Math. Proc. Camb. Phil. Soc. (2022), **172**, 655–684
doi:10.1017/S0305004121000360
First published online 7 July 2021

Random fractals and their intersection with winning sets

BY YIFTACH DAYAN

Department of Mathematics, Tel Aviv University, Tel-Aviv, 69978 *Israel.*

Current address: Mathematics Department, Technion, Haifa, 32000 Israel
e-mail: iftahday@gmail.com

(*Received* 29 *August* 2018; *revised* 8 *February* 2021; *accepted* 10 *February* 2021)

Abstract

We show that fractal percolation sets in $\mathbb{R}^d$ almost surely intersect every hyperplane absolutely winning (HAW) set with full Hausdorff dimension. In particular, if $E \subset \mathbb{R}^d$ is a realisation of a fractal percolation process, then almost surely (conditioned on $E \neq \emptyset$), for every countable collection $(f_i)_{i\in\mathbb{N}}$ of C^1 diffeomorphisms of $\mathbb{R}^d$, $\dim_H\left(E \cap \left(\bigcap_{i\in\mathbb{N}} f_i\left(\mathrm{BA}_d\right)\right)\right) = \dim_H(E)$, where BA_d is the set of badly approximable vectors in $\mathbb{R}^d$. We show this by proving that E almost surely contains hyperplane diffuse subsets which are Ahlfors-regular with dimensions arbitrarily close to $\dim_H(E)$.

We achieve this by analysing Galton–Watson trees and showing that they almost surely contain appropriate subtrees whose projections to $\mathbb{R}^d$ yield the aforementioned subsets of E. This method allows us to obtain a more general result by projecting the Galton–Watson trees against any similarity IFS whose attractor is not contained in a single affine hyperplane. Thus our general result relates to a broader class of random fractals than fractal percolation.

2020 Mathematics Subject Classification: 28A80, 60J80, 60D05, 11K60 (Primary); 37C45, 05C80 (Secondary)

1. *Introduction*

1·1. *The set BA_d.*

The field of Diophantine approximations deals with approximations of real numbers and vectors by rationals, where the idea is to keep the denominators as small as possible. A theorem by Dirichlet implies that for every $v \in \mathbb{R}^d$, there exist infinitely many $(P, q) \in \mathbb{Z}^d \times \mathbb{N}$, such that

$$\left\| v - \frac{P}{q} \right\|_\infty < \frac{1}{q^{1+\frac{1}{d}}}.$$

This result leads to one of the key definitions in the field - the badly approximable vectors.

© The Author(s), 2021. Published by Cambridge University Press on behalf of Cambridge Philosophical Society.
https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Definition 1·1. A vector $v \in \mathbb{R}^d$ is called *badly approximable* if there exists some $c > 0$, s.t. for every $(P, q) \in \mathbb{Z}^d \times \mathbb{N}$,

$$\left\| v - \frac{P}{q} \right\| \geq \frac{c}{q^{1+\frac{1}{d}}}.$$

The set of all badly approximable vectors in $\mathbb{R}^d$ is denoted by BA_d.

Throughout this paper, $\|\cdot\|$ is the Euclidean norm, which is the only norm on $\mathbb{R}^d$ to be considered from this point forward. Given $x \in \mathbb{R}^d$ and $r > 0$,

$$B_r(x) = \left\{ y \in \mathbb{R}^d : \|x - y\| < r \right\},$$

and finally, given a set $S \subseteq \mathbb{R}^d$, and $\varepsilon > 0$, $S^{(\varepsilon)}$ is the ε-*neighbourhood* of S defined by $S^{(\varepsilon)} = \bigcup_{x \in S} B_\varepsilon(x)$. Note that using any other norm in Definition 1·1 would result in an equivalent definition.

The set BA_d is one of the most intensively investigated sets in the field of Diophantine approximations. It is well known that BA_d has Lebesgue measure 0. On the other hand, it has Hausdorff dimension d, which makes it reasonable to surmise that it intersects various kinds of fractal sets. Indeed, in recent years there has been a lot of interest, and many results, about the intersection of BA_d with fractal sets. A key result in this line of research is due to Broderick, Fishman, Kleinbock, Reich and Weiss [3], which deals with the intersection of BA_d with a certain kind of fractals called *hyperplane diffuse*.

Definition 1·2. Given $\beta > 0$, a closed set $K \subseteq \mathbb{R}^d$ is called *hyperplane* β - *diffuse* if the following holds:

$\exists \xi_0 > 0,\ \forall \xi \in (0,\ \xi_0),\ \forall x \in K,\ \forall \mathcal{L} \subset \mathbb{R}^d$ affine hyperplane,

$$K \cap B_\xi(x) \setminus \mathcal{L}^{(\beta\xi)} \neq \emptyset.$$

A set is called *hyperplane diffuse* if it is hyperplane β - diffuse for some β.

This turns out to be a quite natural property for fractals and many interesting fractals are known to be hyperplane diffuse, especially when they have some self similarity. see [6, theorems 1·3–1·5] for some examples.

In [3] it was shown that if $K \subset \mathbb{R}^d$ is hyperplane diffuse, then $\dim_H(K \cap \mathrm{BA}_d) > 0$. Moreover, if K is also Ahlfors-regular (defined below) then $\dim_H(K \cap \mathrm{BA}_d) = \dim_H(K)$.

Definition 1·3. For any $\delta > 0$, a measure μ on $\mathbb{R}^d$ is called δ-Ahlfors-regular, if $\exists c_1, c_2 > 0$ s.t. $\forall \rho \in (0, 1)$, $\forall x \in \mathrm{supp}(\mu)$,

$$c_1 \rho^\delta \leq \mu\left(B_\rho(x)\right) \leq c_2 \rho^\delta$$

and Ahlfors-regular if it is δ-Ahlfors-regular for some $\delta > 0$. A set $K \subset \mathbb{R}^d$ is called Ahlfors-regular (resp. δ-Ahlfors-regular) if there exists an Ahlfors-regular (resp. δ-Ahlfors-regular) measure μ on $\mathbb{R}^d$ s.t. $\mathrm{supp}(\mu) = K$.

This result became a main tool for studying intersections of BA_d with fractals. The above statement is in fact a corollary of a more general theorem, where the set BA_d is replaced by an arbitrary hyperplane absolute winning (HAW) set. These are sets which are winning in

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

a certain game called the hyperplane absolute game, which we shall describe in Section 3. Note that the set BA_d is HAW [3]. Thus, the more general theorem is the following.

THEOREM 1·4. ([3]). *Let $K \subset \mathbb{R}^d$ be hyperplane diffuse. Then there exists a constant $C > 0$, s.t. $\forall S \subseteq \mathbb{R}^d$ HAW, $\dim_H(K \cap S) > C$. Moreover, if K is Ahlfors-regular then* $\dim_H(K \cap S) = \dim_H(K)$.

Two important properties of HAW sets are the following:

THEOREM 1·5 [3, proposition 2·3].

(i) *Any countable intersection of HAW sets is HAW.*
(ii) *Any image of a HAW set under a C^1 diffeomorphism of $\mathbb{R}^d$ is HAW.*

Theorem 1·5 implies for example that if $K \subseteq \mathbb{R}^d$ is hyperplane diffuse, then for every sequence $(f_n)_{n\in\mathbb{N}}$ of C^1 diffeomorphisms of $\mathbb{R}^d$, the intersection $K \cap \left(\bigcap_{i\in\mathbb{N}} f_i(BA_d)\right)$ has positive Hausdorff dimension, and if K is also Ahlfors-regular then the Hausdorff dimension of the intersection is maximal, i.e., equal to $\dim_H(K)$.

1·2. *Random fractals*

In this paper we deal with a natural model of random fractals which we will refer to as Galton–Watson fractals. This model may be described as follows. Suppose we are given a finite IFS $\Phi = \left\{\varphi_i : \mathbb{R}^d \to \mathbb{R}^d\right\}_{i\in\Lambda}$ of contracting similarity maps with attractor K (these notions will be explained in more detail in Subsection 2·3. See also [9] for a good exposition of this topic). Φ defines a coding map $\gamma_\Phi : \Lambda^{\mathbb{N}} \to \mathbb{R}^d$ given by $\gamma_\Phi(i) = \bigcap_{n=1}^{\infty} \varphi_{i_1} \circ \ldots \circ \varphi_{i_n}(K)$. Note that $\gamma_\Phi\left(\Lambda^{\mathbb{N}}\right) = K$. Let W be a random variable taking values in the finite set 2^Λ. We construct a Galton–Watson tree by iteratively choosing at random the children of each element of the tree as realisations of independent copies of W, starting from the root, namely $\emptyset$. By concatenating each child to its parent, this defines a random subset of the symbolic space $\Lambda^{\mathbb{N}}$ which we then project using γ_Φ to yield a random fractal $E \subset \mathbb{R}^d$ (which is contained in K). See Figure 1 for an illustrative example.

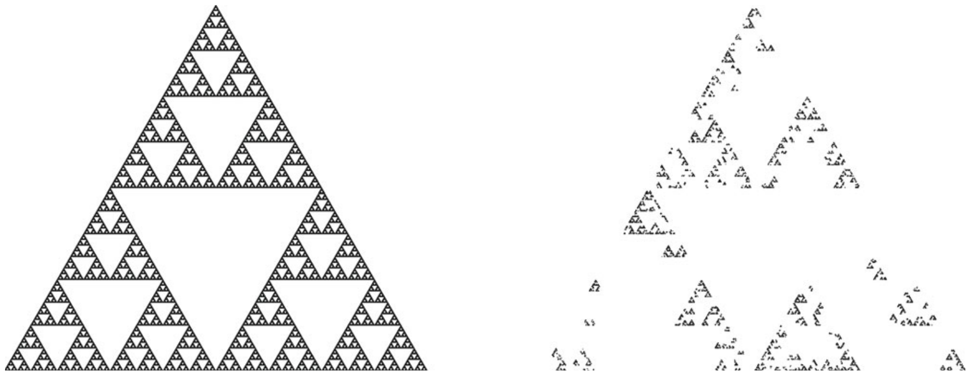
Throughout the paper, we shall always assume that $\forall i \in \Lambda$, $\mathbb{P}(i \in W) > 0$. Note that it is possible that at some level of the tree, no element survives and the process dies out. If this occurs we say that the process is extinct, and the resulting limit set is $E = \emptyset$. It is a well known fact that unless $|W| = 1$ almost surely, $\mathbb{E}(|W|) \leq 1 \iff E = \emptyset$ a.s. (see e.g. [15, proposition 5·1]). The case $\mathbb{E}(|W|) > 1$ is called *supercritical* and we shall assume this property throughout this paper. Another important fact is that if Φ satisfies the open set condition (abbreviated to OSC and will be defined in Subsection 2·3), then a.s. conditioned on nonextinction $\dim_H(E) = \delta$ where δ is the unique number satisfying

$$\mathbb{E}\left(\sum_{i\in W} r_i^\delta\right) = 1,$$

and r_i is the contraction ratio of φ_i for each $i \in \Lambda$.

A specific example of Galton–Watson fractals that the reader should keep in mind is that of fractal percolation (AKA Mandelbrot percolation) which we now describe. Fix some $p \in [0, 1]$ and some integer $b \geq 2$. Let $E_0 \subseteq \mathbb{R}^d$ be the unit cube. Divide E_0 to b^d closed subcubes

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press



The Sierpinski triangle The Galton-Watson fractal

Fig. 1. An approximate realisation of a Galton–Watson fractal which is constructed using an Iterated Function System (IFS) whose attractor is the Sierpinski triangle, where W has the distribution of a Bernoulli process on $\{0, 1\}^{\Lambda}$ with parameter $p = 0.8$.

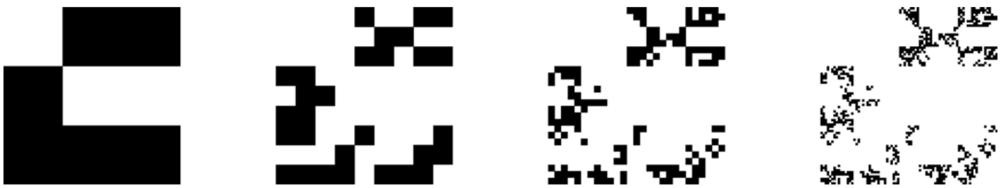


Fig. 2. A realisation of the first 4 steps of a fractal percolation process in $\mathbb{R}^2$ with $b = 3$ and $p = 0.6$.

of equal volume. Now, independently, retain each subcube with probability p or discard it with probability $1 - p$. Let E_1 be the union of all surviving subcubes. Next, for each surviving subcube in E_1 we follow the same procedure. The union of all surviving subcubes in this step will be denoted by E_2. We continue in the same fashion to produce a nested sequence $E_0 \supseteq E_1 \supseteq E_2 \supseteq \cdots$ where each set E_i is the union of the surviving subcubes of level i of the process. Eventually we take $E = \bigcap_{i \in \mathbb{N}} E_i$. See Figure 2 for an example.

All the notions raised in this subsection will be defined in a more formal and detailed manner in Section 2.

1·3. *Main result and applications*

1·3·1. *Main theorem*

A condition which will recur in this paper is that a Galton–Watson fractal is not a.s. contained in an affine hyperplane. Such a Galton–Watson fractal will be referred to as *non-planar*. Since we make the assumption that $\forall i \in \Lambda$, $\mathbb{P}(i \in W) > 0$, non-planarity is essentially a property of the underlying IFS. More precisely, if E is a supercritical Galton–Watson fractal then E is non-planar iff the attractor of the underlying IFS is not contained in an affine hyperplane. This fact as well as some other equivalent conditions to non-planarity are proved in Proposition 3·9. Note that by definition non-planar Galton-Watson fractals are supercritical.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

The main theorem of this paper is the following:

THEOREM 1·6. *Let E be a non-planar Galton–Watson fractal w.r.t. a similarity IFS Φ. Then a.s. conditioned on nonextinction,*

$$\exists C > 0,\ \forall S \subset \mathbb{R}^d \text{ HAW},\ \dim_H (E \cap S) > C.$$

Moreover, if in addition Φ satisfies the OSC, then a.s. conditioned on nonextinction,

$$\forall S \subset \mathbb{R}^d \text{ HAW},\ \dim_H (E \cap S) = \dim_H (E)\,.$$

The reader should pay special attention to the order of the quantifiers in Theorem 1·6 (a.s. $\forall S \subset \mathbb{R}^d$...) which is the stronger form as the collection of HAW sets is uncountable.

Since supercritical fractal percolation sets are non-planar and they correspond to IFSs which satisfy the OSC, an immediate corollary of Theorem 1·6 is the following:

COROLLARY 1·7. *Let $E \subseteq \mathbb{R}^d$ be a limit set of a supercritical fractal percolation process. Then a.s., conditioned on non-extinction, for every HAW set $S \subseteq \mathbb{R}^d$,*

$$\dim_H (E \cap S) = \dim_H (E)\,.$$

The proof of Theorem 1·6 is interesting mainly because in many cases (fractal percolation for example) the Galton–Watson fractal is a.s. not hyperplane diffuse (see [**7**, corollaries A·9, A·10]). Therefore in order to prove Theorem 1·6 we prove the following:

THEOREM 1·8. *Let E be a non-planar Galton–Watson fractal w.r.t. a similarity IFS Φ. Then a.s. conditioned on nonextinction, E contains a hyperplane diffuse subset. Moreover, if Φ satisfies the OSC, then a.s. conditioned on nonextinction, E contains a sequence of subsets $(D_n)_{n\in\mathbb{N}}$, s.t. for each $n \in \mathbb{N}$, $D_n \subseteq E$ is hyperplane diffuse and Ahlfors-regular, and* $\dim_H (D_n) \nearrow \dim_H (E)$.

