BRANCHING RANDOM WALK WITH EXPONENTIALLY DECREASING STEPS, AND STOCHASTICALLY SELF-SIMILAR MEASURES

ITAI BENJAMINI, ORI GUREL-GUREVICH, AND BORIS SOLOMYAK

ABSTRACT. We consider a Branching Random Walk on $\mathbb{R}$ whose step size decreases by a fixed factor, $0 < \lambda < 1$, with each turn. This process generates a random probability measure on $\mathbb{R}$, that is, the limit of uniform distribution among the 2^n particles of the n-th step. We present an initial investigation of the limit measure and its support. We show, in particular, that (1) for almost every $\lambda > 1/2$ the limit measure is almost surely (a.s.) absolutely continuous with respect to the Lebesgue measure, but for Pisot $1/\lambda$ it is a.s. singular; (2) for all $\lambda > (\sqrt{5} - 1)/2$ the support of the measure is a.s. the closure of its interior; (3) for Pisot $1/\lambda$ the support of the measure is "fractured": it is a.s. disconnected and the components of the complement are not isolated on both sides.

1. INTRODUCTION AND STATEMENT OF RESULTS

A *Branching Random Walk (BRW)* on X is a random map from the complete infinite binary tree, $T = \{1, 2\}^*$ into X. We shall consider a symmetric BRW on $\mathbb{R}$ with exponentially decreasing steps, defined as follows. Start with a single particle at 0. At each step each particle multiplies to two particles, and each independently takes a step of size λ^{n-1} to either direction with equal probabilities. Another equivalent useful formulation is this: for each vertex v of the binary tree let a_v be equal to $+1$ or -1 with probability $\frac{1}{2}$ independently. Using these lotteries we define the BRW function to be $f(v) = \sum_{n=1}^{|v|} a_{v|n} \lambda^n$ where $v|n = v_1 \ldots v_n$. This function is extended to the boundary of the tree, $\partial T = \{1, 2\}^{\mathbb{N}}$, in the obvious manner, $f(v) = \sum_{n=1}^{\infty} a_{v|n} \lambda^n$. Let $m = (\frac{1}{2}, \frac{1}{2})^{\mathbb{N}}$ be the standard uniform measure on ∂T. Define μ to be the image of m under f, i.e. $\mu(E) = m(f^{-1}(E))$. This is a random measure on the line, which depends on the choice of signs on the tree. We

are also interested in the properties of the compact support S of μ which is clearly the image of f: $S = f(\partial T)$.

One can view μ as a stochastically self-similar measure and S as a stochastically self-similar set with respect to the appropriate transformations. More precisely, let $F_1(x) = \lambda x + \lambda$ and $F_2(x) = \lambda x - \lambda$, and let

$$\mathcal{F} := \frac{1}{4}\Big(\delta_{(F_1,F_1)} + \delta_{(F_1,F_2)} + \delta_{(F_2,F_1)} + \delta_{(F_2,F_2)}\Big)$$

be a distribution on the pairs of similitudes in $\mathbb{R}$. Then

$$\mu \stackrel{d}{=} \sum_{i=1}^{2} \frac{1}{2}\left(\mu^{(i)} \circ \mathcal{F}_i^{-1}\right),$$

where $(\mathcal{F}_1, \mathcal{F}_2)$ is a random vector of similitudes distributed according to $\mathcal{F}$, and $\mu^{(i)}$, $i = 1, 2$, are i.i.d. copies of μ independent of $(\mathcal{F}_1, \mathcal{F}_2)$. The symbol $\stackrel{d}{=}$ denotes equality in distribution.

There is a large literature on stochastically self-similar sets and measures: Falconer [12], Graf [15], Mauldin and Williams [22] investigated random fractal sets, and U. Zähle [35], Patzshke and U. Zähle [25], Arbeiter [1, 2], Olsen [24], and Hutchinson and Rüschendorf [18] developed the theory of random fractal measures. Existence, uniqueness, and convergence results have been established under very general assumptions, but results on dimension were obtained mostly under some separation ("non-overlapping") conditions. Our case is inherently overlapping for every $\lambda > 0$, since there is a positive probability of having $\mathcal{F}_1 = \mathcal{F}_2$. Overlapping is also allowed in [1, Prop. 6.4], where the translation parts of the similitudes have i.i.d. absolutely continuous distributions with a bounded density, and dimension formulas are obtained which hold a.s. In the recent work by T. Jordan, M. Pollicott, and K. Simon [19], stochastically self-affine sets and measures (with overlaps) are studied, which of course, includes stochastically self-similar ones as a special case. A.s. formulas for the dimension and a.s. absolute continuity are established there under appropriate assumptions. Both in [19] and in the earlier work [28], where a different "overlapping" random model was investigated, the distributions of the vectors of similitudes are absolutely continuous. In our model, on the other hand, these distributions are discrete. This puts it closer to the infinite Bernoulli convolution measures, extensively studied since the 1930's (see [27]).

Of course, our model can be generalized in many ways: instead of the binary tree one can consider an arbitrary rooted tree, instead of the random variables with values ± 1 one can take more general discrete random variables, and one can

consider projections of different measures on the boundary of the tree. Other possible generalizations are mentioned in Section 5.

In Section 2 we adopt an "intermediate-general" viewpoint: $T = \{1, \ldots, \ell\}^*$ is the ℓ-regular tree for $\ell \geq 2$. (On a deterministic non-regular tree we loose stochastic self-similarity; although some results extend to that case, we don't consider it here.) The random variables $\{a_v\}_{v \in T}$ at the vertices are i.i.d. with a discrete distribution $\eta = \sum_{d \in D} p_d \delta_d$. Here $D \subset \mathbb{R}$ is a finite set, which will be called a set of "digits." First we show the "pure types law": for a fixed λ, the measure μ is either absolutely continuous (a.c.) with respect to the Lebesgue measure, or purely singular, almost surely. This is a simple consequence of uniqueness. Then we adapt the approach of Bluhm [6] to obtain estimates of the expectation of some quantities which involve $|\widehat{\mu}(t)|^2$ in terms of the corresponding quantities for the associated deterministic self-similar measure. More precisely, consider the probability measure ν, which is the unique solution of the equation

$$(1.1) \qquad \nu = \sum_{d \in D} p_d (\nu \circ F_d^{-1}), \quad \text{where } F_d(x) = \lambda(x + d),$$

see [17]. It is easy to see that ν is the distribution of the random sum $\sum_{n=1}^{\infty} b_n \lambda^n$ where the coefficients b_n are i.i.d. with the distribution η. It follows that the Fourier transform of ν may be computed as follows:

$$(1.2) \qquad \widehat{\nu}(t) = \prod_{n=1}^{\infty} \widehat{\eta}(t\lambda^n).$$

The classical Bernoulli convolution arises this way if we take $D = \{0, 1\}$ and $p_0 = p_1 = \frac{1}{2}$ (or any other two distinct digits).

Theorem 1.1. *Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the ℓ-regular tree distributed as η. Let ν be the deterministic self-similar measure given by (1.1). Then*

  (i) $\mathbb{E} |\widehat{\mu}(t)|^2 \geq |\widehat{\nu}(t)|^2$ *for all $t \in \mathbb{R}$;*

  (ii) *For any $\gamma \geq 0$, with $\lambda^{1+2\gamma} > \frac{1}{\ell}$, there exist constants C_1 and C_2 such that*

$$(1.3) \qquad \mathbb{E} \left(\int_{\mathbb{R}} |\widehat{\mu}(t)|^2 |t|^{2\gamma} \, dt \right) \leq C_1 + C_2 \int_{\mathbb{R}} |\widehat{\nu}(t)|^2 |t|^{2\gamma} \, dt.$$

Following [26], we use homogeneous Sobolev norms

$$\|\nu\|_{2,\gamma}^2 = \int_{\mathbb{R}} |\widehat{\nu}(t)|^2 |t|^{2\gamma} \, dt.$$

Finiteness of $\|\nu\|_{2,\gamma}$ for $\gamma > 0$ means that ν has γ (fractional) derivatives in L^2.

