Kakeya-type sets over Cantor sets of directions in $\mathbb{R}^{d+1}$

Edward Kroc and Malabika Pramanik

Abstract

Given a Cantor-type subset $\Omega$ of a smooth curve in $\mathbb{R}^{d+1}$, we construct examples of sets that contain unit line segments with directions from $\Omega$ and exhibit analytical features similar to those of classical Kakeya sets of arbitrarily small $(d + 1)$-dimensional Lebesgue measure. The construction is based on probabilistic methods relying on the tree structure of $\Omega$, and extends to higher dimensions an analogous planar result of Bateman and Katz [4]. In particular, the existence of such sets implies that the directional maximal operator associated with the direction set $\Omega$ is unbounded on $L^p(\mathbb{R}^{d+1})$ for all $1 \leq p < \infty$.

Contents

1 Introduction 2
  1.1 Background ....................................................... 2
  1.2 Results .......................................................... 5

2 Overview of the proof of Theorem 1.2 8
  2.1 Steps of the proof and layout ..................................... 8
  2.2 Construction of a Kakeya-type set ................................ 9

3 Geometric Facts ...................................................... 11

4 Rooted, labelled trees .............................................. 16
  4.1 The terminology of trees .......................................... 16
  4.2 Encoding bounded subsets of the unit interval by trees ....... 18
  4.3 Encoding higher dimensional bounded subsets of Euclidean space by trees 20

5 Electrical circuits and percolation on trees ....................... 22
  5.1 The percolation process associated to a tree .................... 22
  5.2 Trees as electrical networks .................................... 23
  5.3 Estimating the survival probability after percolation ........ 25

2010 Mathematics Subject Classification. 28A75, 42B25 (primary), and 60K35 (secondary).
1 Introduction

1.1 Background

A Kakeya set (also called a Besicovitch set) in $\mathbb{R}^{d+1}$ is a set that contains a unit line segment in every direction. The study of such sets spans approximately a hundred years. The first major analytical result in this area, due to Besicovitch [5], shows that there exist Kakeya sets with Lebesgue measure zero. Over the past forty-plus years, dating back at least to the work of Fefferman [10], the study of Kakeya sets has been a simultaneously fruitful and vexing endeavor. On one hand its applications have been found in many deep and diverse corners of analysis, PDEs, additive combinatorics and number theory. On the other hand, certain fundamental questions concerning the size and dimensionality of such sets have eluded complete resolution.

In order to obtain quantitative estimates for analytical purposes, it is often convenient to work with the $\delta$-neighborhood of a Kakeya set, rather than the set itself. Here $\delta$ is an arbitrarily small positive constant. The $\delta$-neighborhood of a Kakeya set is therefore an object that consists of many thin $\delta$-tubes. A $\delta$-tube is by definition a cylinder of unit axial length and spherical cross-section of radius $\delta$. The defining property of a zero measure Kakeya set dictates that the volume of its $\delta$-neighborhood goes to zero as $\delta \to 0$, while the sum total of the sizes of these tubes is roughly a positive absolute constant. Indeed, a common construction of thin Kakeya sets in the plane (see for example [23, Chapter 10]) relies on the following fact: given any $\epsilon > 0$, there exists an integer $N \geq 1$ and a collection of $2^{-N}$-tubes, i.e., a family of $1 \times 2^{-N}$ rectangles, $\{P_t : 1 \leq t \leq 2^N\}$ in $\mathbb{R}^2$ such that

$$\left| \bigcup_t P_t \right| < \epsilon, \quad \text{and} \quad \sum_t |\tilde{P}_t| = 1. \quad (1.1)$$
Here $| \cdot |$ denotes Lebesgue measure (in this case two-dimensional), and $\tilde{P}_t$ denotes the “reach” of the tube $P_t$, namely the tube obtained by translating $P_t$ by two units in the positive direction along its axis. While it is not known that every Kakeya set in two or higher dimensions shares a similar feature, the ones that do have found repeated applications in analysis. Fundamental results have relied on the existence of such sets, for example the lack of differentiation for integral averages over parallelepipeds of arbitrary orientation, and the counterexample of the ball multiplier [23, Chapter 10]. The property described above continues to be the motivation for the Kakeya-type sets that we will study in the present paper.

**Definition 1.1.** For $d \geq 1$, we define a set of directions $\Omega$ to be a compact subset of $\mathbb{R}^{d+1}$. We say that a tube in $\mathbb{R}^{d+1}$ has orientation $\omega \in \Omega$ or a tube is oriented in direction $\omega$ if its axis is parallel to $\omega$. We say that $\Omega$ admits Kakeya-type sets if one can find a constant $C_0 \geq 1$ such that for any $N \geq 1$, there exists $\delta_N > 0$, $\delta_N \to 0$ as $N \to \infty$ and a collection of $\delta_N$-tubes $\{P_t^{(N)}\} \subset \mathbb{R}^{d+1}$ with orientations in $\Omega$ with the following property:

$$\text{if } E_N := \bigcup_t P_t^{(N)}, \quad E_N^*(C_0) := \bigcup_t C_0 P_t^{(N)}, \text{ then } \lim_{N \to \infty} \frac{|E_N^*(C_0)|}{|E_N|} = \infty. \quad (1.2)$$

Here $| \cdot |$ denotes $(d+1)$-dimensional Lebesgue measure, and $C_0 P_t^{(N)}$ denotes the tube with the same centre, orientation and cross-sectional radius as $P_t^{(N)}$, but $C_0$ times its length. We will refer to $\{E_N : N \geq 1\}$ as sets of Kakeya-type.

Specifically in this paper, we will be concerned with certain subsets of a curve, either on the sphere $\mathbb{S}^d$, or equivalently on a hyperplane at unit distance from the origin, that admit Kakeya-type sets.

Kakeya and Kakeya-type sets of zero measure have intrinsic structural properties that continually prove useful in an analytical setting. The most important of these properties is arguably the so-called stickiness property, originally observed by Wolff [25]. Roughly speaking, if a Kakeya-type set is a collection of many overlapping line segments, then stickiness dictates that the map which sends a direction to the line segment in the set with that direction is almost Lipschitz, with respect to suitably defined metrics. Another way of expressing this is that if the origins of two overlapping $\delta$-tubes are positioned close together, then the angle between these thickened line segments must be small, resulting in the intersection taking place far away from the respective bases. This idea, which has been formalized in several different ways in the literature [25], [14], [15], [13], will play a central role in our results, as we will discuss in Section 5.

Geometric and analytic properties of Kakeya and Kakeya-type sets are often studied using a suitably chosen maximal operator. Conversely, certain blow-up behavior for such operators typically follow from the existence of such sets. We introduce two such well-studied operators for which the existence of Kakeya-type sets implies unboundedness.

Given a set of directions $\Omega$, consider the directional maximal operator $D_\Omega$ defined by

$$D_\Omega f(x) := \sup_{\omega \in \Omega} \sup_{h > 0} \frac{1}{2h} \int_{-h}^{h} |f(x + \omega t)| dt, \quad (1.3)$$
where \( f : \mathbb{R}^{d+1} \to \mathbb{C} \) is a function that is locally integrable along lines. Also, for any locally integrable function \( f \) on \( \mathbb{R}^{d+1} \), consider the Kakeya-Nikodym maximal operator \( M_\Omega \) defined by

\[
M_\Omega f(x) := \sup_{\omega \in \Omega} \sup_{P \ni x} \frac{1}{|P|} \int_P |f(y)| dy,
\]

where the inner supremum is taken over all cylindrical tubes \( P \) containing the point \( x \), oriented in the direction \( \omega \). The tubes are taken to be of arbitrary length \( l \) and have circular cross-section of arbitrary radius \( r \), with \( r \leq l \). If \( \Omega \) is a set with nonempty interior, then due to the existence of Kakeya sets with \((d+1)\)-dimensional Lebesgue measure zero \([5]\), \( D_\Omega \) and \( M_\Omega \) are both unbounded as operators on \( L^p(\mathbb{R}^{d+1}) \) for all \( 1 \leq p < \infty \). More generally, if \( \Omega \) admits Kakeya-type sets, then these operators are unbounded on \( L^p(\mathbb{R}^{d+1}) \) for all \( 1 \leq p < \infty \) (see Section 1.2 below).

The complementary case when \( \Omega \) has empty interior has been studied extensively in the literature. It is easy to see that the operators in (1.3) and (1.4) exhibit a kind of monotonicity: if \( \Omega \subset \Omega' \), then \( D_\Omega f(x) \leq D_{\Omega'} f(x) \) and \( M_\Omega f(x) \leq M_{\Omega'} f(x) \), for any suitable function \( f \). Since these operators are unbounded when \( \Omega' = \) the unit sphere \( S^d \), treatment of the positive direction – identifying “small” sets of directions \( \Omega \) for which these operators are bounded on some \( L^p \) – has garnered much attention \([20, 6, 22, 21]\). These types of results rely on classical techniques in \( L^p \)-theory, such as square function estimates, Littlewood-Paley theory and almost-orthogonality principles.

For a general dimension \( d \geq 1 \), Nagel, Stein and Wainger \([20]\) showed that \( D_\Omega \) is bounded on all \( L^p(\mathbb{R}^{d+1}) \), \( 1 < p \leq \infty \), when \( \Omega = \{(v_1^d, \ldots, v_{d+1}^d) : i \geq 1\} \). Here 0 < \( a_1 < \cdots < a_{d+1} \) are fixed constants, and \( \{v_i : i \geq 1\} \) is a sequence obeying 0 < \( v_i+1 \leq \lambda v_i \) for some lacunary constant 0 < \( \lambda < 1 \). Carbery \([6]\) showed that \( D_\Omega \) is bounded on all \( L^p(\mathbb{R}^{d+1}) \), \( 1 < p \leq \infty \), in the special case when \( \Omega \) is the set given by the \((d+1)\)-fold Cartesian product of a geometric sequence, namely \( \Omega = \{(r^k_1, \ldots, r^k_{d+1}) : k_1, \ldots, k_{d+1} \in \mathbb{Z}^+\} \) for some 0 < \( r < 1 \). Very recently, Parcet and Rogers \([21]\) generalized an almost-orthogonality result of Alfonseca \([1]\) to extend the boundedness of \( D_\Omega \) on all \( L^p(\mathbb{R}^{d+1}) \), \( 1 < p \leq \infty \), for sets \( \Omega \) that are lacunary of finite order, defined in a suitable sense. Building upon previous work of Alfonseca, Soria, and Vargas \([2, 4, 3]\) and Nagel, Stein and Wainger \([20]\), the recent result of Parcet and Rogers \([21]\) recovers those of its predecessors.

Aside from this set of positive results with increasingly weak hypotheses, there has also been much development in the negative direction, pioneered by Bateman, Katz and Vargas \([24, 12, 4, 3]\). Of special significance to this article is the work of Bateman and Katz \([4]\), where the authors establish that \( D_\Omega \) is unbounded in \( L^p(\mathbb{R}^2) \) for all \( 1 \leq p < \infty \) if \( \Omega = \{(\cos \theta, \sin \theta) : \theta \in C_{1/3}\} \), where \( C_{1/3} \) is the Cantor middle-third set. A crowning success of the methodology of \([4]\) combined with the aforementioned work in the positive direction (in particular \([1]\)) is a result by Bateman \([3]\) that gives a complete characterization of the \( L^p \)-boundedness of \( D_\Omega \) and \( M_\Omega \) in the plane, while also describing all direction sets \( \Omega \) that admit planar sets of Kakeya-type. The distinctive feature of this latter body of work \([4, 3]\) dealing with the negative point of view is the construction
of counterexamples using a random mechanism that exploits the property of stickiness. We too adopt this approach to construct Kakeya-type sets in \( \mathbb{R}^{d+1} \), \( d \geq 2 \) consisting of tubes whose orientations lie along certain subsets of a curve on the hyperplane \( \{1\} \times \mathbb{R}^d \).

1.2 Results

As mentioned above, Bateman and Katz [4] establish the unboundedness of \( D_{\Omega} \) and \( M_{\Omega} \) on \( L^p(\mathbb{R}^2) \), for all \( p \in [1, \infty) \), when \( \Omega = \{(\cos \theta, \sin \theta) : \theta \in \mathcal{C}_{1/3}\} \) by constructing suitable Kakeya-type sets in the plane. In this paper, we extend their result to the general \((d + 1)\)-dimensional setting. To this end, we first describe what we mean by a Cantor set of directions in \((d + 1)\) dimensions.

Fix some integer \( M \geq 3 \). Construct an arbitrary Cantor-type subset of \([0, 1)\) as follows.

- Partition \([0, 1]\) into \( M \) subintervals of the form \([a, b]\), all of equal length \( M^{-1} \).
  - Among these \( M \) subintervals, choose any two that are not adjacent (i.e., do not share a common endpoint); define \( \mathcal{C}^{[1]}_M \) to be the union of these chosen subintervals, called first stage basic intervals.

- Partition each first stage basic interval into \( M \) further (second stage) subintervals of the form \([a, b]\), all of equal length \( M^{-2} \). Choose two non-adjacent second stage subintervals from each first stage basic one, and define \( \mathcal{C}^{[2]}_M \) to be the union of the four chosen second stage (basic) intervals.

- Repeat this procedure \textit{ad infinitum}, obtaining a nested, non-increasing sequence of sets. Denote the limiting set by \( \mathcal{C}_M \):

\[
\mathcal{C}_M = \bigcap_{k=1}^{\infty} \mathcal{C}^{[k]}_M.
\]

We call \( \mathcal{C}_M \) a \textit{generalized Cantor-type set} (with base \( M \)).

While conventional uniform Cantor sets, such as the Cantor middle-third set, are special cases of generalized Cantor-type sets, the latter may not in general look like the former. In particular, sets of the form \( \mathcal{C}_M \) need not be self-similar, although the actual sequential selection criterion leading up to their definition will be largely irrelevant for the content of this article. It is well-known (see [9] Chapter 4) that such sets have Hausdorff dimension at most \( \log 2 / \log M \). By choosing \( M \) large enough, we can thus construct generalized Cantor-type sets of arbitrarily small dimension.

In this paper, we prove the following.

\textbf{Theorem 1.2.} Let \( \mathcal{C}_M \subset [0, 1] \) be a generalized Cantor-type set described above. Let \( \gamma : [0, 1] \to \{1\} \times [-1, 1]^d \) be an injective map that satisfies a bi-Lipschitz condition

\[
\forall \ x, y, \ c|x - y| \leq |\gamma(x) - \gamma(y)| \leq C|x - y|,
\]

for some absolute constants \( 0 < c < 1 < C < \infty \). Set \( \Omega = \{\gamma(t) : t \in \mathcal{C}_M\} \). Then
(a) the set $\Omega$ admits Kakeya-type sets,

(b) the operators $D_\Omega$ and $M_\Omega$ are unbounded on $L^p(\mathbb{R}^{d+1})$ for all $1 \leq p < \infty$.

The condition in Theorem 1.2 that $\gamma$ satisfies a bi-Lipschitz condition can be weakened, but it will help in establishing some relevant geometry. Throughout this exposition, it is instructive to envision $\gamma$ as a smooth curve on the plane $x_1 = 1$, and we recommend the reader does this to aid in visualization. Our underlying direction set of interest $\Omega = \gamma(C_M)$ is essentially a Cantor-type subset of this curve.

The main focus of this article, for reasons explained below, is on (a), not on (b). Indeed, the implication (a) $\implies$ (b) is well-known in the literature; if $f = 1_{E_N}$, where $E_N$ is as in (1.2), then there exists a constant $c_0 = c_0(d,C_0) > 0$ such that

$$\min\{D_\Omega f(x), M_\Omega f(x)\} \geq c_0 \quad \text{for } x \in E_N^*(C_0). \quad (1.6)$$

This shows that

$$\min\{|\!|\!|D_\Omega|||_{p \to \infty}, |\!|\!|M_\Omega|||_{p \to \infty}\} \geq c_0 \left(\frac{|E_N^*(C_0)|}{|E_N|}\right)^{ \frac{1}{p} }, \quad \text{which } \to \infty \text{ if } 1 \leq p < \infty.$$

On the other hand, the condition (a) of Theorem 1.2 is not a priori strictly necessary in order to establish part (b) of the theorem. Suppose that $\{G_N : N \geq 1\}$ and $\{G_N : N \geq 1\}$ are two collections of sets with $|\tilde{G}_N|/|G_N| \to \infty$, enjoying the additional property that for any point $x \in \tilde{G}_N$, there exists a finite line segment originating at $x$ and pointing in a direction of $\Omega$, which spends at least a fixed positive proportion of its length in $G_N$. By an easy adaptation of the argument in (1.6), the sequence of test functions $f_N = 1_{G_N}$ would then prove the claimed unboundedness of $D_\Omega$. Kakeya-type sets, if they exist, furnish one such family of test functions with $G_N = E_N$ and $\tilde{G}_N = E_N^*$. In [21], Parcet and Rogers construct, for certain examples of direction sets, families of sets $G_N$ that supply a different class of test functions sufficient to prove unboundedness of the associated directional maximal operators. Similar constructions could in principle be applied in our situation as well to establish the unboundedness of directional maximal operators associated with our sets of interest. However, a set as constructed in [21] is typically a Cartesian product of a planar Kakeya-type set with a cube, and as such not of Kakeya-type according to Definition 1.1. In particular, it consists of rectangular parallelepipeds with possibly different sidelengths, with these sides not necessarily pointing in a direction from the underlying direction set $\Omega$, although there are line segments with orientation from $\Omega$ contained within them. Further, in contrast with Definition 1.1, $\tilde{G}_N$ need not be obtained by translating $G_N$ along its longest side.

The reason for considering Kakeya-type sets in this paper is twofold. First, they appear as natural generalizations of a classical feature of planar Kakeya set constructions, as explained in (1.1). Studying higher dimensional extensions of this phenomenon is of interest in its own right, and this article provides a concrete illustration of a sparse set of directions that gives rise to a similar phenomenon. Perhaps more importantly, we use the special direction sets in this paper as a device for introducing certain machinery
whose scope reaches beyond these examples. As discussed in Section 1.1, the problem of determining a characterization of direction sets $\Omega$ that give rise to $L^p$-bounded maximal operators $D_\Omega$ and $M_\Omega$ has garnered much attention. In [21], Parcet and Rogers obtain a positive result for such operators to be bounded on all Lebesgue spaces in general dimensions under certain hypotheses involving lacunarity, and conjecture that their condition is essentially sharp. The counterexamples in [21] mentioned above were furnished as supporting evidence for this claim. We address this conjecture in [16]. The property of admittance of Kakeya-type sets in the sense of Definition 1.1 turns out to be a critical feature of this study, and indeed equivalent to the unboundedness of directional maximal operators. In addition to the framework introduced in [4], the methods developed in the present article, specifically the investigation of root configurations and slope probabilities in Sections 7 and 8 are central to the analysis in [16]. While the consideration of general direction sets in [16] necessarily involves substantial technical adjustments, many of the main ideas of that analysis can be conveyed in the simpler setting of the Cantor example that we treat here. As such, we recommend that the reader approach the current paper as a natural first step in the process of understanding properties of direction sets that give rise to unbounded directional and Kakeya-Nikodym maximal operators on $L^p(\mathbb{R}^{d+1})$

Acknowledgements

The second author would like to thank Gordon Slade of the Department of Mathematics at the University of British Columbia for a helpful discussion on percolation theory. The research was partially supported by an NSERC Discovery Grant.

2 Overview of the proof of Theorem 1.2

2.1 Steps of the proof and layout

The basic structure of the proof is modeled on [4], with some important distinctions that we point out below. Our goal is to construct a family of tubes rooted on the hyperplane $\{0\} \times [0,1)^d$, the union of which will eventually give rise to the Kakeya-type set. The slopes of the constituent tubes will be assigned from $\Omega$ via a random mechanism involving stickiness akin to the one developed by Bateman and Katz [4]. The description of this random mechanism is in Section 6 with the required geometric and probabilistic background collected en route in Sections 3, 4 and 5. The essential elements of the construction, barring the details of the slope assignment, have been laid out in Section 2.2 below. The main estimates leading to the proof of Theorem 1.2 are (2.5) and (2.6) in Proposition 2.1 in this section. Of these the first, a precise version of which is available in Proposition 6.4, provides a lower bound of $a_N = \sqrt{\log N}/N$ on the size of the part of the tubes lying near the root hyperplane. The second inequality, also quantified in Proposition 6.4, yields an upper bound of $b_N = 1/N$ for the portion away from it. The disparity in the relative sizes of these two parts is the desired conclusion of Theorem 1.2
The language of trees was a key element in the random construction of [3, 4]. We continue to adopt this language, introducing the relevant definitions in Section 4 and providing some detail on the connection between the geometry of Ω and a tree encoding it. Specifically, the notion of Bernoulli percolation on trees plays an important role in the proof of (2.6) with \( b_N = \frac{1}{N} \), as it did in the two-dimensional setting. The higher-dimensional structure of Ω does however result in minor changes to the argument, and the general percolation-theoretic facts necessary for handling (2.6) have been compiled in Section 5. Other probabilistic estimates specific to the random mechanism of Section 6 and central to the derivation of (2.5) are separately treated in Section 7. The proof is completed in Sections 8 and 9.

