Prob. Th. Rel. Fields 78, 335–355 (1988)

Probability Theory and Related Fields

Superbranching Processes and Projections of Random Cantor Sets

F.M. Dekking[1] and G.R. Grimmett[2]*

[1] Department of Mathematics, Delft University of Technology, Julianalaan 132, 2628 BL Delft, The Netherlands

[2] School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom

Abstract. We study sequences $(X_0, X_1, \ldots)$ of random variables, taking values in the positive integers, which grow faster than branching processes in the sense that $X_{m+n} \geqq \sum_{i=1}^{X_m} X_n(m, i)$, for $m, n \geqq 0$, where the $X_n(m, i)$ are distributed as X_n and have certain properties of independence. We prove that, under appropriate conditions, $X_n^{1/n} \to \lambda$ almost surely and in L^1, where $\lambda = \sup_n E(X_n)^{1/n}$. Our principal application of this result is to study the Lebesgue measure and (Hausdorff) dimension of certain projections of sets in a class of random Cantor sets, being those obtained by repeated random subdivisions of the M-adic subcubes of $[0, 1]^d$. We establish a necessary and sufficient condition for the Lebesgue measure of a projection of such a random set to be non-zero, and determine the box dimension of this projection.

1. Introduction

The generation sizes of a branching process form a sequence $X_0, X_1, \ldots$ of random variables satisfying the distributional relation

$$X_{m+n} \stackrel{d}{=} \sum_{i=1}^{X_m} X_n(i). \tag{1.1}$$

where the $X_n(i)$ are distributed as X_n. Here we think of the $X_n(i)$ as the number of descendants of the i^{th} member of the m^{th} generation which belong to the $(m+n)^{\text{th}}$ generation. The asymptotic behaviour of X_n for large n is well understood. For example, it follows from (1.1) that, if $X_0 = 1$ and $\mu = EX_1 < \infty$, then

* Work done partly whilst visiting Cornell University with the aid of a Fulbright travel grant

$(Y_n, \mathscr{F}_n)$ is a martingale where $Y_n = X_n \mu^{-n}$ and $\mathscr{F}_n$ is the σ-field generated by $X_0, X_1, \ldots, X_n$. Consequently the limit

$$W = \lim_{n \to \infty} X_n \mu^{-n} \tag{1.2}$$

exists almost surely, and it may be shown that

$$P(W=0) = \begin{cases} 1 & \text{if } E(X_1 \log X_1) = \infty, \\ \delta & \text{if } E(X_1 \log X_1) < \infty, \end{cases}$$

where δ is the probability that the process becomes extinct (see, for example, Athreya and Ney (1972)). In particular, we have that, as $n \to \infty$.

$$X_n^{1/n} \to \begin{cases} \mu & \text{on } S, \\ 0 & \text{off } S, \end{cases} \tag{1.3}$$

almost surely, where S is the event that the process does not become extinct.

There are many branching-type processes which fail to satisfy (1.1), but which satisfy instead a "superbranching" inequality of the form

$$X_{m+n} \overset{d}{\geqq} \sum_{i=1}^{X_m} X_{n,m}(i) \quad \text{for } m, n \geqq 0, \tag{1.4}$$

for suitably defined random variables $X_{n,m}(i)$, having the same distribution as X_n for each i. Here $U \overset{d}{\geqq} V$ means $P(U > x) \geqq P(V > x)$ for all real x. A simple example is a branching process with immigration, for appropriately chosen immigration rates; other examples may easily be found, such as in the study of branching random walks (see Biggins (1977, 1979)). In the next section we present a formal definition of a "superbranching" process and give various concrete examples. Of particular interest to us in this paper is a superbranching process arising in the study of random fractal sets, and we shall return to this example later in the introduction.

Our principal theoretical result is an analogue to (1.3) for processes $(X_n : n \geqq 0)$ satisfying a slightly stronger property than (1.4). Specifically we shall prove that for such a process a limit theorem of the form

$$X_n^{1/n} \to \lambda \quad \text{a.s.} \quad \text{on } \{\limsup_{n \to \infty} X_n > 0\} \tag{1.5}$$

is valid, where $\lambda = \sup_n E(X_n)^{1/n}$.

The primary motivation for our interest in such a question lies in an application to labelled branching processes, thence to answer a question concerning the (Hausdorff) dimension of certain projections of randomly-generated Cantor-like sets. We describe this first by an example. Define $C_0 = [0, 1] \times [0, 1]$, the closed unit square in the plane. We think of C_0 as the union of the four closed squares $C(i,j)$ for $i,j=0,1$, where

$$C(i,j)=[\tfrac{1}{2}i, \tfrac{1}{2}(i+1)]\times[\tfrac{1}{2}j, \tfrac{1}{2}(j+1)].$$

Using some specified random mechanism μ, we delete some (or all) of these $C(i,j)$, writing C_1 for the union of those remaining. We now repeat this process of deletion to each of these remaining subsquares: divide each into four squares with side-length $\frac{1}{4}$ in the obvious way and use the random mechanism μ again in deciding which of these smaller squares to delete. Each application of μ is independent of all previous applications and all applications in other squares. We write C_2 for the union of the remaining squares of side-length $\frac{1}{4}$. Continuing in the obvious way, we obtain a decreasing sequence $C_0, C_1, C_2, \dots$ of closed sets with limit $C=\bigcap_{n=0}^{\infty} C_n$. Let γ be the mean number of subsquares of C_0 which remain after the first application of μ. If $\gamma>1$ then $P(S)>0$ where S is the event that $C_n \neq \emptyset$ for all n, or equivalently that C is non-empty. Various authors have shown that the Hausdorff dimension of C is $\log_2 \gamma$ almost surely on S (see Peyrière (1978), Hawkes (1981), Falconer (1986, 1987), Mauldin and Williams (1986), Graf (1987)). Let πC be the projection of C onto the x-axis:

$$\pi C=\{x\in[0,1]:(x,y)\in C \quad \text{for some } y\}.$$

Then πC is a closed subset of $[0,1]$, and we shall show in Theorem 9 how the limit theorem for superbranching processes may be used to compute the box-dimension (also called capacity, logarithmic density, Kolmogorov-Tihomirov dimension) of πC.

The random Cantor sets we consider here are called random curds in Mandelbrot (1983). Our Theorem 10 may be considered as a general quantification of a remark in Mandelbrot (1983, p. 218) on the size of projections of his implementation of Hoyle's galaxy model as a random curd.

All logarithms are to base e unless otherwise indicated.

2. Superbranching Processes

Let $\underline{X}=(X_n\colon n\geqq 0)$ be a sequence of random variables taking values in the non-negative integers. We shall assume throughout that $X_0=1$, but this assumption is not essential.

We say that $\underline{X}$ satisfies the *weak superbranching inequality if*

$$X_{m+n}\overset{d}{\geqq}\sum_{i=1}^{X_m} X_{n,m}(i) \qquad \text{for } m,n\geqq 0, \tag{2.1}$$

where $U\overset{d}{\geqq}V$ means $P(U>x)\geqq P(V>x)$ for all x, and $(X_{n,m}(i)\colon i\geqq 1)$ are independent random variables which are independent of X_m and are distributed in the same manner as X_n.

We say that $\underline{X}$ satisfies the *strong superbranching inequality* if, whenever $m \geqq 0$ and we are given that $X_i = x_i$ for $0 \leqq i \leqq m$, then

$$X_{m+n} \overset{d}{\geqq} \sum_{i=1}^{x_m} X_n(i) \quad \text{for } n \geqq 0, \tag{2.2}$$

where $(X_n(i): i \geqq 1)$ are independent random variables which are distributed as X_n and which are independent of $X_0, X_1, \ldots, X_m$. If $\underline{X}$ satisfies the strong superbranching inequality, then $\underline{X}$ satisfies the weak inequality also.

