THE HAUSDORFF DIMENSION OF A CLASS OF RANDOM SELF-SIMILAR FRACTAL TREES

D. A. CROYDON,* *University of Warwick*

Abstract

In this article a collection of random self-similar fractal dendrites is constructed, and their Hausdorff dimension is calculated. Previous results determining this quantity for random self-similar structures have relied on geometrical properties of an underlying metric space or the scaling factors being bounded uniformly away from 0. However, using a percolative argument, and taking advantage of the tree-like structure of the sets considered here, it is shown that conditions such as these are not necessary. The scaling factors of the recursively defined structures in consideration form what is known as a multiplicative cascade, and results about the height of this random object are also obtained.

Keywords: Hausdorff dimension; multiplicative cascade; random tree; self-similar fractal

2000 Mathematics Subject Classification: Primary 60D05
Secondary 60J80; 28A80

1. Introduction

There is now a substantial literature focusing on the geometrical and analytical properties of self-similar fractals, which are commonly described as the unique nonempty compact subset $K \subseteq X$ satisfying $K = \bigcup_{i \in S} \psi_i(K)$, where $(\psi_i)_{i \in S}$ is a finite collection of contractions on an underlying complete metric space (X, d). The existence and uniqueness of K is guaranteed by an extension of the usual contraction principle for complete metric spaces; see [16, Theorem 1.1.4], for example. A fundamental problem in this area is to calculate the Hausdorff dimension, $\dim_H K$, of the self-similar fractal K, and in a wide class of examples it is now known (see [16, Corollary 1.5.9]) that $\dim_H K$ is the unique positive α solving $\sum_{i \in S} r_i^\alpha = 1$, where $(r_i)_{i \in S}$ are the contraction ratios of $(\psi_i)_{i \in S}$. Additionally, various stochastic versions of this result have been investigated. For example, when the underlying metric space (X, d) is a finite-dimensional Euclidean space, in [20] a random self-similar set K satisfying $K = \bigcup_{i \in S} w(i) K_i$, where $(w(i))_{i \in S}$ is a random (finite or countable) collection of scaling factors and $(K_i)_{i \in S}$ are independent copies (up to translation) of K, independent of $(w(i))_{i \in S}$, was constructed and (assuming that K is nonempty and $(w(i)K_i)_{i \in S}$ fulfils a nonoverlapping condition) the Hausdorff dimension of K was shown to be P-almost surely (a.s.) equal to the unique positive α solving

$$\mathrm{E} \sum_{i \in S} w(i)^\alpha = 1, \tag{1}$$

which obviously reduces to the deterministic equation when $w(i) = r_i$, P-a.s. Note that we are assuming that all the random variables are defined on an underlying probability space
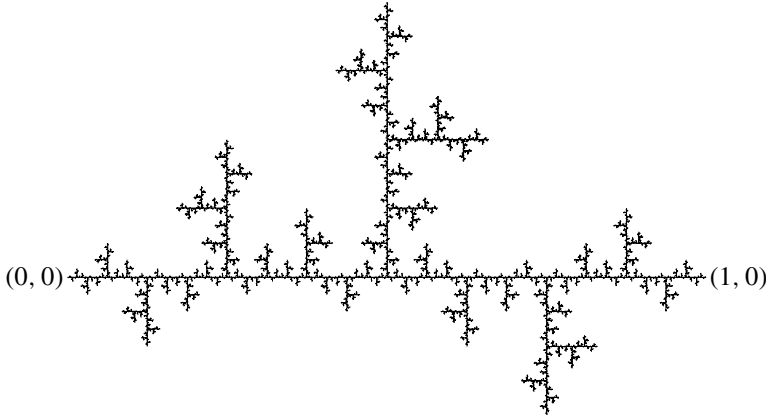

FIGURE 1: Self-similar dendrite T.

with probability measure P, and E is the expectation under P. A similar result was proved independently in [8].

The aim of this article is to obtain similar Hausdorff dimension results to those discussed above for a class of random self-similar trees that do not fit into the usual framework for random self-similar sets. First though, it is necessary to deal with questions concerning their construction. Our starting point, rather than to define a random set directly, is to consider a fixed fractal subset of $\mathbb{R}^2$ and build a random metric upon it. More precisely, for $(x, y) \in \mathbb{R}^2$, set

$$\psi_1(x, y) := \tfrac{1}{2}(1 - x, y), \qquad \psi_2(x, y) := \tfrac{1}{2}(1 + x, -y), \qquad \psi_3(x, y) := (\tfrac{1}{2} + cy, cx),$$

for some constant $c \in (0, \tfrac{1}{2})$, and define T to be the unique nonempty compact set satisfying $T = \bigcup_{i \in S} \psi_i(T)$, where we now define $S := \{1, 2, 3\}$. The set T, shown in Figure 1, is easily checked to be a *dendrite*, by which we mean that it is an arc-wise connected topological space containing no subset homeomorphic to the circle. Although the Euclidean metric is important for its construction, we are only interested in T as a topological space. Indeed, the Hausdorff dimension of T with respect to the intrinsic random metric we construct upon it can be strictly larger than 2.

Heuristically, we build a random metric on T by first supposing that the distance between the 'edge' from $(0, 0)$ to $(1, 0)$ is of length 1. We then replace this by three randomly scaled copies of the edge with new lengths given by $(w(i))_{i \in S}$ (see Figure 2) and continue inductively to replace edges independently of each other by triples of scaled edges, with the relevant scaling factors having the same distribution as $(w(i))_{i \in S}$. As the number of inductive steps increases our discrete approximations eventually fill out a dense subset of T, and (under certain distributional conditions on the scaling factors) calculating the 'limiting distance' between points yields a metric R on T such that (T, R) is a compact metric space, and the topology induced by R on T is the same as the original (Euclidean) one, P-a.s. See Section 3 for full details.

Importantly, we do not assume that $w(1) + w(2) \equiv 1$, P-a.s., and, as a result of this, the distance in (T, R) between $(0, 0)$ and $(1, 0)$ depends, in general, on all of the steps in the inductive procedure. For reasons related to the construction of the metric R as a so-called resistance metric, we call this limiting distance between $(0, 0)$ and $(1, 0)$, $R_\varnothing$, a resistance perturbation. Similarly, the distance between $(0, 0)$ and $(\tfrac{1}{2}, 0)$ in (T, R) is not simply $w(1)$, but
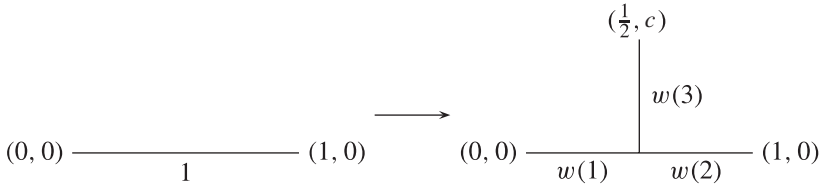
FIGURE 2: Edge replacement procedure.

equal to $w(1)R_1$, where R_1 is the resistance perturbation associated with this 'edge' of T and has the same distribution as $R_\varnothing$. The fact that T is a dendrite means that we can characterise these resistance perturbations in a convenient way that allows us to deduce their distributional properties and perform calculations with them.

In Section 4 we provide conditions on $(w(i))_{i \in S}$ that allow us to deduce the Hausdorff dimension of (T, R). Proving that α, as defined by (1), is an upper bound for the Hausdorff dimension is straightforward, and requires no further conditions than those used for the construction of (T, R). Conversely, when obtaining a lower bound for the dimension of a fractal defined in a recursive fashion, it is often a problem when parts of the fractal become small too quickly, and this is the case here. For the proof in [20] of the result described above, the underlying Euclidean geometry of the random sets being considered is critical, but since our fractal trees are not embedded into any fixed metric space there is no easy translation of this argument to our setting. Another common assumption for proving a Hausdorff dimension lower bound and related results is that the scaling factors are bounded uniformly away from zero [8], [12]. In Theorem 5, below, we show one application of this condition in our setting. More interestingly, though, is that knowledge of the geometry of the trees in consideration allows us to avoid a uniform lower bound; in fact, we shall require only that clusters of small scaling factors are not too large. We construct a random graph approximation to T and use a percolation argument upon this to show that this is the case when the scaling factors are independent and their distributions satisfy a simple polynomial tail bound at 0; see Theorem 6, below.

Let us complete the discussion of our results related to random self-similar dendrites by outlining a pair of examples. Firstly, one choice of random scaling factors that does not fit into any previously studied set-up but satisfies conditions that allow us to calculate the Hausdorff dimension of the fractal (T, R) is if we suppose that $(w(i))_{i \in S}$ are independent and identically distributed (i.i.d.) $U(0, 1)$ random variables. It is easy to check that the Hausdorff dimension is 2 in this case. Note, in particular, that $w(1) + w(2) \neq 1$, P-a.s.; consequently, there are indeed nontrivial resistance perturbations in this case. Secondly, although we cannot calculate the Hausdorff dimension using the techniques of this article, when $(w(i))_{i \in S}$ are the square-roots of a Dirichlet $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ triple, the construction of (T, R) is of interest in its own right. The reason for this is explained in detail in [7], where it was proved that in this case (T, R) is a version of the continuum random tree of Aldous, which is an important random dendrite with connections to many other stochastic tree-like objects [1].

The family of random scaling factors that we use to construct the random metric space (T, R) form what is known as a multiplicative cascade, which is a probabilistic structure that has been studied extensively; see [9] and [17]–[19]. Much of this previous work has concentrated on investigating properties of an associated tree-martingale limit (see Section 2 for a definition), and we add to this body of knowledge by proving a tail bound at 0 for this random variable (Proposition 1). We also define the height of a (generalised) multiplicative cascade, and derive

a simple condition that yields the finiteness of this quantity and its moments; see Corollary 1 and Theorem 2.

Finally, let us remark that, although only one particular self-similar fractal dendrite is considered in this article, at the expense of some additional notation, relatively minor refinements of the proofs used here allow the results we obtain to be extended to any post-critically finite self-similar fractal dendrite; see [16] for a definition of a post-critically finite self-similar fractal. A treatment of the general case appears in [6].

2. Multiplicative cascade results

We first introduce an address space to label various objects in the discussion. Fix a finite index set S, let $N := |S|$, and define the 'shift-space' of infinite sequences $\Sigma := S^{\mathbb{N}}$. For $n \geq 0$, the corresponding finite sequences are denoted by

$$\Sigma_n := S^n \quad \text{and} \quad \Sigma_* := \bigcup_{m=0}^{\infty} \Sigma_m,$$

where $S^0 := \{\varnothing\}$. For $i \in \Sigma_m$, $j \in \Sigma_n$, and $k \in \Sigma$, we write $ij = (i_1, \ldots, i_m, j_1, \ldots, j_n)$ and $ik = (i_1, \ldots, i_m, k_1, k_2 \ldots)$. For $i \in \Sigma_*$, we denote by $|i|$ the integer n such that $i \in \Sigma_n$ and call this the length of i. For $i \in \Sigma_n \cup \Sigma$, $n \geq m$, the truncation of i to length m is written as $i \mid m = (i_1, \ldots, i_m)$.