Of the two estimates (2.5) and (2.6) necessary for the Kakeya-type construction, the first is the most significant contribution of this paper. A deterministic analogue of (2.5) was used in [3, 4], where a similar lower bound for the size of the Kakeya-type set was obtained for every slope assignment \( \sigma \) in a certain measure space. The counting argument that led to this bound fails to produce the necessary estimate in higher dimensions and is replaced here by a probabilistic statement that suffices for our purposes. More precisely, the issue is the following. A large lower bound on a union of tubes follows if they do not have significant pairwise overlap among themselves; i.e. if the total size of pairwise intersections is small. In dimension two, a good upper bound on the size of this intersection was available uniformly in every sticky slope assignment. Although the argument that provided this bound is not transferable to general dimensions, it is still possible to obtain the desired bound with large probability. A probabilistic statement similar to but not as strong as (2.5) can be derived relatively easily via an estimate on the first moment of the total size of random pairwise intersections. Unfortunately, this is still not sharp enough to yield the disparity in the sizes of the tubes and their translated counterparts necessary to claim the existence of a Kakeya-type set. To strengthen the bound, we need a second moment estimate on the pairwise intersections. Both moment estimates share some common features; for instance,

- Euclidean distance relations between roots and slopes of two intersecting tubes,
- interplay of the above with the relative positions of the roots and slopes within the respective trees that they live in, which affects the slope assignments.

However, the technicalities are far greater for the second moment compared to the first. In particular, for the second moment we are naturally led to consider not just pairs, but triples and quadruples of tubes, and need to evaluate the probability of obtaining pairwise intersections among these. Not surprisingly, this probability depends on the structure of the root tuple within its ambient tree. It is the classification of these root configurations, computation of the relevant probabilities and their subsequent application to the estimation of expected intersections that we wish to highlight as the main contributions of this article.
2.2 Construction of a Kakeya-type set

We now choose some integer $M \geq 3$ and a generalized Cantor-type set $C_M \subseteq [0,1)$ as described in Section 1.2 and fix these items for the remainder of the article. We also fix an injective map $\gamma : [0,1] \to \{1\} \times [-1,1]^d$ satisfying the bi-Lipschitz condition in (1.3). These objects then define a fixed set of directions $\Omega = \{\gamma(t) : t \in C_M\} \subseteq \{1\} \times [-1,1]^d$.

Next, we define the thin tube-like objects that will comprise our Kakeya-type set. Fix an arbitrarily large integer $N \geq 1$, typically much bigger than $M$. Let $\{Q_t : t \in \mathbb{T}_N\}$, parametrized by the index set $\mathbb{T}_N$, be the collection of disjoint $d$-dimensional cubes of sidelength $M^{-N}$ generated by the lattice $M^{-N}\mathbb{Z}^d$ in the set $\{0\} \times [0,1)^d$. More specifically, each $Q_t$ is of the form

$$Q_t = \{0\} \times \prod_{i=1}^d \left[ \frac{j_i}{M^N}, \frac{j_i+1}{M^N} \right),$$

for some $j = (j_1, \ldots, j_d) \in \{0, 1, \ldots, M^N - 1\}^d$, so that $\#(\mathbb{T}_N) = M^{Nd}$. For technical reasons, we also define $\tilde{Q}_t$ to be the $\kappa_d$-dilation of $Q_t$ about its center point, where $\kappa_d$ is a small, positive, dimension-dependent constant. The reason for this technicality, as well as possible values of $\kappa_d$, will soon emerge in the sequel, but for concreteness choosing $\kappa_d = d^{-d}$ will suffice.

Recall that the $N$th iterate $C_M^{[N]}$ of the Cantor construction is the union of $2^N$ disjoint intervals each of length $M^{-N}$. We choose a representative element of $C_M$ from each of these intervals, calling the resulting finite collection $D_M^{[N]}$. Clearly $\text{dist}(x, D_M^{[N]}) \leq M^{-N}$ for every $x \in C_M$. Set

$$\Omega_N := \gamma(D_M^{[N]}),$$

so that $\text{dist}(\omega, \Omega_N) \leq CM^{-N}$ for any $\omega \in \Omega$, with $C$ as in (1.5).

For any $t \in \mathbb{T}_N$ and any $\omega \in \Omega_N$, we define

$$\mathcal{P}_{t,\omega} := \left\{ r + s\omega : r \in \tilde{Q}_t, \ 0 \leq s \leq 10C_0 \right\},$$

where $C_0$ is a large constant to be determined shortly (for instance, $C_0 = d^d c^{-1}$ will work, with $c$ as in (1.3)). Thus the set $\mathcal{P}_{t,\omega}$ is a cylinder oriented along $\omega$. Its (vertical) cross-section in the plane $x_1 = 0$ is the cube $\tilde{Q}_t$. We say that $\mathcal{P}_{t,\omega}$ is rooted at $Q_t$. While $\mathcal{P}_{t,\omega}$ is not strictly speaking a tube as defined in the introduction, the distinction is negligible. In particular, $\mathcal{P}_{t,\omega}$ contains and is contained in constant multiples of $\delta$-tubes with $\delta = \kappa_d \cdot M^{-N}$. By a slight abuse of terminology but no loss of generality, we will henceforth refer to $\mathcal{P}_{t,\omega}$ as a tube.

If a slope assignment $\sigma : \mathbb{T}_N \to \Omega_N$ has been specified, we set $P_{t,\sigma} := \mathcal{P}_{t,\sigma(t)}$. Thus $\{P_{t,\sigma} : t \in \mathbb{T}_N\}$ is a family of tubes rooted at the elements of an $M^{-N}$-fine grid in $\{0\} \times [0,1)^d$, with essentially uniform length in $t$ that is bounded above and below by fixed absolute constants. Two such tubes are illustrated in Figure 1. For the remainder, we set

$$K_N(\sigma) := \bigcup_{t \in \mathbb{T}_N} P_{t,\sigma}. $$

(2.4)
Figure 1: Two typical tubes $P_{t_1,\sigma}$ and $P_{t_2,\sigma}$ rooted respectively at $t_1$ and $t_2$ in the $\{x_1 = 0\}$–coordinate plane.

For a certain choice of slope assignment $\sigma$, this collection of tubes will be shown to generate a Kakeya-type set in the sense of Definition 1.1. This particular slope assignment will not be explicitly described, but rather inferred from the contents of the following proposition.

**Proposition 2.1.** For any $N \geq 1$, let $\Sigma_N$ be a finite collection of slope assignments from the lattice $\mathbb{T}_N$ to the direction set $\Omega_N$. Every $\sigma \in \Sigma_N$ generates a set $K_N(\sigma)$ as defined in (2.4). Denote the power set of $\Sigma_N$ by $\mathcal{P}(\Sigma_N)$.

Suppose that $(\Sigma_N, \mathcal{P}(\Sigma_N), \Pr)$ is a discrete probability space equipped with the probability measure $\Pr$, for which the random sets $K_N(\sigma)$ obey the following estimates:

\[
\Pr\left(\sigma : |K_N(\sigma) \cap [0,1] \times \mathbb{R}^d| \geq a_N \right) \geq \frac{3}{4}, \quad (2.5)
\]

and

\[
\mathbb{E}_\sigma|K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d| \leq b_N, \quad (2.6)
\]

where $C_0 \geq 1$ is a fixed constant, and $\{a_N\}, \{b_N\}$ are deterministic sequences satisfying

\[
\frac{a_N}{b_N} \to \infty, \quad \text{as} \quad N \to \infty.
\]

Then $\Omega$ admits Kakeya-type sets.

**Proof.** Fix any integer $N \geq 1$. Applying Markov’s Inequality to (2.6), we see that

\[
\Pr\left(\sigma : |K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d| \geq 4b_N \right) \leq \frac{\mathbb{E}_\sigma|K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d|}{4b_N} \leq \frac{1}{4},
\]

10
The inequality (3.1) follows since \( |x| \) exist \( x \). The proof is described in the diagram below. If \( \sigma \) defined as in (2.3). If there exists \( p \), then

\[
E \quad \text{are not limited to our specific arrangement or description of tubes.}
\]

within the proof of Theorem 1.2. Nonetheless, they are really general observations that two intersecting tubes in Euclidean space. These facts will be used in several instances in this section, we will take the opportunity to establish some geometric facts about

\[
3 \quad \text{Geometric Facts}
\]

return to a more concrete formulation of the required estimates in Proposition 6.4.

Combining this estimate with (2.5), we find that

\[
\begin{align*}
\Pr \left( \{ \sigma : |K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d| \leq 4b_N \} \right) & \geq \frac{3}{4}.
\end{align*}
\]

(2.7)

We may therefore choose a particular \( \sigma \in \Sigma_N \) for which the size estimates on \( K_N(\sigma) \) given by (2.5) and (2.7) hold simultaneously. Set

\[
E_N := K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d,
\]

so that \( E_N(2C_0 + 1) \supseteq K_N(\sigma) \cap [0, 1] \times \mathbb{R}^d \).

Then \( E_N \) is a union of \( \delta \)-tubes oriented along directions in \( \Omega_N \subset \Omega \) for which

\[
\frac{|E_N(2C_0 + 1)|}{|E_N|} \geq \frac{a_N}{4b_N} \rightarrow \infty, \quad \text{as} \quad N \rightarrow \infty,
\]

by hypothesis. This shows that \( \Omega \) admits Kakeya-type sets, per condition (1.2).

Proposition 2.1 proves part (a) of our Theorem 1.2. The implication \( (a) \implies (b) \) has already been discussed in Section 1.2. The remainder of this paper is devoted to establishing a proper randomization over slope assignments \( \Sigma_N \) that will then allow us to verify the hypotheses of Proposition 2.1 for suitable sequences \( \{a_N\} \) and \( \{b_N\} \). We return to a more concrete formulation of the required estimates in Proposition 6.4.

3 Geometric Facts

In this section, we will take the opportunity to establish some geometric facts about two intersecting tubes in Euclidean space. These facts will be used in several instances within the proof of Theorem 1.2. Nonetheless, they are really general observations that are not limited to our specific arrangement or description of tubes.

Lemma 3.1. For \( v_1, v_2 \in \Omega_N \) and \( t_1, t_2 \in T_N, t_1 \neq t_2 \), let \( \mathcal{P}_{t_1,v_1} \) and \( \mathcal{P}_{t_2,v_2} \) be the tubes defined as in (2.3). If there exists \( p = (p_1, \cdots, p_{d+1}) \in \mathcal{P}_{t_1,v_1} \cap \mathcal{P}_{t_2,v_2} \), then the inequality

\[
|\text{cen}(Q_{t_2}) - \text{cen}(Q_{t_1}) + p_1(v_2 - v_1)| \leq 2\kappa_d \sqrt{d}M^{-N},
\]

(3.1)

holds, where \( \text{cen}(Q) \) denotes the centre of the cube \( Q \).

Proof. The proof is described in the diagram below. If \( p \in \mathcal{P}_{t_1,v_1} \cap \mathcal{P}_{t_2,v_2} \), then there exist \( x_1 \in \hat{Q}_{t_1}, x_2 \in \hat{Q}_{t_2} \) such that \( p = x_1 + p_1 v_1 = x_2 + p_1 v_2 \), i.e., \( p_1(v_2 - v_1) = x_1 - x_2 \). The inequality (3.1) follows since \( |x_i - \text{cen}(Q_{t_i})| \leq \kappa_d \sqrt{d}M^{-N} \) for \( i = 1, 2 \).
Figure 2: A simple triangle is defined by two rooted tubes, $\mathcal{P}_{t_1,v_1}$ and $\mathcal{P}_{t_2,v_2}$, and any point $p$ in their intersection.

The inequality in (3.1) provides a valuable tool whenever an intersection takes place. For the reader who would like to look ahead, the Lemma 3.1 will be used along with Corollary 3.2 to establish Lemma 8.5. The following Corollary 3.3 will be needed for the proofs of Lemmas 8.6 and 8.10.

**Corollary 3.2.** Under the hypotheses of Lemma 3.1 and for $\kappa_d > 0$ suitably small,

$$|p_1(v_2 - v_1)| \geq \kappa_d M^{-N}.$$  \hspace{1cm} (3.2)

**Proof.** Since $t_1 \neq t_2$, we must have $|\text{cen}(Q_{t_1}) - \text{cen}(Q_{t_2})| \geq M^{-N}$. Thus an intersection is possible only if

$$p_1|v_2 - v_1| \geq |\text{cen}(Q_{t_2}) - \text{cen}(Q_{t_1})| - 2\kappa_d \sqrt{d} M^{-N} \geq (1 - 2\kappa_d \sqrt{d}) M^{-N} \geq \kappa_d M^{-N},$$

where the first inequality follows from (3.1) and the last inequality holds for an appropriate selection of $\kappa_d$. \hfill \Box

**Corollary 3.3.** If $t_1 \in \mathbb{T}_N$, $v_1, v_2 \in \Omega_N$ and a cube $Q \subseteq \mathbb{R}^{d+1}$ of sidelength $C_1 M^{-N}$ with sides parallel to the coordinate axes are given, then there exists at most $C_2 = C_2(C_1)$ choices of $t_2 \in \mathbb{T}_N$ such that $\mathcal{P}_{t_1,v_1} \cap \mathcal{P}_{t_2,v_2} \cap Q \neq \emptyset$.

**Proof.** As $p = (p_1, \cdots, p_{d+1})$ ranges in $Q$, $p_1$ ranges over an interval $I$ of length $C_1 M^{-N}$. If $p \in \mathcal{P}_{t_1,v_1} \cap \mathcal{P}_{t_2,v_2} \cap Q$, the inequality (3.1) and the fact $\text{diam}(\Omega) \leq \text{diam}([1] \times [-1,1]^d) = 2\sqrt{d}$ implies

$$|\text{cen}(Q_{t_2}) - \text{cen}(Q_{t_1}) + \text{cen}(I)(v_2 - v_1)| \leq |(p_1 - \text{cen}(I))(v_2 - v_1)| + 2\kappa_d \sqrt{d} M^{-N}$$
\[
= 2\sqrt{d}(C_1 + \kappa_d)M^{-N},
\]
restricting \(\text{cen}(Q_t)\) to lie in a cube of sidelength \(2\sqrt{d}(C_1 + \kappa_d)M^{-N}\) centred at \(\text{cen}(Q_t) - \text{cen}(I)(v_2 - v_1)\). Such a cube contains at most \(C_2\) sub-cubes of the form (2.1), and the result follows.

A recurring theme in the proof of Theorem 1.2 is the identification of a criterion that ensures that a specified point lies in the Kakeya-type set \(K_N(\sigma)\) defined in (2.4). With this in mind, we introduce for any \(p = (p_1, p_2, \ldots, p_{d+1}) \in [0, 10C_0] \times \mathbb{R}^d\) a set

\[
\text{Poss}(p) := \{Q_t : t \in T_N, \text{ there exists } v \in \Omega_N \text{ such that } p \in P_{t,v}\}. \quad (3.3)
\]
This set captures all the possible \(M^{-N}\)-cubes of the form (2.1) in \([0, 10C_0] \times \mathbb{R}^d\) such that a tube rooted at one of these cubes has the potential to contain \(p\), provided it is given the correct orientation. Note that \(\text{Poss}(p)\) is independent of any slope assignment \(\sigma\).

Depending on the location of \(p\), \(\text{Poss}(p)\) could be empty. This would be the case if \(p\) lies outside a large enough compact subset of \([0, 10C_0] \times \mathbb{R}^d\), for example. Even if \(\text{Poss}(p)\) is not empty, an arbitrary slope assignment \(\sigma\) may not endow any \(Q_t\) in \(\text{Poss}(p)\) with the correct orientation.

In the next lemma, we list a few easy properties of \(\text{Poss}(p)\) that will be helpful later, particularly during the proof of Lemma 9.3. Lemma 3.4 establishes the main intuition behind the \(\text{Poss}(p)\) set, as we give a more geometric description of \(\text{Poss}(p)\) in terms of an affine copy of the direction set \(\Omega_N\). This is illustrated in Figure 3 for a particular choice of directions \(\Omega_N\).

**Lemma 3.4.** (a) For any slope assignment \(\sigma\),

\[
\{Q_t : t \in T_N, p \in P_{t,\sigma}\} \subseteq \text{Poss}(p).
\]

(b) For any \(p \in [0, 10C_0] \times \mathbb{R}^d\),

\[
\text{Poss}(p) = \{Q_t : \tilde{Q}_t \cap (p - p_1\Omega_N) \neq \emptyset\} \quad (3.4)
\]

\[
\subseteq \{Q_t : t \in T_N, Q_t \cap (p - p_1\Omega_N) \neq \emptyset\}. \quad (3.5)
\]

Note that the set in (3.4) could be empty, but the one in (3.5) is not.

**Proof.** If \(p \in P_{t,\sigma}\), then \(p \in P_{t,\sigma(t)}\) with \(\sigma(t)\) equal to some \(v \in \Omega\). Thus \(P_{t,v}\) contains \(p\) and hence \(Q_t \in \text{Poss}(p)\), proving part (a). For part (b), we observe that \(p \in P_{t,v}\) for some \(v \in \Omega_N\) if and only if \(p - p_1v \in \tilde{Q}_t\), i.e., \(\tilde{Q}_t \cap (p - p_1\Omega_N) \neq \emptyset\). This proves the relation (3.5). The containment in (3.4) is obvious.

We will also need a bound on the cardinality of \(\text{Poss}(p)\) within a given cube, and on the cardinality of possible slopes that give rise to indistinguishable tubes passing through a given point. We now prescribe these. Lemmas 3.5 and 3.6 are not technically needed for the remainder, but can be viewed as steps toward establishing Lemma 3.7 which will prove critical throughout Section 9. Not surprisingly, the Cantor-like construction of \(\Omega\) plays a role in all these estimates.
Figure 3: Figure (a) depicts the cone generated by a second stage Cantor construction, $\Omega_2$, on the set of directions given by the curve $\{(1, t, t^2) : 0 \leq t \leq C\}$ in the $\{1\} \times \mathbb{R}^2$ plane. In Figure (b), a point $p = (p_1, p_2, p_3)$ has been fixed and the cone of directions has been projected backward from $p$ onto the coordinate plane, $p - p_1\Omega_2$. The resulting $\text{Poss}(p)$ set is thus given by all cubes $Q_t$, $t \in T_N$ such that $\tilde{Q}_t$ intersects a subset of the curve $\{(0, p_2 - p_1 t, p_3 - p_1 t^2) : 0 \leq t \leq C\}$.

Lemma 3.5. Given $C_0, C_1 > 0$, there exists $C_2 = C_2(C_0, C_1, M, d) > 0$ with the following property. Let $p = (p_1, \cdots, p_{d+1}) \in (0, 10C_0] \times \mathbb{R}^d$, and $Q$ be any cube in $\{0\} \times [0, 1]^d$ with sidelength in $[M^{-\ell}, M^{-\ell+1})$ for some $\ell \leq N - 1$. Then

$$\#\{Q_t : t \in T_N, Q_t \cap Q \neq \emptyset, \text{dist}(Q_t, p - p_1\Omega_N) \leq C_1M^{-N}\} \leq C_22^{N-\ell}. \quad (3.6)$$

Proof. Let $j \in \mathbb{Z}$ be the index such that $M^{-j} \leq p_1 < M^{-j+1}$. By scaling, the left hand side of (3.6) is comparable to (i.e., bounded above and below by constant multiples of) the number of $p_1^{-1}M^{-N}$-separated points lying in

$$Q' := \{x \in p_1^{-1}Q : \text{dist}(x, p_1^{-1}p - \Omega_N) \leq C_1p_1^{-1}M^{-N}\}.$$

But $p_1^{-1}p - \Omega_N = (1, c) - \Omega_N$ is an image of $\Omega_N$ following an inversion and translation. This implies that there is a subset $\Omega'_N$ of $\Omega_N$, depending on $p$ and $p_1^{-1}Q$ and with diameter $O(M^{j-\ell})$, such that $Q'$ is contained in a $O(M^{j-N})$-neighborhood of $-\Omega'_N + (1, c)$.

The number of $M^{j-N}$-separated points in $Q'$ is comparable to that in $\Omega'_N$. 

