral dimension $3/2$ [1, 279].)

At the upper-critical dimension itself ($d = d_c$), it is expected that mean-field behaviour *almost* holds, with many quantities of interest expected to exhibit a polylogarithmic correction to their mean-field scaling. It is this regime that provides the focus of this paper, in which we determine the precise order of the polylogarithmic corrections to scaling for the geometric and spectral properties of the uniform spanning tree (UST) at its upper-critical dimension $d_c = 4$. The particular polylogarithmic corrections we compute are those governing the volume of balls, the resistance across them, and the return probabilities, range, displacement and exit times of random walks on the tree. Most of our work goes into estimating the volume growth and resistance growth of the 4d UST, with the associated random walk estimates following straightforwardly by techniques developed in [49, 227] that are by now rather standard. (The relevant proofs are presented in a self-contained way in Section 6.3.3.) We believe that this is the first time that polylogarithmic corrections to Alexander-Orbach behaviour have been computed for the random walk on a random fractal at the upper-critical dimension. Following [304], which computes the exact polylogarithmic corrections to a random walk on the four-dimensional random walk trace, we also believe that our work is the second time such polylogarithmic corrections to random walk behaviour at the upper critical dimension have been computed for any model. Partial progress on this problem for other models includes [208] (see also [209]) in which the existence of a non-trivial polylogarithmic correction to resistance growth is established for oriented branching random walk in $\mathbb{Z}^6 \times \mathbb{Z}_+$.

6.1.1 The uniform spanning tree

Over the last thirty years, the uniform spanning tree has emerged as a model of central importance throughout probability theory, with close connections to many other topics including electrical networks [93, 221], loop-erased random walk [68, 231, 321], the dimer model [72, 215], the Abelian sandpile model [76, 193, 210, 211, 266] and the random cluster model [164, 198].

We now very briefly introduce the model, referring the reader to e.g. [42, 193, 259] for further background. The **uniform spanning tree** of a finite connected graph is defined by choosing a spanning tree (i.e. a connected subgraph that contains every vertex and no cycles) of the graph uniformly at random. Pemantle [288] proved that there is a well-defined infinite volume limit of the uniform spanning tree of the hypercubic lattice $\mathbb{Z}^d$ which does not depend on the boundary conditions used when taking the limit and which is connected a.s. if and only if $d \leq 4$ (see also [68]). This infinite volume limit is known as the **uniform spanning tree** of $\mathbb{Z}^d$ when $d \leq 4$ and the **uniform spanning forest** of $\mathbb{Z}^d$ when $d \geq 5$. The critical dimension $d = 4$ is characterized by the UST just barely managing to be connected, with two points at Euclidean distance n typically connected by a path of Euclidean diameter much[1] larger than n and with the length of the path in the tree connecting two neighbouring vertices having an extremely heavy $(\log n)^{-1/3}$ tail [235]. This heavy tail on the probability of an abnormally long connection, and the related fact that the length of a loop-erased random walk in four dimensions is only very weakly concentrated, is responsible for much of the technical difficulties encountered in the paper. For example, it makes it difficult to justify the important heuristic that the volume of the intrinsic n-ball in the tree comes mostly from 'typical' points for which the tree-geodesic to the origin has Euclidean diameter of order $n^{1/2}(\log n)^{1/6}$.

6.1.2 Distributional asymptotic notation

To facilitate a clean presentation of our main results, we use *distributional asymptotic notation* (a.k.a. "big-O and little-o in probability" notation). Since this notation is not at all standard in

[1]Heuristic calculations suggest that the path connecting two distant points x and y has Euclidean diameter distributed approximately like $\|x-y\|^{1+Z}$ where Z is an exponential random variable.

probability theory[2], let us take a moment to explain how it is used. We hope the reader will find this diversion worthwhile after seeing how clean the statements of our main theorems are compared with similar results in the literature, and consider using this notation in their own work.

Before introducing this notation, let us first briefly introduce standard (deterministic) asymptotic notation as we use it. We write $\asymp$, $\succeq$, and $\preceq$ for equalities and inequalities holding to within positive multiplicative constants, so that if f and g are non-negative then "$f(n) \preceq g(n)$ for every $n \geq 1$" means that there exists a positive constant C such that $f(n) \leq Cg(n)$ for every $n \geq 1$. (We will often drop the "for every $n \geq 1$" and write simply "$f(n) \preceq g(n)$" when doing so does not cause confusion.) We use Landau's asymptotic notation similarly, so that $f(n) = O(g(n))$, $f(n) = \Omega(g(n))$, and $f(n) = \Theta(g(n))$ mean the same thing as $f(n) \preceq g(n)$, $f(n) \succeq g(n)$, and $f(n) \asymp g(n)$ respectively, while $f(n) = o(g(n))$ means that $f(n)/g(n) \to 0$ as $n \to \infty$. More complicated expressions can be obtained by putting this notation inside functions, so that e.g. $f(n) = O(e^{n-o(n^{1/2})})$ means that there exists a non-negative function $h(n)$ with $n^{-1/2}h(n) \to 0$ and a positive constant C such that $f(n) \leq Ce^{n-h(n)}$ for every $n \geq 1$. Implicit constants and functions given by this notation will always be non-negative, and we denote quantities of uncertain sign using $\pm O$, $\pm o$, etc. (While this is not completely standard, it greatly increases the expressive power of the notation.) Be careful to note that when forming such compound expressions, Θ should always be interpreted as the conjunction of O and Ω, so that "$f(n) = \Theta(e^{n-o(n)})$" means the same thing as "$f(n) = O(e^{n-o(n)})$ and $f(n) = \Omega(e^{n-o(n)})$", which means that there exist positive constants c and C and *possibly distinct* non-negative functions h^+ and h^- with $\lim_{n\to\infty} h^+(n) = \lim_{n\to\infty} h^-(n) = 0$ such that $f(n) \leq Ce^{n-h^+(n)}$ and $f(n) \geq ce^{n-h^-(n)}$. Whenever we use asymptotic notation, we can add a qualifier such as "as $n \to \infty$" to mean that the inequalities in question hold only for sufficiently large n; this will typically be used to avoid worrying about expressions such as $\log\log n$ being undefined or negative for small values of n.

We use boldface characters to apply this notation in settings where the relevant bounds are guaranteed only to hold with high probability, rather than deterministically. Given two sequences of (possibly deterministic) non-negative random variables (X_n) and (Y_n) defined

[2]Indeed, it is surprising how non-standard this notation is in the probability theory literature given how useful it is. The use of this notation has previously been advocated by Janson [205] who uses the notation O_p, o_p, Θ_p etc. rather than $\mathbf{O}, \mathbf{o}, \boldsymbol{\Theta}$ as we write here. We use bold letters rather than p subscripts since e.g. O_p would typically be used in probability to denote a deterministic upper bound whose implicit constants depend on a parameter p, and we wish to avoid clashing with this existing notational convention. The particular choice to use bold characters was made since LaTeX includes bold fonts for Greek characters by default while e.g. \mathscr{\Theta} and \mathcal{\Theta} are not defined.

on the same probability space, we write

$$
\begin{aligned}
X_n &= \mathbf{O}(Y_n) && \text{to mean that} && \lim_{\lambda\to\infty} \sup_n \mathbb{P}(X_n \geq \lambda Y_n) = 0,\\
X_n &= \boldsymbol{\Omega}(Y_n) && \text{to mean that} && \lim_{\lambda\to\infty} \sup_n \mathbb{P}(X_n \leq \lambda^{-1} Y_n) = 0,\\
X_n &= \boldsymbol{\Theta}(Y_n) && \text{to mean that} && X_n = \mathbf{O}(Y_n) \text{ and } X_n = \boldsymbol{\Omega}(Y_n), \text{ and}\\
X_n &= \mathbf{o}(Y_n) && \text{to mean that} && \lim_{n\to\infty} \mathbb{P}(X_n \geq \varepsilon Y_n) = 0 \text{ for every } \varepsilon > 0.
\end{aligned}
$$

In other words, $X_n = \mathbf{O}(Y_n)$ and $Y_n = \boldsymbol{\Omega}(X_n)$ both mean that $\{X_n/Y_n\}$ is tight in $[0,\infty)$, $X_n = \boldsymbol{\Theta}(Y_n)$ means that $\{X_n/Y_n\}$ is tight in $(0,\infty)$, and $X_n = \mathbf{o}(Y_n)$ means that X_n/Y_n converges to zero in probability. As in the deterministic case, we can add a qualifier "as $n \to \infty$" to mean that there exists $n_0 < \infty$ such that the relevant inequalities hold between X_n and Y_n provided that $n \geq n_0$. Let us stress again that, as in the deterministic case, the random variables denoted implicitly by our use of asymptotic notation are always taken to be non-negative. When we wish to apply this notation to quantities of uncertain sign we use $\pm\mathbf{O}$, $\pm\mathbf{o}$, etc. as appropriate.

Like in the deterministic case, this notation really begins to shine when forming more complicated compound expressions. Again, we warn the reader that in such an expression, the implicit random variables (e.g. those appearing in an exponent) may be different in the upper and lower bounds. Indeed this will usually be the case in our applications. To give a contrived example in which all these conventions come into force, "$X_n = \boldsymbol{\Theta}(\exp[n + \mathbf{O}((\log n)^{\mathbf{O}(1)}) \pm \mathbf{o}(\log\log n)])$ as $n \to \infty$" is equivalent to the statement that there exists $n_0 < \infty$ and sequences of non-negative random variables (A_n^-), (A_n^+), (B_n^-), (B_n^+), (C_n^-), and (C_n^+) and real-valued sequences of random variables (D_n^-) and (D_n^+) such that (A_n^-) is tight in $(0,\infty]$, (A_n^+), (B_n^-), (B_n^+), (C_n^-), and (C_n^+) are tight in $[0,\infty)$, (D_n^-) and (D_n^+) converge to zero in probability, and

$$
A_n^- e^{n + B_n^- (\log n)^{C_n^-} + D_n^- \log\log n} \leq X_n \leq A_n^+ e^{n + B_n^+ (\log n)^{C_n^+} + D_n^+ \log\log n} \qquad \text{for every } n \geq n_0.
$$

Note the incredible economy we have achieved by writing this complicated condition in the simple form "$X_n = \boldsymbol{\Theta}(\exp[n + \mathbf{O}((\log n)^{\mathbf{O}(1)}) \pm \mathbf{o}(\log\log n)])$ as $n \to \infty$"!

Remark 7. As with deterministic asymptotic notation, there are many useful elementary notational identities. Of these, we will repeatedly use that for any sequence of random variables $(X_n)_{n\geq 0}$ if $X_n = \mathbf{o}(Y_n)$ then $X_n = \mathbf{O}(Y_n)$, and if $X_n = \mathbf{O}(Y_n(\log n)^{\delta})$ for all $\delta > 0$, then $X_n = \mathbf{O}(Y_n(\log n)^{o(1)})$. Similarly, if $X_n = \boldsymbol{\Omega}(Y_n(\log n)^{-\delta})$ for all $\delta > 0$, then $X_n = \boldsymbol{\Omega}(Y_n(\log n)^{-o(1)})$.

6.1.3 Statement of results

We now state our main results. We begin with our results on the volumes of intrinsic balls, the proof of which occupies the majority of the paper.

Theorem 57 (Volume growth)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and for each $n \geq 0$ let $\mathfrak{B}(n) = \mathfrak{B}(0,n)$ denote the intrinsic ball of radius n around the origin in $\mathfrak{T}$. The volume of $\mathfrak{B}(n)$ satisfies the distributional asymptotics*

$$|\mathfrak{B}(n)| = \boldsymbol{\Theta}\left(\frac{n^2}{(\log n)^{1/3-o(1)}}\right) \qquad \textit{and} \qquad \mathbb{E}\,|\mathfrak{B}(n)| = \Theta\left(\frac{n^2}{(\log n)^{1/3-o(1)}}\right)$$

as $n \to \infty$. Moreover, letting $\Lambda(r)$ denote the ℓ^∞ ball of radius r around the origin in $\mathbb{Z}^4$ for each $r \geq 0$, we have that

$$\lim_{n\to\infty} \mathbf{P}\left(\mathfrak{B}(n) \subseteq \Lambda\left(n^{1/2}(\log n)^{1/6+\delta}\right)\right) = 1$$

for every $\delta > 0$.

Recall that in high dimensions the components of the uniform spanning forest have quadratic volume growth $|\mathfrak{B}(n)| = \boldsymbol{\Theta}(n^2)$ [48, 193], so that the behaviour in four dimensions differs from the high-dimensional behaviour by a polylogarithmic factor as expected.

The proofs of both the upper and lower bounds of Theorem 57 rely on Wilson's algorithm [68, 321] to express properties of the tree in terms of properties of loop erased random walks. Accordingly, they also both rely on an understanding of the behaviour of the loop-erased random walk in four dimensions developed in [229, 233, 235], with the proof of the lower bounds also relying on the control of the *capacity* of the loop-erased walk developed in [30, 197]. The proof of the upper bound also uses a generalisation of the method of *typical times* introduced in [197], a very useful technical tool that allows us to circumvent several issues that arise from the fact that the length of a four-dimensional loop-erased random walk is only very weakly concentrated. (The use of this machinery is also responsible for the presumably unnecessary subpolylogarithmic $(\log n)^{\pm o(1)}$ errors appearing throughout our results.)

We now turn to our results concerning the *random walk* on the four-dimensional UST. We write $\mathbb{P}$ and $\mathbb{E}$ for probabilities and expectations taken with respect to the joint law of the UST $\mathfrak{T}$ on $\mathbb{Z}^4$ and the random walk $X = (X_n)_{n\geq 0}$ on $\mathfrak{T}$ started at the origin, and write $\mathbf{P}^{\mathfrak{T}}$ and $\mathbf{E}^{\mathfrak{T}}$ for probabilities and expectations taken with respect to the conditional law of X given $\mathfrak{T}$. We write $p_n^{\mathfrak{T}}(x,y)$ for the transition probabilities of a random walk on the uniform spanning tree $\mathfrak{T}$ conditional on $\mathfrak{T}$, write τ_n for the time taken for the random walk to hit the complement of the intrinsic ball of radius of n, and write $d_{\mathfrak{T}}$ for the intrinsic distance on $\mathfrak{T}$.

Theorem 58 (Random walk asymptotics)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and let $X = (X_n)_{n\geq 0}$ be the simple random walk on $\mathfrak{T}$ started at the origin. The following distributional asymptotic expressions hold as $n \to \infty$:*

$$\text{Intrinsic displacement :} \qquad d_{\mathfrak{T}}(X_0, X_n),\ \max_{0\leq i\leq n} d_{\mathfrak{T}}(X_0, X_n) = \boldsymbol{\Theta}\left(n^{\frac{1}{3}} (\log n)^{\frac{1}{9}-o(1)} \right) \tag{6.1}$$

$$\text{Extrinsic displacement :} \qquad \max_{0\leq i\leq n} \|X_i\|_\infty = \boldsymbol{\Theta}\left(n^{\frac{1}{6}} (\log n)^{\frac{2}{9}+o(1)} \right) \tag{6.2}$$

$$\text{Return probabilities :} \qquad p_{2n}^{\mathfrak{T}}(0,0) = \boldsymbol{\Theta}\left(\frac{1}{n^{\frac{2}{3}}} (\log n)^{\frac{1}{9}-o(1)} \right) \tag{6.3}$$

$$\text{Range :} \qquad \#\{X_m : 0 \leq m \leq n\} = \boldsymbol{\Theta}\left(n^{\frac{2}{3}} \frac{1}{(\log n)^{\frac{1}{9}\pm o(1)}} \right) \tag{6.4}$$

$$\text{Hitting times :} \qquad \tau_n, \mathbf{E}^{\mathfrak{T}}[\tau_n] = \boldsymbol{\Theta}\left(n^3 \frac{1}{(\log n)^{\frac{1}{3}-o(1)}} \right) \tag{6.5}$$

Remark 8. It is reasonably straightforward to adapt the proofs of [193] to prove that, in four dimensions, all the quantities we consider here satisfy Alexander-Orbach asymptotics up to $(\log n)^{\pm O(1)}$ factors. Identifying the *correct* powers of log is significantly more difficult and is the primary contribution of this paper.

As mentioned above, the behaviour of the random walk on the uniform spanning tree has previously been studied in dimensions $d = 2$ [44, 51] $d = 3$ [27], and $d \geq 5$ [193], with the two cases $d = 2$ and $d = 3$ presenting unique challenges that are largely distinct from those associated to the critical dimension $d = d_c = 4$ considered here. While we are the first to study the polylogarithmic corrections to the volume of balls and the behaviour of random walks on the UST at $d = 4$, our work builds upon the substantial literature studying other aspects of the 4d UST, the highlights of which include [197, 229, 232, 233, 235, 301]. Our work is influenced most strongly by the recent work of Sousi and the second author [197];

we rely on both the results proven and the techniques developed in that paper in numerous ways.

Following Kumigai-Misumi [227], which collects and generalises results of [43, 49, 50], estimates of the form proven in Theorem 58 can all be deduced from the volume growth estimates of Theorem 57 together with estimates on the *effective resistance* between the origin and the boundary of a ball in the tree. The relevant effective resistance estimates will in turn be deduced from Theorem 57 together with the asymptotics of the *intrinsic arm probability* computed in [197]. We let $\mathscr{R}_{\mathrm{eff}}(A \leftrightarrow B; G)$ denote the effective resistance between sets $A, B \subseteq V[G]$ in the graph G, where we assign unit resistance to each edge $e \in E[G]$, so that if $\deg_{\mathfrak{T}}(0)$ denotes the degree of 0 in $\mathfrak{T}$ then $\mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow \partial\mathfrak{B}(0,n); \mathfrak{T})^{-1} := \deg_{\mathfrak{T}}(0)\mathbf{P}^{\mathfrak{T}}($hit $\partial\mathfrak{B}(0,n)$ before returning to 0$)$. Background on effective resistances can be found in e.g. [226, 259].

Theorem 59 (Effective resistance)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and for each $n \geq 0$ let $\partial\mathfrak{B}(n) = \partial\mathfrak{B}(0,n)$ denote the set of vertices with distance exactly n from the the origin in $\mathfrak{T}$. Then*

$$\mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow \partial\mathfrak{B}(0,n); \mathfrak{T}) = n(\log n)^{-\mathbf{o}(1)}$$

as $n \to \infty$.

Note that the linear upper bound $\mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow \partial\mathfrak{B}(0,n)) \leq n$ is trivial and holds for any graph. Together with existing results in other dimensions [27, 44, 193], Theorem 59 shows that the UST has (approximately) linear effective resistance growth in *every* dimension. As will be clear from the proof, this is a consequence of the scaling relation

$$\mathbf{P}(\text{the past of the origin has intrinsic diameter} \geq n) \approx \frac{n}{\text{typical volume of an intrinsic } n\text{-ball}}, \tag{6.6}$$

which also holds in every dimension. Here, the **past** of the origin is the union of the origin and the finite connected component of the UST left when the origin is deleted; estimating the probability that the past is large in various senses is the main subject of [197], which in particular establishes up-to-constants estimates on the left hand side of (6.6). Currently, however, there is no direct proof of this scaling relation, which in four dimensions is verified only by computing the two sides separately in [197] and the present paper. It would be very interesting to have a direct and general proof of this relation in all dimensions that worked without computing either quantity.

While Theorems 57 and 59 are sufficient to compute the exact logarithmic corrections to the asymptotic properties of the random walk on the UST using the methods of [48, 227] as discussed above, we will also show that a significantly stronger bound on the displacement of the random walk can be proven using the Markov-type method pioneered in the work of Lee and coauthors [127, 149, 244, 245].

Theorem 60 (Sharp upper bounds on the mean-squared displacement)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and let $X = (X_n)_{n\geq 0}$ be the simple random walk on $\mathfrak{T}$ started at the origin. Then*

$$\mathbb{E}\left[\max_{0\leq i\leq n} d_{\mathfrak{T}}(X_0,X_i)^2\right] \preceq n^{2/3}(\log n)^{2/9} \tag{6.7}$$

for every $n \geq 2$.

The specific argument used to prove this theorem is inspired closely by the work of Ganguly and Lee [149]. Briefly, the idea is to use the universal Markov-type inequality for weighted metrics on trees [127] to prove a *diffusive* upper bound for the random walk with respect to a modified metric supported only on vertices of the tree whose past has large intrinsic diameter, then deduce the desired subdiffusive estimate in the original metric. The tail bounds on the intrinsic diameter of the past of the origin proven in [197] are precisely what is needed to carry this argument through. In particular, the proof of Theorem 60 does *not* rely on Theorem 57 or the theory of typical times, allowing us to avoid the $(\log n)^{\pm o(1)}$ present in the statement of that theorem.

Remark 9. As discussed in [49, Example 2.6], although the *typical* displacement of the random walk can always be controlled in terms of volume growth and resistance growth, it is possible in general for the displacement not to be uniformly integrable, so that its mean grows significantly faster than its median. As such, the second moment estimate provided by Theorem 60 is significantly stronger than what can be deduced directly from Theorems 57 and 59 by the techniques of [49, 227].

6.2 Intrinsic volume growth

In this section we prove Theorem 57. The upper and lower bounds of the theorem, which use completely different techniques, are proven in Sections 6.2.1 and 6.2.2 respectively. Both parts of the proof will utilize the connections between the uniform spanning tree and the loop-erased random walk implied by Wilson's algorithm, and so to proceed we must provide notation for the loop-erased random walk and some related quantities.

Loop-erased random walk. For each $-\infty \leq n \leq m \leq \infty$, let $L(n,m)$ be the graph with vertex set $\{i \in \mathbb{Z} : n \leq i \leq m\}$ and edge set $\{\{i,i+1\} : n \leq i \leq m-1\}$. A path is then a multigraph homomorphism from $L(n,m)$ to the hypercubic lattice $\mathbb{Z}^4$ for some $-\infty \le n \le m \le \infty$. We write $w_i = w(i)$ for the vertex visited at time i. For $n \le b \le m$, we write w^b for the restriction of w to $[n,b]$, and call w^b the path **stopped at** b. In particular, given a random walk X, we will often use the notation X^T for a random walk stopped at some possibly random time T. A path is said to be **transient** if it visits every vertex of $\mathbb{Z}^4$ at most finitely many times. In particular, finite paths are always transient. Given a transient path $w : L(0,m) \to \mathbb{Z}^4$, we recursively define the sequence of times $\ell_n(w)$ by $\ell_0(w) = 0$, and

$$\ell_{n+1}(w) = 1 + \max\{k : w_k = w_{\ell_n}\},$$

where we terminate the sequence the first time $\max\{k : w_k = w_{\ell_n}\} = m$ when $m < \infty$. The loop-erasure of w is then the path induced by the sequence of neighbouring vertices

$$\mathrm{LE}(w)_i = w_{\ell_i(w)}.$$

We will also need the quantity

$$\rho_n(w) = \max\{m \ge 0 : \ell_m(w) \ge n\},$$

which for each $n \ge 0$ counts the number of points up to time n which are not erased when computing the loop-erasure of w, so that $(\ell_n)_{n\ge 0}$ and $(\rho_n)_{n\ge 0}$ are inverses of each other in the sense that

$$\ell_n(w) \le m \quad \text{if and only if} \quad \rho_m(w) \ge n,$$

for every $n,m \ge 0$.

The loop-erasure of a simple random walk is known as the **loop-erased random walk**. The theory of loop-erased random walk was both introduced and developed extensively by Lawler [231], whose results on the four-dimensional loop-erased random walk [233, 235], including his joint work with Sun and Wu [229], play an extensive role in this paper both directly and through inputs to [197]. Given a random walk X, we will usually abbreviate $\ell_n = \ell_n(X)$ and $\rho_n = \rho_n(X)$. It will also be convenient to define the notation

$$\mathrm{LE}_\infty(X^n) := \mathrm{LE}(X)^{\rho_n}$$

for $n \ge 0$, giving the component of the infinite loop erasure $\mathrm{LE}(X)$ which is contributed by the first n steps of the random walk X. We emphasise that the brackets of $\mathrm{LE}_\infty(X^n)$ do *not* indicate that $\mathrm{LE}_\infty(X^n)$ is a function of just X^n. The following concentration estimates of

Lawler [233, 235], as stated in [197, Theorem 2.2], will be used repeatedly throughout the the paper.

Theorem 61 ([197], Theorem 2.2)**.** *Let X be a simple random walk on $\mathbb{Z}^4$, then*

$$\mathbf{P}\left(\left|\frac{\rho_n}{n(\log n)^{-1/3}}-1\right|>\varepsilon\right)\preceq_\varepsilon \frac{\log\log n}{(\log n)^{2/3}} \qquad \textit{and hence}$$
$$\mathbf{P}\left(\left|\frac{\ell_n}{n(\log n)^{1/3}}-1\right|>\varepsilon\right)\preceq_\varepsilon \frac{\log\log n}{(\log n)^{2/3}},$$

for every $\varepsilon>0$ and $n\geq 3$.

Wilson's algorithm rooted at infinity [68, 321] allows us to build a sample of the UST of $\mathbb{Z}^4$ (or any other transient graph) out of loop-erased random walks. This algorithm is very important to most analyses of the UST. We will assume that the reader is already familiar with Wilson's algorithm, referring them to e.g. [259] for background otherwise.

Finally, let us introduce notation concerning the geometry of $\mathbb{Z}^4$ and the tree $\mathfrak{T}$. We write $\|x\|$ for the ℓ^∞ norm of $x\in\mathbb{Z}^4$ and write $\Lambda(x,r)$ for the ℓ^∞ ball around $x\in\mathbb{Z}^d$ of radius r. For convenience, we will write $\Lambda(r)$ for $\Lambda(0,r)$. For each $x\in\mathbb{Z}^4$ and $r\geq 1$, $\mathfrak{B}(x,r)$ will denote the intrinsic ball of radius r around x in $\mathfrak{T}$, with $\mathfrak{B}(r):=\mathfrak{B}(0,r)$. For each pair of vertices $x,y\in\mathbb{Z}^4$ we write $\Gamma(x,y)$ for the unique simple path between x and y in $\mathfrak{T}$, which is well-defined since the UST of $\mathbb{Z}^4$ is a.s. connected [68, 288], and write $\Gamma(x,\infty)$ for the **future** of x in $\mathfrak{T}$, i.e. the unique infinite simple path in $\mathfrak{T}$ with x as an endpoint, which is well-defined since the UST of $\mathbb{Z}^4$ is one-ended a.s. [68, 288]. Given two vertices $x,y\in\mathbb{Z}^4$ we will denote by $x\vee y=y\vee x$ the unique point at which the futures of x and y in $\mathfrak{T}$ first intersect.

The **past** of a vertex v in the uniform spanning tree $\mathfrak{T}$, denoted[3] $\mathfrak{P}(v)$, is the union of the vertex and the finite components that are disconnected from infinity when the vertex is deleted from $\mathfrak{T}$. We write $\mathfrak{P}(v,n)$ for $\mathfrak{P}(v)\cap\mathfrak{B}(v,n)$ and write $\partial\mathfrak{B}(v,n)$ for the set of vertices in $\mathfrak{T}$ at intrinsic distance exactly n from v. Further discussion of the basic topological features of the UST used here can be found in [259, Chapter 10].

6.2.1 Upper bounds

In this section we prove the following two propositions, which establish the upper bounds of Theorem 57. Throughout this section we will write $\asymp$, $\preceq$, and $\succeq$ with subscripts such as δ and p to mean that the implicit constants are allowed to depend on these parameters.

[3]This character is \mathfrak{P}.

Proposition 62. *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$. Then*

$$\mathbb{E}|\mathfrak{B}(n)| = O\left(\frac{n^2}{(\log n)^{1/3-o(1)}}\right)$$

as $n \to \infty$.

Proposition 63. *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and let $\delta > 0$. Then*

$$\mathbb{P}\left(\mathfrak{B}(n) \not\subseteq \Lambda\left(n^{1/2}(\log n)^{1/6+\delta}\right)\right) \preceq_\delta \frac{\log\log n}{(\log n)^{2/3}}$$

for every $n \geq 3$.

Both of these results will be proven using the following supporting technical proposition, which bounds in expectation the amount of the volume of intrinsic balls which come from paths of atypical diameter.

Proposition 64. *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$, let $\delta > 0$ and let $p \geq 1$. Then*

$$\mathbb{E}\left|\left\{x \in \mathfrak{B}(n) : \Gamma(0,x) \not\subseteq \Lambda\left(n^{1/2}(\log n)^{1/6+\delta}\right)\right\}\right| \preceq_{p,\delta} \frac{n^2}{(\log n)^p}$$

for every $n \geq 2$.

The expected intrinsic volume bound of Proposition 62 follows immediately from Proposition 64 together with [197, Proposition 7.3], which provides a tight upper bound on the number of points connected to the origin inside an *extrinsic* box of a given radius.

Proposition 65 ([197], Proposition 7.3)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$. Then*

$$\mathbb{E}\left|\{x \in \mathbb{Z}^4 : \Gamma(0,x) \subseteq \Lambda(r)\}\right| \preceq \frac{r^4}{\log r}$$

for every $r \geq 2$.

Proof of Proposition 62. Fix $\delta > 0$, $p \geq 1$ and $n \geq 4$. We have trivially that

$$\mathbb{E}|\mathfrak{B}(n)| \leq \mathbb{E}\left|\{x \in \mathfrak{B}(n) : \Gamma(0,x) \not\subset \Lambda(n^{1/2}(\log n)^{1/6+\delta})\}\right| + \mathbb{E}\left|\{x \in \mathbb{Z}^d : \Gamma(0,x) \subseteq \Lambda(n^{1/2}(\log n)^{1/6+\delta})\}\right|.$$

Applying Proposition 64 to the first term on the right hand side and Proposition 65 to the second yields that

$$\mathbb{E}|\mathfrak{B}(n)| \preceq_{p,\delta} \frac{n^2}{(\log n)^{1/3-2\delta}} + \frac{n^2}{(\log n)^p},$$

which implies the claim since $\delta > 0$ and $p \geq 1$ were arbitrary. □

The deduction of Proposition 63 from Proposition 64 requires a more involved argument using further results of [197] and is given after the proof of Proposition 64.

To prove Proposition 64 we need to be able to relate balls in the extrinsic metric (i.e. the ℓ^∞ metric on $\mathbb{Z}^4$) to balls in the intrinsic metric. Intuitively, since paths in the UST are distributed as loop-erased random walks and since length-n loop-erased random walks in $\mathbb{Z}^4$ are typically generated by simple random walks of length roughly $n(\log n)^{1/3}$ [235], we expect that intrinsic paths of length n in the UST should have extrinsic diameter concentrated around $n^{1/2}(\log n)^{1/6}$. Unfortunately, however, the concentration estimates that are available for the length of loop-erased random walks are far too weak to directly rule out that most of the volume of the intrinsic ball comes from paths of atypically large diameter. We circumvent this problem using a generalization of the *typical time* methodology of [197, Section 8], originally introduced to prove tail estimates on the extrinsic radius of the past of the origin: we will use typical times to subsume balls in the intrinsic metric by balls of an appropriate radius in the extrinsic metric.

Typical times. We now detail the generalised typical time methodology that we use. Given points $x, y \in \mathbb{Z}^4$ and a simple path γ starting at x and ending at y, let X be a random walk started conditioned to hit y and to have loop erasure γ when it first hits y. Roughly speaking, the typical time $T(\gamma)$ of γ is defined to be the typical length of the walk X under this conditional distribution; an important part of the theory is that this length is concentrated around the typical time $T(\gamma)$ under mild conditions on the path γ. Our proofs will apply a slight generalization of this notion, which we now introduce. Instead of stopping the walk at a single point y, we introduce disjoint sets $A, B \subset \mathbb{Z}^d$ and define the (A,B)**-typical time** $T_{A,B}(\eta)$ of a simple path γ starting at x, ending when it first hits A, and avoiding B to be

$$T_{A,B}(\gamma) := \mathbf{E}_x\left[\sum_{i=1}^{|\gamma|}\Big(\ell_i(X^{\tau_A}) - \ell_{i-1}(X^{\tau_A})\Big) \wedge |\gamma| \;\middle|\; \tau_A < \infty, \tau_A < \tau_B, \mathrm{LE}(X^{\tau_A}) = \gamma\right],$$

where $\mathbf{E}_x$ denotes expectation with respect to the law of a simple random walk X on $\mathbb{Z}^4$ started at $X_0 = x$, and where the times $\ell_i(X^{\tau_A})$ are from the definition of the loop-erasure of

X^{τ_A}, so that

$$\tau_A = \ell_{|\gamma|}(X^{\tau_A}) = \sum_{i=1}^{|\gamma|}\Big(\ell_i(X^{\tau_A}) - \ell_{i-1}(X^{\tau_A})\Big)$$

when $\mathrm{LE}(X^{\tau_A}) = \gamma$. We will use boldface to denote probabilities and expectations taken with respect to the law of a simple random walk throughout the paper, so that $\mathbf{P}_x$ will denote probability with respect to the law of a simple random walk started at time 0 at vertex x. We remark that for paths γ which hit A and avoid B we have that $T_{A,B}(\gamma) = T_{A\cup B,\emptyset}(\gamma)$, where we define $\tau_\emptyset = \infty$, and that the usual typical time as defined in [197] is given by $T(\eta) = T_{\{\eta_n\},\emptyset}(\eta)$ when η has length n.

The following Lemma extends [197, Lemma 8.2] to (A,B)-typical times. The proof is identical to the proof of that lemma and is omitted.

Lemma 66. *There exists a constant C such that if $x \in \mathbb{Z}^4$, A,B are disjoint subsets of $\mathbb{Z}^4$, and γ is a simple path of length $n \geq 0$ from x to A which does not intersect B, then*

$$\mathbf{P}_x\Big(|\tau_A - T_{A,B}(\gamma)| > \lambda n \;\Big|\; \tau_A < \infty, \tau_A < \tau_B, \mathrm{LE}(X^{\tau_A}) = \gamma\Big) \leq \frac{C}{\lambda},$$

for every $\lambda \geq 1$.

As explained in detail in [197, Section 8], for most paths of interest the typical time $T(\gamma)$ is significantly larger than $|\gamma|$, so that Lemma 66 can indeed be thought of as a concentration estimate, justifying the use of the 'typical time' terminology. Indeed, when γ is a loop-erased random walk of length n its typical time will usually be of order $n(\log n)^{1/3}$. For an arbitrary path γ of length $n \geq 1$ the best bounds are of the form

$$n \preceq T(\gamma) \preceq n\log(n+1); \tag{6.8}$$

the lower bound is trivial while the upper bound follows by bounding the distribution of the length of the loop $\ell_i(X^{\tau_A}) - \ell_{i-1}(X^{\tau_A})$ by that of the length of an unconditioned simple random walk loop in $\mathbb{Z}^4$. The upper bound is sharp when γ is a straight line, while the lower bound is sharp when γ is a space-filling curve.

As in [197], we bound typical times by a simpler functional that is easier to work with. If γ has length n, we define $A_i(\gamma) = \sum_{k=1}^n \frac{1}{k}\mathrm{Esc}_k(\gamma^i)^2$, where given a finite path η of length m and an integer $k \geq 1$ the k-step **escape probability** $\mathrm{Esc}_k(\eta)$ is defined by $\mathrm{Esc}_k(\eta) = \mathbf{P}_{\eta_m}(X^k \cap \eta^{m-1} = \emptyset)$. We then define

$$\widetilde{T}(\gamma) := \sum_{i=0}^{n-1} A_i(\gamma).$$

It follows from the same calculations used to derive the analogous bound for the ordinary hitting time on [197, Page 69] that

$$\widetilde{T}(\gamma) \succeq T_{A,B}(\gamma) \tag{6.9}$$

for every path γ and every pair of disjoint sets $A,B \subseteq \mathbb{Z}^4$. For a given $0 < \delta \leq 1$, we say that a finite path γ of length $n \geq 0$ is **δ-good** if

$$\sum_{i=0}^{n-1} A_i(\gamma)\mathbb{1}\big(A_i \geq (\log n)^{1/3+\delta}\big) \leq \delta n,$$

and say it is **δ-bad** otherwise. If γ is a δ-good path of length $n \geq 2$, then

$$\widetilde{T}(\gamma) \leq \delta n + \sum_{i=0}^{n-1} A_{i,n}(\gamma)\mathbb{1}\big(A_i < (\log n)^{1/3+\delta}\big) \preceq n(\log n)^{1/3+\delta}. \tag{6.10}$$

We will apply [197, Lemma 8.5], which is based on the work of Lawler, Sun, and Wu [229], and states that the loop erasure of a random walk is highly unlikely to be bad.

Lemma 67 ([197], Lemma 8.5)**.** *Let $\delta > 0$ and $p \geq 0$ and let X be simple random walk on $\mathbb{Z}^4$. Then*

$$\frac{1}{n}\sum_{k=0}^{n} \mathbf{P}_0\Big(\mathrm{LE}(X^k) \textit{ is } \delta\textit{-bad}\Big) \preceq_{\delta,p} \frac{1}{(\log n)^p},$$

for every $n \geq 2$.

We now apply this machinery to prove Proposition 64. We will also use the mass-transport principle for $\mathbb{Z}^4$, which states that if $f : \mathbb{Z}^d \times \mathbb{Z}^d \to [0,\infty]$ is a diagonally invariant function, meaning that $f(x,y) = f(x+z,y+z)$ for every $x,y,z \in \mathbb{Z}^4$, then $\sum_{x\in\mathbb{Z}^d} f(0,x) = \sum_{x\in\mathbb{Z}^d} f(-x,0) = \sum_{x\in\mathbb{Z}^d} f(x,0)$.