We define a *multiplicative cascade* to be a family of random variables $(w(i))_{i \in \Sigma_* \backslash \{\varnothing\}}$ which take values in $[0, 1]$ such that, for $i \in \Sigma_* \backslash \{\varnothing\}$, the N-tuples $(w(ij))_{j \in S}$ are independent copies of $(w(j))_{j \in S}$. The multiplicative cascade has a naturally associated filtration $(\mathcal{F}_n)_{n \geq 0}$, defined by $\mathcal{F}_n := \sigma(w(i) : |i| \leq n)$. Throughout our arguments, we use the function

$$F(\theta) := \mathrm{E}\left(\sum_{i \in S} w(i)^{\theta}\right) \quad \text{for } \theta > 0,$$

which is decreasing and continuous and satisfies

$$F(\theta) \to \sum_{i \in S} \mathrm{P}(w(i) = 1) \quad \text{as } \theta \to \infty. \tag{2}$$

Furthermore, we introduce the notation $l(i) := w(i \mid 1) w(i \mid 2) \cdots w(i)$ for $i \in \Sigma_* \backslash \{\varnothing\}$, and $l(\varnothing) = 1$; and use this to define the so-called *tree-martingale* (this term was coined in [9]), $(M^{\theta}(n))_{n \geq 0}$, by $M^{\theta}(n) := \sum_{i \in \Sigma_n} l(i)^{\theta} F(\theta)^{-n}$. It is straightforward to check that, for each $\theta > 0$, $(M^{\theta}(n))_{n \geq 0}$ is an $(\mathcal{F}_n)_{n \geq 0}$-martingale. In particular, $\mathrm{E}(M^{\theta}(n)) = \mathrm{E}(M^{\theta}(0)) = 1$. By the almost sure martingale convergence theorem, this implies that $M^{\theta}(n) \to M^{\theta}$ as $n \to \infty$, P-a.s., for some random variable M^{θ} with $\mathrm{E} M^{\theta} \in [0, 1]$. By relabelling, we have the same distributional properties for the random variables $(M_i^{\theta})_{i \in \Sigma_*}$ defined by

$$M_i^{\theta} = \lim_{n \to \infty} \frac{\sum_{j \in \Sigma_n} l(i, j)^{\theta}}{l(i)^{\theta} F(\theta)^n}. \tag{3}$$

It is not difficult to check that, for each n, $(M_i^{\theta})_{i \in \Sigma_n}$ is a collection of i.i.d. random variables, independent of $\mathcal{F}_n$. Also, the following identity holds:

$$M^{\theta} = \sum_{i \in \Sigma_n} \frac{l(i)^{\theta}}{F(\theta)^n} M_i^{\theta}. \tag{4}$$

To generalise our cascade model, we introduce random perturbations of the $l(i)$, denoted by $(X_i)_{i \in \Sigma_*}$. We shall assume that the X_i are identically distributed nonnegative random variables satisfying $E(X_i^\theta) < \infty$ for every $\theta > 0$, and that $X_i \perp \mathcal{F}_{|i|}$ for all $i \in \Sigma_*$, where $\perp$ is taken to mean 'is independent of'. The reason for the introduction of the factors $(X_i)_{i \in \Sigma_*}$ will become apparent later, since perturbations with these properties arise naturally in the construction of our self-similar dendrite.

We can consider the cascade model as a weighted graph tree, rooted at $\varnothing$ with vertex set Σ_* and edge set $\{\{i, i|(|i| - 1)\} : i \in \Sigma_* \setminus \{\varnothing\}\}$; where the edge $\{i, i \mid (|i| - 1)\}$ has weight $l(i)X_i$. For two vertices in Σ_*, we define the distance between them to be the sum of edge distances along the shortest path in the graph. We then define the *height* of the tree to be

$$H = \sup_{i \in \Sigma} \sum_{n=0}^{\infty} l(i \mid n) X_{i \mid n}.$$

The usual definition of tree height (the supremum of distances of vertices from the root) actually has the sum index starting from 1, but we shall find this slightly adjusted definition more useful in later sections.

Our first main result about multiplicative cascades is Theorem 1, below, a simple corollary of which gives a sufficient condition for the expected height $E H$ to be finite. In Theorem 2, below, we deal with the unperturbed case and show that the condition is necessary in this case. We start by estimating how fast the edge lengths $l(i)X_i$ decay as $|i| \to \infty$.

Lemma 1. *Suppose that* $\sum_{i \in S} P(w(i) = 1) < 1$ *and fix* $d \geq 0$, *then*

(a) *there exist constants* $c < \infty$ *and* $\alpha_1 \in (0, 1)$ *such that*

$$E\left(\left(\sup_{i \in \Sigma_n} l(i)X_i\right)^d\right) \leq c\alpha_1^n \quad \text{for all } n \in \mathbb{N},$$

(b) *there exists a positive, finite random variable* A *and (deterministic)* $\alpha_2 \in (0, 1)$ *such that*

$$\sup_{i \in \Sigma_n} l(i)X_i \leq A\alpha_2^n \quad \text{for all } n \in \mathbb{N}, \text{P-a.s.}$$

Proof. To prove part (a) we first look for bounds on the tail of the distribution of $\sup_{i \in \Sigma_n} l(i)$. Applying Markov's inequality, the definition of $M^\theta(n)$, and the independence assumption of the X_is we obtain, for $\lambda, \theta > 0$,

$$P\left(\sup_{i \in \Sigma_n} l(i)X_i \geq \lambda\right) \leq P\left(\sum_{i \in \Sigma_n} l(i)^\theta X_i^\theta \geq \lambda^\theta\right) \leq \lambda^{-\theta} E(X_\varnothing^\theta) F(\theta)^n. \tag{5}$$

The condition $\sum_{i \in S} P(w(i) = 1) < 1$ and (2) imply that we can find $\theta_0 > d$ large enough so that $F(\theta_0) < 1$. Let $x := \|X_\varnothing\|_{\theta_0}$ (which is less than ∞ by assumption) and define $\lambda_n := x F(\theta_0)^{n/\theta_0}$, which is less than 1 for $n \geq n_0$ for some $n_0 \in \mathbb{N}$. Assume, for now, that $n \geq n_0$. For $\lambda \geq \lambda_n$, the upper bound in (5) is less than or equal to 1 and so is nontrivial and,

for $\lambda < \lambda_n$, we merely use the fact that we are trying to bound a probability to deduce that

$$
\begin{aligned}
\mathrm{E}\left(\left(\sup_{i \in \Sigma_n} l(i) X_i\right)^d\right) &= \int_0^\infty d\lambda^{d-1} \, \mathrm{P}\left(\sup_{i \in \Sigma_n} l(i) X_i \geq \lambda\right) d\lambda \\
&\leq \int_0^{\lambda_n} d\lambda^{d-1} \, \mathrm{d}\lambda + \int_{\lambda_n}^\infty d\lambda^{d-1-\theta_0} x^{\theta_0} F(\theta_0)^n \, \mathrm{d}\lambda \\
&= \frac{x^d \theta_0}{\theta_0 - d} F(\theta_0)^{dn/\theta_0}.
\end{aligned}
$$

Hence, taking $\alpha_1 = F(\theta_0)^{d/\theta_0}$ and c suitably large gives us part (a) of the lemma.

To prove part (b) we again look to bound the tail probability of $\sup_{i \in \Sigma_n} l(i) X_i$. We proceed as above to obtain the bound

$$
\mathrm{P}(\sup_{i \in \Sigma_n} l(i) X_i \geq \lambda^n) \leq \lambda^{-n\theta} \, \mathrm{E}(X_\varnothing^\theta) F(\theta)^n.
$$

If we fix $\theta = \theta_0$ we can find a $\lambda_0 \in (0, 1)$ such that $\lambda_0^{-\theta_0} F(\theta_0) < 1$ and so

$$
\sum_{n=0}^\infty \mathrm{P}\left(\sup_{i \in \Sigma_n} l(i) X_i \geq \lambda_0^n\right) \leq \mathrm{E}(X_\varnothing^\theta) \sum_{n=0}^\infty (\lambda_0^{-\theta_0} F(\theta_0))^n < \infty.
$$

An application of the Borel–Cantelli lemma then gives us part (b) of our lemma.

In proving Theorem 1 we will apply the following elementary lemma, which we state without proof.

Lemma 2. *Let $(x_n)_{n \geq 0}$ be a sequence of nonnegative real numbers, then*

$$
\left(\sum_{n=0}^\infty x_n\right)^d \leq c \sum_{n=0}^\infty x_n^d (1 + n)^d \quad \text{for all } d \geq 1,
$$

where c is a constant depending only on d.

Theorem 1. *Let $\sum_{i \in S} \mathrm{P}(w(i) = 1) < 1$ and $d \geq 0$, then*

$$
\mathrm{E}\left(\left(\sum_{n=0}^\infty \sup_{i \in \Sigma_n} l(i) X_i\right)^d\right) < \infty.
$$

Proof. For $d \geq 1$, applying Lemmas 1 and 2 yields

$$
\mathrm{E}\left(\left(\sum_{n=0}^\infty \sup_{i \in \Sigma_n} l(i) X_i\right)^d\right) \leq c_1 \, \mathrm{E}\left(\sum_{n=0}^\infty \left(\sup_{i \in \Sigma_n} l(i) X_i\right)^d (1 + n)^d\right) \leq c_2 \sum_{n=0}^\infty (1 + n)^d \alpha_1^n,
$$

which is clearly finite. For $d \in [0, 1)$, the result follows from the case in which $d = 1$ by applying the inequality $x^d \leq 1 + x$ for all $x \geq 0$.

Corollary 1. *Let $\sum_{i \in S} \mathrm{P}(w(i) = 1) < 1$ and $d \geq 0$, then $\mathrm{E}(H^d) < \infty$.*

Proof. After noting that

$$\sup_{i \in \Sigma} \sum_n l(i \mid n) X_{i \mid n} \leq \sum_n \sup_{i \in \Sigma_n} l(i) X_i,$$

this result follows from Theorem 1.

We now show how, for unperturbed cascades, the condition $\sum_{i \in S} P(w(i) = 1) < 1$ is also necessary for finite moments of H.

Theorem 2. *Assume that $X_i \equiv 1$ for all $i \in \Sigma_*$ then*

(a) *if $\sum_{i \in S} P(w(i) = 1) < 1$, $E H^d < \infty$ for all $d \in \mathbb{R}$,*

(b) *if $\sum_{i \in S} P(w(i) = 1) = 1$, $E H = \infty$,*

(c) *if $\sum_{i \in S} P(w(i) = 1) > 1$, $P(H = \infty) > 0$.*

Proof. Assume that $\sum_{i \in S} P(w(i) = 1) < 1$. Clearly, if $X_i \equiv 1$ for all $i \in \Sigma_*$ then $(X_i)_{i \in \Sigma_*}$ satisfies the conditions that enable us to apply the previous results of this section. Hence, if $d \geq 0$, $E(H^d) < \infty$ follows from Corollary 1. We note that, because $l(\varnothing) = 1$, $H \geq 1$. Hence, for $d < 0$, $E(H^d) \leq 1 < \infty$. Thus, part (a) holds.