We are interested in the asymptotic properties of a sequence $\underline{X}$ which satisfies one of the above inequalities. In advance of describing such results, we give some examples.

(i) *Branching process.* The generation numbers of a Galton-Watson branching process satisfy the strong superbranching inequality with equality in (2.2).

(ii) *Branching process with immigration.* Certain schemes for immigration into a branching process do not disturb the strong superbranching inequality. A simple such example is as follows. Let f be a mapping from $\{0, 1, 2, \ldots\}$ into itself satisfying $f(u+v) \geqq f(u) + f(v)$. To each generation of the branching process we add $f(k)$ immigrants where k is the number of natural children of the members of the previous generation.

(iii) *Birth process with superadditive birth rates.* Consider a birth process in continuous time with a single founder at time 0, and assume that the birth rates $\lambda_1, \lambda_2, \ldots$ satisfy $\lambda_u > 0$ and $\lambda_{u+v} \geqq \lambda_u + \lambda_v$ for $u, v \geqq 1$. Writing X_n for the size of the process at time n, we see that $\underline{X} = (X_n: n \geqq 0)$ satisfies the strong superbranching inequality.

(iv) *Branching random walk.* Suppose that particles inhabit the real line and reproduce in the following way. At time 0 there is a single particle at the origin. At time 1 this particle is replaced by a collection of particles distributed about points of $\mathbb{R}$ chosen at random. At time 2, each of these particles is replaced by a collection of particles positioned at points with the same distribution relative to their parent as the position of the children of the unique founder. Each particle reproduces in this way, giving rise to children at points chosen according to the same measure but independent of all other families. Suppose that $0 \leqq \alpha \leqq \beta$ and let X_n be the number of members of the nth generation positioned within the interval $[n\alpha, n\beta]$. It is not difficult to see that $\underline{X} = (X_n: n \geqq 0)$ satisfies the strong superbranching inequality. The example may be generalized to deal with the contents of intervals of the form $[\alpha(n), \beta(n)]$ where $0 \leqq \alpha(n) \leqq \beta(n) \leqq \infty$ for all n, and $(\alpha(n): n \geqq 0)$ is a subadditive sequence and $(\beta(n): n \geqq 0)$ is superadditive.

(v) *Labelled branching process.* Let V be a set of labels, and suppose that we are provided with a branching process with a single founder and non-empty families, in which each member has some family of random size N, the members of which receive randomly chosen labels $L_1, L_2, \ldots, L_N$ from V. We assume that the vector $(N; L_1, L_2, \ldots, L_n)$ is chosen according to some specified probability measure on the appropriate space, independently of all previous family-sizes and label sets. The founder member remains unlabelled. Writing $L(x)$ for the label of member x of the process, we see that with each member x there is associated a sequence $L(x_1), L(x_2), \ldots, L(x)$ of labels, where $\rho, x_1, x_2, \ldots, x$ is the unique path in the family tree of the process joining the root ρ to x. Thinking about V as an alphabet, we call the sequence $L(x_1), L(x_2), \ldots, L(x)$ the *word* $w(x)$ associated with x. If x is in the nth generation of the process then $w(x)$ has length n. Let $X_0=1$ and write X_n for the number of distinct words having length n. Then $\underset{\sim}{X}=(X_n: n \geqq 0)$ satisfies the strong superbranching inequality. To see this, suppose that $X_i=x_i$ for $0 \leqq i \leqq m$, and let $\pi_1, \pi_2, \ldots, \pi_{x_m}$ be members of the mth generation whose corresponding words of length m are distinct. Let $X_n(i)$ be the number of distinct words of length n in the subtree of the process with root at π_i.

It is not difficult to see that

$$X_{m+n} \overset{d}{\geqq} \sum_{i=1}^{x_m} X_n(i)$$

as required.

Labelled branching processes provide a general setting for branching random walks. Suppose for example that $V=\mathbb{R}$, and with the point x in the process we associate the real number $p(x)=\Sigma_i L(x_i)$, the sum of the labels in the word corresponding to x. We may think of $p(x)$ as the position of x in the associated branching random walk.

3. Limit Theorems for Superbranching Processes

We have limit theorems for sequences satisfying both the weak and strong superbranching inequalities. In the weak version of such limit theorems, we shall assume also that the sequences in question cannot become extinct. We shall study sequences $(X_n: n \geqq 0)$ with $X_0=1$, although the assumption that $X_0=1$ is not vital.

Theorem 1. *Let $\underset{\sim}{X}$ satisfy the weak superbranching inequality, and suppose that $P(X_1 \geqq 1)=1$. If $\lambda=\sup_n E(X_n)^{1/n}$ satisfies $\lambda<\infty$ then $\lambda \geqq 1$ and, as $n \to \infty$,*

$$X_n^{1/n} \to \lambda \qquad \text{a.s. and in } L^1. \tag{3.1}$$

In the subsequent parts of this paper we shall make use of the following corollary.

Theorem 2. *Under the hypotheses of Theorem 1, we have that*

$$\frac{1}{n} \log X_n \to \log \lambda \qquad \text{a.s. and in } L^1. \tag{3.2}$$

Our third result concerns the strong inequality.

Theorem 3. *Let $\underset{\sim}{X}$ satisfy the strong superbranching inequality, and suppose that $EX_1>0$. If $\lambda=\sup_n E(X_n)^{1/n}$ satisfies $1<\lambda<\infty$ then, as $n \to \infty$,*

$$X_n^{1/n} \to \lambda \qquad \text{a.s. on } \{\limsup_{n\to\infty} X_n>0\}. \tag{3.3}$$

The last theorem asserts that $X_n^{1/n} \to \lambda$ a.s. on the event $\{X_n \geqq 1$ infinitely often$\}$. It is clear that $P(X_n=0$ for all large $n)=1$ whenever $\underset{\sim}{X}$ is an integer sequence with $\sup_n E(X_n)^{1/n} < 1$. In the proof of Theorem 1 we shall use the following lemma (whose proof will be postponed to the end of this section), the conclusion of which is doubtless well known.

Lemma 4. *Let* $(B(n): n \geqq 0)$ *be the generation sizes of a Galton-Watson branching process with* $B(0)=1$ *and whose family-sizes are non-zero with mean* μ *satisfying* $1<\mu<\infty$. *If* $\alpha<\mu$, *there exists* $\beta(\alpha)<1$ *such that*

$$P(B(n)<\alpha^n)<\beta(\alpha)^n \quad \text{for all } n.$$

Proof of Theorem 1. We write $\mu_n = E(X_n)$, and note from (2.1) that

$$\mu_{m+n} \geqq \mu_m \mu_n \quad \text{for } m, n \geqq 0. \tag{3.4}$$

Since $\mu_1 > 0$, we have that $\mu_n > 0$ for all n. We apply the subadditive limit theorem to (3.4) to find that

$$\mu_n^{1/n} \to \lambda \quad \text{as } n \to \infty \tag{3.5}$$

where $\lambda = \sup_m E(X_m)^{1/m}$. We note that

$$\mu_m \leqq \lambda^m \quad \text{for all } m. \tag{3.6}$$

If $\eta > \lambda$ then

$$P(X_n \geqq \eta^n) \leqq \eta^{-n} E(X_n) \leqq (\lambda/\eta)^n.$$

giving by the Borel-Cantelli lemma that

$$\limsup_{n\to\infty} X_n^{1/n} \leqq \lambda \quad \text{a.s.} \tag{3.7}$$

For the lower bound on $X_n^{1/n}$ we have to work a little harder. We may assume that $\lambda > 1$ since otherwise $\mu_n = 1$ for all n and the result is trivial. For each $N \geqq 1$, let $(B_N(n): n \geqq 0)$ be the generation sizes of a branching process with a single founder and family-sizes distributed as X_N. We have from the weak superbranching inequality that

$$X_{Nk} \overset{d}{\geqq} B_N(k) \quad \text{for all } N, k \geqq 1. \tag{3.8}$$