1·3·2. *Application to BA_d*

Applying Theorem 1·6 to BA_d, together with Theorem 1·5, yields the following immediate corollary.

COROLLARY 1·9. *Let E be a non-planar Galton–Watson fractal w.r.t. a similarity IFS Φ. Then a.s. conditioned on nonextinction, there exists a constant $C > 0$ s.t. for every $(f_i)_{i\in\mathbb{N}}$ a sequence of C^1 diffeomorphisms of $\mathbb{R}^d$,*

$$\dim_H \left(E \cap \left(\bigcap_{i\in\mathbb{N}} f_i\,(BA_d)\right)\right) > C.$$

Moreover, if in addition Φ satisfies the OSC, then a.s. conditioned on nonextinction, for every $(f_i)_{i\in\mathbb{N}}$ a sequence of C^1 diffeomorphisms of $\mathbb{R}^d$,

$$\dim_H \left(E \cap \left(\bigcap_{i\in\mathbb{N}} f_i\,(BA_d)\right)\right) = \dim_H (E)\,.$$

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

1·3·3. *Absolutely non-normal numbers and a generalisation*

Since Theorem 1·6 deals with any HAW set, one may consider other interesting sets which are known to be HAW. One such set is the set of *absolutely non-normal numbers.*

Definition 1·10. Let $2 \le a \in \mathbb{N}$. For $x \in \mathbb{R}$, let $(x_1, x_2, \ldots)$ be the digital expansion of the fractional part of x in base a. Then x is *normal* to base a if $\forall n \in \mathbb{N}$, for every word $\omega \in \{0, 1, \ldots, a-1\}^n$,

$$\lim_{N\to\infty} \frac{1}{N} (\#\text{ occurrences of } \omega \text{ in } x_1, x_2, \ldots, x_N) = a^{-n}.$$

Following [5], x will be called an *absolutely non-normal number* if it is normal to no base $2 \le a \in \mathbb{N}$. By ergodicity of Bernoulli shifts, the set of numbers in the unit interval which are normal to every integer base has Lebesgue measure 1. However, in [3, theorem 2·6] following the ideas of Schmidt [21], it was shown that the set of absolutely non-normal numbers is HAW. In fact, a stronger result was proved - the set of points whose orbit under multiplication by any positive integer (mod 1) is not dense is HAW.

A generalisation of this for higher dimensions is given by the following. Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d - dimensional torus, and let $\pi : \mathbb{R}^d \to \mathbb{T}^d$ be the projection map. For every matrix $R \in GL_d(\mathbb{Q})$ with integer entries, and every point $y \in \mathbb{T}^d$, we shall denote

$$\mathcal{E}(R, y) = \left\{ x \in \mathbb{R}^d : y \notin \overline{\left\{\pi\left(R^k x\right) : k \in \mathbb{N}\right\}} \right\}.$$

PROPOSITION 1·11 [3, theorem 2·6]. *For every nonsingular semisimple matrix with integer entries $R \in GL_d(\mathbb{Q})$, and every point $y \in \mathbb{T}^d$, $\mathcal{E}(R, y)$ is HAW.*

In particular, lifting to $\mathbb{R}^d$ the set of points whose orbit under R is not dense in $\mathbb{T}^d$, yields a HAW set. A further generalisation of this theorem which relates to lacunary sequences of matrices may be found in [2, theorem 1·3].

COROLLARY 1·12. *Let E be a non-planar Galton–Watson fractal w.r.t. a similarity IFS Φ. Then a.s. conditioned on non-extinction, $\exists C > 0$ s.t. for every sequence of nonsingular semisimple matrices with integer entries $R_i \in GL_d(\mathbb{Q})$, every sequence of points $y_i \in \mathbb{T}^d$, and every sequence $(f_i)_{i\in\mathbb{N}}$ of C^1 diffeomorphisms of $\mathbb{R}^d$,*

$$\dim_H \left(E \cap \left(\bigcap_{i\in\mathbb{N}} f_i\left(\mathcal{E}(R_i, y_i)\right) \right) \right) > C.$$

Moreover, if Φ satisfies the OSC, then a.s. conditioned on nonextinction, for every sequences R_i, y_i and f_i as above,

$$\dim_H \left(E \cap \left(\bigcap_{i\in\mathbb{N}} f_i\left(\mathcal{E}(R_i, y_i)\right) \right) \right) = \dim_H(E).$$

Note that in the special case of $d = 1$, under the above conditions, a.s. conditioned on nonextinction, the Hausdorff dimension of the absolutely non-normal numbers in E is bounded from below by some positive constant, and in case Φ satisfies the OSC, this dimension is equal to $\dim_H(E)$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

1·4. *Known results*

In the special case of fractal percolation, a weaker version of Theorem 1·6 may be derived by known results. This goes through the following theorem by Hawkes [**11**] (see also [**18**, theorem 9·5]).

THEOREM 1·13. *Let E be a limit set of a supercritical fractal percolation process with parameters b,p. Let* $A \subset [0, 1]^d$ *be a fixed set s.t.* $\dim_H(A) + \log_b p > 0$. *Then*

$$\textit{esssup} \dim_H(A \cap E) = \dim_H A + \log_b p.$$

Using Hawkes' theorem it is not hard to get the following.

THEOREM 1·14. *Let E be a limit set of a supercritical fractal percolation process and let* $S \subseteq \mathbb{R}^d$ *be a HAW set. Then a.s. conditioned on non-extinction,*

$$\dim_H(E \cap S) = \dim_H(E).$$

The proof of Theorem 1·14 follows immediately from Theorem 1·13 once the following general observation about HAW sets is made (see Remark 3·1): Let $S \subseteq \mathbb{R}^d$ be HAW, and consider the set

$$\tilde{S} = \bigcap_{(r,q)\in\mathbb{Q}\times\mathbb{Q}^d} rS + q.$$

$\tilde{S}$ is also HAW, it is invariant under rational scaling and translations, and is contained in S.

The proof of Theorem 1·14 for S follows from the proof for $\tilde{S}$ which is now left as an exercise for the reader.

Remark 1·15. Corollary 1·7 is stronger than Theorem 1·14 due to different order of the quantifiers, where Corollary 1·7 provides information about intersections of the random fractal in question with every HAW set simultaneously, which may not be obtained directly from Theorem 1·14 since the collection of HAW sets is uncountable.

1·5. *Structure of the paper*

The main goal of this paper is to prove Theorem 1·6. It is proved as a corollary of Theorem 1·8 which will be the focus of this paper. The structure of the paper is as follows: in Section 2 we define trees as subsets of a symbolic space. Then, we turn to the random setup and define Galton–Watson trees. We introduce some background and preliminary results. Then, geometry comes into play and we present the projection of trees to the Euclidean space. We introduce IFSs and the special case of fractal percolation. In Section 3 we define the hyperplane absolute game and describe some related results. We then study the hyperplane diffuse property in the context of iterated function systems and Galton–Watson fractals. In Section 4, after some required preparations, we prove Theorem 1·8. Section 5 provides some additional analysis of one of the tools that are used in the proof of Theorem 1·8, and is included here for completeness.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

2. *Galton–Watson processes*

2·1. *Preliminaries - symbolic spaces and trees*

We shall now fix some notations regarding the symbolic spaces we are about to use. Let $\mathbb{A}$ be some finite set considered as the alphabet. Denote $\mathbb{A}^* = \{\emptyset\} \cup \bigcup_{n\in\mathbb{N}} \mathbb{A}^n$, this is the set of all finite words in the alphabet $\mathbb{A}$, with $\emptyset$ representing the word of length 0. Given a word $i \in \mathbb{A}^*$, we use subscript indexing to denote the letters comprising i, so that $i = i_1 \cdots i_n$ where $i_k \in \mathbb{A}$ for $k = 1, ..., n$. $\mathbb{A}^*$ is considered as a semigroup with the concatenation operation $(i_1 \cdots i_n) \cdot (j_1 \cdots j_m) = (i_1 \cdots i_n, j_1 \cdots j_m)$ and with $\emptyset$ the identity element. The dot notation will usually be omitted so that the concatenation of two words $i, j \in \mathbb{A}^*$ will be denoted simply by ij. We will also consider the action of $\mathbb{A}^*$ on $\mathbb{A}^{\mathbb{N}}$ by concatenations denoted in the same way. We put a partial order on $\mathbb{A}^* \cup \mathbb{A}^{\mathbb{N}}$ by defining $\forall i \in \mathbb{A}^*$, $\forall j \in \mathbb{A}^* \cup \mathbb{A}^{\mathbb{N}}$, $i \leq j$ iff $\exists k \in \mathbb{A}^* \cup \mathbb{A}^{\mathbb{N}}$ with $ik = j$, that is to say $i \leq j$ iff i is a prefix of j. Given any $i \in \mathbb{A}^*$ we shall denote the length of i by $|i| = n$ where n is the unique integer with the property $i \in \mathbb{A}^n$. Given any $i \in \mathbb{A}^*$, the corresponding cylinder set in $\mathbb{A}^{\mathbb{N}}$ is defined as $[i] = \left\{ j \in \mathbb{A}^{\mathbb{N}} : i < j \right\}$.

Definition 2·1. A subset $T \subseteq \mathbb{A}^*$ will be called a *tree* with alphabet $\mathbb{A}$ if $\emptyset \in T$, and for every $a \in T$, $\forall b \in \mathbb{A}^*$, $b \leq a \Rightarrow b \in T$. We shall denote $T_n = T \cap \mathbb{A}^n$ for every $n \geq 1$, and $T_0 = \{\emptyset\}$ so that $T = \bigcup_{n\geq 0} T_n$. We also denote for each $a \in T$, $W_T(a) = \{i \in \mathbb{A} : ai \in T\}$. The *boundary* of a tree T is denoted by ∂T and is given by

$$\partial T = \left\{ a \in \mathbb{A}^{\mathbb{N}} : \forall n \in \mathbb{N},\ a_1 \cdots a_n \in T_n \right\}.$$

The set of all trees with alphabet $\mathbb{A}$ will be denoted by $\mathscr{T}_{\mathbb{A}} \subset 2^{\mathbb{A}^*}$. A *subtree* of $T \in \mathscr{T}_{\mathbb{A}}$ is any tree $T' \in \mathscr{T}_{\mathbb{A}}$ *s.t.* $T' \subseteq T$.

We continue with a few more definitions which will come in handy in what follows. Given any set $A \subseteq \mathbb{A}^*$ and $a \in A$, we denote $A^a = \{j \in \mathbb{A}^* : aj \in A\}$. Given a tree $T \in \mathscr{T}_{\mathbb{A}}$ and $a \in T$ some vertex of T, $T^a \in \mathscr{T}_{\mathbb{A}}$, and is called the *descendants tree of a*. The *length* of a tree $T \subseteq \mathbb{A}^*$ is defined by length $(T) = \sup\{n \in \mathbb{N} \setminus \{0\} : T_{n-1} \neq \emptyset\}$ and takes values in $\mathbb{N} \cup \{\infty\}$. A basic observation in this context is that $\forall T \in \mathscr{T}_{\mathbb{A}}$ with length $(T) = \infty$, $\partial T \neq \emptyset$.

Definition 2·2. A finite set $\Pi \subset \mathbb{A}^*$ is called a *section* if $\bigcup_{i\in\Pi} [i] = \mathbb{A}^{\mathbb{N}}$ and the union is a disjoint union. Given a tree $T \in \mathscr{T}_{\mathbb{A}}$ and a section $\Pi \subset \mathbb{A}^*$ we denote $T_\Pi = T \cap \Pi$.

An obvious example of a section is $\mathbb{A}^n$ for any $n \in \mathbb{N}$. Sections will play an important role in the proof of Theorem 1·8.

2·2. *The random setup - Galton–Watson processes*

Definition 2·3. Let $\mathbb{A}$ be some finite alphabet. Let W be a random variable with values in $2^{\mathbb{A}}$. Let $(W_a)_{a\in\mathbb{A}^*}$ be a (countable) collection of independent copies of W. We now define inductively:

(i) $T_0 = \{\emptyset\}$;

(ii) for $n \geq 1$, $T_n = \bigcup_{a\in T_{n-1}} \{aj : j \in W_a\} \subseteq \mathbb{A}^n$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

If at some point $T_n = \emptyset$, then for every $l > n$, $T_l = \emptyset$ and we shall say that *extinction* occurred. Finally we denote $T = \bigcup_{n\geq 0} T_n$. We shall call the process $T_0, T_1, T_2, \ldots$ (and T as well) a *Galton–Watson process* with alphabet $\mathbb{A}$ and *offspring distribution* W. We consider T as a tree and refer to it as a *Galton–Watson Tree (GWT)*. Note that T is a random variable determined by the random variables $(W_a)_{a\in\mathbb{A}^*}$. As mentioned in the Introduction, we shall make the assumption that $\forall i \in \mathbb{A},\ \mathbb{P}(i \in W) > 0$ (otherwise we may take a smaller alphabet without affecting the law of T).

By definition $\mathscr{T}_{\mathbb{A}} \subset 2^{\mathbb{A}^*}$. As sets, $2^{\mathbb{A}^*} \approx \prod_{n=0}^{\infty} 2^{\mathbb{A}^n}$ with the convention that $\mathbb{A}^0 = \{\emptyset\}$, thus $2^{\mathbb{A}^*}$ may be endowed with the product topology of $\prod_{n=0}^{\infty} 2^{\mathbb{A}^n}$ which is metrisable, separable and compact (where each $2^{\mathbb{A}^n}$ carries the discrete topology). With this topology, $\mathscr{T}_{\mathbb{A}}$ is a closed subset of $2^{\mathbb{A}^*}$, and from this point forward $\mathscr{T}_{\mathbb{A}}$ will carry the topology induced by $2^{\mathbb{A}^*}$.

Given a finite tree $L \subset \mathbb{A}^*$, let $[L] \subset \mathscr{T}_{\mathbb{A}}$ be defined by

$$[L] = \left\{S \in \mathscr{T}_{\mathbb{A}} : \forall n \in \mathbb{N},\ L_n \neq \emptyset \implies S_n = L_n\right\}.$$

These sets form a basis for the topology of $\mathscr{T}_{\mathbb{A}}$ and generate the Borel σ-algebra on $\mathscr{T}_{\mathbb{A}}$ which we denote by $\mathscr{B}$. By Kolmogorov's extension theorem the Galton–Watson process yields a unique Borel measure on $\mathscr{T}_{\mathbb{A}}$ which we denote by $\mathscr{GW}$ and is the distribution of the random variable T. The careful reader will notice that all the events in this paper whose probability is analyzed are in $\mathscr{B}$. For any measurable property $\mathscr{T}' \subseteq \mathscr{T}_{\mathbb{A}}$, the notation $\mathbb{P}\left(T \in \mathscr{T}'\right)$ means $\mathscr{GW}\left(\mathscr{T}'\right)$.