14
Suppose first that \( j \leq \ell \). If \( C' \subseteq C_M^{[N]} \) is defined by the requirement \( \Omega_N' = \gamma(C') \), then (1.5) implies that \( \text{diam}(C') = O(M^{j-\ell}) \). Thus \( C' \) is contained in at most \( O(1) \) intervals of length \( M^{j-\ell} \) chosen at step \( \ell \) in the Cantor-type construction. Each chosen interval at the \( k \)th stage gives rise to two chosen subintervals at the next stage, with their centres being separated by at least \( M^{-k-1} \). So the number of \( M^{j-N} \)-separated points in \( C' \), and hence \( \gamma(C') \) is \( O(2(N-j)-(\ell-j)) = O(2^{N-\ell}) \) as claimed. The case \( j \geq \ell \) is even simpler, since the number of \( M^{j-N} \)-separated points in \( C' \) is trivially bounded by \( 2^{N-j} \leq 2^{N-\ell} \).

Lemma 3.6. Fix \( t \in \mathbb{T}_N \) and \( p = (p_1, \cdots, p_{d+1}) \in [M^{-\ell},M^{-\ell+1}] \times \mathbb{R}^d \), for some \( 0 \leq \ell \ll N \). Let \( Q \) be a cube centred at \( p \) of sidelength \( C_1 M^{-N} \). Then

\[
\# \{ v \in \Omega_N : Q \cap P_{t,v} \neq \emptyset \} \leq C_2 2^{\ell}.
\]

**Proof.** If both \( P_{t,v} \) and \( P_{t,v'} \) have nonempty intersection with \( Q \), then there exist \( q = (q_1, \cdots, q_{d+1}), q' = (q'_1, \cdots, q'_{d+1}) \in Q \) such that both \( q - q_1 v \) and \( q' - q'_1 v' \) land in \( Q_t \). Thus,

\[
p_1 |v - v'| \leq |(q - p_1 v) - (q' - p_1 v')| + |q - q'| \\
\leq |(q - q_1 v) - (q' - q'_1 v')| + |q_1 - p_1||v| + |q'_1 - p_1||v'| + |q - q'| \\
\leq (\kappa_d \sqrt{d} + 10C_1 \sqrt{d}) M^{-N}.
\]

In other words, \( |v - v'| \leq (10C_1 + \kappa_d) \sqrt{d} M^{\ell-N} \). Recalling that \( v = \gamma(\alpha) \) and \( v' = \gamma(\alpha') \) for some \( \alpha, \alpha' \in D^{[N]}_M \), combining the last inequality with (1.5) implies that \( |\alpha - \alpha'| \leq C_2 M^{\ell-N} \). Thus there is a collection of at most \( O(1) \) chosen intervals at step \( N - \ell \) of the Cantor-type construction which \( \alpha \) (and hence \( \alpha' \)) can belong to. Since each interval gives rise to two chosen intervals at the next stage, the number of possible \( \alpha \) and hence \( v \) is \( O(2^\ell) \).

A slight modification of the proof above yields a stronger conclusion, stated below, when \( p \) is far away from the root hyperplane. We will return to this result several times in the sequel (see for example Lemma 6.6 for a version of it in the language of trees), and make explicit use of it in Section 9, specifically in the proofs of Lemmas 9.1 and 9.2.

**Lemma 3.7.** There exists a constant \( C_0 \geq 1 \) with the following properties.

(a) For any \( p \in [C_0, C_0 + 1] \times \mathbb{R}^d \) and \( t \in \mathbb{T}_N \), there exists at most one \( v \in \Omega_N \) such that \( p \in P_{t,v} \). In other words, for every \( Q_t \) in \( \text{Poss}(p) \), there is exactly one \( \delta \)-tube rooted at \( t \) that contains \( p \).

(b) For any \( p \) as in (a), and \( Q_t, Q_{t'} \in \text{Poss}(p) \), let \( v = \gamma(\alpha), v' = \gamma(\alpha') \) be the two unique slopes in \( \Omega_N \) guaranteed by (a) such that \( p \in P_{t,v} \cap P_{t',v'} \). If \( k \) is the largest integer such that \( Q_t \) and \( Q_{t'} \) are both contained in the same cube \( Q \subseteq \{0\} \times [0,1]^d \) of sidelength \( M^{-k} \) whose corners lie in \( M^{-k} \mathbb{Z}^d \), then \( \alpha \) and \( \alpha' \) belong to the same \( k \)th stage basic interval in the Cantor construction.
Proof. (a) Suppose \( v, v' \in \Omega_N \) are such that \( p \in \mathcal{P}_{t,v} \cap \mathcal{P}_{t,v'} \). Then \( p - p_1 v \) and \( p - p_1 v' \) both lie in \( \bar{Q}_t \), so that \( p_1 |v - v'| \leq \kappa d \sqrt{d} M^{-N} \). Since \( p_1 \geq C_0 \) and \((1.5)\) holds, we find that
\[
|\alpha - \alpha'| \leq \frac{\kappa d \sqrt{d}}{c C_0} M^{-N} < M^{-N},
\]
where the last inequality holds if \( C_0 \) is chosen large enough. Let us recall from the description of the Cantor-like construction in Section \((1.2)\) that any two basic \( r \)th stage intervals are non-adjacent, and hence any two points in \( C_M \) lying in distinct basic \( r \)th stage intervals are separated by at least \( M^{-r} \). Therefore the inequality above implies that both \( \alpha \) and \( \alpha' \) belong to the same basic \( N \)th stage interval in \( C^{[N]}_M \). But \( D^{[N]}_M \) contains exactly one element from each such interval. So \( \alpha = \alpha' \) and hence \( v = v' \).

(b) If \( p \in \mathcal{P}_{t,v} \cap \mathcal{P}_{t,v'} \), then \( p_1 |v - v'| \leq \text{diam}(\bar{Q}_t \cup \bar{Q}_v) \leq \text{diam}(Q) = \sqrt{d} M^{-k} \). Applying \((1.3)\) again combined with \( p_1 \geq C_0 \), we find that \( |\alpha - \alpha'| \leq \frac{\sqrt{d}}{c C_0} M^{-k} < M^{-k} \), for \( C_0 \) chosen large enough. By the same property of the Cantor construction as used in \((\alpha)\), we obtain that \( \alpha \) and \( \alpha' \) lie in the same \( k \)th stage basic interval in \( C^{[k]}_M \).

\( \Box \)

4 Rooted, labelled trees

4.1 The terminology of trees

An undirected graph \( G := (V, E) \) is a pair, where \( V \) is a set of vertices and \( E \) is a symmetric, nonreflexive subset of \( V \times V \), called the edge set. By symmetric, here we mean that the pair \((u, v) \in E\) is unordered; i.e. the pair \((u, v)\) is identical to the pair \((v, u)\). By nonreflexive, we mean \( E \) does not contain the pair \((v, v)\) for any \( v \in V \).

A path in a graph is a sequence of vertices such that each successive pair of vertices is a distinct edge in the graph. A finite path (with at least one edge) whose first and last vertices are the same is called a cycle. A graph is connected if for each pair of vertices \( v \neq u \), there is a path in \( G \) containing \( v \) and \( u \). We define a tree to be a connected undirected graph with no cycles.

All our trees will be of a specific structure. A rooted, labelled tree \( T \) is one whose vertex set is a nonempty collection of finite sequences of nonnegative integers such that if \( \langle i_1, \ldots, i_n \rangle \in T \), then

(i.) for any \( k, 0 \leq k \leq n \), \( \langle i_1, \ldots, i_k \rangle \in T \), where \( k = 0 \) corresponds to the empty sequence, and

(ii.) for every \( j \in \{0, 1, \ldots, i_n\} \), we have \( \langle i_1, \ldots, i_{n-1}, j \rangle \in T \).

We say that \( \langle i_1, \ldots, i_{n-1} \rangle \) is the parent of \( \langle i_1, \ldots, i_{n-1}, j \rangle \) and that \( \langle i_1, \ldots, i_{n-1}, j \rangle \) is the \( (j + 1) \)th child of \( \langle i_1, \ldots, i_{n-1} \rangle \). If \( u \) and \( v \) are two sequences in \( T \) such that \( u \) is a child of \( v \), or a child’s child of \( v \), or a child’s child’s child of \( v \), etc., then we say that \( u \)
is a descendant of \( v \) (or that \( v \) is an ancestor of \( u \)), and we write \( u \subset v \) (see the remark below). If \( u = (i_1, \ldots, i_m) \in \mathcal{T} \), \( v = (j_1, \ldots, j_n) \in \mathcal{T} \), \( m \leq n \), and neither \( u \) nor \( v \) is a descendant of the other, then the youngest common ancestor of \( u \) and \( v \) is the vertex in \( \mathcal{T} \) defined by

\[
D(u, v) = D(v, u) := \begin{cases} 
\emptyset, & \text{if } i_1 \neq j_1 \\
\langle i_1, \ldots, i_k \rangle & \text{if } k = \max\{l : i_l = j_l\}. 
\end{cases} \tag{4.1}
\]

One can similarly define the youngest common ancestor for any finite collection of vertices.

**Remark:** At first glance, using the notation \( u \subset v \) to denote when \( u \) is a descendant of \( v \) may seem counterintuitive, since \( u \) is a descendant of \( v \) precisely when \( v \) is a subsequence of \( u \). However, we will soon be identifying vertices of rooted labelled trees with certain nested families of cubes in \( \mathbb{R}^d \). Consequently, as will become apparent in the next two subsections, \( u \) will be a descendant of \( v \) precisely when the cube associated with \( u \) is contained within the cube associated with \( v \).

We designate the empty sequence \( \emptyset \) as the root of the tree \( \mathcal{T} \). The sequence \( \langle i_1, \ldots, i_n \rangle \) should be thought of as the vertex in \( \mathcal{T} \) that is the \((i_n + 1)\)th child of the \((i_{n-1} + 1)\)th child, \ldots., of the \((i_1 + 1)\)th child of the root. All unordered pairs of the form \( \langle\langle i_1, \ldots, i_{n-1}\rangle, \langle i_1, \ldots, i_{n-1}, i_n\rangle\rangle \) describe the edges of the tree \( \mathcal{T} \). We say that the edge originates at the vertex \( \langle i_1, \ldots, i_{n-1} \rangle \) and that it terminates at the vertex \( \langle i_1, \ldots, i_{n-1}, i_n \rangle \). Note that every vertex in the tree that is not the root is uniquely identified by the edge terminating at that vertex. Consequently, given an edge \( e \in \mathcal{E} \), we define \( v(e) \) to be the vertex in \( \mathcal{V} \) at which \( e \) terminates. The vertex \( \langle i_1, \ldots, i_n \rangle \in \mathcal{T} \) also prescribes a unique path, or my, from the root to this vertex:

\[
\emptyset \rightarrow \langle i_1 \rangle \rightarrow \langle i_1, i_2 \rangle \rightarrow \cdots \rightarrow \langle i_1, i_2, \ldots, i_n \rangle.
\]

We let \( \partial \mathcal{T} \) denote the collection of all rays in \( \mathcal{T} \) of maximal (possibly infinite) length. For a fixed vertex \( v = \langle i_1, \ldots, i_m \rangle \in \mathcal{T} \), we also define the subtree (of \( \mathcal{T} \)) generated by the vertex \( v \) to be the maximal subtree of \( \mathcal{T} \) with \( v \) as the root; i.e. it is the subtree

\[
\{\langle i_1, \ldots, i_m, j_1, \ldots, j_k \rangle \in \mathcal{T} : k \geq 0\}.
\]

The height of the tree is taken to be the supremum of the lengths of all the sequences in the tree. Further, we define the height \( h(\cdot) \), or level, of a vertex \( \langle i_1, \ldots, i_n \rangle \) in the tree to be \( n \), the length of its identifying sequence. All vertices of height \( n \) are said to be members of the \( n \)th generation of the root, or interchangeably, of the tree. More explicitly, a member vertex of the \( n \)th generation has exactly \( n \) edges joining it to the root. The height of the root is always taken to be zero.

If \( \mathcal{T} \) is a tree and \( n \in \mathbb{Z}^+ \), we write the truncation of \( \mathcal{T} \) to its first \( n \) levels as \( \mathcal{T}_n = \{\langle i_1, \ldots, i_k \rangle \in \mathcal{T} : 0 \leq k \leq n\} \). This subtree is a tree of height at most \( n \). A tree is called locally finite if its truncation to every level is finite; i.e. consists of finitely
many vertices. All of our trees will have this property. In the remainder of this article, when we speak of a tree we will always mean a locally finite, rooted labelled tree, unless otherwise specified.

Roughly speaking, two trees are isomorphic if they have the same collection of rays. To make this precise we define a special kind of map between trees that will turn out to be very important for us later.

**Definition 4.1.** Let $T_1$ and $T_2$ be two trees with equal (possibly infinite) heights. Let $\sigma: T_1 \to T_2$; we call $\sigma$ sticky if

- for all $v \in T_1$, $h(v) = h(\sigma(v))$, and
- $u \subset v$ implies $\sigma(u) \subset \sigma(v)$ for all $u, v \in T_1$.

We often say that $\sigma$ is sticky if it preserves heights and lineages.

A one-to-one and onto sticky map between two trees, whose inverse is then automatically sticky, is an isomorphism and the two trees are said to be isomorphic; we will write $T_1 \cong T_2$. Two isomorphic trees can be treated as essentially identical objects.

4.2 Encoding bounded subsets of the unit interval by trees

The language of rooted labelled trees is especially convenient for representing bounded sets in Euclidean spaces. This connection is well-studied in the literature. We refer the interested reader to [19] for more information.

We start with $[0, 1) \subset \mathbb{R}$. Fix any positive integer $M \geq 2$. We define an $M$-adic rational as a number of the form $i/M^k$ for some $i \in \mathbb{Z}, k \in \mathbb{Z}^+,$ and an $M$-adic interval as $[i \cdot M^{-k}, (i+1) \cdot M^{-k})$. For any nonnegative integer $i$ and positive integer $k$ such that $i < M^k$, there exists a unique representation

$$i = i_1 M^{k-1} + i_2 M^{k-2} + \cdots + i_{k-1} M + i_k,$$

(4.2)

where the integers $i_1, \ldots, i_k$ take values in $\mathbb{Z}_M := \{0, 1, \ldots, M - 1\}$. These integers should be thought of as the “digits” of $i$ with respect to its base $M$ expansion. An easy consequence of (4.2) is that there is a one-to-one and onto correspondence between $M$-adic rationals in $[0, 1)$ of the form $i/M^k$ and finite integer sequences $\langle i_1, \ldots, i_k \rangle$ of length $k$ with $i_j \in \mathbb{Z}_M$ for each $j$. Naturally then, we define the tree of infinite height

$$T([0, 1); M) = \{ \langle i_1, \ldots, i_k \rangle : k \geq 0, i_j \in \mathbb{Z}_M \}.$$  

(4.3)

The tree thus defined depends of course on the base $M$; however, if $M$ is fixed, as it will be once we fix the direction set $\Omega = \gamma(C_M)$ (see Section 1.2), we will omit its usage in our notation, denoting the tree $T([0, 1); M)$ by $T([0, 1))$ instead.

Identifying the root of the tree defined in (4.3) with the interval $[0, 1)$ and the vertex $\langle i_1, \ldots, i_k \rangle$ with the interval $[i \cdot M^{-k}, (i+1) \cdot M^{-k})$, where $i$ and $\langle i_1, \ldots, i_k \rangle$ are related by (4.2), we observe that the vertices of $T([0, 1); M)$ at height $k$ yield a partition of $[0, 1)$ into $M$-adic subintervals of length $M^{-k}$. This tree has a self-similar structure: every
vertex of $T([0, 1]; M)$ has $M$ children and the subtree generated by any vertex as the root is isomorphic to $T([0, 1]; M)$. In the sequel, we will refer to such a tree as a full $M$-adic tree.

Any $x \in [0, 1)$ can be realized as the intersection of a nested sequence of $M$-adic intervals, namely

$$\{x\} = \bigcap_{k=0}^{\infty} I_k(x),$$

where $I_k(x) = [i_k(x) \cdot M^{-k}, (i_k(x) + 1) \cdot M^{-k})$. The point $x$ should be visualized as the destination of the infinite ray $\emptyset \to \langle i_1(x) \rangle \to \langle i_1(x), i_2(x) \rangle \to \cdots \to \langle i_1(x), i_2(x), \ldots, i_k(x) \rangle \to \cdots$ in $T([0, 1]; M)$. Conversely, every infinite ray $\emptyset \to \langle i_1 \rangle \to \langle i_1, i_2 \rangle \to \langle i_1, i_2, i_3 \rangle \cdots$ identifies a unique $x \in [0, 1)$ given by the convergent sum

$$x = \sum_{j=1}^{\infty} \frac{i_j}{M^j}.$$

Thus the tree $T([0, 1]; M)$ can be identified with the interval $[0, 1)$ exactly. Any subset $E \subseteq [0, 1)$ is then given by a subtree $T(E; M)$ of $T([0, 1]; M)$ consisting of all infinite rays that identify some $x \in E$. As before, we will drop the notation for the base $M$ in $T(E; M)$ once this base has been fixed.

Any truncation of $T(E; M)$, say up to height $k$, will be denoted by $T_k(E; M)$ and should be visualized as a covering of $E$ by $M$-adic intervals of length $M^{-k}$. More precisely, $(i_1, \ldots, i_k) \in T_k(E; M)$ if and only if $E \cap [i \cdot M^{-k}, (i + 1) \cdot M^{-k}) \neq \emptyset$, where $i$ and $(i_1, \ldots, i_k)$ are related by (4.2).

We now state and prove a key structural result about our sets of interest, the generalized Cantor sets $C_M$.

**Proposition 4.2.** Fix any integer $M \geq 3$. Define $C_M$ as in Section 1.2. Then

$$T(C_M; M) \cong T([0, 1]; 2).$$

That is, the $M$-adic tree representation of $C_M$ is isomorphic to the full binary tree, illustrated in Figure 4.

**Proof.** Denote $T = T(C_M; M)$ and $T' = T([0, 1]; 2)$. We must construct a bijective sticky map $\psi : T \to T'$. First, define $\psi(v_0) = v'_0$, where $v_0$ is the root of $T$ and $v'_0$ is the root of $T'$.

Now, for any $k \geq 1$, consider the vertex $\langle i_1, i_2, \ldots, i_k \rangle \in T$. We know that $i_j \in \mathbb{Z}_M$ for all $j$. Furthermore, for any fixed $j$, this vertex corresponds to a $k$th level subinterval
Figure 4: A pictorial depiction of the isomorphism between a standard middle-thirds Cantor set and its representation as a full binary subtree of the full base $M = 3$ tree.

of $c_M^{[k]}$. Every such $k$-th level interval is replaced by exactly two arbitrary $(k+1)$-th level subintervals in the construction of $c_M^{[k+1]}$. Therefore, there exists $N_1 := N_1(\langle i_1, \ldots, i_k \rangle)$, $N_2 := N_2(\langle i_1, \ldots, i_k \rangle) \in \mathbb{Z}_M$, with $N_1 < N_2$, such that $\langle i_1, \ldots, i_k, i_{k+1} \rangle \in \mathcal{T}$ if and only if $i_{k+1} = N_1$ or $N_2$. Consequently, we define

$$\psi(\langle i_1, i_2, \ldots, i_k \rangle) = \langle l_1, l_2, \ldots, l_k \rangle \in \mathcal{T}'$$

where

$$l_{j+1} = \begin{cases} 0 & \text{if } i_{j+1} = N_1(\langle i_1, \ldots, i_j \rangle), \\ 1 & \text{if } i_{j+1} = N_2(\langle i_1, \ldots, i_j \rangle). \end{cases}$$

The mapping $\psi$ is injective by construction and surjectivity follows from the binary selection of subintervals at each stage in the construction of $c_M$. Moreover, $\psi$ is sticky by (4.4).

The following corollary is an easy consequence of the above and left to the reader.

Corollary 4.3. Recall the definition of $D_M^{[N]}$ from Section 2.2. Then

$$\mathcal{T}_N(D_M^{[N]}; M) \sim \mathcal{T}_N([0, 1); 2).$$

Proposition 4.2 and Corollary 4.3 guarantee that the tree encoding our set of directions will retain a certain binary structure. This fact will prove vital to establishing Theorem 1.2.