We fix u such that $1<u<\lambda$, and we choose N such that $E(X_N) > u^N$; this is possible by the definition of λ. Now

$$P(X_{Nk} \leqq u^{Nk}) \leqq P(B_N(k) \leqq u^{Nk}) = P(B_N(k) \leqq \alpha^k) \tag{3.9}$$

where $\alpha = u^N < E(X_N)$. We apply Lemma 4 to deduce that

$$\sum_k P(X_{Nk} \leqq u^{Nk}) < \infty \quad \text{for all } u < \lambda. \tag{3.10}$$

From the assumption that $P(X_1 \geqq 1) = 1$ and the weak superbranching inequalities we have that $X_n \overset{d}{\leqq} X_m$ if $n \leqq m$, giving that

$$\begin{aligned}\sum_n P(X_n \leqq v^n) &= \sum_k \sum_{j=1}^{N} P(X_{Nk+j} \leqq v^{Nk+j}) \\ &\leqq \sum_k \sum_{j=1}^{N} P(X_{Nk+j} \leqq v^{N(k+1)}) \leqq N \sum_k P(X_{Nk} \leqq v^{N(k+1)}) < \infty\end{aligned}$$

whenever $1 \leqq v < \lambda$. By the Borel-Cantelli lemma,

$$\liminf_{n\to\infty} X_n^{1/n} \geqq \lambda \quad \text{a.s.} \tag{3.11}$$

as required.

To prove L^1 convergence, we define

$$Y_n = \inf_{m \geqq n} X_m^{1/m}.$$

Note that $(Y_n: n \geqq 0)$ is monotone with limit given by $Y_n \uparrow \lambda$ a.s. Thus, by monotone convergence,

$$E(Y_n) \uparrow \lambda \quad \text{as } n \to \infty. \tag{3.12}$$

From (3.11), Fatou's lemma, Jensen's inequality, and (3.5),

$$\begin{aligned}\lambda \leqq E(\liminf_{n\to\infty} X_n^{1/n}) &\leqq \liminf_{n\to\infty} E(X_n^{1/n}) \\ &\leqq \limsup_{n\to\infty} E(X_n^{1/n}) \\ &\leqq \limsup_{n\to\infty} E(X)^{1/n} = \lambda,\end{aligned}$$

giving that

$$E(X_n^{1/n}) \to \lambda \quad \text{as } n \to \infty. \tag{3.13}$$

Thus

$$\begin{aligned}E|X_n^{1/n} - \lambda| &\leqq E|X_n^{1/n} - Y_n| + E|Y_n - \lambda| \\ &= E(X_n^{1/n}) - E(Y_n) + \lambda - E(Y_n) \\ &\to 0 \quad \text{as } n \to \infty\end{aligned}$$

by (3.12) and (3.13). This completes the proof. □

Proof of Theorem 2. The almost sure convergence is immediate from Theorem 1. Convergence in L^1 is almost immediate. We write $Z_n = X_n^{1/n}$. If $\lambda = 1$ then $P(X_n = 1) = 1$ for all n and the result is trivial. Suppose therefore that $\lambda > 1$. On the interval $[\lambda^{-1}, \infty)$ it is the case that

$$|\log x| \leqq \frac{\lambda \log \lambda}{\lambda - 1} |x - 1|.$$

Thus, as $n \to \infty$

$$\begin{aligned}E\left|\frac{1}{n}\log X_n - \log \lambda\right| &= E|\log(Z_n/\lambda)| \\ &\leqq \frac{\log\lambda}{\lambda - 1} E|Z_n - \lambda| \to 0. \quad \square\end{aligned}$$

Proof of Theorem 3. Suppose now that $\underline{X}$ satisfies the strong superbranching inequality with $1 < \lambda < \infty$. Then $\underline{X}$ satisfies the weak inequality also, so that the proof of Theorem 1 gives us that $\mu_n = E(X_n)$ satisfies $\mu_n^{1/n} \to \lambda$ as $n \to \infty$, and $\mu_n \leqq \lambda^n$ for all n. Also

$$\limsup_{n\to\infty} X_n^{1/n} \leqq \lambda \qquad \text{a.s.} \tag{3.14}$$

as in (3.7). It remains to prove that

$$P(\liminf_{n\to\infty} X_n^{1/n} \geqq \lambda, A) = P(A) \tag{3.15}$$

where $A = \{X_n \geqq 1 \text{ infinitely often}\}$. We prove this in a sequence of lemmas.

Lemma 5. *Let L be a positive integer and define $T_L = \min\{n: X_n \geqq L\}$. Then $P(T_L < \infty, A) = P(A)$.*

Proof. Let N be a positive integer such that $EX_N > 1$, and let $(B_N(n): n \geqq 0)$ be the generation sizes of a Galton-Watson branching process with $B_N(0) = 1$ and family-sizes distributed as X_N. The process $\underline{B}_N$ is supercritical, and thus there exists k such that

$$\eta = P(B_N(k) < L) < 1; \tag{3.16}$$

we choose k accordingly and define η by (3.16).

Now, we may construct an increasing random sequence $(J_i: i \geqq 1)$ of integers as follows. We set $J_1 = \min\{m: X_m \geqq 1\}$ and $J_{i+1} = \min\{m: m > J_i + Nk \text{ and } X_m \geqq 1\}$. We note that J_i is a stopping time for $\underline{X}$, for all i. It can happen that $J_i = \infty$ for some i, in which case $J_j = \infty$ for all $j \geqq i$; this is impossible on A. We write $A_I = \{J_I < \infty\}$ for I a positive integer and we note that

$$A = \lim_{I\to\infty} A_I.$$

Thus

$$P(X_n < L \text{ for all } n, A) \leqq P(X_{J_i+Nk} < L \text{ for all } i, A) \tag{3.17}$$
$$= \lim_{I\to\infty} \pi(I)$$

where

$$\pi(I) = P(X_{J_i+Nk} < L \text{ for } 1 \leqq i \leqq I, A_I).$$

However,

$$\pi(I) = \sum_{j=0}^{\infty} P(X_{J_i+Nk} < L \text{ for } 1 \leqq i \leqq I, A_I, J_I = j)$$

$$= \sum_{j=0}^{\infty} P(X_{j+Nk} < L, C_I(j))$$

where

$$C_I(j) = \{X_{J_i+Nk} < L \text{ for } 1 \leqq i < I, A_{I-1}, J_I = j\}$$

is an event which depends only on $X_0, X_1, \ldots, X_j$. On the event $\{J_I=j\}$ we have that $X_j \geqq 1$, and we may use the superbranching inequality and (3.16) to deduce that

$$P(X_{j+Nk}<L,\, C_I(j)) \leqq \eta P(C_I(j)).$$

Hence

$$\pi(I) \leqq \eta \sum_{j=0}^{\infty} P(C_I(j)) \leqq \eta \pi(I-1). \tag{3.18}$$