Corollary 1.2. *Let μ and ν be as in Theorem 1.1.*

(i) *If $\lim_{|t|\to\infty} |\widehat{\nu}(t)| \neq 0$, then μ is a.s. singular.*

(ii) *If $\mathbb{E} \int_{\mathbb{R}} |\widehat{\mu}(t)|^2 \, dt < \infty$, then ν is absolutely continuous with a density in $L^2(\mathbb{R})$.*

(iii) *If ν is absolutely continuous with a density in $L^2(\mathbb{R})$ and $\lambda > \frac{1}{\ell}$, then μ is a.s. absolutely continuous with a density in $L^2(\mathbb{R})$.*

(iv) *If $\|\nu\|_{2,\gamma} < \infty$ and $\lambda^{1+2\gamma} > \frac{1}{\ell}$, then $\|\mu\|_{2,\gamma} < \infty$ a.s.*

Using the results available for deterministic self-similar measures and Bernoulli convolutions with overlaps (see [8, 9, 34, 27, 33] and references therein), we obtain a lot of information on the random measure μ. In particular, we have the following

Corollary 1.3. *Suppose that T is ℓ-regular for $\ell \geq 2$ and the random variables at the vertices are i.i.d. with the distribution $\eta = \frac{1}{m}(\delta_{d_1} + \cdots + \delta_{d_m})$ for $m \geq 2$, where the digits d_j have uniform spacing: $d_j = d_1 + (j-1)h$ for some $h > 0$. Let μ be the corresponding BRW with steps of size λ^n. Then*

(i) *μ is a.s. singular for $\lambda = 1/\theta$, where $\theta < m$ is a Pisot number.*

(ii) *μ is a.s. absolutely continuous for a.e. $\lambda \in (\max\{\frac{1}{\ell}, \frac{1}{m}\}, 1)$;*

Recall that a Pisot number is an algebraic integer $\theta > 1$ whose conjugates (i.e. other zeros of the minimal polynomial) are strictly less than one in absolute value.

Remarks. 1. Looking at the n-th level, we see that S is covered by $\min\{\ell^n, m^n\}$ intervals of size $\sim \lambda^n$. Thus, if $\lambda < \max\{\frac{1}{\ell}, \frac{1}{m}\}$, then S has Hausdorff dimension less than one and hence μ is singular (surely, not just almost surely).

2. Corollary 1.2 shows that the a.s. properties of the stochastically self-similar measure μ and those of its deterministic counterpart are closely related. There is a heuristic principle that putting more randomness into the model increases the likelihood of absolute continuity. In our model the randomness is fairly "mild," so that the number-theoretic phenomena associated with Pisot numbers are preserved (unlike the models in [28, 19]).

3. Corollary 1.2(ii) opens a possibility of applications in the other direction; although this may be far-fetched, any progress in the problem of determining precisely for which λ the Bernoulli convolution measure is absolutely continuous (see [27]), would be interesting.

Finally, we state a corollary which gives additional information for the most basic case $m = 2$ and $\ell = 2$, using the results available for classical Bernoulli convolutions.

Motivated by [13], we say that $\theta > 1$ is a *Garsia number* if it is an algebraic integer whose minimal polynomial has all zeros greater than one in absolute value and the constant term ± 2. Examples of such polynomials include $x^n - 2$ for $n \geq 2$, $x^{n+p} - x^n - 2$ for $p, n \geq 1$ and $\max\{p, n\} \geq 2$, $x^3 - 2x - 2$, etc., see [13].

We write dim to denote the Hausdorff dimension.

Corollary 1.4. *Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$.*

(i) *Suppose $\theta = 1/\lambda \in (1, 2)$ is a Garsia number. Then μ is a.c. with a density in L^2 almost surely.*

(ii) *There exist $a_k < 1$, $a_k \to 1$, such that μ is a.c. with a k times differentiable density almost surely, for a.e. $\lambda \in (a_k, 1)$.*

(iii) *There exists $C > 0$ such that for all $\varepsilon > 0$ we have $\dim\{\lambda \in (\frac{1}{2} + \varepsilon, 1) : \mu$ is a.s. singular $\} < 1 - C\varepsilon$.*

It should be noted that the estimate in (iii) is the best known, but probably not the best possible. The reciprocals of Pisot numbers are the only known parameters in $(\frac{1}{2}, 1)$ for which the Bernoulli convolution measure is singular.

In Sections 3 and 4 we study the topological properties of the support, restricting ourselves to the case of the binary tree and uniform Bernoulli (2-digit) random variables, as in the beginning of the Introduction. Here the case of classical Bernoulli convolutions cannot serve as a guide, since for them the support of the measure is an interval whenever $\lambda \geq \frac{1}{2}$.

It is more convenient to use the digits $0, 1$ rather than ± 1 (this is obtained by a linear change of variables). Then all elements of $S = \mathrm{supp}(\mu)$ are of the form $\sum_{n=1}^{\infty} a_n \lambda^n$ for $a_n \in \{0, 1\}$. Let $I := [0, \frac{\lambda}{1-\lambda}]$ and note that $S \subset I$. Also note that if all infinite words in $\{0, 1\}^{\mathbb{N}}$ can be "read off" the tree T from the root, then $S = I$. This is due to $\lambda > \frac{1}{2}$ and the fact that every $x \in I$ has an expansion in base λ with digits 0,1. The expansion is, in general, non-unique, so the condition for $S = I$ is only sufficient.

For $x \in I$ consider the set of all infinite words giving an expansion of x in base λ:

$$\mathcal{E}_\lambda(x) = \left\{ a \in \{0, 1\}^{\mathbb{N}} : \ x = \sum_{n=1}^{\infty} a_n \lambda^n \right\}.$$

Questions about the size of $\mathcal{E}_\lambda(x)$ have been studied, see Erdős, Joó and Komornik [10], Glendinning and Sidorov [14], and references therein. In particular, in [10]

it is proved that for all $\lambda > g := \frac{\sqrt{5}-1}{2}$ and $x \in (0, \frac{\lambda}{1-\lambda})$ the set $\mathcal{E}_\lambda(x)$ has the cardinality of continuum, and its Hausdorff dimension in the natural metric on ∂T is positive. On the other hand, for $\lambda < g$ there are $x \in (0, \frac{\lambda}{1-\lambda})$ having a unique expansion. More precisely, let $\Psi_\lambda := \{x \in (0, \frac{\lambda}{1-\lambda}) : \#\mathcal{E}_\lambda(x) = 1\}$. In [14] it is proved, in particular, that Ψ_λ is countably infinite for $\lambda \in (\beta, g)$ and is uncountable for $\lambda \in (1/2, \beta]$, where $\beta \approx 0.559852...$ is the "Komornik-Loreti constant" [20].

Returning to our problem, we note that $x \in S$ if and only if there exists $a \in \mathcal{E}_\lambda(x)$ which can be "read off" from the root of T. The questions about "hitting" a given subset of the sequence space by infinite words seen along the paths of a tree for a tree-indexed process were considered by many authors. A set is called *nonpolar* or *polar* according to whether or not it is hit with positive probability. Thus, $x \in S$ with positive probability if and only if $\mathcal{E}_\lambda(x)$ is nonpolar, and by a result of Evans [11, Th. 2], this is equivalent to $\mathcal{E}_\lambda(x)$ having positive logarithmic capacity in the standard metric on $\{0,1\}^\mathbb{N}$. In particular, singletons are polar. (This is also easy to see directly, since for a given word $a \in \{0,1\}^\mathbb{N}$, the random subtree consisting of the edges along which we see the beginning of a, starting from the root, is a Galton-Watson process with offspring distribution $\frac{1}{4}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{4}\delta_2$. It is a critical branching process which dies out a.s.) It follows that for any $x \in \Psi_\lambda$, and more generally, for any x having at most countable many expansions in base λ with digits 0 and 1, almost surely $x \notin S$. We summarize this discussion in the following proposition.

Proposition 1.5. *Let $\lambda \in (\frac{1}{2}, g]$. Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$. Let $I = [0, \frac{\lambda}{1-\lambda}]$ and suppose that the set of $x \in I$ for which $\mathcal{E}_\lambda(x)$ has zero logarithmic capacity (e.g. if it is countable) is dense in I. Then S is totally disconnected.*

Corollary 1.6. *If $\lambda = g$, then S is totally disconnected.*

On the other hand, if $\mathcal{E}_\lambda(x)$ has positive dimension, then $x \in S$ with positive probability. Using the methods of Benjamini and Kesten [4], who investigated when *all* words can be seen from finitely many vertices, we obtain the following (see Section 3):

Theorem 1.7. *Suppose that $\lambda \in (g, 1)$, where $g = \frac{\sqrt{5}-1}{2}$ is the golden ratio. Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary*

tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$. Then $S = \text{supp}(\mu)$ has nonempty interior, and is the closure of its interior, a.s.

The following easy statement is included for completeness. Below $|S|$ denotes the Lebesgue measure of S.

Proposition 1.8. *Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$. If $\lambda \leq \frac{1}{2}$, then $|S| = 0$ a.s.*

It is also standard that for $\lambda = \frac{1}{2}$ the set S has Hausdorff dimension equal to one. This can be deduced directly or from [11] (which implies that S hits any subset of I of positive dimension with positive probability), combined with [21], or, alternatively, from [16].

Open questions. For which $\lambda \in (\frac{1}{2}, g)$ does the set S contain intervals almost surely? For which $\lambda \in (\frac{1}{2}, g)$ is S totally disconnected a.s.? What are the a.s. values of Hausdorff dimension and Lebesgue measure of S for the golden ratio case $\lambda = g$?

We remark that there are many open problems concerning the interior of self-similar sets, both deterministic (see [30]) and random (see [28]).