For each $n \geq 0$, we denote $Z_n = |T_n|$. We note that the usual definition of a Galton–Watson process (as defined e.g. in [**15**]) would be the random process $(Z_n)_{n\geq 1}$, but in our case it is important to keep track of the labels as later on we are going to project these trees to the Euclidean space (in the beginning of Subsection 2·3). Nevertheless, in some cases where the labels aren't important we shall refer to the process $(Z_n)_{n\geq 1}$ as a Galton–Watson process as well.

Given a Galton–Watson process, we shall denote $m = \mathbb{E}(Z_1)$. It is a basic fact that for every $n \geq 1$, $\mathbb{E}(Z_n) = m^n$. As mentioned in Section 1, the process is called supercritical when $m > 1$, in which case $\mathbb{P}$ (nonextinction) > 0.

The following is a basic result in the theory of Galton–Watson processes ([**13**], see also [**15**, section 5·1]).

THEOREM 2·4. (Kesten–Stigum). *Let $(Z_k)_{k=1}^{\infty}$ be a supercritical Galton–Watson process, then Z_k/m^k converges a.s. (as $k \to \infty$) to a random variable L, where $\mathbb{E}(L) = 1$ and $L > 0$ a.s. conditioned on nonextinction.*

An immediate corollary of Theorem 2·4 is the following.

COROLLARY 2·5. *Let $(Z_k)_{k=1}^{\infty}$ be a supercritical Galton–Watson process, then the following holds.*

$$\forall \varepsilon > 0,\ \exists c > 0,\ \exists K_0 \in \mathbb{N},\ \forall k > K_0,\ \mathbb{P}\left(\frac{Z_k}{m^k} > c \mid \textit{nonextinction}\right) > 1 - \varepsilon.$$

The proof of Corollary 2·5 is standard and is left as an exercise to reader.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

The next proposition is a basic property of GWTs. Note that if T is a GWT, then for every $v \in T$, the tree T^v is itself a GWT with the same offspring distribution as T, and that for $v, w \in T$ which are not descendants of each other, T^v and T^w are independent.

PROPOSITION 2·6. *Let T be a supercritical Galton–Watson tree with alphabet $\mathbb{A}$ and let $\mathscr{T}' \subseteq \mathscr{T}_{\mathbb{A}}$ be a measurable subset. Suppose that $\mathbb{P}\left(T \in \mathscr{T}'\right) > 0$, then a.s. conditioned on nonextinction, there exist infinitely many $v \in T$ s.t. $T^v \in \mathscr{T}'$.*

Proof. By Corollary 2·5, given some $\varepsilon > 0$, there exists a constant $c > 0$ s.t.

$$\mathbb{P}\left(|T_k| > cm^k \mid \text{nonextinction}\right) > 1 - \varepsilon$$

whenever k is large enough. Denote $\rho = \mathbb{P}\left(T \in \mathscr{T}'\right)$. Given any $M > 0$,

$$\begin{aligned}\mathbb{P}\left(\left|\left\{v \in T_k : T^v \in \mathscr{T}'\right\}\right| < M \mid \text{nonextinction}, \ |T_k| > cm^k\right) \\ \leq \mathbb{P}\left(\text{Bin}\left(cm^k, \ \rho\right) < M\right) \\ \leq \frac{\rho\left(1-\rho\right) cm^k}{\left[\rho cm^k - M\right]^2} \xrightarrow[k\to\infty]{} 0.\end{aligned}$$

In the last inequality we used Chebyshev's inequality assuming that k is large enough so that $\rho cm^k > M$.

2·3. *IFSs and projections to $\mathbb{R}^d$*

An *iterated function system (IFS)* is a finite collection $\{\varphi_i\}_{i\in\Lambda}$ of self maps of $\mathbb{R}^d$ which are Lipschitz continuous with Lipschitz constants smaller than 1. It is one of the most basic results in fractal theory (due to Hutchinson [12]) that every IFS $\{\varphi_i\}_{i\in\Lambda}$ gives rise to a unique nonempty compact set $K \subset \mathbb{R}^d$ which satisfies the equation $K = \bigcup_{i\in\Lambda} \varphi_i K$. The set K is called the *attractor* of the IFS.

A map $f : \mathbb{R}^d \to \mathbb{R}^d$ is called a *contracting similarity* if there exists a constant $r \in (0, 1)$, referred to as the *contraction ratio* of f, s.t. $\forall x, y \in \mathbb{R}^d$, $\|f(x) - f(y)\| = r\|x - y\|$, so that f is a composition of a scaling by factor r, an orthogonal transformation and a translation. In this paper we shall only discuss IFSs which are formed by contracting similarity maps. Such IFSs will be referred to as *similarity IFSs.*

When analysing a similarity IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ it is natural to work in the symbolic spaces $\Lambda^{\mathbb{N}}$ and Λ^*. In view of the setup above, in the abstract setting of trees with alphabet $\mathbb{A}$, we will often assign weights to the alphabet $\{r_i\}_{i\in\mathbb{A}}$. These weights will correspond to the contraction ratios of similarity maps and therefore we shall always assume that $r_i \in (0, 1)$ for every $i \in \mathbb{A}$.

Given an IFS $\{\varphi_i\}_{i\in\Lambda}$, the identification between the symbolic spaces $\Lambda^{\mathbb{N}}$, Λ^* and the Euclidean space is made via the coding map $\gamma_\Phi : \Lambda^{\mathbb{N}} \to \mathbb{R}^d$ which is given by

$$\gamma_\Phi(j) = \bigcap_{n=1}^{\infty} \varphi_{j_1\dots j_n}(K), \tag{2·1}$$

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

where $\varphi_{j_1\ldots j_n} = \varphi_{j_1} \circ \ldots \circ \varphi_{j_n}$. It may be easily seen that $K = \gamma_\Phi\left(\Lambda^{\mathbb{N}}\right)$. Moreover, given a tree $T \in \Lambda^*$, we may project the boundary of T to the Euclidean space using γ_Φ, where

$$\gamma_\Phi(\partial T) = \bigcup_{j\in\partial T}\bigcap_{n=1}^{\infty}\varphi_{j_1\ldots j_n}(K) = \bigcap_{n=1}^{\infty}\bigcup_{i\in T_n}\varphi_i K. \tag{2·2}$$

Note that for every compact set $F \subset \mathbb{R}^d$ s.t. $\forall i \in \Lambda,\ \varphi_i F \subseteq F$, we have

$$K = \bigcap_{n=1}^{\infty}\bigcup_{i\in\Lambda^n}\varphi_i F,$$

hence $K \subseteq F$ and the decreasing sequence of sets $\left(\bigcup_{i\in\Lambda^n}\varphi_i F\right)_{n=1}^{\infty}$ may be thought of as approximating K. Since $K \subseteq F$, we may replace K with F in equations (2·1), (2·2) and the equations will remain true.

An IFS $\{\varphi_i\}_{i\in\Lambda}$ satisfies the *open set condition (OSC)* if there exists some nonempty open set $U \subset \mathbb{R}^d$ s.t. $\varphi_i U \subseteq U$ for every $i \in \Lambda$, and $\varphi_i U \cap \varphi_j U = \emptyset$ for distinct $i,\ j \in \Lambda$. A set U satisfying these conditions will be called an *OSC set for* Φ. In case an IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ with contraction ratios $\{r_i\}_{i\in\Lambda}$ satisfies the open set condition, it is well known[1] that the Hausdorff dimension of the attractor of Φ is the unique number δ which satisfies the equation $\sum_{i\in\Lambda} r_i^\delta = 1$. For convenience, $\forall i = i_1 \cdots i_n \in \Lambda^*$ we denote $r_i = r_{i_1}\cdots r_{i_n}$ which is the contraction ratio of the map φ_i. We also denote $r_{min} = \min\{r_i : i \in \Lambda\}$ and $r_{max} = \max\{r_i : i \in \Lambda\}$.

We shall now turn to the probabilistic setup.

Definition 2·7. Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS, and let W be some random variable with values in 2^Λ. Let T be a GWT with alphabet Λ and offspring distribution W, and finally let E be the random set $E = \gamma_\Phi(\partial T)$. The random set E will be called a *Galton–Watson fractal (GWF)* w.r.t. the IFS Φ and offspring distribution W. We shall always assume that $\mathbb{E}(|W|) > 1$ so that the Galton–Watson process is supercritical.

The following theorem is due to Falconer [**8**] and Mauldin and Williams [**16**]. See also [**15**, theorem 15·10] for another elegant proof.

THEOREM 2·8. *Let E be a Galton–Watson fractal w.r.t. a similarity IFS* $\Phi = \{\varphi_i\}_{i\in\Lambda}$ *satisfying the OSC, with contraction ratios* $\{r_i\}_{i\in\Lambda}$ *and offspring distribution W. Then a.s. conditioned on nonextinction,* $\dim_H E = \delta$ *where* δ *is the unique number satisfying*

$$\mathbb{E}\left(\sum_{i\in W} r_i^\delta\right) = 1.$$

2·4. *Fractal percolation*

Fractal percolation (which was already described in Subsection 1·2) is an important special case of GWFs. Its definition involves the following natural proability distribution on finite sets.

[1]This was first proved by Moran for self-similar sets without overlaps in 1946 (see [**17**]). The form stated here assuming the OSC was first proved by Hutchinson in 1981 (see [**12**]).

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Definition 2·9. Let A be some finite set, and $p \in [0, 1]$. A random subset $Y \subseteq A$ is said to have a binomial distribution with parameter p if

$$\forall B \subseteq A,\ \mathbb{P}\left(Y = B\right) = p^{|B|}\left(1 - p\right)^{|A \setminus B|}.$$

In this case we denote $Y \sim \text{Bin}\left(A,\ p\right)$.[2]

In the framework described above, fractal percolation in $\mathbb{R}^d$ may be defined as follows. Fix an integer $b \geq 2$, and $p \in (0, 1)$. Denote $\Lambda = \{0,\ 1,\ \ldots,\ b - 1\}^d$, and consider the IFS $\Phi = \{\varphi_i\}_{i \in \Lambda}$ where for every $i \in \Lambda$, $\varphi_i : \mathbb{R}^d \to \mathbb{R}^d$ is the similarity map given by $\varphi_i(x) = (1/b)x + i/b$. Finally, take E to be the GWF w.r.t. the IFS Φ and offspring distribution $W \sim \text{Bin}\left(\Lambda,\ p\right)$. E is then considered a fractal percolation set.

Note that the IFS Φ clearly satisfies the OSC. Hence, in the supercritical case, Theorem 2·8 implies that a.s. conditioned on nonextinction,

$$\dim_H E = \log_b\left(m\right),$$

where $m = pb^d = \mathbb{E}\left(|W|\right)$.

3. *Hyperplane diffuse sets*

3·1. *The hyperplane absolute game*

The *hyperplane absolute game,* developed in [**3**], is a useful variant of Schmidt's game which was invented by W. Schmidt in [**21**] and became a main tool for the study of BA_d.

The hyperplane absolute game is played between two players, Bob and Alice and has one fixed parameter $\beta \in (0, 1/3)$. Bob starts by defining a closed ball $B_1 = B_{\rho_1}\left(x_1\right) \subset \mathbb{R}^d$. Then, for every $i \in \mathbb{N}$, after Bob has chosen a ball $B_i = B_{\rho_i}\left(x_i\right)$, Alice chooses an affine hyperplane $\mathcal{L}_i \subset \mathbb{R}^d$ and an $\varepsilon_i \in (0, \beta\rho_i)$, and removes the ε_i-neighbourhood of $\mathcal{L}_i$ denoted by $A_i = \mathcal{L}_i^{(\varepsilon_i)}$ from B_i. Then Bob chooses his next ball $B_{i+1} = B_{\rho_{i+1}}\left(x_{i+1}\right) \subset B_i \setminus A_i$ with the restriction on the radius $\rho_{i+1} \geq \beta\rho_i$. The game continues ad infinitum. A set $S \subset \mathbb{R}^d$ is called *hyperplane absolute winning* (HAW) if for every $\beta \in (0, 1/3)$, Alice has a strategy guaranteeing that $\bigcap_{n=1}^{\infty} B_n$ intersects S. Note that existence of such a strategy for some $\beta \in (0, 1/3)$, implies the existence of a strategy for every $\beta' \in (\beta, 1/3)$.

Many interesting sets are known to be HAW (see e.g. [**1**, **10**, **19**]), including the set BA_d [**3**, theorem 2·5]. Note that HAW sets in $\mathbb{R}^d$ are always dense and have Hausdorff dimension d. Also, as stated in Theorem 1·5, the HAW property is preserved under countable intersections and C^1 diffeomorphisms, which make these sets "large". The following observation may be found useful.

Remark 3·1. Let $S \subseteq \mathbb{R}^d$ be HAW. Then for every countable group G of C^1 diffeomorphisms of $\mathbb{R}^d$, S contains a G - invariant set $\tilde{S}$ which is also HAW. Indeed, we may take $\tilde{S} = \bigcap_{g \in G} gS$, and by Theorem 1·5, $\tilde{S}$ is itself HAW. This property may be useful in some cases, and as an example we already saw a use for this property in Theorem 1·14.

Although HAW sets are "large" in the senses mentioned above, as in the case of BA_d, HAW sets may have Lebesgue measure 0.

[2] Note that the notation $\text{Bin}\left(\cdot,\ \cdot\right)$ will also be used for the usual binomial distribution as well, where the first argument will be an integer and not a set.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

One of the main features of HAW sets is given in Theorem 1·4, which states, generally speaking, that HAW sets intersect hyperplane diffuse sets.

3·2. *Diffuseness in IFSs*

Definition 3·2. Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS, and let $F \subset \mathbb{R}^d$ be a non-empty compact set s.t. $\forall i \in \Lambda$, $\varphi_i F \subseteq F$. For any $c > 0$, we say that Φ is (F, c)-diffuse if $\forall \mathcal{L} \subseteq \mathbb{R}^d$ affine hyperplane $\exists i \in \Lambda$ s.t. $\varphi_i F \cap \mathcal{L}^{(c)} = \emptyset$. Moreover, we say that Φ is F-diffuse if it is (F, c)-diffuse for some $c > 0$.

Obviously, the attractor of Φ is a natural candidate for F in the definition above.

LEMMA 3·3. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS whose attractor is denoted by K. Let $F \subset \mathbb{R}^d$ be as above, and $c > 0$.*

(i) If Φ is (F, c)-diffuse, then Φ is also (K, c)-diffuse.
(ii) If Φ is (K, c)-diffuse, then $\forall c' \in (0, c)$, for a large enough n, the IFS $\{\varphi_i\}_{i\in\Lambda^n}$ is (F, c')-diffuse.

Proof. (i) is trivial since $K \subseteq F$. (ii) follows directly from the fact that the decreasing sequence of sets $\bigcup_{i\in\Lambda^n} \varphi_i F$ converges to K as $n \to \infty$.

In view of the above, we say that Φ is c-diffuse if it is (K, c)-diffuse where K is the attractor of Φ, and we say that Φ is diffuse if it is c-diffuse for some $c > 0$.