Hence $\pi(I) \to 0$ as $I \to \infty$, giving by (3.17) that Lemma 5 is true. □

Lemma 6. *Let $(B_1(n)\colon n \geqq 0)$ be the generation sizes of a Galton-Watson process with family-sizes distributed as X_1. Suppose that $X_i = x_i$ for $0 \leqq i \leqq m$. Then $(X_n\colon n \geqq m)$ is a collection of random variables which are jointly no smaller in distribution than the distribution of $(B_1(n)\colon n \geqq m)$ conditional on $B_1(m) = x_m$.*

Proof. This is a consequence of the strong superbranching inequality. It is easy to see by induction on k that

$$\begin{aligned} &P(X_{m+i} \geqq y_{m+i} \quad \text{for } 0<i\leqq k \mid X_j = x_j \text{ for } 0 \leqq j \leqq m) \\ &\geqq P(B_1(m+i) \geqq y_{m+i} \quad \text{for } 0<i\leqq k \mid B_1(m) = x_m) \end{aligned}$$

for all k. □

Lemma 7. *Suppose $\mu_1 = E(X_1) > 1$. Then*

$$P(\liminf_{n\to\infty} X_n^{1/n} \geqq \mu_1,\, A) = P(A). \tag{3.19}$$

Proof. This is the principal calculation. Let $(B_1(n)\colon n \geqq 0)$ be as given in Lemma 6 with $B_1(0)=1$. Choose v such that $1<v<\mu_1$ and note that, for each i,

$$P(B_1(n) < v^{n+i} \text{ infinitely often}) \leqq \eta \tag{3.20}$$

where η is the probability that the branching process $\underline{B}_1$ becomes extinct; we have that $\eta<1$ since $\mu_1>1$.

Let L be a positive integer, and define $T_L = \min\{n\colon X_n \geqq L\}$ as in Lemma 5. Then

$$\begin{aligned} &P(X_n \geqq v^n \quad \text{for all large } n,\, T_L<\infty) \\ &\quad \geqq P(X_n \geqq v^n \quad \text{for all large } n,\, T_L \leqq k) \\ &\quad \geqq \sum_{i=0}^{k} P(T_L = i)\, P(B_1(n) \geqq v^n \quad \text{for all large } n \mid B_1(i) \geqq L) \end{aligned}$$

by Lemma 6, the strong superbranching inequality, and the fact that $X_i \geqq L$ if $T_L = i$. Thus

$$\begin{aligned} &P(X_n \geqq v^n \quad \text{for all large } n,\, T_L<\infty) \\ &\quad \geqq \sum_{i=0}^{k} P(T_L=i)(1 - P(B_1(n-i) < v^n \text{ infinitely often})^L) \\ &\quad = (1-\eta^L)\, P(T_L \leqq k) \quad \text{by (3.20)} \\ &\quad \to (1-\eta^L)\, P(T_L<\infty) \quad \text{as } k\to\infty. \end{aligned}$$

However $P(A \setminus \{T_L < \infty\}) = 0$ by Lemma 5, and hence

$$P(X_n \geqq v^n \quad \text{for all large } n, A) \geqq P(A) - \eta^L P(T_L < \infty) \to P(A) \quad \text{as } L \to \infty.$$

Thus

$$\liminf_{n \to \infty} X_n^{1/n} \geqq v \quad \text{a.s. on } A$$

for all $v < \mu_1$ and the lemma is shown. □

The theorem has nearly been proved. Let N be a positive integer such that $E(X_N) > 1$. If A holds then one of the events $A_0, A_1, \ldots, A_{N-1}$ holds, where

$$A_i = \{X_{Nk+i} \geqq 1 \text{ for infinitely many values of } k\}.$$

We claim that if A_i holds for some i then A_j holds for all j a.s., whence it follows that

$$P(A \Delta \{A_0 \cap A_1 \cap \ldots \cap A_{N-1}\}) = 0. \tag{3.21}$$

We prove this as follows. Suppose that A_i holds. As in the proof of Lemma 5, we define $J_1 = \min\{k: X_{Nk+i} \geqq 1\}$ and $J_{r+1} = \min\{k > J_r: X_{Nk+i} \geqq 1\}$. Finally we write $A_{i,k}$ for the event $\{X_n \geqq 1$ for all n satisfying $Nk + i \leqq n < N(k+1) + i\}$. We shall show that (on A_i) the event $A_{i,k}$ occurs for infinitely many values of k a.s., implying that A_j occurs for every j a.s. This is not difficult to accomplish, by adapting the proof of Lemma 5. We have as in (3.17) that

$$\begin{aligned} &P(A_{i,k} \text{ occurs for no value of } k, A_i) \\ &\leqq P(A_{i,J_r}^c \text{ for all } r, A_i) = \lim_{I \to \infty} \zeta(I), \end{aligned} \tag{3.22}$$

where

$$\zeta(I) = P(A_{i,J_r}^c \text{ for } 1 \leqq r \leqq I, J_I < \infty\}.$$

As in the proof of Lemma 5 we may use the superbranching inequality, and particularly Lemma 6 (replacing (3.16) by $\eta = P(B_1(n) = 0$ for some n, $1 \leqq n < N)$, where $(B_1(n): n \geqq 0)$ is a branching process with $B_1(0) = 1$ and family-sizes distributed as X_1), to conclude that $\zeta(I) \to 0$ as $I \to \infty$. It now follows immediately from (3.22) that $A_{i,k}$ occurs for infinitely many values of k a.s. on A_i, since otherwise there exists m such that $Y_r = X_{M+r} (r \geqq 0)$, defines a superbranching process for which the event corresponding to A_i occurs, but that corresponding to $\{A_{i,k}$ for some $k\}$ does not. Equation (3.21) follows. Finally let $0 \leqq j < N$, and define $Y_k = X_{Nk+j}$ for $0 \leqq k < \infty$. If A occurs, then by (3.21) A_j occurs a.s., and we may apply Lemma 7 to the sequence $\underline{Y}$ to find that

$$P(\liminf_{k \to \infty} X_{Nk+j}^{1/k} \geqq E(X_N), A) = P(A);$$

we have used the fact that $\underline{Y}$ satisfies the superbranching inequality and $E(Y_1 \mid Y_0=1) \geqq E(X_N) > 1$. It is not necessarily the case that $Y_0=1$, but we never used the fact that $X_0=1$ in the proof of Lemma 7. Thus

$$\liminf_{k\to\infty} X_{Nk+j}^{1/Nk} \geqq E(X_N)^{1/N} \qquad \text{a.s. on } A.$$

This is valid for all j with $0 \leqq j < N$, giving that

$$\liminf_{n\to\infty} X_n^{1/n} \geqq E(X_N)^{1/N} \qquad \text{a.s. on } A.$$

We let $N\to\infty$ to deduce that (3.15) holds, as required. □

Proof of Lemma 4. First we prove that $P(B(n)<an)$ decays at least geometrically as $n\to\infty$, when a is small enough. Fix a such that $0<a<1$. We note that $B(k) \leqq B(k+1)$ for all k, so that there exist at least $n(1-a)$ values of k such that $0 \leqq k < n$ and $B(k)=B(k+1)$, whenever $B(n)<an$. However

$$P(B(k)=B(k+1)) \leqq P(B(1)=1) \qquad \text{for all } k,$$

giving that

$$P(B(n)<an) \leqq P(Y_n \geqq n(1-a))$$

where Y_n has the binomial distribution with parameters n and $p=P(X_1=1)$. We use Markov's inequality in the usual way to find that there exists $\gamma(a) \geqq 0$ such that

$$P(B(n)>an) < \exp(-n\gamma(a)) \qquad \text{for all } n,$$

and furthermore $\gamma(a)>0$ if $a<1-p$.