In Section 4 we turn to the question of connectedness of the support S. It is obvious that S has a positive probability of being disconnected for any tree and any distribution η at the vertices. But is S disconnected almost surely? For some trees this is not the case. For instance, if T is a 4-regular tree and $\eta = \frac{1}{2}(\delta_0 + \delta_1)$, then there is a positive probability of seeing all the words in $\{0,1\}^{\mathbb{N}}$ from the root [4, Cor. 6.3], hence a positive probability of having $S = [0, \frac{\lambda}{1-\lambda}]$ for all λ. In the case of the binary tree it seems that S should be a.s. disconnected, but we could only prove it in a special case.

Theorem 1.9. *Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$. If $\lambda = 1/\theta$, where θ is a Pisot number, then $S = \text{supp}(\mu)$ is disconnected and has no isolated gaps, almost surely. More precisely, let (α, β) be a component of $\mathbb{R} \setminus S$. If $\alpha > -\infty$, then $(\alpha - \varepsilon, \alpha) \not\subset S$ for every $\varepsilon > 0$, and if $\beta < \infty$, then $(\beta, \beta + \varepsilon) \not\subset S$ for every $\varepsilon > 0$.*

We remark that there are infinitely many λ's for which both Theorem 1.7 and Theorem 1.9 apply, see [5]. These are $\lambda \in (g, 1)$ for which $1/\lambda$ is Pisot. For the corresponding random measure μ we know that (i) μ is singular a.s.; (ii) the support of μ is a.s. the closure of a countable union of intervals, with the property that every

gap is accumulated by gaps on both sides. Such sets were called *M-cantorvals* by Mendes and Oliveira [23]; they often appear as arithmetic sums of (deterministic) Cantor sets.

Open questions. For which $\lambda > 1/2$ is the set S a.s. disconnected? If S is disconnected with infinitely many gaps, what can be said about the distribution of lengths of the complementary intervals?

Figure. We show the approximate "densities" of the random measure μ for different simulations and for several interesting values of λ. For comparison, we include pictures of the corresponding (approximations to) Bernoulli convolutions. To be precise, we generated 2^{20} points using a simulation of the BRW, subdivided the interval $[0, \frac{\lambda}{1-\lambda}]$ into $2^{10} = 1024$ intervals of equal size, and plotted the corresponding histogram (this explains the numbers on the axes). The parameters $\lambda = .565198\ldots$ and $\lambda = 2^{-1/2}$ are reciprocals of Garsia numbers, so by Corollary 1.4(i), the random measure μ is a.c. with a density in L^2 almost surely. The parameter $\lambda = .618034\ldots$ is the reciprocal of the golden ratio, a Pisot number, and $\lambda = .754877\ldots$ is the reciprocal of the smallest Pisot number. Thus, the corresponding random measures are singular and have fractured support, even totally disconnected in the case of the golden ratio, almost surely. (These phenomena, however, are not visible in our figures, since the gaps are likely to be very small for larger values of λ.)
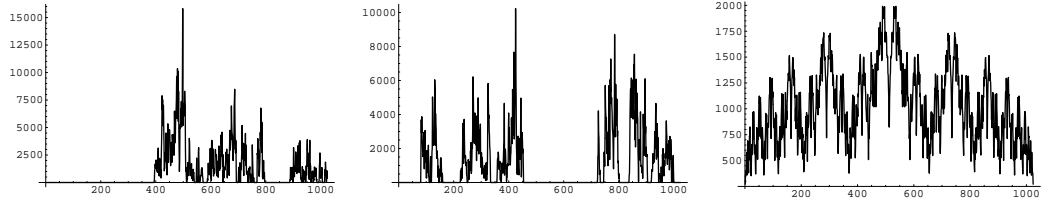
We emphasize again that there is a positive probability of having disconnected support of the random measure μ for all $\lambda < 1$, and this is a major difference with Bernoulli convolutions.
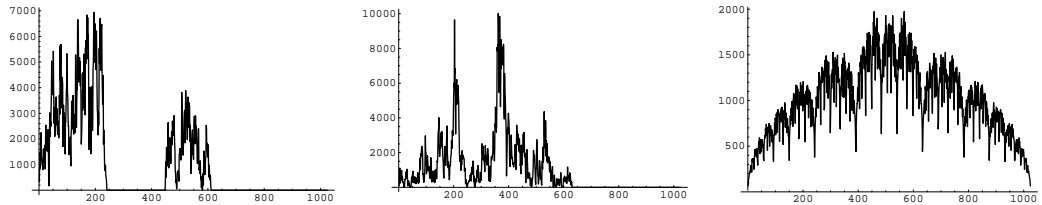
2. Properties of the measure

Here we prove Theorem 1.1 and the corollaries. Recall that $T = \{1, \ldots, \ell\}^*$ is the regular ℓ-ary tree, with $\ell \geq 2$. Let $\eta = \sum_{d \in D} p_d \delta_d$ be a probability distribution on a finite set D and suppose that $\{a_v\}_{v \in T}$ are i.i.d. with the distribution of η. Fix $\lambda \in (0, 1)$ and let μ be the random measure on $\mathbb{R}$ arising from the BRW with steps of size λ^n at time n and the behavior of particles governed by η.

We can also define μ using approximating measures at level n. Let $\mathcal{A} = \{1, \ldots, \ell\}$. Define a random measure μ_n by
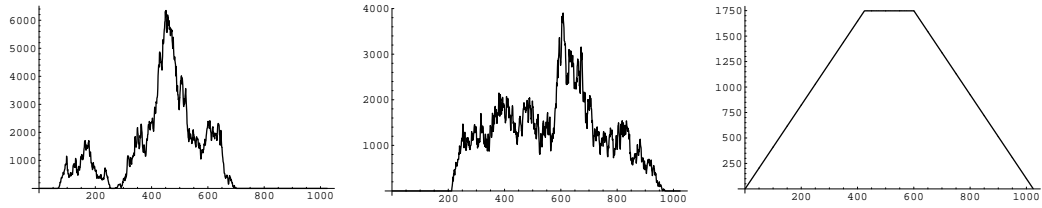
$$(2.1) \qquad \mu_n := \sum_{v \in \mathcal{A}^n} \frac{1}{\ell^n} \delta_{f(v)} \quad \text{where} \quad f(v) := \sum_{j=1}^{n} a_{v|j} \lambda^j.$$
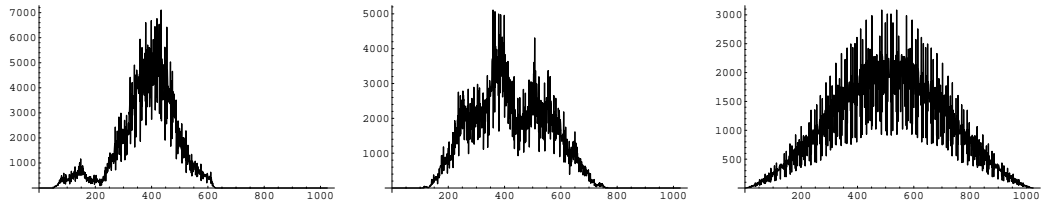
$$\lambda = .565198\ldots$$

$$\lambda = .618034\ldots$$

$$\lambda = 2^{-1/2}$$

$$\lambda = .754877\ldots$$

simulation (a)   simulation (b)   Bernoulli convolution

Theorem 2.1 (Arbeiter [2]). *The measures μ_n converge weakly to a random probability measure μ almost surely.*

Now consider the space $\mathrm{Sim}(\mathbb{R})$ of contracting similitudes on $\mathbb{R}$. Let $\phi(z) = \lambda x + z$ be a map from $\mathbb{R}$ to $\mathrm{Sim}(\mathbb{R})$. Now define a probability measure $\Phi := \otimes_{i=1}^{\ell}(\eta_i \circ \phi^{-1})$ on the Borel σ-algebra of $\mathrm{Sim}(\mathbb{R})^{\ell}$, where η_i are independent copies of η. The

measure Φ is the distribution of a random vector $(\mathcal{F}_1, \ldots, \mathcal{F}_\ell)$ of similitudes on $\mathbb{R}$ with contraction ratio λ and translation vectors distributed according to η.

Theorem 2.2 (Arbeiter [2]). *The random measure μ is stochastically self-similar with respect to Φ. More precisely,*

$$(2.2) \qquad \mu \overset{d}{=} \sum_{i=1}^{\ell} \frac{1}{\ell} \left(\mu^{(i)} \circ \mathcal{F}_i^{-1} \right),$$

where $(\mathcal{F}_1, \ldots, \mathcal{F}_\ell)$ is a random vector of similitudes distributed according to Φ, and $\mu^{(i)}$, $i = 1, \ldots, \ell$ are i.i.d. copies of μ independent of $(\mathcal{F}_1, \ldots, \mathcal{F}_\ell)$.

Next we derive the "pure types law" for our measure, which is analogous to the classical deterministic case.