We now construct a Galton–Watson branching process related to our tree. Given $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$, we define $\tilde{w}(i) := \mathbf{1}_{\{w(i)=1\}}$, where $\mathbf{1}_{\{\cdot\}}$ is the indicator function, and

$$\tilde{l}(i) := \prod_{n=1}^{|i|} \tilde{w}(i \mid n).$$

It is then easy to check that if

$$Z_n := \sum_{i \in \Sigma_n} \tilde{l}(i),$$

then $(Z_n)_{n \geq 0}$ is a Galton–Watson process. Importantly, if $Z_n > 0$ then we must have $\tilde{l}(i) = 1$ for some $i \in \Sigma_n$, and so $\tilde{l}(i \mid m) = 1$ for $1 \leq m \leq n$. Consequently, $H \geq n + 1$. This means that $\{Z_n > 0\} \subseteq \{H \geq n + 1\}$ and so we can use known results about the extinction of the Galton–Watson process to infer results about H. In particular, it may be calculated that

$$E Z_1 = \sum_{i \in S} P(w(i) = 1).$$

To prove part (c) we note that if $\sum_{i \in S} P(w(i) = 1) > 1$ then $E Z_1 > 1$ and so the branching process is supercritical and survives with positive probability (for a proof of this see, for example, [3]). Thus, we have $Z_n > 0$ for all $n \geq 0$, with strictly positive probability and this implies that

$$P(H = \infty) = P(H \geq n, \text{ for all } n \geq 0) > 0.$$

Now assume that $E Z_1 = \sum_{i \in S} P(w(i) = 1) = 1$. It is a standard result that the extinction time of a critical Galton–Watson process with finite offspring variance has an infinite expectation. If X is the extinction time of our Galton–Watson process then we have $\{X > n\} = \{Z_n > 0\} \subseteq \{H \geq n + 1\}$ and so $H > X$. Since $E X = \infty$, it follows that part (b) holds.

Remark 1. At criticality, the Galton–Watson process exhibits two types of behaviour. First, there is the trivial case when $Z_1 = 1$, P-a.s. This implies that $Z_n = 1$ for all n, and so the process survives. This is mirrored in the multiplicative cascade when we have $\sum_{i \in S} P(w(i) = 1) = 1$ and $P(\sup_{i \in S} w(i) = 1) = 1$. In this case $H = \infty$, P-a.s. In the nontrivial case, $P(Z_1 = 1) < 1$, the Galton–Watson process becomes extinct with probability 1. It follows that we can also find a multiplicative cascade with $\sum_{i \in S} P(w(i) = 1) = 1$ and $H < \infty$, P-a.s. The problem of whether we have $H < \infty$ in the general nontrivial case is left open.

Remark 2. If $N = \infty$, there exist random variables with $\sum_{i \in S} P(w(i) = 1) < 1$ and also $\sup_{i \in S} w(i) = 1$, P-a.s. In this case we have $H = \infty$, P-a.s., and so Theorem 2 does not hold in general when $N = \infty$.

Remark 3. The condition $w(i) \in [0, 1]$, P-a.s., is not essential. For example, if we let $(w(i))_{i \in S}$ be i.i.d. U$[0, x]$ random variables and $N \geq 2$, then by mimicking the proofs of Lemma 1 and Theorem 2 it can be shown that $\mathrm{E}\, H^d < \infty$ for some values of x strictly greater than 1 (the largest x for which this is true is $1 + O(N^{-1})$).

The right-tail of the distribution of the tree-martingale limit M^θ has been considered by various authors, including Liu, who has proved a widely applicable result demonstrating exponential tails [17]. However, we will later require that M^θ also has finite negative moments of some order for certain values of θ, and so we will also need information about its tail at 0. The aim of the remainder of this section is to provide some simple conditions on the distribution of $(w(i))_{i \in S}$ that allow us to deduce suitable estimates on the distribution of M^θ.

Let us start by assuming that $N > 1$, $\sum_{i \in S} P(w(i) = 1) < 1$, and $P(w(i) = 0) = 0$ for every i. It follows that $F(\theta)$ is a strictly decreasing continuous function with $F(0) = N > 1$ and

$$F(\theta) \to \sum_{i \in S} P(w(i) = 1) < 1 \quad \text{as } \theta \to \infty.$$

Hence, there exists a unique, strictly positive solution to the equation $F(\theta) = 1$. We shall denote this solution by α. In the next lemma we present a few simple properties of M^θ when $\theta \leq \alpha$ that we will apply later.

Lemma 3. *Let $N > 1$, $\sum_{i \in S} P(w(i) = 1) < 1$, and $P(w(i) = 0) = 0$ for every i, and let α be the unique solution to $F(\theta) = 1$. If $\theta \leq \alpha$, then*

(a) $P(M^\theta \in (0, \infty)) = 1$,

(b) $\mathrm{E}((M^\theta)^d) < \infty$ *for all $d \geq 0$.*

Proof. For $\theta \leq \alpha$, we can apply the results of [17] to deduce that $M^\theta(n) \to M^\theta$ in mean, and so $\mathrm{E}\, M^\theta = 1$. Furthermore, the decomposition of M^θ given by (4) allows us to deduce that $P(M^\theta = 0)^N = P(M^\theta = 0)$; thus, $P(M^\theta = 0) \in \{0, 1\}$. Combining these two facts immediately implies part (a). Part (b) is also proved in [17].

Our main result for the tail at 0 of M^θ is proved using the methods of Barlow and Bass [5]. Their arguments allow a polynomial estimate for the distribution function of the limit random variable to be improved to an exponential bound. We shall prove our main result only in the case in which $\theta = \alpha$, as this eases notation by removing the factors $F(\theta)$ (recall that $F(\alpha) = 1$). However, the techniques used will actually apply whenever $\theta \leq \alpha$. The first step is to deduce the necessary polynomial bound.

Lemma 4. *If $(w(i))_{i \in S}$ satisfies the conditions of Lemma 3, and we fix $\beta > 0$ and $\varepsilon \in (0, 1)$, then there exists a constant c such that $P(M^\alpha \leq x) \leq \varepsilon + cx^\beta$ for every $x \geq 0$.*

Proof. Fix $\beta > 0$ and $\varepsilon \in (0, 1)$. By Lemma 3, the tree-martingale limit M^α is nonzero, P-a.s. Hence, $P(M^\alpha \leq x) \to 0$ as $x \to 0$. In particular, there exists an $x_0 > 0$ such that $P(M^\alpha \leq x_0) \leq \varepsilon$. Thus, for $x \geq 0$, $P(M^\alpha \leq x) \leq \varepsilon + x_0^{-\beta} x^\beta$.

We can now prove the exponential tail bound at 0 for M^α under the additional assumptions of independence and finiteness of negative moments of the $(w(i))_{i \in S}$.

Proposition 1. *If $(w(i))_{i \in S}$ satisfies the conditions of Lemma 3, are independent, and also $\max_{i \in S} E(w(i)^{-\beta}) < \infty$ for some $\beta > 0$, then, for some constants $c, \gamma \in (0, \infty)$, we have*

$$P(M^\alpha \leq x) \leq \exp(-cx^{-\gamma}) \quad \text{for all } x \geq 0.$$

Proof. Let $\beta > 0$ be a constant for which $\sup_{i \in S} E(w(i)^{-\beta}) < \infty$ and fix $\varepsilon \in (0, 1)$. By Lemma 4, we can find a constant c_1 such that $P(M^\alpha \leq x) \leq \varepsilon + c_1 x^\beta$ for all $x \geq 0$. Applying the relevant independence assumptions and the fact that $M_i^\alpha \sim M^\alpha$, we can deduce from this that, for all $x \geq 0$ and $i \in \Sigma_n$,

$$P(l(i)^\alpha M_i^\alpha \leq x) = E(P(l(i)^\alpha M_i^\alpha \leq x \mid \mathcal{F}_n)) \leq E\left(\varepsilon + c_1 \frac{x^{\beta/\alpha}}{l(i)^\beta}\right) \leq \varepsilon + c_1 c_2^n x^{\beta/\alpha},$$

where $c_2 := \max_{i \in S} E(w(i)^{-\beta}) \vee (N + 1)$. By writing $M^\alpha = \sum_{i \in \Sigma_n} l(i)^\alpha M_i^\alpha$, we can check that the conditions of [5, Lemma 1.1], hold, which implies that

$$P(M^\alpha \leq x) \leq \exp(c_3(c_2 N)^{n/2} x^{\beta/2\alpha} + N^n \ln \varepsilon) \quad \text{for all } x \geq 0, \tag{6}$$

for some constant c_3. We now look to choose n in a way that will give us the control we require over this bound. Define $n_0 = n_0(x)$ to be the unique solution to $(c_2/N)^{n_0/2} = -\ln \varepsilon / x^{\beta/2\alpha} c_3$, and then set $n = \lfloor n_0 - 1 \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer-part function. We have $c_2 > N$ and so we can find $c_4 > 0$ such that $N c_2^{-1} \leq (1 - c_4)^2$. Consequently, because $n - n_0 \in (-2, -1]$, we have

$$(c_2 N)^{(n-n_0)/2} - N^{n-n_0} = N^{n-n_0}((c_2 N^{-1})^{(n-n_0)/2} - 1) \leq -c_4 N^{-2}.$$

By the choice of n_0, our upper bound (6) now becomes

$$
\begin{aligned}
\ln P(M^\alpha \leq x) &\leq c_3(c_2 N)^{n_0/2} x^{\beta/2\alpha} (c_2 N)^{(n-n_0)/2} + N^{n_0} N^{n-n_0} \ln \varepsilon \\
&= -N^{n_0}((c_2 N)^{(n-n_0)/2} - N^{n-n_0}) \ln \varepsilon \\
&\leq c_4 N^{n_0-2} \ln \varepsilon \\
&= c_4 N^{-2} \ln \varepsilon \left(\frac{-\ln \varepsilon}{x^{\beta/2\alpha} c_3}\right)^{2 \ln N/(\ln c_2 - \ln N)},
\end{aligned}
$$

and the result follows.

The final result of this section is a simple corollary of Proposition 1.

Corollary 2. *Under the assumptions of Proposition 1, $E((M^\alpha)^{-d}) < \infty$ for every $d \geq 0$.*

3. Random self-similar fractal dendrites

In this and the following section we fix $S := \{1, 2, 3\}$ and T to be the deterministic dendrite described in Section 1. For a wide class of self-similar fractals, including T, there is now a well-established approximation procedure for defining an intrinsic quadratic form and associated resistance metric on the relevant space, which we briefly outline in our specific case for the purpose of introducing notation. See [4] and [16] for more details.