Suppose that $0<\alpha<\mu$. Let $0<\beta<1$ and $0<a<1$; we shall choose β and a shortly. In the following, we shall occasionally write non-integral quantities in places where integral quantities are required; this minor aberration is easily corrected but the notation becomes less simple. We have that

$$\begin{aligned} P(B(n)<\alpha^n) &\leqq P(B(n)<\alpha^n \mid B(\beta n) \geqq an)\, P(B(\beta n) \geqq an) + P(B(\beta n)<an) \\ &\leqq P(B(n(1-\beta))<\alpha^n)^{an} + P(B(\beta n)<an). \end{aligned}$$

We pick β and a such that

$$\alpha^{(1-\beta)^{-1}} < \mu, \qquad a/\beta < 1-p$$

and recall from (1.3) that $P(B(n)<\delta^n)\to 0$ as $n\to\infty$ if $\delta<\mu$, to find that

$$P(B(n)<\alpha^n) \leqq P(B(m) \leqq \alpha^{m(1-\beta)^{-1}})^{an} + P(B(r)<r(a/\beta))$$

decays at least geometrically in n as $n\to\infty$; we have written $m=(1-\beta)n$ and $r=\beta n$ here. The proof is complete. □

Remark. Let $(X_n : n \geqq 0)$ be a sequence of random variables satisfying the strong superbranching inequality. Then in general there will be no limit theorem for the sequence $(X_n/EX_n : n \geqq 0)$. This is demonstrated by the following example due to Harry Kesten. Let $(Y_n: n \geqq 0)$ be any sequence of random variables, such that each Y_n takes values in $[\frac{1}{2}, 1]$, and let $(a_n: n \geqq 0)$ be a sequence of positive real numbers satisfying $a_{n+m} \geqq 2 a_n a_m$ for all $n, m \geqq 0$ (e.g., $a_n = 2^{n-1}$). Then $\underline{X}$ defined by $X_n = a_n Y_n$ for $n \geqq 0$ satisfies the strong superbranching inequality, but we have assumed extremely little about the distribution of $X_n/EX_n = Y_n/EY_n$.

4. Labelled Branching Processes and Random Cantor Sets

We establish next the appropriate terminology for labelled branching processes and associated random Cantor sets, more or less following the scheme of Falconer (1986) (see also Neveu (1986)).

We begin with the concept of a labelled branching process. For any set S, we write $S^* = \{\Lambda\} \cup S^2 \cup \ldots$ for the set of all finite ordered sequences of numbers of S, where Λ denotes the empty sequence. We write $\mathbb{N} = \{1, 2, \ldots\}$ and, for $\underline{i} = (i_1, i_2, \ldots, i_n) \in \mathbb{N}^n$ and $\underline{j} = (j_1, j_2, \ldots, j_m) \in \mathbb{N}^m$, we write $(\underline{i}, \underline{j})$ for the sequence $(i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_m) \in \mathbb{N}^{n+m}$.

A *branching tree* T is a subset of $\mathbb{N}^*$ defined as follows. First, $\Lambda \in \mathrm{T}$. There exists a non-negative integer $N(\Lambda)$ such that $i \in \mathbb{N}$ belongs to T if and only if $1 \leqq i \leqq N(\Lambda)$. Suppose that we have determined which points in $\{\Lambda\} \cup \mathbb{N} \cup \ldots \cup \mathbb{N}^n$ belong to T. If $\underline{i}_n \in \mathbb{N}^n$ and $i \in \mathbb{N}$, then there exists a non-negative integer $N(\underline{i}_n)$ with the property that $\underline{i}_{n+1} = (\underline{i}_n, i)$ belongs to T if and only if $\underline{i}_n \in T$ and $1 \leqq i \leqq N(\underline{i}_n)$. If this holds, we call $\underline{i}_{n+1}$ *a child* of $\underline{i}_n$. We may turn T into a graph with vertex set T by joining two vertices whenever one is a child of the other. T then becomes a rooted tree.

Now let V be a set, called the *label space*. To each point in T we assign a label from V, and we call the resulting tree a *labelled tree*. In this labelled tree, each point $\underline{i}$ has a certain number $N(\underline{i})$ of children, and these offspring have labels, say $L_1, L_2, \ldots, L_{N(\underline{i})}$ in lexicographic order. We may write

$$Z(\underline{i}) = (N(\underline{i});\ L_1, L_2, \ldots, L_{N(\underline{i})}),$$

a vector which describes the labelled offspring of $\underline{i}$.

Next we introduce randomness. Let

$$Z = (N;\ L_1, L_2, \ldots, L_N) \tag{4.1}$$

be a random vector, where N is a (random) non-negative integer and $L_1, L_2, \ldots, L_N$ are (random) labels chosen from V. Suppose that $\{Z(\underline{i}): \underline{i} \in \mathbb{N}^*\}$ is a family of independent copies of Z. We call the resulting (random) labelled tree T a *labelled branching process*. We speak of this process as having "typical offspring distributed as Z".

Consider now a labelled branching process with offspring distributed as Z. To each point $\underline{i}_n = (i_1, i_2, \ldots, i_n) \in T$, where $n \geqq 1$, there corresponds a unique path $\Lambda, \underline{i}_1, \underline{i}_2, \ldots, \underline{i}_{n-1}, \underline{i}_n$ joining Λ to $\underline{i}_n$. Writing $L(\underline{i})$ for the label of the point

$\underline{i}$ in T, this path gives rise to a sequence $(L(\underline{i}_1), L(\underline{i}_2), \ldots, L(\underline{i}_n))$ of labels; we call this sequence the *word* associated with $\underline{i}_n$, and write this word as

$$w(\underline{i}_n) = L(\underline{i}_1)\, L(\underline{i}_2) \ldots L(\underline{i}_n).$$

We write $T_n = \{\underline{i}_n = (i_1, i_2, \ldots; i_n): \underline{i}_n \in T\}$ for the set of members of the nth generation of the process, and write $W_n = \{w(\underline{i}): \underline{i} \in T_n\}$ for the set of distinct words of length n in the process.

If V is a vector space, the above labelled branching process is a branching random walk in V, in which the particle $\underline{i}$ at the point $p(\underline{i})$ gives birth to $N(\underline{i})$ children which are positioned at the points $\{p(\underline{i}) + L_j: 1 \leqq j \leqq N(\underline{i})\}$ where L_j is the label of the jth child of $\underline{i}$.

Next, we recall how labelled branching processes give rise to certain random Cantor sets in $\mathbb{R}^d$. We fix positive integers $d \geqq 1$ and $M \geqq 2$, and we write $\mathbb{N}(M) = \{0, 1, 2, \ldots, M-1\}$. Let T be a labelled branching process with label space $V = \mathbb{N}(M)^d$ and labelled offspring distributed as Z. With each word in V^* we associate a M-adic cube in $\mathbb{R}^d$ in the following way. With the empty word Λ, we associate the unit cube $I_0(\Lambda) = [0, 1]^d$. If $w = l_1\, l_2 \ldots l_n$ is a word of length $n \geqq 1$ and each $l_j \in \mathbb{N}(M)^d$ is written coordinatewise as $l_j = (l_j(1), l_j(2), \ldots, l_j(d))$ where $0 \leqq l_j(i) < M$ for all i, j, then we associate with w the cube

$$I_n(w) = \prod_{i=1}^{d} \left[\sum_{j=1}^{n} l_j(i)\, M^{-j}, \; \sum_{j=1}^{n} l_j(i)\, M^{-j} + M^{-n} \right]. \tag{4.2}$$

The process T determines a sequence $G_0, G_1, \ldots$ of closed random subsets of $[0, 1]^d$ defined by

$$G_0 = [0, 1]^d, \qquad G_n = \bigcup_{w \in W_n} I_n(w).$$

If $\underline{i}_{n+1} = (\underline{i}_n, i)$ belongs to T then $w(\underline{i}_{n+1}) = w(\underline{i}_n)\, L(i)$, giving that $I_{n+1}(w(\underline{i}_{n+1})) \subseteq I_n(w(\underline{i}_n))$, so that $G_{n+1} \subseteq G_n$ for all n. We are mainly interested in the limit set

$$G = \lim_{n \to \infty} G_n = \bigcap_{n=0}^{\infty} G_n,$$

which we term the random M-adic Cantor set associated with the labelled branching process.