4.3 Encoding higher dimensional bounded subsets of Euclidean space by trees

The approach to encoding a bounded subset of Euclidean space by a tree extends readily to higher dimensions. For any $\mathbf{i} = \langle j_1, \ldots, j_d \rangle \in \mathbb{Z}^d$ such that $\mathbf{i} \cdot M^{-k} \in [0, 1)^d$, we can
apply (1.2) to each component of \( i \) to obtain
\[
\frac{i}{M^k} = \frac{i_1}{M} + \frac{i_2}{M^2} + \cdots + \frac{i_k}{M^k},
\]
with \( i_j \in \mathbb{Z}_M^d \) for all \( j \). As before, we identify \( i \) with \( \langle i_1, \ldots, i_k \rangle \).

Let \( \phi : \mathbb{Z}_M^d \to \{0, 1, \ldots, M^d - 1\} \) be an enumeration of \( \mathbb{Z}_M^d \). Define the full \( M^d \)-adic tree
\[
T([0,1]^d; M, \phi) = \left\{ \langle \phi(i_1), \ldots, \phi(i_k) \rangle : k \geq 0, \ i_j \in \mathbb{Z}_M^d \right\}.
\] (4.5)
The collection of \( k \)th generation vertices of this tree may be thought of as the \( d \)-fold Cartesian product of the \( k \)th generation vertices of \( T([0,1); M]) \). For our purposes, it will suffice to fix \( \phi \) to be the lexicographic ordering, and so we will omit the notation for \( \phi \) in (4.5), writing simply, and with a slight abuse of notation,
\[
T([0,1]^d; M) = \left\{ \langle i_1, \ldots, i_k \rangle : k \geq 0, \ i_j \in \mathbb{Z}_M^d \right\}.
\] (4.6)
As before, we will refer to the tree in (4.6) by the notation \( T([0,1)^d) \) once the base \( M \) has been fixed.

By a direct generalization of our one-dimensional results, each vertex \( \langle i_1, \ldots, i_k \rangle \) of \( T([0,1)^d; M) \) at height \( k \) represents the unique \( M \)-adic cube in \([0,1)^d \) of sidelength \( M^{-k} \), containing \( i \cdot M^{-k} \), of the form
\[
\left[ \frac{j_1}{M^k}, \frac{j_1 + 1}{M^k} \right) \times \cdots \times \left[ \frac{j_d}{M^k}, \frac{j_d + 1}{M^k} \right).
\]
As in the one-dimensional setting, any \( x \in [0,1)^d \) can be realized as the intersection of a nested sequence of \( M \)-adic cubes. Thus, we view the tree in (4.6) as an encoding of the set \([0,1)^d \) with respect to base \( M \). As before, any subset \( E \subseteq [0,1)^d \) then corresponds to a subtree of \( T([0,1)^d; M) \).

The connection between sets and trees encoding them leads to the following easy observations that we record for future use in Lemma 9.3.

**Lemma 4.4.** Let \( \Omega_N \) be the set defined in (2.2).

(a) Given \( \Omega_N \), there is a constant \( C_1 > 0 \) (depending only on \( d \) and \( C, c \) from (1.5)) such that for any \( 1 \leq k \leq N \), the number of \( k \)th generation vertices in \( T_N(\Omega_N^d); M) \) is \( \leq C_1 2^k \).

(b) For any compact set \( \mathcal{K} \subseteq \mathbb{R}^{d+1} \), there exists a constant \( C(\mathcal{K}) > 0 \) with the following property. For any \( x = (x_1, \ldots, x_{d+1}) \in \mathcal{K} \), and \( 1 \leq k \leq N \), the number of \( k \)th generation vertices in \( T_N(E(x); M) \) is \( \leq C(\mathcal{K}) 2^k \), where \( E(x) := (x - x_1 \Omega_N) \cap \{0\} \times [0,1)^d \).

**Proof.** There are exactly \( 2^k \) basic intervals of level \( k \) that comprise \( C_M^{(k)} \). Under \( \gamma \), each such basic interval maps into a set of diameter at most \( CM^{-k} \). Since \( \Omega_N = \gamma(D_M^{(N)}) \subseteq \mathcal{K} \),
\( \gamma(C^{[k]}_M) \), the number of \( k \)th generation vertices in \( T_N(\Omega_N; M) \), which is also the number of \( k \)th level \( M \)-adic cubes needed to cover \( \Omega_N \), is at most \( C_12^k \). This proves (a).

Let \( Q \) be any \( k \)th generation \( M \)-adic cube such that \( Q \cap \Omega \neq \emptyset \). Then on one hand, \((x - x_1Q) \cap (x - x_1\Omega_N) \neq \emptyset\); on the other hand, the number of \( k \)th level \( M \)-adic cubes covering \((x - x_1Q)\) is \( \leq C(K) \), and part (b) follows.

**Notation:** We end this section with a notational update. In light of the discussion above and for simplicity, we will henceforth identify a vertex \( u = \langle i_1, i_2, \ldots, i_k \rangle \in T(\{0, 1\}^d) \) with the corresponding cube \( \{0\} \times u \) lying on the root hyperplane \( \{0\} \times [0, 1)^d \). In this parlance, a vertex \( t \in T_N([0, 1)^d) \) of height \( N \) is the same as a root cube \( Q_t \) (or \( \tilde{Q}_t \)) defined in (2.1), and the notation \( t \subseteq u \) stands both for set containment as well as tree ancestry.

5 Electrical circuits and percolation on trees

5.1 The percolation process associated to a tree

The proof of Theorem 1.2 will require consideration of a special probabilistic process on certain trees called a (bond) percolation. Imagine a liquid that is poured on top of some porous material. How will the liquid flow - or percolate - through the holes of the material? How likely is it that the liquid will flow from hole to hole in at least one uninterrupted path all the way to the bottom? The first question forms the intuition behind a formal percolation process, whereas the second question turns out to be of critical importance to the proof of Theorem 1.2; this idea plays a key role in establishing the planar analogue of that theorem in Bateman and Katz [4], and again in the more general framework of [3].

Although it is possible to speak of percolation processes in far more general terms (see [11]), we will only be concerned with a percolation process on a tree. Accordingly, given some tree \( T \) with vertex set \( V \) and edge set \( E \), we define an edge-dependent Bernoulli (bond) percolation process to be any collection of random variables \( \{X_e : e \in E\} \), where \( X_e \) is Bernoulli(\( p_e \)) with \( p_e < 1 \). The parameter \( p_e \) is called the survival probability of the edge \( e \). We will always be concerned with a particular type of percolation on our trees: we define a standard Bernoulli(\( p \)) percolation to be one where the random variables \( \{X_e : e \in E\} \) are mutually independent and identically distributed Bernoulli(\( p \)) random variables, for some \( p < 1 \). In fact, for our purposes, it will suffice to consider only standard Bernoulli(\( \frac{1}{2} \)) percolations.

Rather than imagining a tree with a percolation process as the behaviour of a liquid acted upon by gravity in a porous material, it will be useful to think of the percolation process as acting more directly on the mathematical object of the tree itself. Given some percolation process on a tree \( T \), we will think of the event \( \{X_e = 0\} \) as the event that we remove the edge \( e \) from the edge set \( E \), and the event \( \{X_e = 1\} \) as the event that we retain this edge; denote the random set of retained edges by \( E^* \). Notice that with this interpretation, after percolation there is no guarantee that \( E^* \), the subset of edges that remain after percolation, defines a subtree of \( T \). In fact, it can be quite likely that the
subgraph that remains after percolation is a union of many disconnected subgraphs of \( T \).

For a given edge \( e \in \mathcal{E} \), we think of \( p = \Pr(X_e = 1) \) as the probability that we retain this edge after percolation. The probability that at least one uninterrupted path remains from the root of the tree to its bottommost level is given by the survival probability of the corresponding percolation process. More explicitly, given a percolation on a tree \( T \), the survival probability after percolation is the probability that the random variables associated to all edges of at least one ray in \( T \) take the value 1; i.e.

\[
\Pr(\text{survival after percolation on } T) := \Pr \left( \bigcup_{R \in \partial T} \cap_{e \in E \cap R} \{X_e = 1\} \right). \tag{5.1}
\]

Estimation of this probability will prove to be a valuable tool in the proof of Theorem 1.2. This estimation will require reimagining a tree as an electrical network.

5.2 Trees as electrical networks

Formally, an electrical network is a particular kind of weighted graph. The weights of the edges are called conductances and their reciprocals are called resistances. In his seminal works on the subject, Lyons visualizes percolation on a tree as a certain electrical network. In [17], he lays the groundwork for this correspondence. While his results hold in great generality, we describe his results in the context of standard Bernoulli percolation on a locally finite, rooted labelled tree only. We briefly review the concepts relevant to our application here.

A percolation process on the truncation of any given tree \( T \) is naturally associated to a particular electrical network. To see this, we truncate the tree \( T \) at height \( N \) and place the positive node of a battery at the root of \( T_N \). Then, for every ray in \( \partial T_N \), there is a unique terminating vertex; we connect each of these vertices to the negative node of the battery. A resistor is placed on every edge \( e \) of \( T_N \) with resistance \( R_e \) defined by

\[
\frac{1}{R_e} = \frac{1}{1 - p_e} \prod_{\emptyset \subset v(e') \subseteq v(e)} p_{e'} . \tag{5.2}
\]

Notice that the resistance for the edge \( e \) is essentially the reciprocal of the probability that a path remains from the root of the tree to the vertex \( v_e \) after percolation. For standard Bernoulli(\( \frac{1}{2} \)) percolation, we have

\[
R_e = 2^{h(v_e) - 1} . \tag{5.3}
\]

One fact that will prove useful for us later is that connecting any two vertices at a given height by an ideal conductor (i.e. one with zero resistance) only decreases the overall resistance of the circuit. This will allow us to more easily estimate the total resistance of a generic tree.
Proposition 5.1. Let $T_N$ be a truncated tree of height $N$ with corresponding electrical network generated by a standard Bernoulli($\frac{1}{2}$) percolation process. Suppose at height $k < N$ we connect two vertices by a conductor with zero resistance. Then the resulting electrical network has a total resistance no greater than that of the original network.

Proof. Let $u$ and $v$ be the two vertices at height $k$ that we will connect with an ideal conductor. Let $R_1$ denote the resistance between $u$ and $D(u,v)$, the youngest common ancestor of $u$ and $v$; let $R_2$ denote the resistance between $v$ and $D(u,v)$. Let $R_3$ denote the total resistance of the subtree of $T_N$ generated by the root $u$ and let $R_4$ denote the total resistance of the subtree of $T_N$ generated by the root $v$. These four connections define a subnetwork of our tree, depicted in Figure 5(a). The connection of $u$ and $v$ by an ideal conductor, as pictured in Figure 5(b), can only change the total resistance of this subnetwork, as that action leaves all other connections unaltered. It therefore suffices to prove that the total resistance of the subnetwork comprising of the resistors $R_1, R_2, R_3$ and $R_4$ can only decrease if $u$ and $v$ are joined by an ideal conductor.

In the original subnetwork, the resistors $R_1$ and $R_3$ are in series, as are the resistors $R_2$ and $R_4$. These pairs of resistors are also in parallel with each other. Thus, we calculate the total resistance of this subnetwork, $R_{\text{original}}$:

$$R_{\text{original}} = \left( \frac{1}{R_1 + R_3} + \frac{1}{R_2 + R_4} \right)^{-1} = \frac{(R_1 + R_3)(R_2 + R_4)}{R_1 + R_2 + R_3 + R_4}. \tag{5.4}$$

After connecting vertices $u$ and $v$ by an ideal conductor, the structure of our subnetwork is inverted as follows. The resistors $R_1$ and $R_2$ are in parallel, as are the resistors $R_3$ and $R_4$. 

Figure 5: (a) The original subnetwork with the resistors $R_1, R_3$ and $R_2, R_4$ in series; (b) the new subnetwork obtained by connecting vertices $u$ and $v$ by an ideal conductor.
and $R_4$, and these pairs of resistors are also in series with each other. Therefore, we calculate the new total resistance of this subnetwork, $R_{\text{new}}$, as

$$R_{\text{new}} = \left( \frac{1}{R_1} + \frac{1}{R_2} \right)^{-1} + \left( \frac{1}{R_3} + \frac{1}{R_4} \right)^{-1}$$

$$= \frac{R_1 R_2 (R_3 + R_4) + R_3 R_4 (R_1 + R_2)}{(R_1 + R_2)(R_3 + R_4)}.$$  \hspace{1cm} (5.5)

We claim that (5.4) is greater than or equal to (5.5). To see this, simply cross-multiply these expressions. After cancellation of common terms, our claim reduces to

$$R_1^2 R_2^2 + R_3^2 R_4^2 \geq 2 R_1 R_2 R_3 R_4.$$  \hspace{1cm} (5.6)

But this is trivially satisfied since $(a - b)^2 \geq 0$ for any real numbers $a$ and $b$. \hfill \Box

### 5.3 Estimating the survival probability after percolation

We now present Lyons’ pivotal result linking the total resistance of an electrical network and the survival probability under the associated percolation process.

**Theorem 5.2** (Lyons, Theorem 2.1 of [18]). Let $T$ be a tree with mutually associated percolation process and electrical network, and let $R(T)$ denote the total resistance of this network. If the percolation is Bernoulli, then

$$\frac{1}{1 + R(T)} \leq \Pr(T) \leq \frac{2}{1 + R(T)},$$

where $\Pr(T)$ denotes the survival probability after percolation on $T$.

We will not require the full strength of this theorem. A reasonable upper bound on the survival probability coupled with the result of Proposition 5.1 will suffice for our applications. For completeness, we state and prove a sufficient simpler version of Theorem 5.2 as essentially formulated by Bateman and Katz [4].

**Proposition 5.3.** Let $M \geq 2$ and let $T$ be a subtree of a full $M$-adic tree. Let $R(T)$ and $\Pr(T)$ be as in Theorem 5.2. Then under Bernoulli percolation, we have

$$\Pr(T) \leq \frac{2}{1 + R(T)}.$$  \hspace{1cm} (5.6)

**Proof.** We will only focus on the case when $R(T) \geq 1$, since otherwise (5.6) holds trivially. We prove this by induction on the height of the tree $N$. When $N = 0$, then (5.6) is trivially satisfied. Now suppose that up to height $N - 1$, we have

$$\Pr(T) \leq \frac{2}{1 + R(T)}.$$

Suppose $T$ is of height $N$. We can view the tree $T$ as its root together with at most $M$ edges connecting the root to the subtrees $T_1, \ldots, T_M$ of height $N - 1$ generated by
the terminating vertices of these edges. If there are \( k < M \) edges originating from the root, then we take \( M - k \) of these subtrees to be empty. Note that by the induction hypothesis, (5.6) holds for each \( T_j \). To simplify notation, we denote

\[
\Pr(T_j) = P_j \quad \text{and} \quad R(T_j) = R_j,
\]

taking \( P_j = 0 \) and \( R_j = \infty \) if \( T_j \) is empty.

Using independence and recasting \( \Pr(T) \) as one minus the probability of not surviving after percolation on \( T \), we have the formula:

\[
\Pr(T) = 1 - \prod_{k=1}^{M} \left( 1 - \frac{1}{2} P_k \right).
\]

Note that the function \( F(x_1, \ldots, x_M) = 1 - (1-x_1/2)(1-x_2/2) \cdots (1-x_M/2) \) is monotone increasing in each variable on \([0, 2]^M\). Now define

\[
Q_j := \frac{2}{1 + R_j}.
\]

Since resistances are nonnegative, we know that \( Q_j \leq 2 \) for all \( j \). Therefore,

\[
\Pr(T) = F(P_1, \ldots, P_M) \\
\leq F(Q_1, \ldots, Q_M) \\
\leq \frac{1}{2} \sum_{k=1}^{M} Q_k.
\]

Here, the first inequality follows by monotonicity and the induction hypothesis. Plugging in the definition of \( Q_k \), we find that

\[
\Pr(T) \leq \sum_{k=1}^{M} \frac{1}{1 + R_k}.
\]

But since each resistor \( R_j \) is in parallel, we know that

\[
\frac{1}{R(T)} = \sum_{k=1}^{M} \frac{1}{1 + R_k}.
\]

Combining this formula with the previous inequality and recalling that \( R(T) \geq 1 \), we have

\[
\Pr(T) \leq \frac{1}{R(T)} \leq \frac{2}{1 + R(T)},
\]

as required. \( \square \)
6 The random mechanism and the property of stickiness

As discussed in the introduction of this paper, the construction of a Kakeya-type set with orientations given by $\Omega$ will require a certain random mechanism. We now describe this mechanism in detail.

In order to assign a slope $\sigma(\cdot)$ to the tubes $R_{t,\sigma} := \mathcal{P}_{t,\sigma(t)}$ given by (2.3), we want to define a collection of random variables $\{X_{(i_1,\ldots,i_k)} : \langle i_1,\ldots,i_k \rangle \in \mathcal{T}([0,1]^d)\}$, one on each edge of the tree used to identify the roots of these tubes. The tree $\mathcal{T}_1([0,1]^d)$ consists of all first generation edges of $\mathcal{T}([0,1]^d)$. It has exactly $M^d$ many edges and we place (independently) a Bernoulli($\frac{1}{2}$) random variable on each edge: $X_{(0)}, X_{(1)}, \ldots, X_{(M^d-1)}$. Now, the tree $\mathcal{T}_2([0,1]^d)$ consists of all first and second generation edges of $\mathcal{T}([0,1]^d)$ and has $M^d + M^{2d}$ many edges and we place (independently) a new Bernoulli($\frac{1}{2}$) random variable on each of the $M^{2d}$ second generation edges. We label these $X_{(i_1,i_2)}$ where $0 \leq i_1, i_2 < M^d$. We proceed in this way, eventually assigning an ordered collection of independent Bernoulli($\frac{1}{2}$) random variables to the tree $\mathcal{T}_N([0,1]^d)$:

$$X_N := \left\{ X_{(i_1,\ldots,i_k)} : \langle i_1,\ldots,i_k \rangle \in \mathcal{T}_N([0,1]^d), \ 1 \leq k \leq N \right\},$$

where $X_{(i_1,\ldots,i_k)}$ is assigned to the unique edge identifying $\langle i_1, i_2, \ldots, i_k \rangle$, namely the edge joining $\langle i_1, i_2, \ldots, i_{k-1} \rangle$ to $\langle i_1, i_2, \ldots, i_k \rangle$. Each realization of $X_N$ is a finite ordered collection of cardinality $M^d + M^{2d} + \cdots + M^{Nd}$ with entries either 0 or 1.

We will now establish that every realization of the random variable $X_N$ defines a sticky map between the truncated position tree $\mathcal{T}_N([0,1]^d)$ and the truncated binary tree $\mathcal{T}_N([0,1];2)$, as defined in Definition 4.1. Fix a particular realization $X_N = x = \{x_{(i_1,\ldots,i_k)}\}$. Define a map $\tau_x : \mathcal{T}_N([0,1]^d) \to \mathcal{T}_N([0,1];2)$, where

$$\tau_x(\langle i_1, i_2, \ldots, i_k \rangle) = \langle x_{(i_1)}, x_{(i_1,i_2)}, \ldots, x_{(i_1,i_2,\ldots,i_k)} \rangle.$$

We then have the following key proposition.

**Proposition 6.1.** The map $\tau_x$ just defined is sticky for every realization $x$ of $X_N$. Conversely, any sticky map $\tau$ between $\mathcal{T}_N([0,1]^d)$ and $\mathcal{T}_N([0,1];2)$ can be written as $\tau = \tau_x$ for some realization $x$ of $X_N$.

**Proof.** Recalling Definition 4.1, we need to verify that $\tau_x$ preserves heights and lineages. By (6.1), any finite sequence $v = \langle i_1, i_2, \ldots, i_k \rangle$ in $\mathcal{T}([0,1]^d)$ is mapped to a sequence of the same length in $\mathcal{T}([0,1];2)$. Therefore $h(v) = h(\tau_x(v))$ for every $v \in \mathcal{T}([0,1]^d)$.

Next suppose $u \supset v$. Then $u = \langle i_1, \ldots, i_h(u) \rangle$, with $h(u) \leq k$. So again by (6.1),

$$\tau_x(u) = \left\langle x_{(i_1)}, \ldots, x_{(i_1,\ldots,i_h(u))} \right\rangle \supset \left\langle x_{(i_1)}, \ldots, x_{(i_1,\ldots,i_h(u))}, \ldots, x_{(i_1,\ldots,i_k)} \right\rangle = \tau_x(v).$$

Thus, $\tau_x$ preserves lineages, establishing the first claim in Proposition 6.1.