Proof of Proposition 62. To prove the proposition, we will show that if $\mathscr{A}$ is any set of simple paths γ with $\gamma_0 = 0$ and with length $|\gamma| \leq n$, then

$$\begin{aligned}\sum_{v\in\mathbb{Z}^4} \mathbb{P}\Big(\Gamma(0,v) \in \mathscr{A},\ \text{and } \Gamma(0,0\wedge v) \subseteq \Lambda(0\wedge v, r)\Big) \\ \preceq_{\delta,p} \sum_{v\in\mathbb{Z}^4} \mathbb{P}\Big(\Gamma(0,v) \in \mathscr{A}, \Gamma(0,0\wedge v) \subseteq \Lambda(0\wedge v,r) \text{ and} \\ \Gamma(v,0\wedge v) \subseteq \Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big) + n^2(\log n)^{-p}\end{aligned} \tag{6.11}$$

for every $0 < \delta \leq 1$, $p \geq 1$, and $n,r \geq 2$. Before proving (6.11), let us first see how it implies the proposition. We must first define some notation. Given a finite path $\gamma = (\gamma_0,\ldots,\gamma_{|\gamma|})$ and a vector x, we define $\gamma+x=(\gamma_0+x,\ldots,\gamma_{|\gamma|}+x)$, and $\gamma^{\leftarrow}=(\gamma_{|\gamma|},\ldots,\gamma_0)$. We extend these operations to sets of paths in the obvious way. Fix $\delta\in(0,1]$, $p\geq 1$ and define the two sets of paths

$$\begin{aligned}\mathscr{A}_0 &= \{\gamma : \gamma \text{ simple}, \gamma_0=0, |\gamma|\leq n, \gamma\nsubseteq \Lambda(0, \quad n^{1/2}(\log n)^{1/6+2\delta})\},\\ \mathscr{A}_0' &= \{\gamma : \gamma \text{ simple}, \gamma_0=0, |\gamma|\leq n, \gamma\nsubseteq \Lambda(\gamma_{|\gamma|}, n^{1/2}(\log n)^{1/6+2\delta})\}.\end{aligned}$$

For any $x\in\mathbb{Z}^d$, writing $\mathscr{A}_0(x)$ for the set of paths $\mathscr{A}_0+x$, we observe that for any path γ with $\gamma_0=x$, $\gamma_{|\gamma|}=0$, we have that

$$\gamma\in\mathscr{A}_0(x) \iff \gamma^{\leftarrow}\in\mathscr{A}_0'. \tag{6.12}$$

With this notation and observation in hand, setting $\mathscr{A}=\mathscr{A}_0$ in (6.11) and taking $r\uparrow\infty$, we get

$$\begin{aligned}\mathbb{E}\left|\{x\in\mathbb{Z}^d : \Gamma(0,x)\in\mathscr{A}_0\}\right| &= \sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0\Big)\\ &\preceq_{\delta,p} \sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0(0) \text{ and } \Gamma(v,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big)+n^2(\log n)^{-p}\\ &= \sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(v,0)\in\mathscr{A}_0(v) \text{ and } \Gamma(0,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big)+n^2(\log n)^{-p}\\ &= \sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0' \text{ and } \Gamma(0,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big)+n^2(\log n)^{-p}\end{aligned} \tag{6.13}$$

for every $0<\delta\leq 1$, $p\geq 1$, and $n\geq 2$, where the second equality follows by an application of the mass-transport principle to exchange the roles of 0 and v, and the third equality follows by (6.12). Applying (6.11) a second time with $\mathscr{A}=\mathscr{A}_0'$ then yields that

$$\begin{aligned}\mathbb{E}\left|\{x\in\mathbb{Z}^d : \Gamma(0,x)\in\mathscr{A}_0\}\right| \preceq_{\delta,p} &\sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0',\\ &\text{and } \Gamma(0,0\wedge v),\Gamma(v,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big)+n^2(\log n)^{-p},\end{aligned}$$

and hence, applying mass-transport a second time,

$$\begin{aligned}
\mathbb{E}\left|\{x\in\mathbb{Z}^d : \Gamma(0,x)\in\mathscr{A}_0\}\right| &\preceq_{\delta,p} n^2(\log n)^{-p} + \\
&\sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0,\ \text{and}\ \Gamma(0,0\wedge v),\Gamma(v,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})\Big)\\
&\preceq_{\delta}\quad n^2(\log n)^{-p} + \sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A}_0,\ \text{and}\ \Gamma(0,v)\subseteq\Lambda(2n^{1/2}(\log n)^{1/6+\delta})\Big).
\end{aligned}$$

If n is sufficiently large that $(\log n)^{\delta} > 2$ then the second term is zero and the claim follows.

It remains to prove (6.11). Fix $0<\delta\le 1, p\ge 1$ and $n,r\ge 2$. Let η be the future of the origin in $\mathfrak{T}$ and write $\mathbb{P}^{\eta}$ and $\mathbb{E}^{\eta}$ for probabilities and expectations taken with respect to the conditional law of $\mathfrak{T}$ given η. Let $\mathscr{I}=\{i\in\{0,\ldots,n\} : \eta[0,i]\subseteq\Lambda(\eta_i,r)\}$, and for any $i\ge 0$ define the restriction $\mathscr{A}|_{x,\eta}^{i}$ to be the set of finite simple paths

$$\mathscr{A}|_{x,\eta}^{i} = \{\gamma : \gamma_0=\eta_i,\ \gamma[1,|\gamma|]\cap\eta=\emptyset,\ \gamma_{|\gamma|}=x,\ \eta[0,i]\oplus\gamma[1,|\gamma|]\in\mathscr{A}\},$$

where for any two finite paths γ, γ', we have $(\gamma_0,\ldots,\gamma_{|\gamma|})\oplus(\gamma'_0,\ldots,\gamma'_{|\gamma'|})=(\gamma_0,\ldots,\gamma_{|\gamma|},\gamma'_0,\ldots,\gamma'_{|\gamma'|})$. In other words, $\mathscr{A}|_{x,\eta}^{i}$ is the set of simple paths (including paths of just a single vertex) beginning at η_i, avoiding the other points of η, and which when concatenated to $\eta[0,i-1]$ yield a path in $\mathscr{A}$ ending at x.

For each $v\in\mathbb{Z}^d$ we can sample from the conditional distribution of the path in $\mathfrak{T}$ connecting v to η using Wilson's algorithm by starting a random walk X and v and loop erasing it when it first hits η. When sampling the path in this manner we have that the event $\{\Gamma(0,v)\in\mathscr{A}_n$ and $\Gamma(0,0\wedge v)\subseteq\Lambda(0\wedge v,r)\}$ occurs if and only if the union of disjoint events

$$\bigcup_{i\in\mathscr{I}}\{\tau_i<\tau_i^c,\ \mathrm{LE}(X^{\tau_i})^{\leftarrow}\in\mathscr{A}|_{v,\eta}^{i}\}$$

occurs, where we write τ_i for the hitting time of η_i and write τ_i^c for the hitting time of $\eta\setminus\{\eta_i\}$, so that

$$\begin{aligned}
\sum_{v\in\mathbb{Z}^4}\mathbb{P}\Big(\Gamma(0,v)\in\mathscr{A},\ \text{and}\ \Gamma(0,0\wedge v)\subseteq\Lambda(0\wedge v,r)\Big)&\\
=\sum_{i\in\mathscr{I}}\sum_{v\in\mathbb{Z}^4}\mathbf{P}_v\big(\tau_i<\tau_i^c,\ \mathrm{LE}(X^{\tau_i})^{\leftarrow}\in\mathscr{A}|_{v,\eta}^{i}\big).&\quad (6.14)
\end{aligned}$$

We remark that the probabilities on the right hand side of (6.14) are themselves random variables given that τ_i, τ_i^c and $\mathscr{A}|_{v,\eta}^{i}$ depend on η. (The law of the simple random walk

$\mathbf{P}_v$ does not depend on η and, since there is no possible ambiguity that $\mathbf{P}$ could denote expectation over the UST, we have chosen to made the dependence implicit.)

Temporarily fixing $i \in \mathscr{I}$, we analyze the inner summation on the right hand side of (6.14) using the union bound

$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v\big(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i\big) \\ \leq \sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-good}) \\ + \sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-bad}). \quad (6.15)$$

If $\mathrm{LE}(X^{\tau_i})$ is δ-good then we have by (6.9) and (6.10) that $T_i := T_{\eta_i,\eta\setminus\{\eta_i\}}(|\mathrm{LE}(X^{\tau_i})|) \leq C_1 n(\log n)^{1/3+\delta}$ for some universal constant C_1, and hence that

$$\mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-good}) \\ \leq \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, T_i \leq C_1 n(\log n)^{1/3+\delta}) \\ \leq \mathbf{P}_v\big(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, |T_i - \tau_i| \geq \lambda n\big) \\ + \mathbf{P}_v\big(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \tau_i \leq C_1 n(\log n)^{1/3+\delta} + \lambda n\big) \quad (6.16)$$

for every $\lambda > 0$. The first term on the right hand side of (6.16) is bounded above by $C_2\lambda^{-1}\mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i)$ for some universal constant C_2 by Lemma 66, so that taking $\lambda = 2C_2$, substituting (6.16) into (6.15) and rearranging yields that

$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v\big(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i\big) \\ \leq 2\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \tau_i \leq C_3 n(\log n)^{1/3+\delta}) \\ + 2\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-bad}), \quad (6.17)$$

where $C_3 = C_3(\delta)$ has been chosen so that $C_1 n(\log n)^{1/3+\delta} + 2C_2 n \leq C_3 n(\log n)^{1/3+\delta}$ for every $n \geq 2$.

We next bound the second term on the right hand side of (6.17). Since the typical time of a length n path is always $O(n \log n)$, it follows by the same argument used to derive (6.17) from (6.16) that there exists a constant C_4 such that

$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-bad}) \\ \leq 2\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i,\ \tau_i \leq C_4 n\log n).$$

Thus, taking a union bound over the possible values of τ_i, we have that

$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-bad}) \\ \leq 2\sum_{v\in\mathbb{Z}^4} \sum_{k=0}^{\lceil C_4 n\log n\rceil} \mathbf{P}_v(X_k = \eta_i, \mathrm{LE}(X^k)\ \delta\text{-bad})$$

and hence by symmetry that

$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i, \mathrm{LE}(X^{\tau_i})\ \delta\text{-bad}) \\ \leq 2\sum_{v\in\mathbb{Z}^4} \sum_{k=0}^{\lceil C_4 n\log n\rceil} \mathbf{P}_{\eta_i}(X_k = v, \mathrm{LE}(X^k)\ \delta\text{-bad}) = 2\sum_{k=0}^{\lceil C_4 n\log n\rceil} \mathbf{P}_{\eta_i}(\mathrm{LE}(X^k)\ \delta\text{-bad}) \\ \preceq_{\delta,p} n(\log n)^{1-p} \quad (6.18)$$

for every $n \geq 2$.

Next, we consider the first term on the right hand side of (6.17). We write $B = \{\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|_{v,\eta}^i,\ \tau_i \leq C_3 n(\log n)^{1/3+\delta}\}$ and wish to estimate $\sum_{v\in\mathbb{Z}^4}\mathbf{P}_v(B)$. To do this, we split the event B according to how far the walk travels before hitting η_i, yielding the union bound

$$\mathbf{P}_v(B) \leq \mathbf{P}_v\big(B, \sup_{0\leq m\leq \tau_i} \|X_m - \eta_i\| \geq n^{1/2}(\log n)^{1/6+\delta}\big) \\ + \mathbf{P}_v\big(B, \sup_{0\leq m\leq \tau_i} \|X_m - \eta_i\| < n^{1/2}(\log n)^{1/6+\delta}\big). \quad (6.19)$$

For the first of these terms, we bound

$$\mathbf{P}_v\big(B, \sup_{m\le\tau_i}\|X_m-\eta_i\|\ge n^{1/2}(\log n)^{1/6+\delta}\big)$$
$$\le \sum_{k=0}^{\lceil C_3 n(\log n)^{1/3+\delta}\rceil}\mathbf{P}_v(X_k=\eta_i,\sup_{m\le k}\|X_m-\eta_i\|\ge n^{1/2}(\log n)^{1/6+\delta}).$$

Summing over v and using time-reversal gives that

$$\begin{aligned}\sum_{v\in\mathbb{Z}^4}\mathbf{P}_v\big(B, \sup_{m\le\tau_i}\|X_m-\eta_i\|\ge n^{1/2}(\log n)^{1/6+\delta}\big)&\\ \le \sum_{k=0}^{\lceil C_3 n(\log n)^{1/3+\delta}\rceil}\mathbf{P}_{\eta_i}\left(\sup_{m\le k}\|X_m-\eta_i\|\ge n^{1/2}(\log n)^{1/6+\delta}\right)&\\ \preceq n(\log n)^{1/3+\delta}\mathbf{P}_0\left(\sup_{m\le\lceil C_3 n(\log n)^{1/3+\delta}\rceil}\|X_m\|\ge n^{1/2}(\log n)^{1/6+\delta}\right)&\\ \preceq n(\log n)^{1/3+\delta}e^{-c_1(\log n)^\delta}\preceq_{\delta,p} n(\log n)^{-p}& \qquad (6.20)\end{aligned}$$

for some constant $c_1>0$, where the first inequality in the last line follows by e.g. the maximal version of Azuma-Hoeffding [269, Section 2].

Substituting the estimates (6.18) and (6.20) into (6.17) in light of (6.19) yields that there exists a constant $C_{\delta,p}$ such that

$$\begin{aligned}\sum_{v\in\mathbb{Z}^4}\mathbf{P}_v(\tau_i<\tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow}\in\mathscr{A}|^i_{v,\eta})&\\ \preceq \sum_{v\in\mathbb{Z}^4}\mathbf{P}_v\big(\tau_i<\tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow}\in\mathscr{A}|^i_{v,\eta}, \sup_{m\le\tau_i}\|X_m-\eta_i\|\le n^{1/2}(\log n)^{1/6+\delta}\big)&\\ +C_{\delta,p}n(\log n)^{-p}.& \qquad (6.21)\end{aligned}$$

Now $\mathrm{LE}(X^{\tau_i})\subseteq(X_m)_{m\le\tau_i}$, and so applying Wilson's algorithm, we have

$$\mathbf{P}_v\big(\tau_i<\tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow}\in\mathscr{A}|^i_{v,\eta}\sup_{m\le\tau_i}\|X_m-\eta_i\|\le n^{1/2}(\log n)^{1/6+\delta}\big)\preceq$$
$$\mathbf{P}^{\eta}(0\wedge v=\eta_i, \Gamma(0,0\wedge v)\in\mathscr{A}|^i_{v,\eta},\ \text{and}\ \Gamma(v,0\wedge v)\subseteq\Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta})).$$

Substituting this inequality into (6.21) and summing over $i \in \mathscr{I}$ yields

$$\sum_{i\in\mathscr{I}}\sum_{v\in\mathbb{Z}^4} \mathbf{P}_v(\tau_i < \tau_i^c, \mathrm{LE}(X^{\tau_i})^{\leftarrow} \in \mathscr{A}|^i_{v,\eta}) \preceq$$
$$\sum_{v\in\mathbb{Z}^4} \mathbf{P}^\eta(\Gamma(0,v) \in \mathscr{A}, \Gamma(0,0\wedge v) \subseteq \Lambda(r) \text{ and } \Gamma(v,0\wedge v) \subseteq \Lambda(0\wedge v, n^{1/2}(\log n)^{1/6+\delta}))$$
$$+ C_{\delta,p} n^2 (\log n)^{-p},$$

since $|\mathscr{I}| \leq n+1$. Substituting this inequality into (6.14) and taking expectations over η yields the claimed inequality (6.11). □

Containment of balls. We now turn our attention to the proof of Proposition 63. We begin by showing that it is very unlikely for $\mathfrak{T}$ to include a crossing of an annulus that it shorter than it should be by a large (i.e. non-sharp) polylogarithmic factor. We write $\partial\Lambda(r)$ for the set of vertices in $\mathbb{Z}^4$ with $\|x\|_\infty = r$.

Lemma 68. *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$ and for each $r,n \geq 1$ let $\mathscr{E}(r,n)$ be the event that there exists a path in $\mathfrak{T}$ from $\partial\Lambda(r)$ to $\partial\Lambda(4r)$ that has length at most n. Then*

$$\mathbf{P}\left(\mathscr{E}\left(r, \lceil r^2(\log r)^{-3}\rceil\right)\right) = \exp\left[-\Omega((\log r)^2)\right]$$

as $r \to \infty$.

Proof of Lemma 68. Fix $r \geq 2$, let $n = \lceil r^2(\log r)^{-3}\rceil$, and write $\mathscr{E} = \mathscr{E}(r,n)$. If $\mathscr{E}$ holds, there must exist a pair of points $x \in \partial\Lambda(r)$ and $y \in \partial\Lambda(4r)$ such that the path connecting x and y in $\mathfrak{T}$ is contained in the box $\Lambda(4r)$ and has length at most n. Considering separately the case that $x \wedge y$ belongs to $\Lambda(2r)$ or not yields the union bound

$$\mathbf{P}(\mathscr{E}) \leq \sum_{y\in\partial\Lambda(4r)}\sum_{z\in\Lambda(2r)} \mathbf{P}(z \in \Gamma(y,\infty), |\Gamma(y,z)| \leq n) +$$
$$\sum_{x\in\partial\Lambda(r)}\sum_{z\in\Lambda(4r)\setminus\Lambda(2r)} \mathbf{P}(z \in \Gamma(x,\infty), |\Gamma(x,z)| \leq n),$$

and using Wilson's algorithm to convert this into a loop-erased random walk quantity yields that

$$\begin{aligned}\mathbf{P}(\mathscr{E}) &\leq \sum_{y\in\partial\Lambda(4r)}\sum_{z\in\Lambda(2r)}\sum_{k=0}^{n}\mathbf{P}_y\left(\mathrm{LE}(X)_k = z\right) + \sum_{x\in\partial\Lambda(r)}\sum_{z\in\Lambda(4r)\setminus\Lambda(2r)}\sum_{k=0}^{n}\mathbf{P}_x\left(\mathrm{LE}(X)_k = z\right)\\ &= \sum_{y\in\partial\Lambda(4r)}\sum_{k=0}^{n}\mathbf{P}_y\left(\mathrm{LE}(X)_k \in \Lambda(2r)\right) + \sum_{x\in\partial\Lambda(r)}\sum_{k=0}^{n}\mathbf{P}_x\left(\mathrm{LE}(X)_k \in \Lambda(4r)\setminus\Lambda(2r)\right)\\ &\preceq r^3 n\mathbf{P}_0\left(\max_{0\leq k\leq n}\|\mathrm{LE}(X)_k\|_\infty \geq r\right). \end{aligned} \tag{6.22}$$

We will bound this probability using the weak L^1 method as introduced in [197, Section 6.2], which can be thought of as a simple special case of the typical time theory. Conditional on the loop-erased random walk $\mathrm{LE}(X)$, we have as in [193, Lemma 5.3] that the sequence of random variables $(\ell_{i+1}(X)-\ell_i(X))_{i\geq 0}$ are conditionally independent and satisfy

$$\mathbf{P}_0(\ell_{i+1}(X)-\ell_i(X) = m \mid \mathrm{LE}(X)) \leq p_{m-1}(0,0) \preceq \frac{1}{m^2}$$

for every $m \geq 1$, and it follows from Vershynin's weak triangle inequality for the weak L^1 norm [318] as explained in [197, Section 6.2] that

$$\mathbf{P}_0(\ell_n(X) \geq m \mid \mathrm{LE}(X)) \preceq \frac{n\log n}{m}$$

for every $n \geq 2$ and $m \geq 1$. As such, there exists a constant C such that

$$\begin{aligned}\mathbf{P}_0\left(\max_{0\leq k\leq n}\|\mathrm{LE}(X)_k\|_\infty \geq r\right) &\leq 2\mathbf{P}_0\left(\max_{0\leq k\leq n}\|\mathrm{LE}(X)_k\|_\infty \geq r,\, \ell_n(X) \leq Cn\log n\right)\\ &\leq 2\mathbf{P}_0\left(\max_{0\leq i\leq Cn\log n}\|X_i\|_\infty \geq r\right)\\ &\preceq \exp\left[-\Omega\left(\frac{r^2}{n\log n}\right)\right] \preceq \exp\left[-\Omega\left((\log r)^2\right)\right].\end{aligned}$$

where we have used the maximal version of Azuma-Hoeffding in the last line [269, Section 2]. The claim follows by substituting this estimate into (6.22) and using that $r^3 n = r^{O(1)} = \exp[o((\log r)^2)]$. □

Before proceeding with the deduction of Proposition 63 from Proposition 64 and Lemma 68, we will first introduce some more tools from [193, 197]. We begin by defining a variant of the uniform spanning tree known as the *0-wired uniform spanning forest*, which was first introduced by Járai and Redig [210] as part of their work on the Abelian sandpile model. Let $(V_n)_{n\geq 0}$ be an exhaustion of $\mathbb{Z}^4$ by finite connected sets. For each $n \geq 0$, let G_n^* be the graph obtained by identifying (a.k.a. wiring) $\mathbb{Z}^4 \backslash V_n$ into a single point denoted by ∂_n. Let G_n^{*0} be the graph obtained by identifying 0 with ∂_n in in G_n^*. The **0-wired uniform spanning forest** is then the weak limit of the uniform spanning trees on G_n^{*0} as $n \to \infty$, which is well-defined and does not depend on the choice of exhaustion [258, §3]. Lyons, Morris and Schramm [258] proved that the component of the origin in the 0-wired forest is finite almost surely, and, since the entire 0-wired forest is stochastically dominated by the uniform spanning tree by [259, Theorem 4.6], and the definitions ensure that every component other than that of the origin is infinite, the rest of the vertices of $\mathbb{Z}^4$ are contained in a single infinite one-ended component almost surely.

The stochastic domination property. We let $\mathfrak{T}_0$ be the component of 0 in the 0-wired UST. Lyons, Morris and Schramm [258, Proposition 3.1] proved that $\mathfrak{T}_0$ stochastically dominates $\mathfrak{P}(0)$, which we recall denotes the past of the origin in $\mathfrak{T}$. In [193], a stronger version of this stochastic domination property was derived, the relevant parts of which we restate below in our context. Given that the UST of $\mathbb{Z}^4$ is connected and one-ended, we can, in a unique manner, add an orientation to each edge in $\mathfrak{T}$ so that each vertex in the tree has exactly one oriented edge emanating from it. By abuse of notation, we denote the resulting oriented tree by $\mathfrak{T}$ as we do in the unoriented case. The oriented 0-wired spanning forest $\mathfrak{F}_0$ is generated similarly, but with the edges in the finite component all oriented towards the origin. Lastly, we generalise the notion of the *past*: given an arbitrary oriented forest F, we define the past of a vertex $v \in F$, denoted $\mathrm{past}_F(v)$, to be the set of vertices u with a directed path γ in F emanating from u and ending at v.

Lemma 69 (Stochastic Domination)**.** *Let $\mathfrak{T}$ be the oriented uniform spanning tree of $\mathbb{Z}^4$, and let $\mathfrak{F}_0$ be the oriented 0-wired uniform spanning forest of $\mathbb{Z}^4$. Let K be a finite set of vertices in $\mathbb{Z}^4$ and let $\Gamma(K) = \cup_{u\in K}\Gamma(u,\infty)$. Then for every increasing event $\mathscr{A} \subseteq \{0,1\}^{E(\mathbb{Z}^4)}$ and we have that*

$$\mathbb{P}\left(\mathrm{past}_{\mathfrak{F}\backslash F(K)}(0) \in \mathscr{A} \mid \Gamma(K)\right) \leq \mathbb{P}\left(\mathfrak{T}_0 \in \mathscr{A}\right).$$

We will also utilize the following result of [197].

Theorem 70 ([197], Theorem 1.6)**.** *Let $\mathfrak{T}_0$ be the component of the origin in the 0-wired uniform spanning tree of $\mathbb{Z}^4$. Then*

$$\mathbb{P}\left(\mathrm{rad}_{\mathrm{ext}}(\mathfrak{T}_0) \geq n\right) \asymp \frac{(\log n)^{1+o(1)}}{n^2}$$

for every $n \geq 2$.

Remark 10. For the proof of Proposition 63 it would suffice to have the weaker bound in which $(\log n)^{1+o(1)}$ is replaced by $(\log n)^{O(1)}$, which is significantly easier to prove.

With these tools in hand we proceed to the proof of Proposition 63.

Proof of Proposition 63. Fix $\delta \in (0,1]$, and fix an integer $n \geq 2$. Let η be the future of the origin in the uniform spanning tree $\mathfrak{T}$ and let $r = \lceil n^{1/2}(\log n)^{1/6+\delta}\rceil$. We write

$$\{\mathfrak{B}(n) \nsubseteq \Lambda(8r)\} \subseteq \mathscr{F} \cup \mathscr{E} \cup \mathscr{A},$$

where $\mathscr{F} = \{\eta[0,n] \nsubseteq \Lambda(r)\}$ is the event that the first n steps of the future are not contained in the box of radius r, $\mathscr{E} = \mathscr{E}(r, \lceil r^2/(\log r)^{-3}\rceil)$ is the event defined in Lemma 68, and $\mathscr{A}$ is the event $\{\mathfrak{B}(n) \nsubseteq \Lambda(8r)\} \setminus (\mathscr{F} \cup \mathscr{E})$. We have already shown in Lemma 68 that the probability of $\mathscr{E}$ is much smaller than required for n sufficiently large. For the event $\mathscr{F}$, we use Wilson's algorithm to compute that

$$\begin{aligned}
\mathbb{P}\left(\mathscr{F}\right) &= \mathbf{P}_0\left(\mathrm{LE}(X)^n \nsubseteq \Lambda(r)\right) \\
&\leq \mathbf{P}_0\left(\ell_n > 2n(\log n)^{1/3}\right) + \mathbf{P}_0\left(\max_{0 \leq k \leq 2n(\log n)^{1/3}} \|X_k\|_\infty > r\right) \\
&\preceq \frac{\log\log n}{(\log n)^{2/3}} + \exp\left[-\Omega((\log n)^{\delta})\right] \preceq_\delta \frac{\log\log n}{(\log n)^{2/3}}
\end{aligned}$$

as required, where the second inequality follows by Theorem 61 for the bound on ℓ_n, and e.g. the maximal version of Azuma-Hoeffding [269, Section 2] for the bound on the displacement of the simple random walk.

We now bound the probability of $\mathscr{A}$. Observe that if $\mathscr{A}$ holds then there exists an integer $0 \leq i \leq n-1$ such that $\mathfrak{P}(\eta_i, n)$ is not contained in $\Lambda(8r)$. Since $\mathscr{E}$ does not hold, we must also have that every crossing of the annulus $\Lambda(4r) \setminus \Lambda(r)$ has length at least $r^2/(\log r)^3$, and it follows that there must exist a collection of at least $r^2/(\log r)^3$ points $y \in (\mathfrak{B}(n) \setminus \eta[0,n]) \cap (\Lambda(4r) \setminus \Lambda(r))$ such that $\mathfrak{P}(y,n)$ has extrinsic diameter at least $4r$. Summing over all possible such points, applying Markov's inequality yields, and using the stochastic domination lemma (Lemma 69) yields that

$$\begin{aligned}
\mathbf{P}(\mathscr{A}) &\leq \frac{(\log r)^3}{r^2} \sum_{y\in\Lambda(4r)\setminus\Lambda(r)} \mathbb{P}(y \in \mathfrak{B}(n)\setminus \eta[0,n] \text{ and } \operatorname{diam}(\mathfrak{P}(y)) \geq 4r)\\
&\leq \frac{(\log r)^3}{r^2} \sum_{y\in\Lambda(4r)\setminus\Lambda(r)} \mathbb{P}\left(y \in \mathfrak{B}(n)\right)\mathbb{P}\left(\operatorname{rad}_{\mathrm{ext}}(\mathfrak{T}_0)\geq 2r\right)\\
&= \frac{(\log r)^3}{r^2}\mathbb{E}\left|\{y\in \mathfrak{B}(n) : y \notin \Lambda(r)\}\right|\mathbb{P}\left(\operatorname{rad}_{\mathrm{ext}}(\mathfrak{T}_0)\geq 2r\right),
\end{aligned}$$

and it follows from Proposition 64 and Theorem 70 that

$$\mathbf{P}(\mathscr{A}) \preceq_{\delta,p} \frac{(\log r)^3}{r^2}\frac{n^2}{(\log n)^p}\cdot\frac{(\log r)^{1+o(1)}}{r^2} \preceq (\log n)^{4-p+o(1)},$$

for every $p \geq 1$. Taking $p = 10$, say, yields a bound that is stronger than required and completes the proof. □

6.2.2 Lower bounds

In this section we prove the following proposition, which implies the lower bounds of Theorem 57. Note that, in contrast to Proposition 62, we do not lose any $(\log n)^{\pm o(1)}$ factors in this bound.

Proposition 71. *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$. Then*

$$|\mathfrak{B}(n)| = \mathbf{\Omega}\left(\frac{n^2}{(\log n)^{1/3}}\right)$$

as $n \to \infty$.

Remark 11. The proof yields the explicit lower tail bound

$$\mathbf{P}\left(|\mathfrak{B}(n)| \leq \frac{n^2}{\lambda(\log n)^{1/3}}\right) \preceq \lambda^{-1/5}$$

for every $n \geq 3$ and $1 \leq \lambda \leq \log n$. Presumably this bound is far from optimal.

We will prove this proposition by estimating the mean and variance of certain random variables that lower bound $|\mathfrak{B}(n)|$. We expect $|\mathfrak{B}(n)|$ to be unconcentrated[4], so its variance should be of the same order as its second moment and applying Chebyshev directly to $|\mathfrak{B}(n)|$

[4]Indeed, it should converge under appropriate rescaling to the volume of a ball in the ICRT, which is not deterministic.

should not be a viable method to prove lower tail bounds. Instead we calculate the mean and variance of a certain 'good' portion of the uniform spanning tree within a certain radius of the spine. We choose this radius according to how deep into the lower tail of the volume we wish to control: the lower we take this radius, the deeper into the tail we bound. The precise meaning of 'good' we will use is engineered precisely to make the later parts of the proof go through cleanly.

Our first task is to set up the relevant definitions. Recall that $\mathbf{P}_z$ denotes the law of a simple random walk X on $\mathbb{Z}^4$ started at z for each $z \in \mathbb{Z}^4$. [197, Theorem 7.4] states that if $\mathbf{P}_{0,\Lambda(r)}$ denotes the joint law of two independent random walks X and Y started at 0 and at a uniform point of $\Lambda(r)$ respectively, then

$$\mathbf{P}_{0,\Lambda(r)}(X \cap Y \cap \Lambda(r) \neq \emptyset) \asymp \frac{1}{\log r} \tag{6.23}$$

for $r \geq 2$. Fix $\alpha > 0$ and $r \geq 2$. We say a path γ in $\mathbb{Z}^4$ is (α, r)**-good** if

$$\sum_{z \in \Lambda(\gamma_0, 6r)} \mathbf{P}_z\big(\text{hit } \gamma \cap \Lambda(\gamma_0, 6r)\big) \leq \alpha \frac{r^4}{\log r},$$

and say that γ is (α, r)**-bad** otherwise. We note that

$$\mathbf{P}_0(X \text{ is } (\alpha, r)\text{-bad}) = \mathbf{P}_{0,\Lambda(6r)}\left(|\Lambda(6r)|\mathbf{P}_{0,\Lambda(6r)}(X \cap Y \cap \Lambda(6r) \neq \emptyset \mid X) > \alpha \frac{r^4}{\log r}\right) \preceq \alpha^{-1} \tag{6.24}$$

by (6.23) and Markov's inequality. Crucially, we also observe that being (α, r)-bad is an increasing property of a path in the sense that if γ and $\tilde{\gamma}$ are two paths satisfying $\gamma_0 = \tilde{\gamma}_0$ and $\gamma \subseteq \tilde{\gamma}$, then $\tilde{\gamma}$ is (α, r)-bad whenever γ is (α, r)-bad. We will apply this to bound the probability that a loop-erased random walk is bad in terms of the probability that the corresponding *simple* random walk is bad.

Condition on the future of the origin $\eta := \Gamma(0, \infty)$ in the uniform spanning tree $\mathfrak{T}$ and for each $x \in \mathbb{Z}^4$ and $r \geq 3$ consider the random set

$$M_\alpha(x, r) = \Big\{ y \in \Lambda(x, 3r) : \Gamma(y, 0 \wedge y) \subseteq \Lambda(x, 3r),\ |\Gamma(y, 0 \wedge y)| \leq \frac{r^2}{(\log r)^{1/3}}, \\ \text{and } \Gamma(y, 0 \wedge y) \text{ is } (\alpha, r)\text{-good} \Big\}.$$

The key step in the proof of Proposition 71 is to bound the conditional mean and variance of $|M_\alpha(x,r)|$ in terms of the *capacity* of η. Here we recall that the **capacity** (a.k.a. conductance to infinity) of a set $A \subseteq \mathbb{Z}^4$ is defined to be

$$\begin{aligned}\mathrm{Cap}(A) = \sum_{a\in A} \deg(a)\mathbf{P}_a(\text{never return to } A \text{ after time zero})\\ = 8\sum_{a\in A}\mathbf{P}_a(\text{never return to } A \text{ after time zero}).\end{aligned}$$

The two relevant estimates are as follows, where we write Var^η for the conditional variance given η:

Proposition 72. *There exist $\alpha_0 > 0$ and $r_0 > 0$ such that if $\alpha \geq \alpha_0$ then*

$$\mathbb{E}^\eta\,|M_\alpha(x,r)| \succeq r^2\mathrm{Cap}(\eta\cap\Lambda(x,r))$$

for every $x \in \mathbb{Z}^4$ and every $r \geq r_0$.

Proposition 73. *For each $\alpha > 0$ we have*

$$\mathrm{Var}^\eta(|M_\alpha(x,r)|) \preceq \alpha\frac{r^6}{\log r}\mathrm{Cap}(\eta\cap\Lambda(x,3r)).$$

for every $x \in \mathbb{Z}^4$ and every $r \geq 2$.

We will require the following variational formula for the capacity proved in [200, Lemma 2.3]. Recall that the Green's function on $\mathbb{Z}^4$ is defined by

$$G(x,y) = \frac{1}{\deg y}\mathbf{E}_x\sum_{n\geq 0}\mathbb{1}(X_n = y) = \frac{1}{8}\mathbf{E}_x\sum_{n\geq 0}\mathbb{1}(X_n = y),$$

where X is a simple random walk on $\mathbb{Z}^4$ and $x,y \in \mathbb{Z}^d$.

Lemma 74. *The capacity of a set $S \subset \mathbb{Z}^4$ can be expressed as*

$$\mathrm{Cap}(S)^{-1} = \inf\left\{\sum_{u,v\in S} G(u,v)\mu(u)\mu(v) : \mu \text{ is a probability measure on } S\right\}. \tag{6.25}$$

Proof of Propostion 72. Fix $x \in \mathbb{Z}^4$, $r \geq 1$ and $\alpha > 0$. We assume that $\mathrm{Cap}(\eta\cap\Lambda(x,r)) > 0$ or else the proposition is trivial. We let $n = \lfloor r^2(\log r)^{-1/3}\rfloor$ and $N = \lfloor \lambda r^2\rfloor$ where $\lambda \in (0,1/2)$ is a parameter that will later be taken to be a small constant. Let V be a uniform random element of $\Lambda(x,3r)$, let $X = (X_m)_{m\geq 0}$ be a random walk started at V, and let $\mathbf{P}$ denote the joint law of V and X. Let σ be the time at which X^N hits $\eta \cap \Lambda(x,3r)$ and let τ be the time X^N first exits $\Lambda(x,3r)$. Each of these stopping times is defined to be infinite if the relevant event does not occur before or at time N. We let μ be a measure which minimises the right hand side of (6.25) when $S = \eta \cap \Lambda(x,r)$ and define the random variable

$$A_r = \mathbb{1}(\sigma < \tau, |\mathrm{LE}(X^\sigma)| \leq n, \mathrm{LE}(X^\sigma) \text{ good}) \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}(X_j = w),$$

where to save on notation we have and will abbreviate (α, r)-good and (α, r)-bad to good and bad respectively. The weight μ is included in the definition of A_r since it makes the second moment of A_r easier to control; this is closely related to the theory of *Martin capacity* as developed in [69]. An application of Wilson's algorithm implies that

$$\mathbb{E}^\eta |M_\alpha(x,r)| \geq \sum_{v \in \Lambda(x,3r)} \mathbf{P}(A_r > 0 \mid V = v) = |\Lambda(x,3r)| \mathbf{P}(A_r > 0),$$

so that to prove the proposition we need only demonstrate that there exists $\alpha_0, r_0 > 0$ such that

$$\mathbf{P}(A_r > 0) \succeq r^{-2} \mathrm{Cap}(\eta \cap \Lambda(x,r))$$

for every $\alpha \geq \alpha_0, r \geq r_0$, where we emphasize that the constant implied by the $\succeq$ on the right hand side is independent of η and r. We do so by proving that

$$\mathbf{E} A_r \succeq r^{-2} \qquad (6.26) \qquad \mathbf{E} A_r^2 \preceq r^{-2} \mathrm{Cap}^{-1}(\eta \cap \Lambda(x,r)) \qquad (6.27)$$

for appropriately large α, r and an appropriately small constant value of λ; once (6.26) and (6.27) are established the claim will follow since, by Cauchy-Schwartz,

$$\mathbb{E}^\eta |M_\alpha(x,r)| \succeq r^4 \mathbf{P}(A_r > 0) \succeq r^4 \frac{\mathbf{E}[A_r]^2}{\mathbf{E}[A_r^2]} \succeq r^2 \mathrm{Cap}(\eta \cap \Lambda(x,r))$$

as claimed.

We begin by lower bounding the expectation of A_r. We decompose A_r as $A_r = E_r - D_r - C_r - B_r$, where

$$B_r = \mathbb{1}(\sigma \geq \tau) \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}(X_j = w),$$

$$C_r = \mathbb{1}(\sigma < \tau, |\mathrm{LE}(X^\sigma)| > n) \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}(X_j = w),$$

$$D_r = \mathbb{1}(\sigma < \tau, |\mathrm{LE}(X^\sigma)| \leq n, \mathrm{LE}(X^\sigma) \text{ bad}) \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}(X_j = w), \qquad \text{and}$$

$$E_r = \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}(X_j = w).$$

The random variable E_r is the μ-mass of the intersections of the random walk with the relevant part of η, i.e. $\eta \cap \Lambda(x,r)$. From E_r, we have subtracted the error term B_r pertaining to the possibility that the walk exits the ball $\Lambda(x,3r)$ before hitting the relevant part of η; the term C_r pertaining to the possibility that that the walk hits the relevant part of η before exiting this ball, but has too long a loop erasure; and finally the term D_r pertaining to the possibility that the walk hits the relevant part of η before exiting this ball and has a suitably short loop erasure, but the loop erasure is *bad*, as defined above.

Lower bounding the expectation of E_r: First, we lower bound the expectation of E_r. We have by time-reversal that

$$\begin{aligned}
\mathbf{E}[E_r] &\geq \frac{1}{|\Lambda(x,3r)|} \sum_{w \in \eta \cap \Lambda(x,r)} \mu(w) \sum_{j=0}^{N} \sum_{v \in \Lambda(x,3r)} \mathbf{P}_v(X_j = w) \\
&\succeq r^{-4} \sum_{w \in \eta \cap \Lambda(x,r)} \mu(w) \sum_{j=0}^{N} \mathbf{P}_w(X_j \in \Lambda(x,3r)) \\
&\succeq r^{-4} \sum_{j=0}^{N} \mathbf{P}_0(X_j \in \Lambda(0,2r)) \succeq r^{-4} \sum_{j=0}^{N} 1 - \frac{j}{4r^2} \geq r^{-4} N \left(1 - \frac{N}{4r^2}\right) \\
&\succeq \lambda r^{-2}(1 - \lambda/4) \succeq \lambda r^{-2},
\end{aligned} \tag{6.28}$$

where the third inequality follows since $\sum_{w \in \eta \cap \Lambda(x,r)} \mu(w) = 1$ and $\Lambda(w,2r) \subset \Lambda(x,3r)$ for $w \in \Lambda(x,r)$, the fourth inequality follows by e.g. the central limit theorem for the simple random walk, and the penultimate inequality holds if $r > 1/\lambda$ (which is just the condition we need to avoid rounding N down to zero).