Proposition 2.3. *The measure μ is either a.s. absolutely continuous, or a.s. pure singular with respect to the Lebesgue measure.*

Proof. Iterating (2.2) we obtain for any $n \geq 1$:

$$\mu \overset{d}{=} \sum_{|v|=n} \frac{1}{\ell^n} \left(\mu^{(v)} \circ \mathcal{F}_v^{-1} \right),$$

where $\mathcal{F}_v$ are some random non-degenerate similitudes $\mathcal{F}_v$ (all having the same contraction rate λ^n) and $\mu^{(v)}$, $|v| = n$, are i.i.d. copies of μ independent of what happens in our tree in the levels $|v| < n$. Consider μ_{sing}, the singular part of μ. Taking the singular part of the last equation yields that $\mu_{\mathrm{sing}}(\mathbb{R})$ is independent of the outcome of any finite number of the random variables on the tree. Thus, $\{\mu_{\mathrm{sing}}(\mathbb{R}) > c\}$ is a 0-1 event for any $c > 0$, hence there exists c such that $\mu_{\mathrm{sing}}(\mathbb{R}) = c$ almost surely. If $c = 0$, then μ is a.s. absolutely continuous. Otherwise, $(1/c)\mu_{\mathrm{sing}}$ is a stochastically self-similar probability measure satisfying the equation (2.2). By the Uniqueness Theorem (see [1] or [18, Th. 3.1]), it equals μ in distribution, so μ is a.s. singular. $\qquad\square$

Proof of Theorem 1.1. We follow Bluhm [6] closely, up to a point. Fix $n \in \mathbb{N}$, then by (2.1),

$$\begin{aligned} \mathbb{E}\, |\widehat{\mu_n}(t)|^2 &= \mathbb{E} \left| \sum_{v \in \mathcal{A}^n} \ell^{-n} \exp(it f(v)) \right|^2 \\ &= \sum_{v, w \in \mathcal{A}^n} \ell^{-2n} \mathbb{E}\, \exp(it(f(v) - f(w))). \end{aligned}$$

Observe that

$$f(v) - f(w) = \sum_{j=k+1}^{n} \lambda^j (a_{v|j} - a_{w|j}),$$

where $k = |v \wedge w|$ and $v \wedge w$ is the longest common initial segment of v and w. Since $a_{v|j}, a_{w|j}$ are i.i.d. with the distribution of η for $j = k+1, \ldots, n$, we obtain

$$
\begin{aligned}
\mathbb{E} \exp(it(f(v) - f(w))) &= \prod_{j=k+1}^{n} \mathbb{E} \exp(it\lambda^j a_{v|j}) \, \mathbb{E} \exp(-it\lambda^j a_{w|j}) \\
&= \prod_{j=k+1}^{n} \widehat{a_{v|j}}(t\lambda^j) \widehat{a_{w|j}}(-t\lambda^j) \\
&= \prod_{j=k+1}^{n} |\widehat{\eta}(t\lambda^j)|^2,
\end{aligned}
$$

by the definition of the Fourier transform. Therefore,

$$
\begin{aligned}
\mathbb{E} |\widehat{\mu_n}(t)|^2 &= \ell^{-2n} \sum_{v,w \in \mathcal{A}^n} \prod_{j=|v \wedge w|+1}^{n} |\widehat{\eta}(t\lambda^j)|^2 \\
&= \ell^{-2n} \sum_{v \in \mathcal{A}^n} \sum_{k=0}^{n} \sum_{w:|v \wedge w|=k} \prod_{j=k+1}^{n} |\widehat{\eta}(t\lambda^j)|^2 \\
(2.3) \qquad &= \ell^{-n} \Big(\sum_{k=0}^{n-1} (\ell-1)\ell^{n-k-1} \prod_{j=k+1}^{n} |\widehat{\eta}(t\lambda^j)|^2 + 1 \Big).
\end{aligned}
$$

In the last line we used that $\#\{w \in \mathcal{A}^n : |v \wedge w| = k\} = (\ell-1)\ell^{n-k-1}$ for $k = 0, \ldots, n-1$, and $|v \wedge w| = n$ only for $v = w$. Since $|\widehat{\eta}(t\lambda^j)| \leq 1$, we obtain from (2.3) and (1.2):

$$
\begin{aligned}
\mathbb{E} |\widehat{\mu_n}(t)|^2 &\geq \ell^{-n} \Big((\ell-1) \sum_{k=0}^{n-1} \ell^{n-k-1} \prod_{j=1}^{n} |\widehat{\eta}(t\lambda^j)|^2 + 1 \Big) \\
&\geq \prod_{j=1}^{n} |\widehat{\eta}(t\lambda^j)|^2 \geq |\widehat{\nu}(t)|^2.
\end{aligned}
$$

By Theorem 2.1, for every t we have $|\widehat{\mu}(t)|^2 = \lim_{n \to \infty} |\widehat{\mu_n}(t)|^2$ hence $\mathbb{E} |\widehat{\mu}(t)|^2 = \lim_{n \to \infty} \mathbb{E} |\widehat{\mu_n}(t)|^2$ since the Fourier transforms are bounded by 1 and the expectation is over a finite measure. This proves part (i) of the theorem.

(ii) Denote $\theta := \lambda^{-1}$. Let s be a natural number, $1 \leq s \leq n-1$, suppose that $t \in [\theta^{s-1}, \theta^s]$, and estimate (2.3) from above:

$$
\begin{aligned}
\mathbb{E}\, |\widehat{\mu_n}(t)|^2 \;\; &\leq \\
&\leq \quad \ell^{-n}\Big(\sum_{k=0}^{s-1}(\ell-1)\ell^{n-k-1}\prod_{j=k+1}^{n}|\widehat{\eta}(t\lambda^j)|^2 + (\ell-1)\sum_{k=s}^{n-1}\ell^{n-k-1}+1\Big) \\
&= \quad (\ell-1)\sum_{k=0}^{s-1}\ell^{-k-1}\prod_{j=k+1}^{n}|\widehat{\eta}(t\lambda^j)|^2 + \ell^{-s}.
\end{aligned}
$$

Letting $n \to \infty$ and using (1.2) we obtain

(2.4) $$\mathbb{E}\, |\widehat{\mu}(t)|^2 \leq \ell^{-s} + (\ell-1)\sum_{k=0}^{s-1}\ell^{-k-1}|\widehat{\nu}(t\lambda^k)|^2.$$

Next we multiply by $t^{2\gamma}$ and integrate over $[\theta^{s-1}, \theta^s]$, keeping in mind that $\lambda = \theta^{-1}$ and

$$\int_{\theta^{s-1}}^{\theta^s}|\widehat{\nu}(t\lambda^k)|^2 t^{2\gamma}\, dt = \theta^{k(1+2\gamma)}\int_{\theta^{s-k-1}}^{\theta^{s-k}}|\widehat{\nu}(t)|^2 t^{2\gamma}\, dt$$

to obtain

$$
\begin{aligned}
\int_{\theta^{s-1}}^{\theta^s}\mathbb{E}\, |\widehat{\mu}(t)|^2 t^{2\gamma}\, dt \;\; &\leq \quad \frac{\theta^{s(1+2\gamma)}-\theta^{(s-1)(1+2\gamma)}}{\ell^s(1+2\gamma)} \\
&\quad + \quad (\ell-1)\sum_{k=0}^{s-1}\frac{\theta^{k(1+2\gamma)}}{\ell^{k+1}}\int_{\theta^{s-k-1}}^{\theta^{s-k}}|\widehat{\nu}(t)|^2 t^{2\gamma}\, dt.
\end{aligned}
$$

Summing over $s = 1, 2, \ldots$, using that $\theta^{1+2\gamma} < \ell$, and exchanging the order of summation yields for some C_1, C_2:

$$\int_1^\infty \mathbb{E}\, |\widehat{\mu}(t)|^2 t^{2\gamma}\, dt \leq C_1 + C_2\int_1^\infty |\widehat{\nu}(t)|^2 t^{2\gamma}\, dt,$$

and this implies (1.3). □

Proof of Corollary 1.2. (i) If μ is not almost surely singular, then it is almost surely absolutely continuous by Proposition 2.3, hence by the Riemann-Lebesgue Lemma we have $\lim_{|t|\to\infty}|\widehat{\mu}(t)|^2 = 0$ almost surely. Since $|\widehat{\mu}(t)| \leq 1$ for all t, this being a probability measure, by the Lebesgue Dominated Convergence Theorem we have $\lim_{|t|\to\infty}\mathbb{E}\,|\widehat{\mu}(t)|^2 = 0$, contradicting the assumption on ν in view of Theorem 1.1(i).

(ii) By Plancherel's Theorem, ν has a density in L^2 if and only if $\int_{\mathbb{R}}|\widehat{\nu}(t)|^2\, dt < \infty$. Now the claim follows from Theorem 1.1(i).

(iii) and (iv) are immediate from Theorem 1.1(ii) and Fubini Theorem. □

Proof of Corollary 1.3. (i) It is well-known that $\lim_{|t|\to\infty} |\widehat{\nu}(t)| \neq 0$ when $1/\lambda > m$ is Pisot. This is due to Erdős [8] for $m = 2$, and the proof easily extends to arbitrary m (see [7]). Now the claim follows from Corollary 1.2(i).