For the second, fix a sticky map $\tau : \mathcal{T}_N([0,1]^d) \to \mathcal{T}_N([0,1];2)$. Define $x_{(i_1)} := \tau(\langle i_1 \rangle)$, $x_{(i_1,i_2)} := \pi_2 \circ \tau(\langle i_1, i_2 \rangle)$, and in general

$$x_{(i_1,\ldots,i_k)} := \pi_k \circ \tau(\langle i_1, i_2, \ldots, i_k \rangle), \quad k \geq 1,$$

where $\pi_k$ is the $k$th projection.
where $\pi_k$ denotes the projection map whose image is the $k$th coordinate of the input sequence. The collection $x = \{x_{(i_1,i_2,\cdots,i_k)}\}$ is the unique realization of $X_N$ that verifies the second claim.  

6.1 Slope assignment algorithm

Recall from Sections 4.2 and 2.2 that $\Omega := \gamma(C_M)$ and $\Omega_N := \gamma(D_M^{[N]})$, where $C_M$ is the generalized Cantor-type set and $D_M^{[N]}$ a finitary version of it. In order to exploit the binary structure of the trees $T(C_M) := T([C_M; M])$ and $T(D_M^{[N]}) := T([D_M^{[N]}; M])$ advanced in Proposition 4.2 and Corollary 4.3, we need to map traditional binary sequences onto the subsequences of $\{0, \ldots, M - 1\}^\infty$ defined by $C_M$ or $D_M^{[N]}$.

**Proposition 6.2.** Every sticky map $\tau$ as in (6.1) that maps $T_N([0,1]^d; M)$ to $T_N([0,1); 2)$ induces a natural mapping $\sigma = \sigma_\tau$ from $T_N([0,1)^d)$ into $\Omega_N$. The maps $\sigma_\tau$ obey a uniform Lipschitz-type condition: for any $t,t' \in T_N([0,1)^d)$, $t \neq t'$,

$$|\sigma_\tau(t) - \sigma_\tau(t')| \leq CM^{-h(D(\tau(t),\tau(t')))},$$

(6.2)

where $C$ is as in (6.1).

**Remark:** While the choice of $D_M^{[N]}$ for a given $C_M^{[N]}$ is not unique, the mapping $\tau \mapsto \sigma_\tau$ is unique given a specific choice. Moreover, if $D_M^{[N]}$ and $D_M$ are two selections of finitary direction sets at scale $M^{-N}$, then the corresponding maps $\sigma_\tau$ and $\sigma_\tau$ must obey

$$|\sigma_\tau(v) - \sigma_\tau(v)| \leq CM^{-h(v)}$$

for every $v \in T_N([0,1^d)$,  

(6.3)

where $C$ is as in (6.1). Thus given $\tau$, the slope in $\Omega$ that is assigned by $\sigma_\tau$ to an $M$-adic cube in $\{0\} \times [0,1)^d$ of side length $M^{-N}$ is unique up to an error of $O(M^{-N})$. As a consequence $P_{t,\sigma_\tau}$ and $P_{t,\sigma_\tau}$ are comparable, in the sense that each is contained in a $O(M^{-N})$-thickening of the other.

**Proof.** There are two links that allow passage of $\tau$ to $\sigma$. The first of these is the isomorphism $\psi$ constructed in Proposition 4.2 that maps $T(C_M; M)$ onto $T([0,1); 2)$. Under this isomorphism, the pre-image of any $k$-long sequence of 0’s and 1’s is a vertex $w$ of height $k$ in $T(C_M; M)$, in other words one of the $2^k$ chosen $M$-adic intervals of length $M^{-k}$ that constitute $C_M^{[k]}$. The second link is a mapping $\Phi : T_N(C_M; M) \rightarrow D_M^{[N]}$ that sends every vertex $w$ to a point in $C_M \cap w$, where, per our notational agreement at the end of Section 4, we have also let $w$ denote the particular $M$-adic interval that it identifies. While the choice of the image point, i.e., $D_M^{[N]}$ is not unique, any two candidates $\Phi, \overline{\Phi}$ satisfy

$$|\Phi(w) - \Phi(w)| \leq \text{diam}(w) = M^{-h(w)}$$

for every $w \in T_N(C_M; M)$.  

(6.4)

We are now ready to describe the assignment $\tau \mapsto \sigma = \sigma_\tau$. Given a sticky map $\tau : T_N([0,1)^d; M) \rightarrow T_N([0,1); 2)$ such that

$$\tau(\langle i_1,i_2,\cdots,i_k \rangle) = \langle X_{(i_1)},\cdots,X_{(i_1,i_2,\cdots,i_k)} \rangle,$$

28
the transformed random variable

\[ Y_{(i_1,i_2,...,i_k)} := \gamma \circ \Phi \circ \psi^{-1} \left( (X_{(i_1)}, X_{(i_1,i_2)}, \ldots, X_{(i_1,i_2,...,i_k)}) \right) \]

associates a random direction in \( \Omega_N = \gamma(D^{[N]}_M) \) to the sequence \( t = \langle i_1, \ldots, i_k \rangle \) identified with a unique vertex \( t \in T_N([0,1]^d) \). Thus, defining

\[ \sigma := \gamma \circ \Phi \circ \psi^{-1} \circ \tau \]  

(6.5)

gives the appropriate (random) mapping claimed by the proposition. The weak Lipschitz condition (6.2) is verified as follows,

\[ |\sigma_\tau(t) - \sigma_\tau(t')| = |\gamma \circ \Phi \circ \psi^{-1} \circ \tau(t) - \gamma \circ \Phi \circ \psi^{-1} \circ \tau(t')| \]
\[ \leq C|\Phi \circ \psi^{-1} \circ \tau(t) - \Phi \circ \psi^{-1} \circ \tau(t')| \]
\[ \leq CM^{-h(D(\psi^{-1} \circ \tau(t), \psi^{-1} \circ \tau(t')))} \]
\[ = CM^{-h(D(\tau(t), \tau(t')))}. \]

Here the first inequality follows from (1.5), the second from the definition of \( \Phi \). The third step uses the fact that \( \psi \) is an isomorphism, so that \( h(D(\tau(t), \tau(t'))) = h(D(\psi^{-1} \circ \tau(t), \psi^{-1} \circ \tau(t'))) \). Finally, any non-uniqueness in the definition of \( \sigma \) comes from \( \Phi \), hence (6.3) follows from (6.4) and (1.5). \( \square \)

The stickiness of the maps \( \tau_\kappa \) is built into their definition (6.1). The reader may be interested in observing that there is a naturally sticky map already introduced in this article, which should be viewed as the inspiration for the construction of \( \tau \) and \( \sigma_\tau \). We refer to the geometric content of Lemma 3.7, which in the language of trees has a particularly succinct reformulation. We record this below.

**Lemma 6.3.** For \( C_0 \) obeying the requirement of Lemma 3.7 and \( p \in [C_0, C_0 + 1] \times \mathbb{R}^d \), let \( \text{Poss}(p) \) be as in (3.3). Let \( \Phi \) and \( \psi \) be the maps used in Proposition 6.2. Then the map \( t \mapsto \beta(t) \) which maps every \( t \in \text{Poss}(p) \) to the unique \( \beta(t) \in [0,1) \) such that

\[ p \in \mathcal{P}_{t,v(t)} \text{ where } v(t) = \gamma \circ \Phi \circ \psi^{-1} \circ \beta(t), \]  

(6.6)

extends as a well-defined sticky map from \( T_N(\text{Poss}(p);M) \) to \( T_N([0,1);2) \).

**Proof.** By Lemma 3.7(a), there exists for every \( t \in \text{Poss}(p) \) a unique \( v(t) \in \Omega_N \) such that \( p \in \mathcal{P}_{t,v(t)} \). Let us therefore define for \( 1 \leq k \leq N \),

\[ \beta(\pi_1(t), \ldots, \pi_k(t)) = (\pi_1 \circ \beta(t), \ldots, \pi_k \circ \beta(t)) \]  

(6.7)

where \( \beta(t) \) is as in (6.3) and as always \( \pi_k \) denotes the projection to the \( k \)th coordinate of an input sequence. More precisely, \( \pi_k(t) \) represents the unique \( k \)th level \( M \)-adic cube that contains \( t \). Similarly \( \pi_k(\beta(t)) \) is the \( k \)th component of the \( N \)-long binary sequence that identifies \( \beta(t) \). The function \( \beta \) defined in (6.7) maps \( T_N(\text{Poss}(p);M) \) to \( T_N([0,1);2) \), and agrees with \( \beta \) as in (6.6) if \( k = N \).
To check that the map is consistently defined, we pick \( t \neq t' \) in Poss\((p)\) with \( u = D(t, t') \) and aim to show that \( \beta(\pi_1(t), \ldots, \pi_k(t)) = \beta(\pi_1(t'), \ldots, \pi_k(t')) \) for all \( k \) such that \( k \leq h(u) \). But by definition (6.6), \( v(t) \) and \( v(t') \) have the property that \( p \in \mathcal{P}_{t, v(t)} \cap \mathcal{P}_{t', v(t')} \). Hence Lemma 3.7(b) asserts that \( \alpha(t) = \gamma^{-1}(v(t)) \) and \( \alpha(t') = \gamma^{-1}(v(t')) \) share the same basic interval at step \( h(u) \) of the Cantor construction. Thus \( \beta(t) = \psi \circ \Phi^{-1} \circ \alpha(t) \) and \( \beta(t') = \psi \circ \Phi^{-1} \circ \alpha(t') \) have a common ancestor in \( T_N([0, 1); 2) \) at height \( h(u) \), and hence \( \pi_k(\beta(t)) = \pi_k(\beta(t')) \) for all \( k \leq h(u) \), as claimed. Preservation of heights and lineages is a consequence of the definition (6.7), and stickiness follows.

6.2 Construction of Kakeya-type sets revisited

As \( \tau \) ranges over all sticky maps \( \tau : T_N([0, 1]^d) \rightarrow T_N([0, 1); 2) \) with \( x \in X_N \), we now have for every vertex \( t \in T_N([0, 1]^d) \) with \( h(t) = N \) a random sticky slope assignment \( \sigma(t) \in \Omega_N \) defined as above. For all such \( t \), this generates a randomly oriented tube \( P_{t, \sigma} \) given by (2.3) rooted at the \( M \)-adic cube \( Q_t \) identified by \( t \), with sidelength \( \kappa_d \cdot M^{-N} \) in the \( x_1 = 0 \) plane. We may rewrite the collection of such tubes from (2.4) as

\[
K_N(\sigma) := \bigcup_{t \in T_N([0, 1]^d) \atop h(t) = N} P_{t, \sigma}.
\]  
(6.8)

On average, a random collection of tubes with the above described sticky slope assignment will comprise a Kakeya-type set, as per (1.2). Specifically, we will show in the next section that the following proposition holds. In view of Proposition 2.1, this will suffice to prove Theorem 1.2.

**Proposition 6.4.** Suppose \((\Sigma_N, \mathcal{P}(\Sigma_N), Pr)\) is the probability space of sticky maps described above, equipped with the uniform probability measure. For every \( \sigma \in \Sigma_N \), there exists a set \( K_N(\sigma) \) as defined in (6.8), with tubes oriented in directions from \( \Omega_N = \gamma(\mathcal{P}^{[N]}_M) \). Then these random sets obey the hypotheses of Proposition 2.1 with

\[
a_N = c_M \sqrt{\log N} \quad \text{and} \quad b_N = C_M \frac{N}{N},
\]  
(6.9)

where \( c_M \) and \( C_M \) are fixed positive constants depending only on \( M \) and \( d \). The content of Proposition 2.1 allows us to conclude that \( \Omega \) admits Kakeya-type sets.

7 Slope probabilities and root configurations

Having established the randomization method for assigning slopes to tubes, we are now in a position to apply this toward the estimation of probabilities of certain events that will be of interest in the next section. Roughly speaking, we wish to compute conditional probabilities that one or more cubes on the root hyperplane are assigned prescribed slopes, provided similar information is available for other cubes.
Lemma 7.1. Let us fix \( v_1, v_2 \in \Omega_N \), so that \( v_1 = \gamma(\alpha_1) \) and \( v_2 = \gamma(\alpha_2) \) for unique \( \alpha_1, \alpha_2 \in D_M^{[N]} \). We also fix \( t_1, t_2 \in T_N([0,1]^d) \), \( h(t_1) = h(t_2) = N \), \( t_1 \not= t_2 \). Let us denote by \( u \in T_N([0,1]^d) \) and \( \alpha \in T_N(D_M^{[N]}) \) the youngest common ancestors of \( (t_1, t_2) \) and \((\alpha_1, \alpha_2)\) respectively, i.e., \( u = D(t_1, t_2) \), \( \alpha = D(\alpha_1, \alpha_2) \). Then

\[
Pr(\sigma(t_2) = v_2 | \sigma(t_1) = v_1) = \begin{cases} 2^{-(N-h(\alpha))} & \text{if } h(u) \leq h(\alpha), \\ 0 & \text{otherwise.} \end{cases} \tag{7.1}
\]

**Proof.** Keeping in mind the slope assignment as described in (6.5), and the stickiness of the map \( \tau \) as given in Proposition 6.1, the proof can be summarized as in Figure 6. Since \( t_1 \) and \( t_2 \) must map to \( v_1 = \gamma(\alpha_1) \) and \( v_2 = \gamma(\alpha_2) \) under \( \sigma = \sigma_\tau \), the sticky map \( \psi^{-1} \circ \tau \) must map \( t_1 \) and \( t_2 \) to the \( N \)th stage basic intervals in the Cantor construction containing \( \alpha_1 \) and \( \alpha_2 \) respectively. Since sticky maps preserve heights and lineages, we must have \( h(\alpha) \geq h(u) \). Assuming this, we simply count the number of distinct edges on the ray defining \( t_2 \) that are not common with \( t_1 \). The map \( \tau \) generating \( \sigma = \sigma_\tau \) is defined by a binary choice on every edge in \( T_N([0,1]^d) \), and the rays given by \( t_1 \) and \( t_2 \) agree on their first \( h(u) \) edges, so we have exactly \( N - h(u) \) binary choices to make. This is precisely (7.1).

![Figure 6: Diagram of the sticky assignment between the two rays defining \( t_1, t_2 \in T_N([0,1]^d) \) and the two rays defining their assigned slopes \( \alpha_1, \alpha_2 \in D_M^{[N]} \). The bold edges defining \( t_1 \) are fixed to map to the corresponding bold edges at the same height defining \( \alpha_1 \). This leaves a binary choice to be made at each of the dotted edges along the path between \( D(t_1, t_2) \) and \( t_2 \). We see that \( t_2 \) is assigned the slope \( v_2 \) under \( \sigma \) if and only if these dotted edges are assigned via \( \Phi \circ \psi^{-1} \circ \tau \) to the dotted edges on the ray defining \( \alpha_2 \).](image)

More explicitly, if \( t_1 = \langle i_1, i_2, \cdots, i_N \rangle \) and \( t_2 = \langle j_1, \cdots, j_N \rangle \), then

\[
\langle i_1, \cdots, i_{h(u)} \rangle = \langle j_1, \cdots, j_{h(u)} \rangle. \tag{7.2}
\]

The event of interest may therefore be recast as

\[
\{ \sigma(t_2) = v_2 | \sigma(t_1) = v_1 \} \tag{7.3}
\]
Lemma 7.2. Let \( N \) be a finite subset of \( \mathbb{N} \) such an event is 2 \(^{-c} \) certain consistency requirements on a sticky-admissible collection (7.5); for example, \( \psi \) of point-slope combinations sticky-admissible configurations that are not compatible with stickiness, let us agree to call a collection \( A \# \) with in Lemma 7.2 that will be critical to the proof of (2.5). In all these cases of interest the probability is zero.

Definition 7.3. Let \( I \) be an ordered tuple of four distinct points in \( T_N([0,1]^d) \) of height \( N \) such that \( \#(A), \#(B) \leq 2 \). As is reasonable to expect, the configuration of the root cubes within the tree \( T_N([0,1]^d) \) plays a role in determining \( k(A,B) \). While there is a large number of possible configurations, we isolate certain structures that will turn out to be generic enough for our purposes.

7.1 Four point root configurations

Definition 7.3. Let \( I = \{(t_1,t_2);(t_1',t_2')\} \) be an ordered tuple of four distinct points in \( T_N([0,1]^d) \) of height \( N \) such that

\[
\tag{7.6}
\Pr(\sigma(t) = \gamma(t') \text{ for all } t \in B \mid \sigma(t) = \gamma(t) \text{ for all } t \in A) = \left(\frac{1}{2}\right)^{k(A,B)},
\]

where \( k(A,B) \) is the number of distinct edges in the tree identifying \( B \) that are not common with the tree identifying \( A \). If \( \{(t,\alpha_t) : t \in A' \} \) is not sticky-admissible, then the probability is zero.

For the remainder of this section, we focus on some special events of the form dealt with in Lemma 7.2 that will be critical to the proof of (2.5). In all these cases of interest \( \#(A), \#(B) \leq 2 \). As is reasonable to expect, the configuration of the root cubes within the tree \( T_N([0,1]^d) \) plays a role in determining \( k(A,B) \). While there is a large number of possible configurations, we isolate certain structures that will turn out to be generic enough for our purposes.
We say that $I$ is in type 1 configuration if exactly one of the following conditions is satisfied:

(a) either $u \cap u' = \emptyset$, or

(b) $u' \subset u$, or

(c) $u = u' = D(t_i, t'_i)$ for all $i, j = 1, 2$

If $I$ satisfying (7.6) is not of type 1, we call it of type 2. An ordered tuple $I$ not satisfying the inequality in (7.6) is said to be of type $j = 1, 2$ if $I' = \{(t'_1, t'_2); (t_1, t_2)\}$ is of the same type.

The different structural possibilities are listed in Figure 7. The advantage of a type

| Type 1 Configurations |
|------------------------|
| (a)                    |
| (b)                    |
| (c)                    |
| (d)                    |
| (e)                    |
| (f)                    |

**Figure 7:** All possible four point configurations of type 1, up to permutations.

1 configuration is that, in addition to being overwhelmingly popular, it allows (up to permutations) an easy computation of the quantity $k(A, B)$ described in Lemma 7.2 if $\#(A) = \#(B) = 2$, $A \cup B = \{t_1, t'_1, t_2, t'_2\}$ and $\#(A \cap \{t_1, t_2\}) = \#(B \cap \{t_1, t_2\}) = 1$.

**Lemma 7.4.** Let $I = \{(t_1, t_2); (t'_1, t'_2)\}$ obeying (7.6) be in type 1 configuration. Let $v_i = \gamma(\alpha_i), v'_i = \gamma(\alpha'_i), i = 1, 2$, be (not necessarily distinct) points in $\Omega_N$. Then there exist two permutations $\{i_1, i_2\}$ and $\{j_1, j_2\}$ of $\{1, 2\}$ such that

$$Pr(\sigma(t_{i_2}) = v_{i_2}, \sigma(t'_{j_2}) = v'_{j_2} | \sigma(t_{i_1}) = v_{i_1}, \sigma(t'_{j_1}) = v'_{j_1}) = \left(\frac{1}{2}\right)^{2N - h(u) - h(u')}.$$  

provided the collection $\{(t_i, \alpha_i), (t'_i, \alpha'_i); i = 1, 2\}$ is sticky-admissible. If the admissibility requirement is not met, then the probability is zero.
Proof. The proof is best illustrated by referring to the above diagram, Figure 7(a). If \( u \cap u' = \emptyset \), then any two permutations will satisfy the conclusion of the lemma, Figure 7(b). In particular, choosing \( i_1 = j_1 = 1, \ i_2 = j_2 = 2 \), we see that the number of edges in \( B = \{ t_2, t'_2 \} \) not shared by \( A = \{ t_1, t'_1 \} \) is \( k(A, B) = (N - h(u)) + (N - h(u')) = 2N - h(u) - h(u'). \) The same argument applies if \( u = u' = D(t_i, t'_j) \) for all \( i,j = 1,2 \), Figure 7(b).