Upper bounding the expectation of B_r: Next, we upper bound the expectation of B_r, which pertains to the possibility that the walk exits the ball $\Lambda(x,3r)$ before hitting the relevant part of η. We have

$$\begin{aligned}\mathbf{E}[B_r] &= \frac{1}{|\Lambda(x,3r)|}\sum_{v\in\Lambda(x,3r)}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\mathbf{P}_v(X_j=w,\sigma\geq\tau)\\&\leq\frac{1}{|\Lambda(x,3r)|}\sum_{v\in\Lambda(x,3r)}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\mathbf{P}_v(X_j=w,\tau\leq j)\\&\preceq r^{-4}\sum_{v\in\Lambda(x,3r)}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\mathbf{P}_w(X_j=v,\tau\leq j)\\&\preceq r^{-4}N\sum_{w\in\eta\cap\Lambda(x,r)}\mu(w)\mathbf{P}_w(\tau\leq N)\preceq\lambda r^{-2}\mathbf{P}_0\left(\sup_{0\leq i\leq N}\|X_i\|_\infty\geq 2r\right),\end{aligned}$$

where the second inequality follows by time reversal of X, and the final inequality holds because the distance between any $w\in\eta\cap\Lambda(x,r)$ and $\partial\Lambda(x,3r)$ is greater than or equal to $2r$. Since $\mathbb{E}_o[\sup_{j\leq i}\|X_j\|^2]\preceq i$ for $i\geq 0$, it follows by Markov's inequality that

$$\mathbf{E}[B_r]\preceq\lambda r^{-2}\frac{N}{r^2}\preceq\lambda^2r^{-2}. \tag{6.29}$$

Upper bounding the expectation of D_r. We now upper bound the expectation of D_r, which pertains to the possibility that the walk hits the relevant part of η before exiting this ball and has a suitably short loop erasure, but the loop erasure is bad. Observe that

$$\begin{aligned}\mathbf{E}[D_r]\leq\mathbf{E}\left[\mathbb{1}(\mathrm{LE}(X^\sigma)\text{ bad})\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\mathbb{1}(X_j=w)\right]&\\ \leq\mathbf{E}\left[\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\mathbb{1}(X\text{ bad},X_j=w)\right]&\\ \preceq r^{-4}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\sum_{v\in\Lambda(x,3r)}\mathbf{P}_v(X\text{ bad},X_j=w)&\\ \leq r^{-4}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=0}^{N}\mu(w)\sum_{v}\mathbf{P}_0(X\text{ bad},X_j=w-v)&\\ =r^{-4}\sum_{j=0}^{N}\mathbf{P}_0(X\text{ bad})\preceq Nr^{-4}\mathbf{P}_0(X\text{ bad})\preceq\alpha^{-1}Nr^{-4}\preceq\alpha^{-1}\lambda r^{-2},&\end{aligned}$$

where the second inequality follows as $\mathrm{LE}(X^\sigma) \subseteq X$, the fourth inequality follows by translation-invariance, and the penultimate inequality follows by (6.24). Combining this inequality with (6.29) and (6.28), we can see that there exist positive constants α_0 and λ_0 such that if $\alpha \geq \alpha_0$, $\lambda = \lambda_0$, and $r \geq 1/\lambda_0$ then

$$\mathbf{E}[E_r - D_r - B_r] \succeq r^{-2}.$$

Thus, to complete the proof of (6.26), it is sufficient to show that $\mathbf{E}[C_r] = o(r^{-2})$.

Upper bounding the expectation of C_r: To bound the final term C_r, which pertains to the possibility that that the walk hits the relevant part of η before exiting this ball, but has too long a loop erasure. We will need some understanding of the *cut times* of a simple random walk. Recall that a time $t \geq 0$ is said to be a **cut time**, or *loop-free time* of the random walk X if $X[0,t]$ and $X(t,\infty)$ are disjoint. We observe that if $0 \leq s \leq t$ are cut times of X then the loop-erasure of X is equal to the concatenation of the loop-erasures of the portions of X before s, between s and t, and after t; this property allows us to decorrelate different parts of the loop-erased random walk. We use the following estimate of Lawler which demonstrates that the random walk on $\mathbb{Z}^4$ has a reasonably good supply of cut times.

Lemma 75 ([233], Lemma 7.7.4)**.** *Let X be simple random walk on $\mathbb{Z}^4$. Then*

$$\mathbf{P}(\textit{there are no cut times between times } n \textit{ and } m) \preceq \frac{\log\log m}{\log m}.$$

for every $3 \leq n \leq m$ *such that* $|n-m| \geq m/(\log m)^6$.

Observe that if $|\mathrm{LE}(X^\sigma)| > n$ then we must have that $\sigma > n$ and that if X has a cut time in $[\sigma - n/4, \sigma]$, then $\left|\mathrm{LE}(X^j)\right| \geq 3n/4$ for every $j \geq \sigma$. Therefore,

$$C_r \leq \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=0}^{N} \mu(w) \mathbb{1}\Big(X_j = w, |\mathrm{LE}(X^\sigma)| > n, n < \sigma \leq N \wedge j\Big) \leq C_r' + C_r'', \qquad (6.30)$$

where

$$C_r' = \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=n+1}^{N} \mu(w) \mathbb{1}\Big(X_j = w, X \text{ has no cut time in } [\sigma, \sigma - n/4], n < \sigma \leq N \wedge j\Big),$$

$$C_r'' = \sum_{w \in \eta \cap \Lambda(x,r)} \sum_{j=n+1}^{N} \mu(w) \mathbb{1}\Big(X_j = w, \left|\mathrm{LE}(X^j)\right| > \frac{3}{4} n\Big).$$

We show that the expectation conditioned on η of both C_r' and C_r'' is $o(r^{-2})$; we begin with the latter. We have

$$\begin{aligned}
\mathbf{E}C_r'' &\le \frac{1}{|\Lambda(x,3r)|}\sum_{w\in\eta\cap\Lambda(x,r)}\mu(w)\sum_{j=n}^{N}\sum_{v\in\Lambda(x,3r)}\mathbf{P}_v\left(X_j=w,\left|\mathrm{LE}(X^j)\right|>\frac{3}{4}n\right)\\
&=\frac{1}{|\Lambda(x,3r)|}\sum_{w\in\eta\cap\Lambda(x,r)}\mu(w)\sum_{j=n}^{N}\sum_{v\in\Lambda(x,3r)}\mathbf{P}_0\left(X_j=w-v,\left|\mathrm{LE}(X^j)\right|>\frac{3}{4}n\right)\\
&\le \frac{1}{|\Lambda(x,3r)|}\sum_{w\in\eta\cap\Lambda(x,r)}\mu(w)\sum_{j=n}^{N}\mathbf{P}_0\left(\left|\mathrm{LE}(X^j)\right|>\frac{3}{4}n\right)\\
&\preceq r^{-4}\sum_{j=n}^{N}\mathbf{P}_0\left(\left|\mathrm{LE}(X^j)\right|>\frac{3}{4}n\right),
\end{aligned}$$

where we used translation invariance in the second line. Observe for each $n\le i\le N$ that if X has a cut time in $[i-i/(\log i)^6,i]$, then $\left|\mathrm{LE}(X^i)\right|\le\left|\mathrm{LE}_\infty(X^i)\right|+i/(\log i)^6$. Therefore,

$$\begin{aligned}
r^4\mathbf{E}C_r'' &\preceq \sum_{i=n}^{N}\mathbf{P}_0\left(\left|\mathrm{LE}(X^i)\right|>\frac{3}{4}n\right)\\
&\le \sum_{i=n}^{N}\mathbf{P}_0(\left|\mathrm{LE}_\infty(X^i)\right|>(3/4)n-i/(\log i)^6)+\mathbf{P}_0\left(X \text{ has no cut times in } [i-i/(\log i)^6 i,i]\right)\\
&\preceq \sum_{i=n}^{N}\mathbf{P}_0(\rho_i>(3/4)n-i/(\log i)^6 i)+c\frac{\log\log i}{\log i}\\
&\preceq N\frac{\log\log n}{\log n}+\sum_{i=n}^{N}\frac{\log\log i}{(\log i)^{2/3}}\preceq N\frac{\log\log n}{(\log n)^{2/3}}=o(r^2)
\end{aligned}\tag{6.31}$$

as required, where the third inequality follows by Lemma 75 and the fourth inequality follows from Theorem 61 and the fact that $\lambda<1/2$. Next, we upper bound the conditional expectation of C_r'. Recalling the definitions $N=\lfloor\lambda r^2\rfloor$ for some $\lambda\in(0,1/2)$ and $n=\lfloor r^2(\log r)^{1/3}\rfloor$, we can calculate that $N\le n(\log n)^{1/3}$ for all $r\ge 2$. Define the sequence of times $T_k=\lceil(1+k/8)n\rceil$ for $k\ge 0$, and observe that for r larger than some universal constant, if $n\le\sigma\le N$ and X has no cut time in $[\sigma-n/4,\sigma]$, then X has no cut times in at least one of the intervals belonging to the family $\{[T_k-T_k/(\log T_k)^6,T_k]:0\le k\le 8\lceil(\log n)^{1/3}\rceil\}$. Therefore, for r larger than some universal constant, we have that

$$C_r'\le\sum_{k=0}^{8\lceil(\log n)^{1/3}\rceil}\sum_{w\in\eta\cap\Lambda(x,r)}\sum_{j=n+1}^{N}\mu(w)\mathbb{1}\left(X_j=w,X \text{ has no cut time in } [T_k-T_k/(\log T_k)^6,T_k]\right).\tag{6.32}$$

We also have by symmetry that

$$\mathbf{P}_x\Big(X_j = y, X \text{ has no cut time in } [T_k - T_k/(\log T_k)^6, T_k]\Big)$$
$$= \mathbf{P}_y\Big(X_j = x, X \text{ has no cut time in } [T_k - T_k/(\log T_k)^6, T_k]\Big)$$

for each $x, y \in \mathbb{Z}^4$ and $j \geq 0$, so that for r larger than some universal constant

$$\begin{aligned}\mathbf{E}C'_r &\leq \frac{N}{|\Lambda(3x,r)|} \sum_{k=0}^{8\lceil(\log n)^{1/3}\rceil} \mathbf{P}_0\Big(X \text{ has no cut time in } [T_k, T_k - T_k/(\log T_k)^6]\Big) \\ &\preceq \lambda r^{-2} \sum_{k=0}^{8\lceil(\log n)^{1/3}\rceil} \frac{\log\log T_k}{\log T_k} \preceq \lambda r^{-2} (\log n)^{1/3} \frac{\log\log n}{\log n} = o(r^{-2}), \quad (6.33)\end{aligned}$$

where the second inequality follows from Lemma 75. We have now shown (6.26), and so to complete the proof we must show (6.27), which upper bounds the second moment of A.

Upper bounding the second moment of A. It is at this stage of the proof that we benefit from defining A in terms of the measure μ. Indeed, we can use the Markov property to compute that

$$\begin{aligned}\mathbf{E}A_r^2 &\leq \mathbf{E}\left[\left(\sum_{i\geq 0} \sum_{w\in\eta\cap\Lambda(x,r)} \mu(w)\mathbb{1}\big(X_i = w\big)\right)^2\right] \\ &\leq 2\mathbf{E}\left[\sum_{i\geq 0} \sum_{w,z\in\eta\cap\Lambda(x,r)} \mu(w)\mu(z)\mathbb{1}\big(X_i = w\big)\sum_{j\geq i}\mathbb{1}\big(X_j = z\big)\right] \\ &\asymp 2\mathbf{E}\left[\sum_{i\geq 0} \sum_{w,z\in\eta\cap\Lambda(x,r)} \mu(w)\mu(z)G(w,z)\mathbb{1}\big(X_i = w\big)\right] \\ &= \frac{2}{|\Lambda(x,3r)|} \sum_{w,z\in\eta\cap\Lambda(x,r)} \mu(w)\mu(z)G(w,z) \sum_{i\geq 0}\sum_{v\in\Lambda(x,3r)} \mathbf{P}_v\big(X_i = w\big),\end{aligned}$$

and hence by time-reversal that

$$\begin{aligned}\mathbf{E}A_r^2 &\preceq r^{-4} \sum_{w,z\in\eta\cap\Lambda(x,r)} \mu(w)\mu(z)G(w,z)\sum_{i\geq 0}\mathbf{P}_w\big(X_i \in \Lambda(x,3r)\big) \\ &\preceq r^{-2} \sum_{w,z\in\eta\cap\Lambda(x,r)} \mu(w)\mu(z)G(w,z) = r^{-2}\mathrm{Cap}^{-1}(\eta\cap\Lambda(x,r)),\end{aligned}$$

where the final inequality follows since the random walk spends at most $O(r^2)$ time in any ball of radius r in expectation (which follows from the Green's function bound $G(x,y) \preceq$

$\|x-y\|_2^{-2}$ for $x \neq y$), and the final equality follows from the definition of μ. This concludes the proof of (6.27) and hence the proof of the proposition. □

We now turn to the proof of the variance estimate of Proposition 73. We will require the following lemma relating the capacity of a set S to the probability that a random walk, started at a uniform position in a ball containing S, hits S. The lemma will follow straightforwardly from [69, Theorem 2.2] and Lemma 74. We prove the result in all dimensions $d \geq 3$ for completeness.

Lemma 76. *Fix a dimension $d \geq 3$, a radius $r \geq 1$, and let $S \subseteq \Lambda(r) := \{x \in \mathbb{Z}^d : \|x\|_\infty \leq r\}$. Let X be a simple random walk on $\mathbb{Z}^d$. Then*

$$\sum_{x \in \Lambda(r)} \mathbf{P}_x(X \text{ hits } S) \asymp r^{d-2} \mathrm{Cap}(S).$$

Proof of Lemma 76. [69, Theorem 2.2] states that for any transient Markov chain $(X_n)_{n \geq 0}$ on a countable state space Ω with initial state ρ and Green's function $G(x,y) = (\deg y)^{-1} \sum_{n \geq 0} \mathbf{P}_x(X_n = y)$, we have that

$$\mathbb{P}_\rho(X \text{ hits } S) \asymp \inf_\mu \left[\sum_{x,y \in S} \mu(x) \frac{G(x,y)}{G(\rho,y)} \mu(y) \right]^{-1},$$

for any subset $S \subseteq \Omega$, where the infimum on the right hand side is taken over probability measures on S. We would like to apply this result with X a simple random walk on state space $\mathbb{Z}^d$, however, we would like the walk to start at a random vertex. To achieve this, we attach a 'ghost vertex' to the state space from which the random walk will start. We set up the transition probabilities from the ghost vertex so that after one step, the walk's distribution on $\mathbb{Z}^d$ is equal to that which we desire.

Define the set $\mathbb{Z}^d_* = \mathbb{Z}^d \cup \{*\}$, where $*$ is the additional ghost vertex, and define the Markov transition kernel p on the state space S by $p(x,y) = \frac{1}{8}\mathbb{1}(x \sim y)$ for $x, y \in \mathbb{Z}^d$ and $p(*,z) = 1/|\Lambda|$ for $z \in \Lambda := \Lambda(r)$. Note that a trajectory of this chain, which we will denote by X, is just a simple random walk on $\mathbb{Z}^d$ when started in $\mathbb{Z}^d$. We observe that

$$\sum_{x \in \Lambda} \mathbf{P}(X \text{ hits } S) = \mathbf{P}_*(X \text{ hits } S) \asymp \inf_\mu \left[\sum_{x,y \in S} \mu(x) \frac{G(x,y)}{G(*,y)} \mu(y) \right]^{-1}, \tag{6.34}$$

for any subset $S \subseteq \Lambda$. An integral comparison yields that

$$G(*,y) \asymp \sum_{x \in \Lambda} \frac{1}{(1 \vee \|x-y\|_\infty)^{d-2}} \asymp r^{d-2},$$

for $y \in \Lambda$, and so

$$\inf_{\mu} \left[\sum_{x,y \in S} \mu(x) \frac{G(x,y)}{G(*,y)} \mu(y) \right]^{-1} \asymp r^{d-2} \inf_{\mu} \left[\sum_{x,y \in S} \mu(x) G(x,y) \mu(y) \right]^{-1}.$$

Substituting this into (6.34) and applying Lemma 74, we get

$$\sum_{x \in \Lambda(r)} \mathbf{P}(X \text{ hits } S) \asymp r^{d-2} \inf_{\mu} \left[\sum_{x,y \in S} \mu(x) G(x,y) \mu(y) \right]^{-1} = r^{d-2} \mathrm{Cap}(S)$$

as claimed. □

Proof of Proposition 73. Given $y,z \in \mathbb{Z}^4$, let Y be a random walk started at y and let Z be an independent random walk started at z and write $\mathbf{P}_{y,z}$ for the joint law of Y and Z. Let σ_1 be the first time Y hits η and let σ_2 be the first time Z hits $\eta \cup \mathrm{LE}(Y^{\sigma_1})$. We continue to write $n = \lceil r^2 (\log r)^{-1/3} \rceil$ as in the previous proof. Abbreviating $M = M_\alpha$, $\Lambda = \Lambda(x, 3r)$, we have by Wilson's algorithm that

$$\begin{aligned}
\mathbb{E}^{\eta}[|M(x,r)|^2] &\le \sum_{y,z \in \Lambda} \mathbb{P}^{\eta}\big(y,z \in M(x,r)\big) \\
&\le \sum_{y,z \in \Lambda} \mathbf{P}_{y,z}(\sigma_1 < \infty, \sigma_2 < \infty, |\mathrm{LE}(Y^{\sigma_1})| \le n, |\mathrm{LE}(Z^{\sigma_2})| \le n, \\
&\qquad \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \mathrm{LE}(Z^{\sigma_2}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}), \mathrm{LE}(Z^{\sigma_2}) \text{ both good}).
\end{aligned} \tag{6.35}$$

Now, on the event that $\sigma_1, \sigma_2 < \infty$, let σ_3 be the time Z first hits $\mathrm{LE}(Y^{\sigma_1})$ and let σ_4 be the time Z first hits η. We split according to whether $\sigma_3 \le \sigma_4$ or $\sigma_4 < \sigma_3$, beginning with the case $\sigma_4 < \sigma_3$. Observing that $\sigma_2 = \sigma_4$ on this event, we obtain

$$\begin{aligned}
&\sum_{y,z\in\Lambda} \mathbf{P}_{y,z}(\sigma_1 < \infty,\, \sigma_2 < \infty,\, |\mathrm{LE}(Y^{\sigma_1})| \le n,\, |\mathrm{LE}(Z^{\sigma_2})| \le n,\\
&\qquad\qquad \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda,\, \mathrm{LE}(Z^{\sigma_2}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}), \mathrm{LE}(Z^{\sigma_2}) \text{ both good, and } \sigma_4 < \sigma_3).\\
&\quad \le \sum_{y,z\in\Lambda} \mathbf{P}_{y,z}(\sigma_1 < \infty,\, \sigma_4 < \infty,\, |\mathrm{LE}(Y^{\sigma_1})| \le n,\, |\mathrm{LE}(Z^{\sigma_4})| \le n,\\
&\qquad\qquad \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda,\, \mathrm{LE}(Z^{\sigma_4}) \subseteq \Lambda, \text{ and } \mathrm{LE}(Y^{\sigma_1}), \mathrm{LE}(Z^{\sigma_4}) \text{ both good}).\\
&\quad = \sum_{y,z\in\Lambda} \mathbf{P}_{y}(\sigma_1 < \infty,\, |\mathrm{LE}(Y^{\sigma_1})| \le n, \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \text{ and } \mathrm{LE}(Y^{\sigma_1}) \text{ good})\\
&\qquad\qquad \cdot \mathbf{P}_{z}(\sigma_4 < \infty,\, |\mathrm{LE}(Y^{\sigma_4})| \le n, \mathrm{LE}(Y^{\sigma_4}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_4}) \text{ good})\\
&\quad = \left[\sum_{y\in\Lambda} \mathbf{P}_{y}(\sigma_1 < \infty, |\mathrm{LE}(Y^{\sigma_1})| \le n, \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}) \text{ good})\right]^2 = \mathbb{E}^\eta[|M(y,r)|]^2,
\end{aligned} \tag{6.36}$$

where the first equality follows by independence of Y and Z conditional on η, and the last follows by an application of Wilson's algorithm. On the other hand, if $\sigma_3 \le \sigma_4$ then $\sigma_2 = \sigma_3$, and so we get

$$\begin{aligned}
&\sum_{y,z\in\Lambda} \mathbf{P}_{y,z}(\sigma_1 < \infty, \sigma_2 < \infty, |\mathrm{LE}(Y^{\sigma_1})| \le n, |\mathrm{LE}(Z^{\sigma_2})| \le n,\\
&\qquad\qquad \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \mathrm{LE}(Z^{\sigma_2}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}), \mathrm{LE}(Z^{\sigma_2}) \text{ both good}, \sigma_3 \le \sigma_4)\\
&\quad \le \sum_{y\in\Lambda} \mathbf{E}_y\left[\mathbb{1}(\sigma_1 < \infty, |\mathrm{LE}(Y^{\sigma_1})| \le n, \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}) \text{ good}) \sum_{z\in\Lambda} \mathbf{P}_{y,z}(\sigma_3 < \infty \mid Y)\right]\\
&\quad \le \alpha \frac{r^4}{\log r} \sum_{y\in\Lambda} \mathbf{P}_y(\sigma_1 < \infty, |\mathrm{LE}(Y^{\sigma_1})| \le n, \mathrm{LE}(Y^{\sigma_1}) \subseteq \Lambda, \mathrm{LE}(Y^{\sigma_1}) \text{ good})\\
&\quad = \alpha \frac{r^4}{\log r} \mathbb{E}^\eta\, |M(x,r)|,
\end{aligned} \tag{6.37}$$

where the final inequality follows by the definition of 'good', and the final equality follows by an application of Wilson's algorithm. Substituting (6.37) and (6.36) into (6.35) with a union bound yields

$$\mathbb{E}^\eta[|M(x,r)|^2] \le \mathbb{E}^\eta[|M(x,r)|]^2 + \alpha \frac{r^4}{\log r} \mathbb{E}^\eta\, |M(x,r)|$$

and hence that

$$\operatorname{Var}^{\eta}(|M(x,r)|) \le \alpha \frac{r^4}{\log r} \mathbb{E}^{\eta} |M(x,r)| . \tag{6.38}$$

Finally we upper bound $\mathbb{E}^{\eta} |M(x,r)|$. We have that

$$\mathbb{E}^{\eta}|M(x,r)| \le \sum_{x\in\Lambda} \mathbf{P}_x(X \text{ hits } \eta \cap \Lambda),$$

so that applying Lemma 76 to the right hand side and plugging the resulting inequality into (6.38) concludes the proof. $\square$

Our next goal is to deduce Proposition 71 from Propositions 72 and 73. To proceed we will need the following result controlling the capacity of the first n steps of a loop-erased random walk which follows easily from [197, Proposition 3.4] and Theorem 61.

Proposition 77. *Let X be a random walk on $\mathbb{Z}^4$ started at the origin. There exists a constant $C > 0$ such that we have*

$$\mathbf{P}\left(\operatorname{Cap}(\operatorname{LE}(X)^n) \le \frac{Cn}{(\log n)^{2/3}}\right) \preceq \frac{\log\log n}{(\log n)^{2/3}},$$

for every $n \ge 2$.

Proof. By [197, Proposition 3.4], we know that there exists a constant c such that

$$\mathbf{P}\left(\operatorname{Cap}(\operatorname{LE}_\infty(X^n)) \le \frac{cn}{\log n}\right) \preceq \frac{1}{(\log n)^{2/3}},$$

for each $n \ge 2$. Fix $\varepsilon \in (0,1/3)$. Employing a union bound and the fact that capacity is increasing, we obtain

$$\begin{aligned}
\mathbf{P}\left(\operatorname{Cap}(\operatorname{LE}(X)^n) \le \frac{Cn}{(\log n)^{2/3}}\right) &= \mathbf{P}\left(\operatorname{Cap}(\operatorname{LE}_\infty(X^{\ell_n})) \le \frac{Cn}{(\log n)^{2/3}}\right) \\
&\le \mathbf{P}\left(\operatorname{Cap}(\operatorname{LE}_\infty(X^{(1-\varepsilon)n(\log n)^{1/3}})) \le \frac{Cn}{(\log n)^{2/3}}\right) \\
&\quad + \mathbf{P}\left(\left|\frac{\ell_n}{n(\log n)^{1/3}} - 1\right| > \varepsilon\right) \\
&\preceq \frac{1}{(\log n)^{2/3}} + \frac{\log\log n}{(\log n)^{2/3}} \preceq \frac{\log\log n}{(\log n)^{2/3}}
\end{aligned}$$

when we choose $C < c(1-\varepsilon)$. $\square$

We will also use the following covering lemma, whose proof we defer to the end of the section.

Lemma 78. *Let S be a finite subset of $\mathbb{Z}^4$, and let $r \geq 1$. Then there exists an integer K and points $\{x_i : 1 \leq i \leq K\} \subseteq \mathbb{Z}^4$ such that the balls $\Lambda(x_i, 3r)$ are disjoint, $\{x_i\}_{1 \leq i \leq K} \subseteq S + \Lambda(r)$, and*

- $\sum_{i=1}^{K} \mathrm{Cap}(S \cap \Lambda(x_i, r)) \geq 3^{-4}\mathrm{Cap}(S)$, *and*
- $\sum_{i=1}^{K} \mathrm{Cap}(S \cap \Lambda(x_i, 3r)) \leq 15^4 \sum_{i=1}^{K} \mathrm{Cap}(S \cap \Lambda(x_i, r))$.

We now have everything we need to complete the proof of Proposition 71 given Lemma 78.

Proof of Proposition 71. Let α_0, r_0 be the constants yielded by Proposition 72, and fix $r \geq r_0 \vee 2$, $\alpha > \alpha_0$. For the remainder of the proof we will abbreviate $M = M_\alpha$. Let $K \geq 1$ and suppose that $\{x_i : 1 \leq i \leq K\} \subseteq \mathbb{Z}^4$ is a set of points such that the family of boxes $(\Lambda(x_i, 3r))_{i=1}^{K}$ are mutually disjoint. We first show that the random variables $|M(x_i, r)|$ are pairwise negatively correlated conditional on η in the sense that

$$\mathbb{E}^\eta\left[|M(x_i,r)| \cdot \left|M(x_j,r)\right|\right] \leq \mathbb{E}^\eta\left[|M(x_i,r)|\right]\mathbb{E}^\eta\left[\left|M(x_j,r)\right|\right]$$

for every $1 \leq i < j \leq K$. Indeed, suppose that $u \in \Lambda(x_i, 3r)$ and $v \in \Lambda(x_j, 3r)$ for some $i \neq j$. We sample the UST conditional on $\eta = \Gamma(0, \infty)$ with Wilson's algorithm, beginning with a random walk X started at u, followed by another walk Y started at v. Let τ_1 be the first time X hits η, let τ_2 be the first time Y hits $\mathrm{LE}(X^{\tau_1}) \cup \eta$, and let τ_2' be the first time Y hits η. Then

$$\begin{aligned}&\mathbb{P}^\eta\left(u \in M(x_i,r),\, v \in M(x_j,r)\right) = \mathbb{P}^\eta\left(u \in M(x_i,r)\right)\\ &\cdot \mathbb{P}^\eta\left(\mathrm{LE}(Y^{\tau_2}) \subseteq \Lambda(x,3r),\, |\mathrm{LE}(Y^{\tau_2})| \leq \frac{r^2}{(\log r)^{1/3}},\, \mathrm{LE}(Y^{\tau_2}) \text{ is } (\alpha,r)\text{-good} \;\middle|\; u \in M(x_i,r)\right).\end{aligned}$$

We have by the definition of $M(x_i, r)$ that if $u \in M(x_i, r)$ then $\mathrm{LE}(X^{\tau_1}) \subseteq \Lambda(x_i, r)$, so that if $\mathrm{LE}(Y^{\tau_2}) \subseteq \Lambda(x, 3r)$ then $\tau_2 = \tau_2'$. It follows that

$$\begin{aligned}&\mathbb{P}^\eta\left(v \in M(x_j,r) \mid u \in M(x_i,r)\right)\\ &\leq \mathbb{P}^\eta\left(\mathrm{LE}(Y^{\tau_2'}) \subseteq \Lambda(x,3r),\, \left|\mathrm{LE}(Y^{\tau_2'})\right| \leq \frac{r^2}{(\log r)^{1/3}},\, \mathrm{LE}(Y^{\tau_2'}) \text{ is } (\alpha,r)\text{-good} \;\middle|\; u \in M(x_i,r)\right)\\ &= \mathbb{P}^\eta\left(\mathrm{LE}(Y^{\tau_2'}) \subseteq \Lambda(x,3r),\, \left|\mathrm{LE}(Y^{\tau_2'})\right| \leq \frac{r^2}{(\log r)^{1/3}},\, \mathrm{LE}(Y^{\tau_2'}) \text{ is } (\alpha,r)\text{-good}\right)\\ &= \mathbb{P}^\eta\left(v \in M(x_j,r)\right)\end{aligned}$$

where the first equality follows because $Y^{\tau_2'}$ is independent from the event $\{u \in M(x_i,r)\}$ conditional on η and where the last equality follows by an application of Wilson's algorithm. The claimed negative correlation of $|M(x_i,r)|$ and $\left|M(x_j,r)\right|$ follows by summing over u and v. Negativity of the correlations immediately implies that

$$\mathrm{Var}^\eta\left(\left|\bigcup_{i=1}^K M(x_i,r)\right|\right) \leq \sum_{1\leq i\leq K} \mathrm{Var}^\eta(|M(x_i,r)|),$$

and we deduce by Chebyshev together with Propositions 72 and 73 that

$$\mathbb{P}^\eta\left(\left|\bigcup_{i=1}^K M(x_i,r)\right| \leq c_1 r^2 \sum_{i=1}^K \mathrm{Cap}(\eta\cap\Lambda(x_i,r))\right) \preceq \frac{r^2}{\log r}\cdot\frac{\sum_{i=1}^K \mathrm{Cap}(\eta\cap\Lambda(x_i,3r))}{\left(\sum_{i=1}^K \mathrm{Cap}\big(\eta\cap\Lambda(x_i,r)\big)\right)^2} \tag{6.39}$$

for some constant $c_1>0$. Note that this estimate holds for any $K\geq 1$ and any collection of points $(x_i)_{i=1}^K$ in $\mathbb{Z}^4$ such that the family of boxes $(\Lambda(x_i,3r))_{i=1}^K$ are mutually disjoint, where we are free to choose K and $(x_i)_{i=1}^K$ as functions of η if we wish. (Of course the points we choose must be conditionally independent of the rest of the UST given η.)

We now want to apply this estimate to prove our lower tail estimate on $|\mathfrak{B}(n)|$. Fix $n\geq 1$, and for each $R\geq 1$, let $\mathscr{A}_R$ be the event that $\|\eta_i\|_\infty \geq 2R$ for every $i\geq n/2$. Observe from the definitions that if $\mathscr{A}_R$ holds and $r\geq 2$ is such that $r^2(\log r)^{-1/3}\leq n/2$ and $3r\leq R$ then

$$|\mathfrak{B}(n)| \geq \left|\bigcup_{i=1}^K M(x_i,r)\right|$$

for any collection of points $x_1,\ldots,x_K$ in $\Lambda(R)$: the definition of the set $M(x_i,r)$ and the choice of r ensures the path connecting x to η is contained in $\Lambda(2R)$ and has length at most $n/2$, while the definition of $\mathscr{A}_R$ ensures that this path meets η within the first $n/2$ steps of η. Thus, choosing these points as a function of η and $r\geq 1$ as in the covering lemma, Lemma 78, we deduce from (6.39) that there exists a constant c_1 such that

$$\mathbb{1}(\mathscr{A}_R)\,\mathbb{P}^\eta\left(|\mathfrak{B}(n)|\leq c_1 r^2 \mathrm{Cap}(\eta\cap\Lambda(R))\right) \preceq \frac{r^2}{\log r}\cdot\frac{1}{\mathrm{Cap}(\eta\cap\Lambda(R))} \tag{6.40}$$

for every $r,R\geq 2$ such that $r^2(\log r)^{-1/3}\leq n/2$ and $3r\leq R$. As such, we have by a union bound that

$$\mathbb{P}\left(|\mathfrak{B}(n)|\leq \frac{c_1 r^2 R^2}{\lambda\log R}\right) \preceq \frac{\lambda r^2\log R}{R^2(\log r)} + \mathbf{P}(\mathscr{A}_R^c) + \mathbf{P}\left(\mathrm{Cap}(\eta\cap\Lambda(R))\leq \frac{R^2}{\lambda\log R}\right) \tag{6.41}$$

for every $r, R \geq 2$ such that $r^2(\log r)^{-1/3} \leq n/2$ and $3r \leq R$ and every $\lambda \geq 1$.

To proceed, we will bound the second and third terms on the right hand side them optimize over the choice of r, R, and λ. To bound $\mathbf{P}(\mathscr{A}_R)$, we use Wilson's algorithm to write

$$\begin{aligned}\mathbf{P}(\mathscr{A}_R^c) &= \mathbf{P}_0(\mathrm{LE}(X)_i \in \Lambda(2R) \text{ for some } i \geq n/2)\\ &\leq \mathbf{P}_0\left(\rho_{\lfloor n/2 \rfloor}(X) \leq \frac{n}{4(\log n)^{1/3}}\right) + \mathbf{P}_0\left(X_j \in \Lambda(2R) \text{ for some } j \geq \frac{n}{4(\log n)^{1/3}}\right)\\ &\preceq \frac{\log\log n}{(\log n)^{2/3}} + \frac{R^2(\log n)^{1/3}}{n},\end{aligned}$$

where the first term has been bounded using Theorem 61 and the second follows by a standard random walk computation (for example, it follows by [193, Lemma 4.4] and Markov's inequality). To bound the second term, we use the union bound

$$\mathbf{P}\left(\mathrm{Cap}(\eta \cap \Lambda(R)) \leq \frac{R^2}{\lambda \log R}\right) \leq \mathbf{P}(\|\eta_i\| \geq R \text{ for some } i \leq k) + \mathbf{P}\left(\mathrm{Cap}(\eta^k) \leq \frac{R^2}{\lambda \log R}\right)$$

for every $R, k \geq 1$ and $\lambda \geq 1$. Using Wilson's algorithm and a further union bound yields that

$$\begin{aligned}\mathbf{P}\left(\mathrm{Cap}(\eta \cap \Lambda(R)) \leq \frac{R^2}{\lambda \log R}\right) &\leq \mathbf{P}_0\left(\ell_k \geq 2k(\log k)^{1/3}\right) +\\ &\mathbf{P}_0(\|X_j\| \geq R \text{ for some } j \leq 2k(\log k)^{1/3}) + \mathbf{P}_0\left(\mathrm{Cap}(\mathrm{LE}(X)^k) \leq \frac{R^2}{\lambda \log R}\right),\end{aligned}$$

and we deduce from Theorem 61, the maximal version of Azuma-Hoeffding [269, Section 2], and Proposition 77 that there exists a positive constant C such that

$$\mathbf{P}\left(\mathrm{Cap}(\eta \cap \Lambda(R)) \leq \frac{R^2}{\lambda \log R}\right) \preceq \frac{\log\log k}{(\log k)^{2/3}} + \exp\left[-\Omega\left(\frac{R^2}{k(\log k)^{1/3}}\right)\right]$$

for every $R, k \geq 1$ such that $k(\log k)^{-2/3} \leq C\lambda^{-1}R^2(\log R)^{-1}$. If $\lambda \leq R^{1/2}$ then the maximal such k is of order $\lambda^{-1}R^2(\log R)^{-1/3}$ and it follows by calculus that

$$\mathbf{P}\left(\mathrm{Cap}(\eta \cap \Lambda(R)) \leq \frac{R^2}{\lambda \log R}\right) \preceq \frac{\log\log R}{(\log R)^{2/3}} + \exp\left[-\Omega(\lambda^{-1})\right]$$

for every $R \geq 3$ and $1 \leq \lambda \leq R^{1/2}$. Putting these estimates together yields that

$$\mathbb{P}\left(|\mathfrak{B}(n)| \leq \frac{c_1 r^2 R^2}{\lambda \log R}\right) \\ \leq \frac{\lambda r^2 \log R}{R^2 \log r} + \frac{\log\log n}{(\log n)^{2/3}} + \frac{R^2 (\log n)^{2/3}}{n^2} + \frac{\log\log R}{(\log R)^{2/3}} + \exp\left[-\Omega(\lambda^{-1})\right]$$

for every $r, R \geq 2$ such that $r^2(\log r)^{-1/3} \leq n/2$ and $3r \leq R$ and every $1 \leq \lambda \leq R^{1/2}$. Letting $\beta \geq 10$, taking $R = \lceil \beta^{-1} n^{1/2} (\log n)^{1/6} \rceil$, $r = \lceil \beta^{-2} n^{1/2} (\log n)^{1/6} \rceil$ and $\lambda = \beta$ yields that if $n \geq \beta^4$ then

$$\mathbb{P}\left(|\mathfrak{B}(n)| \leq \frac{c_2 n^2}{\beta^5 (\log n)^{1/3}}\right) \preceq \beta^{-1} + \frac{\log\log n}{(\log n)^{2/3}} + \beta^{-2} + \frac{\log\log n}{(\log n)^{2/3}} + \exp\left[-\Omega(\beta^{-1})\right] \\ \preceq \beta^{-1} + \frac{\log\log n}{(\log n)^{2/3}},$$

which implies the claim. □

It remains to prove our covering lemma for the capacity, Lemma 78. The proof, which exhibits and analyzes a greedy algorithm for constructing the desired set of balls, follows a standard strategy for proving covering lemmas of similar form.

Proof of Lemma 78. Consider the set of centres $\mathscr{C} = \{x \in (2r+1)\mathbb{Z}^4 : \mathrm{Cap}(\Lambda(x,r) \cap S) > 0\}$ and the partition of $\mathbb{Z}^4$ defined by $\mathscr{B} = \{\Lambda(x,r) : x \in \mathscr{C}\}$. Note that $x \in S + \Lambda(x,r)$ for each $x \in \mathscr{C}$, since otherwise the box $\Lambda(x,r)$ would not contain any points of S. Given $x \in \mathscr{C}$, we write $A[x] = \{y \in \mathscr{C} : \|y - x\|_\infty \leq 2r+1\}$ for the set of centres in $\mathscr{C}$ equal to or adjacent to x. We note the crude bound $\#A[x] \leq 3^4$. Similarly, we write $A^2[x] = \cup_{y \in A[x]} A[y]$ and note that $\#A^2[x] \leq 5^4$.