We note that $G = \emptyset$, the empty set, if and only if the underlying branching process, with family-sizes distributed as N, is ultimately extinct.

The classical Cantor set is the set G obtained when $d = 1$, $M = 3$, and $Z = (2; 0, 2)$ almost surely. A less trivial example, to which we shall return later, has $d = 2$, $M = 2$, and

$$P(Z = (2; (1, 0), (1, 1))) = 1 - p$$

$$P(Z = (3; (0, 0), (0, 1), (1, 0))) = p,$$

where $0<p<1$.

Finally, we call such a labelled branching process a *M^d-adic tree* if $N=M^d$ and $\{L_1, L_2, \ldots, L_N\}=\mathbb{N}(M)^d$ almost surely.

5. Projections of Random M-adic Cantor Sets

For integers d and e satisfying $1\leqq e\leqq d$ and a vector $v=(v_1, v_2, \ldots, v_d)$, we write $\pi_e v=(v_1, v_2, \ldots, v_e)$ for the projection of v onto the subspace generated by the first e coordinates of the vector.

If G is a subset of $\mathbb{R}^d$, then $\pi_e G$ is the projection of G onto the e-dimensional subspace of $\mathbb{R}^d$ generated by the first e coordinates. A few general results are known about the Hausdorff dimensions of the collection of orthogonal projections of G over the set of all e-dimensional subspaces of $\mathbb{R}^d$; see Marstrand (1954) or Falconer (1985). We are interested here in the case when G is the random M-adic Cantor set generated by a labelled branching process with label space $\mathbb{N}(M)^d$, and with the single projection π_e. In this section we address the question of determining necessary and sufficient conditions for the property that $\pi_e G$ has non-zero e-dimensional Lebesgue measure. In the next section, we study the Hausdorff dimension of $\pi_e G$.

Let $V=\mathbb{N}(M)^d$ as before, and consider a labelled branching process $LBP(d)$ with labels in V and offspring distributed as $Z=(N; L_1, L_2, \ldots, L_N)$. For a given vector $s\in\mathbb{N}(M)^e$, we write $Z(s)$ for the number of labelled offspring of the hypothetical family with size N and labels $L_1, L_2, \ldots, L_N$ whose labels project onto s:

$$Z(s)=\sum_{i=1}^{N} 1_s(L_i) \tag{5.1}$$

where, for $l\in\mathbb{N}(M)^d$, $1_s(l)$ equals 1 or 0 depending on whether or not $\pi_e l=s$. We write

$$m_s=E(Z(s)) \tag{5.2}$$

for the mean number of such offspring, and

$$m=\prod_{s\in\mathbb{N}(M)^e} m_s, \tag{5.3}$$

noting that $m\leqq 1$ if and only if the geometric mean of the m_s is less than or equal to 1.

The labelled branching process LBP(d) gives rise to a projected labelled branching process, LBP(e) say, obtained by projecting each label in LBP(d) onto its first e coordinates. Let G be the random M-adic Cantor set generated by LBP(d). It is easy to see that $\pi_e G$ is the random M-adic Cantor set generated by LBP(e).

If LBP(e) is a M^e-adic tree, then $\pi_e G=[0, 1]^e$ almost surely, and we exclude this special case from the next theorem. We say that a labelled branching process becomes extinct if the associated tree is finite. The probability of extinction is the smallest root in [0, 1] of the equation $x=E(x^N)$. The following three statements are equivalent:

(i) LBP(d) becomes extinct,

(ii) $G=\emptyset$, the empty set,

(iii) $\pi_e G=\emptyset$, the empty set.

We write λ_e for e-dimensional Lebesgue measure.

Theorem 8. *Let G be the random M-adic Cantor set generated by* LBP(d) *and $\pi_e G$ the corresponding projected set generated by* LBP(e). *Suppose that* LBP(e) *is not a M^e-adic tree. Then $\lambda_e(\pi_e G)=0$ almost surely if $m\leqq 1$. If $m>1$ then $\lambda_e(\pi_e G)>0$ almost surely whenever $G\neq\emptyset$.*

Proof. As in Dekking (1987) this proof makes use of branching processes in varying and in random environments.

Writing W_n as before for the set of distinct words of length n in LBP(d) and

$$\pi_e W_n=\{\pi_e w\colon w\in W_n\}$$

for the set of words of length n in LBP(e), we have that $F_n=\pi_e G_n$ is given by

$$F_n=\bigcup_{s\in\pi_e W_n} I_n(s).$$

Certainly $F_n\downarrow F$ as $n\to\infty$, where $F=\pi_e G$, and so

$$E(\lambda_e(F_n))\to E(\lambda_e(F)) \qquad \text{as } n\to\infty, \tag{5.4}$$

by monotone convergence. Furthermore

$$\lambda_e(F_n)=M^{-ne}|\pi_e W_n|, \tag{5.5}$$

where $|S|$ denotes the cardinality of the set S. Thus

$$E(\lambda_e(F_n))=M^{-ne}\sum_w P(w\in\pi_e W_n), \tag{5.6}$$

where the sum is over all words w of length n in the alphabet $\mathbb{N}(M)^e$. As in Dekking (1987), we may relate this expression to the survival probability of a certain branching process in a random environment. Specifically, consider first a branching process in a varying environment in which there are M^e distinct possible environments labelled by $\mathbb{N}(M)^e$: we write this environment space as $\{\eta(s)\colon s\in\mathbb{N}(M)^e\}$. If the environment of the process in a given generation is $\eta(s)$ then the families of the members of that generation are distributed as $Z(s)$ in (5.1). For a given sequence $\eta(s_1)$, $\eta(s_2)$, … of environments, we write $Z(\eta(s_1), \eta(s_2), \ldots)$ for a branching process in a varying environment with initial size 1 and environment $\eta(s_i)$ in generation $i-1$; we write $Z_n(\eta(s_1), \eta(s_2), \ldots, \eta(s_n))$ for the size of the nth generation of this process. We may now see that

$$P(s\in\pi_e W_n)=P(Z_n(\eta(s_1), \eta(s_2), \ldots, \eta(s_n))>0) \tag{5.7}$$

for words $s=s_1 s_2 \dots s_n$ of length n in the alphabet $\mathbb{N}(M)^e$. Thus, from (5.6),

$$E(\lambda_e(F_n)) = M^{-ne} \sum_s P(Z_n(\eta(s_1), \dots, \eta(s_n)) > 0)$$
$$= M^{-ne} \sum_s P(\tilde{Z}_n > 0 \mid \zeta_i = \eta(s_{i+1}) \text{ for } 0 \leqq i < n)$$

where $\tilde{Z}_n$ is the size of the n^{th} generation of a branching process $\tilde{Z}$ in a random environment, having random environments $\zeta_0, \zeta_1, \dots,$ and a single founder member; as before the summations are over all words $s=s_1 s_2 \dots s_n$ of length n in the alphabet $\mathbb{N}(M)^e$. We are at liberty to specify how the environments of $\tilde{Z}$ are chosen, and we shall suppose that they are independent identically distributed environments chosen uniformly from the set $\{\eta(s): s \in \mathbb{N}(M)^e\}$. With this choice, the previous conditional probability may be written in terms of an absolute probability instead to find that

$$E(\lambda_e(F_n)) = P(\tilde{Z}_n > 0). \tag{5.8}$$

We call $\tilde{Z}$ the branching process in a random environment associated with LBP(e).