Given an IFS $\{\varphi_i\}_{i\in\Lambda}$ in the background, we shall call a finite subset $A \subset \Lambda^*$ diffuse (resp. c-diffuse and (F, c)-diffuse) if the IFS $\{\varphi_i\}_{i\in A}$ is diffuse (resp. c-diffuse and (F, c)-diffuse). Moreover, given a tree $T \in \mathscr{T}_\Lambda$, we say that T is diffuse (resp. c-diffuse and (F, c)-diffuse) if for each $i \in T$, $W_T(i)$ is diffuse (resp. c-diffuse and (F, c)-diffuse). Note that a tree $T \in \mathscr{T}_\Lambda$ is diffuse iff it is c-diffuse for some $c > 0$.

The following lemma will be useful in what follows.

LEMMA 3·4. *$\forall A \subseteq \mathbb{R}^d$, A is contained in an affine hyperplane $\Leftrightarrow$*

$$\inf\left\{\varepsilon > 0 : \exists \mathcal{L} \subset \mathbb{R}^d \textit{ affine hyperplane s.t. } A \subseteq \mathcal{L}^{(\varepsilon)}\right\} = 0.$$

Proof. The implication $(\Rightarrow)$ is trivial. For the other direction, assume that A is not contained in an affine hyperplane. Then there exist $x_1, \ldots, x_{d+1} \in A$ which are not contained in a single affine hyperplane, that is to say that the vectors $v_1 = x_2 - x_1, \ldots, v_d = x_{d+1} - x_1$ are linearly independent, so the matrix $M = \begin{pmatrix} - v_1 - \\ \ldots \\ - v_d - \end{pmatrix}$ is nonsingular. Since $\det(\cdot)$ is a continuous function, small perturbations of M are still nonsingular, so for $\varepsilon > 0$ small enough, no affine hyperplane intersects all the balls $B_\varepsilon(x_i)$ for $i = 1, \ldots, d + 1$, and there is no affine hyperplane $\mathcal{L} \subset \mathbb{R}^d$, s.t. $A \subseteq \mathcal{L}^{(\varepsilon)}$.

PROPOSITION 3·5. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS and let $F \subset \mathbb{R}^d$ be a nonempty compact set s.t. $\varphi_i F \subseteq F$ for every $i \in \Lambda$. Then Φ is (F, c)-diffuse for some $c > 0 \Leftrightarrow$ for every affine hyperplane $\mathcal{L} \subseteq \mathbb{R}^d$, there exists some $i \in \Lambda$ s.t. $\varphi_i F \cap \mathcal{L} = \emptyset$.*

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Proof. The implication ($\Rightarrow$) is true by definition. For the other direction, assume that Φ is not diffuse, i.e., Φ is not (F, c)-diffuse for any c. Take any decreasing sequence (c_n) s.t. $c_n \searrow 0$. Then there exists a sequence of affine hyperplanes $(\mathcal{L}_n)_{n\in\mathbb{N}}$ s.t.

$$\forall n \in \mathbb{N},\ \forall i \in \Lambda,\ \varphi_i F \cap \mathcal{L}_n^{(c_n)} \neq \emptyset.$$

Let $C \subset \mathbb{R}^d$ be a closed ball containing $F^{(c_0)}$ (hence intersecting all the affine hyperplanes $\mathcal{L}_n$), and denote $\mathcal{L}'_n = \mathcal{L}_n \cap C$ for every n. Let Ω_C be the space of non-empty compact subsets of C. Equipped with the Hausdorff metric d_H, this is a compact space. Hence, taking a subsequence we may assume that $\mathcal{L}'_n \to A$ for some $A \in \Omega_C$. By Lemma 3·4, A is contained in some affine hyperplane $\mathcal{L}$. Now, given $\varepsilon > 0$, taking n large enough s.t. $d_H\left(\mathcal{L}'_n, A\right) < \varepsilon/2$ and $c_n < \varepsilon/2$, we obtain $\mathcal{L}^{(\varepsilon)} \cap \varphi_i F \neq \emptyset$ for every $i \in \Lambda$. Since this is true for every $\varepsilon > 0$ and each $\varphi_i F$ is closed, this implies that $\mathcal{L} \cap \varphi_i F \neq \emptyset$ for every $i \in \Lambda$.

The following Proposition relates the concept of diffuseness of IFSs with that of diffuseness of subsets of $\mathbb{R}^d$ as defined in Definition 1·2.

PROPOSITION 3·6. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS with contracting ratios $\{r_i\}_{i\in\Lambda}$ and attractor K. Let $T \in \mathscr{T}_\Lambda$ be a (K, c)-diffuse tree, then $\gamma_\Phi(\partial T)$ is hyperplane $c(r_{min}/diam(K))$-diffuse.*

Proof. Denote $E = \gamma_\Phi(\partial T)$ and $\Delta = \mathrm{diam}(K)$. Assume we are given $\xi \in (0,\ \Delta \cdot r_{min})$, $x \in E$, and an affine hyperplane $\mathcal{L} \subset \mathbb{R}^d$. Let $i = i_1 \cdots i_n \in \Lambda^n$ be a finite word s.t. $x \in \varphi_i K$ and $(\xi/\Delta) r_{min} < r_i \leq \xi/\Delta$ (in order to find such i, let $k = k_1 k_2 \cdots \in \Lambda^{\mathbb{N}}$ be s.t. $\gamma_\Phi(k) = x$, and let $n \in \mathbb{N}$ be the unique integer s.t. $r_{k_1} \cdots r_{k_n} \leq \xi/\Delta < r_{k_1} \cdots r_{k_{n-1}}$, then take $i = k_1 \ldots k_n \in \Lambda^n$). Since $\varphi_i^{-1}(\mathcal{L})$ is still an affine hyperplane, there exists some $j \in W_T(i)$ s.t. $\varphi_j K \cap \left(\varphi_i^{-1}(\mathcal{L})\right)^{(c)} = \emptyset$. Applying φ_i we get that $\varphi_{ij} K \cap \mathcal{L}^{(r_i c)} = \emptyset$. Since $r_i \leq \xi/\Delta$, $\mathrm{diam}(\varphi_i K) = r_i \Delta \leq \xi$ and therefore $\varphi_{ij} K \subseteq \varphi_i K \subseteq B_\xi(x)$. Noting that $\varphi_{ij} K \cap K \neq \emptyset$ we have shown that $K \cap B_\xi(x) \setminus \mathcal{L}^{(r_i c)} \neq \emptyset$, and since $r_i > (\xi/\Delta) r_{min}$ we are done.

Proposition 3·6 implies in particular that whenever a similarity IFS is diffuse, its attractor is hyperplane diffuse.

The following theorem was proved in [**14**].

THEOREM 3·7 [**14**, theorem 2·3]. *Let Φ be a similarity IFS satisfying the OSC whose attractor K is not contained in an affine hyperplane, then K is hyperplane diffuse and Ahlfors-regular.*

Note that instead of the condition that K is not contained in a single affine hyperplane, the original condition in [**14**, theorem 2·3] is that no finite collection of affine hyperplanes is preserved by Φ (such an IFS is referred to as *irreducible*), but it turns out that these two conditions are in fact equivalent (regardless of the OSC). This fact is proved in [**4**, proposition 3·1].

In the following theorem we show another equivalent condition which will be useful in what follows. Note that in the following theorem we don't assume the OSC holds as opposed to Theorem 3·7.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

THEOREM 3·8. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS in $\mathbb{R}^d$ with attractor K. The following are equivalent:*

(i) *K is not contained in an affine hyperplane;*
(ii) *there exists a diffuse section $\Pi \subset \Lambda^*$ (i.e., such that the IFS $\{\varphi_i\}_{i\in\Pi}$ is diffuse);*
(iii) *K is hyperplane diffuse.*

Proof. (i)⇒(ii). Assume that (ii) does not hold, so that every section Π is not diffuse. So given some $\varepsilon > 0$, let Π be a section s.t. $\forall i \in \Pi$, $\operatorname{diam}(\varphi_i K) < \varepsilon/2$. Since Π is not $\varepsilon/2$-diffuse, there is some affine hyperplane $\mathcal{L}$, s.t. $\mathcal{L}^{(\varepsilon/2)} \cap \varphi_i K \neq \emptyset$ for every $i \in \Pi$. Since $\operatorname{diam}(\varphi_i K) < \varepsilon/2$ this implies that $\forall i \in \Pi$, $\varphi_i K \subset \mathcal{L}^{(\varepsilon)}$, hence $K \subset \mathcal{L}^{(\varepsilon)}$. Taking ε to 0 implies that K is a.s. contained in an affine hyperplane by Lemma 3·4.
(ii)⇒(iii). This follows immediately from Proposition 3·6.
(iii)⇒(i). Follows from the definition of the hyperplane diffuse property.

One of the conditions of Theorem 1·6 is that the GWF is non-planar. We now list a few equivalent conditions to non-planarity of supercritical GWFs.

PROPOSITION 3·9. *Let E be a supercritical GWF w.r.t. a similarity IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ and offspring distribution W, and let T be the corresponding GWT. Denote by K the attractor of Φ. The following conditions are equivalent.*

(i) *E is non-planar;*
(ii) $\mathbb{P}$ (*E is not contained in an affine hyperplane*| *nonextinction*) $= 1$;
(iii) *K is not contained in an affine hyperplane;*
(iv) $\exists \Pi \subseteq \Lambda^*$ *section, and* $c > 0$ *s.t.* $\mathbb{P}$ (T_Π *is* (K, c)-*diffuse*) > 0.

Proof. First note that since there are only countably many sections (for every n there is only a finite number of sections of size n), (iv) is equivalent to the following statement:

$$\mathbb{P}\left(\exists \Pi \subseteq \Lambda^* \text{ section, and } c{>}0 \text{ s.t. } T_\Pi \text{ is } (K, c)\text{-diffuse}\right) > 0.$$

(i)⇒(ii). Follows from Proposition 2·6, namely, since

$$\mathbb{P}\left(E \text{ is not contained in an affine hyperplane}\right) > 0,$$

almost surely given nonextinction $\exists v \in T$ s.t. $\gamma_\Phi(\partial T^v)$ is not contained in an affine hyperplane, which implies that E is not contained in an affine hyperplane.
(ii)⇒(iii). Trivial
(iii)⇒(iv). We first prove the following claim by induction.

Claim For every integer $0 \le k \le d-1$ there exists $\left\{i^1, \ldots, i^{k+2}\right\} \subseteq \Lambda^*$ s.t. the following hold:

(a) $\forall i, j \in \left\{i^1, \ldots, i^{k+2}\right\}$, $\varphi_i K \cap \varphi_j K = \emptyset$ (hence $i \not\geq j$ and $j \not\geq i$);
(b) $\mathbb{P}\left(\left\{i^1, \ldots, i^{k+2}\right\} \subseteq T\right) > 0$;
(c) for every integer $0 \le n \le k$, for every n-dimensional affine subspace $\mathcal{L} \subseteq \mathbb{R}^d$ that intersects $\varphi_{i^1} K, \ldots, \varphi_{i^{n+1}} K$, we have $\varphi_{i^{n+2}} K \cap \mathcal{L} = \emptyset$.

Proof of claim. For $k = 0$: since the process is supercritical, there exist $i, j \in \Lambda$ s.t. $\mathbb{P}(\{i, j\} \subseteq W) > 0$. Therefore, for some $i' > i$ and $j' > j$, $\varphi_{i'} K \cap \varphi_{j'} K = \emptyset$ and all Three conditions are fulfilled by $\{i', j'\}$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Assume $\{i^1, \ldots, i^{k+1}\} \subseteq \Lambda^*$ satisfies (a), (b), (c) for $k-1$. By the case $k=0$, there are a, $b > i^{k+1}$ s.t. $\varphi_a K \cap \varphi_b K = \emptyset$, and $\mathbb{P}(\{a, b\} \subseteq T) > 0$. Pick any $x_1 \in \varphi_{i^1} K, \ldots, x_k \in \varphi_{i^k} K$, $x_{k+1} \in \varphi_a K$. By assumption $x_1, \ldots, x_{k+1}$ are affinely independent, and thus span a unique k-dimensional affine subspace $\mathcal{L} \subset \mathbb{R}^d$. By Theorem 3·8, Φ has some diffuse section, so there is some $j \in \Lambda^*$ s.t. $\varphi_{bj} K \cap \mathcal{L} = \emptyset$. Therefore, $\varphi_{bj} K$ does not intersect small enough perturbations of $\mathcal{L}$ as well. So there are $j^1 > i^1, \ldots, j^k > i^k$, $j^{k+1} > a$ s.t. $\varphi_{bj} K \cap \mathcal{L}' = \emptyset$ for every k-dimensional affine subspace $\mathcal{L}'$ which intersects the sets $\varphi_{j^1} K, \ldots, \varphi_{j^{k+1}} K$. Denoting $j^{k+2} = bj$, the set $\{j^1, \ldots, j^{k+2}\}$ satisfies conditions (a), (b), (c) for k.

Let $i^1, \ldots, i^{d+1} \in \Lambda^*$ be the elements whose existence is guaranteed by the claim for $k = d-1$. By property (a), there is some section $\Pi \subset \Lambda^*$ s.t. $i^1, \ldots, i^{d+1} \in \Pi$. By property (b), $\mathbb{P}\left(\{i^1, \ldots, i^{d+1}\} \subseteq T_\Pi\right) > 0$. And by property (c) combined with Proposition 3·5, $\{i^1, \ldots, i^{d+1}\}$ is (K, c)-diffuse for some $c > 0$. Since $\mathbb{P}\left(\{i^1, \ldots, i^{d+1}\} \subseteq T_\Pi\right) > 0$, then $\mathbb{P}\left(\{i^1, \ldots, i^{d+1}\}^n \subseteq T_{\Pi^n}\right) > 0$. To summarise, we have found a section $\Pi^n \subset \Lambda^*$ and $c > 0$ s.t. $\mathbb{P}(T_{\Pi^n} \text{ is } (K, c)\text{-diffuse}) > 0$.

(iv)⇒(i). By Lemma 3·3 and Theorem 3·8, Assuming (iv) implies that there exists some section Π s.t.

$$\mathbb{P}\left(\text{the attractor of } \{\varphi_i\}_{i \in T_\Pi} \text{ is not contained in an affine hyperplane}\right) > 0.$$

Consider the GWF with IFS $\{\varphi_i\}_{i \in T_\Pi}$ and offspring distribution $\sim T_\Pi$. Obviously it has the same law as E. So without loss of generality we may assume that

$$\mathbb{P}\left(\text{the attractor of } \{\varphi_i\}_{i \in W} \text{ is not contained in an affine hyperplane}\right) > 0$$

and take Λ instead of Π for convenience of notations. Let $A \subseteq \Lambda$ be s.t. the attractor of $\{\varphi_i\}_{i \in A}$, which we denote by K_A, is not contained in an affine hyperplane and $\mathbb{P}(W = A) > 0$. Since K_A is not contained in an affine hyperplane, by Lemma 3·4 there exists some $\varepsilon > 0$ s.t. $\forall \mathcal{L} \subset \mathbb{R}^d$ affine hyperplane, $K_A \nsubseteq \mathcal{L}^{(\varepsilon)}$. Therefore, taking $n \in \mathbb{N}$ large enough, for every affine hyperplane $\mathcal{L}$ there exists some $i \in A^n$ s.t. $\varphi_i K \cap \mathcal{L}^{(\varepsilon/2)} = \emptyset$. Now, since $\mathbb{P}(W = A) > 0$, there exists a positive probability that $A^n \subset T_n$ and for every $i \in A^n$, T^i is infinite. Obviously, in this case E is not contained in an affine hyperplane.