We turn to the remaining case where \( u' \subseteq u \). Here there are several possibilities for the relative positions of \( t_1, t_2 \). Suppose first that there is no vertex \( w \) on the ray joining \( u \) and \( u' \) with \( h(u) < h(w) < h(u') \) such that \( w \) is an ancestor of \( t_1 \) or \( t_2 \). This means that the rays of \( t_1, t_2 \) and \( u' \) follow disjoint paths starting from \( u \), so any choice of permutation suffices, Figure 7(c). Suppose next that there is a vertex \( w \) on the ray joining \( u \) and \( u' \) with \( h(u) < h(w) < h(u') \) such that \( w \) is an ancestor of exactly one of \( t_1, t_2 \), but no descendant of \( w \) on this path is an ancestor of either \( t_1 \) or \( t_2 \), Figure 7(d). In this case, we choose \( t_{i_1} \) to be the unique element of \( \{ t_1, t_2 \} \) whose ancestor is \( w \). Note that the ray for \( t_{i_2} \) must have split off from \( u \) in this case. Any permutation of \( \{ t'_1, t'_2 \} \) will then give rise to the desired estimate. If neither of the previous two cases hold, then exactly one of \( \{ t_1, t_2 \} \), say \( t_{i_1} \), is a descendant of \( u' \). If \( u' = D(t_{i_1}, t'_1) \) for both \( j = 1,2 \), then again any permutation of \( \{ t'_1, t'_2 \} \) works, Figure 7(e). Thus the only remaining scenario is where there exists exactly one element in \( \{ t'_1, t'_2 \} \), call it \( t'_{j_1} \), such that \( h(D(t_{i_1}, t'_{j_1})) > h(u') \). In this case, we choose \( A = \{ t_{i_1}, t'_1 \} \) and \( B = \{ t_{i_2}, t'_2 \} \), Figure 7(f). All cases now result in \( k(A, B) = 2N - h(u) - h(u') \), completing the proof. \( \square \)

Lemma 7.5. Let \( I = \{ (t_1, t_2); (t'_1, t'_2) \} \) obey \( 7.6 \) be in type 2 configuration. Then there exist permutations \( \{ i_1, i_2 \} \) and \( \{ j_1, j_2 \} \) of \( \{1,2\} \) for which we have the relations

\[
\begin{align*}
    u_1 \subseteq u, \ u_2 \subseteq u & \text{ with } h(u) \leq h(u_1) \leq h(u_2), \quad \text{where} \\
    u_1 = D(t_{i_1}, t'_{j_1}), \ u_2 = D(t_{i_2}, t'_{j_2}),
\end{align*}
\]

and for which the following equality holds:

\[
Pr(\sigma(t_{i_1}) = v_{i_1}, \sigma(t'_{j_1}) = v'_{j_1} \mid \sigma(t_{i_2}) = v_{i_2}, \sigma(t'_{j_2}) = v'_{j_2}) = \left( \frac{1}{2} \right)^{2N - h(u) - h(u_1)}
\]

for any choice of slopes \( v_1, v'_1, v_2, v'_2 \in \Omega_N \) for which \( \{(t_1, \alpha_i), (t'_1, \alpha'_i); i = 1,2\} \) is sticky-admissible.

Proof. Since \( I \) is of type 2, we know that \( u = u' \), and hence all pairwise youngest common ancestors of \( \{ t_1, t'_1, t_2, t'_2 \} \) must lie within \( u \), but that there exist \( i, j \in \{1,2\} \) such that \( h(D(t_i, t'_j)) > h(u) \). Let us set \( (i_2, j_2) \) to be a tuple for which \( h(D(t_{i_2}, t'_{j_2})) \) is maximal. The height inequalities and containment relations are now obvious, and Figure 8 shows that \( k(A, B) = (N - h(u)) + (N - h(u_1)) \) if \( A = \{ t_{i_2}, t'_{j_2} \} \) and \( B = \{ t_{i_1}, t'_{j_1} \} \). \( \square \)
Type 1 Configurations

\[(a) \quad u = u', \quad u_1 \quad u_2 \quad t_1 \quad t_2 \quad t'_1 \quad t'_2\]

\[(b) \quad u = u', \quad u_1 \quad u_2 \quad t_1 \quad t_2 \quad t'_1 \quad t'_2\]

\[(c) \quad u = u', \quad u_1 \quad u_2 \quad t_1 \quad t_2 \quad t'_1 \quad t'_2\]

Figure 8: All possible four point configurations of type 2, up to permutations.

7.2 Three point root configurations

The arguments in the previous section simplify considerably when there are three root cubes instead of four. Since the proofs here are essentially identical to those presented in Lemmas 7.3 and 7.4, we simply record the necessary facts with the accompanying diagram of Figure 9, leaving their verification to the interested reader.

Definition 7.6. Let \(I = \{(t_1, t_2); (t_1', t_2')\}\) be an ordered tuple of three distinct points in \(T_N([0,1]^d)\) of height \(N\) such that \(h(u) \leq h(u')\), where \(u = D(t_1, t_2)\), \(u' = D(t_1', t_2')\). We say that \(I\) is in type 1 configuration if exactly one of the following two conditions holds:

(a) \(u' \subseteq u\), or

(b) \(u = u' = D(t_2, t_2')\).

Else \(I\) is of type 2, in which case one necessarily has \(u = u'\) and \(u_2 = D(t_2, t_2')\) obeys \(u_2 \subseteq u\). If \(h(u) > h(u')\), then the type \(I\) is the same as that of \(I' = \{(t_1, t_2'); (t_1, t_2)\}\).

Type 1

\[(u, u', t_1, t_2, t'_1, t'_2)\]

Type 2

\[(t_1, t_2, t_2', t'_1, t'_2)\]

Figure 9: Structural possibilities for three point root configurations

Lemma 7.7. Let \(I = \{(t_1, t_2); (t_1', t_2')\}\) be any three-point configuration with \(h(u) \leq h(u')\) in the notation of Definition 7.6, and let \(v_1 = \gamma(\alpha_1), v_2 = \gamma(\alpha_2), v_2' = \gamma(\alpha_2')\) be slopes in
\[ \Omega_N. \text{ Then} \]

\[ \Pr(\sigma(t_2) = v_2, \sigma(t'_2) = v'_2 | \sigma(t_1) = v_1) = \begin{cases} \left(\frac{1}{2}\right)^{2N - h(u) - h(u')} & \text{if } I \text{ is of type 1,} \\
\left(\frac{1}{2}\right)^{2N - h(u) - h(u_2)} & \text{if } I \text{ is of type 2,} \end{cases} \]

provided the point-slope combination \{(t_1, \alpha_1), (t_2, \alpha_2), (t'_2, \alpha'_2)\} is sticky-admissible.

8 Proposition 6.4: Proof of the lower bound (2.5)

If a collection of many thin tubes is to have a large volume, then it is sensible to expect that the intersection of most pairs of tubes should be small. The following measure-theoretic lemma of Bateman and Katz [4] quantifies this phenomenon generally.

Lemma 8.1 ([4], Proposition 2, page 75). Suppose \((X, \mathcal{A}, \mu)\) is a measure space and \(A_1, \ldots, A_n \in \mathcal{A}\) are sets with \(\mu(A_j) = \alpha\) for every \(j\). If

\[ \sum_{i=1}^{n} \sum_{j=1}^{n} \mu(A_i \cap A_j) \leq L, \]

then

\[ \mu \left( \bigcup_{i=1}^{n} A_i \right) \geq \frac{\alpha^2 n^2}{16L}. \]

We defer the proof of this fact to reference [3] or [4]. Using it, we reduce the derivation of inequality (2.5) with the \(a_N\) specified in (6.9) via the following lemma. Throughout this subsection, all probability statements are understood to take place on the probability space \((\Sigma_N, \mathcal{F}(\Sigma_N), \Pr)\) identified in Proposition 6.4.

Proposition 8.2. Fix integers \(N\) and \(R\) with \(N \gg M\) and \(N - \frac{1}{10} \log M N \leq R \leq N - 10\). Define \(P^*_{t,\sigma,R}\) to be the portion of \(P_t,\sigma\) contained in the vertical slab \([M^{R-N}, M^{R+1-N}] \times \mathbb{R}^d\). Then

\[ \mathbb{E}_\sigma \left[ \sum_{t_1 \neq t_2} \left| P^*_{t_1,\sigma,R} \cap P^*_{t_2,\sigma,R} \right| \right] \lesssim N M^{-2N + 2R}, \tag{8.1} \]

where the implicit constant depends only on \(M\) and \(d\).

If one can show that with large probability and for all \(R\) specified in Proposition 8.2, the quantity \(\sum_{t_1 \neq t_2} \left| P^*_{t_1,\sigma,R} \cap P^*_{t_2,\sigma,R} \right|\) is bounded above by the right hand side of (8.1), then Lemma 8.1 would imply (2.5) with \(a_N = \sqrt{\log N}/N\). Unfortunately, (8.1) only shows this on average for every \(R\), and hence is too weak a statement to permit such a conclusion. However, with some additional work we are able to upgrade the statement in Proposition 8.2 to a second moment estimate, given below. While still not as strong as the statement mentioned above, this suffices for our purposes with a smaller choice of \(a_N\).
Proposition 8.3. Under the same hypotheses as Proposition 8.2, there exists a constant $C_{M,d} > 0$ such that
\[
\mathbb{E}_{\sigma} \left[ \left( \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \right)^2 \right] \leq C_{M,d}^2 \left( N M^{-2N+2R} \right)^2. \tag{8.2}
\]

Corollary 8.4. Proposition 8.3 implies (2.5) with $a_N$ as in (6.9).

Proof. Fix a small constant $c_1 > 0$ such that $2 c_1 < \frac{1}{10}$. By Chebyshev’s inequality, (8.2) implies that there exists a large constant $C_{M,d} > 0$ such that for every $R$ with $c_1 \log N \leq N - R \leq 2 c_1 \log N$,
\[
\Pr \left( \left\{ \sigma : \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \geq 2 C_{M,d} N \sqrt{\log N M^{-2N+2R}} \right\} \right)
\leq \frac{\mathbb{E}_{\sigma} \left[ \left( \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \right)^2 \right]}{(2 C_{M,d} N \sqrt{\log N M^{-2N+2R}})^2}
\leq \frac{1}{4 \log N}.
\]

Therefore,
\[
\Pr \left( \bigcup_{N - R = c_1 \log N} \left\{ \sigma : \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \geq C_{M,d} N \sqrt{\log N M^{-2N+2R}} \right\} \right)
\leq \frac{c_1 \log N}{4 \log N} < \frac{1}{4}.
\]

In other words, for a class of $\sigma$ with probability at least $\frac{3}{4}$,
\[
\sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \leq C_{M,d} N \sqrt{\log N M^{-2N+2R}}
\]

for every $N - R \in [c_1 \log N, 2 c_1 \log N]$. For such $\sigma$ and the chosen range of $R$, we apply Lemma 8.1 with $A_t = P_{t, \sigma, R}^*$, $n = M^N d$, for which $\alpha = C_{d} M^{R-N} N^{-d}$, and
\[
\sum_{t_1, t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| = \left[ \sum_{t_1 = t_2} + \sum_{t_1 \neq t_2} \right] \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right|
\leq \alpha n + C_{M,d} N \sqrt{\log N M^{-2N+2R}}
\leq M^{R-N} N \sqrt{\log N M^{-2N+2R}}
\leq N \sqrt{\log N M^{-2N+2R}} =: L.
\]

The last step above uses the specified range of $R$. Lemma 8.1 now yields that
\[
\left| \bigcup_{t} P_{t, \sigma, R}^* \right| \geq \frac{(M^{R-N})^2}{L} \sim \frac{1}{N \sqrt{\log N}}
\]
for every $N - R \in [c_1 \log N, 2c_1 \log N]$. Since $\{ \cup_t P^*_{t,\sigma,R} : R \geq 0 \}$ is a disjoint collection, we obtain

$$|K_N(\sigma) \cap [0,1] \times \mathbb{R}^d| \geq \sum_{R = N - 2c_1 \log N}^{N - c_1 \log N} \left| \bigcup_t P^*_{t,\sigma,R} \right| \gtrsim \log N \frac{1}{N \sqrt{\log N}} = a_N,$$

which is the desired conclusion $\Box$.

8.1 Proof of Proposition 8.2

Thus, we are charged with proving Proposition 8.3. We will prove Proposition 8.2 first, since it involves many of the same ideas as in the proof of the main proposition, but in a simpler setting. We will need to take advantage of several geometric facts, counting arguments and probability estimates prepared in Sections 3 and 7 that will be described shortly. For now, we prescribe the main issues in establishing the bound in (8.1).

Proof. Given $N$ and $R$ as in the statement of the proposition, we decompose the slab $[M^R - N, M^{R+1} - N] \times \mathbb{R}^d$ into thinner slices $Z_k$, where

$$Z_k := \left[ \frac{k}{MN}, \frac{k + 1}{MN} \right] \times \mathbb{R}^d, \quad M^R \leq k \leq M^{R+1} - 1.$$

Setting $P^*_{t,\sigma,k} := P^*_{t,\sigma} \cap Z_k$, we observe that $P^*_{t,\sigma,R}$ is an essentially disjoint union of $\{P^*_{t,\sigma,k}\}$. Since $P^*_{t,\sigma,R}$ is transverse to $Z_k$, we arrive at the estimate

$$\sum_{t_1 \neq t_2} |P^*_{t_1,\sigma,R} \cap P^*_{t_2,\sigma,R}| = \sum_{M^R \leq k \leq M^{R+1}} \sum_{t_1 \neq t_2} |P^*_{t_1,\sigma,k} \cap P^*_{t_2,\sigma,k}| \lesssim M^{-(d+1)N} \sum_{M^R \leq k \leq M^{R+1}} \sum_{t_1 \neq t_2} T_{t_1 t_2}(k) \lesssim M^{-(d+1)N} \sum_{M^R \leq k \leq M^{R+1}} \sum_{u \in T_N((0,1)^d)} \sum_{h(u) < N} T_{t_1 t_2}(k),$$

where $T_{t_1 t_2}(k)$ is a random variable that equals one if $P^*_{t_1,\sigma,k} \cap P^*_{t_2,\sigma,k} \neq \emptyset$, and is zero otherwise. At the last step in the above string of inequalities, we have further stratified the sum in $(t_1, t_2)$ in terms of their youngest common ancestor $u = D(t_1, t_2)$ in the tree $T_N((0,1)^d)$, with the index set $S_u$ of the innermost sum being defined by

$$S_u := \left\{ (t_1, t_2) : t_1, t_2 \in T_N((0,1)^d), h(t_1) = h(t_2) = N, \ D(t_1, t_2) = u \right\}.$$

We will prove below in Lemma 8.8 that

$$\mathbb{E}_{\sigma} \left[ \sum_{(t_1, t_2) \in S_u} T_{t_1 t_2}(k) \right] \lesssim M^{R-N} M^{-dh(u)+Nd} = M^{R-dh(u)+N(d-1)}.$$
Plugging this expected count into the last step of (8.4) and simplifying, we obtain
\[
\mathbb{E}_\sigma \left[ \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \right] \lesssim \sum_{M^R \leq k < M^{R+1} - 1} M^{R-2N} \sum_{u \in T_N([0,1]^d)} M^{-dh(u)}
\]
\[
\lesssim \sum_{M^R \leq k < M^{R+1} - 1} M^{R-2N} N \lesssim N M^{2R-2N},
\]
which is the estimate claimed by Proposition 8.2. At the penultimate step, we have used the fact that there are \(M^{dr}\) vertices \(u\) in \(T_N([0,1]^d)\) of height \(r\), resulting in
\[
\sum_u M^{-dh(u)} = \sum_{0 \leq r < N} M^{-dr} M^{dr} = N.
\]

(8.6)

8.2 Proof of Proposition 8.3

Proof. To establish (8.2), we take a similar route, with some extra care in summing over the (now more numerous) indices. Squaring the expression in (8.3), we obtain
\[
\left[ \sum_{t_1 \neq t_2} \left| P_{t_1, \sigma, R}^* \cap P_{t_2, \sigma, R}^* \right| \right]^2 \leq M^{-2(d+1)N} \sum_{k,k' \in [M^R,M^{R+1})} \sum_{t_1 \neq t_2, t'_1 \neq t'_2} \sum T_{t_1, t_2}(k) T_{t'_1, t'_2}(k')
\]
\[
\leq \mathcal{G}_2 + \mathcal{G}_3 + \mathcal{G}_4,
\]
where the index \(i\) in \(\mathcal{G}_i\) corresponds to the number of distinct points in the tuple \(\{(t_1, t_2); (t'_1, t'_2)\}\). More precisely, for \(i = 2, 3, 4\),
\[
\mathcal{G}_i := M^{-2(d+1)N} \sum_{k,k' \in [M^R,M^{R+1})} \sum_{t_1 \neq t_2, t'_1 \neq t'_2} \sum T_{t_1, t_2}(k) T_{t'_1, t'_2}(k'), \quad \text{where}
\]
\[
\mathcal{I}_i := \left\{ \mathcal{I} = \{(t_1, t_2); (t'_1, t'_2)\} \mid t_j, t'_j \in T_N([0,1]^d), h(t_j) = h(t'_j) = N \forall j = 1, 2, \right. \left. t_1 \neq t_2, t'_1 \neq t'_2, \#(\{(t_1, t_2); (t'_1, t'_2)\}) = i \right\}. \quad (8.9)
\]
The main contribution to the left hand side of (8.2) will be from \(\mathbb{E}_\sigma(\mathcal{G}_4)\), and we will discuss its estimation in detail. The other terms, whose treatment will be briefly sketched, will turn out to be of smaller size.

We decompose \(\mathcal{I}_4 = \mathcal{I}_{41} \cup \mathcal{I}_{42}\), where \(\mathcal{I}_{4j}\) is the collection of 4-tuples of distinct points \(\{(t_1, t_2); (t'_1, t'_2)\}\) that are in configuration of type \(j = 1, 2\), as explained in Definition 7.3. This results in a corresponding decomposition \(\mathcal{G}_4 = \mathcal{G}_{41} + \mathcal{G}_{42}\). For \(\mathcal{G}_{41}\), we further stratify the sum in terms of \(u = D(t_1, t_2)\) and \(u' = D(t'_1, t'_2)\), where we may assume without loss of generality that \(h(u) \leq h(u')\). Thus,
\[
\mathbb{E}_\sigma(\mathcal{G}_{41}) = \sum_{k,k' \in [M^R,M^{R+1})} \sum_{u,u' \in T_N([0,1]^d)} \mathbb{E}_\sigma(\mathcal{G}_{41}(u, u'; k, k')) \quad \text{where}
\]
(8.10)
\( \mathcal{G}_{41}(u, u'; k, k') := M^{-2(d+1)N} \sum_{I \in \mathcal{J}_{41}(u, u')} T_{t_1t_2}(k)T_{t_1't_2'}(k'), \) and

\[ \mathcal{J}_{41}(u, u') := \{ I \in \mathcal{J}_{41} : u = D(t_1, t_2), u' = D(t_1', t_2') \}. \]

In Lemma 8.9 below, we will show that

\[ \mathbb{E}_\sigma[\mathcal{G}_{41}(u, u'; k, k')] \lesssim M^{-2(d+1)N} M^{2R-d(h(u)+h(u'))+2N(d-1) - 2R-4N-d(h(u)+h(u'))}. \]  \tag{8.11}

Inserting this back into (8.10), we now follow the same summation steps that led to (8.1) from (8.5). Specifically, applying (8.6) twice, we obtain

\[ \mathbb{E}_\sigma(\mathcal{G}_{41}) \lesssim M^{2R-4N} \sum_{k,k'} \sum_{u,u'} M^{-d(h(u)+h(u'))} \lesssim \sum_{k,k'} N^2 M^{2R-4N} \lesssim N^2 M^{4R-4N}, \]

which is the right hand side of (8.2).