We will construct the sequence $(x_i)_{i=1}^K$ using a greedy algorithm. By subadditivity of capacity (which is an immediate consequence of the variational principle of Lemma 74), we know that

$$\mathbf{P}i := \sum_{x \in \mathscr{C}} \mathrm{Cap}(S \cap \Lambda(x,r)) \geq \mathrm{Cap}(S). \tag{6.42}$$

Define the list of centres $(x_i)_{i \geq 0} \subseteq \mathscr{C}$ as follows. Let $\mathscr{C}_0 = \mathscr{C}$, and for $i \geq 0$ such that $\mathscr{C}_i \neq \emptyset$, let

$$x_i = \arg\max\{\mathrm{Cap}(\Lambda(x,r) \cap S) : x \in \mathscr{C}_i\}; \qquad \mathscr{C}_{i+1} = \mathscr{C}_i \setminus A[x_i],$$

Write $I = \inf\{i \geq 0 : \mathscr{C}_i = \emptyset\}$ and define $\kappa_i = \mathrm{Cap}(\Lambda(x_i,r) \cap S)$ for $0 \leq i < I$. We claim that

$$\sum_{0 \leq i \leq n} \mathrm{Cap}(S \cap \Lambda(x_i, 3r)) \leq 15^4 \sum_{0 \leq j \leq n} \kappa_j \qquad \text{for every } n < I. \tag{6.43}$$

Fix $0 \le i < I$. We note that for any $y \in A[x_i]$, there exists a unique $0 \le j \le i$ such that $y \in \mathscr{C}_i \setminus \mathscr{C}_{i+1}$. By definition of κ_j and x_j, it must then hold that $\mathrm{Cap}(\Lambda(y,r) \cap S) \le \kappa_j$. By subadditivity of capacity, we can therefore write

$$\mathrm{Cap}(\Lambda(x_i,3r) \cap S) \le \sum_{y \in A[x_i]} \mathrm{Cap}(\Lambda(y,r) \cap S) \le \sum_{y \in A[x_i]} \sum_{j \le i} \kappa_j \mathbb{1}(y \in \mathscr{C}_j \setminus \mathscr{C}_{j+1}).$$

Observing that $\mathscr{C}_j \setminus \mathscr{C}_{j+1} \subseteq A[x_j]$ for $j < I$, we get

$$\mathrm{Cap}(\Lambda(x_i,3r) \cap S) \le \sum_{j \le i} \kappa_j \left|A[x_i] \cap A[x_j]\right|.$$

By switching the order of summation, we have

$$\sum_{0 \le i \le n} \mathrm{Cap}(\Lambda(x_i,3r) \cap S) \le \sum_{0 \le j \le n} \kappa_j \sum_{j \le i \le n} \left|A[x_i] \cap A[x_j]\right|.$$

Finally, $\left|A[x_i] \cap A[x_j]\right| \le |A[x_i]| \le 3^4$, and if $\left|A[x_j] \cap A[x_i]\right| \ne 0$, then $x_i \in A^2[x_j]$. The x_i are all distinct, and there are at most 5^4 elements in $A^2[x_j]$, and so the summations over i on the right hand side are bounded above by $3^4 \times 5^4 = 15^4$, thus proving the claim (6.43).

Next, observe that for $i \ge 0$

$$\sum_{x \in \mathscr{C}_i} \mathrm{Cap}(\Lambda(x,r) \cap S) \ge \mathbf{P}i - 3^4 \sum_{0 \le j \le i-1} \kappa_j.$$

Indeed, at stage i in the algorithm we remove at most 3^4 centres from $\mathscr{C}_i$ to give $\mathscr{C}_{i+1}$, and for each of these centres x, we must have $\mathrm{Cap}(S \cap \Lambda(x,r)) \le \kappa_i$. Putting $i = I$ in the above equation gives

$$3^4 \sum_{0 \le j < I} \kappa_j \ge \mathbf{P}i,$$

and so by (6.42), we have $\sum_{0 \le j < I} \kappa_j \ge 3^{-4} \mathrm{Cap}(S)$. □

Remark 12. Note that the proof of Lemma 78 does not use any properties of the capacity other than subadditivity and non-negativity, so that a similar covering lemma holds for any subadditive, non-negative set function.

6.3 Random walk

We now apply our main geometric theorem, Theorem 57, to study the behaviour of the random walk on the 4d UST. We begin by applying our results together with those of [197] to prove our effective resistance estimate, Theorem 59, in Section 6.3.1. In Section 6.3.2 we review the theory of Markov-type inequalities and prove our upper bound on the mean-squared displacement, Theorem 60. Finally, in Section 6.3.3 we show how the remaining estimates of Theorem 58 can be deduced from these estimates using the methods of [49, 227].

6.3.1 Effective Resistance

In this section we prove Theorem 59. The upper bound is trivial since resistances are always bounded by distances, so we focus on the lower bound. We will employ [197, Lemma 8.3] which we reproduce here. Let $\mathscr{C}_{\text{eff}}(A \leftrightarrow B; G) = \mathscr{R}_{\text{eff}}(A \leftrightarrow B; G)^{-1}$ denote the effective conductance between sets $A, B \subseteq V[G]$.

Lemma 79 [193], Lemma 8.3)**.** *Let T be a tree, let v be a vertex of T, and let $N_v(n,k)$ be the number of vertices $u \in \partial B(v,k) := B(v,k) \setminus B(v,k-1)$ at distance k from v such that u lies on a geodesic in T from v to $\partial B(v,n)$. Then*

$$\mathscr{C}_{\text{eff}}(v \leftrightarrow \partial B(v,n); T) \leq \frac{1}{k} N_v(n,k)$$

for every $1 \leq k \leq n$.

We will also use the following theorem of [197] concerning the tail of the intrinsic radius of the past.

Theorem 80 [197], Theorem 1.1)**.** *Let $\mathfrak{T}$ be the uniform spanning tree of $\mathbb{Z}^4$. Then*

$$\mathbb{P}(\mathfrak{P}(0,n) \neq \emptyset) \asymp \frac{(\log n)^{1/3}}{n}$$

for every $n \geq 1$.

We now apply these results together with Theorem 57 to prove Theorem 59.

Proof of Theorem 59. Fix $\lambda > 0$ and $\delta \in (0,1]$. For each $0 \leq m \leq n$, let $K(n,m)$ be the set of vertices $u \in \partial \mathfrak{B}(0,m)$ that lie on a geodesic from v to $\partial \mathfrak{B}(v,n)$ and let $K'(n,m)$ be the set of vertices $u \in \partial \mathfrak{B}(v,m)$ such that $\mathfrak{P}(u, n-m) \neq \emptyset$. We observe that $K(n,m) \setminus K'(n,m)$ contains at most one vertex, namely the unique vertex in $\partial \mathfrak{B}(v,m)$ which lies in the future of v, and so, by Lemma 79, we have

$$\mathscr{C}_{\text{eff}}(v \leftrightarrow \partial \mathfrak{B}(0,n); \mathfrak{T}) \leq \frac{1}{m} |K(n,m)| \leq \frac{1}{m} + \frac{1}{m} \left|K'(n,m)\right|$$

for each $1 \leq m \leq n$. Averaging this gives us that

$$\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial \mathfrak{B}(0,3n); \mathfrak{T}) \preceq \frac{1}{n} + \frac{1}{n^2} \sum_{m=n}^{2n} \left| K'(3n,m) \right|,$$

for each $n \geq 1$. Now, for each $n \geq 1$, the sets $(K'(n,m))_{n \leq m \leq 2n}$ are pairwise disjoint and their union satisfies

$$\bigcup_{n \leq m \leq 2n} K'(n,m) \subseteq \{u \in \mathbb{Z}^4 : u \in \mathfrak{B}(0,2n), \mathfrak{P}(u,n) \neq \emptyset\},$$

and so

$$\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial \mathfrak{B}(0,3n); \mathfrak{T}) \preceq \frac{1}{n} + \frac{1}{n^2} \sum_{u \in \mathbb{Z}^4} \mathbb{1}\big(u \in \mathfrak{B}(0,2n), \mathfrak{P}(u,n) \neq \emptyset\big).$$

Multiplying both sides by the indicator function $\mathbb{1}(|\mathfrak{B}(0,4n)| \leq \lambda^{1/2} n^2 (\log n)^{-1/3+\delta})$ and taking expectations gives

$$\mathbb{E}\left[\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial \mathfrak{B}(0,3n); \mathfrak{T}) \mathbb{1}\left(|\mathfrak{B}(0,4n)| \leq \frac{\lambda^{1/2} n^2}{(\log n)^{1/3-\delta}}\right)\right]$$
$$\preceq \frac{1}{n} + \frac{1}{n^2} \sum_{u \in \mathbb{Z}^d} \mathbb{P}\left(u \in \mathfrak{B}(0,2n), \mathfrak{P}(u,n) \neq \emptyset, |\mathfrak{B}(0,4n)| \leq \frac{\lambda^{1/2} n^2}{(\log n)^{1/3-\delta}}\right),$$

and applying the mass-transport principle to exchange the roles of 0 and u yields that

$$\begin{aligned}
&\mathbb{E}\left[\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial \mathfrak{B}(0,3n); \mathfrak{T}) \mathbb{1}\left(|\mathfrak{B}(0,4n)| \leq \frac{\lambda^{1/2} n^2}{(\log n)^{1/3-\delta}}\right)\right] \\
&\qquad \preceq \frac{1}{n} + \frac{1}{n^2} \sum_{u \in \mathbb{Z}^d} \mathbb{P}\left(0 \in \mathfrak{B}(u,2n), \mathfrak{P}(0,n) \neq \emptyset, |\mathfrak{B}(u,4n)| \leq \frac{\lambda^{1/2} n^2}{(\log n)^{1/3-\delta}}\right) \\
&\qquad \leq \frac{1}{n} + \frac{1}{n^2} \mathbb{E}\left[|\mathfrak{B}(0,2n)| \, \mathbb{1}\left(|\mathfrak{B}(0,2n)| \leq \frac{\lambda^{1/2} n^2}{(\log n)^{1/3-\delta}}, \mathfrak{P}(0,n) \neq \emptyset\right)\right] \\
&\qquad \preceq \frac{1}{n} + \frac{\lambda^{1/2}}{(\log n)^{1/3-\delta}} \mathbb{P}\big(\mathfrak{P}(0,n) \neq \emptyset\big) \preceq \lambda^{1/2} \frac{(\log n)^{\delta}}{n}, \qquad (6.44)
\end{aligned}$$

where the final inequality follows from Theorem 80. Now by a union bound, we have

$$\mathbb{P}\left(\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial\mathfrak{B}(0,3n);\mathfrak{T}) > \lambda\frac{(\log n)^{\delta}}{n}\right) \leq \mathbb{P}\left(|\mathfrak{B}(0,4n)| > \frac{\lambda^{1/2}n^2}{(\log n)^{1/3-\delta}}\right) + \mathbb{P}\left(\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial\mathfrak{B}(0,3n);\mathfrak{T})\mathbb{1}\left(|\mathfrak{B}(0,4n)| \leq \frac{\lambda^{1/2}n^2}{(\log n)^{1/3-\delta}}\right) > \lambda\frac{(\log n)^{\delta}}{n}\right).$$

Applying Markov's inequality to each term on the right hand side and using (6.44) and Theorem 57 to estimate the relevant expectations yields that

$$\mathbb{P}\left(\mathscr{C}_{\mathrm{eff}}(v \leftrightarrow \partial\mathfrak{B}(0,3n);\mathfrak{T}) > \lambda\frac{(\log n)^{\delta}}{n}\right) \preceq_{\delta} \lambda^{-1}\lambda^{1/2} + \lambda^{-1} \preceq \lambda^{-1/2}, \tag{6.45}$$

and the claim follows since $\lambda, \delta > 0$ were arbitrary. □

6.3.2 Upper bounds on displacement via Markov-type inequalities

In this section, we will use Markov-type inequalities [40, 127, 281] together with the results of [197] to prove Theorem 60, which establishes sharp upper bounds on the expectation of the squared maximal intrinsic displacement of a random walk on the 4d UST. Markov-type inequalities were first introduced by Ball [40] in the context of the Lipschitz extension problem, and have since been found to have many important applications to the study of random walk [149, 174, 244, 245, 292]. Our work is particularly influenced by that of James Lee and his coauthors [127, 149, 244, 245], who pioneered the use of Markov-type inequalities to prove sharp subdiffusive estimates for random walks on fractals. We begin by quickly reviewing the general theory, including in particular the extension of the universal Markov-type inequality for planar graphs of Ding, Lee, and Peres [127] to *unimodular hyperfinite* planar graphs established in [174].

Unimodular weighted graphs. A **vertex-weighted graph** is a pair (G,ω) consisting of a graph G and a **weighting** on G, that is a function $\omega : V[G] \to [0,\infty)$. We define the **weighted graph distance** between vertices x,y of a weighted graph (G,ω) by

$$d^G_{\omega}(x,y) = \inf_{x=u_0\sim\cdots\sim u_n=y, n\in\mathbb{N}} \sum_{i=1}^{n} \frac{1}{2}\big(\omega(u_i)+\omega(u_{i-1})\big).$$

Let $\mathscr{G}^{\omega}_{\bullet}$ be the space of triples (G,ω,ρ), where (G,ω) is a locally finite vertex-weighted graph, and $\rho \in V[G]$ is a vertex known as the **root** vertex. The space $\mathscr{G}^{\omega}_{\bullet}$ is equipped with the Borel sigma algebra induced by the natural generalisation of the Benamini-Schramm local topology [6, 110] in which two rooted, weighted graphs are considered to be close if there exist large graph-distance balls around their roots for which their respective balls admit a graph isomorphism that approximately preserves the weights. The details of this construction are not important to us and can be found in e.g. [110, Section 1.2]. Similarly, we also have the space $\mathscr{G}_{\bullet\bullet}^{\omega}$ of vertex-weighted graphs with an *ordered pair* of distinguished vertices. We say that a random variable (G,ω,ρ) taking values in $\mathscr{G}_{\bullet}^{\omega}$ is a **unimodular vertex-weighted graph** if it satisfies the **mass-transport principle**, i.e. if

$$\mathbb{E}\left[\sum_{v\in V[G]} F(G,\omega,\rho,v)\right] = \mathbb{E}\left[\sum_{v\in V[G]} F(G,\omega,v,\rho)\right]$$

for each Borel measurable function $F : \mathscr{G}_{\bullet\bullet}^{\omega} \to [0,\infty)$. Unweighted unimodular random graphs are defined similarly; we refer the reader to [6, 110] for a more in-depth discussion of the local topology and unimodularity. These notions are relevant to our setting since if K is the component of the origin in some translation-invariant random subgraph of $\mathbb{Z}^d$ then (K,ρ) always defines a unimodular random rooted graph, so that, in particular, $(\mathfrak{T},0)$ is a unimodular random rooted graph when $\mathfrak{T}$ is the UST of $\mathbb{Z}^4$. Moreover, if the weight $\omega : \mathbb{Z}^4 \to [0,\infty)$ is computed from $\mathfrak{T}$ in a translation-equivariant way then the resulting weighted random rooted graph $(\mathfrak{T},\omega,0)$ is also unimodular, as can be seen by applying the usual mass-transport principle on $\mathbb{Z}^4$ to the expectations $\mathbb{E}F(\mathfrak{T},\omega,x,y)$.

Markov-type inequalities. A metric space $\mathscr{X} = (\mathscr{X},d)$ is said to have **Markov-type** 2 with constant $c<\infty$ if for every finite set S, every irreducible reversible Markov chain M on S, and every function $f : S \to \mathscr{X}$ the inequality

$$\mathbb{E}\left[d\big(f(Y_0),f(Y_n)\big)^2\right] \leq c^2 n\mathbb{E}\left[d\big(f(Y_0),f(Y_1)\big)^2\right]$$

holds for every $n\geq 0$, where $(Y_i)_{i\geq 0}$ is a trajectory of the Markov chain M with Y_0 distributed as the stationary measure of M. Similarly, a metric space $\mathscr{X} = (\mathscr{X},d)$ is said to have **maximal Markov-type** 2 with constant $c<\infty$ if for every finite set S and every irreducible reversible Markov chain M on S, and every function $f : S \to \mathscr{X}$, we have that

$$\mathbb{E}\left[\max_{0\leq i\leq n} d\big(f(Y_0),f(Y_i)\big)^2\right] \leq c^2 n\mathbb{E}\left[d\big(f(Y_0),f(Y_1)\big)^2\right]$$

for each $n\geq 0$, where, as before, $(Y_i)_{i\geq 0}$ is a trajectory of the Markov chain M with Y_0 distributed as the stationary measure of M.

It is proved in [127] that there exists a universal constant C such that every vertex-weighted planar graph has Markov-type 2 with constant C; in fact their proof also establishes the existence of a universal constant C such that every weighted planar graph has *maximal* Markov-type 2 with constant C as explained in [174, Proposition 2.4]. (Presumably this fact is significantly easier to establish for trees than for other planar graphs, but we are not aware of an appropriate reference.)

We now describe the consequences of this theorem for unimodular random planar graphs. We must first define what it means for a unimodular random rooted graph to be *hyperfinite*. A **percolation** on a unimodular random rooted graph (G,ρ) is a labelling η of the edge set of G by the elements $0,1$ such that the resultant edge-labelled graph (G,η,ρ) is unimodular. We think of the percolation η as a random subgraph of G, where each edge is labelled 1 if it is included in the subgraph and 0 otherwise, and denote the connected component of ρ in this subgraph as $K_\eta(\rho)$. We say a percolation is finitary if $K_\eta(\rho)$ is almost surely finite, and say a unimodular random rooted graph (G,ρ) is **hyperfinite** if there exists an increasing sequence of finitary percolations $(\eta_n)_{n\geq 1}$ such that $\cup_{n\geq 1} K_{\eta_n}(\rho) = V[G]$ almost surely. The component of the origin in a translation-invariant random subgraph of $\mathbb{Z}^d$ is always hyperfinite as can be seen by taking a random hierarchical partition of $\mathbb{Z}^d$ into dyadic boxes. The following proposition appears as [174, Corollary 2.5].

Proposition 81. *Let (G,ρ) be a hyperfinite, unimodular random rooted graph with $\mathbb{E}[\deg(\rho)] < \infty$ that is almost surely planar, and suppose that ω is a vertex-weighting of G such that (G,ω,ρ) is a unimodular vertex-weighted graph. If Y is a random walk on G started at ρ then*

$$\mathbb{E}\left[\deg(\rho)\max_{0\leq i\leq n} d^G_\omega\big(Y_0,Y_i\big)^2\right] \leq C^2 n\mathbb{E}\left[\deg(\rho)\omega(\rho)^2\right],$$

for each $n\geq 1$, where C is a universal constant.

We now apply this proposition to prove Theorem 60.

Proof of Theorem 60. Let $r\geq 1$ be a parameter to be optimized over shortly. Seeing as the UST of $\mathbb{Z}^d$ is unimodular, hyperfinite (being a translation-invariant percolation processes on $\mathbb{Z}^d$) and planar, we can apply Proposition 110 to the vertex weight

$$\omega_r(v) = \mathbb{1}(\mathfrak{P}(v,r)\neq\emptyset),$$

which makes $(\mathfrak{T},\omega_r,0)$ unimodular since it is computed as a translation-equivariant function of $\mathfrak{T}$. This particular choice of weight is inspired by that used by Ganguly and Lee in [149].

Writing $d_r = d^{\mathfrak{T}}_{\omega_r}$ and using the fact that $\mathfrak{T}$ has degrees uniformly bounded below by 1 and above by 8, we get that

$$\mathbb{E}\left[\max_{0\leq i\leq n} d_r\big(Y_0,Y_i\big)^2\right] \leq 8C^2 n\,\mathbb{P}\left(\mathfrak{P}(0,r)\neq\emptyset\right) \tag{6.46}$$

for each $r,n\geq 1$. We next claim that

$$d_{\mathfrak{T}}(u,v)\leq 4r+4d_r(u,v) \qquad \text{for every } u,v\in\mathfrak{T} \text{ and } r\geq 1. \tag{6.47}$$

Indeed, let $u,v\in\mathbb{Z}^4$ and suppose that $d_{\mathfrak{T}}(u,v)\geq 4r$, the claimed inequality being trivial otherwise. Let w be the vertex at which the futures of u and v meet. At least one of the inequalities $d_{\mathfrak{T}}(u,w)\geq \frac{1}{2}d_{\mathfrak{T}}(u,v)$ or $d_{\mathfrak{T}}(v,w)\geq\frac{1}{2}d_{\mathfrak{T}}(u,v)$ holds, and we may assume without loss of generality that $d_{\mathfrak{T}}(u,w)\geq\frac{1}{2}d_{\mathfrak{T}}(u,v)\geq 2r$. Since u belongs to the past of each of the vertices in the $\mathfrak{T}$-geodesic connecting u to w, all the vertices in the second half of this geodesic must have past of intrinsic diameter at least r, so that $d_r(u,w)\geq\frac{1}{2}d_{\mathfrak{T}}(u,w)$ and hence that $d_r(u,v)\geq\frac{1}{4}d_{\mathfrak{T}}(u,v)$ as required. It follows from (6.47) together with (6.46) that

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq i\leq n} d_{\mathfrak{T}}\big(Y_0,Y_i\big)^2\right] &\leq 32r^2+2\mathbb{E}\left[\max_{0\leq i\leq n} d_r\big(Y_0,Y_i\big)^2\right]\\ &\preceq r^2+n\,\mathbb{P}\left(\mathfrak{P}(0,r)\neq\emptyset\right)\preceq r^2+\frac{n(\log r)^{1/3}}{r}\end{aligned}$$

for every $r,n\geq 1$, where we applied Theorem 80 in the third inequality, and taking $r=\lceil n^{1/3}(\log n)^{1/9}\rceil$ yields that

$$\mathbb{E}\left[\max_{0\leq i\leq n} d_{\mathfrak{T}}\big(Y_0,Y_i\big)^2\right]\preceq n^{2/3}(\log n)^{2/9}$$

for every $n\geq 2$ as claimed. □

Remark 13. This method also gives sharp upper bounds in dimensions $d\geq 5$: applying [193, Theorem 1.2] in place of Theorem 80, it yields that if $d\geq 5$, $\mathfrak{T}$ is the component of the origin in the uniform spanning forest of $\mathbb{Z}^d$, and Y is a random walk on $\mathfrak{T}$ started at 0, then

$$\mathbb{E}\left[\max_{0\leq i\leq n} d_{\mathfrak{T}}\big(Y_0,Y_i\big)^2\right]\preceq n^{2/3}$$

for every $n\geq 0$. This is stronger than the displacement upper bounds proven in [193], which were based on the results of [50].

6.3.3 Proof of Theorem 58

In this section we use all of the previous results to compute logarithmic corrections to the asymptotic behaviour of the displacement, exit times, return probabilities and range of the simple random walk on the uniform spanning tree. We will draw heavily on the methods of [227], which generalizes and synthesizes the earlier works [43, 49, 50]. Note that we must rederive all our results from the methods of [227] rather than simply quote their results since, as stated, these results do not allow for non-matching upper and lower bounds.

Remark 14. In this proof we will often use our big-O in probability notation on random variables indexed by more than one variable (e.g. n and r). When we write an expression $X_{n,r} = \mathbf{O}(Y_{n,r})$ of this form, it means that the entire family of associated random variables indexed by both n and r is tight.

Proof of Theorem 58. We recall that $\mathbf{E}_x^{\mathfrak{T}}$ denotes expectation with respect to the law of a simple random walk X on $\mathfrak{T}$ started at $x \in \mathbb{Z}^4$ conditional on $\mathfrak{T}$, and write $\mathbf{P}_x^{\mathfrak{T}}$ for the corresponding probability measure. Where clear from context, we will write $\mathbb{P}$ for the joint law and expectation of the uniform spanning tree and a random walk on the tree started at the origin, and similarly will write $\mathbb{E}$ for expectation with respect to this joint law.

Heat-kernel upper bound: [227, Proof of Proposition 3.1(a)] implies that

$$p_{2n}^{\mathfrak{T}}(0,0) + p_{2n+1}^{\mathfrak{T}}(0,0) \preceq \frac{1}{|\mathfrak{B}(0,R)|} \vee \frac{R}{n}$$

for every $n, R \geq 1$. Taking $R = n^{1/3}(\log n)^{1/9}$ and applying the volume lower bound of Theorem 57 therefore yields that

$$p_{2n}^{\mathfrak{T}}(0,0) = \mathbf{O}\left(\frac{(\log n)^{1/9}}{n^{2/3}}\right) \tag{6.48}$$

for every $n \geq 2$.

Intrinsic displacement lower bound: We have by Cauchy-Schwarz that

$$\begin{aligned}\mathbf{P}_0^{\mathfrak{T}}(d_{\mathfrak{T}}(0,X_n) \leq r) = \sum_{v \in \mathfrak{B}(0,r)} p_n^{\mathfrak{T}}(o,v) &\leq |\mathfrak{B}(0,r)|^{1/2} \left(\sum_{v \in \mathfrak{B}(R)} p_n^{\mathfrak{T}}(o,v)^2\right)^{1/2} \\ &\preceq |\mathfrak{B}(0,r)|^{1/2} p_{2n}^{\mathfrak{T}}(0,0)^{1/2} = \mathbf{O}\left(\frac{r}{(\log r)^{1/6-o(1)}} \cdot \frac{1}{n^{\frac{1}{3}}} (\log n)^{\frac{1}{18}}\right) \end{aligned} \tag{6.49}$$

for every $n, r \geq 1$, where we have applied the volume upper bound on Theorem 57, and the previously derived heat-kernel upper bounds. If we take $r = n^{1/3}(\log n)^{1/9-\delta}$ for some $\delta > 0$, then the expression appearing inside the $\mathbf{O}$ is $o(1)$, and, since this holds for every $\delta > 0$ (with implicit constants depending on δ), it follows that $d_{\mathfrak{T}}(0, X_n) = \boldsymbol{\Omega}(n^{1/3}(\log n)^{1/9-o(1)})$ for every $n \geq 2$ as claimed.

Intrinsic displacement upper bound: The estimate

$$d_{\mathfrak{T}}(X_0, X_n) \leq \max_{0 \leq m \leq n} d_{\mathfrak{T}}(X_0, X_n) = \mathbf{O}\left(n^{1/3}(\log n)^{1/9}\right)$$

follows immediately from Theorem 60.

Heat-kernel lower bound: Fix $\delta > 0$ and let $R = n^{1/3}(\log n)^{1/9+\delta}$. Using the same Cauchy-Schwarz argument as in (6.49), it follows from the intrisic displacement upper bounds of Theorem 60 and the volume lower bounds of Theorem 57 that there exists N_δ such that

$$p_{2n}^{\mathfrak{T}}(o,o) \geq \frac{(1 - \mathbf{P}^{\mathfrak{T}}(d_{\mathfrak{T}}(o, X_n) > R))^2}{|\mathfrak{B}(0,R)|} = \frac{1 - \mathbf{o}(1)}{\mathbf{O}\big(R^2(\log R)^{-1/3}\big)} = \frac{1 - \mathbf{o}(1)}{\mathbf{O}\big(n^{2/3}(\log n)^{-1/9+2\delta}\big)}$$

for every $n \geq N_\delta$, and the claim follows since $\delta > 0$ was arbitrary.

Exit time upper bound: [227, Equation 3.7] implies that

$$\mathbf{E}_0^{\mathfrak{T}}[\tau_R] \leq \mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow \mathfrak{B}(0,R)^c; \mathfrak{T})\, |\mathfrak{B}(0,R)| \leq R\, |\mathfrak{B}(0,R)|$$

for every $R \geq 1$, and applying Theorem 57 yields that

$$\mathbf{E}^{\mathfrak{T}}[\tau_R] = \mathbf{O}\left(\frac{R^3}{(\log R)^{1/3-o(1)}}\right) \qquad \text{and hence that} \qquad \tau_R = \mathbf{O}\left(\frac{R^3}{(\log R)^{1/3-o(1)}}\right)$$

for every $R \geq 2$.

Exit time lower bound: Fix $R \geq 1$, and let $\beta > 0$, $n = R^3/(\log R)^{1/3}$. Applying Theorem 60, we have

$$\mathbb{P}(\tau_R \leq \beta n) = \mathbb{P}\left(\max_{0 \leq i \leq \beta n} d_{\mathfrak{T}}(o, X_i)^2 \geq R^2\right) = O\left(\frac{\beta^{2/3} n^{2/3} (\log n)^{2/9}}{R^2}\right) = O(\beta^{2/3}),$$

and so $\tau_R = \boldsymbol{\Omega}(R^3/(\log R)^{1/3})$. The relation $\mathbf{E}^{\mathfrak{T}}[\tau_R] = \boldsymbol{\Omega}(R^3/(\log R)^{1/3})$ then follows.

Extrinsic displacement upper bound: Let $R \geq 1$ and fix $\delta > 0$. We have already established that

$$\max_{0 \leq m \leq n} d_{\mathfrak{T}}(X_0, X_n) = \mathbf{O}\left(n^{1/3}(\log n)^{1/9}\right),$$

and Theorem 57 tells us that

$$\mathfrak{B}(n) \subseteq \Lambda\left(n^{1/2}(\log n)^{1/6+\mathbf{o}(1)}\right) \qquad \text{as } n \to \infty.$$

Combining these two facts gives us

$$\max_{0 \leq m \leq n} \|X_m\|_\infty = \mathbf{O}\left(n^{\frac{1}{6}}(\log n)^{\frac{2}{9}+o(1)}\right) \qquad \text{as } n \to \infty,$$

as required.

Extrinsic displacement lower bound: Let $R \geq 1$. Exploiting the tree structure of $\mathfrak{T}$, we note that if $\max_{m \leq n} \|X_m\|_\infty \leq R$, then $\Gamma(0, X_m) \subseteq \Lambda(R)$. Thus, arguing as in (6.49), we have that

$$\begin{aligned}\mathbf{P}^{\mathfrak{T}}\left(\max_{m \leq n} \|X_m\|_\infty \leq R\right) &\preceq \left|\{x \in \mathbb{Z}^d : \Gamma(0,x) \subseteq \Lambda(R)\}\right|^{1/2} p^{\mathfrak{T}}_{2n}(o,o)^{1/2} \\ &= \mathbf{O}\left(\frac{R^2}{(\log R)^{1/2}} \cdot \frac{(\log n)^{1/18}}{n^{1/3}}\right),\end{aligned}$$

where the we have applied Proposition 65 and heat kernel upper bound (6.48) in the last line. This implies that $\max_{m \leq n} \|X_m\|_\infty = \boldsymbol{\Omega}(n^{1/6}(\log n)^{2/9})$ as claimed.

Range upper bound: Fix $\delta > 0$. For $n \geq 1$, let $D_n = \max_{0 \leq i \leq n} d_{\mathfrak{T}}(0, X_i)$. Applying displacement upper bounds and the volume upper bounds of Theorem 57, we have that

$$|\{X_m : 0 \leq m \leq n\}| \leq |\mathfrak{B}(D_n)| = \left|\mathfrak{B}(\mathbf{O}(n^{1/3}(\log n)^{1/9}))\right| = \mathbf{O}\left(\frac{n^{2/3}}{(\log n)^{1/9-o(1)}}\right)$$

as $n \to \infty$ as required.

Range lower bound: Fix $R \geq 1$, $\delta > 0$ and write $\mathfrak{B} = \mathfrak{B}(R)$.
Let $g_R(x,y) = (\deg_{\mathfrak{T}} y)^{-1} \mathbf{E}^{\mathfrak{T}}_x[\sum_{0 \leq i \leq \tau_R} \mathbb{1}(X_n = y)]$ and let $p(y) = g_R(0,y)/g_R(y,y)$ be the probability that a random walk started at $0 \in \mathfrak{T}$ hits y before exiting $\mathfrak{B}$. For each $y \in \mathfrak{B}' := \mathfrak{B}(\lfloor R/(\log R)^\delta \rfloor)$, we have $\mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow y; \mathfrak{T}) \leq R/(\log R)^\delta$, so that if the event $A = \{\mathscr{R}_{\mathrm{eff}}(0 \leftrightarrow$

$\mathfrak{B}^c; \mathfrak{T}) \geq R/(\log R)^{\delta/2}\}$ holds then

$$\inf_{y \in \mathfrak{B}'} \mathscr{R}_{\text{eff}}(y \leftrightarrow \mathfrak{B}^c; \mathfrak{T}) \geq \inf_{y \in \mathfrak{B}'} \left[\mathscr{R}_{\text{eff}}(0 \leftrightarrow \mathfrak{B}^c; \mathfrak{T}) - \mathscr{R}_{\text{eff}}(0 \leftrightarrow y; \mathfrak{T}) \right]$$
$$\geq R/(\log R)^{\delta/2} - R/(\log R)^{\delta} = \Omega(R/(\log R)^{\delta/2}).$$

Now for each $y \in \mathfrak{B}$ we have the following inequality which was derived for general graphs in [227, Proof of Proposition 3.2(b)]:

$$|1 - p(y)|^2 \leq \mathscr{R}_{\text{eff}}(0 \leftrightarrow y; \mathfrak{T}) \mathscr{R}_{\text{eff}}(y \leftrightarrow \mathfrak{B}^c; \mathfrak{T})^{-1}.$$

Taking the supremum over $y \in \mathfrak{B}' \subset \mathfrak{B}$ yields

$$\sup_{y \in \mathfrak{B}'} |1 - p(y)|^2 \leq \frac{R}{(\log R)^{\delta}} \cdot \sup_{y \in \mathfrak{B}'} \mathscr{R}_{\text{eff}}(y \leftrightarrow \mathfrak{B}^c; \mathfrak{T})^{-1} = O((\log R)^{-\delta/2})$$

on the event A. For each $R \geq 1$, consider the random variable $U_R = |\{X_i : 0 \leq i \leq \tau_R\} \cap \mathfrak{B}'|$. Then

$$\mathbf{E}_0^{\mathfrak{T}}[U_R] \geq \mathbf{E}_0^{\mathfrak{T}} \Big[\sum_{x \in \mathfrak{B}'} \mathbb{1}(X \text{ hits } x \text{ before exiting } \mathfrak{B}) \Big] =$$
$$\sum_{y \in \mathfrak{B}'} p(y) \geq \mathbb{1}(A)(1 - O((\log R)^{-\delta/2})) \left| \mathfrak{B}' \right|. \quad (6.50)$$

Now

$$\mathbb{P}\left(\frac{U_R}{|\mathfrak{B}'|} \leq 1/2 \right)$$
$$\leq \mathbb{E}\left[\mathbf{P}^{\mathfrak{T}} \left(A, \frac{U_R}{|\mathfrak{B}'|} \leq 1/2 \right) \right] + \mathbb{P}(A^c) = \mathbb{E}\left[\mathbf{P}^{\mathfrak{T}} \left(\mathbb{1}(A) \Big(1 - \frac{U_R}{|\mathfrak{B}'|} \Big) \geq 1/2 \right) \right] + \mathbb{P}(A^c),$$

and so applying (6.50) with Markov's inequality to the conditional probability inside the expectation gives

$$\mathbb{P}\left(\frac{U_R}{|\mathfrak{B}'|} \leq 1/2 \right) \leq O((\log R)^{-\delta/2}) \mathbb{P}(A) + \mathbb{P}(A^c) = o(1)$$

as $R \to \infty$, where the fact that $\mathbb{P}(A^c) \to 0$ as $R \to \infty$ follows from Corollary 59. The claim follows since $|\mathfrak{B}'| = \boldsymbol{\Omega}(R^2 (\log R)^{-1/3 - 2\delta})$, $\tau_R = \mathbf{O}(R^3 (\log R)^{-1/3 + o(1)})$, and $\delta > 0$ was arbitrary. $\square$

Chapter 7

[F] Uniqueness of the infinite tree in low-dimensional random forests

Abstract. The *arboreal gas* is the random (unrooted) spanning forest of a graph in which each forest is sampled with probability proportional to $\beta^{\#\text{edges}}$ for some $\beta \geq 0$, which arises as the $q \to 0$ limit of the Fortuin-Kastelyn random cluster model with $p = \beta q$. We study the infinite-volume limits of the arboreal gas on the hypercubic lattice $\mathbb{Z}^d$, and prove that when $d \leq 4$, any translation-invariant infinite volume Gibbs measure contains at most one infinite tree almost surely. Together with the existence theorem of Bauerschmidt, Crawford and Helmuth (2021), this establishes that for $d = 3,4$ there exists a value of β above which subsequential weak limits of the β-arboreal gas on tori have exactly one infinite tree almost surely. We also show that the infinite trees of any translation-invariant Gibbs measure on $\mathbb{Z}^d$ are one-ended almost surely in every dimension. The proof has two main ingredients: First, we prove a resampling property for translation-invariant arboreal gas Gibbs measures in every dimension, stating that the restriction of the arboreal gas to the trace of the union of its infinite trees is distributed as the uniform spanning forest on this same trace. Second, we prove that the uniform spanning forest of any translation-invariant random connected subgraph of $\mathbb{Z}^d$ is connected almost surely when $d \leq 4$. This proof also provides strong heuristic evidence for the conjecture that the supercritical arboreal gas contains infinitely many infinite trees in dimensions $d \geq 5$. Along the way, we give the first systematic and axiomatic treatment of Gibbs measures for models of this form including the random cluster model and the uniform spanning tree.

7.1 Introduction

For each $\beta \geq 0$, the **β-arboreal gas** (a.k.a. the *weighted uniform forest model*) on a finite undirected graph $G = (V,E)$ is a random subgraph A of G with probability mass function

$$\mathbb{P}_\beta(A=F) = \begin{cases} (1/Z_\beta)\beta^{|F|} & F \subseteq G \text{ is a spanning forest} \\ 0 & \text{otherwise} \end{cases}, \qquad Z_\beta = \sum_{F \subseteq G \text{ a spanning forest}} \beta^{|F|}, \tag{7.1}$$

where $|F|$ denotes the cardinality of the edge set of F and a *spanning forest* of G is an acyclic subgraph of G containing every vertex. Equivalently, the law of A is equal to the law of Bernoulli percolation on G with parameter $p = \beta/(1+\beta)$ conditioned to be acyclic. It is also equal to the $q \to 0$ limit of the q-state random cluster model with p/q converging to β [199, 289], while its $\beta \to \infty$ limit is equal to the uniform spanning tree when G is connected. (When $\beta = 1$, the model is a uniform random spanning forest of G; this value of the parameter plays no special role in our analysis.) The arboreal gas is also closely related to various supersymmetric spin systems, which has led it to receive substantial attention in the physics literature [95–97, 117]. Despite these connections, there are very few tools available to study the model and several very basic conjectures about its behaviour have remained open for twenty years [168]. See [56, 310] for surveys of the model and its connections to other topics.

Interest in the arboreal gas has grown significantly in recent years following the breakthrough works of Bauerschmidt, Crawford, Helmuth and Swan [55] and Bauerschmidt, Crawford and Helmuth [54], who studied the model's percolation phase transition through the lens of spontaneous symmetry breaking in an equivalent supersymmetric hyperbolic sigma model: In [55] they proved that the arboreal gas on $\mathbb{Z}^2$ never contains any infinite trees for any finite $\beta < \infty$, while in [54] they proved that the arboreal gas on $\mathbb{Z}^d$ contains infinite trees for sufficiently large values of β when $d \geq 3$. (Stochastic domination by percolation easily implies that the arboreal gas does not contain infinite trees for *small* values of β in any dimension.) Since it remains open whether the arboreal gas is stochasically monotone in β or in its boundary conditions, one must be careful to note some important subtleties in both statements: it is unclear whether there exist "canonical" definitions of the "infinite-volume arboreal gas" on $\mathbb{Z}^d$, and it is also unknown whether the existence of an infinite tree is monotone in β. A more precise statement of the results of [54, 55] is that any subsequential infinite-volume limit of the model on $\mathbb{Z}^2$ (with arbitrary boundary conditions) does not contain an infinite tree, while for $d \geq 3$ there exists $\beta_0 = \beta_0(d)$ such that if $\beta > \beta_0(d)$ then any subsequential limit of the model on large d-dimensional tori contain at least one infinite tree almost surely. The authors also establish strong quantitative control of the model, showing in particular that the finite-cluster two-point function continues to display critical-like behaviour in the supercritical regime. (Similar phenomena have also been shown to occur for the arboreal gas on the complete graph [255, 267] and on regular trees with wired boundary conditions [132, 296], where the analysis of the critical-like behaviour of finite/non-giant clusters is more complete.)