First, define the boundary of T to be the two-point set $V^0 := \{(0, 0), (1, 0)\}$, and define an initial Dirichlet form by $D(f, f) := (f(0, 0) - f(1, 0))^2$ for $f \in C(V^0)$, where, for a countable set A, we denote $C(A) := \{f \colon A \to \mathbb{R}\}$. Furthermore, we introduce an increasing family of subsets of T by setting

$$V^n := \bigcup_{i \in \Sigma_n} \psi_i(V^0),$$

where, for $i \in \Sigma_n$, $\psi_i := \psi_{i_1} \circ \cdots \circ \psi_{i_n}$. Given the quadratic form D and a set of scaling factors $(r_i)_{i \in S}$, with $r_i > 0$ for each $i \in S$, we can define a Dirichlet form on each of the V^n by setting, for $n \geq 1$,

$$\mathcal{E}^n(f, f) := \sum_{i \in \Sigma_n} \frac{1}{r_i} D(f \circ \psi_i, f \circ \psi_i) \quad \text{for all } f \in C(V^n), \tag{7}$$

where $r_i := r_{i_1} \cdots r_{i_n}$ for $i \in \Sigma_n$. To establish the existence of a nontrivial limit as $n \to \infty$, we need to place some restrictions on the choice of $(r_i)_{i \in S}$ so that the family $\{(V^n, \mathcal{E}^n)\}_{n \geq 0}$ is compatible in the sense that the trace of $\mathcal{E}^{n+1}$ on V^n is precisely $\mathcal{E}^n$; see [16, Definition 2.2.1]. Some elementary algebra shows that, in our case, $\{(V^n, \mathcal{E}^n)\}_{n \geq 0}$ is compatible if and only if we assume that $r_1 + r_2 = 1$, and in this case we can take a limit in a sensible way. Specifically, assume that $r_1 + r_2 = 1$ and let

$$\mathcal{E}^*(f, f) := \lim_{n \to \infty} \mathcal{E}^n(f, f) \quad \text{for all } f \in \mathcal{F}^*, \tag{8}$$

where $\mathcal{F}^*$ is the set of functions on the countable set $V^* := \bigcup_{n \geq 0} V^n$ for which this limit exists finitely. Note that we have abused the notation slightly by using the convention that if a form $\mathcal{E}$ is defined for functions on a set A and f is a function defined on $B \supseteq A$, then we write $\mathcal{E}(f, f)$ to mean $\mathcal{E}(f|_A, f|_A)$. The resulting quadratic form $(\mathcal{E}^*, \mathcal{F}^*)$ is actually a resistance form (see [16, Definition 2.3.1]), and we can use it to define a (resistance) metric R^* on V^* by

$$R^*(x, y)^{-1} = \inf\{\mathcal{E}^*(f, f) : f \in \mathcal{F}^*, f(x) = 0, f(y) = 1\} \tag{9}$$

for $x, y \in V^*$, $x \neq y$, and setting $R^*(x, x) = 0$. Finally, if we also assume that $r_i < 1$ for each $i \in S$ then R^* extends uniquely to a metric R on T such that (T, R) is the completion of (V^*, R^*), and moreover the topology induced by R on T is the same as that induced by the original (Euclidean) metric.

We will now explain how to randomise the above construction. The scaling factors that we will use to define a sequence of Dirichlet forms on the subsets $(V^n)_{n \geq 0}$ of T will form a multiplicative cascade and, identifying notation with the previous section, be denoted by $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$. In addition to the independence assumptions that we introduced in Section 2 for a multiplicative cascade, we will further suppose that the random variables $w(i)$ are nonzero, P-a.s. The following assumptions will also be useful, and, for clarity, we will explicitly state when we apply these.

Assumption 1. *We have*

  (a) $E(w(1) + w(2)) = 1$,

  (b) $\sum_{i \in S} P(w(i) = 1) < 1$.

Although we would like to simply replace the deterministic scaling factors r_i by the random variables $w(i)$ in a formula similar to (7), a sequence of forms defined in this way will not be compatible in general and taking limits would not be straightforward. To deal with this problem, we introduce another collection of random variables

$$R_i := \lim_{n \to \infty} \sum_{j \in \{1,2\}^n} \frac{l(ij)}{l(i)}, \qquad i \in \Sigma_*,$$

which we shall call *resistance perturbations*. Clearly these are identically distributed, and, by appealing to the independence properties of $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$, various questions regarding the convergence and distribution of the $(R_i)_{i \in \Sigma_*}$ may be answered by multiplicative cascade techniques as discussed in Section 2. In particular, under Assumption 1(a) we find that the limit defining R_i exists in $(0, \infty)$, P-a.s., $E\, R_i^d < \infty$ for every $d \geq 0$, and that $R_i = w(i1)R_{i1} + w(i2)R_{i2}$ for every $i \in \Sigma_*$. Note that Assumption 1(a) is actually necessary for the nontriviality of the R_is, because if $E(w(1) + w(2)) \neq 1$ then $R_i \in \{0, \infty\}$, P-a.s.

Given a multiplicative cascade of scaling factors satisfying Assumption 1(a) we define a random sequence of Dirichlet forms on the vertex sets $(V^n)_{n \geq 1}$ by

$$\mathcal{E}^n(f, f) := \sum_{i \in \Sigma_n} \frac{1}{l(i)R_i} D(f \circ \psi_i, f \circ \psi_i) \quad \text{for all } f \in C(V^n),$$

for $n \geq 1$, where, as before, $D(f, f) := (f(0, 0) - f(1, 0))^2$ for $f \in C(V^0)$. In analysing this quadratic form, it is natural to consider the graph (V^n, E^n), where we define $E^n := \{\psi_i(V^0) : i \in \Sigma_n\}$. The reason for this is that if we place a resistance of $l(i)R_i$ along each edge $\psi_i(V^0)$, then $\mathcal{E}^n$ is the energy operator associated with the resulting electrical network. It is elementary to check that (V^n, E^n) is a graph tree for every $n \geq 0$ by induction, noting that (V^{n+1}, E^{n+1}) is constructed by joining three graph trees with the same structure as (V^n, E^n) at a single vertex. This fact is convenient as it means that to calculate the resistance between two points in V^n we can simply apply the series law and sum the resistances of the edges on the (unique shortest) path between them in (V^n, E^n). We can use this result to show that the sequence $\{(V^n, \mathcal{E}^n)\}_{n \geq 0}$ is a compatible sequence in the sense described above.

Lemma 5. *If Assumption 1(a) holds, then the sequence $\{(V^n, \mathcal{E}^n)\}_{n \geq 0}$ is compatible,* P-*a.s.*

Proof. To prove the result it will suffice to prove that the resistance between $\psi_i(0, 0)$ and $\psi_i(1, 0)$ in the electrical network on (V^{n+1}, E^{n+1}) is equal to $l(i)R_i$ for every $i \in \Sigma_n$. The path between $\psi_i(0, 0)$ and $\psi_i(1, 0)$ in (V^{n+1}, E^{n+1}) is given by the edges $\psi_{i1}(V^0)$ and $\psi_{i2}(V^0)$. Thus, the resistance between $\psi_i(0, 0)$ and $\psi_i(1, 0)$ in the level $n + 1$ network is $l(i1)R_{i1} + l(i2)R_{i2}$, which is equal to $l(i)R_i$ and so the proof is complete.

As a consequence of Lemma 5, if Assumption 1(a) holds, we can proceed as in the deterministic case to define $(\mathcal{E}^*, \mathcal{F}^*)$ on V^* by taking the limit of the sequence $(\mathcal{E}^n)_{n \geq 1}$, as in (8). By [16, Theorem 2.2.6], this is a resistance form and we can define a resistance metric R^* on

V^* by (9), P-a.s. Note that the compatibility of the sequence of forms used to define R^* implies that

$$R^*(\psi_i(0,0), \psi_i(1,0)) = l(i)R_i \quad \text{for all } i \in \Sigma_*. \tag{10}$$

To complete the construction of our random metric on T it remains to show that R^* can be extended uniquely to a metric R on T such that (T, R) is the completion of (V^*, R^*), and that the topology induced by R on T is the same as that induced by the Euclidean metric. Before demonstrating that this is the case, we prove some preliminary results about the diameter of sets of the form $(\psi_i(V^*))_{i \in \Sigma_*}$ in the metric R^*. Let us start with a simple chaining lemma.

Lemma 6. *If $x \in V^0$ and $y \in V^n$, then we can find a sequence $x_0, \dots, x_m$, with $x_0 = x$, $x_m = y$, and $\{x_{l-1}, x_l\} \in \bigcup_{n'=0}^n E^{n'}$ for $l \in \{1, \dots m\}$, and such that, for $n' \leq n$, $\{x_{l-1}, x_l\} \in E^{n'}$ for at most two of the $l \in \{1, \dots m\}$.*

Proof. The proof of this lemma is elementary and we present only an outline here. If $y \in V^n$ then we can connect y to a vertex of V^{n-1} using not more than two edges in E^n. Proceeding inductively gives us (the reverse of) a sequence with the desired properties.

Throughout the remainder of the article, for a subset A of a metric space (X, d), we shall denote the diameter of A by $\mathrm{diam}_d A := \sup\{d(x, y) : x, y \in A\}$.

Lemma 7. *If we assume that Assumption 1 holds, then $\mathrm{E}((\mathrm{diam}_{R^*} V^*)^d) < \infty$ for all $d \geq 0$.*

Proof. Let $x \in V^0$ and $y \in V^*$. Necessarily, $y \in V^n$ for some $n \geq 0$. Thus, the description of paths in V^* that was proved in Lemma 6 and the triangle inequality imply that

$$R^*(x, y) \leq 2 \sum_{m=0}^n \sup_{i \in \Sigma_m} R^*(\psi_i(0,0), \psi_i(1,0)) = 2 \sum_{m=0}^n \sup_{i \in \Sigma_m} l(i)R_i,$$

where the equality follows from (10). Consequently, the diameter of (V^*, R^*) is bounded above by $4 \sum_{m=0}^\infty \sup_{i \in \Sigma_m} l(i)R_i$. Applying this estimate, the result can be deduced from Theorem 1.

We now introduce random variables $(W_i)_{i \in \Sigma_*}$ to represent the normalised diameters of the sets $(\psi_i(V^*))_{i \in \Sigma_*}$. In particular, set $W_i := l(i)^{-1}\mathrm{diam}_{R^*}\psi_i(V^*)$ whenever the resistance metric R^* is defined. The definition of the σ-algebras $\mathcal{F}_n := \sigma(w(i) : |i| \leq n)$ can be recalled from Section 2.

Lemma 8. *If Assumption 1 holds, then $(W_i)_{i \in \Sigma_*}$ are identically distributed and satisfy $\mathrm{E}(W_i^d) < \infty$ for every $i \in \Sigma_*$, $d > 0$, and $W_i \perp \mathcal{F}_{|i|}$ for every $i \in \Sigma_*$.*

Proof. The self-similarity of T and the multiplicative cascade $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$ immediately implies that W_i has the same distribution as $\mathrm{diam}_{R^*} V^*$ for every $i \in \Sigma_*$. From this we obtain the first claim of the lemma and, when combined with Lemma 7, finite moments of the W_i. The remaining result is a straightforward application of the independence properties of the multiplicative cascade $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$.

It is now easy to show that the diameters of the sets $(\psi_i(V^*))_{i \in \Sigma_*}$ decrease to 0 uniformly as $|i| \to \infty$.

Lemma 9. *If Assumption 1 holds, then $\sup_{i \in \Sigma_n} \mathrm{diam}_{R^*}\psi_i(V^*) \to 0$, P-a.s.*

Proof. This result is a simple corollary of the definition of $(W_i)_{i \in \Sigma_*}$, Lemma 8, and the corresponding multiplicative cascade result, Lemma 1(b).

This uniform decay of the diameter of the sets $(\psi_i(V^*))_{i \in \Sigma_*}$ allows us to extend the definition of R^* to the whole of T. Note that T is the closure of V^* with respect to the Euclidean metric (see [16, Theorem 1.1.7]), which we will now denote by ρ. Hence, for any $x, y \in T$, there exist sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ in V^* with $\rho(x_n, x) \to 0$ and $\rho(y_n, y) \to 0$. Define

$$R(x, y) := \lim_{n \to \infty} R^*(x_n, y_n).$$

Before stating and proving Theorem 3, in which we show that this is a sensible definition for R, we introduce the notation $T_i := \psi_i(T)$ and, for $x \in T$,

$$T_n(x) := \bigcup \{T_i : i \in \Sigma_n, x \in T_i\}, \qquad \tilde{T}_n(x) := \bigcup \{T_i : i \in \Sigma_n, T_i \cap T_n(x) \neq \varnothing\}. \quad (11)$$

We will also apply the following result of [16, Proposition 1.3.6].