Next we turn to \( \mathcal{G}_{42} \). Motivated by the configuration type, and after permutation of \( \{t_1, t_2\} \) and of \( \{t_1', t_2'\} \) if necessary (so that the conclusion of Lemma 7.5 holds), we stratify this sum in terms of \( u = u' = D(t_1, t_2) = D(t_1', t_2'), u_1 = D(t_1, t_1'), u_2 = D(t_2, t_2') \), writing

\[ \mathcal{G}_{42} = \sum_{k,k'} \sum_{u,u_1,u_2 \in T_N((0,1)^d)} \mathcal{G}_{42}(u, u_1, u_2; k, k'), \]

where

\[ \mathcal{G}_{42}(u, u_1, u_2; k, k') := M^{-2(d+1)N} \sum_{I \in \mathcal{J}_{42}(u, u_1, u_2)} T_{t_1t_2}(k)T_{t_1't_2'}(k'), \]

\[ \mathcal{J}_{42}(u, u_1, u_2) := \{ I \in \mathcal{J}_{42} : u = D(t_1, t_2) = D(t_1', t_2'), u_1 = D(t_1, t_1'), u_2 = D(t_2, t_2') \}. \]  \tag{8.12}

for given \( u_1, u_2 \subseteq u \) with \( h(u) \leq h(u_1) \leq h(u_2) \). For such \( u, u_1, u_2 \), we will prove in Lemma 8.10 below that

\[ \mathbb{E}_\sigma(\mathcal{G}_{42}(u, u_1, u_2; k, k')) \lesssim M^{-2N-2dh(u_2)}. \]  \tag{8.13}

Accepting this estimate for the time being, we complete the estimation of \( \mathbb{E}_\sigma(\mathcal{G}_{42}) \) as follows,

\[ \mathbb{E}_\sigma(\mathcal{G}_{42}) \lesssim \sum_{k,k'} \sum_{u,u_1,u_2} M^{-2N-2dh(u_2)} \lesssim M^{-2N} \sum_{k,k'} \sum_{u \subseteq u} M^{-2dh(u_2)} \sum_{u \subseteq u} \sum_{h(u) \leq h(u)} 1 \]

40
\[ M^{-2N} \sum_{k,k' \in \mathbb{Z}} \sum_{u \leq u_2} M^{-2dh(u_2)} \left[ M^{d(h(u_2) - h(u))} \right] \]  
(8.14)

\[ M^{-2N} \sum_{k,k' \in \mathbb{Z}} \sum_{u} M^{-dh(u)} \sum_{u_2 \leq u} M^{-2dh(u_2)} \]  
(8.15)

\[ NM^{-2N} \sum_{k,k' \in \mathbb{Z}} \sum_{u} M^{-2dh(u)} \]  
(8.16)

For the range \( N - R \leq 1/2 \log_M N \) assured by Proposition 8.3, the last quantity above is smaller than \( (NM^{2R-2N})^2 \). The string of inequalities displayed above involve repeated applications of the fact used to prove (8.6), namely here are \( M^d j - dh(u) \) cubes of sidelength \( M^{-j} \) contained in \( u \). Thus the estimates

\[ \sum_{u_1 \leq u} 1 \leq \sum_{j = h(u)} M^{d(j - h(u))} \lesssim M^{d(h(u_2) - h(u))}, \]

\[ \sum_{u_2 \leq u} M^{-dh(u_2)} \lesssim \sum_{N \geq j \geq h(u)} M^{-dj} M^{d(j - h(u))} \lesssim NM^{-dh(u)}, \]  
and

\[ \sum_{u} M^{-2dh(u)} = \sum_{j=0}^{N} M^{dj} M^{-2dj} \]  
(8.17)

\[ \sum_{j=0}^{N} M^{dj} \lesssim 1 \]

were used in (8.14), (8.15) and (8.16) respectively, completing the estimation of \( \mathbb{E}(\mathcal{S}_4) \).

Arguments similar to and in fact simpler than those above lead to the following estimates for \( \mathbb{E}(\mathcal{S}_3) \) and \( \mathbb{E}(\mathcal{S}_2) \):

\[ \mathbb{E}(\mathcal{S}_3) = \mathbb{E}(\mathcal{S}_{31}) + \mathbb{E}(\mathcal{S}_{32}) \]

\[ \lesssim NM^{3R-3N} + M^{3R-3N} \lesssim NM^{3R-3N}, \]  
and

\[ \mathbb{E}(\mathcal{S}_2) \lesssim NM^{3R-(d+3)N}. \]  
(8.18)

Here without loss of generality and after a permutation if necessary, we have assumed that \( I = \{(t_1, t_2); (t_1, t'_2)\} \in \mathcal{I}_2 \), with \( h(D(t_1, t_2)) \leq h(D(t_1, t'_2)) \). The subsum \( \mathcal{S}_{3i} \) then corresponds to tuples \( I \) that are in type \( i \) configuration in the sense of Definition 7.6. There is only one possible configuration of pairs in \( \mathcal{I}_2 \). The derivation of the expectation estimates (8.17) and (8.18) closely follow the estimation of \( \mathcal{S}_4 \), with appropriate adjustments in the probability counts; for instance, (8.17) uses Lemma 8.7 and (8.18) uses Lemma 8.11. To avoid repetition, we leave the details of (8.17) and (8.18) to the reader, noting that the right hand term in each case is dominated by \( (NM^{2R-2N})^2 \) by our conditions on \( R \). \( \square \)

### 8.3 Expected intersection counts

It remains to establish (8.9), (8.11) and (8.13). The necessary steps for this are laid out in the following sequence of lemmas. Unless otherwise stated, we will be using the
Lemma 8.5. Fix $Z_k$. Let us define $A_u = A_u(k)$ to be the (deterministic) collection of all $t_1 \in T_N([0,1]^d)$, $h(t_1) = N$ that are contained in the cube $u$ and whose distance from the boundary of some child of $u$ is $\leq kM^{-N-h(u)}$.

For $t_1 \in A_u$, let $B_{t_1} = B_{t_1}(k)$ denote the (also deterministic) collection of $t_2 \in T_N([0,1]^d)$ with $h(t_2) = N$ and $D(t_1,t_2) = u$ such that the distance between the centres of $t_1$ and $t_2$ is $\leq kM^{-N-h(u)}$.

(a) Then for any slope assignment $\sigma$, the random variable $T_{t_1t_2}(k) = 0$ unless $t_1 \in A_u$ and $t_2 \in B_{t_1}$. In other words,

$$\sum_{(t_1,t_2)\in S_u} T_{t_1t_2}(k) = \sum_{t_1 \in A_u} \sum_{t_2 \in B_{t_1}} T_{t_1t_2}(k),$$

so that

$$\mathbb{E}_\sigma\left[\sum_{(t_1,t_2)\in S_u} T_{t_1t_2}(k)\right] = \sum_{t_1 \in A_u} \mathbb{E}_\sigma\left[\sum_{t_2 \in B_{t_1}} T_{t_1t_2}(k)\right]. \quad (8.19)$$

(b) The description of $A_u$ yields the following bound on its cardinality:

$$\#(A_u) \lesssim \left(\frac{k}{MN}\right) M^{d(N-h(u))} \lesssim M^{R-dh(u)+(d-1)N}.$$

\[\text{Figure 10: A diagram of } A_u \text{ when } d = 2, M = 3. \] Here the largest square is $u$. The shaded area depicts $A_u$. The finest squares are the root cubes contained in $A_u$. \]
Proof. We observe that $T_{t_1 t_2}(k) = 1$ if and only if there exists a point $p = (p_1, \ldots, p_{d+1}) \in Z_k$ and $v_1, v_2 \in \Omega_N$ such that $p \in \mathcal{P}_{t_1, v_1} \cap \mathcal{P}_{t_2, v_2}$, and $\sigma(t_1) = v_1, \sigma(t_2) = v_2$. By Lemma 3.1 this implies that

$$|\operatorname{cen}(t_1) - \operatorname{cen}(t_2) + p_1(\sigma(t_1) - \sigma(t_2))| \leq 2\kappa_d \sqrt{d} M^{-N},$$

(8.20)

where $\operatorname{cen}(t_i)$ denotes the centre of the cube $t_i$. For $p \in Z_k$, (8.20) yields

$$|\operatorname{cen}(t_1) - \operatorname{cen}(t_2)| \leq p_1|\sigma(t_1) - \sigma(t_2)| + 2\kappa_d \sqrt{d} M^{-N} \lesssim p_1|\sigma(t_1) - \sigma(t_2)|$$

$$\lesssim \left(\frac{k + 1}{M^N}\right)|\sigma(t_1) - \sigma(t_2)| \lesssim \left(\frac{k}{M^N}\right) M^{-h(D(\tau(t_1), \tau(t_2)))}$$

$$\lesssim kM^{-N-h(u)}.$$  

(8.21)

The second inequality in the steps above follows from Corollary 3.2, the third from the definition of $Z_k$ and the fourth from the property (6.2) of the slope assignment. Here $\tau$ is the unique sticky map that generates $\sigma$, as specified in Proposition 6.2. Since $\tau$ preserves heights and lineages, $h(D(\tau(t_1), \tau(t_2))) \geq h(D(t_1, t_2)) = h(u)$, and the last step follows.

The inequality in (8.21) implies that $T_{t_1 t_2}(k) = 0$ unless $t_2 \in B_{t_1}$. Further, $t_1, t_2$ lie in distinct children of $u$, so $t_1$ must satisfy

$$\operatorname{dist}(t_1, \partial u') \lesssim \frac{k}{M^N} M^{-h(u)}$$

for some child $u'$ of $u$.

To allow for the existence of some $t_2$ obeying (8.21). This means $t_1 \in A_u$, proving (a).

For (b) we observe that $u$ has $M^d$ children. The Lebesgue measure of the set

$$\bigcup_{u'} \left\{x \in u' : \operatorname{dist}(x, \partial u') \lesssim kM^{-N-h(u)}, u' \text{ is a child of } u \right\}$$

(8.22)

is therefore $\lesssim (M^d)kM^{-N-h(u)} M^{-(d-1)h(u)}$. The cardinality of $A_u$ is comparable to the number of $M^{-N}$-separated points in the set (8.22), and (b) follows.

Our next task is to make further reductions to the expression on the right hand side of (8.19) that will enable us to invoke the probability estimates from Section 7. To this end, let us fix $Z_k, t_1 \in A_u(k), v_1 = \gamma(\alpha_1) \in \Omega_N$, and define a collection of point-slope pairs

$$\mathcal{E}_u(t_1, v_1; k) := \left\{(t_2, v_2) : t_2 \in \mathcal{T}_N([0, 1]^d) \cap B_{t_1}, v_2 = \gamma(\alpha_2) \in \Omega_N, h(t_2) = h(\alpha_2) = N, u = D(t_1, t_2), \mathcal{P}_{t_1, v_1} \cap \mathcal{P}_{t_2, v_2} \cap Z_k \neq \emptyset, h(D(\alpha_1, \alpha_2)) \geq h(u) \right\}.$$  

(8.23)

Thus $\mathcal{E}_u(t_1, v_1; k)$ is non-random as well. The significance of this collection is clarified in the next lemma.
Lemma 8.6. For \((t_2, v_2) \in \mathcal{E}_u(t_1, v_1; k)\) described as in (8.23), define a random variable \(T_{t_2v_2}(t_1, v_1; k)\) as follows:

\[
T_{t_2v_2}(t_1, v_1; k) := \begin{cases} 
1 & \text{if } \sigma(t_2) = v_2, \\
0 & \text{otherwise.}
\end{cases}
\]  

(8.24)

(a) The random variables \(T_{t_1t_2}(k)\) and \(T_{t_2v_2}(t_1, v_1; k)\) are related as follows: given \(\sigma(t_1) = v_1\),

\[
T_{t_1t_2}(k) = \sup \{T_{t_2v_2}(t_1, v_1; k) : (t_2, v_2) \in \mathcal{E}_u(t_1, v_1; k)\}.
\]  

(8.25)

In particular under the same conditional hypothesis \(\sigma(t_1) = v_1\), one obtains the bound

\[
T_{t_1t_2}(k) \leq \sum_{(t_2,v_2) \in \mathcal{E}_u(t_1,v_1;k)} T_{t_2v_2}(t_1,v_1;k),
\]  

(8.26)

which in turn implies

\[
\mathbb{E}_\sigma \left[ \sum_{t_2 \in B_1} T_{t_1t_2}(k) \big| \sigma(t_1) = v_1 \right] \leq \sum_{(t_2,v_2) \in \mathcal{E}_u(t_1,v_1;k)} \mathbb{P}(\sigma(t_2) = v_2 \big| \sigma(t_1) = v_1).
\]  

(8.27)

(b) The cardinality of \(\mathcal{E}_u(t_1, v_1; k)\) is \(\lesssim 2^{N-h(u)}\).

Proof. We already know from Lemma 8.5 that \(T_{t_1t_2}(k) = 0\) unless \(t_2 \in B_1\). Further, if \(\sigma(t_1) = v_1\) is known, then it is clear that \(T_{t_1t_2}(k) = 1\) if and only if there exists \(v_2 \in \Omega_N\) such that \(P_{t_1,v_1} \cap P_{t_2,v_2} \cap Z_k \neq \emptyset\) and \(\sigma(t_2) = v_2\). But this means that the sticky map \(\tau\) that generates \(\sigma\) must map \(t_2\) to the \(N\)-long binary sequence that identifies \(\alpha_2\). Stickiness dictates that \(h(D(\alpha_1, \alpha_2)) = h(D(\tau(t_1), \tau(t_2))) \geq h(D(t_1, t_2)) = h(u)\), explaining the constraints that define \(\mathcal{E}_u(t_1, v_1; k)\). Rephrasing the discussion above, given \(\sigma(t_1) = v_1\), the event \(T_{t_1t_2}(k) = 1\) holds if and only if there exists \(v_2 \in \Omega_N\) such that \(\sigma(t_2) = v_2\). This is the identity claimed in (8.25) of part (a). The bound in (8.26) follows easily from (8.25) since the supremum is dominated by the sum. The final estimate (8.27) in part (b) follows by taking conditional expectation of both sides of (8.26), and observing that \(\mathbb{E}_\sigma(T_{t_2v_2}(t_1,v_1;k) | \sigma(t_1) = v_1) = \mathbb{P}(\sigma(t_2) = v_2 | \sigma(t_1) = v_1)\).

We turn to (b). If \(v_2 \in \Omega_N\) is fixed, then it follows from Corollary 4.3 (taking \(Q\) in that corollary to be the cube of sidelength \(O(M^{-N})\) containing \(P_{t_1,v_1} \cap Z_k\) that there exist at most a constant number of choices of \(t_2\) such that \(t_2, v_2 \in \mathcal{E}_u(t_1, v_1; k)\). But by Corollary 4.3 the number of points \(\alpha_2 \in D^{|N|}_M\) (and hence slopes \(v_2 \in \Omega_N\) that obey \(h(D(\alpha_1, \alpha_2)) \geq h(u)\) is no more than \(2^{N-h(u)}\), proving the claim.

The same argument above applied twice yields the following conclusion, the verification of which is left to the reader.
Corollary 8.7. Given $t_1 \in A_u(k)$, $t_1' \in A_u(k')$, $v_1, v_1' \in \Omega_N$, define $E_u(t_1, v_1; k)$ and $E_{u'}(t_1', v_1'; k')$ as in (8.23) and the random variables $T_{t_1 v_2}(t_1, v_1; k)$, $T_{t_1' v_2'}(t_1', v_1'; k')$ as in (8.24). Then given $\sigma(t_1) = v_1$ and $\sigma(t_1') = v_1'$,

$$\sum_{t_2 \in B_1} T_{t_1 t_2}(k) T_{t_1' v_2'}(k') \leq \sum_{t_2 \in B_1} T_{t_2 v_2}(t_1, v_1; k) T_{t_1' v_2'}(t_1', v_1'; k'),$$

where the notation $\sum^*$ represents the sum over all indices $\{(t_2, v_2); (t_1', v_1')\} \in E_u(t_1, v_1; k) \times E_{u'}(t_1', v_1'; k').$

We are now ready to establish the key estimates in the proofs of Propositions 8.2 and 8.3.

Lemma 8.8. The estimate in (8.5) holds.

Proof. We combine the steps outlined in Lemmas 8.5, 8.6 and 7.1. By Lemma 8.5(a),

$$E_{\sigma} \left[ \sum_{(t_1, t_2) \in S_u} T_{t_1 t_2}(k) \right] = \sum_{t_1 \in A_u} E_{\sigma} \left[ \sum_{t_2 \in B_1} T_{t_1 t_2}(k) \right] \leq \sum_{t_1 \in A_u} E_{\sigma} \left[ \sum_{t_2 \in B_1} T_{t_1 t_2}(k) \sigma(t_1) = v_1 \right].$$

(8.28)

Applying (8.27) from Lemma 8.6 followed by Lemma 7.1, we find that the inner expectation above obeys the bound

$$E_{\sigma} \left[ \sum_{t_2 \in B_1} T_{t_1 t_2}(k) \sigma(t_1) = v_1 \right] \leq \sum_{(t_2, v_2) \in E_u(t_1, v_1; k)} \Pr(\sigma(t_2) = v_2 | \sigma(t_1) = v_1) \leq \#(E_u(t_1, v_1; k)) \times 2^{-N + h(u)} \leq 2^{N - h(u)} \times 2^{-N + h(u)} \lesssim 1,$$

uniformly in $v_1$. Inserting this back into (8.28), we arrive at

$$E_{\sigma} \left[ \sum_{(t_1, t_2) \in S_u} T_{t_1 t_2}(k) \right] \lesssim \#(A_u),$$

which according to Lemma 8.5(b) is the bound claimed in (8.5).

Lemma 8.9. The estimate in (8.11) holds.

Proof. The proof of (8.11) shares many similarities with that of Lemma 8.8, except that there are now two copies of each of the objects appearing in the proof of (8.5) and the
probability estimate comes from Lemma 7.4 instead of Lemma 7.1. We outline the main steps below.

In view of Lemma 7.4 and after a permutation of \((t_1, t_2)\) and of \((t'_1, t'_2)\) if necessary, we may assume that for every \(I = \{(t_1, t_2); (t'_1, t'_2)\} \in \mathcal{J}_{41}(u, u')\),

\[
\Pr(\sigma(t_2) = v_2, \sigma(t'_2) = v'_2 | \sigma(t_1) = v_1, \sigma(t'_1) = v'_1) = \left(\frac{1}{2}\right)^{2N-h(u)-h(u')}.
\]

(8.29)

Now,

\[
\mathbb{E}_s(\mathcal{S}_{41}(u, u'; k, k') \leq M^{-2(d+1)} \mathbb{E}_s \left[ \sum_{I \in \mathcal{J}_{41}(u, u')} T_{t_1t_2}(k)T_{t'_1t'_2}(k') \right]
\]

\[
= M^{-2(d+1)} \sum_{t_1 \in \mathcal{A}_s(k)} \mathbb{E}_{v_1,v'_1} \mathbb{E}_s \left[ \sum_{t_2 \in \mathcal{B}_1} T_{t_1t_2}(k)T_{t'_1t'_2}(k') | \sigma(t_1) = v_1, \sigma(t'_1) = v'_1 \right]
\]

\[
\leq M^{-2(d+1)} \left( \frac{kk'}{M^{2N}} M^{d(2N-h(u)-h(u'))} \right) \leq M^{2R-4N-d(h(u)+h(u'))},
\]

since according to Corollary 8.7,

\[
\mathbb{E}_s \left[ \sum_{(t_2, t'_2) \in \mathcal{B}_1 \times \mathcal{B}_1'} T_{t_1t_2}(k)T_{t'_1t'_2}(k') | \sigma(t_1) = v_1, \sigma(t'_1) = v'_1 \right]
\]

\[
\leq \mathbb{E}_s \left[ \sum_{(t_2, t'_2) \in \mathcal{B}_1 \times \mathcal{B}_1'} T_{t_2t'_2}(t_1, v_1; k)T_{t'_2t'_2}(t'_1, v'_1; k') | \sigma(t_1) = v_1, \sigma(t'_1) = v'_1 \right]
\]

\[
\leq \sum \Pr(\sigma(t_2) = v_2, \sigma(t'_2) = v'_2 | \sigma(t_1) = v_1, \sigma(t'_1) = v'_1)
\]

\[
\leq \left( \frac{2^{N-h(u)}}{\#(\mathcal{E}(t_1, v_1; k))} \right) \times \left( \frac{2^{N-h(u')}}{\#(\mathcal{E}(t_1, v'_1; k'))} \right) \times \left( \frac{2^{-2N-h(u)+h(u')}}{\#(\mathcal{E}(t'_1, v'_1; k'))} \right)
\]

\[
\leq 1, \quad \text{uniformly in } v_1, v'_1.
\]

The proof is therefore complete. \(\square\)

**Lemma 8.10.** The estimate in (8.13) holds.

*Proof.* The proof of (8.13) is similar to (8.11), and in certain respects simpler. But the configuration type dictates that we set up a different class \(\mathcal{E}^*\) of point-slope tuples that will play a role analogous to \(\mathcal{E}(t_1, v_1; k)\) in the preceding lemmas. Recall the structure of a type 2 configuration from Figure 8 and the definition of \(\mathcal{J}_{42}(u, u_1, u_2)\) from (8.12). Given root cubes \(t_2, t'_2\), and \(u, u_1, u_2 \in \mathcal{T}_N([0, 1]^d)\) with the property that

\[
u_1 \subseteq u, \quad u_2 \subset u, \quad u_2 = D(t_2, t'_2), \quad h(u) \leq h(u_1) \leq h(u_2) \leq N = h(t_2) = h(t'_2),\]

46
and slopes $v_2 = \gamma(\alpha_2), v_2' = \gamma(\alpha_2') \in \Omega_N$, we define $E^*$ (depending on all these objects) to be the following collection of root-slope tuples:

\[
E^* := \left\{ (t_1, v_1); (t_1', v_1') \mid \begin{array}{l}
\| = \{(t_1, t_2); (t_1', t_2') \in \mathcal{I}_{42}(u, u_1, u_2), \\
v_1 = \gamma(\alpha_1), v_1' = \gamma(\alpha_1') \text{ for some } \alpha_1, \alpha_1' \in D^N_M, \\
\mathcal{P}_{t_1, v_1} \cap \mathcal{P}_{t_2, v_2} \cap Z_k \neq \emptyset, \mathcal{P}_{t_1', v_1'} \cap \mathcal{P}_{t_2', v_2'} \cap Z_{k'} \neq \emptyset,
\end{array}
\right\} (8.30)
\]

The relevance of $E^*$ is this: if $\sigma(t_2) = v_2$ and $\sigma(t_2') = v_2'$ are given, then $T_{t_1, t_2}(k)T_{t_1', t_2'}(k') = 0$ unless there exist $v_1, v_1' \in \Omega_N$ with $\{(t_1, v_1); (t_1', v_1')\} \in E^*$ and $\sigma(t_1) = v_1, \sigma(t_1') = v_1'$.