The analysis of [54, 55] tells us nothing about the *number* of infinite trees in the arboreal gas, which is the main subject of this paper. The analogous question has, however, been extensively studied for the *uniform spanning tree*. Indeed, the seminal paper of Pemantle [288] established that the uniform spanning tree of $\mathbb{Z}^d$ has a well-defined infinite-volume limit that is independent of the choice of boundary conditions and that is almost surely connected, i.e. a single tree, if and only if $d \leq 4$. This theorem was greatly generalized by Benjamini, Lyons, Peres, and Schramm [68] who proved that the *wired uniform spanning forest* (i.e. the infinite-volume limit of the uniform spanning tree with wired boundary conditions) of an infinite graph G is connected almost surely if and only if two independent random walks on G intersect infinitely often. This is known to occur for $G = \mathbb{Z}^d$ if and only if $d \leq 4$ by a classical theorem of Erdös and Taylor [134]. Since the uniform spanning tree is the $\beta \to \infty$ limit of the arboreal gas, it is natural to conjecture (see [54, Page 8]) that the same transition from uniqueness to non-uniqueness in four dimensions holds for the arboreal gas as in the uniform spanning tree.

In this paper we verify the low-dimensional case of this conjecture. Our proof also lends strong heuristic evidence to the high-dimensional case as we discuss later in the introduction.

Theorem 82. *For each $\beta > 0$ and $d \in \{3,4\}$, every translation-invariant β-arboreal gas Gibbs measure on the Euclidean lattice $\mathbb{Z}^d$ is supported on configurations that have at most one infinite tree.*

Here, *an arboreal gas Gibbs measure* on $\mathbb{Z}^d$ is any subsequential weak limit of arboreal gas measures on finite subgraphs of $\mathbb{Z}^d$ with (possibly random) boundary conditions; such Gibbs measures always exist by compactness, and *translation-invariant* Gibbs measures always exist by taking e.g. subsequential limits of the model with periodic boundary conditions. Let us stress that the structure of the set of Gibbs measures for the arboreal gas is very poorly understood, and, unlike the uniform spanning tree and ($q \geq 1$) random cluster model, it is not clear whether the free and wired infinite-volume measures are well-defined independently of the choice of exhaustion, or, for that matter, whether there is more than one Gibbs measure for the model at any value of β. Indeed, an important contribution of our paper is to develop the first systematic, axiomatic treatment of Gibbs measures for models of this form (where the weight of a configuration depends on its connectivity properties), as discussed in more detail below.

Remark 15. The proof of Theorem 82 also applies in dimensions $d \leq 2$, but the result is vacuous in this case since the model has no infinite clusters for any $\beta < \infty$ by the results of [55]. (While the main theorem of that paper is written only for subsequential limits of the model with free boundary conditions, the proof applies with arbitrary boundary conditions).

Theorem 82 has the following corollary in conjunction with the aforementioned results of [54] (translation-invariance being an automatic feature of subsequential limits of automorphism-invariant models on tori).

Corollary 83. *Fix a dimension $d \in \{3.4\}$ and $\beta > 0$, and for each $n \geq 1$ let $\mathbb{P}_n$ be a β-arboreal gas measure on the d-dimensional torus of side length n. There exists a constant $\beta_0 = \beta_0(d) > 0$ such that if $\beta > \beta_0$ then every subsequential weak limit of the sequence $(\mathbb{P}_n)_{n\geq 1}$ is supported on configurations that contain a unique infinite tree.*

Remark 16. Theorem 82 also implies an analogue of Corollary 83 for (subsequential) double limits of the model on the torus with an external field as considered in [54], where one first sends the size of the torus to infinity and then takes the external field to zero. This is because any such subsequential limit is a translation-invariant Gibbs measure for the model, as follows from a straightforward modification of the proof of Proposition 87.

About the proof. We now briefly overview the proof of Theorem 82. Unlike [54, 55], which exploit a non-probabilistic equivalence between the arboreal gas and a supersymmetric sigma model, our methods are purely probabilistic. Our argument can be divided into two parts, which we now describe in turn. Both parts of the proof lead to intermediate results of independent interest.

Augmented Gibbs measures and the resampling property. The first part of the paper, which is valid in any dimension, establishes a relationship between the infinite trees in the arboreal gas and the wired uniform spanning forest of a certain random subgraph of $\mathbb{Z}^d$. This part of the paper is mostly ergodic-theoretic in nature, and works by studying the properties of the space of translation-invariant Gibbs measures.

Theorem 84. *Let $d \geq 1$ and $\beta > 0$ and let A be distributed as a translation-invariant β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. If we define I_∞ to be the set of vertices that belong to the infinite components of A and define $\mathrm{Tr}(I_\infty)$ to be the subgraph of $\mathbb{Z}^d$ induced by I_∞ then the following hold:*

1. $\mathrm{Tr}(I_\infty)$ *is connected almost surely.*

2. *The conditional distribution of the restriction of A to $\mathrm{Tr}(I_\infty)$ given I_∞ and the restriction of A to $\mathrm{Tr}(I_\infty^c)$ is almost surely equal to the law of the wired uniform spanning forest of $\mathrm{Tr}(I_\infty)$.*

The second part of this theorem can be rephrased equivalently in terms of resampling: If we first sample the arboreal gas A then take F' to be a random variable sampled according to the law of the wired uniform spanning forest on $\mathrm{Tr}(I_\infty)$, then the forest formed from A by deleting all the infinite trees of A and adding in the trees of F' has the same distribution as A itself.

In the process of proving this theorem we develop a new axiomatic framework for infinite-volume Gibbs measures of the arboreal gas, with the usual DLR theory of Gibbs measures *not* being applicable to the arboreal gas due to a failure of 'quasilocality' of the Hamiltonian. Our replacement for this theory, which is developed in Section 7.2, revolves around what we term *augmented subgraphs*. Roughly speaking, this means that we enrich our random variables so that they include information about which vertices are connected to each other – possibly "through infinity" – outside of each finite set. We remark that previous papers on related models including the random cluster model and the uniform spanning tree have sidestepped the development of such a framework (in part because they tend to be focused on the free and wired measures, which we do not know are well-defined for the arboreal gas), and we are optimistic that the framework we develop will also be useful in the future study of those models. See Remarks 21 and 23 for further discussion.

It will already be clear to experts that the first part of Theorem 84 is a kind of Burton-Keane [94] theorem for the induced subgraph $\mathrm{Tr}(I_\infty)$. More interestingly, the second part of the theorem also hides a second Burton-Keane argument 'under the hood': To prove it, we first show that a similar resampling theorem holds where one replaces I_∞ by the infinite classes of the *augmented* connectivity relation (so that, *a priori*, one must sample the wired uniform spanning forest separately on the trace of each such class), before employing an "augmented" Burton-Keane argument to prove that there is in fact only one infinite augmented connectivity class almost surely.

This argument clearly demonstrates the utility of our perspective on the arboreal gas in terms of augmented subgraphs and augmented Gibbs measures. A further demonstration is given by the following theorem on the almost-sure one-endedness of infinite trees in the arboreal gas, which drops out neatly once the surrounding framework has been established. Here, an infinite tree is said to be **one-ended** if there is exactly one infinite simple path starting at each vertex. The same theorem has also been established for the uniform spanning tree via very different methods [68, 288].

Theorem 85. *Let $d \geq 1$ and $\beta > 0$ and let A be distributed as a translation-invariant β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. Then every infinite tree in A is one-ended almost surely.*

Remark 17. Theorem 84 allows us to import 'for free' various ergodic-theoretic theorems from the uniform spanning tree to the arboreal gas. For example, the indistinguishability theorem of [194] can be immediately applied to get that the infinite trees of the arboreal gas are indistinguishable when they exist, and a similar statement holds for the "multicomponent indistinguishability theorem" of [192]. This may be useful for studying more refined properties of the arboreal gas in high dimensions, as the multicomponent indistinguishability theorem plays an important role in the study of the adjacency structure of trees in the high-dimensional uniform spanning forest [67, 196].

Connectivity of the UST in low-dimensional unimodular random graphs. Theorem 84 reduces the study of the infinite trees in the arboreal gas to the study of the uniform spanning forest of the induced subgraph $\mathrm{Tr}(I_\infty)$, which is a translation-invariant random subgraph of $\mathbb{Z}^d$. When $d \geq 3$ and β is very large we have by the results of [54] that I_∞ has density very close to 1 (at least for subsequential limits of the arboreal gas on tori), so that it is reasonable to think of $\mathrm{Tr}(I_\infty)$ as a "small perturbation" of the original hypercubic lattice $\mathbb{Z}^d$. It seems very unlikely that this small perturbation would lead to any drastic difference in the behaviour of the random walk, which supports the conjecture that the number of infinite trees in the trees in the arboreal gas and uniform spanning tree should be the same, at least for β very large. Unfortunately it *is* possible in general for a high-density translation-invariant random induced subgraph of $\mathbb{Z}^d$ to have very different large-scale random walk behaviour than that of the full lattice, so that to implement this argument rigorously in the high-dimensional case one must use features of the arboreal gas beyond its translation invariance. The problem is made particularly delicate by the slow decay of correlations in the model [54], which make it difficult to compare $\mathrm{Tr}(I_\infty)$ to a better-understood model such as Bernoulli site percolation.

While we have not yet been able to circumvent this problem in the high-dimensional case, the low-dimensional case is more tractable since, informally, "the monotonicity goes in the right direction": we think of the connectivity of the wired uniform spanning forest (which, as previously mentioned, is equivalent to two independent random walks intersecting infinitely often almost surely) as a "small graph" property, so that it is plausibly preserved when taking "reasonable" subgraphs. Unfortunately, despite this intuition, it is still not literally true that every connected subgraph of $\mathbb{Z}^d$ has a connected wired uniform spanning forest when $d \leq 4$. Indeed, the subgraph of $\mathbb{Z}^3$ induced by the union of the origin with the two half-spaces $\{(x,y,z) : x > 0\}$ and $\{(x,y,z) : x < 0\}$ has two components in its wired uniform spanning forest almost surely. Moreover, it follows from a theorem of Thomassen [311, Theorem 3.3] that $\mathbb{Z}^d$ contains a transient tree for every $d \geq 3$, and it is easily seen that the wired uniform spanning forest of any such tree has infinitely many components almost surely.

The second part of the paper, which is specific to the low-dimensional case, establishes that, in contrast to these examples, the wired uniform spanning forest is always connected almost surely in any *translation-invariant* random subgraph of $\mathbb{Z}^d$ when $d \leq 4$. We state a simple special case of the relevant theorem now, with a significant generalization given in Theorem 106.

Theorem 86. *Let $d \leq 4$, let S be a translation-invariant random subset of $\mathbb{Z}^d$ and let* $\mathrm{Tr}(S)$ *be the subgraph of $\mathbb{Z}^d$ induced by S. Then the wired uniform spanning forest of each infinite connected component of* $\mathrm{Tr}(S)$ *is connected almost surely.*

The proof of this theorem draws mostly on random walk techniques, and is inspired in particular by previous work on *collisions* of random walks in unimodular random graphs [176, 195].

Remark 18. Translation-invariant random subgraphs of $\mathbb{Z}^d$ do *not* always have disconnected wired uniform spanning trees when $d \geq 5$, even when these graphs are induced by connected sets of vertices. (Indeed, starting with a random space-filling curve one can construct such a translation-invariant random induced subgraph that is a.s. rough-isometric to $\mathbb{Z}$.) This suggests that a more delicate approach is required to understand the number of infinite trees in the high-dimensional arboreal gas.

Remark 19. We believe that the theory we develop in this paper can be applied with minor modifications to prove analogous uniqueness theorems for a number of similar random forest models in dimensions $d \leq 4$. For example, it should apply to the variant of the arboreal gas in which the forest is required to contain at most one non-singleton component, which is a kind of 'dilute spanning tree' model.[1] Indeed, this model should actually be significantly simpler to study via our methods than the arboreal gas, since (in the language of Section 7.2) its Gibbs augmentations trivially have at most one non-singleton augmented connectivity class almost surely. The main (easily addressed) complication is that the definition of an augmented Gibbs measure needs to be modified so that the random variables are also enriched with the data of which finite subgraphs have a non-singleton component in their complement, and which boundary vertices (if any) belong to this component. We do not pursue such generalizations further in this paper.

[1] This model always has infinite-volume limits containing infinite trees when $\beta > 1$, even when $d = 1$. Indeed, in this regime the contribution to the partition function from a single spanning tree is larger than that from all configurations with a sublinear number of edges, so that most the contribution to the partition function comes from configurations with a linear number of edges. The actual critical value should be smaller than 1. This is related to the results of [120].

Remark 20. All our methods generalize immediately to arbitrary transitive graphs of at most four-dimensional volume growth. The resampling theorem, Theorem 84, can be extended much more generally to every *amenable* transitive graph. One noteworthy consequence of this is as follows: In [55], Bauerschmidt, Crawford, Helmuth, and Swan prove that the arboreal gas on $\mathbb{Z}^2$ cannot have a *unique* infinite tree for any $\beta < \infty$ (since the probability that x is connected to y is small when $x - y$ is large), then deduce that there are no infinite trees almost surely using a Burton-Keane argument *on the model's planar dual*. The first part of their argument does not use planarity, and also applies to quasi-transitive graphs such as slabs which are quasi-isometric to $\mathbb{R}^2$ but not planar. An appropriate generalization of our Theorem 82 can be used to replace the second part of their argument, so that the entire result holds without planarity.

7.2 Gibbs measures and augmented subgraphs

In this paper, we are primarily concerned with weak limits of finite-volume arboreal gas measures on infinite graphs G. In order to proceed, it is desirable to have an axiomatic characterization of these infinite-volume measures, which will make it easier to apply ergodic-theoretic arguments. Unfortunately, the usual DLR–Gibbs theory (as described in e.g. [145, 241]) is not applicable to these measures: given a limit measure μ, a random variable $A \sim \mu$ and a finite box $H \subset G$, the law of the restriction of A to H conditioned on $A \cap H^c$ cannot, *a priori*, be expressed as a function of $A \cap H^c$. This is because when we take the limit, connectivity information is lost and we do not know which infinite trees in A should be regarded as connected "through infinity" to which other infinite trees.

In this section, we develop an *augmented Gibbs framework* which rectifies this problem. A central idea is to make the appropriate long-range connectivity information available locally by enriching the space that our random variables are defined in. In the next section, we use this framework to prove the resampling property for translation-invariant Gibbs measures, Theorem 84.

Remark 21. As mentioned earlier, we believe that the theory of augmented Gibbs measures we develop here should be useful to the study of other probabilistic statistical physics models such as the uniform spanning tree and random cluster model, which are also incompatible with the standard DLR framework for the same reasons as in our setting. Indeed, is is notable that no abstract theory of Gibbs measures has previously been developed for these models despite their broad popularity. For example, in Glazman and Manolescu's work on the structure of the set of Gibbs measures for the random cluster model on $\mathbb{Z}^2$ [159], the authors consider only an (*a priori*) special class of Gibbs measures in which infinite clusters are always considered to be connected at infinity. As discussed in [159, Remark 1.5], considering only this restricted class of Gibbs measures has various downsides, including that this class is not (*a priori*) preserved under planar duality. Our definition of Gibbs measures for models of this form is given strong justification by the fact that it coincides with the set of all possible limits of the models in finite-volume, with arbitrary boundary conditions, and is more general than that of [159]. The two notions can be shown to coincide for the random cluster model in the translation-invariant case, but it is currently unclear whether the two notions will coincide without the assumption of translation-invariance. For the uniform spanning tree, a version of the Gibbs property was proposed by Sheffield [303], which has the non-standard property that it describes the conditional distribution of the restriction of the tree to a finite set given both what is outside the set *and how the points on the boundary of the set are connected inside the set*; our definition is more standard in that it describes the distribution of what is inside the set given information only about what is outside. Further discussion of how our theory applies to the UST appears in Remarks 22 and 23.

7.2.1 Definitions

We begin by setting up some necessary notation which will be used throughout the rest of the paper before defining augmented subgraphs and arboreal gas Gibbs measures.

Graph notation. For any graph $G=(V,E)=(V[G],E[G])$, and vertices $u,v\in V[G]$, we write $u\sim_G v$ if $\{u,v\}\in E[G]$, write $u\overset{G}{\longleftrightarrow} v$ if the vertices u and v are in the same connected component of G, and write $G(v)$ for the connected component of G containing v. For any graph G, write $\mathscr{S}(G)$ for the set of subgraphs of G (which we take to be pairs of subsets of V and E) and write $\mathscr{S}^f(G)$ for the set of finite subgraphs of G. We will always assume that all graphs G are *locally finite*, meaning that all their vertex degrees are finite. For any graph G, an increasing sequence of finite subgraphs of G whose union is the entire graph is called an **exhaustion** of G.

Finite-volume arboreal gas Gibbs measures. Let $G=(V,E)$ be a countable, locally finite graph $G=(V,E)$ and let $H\subset G$ be a finite subgraph of G. We define the **inner vertex boundary** ∂H to be the set of vertices of H that are incident to an edge of G that does not belong to H. (If H is an *induced* subgraph of G then ∂H is equal to the set of vertices of H that are adjacent to a vertex of $V[G]\setminus V[H]$.) For each set S we write $\mathscr{P}[S]$ for the set of equivalence relations on S, which we encode as functions $\phi: S\times S\to\{0,1\}$ such that $\phi(x,y)=1$ if and only if x and y are in the same equivalence class. For each $\phi\in\mathscr{P}(\partial H)$ and subgraph $H'\subseteq H$, we write H'/ϕ for the graph constructed by taking H' and identifying the sets of vertices in $V[H']\cap\partial H$ which belong to the same equivalence class of ϕ, deleting any self-loops created by this identification. These equivalence relations will serve as boundary conditions, keeping track of connectivity outside of H. We write $\mathscr{F}(H)$ for the set of spanning forests of H, i.e. the set of acyclic subgraphs of H containing every vertex of H and, given an equivalence relation $\phi \in \mathscr{P}[\partial H]$, we say a forest $F \in \mathscr{F}(H)$ **extends** ϕ if F/ϕ is acyclic. We write $\mathscr{F}(H,\phi) = \mathscr{F}(G,H,\phi)$ for the subset of forest subgraphs of H which extend ϕ and say that such a forest is an (H,ϕ)**-maximal spanning forest** if it contains every vertex of H and there is no edge in $E[H]$ which can be added to F to yield another element of $\mathscr{F}(H,\phi)$. We write $\mathscr{F}_T(H,\phi)$ for the set of (H,ϕ)-maximal spanning forests; when H/ϕ is connected, maximal spanning forests of H/ϕ are the same thing as spanning trees of H/ϕ.

For each $\beta \in [0,\infty)$, we define the **finite-volume β-arboreal gas Gibbs measure** on a finite subgraph H of G with **boundary condition** $\phi \in \mathscr{P}(\partial H)$ by

$$\mathbb{P}^{\phi}_{H,\beta}(F) = \mathbb{P}^{\phi}_{G,H,\beta}(F) = \begin{cases} (1/Z^{\phi}_{\beta})\beta^{|F|} & F \in \mathscr{F}(H,\phi) \\ 0 & \text{otherwise} \end{cases}, \qquad Z^{\phi}_{\beta} = \sum_{F \in \mathscr{F}(H,\phi)} \beta^{|F|}.$$

(In particular, when $\beta = 0$ this measure puts all its mass on the subgraph of H with no edges.) We remark that if every equivalence class of ϕ contains just a single element then this measure coincides with the free arboreal gas measure on H. We also define the **finite-volume ∞-arboreal gas Gibbs measure** on H with boundary condition ϕ by

$$\mathbb{P}^{\phi}_{H,\infty}(F) = \mathbb{P}^{\phi}_{G,H,\infty}(F) = \begin{cases} |\mathscr{F}_T(H,\phi)|^{-1} & F \in \mathscr{F}_T(H,\phi) \\ 0 & \text{otherwise,} \end{cases}$$

which is the weak limit of $\mathbb{P}^{\phi}_{H,\beta}$ as $\beta \to \infty$ and can be identified with the uniform measure on maximal spanning forests of H/ϕ. In particular, when H/ϕ is connected, this measure can be identified with the uniform spanning tree measure on H/ϕ. More generally, given $\beta \in [0,\infty]$, a finite subgraph $H \in \mathscr{S}^f(G)$, and a probability measure ν on $\mathscr{P}(\partial H)$, we write $\mathbb{P}^{\nu}_{H,\beta}$ for the measure with probability mass function

$$\mathbb{P}^{\nu}_{H,\beta}(F) = \sum_{\varphi \in \mathscr{P}(\partial H)} \nu(\varphi)\mathbb{P}^{\phi}_{H,\beta}(F),$$

which we call a **finite-volume β-arboreal gas Gibbs measure** with boundary condition ν. Probabilistically, this measure is the law of the configuration obtained by first sampling a random boundary condition according to the (arbitrary) distribution ν, then sampling the arboreal gas with this boundary condition. Considering random boundary conditions in this way has the advantage that it automatically makes all the sets of measures we consider convex.

The finite-volume version of the Gibbs property for these measures is as follows: Given a finite subgraph H and a probability measure ν on the set of equivalence relations on ∂H, let ϕ be a random variable with law ν and, given ϕ, let A be a random variable with conditional law $\mathbb{P}^{\phi}_{H,\beta}$, so that A has marginal law $\mathbb{P}^{\nu}_{H,\beta}$. If H' is a subgraph of H and we define an equivalence relation $\Phi(H')$ on $\partial H'$ by taking u and v to be in the same class of $\Phi(H')$ if they are connected in $(A \setminus E[H'])/\phi$, then

$$\mathbb{P}^{\nu}_{H,\beta}(A \cap H' = \cdot \mid A \setminus E[H'], \phi) = \mathbb{P}^{\Phi(H)}_{H',\beta}(A = \cdot). \tag{7.2}$$

In words, the conditional law of $A \cap H'$ given $A \setminus E[H']$ and ϕ is equal to $\mathbb{P}^{\Phi(H)}_{H',\beta}$. This identity is an immediate consequence of the definitions, and encapsulates the intuition that what happens outside of H' affects the distribution of A inside H' only in so far as it determines which boundary vertices of H' are connected to each other *outside* of H'. Note that (7.2) is exactly the same Gibbs property enjoyed by the random cluster model; most of the theory we develop in the rest of this section will also apply straightforwardly to any other model satisfying this same form of the Gibbs property in finite volume.

We now move on to defining the space of *augmented subgraphs*, which allow us to meaningfully extend the Gibbs property (7.2) to infinite-volume measures. To avoid trivialities, we take care to make sure all relevant definitions continue to work as expected in the case that G is finite or disconnected.

The space of augmented subgraphs. Let $G = (V, E)$ be a locally finite graph. We define an **augmented subgraph** of G to be a pair (S, Φ) where S is a subgraph of G and Φ is a collection $(\Phi(H) : H \in \mathscr{S}^f(G))$, where $\Phi(H)$ is an equivalence relation on ∂H for each $H \in \mathscr{S}^f(G)$, satisfying the consistency condition

$$\begin{gathered}\text{For every } H, K \in \mathscr{S}^f(G) \text{ with } H \subset K \text{ and } u, v \in \partial H,\\ u \text{ and } v \text{ are related in } \Phi(H) \text{ if and only if they are connected in } (S \cap K \setminus E[H])/\Phi(K),\end{gathered} \tag{Con}$$

where vertices that do not belong to a subgraph are considered to not be connected to any other vertex in that subgraph. We interpret $\Phi(H)$ as dictating connectivity *outside* of H: the consistency condition states that if two vertices in the boundary of $H \subseteq K$ are connected outside of H according to $\Phi(H)$, then these two vertices must also be connected outside of H according to $S \cap K$ and $\Phi(K)$, and vice versa. Given an augmented subgraph (S, Φ) of G, we define the **augmented connectivity relation** by $u \xleftrightarrow{(S,\Phi)} v := \Phi(\{u, v\})(u, v)$, where

here $\{u,v\}$ is the graph consisting of the vertices u,v and no edges, so that, by consistency,

$$u \xleftrightarrow{(S,\Phi)} v \text{ if and only if } u \text{ is connected to } v \text{ in } (H\cap S)/\Phi(H)$$

for each (and hence every) finite subgraph $H \in \mathscr{S}^f(G)$ containing both u and v.

We write $\mathscr{A}(G)$ for the space of augmented subgraphs of G, which we endow with its natural product topology and associated Borel sigma-algebra, so that $\mathscr{A}(G)$ is compact and the projection map $\pi : \mathscr{A}(G) \to \mathscr{S}(G)$ defined by $\pi : (S,\Phi) \mapsto S$ is continuous. We call an augmented subgraph (S,Φ) with underlying subgraph S an **augmentation** of S, and call Φ the **boundary map** of the augmentation (S,Φ). Every subgraph S of G admits boundary maps $\Phi_{\text{free}} = \Phi^S_{\text{free}}$ and $\Phi_{\text{wired}} = \Phi^S_{\text{wired}}$ defined by

$$\Phi_{\text{free}}(H)(u,v) = 1 \iff u \text{ and } v \text{ are connected in } S\setminus E[H] \tag{7.3}$$

and

$$\Phi_{\text{wired}}(H)(u,v) = 1 \iff \begin{array}{c} u \text{ and } v \text{ are connected in } S\setminus E[H] \text{ or both} \\ \text{belong to infinite connected components of } S\setminus E[H], \end{array} \tag{7.4}$$

which are distinct whenever S has more than one infinite connected component or more than one end. We call the resulting augmentations (A,Φ_{free}) and (A,Φ_{wired}) the **free** and **wired** augmentations of A. (We warn the reader that the relationship between these augmentations and the usual terminology for free and wired Gibbs measures for the uniform spanning tree is not as straightforward as one might hope; see Remark 23.) These augmentations are extremal in the sense that the equivalence classes of an arbitrary augmentation contain those of the free augmentation and are contained in those of the wired augmentation. In general a subgraph may admit a very large number of distinct augmentations.

Augmentations are determined by their tails. We now discuss a key property of augmented subgraphs that will be used throughout our analysis. Let (S,Φ) be an augmented subgraph of a locally finite graph G. The consistency property implies that if H and H' are two finite subgraphs of G with $H \subseteq H'$, then $\Phi(H)$ is determined by $\Phi(H')$ and S. In particular, if for each finite subgraph H of G we define

$$\Phi_H = (\Phi(K) : K \text{ is a finite subgraph of } G \text{ containing } H),$$

then the full augmented subgraph (S,Φ) is completely determined by the pair (S,Φ_H) for each finite subgraph H of G. *This gives us a well-defined notion of what it means to add or delete finitely many edges from an augmented subgraph* (S,Φ)*:* Given an augmented subgraph (S,Φ) and two disjoint finite sets of edges A and B, we define an augmented subgraph $(S,\Phi)\cup A\setminus B$ by taking H to be a finite subgraph of G containing both A and B and extending $(S\cup A\setminus B,\Phi_H)$ to a full augmented subgraph by consistency; it is easily verified that this definition does not depend on the choice of finite subgraph H.

Infinite-volume arboreal gas Gibbs measures. We now define infinite-volume Gibbs measures for the arboreal gas. (NB: Although we emphasize the infinite-volume case, the definition also works in finite volume.) Given a random augmented subgraph (A,Φ) of a countable, locally finite graph G and a finite subgraph H of G, we write $\mathscr{G}_H$ for the sigma-algebra generated by $A\setminus E[H]$ and Φ_H, which represents the data of the augmented subgraph that is determined 'outside of H'.

Definition 3. *Let G be a countable, locally finite graph and fix $\beta\in[0,\infty]$. We say that a probability measure $\mathbb{P}_\beta$ on $\mathscr{F}(G)$ is a **β-arboreal gas Gibbs measure of** G if there exists a probability measure $\mathbb{Q}_\beta$ on $\mathscr{A}(G)$, such that the following hold:*

1. *The pushforward $\pi_*\mathbb{Q}_\beta$ is equal to $\mathbb{P}_\beta$. In other words, if $(A,\Phi)\sim\mathbb{Q}_\beta$ then $A\sim\mathbb{P}_\beta$.*
2. *If (A,Φ) is a random variable distributed as $\mathbb{Q}_\beta$ and H is a finite subgraph of G, then the conditional law of $A\cap H$ given $\mathscr{G}_H$ is almost surely equal to $\mathbb{P}_{H,\beta}^{\Phi(H)}$.*

*We will refer to the second property as the **augmented Gibbs property**. We call any measure $\mathbb{Q}_\beta$ which satisfies these two properties a **Gibbs augmentation** of $\mathbb{P}_\beta$, and call any measure $\mathbb{Q}_\beta$ on $\mathscr{A}(G)$ satisfying the second of these two properties an **augmented β-arboreal gas Gibbs measure**.*

We will often refer to ∞-arboreal gas Gibbs measures as **uniform spanning tree Gibbs measures** or **uniform maximal spanning forest Gibbs measures** (the former terminology not always being appropriate when G is not connected).

This axiomatic definition has the advantage that it is well-suited to ergodic-theoretic techniques. That it is an *appropriate* definition is justified by the following alternative characterisation of infinite-volume arboreal gas measures, as presented in the introduction.

Proposition 87. *Let G be an infinite, countable, locally finite graph. For each $\beta\in[0,\infty]$, the β-arboreal gas Gibbs measures of G are exactly the subsequential weak limits of finite-volume β-arboreal gas Gibbs measures – with possibly random boundary conditions – on exhaustions of G.*

We note that for any $\beta\in[0,\infty]$, any exhaustion $(H_n)_{n\geq 0}$ of G and any sequence of probability measures on boundary conditions $(\nu_n)_{n\geq 1}$, the sequence of measures $(\mathbb{P}_{H_n,\beta}^{\nu_n})_{n\geq 1}$ will always have at least one subsequential weak limit by compactness of $\mathscr{A}(G)$.

Proof of Proposition 87. Fix $\beta \in [0,\infty]$. We first check that any β-arboreal gas Gibbs measure $\mathbb{P}_\beta$ is a subsequential weak limit of finite-volume β-arboreal gas Gibbs measures with possibly random boundary conditions. Let (A,Φ) be a random variable with the law of a Gibbs augmentation of $\mathbb{P}_\beta$ and let $(H_n)_{n\geq 1}$ be any exhaustion of G. By the Gibbs property, the law of A restricted to H_n is equal to the law of $\mathbb{P}^{\nu_n}_{H_n,\beta}$, where ν_n is the law of $\Phi(H_n)$, and so the weak limit of the sequence $(\mathbb{P}^{\nu_n}_{H_n,\beta})_{n\geq 1}$ of finite volume β-arboreal gas Gibbs measures with random boundary conditions is equal to $\mathbb{P}_\beta$.

We now show the converse. Let $(H_n)_{n\geq 1}$ be an exhaustion of G, let $(\nu_n)_{n\geq 1}$ be a sequence of probability measures on equivalence relations on ∂H_n, and suppose that the sequence $(\mathbb{P}^{\nu_n}_{H_n,\beta})$ converges to some limit measure $\mathbb{P}_\beta$. For each $n \geq 1$ let ϕ_n be an equivalence relation on ∂H_n with law ν_n, let A_n be a random variable with conditional law $\mathbb{P}^{\phi_n}_{H_n,\beta}$ (so that A_n has marginal law $\mathbb{P}^{\nu_n}_{H_n,\beta}$), and for each finite subgraph H of G define an equivalence relation $\Phi_n(H)$ on ∂H by setting

$$\Phi_n(H)(u,v) = \begin{cases} \mathbb{1}(u \text{ and } v \text{ are connected in } A_n/\phi_n(H)) & H \subseteq H_n \\ 1 & \text{otherwise.} \end{cases}$$

By compactness, taking a subsequence if necessary, (A_n,Φ_n) converges weakly to some random variable (A,Φ), where A has law $\mathbb{P}_\beta$. Using (7.2), one can check from the definitions that Φ is almost surely an augmentation of A and that the law of (A,Φ) is a Gibbs augmentation of $\mathbb{P}_\beta$, completing the proof. □

The uniform spanning tree. Let G be an infinite, connected, locally finite graph. For each finite subgraph H of G, we define the **free boundary condition** $\mathrm{f} = \mathrm{f}_H \in \mathscr{P}(\partial H)$ to be the equivalence relation whose classes all have cardinality one and define **wired boundary condition** $\mathrm{w} = \mathrm{w}_H$ on H to be the equivalence relation on ∂H in which all points are related. It was proven implicitly by Pemantle [288] that if $(H_n)_{n\geq 1}$ is any exhaustion of G by finite subgraphs then the two sequences $(\mathbb{P}^{\mathrm{f}}_{H_n,\infty})_{n\geq 1}$ and $(\mathbb{P}^{\mathrm{w}}_{H_n,\infty})_{n\geq 1}$ have well-defined weak limits that do not depend on the choice of exhaustion $(H_n)_{n\geq 1}$; these limits are known as the **free** and **wired uniform spanning forest** measures on G. It follows from the $\beta = \infty$ case of Proposition 87 that if G is a connected, locally finite graph then the free and wired uniform spanning forests on G are indeed Gibbs measures for the uniform spanning tree on G. Moreover, these two measures are always stochastically maximal and minimal among the set of all Gibbs measures for the uniform spanning tree on G as made precise in the following lemma.

Lemma 88. *Let G be a connected, locally finite graph and let $\mathbb{P}$ be a Gibbs measure for the uniform spanning tree on G. Then $\mathbb{P}$ is stochastically dominated by the free uniform* *spanning forest on G and stochastically dominates the wired uniform spanning forest on G. In particular, if the free and wired uniform spanning forest of G coincide then G has a unique Gibbs measure for the uniform spanning tree.*

Proof. Let $(V_n)_{n\geq 1}$ be an increasing sequence of subsets of $V[G]$ converging to $V[G]$, and for each $n \geq 1$, let $H_n = \mathrm{Tr}[V_n]$ be the subgraph of G induced by V_n. It follows from the negative associated theorem of Feder and Mihail [139] (see also [259, Theorem 4.6 and Exercise 10.8]) that the measure $\mathbb{P}^{\phi}_{H_n,\infty}$ is stochastically decreasing in ϕ in the sense that if ϕ_1, ϕ_2 are two equivalence relations with ϕ_1 a refinement of ϕ_2 then $\mathbb{P}^{\phi_1}_{H_n,\infty}$ stochastically dominates $\mathbb{P}^{\phi_2}_{H_n,\infty}$. It follows in particular that every measure of the form $\mathbb{P}^{V}_{H_n,\infty}$ is stochastically dominated by $\mathbb{P}^{\mathrm{f}}_{H_n,\infty}$ and stochastically dominates $\mathbb{P}^{\mathrm{w}}_{H_n,\infty}$. The claim follows by taking limits in light of this and Proposition 87. □

Remark 22. Pemantle [288] established implicitly that the free and wired uniform spanning forests of $\mathbb{Z}^d$ coincide for every $d \geq 1$. In general, a graph G has a unique Gibbs measure for the uniform spanning tree if and only if it does not admit any non-constant harmonic functions of finite Dirichlet energy [68], which holds in particular for every amenable transitive graph [259, Corollary 10.9] as well as in many nonamenable examples. See [259, Chapter 10] for detailed background.

Remark 23. Naively, one might like to say that the augmentation we need to put on the free uniform spanning forest to make its law into an augmented Gibbs measure is precisely the free augmentation as defined in (7.3), while the augmentation we need to wired uniform spanning forest to make its law into an augmented Gibbs measure is precisely the wired augmentation as defined in (7.4). This intuition is correct when G is, say, a 3-regular tree, but is false in general. Indeed, consider the hypercubic lattice $\mathbb{Z}^d$, where the free and wired uniform spanning forest measures coincide for every dimension $d \geq 1$ as discussed above. In one dimension (where the spanning tree is just the entire line), the correct augmentation to place on the infinite-volume uniform spanning tree is the free augmentation; using the wired augmentation does not work, since under this augmentation the conditional probability that any edge is present given that all other edges are present would be zero, not one. In dimensions two to four the infinite-volume limit is supported on configurations with a single one-ended tree, and there is no choice in how to define the augmentation. In dimension five and higher, where there are infinitely many one-ended trees, the correct augmentation to use is the wired augmentation; using the free augmentation does not work since the Gibbs property would imply that an edge connecting two distinct infinite trees must be present with probability 1. (In other examples, such as the free uniform spanning forest on the free product $\mathbb{Z}^5 * \mathbb{Z}_2$, neither the free nor the wired augmentations are appropriate.) As a historical note, let us remark that this subtlety in how to correctly define the Gibbs property for uniform spanning forests led to an error in the work of Burton and Pemantle [93] which was not discovered until a decade later by Lyons [257] and corrected in the work of Sheffield [303].

7.2.2 Translation-invariant Gibbs measures

In this section we refine our focus to translation-invariant Gibbs measures on $\mathbb{Z}^d$. In particular, we will discuss how each such Gibbs measure can be decomposed in terms of *extremal* translation-invariant Gibbs measures, which have better ergodicity properties. In the usual DLR–Gibbs formalism for (quasi)local systems such as the Ising model, it is a standard result that any Gibbs measure can be decomposed as a mixture of *tail-trivial* Gibbs measures, which assign probability 0 or 1 to any event in the tail-sigma algebra. Indeed, in this framework, the tail-trivial Gibbs measures are exactly the extremal points of the convex set of Gibbs measures and so the desired decomposition is an immediate corollary of Choquet's theorem. An analogous result also holds for translation-invariant Gibbs measures (see Remark 24), which can always be decomposed into a mixture of *ergodic* translation-invariant Gibbs measures; these are the measures that assign probability 0 or 1 to all translation-invariant events. While the first of these results translates directly to our setting, we were not able to prove the direct analogue of the second result, and instead prove a slightly weaker result that will suffice for our later applications.

Tail triviality. We begin by discussing *tail triviality*, where the relevant theory holds for arbitrary graphs. Let G be a countable, locally finite graph, and recall that for each finite subgraph H of G we define $\mathscr{G}_H$ to be the sigma-algebra of Borel sets E in $\mathscr{A}(G)$ such that an augmented subgraph (S,Φ)'s belonging to E is determined by $S\setminus H$ and $\Phi_H := (\Phi(H') : H'$ a finite subgraph of G containing $H)$. We define the **tail sigma-algebra** $\mathscr{T}$ on $\mathscr{A}(G)$ to be the intersection $\bigcap_H \mathscr{G}_H$ taken over all finite subgraphs H of G.