Lemma 10. *The collection $(T_n(x))_{n \in \mathbb{N}}$ forms a base of neighbourhoods of x with respect to ρ.*

Theorem 3. *Assume that Assumption 1 holds. Then R is a well-defined metric on T, topologically equivalent to the Euclidean metric,* P*-a.s.*

Proof. Under the assumptions of the theorem, the argument that we give holds P-a.s. Let $x, y \in T$ and suppose that there exist, for $m = 1, 2$, sequences (x_n^m) and (y_n^m) in V^* such that $\rho(x_n^m, x) \to 0$ and $\rho(y_n^m, y) \to 0$. Fix $\varepsilon > 0$. By Lemma 9, we can choose $n_0 \geq 0$ such that $\sup_{i \in \Sigma_n} \operatorname{diam}_{R^*} \psi_i(V^*) < \varepsilon/2$ for every $n \geq n_0$. Furthermore, we can use the convergence of the sequences and Lemma 10 to show that there exists $n_1 \geq n_0$ such that $x_n^m \in T_{n_0}(x)$ and $y_n^m \in T_{n_0}(y)$, for $m = 1, 2$ and $n \geq n_1$. Thus,

$$|R^*(x_n^1, y_n^1) - R^*(x_{n'}^m, y_{n'}^m)| \leq R^*(x_n^1, x_{n'}^m) + R^*(y_n^1, y_{n'}^m) < \varepsilon$$

for $m = 1, 2$ and $n, n' \geq n_1$. Taking $m = 1$, this implies that $(R^*(x_n^1, y_n^1))_{n \geq 0}$ is a Cauchy sequence and has a limit. Taking $m = 2$ and $n' = n$, this implies that the limit is unique and so the function R is well-defined on $T \times T$. It immediately follows, from the fact that R^* is a metric on V^*, that R is positive, symmetric, and satisfies the triangle inequality. To prove that R is a metric on T, it remains to show that $R(x, y) = 0$ implies that $x = y$. We shall prove the stronger claim, i.e. that $R(x_n, y) \to 0$ implies that $\rho(x_n, y) \to 0$.

Suppose that $(x_n)_{n \geq 0}$ is a sequence in T with $R(x_n, y) \to 0$ for some $y \in T$. Fix $\varepsilon > 0$, and choose n_0 such that $2^{1-n_0} \operatorname{diam}_\rho T < \varepsilon$. For $z \notin \tilde{T}_{n_0}(y)$, we must have $z \in T_i$ and $y \in T_j$, for some $i, j \in \Sigma_{n_0}$ with $T_i \cap T_j = \varnothing$. For any $z' \in \psi_i(V^*)$ and $y' \in \psi_j(V^*)$, using the additivity along paths of the metric R^* and the fact that the sets $(T_k)_{k \in \Sigma_{n_0}}$ only intersect at vertices of V^{n_0} (see [16, Proposition 1.3.5]), it is possible to show that

$$R(z', y') = R^*(z', y') \geq \inf_{k \in \Sigma_{n_0}} R^*(\psi_k(0, 0), \psi_k(1, 0)) =: c.$$

It follows that $R(z, y) \geq c$. Since $c > 0$ and $R(x_n, y) \to 0$, there exists an n_1 such that $R(x_n, y) < c$ for all $n \geq n_1$. Consequently, $x_n \in \tilde{T}_{n_0}(y)$ for $n \geq n_1$. By our choice of n_0 and the fact that the contraction ratios of $(\psi_i)_{i \in S}$ (with respect to ρ) are no bigger than $\frac{1}{2}$, this implies that $\rho(x_n, y) < \varepsilon$ for $n \geq n_1$. Hence, $\rho(x_n, y) \to 0$.

To prove the equivalence of the metrics, it remains to show that, for all sequences $(x_n)_{n \in \mathbb{N}}$ in T with $\rho(x_n, x) \to 0$ for some $x \in T$, we have $R(x_n, x) \to 0$. We note that if $y \in T_i$ then there exists a sequence $(y_n)_{n \in \mathbb{N}} \in \psi_i(V^*)$ with $\rho(y_n, y) \to 0$. Consequently,

$$\operatorname{diam}_R T_i = \sup_{x, y \in T_i} R(x, y) = \sup_{x, y \in \psi_i(V^*)} R^*(x, y) = \operatorname{diam}_{R^*} \psi_i(V^*). \quad (12)$$

Applying this fact and Lemma 9, we find that, given $\varepsilon > 0$, there exists an n_0 such that $\sup_{i \in \Sigma_{n_0}} \operatorname{diam}_R T_i < \varepsilon$. By Lemma 10, we find that $x_n \in T_{n_0}(x)$ for all $n \geq n_1$, for some n_1. It follows that $R(x_n, x) < \varepsilon$ for all $n \geq n_1$, and so $R(x_n, x) \to 0$ as desired.

As noted above, the dendrite T is the closure of V^* with respect to ρ. Thus, under Assumption 1, Theorem 3 implies that (T, R) is the completion of (V^*, R) and is a dendrite, P-a.s. Furthermore, because T is a dendrite, it is possible to check, from the definition of R^* as a resistance metric, that R must be a shortest path metric (additive along paths); see [15].

4. The Hausdorff dimension of T

For a metric space, (X, d), the *Hausdorff dimension* of $F \subseteq X$ is defined by

$$\dim_H(F) = \inf\{s \colon \mathcal{H}^s(F) < \infty\} = \sup\{s \colon \mathcal{H}^s(F) > 0\}, \tag{13}$$

where

$$\mathcal{H}^s(F) := \lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{\infty} (\operatorname{diam}_d U_i)^s \colon \{U_i\}_i \text{ is a } \delta\text{-cover of } F \right\}$$

and, for $\delta > 0$, we call a finite or countable family of sets $(U_i)_{i=1}^{\infty}$ a δ-*cover* of $F \subseteq X$ if $\operatorname{diam}_d U_i \leq \delta$ for all i and $F \subseteq \bigcup_i U_i$. If we suppose that Assumption 1 holds, then by Theorem 3 we can construct the metric space (T, R), P-a.s. In this section we will present further conditions on the scaling factors $(w(i))_{i \in S}$ that allow us to calculate the Hausdorff dimension of (T, R) to be constant and equal to α, P-a.s., where α, as in Lemma 3, is defined to be the unique positive solution to $F(\theta) = 1$ for the function $F(\theta) := \mathrm{E}(\sum_{i \in S} w(i)^{\theta})$. Note that Assumption 1 implies that $\alpha \in (1, \infty)$. We start by demonstrating the upper bound, for which Assumption 1 is sufficient.

Theorem 4. *Suppose that Assumption 1 holds, then* $\dim_H(T) \leq \alpha$, *P-a.s.*

Proof. The representation of $\operatorname{diam}_R T_i$ in (12) allows us to apply Lemma 9 to deduce that, for large n, $(T_i)_{i \in \Sigma_n}$ is a δ-cover for T. Thus,

$$\mathrm{E}(\mathcal{H}^{\theta}(T)) \leq \mathrm{E}\left(\liminf_{n \to \infty} \sum_{i \in \Sigma_n} (\operatorname{diam}_R T_i)^{\theta} \right) \leq \liminf_{n \to \infty} \mathrm{E}\left(\sum_{i \in \Sigma_n} (l(i) W_i)^{\theta} \right),$$

where we have applied Fatou's lemma, (12), and the definition of W_i to obtain the second inequality. By Lemma 8, the expectation appearing in the right-hand side is equal to $F(\theta)^n \mathrm{E}(W_{\varnothing}^{\theta})$, and the second of these factors is finite. Furthermore, for $\theta > \alpha$, $F(\theta) < F(\alpha) = 1$, and from this we can deduce that $\mathcal{H}^{\theta}(T) = 0$, P-a.s. The result follows using the characterisation of the Hausdorff dimension in (13). $\qquad\qquad\square$

To prove that the α defined above is also a lower bound for the Hausdorff dimension of (T, R), we need to make more restrictive assumptions on the scaling factors, and we will derive the result in two special cases only. We proceed by applying a standard density result (see [10, Proposition 4.9]), and the first step in doing this involves constructing a suitable measure on T. The measure on T that will be useful for our purposes will be a natural stochastically self-similar measure, and to construct it we proceed as in [12], initially defining a measure on the shift-space Σ and then applying the natural projection onto T. To characterise a measure on Σ, it is sufficient to define it on the cylinder sets, $i\Sigma := \{ij \colon j \in \Sigma\}$ for $i \in \Sigma_*$. First, let M_i^{θ}

be defined by (3) and also $M^\theta := M^\theta_\varnothing$. Under Assumption 1 we can apply Lemma 3 to obtain, for $\theta \le \alpha$, $P(M_\theta \in (0, \infty)) = 1$. Thus, if Assumption 1 holds, $\theta \le \alpha$, and we define $\tilde{\mu}^\theta$ by

$$\tilde{\mu}^\theta(i\Sigma) = \frac{M^\theta_i l(i)^\theta}{M^\theta F(\theta)^n}, \qquad i \in \Sigma_n,$$

then it is easy to apply the decomposition identity (4) to check that $\tilde{\mu}^\theta$ can be extended uniquely to a probability measure on Σ. Taking $\theta = \alpha$ removes the dependency on the index length, which suggests that this is the natural exponent to choose.

The projection π from Σ to T that we apply is obtained from the fact that, for each $i \in \Sigma$, there exists a unique $\pi(i) \in T$ such that $\{\pi(i)\} = \bigcap_{n=1}^\infty T_{i \mid n}$; see [16, Proposition 1.3.3]. It is the measure $\mu^\alpha := \tilde{\mu}^\alpha \circ \pi^{-1}$ that we will utilise in proving lower bound results for the Hausdorff dimension of T. Since the map π is standard, we will not discuss its measurability (it is in fact a continuous surjection [16]); however, we will note that it is possible to check that μ^α is a nonatomic Borel probability measure on (T, R), which satisfies

$$\mu^\alpha(T_i) = \frac{M^\alpha_i l(i)^\alpha}{M^\alpha} \tag{14}$$

for $i \in \Sigma_*$ (at least P-a.s. when Assumption 1 holds).

So far we have been able to use the fixed graphs (V^n, E^n) to approximate T. The problem with these discretisations is that the lengths $l(i)R_i$ of edges within the graphs are, in general, widely varying as $n \to \infty$. To try to limit this effect, it will be useful to consider graph approximations to T for which we have some uniform control over the edge lengths. This technique was also applied in [11, Section 4], for example. Let us start by saying that $\Lambda \subseteq \Sigma_*$ is a cut-set if, for every $i \in \Sigma$, there exists a unique $j \in \Lambda$ with $i \mid |j| = j$, and there exists an n such that $|j| \le n$ for all $j \in \Lambda$. This final condition is included to ensure that there is only a countable number of cut-sets. In particular, we will be interested in the random cut-sets $(\Sigma_\delta)_{\delta > 0}$ defined by

$$\Sigma_\delta := \{i : l(i) \le \delta < l(i \mid (|i| - 1))\}.$$

Note that if Assumption 1(b) holds, then Lemma 1(b) guarantees that this is indeed a cut-set for all $\delta > 0$, P-a.s. We introduce the corresponding graphs (V^δ, E^δ), defined by

$$V^\delta := \bigcup_{i \in \Sigma_\delta} \psi_i(V^0) \quad \text{and} \quad E^\delta := \{\psi_i(V^0) : i \in \Sigma_\delta\}.$$

It is easy to check that, for each $\delta > 0$, (V^δ, E^δ) is a graph tree, P-a.s. Heuristically, to construct (V^δ, E^δ) we start with (V^0, E^0) and if there is an edge $\psi_i(V^0)$ in the graph with $l(i)$ greater than δ, then replace it by the three edges $(\psi_{ij}(V^0))_{j \in S}$ and continue until there are no edges left to replace.