A discussion about hyperplane diffuseness of GWFs may be found in the Appendix of [7], where it is shown that in many cases, GWFs are a.s. not hyperplane diffuse. In particular, fractal percolation sets are almost surely not hyperplane diffuse.

4. *Proof of main result*

4·1. *Main ideas of the proof*

In order to help understanding the main ideas of the proof of Theorem 1·8, we now sketch them informally for the special case of fractal percolation. Let E be a (supercritical) fractal percolation set with parameters $b \geq 2$ and $p \in (0, 1)$, as defined in subsection 2·4, and let T be the underlying Galton–Watson tree. In order to find the sets $D_n \subseteq E$, we find appropriate subtrees of T and project them using the coding map.

Choose some $c \in (1, m)$, where $m = pb^d$. For some very large $k \in \mathbb{N}$ consider the tree $T_{(k)}$ - the tree with alphabet Λ^k given by $T_{(k)} = \bigcup_{n \geq 0} T_{k \cdot n}$. This is a Galton–Watson tree, but its offspring distribution is no longer binomial. By Theorem 2·4, every vertex of $T_{(k)}$ is expected

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

to have approximately m^k children. Since k is very large, c^k is much smaller than m^k and so the probability that a vertex of $T_{(k)}$ has less than c^k children approaches 0 as $k \to \infty$.

Moreover, the probability that a vertex of $T_{(k)}$ has all its children aligned on some affine hyperplane also tends to 0 as $k \to \infty$. Indeed, going back 2 generations[3] into the past, T_{k-2} has approximately m^{k-2} surviving vertices. Since m^{k-2} is very large there is a large probability that at least one of these vertices has all its level 2 descendants survive, which implies that no affine hyperplane intersects all the cubes corresponding to the surviving children.

From the above, one may deduce (not trivially) that when k is large enough, there is a positive probability that $T_{(k)}$ contains a subtree S with the following 2 properties:

(i) each element of S has exactly c^k children;
(ii) S is diffuse.

These 2 properties imply that S projects to an Ahlfors-regular set of Hausdorff dimension $\log_b(c)$, which is hyperplane diffuse. Since there is a positive probability that $T_{(k)}$ contains such a subtree, almost surely there is some vertex $v \in T_{(k)}$ such that $\left(T_{(k)}\right)^v$ contains a subtree with these properties.

Finally, we take $c \nearrow m$ through some sequence to ensure that the Hausdorff dimension of the resulting sets approaches $\log_b(m)$, which is almost surely the Hausdorff dimension of E.

The main difficulty in the general setting in comparison to fractal percolation arises from allowing different maps in the IFS to have different contraction ratios, in which case the trees $T_{(k)}$ have very different weights assigned to the vertices of each level, and so a-ary subtrees of these trees would project to sets that are not regular enough. In order to deal with this problem we need to define an analogue notion of the trees $T_{(k)}$, but along sections where all vertices have approximately the same weights.

4·2. *Sections*

We first note that for an IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$, given a section $\Pi \subset \Lambda^*$, we may think of the IFS $\{\varphi_i\}_{i\in\Pi}$, which obviously has the same attractor as Φ.

LEMMA 4·1. *Let T be a GWT with alphabet $\mathbb{A}$ and offspring distribution W, and with weights $\{r_i\}_{i\in\mathbb{A}}$. Assume that $\mathbb{E}\left(\sum_{i\in W} r_i^\delta\right) = 1$. Then for every section $\Pi \subset \mathbb{A}^*$,*

$$\mathbb{E}\left(\sum_{i\in T_\Pi} r_i^\delta\right) = 1.$$

The proof of the lemma may be carried out by induction on the size of Π and is left as an exercise for the reader.

Definition 4·2. Given an alphabet $\mathbb{A}$ with weights $\{r_i\}_{i\in\mathbb{A}}$ and a positive number $\rho \in (0, r_{min})$, we denote by Π_ρ the section given by

$$\Pi_\rho = \left\{i \in \mathbb{A}^* : r_i \le \rho < r_{i_1} \cdots r_{i_{|i|-1}}\right\}.$$

[3]For $b > 2$ it is enough to go back 1 generation.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Note that $\forall i \in \Pi_\rho,\ \rho \cdot r_{min} < r_i \le \rho$ (recall that $r_{min} = \min\{r_i : i \in \mathbb{A}\}$)

LEMMA 4·3. *Let T be a supercritical GWT with alphabet $\mathbb{A}$, weights $\{r_i\}_{i\in\mathbb{A}}$, and some offspring distribution W, and let δ satisfy $\mathbb{E}\left(\sum_{i\in W} r_i^\delta\right) = 1$. Then $\forall\alpha < \delta$, a.s. conditioned on nonextinction there exists some $\rho_0 > 0$ s.t. $\forall\rho < \rho_0$, $\left|T_{\Pi_\rho}\right| > 1/\rho^\alpha$.*

In order to prove the lemma we need the following theorem by Falconer [8].

THEOREM 4·4. *Let T be a GWT with alphabet $\mathbb{A}$, weights $\{r_i\}_{i\in\mathbb{A}}$, and offspring distribution W. Given $\nu > 0$, the following statements hold*:

(i) $\mathbb{E}\left(\sum_{i\in W} r_i^\nu\right) \le 1 \implies$ *either* $\sum_{i\in W} r_i^\nu = 1$ *a.s. or* $\inf_{\text{section } \Pi\subset\mathbb{A}^*} \sum_{i\in T_\Pi} r_i^\nu = 0$ *a.s;*

(ii) $\mathbb{E}\left(\sum_{i\in W} r_i^\nu\right) > 1 \implies \inf_{\text{section } \Pi\subset\mathbb{A}^*} \sum_{i\in T_\Pi} r_i^\nu > 0$ *a.s. conditioned on nonextinction.*

We note that Falconer's Theorem is in fact more general than stated above and may be applied in cases were the weights themselves are random variables (see [15, theorem 5·35]). Also note that Falconer's theorem is the main ingredient in the proof given in [15] of Theorem 2·8.

Proof of Lemma 4·3. Given $\alpha < \delta$, fix some $\alpha' \in (\alpha, \delta)$. If the lemma is false, then with positive probability there exists a decreasing sequence $\rho_n \searrow 0$ s.t. $\left|T_{\Pi_{\rho_n}}\right| \le 1/\rho_n^\alpha$ for every n. In this case, for each n, we have

$$\sum_{i\in T_{\Pi_{\rho_n}}} r_i^{\alpha'} \le \frac{1}{\rho_n^\alpha} \cdot \rho_n^{\alpha'} = \rho_n^{\alpha'-\alpha} \xrightarrow[n\to\infty]{} 0$$

which contradicts (2) of Theorem 4·4 since $\alpha' < \delta$ implies that $\mathbb{E}\left(\sum_{i\in W} r_i^{\alpha'}\right) > 1$.

COROLLARY 4·5. *Let T and δ be as in the previous lemma. Then for every $\alpha < \delta$, $\forall\varepsilon > 0$, $\exists\rho_0 > 0$ s.t. $\forall\rho < \rho_0$*

$$\mathbb{P}\left(\left|T_{\Pi_\rho}\right| > \frac{1}{\rho^\alpha} \mid \textit{nonextinction}\right) > 1 - \varepsilon.$$

Proof. Denote by A_n the event: $\forall\rho < n^{-1}$, $\left|T_{\Pi_\rho}\right| > \rho^{-\alpha}$. By Lemma 4·3

$$\mathbb{P}\left(\bigcup_{n=1}^\infty A_n \mid \text{nonextinction}\right) = 1.$$

Since $(A_n)_{n=1}^\infty$ is an increasing sequence of events, $\mathbb{P}(A_n \mid \text{nonextinction}) \nearrow 1$. Taking $\rho_0 = n_0^{-1}$ with n_0 large enough we finish the proof.

Remark 4·6. Corollary 4·5 may be seen as a version of Corollary 2·5 for sections of the type Π_ρ, where Corollary 2·5 deals with sections of the type $\mathbb{A}^k$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

We conclude this subsection with the following lemma which is a standard application of the open set condition.

LEMMA 4·7. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS with contraction ratios $\{r_i\}_{i\in\Lambda}$, which satisfies the open set condition. Let K be the attractor of Φ. Then there exists some constant $C > 0$, s.t. $\forall x \in \mathbb{R}^d$, for any $\rho \in (0, r_{min})$, $\left|\left\{a \in \Pi_\rho : \varphi_a K \cap B_\rho(x) \neq \emptyset\right\}\right| < C$.*

Proof. Let $U \subset \mathbb{R}^d$ be an OSC set for Φ. Denote $\delta = \mathrm{diam}(U)$ and $\Delta = \mathrm{diam}(K)$. Note that since $\varphi_i \overline{U} \subseteq \overline{U}$ for every $i \in \Lambda$, $K \subseteq \overline{U}$. Hence, since the ball $B_\rho(x)$ is open, it is enough to show that $\left|\left\{a \in \Pi_\rho : \varphi_a U \cap B_\rho(x) \neq \emptyset\right\}\right| < C$ for some constant C.

Fix some open ball $B \subseteq U$ of radius $\varepsilon > 0$. For any $a \in \Pi_\rho$, $\varphi_a B$ is a ball of radius $r_a\varepsilon$. Since $r_a\varepsilon > \rho r_{min}\varepsilon$, we have

$$\mathrm{Vol}(\varphi_a U) \geq \mathrm{Vol}(\varphi_a B) > C_1\rho^d,$$

where $C_1 > 0$ is some constant which depends only on $\varepsilon,\ r_{min},\ d$.

On the other hand, for any $a \in \Pi_\rho$, $\mathrm{diam}(\varphi_a U) = r_a\delta \leq \rho\delta$. Hence, if $\varphi_a U \cap B_\rho(x) \neq \emptyset$, then $\varphi_a U \subseteq B_{\rho(1+\delta)}(x)$. Note that $\mathrm{Vol}\left(B_{\rho(1+\delta)}(x)\right) = C_2\rho^d$ where $C_2 > 0$ is some constant which depends only on δ and d.

Since all the sets $\{\varphi_a U\}_{a\in\Pi_\rho}$ are disjoint, we have

$$\left|\left\{a \in \Pi_\rho : \varphi_a U \cap B_\rho(x) \neq \emptyset\right\}\right| < \frac{C_2}{C_1}.$$

4·3. $*$-*trees*

For the proof of Theorem 1·8, we need to make the definition of trees a bit more flexible, allowing the building blocks of the tree to be strings in the alphabet $\mathbb{A}$ instead of just letters.

First, we introduce the following notation: given a subset $S \subseteq \mathbb{A}^*$, we define the function $h_S : S \to \mathbb{N} \cup \{0\}$ by

$$\forall i \in S,\ h_S(i) = |\{j \in S : j < i\}|.$$

The value $h_S(i)$ is referred to as the *height of* i. The subscript after h may be omitted whenever the context is believed to be clear.

Definition 4·8. Let $\mathbb{A}$ be a finite alphabet. A subset $S \subset \mathbb{A}^*$ will be called a $*$-*tree* with alphabet $\mathbb{A}$ if the following conditions hold:

(i) $\emptyset \in S$;
(ii) $\forall \emptyset \neq i \in S$, there exists a unique $j \in S$, s.t. $j < i$ and $h(j) = h(i) - 1$;
(ii) $\forall n \in \mathbb{N},\ \left|h^{-1}(n)\right| < \infty$.

We denote $S_n = h^{-1}(n)$. For each $i \in S$ we denote $W_S(i) = \left\{j \in \mathbb{A}^* : ij \in S_{h(i)+1}\right\}$. The boundary of S is defined by $\partial S = \left\{i \in \mathbb{A}^{\mathbb{N}} : \forall n \in \mathbb{N},\ \exists j \in S_n,\ j < i\right\}$. Given a $*$-tree S and some vertex $i \in S$, $S^i = \{j \in \mathbb{A}^* : ij \in S\}$ is a $*$-tree and will be referred to as the descendants tree of i.

Obviously, every tree $T \in \mathscr{T}_{\mathbb{A}}$ is a $*$-tree with alphabet $\mathbb{A}$, and $h_T(i) = |i|$ for every $i \in T$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Definition 4·9. Let S be a $*$-tree with alphabet $\mathbb{A}$. A *$*$-subtree* of S is any $*$-tree Q with alphabet $\mathbb{A}$ s.t. $Q \subseteq S$ and for every $i \in Q$, $\{j \in S : j < i\} \subset Q$ (this condition ensures that $h_Q = h_S \restriction_Q$).

We now define the compression of trees along sections.

Definition 4·10. Let T be a tree with alphabet $\mathbb{A}$. Let $(\Pi_n)_{n=1}^{\infty}$ be a sequence of sections s.t. for every n, $\forall i \in \Pi_{n+1}$, $\exists j \in \Pi_n$, s.t. $j < i$. Then *the compression of T along the sections* $(\Pi_n)_{n=1}^{\infty}$ is defined to be the $*$-tree S with alphabet $\mathbb{A}$ given by $S = \bigcup_{n=0}^{\infty} T_{\Pi_n}$, where we define $\Pi_0 = \{\emptyset\}$.

Note that $S_n = T_{\Pi_n}$ for every n, and that $\partial S = \partial T$. Now, suppose that T is a tree with alphabet $\mathbb{A}$ and weights $\{r_i\}_{i\in\mathbb{A}}$, and let $\rho \in (0, r_{min})$ be some positive number. The compression of T along the sections $\left(\Pi_{\rho^n}\right)_{n=1}^{\infty}$ will be denoted by $T_{(\rho)}$.

Definition 4·11. Given an alphabet $\mathbb{A}$ with weights $\{r_i\}_{i\in\mathbb{A}}$, $\rho \in (0, r_{min})$ and $i \in \bigcup_{n=1}^{\infty} \Pi_{\rho^n}$, we denote $a_\rho(i) := r_i/\rho^{n_\rho(i)}$, where $n_\rho(i)$ is the unique $n \in \mathbb{N}$ s.t. $i \in \Pi_{\rho^n}$. We also denote $n_\rho(\emptyset) = 0$ and $a_\rho(\emptyset) = 1$.

PROPOSITION 4·12. *Let $\mathbb{A}$ be an alphabet with weights $\{r_i\}_{i\in\mathbb{A}}$, and let $\rho \in (0, r_{min})$. Then for every $i \in \bigcup_{n=1}^{\infty} \Pi_{\rho^n} \cup \{\emptyset\}$, $m \in \mathbb{N}$, and $j \in \mathbb{A}^*$,*

$$ij \in \Pi_{\rho^{n_\rho(i)+m}} \iff j \in \Pi_{\frac{\rho^m}{a_\rho(i)}}.$$

Proof. For $i = \emptyset$ the claim is trivial. Given $i \in \Pi_{\rho^n}$ for some $n \geq 1$, for every $j \in \mathbb{A}^*$,

$$\begin{aligned} ij \in \Pi_{\rho^{n+m}} &\iff r_i r_j \leq \rho^{n+m} < r_i r_{j_1} \cdots r_{j_{|j|-1}} \\ &\iff r_j \leq \frac{\rho^m}{a_\rho(i)} < r_{j_1} \cdots r_{j_{|j|-1}} \\ &\iff j \in \Pi_{\frac{\rho}{a_\rho(i)}} \end{aligned}$$

The following Proposition is an immediate consequence of Proposition 4·12.