Lemma 89. *Let $G=(V,E)$ be a countable, locally finite graph, let $\beta\in[0,\infty]$, and let $\mathbb{Q}_\beta$ be an augmented β-arboreal gas Gibbs measure on G. If $X\in\mathscr{T}$ is a tail event with $\mathbb{Q}_\beta(X)>0$, then the conditional measure $\mathbb{Q}_\beta(\,\cdot\,|X)$ is an augmented β-arboreal gas Gibbs measure on G.*

Proof of Lemma 89. Let $\mathbb{Q}_X := \mathbb{Q}_\beta(\,\cdot\,|X)$. Since X is $\mathscr{G}_H$ measurable for each finite subgraph H of $\mathbb{Z}^d$, we have for each such subgraph and each subgraph F of H that

$$\mathbb{Q}_X(A\cap H=F\mid\mathscr{G}_H)=\mathbb{Q}_\beta(A\cap H=F\mid\mathscr{G}_H)\qquad\text{a.s.}$$

and hence by the augmented Gibbs property of $\mathbb{Q}_\beta$ that

$$\mathbb{Q}_X(A \cap H = F \mid \mathscr{G}_H) = \mathbb{P}_{H,\beta}^{\Phi(H)}(A \cap H = F) \qquad \text{a.s.}$$

for every finite subgraph H of $\mathbb{Z}^d$ and every subgraph F of H, which is precisely the augmented Gibbs property for $\mathbb{Q}_X$. $\square$

Corollary 90. *Let $G = (V,E)$ be a countable, locally finite graph and let $\beta \in [0,\infty]$. Every extremal element of the convex set of augmented β-arboreal gas Gibbs measures on G is tail-trivial in the sense that it gives every tail event probability* 0 *or* 1.

Proof of Corollary 90. If $\mathbb{Q}_\beta$ is a β-arboreal gas Gibbs measure and $X \in \mathscr{T}$ is such that $\mathbb{Q}(X) \in (0,1)$ then, by Lemma 89, we can write $\mathbb{Q}_\beta$ as a convex combination of β-arboreal gas Gibbs measures $\mathbb{Q}(\cdot) = \mathbb{Q}(\cdot|X)\mathbb{Q}(X) + \mathbb{Q}(\cdot|X^c)\mathbb{Q}(X^c)$. Clearly $\mathbb{Q}(\cdot|X)$ and $\mathbb{Q}(\cdot|X^c)$ are non-identical as they each assign a different probability to X, so that $\mathbb{Q}_\beta$ is not extremal. $\square$

Let $\mathscr{M}_\beta = \mathscr{M}_\beta(G)$ denote the set of all augmented β-arboreal gas Gibbs measures on G. Since $\mathscr{M}_\beta$ is a compact convex subspace of the space of all signed measures on $\mathscr{A}(\mathbb{Z}^d)$, which is a locally-convex topological vector space with respect to the weak (a.k.a. weak*) topology, we may apply Choquet's theorem [306] to get that for each $\mathbb{Q}_\beta \in \mathscr{M}_\beta$ there exists a measure ν on the set of extremal points $\mathrm{ext}(\mathscr{M}_\beta)$ such that

$$\mathbb{Q}_\beta(\cdot) = \int_{\mathrm{ext}(\mathscr{M}_\beta)} \mathbb{Q}'_\beta(\cdot)\, d\nu(\mathbb{Q}'_\beta).$$

Probabilistically, this means that every augmented β-arboreal gas Gibbs measure can be sampled by first sampling a random *tail trivial* augmented β-arboreal gas Gibbs measure of appropriate distribution, then sampling from this random tail-trivial measure. Unfortunately this result has limited applicability to our setting since we are interested primarily in the translation-invariant case, and it is not guaranteed that a translation-invariant augmented Gibbs measure decomposes as a mixture of *translation-invariant* tail-trivial augmented Gibbs measures.

Remark 24. One can use the Krein-Milman theorem [306] to prove that every extremal β-arboreal gas Gibbs measure can be expressed as a weak limit over finite-volume Gibbs measures with *non-random* boundary conditions. We omit the details of these arguments since we are interested primarily in the translation-invariant setting.

Translation invariance and ergodicity. We now fix a dimension $d \geq 2$ and, as usual, abuse notation by writing $\mathbb{Z}^d$ both for the set of d-tuples of integers and the hypercubic lattice considered as a graph, writing E_d for the associated set of nearest-neighbour edges in $\mathbb{Z}^d$. For each $x \in \mathbb{Z}^d$, we define the translation operator τ_x on subgraphs of $\mathbb{Z}^d$ as

$$\tau_x\big((V,E)\big) = (\{v+x : v \in V\}, \{\{v_1+x, v_2+x\} : \{v_1,v_2\} \in E\}).$$

For each $x \in \mathbb{Z}^d$, τ_x also acts on augmented subgraphs via $\tau_{-x}(S,\Phi) = (\tau_{-x}S, \tau_{-x}\Phi)$ where $[\tau_{-x}\Phi](H)(u,v) = \Phi(H+x)(u+x,v+x)$. **Translation-invariant events** in, and **translation-invariant measures** on $\mathscr{S}(G)$ and $\mathscr{A}(G)$ are then defined as expected with respect to these operations. We write $\mathscr{I}$ for the sigma-algebra of translation-invariant events in $\mathscr{A}$ and write $\mathscr{I}_S$ for the sigma-algebra of translation-invariant events in $\mathscr{A}$ depending only on the subgraph coordinate (that is, for which any two augmentations of the same subgraph either both belong to the event or both belong to its complement).

The following lemma implies that if we wish to study translation-invariant Gibbs measures, it suffices to consider translation-invariant *augmented* Gibbs measures.

Lemma 91. *Fix $d \geq 1$, $\beta \in (0,\infty]$, and let $\mathbb{P}_\beta$ be a β-arboreal gas infinite-volume Gibbs measure on $\mathbb{Z}^d$. Then $\mathbb{P}_\beta$ is translation-invariant if and only if it admits a translation-invariant Gibbs augmentation.*

Proof. The 'if' direction is trivial; we focus on the 'only if' direction, which follows from the amenability of $\mathbb{Z}^d$. Let $\mathbb{P}$ be a translation-invariant infinite volume Gibbs measure and let (A,Φ) have the law of an augmentation of $\mathbb{P}$. For each $n \geq 1$, let V_n be a uniformly chosen vector in $\Lambda(n)$, and consider the sequence of random variables $(\tau_{V_n}A, \tau_{V_n}\Phi)_{n\geq 1}$. Taking a subsequential weak limit yields a translation-invariant random variable (A',Φ') whose law is a Gibbs augmentation of $\mathbb{P}$. (Alternatively, one can check that for each β-arboreal gas Gibbs measure $\mathbb{P}_\beta$ on $\mathbb{Z}^d$, the set of Gibbs augmentations of $\mathbb{P}_\beta$ is a weakly compact convex subset of the space of probability measures on augmented subgraphs of $\mathbb{Z}^d$. When $\mathbb{P}_\beta$ is translation-invariant this set is fixed by the action of $\mathbb{Z}^d$, and therefore must contain a fixed point since $\mathbb{Z}^d$ is amenable.) $\square$

We write $\mathscr{M}_\beta^T = \mathscr{M}_\beta^T(\mathbb{Z}^d)$ for the set of translation-invariant β-arboreal gas Gibbs measures on $\mathbb{Z}^d$, which is a weakly closed, convex set of the space of all signed measures on $\mathscr{A}(\mathbb{Z}^d)$. Applying Choquet's theorem as above yields that every element of $\mathscr{M}_\beta^T$ can be written as a mixture of its extremal points: For each $\mathbb{Q}_\beta \in \mathscr{M}_\beta^T$ there exists a measure ν on the set of extremal points $\mathrm{ext}(\mathscr{M}_\beta^T)$ such that

$$\mathbb{Q}_\beta(\cdot) = \int_{\mathrm{ext}(\mathscr{M}_\beta^T)} \mathbb{Q}'_\beta(\cdot)\, d\nu(\mathbb{Q}'_\beta).$$

In the standard quasilocal DLR–Gibbs theory, one would then argue that every element of $\mathrm{ext}(\mathscr{M}_\beta^T)$ is ergodic, meaning that it assigns probability 0 or 1 to every invariant event in $\mathscr{A}$. Unfortunately, the standard proof of this fact breaks down in our setting. More specifically, it is not clear whether the translation-invariant sigma-algebra is always contained in the completion of the tail sigma-algebra. Nevertheless, we do still have that extremal translation-invariant Gibbs measures are trivial on the intersection of the tail and invariant sigma algebras:

Lemma 92. *Fix $d \geq 1$ and $\beta \in [0,\infty]$. If $\mathbb{Q}_\beta \in \mathscr{M}_\beta^T$ is a translation-invariant augmented β-arboreal gas Gibbs measure and $X \subseteq \mathscr{A}$ is an event belonging to the $\mathbb{Q}_\beta$-completions of both $\mathscr{T}$ and $\mathscr{I}$ with $\mathbb{Q}_\beta(X) > 0$ then $\mathbb{Q}_\beta(\cdot|X)$ is also a translation-invariant β-arboreal gas Gibbs measure.*

Proof. Since X is in the completion of $\mathscr{T}$, there exists an event $X' \in \mathscr{T}$ with $\mathbb{Q}_\beta(X\Delta X') = 0$ and hence with $\mathbb{Q}_\beta(\cdot|X) = \mathbb{Q}_\beta(\cdot|X')$, so that Lemma 89 implies that $\mathbb{Q}_\beta(\cdot|X)$ is an augmented β-arboreal gas Gibbs measure. Similarly, since X is in the completion of $\mathscr{I}$, there exists an event $X'' \in \mathscr{I}$ such that $\mathbb{Q}_\beta(\cdot|X) = \mathbb{Q}_\beta(\cdot|X'')$, and one may verify from the definitions that $\mathbb{Q}_\beta(\cdot|X'')$ is translation-invariant since both $\mathbb{Q}_\beta$ and X'' are. □

Corollary 93. *Fix $d \geq 1$ and $\beta \in [0,\infty]$. If $\mathbb{Q}_\beta \in \mathrm{ext}(\mathscr{M}_\beta^T)$ is an extremal translation-invariant augmented β-arboreal gas Gibbs measure and $X \subseteq \mathscr{A}$ is an event belonging to the $\mathbb{Q}_\beta$-completions of both $\mathscr{T}$ and $\mathscr{I}$ then $\mathbb{Q}_\beta(X) \in \{0,1\}$.*

This corollary together with the next lemma implies that the sigma-algebra $\mathscr{I}_S$ of translation-invariant events that are insensitive to the choice of augmentation is always trivial for any extremal translation-invariant augmented Gibbs measure. This is a (slightly unsatisfactory) analogue of the statement in the standard DLR–Gibbs theory that extremal translation invariant measures are ergodic.

Lemma 94. *Fix $d \geq 1$, $\beta \in (0,\infty]$, and let $\mathbb{Q}_\beta$ be a translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. Then $\mathscr{I}_S$ is contained in the $\mathbb{Q}_\beta$-completion of $\mathscr{T}$. That is, for any translation-invariant $X \in \mathscr{I}_S$, there exists $Y \in \mathscr{T}$ such that $\mathbb{Q}_\beta(X\Delta Y) = 0$.*

Proof of Lemma 94. Let (A,Φ) be distributed as $\mathbb{Q}_\beta$ and for each $n \geq 1$ let Λ_n be the box $[-n,n]^d$ considered as a subgraph of $\mathbb{Z}^d$. By definition of the product Borel sigma-algebra, $\sigma(A)$ is generated by the union $\bigcup_H \sigma(A \cap H)$, where this union is taken over all finite subgraphs H of $\mathbb{Z}^d$. Since $\mathscr{I}_S = \mathscr{I} \cap \sigma(A) \subseteq \sigma(A)$, it follows from the Dynkin $\pi - \lambda$ theorem that for every event $X \in \mathscr{I}_S$ and every $\varepsilon > 0$ there exists a finite subgraph H of $\mathbb{Z}^d$ and an event $X' \in \sigma(A \cap H)$ such that $\mathbb{Q}_\beta(X'\Delta X) \leq \varepsilon$. Fix an event $X \in \mathscr{I}_S$ and for each $n \geq 1$ let $H_n \in \mathscr{S}^f(\mathbb{Z}^d)$ and $X_n \in \sigma(A \cap H_n)$ be such that $\mathbb{Q}_\beta(X \Delta X_n) \leq 2^{-n}$. For each $n \geq 1$, let $X'_n = \tau_{x_n}(X_n)$, where $x_n \in \mathbb{Z}^d$ is such that $\tau_{x_n}(H_n)$ is disjoint from Λ_n. We observe that $\mathbb{Q}(X \Delta X'_n) = \mathbb{Q}(X \Delta X_n) \leq 2^{-n}$ by translation-invariance of X and $\mathbb{Q}_\beta$, and moreover that $X'_n \in \sigma(A \setminus \Lambda_n) \subseteq \mathscr{G}_{\Lambda_n}$ for every $n \geq 1$. Letting $X'' = \limsup X'_n := \cap_{n \geq 1} \cup_{m \geq n} X'_m$ be the event that infinitely many of the events X'_n hold, we have that $X'' \in \mathscr{T}$ and that

$$\mathbb{Q}(X \cap X'') \leq \mathbb{Q}\Big(X \Delta X'_n \text{ holds for infinitely many } n\Big) \leq \lim_{n \to \infty} \sum_{m \geq n} 2^{-m} = 0,$$

which completes the proof. □

Remark 25. This proof does *not* straightforwardly extend to show that $\mathscr{I}$ is contained in the completion of $\mathscr{T}$ due to the long-range dependencies encoded in the boundary map. It would be possible to run the proof if one knew that $\sigma(A)$ and $\mathscr{T}$ together generate the entire sigma algebra on $\mathscr{A}(G)$, but this seems to be a surprisingly subtle matter.

We deduce the following immediate corollary.

Corollary 95. *Fix $d \geq 1$ and $\beta \in [0, \infty]$. If $\mathbb{Q}_\beta \in \mathrm{ext}(\mathscr{M}_\beta^T)$ is an extremal translation-invariant augmented β-arboreal gas Gibbs measure then $\pi_* \mathbb{Q}_\beta$ is an ergodic translation-invariant β-arboreal gas Gibbs measure.*

Remark 26. We will later prove in Corollary 103 that if (A, Φ) is distributed as an a translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$ with $\beta < \infty$, the boundary map Φ is almost surely equal to the wired boundary map associated to A, and hence coincides a.s. with a measurable function of A. Moreover, the boundary map also coincides a.s. with a measurable function of A in the case $\beta = \infty$ as discussed in Remark 23. As such, it follows *a posteriori* (see Corollary 104) that the completions of the sigma-algebras $\mathscr{I}$ and $\mathscr{I}_S$ are equal, and hence that every measure in $\mathrm{ext}(\mathscr{M}_\beta^T)$ is ergodic. Let us stress however that this proof uses specific properties of the arboreal gas (and, implicitly, the amenability of $\mathbb{Z}^d$), in contrast to the other proofs of this section which apply without change to a very large class of models with connection-based interactions. Moreover, the logical structure of the paper means that we cannot assume true ergodicity in the proof of Theorem 84 since this ergodicity is established only at the very end of Section 7.3.

Remark 27. It follows by standard arguments that the extremal elements of the set of *all* translation-invariant measures on $\mathscr{A}(\mathbb{Z}^d)$ are ergodic, and hence by Choquet theory that every translation-invariant measure on $\mathscr{A}(\mathbb{Z}^d)$ can be written as a mixture of ergodic translation-invariant measures. This statement is of limited use to us since we prefer to stay within the class of augmented arboreal gas Gibbs measures.

7.3 Proof of Theorems 84 and 85

In this section we use the framework developed in the previous section to prove Theorems 84 and 85. We begin with Theorem 84, whose proof is split into two propositions. The first, proven in Section 7.3.1, establishes a 'local' version of the same resampling theorem that does not require the symmetry of $\mathbb{Z}^d$, while the second, proven in Section 7.3.2, establishes the basic qualitative features of the augmented connectivity relation for augmented arboreal gas Gibbs measures on $\mathbb{Z}^d$. As a part of the proof of Section 7.3.2 we prove Theorem 85, which states that all the infinite trees in the arboreal gas are one-ended almost surely.

7.3.1 Resampling without symmetry

In this section we prove the following proposition, which establishes a very general version of the resampling property that does not require any symmetry assumptions on the graph or the measure. This proposition is inspired in part by the UST resampling theorem of Lyons, Peres, and Sun [262].

Proposition 96. *Let $G=(V,E)$ be a connected, locally finite graph, let o be a vertex of G, and let (A,Φ) be distributed as an augmented β-arboreal gas Gibbs measure on G. Let $I_o=\{x\in V : o \xleftrightarrow{(A,\Phi)} x\}$ and let $\mathrm{Tr}(I_o)$ be the subgraph of G induced by I_o. Then the conditional distribution of the restriction of A to I_o given I_o and the restriction of A to the complement of I_o is almost surely equal to some Gibbs measure for the uniform maximal spanning forest on $\mathrm{Tr}(I_o)$, where the choice of Gibbs measure may be random.*

Proof of Proposition 96. We begin by observing that a related resampling property holds in finite volume. Let H be a finite subgraph of G, so that $\Phi(H)$ is an equivalence relation on ∂H. For each forest $F\in\mathscr{F}(H,\Phi(H))$, let $T_o[F]$ be the connected component of o in F considered as a subgraph of $H/\Phi(H)$, let $I_o[F]=I_o^{H,\Phi(H)}[F]$ be the vertex set of $T_o[F]$, and let $\mathrm{Tr}(I_o[F])$ be the subgraph of $H/\Phi(H)$ induced by $I_o[F]$. We make three observations. First, note that $T_o[F]$ is always a spanning tree of $\mathrm{Tr}(I_o[F])$. Second, note that if we let T' be any other spanning tree of $\mathrm{Tr}(I_o[F])$ and let F' be formed from F by deleting $T_o[F]$ and adding T', then $I_o[F']=I_o[F]$. Finally, we observe that the probability $\mathbb{P}_{H,\beta}^{\phi_n}$ assigns to forests $F\in\mathscr{F}(H,\phi_n)$ depends only on the cardinality of their edge sets, so that $\mathbb{P}_{H,\beta}^{\phi_n}(F)=\mathbb{P}_{H,\beta}^{\phi_n}(F')$. Putting these observations together gives that if $F\sim\mathbb{P}_{\Lambda_n,\beta}^{\Phi(H)}$, then conditional on $I_o[F]$ and the restriction of F to the complement of $I_o[F]$, the restriction of F to $I_o[F]$ is distributed as the uniform spanning tree on $\mathrm{Tr}(I_o[F])/\Phi(H)$.

By the augmented Gibbs property, it follows that the conditional distribution of the restriction of A to $I_o[A\cap H]=I_o^{H,\Phi(H)}[F]$ given $\mathscr{G}_H$, $I_o[A\cap H]$, and the restriction of A to the complement of $I_o[A\cap H]$ is almost surely equal to the uniform spanning tree measure on $\mathrm{Tr}(I_o[A\cap H])/\Phi(H)$. In particular, this conditional distribution depends only on $\Phi(H)$ and $I_o[A\cap H]$. Moreover, the consistency property of the boundary map Φ implies that $I_o[A\cap H]=I_o^{H,\Phi(H)}[A\cap H]$ is equal to the intersection of I_o with the vertex set of H. Thus, if for each finite subgraph H of G we define $\mathscr{F}_H$ to be the sigma-algebra generated by $\mathscr{G}_H$, $I_o\cap V[H]=I_o^{H,\Phi(H)}[A\cap H]$, and the restriction of A to the complement of I_o, then the conditional law of the restriction of A to $I_o\cap V[H]$ given $\mathscr{F}_H$ is a.s. equal to the uniform spanning tree measure on $\mathrm{Tr}(I_o\cap H)/\Phi(H)$. Since this law depends only on $I_o\cap H$ and $\Phi(H)$, it follows that the conditional law of the restriction of A to $I_o\cap V[H]$ given I_o and the restriction of A to the complement of I_o is almost surely of the form $\mathbb{P}^{\nu}_{\mathrm{Tr}(I_o\cap V[H]),\infty}$ for some probability measure ν on the boundary of $\mathrm{Tr}(I_o\cap V[H])$ in $\mathrm{Tr}(I_o)$, where the measure ν is determined by the conditional distribution of $\Phi(H)$ given this information. Taking a limit as H exhausts G and using Proposition 87 yields the claim. □

7.3.2 The structure of the augmented connectivity relation

In this section we prove the following proposition about the structure of the augmented connectivity relation in a translation-invariant arboreal gas Gibbs measure on $\mathbb{Z}^d$ and then deduce Theorem 84 from this proposition together with Proposition 96.

Proposition 97. *Let $d\geq 1$ and $\beta\in[0,\infty)$ and let (A,Φ) be distributed as a translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. The following hold:*

1. *The augmented connectivity relation $\xleftrightarrow{(A,\Phi)}$ has at most one infinite equivalence class a.s.*
2. *If the augmented connectivity relation $\xleftrightarrow{(A,\Phi)}$ has an infinite equivalence class, then the subgraph of $\mathbb{Z}^d$ induced by this equivalence class is connected a.s.*

It suffices to prove this in the case that the law of (A,Φ) is extremal in $\mathscr{M}_\beta^T$, taking a decomposition in terms of such extremal measures otherwise.

The proof of Proposition 97 will make use of the following important fact, which follows from the work of Aldous and Lyons [6] as explained in detail in [28, Section 3] and which is closely related to the classical work of Burton and Keane [94].

Proposition 98. *Let $d\geq 1$ and let S be a translation-invariant random subgraph of $\mathbb{Z}^d$. Then every connected component of S has at most two ends almost surely.*

Fix $\beta\in(0,\infty)$, and $d\geq 2$ and let $\mathbb{Q}$ denote an extremal β-arboreal gas augmented Gibbs measure on $\mathbb{Z}^d$, and let $(A,\Phi)\sim\mathbb{Q}$. The Gibbs property tells us that for any $H\in\mathscr{S}^f(G)$, we have that

$$(A,\Phi) \sim (A,\Phi) \cup F \setminus E[H],$$

F has conditional law $\mathbb{P}_{H,\beta}^{\Phi(H)}$ given (A,Φ). Since $\beta \in (0,\infty)$, this implies in particular that, conditional on $A \setminus E[H]$ and $\Phi(H)$, there is a.s. a positive probability that $A \cap E[H] = F'$ for any forest $F' \in \mathscr{F}(H,\Phi(H))$. This leads in particular to the following lemma.

Lemma 99. *Fix $d \geq 2$, $\beta \in (0,\infty)$, let $\mathbb{Q}_\beta$ be an augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$, and let (A,Φ) be distributed as $\mathbb{Q}$.*

1. *If H is a finite subgraph of $\mathbb{Z}^d$ then*

$$\mathbb{Q}_\beta(H \cap A = \emptyset \mid \mathscr{G}_H) > 0 \qquad \text{a.s.} \tag{7.5}$$

2. *If H is a finite* connected *subgraph of $\mathbb{Z}^d$ then*

$$\mathbb{Q}_\beta(\text{all vertices of } H \text{ belong to the same augmented connectivity class} \mid \mathscr{G}_H) > 0 \quad \text{a.s.} \tag{7.6}$$

We refer to the property (7.5) of $\mathbb{Q}_\beta$ as **deletion tolerance** and the property (7.6) as **merge tolerance**.

Proof of Lemma 99. The deletion tolerance property (7.5) is an immediate consequence of the augmented Gibbs property since $\beta < \infty$. We now turn to the merge tolerance property (7.6). Since H is connected, $H/\Phi(H)$ is connected and therefore admits at least one spanning tree, which is given positive mass by the conditional measure $\mathbb{P}_{H,\beta}^{\Phi(H)}$ since $\beta > 0$. On the event that the restriction of the arboreal gas to H is equal to such a spanning tree, all vertices of H belong to the same augmented connectivity class. □

The proofs in the remainder of this section and in the next will generally proceed by assuming that (A,Φ) satisfies a certain property with positive probability and then attempting to derive a contradiction. We will use the above observation to make local edits to (A,Φ), stitching together or separating infinite subgraphs as appropriate. Either ergodicity of $\pi_*\mathbb{Q}$, Proposition 98, or a combination thereof will then be used to generate the desired contradictions.

Remark 28. Several of the proofs in this section are of a similar flavour to those of [188, 194, 312], which studied uniform spanning forests using a property known as *update tolerance* or *weak insertion tolerance*. There are however several important differences: 1) We need to understand the structure of the augmented connectivity relation, which was not a feature

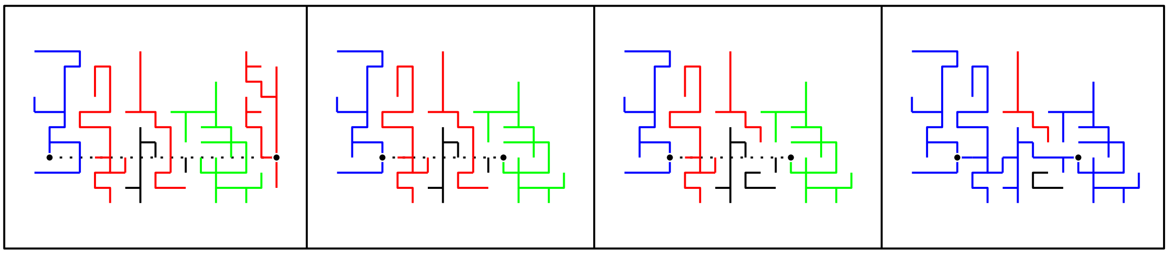


Fig. 7.1 Schematic illustration of the proof of Lemma 100. Infinite augmented connectivity classes are represented by colours, finite classes are black. Far left: a path γ (dotted line) intersecting three distinct infinite augmented connectivity classes. Centre left: By shortening γ if necessary, we may assume that γ intersects exactly three distinct infinite augmented connectivity classes, two of which it intersects only at its endpoints. Centre right: By deleting finitely many edges from the configuration if necessary, we can make it so that each infinite augmented connectivity class intersecting γ contains exactly one A-component intersecting γ. Far right: Using Lemma 99, we may glue together the components intersecting γ to create a component with three or more ends, contradicting Proposition 98.

of those works. 2) Since $\beta < \infty$, we can use deletion tolerance to simplify several steps. 3) Our augmented Gibbs framework allows us to put many of the *ad hoc* seeming parts of those papers on a more robust conceptual footing.

We now begin the proof of Proposition 97 in earnest. We begin by proving that $\xleftrightarrow{(A,\Phi)}$ has at most two infinite equivalence classes almost surely.

Lemma 100. *Fix $d \geq 2$, $\beta \in [0,\infty)$, let $\mathbb{Q}_\beta$ be an extremal translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$, and let (A,Φ) be distributed as $\mathbb{Q}$. Then the augmented connectivity relation $\xleftrightarrow{(A,\Phi)}$ has at most two infinite equivalence classes almost surely.*

Proof of Lemma 100. An illustration of the proof is given in Figure 7.1. The claim is trivial for $\beta = 0$, so we restrict to the case $\beta > 0$. Suppose for contradiction that the event

$$E_1 = \left\{ \xleftrightarrow{(A,\Phi)} \text{ has three or more infinite equivalence classes} \right\}$$

has positive probability. For each $x \in \mathbb{Z}^d$, write $[x]$ for the equivalence class of x under the augmented connectivity relation. Because $\mathbb{Q}_\beta(E_1) > 0$, there must exist three vertices x, y, and z such that $[x]$, $[y]$, and $[z]$ are all distinct with positive probability. Fix three such vertices $x, y, z \in \mathbb{Z}^d$ and let E_2 be the event that this occurs. Since $\mathbb{Z}^d$ is 2-connected, there exists a simple path γ in $\mathbb{Z}^d$ passing through x, y, and z. In particular, there must exist a finite simple path γ that intersects at least three distinct infinite equivalence classes of the augmented connectivity relation with positive probability. Reducing the length of γ if necessary, we may assume that, with positive probability, γ intersects at least three infinite equivalence classes of the augmented connectivity relation, two of which it intersects only at its endpoints. Denote this event by E_2. Using deletion tolerance, it follows that, with positive probability, γ intersects exactly three infinite clusters of A, all of which belong to distinct augmented equivalence classes, and with two of these clusters intersecting γ only at its endpoints. Indeed, denoting this event by E_3, we note that if E_2 occurs but E_3 does not, so that the infinite augmented connectivity class $\mathscr{C}$ intersecting the interior of γ contains multiple infinite A-components intersecting γ, then we can modify the configuration to make E_3 occur by choosing one of the infinite A components that belongs to $\mathscr{C}$ and intersects γ, and deleting from A all edges that are incident to γ and belong to an A-component that belongs to $\mathscr{C}$ but is not equal to the one component we chose to keep. Using merge tolerance allows us to glue together these three infinite A-components into a single infinite cluster by modifying A on γ in a way that preserves absolute continuity, and doing so creates a three ended component. Thus, there is a positive probability that A contains a tree with at least three ends. Since A is translation-invariant this contradicts Proposition 98, and so $\xleftrightarrow{(A,\Phi)}$ has at most two infinite equivalence classes almost surely. $\square$

The next step of the proof of Proposition 97 is to prove Theorem 85, which states that every infinite component of any translation-invariant β-arboreal gas Gibbs measure is one-ended almost surely for every $d \geq 1$ and $\beta \in (0,\infty)$.

Proof of Theorem 85. The claim is trivial if $\beta = 0$ or $d = 1$ so we may assume that $\beta > 0$ and $d \geq 2$. It suffices to prove the claim for measures of the form $\mathbb{P} = \pi_* \mathbb{Q}$ where $\mathbb{Q} = \mathbb{Q}_\beta$ is an extremal translation-invariant β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. Let $(A,\Phi) \sim \mathbb{Q}$. By Proposition 98, all trees in A have at most two ends almost surely, so we need only rule out the existence of two-ended trees. Note that if $e = \{x,y\}$ is an edge of $\mathbb{Z}^d$ we have by the augmented Gibbs property that

$$\mathbb{Q}(e \in A \mid \mathscr{G}_e) = \frac{\beta}{1+\beta} \mathbb{1}\left(\Phi(e)(x,y) = 0\right),$$

where we abuse notation to identify e with the subgraph of $\mathbb{Z}^d$ having $\{x,y\}$ as its only vertices and e as its only edge. Thus, we must have that $\Phi(e)(x,y) = 0$ almost surely for every edge $e = \{x,y\} \in A$. It follows that, almost surely, if A contains a two-ended tree T and e is an edge of T such that $T \setminus e$ has two infinite connected components, then $(A,\Phi) \setminus e$ has one more infinite augmented connectivity class than (A,Φ) (where we allow both augmented subgraphs to have infinitely many infinite augmented connectivity classes in this statement). Thus, it follows by deletion tolerance that if A has at least n two-ended components with positive probability then (A,Φ) has at least $n+1$ infinite augmented equivalence classes with positive probability. Together with Lemma 100, this implies that A has at most one two-ended component almost surely. On the other hand, if A has exactly one two-ended component with positive probability then we have by deletion tolerance that A has no two-ended components with positive probability. Since $\mathbb{Q}$ is extremal the law of A is ergodic by Corollary 95, and since the event that A does not have any two-ended components is translation-invariant it must have probability 1. $\square$

We next deduce that there is at most one infinite augmented connectivity class almost surely.

Lemma 101. *Fix $d \geq 2$, $\beta \in (0,\infty)$, let $\mathbb{Q}_\beta$ be an extremal translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$, and let (A,Φ) be distributed as $\mathbb{Q}_\beta$. Then the augmented connectivity relation $\xleftrightarrow{(A,\Phi)}$ has at most one infinite equivalence class almost surely.*

Proof of Lemma 101. Suppose for contradiction that (A,Φ) has two infinite augmented connectivity classes with positive probability. Letting H be a finite subgraph of $\mathbb{Z}^d$ that intersects both infinite equivalence classes with positive probability, we can use the merge tolerance of (A,Φ) to deduce that, with positive probability, (A,Φ) has a single infinite augmented equivalence class but $(A,\Phi)\setminus H$ does not. On this event there must exist an infinite component of A with more than one end, contradicting Theorem 85. $\square$

To complete the proof of Proposition 97, we show that the induced subgraph $\mathrm{Tr}(I_\infty)$ is connected a.s.

Lemma 102. *Fix $d \geq 2$, $\beta \in (0,\infty)$, let $\mathbb{Q}_\beta$ be an extremal augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$, let $(A,\Phi) \sim \mathbb{Q}_\beta$ and let I_∞ be the set of vertices of $\mathbb{Z}^d$ belonging to infinite clusters of A. If I_∞ is non-empty then the induced subgraph* $\mathrm{Tr}(I_\infty)$ *is connected almost surely.*

Proof of Lemma 102. The proof is similar to that of Lemma 101, but instead of attempting to connect infinite trees, we need (and, given Lemma 101, can) only connect their traces. Suppose for contradiction that the event

$$E_1 = \{\mathrm{Tr}(I_\infty) \text{ has three or more connected components}\},$$

has positive probability. We will connect up the traces of three infinite trees from different components of Tr to give a component with at least three ends. Because $\mathbb{Q}(E_1) > 0$, there exists a finite subgraph H of $\mathbb{Z}^d$ that intersects at least three distinct infinite clusters of $\mathrm{Tr}(I_\infty)$ with positive probability. Using merge tolerance to force all elements of H to belong to the same augmented connectivity cluster, it follows that, with positive probability, $\mathrm{Tr}(I_\infty(A))$ has a single component intersecting H but $\mathrm{Tr}(I_\infty(A \setminus H))$ has at least three infinite components intersecting H. On this event we must have that $\mathrm{Tr}(I_\infty(A))$ contains a component with at least three ends. However Tr is connected and translation-invariant and so this contradicts Proposition 98, and so almost surely Tr has at most two infinite connected components almost surely. □

We are now ready to conclude the proofs of Proposition 97 and Theorem 84.

Proof of Proposition 97. It suffices to consider the case that $\beta > 0$ and $d \geq 2$, the remaining cases being trivial. We may also assume that the law of (A, Φ) is extremal, taking an extremal decomposition otherwise. Once these reductions are made, the claims of Proposition 97 are exactly those of Lemmas 100 and 102. □

Proof of Theorem 84. Let $\mathbb{P}$ be a translation-invariant β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. Lemma 94 tell us that we can find a measure $\mathbb{Q}$ which is a translation-invariant augmentation thereof. Let $(A, \Phi) \sim \mathbb{Q}$. Propositions 96 and 97 together imply that $\mathrm{Tr}(I_\infty)$ is a.s. connected and that the conditional distribution of the restriction of A to $\mathrm{Tr}(I_\infty)$ given I_∞ and the restriction of A to $\mathrm{Tr}(I_\infty^c)$ is almost surely equal to some (possibly random) Gibbs measure for the uniform spanning tree on $\mathrm{Tr}(I_\infty)$. On the other hand, since $\mathrm{Tr}(I_\infty)$ is a translation-invariant random subgraph of $\mathbb{Z}^d$, it is a hyperfinite unimodular random rooted graph. As such, the results of Aldous and Lyons [6, Proposition 8.14] imply that its free and wired uniform spanning forests coincide, and hence that it has a unique Gibbs measure for the uniform spanning tree by Lemma 88. This completes the proof. □

We end this section by observing the following corollary of Proposition 97 and Theorem 85.

Corollary 103. *Let $d \geq 1$ and $\beta \in [0, \infty)$ and let (A, Φ) be distributed as a translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$. Then the augmented subgraph (A, Φ) is almost surely equal to the wired augmentation of A as defined in* (7.4).

Corollary 103 implies in particular that the completions of the sigma-algebras $\mathscr{I}$ and $\mathscr{I}_S$ coincide, which implies the following corollary in conjunction with Corollary 95.

Corollary 104. *Every extremal translation-invariant augmented β-arboreal gas Gibbs measure on $\mathbb{Z}^d$ is ergodic for every $d \geq 1$ and $\beta \in [0, \infty]$.*

7.4 Random walk intersections in unimodular random graphs

In this section we prove Theorem 86, which states that uniform spanning trees of unimodular random rooted subgraphs of $\mathbb{Z}^d$ are connected almost surely when $d \leq 4$; by the results of Benjammini, Lyons, Peres and Schramm [68, 261] this is equivalent to the statement that two independent random walks on such a graph intersect infinitely often almost surely. This property is known as the **infinite intersection property**. The proof is a combination of two results. First, in Section 7.4.1, we establish, for general unimodular random rooted graphs whose degree has finite second moment, that two random walks intersect infinitely often almost surely if and only if their *expected* number of intersections *conditional on the rooted graph and one of the two walks* is infinite almost surely. Then, in Section 7.4.2, we show that this condition is satisfied for random walks on unimodular subgraphs of $\mathbb{Z}^d$ for $d \leq 4$ using the theory of *Markov-type inequalities*.

Before getting started with the proof, we quickly review some relevant definitions and state a generalization of Theorem 86.

Unimodular random rooted graphs. A **rooted graph** is a pair (G, ρ) where G is a connected, locally finite graph and ρ is a distinguished vertex of G known as the root vertex; an isomorphism of graphs is an isomorphism of rooted graphs if it preserves the root. We define $\mathscr{G}_\bullet$ to be the space of isomorphism classes of rooted graphs, which is equipped with the Borel sigma algebra induced by the *local topology* [6, 110], in which two elements of $\mathscr{G}_\bullet$ are considered to be close if there exist large graph-distance balls around their root vertices which admit a graph isomorphism that preserves the root. The details of this construction are not important to us and can be found in e.g. [110, Section 1.2]. Similarly, we also have the space $\mathscr{G}_{\bullet\bullet}$ of (isomorphism classes of) doubly-rooted graphs (G, ρ_1, ρ_2), with an ordered pair of distinguished root vertices $\rho_1, \rho_2 \in V[G]$. We say that a random variable (G, ρ) taking values in $\mathscr{G}_\bullet$ is **unimodular** if it satisfies the **mass-transport principle**, meaning that

$$\mathbb{E}\left[\sum_{v \in V[G]} F(G, \rho, v)\right] = \mathbb{E}\left[\sum_{v \in V[G]} F(G, v, \rho)\right]$$

for every Borel measurable function $F : \mathscr{G}_{\bullet\bullet} \to [0, \infty)$.

Next we define the space of *rooted subgraphs of* $\mathbb{Z}^d$; this definition is not standard. For any connected graph G and $d \geq 1$, we say the function $\phi : V[G] \times V[G] \to \mathbb{Z}^d$ is an **embedding** of G into $\mathbb{Z}^d$ if $\phi(u,w) = \phi(u,v) + \phi(v,w)$ for every $u, v, w \in \mathbb{Z}^d$ (i.e. if ϕ is an additive cocyle), $\phi(u,w) = 0$ if and only if $u = w$, and $\|\phi(u,w)\|_\infty = 1$ if $\{u,w\} \in E[G]$.

A **rooted subgraph of** $\mathbb{Z}^d$ is then a tuple (G, ϕ, ρ), where G, ρ are as before, and ϕ is an embedding of G into $\mathbb{Z}^d$. We denote the space of isomorphism classes of rooted subgraphs of $\mathbb{Z}^d$ by $\mathscr{S}_\bullet(\mathbb{Z}^d)$, which we endow with the Borel sigma algebra corresponding to the local topology, where for two elements to be close, the embeddings now also have to coincide in a large ball. Defining the space of doubly-rooted subgraphs $\mathscr{S}_{\bullet\bullet}(\mathbb{Z}^d)$ similarly, we say that a random tuple (G, ϕ, ρ) is unimodular if

$$\mathbb{E}\left[\sum_{v \in V[G]} F(G, \phi, \rho, v)\right] = \mathbb{E}\left[\sum_{v \in V[G]} F(G, \phi, v, \rho)\right]$$

for every Borel measurable function $F : \mathscr{S}_{\bullet\bullet}(\mathbb{Z}^d) \to [0, \infty)$.