(ii) By the result of Simon and Tóth [33], which extended [34] to the case $m > 2$, the self-similar measure ν is a.c. with a density in L^2 for a.e. $\lambda \in (\frac{1}{m}, 1)$. Now the claim follows from Corollary 1.2(iii). $\square$

Proof of Corollary 1.4. (i) follows from [13] and Corollary 1.2(iii).

(ii) follows from [26, Lem. 5] (see also [9] for the classical, but less sharp result) and Corollary 1.2(iv).

(iii) follows from [26, Section 5]. $\square$

3. Intervals in the support

In this section we prove Theorem 1.7. Here we return to our most basic set-up: BRW with two equally likely digits on the binary tree. As already mentioned, it is convenient to use the digits 0,1, so we have i.i.d. random variables distributed as $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$, at each vertex of the tree.

First we introduce some notation. For $a = a_1 a_2 \ldots$, a finite or infinite word in the alphabet $\{0, 1\}$, denote

$$\xi(a) = \xi(a_1 a_2 \ldots) = \sum_{j=1}^{\infty} a_j \lambda^j \stackrel{\lambda}{=} .a_1 a_2 \ldots$$

Recall that $g = \frac{\sqrt{5}-1}{2}$, so that $1 = g + g^2 = \sum_{j=2}^{\infty} g^j$. Since $\lambda > g$, there exists $\ell \geq 3$ such that

(3.1) $1 < \lambda^2 + \cdots + \lambda^\ell.$

We fix ℓ for the rest of the proof. Let

$$\mathcal{U} := (\alpha, \beta), \quad \text{where} \quad \alpha = \frac{\lambda^\ell}{1 - \lambda^\ell}, \quad \beta = \frac{\lambda}{1 - \lambda} - \frac{\lambda^\ell}{1 - \lambda^\ell}.$$

Note that

$$\alpha \stackrel{\lambda}{=} .(0^{\ell-1}1)^\infty, \quad \beta \stackrel{\lambda}{=} .(1^{\ell-1}0)^\infty.$$

For $a \in \{0, 1\}^n$, let

$$\mathcal{U}_a := \xi(a) + \lambda^n \mathcal{U}.$$

Lemma 3.1. *There exists $c > 0$ such that for any subinterval $J \subset \mathcal{U}$, with $|J| \leq c$, there exist $a, a' \in \{0, 1\}^{\ell-1}$ satisfying*

$$J \subset \mathcal{U}_{a0} \cap \mathcal{U}_{a'1}.$$

Proof. The idea comes from [10, Th. 3] which shows that given any $x \in \mathcal{U}$, we can fix an arbitrary sequence $\{a_{kl}\}_{k \geq 1} \in \{0, 1\}^\infty$ and obtain an expansion $x \overset{\lambda}{=} .a_1 a_2 \ldots$ with the digits 0,1 by the "greedy algorithm." It is enough to show that

$$(3.2) \qquad \mathcal{U} \subset \Big(\bigcup_{a \in \{0,1\}^{\ell-1}} \mathcal{U}_{a0} \Big) \cap \Big(\bigcup_{a \in \{0,1\}^{\ell-1}} \mathcal{U}_{a1} \Big).$$

Indeed, since the intervals are open, it will follow that there exists a positive c as desired.

Let $a \in \{0, 1\}^{\ell-1}$ be non-maximal, that is, $a \neq 1^{\ell-1}$. We claim that there exists $a' \in \{0, 1\}^{\ell-1}$ such that $\xi(a) < \xi(a')$ and

$$(3.3) \qquad \mathcal{U}_{aj} \cap \mathcal{U}_{a'j} \neq \emptyset, \quad \text{for } j = 0, 1.$$

This is equivalent to showing that

$$(3.4) \qquad \xi(a') - \xi(a) < |\mathcal{U}_{aj}| = \lambda^{-\ell} |\mathcal{U}|.$$

If a ends with 0, consider a' which ends with 1, but otherwise agrees with a. Then $\xi(a') - \xi(a) = \lambda^{\ell-1}$. If a ends with 1, then a ends with 01^p for some $1 \leq p \leq \ell - 2$ since a is non-maximal. Consider the greedy expansion of 1 in base λ:

$$1 = d_1 \lambda + d_2 \lambda^2 + \cdots$$

It has the property that $d_j \in \{0, 1\}$ for all j and $1 - \sum_{j=1}^n d_j \lambda^j < \lambda^n$ for all $n \geq 1$. Let

$$w = 1, (1 - d_1), \ldots, (1 - d_p)$$

and consider $a' \in \{0, 1\}^{\ell-1}$ which ends with w, but otherwise agrees with a. Then

$$
\begin{aligned}
\xi(a') - \xi(a) &= \lambda^{(\ell-1)-(p+1)}(\xi(w) - \xi(1^p)) \\
&= \lambda^{\ell-p-1}\Big(1 - \sum_{j=1}^p d_j \lambda^j\Big) \in (0, \lambda^{\ell-1}).
\end{aligned}
$$

In both cases we obtain that $\xi(a') - \xi(a) \leq \lambda^{\ell-1}$. Note that $1 - \lambda^\ell < \lambda^2 + \cdots + \lambda^\ell - \lambda^{\ell+1}$ by (3.1), hence

$$\lambda^{-1} < \frac{\lambda + \cdots + \lambda^{\ell-1} - \lambda^\ell}{1 - \lambda^\ell} = \frac{\lambda}{1 - \lambda} - \frac{2\lambda^\ell}{1 - \lambda^\ell} = |\mathcal{U}| = \lambda^{-\ell}|\mathcal{U}_{aj}|.$$

Therefore,

$$\xi(a') - \xi(a) \leq \lambda^{\ell-1} < |\mathcal{U}_{aj}|,$$

which verifies (3.4) and hence (3.3). It remains to note that α (the left endpoint of $\mathcal{U}$) is the left endpoint of $\mathcal{U}_{0^{\ell-1}1}$ and is covered by $\mathcal{U}_{0^\ell}$; similarly, β is the right endpoint of $\mathcal{U}_{1^{\ell-1}0}$ and is covered by $\mathcal{U}_{1^\ell}$. This, together with (3.3), implies (3.2), and the lemma is proved. $\qquad\square$

Corollary 3.2. *There exists a constant*

$$(3.5) \qquad\qquad 0 < c < \lambda^\ell |\mathcal{U}|/4$$

such that for any $n \geq 1$ and any $a \in \{0,1\}^{(n-1)\ell}$, if $J \subset \mathcal{U}_a$, with $|J| \leq c\lambda^{n\ell}$, then there exist $a', a'' \in \{0,1\}^{\ell-1}$ satisfying

$$J \subset \mathcal{U}_{aa'0} \cap \mathcal{U}_{aa''1}.$$

Proof. For $n = 1$ this is just Lemma 3.1 (of course, we can always impose an upper bound on c). For $n > 1$ this follows from Lemma 3.1 and the definitions by rescaling, with the same c. $\qquad\square$

Proof of Theorem 1.7. We are going to use a variant of the argument from [4]. In order to set it up, we need to consider another family of intervals, which we denote J_w and which should not be confused with the intervals $\mathcal{U}_a$. Fix $p \in \mathbb{N}$ such that

$$(3.6) \qquad\qquad p^{-1} < \min\{c(1-\lambda)/\lambda, \lambda^\ell\}$$

where c is from Lemma 3.1. We subdivide the interval $[0, \frac{\lambda}{1-\lambda}]$ into p closed subintervals of equal length and denote them J_i, $i = 0, \ldots, p-1$. Then continue the subdivision and denote the intervals of level n by J_w for $w \in \{0, \ldots, p-1\}^n$. We have $|J_w| = \frac{\lambda}{1-\lambda}p^{-|w|}$, so in view of (3.6) and (3.5),

$$(3.7) \qquad |J_w| < c\lambda^{\ell(n-1)} < |\mathcal{U}_a|/4 \quad \text{for } |w| = n \text{ and } |a| = n\ell.$$

For a vertex σ of the binary tree denote by $a(\sigma)$ the (random) finite word of 0's and 1's which we see on the path from the root to σ. Fix ε such that

$$(3.8) \qquad\qquad 0 < \varepsilon < 2\lambda - 1 \quad \text{and} \quad (1+\varepsilon)^\ell < 2.$$

Now we follow the argument of [4, pp. 1046-47] very closely. For each $n \geq 1$ and $w \in \{0, \ldots, p-1\}^n$ define the event

$$
\begin{aligned}
H_n(w) \quad &:= \quad \{\text{there are at least } (1+\varepsilon)^{n\ell} \text{ vertices} \\
&\qquad \sigma, \text{ with } |\sigma| = n\ell, \text{ such that } J_w \subset \mathcal{U}_{a(\sigma)}\}.
\end{aligned}
$$

(3.9)