We first set about obtaining a bound on the size of $E^*$ that we will need momentarily. Stickiness dictates that $h(D(\alpha_1, \alpha_2)) \geq h(u)$, and that $\alpha_1$ is an $N$th level descendant of $\alpha$, the ancestor of $\alpha_2$ at height $h(u)$. Thus the number of possible $\alpha_1$ (and hence $v_1$) is $\leq 2^{N-h(u)}$, by Corollary 4.3. Again by stickiness, $h(D(\alpha_1, \alpha_1')) \geq h(u_1)$, so for a given $\alpha_1$, the number of $\alpha_1'$ (hence $v_1'$) is no more than the number of possible descendants of $\alpha^*$, the ancestor of $\alpha_1$ at height $h(u_1)$. This number is thus $\leq 2^{N-h(u_1)}$. Once $v_1, v_1'$ have been fixed (recall that $v_2, v_2', t_2, t_2'$ are already fixed), it follows from Corollary 5.3 that the number of $t_1, t_1'$ obeying the intersection conditions in (8.30) is $\leq 1$. Combining these, we arrive at the following bound on the cardinality of $E^*$:

\[
\#(E^*) \lesssim (2^{N-h(u)}) (2^{N-h(u_1)}) = 2^{2N-h(u)-h(u_1)} .
\]

(8.31)

We use this bound on the size of $E^*$ to estimate a conditional expectation, essentially the same way as in the previous two lemmas.

\[
\mathbb{E}_\sigma \left[ \sum_{\sigma(t_1) = v_1, \sigma(t_1') = v_1'} T_{t_1, t_2}(k)T_{t_1', t_2'}(k') | \sigma(t_2) = v_2, \sigma(t_2') = v_2' \right]
\]

\[
= \sum_{E^*} \mathbb{P}(\sigma(t_1) = v_1, \sigma(t_1') = v_1' | \sigma(t_2) = v_2, \sigma(t_2') = v_2') \lesssim \#(E^*) \left( \frac{1}{2} \right)^{2N-h(u)-h(u_1)} \lesssim 1 ,
\]

(8.32)

where the last step follows by combining Lemma 7.5 with (8.31). As a result, we obtain

\[
\mathbb{E}_\sigma(\mathcal{S}_{42}(u, u_1, u_2; k, k'))
\]

\[
= M^{-2(d+1)N} \mathbb{E}_\sigma \left[ \sum_{t_1, t_1' \in \mathcal{I}_{42}(u, u_1, u_2)} T_{t_1, t_2}(k)T_{t_1', t_2'}(k') \right]
\]

\[
\leq M^{-2(d+1)N} \sum_{t_2, t_2' \subseteq u_2} \mathbb{E}_{v_2, v_2'} \mathbb{E}_\sigma \left[ \sum_{t_1, t_1'} T_{t_1, t_2}(k)T_{t_1', t_2'}(k') | \sigma(t_2) = v_2, \sigma(t_2') = v_2' \right]
\]

\[
\lesssim M^{-2(d+1)N} \sum_{t_2, t_2' \subseteq u_2} 1
\]

47
\[ M^{-2(d+1)N} (M^{-dh(u_2)+Nd})^2, \]

where the estimate from \( \text{(8.32)} \) has been inserted in the third step above. The final expression is the bound claimed in \( \text{(8.13)} \).

\section{Proposition 6.4: Proof of the upper bound \( (2.6) \)}

Using the theory developed in Section 5, we can establish inequality \( (2.6) \) with \( b_N = C_M/N \) as in Proposition 6.4 with relative ease. For \( x \in \mathbb{R}^{d+1} \), we write \( x = (x_1, \bar{x}), \) where \( \bar{x} = (x_2, \ldots, x_{d+1}). \) Since the Kakeya-type set defined by \( (6.8) \) is contained in the parallelepiped \( [C_0, C_0 + 1] \times [-2C_0, 2C_0]^d \), we may write

\[
\mathbb{E}_\sigma \left[ K_N(\sigma) \cap [C_0, C_0 + 1] \times \mathbb{R}^d \right] = \mathbb{E}_\sigma \left( \int_{C_0}^{C_0+1} \int_{[-2C_0,2C_0]^d} 1_{K_N(\sigma)}(x_1, \bar{x}) d\bar{x} dx_1 \right)
= \int_{C_0}^{C_0+1} \int_{[-2C_0,2C_0]^d} \mathbb{E}_\sigma \left( 1_{K_N(\sigma)}(x_1, \bar{x}) \right) d\bar{x} dx_1
= \int_{C_0}^{C_0+1} \int_{[-2C_0,2C_0]^d} \mathbb{P}(x) d\bar{x} dx_1, \tag{9.1}
\]

where \( \mathbb{P}(x) \) denotes the probability that the point \((x_1, \bar{x})\) is contained in the set \( K_N(\sigma) \). To establish inequality \( (2.6) \) then, it suffices to show that this probability is bounded by a constant multiple of \( 1/N \), the constant being uniform in \( x \in [C_0, C_0 + 1] \times \mathbb{R}^d \).

Let us recall the definition of \( \text{Poss}(x) \) from \( \text{(3.3)} \). We would like to define a certain percolation process on the tree \( T_N(\text{Poss}(x)) \) whose probability of survival can majorize \( \mathbb{P}(x) \). By Lemma 5.7(a), there corresponds to every \( t \in \text{Poss}(x) \) exactly one \( v(t) \in \Omega_N \) such that \( P_{t,v(t)} \) contains \( x \). Let us also recall that \( v(t) = \gamma(\alpha(t)) \) for some \( \alpha(t) \in D_M^{[N]} \). By Corollary 4.3, \( \alpha(t) \) is uniquely identified by \( \beta(t) := \psi(\alpha(t)) \), which is a deterministic sequence of length \( N \) with entries 0 or 1. Here \( \psi \) is the tree isomorphism described in Lemma 4.2.

Given a slope assignment \( \sigma = \sigma_r \) generated by a sticky map \( \tau : T_N([0,1]^d) \rightarrow T_N([0,1]; 2) \) as defined in Proposition 6.2 and a vertex \( t = \langle i_1, \ldots, i_N \rangle \in T_N(\text{Poss}(x)) \) with \( h(t) = N \), we assign a value of 0 or 1 to each edge of the ray identifying \( t \) as follows. Let \( e \) be the edge identified by the vertex \( \langle i_1, i_2, \ldots, i_k \rangle \). Set

\[
Y_e := \begin{cases} 1 & \text{if } \pi_k(\tau(t)) = \pi_k(\beta(t)), \\ 0 & \text{if } \pi_k(\tau(t)) \neq \pi_k(\beta(t)). \end{cases} \tag{9.2}
\]

To clarify the notation above, recall that both \( \tau(t) \) and \( \beta(t) \) are \( N \)-long binary sequences, and \( \pi_k \) denotes the \( k \)th component of the input. Though the definition of \( Y_e \) suggests a potential conflict for different choices of \( t \), our next lemma confirms that this is not the case.

\textbf{Lemma 9.1.} The description in \( \text{(9.2)} \) is consistent in \( t \); i.e., it assigns a uniquely defined binary random variable \( Y_e \) to each edge of \( T_N(\text{Poss}(x)) \). The collection \( \{Y_e\} \) is independent and identically distributed as Bernoulli(\( \frac{1}{2} \)) random variables.
Proof. Let \( t, t' \in \mathcal{T}_N(\text{Poss}(x)) \), \( h(t) = h(t') = N \). Set \( u = D(t,t') \), the youngest common ancestor of \( t \) and \( t' \). In order to verify consistency, we need to ascertain that for every edge \( e \) in \( \mathcal{T}_N(\text{Poss}(x)) \) leading up to \( u \) and for every sticky map \( \tau \), the prescription \((9.2)\) yields the same value of \( Y_e \) whether we use \( t \) or \( t' \). Rephrasing this, it suffices to establish that

\[
\pi_k(\tau(t)) = \pi_k(\tau(t')) \quad \text{and} \quad \pi_k(\beta(t)) = \pi_k(\beta(t')) \quad \text{for all} \ 0 \leq k \leq h(u). \tag{9.3}
\]

Both equalities are consequences of the height and lineage-preserving property of sticky maps, by virtue of which

\[
h(D(t,t')) \leq \min\{h(D(\tau(t), \tau(t'))), h(\beta(t), \beta(t'))\}.
\]

Of these, stickiness of \( \tau \) has been proved in Proposition\[6.1\] The unambiguous definition and stickiness of \( \beta \) has been verified in Lemma\[3.7\].

For the remainder, we recall from Section\[6\] (see the discussion preceding Proposition\[6.1\] that for \( t = (i_1, i_2, \cdots, i_N) \), the projection \( \pi_k(\tau(t)) = X_{i_1, \cdots, i_k} \) is a Bernoulli\((\frac{1}{2})\) random variable, so \( \Pr(Y_e = 1) = \frac{1}{2} \). Further the random variables \( Y_e \) associated with distinct edges \( e \) in \( \mathcal{T}_N(\text{Poss}(x)) \) are determined by distinct Bernoulli random variables of the form \( X_{i_1, \cdots, i_k} \). The stated independence of the latter collection implies the same for the former.

Thus the collection \( Y_N = \{Y_e\}_e \in \mathcal{E} \) defines a Bernoulli percolation on \( \mathcal{T}_N(\text{Poss}(x)) \), where \( \mathcal{E} \) is the edge set of \( \mathcal{T}_N(\text{Poss}(x)) \). As described in Section\[5.1\] the event \( \{Y_e = 0\} \) corresponds to the removal of the edge \( e \) from \( \mathcal{E} \), and the event \( \{Y_e = 1\} \) corresponds to retaining this edge.

**Lemma 9.2.** Let \( \Pr(x) = \Pr\{\tau : x \in K_N(\sigma_\tau)\} \) be as in \((9.1)\), and \( \{Y_e\} \) as in \((9.2)\).

(a) For any \( x \in [C_0, C_0 + 1] \times \mathbb{R}^d \), the event \( \{\tau : x \in K_N(\sigma_\tau)\} \) is contained in

\[
\{\tau : \exists \text{ a full-length ray in } \mathcal{T}_N(\text{Poss}(x)) \text{ that survives percolation via } \{Y_e\}\}. \tag{9.4}
\]

(b) As a result,

\[
\Pr(x) \leq \Pr(\text{survival after percolation on } \mathcal{T}_N(\text{Poss}(x))).
\]

**Proof.** It is clear that \( x \in K_N(\sigma_\tau) \) if and only if there exists \( t \in \text{Poss}(x) \) such that \( \sigma_\tau(t) = v(t) \), where \( v(t) \) is the unique slope in \( \Omega_N \) prescribed by Lemma\[5.7\] for which \( x \in \mathcal{P}_{t,v(t)} \). In other words, we have

\[
\{\tau : x \in K_N(\sigma_\tau)\} = \bigcup\{\sigma(t) = v(t) : t \in \text{Poss}(x)\} = \bigcup\{\tau(t) = \beta(t) : t \in \text{Poss}(x)\}, \tag{9.5}
\]

where the last step follows from the preceding one by unraveling the string of bijective mappings \( \gamma^{-1}, \Phi^{-1} \) and \( \psi \) (described in Proposition\[6.2\]) that leads from \( \sigma(t) \) to \( \tau(t) \),
and which incidentally also generates \( \beta(t) = \langle j_1, \ldots, j_N \rangle \in \mathcal{T}([0, 1); 2) \) from \( v(t) \). Since \( t \) is identified by some sequence \( \langle i_1, i_2, \ldots, i_N \rangle \), we have its associated random binary sequence

\[
\tau(t) = \langle X_{i_1}, X_{i_1,i_2}, \ldots, X_{i_1,i_2,\ldots,i_N} \rangle \in \mathcal{T}_N([0, 1); 2).
\]

Using this, we can rewrite (9.5) as follows:

\[
\bigcup_{t \in \text{Poss}(x)} \{ \sigma(t) = v(t) \} = \bigcup_{t \in \text{Poss}(x)} \{ \langle X_{i_1}, X_{i_1,i_2}, \ldots, X_{i_1,i_2,\ldots,i_N} \rangle = \langle j_1, j_2, \ldots, j_N \rangle \} = \bigcup_{t \in \text{Poss}(x)} \bigcap_{k=1}^{N} \{ X_{i_1,\ldots,i_k} = j_k \} = \bigcup_{R \in \partial \mathcal{T} \cap \mathcal{E} \cap R} \bigcap_{e \in \mathcal{E} \cap \mathcal{R}} \{ Y_e = 1 \}. \tag{9.6}
\]

In the above steps we have set \( \mathcal{T} := \mathcal{T}_N(\text{Poss}(x)) \) for brevity and let \( \mathcal{E} \) be the edge set of \( \mathcal{T} \). The last step uses (9.2), and the final event is the same as the one in (9.4). Using (9.6), we have

\[
\Pr(x) \leq \Pr \left( \bigcup_{R \in \partial \mathcal{T} \cap \mathcal{E} \cap \mathcal{R}} \bigcap_{e \in \mathcal{E} \cap \mathcal{R}} \{ Y_e = 1 \} \right). \tag{9.7}
\]

This last expression is obviously equivalent to the righthand side of (5.1), verifying the second part of the lemma.

Our next task is therefore to estimate the survival probability of \( \mathcal{T}_N(\text{Poss}(x)) \) under Bernoulli(\( \frac{1}{2} \)) percolation. For this purpose and in view of the discussion in Section 5.3, we should visualize \( \mathcal{T}_N(\text{Poss}(x)) \) as an electrical circuit, the resistance of an edge terminating at a vertex of height \( k \) being \( 2^k - 1 \), per equation (5.2). Let us denote by \( R(\text{Poss}(x)) \) the resistance of the entire circuit. In light of the theorem of Lyons, restated in the form of Proposition 5.3, it suffices to establish the following lemma.

**Lemma 9.3.** With the resistance of \( \text{Poss}(x) \) defined as above, we have

\[
R(\text{Poss}(x)) \gtrsim N. \tag{9.8}
\]

**Proof.** We begin by constructing a different electrical network from the one naturally associated to our tree \( \text{Poss}(x) \). For every \( k \geq 1 \), we connect all vertices at height \( k \) by an ideal conductor to make one node \( V_k \), as in Figure 11. Call this new circuit \( E \).

The resistance of \( E \) cannot be greater than the resistance of the original circuit, by Proposition 5.1. Now fix \( k, 1 \leq k \leq N \), and let \( R_k \) denote the resistance between...
V_{k-1}$ and $V_k$. The number of edges between $V_{k-1}$ and $V_k$ is equal to the number $N_k$ of $k$th generation vertices in $\mathcal{T}_N(\text{Poss}(x))$. Recalling the containment (3.3) from Lemma 3.4 we find that $N_k$ is bounded above by $\overline{N}_k$, the number of $k$th level vertices in $\mathcal{T}_N(\{0\} \times [0,1]^d \cap (x-x_1\Omega_N))$. By Lemma 4.4(b), $\overline{N}_k \lesssim 2^k$, where the implicit constant is uniform in $x \in [C_0, C_0 + 1] \times [-2C_0, 2C_0]^d$. Thus,

$$\frac{1}{R_k} = \sum_{i=1}^{N_k} \frac{1}{2^{k-1}} = \frac{N_k}{2^{k-1}} \lesssim \frac{\overline{N}_k}{2^k} \lesssim 1,$$

and this holds for any $1 \leq k \leq N$. Since the resistors $\{R_k\}_{k=1}^N$ are in series, $R(\text{Poss}(x)) \geq R(E) = \sum_{k=1}^N R_k \gtrsim N$, establishing inequality (9.8).

Combining Lemmas 9.2 and 9.3 with Proposition 5.3 gives us the desired bound of $\lesssim 1/N$ on (9.1). This completes the proof of inequality (2.6), and so too Proposition 6.4.

References

[1] A. Alfonseca. \textit{Strong type inequalities and an almost-orthogonality principle for families of maximal operators along directions in $\mathbb{R}^2$}. J. London Math. Soc., 67, No. 1, 208-218 (2003).

[2] A. Alfonseca, F. Soria, A. Vargas. \textit{A remark on maximal operators along directions in $\mathbb{R}^2$}. Math. Res. Lett., 10, No. 1, 41-49 (2003).

[3] M. Bateman. \textit{Kakeya sets and directional maximal operators in the plane}. Duke Math. J., 147, No. 1, 55-77 (2009).

[4] M. Bateman, N.H. Katz. \textit{Kakeya sets in Cantor directions}. Math. Res. Lett., 15, No. 1, 73-81 (2008).
[5] A.S. Besicovitch. *On Kakeya’s problem and a similar one.* Mat. Zeitschrift, 27, No. 1, 312-320 (1928).

[6] A. Carbery. *Differentiation in lacunary directions and an extension of the Marcinkiewicz multiplier theorem.* Ann. Inst. Four., 38, No. 1, 157-168 (1988).

[7] A. C´ordoba. *The Kakeya maximal function and spherical summation multipliers.* Amer. J. Math., 99, No. 1, 1-22 (1977).

[8] J. Duoandikoetxea, A. Vargas. *Directional operators and radial functions on the plane.* Ark. Mat. 33, 281-291 (1995).

[9] K. Falconer. *Fractal Geometry: mathematical foundations and applications,* 2nd edition. John Wiley & Sons (2003).

[10] C. Fefferman. *The multiplier problem for the ball.* Ann. Math, 94, No. 2, 330-336 (1971).

[11] G. Grimmett. *Percolation,* 2nd edition. Grundlehren der math. Wissenschaften, Vol. 321 (1999).

[12] N.H. Katz. *A countereexample for maximal operators over a Cantor set of directions.* Mat. Res. Let. 3, 527-536 (1996).

[13] N.H. Katz, I. Laba, T. Tao. *An improved bound on the Minkowski dimension of Besicovitch sets in \( \mathbb{R}^3 \).* Ann. of Math. (2) 152, 383-446 (2000).

[14] N.H. Katz, T. Tao. *New bounds for Kakeya problems.* J. Analy. Mat. 87, 231-263 (2002).

[15] L. Kolasa, T. Wolff. *On some variants of the Kakeya problem.* Pacific J. Math. 190, 111-154 (1999).

[16] E. Kroc, M. Pramanik. *Lacunarity, Kakeya-type sets and directional maximal operators.* [preprint available at http://www.arxiv.org](http://www.arxiv.org)

[17] R. Lyons. *Random walks and percolation on trees.* Ann. Prob., 18, 931-958 (1990).

[18] R. Lyons. *Random walks, capacity, and percolation on trees.* Ann. of Prob., 20, (1992).

[19] R. Lyons, Y. Peres. *Probability on Trees and Networks.* Cambridge Univ. Press, mypage.iu.edu/~rdlyons/prbtree/prbtree.html (in preparation).

[20] A. Nagel, E.M. Stein, S. Wainger. *Differentiation in lacunary directions.* Proc. Natl. Acad. Sci. US, 75, No. 3, 1060-1062 (1978).

[21] J. Parcet, K.M. Rogers. *Differentiation of integrals in higher dimensions.* Proc. Natl. Acad. Sci. US, 110, No. 13, 4941-4944 (2013).
[22] P. Sjögren, P. Sjölin. *Littlewood-Paley decompositions and Fourier multipliers with singularities on certain sets*. Ann. Inst. Four., 31, 157-175 (1981).

[23] E. Stein. *Harmonic Analysis: real-variable methods, orthogonality and oscillatory integrals*. Princeton mathematical series; 43. Princeton University Press, 1995.

[24] A. Vargas. *A remark on a maximal function over a Cantor set of directions*. Rend. Circ. Mat. Palermo 44, 273-282 (1995).

[25] T. Wolff. *An improved bound for Kakeya type maximal functions*. Rev. Math. Iberoamericana, 11, No. 3, 651-674 (1995).

Edward Kroc
University of British Columbia, Vancouver, Canada.
Electronic address: ekroc@math.ubc.ca

Malabika Pramanik
University of British Columbia, Vancouver, Canada.
Electronic address: malabika@math.ubc.ca