Lemma 105. *If ω is a translation-invariant random subgraph of $\mathbb{Z}^d$, K_0 denotes the cluster of the origin in ω, and we define a cocyle $\phi : V[K_0] \times V[K_0] \to \mathbb{Z}^d$ by $\phi(u, v) = u - v$, then $(K_0, \phi, 0)$ is a unimodular random rooted subgraph of $\mathbb{Z}^d$.*

Proof. The translation-invariance of the model implies that if $F : \mathscr{S}_{\bullet\bullet}(\mathbb{Z}^d) \to [0, \infty)$ is measurable then $F'(u, v) = \mathbb{E}[F(K_u, \phi, u, v)]$ satisfies $F'(u + x, v + x) = F'(u, v)$ for every $u, v, x \in \mathbb{Z}^d$, and the claim follows from the usual mass transport principle for $\mathbb{Z}^d$. $\square$

Since unimodularity is preserved by conditioning on re-rooting invariant events, it follows that $(K_0, \phi, 0)$ remains unimodular when we condition on it having size n for any $n \in \mathbb{N} \cup \{\infty\}$ for which the relevant probability is positive. As such, Theorem 86 follows from the following more general theorem. (Examples of unimodular random rooted subgraphs of $\mathbb{Z}^d$ that do *not* arise as a cluster in a translation-invariant model include the incipient infinite percolation cluster and the trace of a doubly-infinite random walk.)

Theorem 106. *Let $d \leq 4$ and let (G, ϕ, ρ) be a unimodular random rooted subgraph of $\mathbb{Z}^d$. Then G has the infinite intersection property almost surely.*

Equivalently, if (G, ϕ, ρ) is a unimodular random rooted subgraph of $\mathbb{Z}^d$ then the uniform spanning forest of G is connected almost surely on the event that G is infinite (the uniform spanning forest of G being a.s. well-defined independently of boundary conditions by the results of [6] as discussed in the proof of Theorem 84).

7.4.1 A criterion for the infinite intersection property

The goal of this subsection is to prove the following general proposition concerning intersections of random walks on general unimodular random rooted graphs.

Proposition 107. *Let (G,o) be a unimodular random rooted graph which is almost surely connected and suppose that the second moment of the degree of the root is finite, i.e. $\mathbb{E}[\deg(o)^2] < \infty$. Let X and Y are two random walks on G, both started at o, that are conditionally independent given (G,o). If*

$$\mathbb{E}\left[\#\{i,j \geq 0 : X_i = Y_j\} \mid (G,o), Y\right] = \infty \qquad \textit{almost surely}$$

then G has the infinite intersection property almost surely.

The proof of this proposition is of a similar flavour to those of [176, 195], which involve *collisions* (where the two walks are at the same location *at the same time*) rather than *intersections* (where the two walks are at the same location but not necessarily at the same time).

We begin by establishing a lemma concerning random walks on *deterministic* graphs. It will be convenient to work with two-sided rather than one-sided random walks. Given a connected, locally finite graph G and two vertices $u, v \in V[G]$, we write $\mathbf{P}^G_{u,v}$ for the joint law of a pair of independent doubly-infinite random walks $(X_n)_{n\in\mathbb{Z}}$ and $(Y_n)_{n\in\mathbb{Z}}$ started at u and v respectively: Concretely, we let X^+, X^-, Y^-, and Y^+ be independent random walks on G, where X^+ and X^- are started at u and Y^+ and Y^- are started at v, and define the two-sided random walks $(X_n)_{n\in\mathbb{Z}}$ and $(Y_n)_{n\in Z}$ by

$$X_n = \begin{cases} X_n^+ & n \geq 0 \\ X_{-n}^- & n \leq 0 \end{cases} \qquad \text{and } Y_n = \begin{cases} Y_n^+ & n \geq 0 \\ Y_{-n}^- & n \leq 0. \end{cases}$$

Given a subset A of $\mathbb{Z}\times\mathbb{Z}$, we write lex-max$A$ for the lexicographical maximum of A when this maximum is well-defined. The following lemma may be thought of as a time-reversal identity for the probabilities of these events.

Lemma 108. *Let $G = (V,E)$ be a transient, connected, locally finite graph, and let o be a vertex of G. Then*

$$\begin{aligned} &\mathbf{P}^G_{o,o}\big(\text{lex-max}\{(i,j) : X_i = Y_j\} = (n,m)\big) \\ &= \sum_{v\in V} \frac{\deg(v)^2}{\deg(o)^2} \mathbf{P}^G_{v,v}(X_{-n} = Y_{-m} = o, \{X_i\}_{i\geq 0} \cap \{Y_j\}_{j>0} = \emptyset, \{X_i\}_{i>0} \cap \{Y_j\}_{j\leq 0} = \emptyset) \end{aligned} \quad (7.7)$$

for every $n,m \geq 0$.

(Here, the event "lex-max$\{(i,j) : X_i = Y_j\} = (n,m)$" implicitly includes the condition that the lexicographical maximum is well-defined.)

Proof of Lemma 108. Fix $n,m \geq 0$ and write

$$B_{n,m} := \Big\{ \text{lex-max}\{(i,j) : X_i = Y_j\} = (n,m) \Big\} = \\ \Big\{ X_n = Y_m, \{X_i\}_{i\geq n} \cap \{Y_j\}_{j>m} = \emptyset, \{X_i\}_{i>n} \cap \{Y_j\}_{j\leq m} = \emptyset \Big\}.$$

Decomposing according to the value of $X_n = Y_m$ yields that

$$\mathbf{P}^G_{o,o}(B_{n,m}) = \sum_{v\in V} \mathbf{P}^G_{o,o}\big(X_n = Y_m = v, \{X_i\}_{i\geq n} \cap \{Y_j\}_{j>m} = \emptyset, \{X_i\}_{i>n} \cap \{Y_j\}_{j\leq m} = \emptyset\big). \quad (7.8)$$

Let $\mathbf{P}^G_o$ denote the marginal law of $(Y_n)_{n\in\mathbb{Z}}$ and abbreviate $\deg(v) = \mathrm{d}(v)$ for each vertex v of G. For each $v \in V[G]$ and each doubly-infinite simple $(x_n)_{n\in\mathbb{Z}}$ path in G with $x_0 = o$ and $x_n = v$ we can compute that

$$\begin{aligned}
&\mathbf{P}^G_o\big(Y_m = v,\, \{x_i\}_{i\geq n} \cap \{Y_j\}_{j>m} = \emptyset,\, \{x_i\}_{i>n} \cap \{Y_j\}_{j\leq m} = \emptyset\big)\\
&= \mathbf{P}^G_o\big(\{x_i\}_{i>n} \cap \{Y_j\}_{j<0} = \emptyset\big)\mathbf{P}^G_o\big(\{x_i\}_{i>n} \cap \{Y_j\}_{0\leq j\leq m} = \emptyset,\, Y_m = v\big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdot \mathbf{P}^G_v\big(\{x_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset\big)\\
&= \mathbf{P}^G_o\big(\{x_i\}_{i>n} \cap \{Y_j\}_{j<0} = \emptyset\big)\left(\frac{\mathrm{d}(v)}{\mathrm{d}(o)}\mathbf{P}^G_v\big(\{x_i\}_{i>n} \cap \{Y_j\}_{0\leq j\leq m} = \emptyset,\, Y_m = o\big)\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdot \mathbf{P}^G_v\big(\{x_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset\big)\\
&= \frac{\mathrm{d}(v)}{\mathrm{d}(o)}\mathbf{P}^G_o\big(\{x_i\}_{i>n} \cap \{Y_j\}_{j<0} = \emptyset\big)\mathbf{P}^G_v\big(\{x_i\}_{i>n} \cap \{Y_j\}_{-m\leq j\leq 0} = \emptyset,\, Y_{-m} = o\big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdot \mathbf{P}^G_v\big(\{x_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset\big)\\
&= \frac{\mathrm{d}(v)}{\mathrm{d}(o)}\mathbf{P}^G_v\big(\{x_i\}_{i>n} \cap \{Y_j\}_{j\leq 0} = \emptyset,\, Y_{-m} = o\big)\mathbf{P}^G_v\big(\{x_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset\big)\\
&= \frac{\mathrm{d}(v)}{\mathrm{d}(o)}\mathbf{P}^G_v\big(Y_{-m} = o,\, \{x_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset,\, \{x_i\}_{i>n} \cap \{Y_j\}_{j\leq 0} = \emptyset\big),
\end{aligned}$$

where the first equality follows by independence of $\{Y_j\}_{j<0}$ and $\{Y_j\}_{j>0}$ and the Markov property of $\{Y_j\}_{j>0}$, the second equality follows by time-reversal for $\{Y_j\}_{j>0}$, the third equality follows as $\{Y_j\}_{j<0}$ and $\{Y_j\}_{j>0}$ are identically distributed, the penultimate inequality follows by the Markov property, and the final equality follows by independence of $\{Y_j\}_{j<0}$ and $\{Y_j\}_{j>0}$. Now, since X and Y are independent, letting $x = X$ gives

$$\begin{aligned}
&\mathbf{P}^G_{o,o}\big(X_n = Y_m = v, \{X_i\}_{i\geq n} \cap \{Y_j\}_{j>m} = \emptyset, \{X_i\}_{i>n} \cap \{Y_j\}_{j\leq m} = \emptyset\big)\\
&\qquad = \frac{\mathrm{d}(v)}{\mathrm{d}(o)}\mathbf{P}^G_{o,v}\big(X_n = v, Y_{-m} = o,\, \{X_i\}_{i\geq n} \cap \{Y_j\}_{j>0} = \emptyset,\, \{X_i\}_{i>n} \cap \{Y_j\}_{j\leq 0} = \emptyset\big),
\end{aligned}$$

and applying a similar time-reversal to X gives that

$$\begin{aligned}\mathbf{P}^G_{o,o}\big(X_n = Y_m = v, \{X_i\}_{i\geq n} \cap \{Y_j\}_{j>m} = \emptyset, \{X_i\}_{i>n} \cap \{Y_j\}_{j\leq m} = \emptyset\big)\\ = \frac{\mathrm{d}(v)^2}{\mathrm{d}(o)^2}\mathbf{P}^G_{v,v}(X_{-n} = Y_{-m} = o, \{X_i\}_{i\geq 0} \cap \{Y_j\}_{j>0} = \emptyset, \{X_i\}_{i>0} \cap \{Y_j\}_{j\leq 0} = \emptyset).\end{aligned}$$

The claim follows by substituting this into (7.8). □

Proof of Proposition 107. The claim holds trivially when G is recurrent, so we may assume that G is transient. Let $(X_n)_{n\in Z}$ and $(Y_n)_{n\in\mathbb{Z}}$ be doubly-infinite random walks started at o that are conditionally independent given (G,o). We assume that

$$\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,o),Y\right] = \infty$$

almost surely and prove that in this case $\#\{i,j\geq 0 : X_i = Y_j\} = \infty$ almost surely.

Recall that $B_{n,m}$ denotes the event that lex-max$\{(n,m) : X_n = Y_m\} = (n,m)$. Multiplying both sides of the identity of Lemma 108 by $\deg(o)^2$, taking expectations and applying the mass-transport principle to the right-hand side gives

$$\begin{aligned}&\mathbb{E}\big[\deg(o)^2 \mathbb{1}(B_{n,m})\big]\\ &\qquad\geq \mathbb{E}\left[\sum_{v\in G}\mathbf{P}^G_{v,v}(X_{-n} = Y_{-m} = o, \{X_i\}_{i\geq 0} \cap \{Y_j\}_{j>0} = \emptyset, \{X_i\}_{i>0} \cap \{Y_j\}_{j\leq 0} = \emptyset)\right]\\ &\qquad= \mathbb{E}\left[\sum_{v\in G}\mathbf{P}^G_{o,o}(X_{-n} = Y_{-m} = v, \{X_i\}_{i\geq 0} \cap \{Y_j\}_{j>0} = \emptyset, \{X_i\}_{i>0} \cap \{Y_j\}_{j\leq 0} = \emptyset)\right],\end{aligned}$$

where we bounded $\deg(v) \geq 1$ in the first line. Summing over $n,m \geq 0$ and using that the events $B_{n,m}$ are disjoint, we obtain that

$$\begin{aligned}\mathbb{E}\left[\#\{i,j\leq 0 : X_i = Y_j\}\mathbb{1}\big(\{X_i\}_{i\geq 0} \cap \{Y_j\}_{j>0} = \emptyset, \{X_i\}_{i>0} \cap \{Y_j\}_{j\leq 0} = \emptyset\big)\right]\\ \leq \mathbb{E}[\deg(o)^2] < \infty.\end{aligned}$$

Conditioning on the random rooted graph (G,o) and the two-sided walk Y, conditional independence of $(X_i)_{i\leq 0}$ and $(X_i)_{i\geq 0}$ yields

$$\mathbb{E}\Bigg[\mathbb{E}\left[\#\{i,j\leq 0 : X_i = Y_j\} \mid (G,o),Y\right]\cdot\mathbb{P}\big(\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j>0}=\emptyset,$$
$$\{X_i\}_{i>0}\cap\{Y_j\}_{j\leq 0}=\emptyset \mid (G,o),Y\big)\Bigg] < \infty$$

Since $\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,o),Y\right] = \infty$ almost surely by assumption, the right hand side can only be finite if

$$\mathbb{P}\big(\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j>0}=\emptyset,\{X_i\}_{i>0}\cap\{Y_j\}_{j\leq 0}=\emptyset\big)=0.$$

Since the two events $\{\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j>0}=\emptyset\}$ and $\{X_i\}_{i>0}\cap\{Y_j\}_{j\leq 0}=\emptyset$ are conditionally independent given (G,o) and X, and since $\mathbb{P}\big(\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j\leq 0}=\emptyset|(G,o),X\big)=\mathbb{P}\big(\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j\geq 0}=\emptyset|(G,o),X\big)$, it follows that

$$\mathbb{P}\big(\{X_i\}_{i\geq 0}\cap\{Y_j\}_{j>0}=\emptyset,\{X_i\}_{i>0}\cap\{Y_j\}_{j\geq 0}=\emptyset\big)=0.$$

In other words, two conditionally independent random walks X and Y started at $o\in G$ will almost surely satisfy $X_n = Y_m$ at some time (n,m) with $n,m\geq 0$ and $(n,m)\neq(0,0)$. Since the random rooted graph (G,o) is unimodular, the same statement holds almost surely for *any* starting vertex $v\in G$ [110, Proposition 11]. Now if X and Y are conditionally independent random walks on G with arbitrary starting vertices, then the Markov properties of the random walks implies that for any $(n,m)\in\mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}$, the processes $(X_i)_{i\geq n}$ and $(Y_i)_{i\geq m}$ are jointly distributed as two conditionally independent random walks on G started at X_n and Y_m respectively. In particular, together with our conclusion above, this implies that for any $n,m\geq 0$, the event

$$\{X_n\neq Y_m\}\cup\{\exists i\geq n,\ j\geq m : (i,j)\neq(n,m) \text{ and } X_i = Y_j\}$$

occurs almost surely. If we now suppose that X and Y start at the same vertex, then we can use this fact inductively to construct two non-decreasing sequences of times $(T_i)_{i\geq 0}$ and $(S_i)_{i\geq 0}$ such that $(S_i+T_i)_{i\geq 0}$ is strictly increasing and $X_{T_i}=Y_{S_i}$ almost surely for every $i\geq 0$. Thus the proposition is proved. $\square$

Remark 29. This proposition certainly does not hold if the unimodularity assumption is removed. For instance, take two copies of $\mathbb{Z}^3$ attached by a single edge: The conditional expectation of the number of intersections is almost surely infinite, but the number of intersections has a positive probability of being finite due to the fact that the random walks may eventually remain in distinct copies of $\mathbb{Z}^3$. We are unsure if the analogous statement holds if we only require that $\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,o)\right] = \infty$ a.s. rather than $\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,o),Y\right] = \infty$ a.s.

Remark 30 (Relaxing the second moment condition). The proof of Proposition 107 shows more generally that if (G,o) is a unimodular random rooted graph with $\mathbb{E}[\deg(o)^\alpha] < \infty$ for some $0 \leq \alpha \leq 2$ and $\mathbb{E}[\sum_{i,j=0}^\infty \mathbb{1}(X_i = Y_j)\deg(Y_j)^{-2+\alpha} \mid (G,o),Y] = \infty$ almost surely then G has the infinite intersection property almost surely.

7.4.2 Proof of Theorems 86 and 106

In this section we complete the proof of Theorems 86 and 106, and hence also of Theorem 82, by proving the following proposition, which implies these theorems in conjunction with Proposition 107 and Theorem 84.

Proposition 109. *Let $1 \leq d \leq 4$, let (G,ϕ,ρ) be a unimodular random random rooted subgraph of $\mathbb{Z}^d$, and let X and Y be two independent random walks on G beginning at ρ. Then*

$$\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,\phi,\rho),Y\right] = \infty \text{ almost surely.}$$

The proof of this proposition will apply the theory of *Markov-type inequalities*, which were first introduced by Ball [40] in the context of the Lipschitz extension problem and have since been found to have many important applications to the study of random walk. We now give a quick review of the parts of the theory most relevant to us, referring the reader to [259, Chapter 13.4] for further background.

Markov-type inequalities. A metric space $\mathscr{X} = (\mathscr{X},d)$ is said to have **Markov-type** 2 with constant $C < \infty$ if for every finite set S, every irreducible reversible Markov chain M on S, and every function $f : S \to \mathscr{X}$ the inequality

$$\mathbb{E}\left[d\big(f(Y_0),f(Y_n)\big)^2\right] \leq C^2 n\mathbb{E}\left[d\big(f(Y_0),f(Y_1)\big)^2\right]$$

holds for every $n \geq 0$, where $(Y_i)_{i\geq 0}$ is a trajectory of the Markov chain M with Y_0 distributed as the stationary measure of M. Similarly, a metric space $\mathscr{X} = (\mathscr{X},d)$ is said to have **maximal Markov-type** 2 with constant $C < \infty$ if for every finite set S and every irreducible reversible Markov chain M on S, and every function $f : S \to \mathscr{X}$, we have that

$$\mathbb{E}\left[\max_{0\leq i\leq n} d\big(f(Y_0),f(Y_i)\big)^2\right] \leq C^2 n\mathbb{E}\left[d\big(f(Y_0),f(Y_1)\big)^2\right] \tag{7.9}$$

for each $n \geq 0$, where, as before, $(Y_i)_{i\geq 0}$ is a trajectory of the Markov chain M with Y_0 distributed as the stationary measure of M. Of particular importance to us will be the fact that $\mathbb{R}$ has maximal Markov-type 2 [259, Theorem 13.15], which implies by projecting onto each coordinate that $\mathbb{R}^d$ has maximal Markov-type 2 with the same constant for each $d \geq 1$, which implies the following inequality for unimodular random rooted subgraphs of $\mathbb{Z}^d$.

Proposition 110. *Let $d \geq 1$ and let (G,ϕ,ρ) be a unimodular random rooted subgraph of $\mathbb{Z}^d$. If Y is a random walk on G started at ρ then*

$$\mathbb{E}\left[\deg(\rho)\max_{0\leq i\leq n}\|\phi(Y_{ik},Y_0)\|_\infty^2\right] \leq C^2 n\mathbb{E}\Big[\deg(\rho)\|\phi(Y_k,Y_0)\|\Big].$$

for each $n,k \geq 1$. Since $\|\phi(Y_k,Y_0)\| \leq 1$ and $1 \leq \deg(\rho) \leq 2d$, it follows in particular that

$$\mathbb{E}\left[\max_{0\leq i\leq n}\|\phi(Y_i,Y_0)\|_\infty^2\right] \leq 2dC^2 n = C_0(d)^2 n$$

for each $n \geq 1$, where $C_0(d) = C\sqrt{2d}$.

Proof of Proposition 110. This follows from the standard maximal Markov type inequality (7.9) by using that unimodular random rooted subgraphs of $\mathbb{Z}^d$ are *hyperfinite*. This means in particular that they can always be written as Benjamini-Schramm limits of *finite* random rooted subgraphs of $\mathbb{Z}^d$, which are finite reversible Markov chains whose stationary measure is proportional to their degree. The details are very similar to the proof of [174, Corollary 2.5] and are omitted. □

Proof of Proposition 109. Fix $\varepsilon \in (0,1)$ and let $C_0 = C_0(d)$ be the constant from Proposition 110. Define constants $c_1 = \sqrt{2C_0/\varepsilon}$ and $c_2 = 2/\varepsilon$, and define sequences of times $t_n = 4^n$, radii $r_n = \lceil c_1 \cdot 2^n \rceil$, and Euclidean boxes $\Lambda_n = [-r_n, r_n]^d \subset \mathbb{R}^d$. Proposition 110 and Markov's inequality give us that for each $n \geq 1$,

$$\mathbb{P}(\{\phi(\rho,Y_i)\}_{i\leq t_n} \subset \Lambda_n) \geq 1-\varepsilon. \tag{7.10}$$

For each subset $A \subseteq \mathbb{Z}_{\geq 0}$ and $v \in \mathbb{Z}^d$, define the random variable $L_A(v) = \sum_{n\in A} \mathbb{1}(Y_n = v)$ giving the number of times i in A such that $Y_i = v$, and define the partial Green's function $G_A(v) = \mathbb{E}^G[L_A(v)]$. We lower bound

$$\mathbb{E}\left[\#\{i,j\geq 0 : X_i = Y_j\} \mid (G,o),Y\right] = \sum_{v\in G} G_{\mathbb{Z}_{\geq 0}}(v)L_{\mathbb{Z}_{\geq 0}}(v) \geq \sum_{n\geq 1}\sum_{v\in\Lambda_n} G_{[t_{n-1},t_n)}(v)L_{[t_{n-1},t_n)}(v), \tag{7.11}$$

where we write $v \in \Lambda_n$ as shorthand for $\phi(\rho, v) \in \Lambda_n$. We aim to show that each sum over Λ_n has good probability to contribute a constant to the total. To this end, for each $n \geq 1$ let $b_n = 2^{-n(d-2)}/(4c_2(4c_1)^d)$ and let

$$U_n = \sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) \mathbb{1}(G_{[t_{n-1},t_n)}(v) < b_n).$$

We can bound

$$\begin{aligned}\sum_{v \in \Lambda_n} G_{[t_{n-1},t_n)}(v) L_{[t_{n-1},t_n)}(v) &\geq \sum_{v \in \Lambda_n} G_{[t_{n-1},t_n)}(v) L_{[t_{n-1},t_n)}(v) \mathbb{1}(G_{[t_{n-1},t_n)}(v) \geq b_n) \\ &\geq b_n \sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) \mathbb{1}(G_{[t_{n-1},t_n)}(v) \geq b_n) = b_n \left[\sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) - U_n \right], \quad (7.12)\end{aligned}$$

and also have trivially that

$$\begin{aligned}\mathbf{E}^G[U_n] = \mathbf{E}^G &\left[\sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) \mathbb{1}(G_{[t_{n-1},t_n)}(v) < b_n) \right] \\ &= \sum_{v \in \Lambda_n} G_{[t_{n-1},t_n)}(v) \mathbb{1}(G_{[t_{n-1},t_n)}(v) < b_n) \leq b_n |\Lambda_n|,\end{aligned}$$

where we write $|\Lambda_n|$ for the number of vertices $v \in G$ such that $\phi(\rho, v) \in \Lambda_n$. Since we also have that $\sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) \geq t_n - t_{n-1}$ on the event that $\{\phi(\rho, Y_i)\}_{i \leq t_n} \subseteq \Lambda_n$, we have by (7.10) and Markov's inequality that

$$\mathbb{P}\left(\sum_{v \in \Lambda_n} L_{[t_{n-1},t_n)}(v) \geq t_n - t_{n-1} \right) \geq 1 - \varepsilon \qquad \text{and} \qquad \mathbb{P}\big(U_n \leq c_2 b_n |\Lambda_n|\big) \geq 1 - \varepsilon \quad (7.13)$$

for every $n \geq 1$. Since $c_2 b_n |\Lambda_n| \leq (t_n - t_{n-1})/2$ by choice of b_n, it follows from this and (7.12) that

$$\mathbb{P}\left(\sum_{v \in \Lambda_n} G_{[t_{n-1},t_n)}(v) L_{[t_{n-1},t_n)}(v) \geq \frac{b_n(t_n - t_{n-1})}{2} \right) \geq 1 - 2\varepsilon$$

for every $n \geq 1$. Now, we also have that $b_n(t_n - t_{n-1})$ is of order $2^{(4-d)n}$ and hence, since $d \leq 4$, that $\frac{b_n(t_n - t_{n-1})}{2}$ is bounded below by a positive constant $c_3 = c_3(\varepsilon)$. Fatou's lemma then implies that

$$\mathbb{P}\left(\sum_{v \in G} G_{\mathbb{Z}_{\geq 0}}(v) L_{\mathbb{Z}_{\geq 0}}(v) = \infty \right) \geq \mathbb{P}\left(\limsup_{n \to \infty} \sum_{v \in \Lambda_n} G_{[t_{n-1},t_n)}(v) L_{[t_{n-1},t_n)}(v) \geq c_3(\varepsilon) \right) \geq 1 - 2\varepsilon$$

for every $\varepsilon > 0$, and the claim follows since $\varepsilon > 0$ was arbitrary. $\square$

Proof of Theorem 106. This is an immediate consequence of Propositions 107 and 109. $\square$

Proof of Theorem 86. This follows immediately from Theorem 106 and the results of Benjammini, Lyons, Peres and Schramm [68, 261]. $\square$

Proof of Theorem 82. This is an immediate consequence of Theorems 84 and 86. $\square$

Remark 31. For $d < 4$ we can substitute the use of the Markov-type inequalities in the proof of Proposition 109 with the Varopoulos-Carne inequality, which implies that the maximal displacement bound $\max_{i \leq n} d(X_0, X_i)$ has order at most $\sqrt{n \log n}$ with high probability on any graph of at most polynomial volume growth. As such, Proposition 109 and thus Theorem 86 generalises easily to unimodular random rooted graphs whose balls have volume $O(n^d)$ for some $d < 4$ (and with $\mathbb{E}[\deg(\rho)^2] < \infty$), without the need to have a unimodular embedding into $\mathbb{Z}^d$. The four-dimensional case is more delicate since this dimension is critical for $\mathbb{Z}^d$ to have the infinite intersection property, with each dyadic scale only contributing $O(1)$ intersections in expectation. We believe that it should be possible to extend Theorem 86 to unimodular random rooted graphs whose balls have volume $O(n^4)$ using the methods of Ganguly, Lee, and Peres [150], who proved that any unimodular random rooted graph of polynomial volume growth satisfies a diffusive estimate at infinitely many scales. To do this, one would need to improve their displacement estimate to a *maximal* displacement estimate of the same order; we do not investigate this here.

References

[1] Addario-Berry, L. (2013). The local weak limit of the minimum spanning tree of the complete graph. *arXiv preprint arXiv:1301.1667*.

[2] Aizenman, M., Burchard, A., Newman, C. M., and Wilson, D. B. (1999). Scaling limits for minimal and random spanning trees in two dimensions. *Random Structures & Algorithms*, 15(3-4):319–367.

[3] Akcoglu, M. A. and del Junco, A. (1975). Convergence of averages of point transformations. *Proc. Amer. Math. Soc.*, 49:265–266.

[4] Aldous, D. (1991). The continuum random tree. I. *Ann. Probab.*, 19(1):1–28.

[Aldous and Fill] Aldous, D. and Fill, J. Reversible Markov chains and random walks on graphs.

[6] Aldous, D. and Lyons, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508.

[7] Aldous, D. J. (1990). The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3(4):450–465.

[8] Aldous, David J., L. R. (2007). Processes on unimodular random networks. *Electronic Journal of Probability [electronic only]*, 12:1454–1508.

[9] Alexander, S. and Orbach, R. (1982). Density of states on fractals:«fractons». *Journal de Physique Lettres*, 43(17):625–631.

[10] Andres, S. (2014). Invariance principle for the random conductance model with dynamic bounded conductances. *Ann. Inst. H. Poincaré Probab. Statist.*, 50:352–374.

[11] Andres, S., Barlow, M. T., Deuschel, J.-D., and Hambly, B. M. (2013). Invariance principle for the random conductance model. *Probab. Theory Related Fields*, 156(3-4):535–580.

[12] Andres, S., Chiarini, A., Deuschel, J.-D., and Slowik, M. (2018). Quenched invariance principle for random walks with time-dependent ergodic degenerate weights. *Ann. Probab.*, 46(1):302–336.

[13] Andres, S., Chiarini, A., and Slowik, M. (2020a). Quenched local limit theorem for random walks among time-dependent ergodic degenerate weights.

[14] Andres, S., Chiarini, A., and Slowik, M. (2021). Quenched local limit theorem for random walks among time-dependent ergodic degenerate weights. *Probab. Theory Related Fields*, 179(3-4):1145–1181.

[15] Andres, S., Deuschel, J.-D., and Slowik, M. (2015a). Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.*, 43(4):1866–1891.

[16] Andres, S., Deuschel, J.-D., and Slowik, M. (2015b). Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.*, 43(4):1866–1891.

[17] Andres, S., Deuschel, J.-D., and Slowik, M. (2016a). Harnack inequalities on weighted graphs and some applications to the random conductance model. *Probab. Theory Related Fields*, 164(3-4):931–977.

[18] Andres, S., Deuschel, J.-D., and Slowik, M. (2016b). Heat kernel estimates for random walks with degenerate weights. *Electron. J. Probab.*, 21:Paper No. 33, 21.

[19] Andres, S., Deuschel, J.-D., and Slowik, M. (2019). Heat kernel estimates and intrinsic metric for random walks with general speed measure under degenerate conductances. *Electron. Commun. Probab.*, 24:Paper No. 5, 17.

[20] Andres, S., Deuschel, J.-D., and Slowik, M. (2020b). Green kernel asymptotics for two-dimensional random walks under random conductances. *Electron. Commun. Probab.*, 25:Paper No. 58, 14.

[21] Andres, S., Deuschel, J.-D., and Slowik, M. (2020c). Green kernel asymptotics for two-dimensional random walks under random conductances. *Electronic Communications in Probability*, 25.

[22] Andres, S., Gantert, N., Schmid, D., and Sousi, P. (2023). Biased random walk on dynamical percolation.

[23] Andres, S. and Neukamm, S. (2019). Berry-Esseen theorem and quantitative homogenization for the random conductance model with degenerate conductances. *Stoch. Partial Differ. Equ. Anal. Comput.*, 7(2):240–296.

[24] Andres, S. and Taylor, P. A. (2021a). Local limit theorems for the random conductance model and applications to the Ginzburg-Landau $\nabla\phi$ interface model. *J. Stat. Phys.*, 182(2):35.

[25] Andres, S. and Taylor, P. A. (2021b). Local limit theorems for the random conductance model and applications to the Ginzburg-Landau $\nabla\phi$ interface model. *arXiv preprint arXiv:1907.05311*, 182(2):Paper No. 35, 35.

[26] Angel, O., Crawford, N., and Kozma, G. (2014). Localization for linearly edge reinforced random walks. *Duke Math. J.*, 163(5):889–921.

[27] Angel, O., Croydon, D. A., Hernandez-Torres, S., and Shiraishi, D. (2021). Scaling limits of the three-dimensional uniform spanning tree and associated random walk. *The Annals of Probability*, 49(6):3032 – 3105.

[28] Angel, O., Hutchcroft, T., Nachmias, A., and Ray, G. (2018). Hyperbolic and parabolic unimodular random maps. *Geom. Funct. Anal.*, 28(4):879–942.

[29] Armstrong, S., Kuusi, T., and Mourrat, J.-C. (2019). *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham.

[30] Asselah, A., Schapira, B., and Sousi, P. (2019). Capacity of the range of random walk on $\mathbb{Z}^4$. *Ann. Probab.*, 47(3):1447–1497.

[31] Avena, L. (2012). Symmetric exclusion as a model of non-elliptic dynamical random conductances. *Electron. Commun. Probab.*, 17:8 pp.

[32] Avena, L., Blondel, O., and Faggionato, A. (2016). A class of random walks in reversible dynamic environments: antisymmetry and applications to the east model. *Journal of Statistical Physics*, 165(1):1–23.

[33] Avena, L., Blondel, O., and Faggionato, A. (2018). Analysis of random walks in dynamic random environments via L^2-perturbations. *Stochastic Processes and their Applications*, 128(10):3490 – 3530.

[34] Bachelier, L. (1900). Theory of speculation: The origins of modern finance. *Francia: Gauthier-Villars*.

[35] Baek, S. K., Minnhagen, P., and Kim, B. J. (2009). Percolation on hyperbolic lattices. *Phys. Rev. E*, 79:011124.

[36] Bakhtin, Y. Y. and Bulinskiĭ, A. V. (1997). Moment inequalities for sums of dependent multi-indexed random variables. *Fundam. Prikl. Mat.*, 3(4):1101–1108.

[37] Balankin, A., Martínez-Cruz, M., Susarrey-Huerta, O., and Damian Adame, L. (2018). Percolation on infinitely ramified fractal networks. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 382(1):12–19.

[38] Balankin, A. S., Martínez-Cruz, M., Álvarez-Jasso, M., Patiño-Ortiz, M., and Patiño-Ortiz, J. (2019). Effects of ramification and connectivity degree on site percolation threshold on regular lattices and fractal networks. *Phys. Lett. A*, 383(10):957–966.

[39] Balka, R., Buczolich, Z., and Elekes, M. (2015). A new fractal dimension: the topological Hausdorff dimension. *Adv. Math.*, 274:881–927.

[40] Ball, K. (1992). Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172.

[41] Barlow, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.*, 32(4):3024–3084.

[42] Barlow, M. T. (2016). Loop erased walks and uniform spanning trees. 34:1–32.

[43] Barlow, M. T., Coulhon, T., and Kumagai, T. (2005). Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs. *Comm. Pure Appl. Math.*, 58(12):1642–1677.

[44] Barlow, M. T., Croydon, D. A., and Kumagai, T. (2017). Subsequential scaling limits of simple random walk on the two-dimensional uniform spanning tree. *The Annals of Probability*, 45(1):4 – 55.

[45] Barlow, M. T., Croydon, D. A., and Kumagai, T. (2021). Quenched and averaged tails of the heat kernel of the two-dimensional uniform spanning tree. *Probability Theory and Related Fields*, 181(1):57–111.

[46] Barlow, M. T. and Deuschel, J.-D. (2010). Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.*, 38(1):234–276.

[47] Barlow, M. T. and Hambly, B. M. (2009). Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab*, 14(1):1–27.

[48] Barlow, M. T. and Járai, A. A. (2019). Geometry of uniform spanning forest components in high dimensions. *Canad. J. Math.*, 71(6):1297–1321.

[49] Barlow, M. T., Járai, A. A., Kumagai, T., and Slade, G. (2008). Random walk on the incipient infinite cluster for oriented percolation in high dimensions. *Comm. Math. Phys.*, 278(2):385–431.

[50] Barlow, M. T. and Kumagai, T. (2006). Random walk on the incipient infinite cluster on trees. *Illinois J. Math.*, 50(1-4):33–65.

[51] Barlow, M. T. and Masson, R. (2011). Spectral dimension and random walks on the two dimensional uniform spanning tree. *Communications in mathematical physics*, 305(1):23–57.

[52] Barlow, M. T., Peres, Y., and Sousi, P. (2012). Collisions of random walks. *Ann. Inst. H. Poincaré Probab. Statist.*, 48(4):922–946.

[53] Bass, H. (1972). The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc. (3)*, 25:603–614.

[54] Bauerschmidt, R., Crawford, N., and Helmuth, T. (2021a). Percolation transition for random forests in $d \geq 3$. *arXiv preprint arXiv:2107.01878*.

[55] Bauerschmidt, R., Crawford, N., Helmuth, T., and Swan, A. (2021b). Random spanning forests and hyperbolic symmetry. *Comm. Math. Phys.*, 381(3):1223–1261.

[56] Bauerschmidt, R. and Helmuth, T. (2021). Spin systems with hyperbolic symmetry: a survey. *arXiv preprint arXiv:2109.02566*.

[57] Bella, P. and Schäffner, M. (2020). Quenched invariance principle for random walks among random degenerate conductances. *Ann. Probab.*, 48(1):296–316.

[58] Bella, P. and Schäffner, M. (2022). Non-uniformly parabolic equations and applications to the random conductance model. *Probab. Theory Related Fields*, 182(1-2):353–397.

[59] Ben Arous, G., Cabezas, M., and Fribergh, A. (2019). Scaling limit for the ant in a simple high-dimensional labyrinth. *Probab. Theory Related Fields*, 174(1-2):553–646.

[60] Ben-Avraham, D. and Havlin, S. (2000). *Diffusion and reactions in fractals and disordered systems*. Cambridge University Press, Cambridge.

[61] Ben-Avraham, D., Havlin, S., and Movshovitz, D. (1984). Infinitely ramified fractal lattices and percolation. *Philos. Mag. B*, 50(2):297–306.

[62] Benedetti, D. (2015). Critical behavior in spherical and hyperbolic spaces. *J. Stat. Mech.: Theory Exp.*, 2015(1):P01002.

[63] Benjamini, I., Gurel-Gurevich, O., and Lyons, R. (2007). Recurrence of random walk traces. *Ann. Probab.*, 35(2):732–738.

[64] Benjamini, I., Gurel-Gurevich, O., and Schramm, O. (2011a). Cutpoints and resistance of random walk paths. *Ann. Probab.*, 39(3):1122–1136.

[65] Benjamini, I. and Hermon, J. (2020). Recurrence of Markov chain traces. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(1):734–759.

[66] Benjamini, I., Kesten, H., Peres, Y., and Schramm, O. (2004). Geometry of the uniform spanning forest: Transitions in dimensions 4,8,12,... *Annals of Mathematics*, 160(2):465–491.

[67] Benjamini, I., Kesten, H., Peres, Y., and Schramm, O. (2011b). Geometry of the uniform spanning forest: transitions in dimensions 4, 8, 12,.... *Selected Works of Oded Schramm*, pages 751–777.

[68] Benjamini, I., Lyons, R., Peres, Y., and Schramm, O. (2001). Uniform spanning forests. *The Annals of Probability*, 29(1):1 – 65.

[69] Benjamini, I., Pemantle, R., and Peres, Y. (1995). Martin capacity for Markov chains. *Ann. Probab.*, 23(3):1332–1346.

[70] Benjamini, I. and Peres, Y. (1994). Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields*, 98(1):91–112.

[71] Benjamini, I. and Schramm, O. (2001). Recurrence of Distributional Limits of Finite Planar Graphs. *Electronic Journal of Probability*, 6(none):1 – 13.

[72] Berestycki, N., Laslier, B., and Ray, G. (2020). Dimers and imaginary geometry. *The Annals of Probability*, 48(1):1–52.

[73] Berg, H. C. (1984). *Random Walks in Biology*. Princeton University Press, Princeton.