PROPOSITION 4·13. *Let T be a GWT with alphabet $\mathbb{A}$ and weights $\{r_i\}_{i\in\mathbb{A}}$, and fix any $\rho \in (0, r_{min})$. Then $\forall i \in \bigcup_{n=1}^{\infty} \Pi_{\rho^n} \cup \{\emptyset\}$, conditioned on $i \in T_{(\rho)}$, $W_{T_{(\rho)}}(i)$ has the same law as $T \cap \Pi_{\frac{\rho}{a_\rho(i)}}$.*

Next, we define random $*$-trees. Let $\mathbb{A}$ be a finite alphabet and let $B \subseteq \mathbb{A}^*$ be a $*$-tree. Let $\{M_x\}_{x\in B}$ be a collection of independent random variables s.t. for every $x \in B$, M_x takes values in the finite set $2^{W_B(x)}$. Define

(i) $S_0 = \emptyset$;

(ii) for $n \geq 1$, $S_n = \bigcup_{i\in S_{n-1}} \{ij : j \in M_i\}$,

and finally take $S = \bigcup_{n=0}^{\infty} S_n$. S is then a random $*$-tree on B with offspring distributions $\{M_x\}_{x\in B}$. Note that every realisation of S is a $*$-subtree of B. There may be elements $i \in B$

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

s.t. $\mathbb{P}(i \in S) = 0$. These elements add no information to the construction. Therefore, given the setup above we denote $B' = \{i \in B : \mathbb{P}(i \in S) > 0\}$. Note that by the construction of S, B' is a $*$-subtree of B. Throughout this paper, we assume that the offspring distributions $\{M_x\}_{x \in B}$ are *bounded*, i.e., there exists some constant $C > 0$ s.t. $\forall x \in B'$, $\mathbb{P}(|M_x| < C) = 1$.

In the graph theoretic perspective, this process may be thought of in the following way: Let G be the graph with vertices B and edges $\{(i, j) : j \in W_B(i)\}$. The graph G is a directed rooted tree in the graph theoretic sense with $\emptyset$ serving as the root. Now, given a realisation of the random variables $\{M_x\}_{x \in B}$, we take S to be the connected component of the subgraph of G, with vertices B and edges $\{(i, j) : j \in M_i\}$, which contains $\emptyset$.

Let T be a GWT with alphabet $\mathbb{A}$, and let $(\Pi_n)_{n=1}^{\infty}$ be a sequence of sections as in Definition 4·10. The compression of T along the sections $(\Pi_n)_{n=1}^{\infty}$ has the law of a random $*$-tree on $B = \bigcup_{n=0}^{\infty} \Pi_n$, where for each $i \in B$, M_i has the law of $T \cap W_B(i)$. In particular, the following is an immediate consequence of Proposition 4·13.

PROPOSITION 4·14. *Let T be a GWT with alphabet $\mathbb{A}$, weights $\{r_i\}_{i \in \mathbb{A}}$ and offspring distribution W. Then $\forall \rho \in (0, r_{min})$, the compressed tree $T_{(\rho)}$ has the law of a random $*$-tree on $B = \bigcup_{n=1}^{\infty} \Pi_{\rho^n} \cup \{\emptyset\}$, with offspring distributions $M_i \sim T_{\Pi_{\frac{\rho}{a_\rho(i)}}}$, $\forall i \in B$.*

Note that since we assume that $\forall i \in \mathbb{A}$, $\mathbb{P}(i \in W) > 0$, we have $B' = B$. It is important to notice that for every $i \in T_{(\rho)}$, $r_{min} < a_\rho(i) \leq 1$. This means that although $T_{(\rho)}$ need not have the structure of a GWT where all the offspring distributions are the same, the offspring distributions of $T_{(\rho)}$ can not vary too much.

4·4. *Existence of subtrees in random *-trees*

The main theorem of this subsection (Theorem 4·17) plays a major role in the proof of Theorem 1·8. It deals with the existence of certain subtrees in random $*$-trees and is inspired by a theorem due to Pakes and Dekking from [**20**] (see also [**15**, theorem 5·29]). In order to state the theorem, we need the following definitions and notations.

Definition 4·15. Let $\mathscr{A}$ be a mapping $\mathscr{A} : D \to 2^{(2^{\mathbb{A}^*})}$ where $D \subseteq \mathbb{A}^*$, with the notation $\mathscr{A}(i) = \mathscr{A}_i$ s.t. $\forall i \in D$, $\emptyset \neq \mathscr{A}_i \subseteq 2^{\mathbb{A}^*}$. A $*$-tree S is called an *$\mathscr{A}$-$*$-tree* if $S \subseteq D$ and $\forall i \in S$, $W_S(i) \in \mathscr{A}_i$. S will be called an *$\mathscr{A}$-$*$-tree of level n* if for every element $i \in S$ of height $< n$, $i \in D$ and $W_S(i) \in \mathscr{A}_i$.

Given a mapping $\mathscr{A}$ as above, and any $x \in \mathbb{A}^*$, we denote by $\mathscr{A}^x$ the mapping $\mathscr{A}^x : D^x \to 2^{(2^{\mathbb{A}^*})}$ given by $\mathscr{A}^x(i) = \mathscr{A}_{xi}$.

Another notation we are going to use is the following: given any set A, and a collection $\mathscr{I} \subseteq 2^A$, we denote $\overline{\mathscr{I}} = \left\{S \in 2^A : \exists X \in \mathscr{I},\ X \subseteq S\right\}$. The collection $\mathscr{I}$ will be called *monotonic* if $\overline{\mathscr{I}} = \mathscr{I}$. Given a mapping $\mathscr{A} : D \to 2^{(2^{\mathbb{A}^*})}$ as in Definition 4·15, we denote by $\overline{\mathscr{A}}$ the mapping $\overline{\mathscr{A}} : D \to 2^{(2^{\mathbb{A}^*})}$ given by $\overline{\mathscr{A}}(i) = \overline{\mathscr{A}_i}$ for every $i \in D$. Note that a $*$-tree S has an $\overline{\mathscr{A}}$-$*$-subtree iff S has an $\mathscr{A}$-$*$-subtree.

LEMMA 4·16. *Let S be a $*$-tree with alphabet $\mathbb{A}$, and let $\mathscr{A} : D \to 2^{(2^{\mathbb{A}^*})}$ be as above. Then S has an infinite $\mathscr{A}$-$*$-subtree $\Leftrightarrow$ $\forall n \in \mathbb{N}$, S has an $\mathscr{A}$-$*$-subtree of level n.*

Proof. The direction $\Rightarrow$ is trivial. For the other direction, let $T^{(n)}$ be $\mathscr{A}$-$*$-subtrees of level n s.t. $T^{(n)}_{n+1} = \emptyset$. Define $T = \bigcup_{n=0}^{\infty} T^{(n)} \subseteq S$. This is a $*$-subtree of S. Define

$T' = \{v \in T : T^v \text{ is infinite}\}$. Then T' is a $*$-subtree of T. In fact, T' is an $\overline{\mathscr{A}}$-$*$-subtree. To see this, notice that for every $v \in T'$, $v \in T^{(n)}$ for infinitely many n. Therefore, T^v has $\mathscr{A}^v$-$*$-subtrees of every level. Let $A = W_T(v) \setminus W_{T'}(v)$, and denote $k = \min\{n \in \mathbb{N} : \forall w \in A,\ (T^{vw})_n = \emptyset\}$. T^v has an $\mathscr{A}^v$-$*$-subtree of level $> k+1$. This subtree can not contain any elements from A, hence $W_{T'}(v) \in \overline{\mathscr{A}_v}$.

Let A be some fixed finite set, and let X be a random subset of A. Given $s \in [0, 1]$, we denote $X^{(s)} = X \cap Y$ where $Y \sim \mathrm{Bin}(A,\ 1-s)$.

THEOREM 4·17. *Let S be a random $*$-tree on the $*$-tree $B \subseteq \mathbb{A}^*$ with bounded offspring distributions $\{M_x\}_{x \in B}$. Let $\mathscr{A} : B' \to 2^{(2^{\mathbb{A}^*})}$ be s.t. $\forall x \in B'$, $\mathscr{A}_x \subseteq 2^{W_{B'}(x)}$ is monotonic. Define $g_{\mathscr{A}} : [0, 1] \to [0, 1]$ by $g_{\mathscr{A}}(s) = \sup_{x \in B'} \mathbb{P}\left(M_x^{(s)} \notin \mathscr{A}_x\right)$. Let s_0 be the smallest fixed point of $g_{\mathscr{A}}$ in $[0, 1]$. Then*

$$\sup_{x \in B'} \mathbb{P}\left(S^x \text{ has no } \mathscr{A}^x\text{-}*\text{-subtree} \mid x \in S\right) \leq s_0. \tag{4·1}$$

Proof. Since the collection $\{M_x\}_{x \in B'}$ is bounded, the collection of functions

$$\left\{\mathbb{P}\left(M_x^{(s)} \notin \mathscr{A}_x\right)\right\}_{x \in B'}$$

is equicontinuous and therefore $g_{\mathscr{A}}$ is continuous. Also, monotonicity of $\mathscr{A}_x$ for every $x \in B'$ implies that $g_{\mathscr{A}}$ is monotonically increasing. These 2 properties of $g_{\mathscr{A}}$ imply that $\lim_{n\to\infty} g_{\mathscr{A}}^n(0)$ is the smallest fixed point of $g_{\mathscr{A}}$, where $g_{\mathscr{A}}^n$ denotes the composition of $g_{\mathscr{A}}$ with itself n times. So $s_0 = \lim_{n\to\infty} g_{\mathscr{A}}^n(0)$. Define

$$q_n = \sup_{x \in B'} \mathbb{P}\left(S^x \text{ has no } \mathscr{A}^x\text{-}*\text{-subtree of level } n \mid x \in S\right).$$

Claim $\forall n \geq 1,\ g_{\mathscr{A}}^n(0) \geq q_n$.

Proof of claim. First, notice that $g_{\mathscr{A}}(0) = q_1$. Now, assume the claim is true for n. So $g_{\mathscr{A}}^n(0) \geq q_n$, which implies by monotonicity of $g_{\mathscr{A}}$ that $g_{\mathscr{A}}^{n+1}(0) \geq g_{\mathscr{A}}(q_n) = \sup_{x \in B'} \mathbb{P}\left(M_x^{(q_n)} \notin \mathscr{A}_x\right)$. For every $x \in B'$,

$$\begin{aligned}
&\mathbb{P}\left(M_x^{(q_n)} \notin \mathscr{A}_x\right) \\
&\qquad \geq \mathbb{P}\left(\left\{a \in W_S(x) : S^{xa} \text{ has an } \mathscr{A}^{xa}\text{-}*\text{-subtree of level } n\right\} \notin \mathscr{A}_x \mid x \in S\right) \\
&\qquad = \mathbb{P}\left(S^x \text{ has no } \mathscr{A}^x\text{-}*\text{-subtree of level } n{+}1 \mid x \in S\right)
\end{aligned}$$

(the inequality uses the fact that q_n is defined as a supremum and the monotonicity of $\mathscr{A}_x$). By taking supremums we obtain $g_{\mathscr{A}}(q_n) \geq q_{n+1}$ which finishes the proof of the claim.

The sequence (q_n) is monotonically increasing and bounded by s_0, so $q = \lim_{n\to\infty} q_n \leq s_0$. We now need the following elementary lemma whose proof is left to the reader:

LEMMA *Let $f_n : A \to \mathbb{R}$ be a sequence of functions s.t. $\forall a \in A$, the sequence $f_n(a)$ is monotonically increasing, and the functions f_n are uniformly bounded. Then*

$$\lim_{n\to\infty} \sup_{a \in A} f_n(a) = \sup_{a \in A} \lim_{n\to\infty} f_n(a).$$

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Combining the above lemma with Lemma 4·16, we obtain that

$$q = \sup_{x \in B'} \mathbb{P}\left(S^x \text{ has no } \mathscr{A}^x\text{-}*\text{-subtree} \mid x \in S\right)$$

which concludes the proof of the theorem.

Remark 4·18. Equality in equation (4·1) need not hold (see Example 5·1). However, in the special case of GWTs, and assuming the map $\mathscr{A}$ is constant, equality does hold and the theorem becomes a generalisation of Pakes–Dekking theorem. see Section 5 for more details.

4·5. *Proof of main theorem*

4·5·1. *Ahlfors-regularity*

LEMMA 4·19. *Let T be a supercritical GWT with alphabet $\mathbb{A}$, weights $\{r_i\}_{i \in \mathbb{A}}$, and offspring distribution W, and let δ satisfy $\mathbb{E}\left(\sum_{i \in W} r_i^\delta\right) = 1$. Then $\forall \alpha \in (0, \delta)$, $\forall s \in (0, 1)$,*

$$\sup_{x \in B_\rho} \mathbb{P}\left(\left|\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)}\right| < \rho^{-\alpha} \mid \textit{nonextinction}\right) \underset{\rho \to 0}{\longrightarrow} 0,$$

where $B_\rho = \{\emptyset\} \cup \left(\bigcup_{n=1}^\infty \Pi_{\rho^n}\right)$.

Proof. First, we recall that $a_\rho(x) \in [r_{min}, 1]$ for every $x \in B_\rho$. Since for any positive constant C, $\mathbb{P}\left(\left|T_{\Pi_\varepsilon}\right| < C \mid \text{nonextinction}\right)$ decreases as ε decreases, we have for any $x \in B_\rho$,

$$\mathbb{P}\left(\left|\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)\right| < \frac{1}{\rho^\alpha} \mid \text{nonextinction}\right) \leq \mathbb{P}\left(\left|T_{\Pi_{\frac{\rho}{r_{min}}}}\right| < \frac{1}{\rho^\alpha} \mid \text{nonextinction}\right).$$

Hence, it is enough to show that $\mathbb{P}\left(\left|\left(T_{\Pi_{\frac{\rho}{r_{min}}}}\right)^{(s)}\right| < \frac{1}{\rho^\alpha} \mid \text{nonextinction}\right) \underset{\rho \to 0}{\longrightarrow} 0.$

Fix some $\beta \in (\alpha, \delta)$. Given any $\varepsilon > 0$, by Corollary 4·5

$$\mathbb{P}\left(\left|T_{\Pi_{\frac{\rho}{r_{min}}}}\right| > \left(\frac{r_{min}}{\rho}\right)^\beta \mid \text{nonextinction}\right) > 1 - \varepsilon$$

whenever ρ is small enough. Given some $s \in (0, 1)$, for a small enough ρ,

$$\left(\frac{r_{min}}{\rho}\right)^\beta > \frac{2}{1-s} \cdot \frac{1}{\rho^\alpha}$$

so that

$$\mathbb{P}\left(\left|T_{\Pi_{\frac{\rho}{r_{min}}}}\right| > \frac{2}{(1-s)\,\rho^\alpha} \mid \text{nonextinction}\right) > 1 - \varepsilon.$$

By Chebyshev's inequality

$$\mathbb{P}\left(\text{Bin}\left(\frac{2}{(1-s)\,\rho^\alpha}, 1-s\right) < \frac{1}{\rho^\alpha}\right) \leq \frac{\frac{2}{(1-s)\rho^\alpha} s\,(1-s)}{\left(\frac{2}{(1-s)\rho^\alpha}\,(1-s) - \frac{1}{\rho^\alpha}\right)^2} = 2s\rho^\alpha \underset{\rho \to 0}{\longrightarrow} 0.$$

Therefore, $\mathbb{P}\left(\left|\left(T_{\Pi_{\frac{\rho}{r_{min}}}}\right)^{(s)}\right| < \frac{1}{\rho^\alpha} \mid \text{nonextinction}\right) \underset{\rho \to 0}{\longrightarrow} 0$ as required.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Remark 4·20. Applying Lemma 4·19 to sections of the form $\mathbb{A}^k$ (setting all the weights to be equal to some $r \in (0, 1)$, and taking $\rho = r^k$), one may obtain that for every sequence $(a_k)_{k\in\mathbb{N}}$ of positive integers with the property $\limsup_{k\to\infty} \sqrt[k]{a_k} < m$,

$$\forall s \in (0, 1), \quad \mathbb{P}\left(\left|T^{(s)}_{\mathbb{A}^k}\right| < a_k\right) \underset{k\to\infty}{\longrightarrow} \mathbb{P}\,(\text{extinction}).$$

From that, using Theorem 4·17, letting the compression of T along the sections $\left(\mathbb{A}^{kn}\right)_{n\in\mathbb{N}}$ be denoted by $T_{\{k\}}$, one may deduce that

$$\mathbb{P}\left(T_{\{k\}} \text{ contains an } a_k\text{-ary subtree} |\text{ nonextinction}\right) \underset{k\to\infty}{\longrightarrow} 1.$$

Since for every k, $T_{\{k\}}$ is itself a GWT (whose offspring distribution is not Binomial), this result is in the spirit of [**15**, proposition 5·31].