We are going to show that

$$
(3.10) \qquad \sum_{n=1}^{\infty} \sum_{w_1 \ldots w_{n+1}} \mathbb{P}\left\{H_n(w_1 \ldots w_n) \setminus H_{n+1}(w_1 \ldots w_{n+1})\right\} < \infty
$$

where the second sum is over all sequences $w \in \{0, \ldots, p-1\}^{n+1}$. First we explain how this implies the desired result. By Borel-Cantelli, (3.10) implies that almost surely for all n sufficiently large ($n \geq N$ where N is random), $H_n(w_1 \ldots w_n) \setminus H_{n+1}(w_1 \ldots w_{n+1})$ does not occur. Let us fix a random configuration (the choice of 0's and 1's), so that N is now fixed. We write $\sim$ to indicate equality up to a (multiplicative) positive constant independent of n. We have $2^{n\ell}$ intervals $\mathcal{U}_{a(\sigma)}$ of level $n\ell$, each of length $\sim \lambda^{n\ell}$, whose union is contained in $[0, \frac{\lambda}{1-\lambda}]$. Note that many of them will likely coincide; they are counted with multiplicity. By the pigeon-hole principle, there are at least $\sim 2^{n\ell}\lambda^{n\ell}$ intervals $\mathcal{U}_{a(\sigma)}$ of level $n\ell$ with a common intersection longer than $|\mathcal{U}_{a(\sigma)}|/2$. For large n the number of these intervals exceeds $(1+\varepsilon)^n$ since $2\lambda > 1+\varepsilon$, hence there is an interval J_w, with $|w| = n$, contained in their intersection by (3.7). Thus, the event $H_n(w)$ occurs, and we can assume that $n \geq N$. By the choice of N, the events $H_j(ww')$, with $j = |ww'|$, occur for all finite extensions of the word w. This implies that a.s.

$$
J_w \subset \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} \bigcup_{|\sigma|=j} \mathcal{U}_{a(\sigma)}.
$$

It is easy to see that the right-hand side of the last formula is contained in S: it consists of points which can be approximated by $\xi(a(\sigma))$ for $|\sigma|$ arbitrarily large, and S is compact. Thus, S contains an interval a.s. The same argument, of course, implies that a.s. there is an interval in the "cylinder" of S obtained by taking a subtree from any given vertex, and since such cylinders are dense in S, it follows that S is the closure of its interior.

It remains to verify (3.10). We continue to follow the scheme of [4, Section 5]. Assume that $H_n(w)$ occurred, so that there exist at least $(1+\varepsilon)^{n\ell}$ vertices $\sigma_1, \ldots, \sigma_r$ at the level $n\ell$ such that $J_w \subset \mathcal{U}_{a(\sigma_j)}$, $1 \leq j \leq r$. Let $w' = ww_{n+1}$. We have $J_{w'} \subset J_w$ and $|J_{w'}| < c\lambda^{n\ell}$. We can apply Corollary 3.2 for $J = J_{w'}$ and each

$\mathcal{U}_{a(\sigma_j)}$ to conclude that there exist $a'_j, a''_j \in \{0,1\}^{\ell-1}$ satisfying

$$J_{w'} \subset \mathcal{U}_{a(\sigma_j)a'_j 0} \cap \mathcal{U}_{a(\sigma_j)a''_j 1}.$$

Conditionally on $H_n(w)$, the following random variables are independent of each other for $j = 1, \ldots, r$:

$$U_n(\sigma_j) \quad := \quad \{\text{number of descendants } \tau \text{ of } \sigma_j \text{ in the } (n+1)\ell\text{-th level}$$
$$\text{with the property that the word seen on the path from}$$
$$\sigma_j \text{ to } \tau \text{ is either } a'_j 0 \text{ or } a''_j 1\}.$$

Moreover,

$$U_n(\sigma_j) \geq 0 \quad \text{and} \quad \mathbb{E}\{U_n(\sigma_j) \mid H_n(w), \sigma_1, \ldots, \sigma_r\} = 2.$$

Indeed, the expected number of times to read off a specific word W in the tree of descendants of any vertex from level $n\ell$ down to level $(n+1)\ell$ equals the number of paths times the probability of seeing W on a given path. This gives $2^\ell \cdot 2^{-\ell} = 1$, (recall that we are in the unbiased case), but since two words are good for us, we get 2. Note that $U_n(\sigma_j) \leq 2^\ell$, so these random variables are uniformly bounded, and one easily checks that the variance satisfies

$$0 < C_1 \leq \mathrm{Var}(U_n(\sigma_j)) \leq C_2 < \infty$$

for some C_1, C_2 independent of n and j. Therefore, by the Large Deviation Estimate (see e.g. [31, Section 7.4]), in view of $(1+\varepsilon)^\ell < 2$, there exists a constant $C_3 = C_3(\varepsilon)$ such that

$$\mathbb{P}\left\{\sum_{j=1}^{r} U_n(\sigma_j) \leq r(1+\varepsilon)^\ell \mid H_n(w), \sigma_1, \ldots, \sigma_r\right\} \leq 2\exp(-C_3 r).$$

However, if $H_n(w)$ occurs, and hence $r \geq (1+\varepsilon)^{n\ell}$, and $\sum_{j=1}^{r} U_n(\sigma_j) > r(1+\varepsilon)^\ell \geq (1+\varepsilon)^{(n+1)\ell}$, then there are at least $(1+\varepsilon)^{(n+1)\ell}$ vertices τ in the $(n+1)\ell$-th level such that $J_{w'} \subset \mathcal{U}_{a(\tau)}$, that is, $H_{n+1}(w')$ occurs. This implies

$$\mathbb{P}\{H_{n+1}(w') \text{ fails} \mid H_n(w)\} \leq 2\exp(-C_3(1+\varepsilon)^{n\ell}),$$

and (3.10) follows. $\qquad\square$

We do not know whether the set S contains an interval a.s. for $\lambda \in (\frac{1}{2}, g)$. However, if we can find an interval which is covered with multiplicity at least two by its images on certain level, then the proof above goes through without any changes. We state this precisely for future reference.

Proposition 3.3. *Let $\lambda \in (\frac{1}{2}, 1)$ and suppose that there exists a nonempty open interval $\mathcal{U}$ and $\ell \in \mathbb{N}$ satisfying the condition (3.2). Let μ be the BRW with steps of size λ^n, with the i.i.d. random variables on the binary tree distributed as $\frac{1}{2}(\delta_0 + \delta_1)$. Then $S = \mathrm{supp}(\mu)$ has nonempty interior, and is the closure of its interior, a.s.*

Proof of Proposition 1.5. As explained in the Introduction, if $\mathcal{E}_\lambda(x)$ has zero logarithmic capacity, then $x \notin S$ a.s. We can choose a countable dense set in I with this property, and the claim follows. $\qquad\square$

Proof of Corollary 1.6. It is known (easy to see, folklore, see e.g. [32, App. A]) that for $\lambda = g$ every point $x \equiv ng^{-1} \pmod 1$ has countably many expansions in base g. Since this is a countable dense set, Proposition 1.5 applies. $\qquad\square$

Now we turn to the proof of Proposition 1.8. Let $I = [0, \frac{\lambda}{1-\lambda}]$ and recall that $S \subset I$. For $v \in T$, let $I(v) = f(v) + \lambda^{|v|}I$. We have

$$(3.11) \qquad S = \bigcap_{n=1}^{\infty} S(n), \quad \text{where} \quad S(n) := \bigcup_{|v|=n} I(v).$$

The sets $S(n)$ form a decreasing nested family.

Proof of Proposition 1.8. Let K_n be the number of distinct words of length n which we see from the root, that is, $K_n = \#\{a(v) : |v| = n\}$. The probability of seeing any given word is $\sim n^{-1}$, since this is the probability of survival of a critical branching process (see e.g. [3, Th. I.9.1]). Summing over all possible words we obtain $\mathbb{E}(K_n) \sim n^{-1}2^n$. It follows that $\mathbb{E}|S(n)| \leq C\lambda^n \mathbb{E}(K_n) \to 0$, as $n \to \infty$, for $\lambda \leq \frac{1}{2}$. Clearly, $\mathbb{E}|S| = \lim_{n\to\infty} \mathbb{E}|S(n)| = 0$, hence $|S| = 0$ almost surely. $\qquad\square$

4. The support is fractured

Here we prove Theorem 1.9. We will use the following fact about Pisot numbers.

Lemma 4.1 (Garsia [13]). *For every $\lambda = 1/\theta$, where θ is Pisot, there is a constant $0 < c_1 < 1$ such that if $a_1, .., a_n$ and $b_1, .., b_n$ are sequences of $0, 1$ and $\sum_{i=1}^{n} a_i\lambda^i \neq \sum_{i=1}^{n} b_i\lambda^i$, then $|\sum_{i=1}^{n} a_i\lambda^i - \sum_{i=1}^{n} b_i\lambda^i| \geq c_1\lambda^n$.*

Proof of Theorem 1.9. We denote by $\mathbf{a} \in \{0,1\}^T$ the outcome of the lotteries on the tree, that is, $\mathbf{a} = \{a_v : v \in T\}$. We also write T_n for the set of vertices at level n and use similar notation for subgraphs of T.