[74] Berger, N. and Biskup, M. (2007). Quenched invariance principle for simple random walk on percolation clusters. *Probability Theory and Related Fields*, 137(1):83–120.

[75] Berger, N., Biskup, M., Hoffman, C. E., and Kozma, G. (2008). Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):374–392.

[76] Bhupatiraju, S., Hanson, J., and Járai, A. A. (2017). Inequalities for critical exponents in d-dimensional sandpiles. *Electronic Journal of Probability*, 22(none):1 – 51.

[77] Biskup, M. (2011). Recent progress on the random conductance model. *Probab. Surv.*, 8:294–373.

[78] Biskup, M. (2020). Extrema of the two-dimensional discrete Gaussian free field. In *Random graphs, phase transitions, and the Gaussian free field*, volume 304 of *Springer Proc. Math. Stat.*, pages 163–407. Springer, Cham.

[79] Biskup, M. and Boukhadra, O. (2012). Subdiffusive heat-kernel decay in four-dimensional i.i.d. random conductance models. *J. Lond. Math. Soc. (2)*, 86(2):455–481.

[80] Biskup, M., Chen, X., Kumagai, T., and Wang, J. (2021). Quenched invariance principle for a class of random conductance models with long-range jumps. *Preprint, available at arXiv:2004.01971*, 180(3-4):847–889.

[81] Biskup, M., Louidor, O., Rozinov, A., and Vandenberg-Rodes, A. (2013). Trapping in the Random Conductance Model. *Journal of Statistical Physics*, 150(1):66–87.

[82] Biskup, M. and Rodriguez, P.-F. (2018). Limit theory for random walks in degenerate time-dependent random environments. *Journal of Functional Analysis*, 274(4):985 – 1046.

[83] Blachère, S. (2003a). Cut times for random walks on the discrete Heisenberg group. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):621–638.

[84] Blachère, S. (2003b). Word distance on the discrete Heisenberg group. *Colloq Math*, 95.

[85] Blachère, S., Haïssinsky, P., and Mathieu, P. (2008). Asymptotic entropy and Green speed for random walks on countable groups. *Ann. Probab.*, 36(3):1134–1152.

[86] Bo-Ming, Y. and Kai-Lun, Y. (1988). Numerical evidence of the critical percolation probability $p_c = 1$ for site problems on sierpinski gaskets. *J. Phys. A: Math. Gen.*, 21(15):3269–3274.

[87] Broder, A. (1989). Generating random spanning trees. In *30th Annual Symposium on Foundations of Computer Science*, pages 442–447.

[88] Bruss, F. T. (1980). A counterpart of the Borel-Cantelli lemma. *J. Appl. Probab.*, 17(4):1094–1101.

[89] Brydges, D. C. and Imbrie, J. Z. (2003). Branched polymers and dimensional reduction. *Ann. of Math. (2)*, 158(3):1019–1039.

[90] Bulterman, R., van der Sommen, F., Zwaan, G., Verhoeff, T., van Gasteren, A., and Feijen, W. (2002). On computing a longest path in a tree. *Inform Process Lett*, 81(2):93–96.

[91] Burdzy, K. and Lawler, G. F. (1990a). Nonintersection exponents for Brownian paths. I. Existence and an invariance principle. *Probab. Theory Related Fields*, 84(3):393–410.

[92] Burdzy, K. and Lawler, G. F. (1990b). Nonintersection exponents for Brownian paths. II. Estimates and applications to a random fractal. *Ann. Probab.*, 18(3):981–1009.

[93] Burton, R. and Pemantle, R. (1993). Local Characteristics, Entropy and Limit Theorems for Spanning Trees and Domino Tilings Via Transfer-Impedances. *The Annals of Probability*, 21(3):1329 – 1371.

[94] Burton, R. M. and Keane, M. (1989). Density and uniqueness in percolation. *Communications in mathematical physics*, 121:501–505.

[95] Caracciolo, S., Jacobsen, J. L., Saleur, H., Sokal, A. D., and Sportiello, A. (2004). Fermionic field theory for trees and forests. *Physical review letters*, 93(8):080601.

[96] Caracciolo, S., Sokal, A. D., and Sportiello, A. (2007). Grassmann integral representation for spanning hyperforests. *Journal of Physics A: Mathematical and Theoretical*, 40(46):13799.

[97] Caracciolo, S., Sokal, A. D., and Sportiello, A. (2017). Spanning forests and $\mathrm{OSP}(n|2m)$-invariant σ-models. *Journal of Physics A: Mathematical and Theoretical*, 50(11):114001.

[98] Cardy, J. (1996). *Scaling and Renormalization in Statistical Physics*. Cambridge Lecture Notes in Physics. Cambridge University Press.

[99] Carne, T. K. (1985). A transmutation formula for Markov chains. *Bull. Sci. Math. (2)*, 109(4):399–405.

[100] Chandler, R., Koplik, J., Lerman, K., and Willemsen, J. F. (1982). Capillary displacement and percolation in porous media. *J. Fluid Mech.*, 119:249–267.

[101] Chayes, J. T., Chayes, L., and Newman, C. M. (1985). The stochastic geometry of invasion percolation. *Commun. Math. Phys.*, 101(3):383–407.

[102] Chen, X. (2016). Gaussian bounds and collisions of variable speed random walks on lattices with power law conductances. *Stochastic Processes and their Applications*, 126(10):3041–3064.

[103] Chen, X. and Chen, D. (2010). Two random walks on the open cluster of $\mathbb{Z}^2$ meet infinitely often. *Science China Mathematics*, 53(8):1971–1978.

[104] Chen, X. and Chen, D. (2011). Some sufficient conditions for infinite collisions of simple random walks on a wedge comb. *Electron. J. Probab.*, 16:1341–1355.

[105] Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of The Royal Society Interface*, 5(25):813–834.

[106] Cornulier, Y. d. (2018). *On the quasi-isometric classification of locally compact groups*, volume 447 of *London Math. Soc. Lecture Note Ser.*, pages 275–342. Cambridge Univ. Press, Cambridge.

[107] Coulhon, T. and Saloff-Coste, L. (1993). Isopérimétrie pour les groupes et les variétés. *Rev. Mat. Iberoamericana*, 9(2):293–314.

[108] Cranston, M. C. and Mountford, T. S. (1991). An extension of a result of Burdzy and Lawler. *Probab. Theory Related Fields*, 89(4):487–502.

[109] Csáki, E., Földes, A., and Révész, P. (2010). On the number of cutpoints of the transient nearest neighbor random walk on the line. *J. Theoret. Probab.*, 23(2):624–638.

[110] Curien, N. (2018). Random graphs: the local convergence point of view. Unpublished lecture notes. Available at https://www.imo.universite-paris-saclay.fr/~nicolas.curien/enseignement.html.

[111] Curien, N., Hutchcroft, T., and Nachmias, A. (2020). Geometric and spectral properties of causal maps. *Journal of the European Mathematical Society*, 22(12):3997–4024.

[112] Dario, P. and Gu, C. (2021). Quantitative homogenization of the parabolic and elliptic Green's functions on percolation clusters. *Ann. Probab.*, 49(2):556–636.

[113] de Gennes, P. G. et al. (1976). La percolation: un concept unificateur. *La recherche*, 7(72):919–927.

[114] de Graaf, W. A. (2007). Classification of 6-dimensional nilpotent Lie algebras over fields of characteristic not 2. *J. Algebra*, 309(2):640–653. NB: Further proof details are given in the arXiv version.

[115] Delmotte, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana*, 15(1):181–232.

[116] Delmotte, T. and Deuschel, J.-D. (2005). On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\Delta\phi$ interface model. *Probability Theory and Related Fields*, 133:358–390.

[117] Deng, Y., Garoni, T. M., and Sokal, A. D. (2007). Ferromagnetic phase transition for the spanning-forest model ($q \to 0$ limit of the Potts model) in three or more dimensions. *Physical review letters*, 98(3):030602.

[118] Derbez, E. and Slade, G. (1997). Lattice trees and super-Brownian motion. *Canad. Math. Bull.*, 40(1):19–38.

[119] Derbez, E. and Slade, G. (1998). The scaling limit of lattice trees in high dimensions. *Comm. Math. Phys.*, 193(1):69–104.

[120] Dereudre, D. (2022). Fully-connected bond percolation on $\mathbb{Z}^d$. *Probability Theory and Related Fields*, 183(1-2):547–579.

[121] Deuschel, J.-D., Nguyen, T. A., and Slowik, M. (2018). Quenched invariance principles for the random conductance model on a random graph with degenerate ergodic weights. *Probab. Theory Related Fields*, 170(1-2):363–386.

[122] Devulder, A., Gantert, N., and Pène, F. (2018). Collisions of several walkers in recurrent random environments. *Electronic Journal of Probability*, 23.

[123] Devulder, A., Gantert, N., and Pène, F. (2019). Arbitrary many walkers meet infinitely often in a subballistic random environment. *Electron. J. Probab.*, 24:25 pp.

[124] Dhar, D. (1990). Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64:1613–1616.

[125] Dhar, D. (1999). The abelian sandpile and related models. *Physica A: Statistical Mechanics and its Applications*, 263(1):4–25. Proceedings of the 20th IUPAP International Conference on Statistical Physics.

[126] Diaconis, P. and Freedman, D. (1980). de Finetti's theorem for Markov chains. *Ann. Probab.*, 8(1):115–130.

[127] Ding, J., Lee, J. R., and Peres, Y. (2013). Markov type and threshold embeddings. *Geom. Funct. Anal.*, 23(4):1207–1229.

[128] Dolgopyat, D., Keller, G., and Liverani, C. (2008). Random walk in Markovian environment. *The Annals of Probability*, 36(5):1676–1710.

[129] Druţu, C. (2002). Quasi-isometry invariants and asymptotic cones. *Internat. J. Algebra Comput.*, 12(1-2):99–135.

[130] Duminil-Copin, H. (2008). Law of the iterated logarithm for the random walk on the infinite percolation cluster. *arXiv:0809.4380*.

[131] Duplantier, B. (1988). Intersections of random walks. A direct renormalization approach. *Comm. Math. Phys.*, 117(2):279–329.

[132] Easo, P. (2022). The wired arboreal gas on regular trees. *Electronic Communications in Probability*, 27:1–10.

[133] Engelking, R. (1978). *Dimension theory*, volume 19 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-Oxford-New York; PWN—Polish Scientific Publishers, Warsaw. Translated from the Polish and revised by the author.

[134] Erdős, P. and Taylor, S. J. (1960). Some intersection properties of random walk paths. *Acta Math. Acad. Sci. Hungar.*, 11:231–248.

[135] Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474.

[136] Fabes, E. B. and Stroock, D. W. (1986). A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338.

[137] Falconer, K. (2014). *Fractal geometry*. John Wiley & Sons, Ltd., Chichester, third edition. Mathematical foundations and applications.

[138] Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, 21(5):55–59.

[139] Feder, T. and Mihail, M. (1992). Balanced matroids. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 26–38.

[140] Fischer, V. and Ruzhansky, M. (2016). *Quantization on nilpotent Lie groups*, volume 314 of *Progress in Mathematics*. Birkhäuser/Springer, [Cham].

[141] Folz, M. (2011). Gaussian upper bounds for heat kernels of continuous time simple random walks. *Electron. J. Probab.*, 16:no. 62, 1693–1722.

[142] Folz, M. (2014). Volume growth and stochastic completeness of graphs. *Trans. Amer. Math. Soc.*, 366(4):2089–2119.

[143] Fortuin, C. M., Kasteleyn, P. W., and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22(2):89–103.

[144] Fredes, L. and Francois Marckert, J. (2021). Models of random subtrees of a graph.

[145] Friedli, S. and Velenik, Y. (2017). *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press.

[146] Funaki, T. (2005). Stochastic interface models. In *Lectures on probability theory and statistics*, volume 1869 of *Lecture Notes in Math.*, pages 103–274. Springer, Berlin.

[147] Gail, M. and Boone, C. (1970). The locomotion of mouse fibroblasts in tissue culture. *Biophysical journal*, 10(10):980—993.

[148] Gallesco, C. (2013). Meeting time of independent random walks in random environment. *ESAIM: Probability and Statistics*, 17:257–292.

[149] Ganguly, S. and Lee, J. R. (2022). Chemical subdiffusivity of critical 2D percolation. *Comm. Math. Phys.*, 389(2):695–714.

[150] Ganguly, S., Lee, J. R., and Peres, Y. (2017). Diffusive estimates for random walks on stationary random graphs of polynomial growth. *Geom. Funct. Anal.*, 27(3):596–630.

[151] Gantert, N., Kochler, M., and Pene, F. (2014). On the recurrence of some random walks in random environment. *arXiv preprint arXiv:1404.3874*.

[152] Gefen, Y., Aharony, A., and Mandelbrot, B. B. (1983a). Phase transitions on fractals. i. quasi-linear lattices. *Journal of Physics A: Mathematical and General*, 16(6):1267–1278.

[153] Gefen, Y., Aharony, A., and Mandelbrot, B. B. (1984a). Phase transitions on fractals. III. infinitely ramified lattices. *Journal of Physics A: Mathematical and General*, 17(6):1277–1289.

[154] Gefen, Y., Aharony, A., Mandelbrot, B. B., and Kirkpatrick, S. (1981). Solvable fractal family, and its possible relation to the backbone at percolation. *Phys. Rev. Lett.*, 47:1771–1774.

[155] Gefen, Y., Aharony, A., Shapir, Y., and Mandelbrot, B. B. (1984b). Phase transitions on fractals. II. sierpinski gaskets. *Journal of Physics A: Mathematical and General*, 17(2):435–444.

[156] Gefen, Y., Mandelbrot, B. B., and Aharony, A. (1980). Critical phenomena on fractal lattices. *Phys. Rev. Lett.*, 45:855–858.

[157] Gefen, Y., Meir, Y., Mandelbrot, B. B., and Aharony, A. (1983b). Geometric implementation of hypercubic lattices with noninteger dimensionality by use of low lacunarity fractal lattices. *Phys. Rev. Lett.*, 50:145–148.

[158] Gendiar, A., Daniška, M., Krčmár, R., and Nishino, T. (2014). Mean-field universality class induced by weak hyperbolic curvatures. *Phys. Rev. E*, 90:012122.

[159] Glazman, A. and Manolescu, I. (2023). Structure of Gibbs measure for planar FK-percolation and Potts models. *arXiv preprint arXiv:2106.02403*, 4(2):209–256.

[160] Gloria, A., Neukamm, S., and Otto, F. (2013). Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics – long version. *MPI Leipzig, preprint 3*.

[161] Gloria, A., Neukamm, S., and Otto, F. (2015). Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515.

[162] Gracey, J. A. (2015). Four loop renormalization of ϕ^3 theory in six dimensions. *Phys. Rev. D*, 92:025012.

[163] Grimmett, G. (1999). *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition.

[164] Grimmett, G. (2006). *The random-cluster model*, volume 333 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin.

[165] Grimmett, G. and Stirzaker, D. (2001). *Probability and random processes*, volume 80. Oxford university press.

[166] Grimmett, G. R., Kesten, H., and Zhang, Y. (1993). Random walk on the infinite cluster of the percolation model. *Probability Theory and Related Fields*, 96(1):33–44.

[167] Grimmett, G. R. and Marstrand, J. M. (1990). The supercritical phase of percolation is well behaved. *Proc. Roy. Soc. London Ser. A*, 430(1879):439–457.

[168] Grimmett, G. R. and Winkler, S. N. (2004). Negative association in uniform forests and connected graphs. *Random Structures Algorithms*, 24(4):444–460.

[169] Gromov, M. (1981). Groups of polynomial growth and expanding maps. *Publ. Math. Inst. Hautes Études Sci.*, 53:53–73.

[170] Gromov, M. (1996). *Carnot-Carathéodory spaces seen from within*, volume 144 of *Progr. Math.*, pages 79–323. Birkhäuser, Basel.

[171] Guivarc'h, Y. (1973). Croissance polynomiale et périodes des fonctions harmoniques. *Bull. Soc. Math. France*, 101:333–379.

[172] Gurel-Gurevich, O. and Nachmias, A. (2013). Recurrence of planar graph limits. *Ann. of Math. (2)*, 177(2):761–781.

[173] Gwynne, E. and Hutchcroft, T. (2020a). Anomalous diffusion of random walk on random planar maps. *Probability Theory and Related Fields*, pages 1–45.

[174] Gwynne, E. and Hutchcroft, T. (2020b). Anomalous diffusion of random walk on random planar maps. *Probab. Theory Related Fields*, 178(1-2):567–611.

[175] Gwynne, E. and Miller, J. (2021). Random walk on random planar maps: spectral dimension, resistance and displacement. *Ann. Probab.*, 49(3):1097–1128.

[176] Halberstam, N. and Hutchcroft, T. (2022). Collisions of random walks in dynamic random environments. *Electron. J. Probab.*, 27:Paper No. 8, 18.

[177] Hara, T. and Slade, G. (1990a). Mean-field critical behaviour for percolation in high dimensions. *Comm. Math. Phys.*, 128(2):333–391.

[178] Hara, T. and Slade, G. (1990b). On the upper critical dimension of lattice trees and lattice animals. *Journal of Statistical Physics*, 59(5):1469–1510.

[179] Hara, T. and Slade, G. (1992). The number and size of branched polymers in high dimensions. *Journal of Statistical Physics*, 67(5):1009–1038.

[180] Hassan, M. K. and Rahman, M. M. (2015). Percolation on a multifractal scale-free planar stochastic lattice and its universality class. *Phys. Rev. E*, 92:040101.

[181] Havlin, S., Ben-Avraham, D., and Movshovitz, D. (1983). Percolation on fractal lattices. *Phys. Rev. Lett.*, 51:2347–2350.

[182] Havlin, S., Ben-Avraham, D., and Movshovitz, D. (1984). Percolation on infinitely ramified fractals. *J. Stat. Phys.*, 36(5):831–841.

[183] Helffer, B. and Sjoestrand, J. (1994). On the correlation for Kac-like models in the convex case. *Journal of Statistical Physics*, 74:349–409.

[184] Hermon, J. and Sousi, P. (2020). A comparison principle for random walk on dynamical percolation. *The Annals of Probability*, 48(6):2952 – 2987.

[185] Holden, N. and Sun, X. (2018). SLE as a mating of trees in euclidean geometry. *Communications in Mathematical Physics*, 364(1):171–201.

[186] Hsu, H.-P., Nadler, W., and Grassberger, P. (2005). Simulations of lattice animals and trees. *Journal of Physics A: Mathematical and General*, 38(4):775–806.

[187] Huang, K. (2013). A critical history of renomalization. *Int. J. Mod. Phys. A*, 28(29):1330050.

[188] Hutchcroft, T. (2016). Wired cycle-breaking dynamics for uniform spanning forests. *Annals of Probability*, 44(6):3879–3892.

[189] Hutchcroft, T. (2018a). Interlacements and the wired uniform spanning forest. *The Annals of Probability*, 46(2):1170–1200.

[190] Hutchcroft, T. (2018b). Universality of high-dimensional spanning forests and sandpiles. *Probability Theory and Related Fields*, 176:533–597.

[191] Hutchcroft, T. (2019). Percolation on hyperbolic graphs. *Geom. Funct. Anal.*, 29(3):766–810.

[192] Hutchcroft, T. (2020a). Indistinguishability of collections of trees in the uniform spanning forest. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 917–927. Institut Henri Poincaré.

[193] Hutchcroft, T. (2020b). Universality of high-dimensional spanning forests and sandpiles. *Probab. Theory Related Fields*, 176(1-2):533–597.

[194] Hutchcroft, T. and Nachmias, A. (2017). Indistinguishability of trees in uniform spanning forests. *Probability Theory and Related Fields*, 168:113–152.

[195] Hutchcroft, T. and Peres, Y. (2015). Collisions of random walks in reversible random graphs. *Electron. Commun. Probab.*, 20:6 pp.

[196] Hutchcroft, T. and Peres, Y. (2019). The component graph of the uniform spanning forest: transitions in dimensions $9, 10, 11, \ldots$. *Probability Theory and Related Fields*, 175(1):141–208.

[197] Hutchcroft, T. and Sousi, P. (2023). Logarithmic corrections to scaling in the four-dimensional uniform spanning tree. *Communications in Mathematical Physics*.

[198] Häggström, O. (1995). Random-cluster measures and uniform spanning trees. *Stochastic Processes and their Applications*, 59(2):267–275.

[199] Jacobsen, J. L., Salas, J., and Sokal, A. D. (2005). Spanning forests and the q-state Potts model in the limit $q \to 0$. *Journal of statistical physics*, 119:1153–1281.

[200] Jain, N. C. and Orey, S. (1973). Some properties of random walk paths. *Journal of Mathematical Analysis and Applications*, 43(3):795–815.

[201] James, N., Lyons, R., and Peres, Y. (2008). A transient Markov chain with finitely many cutpoints. In *Probability and statistics: essays in honor of David A. Freedman*, volume 2 of *Inst. Math. Stat. (IMS) Collect.*, pages 24–29. Inst. Math. Statist., Beachwood, OH.

[202] James, N. and Peres, Y. (1997). Cutpoints and exchangeable events for random walks. *Theory of Probability & Its Applications*, 41(4):666–677.

[203] Janse van Rensburg, E. J. and Madras, N. (1992). A nonlocal Monte Carlo algorithm for lattice trees. *J. Phys. A: Math. Gen.*, 25(2):303–333.

[204] Janse van Rensburg, E. J. and Rechnitzer, A. (2003). High precision canonical Monte Carlo determination of the growth constant of square lattice trees. *Phys. Rev. E*, 67:036116.

[205] Janson, S. (2011). Probability asymptotics: notes on notation. Unpublished note, available at arXiv:1108.3924.

[206] Janson, S., Łuczak, T., and Rucinski, A. (2000). *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York.

[207] Járai, A. A. (2018). Sandpile models. *Probability Surveys*, 15(none):243 – 306.

[208] Járai, A. A. and Mata López, D. (2022). Logarithmic correction to resistance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(3):1775–1807.

[209] Járai, A. A. and Nachmias, A. (2014). Electrical resistance of the low dimensional critical branching random walk. *Comm. Math. Phys.*, 331(1):67–109.

[210] Járai, A. A. and Redig, F. (2008). Infinite volume limit of the abelian sandpile model in dimensions $d \geq 3$. *Probab. Theory Related Fields*, 141(1-2):181–212.

[211] Járai, A. A. and Werning, N. (2014). Minimal configurations and sandpile measures. *J. Theoret. Probab.*, 27(1):153–167.

[212] Joag-Dev, K., Perlman, M. D., and Pitt, L. D. (1983). Association of normal random variables and slepian's inequality. *Ann. Probab.*, 11(2):451–455.

[213] Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables with applications. *Ann. Statist.*, 11(1):286–295.

[214] Kaĭmanovich, V. A. and Vershik, A. M. (1983). Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, 11(3):457–490.

[215] Kenyon, R. (2000). The asymptotic determinant of the discrete Laplacian. *Acta Mathematica*, 185(2):239 – 286.

[216] Kenyon, R. and Winkler, P. (2009a). Branched polymers. *The American Mathematical Monthly*, 116(7):612–628.

[217] Kenyon, R. and Winkler, P. (2009b). Branched polymers. *Amer. Math. Monthly*, 116(7):612–628.

[218] Kesten, H. (1986a). The incipient infinite cluster in two-dimensional percolation. *Probab. Theory Related Fields*, 73(3):369–394.

[219] Kesten, H. (1986b). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487.

[220] Kipnis, C. and Varadhan, S. R. S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19.

[221] Kirchhoff, G. (1847). Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Ann. Phys. und Chem.*, (72):497–508.

[222] Kirchoff, G. R. (1847). Über die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik*, 148(2):497–508.

[223] Kozma, G. and Nachmias, A. (2009). The Alexander-Orbach conjecture holds in high dimensions. *Invent. Math.*, 178(3):635–654.

[224] Krcmar, R., Gendiar, A., Ueda, K., and Nishino, T. (2008). Ising model on a hyperbolic lattice studied by the corner transfer matrix renormalization group method. *J. Phys. A*, 41(12):125001, 8.

[225] Krishnapur, M. and Peres, Y. (2004). Recurrent Graphs where Two Independent Random Walks Collide Finitely Often. *Electronic Communications in Probability*, 9(none):72 – 81.

[226] Kumagai, T. (2014). *Random walks on disordered media and their scaling limits*, volume 2101 of *Lecture Notes in Mathematics*. Springer, Cham. Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010, École d'Été de Probabilités de Saint-Flour.

[227] Kumagai, T. and Misumi, J. (2008). Heat kernel estimates for strongly recurrent random walk on random media. *J. Theoret. Probab.*, 21(4):910–935.

[228] Kumagai, T. and Nakamura, C. (2016). Laws of the iterated logarithm for random walks on random conductance models. In *Stochastic analysis on large scale interacting systems*, RIMS Kôkyûroku Bessatsu, B59, pages 141–156. Res. Inst. Math. Sci. (RIMS), Kyoto.

[229] Lawler, G., Sun, X., and Wu, W. (2019). Four-dimensional loop-erased random walk. *Ann. Probab.*, 47(6):3866–3910.

[230] Lawler, G. F. (1979). *A SELF-AVOIDING RANDOM WALK*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–Princeton University.

[231] Lawler, G. F. (1980). A self-avoiding random walk. *Duke Math. J.*, 47(3):655–693.

[232] Lawler, G. F. (1986). Gaussian behavior of loop-erased self-avoiding random walk in four dimensions. *Duke Math. J.*, 53(1):249–269.

[233] Lawler, G. F. (1991). *Intersections of random walks*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA.

[234] Lawler, G. F. (1992). Escape probabilities for slowly recurrent sets. *Probab. Theory Related Fields*, 94(1):91–117.

[235] Lawler, G. F. (1995). The logarithmic correction for loop-erased walk in four dimensions. In *Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993)*, number Special Issue, pages 347–361.

[236] Lawler, G. F. (1996a). Cut times for simple random walk. *Electron. J. Probab.*, 1:no. 13, approx. 24 pp.

[237] Lawler, G. F. (1996b). Hausdorff dimension of cut points for Brownian motion. *Electron. J. Probab.*, 1:no. 2, approx. 20 pp.

[238] Lawler, G. F. (2013). *Intersections of random walks*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York. Reprint of the 1996 edition.

[239] Lawler, G. F., Schramm, O., and Werner, W. (2004). Conformal invariance of planar loop-erased random walks and uniform spanning trees. *The Annals of Probability*, 32(1B):939 – 995.

[240] Le Donne, E. (2017). A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries. *Anal. Geom. Metr. Spaces*, 5(1):116–137.

[241] Le Ny, A. (2008). Introduction to (generalized) Gibbs measures. *Ensaios Matemáticos*, 15(1-126):7.

[242] Leath, P. L. (1976a). Cluster shape and critical exponents near percolation threshold. *Phys. Rev. Lett.*, 36:921–924.

[243] Leath, P. L. (1976b). Cluster size and boundary distribution near percolation threshold. *Phys. Rev. B*, 14(11):5046–5055.

[244] Lee, J. R. (2020). Relations between scaling exponents in unimodular random graphs. *arXiv preprint arXiv:2007.06548*.

[245] Lee, J. R. (2021). Conformal growth rates and spectral geometry on distributional limits of graphs. *Ann. Probab.*, 49(6):2671–2731.

[246] Lejay, A. (2003). Simulating a diffusion on a graph. Application to reservoir engineering. *Monte Carlo Methods Appl.*, 9(3):241–255.

[247] Lenormand, R. and Bories, S. (1980). Description d'un mécanisme de connexion de liaison destiné à l'étude du drainage avec piégeage en milieu poreux. *Comptes Rendus de l'Académie des Sciences*, B 291:279.

[248] Lévy, P. (1954). *Théorie de l'addition des variables aléatoires*. Monographies des probabilités ; calcul des probabilités et ses applications. Gauthier-Villars, Paris, 2. ed. edition.

[249] Liggett, T. M. (2012). *Interacting particle systems*, volume 276. Springer Science & Business Media.

[250] Lin, Z.-Q. and Yang, Z. (1997). Thresholds and universality of the site percolation on the Sierpinski carpets. *Commun. Theor. Phys.*, 27(2):145–152.

[251] Llosa Isenrich, C., Pallier, G., and Tessera, R. (2023). Cone-equivalent nilpotent groups with different Dehn functions. *Proc. Lond. Math. Soc. (3)*, 126(2):704–789.

[252] Lo, C. H., Menshikov, M. V., and Wade, A. R. (2021). Cutpoints of non-homogeneous random walks. *ALEA - Latin American Journal of Probability and Mathematical Statistics*.

[253] Lorenz, C. D. and Ziff, R. M. (1998). Precise determination of the bond percolation thresholds and finite-size scaling corrections for the sc, fcc, and bcc lattices. *Phys. Rev. E*, 57:230–236.

[254] Lubensky, T. C. and Isaacson, J. (1979). Statistics of lattice animals and dilute branched polymers. *Phys. Rev. A*, 20:2130–2146.

[255] Łuczak, T. and Pittel, B. (1992). Components of random forests. *Combinatorics, Probability and Computing*, 1(1):35–52.

[256] Lupu, T. (2016). From loop clusters and random interlacements to the free field. *Ann. Probab.*, 44(3):2117–2146.

[257] Lyons, R. (2005). Asymptotic enumeration of spanning trees. *Combinatorics, Probability and Computing*, 14(4):491–522.

[258] Lyons, R., Morris, B., and Schramm, O. (2008). Ends in Uniform Spanning Forests. *Electronic Journal of Probability*, 13:1702 – 1725.

[259] Lyons, R. and Peres, Y. (2016). *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York.

[260] Lyons, R. and Peres, Y. (2021). Poisson boundaries of lamplighter groups: proof of the Kaimanovich-Vershik conjecture. *J. Eur. Math. Soc. (JEMS)*, 23(4):1133–1160.

[261] Lyons, R., Peres, Y., and Schramm, O. (2003). Markov chain intersections and the loop-erased walk. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 39(5):779–791.

[262] Lyons, R., Peres, Y., and Sun, X. (2020). Induced graphs of uniform spanning forests. *Annales de L'Institut Henri Poincare Section (B) Probability and Statistics*, 56(4):2732–2744.

[263] Lyons, T. J. and Zheng, W. A. (1988). A crossing estimate for the canonical process on a Dirichlet space and a tightness result. Number 157-158, pages 249–271.

[264] Ma, S.-k. (1973). Introduction to the renormalization group. *Rev. Mod. Phys.*, 45:589–614.

[265] Madras, N. and Wu, C. C. (2010). Trees, animals, and percolation on hyperbolic lattices. *Electron. J. Probab.*, 15:no. 66, 2019–2040.

[266] Majumdar, S. N. and Dhar, D. (1992). Equivalence between the abelian sandpile model and the $q \to 0$ limit of the potts model. *Physica A*, (185):129–145.

[267] Martin, J. B. and Yeo, D. (2018). Critical random forests. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(2):913–960.

[268] Mathieu, P. and Piatnitski, A. (2007). Quenched invariance principles for random walks on percolation clusters. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2085):2287–2307.

[269] McDiarmid, C. (1998). *Concentration*, pages 195–248. Springer Berlin Heidelberg, Berlin, Heidelberg.

[270] Merkl, F., Öry, A., and Rolles, S. W. W. (2008). The 'magic formula' for linearly edge-reinforced random walks. *Statist. Neerlandica*, 62(3):345–363.

[271] Mertens, S. and Moore, C. (2017). Percolation thresholds in hyperbolic lattices. *Phys. Rev. E*, 96(4):042116, 13.

[272] Mertens, S. and Moore, C. (2018). Percolation thresholds and Fisher exponents in hypercubic lattices. *Phys. Rev. E*, 98:022120.

[273] Millán, A. P., Gori, G., Battiston, F., Enss, T., and Defenu, N. (2021). Complex networks with tuneable spectral dimension as a universality playground. *Phys. Rev. Research*, 3:023015.

[274] Monceau, P. and Hsiao, P.-Y. (2004a). Direct evidence for weak universality on fractal structures. *Physica A*, 331(1):1–9.

[275] Monceau, P. and Hsiao, P.-Y. (2004b). Percolation transition in fractal dimensions. *Phys. Lett. A*, 332(3):310–319.

[276] Monti, R. and Rickly, M. (2005). Geodetically convex sets in the Heisenberg group. *J. Convex Anal.*, 12(1):187–196.

[277] Morris, B. and Peres, Y. (2005). Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields*, 133(2):245–266.

[278] Mourrat, J.-C. and Otto, F. (2015). Anchored Nash inequalities and heat kernel bounds for static and dynamic degenerate environments. *Journal of Functional Analysis*, 270.

[279] Nachmias, A. and Tang, P. (2022). The wired minimal spanning forest on the poisson-weighted infinite tree. *arXiv preprint arXiv:2207.09305*.

[280] Naddaf, A. and Spencer, T. (1997). On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, 183(1):55–84.

[281] Naor, A., Peres, Y., Schramm, O., and Sheffield, S. (2006). Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197.

[282] Newman, C. M. (1980). Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.*, 74(2):119–128.

[283] Newman, M. E. J. and Ziff, R. M. (2001). Fast Monte Carlo algorithm for site or bond percolation. *Phys. Rev. E*, 64:016706.

[284] Norris, J. R. (1997). *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

[285] Pansu, P. (1983). Croissance des boules et des géodésiques fermées dans les nilvariétés. *Ergodic Theory Dynam. Systems*, 3(3):415–445.

[286] Parisi, G. and Sourlas, N. (1981). Critical behavior of branched polymers and the lee-yang edge singularity. *Phys. Rev. Lett.*, 46:871–874.

[287] Paul, G., Ziff, R. M., and Stanley, H. E. (2001). Percolation threshold, Fisher exponent, and shortest path exponent for four and five dimensions. *Phys. Rev. E*, 64:026115.

[288] Pemantle, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *The Annals of Probability*, 19(4):1559–1574.

[289] Pemantle, R. (2000). Towards a theory of negative dependence. *Journal of Mathematical Physics*, 41(3):1371–1390.

[290] Peres, Y., Sousi, P., and Steif, J. E. (2018). Quenched exit times for random walk on dynamical percolation. *Markov Processes and Related Fields*, 24(5):715–731.

[291] Peres, Y., Sousi, P., and Steif, J. E. (2020). Mixing time for random walk on supercritical dynamical percolation. *Probability Theory and Related Fields*, 176(3):809–849.

[292] Peres, Y., Stauffer, A., and Steif, J. E. (2015). Random walks on dynamical percolation: mixing times, mean squared displacement and hitting times. *Probab. Theory Related Fields*, 162(3-4):487–530.

[293] Peres, Y. and Zheng, T. (2020). On groups, slow heat kernel decay yields Liouville property and sharp entropy bounds. *Int. Math. Res. Not. IMRN*, (3):722–750.

[294] Pólya, G. (1921). Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, 84(1-2):149–160.

[295] Pólya, G. and Boas, R. (1984). Collected papers.

[296] Ray, G. and Xiao, B. (2022). Forests on wired regular trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 19(1):1035–1043.

[297] Redig, F., Saada, E., and Sau, F. (2020). Symmetric simple exclusion process in dynamic environment: hydrodynamics. *Electronic Journal of Probability*, 25:Paper No. 138, 47.

[298] Regnault, J. (1863). *Calcul des chances et philosophie de la bourse*. Pilloy.

[299] Rosenthal, H. P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel Journal of Mathematics*, 8:273–303.

[300] Schramm, O. (1999). Scaling limits of loop-erased random walks and uniform spanning trees. *Israel Journal of Mathematics*, 118:221–288.

[301] Schweinsberg, J. (2009). The loop-erased random walk and the uniform spanning tree on the four-dimensional discrete torus. *Probability Theory and Related Fields*, 144(3-4):319–370.

[302] Shao, Q.-M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *Journal of Theoretical Probability*, 13:343–356.

[303] Sheffield, S. (2006). Uniqueness of maximal entropy measure on essential spanning forests. *The Annals of Probability*, pages 857–864.

[304] Shiraishi, D. (2012). Exact value of the resistance exponent for four dimensional random walk trace. *Probab. Theory Related Fields*, 153(1-2):191–232.

[305] Shiraishi, D. and Watanabe, S. (2022). Volume and heat kernel fluctuations for the three-dimensional uniform spanning tree. *arXiv e-prints*, page arXiv:2211.15031.

[306] Simon, B. (2011). *Convexity: an analytic viewpoint*, volume 187. Cambridge University Press.

[307] Slade, G. (2006). *The Lace Expansion for Lattice Trees*, pages 77–86. Springer Berlin Heidelberg, Berlin, Heidelberg.

[308] Spitzer, F. (1970). Interaction of Markov processes. *Advances in Mathematics*, 5(2):246 – 290.

[309] Stein, R. (1979). Polymer solution properties. part ii. hydrodynamics and light scattering. *Journal of Polymer Science: Polymer Letters Edition*, 17(2):105–105.

[310] Swan, A. (2021). *Superprobability on graphs*. PhD thesis, University of Cambridge.

[311] Thomassen, C. (1992). Isoperimetric Inequalities and Transient Random Walks on Graphs. *The Annals of Probability*, 20(3):1592 – 1600.

[312] Timár, Á. (2018). Indistinguishability of the components of random spanning forests. *The Annals of Probability*, 46(4):2221–2242.

[313] Trofimov, V. I. (1985). Automorphism groups of graphs as topological groups. *Mat. Zametki*, 38(3):378–385, 476.

[314] van Engelenburg, D. and Hutchcroft, T. (2023). The number of ends in the uniform spanning tree for recurrent unimodular random graphs.

[315] Vandewalle, N. and Ausloos, M. (1997). Construction and properties of fractal trees with tunable dimension: The interplay of geometry and physics. *Phys. Rev. E*, 55:94–98.

[316] Varopoulos, N. T. (1985a). Isoperimetric inequalities and markov chains. *Journal of Functional Analysis*, 63(2):215–239.

[317] Varopoulos, N. T. (1985b). Long range estimates for Markov chains. *Bull. Sci. Math. (2)*, 109(3):225–252.

[318] Vershynin, R. (2012). Weak triangle inequalities for weak L^1 norm. Unpublished note. Available at https://www.math.uci.edu/~rvershyn/papers/weak-L1.pdf.

[319] Weiss, G. H. and Rubin, R. J. (1982). *Random Walks: Theory and Selected Applications*, pages 363–505. John Wiley & Sons, Ltd.

[320] Wilkinson, D. and Willemsen, J. F. (1983). Invasion percolation: a new form of percolation theory. *Journal of Physics A Mathematical General*, 16(14):3365–3376.

[321] Wilson, D. B. (1996). Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pages 296–303. ACM, New York.

[322] Xun, Z. and Ziff, R. M. (2020). Precise bond percolation thresholds on several four-dimensional lattices. *Phys. Rev. Research*, 2:013067.

[323] You, S. and Janse van Rensburg, E. J. (1998). Critical exponents and universal amplitude ratios in lattice trees. *Phys. Rev. E*, 58:3971–3976.

[324] Zhang, Z., Hou, P., Fang, S., Hu, H., and Deng, Y. (2021). Critical exponents and universal excess cluster number of percolation in four and five dimensions. *Phys. A*, 580:126124.

[325] Zinn-Justin, J. (2002). *Quantum Field Theory and Critical Phenomena*. Oxford University Press.