Suppose now that LBP(e) is not a M^e-adic tree, so that it is not the case that $P(Z(s)=1 \text{ for all } s)=1$. By the result of Athreya and Karlin (1971), under this condition it is the case that $P(\tilde{Z}_n > 0) \to 0$ as $n \to \infty$ if and only if $E(\log E(\tilde{Z}_1 \mid \zeta_0)) \leqq 0$. However, in our case, from (5.2) and (5.3),

$$E(\log E(\tilde{Z}_1 \mid \zeta_0)) = \sum_{s \in \mathbb{N}(M)^e} M^{-e} \log m_s$$
$$= M^{-e} \log m,$$

giving by (5.4) and (5.8) that $E(\lambda_e(F))=0$ if and only if $m \leqq 1$. Suppose now that $m>1$, so that $E(\lambda_e(F))>0$, and write $\delta = P(\lambda_e(F)=0)$; we note that $\delta<1$. We shall show that δ is a root of the equation

$$x = E(x^N); \tag{5.9}$$

combined with the fact that $\delta<1$, this will imply that δ equals the probability that LBP(d) becomes extinct, and the theorem will be proved. To show that δ satisfies (5.9), we argue in the usual way based on the natural recursion of LBP(d). We may write

$$F = \bigcup_{j=1}^{N} F(j) \tag{5.10}$$

where $F(j)$ is the random M-adic Cantor set obtained from the subtree of LBP(e) having as root the jth child of the root. Apart from a scale factor and a translation. $F(j)$ has the same distribution as F, so that $P(\lambda_e(F(j))=0) = P(\lambda_e(F)=0) = \delta$.

Furthermore, from (5.10), $\lambda_e(F)=0$ if and only if $\lambda_e(F(j))=0$ for all j. By the independence of $F(1), F(2), \ldots, F(N)$,

$$\delta = P(\lambda_e(F)=0)$$
$$= \sum_{n=0}^{\infty} P(N=n)\, P(\lambda_e(F)=0)^n = E(\delta^N).$$

This completes the proof of Theorem 8. □

Remark. We have proved in Theorem 8 that, if $m>1$, then $W=\lim_{n\to\infty} M^{-ne}|\pi_e W_n|$ exists a.s., and is (a.s.) strictly positive on the event that the process does not terminate. Harry Kesten has shown a corresponding result without the assumption on m: he has proved that $W=\lim_{n\to\infty} |\pi_e W_n|/E|\pi_e W_n|$ exists a.s. and is (a.s.) strictly positive on the event that the process is not extinct.

6. Dimensions of Random M-adic Canter Sets

We write $D_H(A)$ and $D_B(A)$ for the Hausdorff dimension and the box dimension of the set A, respectively, recalling that if A is a bounded subset of $\mathbb{R}^d$ then $D_B(A)$ is defined by

$$D_B(A) = \limsup_{r\to 0} \frac{\log N_r(A)}{\log(1/r)}$$

where $N_r(A)$ is the smallest number of closed balls in $\mathbb{R}^d$ with radius r whose union covers A. It is not hard to see that this is equivalent to

$$D_B(A) = \limsup_{n\to\infty} \frac{\log \nu_n^M(A)}{n \log M} \tag{6.1}$$

where $\nu_n^M(A)$ is the number of cubes $I_n(w)$, with $w=l_1 \ldots l_n$ ranging over the words of length n with $l_j \in \mathbb{N}(M)^d$ (see (4.2)) which intersect A. Box dimension is also known as Bouligand dimension, ε-entropy, entropy dimension, Pontryagin-Snirelman dimension, and so on (see Tricot (1981)). Clearly $D_H(A) \leq D_B(A)$ for all A and quite often these dimensions are equal (see e.g., Hawkes (1974)).

Let G be the random M-adic Cantor set generated by the labelled branching process LBP(d) with labelled offspring distributed as $Z=(N; L_1, L_2, \ldots, L_N)$. Several authors have proved that the Hausdorff dimension $D_H(G)$ of G satisfies $D_H(G)=D_B(G)=\log(EN)/\log M$ almost surely; see Peyrière (1978), Hawkes (1981), Falconer (1986, 1987), Mauldin and Williams (1986), Graf (1987). The common tool in this work is the martingale convergence theorem. We are interested here in the dimension of the projected set $\pi_e G$, where $1 \leq e \leq d$. The limit theorems of Sect. 3 are suited for determining the box dimension of $\pi_e G$.

Theorem 9. *Let G be the random M-adic Cantor set generated by* LBP(d) *and $\pi_e G$ the corresponding projected set generated by* LBP(e). *If $P(N=0)=0$, then*

$$D_B(\pi_e G) = \inf_{t\in[0,1]} \frac{\log\left(\sum_{s\in\mathbb{N}(M)^e} m_s^t\right)}{\log M} \quad \text{a.s.} \tag{6.2}$$

where m_s is the expected number of labelled offspring projecting onto $s\in\mathbb{N}(M)^e$.

The condition $P(N=0)=0$ is not essential to the argument underlying the proof of this theorem, and this condition may be removed in the usual way. Suppose that $P(N=0)>0$, so that the probability δ that LBP(d) becomes extinct satisfies $\delta>0$. Clearly $\pi_e G=\emptyset$ if LBP(d) becomes extinct. On the other hand, if $\delta<1$ then, on the event that LBP(d) does not become extinct, the conclusion of Theorem 9 enables us to calculate the almost sure box dimension of $\pi_e G$ by replacing LBP(d) by a suitably amended branching process in which extinction is impossible. For suppose that LBP(d) does not become extinct. The limit set G is unchanged if we remove from LBP(d) all those points which do not give rise to infinite lines of descent. The resulting family tree (for the moment we do not consider the labels of the points) has the same distribution as a branching process with family-size probability generating function

$$\hat{f}(x) = \frac{f((1-\delta)x+\delta)-\delta}{1-\delta}, \quad \text{where } f(x)=E(x^N) \tag{6.3}$$

(see, e.g., Athreya and Ney (1972, p. 47)). The families of this new branching process are non-empty and a typical family size N' satisfies the condition of Theorem 9: $P(N'=0)=0$. After we have deleted from LBP(d) those points with only finitely many descendents, there remains a copy of this new branching process, the points of which are labelled randomly from $\mathbb{N}(M)^d$. The offspring in a typical family have labels $L'_1, L'_2, \ldots, L'_{N'}$, where these labels depend only on N' and the original family size N; thus the new branching process, taken in conjunction with its random labels, constitutes a labelled branching process with label space $\mathbb{N}(M)^d$ and typical labelled offspring $(N'; L'_1, L'_2, \ldots, L'_{N'})$. As noted before, the random M-adic Cantor set generated by this process is just G, and so $D_B(\pi_e G)$ may be ascertained by applying Theorem 9 to this process.

As a corollary to Theorem 9 we have the following result which characterizes those sets for which the dimension does not decrease after projection.