It will be important to be able to estimate the μ^α-measure of balls of the form $B_R(x, \delta) := \{y \in T : R(x, y) < \delta\}$. To do this, we will use collections of the sets $(T_i)_{i \in \Sigma_\delta}$ to cover the balls $B_R(x, \delta)$. In a slight change of notation from (11), for $x \in T$, define

$$T_\delta(x) := \bigcup \{T_i : i \in \Sigma_\delta, x \in T_i\}$$

and a larger neighbourhood of x by

$$\tilde{T}_{\delta, \varepsilon}(x) := \bigcup \{T_i : i \in \Sigma_\delta, R(T_i, T_\delta(x)) < \delta\varepsilon\}.$$

The number of sets making up this union is $N_{\delta,\varepsilon}(x) := \#\{i \in \Sigma_\delta : T_i \subseteq \tilde{T}_{\delta,\varepsilon}(x)\}$. It is clear that $B_R(x,\delta) \subseteq \tilde{T}_{\delta/\varepsilon,\varepsilon}(x)$, and recalling from (14) that $\mu^\alpha(T_i) = M_i^\alpha/(M^\alpha l(i)^\alpha)$ for $i \in \Sigma_*$, we therefore obtain

$$\mu^\alpha(B_R(x,\delta)) \leq (M^\alpha)^{-1}\varepsilon^{-\alpha}\delta^\alpha N_{\delta/\varepsilon,\varepsilon}(x) \sup_{i \in \Sigma_{\delta/\varepsilon}} M_i^\alpha. \tag{15}$$

To complete the argument, we will analyse the behaviour of $\sup_{i \in \Sigma_\delta} M_i^\alpha$ and $N_{\delta,\varepsilon}(x)$. In bounding the first of these terms, we require control over the growth of the mean of $|\Sigma_\delta|$. The next lemma provides this using a related age-dependent branching process.

Lemma 11. *Under Assumption 1(a),* $\mathrm{E}\,|\Sigma_\delta| \leq c\delta^{-\alpha}$ *for some constant* c.

Proof. Consider the following branching process. Start at time 0 with one particle, labelled $\varnothing$. A particle, i, has three children at times $(\sigma_i - \ln w(ij))_{j \in S}$, where $\sigma_i = -\ln l(i)$ is the birth time of i. Label the child born to i at $\sigma_i - \ln w(ij) \equiv -\ln l(ij)$ by ij, noting that children may not be labelled in birth order. It is not necessary to define the time of dying explicitly in this proof. The independence assumptions on the $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$ means that this setup describes a general branching process in the sense of [13, Chapter 6]; furthermore, if Y_t is defined to be the random variable counting the births before time t then it is easy to check that $|\Sigma_\delta| \leq 3Y_{-\ln\delta}$. Noting that the Malthusian parameter for the branching process is precisely the α defined by $\mathrm{E}(\sum_{i \in S} w(i)^\alpha) = 1$, standard arguments then show that $\mathrm{E}\,Y_t \leq c_1 e^{\alpha t}$ for some constant c_1. Combining these facts yields the lemma. $\qquad\blacksquare$

We now proceed with bounding the growth of $\sup_{i \in \Sigma_\delta} M_i^\alpha$ as $\delta \to 0$. To allow us to apply Borel–Cantelli arguments to deduce P-a.s. such as this, it is useful to choose a particular subsequence of δs to investigate. Henceforth, we consider $(\delta_n)_{n \geq 0}$ defined by $\delta_n := e^{-n}$.

Lemma 12. *If Assumption 1(a) holds and* $\beta > 0$, *then* $\lim_{n \to \infty} \delta_n^\beta \sup_{i \in \Sigma_{\delta_n}} M_i^\alpha = 0$, *P-a.s.*

Proof. If $\mathrm{P}(\sum_{i \in S} w(i) = 1) = 1$ then $M_i^\alpha \equiv 1$, P-a.s. for all i and so the result is obvious. Now assume that $\mathrm{P}(\sum_{i \in S} w(i) = 1) < 1$. For each $i \in \Sigma_*$, define a subset of Σ_* by $\Sigma_i := \{ik : k \in \Sigma_*\} \setminus \{i\}$ and related σ-algebras $\mathcal{F}_i := \sigma(w(j) : j \in \Sigma_i)$ and $\mathcal{G}_i := \sigma(w(j) : j \in \Sigma_* \setminus \Sigma_i)$. By definition, we have $\mathcal{F}_i \perp \mathcal{G}_i$. It is straightforward to check that M_i^α is $\mathcal{F}_i$-measurable and $\{\Sigma_\delta = \Lambda\} \in \mathcal{G}_i$ for any cut-set Λ containing i. Thus, for $i \in \Lambda$, with Λ a cut-set, we have

$$\mathrm{P}(M_i^\alpha > \eta, \Sigma_\delta = \Lambda) = \mathrm{P}(\Sigma_\delta = \Lambda)\,\mathrm{P}(M^\alpha > \eta) \quad \text{for } \eta \geq 0.$$

From this we deduce, using the countability of cut-sets, that

$$\begin{aligned}
\mathrm{P}\left(\sup_{i \in \Sigma_\delta} M_i^\alpha > \eta\right) &= \sum_{\Lambda:\,\Lambda \text{ a cutset}} \mathrm{P}\left(\sup_{i \in \Sigma_\delta} M_i^\alpha > \eta, \Sigma_\delta = \Lambda\right) \\
&\leq \sum_{\Lambda:\,\Lambda \text{ a cutset}} \sum_{i \in \Lambda} \mathrm{P}(M_i^\alpha > \eta, \Sigma_\delta = \Lambda) \\
&= \sum_{\Lambda:\,\Lambda \text{ a cutset}} \sum_{i \in \Lambda} \mathrm{P}(M^\alpha > \eta)\,\mathrm{P}(\Sigma_\delta = \Lambda) \\
&= \mathrm{P}(M^\alpha > \eta) \sum_{\Lambda:\,\Lambda \text{ a cutset}} |\Lambda|\,\mathrm{P}(\Sigma_\delta = \Lambda) \\
&= \mathrm{P}(M^\alpha > \eta)\,\mathrm{E}\,|\Sigma_\delta|.
\end{aligned}$$

Since M^α is the limit of a multiplicative cascade, we can check the conditions of [17, Theorem 2.1] to deduce that its distribution satisfies $P(M_\alpha > \eta) \leq \exp(-c_1 \eta^\gamma)$ for some constants c_1 and γ. Applying this and Lemma 11, we obtain

$$P(\sup_{i \in \Sigma_\delta} M_i^\alpha > \eta) \leq c_2 \exp(-c_1 \eta^\gamma) \delta^{-\alpha},$$

and the result can subsequently be obtained using the Borel–Cantelli lemma.

We are now in a position to be able to prove the lower Hausdorff dimension bound in our first special case. To prove the result we assume that $w(1) + w(2) = 1$, P-a.s., which is a strengthening of Assumption 1(a), and also implies that the resistance perturbations are precisely 1, thereby eliminating one random variable from our consideration. We also assume that Assumption 1(b) holds to allow us to construct the metric space (T, R), P-a.s. The assumption that is most specifically related to the problems which arise in the computation of a lower bound for the Hausdorff dimension, however, is a uniform lower bound for $(w(i))_{i \in S}$. Calculations of this kind become difficult if parts of the fractal become, in some sense, too small too quickly, and by bounding the scaling factors uniformly below, we are able to prevent this from occurring here.

Theorem 5. *Suppose that $w(1) + w(2) = 1$, P-a.s., Assumption 1(b) holds, and there exists an $\varepsilon > 0$ such that $P(w(i) > \varepsilon, \ i \in S) = 1$, then $\dim_H(T) = \alpha$, P-a.s.*

Proof. From [16, Proposition 1.3.5], we know that the intersection of distinct sets from $(T_i)_{i \in \Sigma_\delta}$ can only happen at points in V^δ. Consequently, if $i, j \in \Sigma_\delta$ then the distance between the sets T_i and T_j is

$$R(T_i, T_j) = \min_{\substack{x \in V_i^0, y \in V_j^0}} \min_{\substack{i_1, \ldots, i_n \in \Sigma^\delta : x_0 \in \psi_{i_1}(V^0), y \in \psi_{i_n}(V^0), \\ \psi_{i_m}(V^0) \cap \psi_{i_{m+1}}(V^0) \neq \varnothing}} \sum_{m=1}^n l(i_m) R_{i_m}, \tag{16}$$

where the second minimum on the right-hand side is the graph distance in V^δ and E^δ between the vertices x and y when an edge of the form $\psi_i(V^0)$ is weighted by $l(i)R_i$, and that this is equal to $R(x, y)$ is a result of the fact that R is additive along paths. Since we have $\delta \varepsilon < l(i)$ for every $i \in \Sigma_\delta$, all the edge weights in (V^δ, E^δ) are greater than $\delta \varepsilon$ (using the fact that $R^i \equiv 1$). Thus, if $R(T_i, T_j) < \delta \varepsilon$, then $R(T_i, T_j) = 0$. This means that, for $x \in T$, every $T_i \subseteq \tilde{T}_{\delta, \varepsilon}(x)$ has a nonzero intersection with some $T_j \subseteq T_\delta(x)$ (where $i, j \in \Sigma_\delta$). A simple consideration of the geometry of T allows us to deduce that we must therefore have $N_{\delta, \varepsilon}(x) \leq 9$ for all $x \in T$, P-a.s. Thus, applying the bound in (15), for $r \in [\varepsilon \delta_{n+1}, \varepsilon \delta_n)$, we obtain

$$\frac{\mu^\alpha(B_R(x, r))}{r^s} \leq 9 \varepsilon^{-s} e^s M_\alpha^{-1} \delta_n^{\alpha-s} \sup_{i \in \Sigma_{\delta_n}} M_\alpha^i.$$

If $s < \alpha$, then Lemma 12 therefore gives $\limsup_{r \to 0} r^{-s} \mu^\alpha(B_R(x, r)) = 0$, and so [10, Proposition 4.9] implies the result.

For the second special case in which we prove a Hausdorff dimension lower bound, we make the following assumption.

Assumption 2. *The random variables $(w(i))_{i \in S}$ are independent and their distributions satisfy the following tail condition. If $p \in (0, 1)$, there exists a constant $\varepsilon > 0$ such that*

$$P(w(i) \leq x\varepsilon \mid w(i) \leq x) \leq p \quad \text{for all } x \in (0, 1], \ i \in S.$$

Again, this is an assumption which stops the fractal getting too small too quickly. Rather than bounding them uniformly below, as in Theorem 5, we assume independence of the scaling factors and restrict the amount of build up of mass close to 0 in the distributions of the scaling factors. This independence allows us to use a percolation-type argument, which enables us to avoid having to impose a uniform lower bound. If $w(i)$ has distribution function Φ, then the inequality in Assumption 2 is equivalent to $\Phi(\varepsilon x) \leq p\Phi(x)$ for every $x \in (0, 1]$. From this, it is easy to see that if Φ is approximately polynomial (i.e. there exist constants c_1 and c_2 such that $c_1 x^n \leq \Phi(x) \leq c_2 x^n$), then Assumption 2 holds. An example of when the build-up of mass is too great for this to hold is the distribution function $\Phi(x) = (1 - \ln x)^{-1}$. The following lemma gives another alternative characterisation of Assumption 2 that will prove to be useful in obtaining negative moments for the resistance perturbations $(R_i)_{i \in \Sigma_*}$.