LEMMA 4·21. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS satisfying the OSC, with contraction ratios $\{r_i\}_{i\in\Lambda}$, and let $T \in \mathscr{T}_\Lambda$ be an infinite tree with alphabet Λ. Let $\rho \in (0, r_{min})$ and $\alpha > 0$ be s.t. $\rho^{-\alpha}$ is an integer, and let S be a $\rho^{-\alpha}$-ary $*$-subtree of $\left(T_{(\rho)}\right)^\omega$ for some $\omega \in T_{(\rho)}$. Then $E = \gamma_\Phi\,(\partial S)$ is α-Ahlfors regular.*

Proof. Construct a probability measure μ on $\Lambda^{\mathbb{N}}$ by equally distributing mass at each level of the tree, i.e., $\forall i \in S,\ \mu\,([i]) = \rho^{h_S(i)\alpha}$. Let ν be the projection of μ to $\mathbb{R}^d$, i.e., $\nu = (\gamma_\Phi)_*\,\mu$. Obviously $\text{supp}\,(\nu) = E$. We show that ν is α-Ahlfors regular.

Fix any $r > 0$ and $x \in E$. Let n be the unique integer s.t. $\rho^{n+1}/a_\rho\,(\omega) < r \le \rho^n/a_\rho\,(\omega)$. By Lemma 4·7 and Proposition 4·12, the ball $B_{\frac{\rho^n}{a_\rho(\omega)}}\,(x)$ intersects at most C of the sets $\{\varphi_i\,(K)\}_{i\in S_n}$, where K is the attractor of Φ and $C > 0$ is some constant not depending on x and n. Therefore,

$$\nu\,(B_r\,(x)) \le \sum_{i\in S_n,\ \varphi_i(K)\cap B_r(x)\neq\emptyset} \mu\,([i]) \le C\cdot\rho^{n\alpha} \le C\rho^{-\alpha}\cdot r^\alpha.$$

On the other hand, let $j \in S_{m+1}$ be s.t. $x \in \varphi_j\,(K)$ and $\rho^{m+1}/a_\rho\,(\omega) < \frac{r}{\Delta} \le \rho^m/a_\rho\,(\omega)$, where $\Delta = \text{diam}\,(K)$. Then $\text{diam}\left(\varphi_j\,(K)\right) = r_j\Delta \le \rho^{m+1}\Delta/a_\rho\,(\omega) < r$, and therefore $\varphi_j\,(K) \subseteq B_r\,(x)$. Hence

$$\nu\,(B_r\,(x)) \ge \nu\left(\varphi_j\,(K)\right) = \rho^{(m+1)\alpha} \ge \left(\frac{\rho a_\rho\,(\omega)}{\Delta}\right)^\alpha\cdot r^\alpha$$

4·5·2. *Diffuseness*

PROPOSITION 4·22. *Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS with an attractor K and contraction ratios $\{r_i\}_{i\in\Lambda}$. Let $T \subset \Lambda^*$ be an infinite tree. Let $T_{(\rho)}$ be the compression of T for some $\rho \in (0, r_{min})$. Let S be a $*$-subtree of $\left(T_{(\rho)}\right)^\omega$ for some $\omega \in T_{(\rho)}$, and assume that there exists some $c > 0$ s.t. $\forall i \in S$, $W_S\,(i)$ is (K, c)-diffuse. Then the limit set $E = \gamma_\Phi\,(S)$ is $\rho c r_{min}/\Delta$-diffuse, where $\Delta = \text{diam}\,(K)$.*

Proof. Fix $\xi \in \left(0, \dfrac{\rho\Delta}{a_\rho\,(\omega)}\right)$, $x \in E$ and an affine hyperplane $\mathcal{L} \subset \mathbb{R}^d$. Let $n \in \mathbb{N}$ be the unique integer s.t. $\rho^n \le \xi a_\rho\,(\omega)/\Delta < \rho^{n-1}$, and let $i \in S_n$ be s.t. $x \in \varphi_i K$. Note that by Proposition 4·12 $i \in \Pi_{\frac{\rho^n}{a_\rho(\omega)}}$, and therefore,

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

$$\frac{r_{min}\rho}{\Delta}\xi < \frac{r_{min}\rho^n}{a_\rho(\omega)} < r_i \le \frac{\rho^n}{a_\rho(\omega)} \le \frac{\xi}{\Delta}.$$

By assumption $\exists j \in W_S(i)$ s.t. $\varphi_j K \cap \left(\varphi_i^{-1}\mathcal{L}\right)^{(c)} = \emptyset$. Applying φ_i we get that $\varphi_{ij}K \cap \mathcal{L}^{(r_i c)} = \emptyset$, hence $\varphi_{ij}K \cap \mathcal{L}^{\left(\frac{r_{min}\rho c}{\Delta}\xi\right)} = \emptyset$. Notice that diam $(\varphi_i K) = r_i\Delta \le \xi$, therefore $\varphi_{ij}K \subset \varphi_i K \subseteq B_\xi(x)$. By assumption $\forall v \in S$, $W_S(v)$ is (K, c)-diffuse and in particular non-empty, so every descendants tree of S is infinite. Hence, $\varphi_{ij}K \cap E \neq \emptyset$ and therefore

$$B_\xi(x) \cap E \setminus \mathcal{L}^{\left(\frac{r_{min}\rho c}{\Delta}\xi\right)} \neq \emptyset.$$

LEMMA 4·23. *Let T be a GWT corresponding to a similarity IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ with offspring distribution W. Assume that $\mathbb{P}(W$ is (F, c)-diffuse$) = \nu > 0$ for some $c > 0$ and F as in Definition 3·2. Then for every section $\Pi \subset \Lambda^*$,*

$$\mathbb{P}(T_\Pi \text{ is } (F, c)\text{-diffuse}) \ge \nu \cdot \mathbb{P}(T_n \neq \emptyset)^{|\Lambda|},$$

where $n = \max\{|i| : i \in \Pi\}$ is the maximal depth of the section Π.

Proof. First we note that if $A \subseteq \Lambda$ is (F, c)-diffuse, then for every finite set $P \subseteq \Lambda^*$ s.t. $\forall i \in A$, $\exists j \in P$, s.t. $j \ge i$, P is also (F, c)-diffuse. Indeed, given an affine hyperplane $\mathcal{L} \subset \mathbb{R}^d$, $\varphi_i F \cap \mathcal{L}^{(c)} = \emptyset$ for some $i \in A$. For some $j \in P$, $j \ge i$ and therefore $\varphi_j F \subseteq \varphi_i F$ which implies $\varphi_j F \cap \mathcal{L}^{(c)} = \emptyset$. Hence, given some section $\Pi \subset \Lambda^*$, if T_1 is (F, c) - diffuse and $\forall i \in T_1$, $\exists j \in T_\Pi$ s.t. $j \ge i$, then T_Π is also (F, c) - diffuse. Since for every $i \in \Lambda$,

$$\mathbb{P}(\exists j \in T_\Pi \text{ s.t. } j \ge i \mid i \in T_1) \ge \mathbb{P}(T_n \neq \emptyset)$$

and these events are independent for different elements of Λ and also independent of the event: T_1 is (F, c) - diffuse, the claim follows.

For the next lemma we need the following notation. Let $\Phi = \{\varphi_i\}_{i\in\Lambda}$ be a similarity IFS with contraction ratios $\{r_i\}_{i\in\Lambda}$ and attractor K. Given $\rho \in (0, r_{min})$, denote $B_\rho = \{\emptyset\} \cup \left(\bigcup_{n=1}^\infty \Pi_{\rho^n}\right)$. For every $c > 0$ and $x \in B_\rho$, denote

$${}^{\rho,c}\mathcal{D}_x = \left\{A \subseteq \Pi_{\frac{\rho}{a_\rho(x)}} : \exists i \in \Pi_{\frac{\rho}{a_\rho(x) r_{min}}},\ \{\varphi_v : iv \in A\} \text{ is } (K, c)\text{-diffuse}\right\}.$$

We denote ${}^{\rho,c}\mathcal{D} : B_\rho \to 2^{\left(2^{\Lambda^*}\right)}$ where ${}^{\rho,c}\mathcal{D}(x) = {}^{\rho,c}\mathcal{D}_x$.

LEMMA 4·24. *Let T be a GWT corresponding to a similarity IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ with contraction ratios $\{r_i\}_{i\in\Lambda}$ and offspring distribution W. Denote by K the attractor of Φ. Assume that $\exists c > 0$ s.t. $\mathbb{P}\left(\{\varphi_i\}_{i\in W}\text{ is } (K, c)\text{-diffuse}\right) = \nu > 0$. Then for every $s \in (0, 1)$,*

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)} \notin {}^{\rho,c}\mathcal{D}_x \mid \text{nonextinction}\right) \xrightarrow[\rho\to 0]{} 0,$$

where $B_\rho = \{\emptyset\} \cup \left(\bigcup_{n=1}^\infty \Pi_{\rho^n}\right)$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Proof. Recall that $a_\rho(x) \in [r_{min}, 1]$ for every $x \in B_\rho$. Fix some $\alpha \in (0, \delta)$ where δ satisfies the equation $\mathbb{E}\left(\sum_{i\in\Lambda} r_i^\delta\right) = 1$. Then by Corollary 4·5, given $\varepsilon > 0$ there exists some $\rho_0 > 0$ s.t. whenever $\rho < \rho_0$,

$$\mathbb{P}\left(\left|T \cap \Pi_{\frac{\rho}{(r_{min})^2}}\right| > (r_{min})^{2\alpha} \cdot \frac{1}{\rho^\alpha}\right) > 1 - \varepsilon.$$

Fix some $x \in B_\rho$. Since $a_\rho(x) \geq r_{min}$, for every $\rho < \rho_0$,

$$\mathbb{P}\left(\left|T \cap \Pi_{\frac{\rho}{a_\rho(x)\cdot r_{min}}}\right| > \frac{1}{\rho^\alpha}\right) > 1 - \varepsilon, \tag{4·2}$$

where the constant $(r_{min})^{2\alpha}$ was removed as it may be absorbed by taking a slightly smaller α and assuming that ρ_0 is small enough.

For every $i \in \Pi_{\frac{\rho}{a_\rho(x)\cdot r_{min}}}$,

$$\mathbb{P}\left(\left\{\varphi_j\right\}_{j\in W_T(i)} \text{ is } (K, c)\text{-diffuse}|\, i \in T\right) = \nu > 0.$$

Denote $V_i = \left\{j \in \Lambda^* : ij \in \Pi_{\frac{\rho}{a_\rho(x)}}\right\}$. V_i is a section, and $\max_{j\in V_i} |j| \leq \left\lceil \log_{r_{max}} (r_{min})^2 \right\rceil =: n_0$. Note that n_0 is independent of i and of x. By Lemma 4·23,

$$\mathbb{P}\left(\left\{\varphi_j\right\}_{j\in T_{V_i}} \text{ is } (K, c)\text{-diffuse}\right) \geq \nu \cdot \mathbb{P}\left(T_{n_0} \neq \emptyset\right)^{|\Lambda|},$$

and given $s \in (0, 1)$,

$$\mathbb{P}\left(\left\{\varphi_j\right\}_{j\in T_{V_i}^{(s)}} \text{ is } (K, c)\text{-diffuse}\right) \geq \nu \cdot \mathbb{P}\left(T_{n_0} \neq \emptyset\right)^{|\Lambda|} \cdot (1 - s)^{\left(|\Lambda|^{n_0}\right)}. \tag{4·3}$$

We denote the right-hand side of inequality (4·3) by ν'. Notice that conditioned on $i \in T$, $\left\{v \in \Lambda^* : iv \in T \cap \Pi_{\frac{\rho}{a_\rho(x)}}\right\} \sim T_{V_i}$. Therefore,

$$\forall i \in \Pi_{\frac{\rho}{a_\rho(x)\cdot r_{min}}},\ \mathbb{P}\left(\left\{v \in \Lambda^* : iv \in T \cap \Pi_{\frac{\rho}{a_\rho(x)}}\right\} \text{ is } (K, c)\text{-diffuse}|\, i \in T\right) \geq \nu'.$$

Using inequality (4·2) we conclude the proof of the lemma.

4·5·3. *Final step*

Proof of Theorem 1·8. Let E be a non-planar GWF and T the corresponding GWT, w.r.t. a similarity IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ whose attractor is denoted by K, and offspring distribution W. First, note that by Proposition 3·9, there exist a section $\Pi \subseteq \Lambda^*$, and $c > 0$ s.t. $\mathbb{P}(T_\Pi \text{ is } (K, c)\text{-diffuse}) > 0$. Since we may consider the IFS $\{\varphi_i\}_{i\in\Pi}$ which has the same attractor as Φ, and the GWF corresponding to the offspring distribution $\sim T_\Pi$ which has the same law as E, there is no loss of generality in assuming that $\mathbb{P}(W \text{ is } (K, c)\text{-diffuse}) > 0$, hence we proceed assuming the latter holds.