Let q be a rational number, and let $\mathcal{P}_q$ be the event that $q \notin S$. It is enough to prove that the endpoints of the component of $\mathbb{R} \setminus S$ containing q have the desired property a.s., conditioned on $\mathcal{P}_q$. We will work with the left endpoints, since the right endpoints are treated exactly the same way. So fix q, assume $\mathcal{P}_q$ holds, and let (α, β) be the component of $\mathbb{R} \setminus S$ containing q.

Let n_0 be the smallest integer such that $q \notin S(n_0)$. By (3.11), there exists such an integer. Note that conditioned on $\mathcal{P}_q$ and on the value of n_0, the distribution of $\mathbf{a}$ for levels greater then n_0 is still the same product measure. This is because the event $q \notin S(n_0)$ depends only on the first n_0 levels of $\mathbf{a}$, and $\mathcal{P}_q$ is simply the union of the corresponding cylinder sets.

For any $n \geq n_0$, let
$$L^{(n)} = \{v \in T_n : f(v) < q\},$$
and let $m^{(n)} = \max\{f(v) : v \in L^{(n)}\}$. Let
$$\mathcal{M}^{(n)} = \{v \in L^{(n)} : f(v) = m^{(n)}\}$$
be the set of vertices where this maximum is achieved. Members of $\mathcal{M}^{(n)}$ are called the *maximal vertices of level n*. Of course, $\mathcal{M}^{(n)}$ is a random subset of T_n.

For $v \in T$ consider the subtree
$$T_v^+ = T_v^+(\mathbf{a}) = \{vw : \ w \in \{1,2\}^*, \ a_{vw_1} = a_{vw_1w_2} = \cdots = a_{vw} = 1\}.$$
In other words, T_v^+ is the cluster of 1's from the vertex v down (we do not make any assumptions on a_v). The number of vertices of T_v^+ at level n is a critical Galton-Watson process since the number of offsprings of a vertex labeled by 1 has the distribution $\frac{1}{4}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{4}\delta_2$, with expectation 1. Such a process dies out with probability 1, so T_v^+ is a.s. finite, for all $v \in T$.

Fix $n \geq n_0$, consider the graph
$$\Gamma^{(n)} = \bigcup \{T_v^+ : \ v \in \mathcal{M}^{(n)}\},$$
and let
$$\tau^{(n)} = \sup\{j : \ \Gamma_j^{(n)} \neq \emptyset\}.$$
We have $\tau^{(n)} < \infty$ a.s. Fix $\ell \in \mathbb{N}$ such that
$$(4.1) \qquad \frac{\lambda^{\ell+1}}{1-\lambda} < c_1,$$
where c_1 is from Lemma 4.1. Consider the event
$$\mathcal{E}^{(n)} := \left\{\tau^{(n)} < \infty \ \& \ |\Gamma_{\tau^{(n)}-j}^{(n)}| = 2^{\ell-j}, \ j = 0, \ldots, \ell\right\}.$$

In other words, $\mathcal{E}^{(n)}$ occurs whenever $\Gamma^{(n)}_{\tau^{(n)}-\ell} = \{\sigma\}$ for some vertex σ and all its $2^{\ell+1} - 2$ descendants down to the level $\tau^{(n)}$ are labeled by 1's, after which $\Gamma^{(n)}$ dies out.

Lemma 4.2. $\exists\, c_2 > 0,\ \forall\, n \in \mathbb{N},\ \mathbb{P}\,(\mathcal{E}^{(n)}) \geq c_2.$

Proof. This is written in measure-theoretic language. Denote $\Omega = \{0,1\}^T$ and recall that $\mathbb{P} = (\frac{1}{2}, \frac{1}{2})^T$, so that $(\Omega, \mathbb{P})$ is the probability space for our lotteries on the tree. We show the dependence on $\mathbf{a}$ in our notation. Let $\Omega_0 = \{\mathbf{a} \in \Omega :\ \tau^{(n)}(\mathbf{a}) < \infty\}$. The fact that all T_v^+, for $v \in \mathcal{M}^{(n)}$, die out a.s. means that $\mathbb{P}\,(\Omega_0) = 1$.

Consider the following transformation on Ω_0. Let $\mathbf{a} \in \Omega_0$ and choose w to be the rightmost (i.e. greatest in the lexicographic order) vertex of $\Gamma^{(n)}$ in the level $\tau^{(n)}(\mathbf{a})$, that is, just before it dying out. By the definition of $\tau^{(n)}(\mathbf{a})$ we have $a_{w1} = a_{w2} = 0$. Then $\mathbf{a}' \in \Omega_0$ is defined as follows:

$$a'_{w1u} = 1 \quad \forall\, u \in \{1,2\}^*, \quad 0 \leq |u| \leq \ell;$$
$$a'_{w1u} = 0 \quad \forall\, u \in \{1,2\}^{\ell+1}.$$

For all other vertices v we let $a'_v = a_v$.

Observe that $\tau^{(n)}(\mathbf{a}') = \tau^{(n)}(\mathbf{a}) + \ell + 1$, and $\mathbf{a}' \in \mathcal{E}^{(n)}$ by construction (note that $\sigma = w1$ is the vertex indicated after the definition of $\mathcal{E}^{(n)}$). Let $\Omega' = \{\mathbf{a}' :\ \mathbf{a} \in \Omega_0\}$. It is enough to prove that

$$(4.2) \hspace{4cm} \mathbb{P}\,(\Omega') \geq c_2 > 0$$

for some constant c_2 which does not depend on n. This follows from the definition of $\mathbb{P}$ as the product measure on Ω and the fact that given $\mathbf{a}'$ we can recover $\mathbf{a}$ except for the descendants of $\sigma 1$ in the levels $\tau^{(n)}(\mathbf{a}) + 1, \ldots, \tau^{(n)}(\mathbf{a}) + \ell + 2$. Thus, we can take $c_2 = 2^{-2^{\ell+2}}$. $\hfill\square$

Lemma 4.3. *If $\mathcal{E}^{(n)}$ occurs, then $(\alpha - \frac{\lambda^n}{1-\lambda}, \alpha) \not\subset S$.*

Proof. To simplify notation, we write $\Gamma = \Gamma^{(n)}$ and $\tau = \tau^{(n)}$. Suppose that $\mathcal{E}^{(n)}$ occurred, so that $\Gamma_{\tau-\ell} = \{\sigma\}$ for some vertex σ, and all its $2^{\ell+1} - 2$ descendants down to level τ are labeled by 1's, after which Γ dies out.

Recall that Γ is the cluster of 1's starting from the maximal vertices at level n. Thus, for $j \leq \tau^{(n)}$ we have $\Gamma_j = \mathcal{M}^{(j)}$. It follows that σ is the only maximal vertex at the level $\tau - \ell$. For any u, with $|u| \geq \ell$,

$$(4.3) \hspace{1cm} f(\sigma u) \geq f(\sigma) + \lambda^{\tau-\ell+1} + \cdots + \lambda^\tau = f(\sigma) + \lambda^{\tau-\ell+1} \frac{1 - \lambda^\ell}{1 - \lambda}.$$

Now suppose $v \in T_{\tau-\ell}$, $v \neq \sigma$. Then v is not maximal, hence for any $u \in \{1, 2\}^*$, if $f(vu) < q$ then

$$(4.4) \qquad f(vu) \leq f(v) + \frac{\lambda^{\tau-\ell+1}}{1-\lambda} \leq f(\sigma) - c_1 \lambda^{\tau-\ell} + \frac{\lambda^{\tau-\ell+1}}{1-\lambda}$$

by Lemma 4.1. Observe that the right-hand side of (4.4) is less than the right-hand side of (4.3) by (4.1), hence we have a gap.

Since σ is a maximal vertex at the level $\tau - \ell$, we have

$$(4.5) \qquad \alpha = \max(S \cap (-\infty, q)) \in \left[f(\sigma) + \lambda^{\tau-\ell+1} \frac{1-\lambda^\ell}{1-\lambda}, f(\sigma) + \lambda^{\tau-\ell+1} \frac{1}{1-\lambda} \right].$$

In view of (4.3), we have $[A - \varepsilon, A] \not\subset S$ for $A = f(\sigma) + \lambda^{\tau-\ell+1} \frac{1-\lambda^\ell}{1-\lambda}$ and $\varepsilon > 0$. Now (4.5) and $\tau \geq n$ imply the desired claim $(\alpha - \frac{\lambda^n}{1-\lambda}, \alpha) \not\subset S$. $\qquad \square$

Conclusion of the proof. Since the critical branching process dies out a.s., we can find a sequence $n_k \uparrow \infty$ such that $n_{k+1} - n_k > \ell$ and with probability greater than $1 - \delta$, for some fixed δ, for all k and all vertices $\sigma \in T_{n_k}$, the random subgraph T_σ^+ dies out before level n_{k+1}. Let $\mathcal{G}$ be the event that this happens (so that $\mathbb{P}(\mathcal{G}) > 1 - \delta$). Conditioned on this event, the events $\mathcal{E}^{(n_k)}$, considered above, are independent. In view of Lemma 4.2, infinitely many of the events $\mathcal{E}^{(n_k)}$ occur a.s., conditioned on $\mathcal{G}$. By Lemma 4.3, this implies that $(\alpha - \varepsilon, \alpha) \not\subset S$ for every $\varepsilon > 0$. Since $\delta > 0$ is arbitrary, the claim of the theorem follows. $\qquad \square$

5. Further Directions

Here we discuss some directions for further research and mention some open questions (in addition to those listed in Section 1).