Lemma 13. *Let X be a $(0, 1]$-valued random variable with distribution function Φ, then the following statements are equivalent.*

(a) *If $p \in (0, 1)$, then there exists an $\varepsilon \in (0, 1)$ such that $\Phi(x\varepsilon) \leq p\Phi(x)$ for $x \in (0, 1]$.*

(b) *There exist constants $\varepsilon \in (0, 1)$ and $\beta > 0$ such that*

$$\mathrm{E}\left(\left(1 - \frac{x^\beta \varepsilon}{X^\beta}\right)\mathbf{1}_{\{X \leq x\}}\right) \geq 0 \quad \textit{for all } x \in (0, 1].$$

Proof. Assume that part (a) holds and fix $p \in (0, 1)$. Choose $\varepsilon \in (0, 1)$ such that $\Phi(x\varepsilon) \leq p\Phi(x)$ holds, and let $\beta > 0$ satisfy $p < \varepsilon^\beta$. Integration by parts yields

$$\mathrm{E}(x^\beta X^{-\beta} \mathbf{1}_{\{X \leq x\}}) = \lim_{\delta \to 0}\left\{[x^\beta y^{-\beta}\Phi(y)]_{y=\delta}^x + \beta \int_\delta^x x^\beta y^{-\beta-1}\Phi(y)\,\mathrm{d}y\right\}.$$

Now, $\Phi(\varepsilon^n) \leq p^n$, and so, for $y \in (\varepsilon^{n+1}, \varepsilon^n]$, we have $y^{-\beta}\Phi(y) \leq \varepsilon^{-\beta(n+1)}p^n$. It follows that, because $p\varepsilon^{-\beta} < 1$, $\lim_{\delta \to 0}[x^\beta y^{-\beta}\Phi(y)]_{y=\delta}^x = \Phi(x)$, which gives an alternative expression for the first term in the limit above. Furthermore,

$$\begin{aligned}
\lim_{\delta \to 0} \beta \int_\delta^x x^\beta y^{-\beta-1}\Phi(y)\,\mathrm{d}y &= \sum_{n=0}^\infty \beta x^\beta \int_{x\varepsilon^{n+1}}^{x\varepsilon^n} y^{-\beta-1}\Phi(y)\,\mathrm{d}y \\
&\leq \beta x^\beta \sum_{n=0}^\infty \int_{x\varepsilon^{n+1}}^{x\varepsilon^n} (x\varepsilon^{n+1})^{-\beta-1} p^n \Phi(x)\,\mathrm{d}y \\
&\leq \beta \Phi(x)\varepsilon^{-\beta-1} \sum_{n=0}^\infty (p\varepsilon^{-\beta})^n \\
&= \frac{\beta\Phi(x)}{\varepsilon^{\beta+1}(1 - p\varepsilon^{-\beta})}.
\end{aligned}$$

Hence, $\mathrm{E}(x^\beta X^{-\beta} \mathbf{1}_{\{X \leq x\}})$ is bounded above by a constant multiple of $\Phi(x)$, and therefore part (b) holds.

Conversely, suppose that part (b) holds for some $\varepsilon \in (0, 1)$ and $\beta > 0$. Fix $p \in (0, 1)$ and define $\varepsilon' := (p\varepsilon)^{1/\beta}$. For $x \in (0, 1]$, we have

$$\Phi(\varepsilon' x) \leq \mathrm{E}\left(\mathbf{1}_{\{X \leq \varepsilon' x\}} \frac{\varepsilon'^\beta x^\beta}{X^\beta}\right) \leq p\,\mathrm{E}\left(\mathbf{1}_{\{X \leq x\}} \frac{\varepsilon x^\beta}{X^\beta}\right) \leq p\Phi(x),$$

which is statement (a).

Lemma 14. *Under Assumptions 1(a) and 2, we have* $E((R_\varnothing)^{-d}) < \infty$ *for every* $d > 0$.

Proof. The result will follow from Corollary 2 if we can show that $E(w(i)^{-\beta}) < \infty$, $i = 1, 2$, for some $\beta > 0$. Under Assumption 2, this is a simple consequence of the previous lemma.

We now use the alternative description of Assumption 2 provided by Lemma 13 to show that the inequality in Assumption 2 still holds if the $w(i)$ are multiplied by the resistance perturbations R_i. We shall use the ε_0 obtained in the following lemma to describe what constitutes a small edge of (V^δ, E^δ).

Lemma 15. *If Assumption 1(a) and 2 hold and* $p \in (0, 1)$, *then there exists an* $\varepsilon_0 > 0$ *such that*

$$P(w(i)R_i \leq \varepsilon_0 x \mid w(i) \leq x) \leq p \quad \text{for all } x \in (0, 1], i \in S.$$

Proof. By Lemma 13, we can choose $\varepsilon, \beta > 0$ such that

$$E\left(\left(1 - \frac{\varepsilon x^\beta}{w(i)^\beta}\right) \mathbf{1}_{\{w(i) \leq x\}}\right) \geq 0 \quad \text{for all } x \in (0, 1], i \in S.$$

Also note that, by Lemma 14, $E((R_i)^{-\beta}) < \infty$. Hence,

$$P(w(i)R_i \leq \varepsilon_0 x, w(i) \leq x) \leq E\left(\left(\frac{\varepsilon_0 x}{w(i)R_i}\right)^\beta \mathbf{1}_{\{w(i) \leq x\}}\right)$$

$$= \varepsilon_0^\beta E(R_\varnothing^{-\beta}) E\left(\frac{x^\beta}{w(i)^\beta} \mathbf{1}_{\{w(i) \leq x\}}\right)$$

$$\leq \frac{\varepsilon_0^\beta}{\varepsilon} E(R_\varnothing^{-\beta}) P(w(i) \leq x),$$

and so the result holds for ε_0 chosen suitably small.

Henceforth we consider p to be a deterministic constant and choose ε_0 so that the claim of Lemma 15 holds. For reasons that will become clear in the proof of Lemma 17, below, it will be convenient to assume that $p \in (0, 3^3/4^4)$.

We now look to bound $N_{\delta, \varepsilon_0}(x)$, and to do so it will be convenient to use the language of percolation theory. We first define the events $(A_i^\delta)_{i \in \Sigma_\delta}$ by

$$A_i^\delta := \{l(i)R_i \leq \varepsilon_0 \delta\}.$$

For $i \in \Sigma_\delta$, we call the set T_i *open* if A_i^δ occurs, and *closed* otherwise. Thus, T_i being open corresponds to $\psi_i(V^0)$ being a small edge in the graph (V^δ, E^δ) (when an edge of this form is weighted by $l(i)R_i$). We will show that the largest cluster of sets from $(T_i)_{i \in \Sigma_\delta}$ which are open is not too large, and explain how this fact gives us a useful estimate for $N_{\delta, \varepsilon_0}(x)$.

Consider the random variable

$$H_\delta := (\Sigma_\delta; (l(i \mid (|i| - 1)))_{i \in \Sigma_\delta}).$$

We shall be conditioning on H_δ, the informal motivation for doing so is the following. In the proof of Lemma 11 we introduced a branching process where the individual i is born at time $-\ln l(i)$. Hence, if we stop the branching process at time $-\ln \delta$ (and cannot see into the future) then we will be able to ascertain the value of H_δ. However, we will not be able to observe the exact values of $l(i)$ or R_i for $i \in \Sigma_\delta$. So, in this sense, we can consider H_δ to be the information about the weighted graph (V^δ, E^δ) available at time $-\ln \delta$.

We now make precise the nature of the percolation-type behaviour that the independence of the $w(i)$s under Assumption 2 induces on the open/closed sets of $(T_i)_{i \in \Sigma_\delta}$. Note that the result provides an upper bound on the probability of a set from $(T_i)_{i \in \Sigma_\delta}$ being open which is independent of δ. This scale-invariance property will be of particular importance for the arguments that follow.

Lemma 16. *Suppose that Assumptions 1 and 2 hold. Let $\delta \in (0, 1)$. Conditionally on H_δ, the sets $(T_i)_{i \in \Sigma_\delta}$ are open/closed independently. Furthermore, for $i \in \Sigma_\delta$, we have*

$$P(A_i^\delta \mid H_\delta) \leq p, \quad \text{P-}a.s.,$$

and, for $s \geq 1$, we have

$$E(s^{\mathbf{1}_{A_i^\delta}} \mid H_\delta) \leq 1 - p + sp, \quad \text{P-}a.s.$$

Proof. Suppose that $i^1, \ldots, i^n$ are distinct elements of Σ_δ. Applying the independence of the $(w(i))_{i \in \Sigma_* \setminus \{\varnothing\}}$, elementary arguments yield

$$P(A_{i^1}^\delta, \ldots, A_{i^n}^\delta \mid H_\delta) = \prod_{m=1}^n P\left(w(i^m)R_{i^m} \leq \frac{\varepsilon_0 \delta}{x} \;\middle|\; w(i^m) \leq \frac{\delta}{x}\right)_{x = l(i^m \mid (|i^m| - 1))}.$$

This implies the independence claim. Consider the case in which $n = 1$, and write $i = i^1$. Since $i \in \Sigma_\delta$, we must have $l(i \mid (|i| - 1)) > \delta$. Hence, we can apply the bound of Lemma 15 to the above expression to obtain $P(A_i^\delta \mid H_\delta) \leq p$, P-a.s. The generating function bound is a simple consequence of this.

We now introduce an algorithm to find the largest cluster of open sets of the form $(T_i)_{i \in \Sigma_\delta}$. We shall work on the graph $(\Sigma_\delta, \Gamma_\delta)$, where the edge set Γ_δ is defined by

$$\Gamma_\delta := \{\{i, j\} \colon i, j \in \Sigma_\delta, T_i \cap T_j \neq \varnothing, T_i, T_j \text{ open}\}.$$

We shall write $\mathcal{C}(i)$ for the component of $(\Sigma_\delta, \Gamma_\delta)$ which contains the vertex i. Clearly, if T_i is closed, then $\mathcal{C}(i) = \{i\}$. The following argument to find the size of the largest cluster is inspired by similar procedures used in [14] to find the size of the largest cluster of a random digraph, and in [2] to find the size of the largest cluster of a complete graph with edge percolation.

Assume that H_δ is known. Let $i \in \Sigma_\delta$ and set $L_0 := \{i\}$ and $D_0 := \varnothing$. For $n \geq 1$, we define L_n and D_n inductively. Assume that we are given L_n and D_n. If $L_n \neq \varnothing$ then pick a vertex $j \in L_n$ (we can assume that there is a deterministic rule for doing this), and set

$$L_{n+1} := L_n \cup \{k \in \Sigma_\delta \colon k \notin L_n \cup D_n, \{j, k\} \in \Gamma_\delta\} \setminus \{j\}, \qquad D_{n+1} := D_n \cup \{j\}.$$

If $L_n = \varnothing$ then set $L_{n+1} := \varnothing$ and $D_{n+1} := D_n$.