Given $\rho > 0$, consider the compressed $*$-tree $T_{(\rho)}$. Recall that by Proposition 4·14, $T_{(\rho)}$ has the law of a random $*$-tree on $B_\rho = \bigcup_{n=1}^{\infty} \Pi_{\rho^n} \cup \{\emptyset\}$ with offspring distributions $M_i \sim T_{\Pi_{\frac{\rho}{a_\rho(i)}}}$, for every $i \in B_\rho$. By Lemma 4·24, for every $s \in (0, 1)$,

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)} \notin {}^{\rho,c}\mathscr{D}_x \,|\, \text{nonextinction}\right) \xrightarrow[\rho\to 0]{} 0.$$

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

Fix any $\alpha \in (0, \delta)$. By Lemma 4·19,

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left|\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)}\right| < \rho^{-\alpha} | \text{ nonextinction}\right) \underset{\rho\to 0}{\longrightarrow} 0.$$

Denoting for every $x \in B_\rho$, ${}^{\rho,\alpha}\mathscr{C}_x = \left\{A \subseteq \Pi_{\frac{\rho}{a_\rho(x)}} : |A| \geq \rho^{-\alpha}\right\}$, we have

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)} \notin {}^{\rho,c}\mathscr{D}_x \cap {}^{\rho,\alpha}\mathscr{C}_x | \text{ nonextinction}\right) \underset{\rho\to 0}{\longrightarrow} 0$$

which implies that

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left(T_{\Pi_{\frac{\rho}{a_\rho(x)}}}\right)^{(s)} \notin {}^{\rho,c}\mathscr{D}_x \cap {}^{\rho,\alpha}\mathscr{C}_x\right) \underset{\rho\to 0}{\longrightarrow} \mathbb{P}\left(\text{extinction}\right).$$

Therefore, we have shown that for every $s \in (0, 1)$,

$$\sup_{x\in B_\rho} \mathbb{P}\left(M_x^{(s)} \notin {}^{\rho,c}\mathscr{D}_x \cap {}^{\rho,\alpha}\mathscr{C}_x\right) \underset{\rho\to 0}{\longrightarrow} \mathbb{P}\left(\text{extinction}\right).$$

This implies that as a function of s, $\sup_{x\in B_\rho} \mathbb{P}\left(M_x^{(s)} \notin {}^{\rho,c}\mathscr{D}_x \cap {}^{\rho,\alpha}\mathscr{C}_x\right)$ has a fixed point <1 whenever ρ is small enough. Denote ${}^{\rho,c}\mathscr{A} : B_\rho \to 2^{(2^{\Lambda^*})}$ given by ${}^{\rho,c}\mathscr{A}\,(x) = {}^{\rho,c}\mathscr{D}_x \cap {}^{\rho,\alpha}\mathscr{C}_x$. Since ${}^{\rho,c}\mathscr{A}\,(x)$ is monotonic for every $x \in B_\rho$, by Theorem 4·17 we obtain

$$\sup_{x\in B_\rho} \mathbb{P}\left(\left(T_{(\rho)}\right)^x \text{ has no } \left({}^{\rho,c}\mathscr{A}\right)^x\text{-}*\text{-subtree}|\ x \in T_{(\rho)}\right) < 1$$

whenever ρ is small enough, and in this case,

$$\mathbb{P}\left(\exists x \in T_{(\rho)} \text{ s.t. } \left(T_{(\rho)}\right)^x \text{ has a } \left({}^{\rho,c}\mathscr{A}\right)^x\text{-}*\text{-subtree}| \text{ nonextinction}\right) = 1. \tag{4·4}$$

Fix ρ small enough s.t. (4·4) holds, and s.t. $\rho^{-\alpha}$ is an integer larger than $|\Lambda|^{n_0}$, where $n_0 = \left\lceil \log_{r_{max}} (r_{min})^2 \right\rceil$ as in the proof of Lemma 4·24. Then every $\left({}^{\rho,c}\mathscr{A}\right)^x$ - $*$ -subtree contains a $\rho^{-\alpha}$-ary $\left({}^{\rho,c}\mathscr{D}\right)^x$-$*$-subtree. Indeed, if $A \in \left({}^{\rho,c}\mathscr{D}\right)^x_v$ for some $v \in B^x_\rho$, we may remove elements of A except for a subset of size at most $|\Lambda|^{n_0}$ and obtain a smaller set which is still in $\left({}^{\rho,c}\mathscr{D}\right)^x_v$.

Now, assume that for some $x \in T_{(\rho)}$, $\left(T_{(\rho)}\right)^x$ has a $\left({}^{\rho,c}\mathscr{A}\right)^x$-$*$-subtree, then by the above, it also contains a $\rho^{-\alpha}$-ary $\left({}^{\rho,c}\mathscr{D}\right)^x$-$*$-subtree S. Denote $D'_\alpha = \gamma_\Phi\,(\partial S)$. Since $\forall v \in B_\rho$, every set in ${}^{\rho,c}\mathscr{D}_v$ is $(K, \rho c)$ - diffuse, by Proposition 4·22, D'_α is hyperplane diffuse, and so is $D_\alpha = \varphi_x\left(D'_\alpha\right) \subseteq E$.

In case Φ satisfies the OSC, Lemma 4·21 implies that D'_α is also α-Ahlfors regular (and so is D_α). Thus, in this case we have shown that for every $\alpha \in (0, \delta)$, a.s. conditioned on nonextinction, there exists a subset $D_\alpha \subseteq E$ which is hyperplane diffuse and α-Ahlfors regular. Taking a sequence $\alpha_n \nearrow \delta$ concludes the proof.

Remark 4·25. The proof of Theorem 1·8 for the case without the OSC could obviously be much shorter since the existence of a ${}^{\rho,c}\mathscr{D}$-$*$-subtree suffices.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

5. *Some remarks on Theorem 4·17*

5·1. *A counterexample for equality in equation (4·1)*

As mentioned in Remark 4·18, equality in equation 4·1 need not hold. We now provide a counterexample.

Example 5·1. Let $\mathbb{A} = \{a_1, a_2, ..., a_n\}$ be an alphabet and let $B = \{\emptyset\} \cup \{a_1, a_2\} \mathbb{A}^* \subset \mathbb{A}^*$ be a $*$-tree. Let $S \subset \mathbb{A}^*$ be a random $*$-tree on B with offspring distributions $\{M_i\}_{i \in B}$ given by:

(i) $\mathbb{P}(M_i = \mathbb{A}) = 1$ for $i \in \{a_1, a_2\}$;
(ii) $\forall i \in B,\ |i| \geq 2 \Rightarrow M_i \sim \text{Bin}(\{a_1, a_2, a_3\}, p)$ where $p \in (0, 1)$ is large enough so that a GWT with alphabet of size 3 and binomial offspring distribution with parameter p has a positive probability, $\alpha > 0$, of containing a binary subtree (by Pakes-Dekking theorem there exists such p);
(iii) $\mathbb{P}(M_\emptyset = \{a_1, a_2\}) = \alpha + \varepsilon,\ \mathbb{P}(M_\emptyset = \{a_1\}) = 1 - (\alpha + \varepsilon)$ for some small $\varepsilon > 0$.

Now, define $\mathscr{A}_x = \{L \subseteq \mathbb{A}^* : |L| \geq 2\}$ for every $x \in B$, so that $\mathscr{A}$-$*$-trees are $*$-trees which contain binary trees.

Choosing n large enough, we may guarantee that the sup in equation (4·1) is realised by every element $x \in B$ with $|x| \geq 2$ and its value is $1 - \alpha$, that is to say that $q = 1 - \alpha$ where q is as defined in the proof of Theorem 4·17. In that proof we have shown that $q \leq s_0$ where s_0 is the smallest fixed point of the function $g_{\mathscr{A}}(s) = \sup_{x \in B'} \mathbb{P}\left(M_x^{(s)} \notin \mathscr{A}_x\right)$ in [0, 1].

Analysing $g_{\mathscr{A}}(q) = \sup_{x \in B'} \mathbb{P}\left(M_x^{(q)} < 2\right)$, one should notice that the sup in the formula for $g_{\mathscr{A}}(q)$ is realised by $x = \emptyset$. This is because $\alpha < \mathbb{P}(\text{Bin}(3, p) \geq 2)$ and ε may be chosen to be arbitrarily small. So $g_{\mathscr{A}}(q) = \mathbb{P}\left(M_\emptyset^{(q)} < 2\right) = 1 - \left((\alpha + \varepsilon) \cdot \alpha^2\right)$ which is strictly larger than q when ε is small enough.

5·2. *The special case of GWT*

We now focus on the special case of GWT, and assume that the mapping $\mathscr{A}$ is constant. We show that in this case equality in equation 4·1 does hold. The following theorem may be considered an extension of Pakes–Dekking theorem and its proof uses the ideas of the proof given in [**15**, theorem 5·29].

Let T be a GWT with alphabet $\mathbb{A}$ and any offspring distribution W. Let $\mathscr{A} \subseteq 2^{\mathbb{A}}$ be some nonempty collection of nonempty subsets of $\mathbb{A}$. Define the function $g_{\mathscr{A}} : [0, 1] \to [0, 1]$ by $g_{\mathscr{A}}(s) = \mathbb{P}\left(W^{(s)} \notin \overline{\mathscr{A}}\right)$. Finally, denote $\tau(\mathscr{A}) = \mathbb{P}(T \text{ has an } \mathscr{A}\text{-subtree})$.

THEOREM 5·2. *With notations as above, $1 - \tau(\mathscr{A})$ is the smallest fixed point of $g_{\mathscr{A}}$ in* [0, 1].

Proof. Note that the following properties hold:

(i) $g_{\mathscr{A}}$ is continuous and monotonically increasing;
(ii) $g_{\mathscr{A}}(1) = 1$;
(iii) $g_{\mathscr{A}}(0) = \mathbb{P}\left(W \notin \overline{\mathscr{A}}\right)$.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

If $g_{\mathscr{A}}(0)=0$, then $\mathbb{P}\left(W \in \overline{\mathscr{A}}\right)=1$ which implies the existence of an $\mathscr{A}$ - subtree a.s., i.e., $1-\tau\left(\mathscr{A}\right)=0$ and the claim follows. Otherwise, we assume that $g_{\mathscr{A}}(0)>0$. Let q_n be the probability that T does not contain an $\mathscr{A}-$ subtree of length n, where $q_0=0$. Then $1-q_n \searrow \mathbb{P}\left(T \text{ has an } \mathscr{A}\text{-subtree of every length}\right)$ and by Lemma 4·16 this is equivalent to $q_n \nearrow 1-\tau\left(\mathscr{A}\right)$.

Claim For every $n \geq 1$, $q_n = g_{\mathscr{A}}\left(q_{n-1}\right)$.

Proof of claim. Denote for every $n \geq 1$, the following random set:

$$V_n = \left\{ v \in T_1 : T^v \text{ has an } \mathscr{A}\text{-subtree of length } n \right\}.$$

For each element $v \in \mathbb{A}$, $\mathbb{P}\left(v \in V_n \mid v \in T_1\right) = 1-q_n$, and for every two distinct elements in $\mathbb{A}$ these events are independent, so $V_n \sim W^{(q_n)}$. Now, since T has an $\mathscr{A}$-subtree of length n iff $V_{n-1} \in \overline{\mathscr{A}}$,

$$q_n = \mathbb{P}\left(V_{n-1} \notin \overline{\mathscr{A}}\right) = \mathbb{P}\left(W^{(q_{n-1})} \notin \overline{\mathscr{A}}\right) = g_{\mathscr{A}}\left(q_{n-1}\right).$$

Since $g_{\mathscr{A}}$ is increasing and continuous, its smallest fixed point is $\lim_{n\to\infty} g_{\mathscr{A}}^n(0)$, where $g_{\mathscr{A}}^n$ denotes the composition of $g_{\mathscr{A}}$ with itself n times (this is a general property of increasing and continuous functions on [0, 1] whose proof is easy and left to the reader). By the claim above, $\lim_{n\to\infty} g_{\mathscr{A}}^n(0) = \lim_{n\to\infty} q_n$ which concludes the proof.

Acknowledgements. This work is a part of the author's doctoral thesis written under the supervision of Prof. Barak Weiss. The author was partially supported by ISF grant 2095/15 and BSF grant 2016256.

REFERENCES

[1] J. AN, L. GUAN and D. KLEINBOCK. Bounded orbits of diagonalisable flows on $SL_3(\mathbb{R})/SL_3(\mathbb{Z})$. *Internat. Math. Res. Notices*, 2015. **(24)** (2015), 13623–13652.

[2] R. BRODERICK, L. FISHMAN and D. KLEINBOCK. Schmidt's game, fractals, and orbits of toral endomorphisms. *Ergodic Theory Dynam. Syst.* **31**(4) (2011), 1095–1107.

[3] R. BRODERICK, L. FISHMAN, D. KLEINBOCK, A. REICH and B. WEISS. The set of badly approximable vectors is strongly C^1 incompressible. *Math. Proc. Camb. Phil. Soc.* **153**(2) (2012), 319–339.

[4] R. BRODERICK, L. FISHMAN and D. SIMMONS. Badly approximable systems of affine forms and incompressibility on fractals. *J. Number Theory* **133**(7) (2013), 2186–2205.

[5] Y. BUGEAUD. *Distribution Modulo One and Diophantine Approximation*. Cambridge Tracts in Math. (Cambridge University Press, 2012).

[6] T. DAS, L. FISHMAN, D. SIMMONS and M. URBAŃSKI. Badly approximable vectors and fractals defined by conformal dynamical systems. ArXiv e-prints, (Mar. 2016).

[7] Y. DAYAN. Diophantine approximations on random fractals. arXiv e-prints, (Jul 2018).

[8] K. J. FALCONER. Random fractals. *Math. Proc. Camb. Phil. Soc.* **100**(3) (1986), 559–582.

[9] K. J. FALCONER. *Fractal Geometry: Mathematical Foundations and Applications* (Wiley, 2013).

[10] L. FISHMAN, D. KLEINBOCK, K. MERRILL and D. SIMMONS. Intrinsic diophantine approximation on manifolds: General theory. *Trans. Amer Math. Soc.* **370**(1) (2018), 577–599.

[11] J. HAWKES. Trees generated by a simple branching process. *J. London Math. Soc.* **24**(2) (1981), 373–384.

[12] J. E. HUTCHINSON. Fractals and self similarity. *Indiana Univ. Math. J.* **30**(5) (1981), 713–747.

[13] H. KESTEN and B. P. STIGUM. A limit theorem for multidimensional Galton–watson processes. *Ann. Math. Statistics*. **37**(5) (1966), 1211–1223.

[14] D. KLEINBOCK, E. LINDENSTRAUSS and B. WEISS. On fractal measures and diophantine approximation. *Selecta Math.* **10**(4) (2005), 479–523.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press

[15] R. LYONS and Y. PERES. *Probability on Trees and Networks* (Cambridge University Press, New York, 2016).
[16] R. D. MAULDIN and S. C. WILLIAMS. Random recursive constructions: Asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.* **295**(1) (1986), 325–346.
[17] P. A. P. MORAN. Additive functions of intervals and hausdorff measure. *Math. Proc. Camb. Phil. Soc*. **42**(1) (1946), 15–23.
[18] P. MÖRTERS and Y. PERES. *Brownian Motion*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, 2010).
[19] E. NESHARIM and D. SIMMONS. Bad(s, t) is hyperplane absolute winning. *Acta Arith.* **164**(2) (2014), 145–152.
[20] A. PAKES and F. DEKKING. On family trees and subtrees of simple branching processes. *J. Theoret. Probab.* **4**(2) (1991), 353–369.
[21] W. M. SCHMIDT. On badly approximable numbers and certain games. *Trans. Amer. Math. Soc.* **123**(1) (1966) 178–199, 1966.

https://doi.org/10.1017/S0305004121000360 Published online by Cambridge University Press