1. The signs on the tree are taken with different probabilities, e.g. $(p, 1 - p)$ on the binary tree. Biased Bernoulli convolutions exhibit "multifractal" behavior, and the thresholds for absolute continuity and L^q density no longer coincide, see [29]. We can expect a similar phenomenon for our random model. Combined with the results of [29], Corollary 1.2 yields sufficient conditions for the a.s. existence of L^2 density, at least for some values of p. Entropy considerations imply that for all $p \neq \frac{1}{2}$, for some $\lambda > \frac{1}{2}$ the measure μ is a.s. supported on a set of zero Lebesgue measure. The methods of Sections 3 and 4 partially extend. For instance, if $p > \frac{1}{2}$, then for $1/\lambda$ Pisot every gap of the support has gaps accumulating to it from the right.

2. The following model was suggested by Károly Simon: fix p and suppose that the random variables on the vertices are independent and have either the distribution $p\delta_1 + (1-p)\delta_{-1}$ or $(1-p)\delta_1 + p\delta_{-1}$, depending on whether the last edge goes left or right. This can be considered as a p-perturbation of the Bernoulli convolution measure, since for $p = 0$ we get the Bernoulli convolution (there is no randomness). For $p = \frac{1}{2}$ we get the model studied in this paper. What happens for $0 < p < \frac{1}{2}$, especially as $p \to 0$?

3. Other graphs (1): instead of the regular rooted tree, consider more general graphs, e.g. $1/4$ of the grid $\mathbb{Z}^2$. There are 2^n paths of length n, but there is less randomness, since there are only $\sim n^2$ edges (put the signs on edges).

4. Other graphs (2): consider the same problem on a random tree, e.g. on a Galton-Watson tree (choosing a random tree is part of the model). This way we recover stochastic self-similarity (in some sense).

5. Other graphs (3): consider the same problem on a random or deterministic graph of polynomial growth. Instead of the steps λ^n consider slowly shrinking steps. E.g. spherically symmetric trees of growth r^d, or critical GW tree conditioned to survive and steps at level n equal to $\pm n^{-1/2}$. Here there is no corresponding Bernoulli convolution.

6. The random variables a_v on the tree are i.i.d. with some a.c. distribution, e.g. Gaussian. (Note that we first choose these steps, fix them, and then consider the BRW with steps $a_v\lambda^{|v|}$, that is, the contraction ratio is deterministic.) On the binary tree, there will be 2^n intervals of size $\sim \lambda^n$, so for $\lambda < 1/2$ we still get singularity. For $\lambda > 1/2$ we should have a.c. almost surely; the "Pisot numbers effect" will be lost. This should follow easily by the methods of [28, 19]. We do not know what to expect for connectedness properties of the support.

Remark added on May 30, 2007. Russ Lyons indicated to us a simpler proof of Corollary 1.2(i) and Corollary 1.3(i). In fact, the expectation of the random measure μ in Theorem 1.1 over all the "lotteries" on the tree is easily seen to be the self-similar measure ν defined by (1.1). Therefore, if ν is singular, then μ is a.s. singular, by Fubini Theorem.

Professor in the Autumn of 2005. He is grateful to the Department of Mathematics at WIS for its hospitality.

References

[1] M. Arbeiter, Random recursive construction of self-similar fractal measures. The noncompact case, *Probab. Th. Rel. Fields* **88** (1991), 497–520.

[2] M. Arbeiter, Construction of random fractal measures by branching processes, *Stochastics and Stochastics Reports* **39** (1992), 195–212.

[3] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer, New York, 1972.

[4] I. Benjamini and H. Kesten, Percolation of arbitrary words in $\{0,1\}^{\mathbb{N}}$, *Annals of Prob.* **23** (1995), 1024–1060.

[5] M. J. Bertin, A. Decomps-Guilloux, M. Grandet-Hugot, M. Pathiaux-Delefosse, J. P. Shreiber, *Pisot and Salem Numbers*. Birkhäuser, 1992.

[6] C. Bluhm, Fourier asymptotics of statistically self-similar measures, *J. Fourier Analysis and Appl.* **5** (1999), 355–362.

[7] J. Borwein and R. Girgensohn, Functional equations and distribution functions, *Results in Math.* **26** (1994), 229–237.

[8] P. Erdős, On a family of symmetric Bernoulli convolutions, *Amer. J. Math.* **61** (1939), 974–975.

[9] P. Erdős, On the smoothness properties of Bernoulli convolutions, *Amer. J. Math.* **62** (1940), 180–186.

[10] P. Erdős, I. Jóo and V. Komornik, Characterization of the unique expansions $1 = \sum_{i=1}^{\infty} q^{-n_i}$ and related problems, *Bull. Soc. Math. Fr.* **118** (1990), 377–390.

[11] S. N. Evans, Polar and nonpolar sets for a tree indexed process, *Ann. Prob.* **20** (1992), 579–590.

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

[13] A. M. Garsia, Arithmetic properties of Bernoulli convolutions, *Trans. Amer. Math. Soc.* **102** (1962), 409–432.

[14] P. Glendinning and N. Sidorov, Unique representations of real numbers in non-integer bases, *Math. Res. Letters* **8** (2001), 535–543.

[15] S. Graf, Statistically self-similar fractals, *Probab. Th. Rel. Fields* **74** (1987), 357–392.

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

[17] J. E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), 713–747.

[18] J. E. Hutchinson and L. Rüschendorf, Random fractal measures via the contraction method, *Indiana Univ. Math. J.* **47** (1998), 471–487.

[19] T. Jordan, M. Pollicott and K. Simon, Hausdorff dimension for randomly perturbed self affine attractors, *Comm. Math. Physics* **270** (2007), 519–544.

[20] V. Komornik and P. Loreti, Unique developments in non-integer bases, *Amer. Math. Monthly* **105** (1998), 636–639.

[21] R. Lyons, Random walks and percolation on trees, *Ann. Probab.* **18** (1990), 931–958.

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

[23] P. Mendes and F. Oliveira, On the topological structure of the arithmetic sum of two Cantor sets, *Nonlinearity* **7** (1994), 329–343.

[24] R. Olsen, *Random Geometrically Graph Directed Self-similar Multifractals*, Pittman Research Notes **307**, Longman, 1994.

[25] N. Patzschke and U. Zähle, Self-similar random measures. IV. The recursive construction of Falconer, Graf, and Mauldin and Williams, *Math. Nachr.* **149** (1990), 285–302.

[26] Y. Peres and W. Schlag, Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions, *Duke Math. J.* **102** (2000), no. 2, 193–251.

[27] Y. Peres, W. Schlag and B. Solomyak, Sixty years of Bernoulli convolutions, *Fractal Geometry and Stochastics* **II**, C. Bandt, S. Graf, and M. Zähle (editors), Progress in Probability Vol. **46**, 39–65, Birkhäuser, 2000.

[28] Y. Peres, K. Simon and B. Solomyak, Absolute continuity for random iterated function systems with overlaps, *J. London Math. Soc.* (2) **74** (2006), 739–756.

[29] Y. Peres and B. Solomyak, Self-similar measures and intersections of Cantor sets, *Trans. Amer. Math. Soc.* **350** (1998), 4065–4087.

[30] Y. Peres and B. Solomyak, Problems on self-similar sets and self-affine sets: an update, *Fractal Geometry and Stochastics* **II**, C. Bandt, S. Graf, and M. Zähle (editors), Progress in Probability Vol. **46**, 95–106, Birkhäuser, 2000.

[31] A. Rényi, *Probability Theory*, North-Holland, 1970.

[32] N. Sidorov and A. Vershik, Ergodic properties of the Erdős measure, the entropy of the golden shift, and related problems, *Monatsh. Math.* **126** (1998), 215–261.

[33] K. Simon and H. Tóth, The absolute continuity of the distribution of random sums with digits $\{0, 1, \ldots, m-1\}$, *Real Anal. Exchange* **30** (2004/05), no. 1, 397–409.

[34] B. Solomyak, On the random series $\sum \pm\lambda^i$ (an Erdős problem), *Annals of Math.* **142** (1995), 611–625.

[35] U. Zäehle, Self-similar random measures. I. Notion, carrying Hausdorff dimension, and hyperbolic distribution, *Probab. Th. Rel. Fields* **80** (1988), 79–100.

Boris Solomyak, Box 354350, Department of Mathematics, University of Washington, Seattle WA 98195, USA

E-mail address: solomyak@math.washington.edu

Itai Benjamini and Ori Gurel-Gurevich, Department of Theoretical Mathematics, Weizmann Institute of Science, Rehovot, 76100, Israel