It is a little unclear from this description as to exactly what the algorithm is doing and so we now try to provide a more intuitive description in terms of a branching process related to Σ_δ. Call i a *live* vertex. For the first step, connect to i all those vertices in Σ_δ that are joined to i by an edge in Γ_δ. Call these vertices live and i *dead*. At an arbitrary stage, pick a live vertex, j, and connect to it all those vertices which we have not yet considered and are connected to j by an edge in Γ_δ. Call the new vertices in our branching process live and j dead. Continue until we have no live vertices to pick from. At the point of termination, the collection of dead vertices contains exactly the vertices of $\mathcal{C}(i)$.

In our notation, L_n represents the live vertices and D_n the dead ones. Since we can pick each vertex in Σ_δ only once in the algorithm, we must have $D_{|\Sigma_\delta|+1} = \mathcal{C}(i)$. However, the algorithm may effectively terminate before this stage, giving $|D_n| = n \wedge \tau$, where $\tau := \inf\{n\colon L_n = \varnothing\}$. Necessarily, $L_{|\Sigma_\delta|+1} = \varnothing$, and so this infimum is well-defined and finite. In particular, we must have $|\mathcal{C}(i)| = \tau$.

Using this algorithm, we are able to obtain a tail estimate for the distribution of $|\mathcal{C}(i)|$, conditional on H_δ. Note that this result is scale-invariant; the tail bound on the size of a cluster does not depend on δ.

Lemma 17. *Suppose that Assumptions 1 and 2 hold and let $\delta \in (0, 1)$. There exists a deterministic constant c, not depending on δ, such that, for $i \in \Sigma_\delta$,*

$$\mathrm{P}(|\mathcal{C}(i)| > n \mid H_\delta) \leq \mathrm{e}^{-cn}, \qquad \mathrm{P}\text{-}a.s.$$

Proof. Choose $i \in \Sigma_\delta$ and use the algorithm described prior to this lemma to construct $(L_n, D_n)_{n \geq 0}$. Given L_n and D_n, the number of new live vertices in the $(n+1)$th step of the algorithm is

$$Z_n := \#\{k \in \Sigma_\delta\colon k \notin L_n \cup D_n, \{j, k\} \in \Gamma_\delta\},$$

if $L_n \neq \varnothing$, where $j = j(L_n)$ is the vertex chosen from L_n in the algorithm, and 0 otherwise. On $\{L_n = \varnothing\}$, for $s \geq 1$,

$$\mathrm{E}(s^{Z_n} \mid H_\delta, L_n, D_n) = 1, \quad \mathrm{P}\text{-a.s.}$$

On $\{L_n \neq \varnothing\}$ with $j = j(L_n)$, using the independence and generating function bound of Lemma 16, for $s \geq 1$, we have

$$\mathrm{E}(s^{Z_n} \mid H_\delta, L_n, D_n) \leq \prod_{\substack{k \in \Sigma_\delta\colon k \notin L_n \cup D_n, \\ T_j \cap T_k \neq \varnothing}} \mathrm{E}(s^{\mathbf{1}_{\{T_k \text{ open}\}}} \mid H_\delta) \leq (1 - p + sp)^4, \quad \mathrm{P}\text{-a.s.}, \quad (17)$$

where for the second inequality we have applied the facts that at most three of the sets $(T_i)_{i \in \Sigma_\delta}$ intersect at any point, and also that T_j only intersects with other elements of $(T_i)_{i \in \Sigma_\delta}$ at points in $\psi_j(V^0)$. Hence, because this upper bound is larger than 1, we have

$$\mathrm{E}(s^{Z_n} \mid H_\delta, L_n, D_n) \leq (1 - p + sp)^4, \quad \mathrm{P}\text{-a.s.}$$

For $n \leq \tau$, we have $|L_n| = |L_{n-1}| + Z_{n-1} - 1$, and so, for $s \geq 1$, we have

$$
\begin{aligned}
\mathrm{E}(s^{|L_n|} \mathbf{1}_{\{|L_n|>0\}} \mid H_\delta) &\leq \mathrm{E}(s^{|L_n|} \mathbf{1}_{\{|L_{n-1}|>0\}} \mid H_\delta) \\
&= \mathrm{E}(s^{|L_{n-1}|} \mathbf{1}_{\{|L_{n-1}|>0\}} \mathrm{E}(s^{Z_{n-1}-1} \mid H_\delta, L_{n-1}, D_{n-1}) \mid H_\delta) \\
&\leq s^{-1}(1 - p + sp)^4 \, \mathrm{E}(s^{|L_{n-1}|} \mathbf{1}_{\{|L_{n-1}|>0\}} \mid H_\delta),
\end{aligned}
$$

where we use inequality (17) for the final bound and we have also used the fact that $\{|L_n| > 0\} = \{\tau > n\}$. Applying this repeatedly yields a P-a.s. upper bound of $s^{-n}(1 - p + sp)^{4n}$ for the expectation considered. Consequently, for $s \geq 1$, we obtain

$$\mathrm{P}(|\mathcal{C}(i)| > n \mid H_\delta) = \mathrm{P}(|L_n| > 0 \mid H_\delta) \leq \mathrm{E}(s^{|L_n|} \mathbf{1}_{\{|L_n|>0\}} \mid H_\delta) \leq s^{-n}(1 - p + ps)^{4n},$$

P-a.s. This is minimised by $s = (1-p)/3p$, which is greater than 1, because of the upper bound we have assumed on p. Substituting for this value of s we obtain an upper bound of $(4^4 3^{-3}(1-p)^3 p)^n$ for $\mathrm{P}(|\mathcal{C}(i)| > n \mid H_\delta)$, and the result follows. $\qquad \blacksquare$

Lemma 17 is easily extended to give a tail estimate for the distribution of the size of the largest component, $\mathcal{C}_\delta := \sup_{i \in \Sigma_\delta} \mathcal{C}(i)$, from which we can prove the following a.s. convergence result.

Lemma 18. *If Assumptions 1 and 2 hold, then* $\limsup_{n \to \infty} n^{-1} \mathcal{C}_{\delta_n} < \infty$, *P-a.s.*

Proof. Applying the conditional tail distribution of Lemma 17, we have

$$
P(\mathcal{C}_\delta > n) = E(P(\mathcal{C}_\delta > n \mid H_\delta)) \leq E\left(\sum_{i \in \Sigma_\delta} P(|\mathcal{C}(i)| > n \mid H_\delta) \right) \leq E(|\Sigma_\delta|) e^{-c_1 n},
$$

and so it is possible to deduce, from Lemma 11, a tail bound of the form $P(\mathcal{C}_\delta > n) \leq c_2 e^{-c_1 n} \delta^{-\alpha}$ for $\mathcal{C}_\delta$. Applying this, a simple Borel–Cantelli argument yields the lemma.

We are now able to prove the lower bound for the Hausdorff dimension of T in the second special case.

Theorem 6. *Assume that Assumptions 1 and 2 hold, then* $\dim_H(T) = \alpha$, *P-a.s.*

Proof. As in (16), the distance between sets of the form $(T_i)_{i \in \Sigma_\delta}$ is the weighted graph distance between the corresponding vertices in (V^δ, E^δ). Hence, if it happens that $R(T_i, T_j) < \delta \varepsilon_0$, then the shortest path between a vertex of V_i^0 and a vertex of V_j^0 contains only edges contained in open sets from $(T_k)_{k \in \Sigma_\delta}$. Thus, if $T_k \subseteq \tilde{T}_{\delta, \varepsilon_0}(x)$ for $x \in T$, then there exist $i \in \Sigma_\delta$ and $j \in \mathcal{C}(i)$ such that $T_k \cap T_j \neq \varnothing$ and $T_i \cap T_\delta(x) \neq \varnothing$. The number of intersections of sets of the form $(T_i)_{i \in \Sigma_\delta}$ was estimated in the proof of Lemma 17, and from this we can deduce that $N_{\delta, \varepsilon_0}(x) \leq 16 \mathcal{C}_\delta$ uniformly in x. Consequently, for $r \in [\varepsilon_0 \delta_{n+1}, \varepsilon_0 \delta_n)$, the bound at (15) implies that

$$
\mu^\alpha(B_R(x, r)) \leq c (M^\alpha)^{-1} r^\alpha \mathcal{C}_{\delta_n} \sup_{i \in \Sigma_{\delta_n}} M_i^\alpha.
$$

On recalling the conclusions of Lemmas 11 and 18, we are able to deduce, from this bound, that, for $s < \alpha$, $\limsup_{r \to 0} r^{-s} \mu^\alpha(B_R(x, r)) = 0$ for every $x \in T$, P-a.s. The result is subsequently obtained by applying [10, Proposition 4.9].

References

[1] ALDOUS, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (London Math. Soc. Lecture Notes Ser. **167**), Cambridge University Press, pp. 23–70.

[2] ALON, N. AND SPENCER, J. H. (2000). *The Probabilistic Method*, 2nd edn. John Wiley, New York.

[3] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.

[4] BARLOW, M. T. (1998). Diffusions on fractals. In *Lectures on Probability Theory and Statistics* (Lecture Notes Math. **1690**), Springer, Berlin, pp. 1–121.

[5] BARLOW, M. T. AND BASS, R. F. (1989). The construction of Brownian motion on the Sierpiński carpet. *Ann. Inst. H. Poincaré Prob. Statist.* **25,** 225–257.

[6] CROYDON, D. A. (2006). Random fractal dendrites. D.Phil Thesis, University of Oxford.

[7] CROYDON, D. A. AND HAMBLY, B. M. (2007). Self-similarity and spectral asymptotics for the continuum random tree. To appear in *Stoch. Process. Appl.*

[8] FALCONER, K. J. (1986). Random fractals. *Math. Proc. Camb. Philos. Soc.* **100,** 559–582.

[9] FALCONER, K. J. (1987). Cut-set sums and tree processes. *Proc. Amer. Math. Soc.* **101,** 337–346.

[10] FALCONER, K. J. (1990). *Fractal Geometry*. John Wiley, Chichester.

[11] HAMBLY, B. M. (1997). Brownian motion on a random recursive Sierpiński gasket. *Ann. Prob.* **25,** 1059–1102.

[12] HAMBLY, B. M. AND JONES, O. D. (2003). Thick and thin points for random recursive fractals. *Adv. Appl. Prob.* **35,** 251–277.

[13] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.

[14] KARP, R. M. (1990). The transitive closure of a random digraph. *Random Structures Algorithms* **1,** 73–93.

[15] KIGAMI, J. (1995). Harmonic calculus on limits of networks and its application to dendrites. *J. Funct. Anal.* **128,** 48–86.

[16] KIGAMI, J. (2001). *Analysis on Fractals* (Camb. Tracts Math. **143**). Cambridge University Press.

[17] LIU, Q. (1996). The growth of an entire characteristic function and the tail probabilities of the limit of a tree martingale. In *Trees* (Progress Prob. **40**), Birkhäuser, Basel, pp. 51–80.

[18] LIU, Q. AND ROUAULT, A. (2000). Limit theorems for Mandelbrot's multiplicative cascades. *Ann. Appl. Prob.* **10,** 218–239.

[19] MANDELBROT, B. (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.* **62,** 331–353.

[20] MAULDIN, R. D. AND WILLIAMS, S. C. (1986). Random recursive constructions: asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.* **295,** 325–346.