Theorem 10. *If $P(N=0)=0$ and $\sum_s m_s \log m_s \leq 0$, then*

$$D_B(\pi_e G) = D_B(G) = \frac{\log EN}{\log M} \quad \text{a.s.;}$$

if $\sum_s m_s \log m_s > 0$ then $D_B(\pi_e G) < D_B(G)$ a.s. *and hence $D_H(\pi_e G) < D_H(G)$* a.s.

As an application of Theorem 9 we return to the example given at the end of Sect. 4. In this example $d=2$, $e=1$, $M=2$, and a typical labelled family is distributed as $Z=(N; L_1, L_2, \ldots, L_N)$ where

$$P(Z=(2; (1,0), (1,1))=1-p,$$
$$P(Z=(3; (0,0), (0,1), (1,0))=p,$$

and $0<p<1$. Easy calculations yield $m_0=2p$, $m_1=2-p$. Thus $m=\sqrt{m_0 m_1} =\sqrt{2p(2-p)}>1$ if and only if $p>1-\frac{1}{2}\sqrt{2}$, giving from Theorem 8 that $\lambda_1(\pi_1 G)>0$ almost surely if (and only if) this holds. If $0<p\leqq 1-\frac{1}{2}\sqrt{2}$ then $m_0+m_1>2$, giving by Jensen's inequality and the convexity of $g(x)=x\log x$ that $m_0 \log m_0+m_1 \log m_1>0$. It follows (cf. Dekking (1987)) that $\inf_{t\in[0,1]} \log(m_0^t + m_1^t)$ equals $H(\beta)$, where

$$\beta=\frac{\log(2p)}{\log(2p/(2-p))}$$

and $H(\beta)=-\beta\log(\beta)-(1-\beta)\log(1-\beta)$.

By Theorem 9,

$$D_B(\pi_e G)=\frac{H(\beta)}{\log 2} \quad \text{a.s.}$$

Proof of Theorem 9. Let $Q_n=|\pi_e W_n|$ be the number of distinct words of length n generated by the first n generations of LBP(e). Remembering (6.1) and the fact that $P(N=0)=0$, we have that $\nu_n^M(\pi_e G)=Q_n$, so that

$$D_B(\pi_e G)=\limsup_{n\to\infty} \frac{\log Q_n}{n\log M} \quad \text{a.s.} \tag{6.4}$$

Recall from Sect. 2 (part (v)) that $(Q_n: n\geqq 0)$ satisfies the strong superbranching inequality. Furthermore, Q_n is no larger than the total number of M-adic subcubes of $[0,1]^e$ with side length M^{-n}, giving that

$$\xi=\lim_{n\to\infty} \frac{1}{n}\log E(Q_n) \tag{6.5}$$

satisfies

$$\xi\leqq e\log M<\infty.$$

Applying Theorem 3, we deduce that

$$\lim_{n\to\infty} \frac{1}{n}\log Q_n=\xi \quad \text{a.s.} \tag{6.6}$$

From the proof of Theorem 8, and particularly (5.5) and (5.8),

$$E(Q_n)=M^{ne}E(\lambda_e(F_n))=M^{ne}P(\tilde{Z}_n>0) \tag{6.7}$$

where F_n is the subset of $[0, 1]^e$ generated by the first n generations of LBP(e) and $\tilde{Z}$ is the branching process in a random environment associated with LBP(e). According to the main result of Dekking (1988),

$$\lim_{n\to\infty} \frac{1}{n} \log P[\tilde{Z}_n > 0] = \log \inf_{t\in[0,1]} E(m(\zeta_0)^t) \tag{6.8}$$

where $m(\zeta_0) = E(\tilde{Z}_1 \mid \zeta_0)$. For the process $\tilde{Z}$ in (5.8),

$$E(E(\tilde{Z}_1 \mid \zeta_0)^t) = \sum_{s\in\mathbb{N}(M)^e} m_s^t M^{-e}. \tag{6.9}$$

From (6.4), (6.5) and (6.6) we have

$$D_B(\pi_e G) = \frac{1}{\log M} \lim_{n\to\infty} \frac{1}{n} \log E(Q_n) \quad \text{a.s.} \tag{6.10}$$

Combined with (6.7), (6.8) and (6.9) this yields (6.2). □

Proof of Theorem 10. As remarked in Dekking (1988), the limit in (6.8) is equal to $E\tilde{Z}_1$ if $\sum_{s\in\mathbb{N}(M)^e} m_s \log m_s \leqq 0$ (and strictly smaller otherwise). However

$$E\tilde{Z}_1 = M^{-e} \sum_{s\in\mathbb{N}(M)^e} m_s = M^{-e} EN,$$

since the N offspring of any point of LBP(d) may be partitioned into subsets according to the projections of their labels onto $\mathbb{N}(M)^e$. Thus, from (6.7), (6.8) and (6.10),

$$D_B(\pi_e G) = \frac{\log EN}{\log M} \quad \text{a.s.}$$

if $\sum_s m_s \log m_s \leqq 0$, and is strictly smaller otherwise. □

Final remark. Recently K. Falconer has found a simple argument to show that Theorem 9 is also true with D_B replaced by D_H.

Acknowledgements. We are grateful to Harry Kesten for his remarks on the superbranching inequalities.

References

Athreya, K.B., Karlin, S.: Branching processes with random environments I: extinction probabilities. Ann. Math. Statist. **42**, 1499–1520 (1971)

Athreya, K.B., Ney, P.E.: Branching processes. Berlin Heidelberg New York: Springer 1972

Biggins, J.D.: Chernoff's theorem in the branching random walk, J. Appl. Probab. **14**, 630–636 (1977)

Biggins, J.D.: Growth rates in the branching random walk, Z. Wahrscheinlichkeitstheor. Verw. Geb. **48**, 17–34 (1979)

Dekking, F.M.: Subcritical branching processes in a two state random environment, and a percolation problem on trees, J. Appl. Probab., **24**, 798–808 (1987)

Dekking, F.M.: On the survival probability of a branching process in a finite state i.i.d. environment, Stoch. Proc. Appl. in press
Falconer, K.J.: The geometry of fractal sets. Cambridge: Cambridge University Press 1985
Falconer, K.J.: Random fractals. Math. Proc. Camb. Phil. Soc. **100**, 559–582 (1986)
Falconer, K.J.: Cut-set sums and tree processes, Proc. Am. Math. Soc. **101**, 337–346 (1987)
Graf, S.: Statistically self-similar fractals. Probab. Th. Rel. Fields **74**, 357–392 (1987)
Hawkes, J.: Hausdorff measure, entropy, and the independence of small sets, Proc. London Math. Soc. (3), **28**, 700–724 (1974)
Hawkes, J.: Trees generated by a simple branching process. J. London Math. Soc. (2), **24**, 373–384 (1981)
Mandelbrot, B.: The fractal geometry of nature. New York: Freeman 1983
Mauldin, R.D., Williams, S.C.: Random recursive constructions: asymptotic geometric and topological properties. Trans. Amer. Math. Soc. **295**, 325–346 (1986)
Marstrand, J.M.: Some fundamental geometrical properties of plane sets of fractional dimensions. Proc. London Math. Soc. (3), **4**, 257–302 (1954)
Neveu, J.: Arbres et processus de Galton-Watson, Ann. Inst. Henri Poincaré **22**, 199–207 (1986)
Peyrière, J.: Mandelbrot random beadsets and birthprocesses with interaction, I.B.M. research report RC-7417 1978
Tricot, C.: Douze définitions de la densité logarithmique. C.R. Acad. Sc. Paric **293**, 549–552 (1981)

Received July 8, 1987; received in revised form February 2, 1988